



**UiT** The Arctic University of Norway

Faculty of Science and Technology, Department of Computer Science

## **EDMON - Electronic Disease Surveillance and Monitoring Network**

A Personalized Health Model-based Digital Infectious Disease Detection Mechanism using Self-Recorded Data from People with Type 1 Diabetes

Ashenafi Zebene Woldaregay

A dissertation for the degree of Philosophiae Doctor – February 2021

*“Whatever you do, work heartily, as for the Lord and not for men, knowing that from the Lord you will receive the inheritance as your reward. You are serving the Lord Christ.”*

***Colossians 3:23-24.***

## Dedication

Dedicated to my mom *Menda Legesse* and, my late father *Zebene Woldaregay* (†), whom I lost during the course of this Ph.D. journey.

## Abstract

Through time, we as a society have been tested with infectious disease outbreaks of different magnitude, which often pose major public health challenges. To mitigate the challenges, research endeavors have been focused on early detection mechanisms through identifying potential data sources, mode of data collection and transmission, case and outbreak detection methods, and further characterization techniques. Driven by the ubiquitous nature of smartphones and wearables equipped with a variety of physiological sensors, the current endeavor is targeted towards individualizing the surveillance effort, where case detection is realized by exploiting self-collected physiological data from wearables and smartphones. A personalized health model, which could be realized with the individual's self-generated data, plays a vital role in the feasibility of the individualized surveillance system. In this regard, the quest is to devise a model that can continuously monitor and screen the individual's health status and be able to detect when the individual becomes ill. However, despite the research effort and promises the data present, still there doesn't exist any implemented system with such capability.

This dissertation aims to demonstrate the concept of a personalized health model as a case detector for outbreak detection purposes by utilizing self-recorded data from people with type 1 diabetes. To realize the stated goal, the dissertation starts by defining a framework for the proposed personalized health model-based digital infectious disease detection system, i.e. Electronic disease surveillance and monitoring network (EDMON). Then, assess the predictive potential of different self-recorded parameters and optimal parameters indicative of infection onset are identified. A personalized health model (case detector) is devised using the identified optimal parameters and its detection capability is assessed. Finally, the dissertation closes by assessing user concerns, expectations, and willingness towards sharing self-collected health-related data to the proposed EDMON system.

The results have shown that infection onset triggers substantial deviations and results in prolonged hyperglycemia regardless of higher insulin injections and fewer carbohydrate consumptions. Per the findings, key parameters such as blood glucose level, amount of insulin injection, amount of carbohydrate consumption, and the ratio of insulin-to-carbohydrate are found to carry high discriminative power. A personalized health model designed based on a one-class classifier and unsupervised methods using some selected input parameters achieved promising detection performance. Experimental results show the superior performance of the one-class classifier and, particular models such as one-class support vector machine, k-nearest neighbor, and, k-means achieved relatively better performance from their respective groups. Besides, the result also revealed the effect of input parameters, data granularity, and sample sizes on model detection performances. On the other hand, people with diabetes have shown willingness to share most parameters, and, however, are reserved for certain parameters, i.e. geographical location (GPS), and signs of infection. In return to data sharing, this patient group expects some kind of incentives like tailored and personalized services, integrated view, and feedback. Further, data sharing concerns include transparency, confidence related to data

security and confidentiality (trust), ownership related to who owns the data, and storage related to where the data is stored.

The results presented have practical significance for understanding the effect of infection episodes amongst people with type 1 diabetes, and the nature of infection triggered deviations incurred on blood glucose dynamics. Further, it also has practical significance for understanding the potential of personalized health model in outbreak detection settings. The added benefit of the personalized health model concept introduced in this dissertation lies in its usefulness beyond the surveillance purpose. In this regard, for example, the presented personalized health model can also be used for other purposes such as to devise decision support tools and learning platforms for the patient to properly manage infection-induced crises. Generally, developing a personalized health model-based digital infectious disease detection system like EDMON, which aims for early detection, i.e. during the incubation period, requires considering various aspects, and the results presented in this dissertation construct evidence that supports the efforts towards building the next generation personalized health model-based-digital infectious disease surveillance systems and provoke further thoughts in this challenging field.

## Acknowledgments

Every beginning has an end, and the time has come to close the chapter of my Ph.D. journey. In this challenging at a time and fascinating journey, I was very blessed to have so many outstanding and bright people around me, who put an everlasting mark on my career as a researcher.

First of all, I would like to express my deepest gratitude to my supervisor, *Professor Gunnar Hartvigsen*, for his unlimited, and continuous guidance throughout the Ph.D. journey. His patient and perseverance through the darkest time of the journey, coupled with his constructive comments and suggestions have enabled me to capitalize on my potential to seek beyond what is known. Besides the science, your jokes, and your efforts to ease tensions during hard times of the Ph.D. are exceptional and I have learned a lot from you. Without you, I can't see myself succeeding in this tough journey. Dear *Gunnar*, you are one of a kind, and I really want to thank you for everything you have done to help me realize my dreams to become an independent researcher.

I would like to thank my co-supervisor, *Professor Eirik Årsand*, for his insightful and expert comments and suggestions to turn this Ph.D. project into a reality. Your calm personality coupled with deep knowledge of the topic was always an asset for me to have you while navigating through the different stages of the dissertation. Dear *Eirik*, I want to thank you for all your support throughout this Ph.D. journey.

I would also like to thank all the outstanding individuals at the health informatics and technology (HI&T) group, previously known as the Medical informatics and telemedicine (MI&T) group, for their wonderful support throughout the Ph.D. journey. Particularly, Ph.D. fellow *André Henriksen*, it was great working with you, and would like to say thank you for the exceptional time we have. Further, I would like to thank my employer, the Department of Computer Science – University of Tromsø (UiT), for creating this opportunity to realize my goal to become an independent researcher and also for providing every assistance I ever asked for.

I'd like to thank my parents, brothers, and my family – *Michael Ashenafi and Tinsea Eshetu*, for all the support, both emotional and spiritual, you provided as I thrive through this Ph.D. journey. Without your assistance and patience, I couldn't see myself succeeding through this tough journey. Dear *Tinsea*, thank you very much for everything and you are truly one of a kind. My lovely son, *Michael*, you have been a challenge and at the same time a reason and motivation to succeed.

Oh Lord I thank you, for you are with me in every walk of my life. “ አስጀምሮ ይህን ስራ ላስጨረስኝ ንጉስ የፈጠረኝ አምላክ የአለም መድሃኒት ፍቃዱ ደግ ነውና እፊቱ በግንባራ ተደፍቼ አመሰግነዋለሁ። ተመስገን!!! ”

# Table of Contents

Dedication .....	iii
Abstract .....	iv
Acknowledgments .....	vi
Table of Contents .....	vii
List of Tables.....	xi
List of Figures .....	xii
List of Publications.....	xvi
1 Introduction .....	1
1.1 Background.....	1
1.2 Motivation .....	3
1.3 Research context.....	4
1.4 Research problem and questions .....	4
1.5 Significance of the study .....	6
1.6 Claimed contribution and included papers .....	7
1.6.1 Contribution of the dissertation.....	7
1.6.2 Other contributions.....	8
1.6.3 Statements of originality .....	9
1.6.4 Defining the red thread.....	12
1.7 Organization of the dissertation.....	13
2 Theoretical Background .....	15
2.1 Blood glucose dynamics .....	15
2.2 Infections incidences in people with diabetes .....	18
2.2.1 Acute illness and glucose metabolism .....	18
2.2.2 Blood glucose control during infection incidences .....	18
2.2.3 The literature on using diabetes profile for infection detection .....	21
2.3 Chapter summary.....	22
3 Materials and Methods .....	23
3.1 A general overview of research approaches .....	23
3.1.1 Overview of approaches used in different phases of the study .....	23
3.2 Overview of the dissertation stage.....	23
3.3 Empirical data analysis and modeling .....	26
3.3.1 Definition of key terms .....	26

3.3.2	Datasets and data sources .....	27
3.3.3	Data re-sampling, imputation and preprocessing .....	29
3.3.4	Empirical data analysis.....	31
3.3.5	Modeling approaches .....	33
3.4	Assessment of concerns, expectations, and willingness.....	34
3.5	Methodology critique .....	35
3.5.1	Empirical data analysis and modeling.....	35
3.5.2	Assessment of concerns, expectations, and willingness .....	36
3.6	Hardware and software tools .....	36
3.7	Chapter summary .....	37
4	EDMON - A Personalized Health Model-based Digital Infectious Disease Detection System: Framework and Challenges .....	39
4.1	Introduction .....	39
4.2	The proposed framework.....	40
4.2.1	The Patient unit .....	40
4.2.2	Computing unit.....	45
4.2.3	The Information dissemination (visualization) unit.....	47
4.3	Motivational and ethical challenges .....	48
4.4	Knowledge summary .....	48
4.5	Chapter summary.....	49
5	Infection Characterization and Parameter Selection .....	51
5.1	Introduction .....	51
5.2	Raw dataset.....	51
5.3	Data preparation and comparative analysis .....	53
5.3.1	Data trend .....	53
5.3.2	Data Distribution .....	64
5.4	Characterizing the observed phenomenon.....	65
5.5	Limitation .....	66
5.6	Knowledge Summary .....	66
5.7	Chapter Summary .....	67
6	Designing a Personalized Health Model .....	69
6.1	Introduction .....	69



6.2	Input features .....	70
6.2.1	Description of input features .....	72
6.3	Practical experiments and results .....	77
6.3.1	Model evaluation.....	78
6.3.2	One-class classifier performance .....	79
6.3.3	Comparison of input features .....	88
6.3.4	Unsupervised method performance.....	91
6.3.5	Comparison of unsupervised versus one-class classifier .....	94
6.3.6	Computational Time.....	95
6.3.7	Practical significance.....	97
6.4	Limitation .....	98
6.5	Knowledge Summary .....	98
6.6	Chapter summary.....	99
7	Concerns, Expectations, and Willingness Towards Sharing Self-Collected Health-Related Data .....	101
7.1	Introduction .....	101
7.2	Data recording related experiences and preference.....	102
7.2.1	Experience and usage of health-tracking technologies .....	102
7.2.2	mHealth apps and features preference .....	103
7.3	Data sharing.....	104
7.3.1	Data sharing concerns and expectations.....	105
7.3.2	Willingness towards anonymous sharing of data.....	106
7.4	Practical Implication.....	108
7.5	Limitation .....	109
7.6	Knowledge summary .....	109
7.7	Chapter summary.....	110
8	Conclusion and Future Works.....	113
8.1	Contribution.....	113
8.2	Main Conclusion.....	115
8.2.1	Main Research Problem – MP .....	118
8.3	Future Works .....	119
9	References .....	122

10 Appendix A - Score plot of the models using the univariate input feature: Ratio of Insulin to Carbohydrate ..... 135

11 Appendix B – Models performance using the univariate input feature: Ratio of Insulin-to-Carbohydrate ..... 152

12 Included Papers ..... 159

# List of Tables

<b>Table 1:</b> Contributions and research questions addressed in the dissertation.....	8
<b>Table 2:</b> Using diabetes profile for detecting infection incidences among people with diabetes. .....	22
<b>Table 3:</b> Key terms used in empirical data analysis and modeling. ....	27
<b>Table 4:</b> Type of datasets and diabetes technologies used for self-management, source [259], Table 1. ....	28
<b>Table 5:</b> Detailed description of participant characteristics, source [259], Table 2. ....	29
<b>Table 6:</b> Description of self-collected user data, source [259], Table 3.....	30
<b>Table 7:</b> Re-sampled, imputed, and pre-processed data, source [259], Table 4.....	31
<b>Table 8:</b> Approaches used in estimating the distribution of insulin-to-carbohydrate ratio, source [259], Textbox 1. ....	32
<b>Table 9:</b> Approaches used in estimating the distribution of blood glucose levels and insulin-to- carbohydrate ratio, source [259], Textbox 2.....	33
<b>Table 10:</b> Average and standard deviation of blood glucose levels, the sum of insulin doses (bolus), and the sum of carbohydrates intake within pre-infection, infection, and post- infection weeks, source [259], Table 5. ....	59
<b>Table 11:</b> Models tested for the proposed personalized health model. BD = boundary and domain-based, DN= density-based, RE = reconstruction-based methods.....	78
<b>Table 12:</b> Bivariate input features - optimal values of thresholds used in performance evaluation. The values given as $T_h$ are the optimal threshold values used for each patient-year, $h$ depicting that particular year, source [258], Table 8. ....	92
<b>Table 13:</b> Univariate input features - optimal values of thresholds used in performance evaluation. The values given as $T_h$ are the optimal threshold values used for each patient-year, $h$ depicting that particular year. ....	93
<b>Table 14:</b> Participants demographics, familiarity with wearable technologies, and data sharing experiences. Na depicts the number of respondents who declined to answer, source [257], Table 1.....	102
<b>Table 15:</b> Participants' willingness towards sharing specific health-related data, source [257], Table 2. ....	107
<b>Table 16:</b> Topics and findings related to data recording and sharing, requirements of the proposed EDMON system, and mitigation options for challenges that arise.....	108

## List of Figures

<b>Figure 1:</b> EDMON system overview, adapted from [259], Figure 14. ....	4
<b>Figure 2:</b> Blood glucose homeostasis (regulation) [10; 208], source Campbell, Figure 25.10. ....	16
<b>Figure 3:</b> Factors affecting blood glucose dynamics. Figure (a) depicts some of the factors that affect blood glucose levels and Figure (b) depicts a grouping of these factors amongst the individuals. ....	17
<b>Figure 4:</b> Depicts what patients discuss on online forums about infection and blood glucose levels, source [78; 141; 142]. Figure (a) depicts type 2 diabetes and Figure (b) depicts type 1 diabetes. ....	21
<b>Figure 5:</b> Progressive stage of the dissertation. ....	24
<b>Figure 6:</b> The proposed framework of EDMON - a personalized health model-based digital infectious disease detection system, adapted from [58; 255; 259], Figure 14. ....	40
<b>Figure 7:</b> Alternative methods for realizing the proposed personalized health model, source [259], Figure 15. ....	46
<b>Figure 8:</b> Example depicting mapping of a detected cluster in a hypothetical outbreak. ....	48
<b>Figure 9:</b> Illustrates a particular patient-year with influenza (flu) infections in the first week of December. ....	52
<b>Figure 10:</b> Illustrates a particular patient-year with influenza (flu) infections in the second week of August. ....	52
<b>Figure 11:</b> A specific regular patient-year depicting blood glucose levels, amount of carbohydrate consumption and insulin injections, and fluctuations of the insulin-to-carbohydrate ratio. Figure (a) illustrates fluctuations within the daily timeframe. Figure (b) illustrates fluctuations within the hourly timeframe. ....	54
<b>Figure 12:</b> A specific regular patient-year depicting blood glucose levels, amount of carbohydrate consumption and insulin injections, and fluctuations of the insulin-to-carbohydrate ratio. Figure (a) illustrates fluctuations within the daily timeframe. Figure (b) illustrates fluctuations within the hourly timeframe. ....	55
<b>Figure 13:</b> A particular patient-year with influenza (flu) incidence within the first week of December, and depicts blood glucose levels, amount of carbohydrate consumption and insulin injections, and fluctuations of the insulin-to-carbohydrate ratio. Figure (a) illustrates fluctuations within the daily timeframe. Figure (b) illustrates fluctuations within the hourly timeframe. ....	57
<b>Figure 14:</b> A particular patient-year with fever-free mild and light cold in the first week of August and mid-February respectively, and influenza (flu) incidence in mid-August. The figures depict blood glucose levels, amount of carbohydrate consumption and insulin injections, and fluctuations of the insulin-to-carbohydrate ratio. Figure (a) illustrates fluctuations within the daily timeframe. Figure (b) illustrates fluctuations within the hourly timeframe. ....	58

<b>Figure 15:</b> Comparison of blood glucose levels during pre-infection week, infection week, and post-infection week. As can be seen, the blood glucose is elevated during the infection week as compared to the pre and post-infection weeks. ....	60
<b>Figure 16:</b> Comparison of insulin (bolus) injections during the pre-infection week, infection week, and post-infection week. As can be seen, the amount of insulin (bolus) intake is elevated during the infection week as compared to the pre and post-infection weeks. ....	62
<b>Figure 17:</b> Comparison of carbohydrate consumptions (grams) during the pre-infection week, infection week, and post-infection week. As can be seen, the amount of Carbohydrate (grams) intake is significantly lower during the infection week as compared to the pre and post-infection weeks. ....	63
<b>Figure 18:</b> Kernel density estimation of a particular patient-year. Figure (a) depicts the univariate kernel estimation. Figure (b) depicts the bivariate kernel estimation. ....	65
<b>Figure 19:</b> Quadrants of wellness in an individual with type 1 diabetes, source [258], Figure 7. ....	72
<b>Figure 20:</b> A daily and hourly plot of typical smoothed regular patient-year depicting the scatter plot of average blood glucose levels vs. the ratio of insulin to carbohydrate. ....	73
<b>Figure 21:</b> A daily and hourly plot of typical smoothed regular patient-year depicting the scatter plot of average blood glucose levels vs. the ratio of insulin to carbohydrate. ....	74
<b>Figure 22:</b> A daily and hourly plot of the typical smoothed patient-year with an event of infection influenza (flu), and depicting the scatter plot of average blood glucose levels vs. the ratio of insulin to carbohydrate. ....	75
<b>Figure 23:</b> A daily and hourly plot of the typical smoothed patient-year with an event of infection influenza (flu), and depicting the scatter plot of average blood glucose levels vs. the ratio of insulin to carbohydrate. ....	76
<b>Figure 24:</b> Characteristics of the input feature, i.e. insulin to carbohydrate ratio, with simulated infection states of varying degree and shape ( $\alpha = 0\%$ , $10\%$ , $20\%$ , $30\%$ , and $40\%$ ). ..	77
<b>Figure 25:</b> Bivariate input feature - The models' median and average F1-score over the daily raw datasets, source partially from [258], Figure 8. ....	81
<b>Figure 26:</b> Bivariate input feature - The models' median and average F1-score over the daily smoothed datasets, source partially from [258], Figure 8. ....	82
<b>Figure 27:</b> Bivariate input feature - The models' median and average F1-score over the hourly smoothed datasets, source partially from [258], Figure 8. ....	83
<b>Figure 28:</b> Univariate input feature - The models' median and average F1-score over the daily raw datasets. ....	84
<b>Figure 29:</b> Univariate input feature - The models' median and average F1-score over the daily smoothed datasets. ....	85
<b>Figure 30:</b> Univariate input feature - The models' median and average F1-score over the daily smoothed datasets. ....	86
<b>Figure 31:</b> Univariate input feature - median and average performance (F1-score) of the models for the five patient years injected with different degree of deviations. ....	88

<b>Figure 32:</b> Daily raw dataset - performance comparison (F1-score) of models using bivariate input, i.e. blood glucose levels and insulin to carbohydrate ratio, and univariate input feature, i.e. insulin to carbohydrate ratio, based on the daily raw dataset. The error bars are given in terms of the overall mean and standard deviation of each model across all the patient-years and infection states. ....	89
<b>Figure 33:</b> Daily smoothed dataset - performance comparison (F1-score) of models using bivariate input, i.e. blood glucose levels and insulin to carbohydrate ratio, and univariate input feature, i.e. insulin to carbohydrate ratio, based on the daily smoothed dataset. The error bars are given in terms of the overall mean and standard deviation of each model across all the patient-years and infection states. ....	90
<b>Figure 34:</b> Hourly smoothed dataset - performance comparison (F1-score) of models using bivariate input, i.e. blood glucose levels and insulin to carbohydrate ratio, and univariate input feature, i.e. insulin to carbohydrate ratio, based on the hourly smoothed dataset. The error bars are given in terms of the overall mean and standard deviation of each model across all the patient-years and infection states.....	90
<b>Figure 35:</b> Bivariate input features - performance comparison (F1-score) of the unsupervised models using bivariate input, i.e. blood glucose levels and insulin to carbohydrate ratio, source [258], Table 8. ....	92
<b>Figure 36:</b> Univariate input features - performance comparison (F1-score) of the unsupervised models using univariate input, i.e. insulin to carbohydrate ratio .....	93
<b>Figure 37:</b> Performance comparison (F1-score) of the unsupervised models using a bivariate input, i.e. blood glucose levels and insulin to carbohydrate ratio versus a univariate input, i.e. insulin to carbohydrate ratio, source partially from [258], Table 8. ....	94
<b>Figure 38:</b> Performance comparison (F1-score) of unsupervised and one-class classifier version of the local outlier factor (LOF) model using both a bivariate input, i.e. blood glucose levels and insulin to carbohydrate ratio, and a univariate input, i.e. insulin to carbohydrate ratio. ....	95
<b>Figure 39:</b> Average running time required for the training and detection phase. Figure (a) depicts the models' time requirement during the training phase. Figure (b) depicts the models' time requirement during the detection phase. The labels in X-axis stands for a quotient of the sample size divided by twenty-four, i.e.10 (10*24), 20 (20*24), 30 (30*24), 40 (40*24), 50 (50*24), 60 (60*24), 70 (70*24), 80 (80*24), 90 (90*24), 100 (100*24), 110 (110*24), 120 (120*24) sample objects. The models' average computational time requirement is depicted on the Y-axis. ....	96
<b>Figure 40:</b> Health-tracking technologies usage and experience among the three groups, source [257], Figure 2 (b).....	103
<b>Figure 41:</b> Preference of mobile health (mHealth) app features, source [257], Figure 1 (b).	104
<b>Figure 42:</b> Participants' concerns towards sharing health data, source [257], Figure 1 (a).	105
<b>Figure 43:</b> Participants' expectations in response to data sharing, source [257], Figure 2 (a). ....	106

## Abbreviations

EDMON = Electronic disease surveillance and monitoring network

CGM = Continuous glucose monitoring

SMBG = Self-management blood glucose

BG = Blood glucose

BD = Boundary and domain-based

Carbs = Carbohydrate

DN = Density based

RE = Reconstruction based

LOF = local outlier factor

COF = Connectivity-based outlier factor

KNN = K-nearest neighbor

V-SVM = one-class support vector machine

SVDD = Support vector data description

IncSVDD = Incremental support vector data description

P = Parzen

NP = Naïve Parzen

PCA = Principal component analysis

SOM = Self-organizing map

MCG = Minimum covariance Gaussian

MOG = Mixture of Gaussian

NN = Nearest neighbor

MST = Minimum spanning tree

ROC = Receiver operating curve

AUC = Area under the receiver operating curve

T = Threshold

## List of Publications

- Paper 1** Woldaregay, A. Z., Launonen, I. K., Årsand, E., Albers, D., Holubová, A., & Hartvigsen, G. (2020). *Toward Detecting Infection Incidence in People With Type 1 Diabetes Using Self-Recorded Data (Part 1): A Novel Framework for a Personalized Digital Infectious Disease Detection System*. *J Med Internet Res*, 22(8), e18911. doi:10.2196/18911
- Paper 2** Woldaregay, A. Z., Arsand, E., Botsis, T., Albers, D., Mamykina, L., & Hartvigsen, G. (2019). *Data-Driven Blood Glucose Pattern Classification and Anomalies Detection: Machine-Learning Applications in Type 1 Diabetes*. *J Med Internet Res*, 21(5), e11030. doi:10.2196/11030
- Paper 3** Woldaregay, A. Z., Launonen, I. K., Albers, D., Igual, J., Årsand, E., & Hartvigsen, G. (2020). *A Novel Approach for Continuous Health Status Monitoring and Automatic Detection of Infection Incidences in People With Type 1 Diabetes Using Machine Learning Algorithms (Part 2): A Personalized Digital Infectious Disease Detection Mechanism*. *J Med Internet Res*, 22(8), e18912. doi:10.2196/18912
- Paper 4** Woldaregay, A. Z. *Automatic Detection of Infection State in Individuals with Type 1 Diabetes Under Free-Living Conditions Using Using Self-Recorded Insulin and Carbohydrate Information (Part 3)*. EARLY DRAFT MANUSCRIPT.
- Paper 5** Woldaregay, A. Z., Henriksen, A., Issom, D. Z., Pfuhl, G., Sato, K., Richard, A., Lovis, C., Arsand, E., Rochat, J., & Hartvigsen, G. (2020). *User Expectations and Willingness to Share Self-Collected Health Data*. *Stud Health Technol Inform*, 270, 894-898. doi:10.3233/SHTI20029



# 1 Introduction

## 1.1 Background

Both in the past and present, we as a society have experienced infectious disease outbreaks of different magnitude often resulting in mass fatalities and casualties. These incidences are either naturally occurring or man-made, e.g. bio-weapons, phenomenon, and known to inflict major public health challenges [52; 237]. In this regard, there have been continuous endeavors to contain and reduce the burden through early and timely detection of the incidences and implementing proper infection control measures to mitigate the infection rate [46]. These efforts are mainly targeted in identifying potential data sources, mode of data collection and transmission, case definitions and detection methods, outbreak detection methods, and other outbreak characterization techniques [46; 52; 236; 237]. The transition from the traditional surveillance system, which mainly exploits high diagnostic precision data from the laboratory, to the state-of-the-art syndromic surveillance system, which relies on pre-diagnostic data has created opportunities to improve the timeliness of detection [55; 102; 170]. The state-of-the-art systems exploit data aggregated at the population level including clinical data, e.g. chief complaint data [227], and other surrogate data such as different search engines, social media, absenteeism, pharmacy drug sells, and other various internet sources [79; 138; 143; 166; 170]. Furthermore, its capability to collect and transmit data in near real-time coupled with the availability of outbreak detection algorithms that can operate in real-time has further strengthened its significance in this regard [46; 227; 262]. However, as compared to the traditional surveillance system, outbreak characterization resulting from the syndromic surveillance system remain a significant challenge since it mainly operates with low-diagnostic precision (pre-diagnostic) data aggregated at the population level [236; 237]. In other words, a syndromic surveillance system lacks a look-back capability, which is a vital mechanism to trace-back to the sick individual for further clarification of the pathogens involved. Actually, it is worth mentioning that look-back capability is challenging by itself given the privacy and confidentiality issues posed.

Currently, the endeavor is targeted towards individualizing the surveillance effort [194], where the case detection is realized by exploiting self-collected physiological data from different wearables and sensors [157; 181]. In this regard, the main drivers are the rapid progress in information and communication technology, and the widespread availability of different smartphones and wearables equipped with a variety of physiological sensors, which created a suitable platform to easily self-track health data [106; 181; 185; 202; 205]. These technologies are increasingly been integrated into our daily life for a variety of reasons ranging from fitness tracking to managing diseases [20; 74; 85; 95; 106; 154; 203; 243]. As a result, a huge amount of data are being recorded on a daily basis that grows at an unprecedented rate [71; 214; 231]. The existence of these data is the cornerstone in the individualization effort that created unbounded possibilities for improving the state-of-the-art disease surveillance systems [44; 194]. Accordingly, the transition in granularity from population-level surveillance into individualized surveillance are becoming feasible and converging at a faster rate. In this regard, for instance, recent studies have shown the feasibility of smartwatches and wearable technology

in monitoring, detecting, and predicting illness [25; 157; 180; 181; 202; 223]. As compared to the existing syndromic surveillance system, individualization of the surveillance has a double benefit, it can be geared to monitor and notify changes in individual health status and simultaneously monitor the status of infectious disease outbreaks among the public [181; 194]. This is a great leap forward because previous systems, e.g. *Google Flu Trends* [86; 138] and *HealthMap* [45; 79], consider only detecting health threats at population levels and don't possess a mechanism to trace back to the sick individual. The potential and promises of the individualization effort lie in the fact that it operates with a secondary source of information mainly data generated for other purposes outside of the healthcare service settings. However, as described earlier, it faces significant challenges concerning user privacy and confidentiality, and ethical issues as it directly deal with individual health-related data [82; 128; 131; 157; 214]. Moreover, another prominent challenge in such a system is the availability of quality user-generated health-related data, and how to motivate and buy user trust to get the required data is a topic that needs to be carefully addressed [91; 126; 172].

A personalized health algorithm, which could be modeled with the individual's self-generated data, plays a vital role in the feasibility of the individualized surveillance system [157; 181]. In this regard, the quest is to devise a model that can continuously monitor and screen the individual's health status and be able to detect when the individual becomes ill in either real-time or near real-time using the individual's self-generated data. After having a successful case detector at an individual level, it becomes straightforward to realize a method for detecting a cluster of infected individuals on a Spatio-temporal basis [181; 237; 262]. As described above, apart from its use in detecting infectious disease outbreak, this type of early warning system can also address the need of having real-time health status information for different groups in the population, including chronic patients, aging populations, people using ambient assisted living, and healthy individuals [5; 12; 13; 72; 94; 114; 155; 193; 196; 243; 268]. However, despite the research effort and promise the data presents, still there doesn't exist any implemented system with such capability [181; 237]. *This dissertation aims to demonstrate the concept of a personalized health model as a case detector for outbreak detection purposes by utilizing self-recorded data from people with type 1 diabetes.* It further examines users' concerns, expectations, and willingness towards sharing self-collected health-related data for the proposed EDMON system. The dissertation starts by defining a framework for the proposed personalized health model-based digital infectious disease detection system. Then, it quantifies the extent of infection triggered deviations on the key parameters of blood glucose dynamics amongst people with type 1 diabetes. Next, it explores the state-of-the-art approaches for classifying blood glucose patterns and detecting anomalies in blood glucose dynamics amongst people with type 1 diabetes giving due emphasis to the cause and effect relationship between parameters used in modeling. Further, it devises a personalized health model (case detector) for continuous monitoring of an individual's health status and automatic detection of infection onset. Finally, the dissertation closes by assessing user concerns, expectations, and willingness towards sharing self-collected health-related data to the proposed system.

## 1.2 Motivation

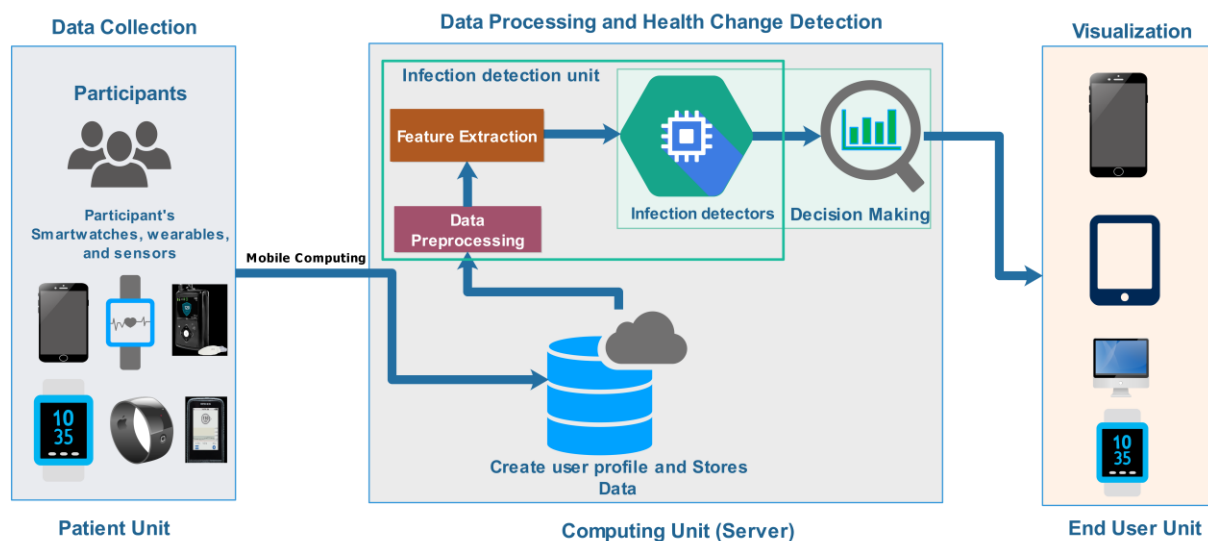
The current tragic event due to the corona-virus (SARS-CoV-2) outbreak, also known as COVID-19, has resulted in mass panic among the society and loss of many lives around the globe [21; 186]. The incident has thought us that we as a society stand far from prepared for such kinds of outbreaks. The first step to reducing the impact of any infectious disease outbreaks is to devise an early detection system, which requires a well-equipped digital infectious disease detection system. As has been tested frequently through time, the existing disease surveillance systems have several shortcomings, which calls for revolutionizing the system by integrating with novel data sources and aberration detection mechanism [54; 157; 170; 181; 199; 237]. The availability of different physiological sensors coupled with a huge amount of data generated daily creates enormous opportunities to advance the state-of-the-art system by driving a lot more research to be conducted to improve the time delay in the existing systems, given the profound implication of time criticality for disease surveillance design. In this regard, the timeliness requirement calls for the integration of new data sources, and novel aberration detection methods with a system that can collect and analyses data in real-time or near real-time [108; 237].

Self-recorded health-related data from people with type 1 diabetes could be one potential choice, which is becoming rich in both quality and quantity as a result of advancement in diabetes technologies [11; 19; 23]. Type 1 diabetes is a chronic disease that results in a lack of blood glucose control as a result of insulin deficiency within the body [56; 64]. This group of patients is expected to follow complex treatment regimens to control their blood glucose levels within the recommended targets including tracking blood glucose levels throughout the day, administrating balanced insulin while considering factors like blood glucose levels, meal intake, physical activity, and other possible factors. This group of patients are estimated to account for more than five percent of all the people diagnosed with diabetes and is expected to grow [161]. Infection onset is known to induce hyperglycemia (elevated blood glucose levels) among people with diabetes [38; 47; 62; 115; 133]. In this regard, there is previous evidence in the literature that shed some light on the possibility of harnessing self-recorded data from people with type 1 diabetes as a potential data source to detect infectious disease outbreaks [17; 30; 31; 34; 36; 38; 90].

In this dissertation, the predictive potential of different self-recorded parameters is investigated and optimal parameters that are indicative of infection onset are identified. A personalized health model is devised using the identified optimal parameters and its detection capability is assessed. It is apparent that blood glucose dynamics are affected by various factors apart from infection incidences, and hence, these optimal features are an important building block for realizing a personalized health model with a minimum and acceptable false alarm rate. Based on these findings, it remains straightforward to realize the proposed EDMON system for assisting in detecting infectious disease outbreaks among the public. In this regard, the main idea here is to exploit self-recorded health-related data from people with type 1 diabetes to assist detection of infectious disease outbreaks among the public.

### 1.3 Research context

This dissertation is part of the “*EDMON-Electronic Disease Surveillance and Monitoring Network*” project, which is a personalized health model-based-digital infectious disease detection system. EDMON aims to exploit self-generated data from various population groups, e.g. healthy and chronic patients, to assist in detecting infectious disease outbreaks among the public. In this regard, this dissertation presents a case study using self-recorded data from people with type 1 diabetes. The proposed EDMON system incorporates five units; *patient unit*, *data repository unit*, *infection detection unit*, *clustering unit (decision making)*, and *information dissemination unit (visualization)*, as shown in **Figure 1**. The patient unit includes a user mobile devices terminal used for collecting and transmitting data to a central server. The data repository unit stores the incoming user data. The infection detection unit tracks the individual health status by analyzing the data in real-time based on a personalized health model. The clustering unit identifies any possible cluster of infected individuals based on a Spatio-temporal analysis. The information dissemination unit provides real-time information regarding the current status of the area under surveillance for the concerned bodies, i.e. ordinary citizens, patients, and public health officials. *This dissertation focuses on topics related to the patient unit and infections detection unit and investigates how to provide a solution for these topics.*



**Figure 1:** EDMON system overview, adapted from [259], Figure 14.

### 1.4 Research problem and questions

The main research problem the dissertation addresses is: **MP - What is the effect of infection incidence on key parameters of blood glucose dynamics amongst people with type 1 diabetes, and how can self-recorded data from this patient group assist in detecting an infectious disease outbreak?** To effectively answer the main research problem, different sub-questions were derived that address part of the main research problem as follows:

- **Q1 – How is a personalized digital infectious disease detection system that collects self-recorded data from participants, analyzes the data, detects deviations on an individual basis, identifies a cluster of individuals, and notifies the status of an outbreak to be designed and implemented?**

**Description:** Pursue to develop a general framework for realizing a personalized health model-based digital infectious disease detection system based on self-collected data from people with type 1 diabetes. It gives an overview of the EDMON system components and the interaction between them. Furthermore, the framework address factors and challenges that need to be addressed during implementation both from the technology and user perspectives.

- **Q2 – What is the effect of infection incidence on key parameters of blood glucose dynamics among people with type 1 diabetes, and which parameters can effectively be used for detecting infection incidences in people with type 1 diabetes?**

**Description:** Pursue to uncover the effect of infection onset on key parameters of blood glucose dynamics. Blood glucose dynamics are affected by numerous factors and it is very important to identify if there exist any other possible confounding parameters, which could perhaps have a similar effect as infection and thereby resulting in a false alarm. This helps me to pinpoint and select potential event indicator parameters from the diabetes profile that can be effectively used in detecting health changes or deviation from normality, i.e. infection onset while reducing false alarms.

- **Q3 –What is the status regarding an infection detection system using self-recorded data from people with type 1 diabetes and to what extent do the existing personalized decision support systems, and blood glucose alarm events applications consider infection incidence and its effect while developing blood glucose anomalies detection algorithms?**

**Description:** Pursue to identify if there is any previously developed and implemented infection detection system using self-recorded data from people with type 1 diabetes and also to understand the extent to which the existing literature incorporates the effect of infection incidences in people with type 1 while modeling personalized blood glucose anomalies, i.e. hyper/hypoglycemia and glycaemic variability, detection algorithms. Further, it is also used to assess and analyses the success of machine learning models in these algorithms for achieving the required task in a real-time, and free-living condition considering different types of input variables, data collection procedures, and devices used. This helps me to figure out the direction of the dissertation in terms of approaching the infection detection tasks.

- **Q4 – How to design and develop a personalized health model that can continuously monitor individual health status and automatically detect infection onset using self-recorded data from people with type 1 diabetes?**

**Description:** Pursue to develop a personalized health model to effectively monitor the individual's health status continually and automatically detect infection onset. The idea behind this is to realize a model that is capable of detecting infection-triggered deviation while leaving changes due to other confounding parameters undetected; thereby reducing false alarms. This

is central to the proposed EDMON system performing as a case detector to screen the individual's health.

- **Q5 – *What are the user concerns and expectations towards sharing self-collected health-related data with the proposed system? and what type of data are they willing to anonymously share?***

**Description:** Pursue to assess and examine user concerns, expectations, and willingness towards sharing self-collected health-related data to the proposed EDMON system. Different factors could affect the user's willingness to share their health-related data unless they perceive that the system could meet their expectations and also address their concerns. These findings are expected to guide on factors that need to be addressed during the design and developments of the patient unit as well as the proposed EDMON system in general.

## **1.5 Significance of the study**

The dissertation is multidisciplinary in its nature and draws facts, concepts, tools, and methods from different fields of study. A personalized health model that can effectively monitor the individual's health status and automatically detect infection onset among people with type 1 diabetes is developed, which can be geared towards both individual and public health benefits. The concept behind the personalized approach can also be generalized towards infection detection in other types of diabetes and also among different population groups. These findings have practical significance in different fields; informatics, public health, medicine, and healthcare.

### **Informatics**

The significance of this dissertation to informatics is based on the models, data analysis and knowledge produced. The approach can be easily extended to people with other types of diabetes, not only type 1 diabetes.

### **Public health**

The significance of this dissertation to public health is related to the approach presented, which can be used to detect the incidence of infection at an individual level and thereby detecting infectious disease outbreaks at the population level. The EDMON system presented in the dissertation can be integrated with existing syndromic surveillance and assist in detecting infectious disease outbreaks and thereby alarming possible outbreaks that could endanger public health.

### **Healthcare**

The significance of the dissertation to healthcare is related to the personalized health model presented since it can be geared towards realizing an application for tailored personalized decision support and learning platform. In this regard, the presented analysis and models can be used to realize a tailored platform to educate individuals on how to respond to infection-induced hyperglycemia.

## **Individuals**

The findings, i.e. analysis and models, presented in this dissertation can also be informative to an individual with type 1 diabetes. Among people with type 1 diabetes, infection occurrence is known to induce significant challenge to self-manage their blood glucose levels. Therefore, information regarding blood glucose evolution, and change in insulin resistance, i.e. insulin to carbohydrate ratio, during infection incidences can be quite helpful for this group of patients.

## **1.6 Claimed contribution and included papers**

### **1.6.1 Contribution of the dissertation**

This dissertation mainly contributes towards the development of a personalized health model-based - digital infectious disease detection system, also known as Electronic Disease Surveillance and Monitoring Network (EDMON), that relied on the detection of infection onset on an individual basis to predict infectious disease outbreaks on the population level using self-recorded data from people with type 1 diabetes, which makes the system the first of its kind. EDMON incorporates a patient unit, data repository unit, infection detection unit, clustering unit, and information dissemination unit. This dissertation presents the following contributions:

- 1) A framework of personalized health model-based digital infectious disease detection system, i.e. EDMON system, for realizing a system that can collect, analyze, detect, and notify the concerned bodies about the ongoing outbreak (C<sub>1</sub>),
- 2) An infection detection unit that can effectively monitor the individual's health status continually and automatically detect infection onset (C<sub>2</sub>, C<sub>3</sub>, C<sub>4</sub>), and
  - A proof of concept towards using key parameters of blood glucose dynamics for detecting infection onset (C<sub>2</sub>). This contribution characterizes the effect of infection onset on key parameters of blood glucose dynamics and put forward optimal parameters for realizing the proposed personalized health model.
  - A personalized health model for detecting infection onset in people with type 1 diabetes using blood glucose, insulin, and carbohydrate information (C<sub>3</sub>). These models require all the three key parameters of blood glucose dynamics.
  - An alternative personalized health model for detecting infection onset in people with type 1 diabetes using only insulin and carbohydrate information (C<sub>4</sub>). The point is to provide alternative models when there is a challenge in getting blood glucose measurements for various reasons. In this regard, the underlying reason could be linked to fear associated with information leakage or data usage by 3rd parties. For example, blood glucose information containing experiences of repeated lows blood glucose level (hypoglycemia), if discovered by a licensing authority, might end up suspending one's driving license.
- 3) Finally presents a list of factors related to user concerns, expectations, and willingness for the successful mass sharing of self-collected health-related data to the proposed EDMON system (C<sub>5</sub>).

The presented system addresses the dissertation research problem and the stated contributions along with the associated research questions are presented in **Table 1**. The third research question (**Q3**) is linked with the background of the dissertation.

**Table 1:**Contributions and research questions addressed in the dissertation.

Contribution		Description	Publication
<b>C<sub>1</sub></b>	<b>Q1</b>	A general framework of a personalized health model-based digital infectious disease detection system for realizing a system that collects, analyses, detects, and notifies the concerned bodies about ongoing outbreaks.	<b>Paper 1</b>
<b>C<sub>2</sub></b>	<b>Q2</b>	A proof of concept towards using key parameters of blood glucose dynamics for detecting infection onset.	<b>Paper 1</b>
<b>BKGND</b>	<b>Q3</b>	Literature review of data-driven blood glucose pattern classification and anomaly detection mechanisms among people with type 1 diabetes.	<b>Paper 2</b>
<b>C<sub>3</sub></b>	<b>Q4</b>	A personalized health model for detecting infection onset in people with type 1 diabetes using blood glucose, insulin, and carbohydrate information.	<b>Paper 3</b>
<b>C<sub>4</sub></b>	<b>Q4</b>	An alternative personalized health model for detecting infection onset in people with type 1 diabetes using only insulin and carbohydrate information.	<b>Paper 4</b>
<b>C<sub>5</sub></b>	<b>Q5</b>	Assessments of user concerns, expectations, and willingness towards sharing self-collected health-related data.	<b>Paper 5</b>

### 1.6.2 Other contributions

Apart from the main contributions, there are other relevant contributions I produced during the work on my Ph.D. project, which is not included in the dissertations.

**Paper 7** Coucheron, S., **Woldaregay, A. Z.**, Årsand, E., Botsis, T., & Hartvigsen, G. (2019). *EDMON - A System Architecture for Real-Time Infection Monitoring and Outbreak Detection Based on Self-Recorded Data from People with Type 1 Diabetes: System Design and Prototype Implementation*. Paper presented at the SHI 2019. Proceedings of the 17th Scandinavian Conference on Health Informatics, November 12-13, 2019, Oslo, Norway.



- Paper 8** Woldaregay, A. Z., Issom, D., Henriksen, A., Marttila, H., Mikalsen, M., Pfuhl, G., Sato, K., Lovis, C., & Hartvigsen, G. (2018). *Motivational Factors for User Engagement with mHealth Apps*. *Studies in Health Technology and Informatics*, 249, 151.
- Paper 9** Yeng, P. K., Woldaregay, A. Z., Solvoll, T., & Hartvigsen, G. (2020). *Cluster Detection Mechanisms for Syndromic Surveillance Systems: Systematic Review and Framework Development*. *JMIR Public Health Surveill*, 6(2), e11512. doi:10.2196/11512.
- Paper 10** Woldaregay, A. Z., Arsand, E., Walderhaug, S., Albers, D., Mamykina, L., Botsis, T., & Hartvigsen, G. (2019). *Data-driven modeling and prediction of blood glucose dynamics: Machine learning applications in type 1 diabetes*. *Artif Intell Med*, 98, 109-134. doi:10.1016/j.artmed.2019.07.007
- Paper 11** Woldaregay, A. Z., Årsand, E., Botsis, T. & Hartvigsen, G. 2017. *An Early Infectious Disease Outbreak Detection Mechanism Based on Self-Recorded Data from People with Diabetes*. *Studies in health technology and informatics*, 245, 619-623.
- Paper 12** Yeng, P. K., Woldaregay, A. Z., & Hartvigsen, G. (2019, November). *K-CUSUM: Cluster Detection Mechanism in EDMON*. In *SHI 2019*. Proceedings of the 17th Scandinavian Conference on Health Informatics, November 12-13, 2019, Oslo, Norway (No. 161, pp. 141-147). Linköping University Electronic Press.
- Paper 13** Tejedor, M., Woldaregay, A. Z., & Godtliebsen, F. (2020). *Reinforcement learning application in diabetes blood glucose control: A systematic review*. *Artificial Intelligence in Medicine*, 104, 101836. doi:https://doi.org/10.1016/j.artmed.2020.101836.
- Paper 14** Giordanengo, A., Årsand, E., Woldaregay, A. Z., Bradway, M., Grottnland, A., Hartvigsen, G., Granja, C., Torsvik, T., & Hansen, A. H. (2019). *Design and Prestudy Assessment of a Dashboard for Presenting Self-Collected Health Data of Patients With Diabetes to Clinicians: Iterative Approach and Qualitative Case Study*. *JMIR Diabetes*, 4(3), e14002. doi:10.2196/14002.

### 1.6.3 Statements of originality

This section puts forward the statement of originality, by stating the relevance of each article to the dissertation along with my contribution.

- Paper 1** Woldaregay, A. Z., Launonen, I. K., Årsand, E., Albers, D., Holubová, A., & Hartvigsen, G. (2020). *Toward Detecting Infection Incidence in People With Type 1 Diabetes Using Self-Recorded Data (Part 1): A Novel Framework for a Personalized Digital Infectious Disease Detection System*. *J Med Internet Res*, 22(8), e18911. doi:10.2196/18911

**Relevance:** The study provides the basis for the development of the proposed personalized digital infectious disease detection system in general and the personalized health model in particular. A proof of concept regarding infection onset and its degree of effect on key parameters of blood glucose dynamics is presented. Thus, it provides information about optimal parameters of blood glucose dynamics to effectively support the strategies need for designing and developing the proposed personalized health model. Moreover, it presents a general framework to effectively guide the development of the proposed personalized health model-based digital infectious disease detection system, i.e. EDMON system. The framework provides a detailed description of the EDMON system components and challenges that needs to be addressed during system implementation. To the best of my knowledge, this is the first study that empirically and numerically quantifies the effect of infection episodes on key parameters of the blood glucose dynamics among people with type 1 diabetes exploiting self-recorded data.

**My Contribution:** I proposed the study, defined the problem, collected the data from the study participants, designed and carried out the experiments, and wrote the manuscript. G. Hartvigsen, I. Launonen, E. Årsand, A. Holubová, and D. Albers provide successive expert input to improve the study, presented some alternative approaches to shape the study, and reviewed the manuscript.

**Paper 2**      **Woldaregay, A. Z.,** Arsand, E., Botsis, T., Albers, D., Mamykina, L., & Hartvigsen, G. (2019). *Data-Driven Blood Glucose Pattern Classification and Anomalies Detection: Machine-Learning Applications in Type 1 Diabetes*. *J Med Internet Res*, 21(5), e11030. doi:10.2196/11030

**Relevance:** The study presented an overview of the state-of-the-art approaches used for classifying blood glucose patterns and detection of anomalies within the framework of personalized decision support and alarm applications for blood glucose control. The review sheds some light on the applicability of different models, challenges that need to be addressed, success and limitations of different input features and diabetes self-management technologies, cause-effect reasoning ability of the implemented system, and the extent of integrating information regarding infection incidence within the implemented systems.

**My Contribution:** I conceived the idea, proposed the study, defined the problem, designed the method, and carried out the search and analysis. G. Hartvigsen, E. Årsand, T. Botsis, D. Albers, and L. Mamykina provide successive expert input in shaping the method design and analysis. I wrote the manuscript based on successive reviews and discussions with co-authors.

**Paper 3**      **Woldaregay, A. Z.,** Launonen, I. K., Albers, D., Igual, J., Årsand, E., & Hartvigsen, G. (2020). *A Novel Approach for Continuous Health Status Monitoring and Automatic Detection of Infection Incidences in People With Type 1 Diabetes Using Machine Learning Algorithms (Part 2): A Personalized Digital Infectious Disease Detection Mechanism*. *J Med Internet Res*, 22(8), e18912. doi:10.2196/18912

**Relevance:** The study provides an approach for detecting infection onset in an individual with type 1 diabetes exploiting self-recorded blood glucose level, insulin, and carbohydrate information. A personalized health model is presented, which can track an individual's health

status in a continuous manner and automatically detects infection onset. The aim of the study was to thoroughly evaluate the performance, necessary sample sizes, and computational time of different machine learning algorithms. To the best of my knowledge, this is the first attempt towards realizing a personalized health model to capture infection episodes among people with type 1 diabetes using self-recorded data.

**My Contribution:** I conceived the idea and proposed the study, designed and performed the experiments. G. Hartvigsen, I. Launonen, E. Årsand, and D. Albers provides successive expert input during the designing of the experiment and proposed some alternative solutions. I wrote the manuscript based on a thorough discussion with co-authors. Further, the co-authors reviewed the manuscript and provide successive comments and suggestions to improve the manuscript.

**Paper 4**      **Woldaregay, A. Z.** *Woldaregay, A. Z. Automatic Detection of Infection State in Individuals with Type 1 Diabetes Under Free-Living Conditions Using Using Self-Recorded Insulin and Carbohydrate Information (Part 3). EARLY DRAFT MANUSCRIPT.* This Manuscript contains overlapping texts and results from the dissertation.

**Relevance:** As an alternative solution, this study presents a personalized health model that exploits self-recorded insulin and carbohydrate information. This alternative provides a greater advantage when there is a lack of information about blood glucose levels. Hence, the study compares with the performance achieved in *Paper 3*. The result demonstrates the potential of the proposed approach.

**My Contribution:** I conceived the study, designed and performed the experiments, and wrote the manuscripts. In this study, the ideal was to further strengthen the result presented in *Paper 3* and to provide alternative ways of achieving the same objective.

**Paper 5**      **Woldaregay, A. Z.,** Henriksen, A., Issom, D. Z., Pfuhl, G., Sato, K., Richard, A., Lovis, C., Årsand, E., Rochat, J., & Hartvigsen, G. (2020). *User Expectations and Willingness to Share Self-Collected Health Data.* *Stud Health Technol Inform, 270,* 894-898. doi:10.3233/SHTI200290

**Relevance:** The study presented a list of findings related to user concerns, expectations, and willingness towards sharing one's self-collected health-related data to the proposed personalized health model-based digital infectious disease detection system, i.e. EDMON system. Given the sensitivity of health data, prior understanding of user concerns and expectations are very important especially when it comes to the sharing of health-related data. To facilitate mass data sharing, it is necessary to properly address user concerns and meet user expectations throughout the EDNON system design and development. In this regard, the study puts forward factors that need to be addressed during system design and implementation so as to buy the user's willingness to engage in sharing data to the proposed system. A survey questionnaire was developed based on the results from a qualitative interview [252] to further understand factors related to user concerns, expectations, and willingness. A survey questionnaire assessing different factors related to user concerns, willingness, and expectations

was distributed to participants and the response was analyzed using statistical techniques. These findings will act as guidelines during the realization of the proposed EDMON system.

**My Contribution:** I, together with the co-authors, has conceived the idea, proposed the study and formulated the problem. Further, contributed through the development of the survey questioners, and posted the survey questionnaire related to diabetes in an appropriate channel in collaboration with E. Årsand. G. Pfuhl conducted the statistical analysis. A. Henriksen and I wrote the manuscript based on a thorough discussion with the co-authors.

#### **1.6.4 Defining the red thread**

This section describes each stage of the dissertation and the connection between them.

***The Initial Phase (Q1) - Development of a general framework of a personalized health model-based digital infectious disease detection system.***

As an initial work, the general framework defines the main components of the proposed personalized health model-based digital infectious disease detection system, i.e. EDMON system, along with the challenge that needs to be addressed during system design and implementation **Paper 1** [259]. The presented framework shapes the dissertation by pinpointing the crucial part of the EDMON system. All the rest of the dissertation topics emanates from the proposed framework.

***The First Phase (Q2) – Nature, extent, and degree of effect of infection incidences on key parameters of blood glucose dynamics amongst people with type 1 diabetes.***

There is evidence in the literature describing the effect of infection incidences in people with type 1 diabetes. Some literature clearly described the potential of blood glucose levels and other diabetes profile parameters as an indicator of infection onset in this patient group. However, there is a knowledge gap when it comes to empirical and numerical evidence depicting the degree of deviation as a result of infection with respect to each diabetes profile. This work uncovers the nature and extent of infection-induced deviations on each key parameter of blood glucose dynamics and puts forward optimal parameters for realizing a personalized health model **Paper 1** [259]. As a result, this piece of work layout the foundation for the development of a personalized health model for detecting infection incidences.

***The Second Phase (Q3) – Personalized decision support systems, blood glucose alarm events applications, and anomalies detection algorithms.***

As a roadmap to the development of a personalized health model, the systematic review is aimed at assessing and analyzing the state-of-the-art approaches used for classifying blood glucose patterns and detecting anomalies within the framework of personalized decision support and alarm applications for blood glucose control. It also tries to assess and understand if there is any literature, which incorporates information related to infection onset and how these algorithms tried to include such information in their process **Paper 2** [253]. This wide piece of work helps me to choose the right modeling approach towards realizing the proposed personalized health model for detecting infection incidences on an individual basis.

### ***The Third Phase (Q4) – Modelling a personalized infection detection algorithm.***

After understanding the state-of-the-art approaches and the associated knowledge gap, this work realizes a personalized health model that can monitor the individual health status in a continuous manner and automatically detects infection onset using the individual's diabetes profile. This piece of work is the core for realizing the proposed personalized health model-based digital infectious disease detection system, i.e. EDMON system.

To begin with, modeling of a personalized health model was carried out using blood glucose levels, insulin, and carbohydrate information as input features to the algorithm *Paper 3* [258]. Furthermore, the possibility of using insulin and carbohydrate information as input features for realizing a personalized health model was explored in *Paper 4*.

### ***The Fourth Phase (Q5) – Assessments of user concerns, expectations, and willingness towards sharing self-collected health-related data.***

As the last stage, this piece of work tries to pinpoint user concerns, expectations, and willingness towards sharing self-collected health-related data to the proposed EDMON system. A system that relies on self-collected data typically requires the timely sharing of these data for further processing to produce precise results. Fulfilling patient expectations, and properly addressing their concerns has a great impact on the successful mass sharing of health-related data to the proposed EDMON system *Paper 5* [257].

## **1.7 Organization of the dissertation**

The entire dissertation is organized as follows; *Chapter two* provides the background of the dissertation; *Chapter three* presents the methods and materials used in the studies; *Chapter four* introduces the proposed framework of the personalized health model-based digital infectious disease detection system along with the associated challenge that needs to be addressed during system design and implementation; *Chapter five* gives a detailed analysis depicting the nature and extent of infection-induced deviations on each key parameter of blood glucose dynamics among people with type 1 diabetes; *Chapter six* presents solutions for a personalized health model that can monitor the individual's health status in a continuous manner and automatically detect infection onset in people with type 1 diabetes; *Chapter seven* puts forward list of findings related to user concerns, expectations, and willingness for long-term and successful sharing of self-collected health-related data. The final chapter, *Chapter Eight*, gives a concluding remark of the dissertation and future works.

This dissertation has a content reuse, and specifically from Paper 1, 2, 3, and 5 and are allowed by the publishers under the Creative Commons Attribution license. In this regard, given the absence of written guidelines from UiT – The Arctic University of Norway, a guideline from the Technical University of Denmark are used, and can be found from <https://www.dtu.dk/english/-/media/DTUdk/Uddannelse/PhD-Udannelse/Dokumenter/DTU-Guidelines-for-avoiding-plagiarism-and-self-plagiarism-in-PhD-thesis-writing.ashx?la=da&hash=3D7503AB200968C246A901BD90C4E0C9A04E3940>. Based on this, a description is provided in the introduction section of each chapters indicating the source paper.



## 2 Theoretical Background

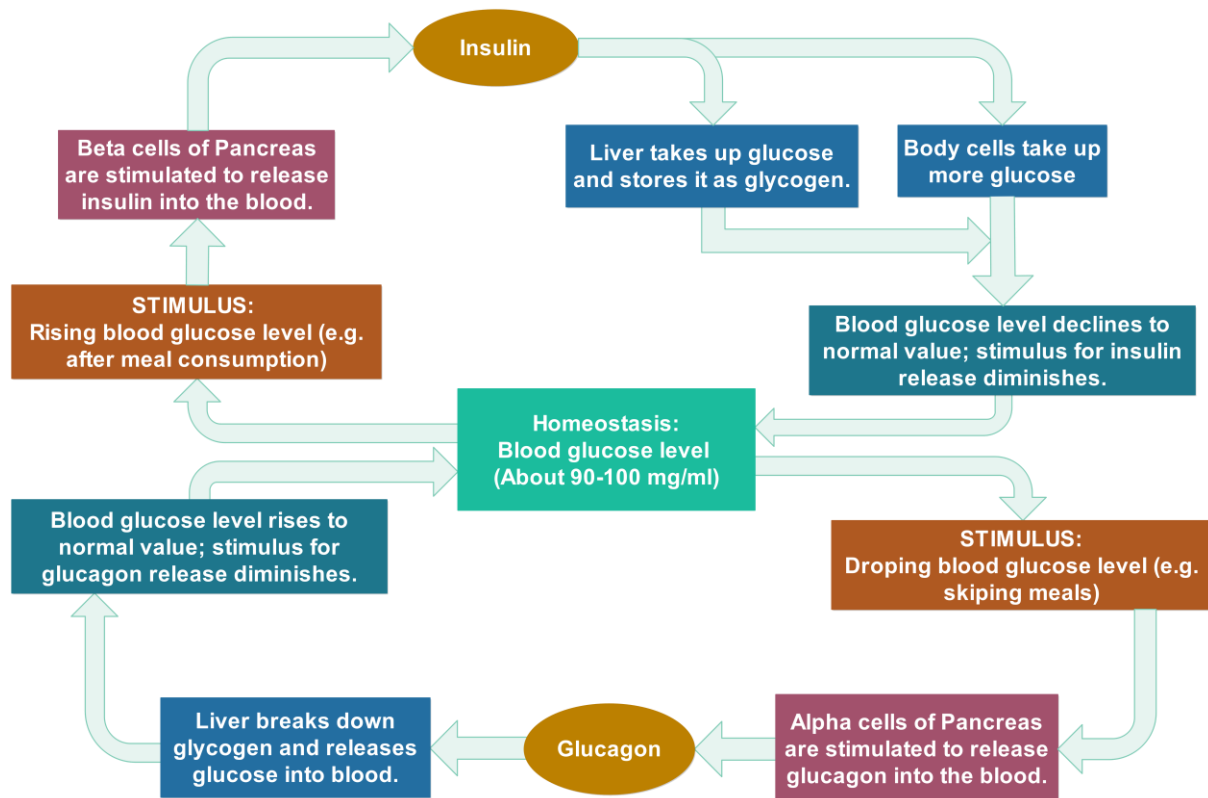
*Synopsis: This chapter puts forward detailed descriptions of the theoretical background of the dissertation. Even if each article has a detailed and broader background section, this part will complement and highlight some theoretical background necessary to understand the work presented in this dissertation. The first half of the chapter provides insight into blood glucose homeostasis and infection incidences, and then proceed to describe existing literature suggesting diabetes profile parameters as infection indicator.*

### 2.1 Blood glucose dynamics

Blood glucose homeostasis is the process by which the body maintains normal blood glucose levels [184]. The pancreas' alpha and beta cells play a vital role in blood glucose homeostasis by secreting insulin and glucagon, which are necessary to regulate blood glucose levels. As shown in **Figure 2** [190], the occurrence of either elevated or lowered blood glucose concentration stimulates the pancreas beta and alpha cells to release insulin and glucagon respectively. For instance, let us consider that there is elevated blood glucose concentration after a meal, and this scenario directly stimulates pancreatic beta cells to release insulin, thereby activating the liver to take up glucose, convert and store it as glycogen through a process known as glycogenesis. In addition, at the same time, the body will be activated to take up and use more glucose. As a result, glucose concentration in the blood lowers to the normal range, and stimulus for insulin release also diminishes, as shown in **Figure 2**. By the same token, let us assume that an individual has skipped a meal and his/her blood glucose levels begin to drop below a certain threshold, and this scenario stimulates pancreatic alpha cells to release glucagon into the blood. Hence, glucagon activates the liver to break down glycogen into glucose through the process called glycogenolysis and release it into the blood. As a result, blood glucose levels rise to the normal range and thereby diminishing the stimulus for glucagon production, as shown in **Figure 2**.

Regarding people with type 1 diabetes, there is a lack of insulin secretion from pancreatic beta cells. This patient group needs to inject a proper amount of exogenous insulin as medication to compensate for the lack of insulin secretion. Mostly, the key diabetes parameter such as the current blood glucose levels, amount of carbohydrate consumption, and physical activity session or exercise load are taken into account to determine the necessary amount of insulin injection. Insulin injection is administered based on either multiple daily injections, which involves several small injections at different times of the day, or continuous subcutaneous insulin infusion, which relies on insulin pump therapy. People with type 1 diabetes usually rely on carbohydrate counting, which involves estimating insulin-to-carbohydrate ratios [210], to accurately determine the necessary amount of insulin during mealtime [84; 167; 210; 228]. The estimated insulin-to-carbohydrate ratio is a function of an individual's insulin sensitivity and factors that affect insulin sensitivity directly or indirectly also affect the ratio. These factors include the time of the day [167], weight, age, menstruation, physical activity, stress, illness, infection, and so many others [48]. During normal situations, the estimated value of the insulin-to-carbohydrate ratio in most cases is normally distributed within the usual range of 0.02, equivalent to 1 unit of insulin to every 50 grams of carbohydrate, to 0.2, equivalent to 1 unit of

insulin to every 5 grams of carbohydrate [158; 167]. However, it should be noted that this normal value, for some individuals could be higher based on their health conditions.

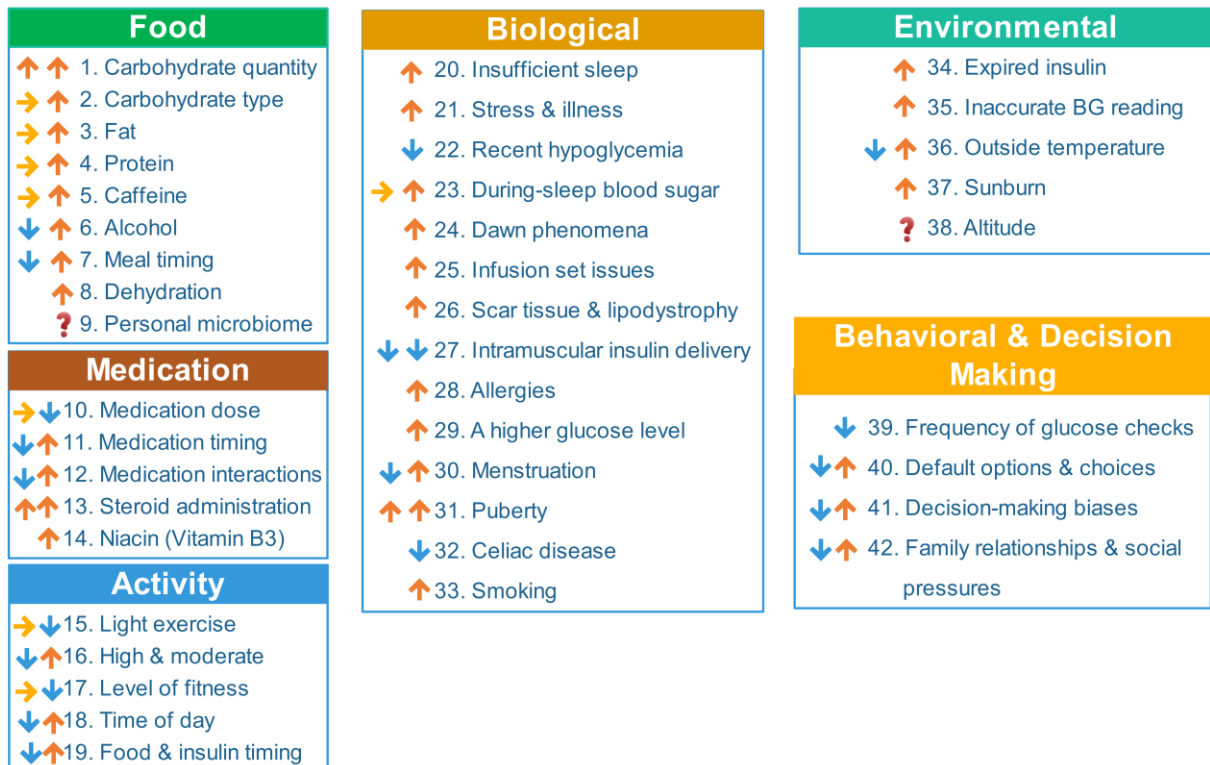


**Figure 2:** Blood glucose homeostasis (regulation) [10; 208], source *Campbell*, Figure 25.10.

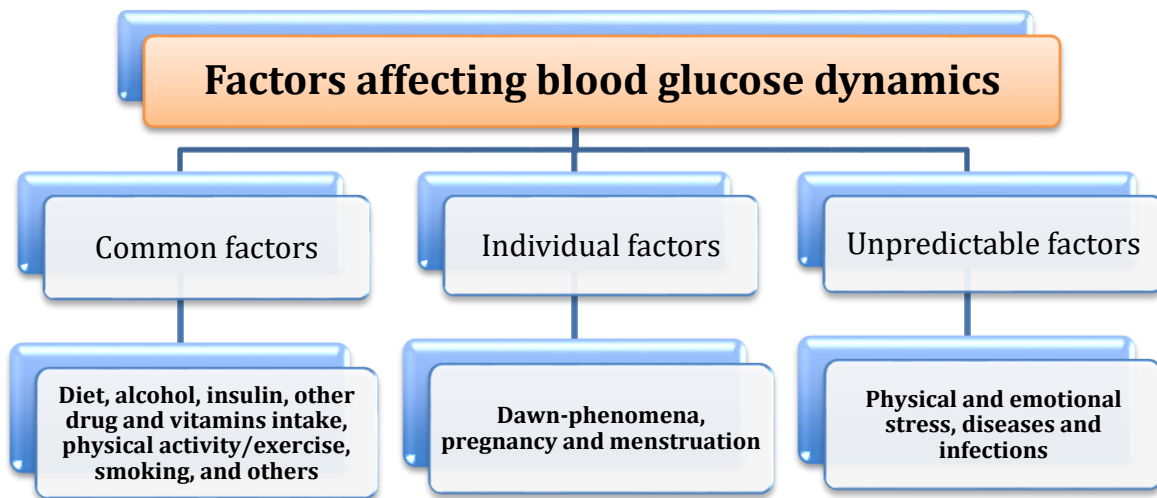
Blood glucose concentration is affected by different factors, which are often designated as common, individual, and unpredictable factors [22; 76], as shown in **Figure 3**, [22].



## 42 Factors That Affect Blood Glucose Levels



a) Some list of factors that affect blood glucose levels, source – diaTribe<sup>1</sup> [43].



b) Grouping of factors affecting blood glucose levels [22].

**Figure 3:** Factors affecting blood glucose dynamics. Figure (a) depicts some of the factors that affect blood glucose levels and Figure (b) depicts a grouping of these factors amongst the individuals.

<sup>1</sup><https://diatribe.org/42factors>

These factors are also grouped into either patient controllable or patient uncontrollable parameters depending on the degree of direct control the patient has to limit the effect of these factors on blood glucose dynamics [253]. From the patient perspective, the patient controllable parameter is known to induce reasonable deviation on blood glucose dynamics. For instance, blood glucose deviation due to meal intake can be approximated and known in advance. However, the effect of patient uncontrollable parameters is difficult to estimate in advance and known to induce unreasonable deviations that usually differs from the typical norm of blood glucose dynamics. For example, blood glucose deviations as a result of infection onset are difficult to manage and often could end up in prolonged hyperglycemia. However, from a blood glucose fluctuations perspective, any incurring blood glucose deviations can be defined as either normal or special cause variations [253]. In this regard, normal cause variations are those induced by common and individual factors. Yet, special cause variations are those induced by unpredictable factors, and patients usually struggle to manage and understand the underlying cause during the incidences. A particular example of special cause variation is depicted during an infection episode, where the individual patient struggles to control the blood glucose levels through frequent administration of insulin [37].

## **2.2 Infections incidences in people with diabetes**

### **2.2.1 Acute illness and glucose metabolism**

During infection episodes, several hormonal changes occur within the body often known as counter-regulatory, which affects blood glucose metabolism. The counter-regulatory hormones, e.g. epinephrine and cortisol, mainly stimulate the production of hepatic glucose, glucose utilization, and lipolysis and thereby elevating the plasma blood glucose concentrations (hyperglycemia) [133; 147; 264].

### **2.2.2 Blood glucose control during infection incidences**

During infection incidences, self-management of diabetes can become very problematic, given the interference of the counter-regulatory hormones, which affects the individual's blood glucose levels and insulin sensitivity. Excess glucose productions and the influence of insulin resistance (less insulin sensitivity) on the blood glucose dynamics are believed to be responsible for creating a shift in the operating point mainly in terms of the insulin-to-carbohydrate ratio [30; 38; 141; 142; 158; 179; 248; 264]. As a result, it is common to observe an abnormal rise in blood glucose levels and different reactions to insulin injection and meal intake. Managing hyperglycemia is critical during infections, where poor self-management can lead to complications including diabetic ketoacidosis and other similar incidences [145]. Generally, it is recommended to frequently measure blood glucose levels, and inject more insulin despite experiencing a feeding problem (loss of appetite) [121; 142; 144; 145]. The optimal management of hyperglycemia episodes during infection incidences needs to balance the amount of insulin and carbohydrate intake, considering a 10 - 40 % raise in the total daily insulin and reduction in the amount of carbohydrate consumption (light foods are recommended) [121; 141; 142; 144; 145]. A typical example of patients discussing the challenges related to infection is illustrated in **Figures 4** below:



Type 2 · Well-Known Member

Messages: 385  
Likes Received: 2,759  
Trophy Points: 158

My BG is between 5-7 normally, when I have any sort of infection my BG runs into double figures 9-13 last week it frustrating but realistically there nothing that I can do, I chatted to my GP last week, it maybe that as a coping strategy I take Glimperide when i have a infection or virus to bring down the number, I ditched my diabetic meds back in November in favour of a LCHF diet, which has been great, my only problem seems to be raised BG in time of illness.

📌 Informative x 2

4y 44w ago

#10



Type 2 (in remission!) · Moderator

Staff Member

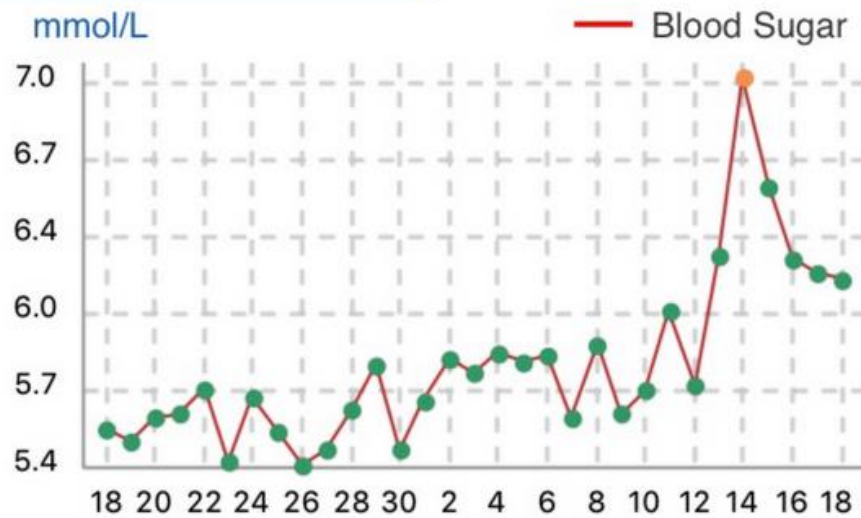
Messages: 11,246  
Likes Received: 14,400  
Trophy Points: 298

I had a flu jab but have been suffering from a chest infection for about a week now, you can clearly see what this has done to my daily average readings 😞


< From 11/18/17 To 12/18/17 >

Blood Sugar

Weight



📌 Informative x 3 🤗 Hug x 1



**Type 2 · Well-Known Member**


Messages: 2,912  
Likes Received: 1,946  
Trophy Points: 198

I have had a really sore throat for the last two days and now my nose is starting to run just tested my BG and it is 9.3 usually I am in the 5/6's so yes it does go up when you are unwell

✔ Agree x 2

4y 44w ago
#17

a) People with type 2 diabetes discussing the effect of infection on blood glucose dynamics.



**Type 1 · Active Member**

Messages: 40  
Likes Received: 33  
Trophy Points: 58

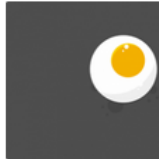
Whenever I have flu my bloody sugar levels go through the roof instantly and I find it very hard to bring them down. Recently though I've been throwing up quite a lot which I put down to just a bug as when I did check my blood sugars they were at 5.3 and rarely changed when I kept checking them every two hours

👍 Like x 4

4y 44w ago
#19

## Flu and high BS

Discussion in 'Ask A Question' started by K · Mar 2, 2018.



**Type 1 · Member**


Messages: 11  
Likes Received: 5  
Trophy Points: 43

Hi,  
I'm a type 1 diabetic and have been diabetic for over 30 years; but thankfully am rarely ill! But I've managed to pick up one of the flu viruses this week. My blood sugars are usually reasonable; I test a lot and also carb count and adjust dosage. But I'm really struggling to bring down the BS; its sitting at between 13-16 and whenever I do a bolus to adjust it barely moves. This means I'm now not actually eating very much as am not finding a window when the levels are sufficiently low. I have increased by lantus (twice daily) by about 20% but little seems to be helping. I understand this is because cortisol released to help recovery also makes the body more insulin resistant. I also know that it will pass. But what I don't know is when I should seek external help; I'm not vomiting or showing any of the other ketoacidosis symptoms but at what point do I worry? How resilient is the body? With all the posts on here saying people are constantly at 5 or 6 I'm feeling like a naughty diabetic.

K

🤗 Hug x 2

2y 42w ago
#1



**· Guest**

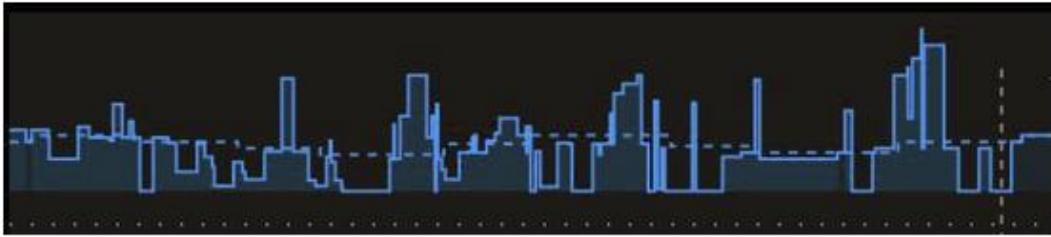
As you say, with a high BG, your body is likely to be insulin resistant.  
Have you tried taking a higher correction bolus?  
If my BG is that high, I need twice as much insulin to correct.  
I also find, when I am ill, I may need 30 or 40% extra basal.

I don't think you are a naughty diabetic and whilst ever high increases your risk of complications, long term high usually refers to weeks rather than a few hours or day.

Focus on getting better. Drink plenty of fluids, rest, test, correct.

👍 Like x 1

2y 42w ago
#2



Showing the net basal adjustments made on day 2 of my norovirus – the dotted line is what my basals usually are, so anything higher than that dotted line is a “high” temp and anything lower is a “low” temp of various sorts.

- When I first started throwing up over the first 8 hours, as is pretty normal for norovirus, I first worried about going low, because obviously my stomach was empty.

**Nope.** I never went lower than about 85 mg/dl. Even when I didn't eat at all for > 24 hours and very little over the course of 5 days.

- b) People with type 1 diabetes discussing the effect of infection episode on blood glucose dynamics.

**Figure 4:** Depicts what patients discuss on online forums about infection and blood glucose levels, source [78; 141; 142]. Figure (a) depicts type 2 diabetes and Figure (b) depicts type 1 diabetes. User's identifiers are removed for privacy purposes.

### 2.2.3 The literature on using diabetes profile for infection detection

As a roadmap to better understand the project, I tried to pinpoint the state-of-the-art studies. In this regard, using parameters from diabetes profile as an indicator of infection incidences have been studied in the literature, as shown in **Table 2** below. The literature ranges from proof of concept study to system description. From the diabetes profile, almost all of the literature considers *blood glucose levels* as indicators of infection episodes. Some other studies also suggested *hba1c* and *white blood cell count* as an alternative indicator of infection episodes, but these parameters have significant availability and feasibility issues given the fact that these parameters are difficult to get the records on a daily basis with the least possible cost. As described above, there are already efforts to prove the relationship between elevated blood glucose levels and infection episodes. However, there is a lack of empirical and numerical evidence describing the nature, extent, and duration of infection-induced deviations on each individual's diabetes profile such as blood glucose levels, insulin intake, and diet consumptions.

**Table 2:** Using diabetes profile for detecting infection incidences among people with diabetes.

Reference	Type	Data Type	Method	Findings
[29; 38]	Proof of concept	Infection evidence and daily glycemic control data of 248 type-2 diabetics.	Logistic regression analysis	Elevated blood glucose levels during the infection episodes
[31; 33; 34; 36]	Analysis and system specification	Diabetes Control and Complications Trial (type-1 diabetes) Informatics for Diabetes Education and Telemedicine (type-2 diabetes).	t-test statistics	After infection, HBA1c values were elevated despite tight blood glucose control
[17]	System description	User ID, blood glucose level, and geographical location	----	----
[90]	System prototype	Blood glucose level	----	----
[135]	System description	Blood glucose levels, illness, and symptom report	----	----
[30]	System Specification	Blood glucose, white blood cell count, and other sources	----	----

## 2.3 Chapter summary

This chapter put forward and discussed four important topics; blood glucose homeostasis, the effect of infection on blood glucose dynamics, blood glucose control during infection incidences, and available evidence suggesting the feasibility of using blood glucose levels and other diabetes profiles as an indicator of infection episode. The connection between prolonged elevated blood glucose levels (hyperglycemia) and infection episode has been known for a long time, however, there are few previous proofs of concept studies that verify the potential of diabetes profile for detecting infection episode. Moreover, there is little effort to capitalize this information towards detecting infection incidences on an individual basis; thereby detecting infectious disease outbreaks on a large scale.

## 3 Materials and Methods

*Synopsis: This chapter puts forward detailed descriptions of the materials and methods used in different sub-phases of the dissertation. It describes the type of data source exploited, approaches used to exploit and clean the dataset, methods used to analyze the effect of infection episode on blood glucose dynamics, approaches used to design a personalized health model, and finally approaches used to identify major factors that can enhance user motivation towards successful engagement and sharing of health-related data. Finer elements of the methodology are also presented in each associated chapters.*

### 3.1 A general overview of research approaches

This dissertation is multidisciplinary in nature and draws diverse concepts, tools, and methods from various disciplines including informatics, public health, and medical research. A multidisciplinary approach has been proved to be effective in addressing complex topics [65; 162; 191; 195]. In this regard, the dissertation uses concepts, tools, and methods from the aforementioned disciplines to develop new methods and approaches that account for a better understanding of infection incidences and their effect on blood glucose levels amongst people with type 1 diabetes and also for realizing a novel personalized health model-based digital infectious disease detection system, which could provide benefits to both individual and the general population.

#### 3.1.1 Overview of approaches used in different phases of the study

In this dissertation, four different methods were used, which are in agreement with the standard scientific practice of attaining new knowledge.

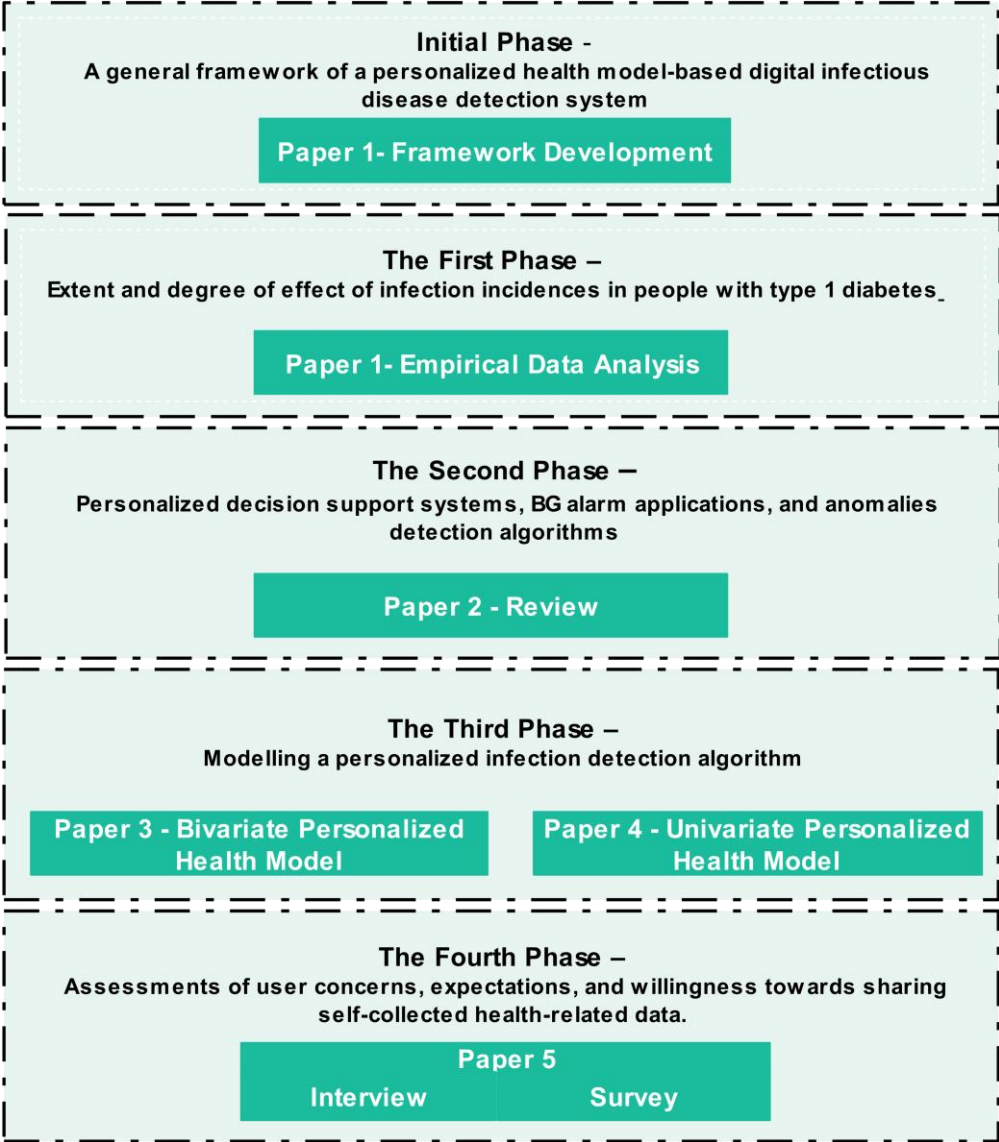
- a) Systematic Review
- b) Retrospective, Numerical and Empirical Data Analysis
- c) Controlled Experiments
- d) Qualitative interview, Quantitative Survey, and Statistical Analysis

The first method, systematic review, is very essential to identify the existing body of knowledge and associated knowledge gaps in the existing literature. The second method, numerical and empirical data analysis, is extremely valuable to assess, analyze, and interpret information from real-world data to draw a valid conclusion. The third method, controlled experiment, is useful to design experiments for the development of machine learning models. The fourth method, qualitative interview, and quantitative survey is used to study and assess factors related to user concerns, expectations, and willingness towards sharing self-collected health-related data to the proposed EDNON system and further explain it through collection and analysis of data.

### 3.2 Overview of the dissertation stage

This dissertation, as shown in **Figure 5**, is consisted of five sub-phases each dealing with a specific sub-question of the original research problem. It should be noted that each sub-phase

is a continuation of the previous sub-phase, and however, some of these sub-phases were conducted in a concurrent manner.



**Figure 5:** Progressive stage of the dissertation.

***The Initial Phase (Q1) - Development of a general framework of a personalized health model-based digital infectious disease detection system***

*Attached with Paper 1*

As an initial phase, the formulation of a general framework was carried out to identify and pinpoint the main components of the proposed personalized health model-based digital infectious disease detection system, i.e. EDMON system. The framework presents and



describes the system architecture, requirements, necessary equipment, communication platform, and other various system components along with the challenge that needs to be addressed during system design and implementation. The proposed framework plays a crucial role as a roadmap to shape the dissertation by pinpointing the crucial part of the system. All the rest of the dissertation topics emanates from the proposed framework.

***The First Phase (Q2) – Nature, extent, and degree of effect of infection incidences on key parameters of blood glucose dynamics amongst people with type 1 diabetes***

*Attached with Paper 1*

After formulating the general framework, the first phase of the dissertation tries to uncover the degree of effect of infection episode on blood glucose dynamics amongst people with type 1 diabetes. Among other things, the study also identifies optimal key parameters of blood glucose dynamics that can successfully be used as input features for the realization of the proposed personalized health model. The study retrospectively analyzed empirical data, i.e. self-collected health data, containing self-reported states of infection episodes. As a result, this piece of work layout the foundation for the development of a personalized health model for detecting infection onset.

***The Second Phase (Q3) – Personalized decision support systems, BG alarm applications, and anomalies detection algorithms***

*Attached with Paper 2*

As a roadmap to the development of a personalized health model, the second phases of the dissertation involves performing a systematic review to assess and analyses the state-of-the-art methods and approaches that focused on realizing an algorithm that performs classification of blood glucose pattern and detection of anomalies within the framework of personalized decision support systems, and blood glucose alarm applications. Among other things, the review also looked at two things; 1) do there exist any previously developed infection detection systems for people with type 1 diabetes? and 2) to what extent do the existing prediction and anomalies detection algorithm incorporates information on infection episode as one of their input parameters? This wide piece of work provides an overview to carefully select the right modeling approach towards realizing the proposed personalized health model.

***The Third Phase (Q4) – Modeling a personalized infection detection algorithm***

*Attached with Paper 3 & 4*

After understanding the state-of-the-art approaches and the associated knowledge gap, the third phase of the dissertation realizes a personalized health model using an individual's self-recorded data (diabetes profile). The model is capable of monitoring an individual's health status in a continuous manner and also can detect infection episodes (health deviation)

automatically. This piece of work is the core for realizing the proposed personalized health model-based digital infectious disease detection system, i.e. EDMON system.

There are two associated papers with this phase; *Paper 3* and *Paper 4*. *Paper 3* presents a personalized health model using bivariate input - *blood glucose levels, and the ratio of insulin-carbohydrate information* as input features to the machine learning algorithm. *Paper 4* provides an alternative implementation but using only one input (univariate) – the *ratio of insulin to carbohydrate information* as an input feature to the algorithm.

### ***The Fourth Phase (Q5) – Assessments of user concerns, expectations, and willingness towards sharing self-collected health-related data.***

*Attached with Paper 5*

As the last stage, this piece of work tries to pinpoint user concerns, expectations, and willingness towards sharing self-collected health-related data to the proposed EDMON system. A real-time system like EDMON that relies on self-collected data typically requires the timely sharing of these data for further processing to produce precise results. In this regard, the study assesses and identifies a list of factors related to user concerns, expectations, and willingness that could impact successful mass sharing of self-collected health-related data to the proposed EDMON system. It is obvious that people face challenges to join a new system for various reasons including individual views and attitudes towards data sharing and perceived benefits, and risks of data sharing, trust issues, fear of privacy, and data breaches, motivations, and a lot of other factors. Hence, fulfilling users' expectations, and properly addressing their concerns can have a great impact on the successful mass sharing of health-related data to the EDMON system.

## **3.3 Empirical data analysis and modeling**

The objective of the empirical data analysis was to study and uncover the effect of infection triggered deviations on the individual's blood glucose dynamics, and, in particular, to identify and select optimal parameters indicative of infection onset to be used for the design and development of the proposed personalized health model. The objective of the modeling part is to devise a personalized computational health model using those previously identified optimal parameters for tracking the individual's health status and automatically detect when the individual becomes sick.

### **3.3.1 Definition of key terms**

The description of terms used throughout the empirical data analysis and modeling are given in **Table 3**. These terms are crucial in understanding the whole design and development of the personalized health model. These terms are described in detail in the theoretical background chapter and presented again here in a more compact and understandable way.

**Table 3:** Key terms used in empirical data analysis and modeling.

Key term	Definition
Blood glucose dynamics	Describes the complex relationship depicting the effect of patient controllable and uncontrollable parameters on blood glucose levels.
Patient controllable parameters	Describes the category of parameters that incorporates factors on which the patient has direct control and can roughly understand their immediate effect on the BG dynamics. For example, diet and physical activity sessions are categorized under this.
Patient uncontrollable parameters	Describes the category of parameters that incorporates factors on which the patient doesn't have direct control and face challenge to understand their immediate effect on the blood glucose levels. For instance, parameters such as emotional stress and illness are categorized under this.
Normal cause deviation	Deviation incurred due to patient controllable factors such as meal-induced hyperglycemia.
Special cause deviation	Deviation incurred due to patient uncontrollable factors including infection episode.
Operating point of blood glucose dynamics	Describes the state of blood glucose level at any time $t$ in response to patient controllable and uncontrollable parameters.

### 3.3.2 Datasets and data sources

The empirical data analysis and modeling were carried out using self-collected health-related data from people with type 1 diabetes. These phases of the dissertation typically require an accurate, rich, and large dataset to validate the concept and further develop the proposed personalized health model. However, getting such kind of dataset poses significant challenges especially when it comes to obtaining data from a large number of participants due to a variety of reasons, i.e. expensive and time-consuming. The dataset used in this dissertation was contributed by three individuals (2 males and 1 female) and incorporates highly precise longitudinal records of 10 patient-years. These participants were proven individuals with high knowledge of their conditions and different diabetes technologies. As shown in **Tables 4 & 6**, the dataset consists of blood glucose reading, insulin injection, ingested carbohydrate, and episodes of acute infection reported by the individuals. Throughout the data collection period, the participants have indicated that they were free of illness or other health complications except for the reported date of infections. As shown in **Table 4**, various diabetes technologies and self-management devices were used throughout the data collection period, such as Dexcom continuous glucose monitor (CGM), Diabetes Diary [16], the xDrip with an app, the Spike app [212], insulin pens, and pumps. From the ten patient-years, the first five patient-years were reported to be regular years and the rest five patient-years were reported to contain at least one or more episodes of acute infections. The reported infection episodes were influenza (flu), mild and light common cold without fever. The study protocol has been reviewed by the Norwegian Regional Committees for Medical Health Research Ethics Northern Norway (REK) (Reference number: 108435). Written consents have been obtained and the participants have donated the datasets. All the data from the participants were anonymized.

**Table 4:** Type of datasets and diabetes technologies used for self-management, source [259], Table 1.

Self-Management			
Patient	BG	Insulin administration	Diet
Subject 1	Self-monitoring blood glucose (SMBG) - Finger pricks recorded in Diabetes Diary mobile app and Dexcom CGM	Insulin Pen ( Multiple boluses (NovoRapid) and one time basal (Lantus) before bed) recorded in Diabetes Diary mobile app	Carbohydrate in grams recorded in Diabetes Diary mobile app
Subject 2	SMBG - Finger pricks recorded in Spike mobile app and Dexcom G4 CGM	Insulin Pen ( Multiple boluses (Humalog) and one-time basal (Toujeo) before bed) recorded in Spike mobile app	Carbohydrate in grams recorded in Spike mobile app
Subject 3	Enlite (Medtronic) CGM and Dexcom G4 CGM	Medtronic MinMed G640 insulin pump (Basal rates profile (Fiasp) and multiple boluses (Fiasp))	Carbohydrate in grams recorded in pump information

During data collection, for trustworthiness, specific and strict inclusion criteria were used to select the dataset included in the study to make valid conclusions. The criteria used in data selection include data reliability, accuracy, completeness, and longer period, i.e. at least more than 3 months. It is obvious that a study that completely relies on user collected data requires a dataset, which is reliable, complete, accurate, and collected over a longer period to produce credible evidence. In this regard, I only include data collected from highly proven and motivated individuals who have advanced knowledge and understanding of the disease condition and its self-management practices including several diabetes-related technologies. For instance, one of the most important issues is carbohydrate counting, which is prone to errors in most cases. In this regard, I make sure that the included participants have long experiences and advanced knowledge of carbohydrate counting, which is referred to as level 3 (advanced) [84]. **Table 5** below provides the detailed characteristics of the participants.

**Table 5:** Detailed description of participant characteristics, source [259], Table 2.

Variables	
Gender	2 males and 1 female
Age	$34 \pm 13.2$ years
Bodyweight	Subject 1 (83 kg), Subject 2 (77kg), Subject 3 (70kg)
HbA1c	Subject 1 (6.0%), Subject 2 (7.3%), Subject 3 (6.2%)
Carbohydrate counting	Level 3 (advanced)

### 3.3.3 Data re-sampling, imputation and preprocessing

As shown in **Table 6**, the self-collected dataset utilized in this dissertation incorporates key parameters of blood glucose dynamics and represented as follows; continuous glucose reading ( $U^{cg}$ ), blood glucose reading ( $U^g$ ), injected bolus insulin ( $U_f^I$ ), injected basal insulin ( $U_s^I$ ), and ingested carbohydrate ( $U^C$ ). Re-sampling of the raw data was carried out at a uniform rate by assigning the individual measurement to its nearest time-bin [69]. For instance, assume a time-bin  $T_b$ , which represent a time interval between 2:00pm-3:00pm of a particular day. Hence, any measurements registered between 2:00pm and 3:00pm are assigned to  $T_b$ . In some cases, the total number of measurements in a specific time-bin could be more than one element and in these cases, either the sum or average of the elements are computed to combine the elements into a single element. In this regard, for a particular time-bin, elements of blood glucose measurements, i.e. SMBG and CGM, are averaged into a single element. However, the sum of elements is used to combine insulin and carbohydrate measurements into a single element in their respective time-bin, as shown in **Table 7**. Within each time-bin, the average blood glucose response to the total amount of carbohydrate consumption and insulin administration is considered.

Generally, data smoothing was performed using a moving average filter after the data is re-sampled as described above. The idea behind the smoothing was to capture only the essential features-long term variations and eliminate features with fast-time scale and short term variations. Selecting the optimal or appropriate window size is very challenging given the complementary issues of better smoothing and the cost associated with a significant delay (shift introduced) [18; 182]. Generally, selecting the proper window size is application dependent. In this regard, for an application that sought after early detection of infection incidences, it is more important to give more emphasis to reducing the inherent delay while selecting the optimal window size.

As shown in **Table 7**, to compute the required insulin-to-carbohydrate ratio, it is necessary to compute the values of the total insulin and carbohydrate for each respective time-bin, and then for each time-bin, divide the computed insulin values with carbohydrate values. In this

dissertation, two time-bins of different lengths were considered; hourly and daily, which defines the scope of the data granularity investigated. For instance, an hourly time-bin signifies that a change in blood glucose dynamics will be tracked at the end of each hour throughout the day. It is obvious that increasing the granularity of the data, i.e. reducing the time-bin to an hourly level, can provide detailed information and helps to look at the data with finer detail. However, it gets tricky and challenging to compute the ratio for narrower time-bin, i.e. hourly resampling, due to frequent zeros and missing values. In this regard, the challenge was mitigated by smoothing the data before computing the ratio and using cubic spline interpolation to estimate the missing blood glucose values.

**Table 6:** Description of self-collected user data, source [259], Table 3.

<b>Subject's Record Variables</b>		
<b>Variable Name</b>	<b>Description</b>	<b>Units</b>
$\mathbb{U}^{cg}$	Continuous Glucose Reading	mg/dl
$\mathbb{U}^g$	Blood Glucose Reading	mg/dl
$\mathbb{U}_f^I$	Injected Insulin (Bolus)	units
$\mathbb{U}_s^I$	Injected Insulin (Basal)	units
$\mathbb{U}^c$	Ingested Carbohydrate	grams

**Table 7:** Re-sampled, imputed, and pre-processed data, source [259], Table 4.

Preprocessed Variables		
Variable Name	Description	Units
$\mathbb{U}_{[t-\Delta t, t]}^{cg}$	Mean Continuous Glucose Reading	mg/dl
$\mathbb{U}_{[t-\Delta t, t]}^g$	Mean Blood Glucose Reading	mg/dl
$\sum_{t-\Delta t}^t \mathbb{U}_f^I$	Sum Injected Insulin (Bolus)	units
$\sum_{t-\Delta t}^t \mathbb{U}_s^I$	Sum Injected Insulin (Basal)	units
$\sum_{t-\Delta t}^t \mathbb{U}^C$	Sum Ingested Carbohydrate	grams
$\frac{\sum_{t-\Delta t}^t \mathbb{U}_f^I}{\sum_{t-\Delta t}^t \mathbb{U}^C}$	Ratio of Insulin (Bolus) to Carbohydrate	units/grams
$\frac{\sum_{t-\Delta t}^t \mathbb{U}_s^I}{\sum_{t-\Delta t}^t \mathbb{U}^C}$	Ratio of Insulin (Basal) to Carbohydrate	units/grams

### 3.3.4 Empirical data analysis

The empirical data analysis aimed to answer two principal questions related to the effect of infection episode on blood glucose dynamics; 1) To what degree does infection affect key blood glucose parameters, and 2) Which parameters can be regarded as optimal parameters, and can be used for developing the proposed personalized health model with a minimum false alarm rate as possible. Pre-selected parameters from the individual's diabetes profile were analyzed including blood glucose levels, carbohydrate, insulin, and insulin to carbohydrate ratio. The investigation was carried out based on a dataset that incorporates infection and non-infection (regular) patient-years, and the non-infection patient-years were used as a baseline to compare and validate the effect of all patient controllable and uncontrollable parameters against the self-reported period of acute infection. To better capture the difference, the data were analyzed at three levels of data granularity (timeframes), i.e. weekly, daily, and hourly. Within these timeframes, the probability distribution and temporal evolution of the pre-selected parameters were analyzed.

During the weekly analysis, a raw dataset was used, and the comparison looks into the pre-infection, infection, and post-infection week's values only. The comparison takes into account the week's daily mean blood glucose values and the sum of carbohydrate and insulin information along with the standard deviation between the days. The comparison was

interpreted based on both the inter and intra-deviations between these groups. For instance, inter-deviation implies the deviation of the infection week blood glucose level as compared to the pre-infection and post-infection week blood glucose values, while the intra-deviation signifies to what extent do the infection week blood glucose level deviates considering the infection week insulin injections and carbohydrate consumptions. The comparative results between the group were depicted using a statistical boxplot.

During the daily and hourly analysis, a smoothed dataset was used, and the comparison looks into the entire patient-years including the period of infection episode. Throughout the patient-year, the mean blood glucose response was analyzed in each time-bin taking into account the total insulin and carbohydrate values along with insulin-to-carbohydrate ratio within the respective time-bin. As described earlier, validation of the infection-induced change was performed by comparing with the non-infection patient-years. Besides, the change in data distribution was analyzed using a kernel density estimator [88; 89; 101]. To this end, a Gaussian adaptive bivariate and univariate kernel density estimator were used, and the procedure is given in **Tables 8 & 9**. To compute the optimal bandwidth, bandwidth selection approaches proposed by Botev et. al. [28] and Bowman et. al. [40; 187] were used for the univariate and bivariate estimators respectively [263].

**Table 8:** Approaches used in estimating the distribution of insulin-to-carbohydrate ratio, source [259], Textbox 1.

<b><u>Approach:</u> - One-dimensional adaptive kernel density estimation</b>
<b>1:</b> Given: Time series datasets of univariate diabetes profile parameter $X \in \mathfrak{D}$ , and Adaptive Kernel density estimator $\mathbb{M}$ – <i>one dimensional</i> ,
<b>2:</b> Remove the reported days of infection from the time series datasets $\mathfrak{D}$ , and form a new dataset $X \in \mathfrak{Q}$ ;
<b>3:</b> Compute the one-dimensional density based on the kernel density estimator $\mathbb{M}$ using $\mathfrak{D}$ and $\mathfrak{Q}$ ;
<b>4:</b> Compare the distribution from $\mathbb{M}$ ;



**Table 9:** Approaches used in estimating the distribution of blood glucose levels and insulin-to-carbohydrate ratio, source [259], Textbox 2.

<b><u>Approach: - Two-dimensional adaptive kernel density estimation</u></b>
<b>1:</b> Given: Time series datasets of bivariate diabetes profile parameters $X, Y \in \mathfrak{D}$ , and Adaptive Kernel density estimator $\mathbb{N}$ – <i>two – dimensional</i> ,
<b>2:</b> Remove the reported days of infection from the time series datasets $\mathfrak{D}$ , and form a new dataset $\mathcal{X}, \mathcal{Y} \in \mathcal{Q}$ ;
<b>3:</b> Compute the two-dimensional density based on the kernel density estimator $\mathbb{N}$ using $\mathfrak{D}$ and $\mathcal{Q}$ ;
<b>4:</b> Compare the distribution from $\mathbb{N}$ ;

### 3.3.5 Modeling approaches

This section aims to realize a personalized health model, which relies on health-related data recorded by the individual with type 1 diabetes and can monitor and screen the individual’s health status in a continuous manner and detect automatically when the individual becomes sick. The input features to the model were selected based on the empirical data analysis described above. The models implemented in this dissertation were MATLAB toolboxes, and include *ddtools*, *prtools*, and *anomaly detection toolbox* [66; 68; 218]. Two categories of approaches were evaluated and compared; one-class classifier [57; 122; 123; 164; 217] and unsupervised method [42; 87; 215]. The models were evaluated based on two attributes; data granularity (daily and hourly), and data nature (raw and smoothed data). In all evaluation scenarios, the frequency of detection is defined by setting the levels of data granularities; hourly and daily analysis.

The dataset was labeled as a set of target and non-target data. All the patient’s data, which are a regular period of the year was set as a target. The period containing the self-reported infection episode was set as a non-target. The one-class classifier models were trained on the target and tested using a dataset containing both the target and non-target data.  $N$  times k-fold stratified cross-validation was used to evaluate the performance of the one-class classifier. To a certain extent, wherever necessary class mitigation procedure was also considered [146]. Regarding the unsupervised method, no data labeling is required, and hence the entire patient-year was presented at once [87; 183].

For performance comparison, three performance metrics were used; *area under the ROC curve (AUC)*, *specificity*, and *F1-Score* [96; 97; 168; 224]. Average (standard deviation) of AUC, specificity, and F1-Score were reported. These metrics are widely used in assessing model performance for one-class and two-class tasks.

- The *area under the ROC curve (AUC)* is the integration of the ROC curve over a range of different thresholds or summing over different misclassification costs. AUC is insensitive to data imbalance and is useful to compare classifiers, however, it is

independent of a single threshold and this poses a challenge to use AUC in a real-world implementation [97].

- *Specificity* can be defined as the ratio of correctly classified non-target sample objects to the total number of non-target sample objects [192]. Hence, it represents the percentage of correctly classified infection state (non-target sample) to the total number of infection days. In my context, it is only used to illustrate the screening power of each model, i.e. accurately classifying the infection days (illness days) as such from the entire infection period.
- *F1-Score* is the harmonic mean of precision and recall, with a value ranging from 0 to 1. A model with a high (1.0) F1-Score indicates high detection performance, i.e. high recall and precision [224]. F1-Score is considered suitable for assessing model performance towards one target class and also when there are unbalanced datasets [96; 97; 124; 168; 192]. This metric is found valuable especially for an application that requires considering both the false positive rate and false-negative rate.

In addition to the performance assessment with the above-mentioned attributes, the models were also compared based on the required sample size to generate acceptable performance, the model computational time, performance associated with dimensions of input features, and performance associated with different degree of deviations in the input features, which could aid towards model selection considering real-world settings.

- The *sample size attributes* describe the minimum set of training sample objects needed for a model to produce an acceptable description. Considering a system that relies on self-recorded data, this attribute sets the lower limit for an individual to join such a system just by fulfilling the minimum data size requirement. This is important mainly because it might be difficult for an individual to accumulate a large set of datasets initially.
- The *computation time attribute* describes how long does a model take to learn and classify sample objects. Considering a system that handles lots of participants with a large dataset, the model's response time is vital in selecting the best model for the task.
- The model performance associated with *input dimensionality* compares the performance gained from using a small number of inputs as possible without affecting the model's performances. This is very important in certain cases where there is a challenge of acquiring all the necessary records.
- Performance associated with different *degrees of simulated deviations* was also used to assess the model's detection performance in response to changes that range from small to large deviations in the input feature. This attribute is essential for choosing the optimal model that can detect deviations that range from small to large values considering the fact that different pathogens could induce different deviations in the blood glucose dynamics.

### **3.4 Assessment of concerns, expectations, and willingness**

This section aims to identify a set of factors that need to be addressed during system design and implementation, the EDMON system in general and the patient unit in particular. These factors are very essential for the successful development of the patient unit in terms of long-term patient

engagement, and the successful sharing of the self-collected data to the proposed EDMON system. A system like EDMON that relies on self-collected data typically requires a timely transfer of these data for further processing to produce precise results. As a requirement, the patient is expected to register accurate and precise data and be able to transfer it promptly. In this regard, fulfilling patient expectations and addressing their concerns could have a great impact on the successful mass sharing of health-related data to the proposed EDMON system

*An exploratory sequential method* [75; 171; 239] was used, where we primarily designed a qualitative interview guide incorporating five separate themes and performed data collection and analysis, and then used the findings from the qualitative study to further inform and rectify the design of the quantitative survey questioners and data collection.

- Goals, attitude, and expectations
- Wearables and sensors usage
- Data integration
- Data sharing
- Social media and entertainment factors

In this regard, initially, a qualitative exploration, i.e. face-to-face interview, was conducted based on a *detailed concept and application scenarios* of the proposed EDMON system in general, and the patient unit, i.e. mHealth app, in particular. Then, the findings from the qualitative interviews were used in refining the quantitative survey questionnaires. As an effort to pinpoint factors that stand out only for people with type 1 diabetes, a comparative analysis with other groups of people was carried out including other chronic patients and healthy individuals. Generally, this phase of the dissertation aims at providing design strategies to develop a system that will be acceptable by the participants.

As described above, among other things, the quantitative survey aimed to focus on the data sharing theme explored in the qualitative study. In this regard, the survey explored user concerns, expectations, and willingness towards sharing self-collected health-related data to the proposed EDMON system. The survey questionnaire was distributed to various internet users and diabetes groups, i.e. Swiss English-speaking cohort of healthy people and both English and Norwegian speaking online diabetes groups. The data were collected between November 2018 and August 2019. Further detailed information about the questionnaire can be found at DataverseNO [105]. Descriptive statistics of various parameters that relate to data sharing while using different wearables and mHealth apps are reported. The study protocol has been reviewed by the Norwegian Regional Committees for Medical Health Research Ethics Northern Norway (REK) (reference number = 2017/562/REK nord) and Norwegian centre for research data (NSD) (reference number = 54558 / 3 / LB). All the participants were asked to consent during the survey and the participants response were anonymized.

## **3.5 Methodology critique**

### **3.5.1 Empirical data analysis and modeling**

Based on the existing body of knowledge, deviation from a normal state into an abnormal state can be detected via developing either a predictive or novelty detection model. In this regard,

this dissertation looks into the possibility of implementing a model for the proposed personalized health model. Regarding the prediction model-based approaches, the literature suggests a shortcoming of the existing state-of-the-art blood glucose prediction models, which often fail to accurately predict above a 30-minute prediction horizon [173; 256]. For this reason, this dissertation focuses on developing a personalized health model relying on novelty detection strategies. The number of subjects and events of infection included in the empirical data analysis and modeling can be another limitation. However, it should be noted that I was able to communicate with a lot of individuals, universities, and an online community such as do it yourself (DIY) community, for a possible donation of a dataset but I didn't receive much positive response. Further, the special requirements that oblige to include a dataset that is reliable, accurate, rich, and long enough have contributed to the failure to collect more dataset, i.e. incorporate at least the three key diabetes parameters (blood glucose level, insulin, and diet), one infection episode, and is longer than three months. For instance, given the strict inclusion criteria, I was forced to reject the dataset received from *Ohio University* for its inaccuracy containing two separate meal registrations entry. I was able to contact the university for clarification on the spotted errors and per the description provided by the university, the error was inherent and the mistake had occurred due to the problem associated with the data collection tools, which allows the participants to record meals registrations in two different places. On top of that, the error was not only having two records of meal registrations at the same time of the day, but these two records also have different values of carbohydrate registrations. It is very challenging to get an accurate, rich, and longer dataset from a large number of participants. It is worth mentioning that different types of infection (pathogens) could have a different effect on blood glucose dynamics but I studied using mainly influenza (Flu) and common cold data and this might be another limitation.

The dataset used in empirical data analysis and modeling didn't include physical activity data and this might also be another limitation. Physical activity could have a significant effect on blood glucose dynamics and its inclusion during modeling could bring a positive effect on the accuracy of the infection detection model.

### **3.5.2 Assessment of concerns, expectations, and willingness**

This sub-study could be benefited more if the number of participants is large enough in all three groups to further strengthen the conclusion, and this could be a limitation. However, it should be noted that for instance the survey questionnaire was distributed in various diabetes forums and online groups and was up for almost ten months. The other possible limitation could be linked with the unbalanced number of participants among the group used during the comparison. However, it should be noted that the comparison was only used to spot outstanding factors for people with type 1 diabetes.

## **3.6 Hardware and software tools**

The dissertation makes uses of different hardware and software tools at various phases of the sub-studies. The following hardware tools were used during data collection, empirical data analysis, and modeling sub-phases. During the data collection phase, participants have used various diabetes management technologies including smartphones, CGM, finger prickers, and

test strips, glucose meters, insulin injection (pens, and pumps). The empirical data analysis and modeling phase of the dissertation were carried out on a Lenovo laptop.

- Lenovo laptop (Intel(R) Core(TM) i7-6600U@2.81GHz, RAM 16 GB, Windows 10 64bits)
- Smartphones
- Finger prickers (lancet) and test strips
- Insulin pens and pumps
- CGM and glucose meters

The following software tools were used during data collection, empirical data analysis, modeling, and quantitative survey sub-phases. During the data collection phase, participants have used various diabetes-related technologies including mobile apps (Diabetes Diary, and Spike apps), and the xDrip with an app. Data extraction tool such as sqlitestudio was used to extract data. The empirical data analysis and modeling phase of the dissertation were carried out using MATLAB. MATLAB toolbox, Prtools, Ddtools, and Anomaly detection toolbox were used during the development of the personalized health model. SPSS statistical software used for the quantitative survey study.

- MATLAB® 2018 a & b (Mathworks, Inc, Natwick, MA).
- SPSS Statistical Software.
- Prtools is a MATLAB toolbox for pattern recognition [68].
- Ddtools is a MATLAB toolbox for data description, outlier, and novelty detection [218].
- Anomaly detection toolbox is a MATLAB toolbox that provides different unsupervised anomaly detection models [66].
- Mobile apps for data collection (Diabetes Diary, and Spike apps) [16; 212].
- Data extraction tools including sqlitestudio-3.1.1.
- Dexcom studio 12.0.4.6.
- EdrawMax is used for drawing figures and flow charts.

### **3.7 Chapter summary**

This chapter presented and discussed an overview of the materials used and the methodology followed in various sub-phases of the dissertation. To answer each research question as part of the main research problem, various sub-phases with distinct approaches were carried out. Moreover, presented a detailed description of key concepts and methodology behind developing the personalized health model along with the approach used to select optimal parameters of blood glucose dynamics indicative of infection onset amongst people with type 1 diabetes. It further presented the approaches used in assessing user concerns, expectations, and willingness for successful sharing of self-collected health-related data to the proposed EDMON system. Finally, it presented the critics of the methodology followed at each sub-phases of the dissertation, and the hardware and software tools used.



## 4 EDMON - A Personalized Health Model-based Digital Infectious Disease Detection System: Framework and Challenges

*Synopsis:* This chapter puts forward the proposed framework for a personalized health model-based digital infectious disease detection system that harness health-related data from people with type 1 diabetes collected on a daily basis to detect infection onset, and thereby utilizing this information to detecting infectious disease outbreak based on a Spatio-temporal cluster detection technique. The framework presents the main components of the proposed EDMON system along with challenges that need to be addressed during system design and implementation. The first half of the chapter presents the components and the latter half discusses the challenges. This chapter provides answers to the first research question (Q1).

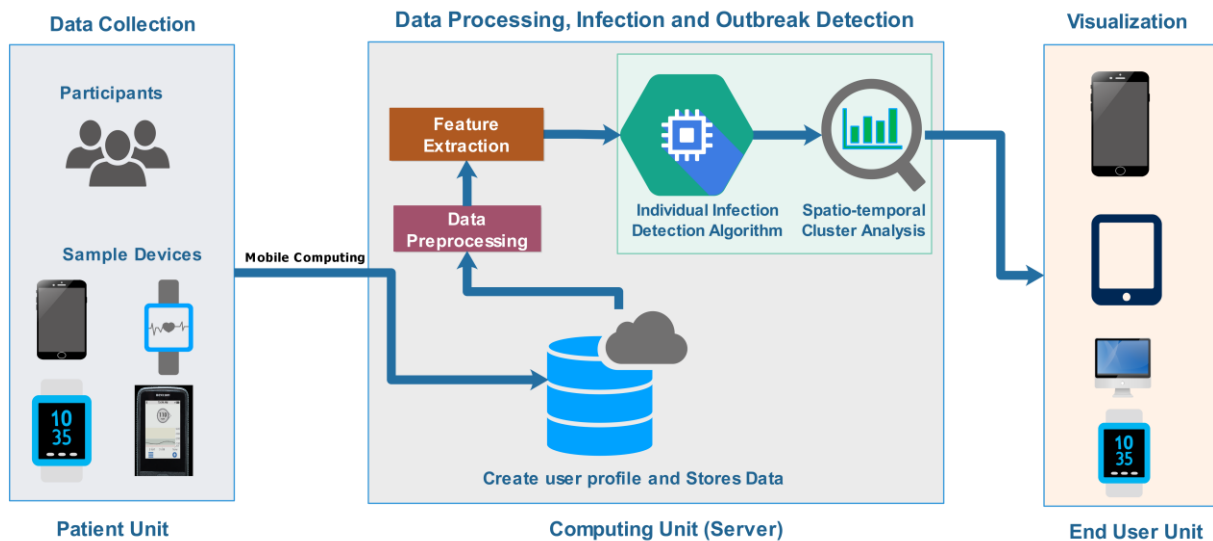
### 4.1 Introduction

Digital infectious disease detection system exploits digital health-related data and technology for the purpose of detecting infectious disease outbreaks [44; 170; 181]. Most of the existing digital infectious disease detection system relies on population-level data and doesn't possess a mechanism to trace back to the individuals who are sick [44; 181; 237]. In this regard, the proposed personalized health model-based digital infectious disease detection system is a pioneer in the field utilizing self-collected health-related data to monitor and screen the individual's health status in a continuous manner and detect infection onset and use this information to detect infectious disease outbreak among the public in a specified region of interest. Hence, this framework introduces a personalized health model concept to perform the task of individual health monitoring and detection. There are very few previous studies that attempt to describe system architecture for detecting infection incidences in people with diabetes [17; 30; 135]. However, these studies provide a system overview considering very limited data types, i.e. blood glucose levels and white blood cell counts.

The most interesting characteristic of the proposed EDMON system, i.e. a personalized health model-based digital infectious disease detection system, is that it can alert individuals about any potential health changes besides the outbreak detection. In other words, a personalized health model designed for outbreak detection can also serve the individuals by providing information that is relevant for decision support during crises. Considering the nature and the way the surveillance data is generated, the digital infectious disease detection system can be grouped as either indicator or event-based system [49; 170; 260]. In this regard, the proposed EDMON system can be regarded as an event-based digital infectious disease detection system. The events in the proposed framework are micro-event and macro-event, where micro-events depict the detection of infection at individual levels and macro-events represent outbreaks at the population level. The framework presented in this chapter is part of the findings presented in *Paper 1* [259].

## 4.2 The proposed framework

The availability of self-recorded health-related data from individuals with type 1 diabetes provides a fertile ground for the proposed EDMON system. In this regard, the proposed framework relies on user collected data and incorporates a *patient unit*, *data repository unit*, *infection detection unit*, *clustering unit*, and *information dissemination unit*, as shown in **Figure 6**. The patient unit is a user mobile devices terminal with a disease surveillance functionality and is used for collecting and transmitting data to a central server. The data repository unit stores the incoming user data. The infection detection unit tracks, monitor, and screen the individual's health status and detect infection onset by analyzing the data in real-time or near real-time based on a personalized health model (a case detector). The presence of outbreaks at any time under the region of interest is detected by the clustering unit, which uses the result from the case detectors as input and identifies any possible cluster of infected individuals based on a Spatio-temporal analysis. The information dissemination unit uses the result from the case detector and clustering unit to provide real-time or near real-time health status information for the individual participants as well as information regarding the current status of the area under surveillance for the concerned bodies, i.e. ordinary citizens, patients, and public health officials. *The rest of the dissertation chapters focus on topics related to the patient unit and infections detection unit and investigates how to provide a solution for these topics.*



**Figure 6:** The proposed framework of EDMON - a personalized health model-based digital infectious disease detection system, adapted from [58; 255; 259], Figure 14.

### 4.2.1 The Patient unit

The patient unit, as shown in **Figure 6**, is consisting of sensors and data, and a mobile health application (mHealth app), where the sensors reading and data are integrated and stored in the mHealth app before being transmitted to the central server.



#### 4.2.1.1 Sensors and data

As part of self-management, people with type 1 diabetes collect different parameters that affect blood glucose dynamics and this group of the patient often record the key parameters; blood glucose levels, insulin administration (bolus and basal), carbohydrate ingestion given in grams, and could also record other optional parameters such as physical activity session or exercise load, heart rate, stress, sleep quality, and others [3; 148].

##### 4.2.1.1.1 Compulsory data and equipment

The proposed EDMON system requires compulsory data and their associated pieces of equipment to perform the day to day operation; these compulsory data include blood glucose level, insulin, carbohydrate, and user's location information and pieces of equipment needed to continuously record these pieces of information.

Blood glucose levels can be measured using either finger pricks based self-monitoring blood glucose (SMBG) or continuous glucose monitoring (CGM). Finger pricks based glucose monitoring requires a lancet (finger pricker), test strips, and glucose meter [127], and various brands on the market can automatically transfer the result into a smartphone app. CGM measures the individual's blood glucose levels in real-time, e.g. every 5 minutes, and continuously throughout the day and 24/7. There are various brands of CGM devices available on the market including Dexcom<sup>2</sup>, FreeStyle<sup>3</sup>, Guardian<sup>4</sup>, Eversense<sup>5</sup>, and others [129]. CGM devices incorporate a sensor, and a transmitter, which estimates the blood glucose values and transmits the result into a receiver, which is usually a smartphone app, pump, or other receivers.

Insulin administration is crucial in diabetes self-management to regulate blood glucose levels. There are different types of insulin based on how quickly they act to lowering blood glucose levels; administered insulin can be fast-acting (bolus), intermediate-acting or, slow-acting (basal) [240; 241]. Insulin administration is carried out based on either multiple daily injections (MDI), or continuous subcutaneous insulin infusion (CSII) [63; 125; 241]. MDI is mostly carried out using insulin pens to inject insulin and there are different brands of insulin pens on the market including Owen Mumford Autopen<sup>6</sup>, Humalog Kwikpen<sup>7</sup>, Novo Nordisk NovoPen 4<sup>8</sup>, and, others. CSII is usually carried out using insulin pumps and there are different brands of insulin pumps on the market including Medtronic<sup>9</sup>, Omnipod, Accu-Chek<sup>10</sup>, Tandem Diabetes Care<sup>11</sup>, and others [139].

Carbohydrate counting is necessary to estimate the amount of insulin intake during mealtime and further action to correct blood glucose levels [84]. Carbohydrate amount measured in grams is estimated for a given meal, which usually depends on the individual patient's expertise related

---

<sup>2</sup><https://www.dexcom.com/home>

<sup>3</sup><https://www.diabetescare.abbott/products.html>

<sup>4</sup><https://www.medtronicdiabetes.com/products/guardian-connect-continuous-glucose-monitoring-system>

<sup>5</sup><https://www.eversenseddiabetes.com/>

<sup>6</sup><https://www.owenmumford.com/en/patients-product/autopen/>

<sup>7</sup><https://www.humalog.com/taking-humalog/using-u100-u200-kwikpen>

<sup>8</sup><https://www.novonordisk.com/>

<sup>9</sup><https://www.medtronicdiabetes.com/home>

<sup>10</sup><https://www.accu-chek.co.uk/insulin-pumps>

<sup>11</sup><https://www.tandemdiabetes.com/home>

to counting carbohydrate contents of each food type and drink. Ideally, this patient group is expected to rely on an advanced (Level 3) [84] carbohydrate counting, which uses the individualized insulin-to-carbohydrate ratios [63; 167; 210; 228] to estimate the insulin amount at a time. The level of knowledge towards carbohydrate counting varies from individual to individual, which determines the quality and accuracy of the registered values. Most importantly, since carbohydrate is registered manually, it is most likely prone to errors.

Apart from these key parameters of the blood glucose dynamic, the proposed EDMON system requires the individual's user location for the purpose of locating a cluster of sick individuals. The user location information can be recorded as either a static or dynamic address and can be in the form of longitude and latitude, postal codes, or any other reference coordinates. The ethical challenges raised by using the individual's health-related data and location information are discussed in the next section.

#### **4.2.1.1.2 Optional data and equipment**

Physical activity sessions, heart rate, stress, sleep quality, and other certain physiological parameters are also important parameters where people with type 1 diabetes could record for the purpose of self-managing their blood glucose levels. These optional datasets can be used to further improve the system detection accuracy if properly recorded on a timely basis. Registration of physical activity sessions or exercise load information is very important along with carbohydrate registration to decide the right amount of insulin intake. Physical activity sessions or exercise load could enhance insulin sensitivity thereby reducing the amount of insulin requirements [27]. There are various wearable sensors that estimate and measure physical activity sessions on the market including Polar<sup>12</sup>, Garmin<sup>13</sup>, Apple Watch<sup>14</sup>, Fitbit<sup>15</sup>, Samsung Watch<sup>16</sup>, and others [73; 103; 104; 213]. Most of the existing wearables also have the capability to record other parameters such as heart rate. In this regard, for example, the following brands incorporate heart rate sensors; Polar, Samsung Watch, Garmin, Scosche Rhythm, Wahoo Tickr Fit, and others. Moreover, there are also wearable that monitor sleep quality including Polar, Fitbit Versa, Oura Ring<sup>17</sup>, Withings Move<sup>18</sup>, and others. There are also wearables that estimate stress levels based on the individual heart rate variability including Garmin, Samsung smartwatches, Apple Watch, Fitbit, Google Wear OS smartwatches<sup>19</sup>, and others.

#### **4.2.1.2 Mobile health (mHealth) app**

The purpose of the mobile health (mHealth) app is to integrate sensor readings and data from various diabetes-related technologies and wearables devices. Automatic data collection features are favored by people with diabetes and whenever possible, the app should support automatic data registrations capability and provide only manual registrations if it is a must [252]. Different

---

<sup>12</sup><https://www.polar.com/en>

<sup>13</sup><https://www.garmin.com/en-US/>

<sup>14</sup><https://www.apple.com/watch/>

<sup>15</sup><https://www.fitbit.com/global/no/home>

<sup>16</sup><https://www.samsung.com/us/watches/>

<sup>17</sup><https://ouraring.com/>

<sup>18</sup><https://www.withings.com/us/en/withings-move>

<sup>19</sup><https://wearos.google.com/#hands-free-help>

sensors reading from blood glucose reading devices, insulin pens and pumps, physical activity and exercise sessions, heart rate, sleep quality, stress and other physiological parameters such as blood pressure, body temperature, and others along with diet and geographical location information should be integrated to the app [58; 255]. There exist a variety of self-management mHealth apps in Google Play and AppStore including the most well-known apps such as Diabetes Diary<sup>20</sup>, Spike App<sup>21</sup>, and mySugr<sup>22</sup>. Hence, the proposed app exhibit similar functionality with these well-known apps and further add surveillance functionality required by the proposed EDMON system. In this regard, the main requirements and functionality of the proposed mHealth app include:

- Automatically record blood glucose levels, both SMBG and CGM.
- Automatically record insulin units, both bolus and basal insulin in a separate record.
- Automatically record physical activities and exercise sessions.
- Automatically record heart rate, emotional stress, and sleep quality in a separate record.
- Enable manual recording of certain parameters, especially for diet information, other medications, and illness status.
- Automatically record user geographical location. For example, a user location can be estimated based on the Global Positioning System (GPS) from the phone upon data registration (Coucheron et al., 2019). User location can be in terms of longitude and latitude [85], postal code address [86], or any other local reference coordinates.
- A timestamp for each data registration.
- Transmit the data to a central server upon each registration. The data packet could include all the above compulsory parameters and the optional parameters if available along with the geographical location and time of registration tagged on it.

#### 4.2.1.3 Communication architecture and protocols

The proposed personalized health model-based digital infectious disease detection system, as shown in **Figure 6**, is a three-tier architecture that contains three different tiers performing different tasks. In this configuration, the personalized health model can be placed either in the *patient unit* or *remote computing unit*. Placing the personalized algorithm in the remote computing server requires all the required user data to be transmitted to the server and this configuration could be prone to degraded accuracy as a result of remote site computations emanating from data transmission requirements [58]. Moreover, it brings challenges in terms of user data security, privacy, and confidentiality issues as a result of migrating the user data from the smartphone to the central server. However, instead of moving the user data to the central server, it is possible to place the personalized health algorithm within the user smartphone app, where only the health status of the individual (as normal, suspicious, and infected) is transmitted to a centralized server for further clustering computation. However, unlike the remote computing unit, placing the personalized health algorithm in the user's smartphone app (patient unit) needs a feasibility study given the fact that the algorithm needs

---

<sup>20</sup><https://play.google.com/store/apps/details?id=no.telemed.diabetesdiary&hl=en&gl=US>

<sup>21</sup><https://spike-app.com/>

<sup>22</sup><https://www.mysugr.com/en/>

to run frequently to identify and detect deviations from normality; thereby could require high power consumption and memory spaces that could affect the user's smartphone.

Most of the existing diabetes equipment and technologies have some form of access to the data. Integrating the data to the mHealth app, transmitting the data to the computing server, and storing the data requires to ensure complete data security, protection of user privacy, and confidentiality throughout the system's data flow [107]. In this regard, it is necessary to use state-of-the-art communication protocols that ensure security, robustness, and privacy and strictly follow compliance with international regulations.

#### **4.2.1.4 Challenges**

The proposed personalized health model-based digital infectious disease detection system highly depends on the accuracy of the data, and entirely on its collection procedure. As described above, one of the main challenges regarding data collection and transmission is related to data accuracy, reliability, data security, privacy and confidentiality, and user acceptance and willingness to share data.

Data quality is the most critical factor that needs to be addressed since accurate analysis in the proposed EDMON system assumes high precision sets of data. An inaccurate dataset could hamper system accuracy and result in unpredictable performance degradation. Data accuracy challenges could emanate during manual data registration into the mHealth app, data integration into the mHealth app, and data transmissions to a remote server. Manual data registration to the mHealth app is prone to errors. For instance, a user can incorrectly record carbohydrate amounts upon registration, which could greatly hamper the accuracy of the personalized health model. Therefore, it is necessary to look for a mechanism that can cross-check values by requesting the user to validate the input especially when an out-of-range input value is recorded. In the same fashion, the necessity of integrating sensors reading into the mHealth app requires dealing with heterogeneous data formats due to multiple vendor involvements. The main challenge in this regard is standardization and interoperability issues that need to be addressed. Apart from these challenges, transmitting the data to a remote server could result in missing, corrupted, and delayed data that could affect the system's accuracy. In this regard, it is necessary to look for a method that ensures the quality of information through an advanced pre-processing and data quality control algorithm [58; 255; 259].

The sensitivity of user-health data is another challenge that needs to be carefully addressed through the entire system's data flow. In this regard, one of the possible approaches could be to de-identify and anonymize the data following international regulating body guidelines such as General Data Protection Regulation (GDPR) and Health Insurance Portability and Accountability Act (HIPAA) [83; 107]. Apart from this, it is also necessary to follow the state-of-the-art privacy-preserving and secure data communication protocols to ensure that data security, privacy, and confidentiality are respected throughout the data collection and transmission phase. It is necessary to understand that the successful design and acceptability of the proposed EDMON system relies on fulfilling those requirements. Moreover, as described above, it is necessary to tag each user data with the user's geographical location to successfully locate a cluster of infected individuals on a Spatio-temporal basis. In this regard, in addition to

the de-identification procedure, it is necessary to look for a robust approach that can strictly hide user identity upon transmission in case if the data is compromised during data transmission.

Finally, a new mHealth app usually faces acceptance challenges by the intended users for a variety of reasons [252]. These factors range from user perception of the app, lack of motivation and trust issues related to using the apps and sharing of the data to the proposed EDMON system, and ease of use associated with the app complexity. Therefore, it is necessary to persuade and motivate the user to continuously engage with the app by buying user trust and answering all their concerns, and addressing ethical and motivational challenges.

## **4.2.2 Computing unit**

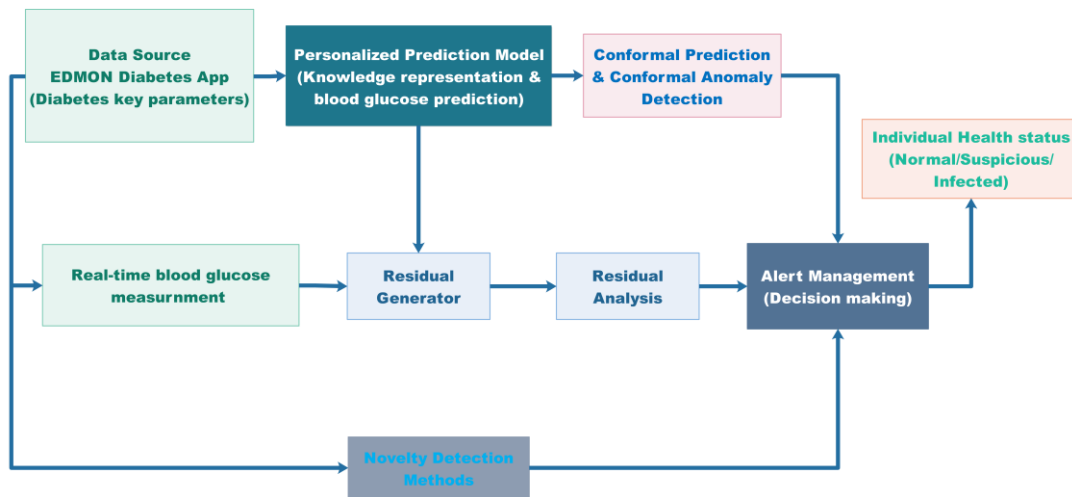
As shown in **Figure 6**, the proposed computing unit performs three important tasks; storing the data, tracking the individual health status, and detecting a cluster of infected individuals based on a Spatio-temporal basis.

### **4.2.2.1 Data repository (database) unit**

The data repository stores the individual's data transmitted from the mHealth app. The database should store each data in a structured format containing user key ID, geographical location, time of data registration, and a record containing all the compulsory and optional datasets. The individual's geographical location could be either static or dynamic address depending on the necessity [58; 259]. The static address can be a home or work address and recorded during user creation and upon individual registration to the system. On the other hand, the dynamic address is updated in real-time during each new data registrations. This type of address is very useful to trace back contacts upon infectious disease outbreaks.

### **4.2.2.2 The Infection detection unit – individual level**

The proposed personalized health model monitors the individual's health status and detect when the individual become sick, which is designated as *mirco-events*. This requires to tracking the individual's health status through developing a personalized health model that learn from the historical diabetes profile and judge the current information. The modeling can be realized through either prediction model [173; 256] or novelty detection methods [51; 70; 153; 178], as shown in **Figure 7**. The prediction model can be implemented as either a residual [99; 100; 254; 266] or conformal approach [112; 136; 137; 204; 211; 233-235]. Similarly, novelty detection can be realized through a supervised, semi-supervised, or unsupervised approaches [51; 70; 153; 253]. In literature, different categories of novelty detection methods have been reported including classification or domain-based [110; 116; 124; 150; 198; 217; 221; 242; 253], statistical techniques based [178], clustering-based [6; 217], distance-based [51; 87], ensemble-based approaches [6; 51; 87; 136; 153; 169; 178; 206; 242; 265; 267], prediction-based, and density-based [42; 169; 215; 267].



**Figure 7:** Alternative methods for realizing the proposed personalized health model, source [259], Figure 15.

The alarm management (decision making) unit processes the output from the personalized health model and assign a status to the individual’s health. Hence, the individual’s health status at any time could be among these three states; normal (0), suspicious (-1), and infected (1) [259]. The normal state depicts when the current readings conform to past knowledge. A suspicious state depicts when the current readings begin deviating but not enough to be designated either as a normal or infected state. An infected state depicts when the current reading deviates completely from past knowledge. The unit also needs to compute and inform the individual about the degree of deviation when detected as being sick. The output from this unit will notify the individuals and also be used as input to the cluster detection analysis.

#### 4.2.2.3 The Clustering unit – population level

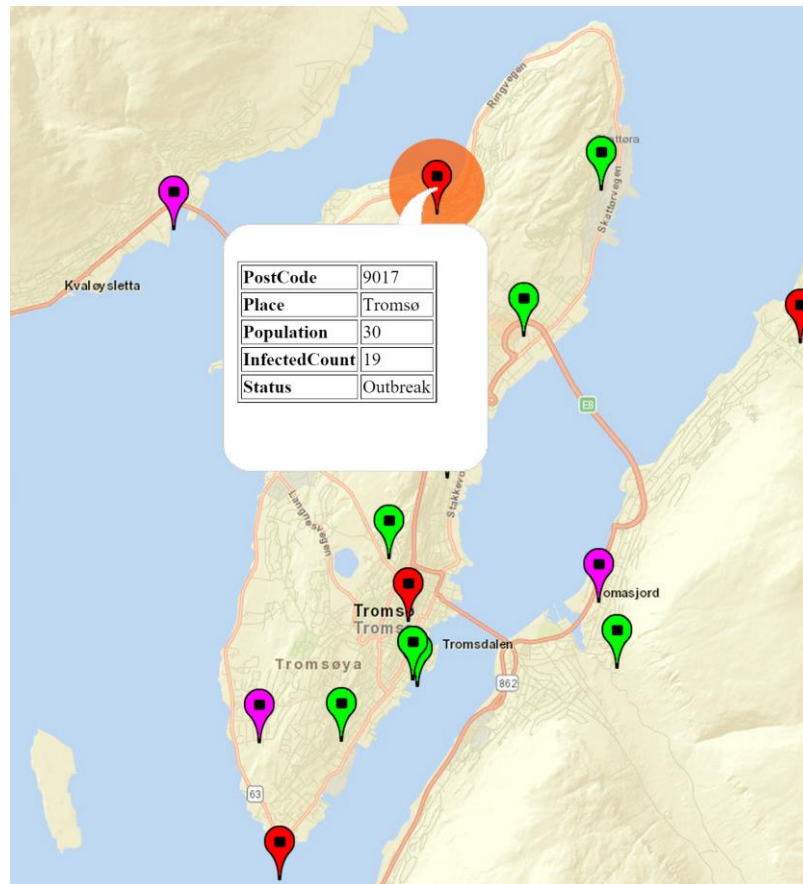
The clustering unit accepts outputs from each individual’s alarm management unit and counts the number of individuals who are reported to be in an infected state. This count data can be used as either a standalone or in conjunction with other data sources, other similar disease surveillance system, to aid in outbreak detection. The following discussion assumes the standalone context and outlines challenges the way forward. In this regard, the proposed cluster detection analysis accepts the individual’s health status, location, and the time of micro-events detection [255] and performs a computation to detect any unexpected rise in either count or rate of infected individuals within a pre-specified region of interest, also known as *macro-events* [259]. As described above, the individual user’s location can be recorded as a static address such as home and postal code address [92] or a dynamic address like longitude and latitude coordinates [67] that can be estimated from the mobile phone’s GPS information upon data registration. The frequency of computation determines the time-step required to repeat the scanning procedures and can be fixed to anything between 1 hour to 24 hrs depending on the system design requirement [259].

Detection of outbreaks can be realized as either spatial, temporal, or Spatio-temporal methods depending on the type of information utilized [261; 262]. According to the literature, the state-of-the-art outbreak detection methods require a certain size of surveillance data from a large

population to draw a statistically valid conclusion [236; 237]. However, certain cases like biological terrorism affect a particular region and relatively small population and outbreak detection methods devised to detect such an incidence requires answering practical challenges related to data sparsity [237]. Data sparsity often arises in a system that tracks a small number of the population. Tracking surveillance data from a small population within a certain period often results in small event counts, which pose a challenge to draw a statistically valid conclusion in practical settings. Furthermore, implementing global detection, which considers the entire region as one, could end up masking local outbreaks affecting a small proportion of the region. By the same token, partitioning the entire region of interest into small equivalent areas could also end up in multiple testing problems [170; 199; 236]. In the same manner, an outbreak detection mechanism that exploits data from personalized surveillance of people with type 1 diabetes faces similar challenges, i.e. relatively small and scattered population. In this case, it is necessary to devise outbreak detection mechanisms, which are less sensitive to geography and sparse data from a small population. The output from this unit is used to notify the responsible individuals through the information dissemination unit.

#### **4.2.3 The Information dissemination (visualization) unit**

As shown in **Figure 6**, the information dissemination unit processes the output from the cluster detection analysis and displays real-time information in an interactive way with the appropriate format to be easily understandable by the end-users. Various data presentation formats such as tables, graphs, and maps can be used for depicting the outbreak information [259; 262]. An example depicting a detected hypothetical outbreak on an interactive map is given in **Figure 8**. The real-time information can be viewed through a dedicated website or app [255]. In this regard, the end-users are consumers of the real-time information for quicker public health actions and individual participant, who wants to know his/her health status along with a degree of deviation from the normal process.



**Figure 8:** Example depicting mapping of a detected cluster in a hypothetical outbreak.

### 4.3 Motivational and ethical challenges

The proposed personalized health model-based digital infectious disease detection system completely relies on user-generated data mainly collected for other purposes and typically requires accurate, reliable, and precise data in a continuous fashion [259]. Getting such kind of data requires motivated users who engage with the patient unit for a longer period of time, and also willing to share the data with the proposed EDMON system. However, it is easier said than done and this requires considering various factors that could enhance user motivation during system design and implementation and also address user's concerns and expectations towards the system [252]. As described above, generally, a new system usually faces acceptance challenges by the intended users for a variety of reasons. These factors range from user perception of the system, lack of motivation especially related to using the mHealth apps and sharing of the data to the system, lack of trust related to data security, privacy, and confidentiality issues, and ease of use associated with system complexity [2; 7; 85; 120; 252]. Therefore, it is necessary to persuade and motivate the user to continuously engage with the system by buying user trust by answering all concerns, ethical, motivational, and other security-related challenges.

### 4.4 Knowledge summary

In the context of the original research question, the added knowledge is presented within the scope of the proposed framework of a personalized health model-based digital infectious disease detection system.



*What do we know about the topic already?*

- Various systems that collect several diabetes-related health data to a remote server for remote diabetes management, monitoring and follow-up systems, data analysis for personalized feedback, and decision making have increasingly been studied and presented in the literature [5; 149; 163]. However, none of these consider detecting infection onset at an individual level.
- Various digital infection detection systems have been studied and presented in literature utilizing various digital data ranging from a query search engine, work and school absenteeism, over the counter pharmacy drug sales, and other sources [44; 170; 232]. However, none of these utilizes a personalized health model to detect infectious disease outbreaks at the population level.
- There are very few previous studies that attempt to describe system architecture for detecting infection incidences in people with diabetes [17; 30; 135]. However, these studies provide the system overview considering very limited data types, i.e. blood glucose levels and white blood cell counts.

*What does this chapter add to our knowledge?*

- Presented a framework that combines self-management practices in people with type 1 diabetes and a disease surveillance system concept to fill the gaps in the existing digital infectious disease detection system via utilizing a personalized health model to detect infection onset at the individual level and thereby utilizing this information to detect outbreaks at the population level.
- The presented framework incorporates different units performing a series of tasks; patient unit, data repository (database) unit, infection detection unit, clustering unit, information visualization unit, and a wireless communication platform.
- Highlights the main challenges in relation to system design and implementation along with ethical, motivational, and data sharing challenges within the scope of the proposed EDMON system.

## **4.5 Chapter summary**

This chapter presents and discussed the framework of the proposed personalized health model-based digital infectious disease detection system. The realized framework identifies five important constituents of the proposed EDMON system, relationships, and task requirements of the components; a patient unit, data repository unit, infection detection unit, clustering unit, and information visualization unit. The *patient unit* is a standalone smartphone app (mHealth app) that integrates and stores different sensor readings, which could be through either manual or automatic recordings. The individual's data-structure should contain two types of data; compulsory and optional data types, where the compulsory data includes blood glucose levels, insulin and carbohydrate registration, geographical location, and time of registration. Currently, most diabetes technologies enable Bluetooth connections to foster the integration of sensor readings. The data stored in the mobile app needs to be transmitted to a *database server*, where it is stored for further processing. However, given the sensitivity of health data, high emphasis

needs to be given to data security, privacy, and confidentiality. In this regard, the transmission and storage of data in a server need to strictly follow major international guidelines, e.g. HIPPA compliance. The *infection detection unit* access the individual's records from the database and execute the personalized health model, which is trained on the individual's historical data, to look for any abnormal deviations promptly. In this regard, the personalized health model can be formulated as blood glucose prediction or novelty detection. At any given time, the output from the personalized health model is the individual's health status coded as normal (0), suspicious (-1), and infected (1). The *clustering unit* accepts the timely health status from each individual participant along with their respective geographical location and time of data registration to perform a Spatio-temporal analysis to detect a group of infected individuals under a region of surveillance. The status of the region or city under surveillance is visualized based on either a standalone smartphone app (mHealth app) or a web-based or both. Besides, this chapter also presents various challenges from both users and technological perspectives for the successful acceptance of the proposed EDMON system; 1) data accuracy related to manual data registration into the mHealth app, data integration into the mHealth app, and data transmissions to a remote server, 2) motivation related to user engagement to the patient unit for a longer period 3), willingness related to user perception concerning data sharing and willingness to share 4) data sparsity related to the statistical significance of outbreak detection alarms derived from a small and sparse population under surveillance.

## 5 Infection Characterization and Parameter Selection

*Synopsis: This chapter puts forward the empirical data analysis conducted to study the effect of infection on key parameters of blood glucose dynamics amongst people with type 1 diabetes. The idea was to investigate, characterize and select optimal parameters for developing the proposed personalized health model. The analysis focuses on the trend and the probability distribution of the key parameters. Different data granularity was considered during the analysis; weekly, daily and hourly timeframes. The first half of the chapter provides an overview of the dataset and proceeds to the trend analysis and probability distribution comparison. Finally, it provides a concluding remark. This chapter provides answers to the second research question (Q2).*

### 5.1 Introduction

As described in the background chapter, several factors disturb blood glucose dynamics [22]. Most of these factors induce predictable disturbances except factors such as physical and emotional stress, diseases, acute illness, severe wounds, infections, and others [253; 256]. In this regard, it is quite necessary to conduct empirical data analysis to understand the confounding nature of the other unpredictable factors along with the predictable ones on the path towards detecting infection onset and also to assess and numerically estimate the impact of infection on the key parameters of blood glucose dynamics. To the best of my knowledge, this is the first study that empirically and numerically quantifies the effect of infection episodes on key parameters of the blood glucose dynamics among people with type 1 diabetes exploiting self-recorded data. To this end, this chapter focus on characterizing the disturbance infection onset creates on the blood glucose dynamics amongst people with type 1 diabetes and thereby choosing optimal parameters for developing the proposed personalized health model. The association between infection and hyperglycemia has been reported in the literature [41; 47; 152; 160]. However, the idea of using self-recorded data for the purpose of detecting outbreaks is a novel and recent phenomenon. In this regard, there are few literature that investigate the potential of using these data sources for disease surveillance purposes [17; 29-36; 38; 90; 98; 134; 135; 209; 255]. For instance, Botsis et.al. [29; 31; 34; 38] reported elevated blood glucose levels and hemoglobin A1c (HbA1c) for the duration of infection as compared to the normal period. Generally, these studies reported the association and suggested the possibility of detecting infection in this patient group by exploiting self-recorded data. Despite these findings, none of this literature investigated the nature and degree of abnormalities triggered by infection onset through systematic analysis of each parameter of blood glucose dynamics. The results presented in this chapter are part of the findings presented in *Paper 1* [259].

### 5.2 Raw dataset

As described in the method chapter, the raw dataset used in the empirical data analysis comprised of self-recorded data containing blood glucose levels, amount of insulin injection, amount of carbohydrate in the diet (grams), and infection episodes. In total, the length of the data is ten patient-years, among which the first half (five years) is regular years with no infection episodes, and the latter half (five years) contains a minimum of one infection episodes per individuals. An exemplar raw data containing influenza (flu) infections are given in **Figures**

9-10. The figures illustrate two patient-year containing the three key parameters of blood glucose dynamics.

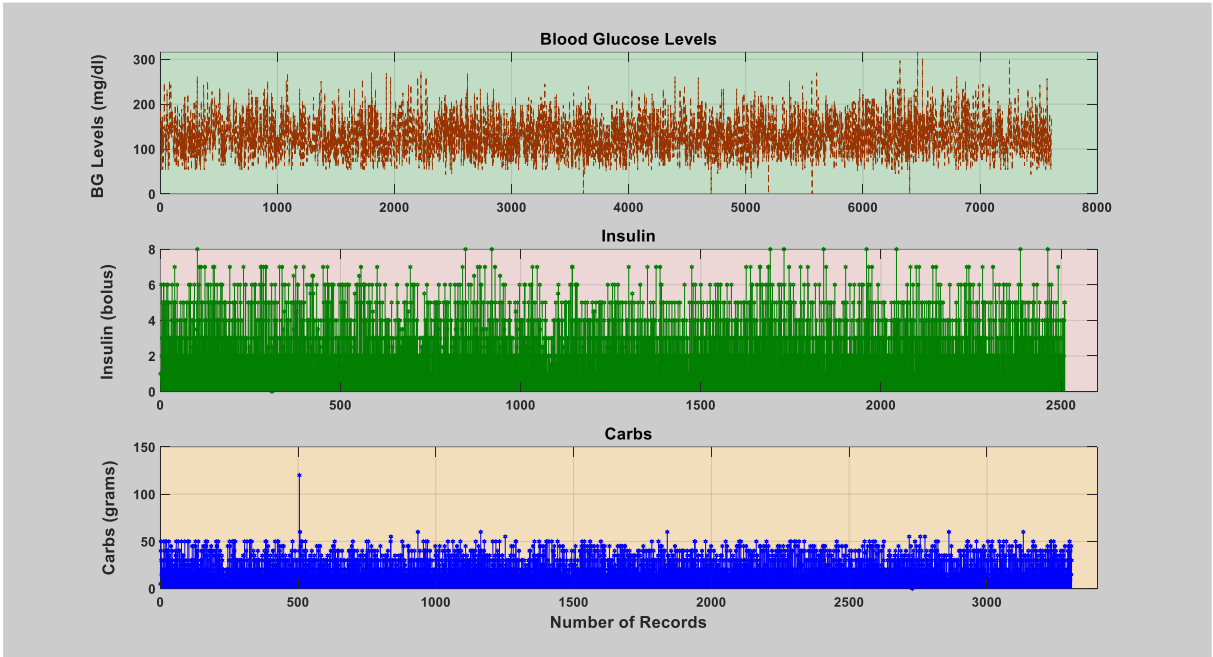


Figure 9: Illustrates a particular patient-year with influenza (flu) infections in the first week of December.

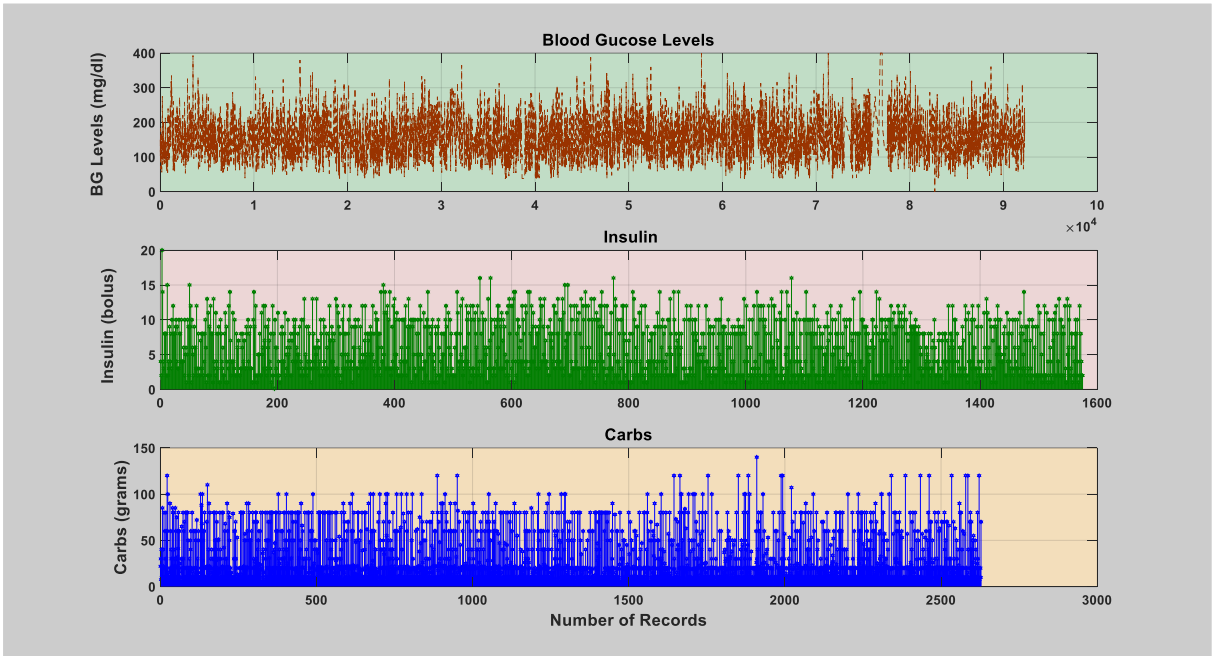


Figure 10: Illustrates a particular patient-year with influenza (flu) infections in the second week of August.

## 5.3 Data preparation and comparative analysis

As described in the method chapter, the raw dataset was first smoothed by a moving average filter. Different length of window sizes was tested to choose the optimal value, and a value of 1, 2, 3, and 4 days was evaluated. Hence, a window size of 2 days (48 hours) was chosen and the choice was much constrained to avoid substantial delays that could arise from using a larger window size. The comparative analysis considers three data granularity (timeframes); hourly, daily, and weekly. Comparison of the first two timeframes uses the smoothed dataset, however, the weekly comparison uses the raw dataset and applies only to the dataset containing infection episodes. Within each timeframe, the comparative analysis investigates the change in average blood glucose levels taking into account the impact of total carbohydrate consumption and insulin intake. It is better to note that, the regular patient-years were aimed at establishing a reference (baseline) knowledge of the individuals under normal period. The trend and distribution of the data were investigated to pinpoint uncharacteristic deviations as a result of the infection episode.

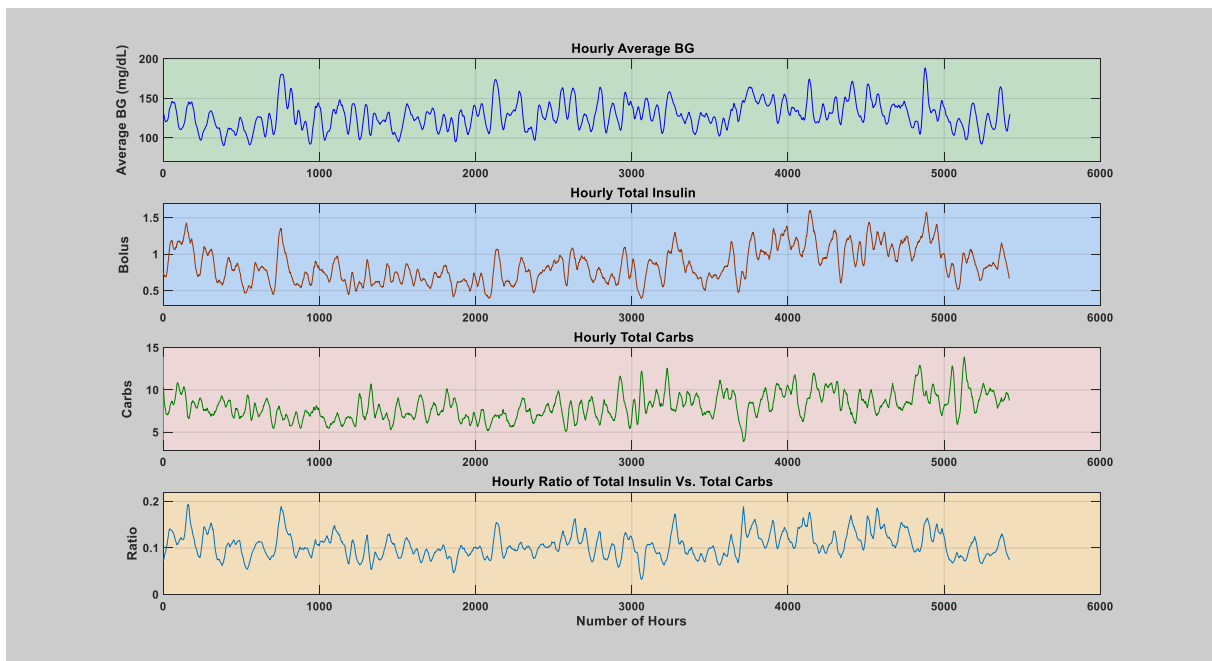
### 5.3.1 Data trend

#### 5.3.1.1 Regular patient-years

The analysis of the regular patient-years with the two timeframes; daily and hourly, exhibit the normal process of blood glucose dynamics, as can be seen from **Figure 11-12**, and more figures can be found in *Appendix 2 of Paper 1* [259]. Insulin is administered in pre, during, and post-meal scenarios taking into account the amount of carbohydrate consumption. Occasionally, correction insulin is also injected depending on the resultant blood glucose levels [26; 175]. The nature of this self-management practice is mainly manifested through the evolution of the insulin-to-carbohydrate ratio [63; 132]. As can be seen from the figures, values of the insulin-to-carbohydrate ratio are relatively stable depicting the normal process of the individual's blood glucose dynamics, where blood glucose regulation is mainly maintained with a balanced consumption of carbohydrates and insulin requirements. As per the findings, the values of the ratio oscillate between 0.05 and 0.2 throughout the entire patient-years. This dynamic but stable characteristic of the ratio signifies the impact of patient-controllable and uncontrollable parameters apart from infection [259].

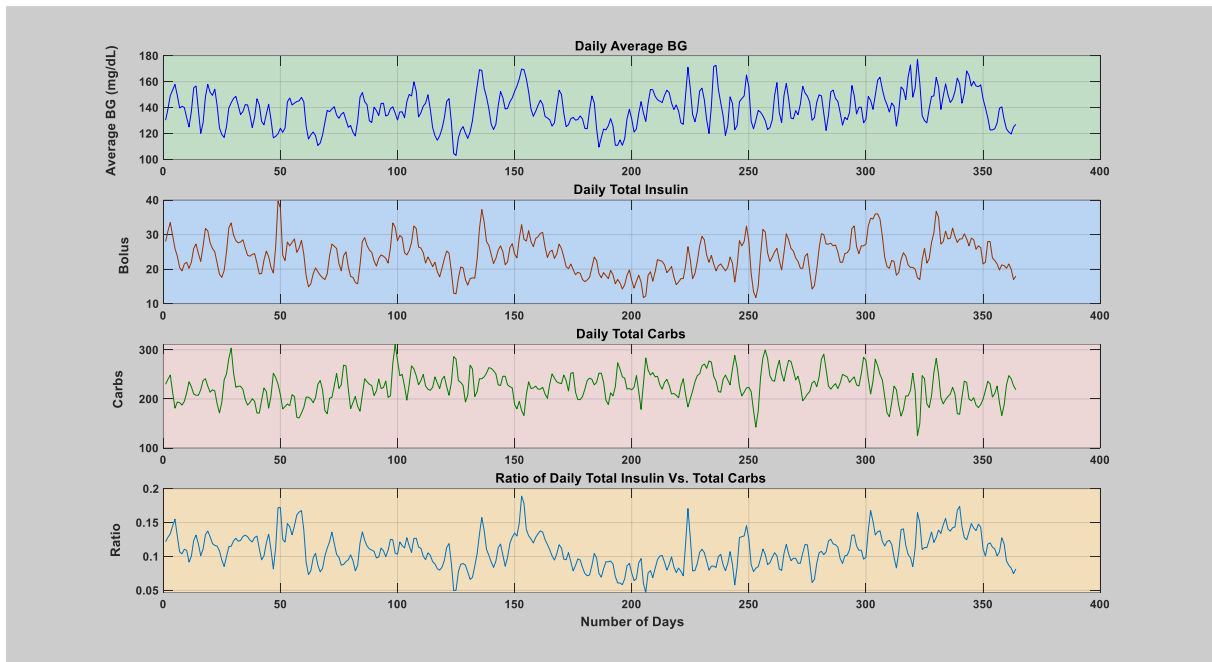


a) A daily plot of the patient-year, source [259] Appendix 2, Figure 1.

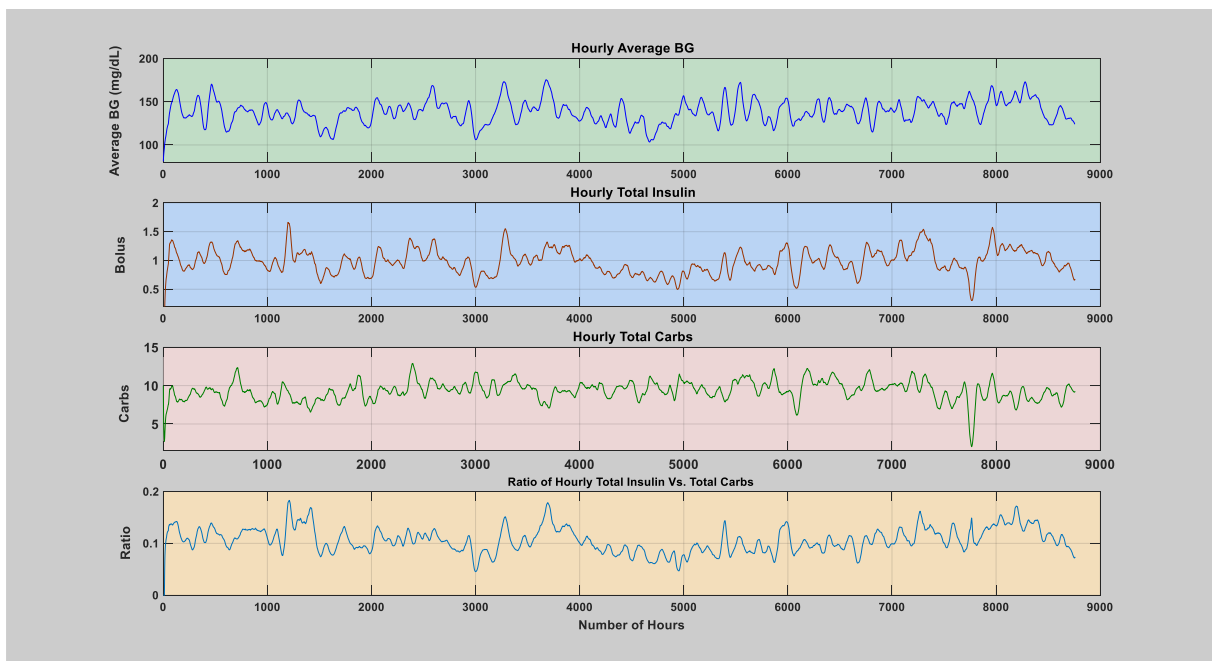


b) An hourly plot of the patient-year, source [259] Appendix 2, Figure 1.

**Figure 11:** A specific regular patient-year depicting blood glucose levels, amount of carbohydrate consumption and insulin injections, and fluctuations of the insulin-to-carbohydrate ratio. Figure (a) illustrates fluctuations within the daily timeframe. Figure (b) illustrates fluctuations within the hourly timeframe.



a) A daily plot of the patient-year, source [259] Appendix 2, Figure 3.



b) An hourly plot of the patient-year, source [259] Appendix 2, Figure 3.

**Figure 12:** A specific regular patient-year depicting blood glucose levels, amount of carbohydrate consumption and insulin injections, and fluctuations of the insulin-to-carbohydrate ratio. Figure (a) illustrates fluctuations within the daily timeframe. Figure (b) illustrates fluctuations within the hourly timeframe.

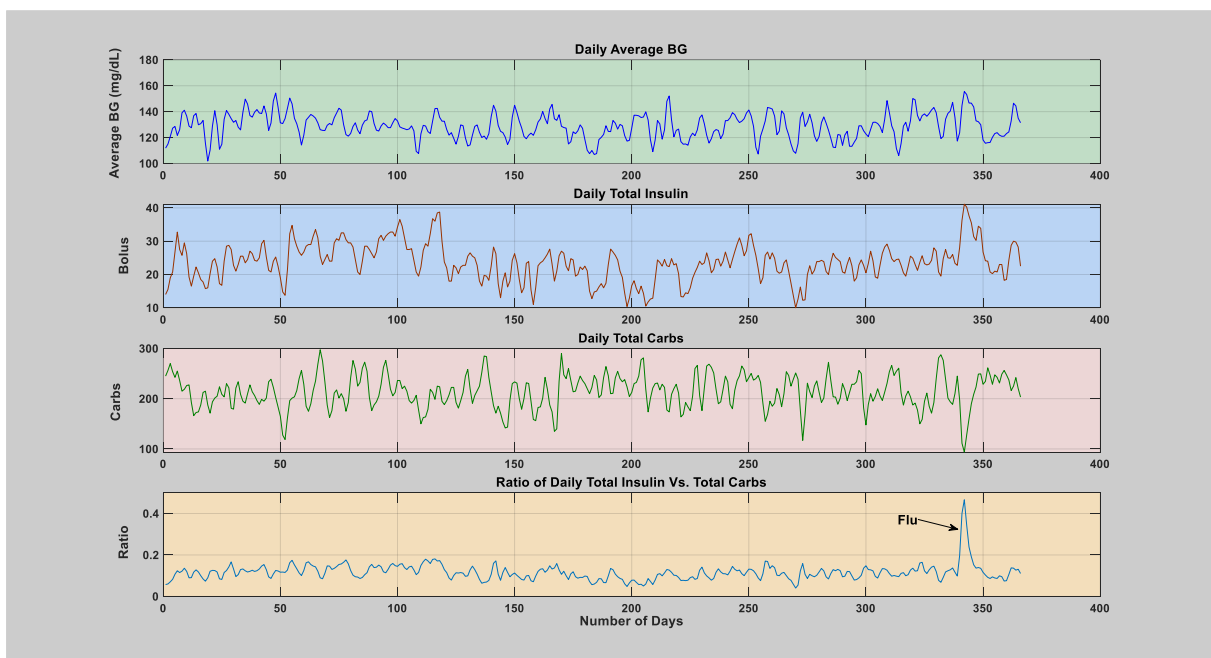
### 5.3.1.2 Patient-years with Infection events

As described above, in the patient-years with infection episodes, apart from the trend, the distribution of data was also compared on a daily and hourly basis. Furthermore, a pre-infection

week, infection week, and post-infection week comparison of parameters were carried to characterize and single out the impact of the infection episode effectively.

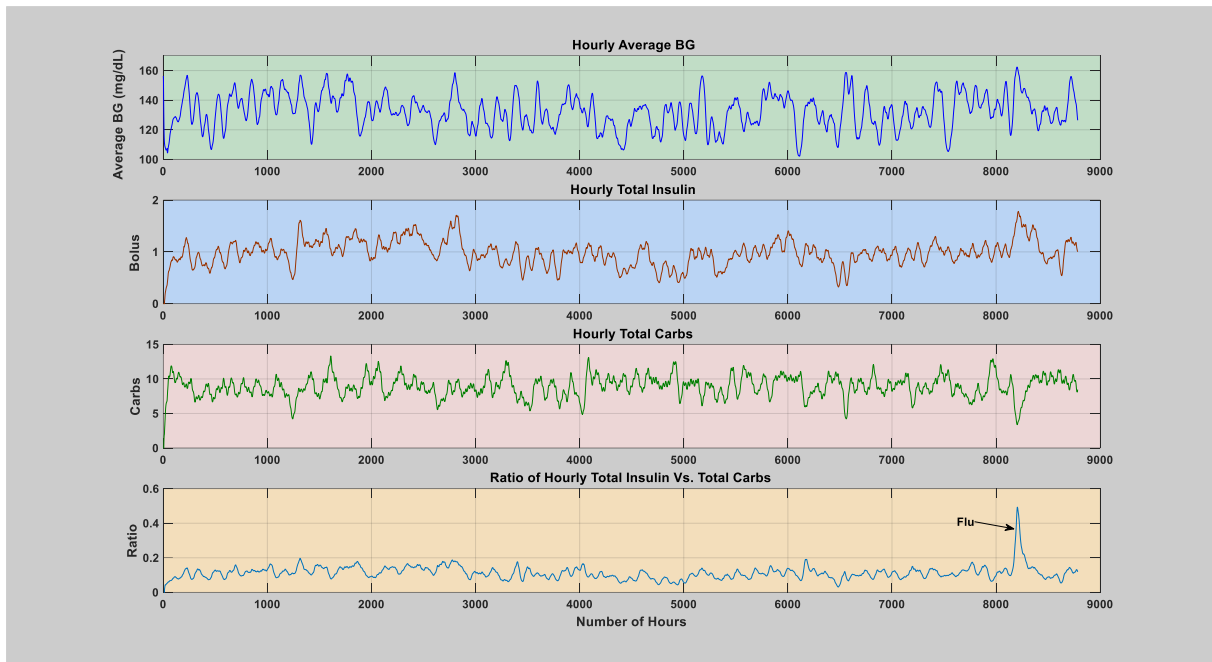
### 5.3.1.2.1 Daily and Hourly Analysis

The analysis of the two timeframes; daily and hourly, reveals the infection-triggered deviations from the normal process of blood glucose dynamics, as can be seen from **Figure 13-14** and more figures can be found in *Appendix 3 of Paper 1* [259]. During normal circumstances, insulin is administered on pre, during, and post-meal scenarios following the amount of carbohydrate consumed and in certain cases, correction insulin is also administered depending on the resultant blood glucose levels. However, this normal process is violated and seems no longer valid throughout the infection episodes [259]. This infection triggered deviation is typically manifested in the values of the insulin-to-carbohydrate ratio. As can be seen from the figures, the values of the ratio are sky-rocketed depicting the abnormal infection triggered situation, where the patient is forced to inject much higher insulin irrespective of consuming carbohydrates. In this situation, much of the insulin intake is correction insulin that is directed towards regulating blood glucose levels, which are disturbed due to infection incidences. As per the findings, the values of the ratio were elevated above the normal values (0.05-0.2), and reach between 0.25 and 0.5 depending on the individual. These findings demonstrate the substantial impact of infection episodes on blood glucose dynamics as compared to the other patient controllable and uncontrollable parameters. The analysis also reveals that other infection episodes such as fever-free mild and light cold as reported by the individual patients have a minor impact on the blood glucose dynamics, particularly the light cold.



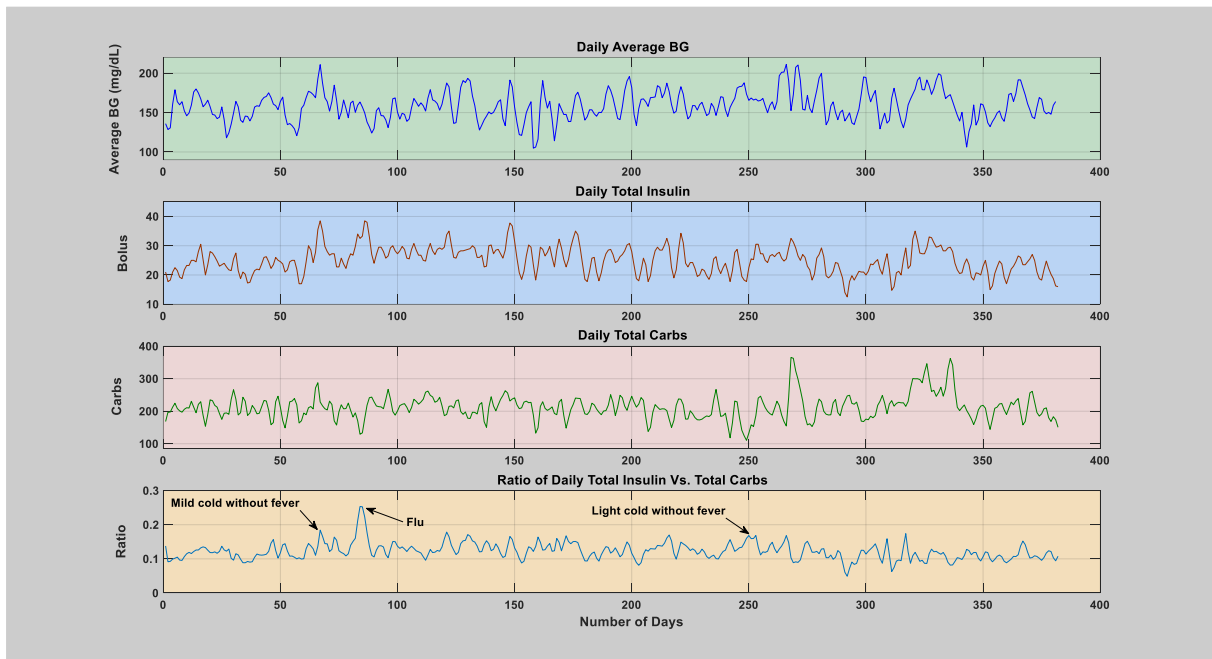
a) A daily plot of the patient-year, source [259], Figure 6.



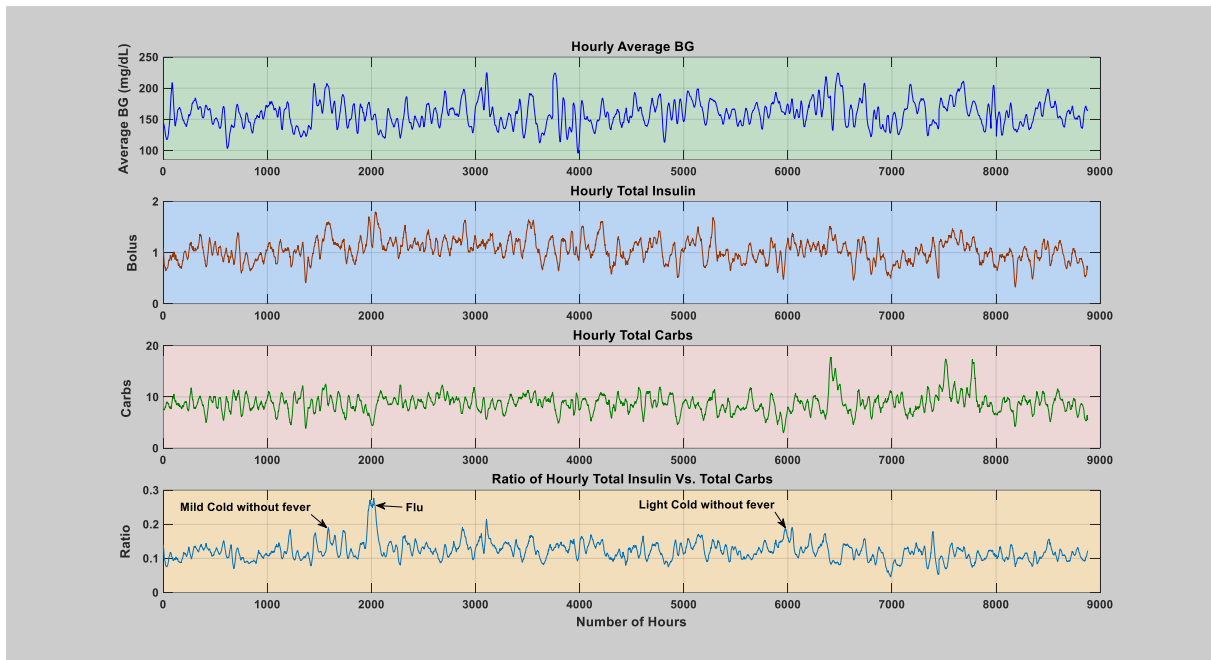


b) An hourly plot of the patient-year, source [259], Figure 7.

**Figure 13:** A particular patient-year with influenza (flu) incidence within the first week of December, and depicts blood glucose levels, amount of carbohydrate consumption and insulin injections, and fluctuations of the insulin-to-carbohydrate ratio. Figure (a) illustrates fluctuations within the daily timeframe. Figure (b) illustrates fluctuations within the hourly timeframe.



a) A daily plot of the patient-year, source [259], Figure 10.



b) An hourly plot of the patient-year, source [259] Appendix 3, Figure 4.

**Figure 14:** A particular patient-year with fever-free mild and light cold in the first week of August and mid-February respectively, and influenza (flu) incidence in mid-August. The figures depict blood glucose levels, amount of carbohydrate consumption and insulin injections, and fluctuations of the insulin-to-carbohydrate ratio. Figure (a) illustrates fluctuations within the daily timeframe. Figure (b) illustrates fluctuations within the hourly timeframe.

### 5.3.1.2.2 Comparison of Pre, Infection, and Post-infection Weeks

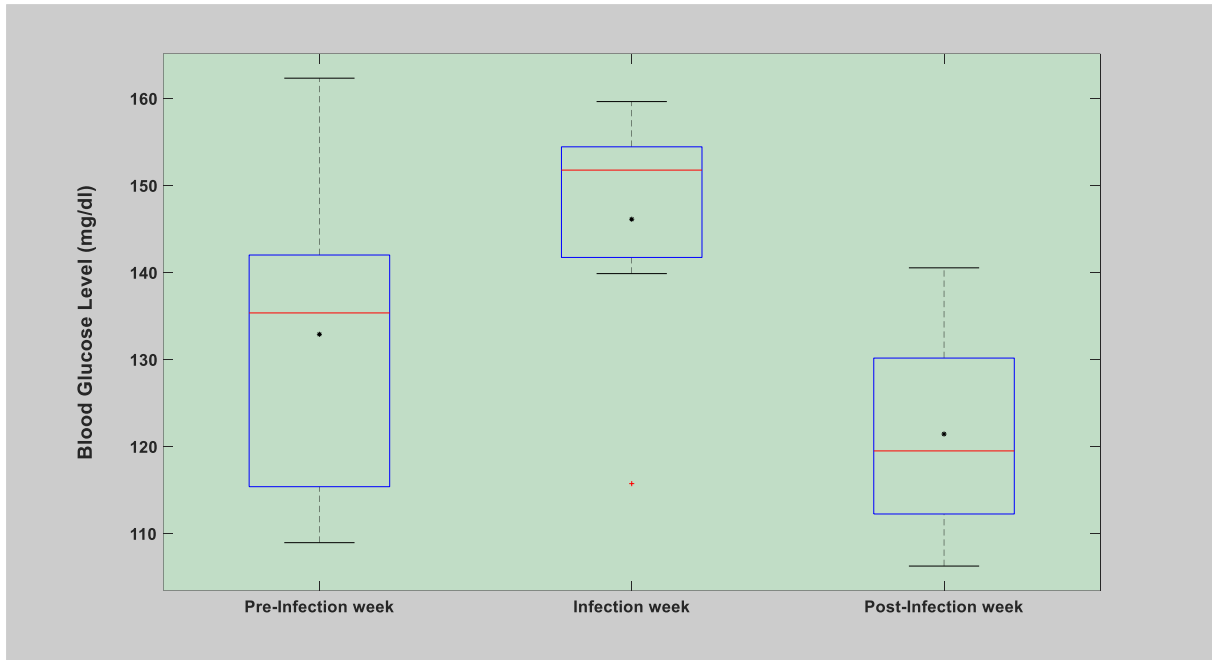
Each week's daily average blood glucose values, the sum of insulin injection, and the sum of carbohydrate consumption were compared within the timespan of the pre-infection week, infection week, and, post-infection weeks. The overall comparison of the infection triggered deviations from the pre-infection and post-infection weeks were depicted with a statistical boxplot incorporating the key parameters. The analysis among other things aims to quantify the infection-triggered deviations compared to the normal pre-infection and post-infection weeks. As per the findings, it can be seen in **Table 10** and **Figures 15-17**, the infection week's blood glucose levels remain elevated even if the individual keeps injecting much more insulin with a smaller amount of carbohydrate consumption. These phenomena violated the normal process of blood glucose dynamics and illustrate the influence of infection (flu) on blood glucose dynamics, which could be associated with the result of the action of infection-triggered glucose production and insulin resistance that develops within the body [152; 159; 259]. Under normal circumstances, when the patient injects much higher insulin and consumes much-reduced carbohydrate, these phenomena push the state of the blood glucose dynamics into the hypoglycemia region. However, this doesn't happen after infection onset, and this could be mainly due to the added effect of infection triggered glucose production and insulin resistance that occurred within the body [259].

**Table 10:** Average and standard deviation of blood glucose levels, the sum of insulin doses (bolus), and the sum of carbohydrates intake within pre-infection, infection, and post-infection weeks, source [259], Table 5.

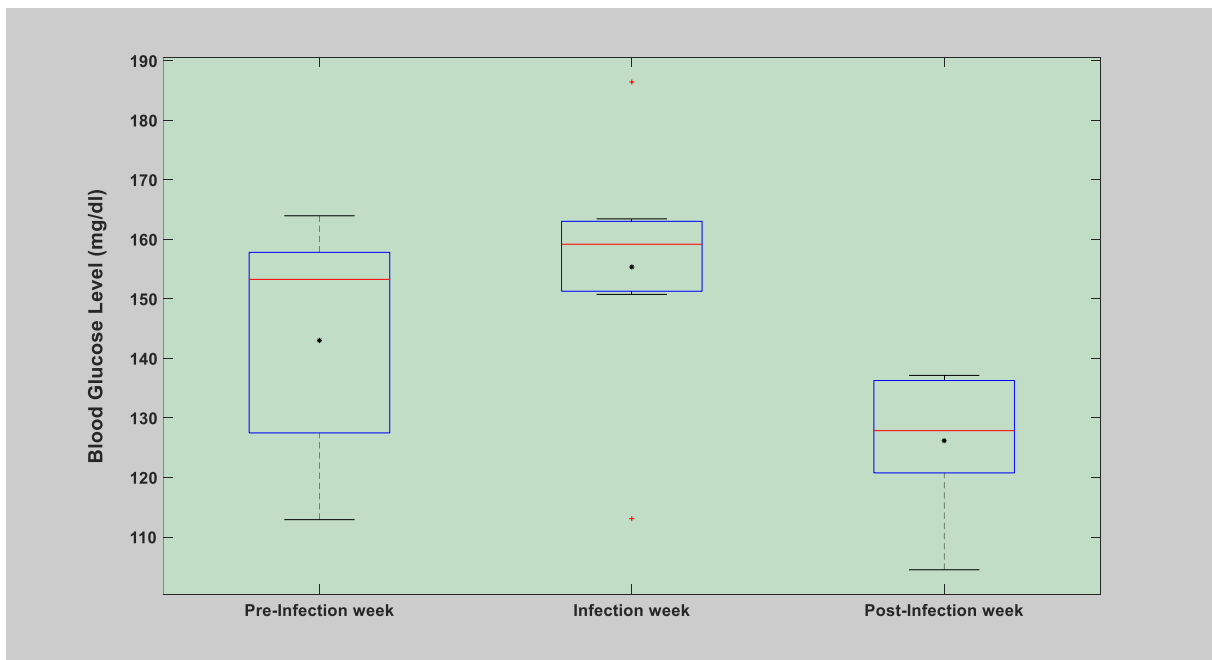
<b>Influenza (flu) incidence in the first patient-year</b>			
<b>Parameter</b>	<b>Pre-infection week</b> ( Mean (SD))	<b>Infection week</b> ( Mean (SD))	<b>Post-infection week</b> ( Mean (SD))
<b>BG (mg/dL)</b>	130.74 (16.89)	141.95 (14.37)	119.16 (7.39)
<b>Total insulin (bolus)</b>	23.39 (4.91)	35.30 (6.11)	21.32 (4.61)
<b>Carbohydrate (grams)</b>	241.11 (57.27)	178.80 (65.69)	241.18 (37.63)
<b>Influenza (flu) incidence in the second patient-year</b>			
<b>BG (mg/dL)</b>	143.01 (19.53)	155.36 (21.99)	126.17 (11.70)
<b>Total insulin (bolus)</b>	28.07 ( 8.85)	41.07 (9.44)	25.36 (6.93)
<b>Carbohydrate (grams)</b>	190.14 (43.93)	161.14 (58.43)	214.57 (34.66)
<b>Influenza (flu) incidence in the third patient-year</b>			
<b>BG (mg/dL)</b>	136.93 (18.58)	144.12 (20.30)	134.18 (11.96)
<b>Total insulin (bolus)</b>	20.08 ( 5.44)	31.50 (10.84)	22.83 (3.86)
<b>Carbohydrate (grams)</b>	178.0 (45.87)	144.83 (37.63)	195.83 (42.59)
<b>Influenza (flu) incidence in the fourth patient-year</b>			
<b>BG (mg/dL)</b>	157.74 (31.12)	161.34 (19.88)	138.57 (19.83)
<b>Total insulin (bolus)</b>	24.43 (5.26)	32.14 (7.01)	29.29 (5.22)
<b>Carbohydrate (grams)</b>	199.06 (53.45)	167.04 (44.94)	226.07 (18.23)
<b>Influenza (flu) incidence in the fifth patient-year</b>			
<b>BG (mg/dL)</b>	135.21 (14.58)	139.88 (15.54)	122.87 (14.49)
<b>Insulin (bolus)</b>	32.80 (4.59)	40.37 (8.31)	33.36 (7.94)
<b>Insulin (basal)</b>	19.20 (1.21)	20.42 (2.06)	18.68 (1.56)
<b>Total Insulin</b>	52.33 (5.14)	61.21 (8.26)	52.46 (8.47)

#### 5.3.1.2.2.1 Blood glucose levels

Blood glucose levels were higher for a longer period throughout the infection period in all the infection weeks. As described earlier, this phenomenon could be linked with the production of excess glucose from the liver as a result of the infection episode [152]. In this regard, the infection week's average blood glucose levels were higher as compared to the pre-infection and post-infection weeks' average blood glucose levels. Numerically, as shown in **Figures 15** and **Table 10**, the average blood glucose levels were higher in all the infection weeks with a percentage of 8.57% and 19.12%, 8.63% and 23.13%, 7.26% and 7.41%, 2.28 and 16.43%, 3.45% and 13.84% as compared to the pre-infection and post-infection weeks' average blood glucose values respectively. More figures can be found in *Appendix 1 of Paper 1* [259].



a) Box-plot of a particular patient-year, source [259], Figure 3.



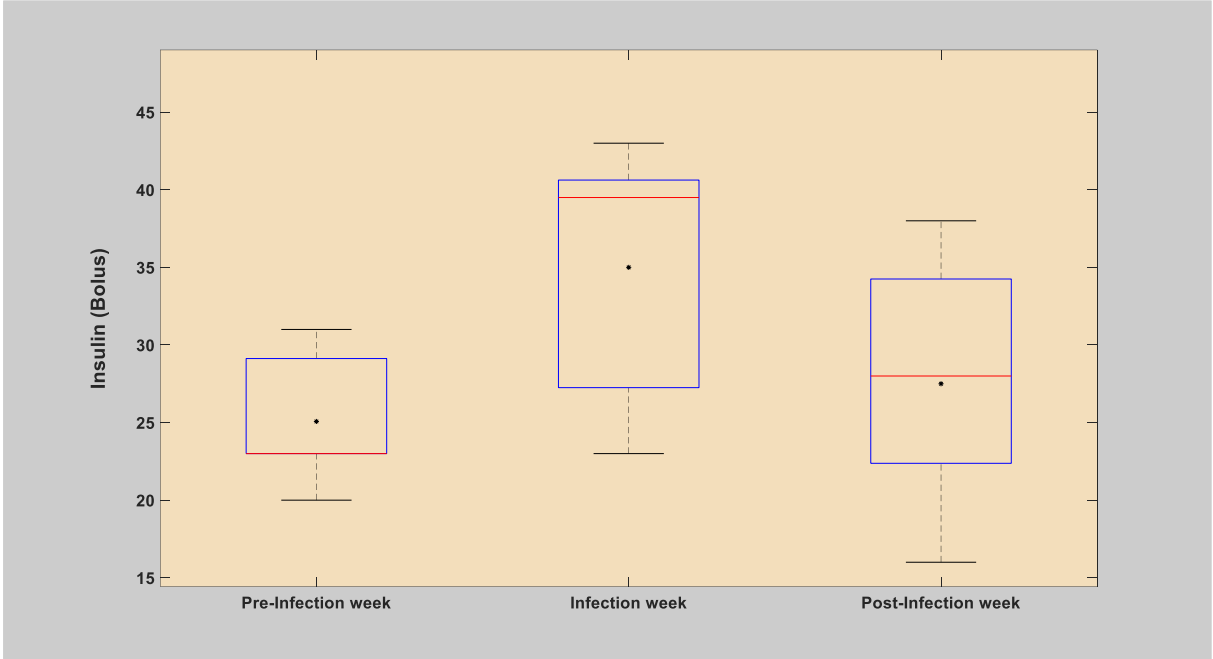
b) Box-plot of a particular patient-year, source [259] Appendix 1, Figure 2.

**Figure 15:** Comparison of blood glucose levels during pre-infection week, infection week, and post-infection week. As can be seen, the blood glucose is elevated during the infection week as compared to the pre and post-infection weeks.

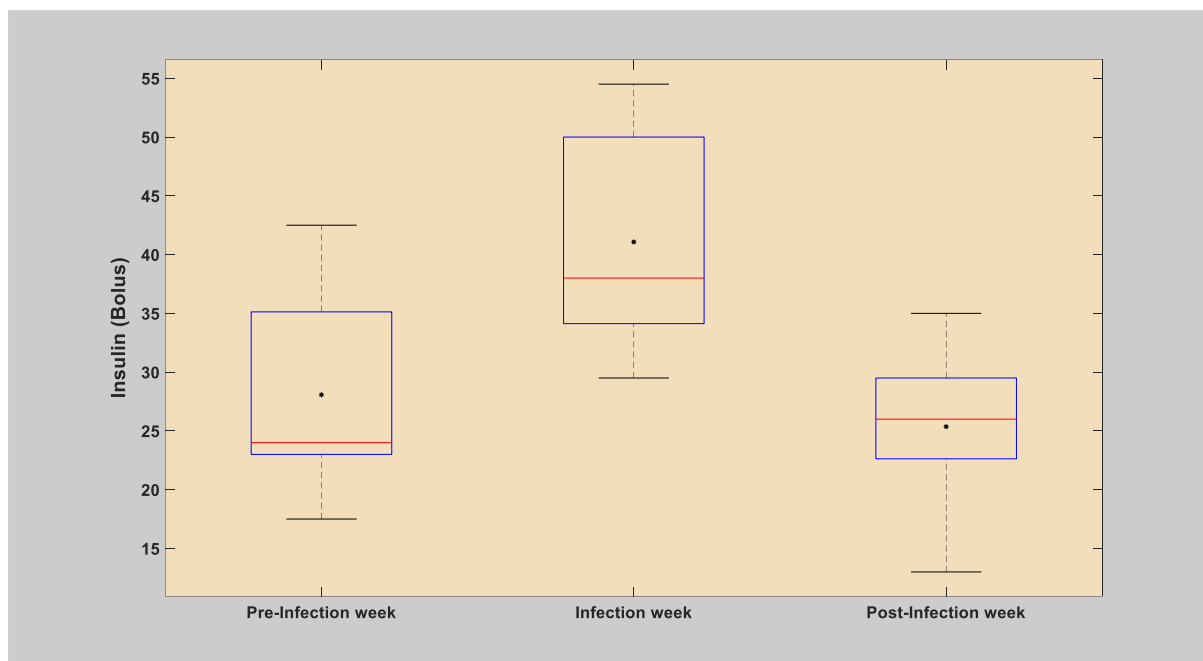
#### 5.3.1.2.2.2 Insulin intake

Insulin (bolus) administered during the infection weeks was much higher as compared to the pre and post-infection weeks. As described earlier, the enhanced insulin intake could be directly

associated with the insulin resistance that develops within the body following infection onset mainly due to the action of the counter-regulatory hormones [152; 159]. In this regard, the infection weeks' insulin injections were much higher as compared to the pre-infection and post-infection weeks' insulin injections. Numerically, as shown in **Figures 16** and **Table 10**, the overall insulin injections were significantly higher in all the infection weeks with a percentage of 50.93% and 65.59%, 46.31% and 61.94%, 56.87% and 37.98%, 31.56% and 9.7%, 23.08% and 21.01% as compared to the pre-infection and post-infection week's overall insulin injections respectively. More figures can be found in *Appendix 1 of Paper 1* [259].



a) Box-plot of a particular patient-year, source [259], Figure 4.



b) Box-plot of a particular patient-year, source [259] Appendix 1, Figure 2.

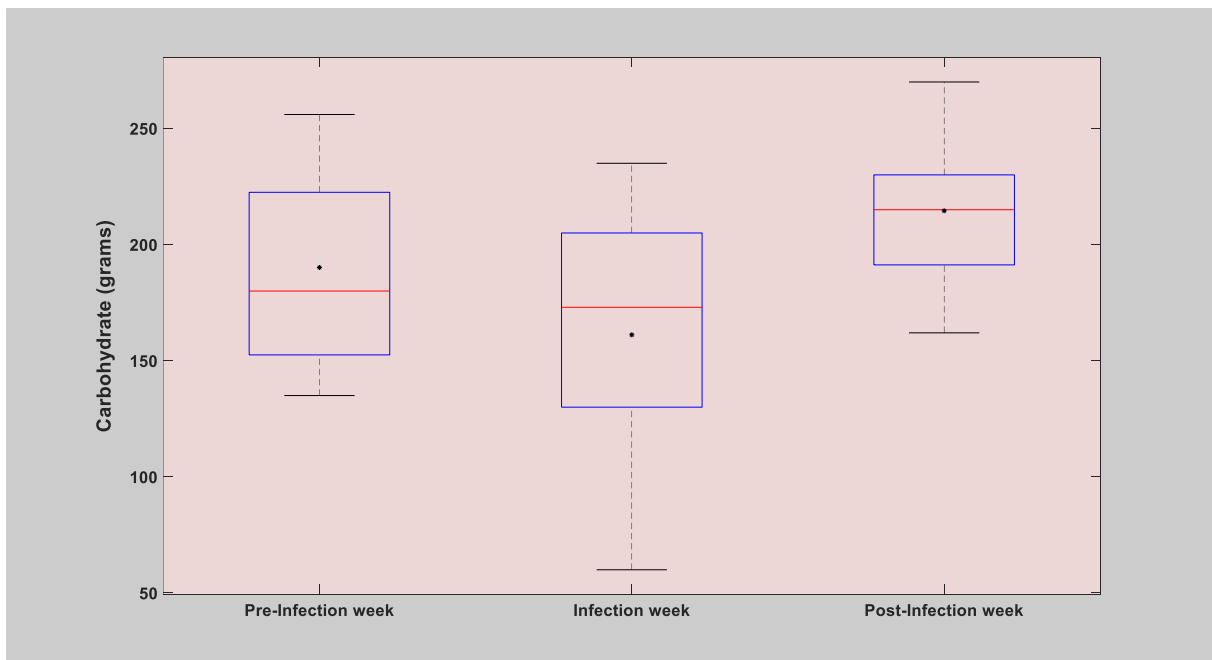
**Figure 16:** Comparison of insulin (bolus) injections during the pre-infection week, infection week, and post-infection week. As can be seen, the amount of insulin (bolus) intake is elevated during the infection week as compared to the pre and post-infection weeks.

#### 5.3.1.2.2.3 Carbohydrate consumption

Carbohydrate (grams) consumed during the infection weeks were much lower as compared to the pre-infection and post-infection weeks' consumption. This phenomenon could be linked either with the patient action to avoid a further crisis of hyperglycemia or lack of appetite as a result of the infection [152; 259]. In this regard, the infection weeks' overall carbohydrate consumptions were much lower as compared to the pre-infection and post-infection weeks' overall carbohydrate consumptions. Numerically, as shown in **Figures 17** and **Table 10**, the overall carbohydrate consumptions were significantly lower in all the infection weeks with a percentage of 25.84% and 25.87%, 15.25% and 24.90%, 18.63% and 26.04%, 16.09% and 35.34% as compared to the pre-infection and post-infection week's overall carbohydrate consumptions respectively. More figures can be found in *Appendix 1 of Paper 1* [259].



a) Box-plot of a particular patient-year, source [259], Figure 5.



b) Box-plot of a particular patient-year, source [259] Appendix 1, Figure 2.

**Figure 17:** Comparison of carbohydrate consumptions (grams) during the pre-infection week, infection week, and post-infection week. As can be seen, the amount of Carbohydrate (grams) intake is significantly lower during the infection week as compared to the pre and post-infection weeks.

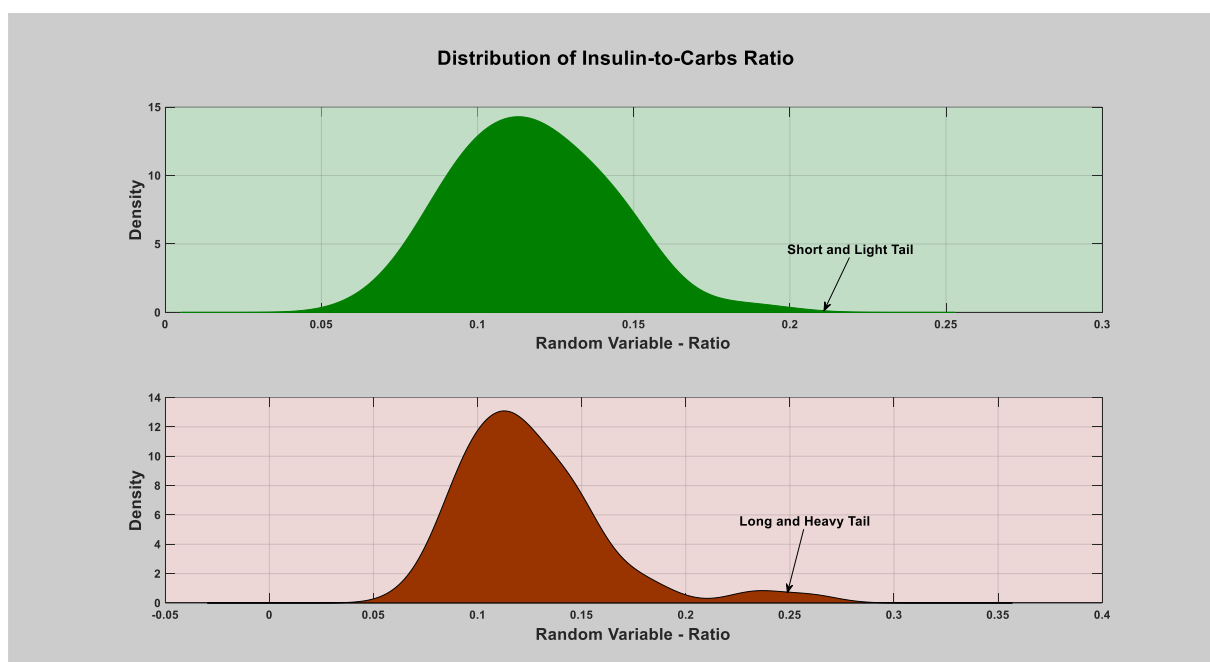
#### 5.3.1.2.2.4 Insulin-to-carbohydrate ratio

The ratio of the insulin-to-carbohydrate reflects the combined phenomena of insulin injections and carbohydrate consumptions. On an individual basis, it depicts the required insulin intake to

offset the action of a single gram of carbohydrate consumed. As described above, the dramatic shift in insulin injections and carbohydrate consumptions following infection episodes is also reflected in the required values of the ratio. During normal circumstances, the value usually fluctuates between 0.05 and 0.2, see **Figure 11-12**. However, as clearly demonstrated in **Figure 13-14**, the values of the insulin-to-carbohydrate ratio significantly raised to a higher value to accompany the individual's requirements following infection onset [259]. In this regard, the infection weeks' values of insulin-to-carbohydrate ratio were much higher as compared to the pre-infection and post-infection weeks' values of the ratio. Numerically, as shown in **Table 10**, the values of insulin-to-carbohydrate ratio were significantly higher in all the infection weeks with a percentage of 125.84%, 144.43%, 93.75%, 70.84% as compared to the pre-infection and post-infection week's normal values. More figures can be found in *Appendix 1 of Paper 1* [259].

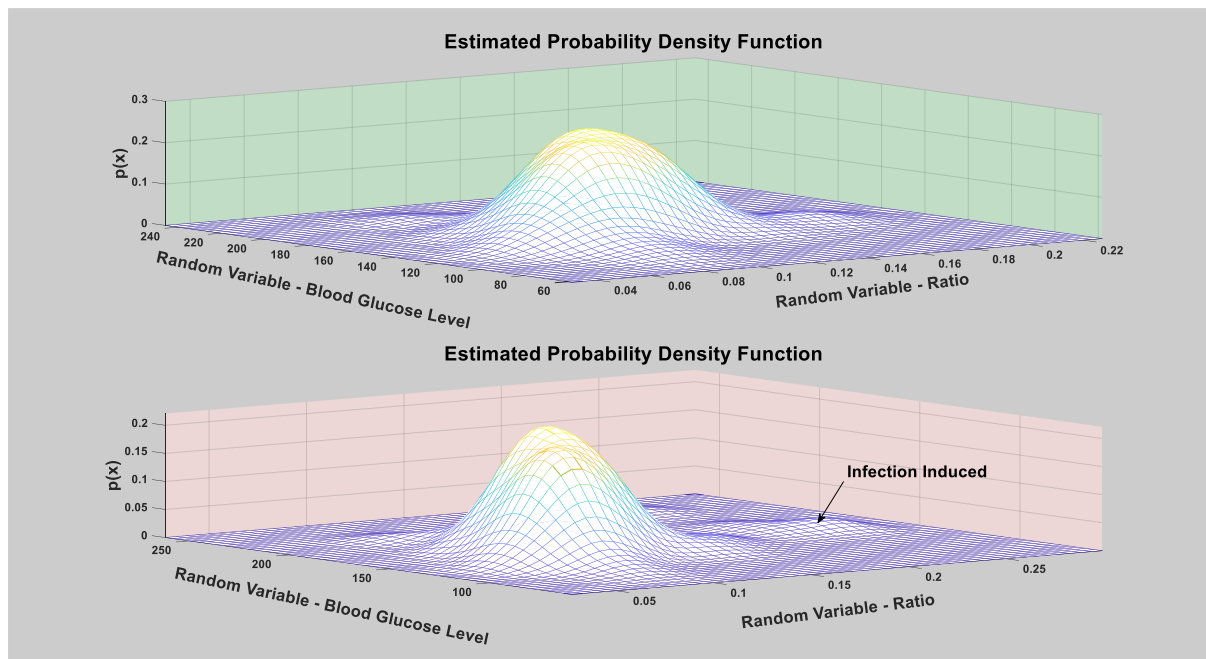
### 5.3.2 Data Distribution

The individual self-management practice is usually related to the individual's diet habits, physical activity or exercise sessions, medication intake, and others. To study the influence of infection on blood glucose dynamics along with the nature and shape of the anomalies generated, the individual's data distribution was estimated using a kernel density estimator and compared with the distribution of the same data after removing the reported infection days. The distribution estimation was conducted for each patient-years with infection episode considering univariate (insulin-to-carbohydrate ratio) and bivariate (insulin-to-carbohydrate ratio versus blood glucose levels) variables. As can be seen from **Figures 18**, data distribution of the infection period generates a longer tail and sparse density, and more figures can be found in *Appendix 1 of Paper 1* [259]. These deviations could correlate with the strength of the individual's immunity, type of infection, and severity (pathogens), and hormones involved [152; 259].



a) Kernel density estimation of daily total insulin (bolus)-to-carbohydrate ratio, source [259] Appendix 1, Figure 12 (a).





b) Kernel density estimation of daily average blood glucose levels and total insulin (bolus)-to-carbohydrate ratio, source [259] Appendix 1, Figure 13 (a).

**Figure 18:** Kernel density estimation of a particular patient-year. Figure (a) depicts the univariate kernel estimation. Figure (b) depicts the bivariate kernel estimation.

## 5.4 Characterizing the observed phenomenon

This section characterizes the observed phenomenon, describes the optimal predictive parameters, and further tries to generalize the findings beyond people with type 1 diabetes into other types of diabetes. As per the findings, infection often triggers a series of shifts in the blood glucose dynamics and mainly with respect to blood glucose levels, insulin and carbohydrate intake, and the ratio of insulin-to-carbohydrate [259]. However, the literature suggests that the extent and degree of abnormalities associated with these infection triggered phenomena often correlate with the strength of the individual's immunity, the category of the pathogens, and the type of hormones [238]. The underlying cause of the hyperglycemia crises upon infection onset is the added effect of patient uncontrollable parameters in addition to the regular patient controllable parameters [152], as given in *Equation 1* below. The effect of the patient's uncontrollable parameter, i.e. counter-regulatory hormone, triggered by infection as a result of stress, is inhibiting insulin production and stimulating glucose production, which is the body mechanism to cope up with infection [159]. Hence, this patient group usually struggles to maintain euglycemia. However, during the regular days, hyperglycemia is mostly triggered by a patient controllable parameter (diet) and usually managed by controlling the key patient controllable parameters, e.g. limiting the amount of carbohydrate intake, injecting proper amount of insulin, and doing adjusted physical activity or exercise session [259]. Generally, this infection triggered phenomena can be characterized by an equation representing the relationship between blood glucose levels and other key parameters that disturb the dynamics along with the added effect of infection onset;

$$\mathbb{B}\mathbb{G}_t = \mathbb{C}\mathbb{H}_t + \mathbb{C}\mathbb{R}\mathbb{H}_t - \varphi \mathbb{I}\mathbb{N}_t - \mathbb{P}\mathbb{A}_t, \quad (1)$$

Where the factor  $\varphi$  depicts the body's level of insulin sensitivity,  $\mathbb{B}\mathbb{G}$  is the average blood glucose values,  $\mathbb{C}\mathbb{H}$  is the sum of carbohydrate intake in grams,  $\mathbb{I}\mathbb{N}$  is the sum of insulin dose,  $\mathbb{P}\mathbb{A}$  is the aggregate level of physical activity or exercises load, and  $\mathbb{C}\mathbb{R}\mathbb{H}$  is the amount of extra glucose generated as a result of the effect of counter-regulatory hormones, i.e. adrenalin and cortisol [259]. As depicted in the equation, infection triggers  $\mathbb{C}\mathbb{R}\mathbb{H}$  induced glucose and diminishes the body's level of insulin sensitivity ( $\varphi$ ), thereby resulting in hyperglycemia despite consuming regular diets and requiring more insulin doses to compensate for the reduction in body's level of sensitivity. Further, to compensate for the effect of  $\mathbb{C}\mathbb{R}\mathbb{H}$ s, a reduction in the amount of carbohydrate consumption is expected [259]. These phenomena are manifested in the analysis of all the individual infection episodes. Hence, the event marker for infection episode can be designated as the occurrence of a dynamic shift from the typical operating point of blood glucose dynamics as manifested by the shift sustained on the ratio of the insulin-to-carbohydrate. Considering the nature of blood glucose metabolism and infection along with the body's physiological reactions, these phenomena can be expected in other types of diabetes, however, the degrees of abnormalities can be varied in between. Based on the evidence presented, infection triggered deviations are manifested on the key parameters, and blood glucose levels, insulin doses, carbohydrate intake, and the ratio of insulin-to-carbohydrate can be taken into account as input features for realizing a personalized health model to detect infection episode in an individual with type 1 diabetes [259]. Generally, the result proved the potential of the presented data to develop the proposed personalized health model-based digital infectious disease detection system. In this regard, the analysis among other things demonstrated the informational value and discriminative power of the data, which obviously can reduce false alarms that could emanate from confounding parameters such as emotional stress.

## 5.5 Limitation

Considering the complexity of blood glucose dynamics, it is obvious that this study could benefit from a larger sample size and could further strengthen the conclusion. However, due to the difficulty of getting accurate and rich datasets containing infection episodes, the study was carried out on the presented datasets. Actually, the inclusion criteria were stiff given the fact that the study expects rich datasets that are accurate enough as well as are required at least to contain one event of infection, which makes the challenge far more difficult. Moreover, the study could also be far more beneficial if the dataset contains physical activity data to further strengthen the conclusion. However, taking into account the nature of blood glucose metabolism, which has more or less common dynamics among each individual, the presented results can conform and be generalized among these patient groups.

## 5.6 Knowledge Summary

In the context of the original research question, the following section presents the added knowledge within the scope of the presented results.

*What do we know about the topic already?*

- The association between infection onset and hyperglycemia episode has long been known.
- Previously, the idea of using blood glucose levels to detect outbreaks and surveillance purposes has been suggested and presented in the literature.
- However, none of these studies characterized and numerically quantifies the impact of infection on the key parameters of blood glucose dynamics and pinpoint optimal parameters with high accuracy.

*What does this chapter add to our knowledge?*

- To the best of my knowledge, this is the first study that empirically analyzes self-recorded data and numerically quantifies the effect of infection episodes on key parameters of blood glucose dynamics among people with type 1 diabetes.
- Infection significantly alters the operating point of the individual's blood glucose dynamics.
- Infection onset triggers elevated blood glucose levels regardless of higher insulin doses and lesser carbohydrate consumptions. This event marker designates the occurrence of a dynamic shift from the typical operating point of blood glucose dynamics as manifested by the shift sustained on the ratio of the insulin-to-carbohydrate following infection onset.
- Discovered and presented a unique parameter, i.e. the ratio of insulin-to-carbohydrate, with excellent informational value and discriminative power to minimize the false alarm.
- Characterization of the blood glucose dynamics reveals optimal parameters for realizing a personalized health model to detect infection onset in an individual with type 1 diabetes.

## **5.7 Chapter Summary**

This chapter presented and discussed results related to the characterization of the impact of infection onset on the individual's blood glucose dynamics. Infection onset brought a substantial shift in the typical operating point of the individual's blood glucose dynamics. As per the findings, infection onset triggers elevated blood glucose levels for a prolonged duration regardless of the individual actions accompanied by higher insulin doses and lesser carbohydrate consumptions to regulate and control the hyperglycemia crisis triggered by the incident. The event marker designates the occurrence of a dynamic shift from the typical operating point of blood glucose dynamics as manifested by the shift sustained on the ratio of the insulin-to-carbohydrate following the infection onset. Hence, this characterization reveals the potential of the key parameters of blood glucose dynamics such as blood glucose level, carbohydrate, insulin, and the ratio of insulin-to-carbohydrate information for realizing the proposed personalized health model towards detecting infection onset in an individual with type 1 diabetes. In this regard, a personalized health model that utilizes these input features can be developed based on either a prediction model or an anomaly and novelty detection methods. As far as my knowledge is concerned, this is the first study that empirically analyses self-recorded

data and numerically quantifies the effect of infection episodes on the key parameters of blood glucose dynamics among people with type 1 diabetes. In conclusion, these findings provide optimal parameters of blood glucose dynamics to support the effort towards realizing a personalized health model, and however, additional large-scale studies might be needed to further strengthen the conclusion.

## 6 Designing a Personalized Health Model

*Synopsis:* This chapter puts forward a solution for realizing the proposed personalized health model, which can describe the individual's blood glucose dynamics from the self-recorded data. The aim was to devise a model for tracking the individual's health status and automatic detection of health changes. The first half of the chapter provides an overview of the input features used in modeling, and the subsequent section provides the model's performance derived from using a bivariate input feature and an alternative solution using a univariate input feature. Finally, it provides a comparative analysis and concluding remark. This chapter provides answers to the third research question (Q3).

### 6.1 Introduction

This chapter focuses on the approaches devised to realize the proposed personalized health model for effective tracking of the individual's health status and automatic detection of infection onset among people with type 1 diabetes. To the best of my knowledge, this is the first attempt towards realizing a personalized health model to capture infection episodes among people with type 1 diabetes using self-recorded data. Apart from this, it is better to note that there was a single previous attempt to capture the stress state among type 1 diabetes under ambulatory settings [76]. The proposed model utilizes those optimal features suggested in *chapter five*, specifically blood glucose levels and the ratio of insulin-to-carbohydrate. As demonstrated in *chapter five*, those input features were selected based on their informational values and discriminative power compared to the baseline normal data [259]. As discussed in *chapter four*, the realization of the proposed personalized health model can be approached either using a predictive model of blood glucose dynamics or anomaly (novelty) detection methods [259]. In the literature, the state-of-the-art blood glucose prediction models are often described to be accurate in a limited setting and prediction horizon [173; 256]. The prediction performance of those models often degrades beyond a 30-minute prediction horizon, which becomes a bottleneck to effectively be used in the proposed settings, i.e. changes within hours and days [165; 256]. An ideal blood glucose predictor requires to incorporate a vast majority of input features that affect blood glucose dynamics, which is difficult to achieve in a practical setting [256]. In this regard, the requirement calls for a personalized health model that can capture the dynamics and be able to detect abnormalities within the context of limited data settings. Consequently, as an option to these drawbacks of prediction models, an anomaly (novelty) detection method was devised to realize the proposed personalized health model. This option can be formulated as supervised, i.e. multi-class classification, semi-supervised, i.e. one-class classification, or unsupervised method [4; 51; 183]. Implementing the proposed model as supervised can be challenging and impractical given the effort required to acquire and labeling the normal and anomaly class [124; 217]. Practically there are several issues involved with the characterization and demarcation of the infection triggered anomaly class boundaries in comparison to the normal class [51; 258]:

- ***Characterizing the effect of different pathogens:*** Demarcating pathogen-specific class boundary requires understanding the effect of each pathogen on blood glucose dynamics, and this calls for gathering pathogen-specific data from a large group of

participants. Such data collection is expensive and time-consuming if not impossible [258].

- ***Effect of the same pathogens among different individuals:*** Apart from the inter-pathogen class boundaries, there still exists variations among the intra-pathogen class boundary that emanates from the difference in individuals' infection resistance. This poses a serious challenge and further complicates the demarcation task [258].
- ***Lack of available data source:*** Even if the above demarcation tasks are accomplished, there still exist challenges related to an imbalanced class problem. Gathering a balanced dataset containing both the normal and abnormal class from the individual participant require waiting for the individual to get infected on multiple occasions and this further poses an implementation challenge [258].

Considering these stated challenges, the other design option is to use a semi-supervised method, i.e. one-class classifier, that relies on learning the regular situations and able to detecting nonconformities with the reference description [51; 113; 217; 219]. This strategy alleviates the above-stated challenges related to characterization and lack of dataset. Apart from this, another design option to the stated challenges could be an unsupervised method that relies on the entire dataset to define and find anomalies within the data [51]. In this regard, this chapter presents the performance of semi-supervised (one-class classifier) and unsupervised methods for the proposed personalized health model. These models utilize univariate (insulin-to-carbohydrate ratio) and bivariate (blood glucose levels versus insulin-to-carbohydrate ratio) as input features. The performance achieved from these models and input features is compared along with the model's computational time and necessary sample size to generate acceptable performance. The first half of the results presented in this chapter, i.e. bivariate inputs, are part of the findings presented in *Paper 3* [258].

## 6.2 Input features

To better understand the input features, it is necessary to comprehend the characteristics of blood glucose dynamics depending on the three key parameters; insulin, carbohydrate, and physical activity or exercise load, that disturb blood glucose levels. Depending on these parameters, the state of wellness of blood glucose dynamics among an individual with type 1 diabetes can be grouped into any of the four quadrants; *carbohydrate action as quadrant 1*, *physical activity action as quadrant 2*, *insulin action as quadrant 3*, and *metabolic change due to infection onset as quadrant 4*.

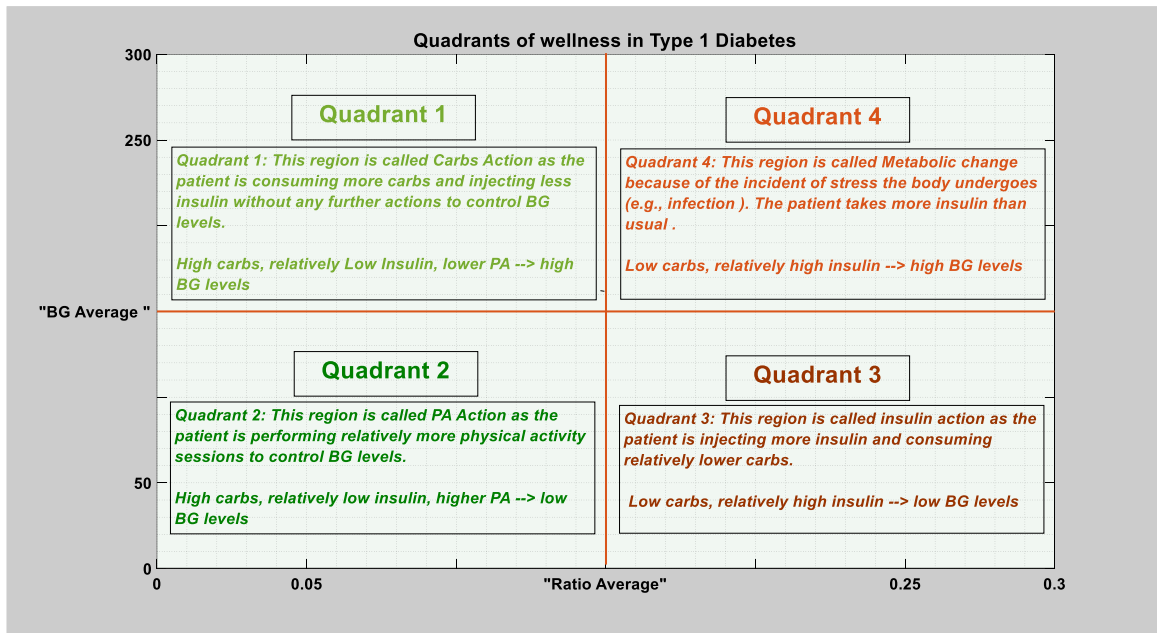
***Carbohydrate action as quadrant 1:*** As shown in **Figure 19**, this quadrant covers the normal region of the blood glucose dynamics, where the patient consumes more carbohydrate without taking required actions, i.e. administering insulin and performing balanced physical activity or exercise sessions, to control or regulate the blood glucose levels. Particular examples of such a situation is a *holiday effect*, where individuals ingest too much carbohydrate during the holiday seasons. Datasets that lie within this region are sparser and less dense.

***Physical activity action as quadrant 2:*** As shown in **Figure 19**, this quadrant covers the normal region of the blood glucose dynamics, where the patient ingests more carbohydrates but favors

performing physical activity or exercise sessions than administering insulin to control the blood glucose levels. It is evident that though the patient ingests more carbohydrates, rigorous physical activity or exercise load could induce hypoglycemia. The sparsity of any particular points that fall within this region is dependent on the extent to which physical activity is used to replace the insulin needs. This sparsity happens since replacing insulin needs with physical activity or exercises session could create a very small value of insulin-to-carbohydrate ratio that fall outside the typical range of 0.05 to 0.2.

***Insulin action as quadrant 3:*** As shown in **Figure 19**, this quadrant also covers a normal region of the blood glucose dynamics, where the patient ingests less carbohydrate and administer more insulin to regulate the blood glucose levels. This is a very dangerous region of the dynamics, since taking more insulin without taking a proper meal could drive the blood glucose dynamics into a dangerous hypoglycemia state that could make the individuals end up being unconscious and sometimes death depending on the degree. The sparsity of the points that fall in this region is dependent on the degree of difference between carbohydrate intake and how much insulin is delivered for such an amount.

***Metabolic change due to infection onset as quadrant 4:*** As shown in **Figure 19**, this quadrant covers the abnormal region of the blood glucose dynamics, where the blood glucose levels remain elevated though the patient administers more insulin and ingests lower carbohydrate to control blood glucose levels. This is abnormal given the fact that these states of inputs should always be expected to put the individuals in quadrant 3 under normal circumstances, which is a dangerous hypoglycemia state. Yet, excessive production of glucose and insulin resistance as a result of infection prevented the blood glucose levels from going into a hypoglycemia state and, therefore, the individual's blood glucose levels remain high. This region only contains sparse data points that reflect the abnormal state of blood glucose dynamics as a result of health changes.



**Figure 19:** Quadrants of wellness in an individual with type 1 diabetes, source [258], Figure

7.

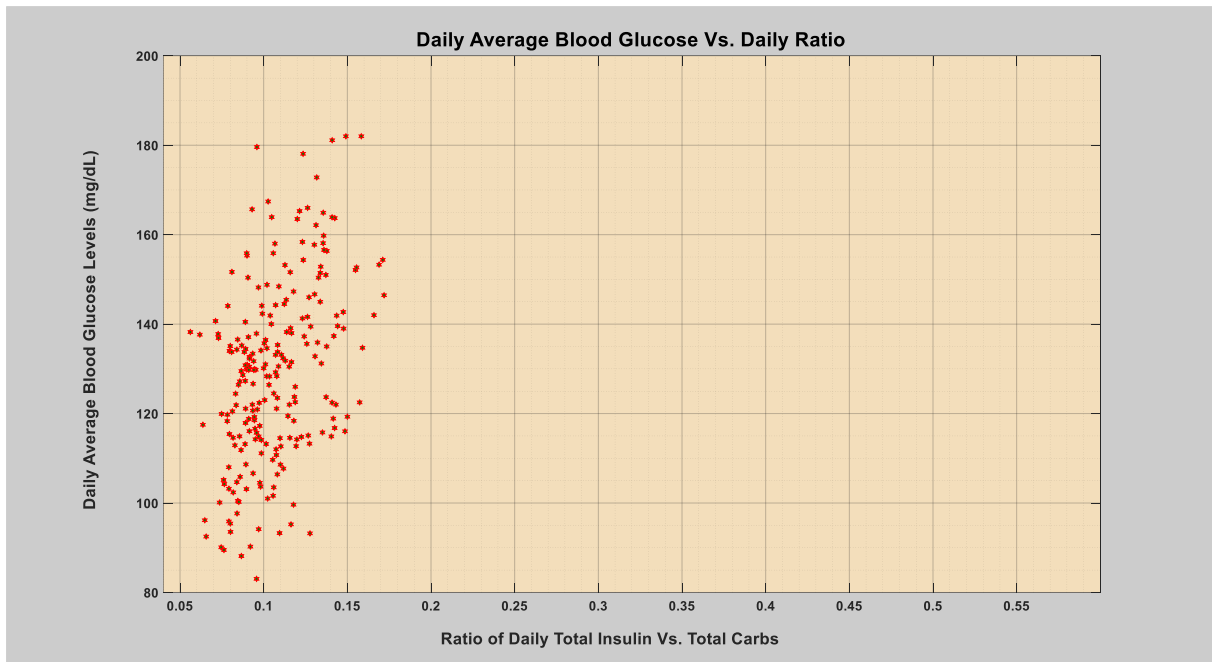
### 6.2.1 Description of input features

This section describes the characteristics of the input features as per the results presented in *chapter five*. As described above, the models' input features were *univariate input*, the ratio of insulin-to-carbohydrate, and *bivariate input*, blood glucose levels vs. the ratio of insulin to carbohydrate. The models' performance was compared based on *data type* - raw and smoothed data, *data granularities* - hourly and daily timeframes, and *data sample sizes* - one, two, three, and four months. The input features were smoothed via a moving average filter of 2 days (48 hours) window.

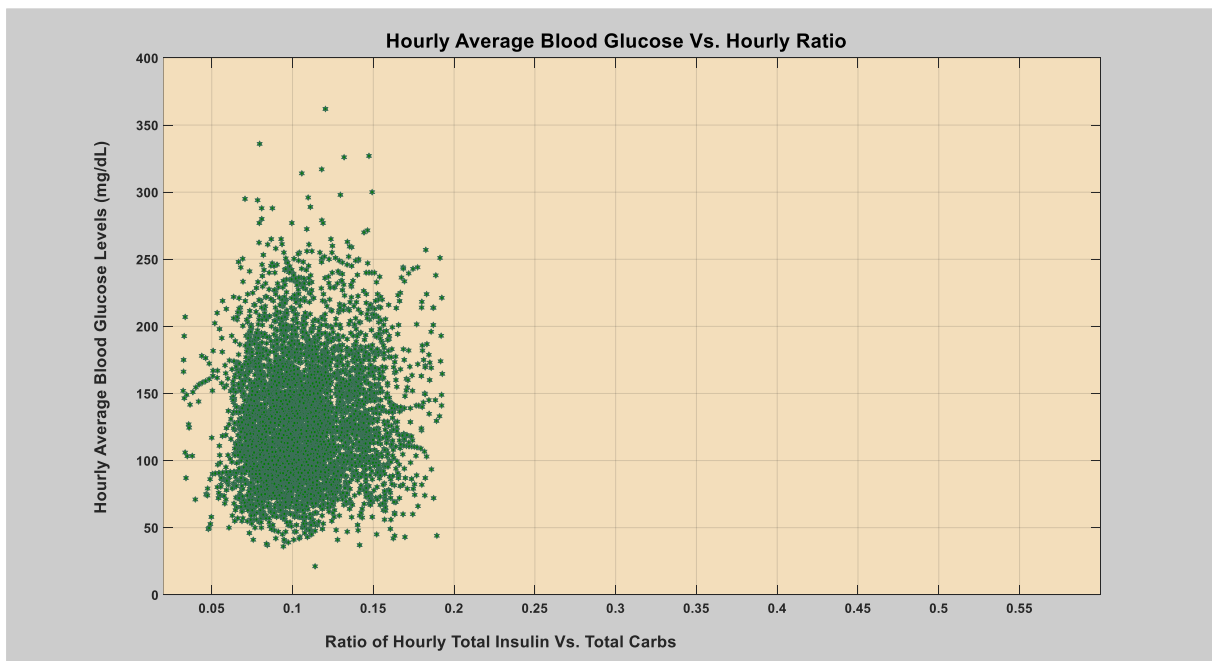
#### 6.2.1.1 Input features - Regular patient-years versus patient-years with infection

The patient years exhibit different characteristics depending on the presence or absence of infection as per the result presented in *chapter five*. Throughout the regular patient-years, the data happens to be bounded but containing a high-density region and some sparse regions with low density, as shown in **Figures 20-21**, and more figures can be found in *Appendix 2 of Paper 3* [258]. In all the figures, the x-axis representing the ratio happens to be bounded between 0.05 and 0.2 depicting the typical behavior of the insulin-to-carbohydrate ratio during the normal operations, as demonstrated in *chapter five*. In regard to the patient-years with infection, the data exhibits similar characteristics compared to the regular patient-years, except that in this data infection days happen to be outside of the typical range of the boundary, as shown in **Figure 22-23**, and more figures can be found in *Appendix 2 of Paper 3* [258]. In this regard, the task of modeling requires to describe those regular days and capture those novel or anomalous days with a minimum false alarm rate.



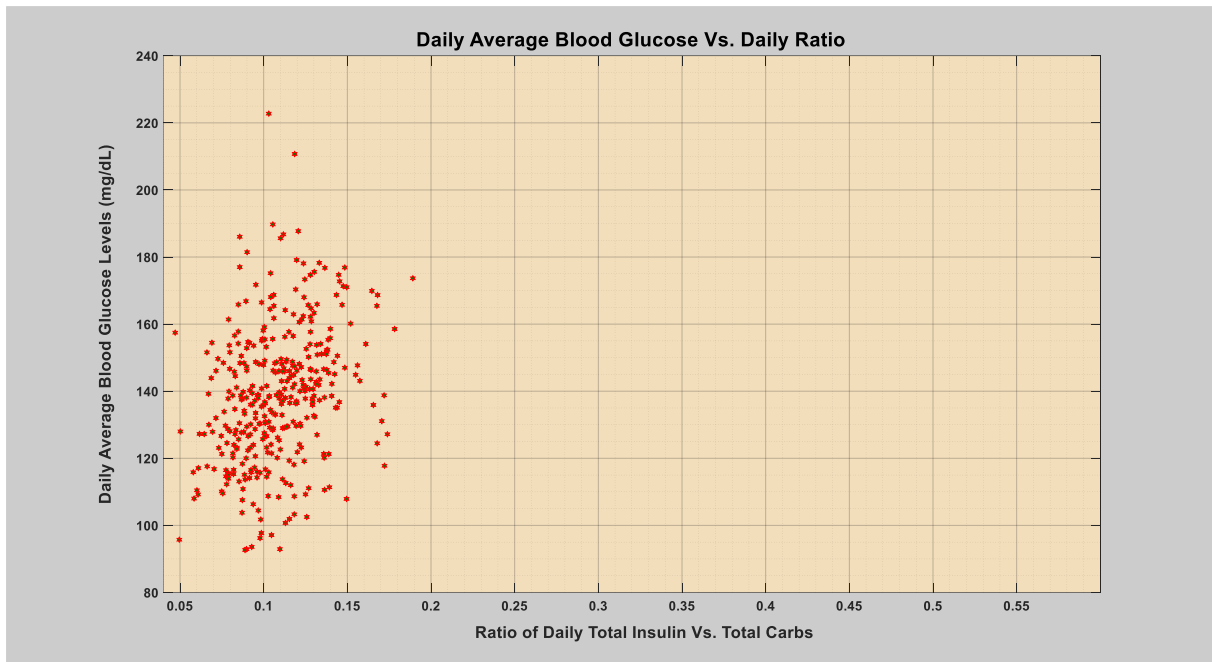


a) A daily plot of typical regular patient-year, source [258], Figure 1.

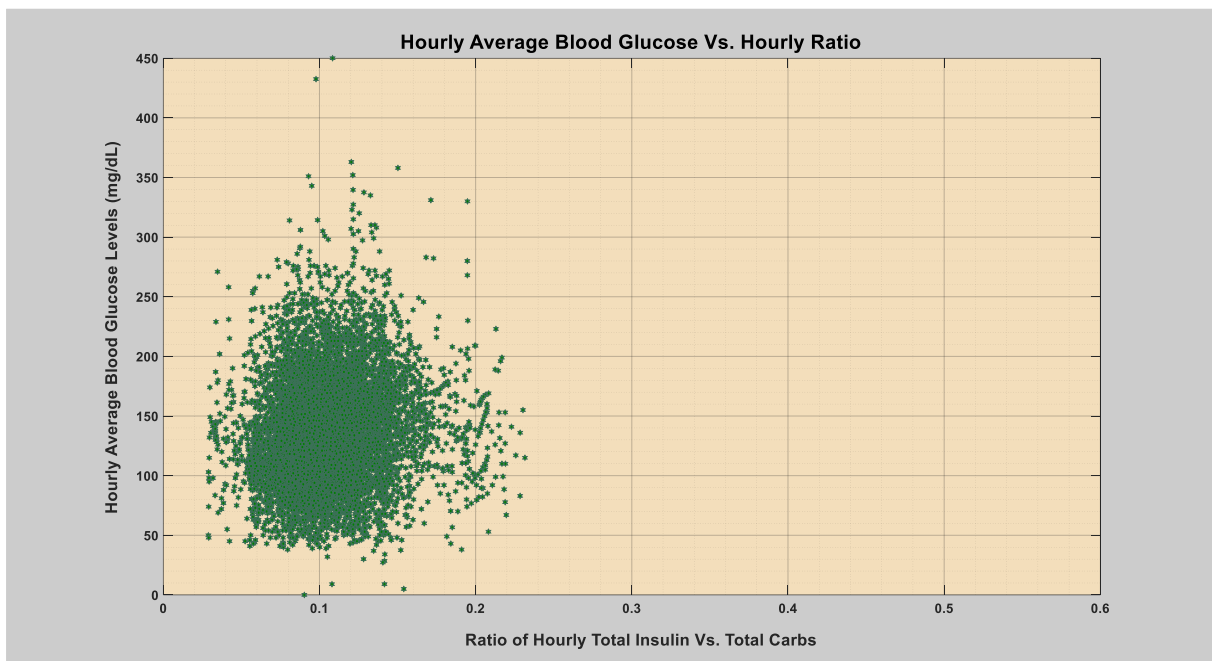


b) An hourly plot of typical regular patient-year, source [258], Figure 2.

**Figure 20:** A daily and hourly plot of typical smoothed regular patient-year depicting the scatter plot of average blood glucose levels vs. the ratio of insulin to carbohydrate.

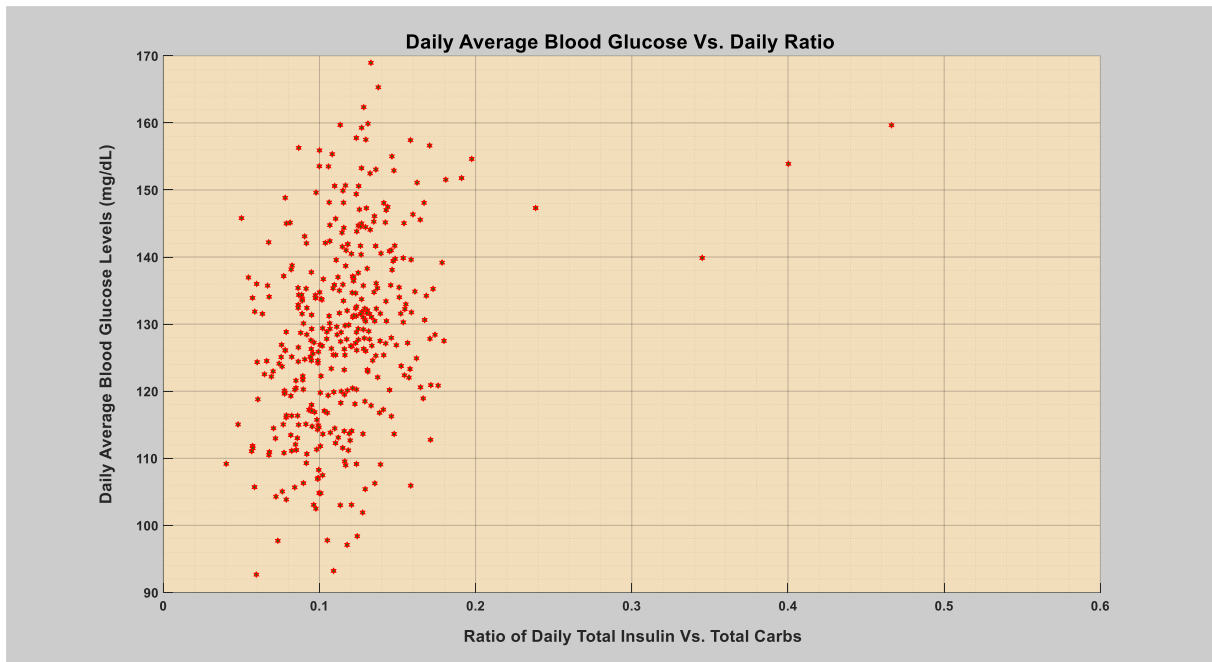


a) A daily plot of typical regular patient-year, source [258] Appendix 2, Figure 4.

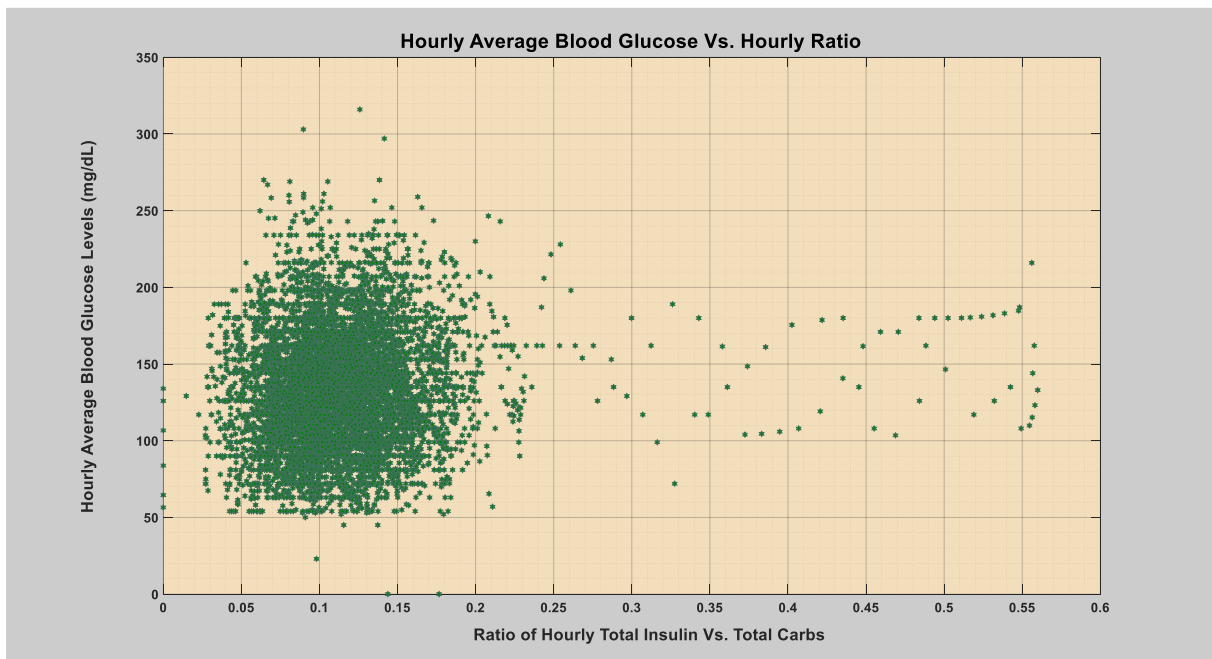


b) An hourly plot of typical regular patient-year, source [258] Appendix 2, Figure 4.

**Figure 21:** A daily and hourly plot of typical smoothed regular patient-year depicting the scatter plot of average blood glucose levels vs. the ratio of insulin to carbohydrate.

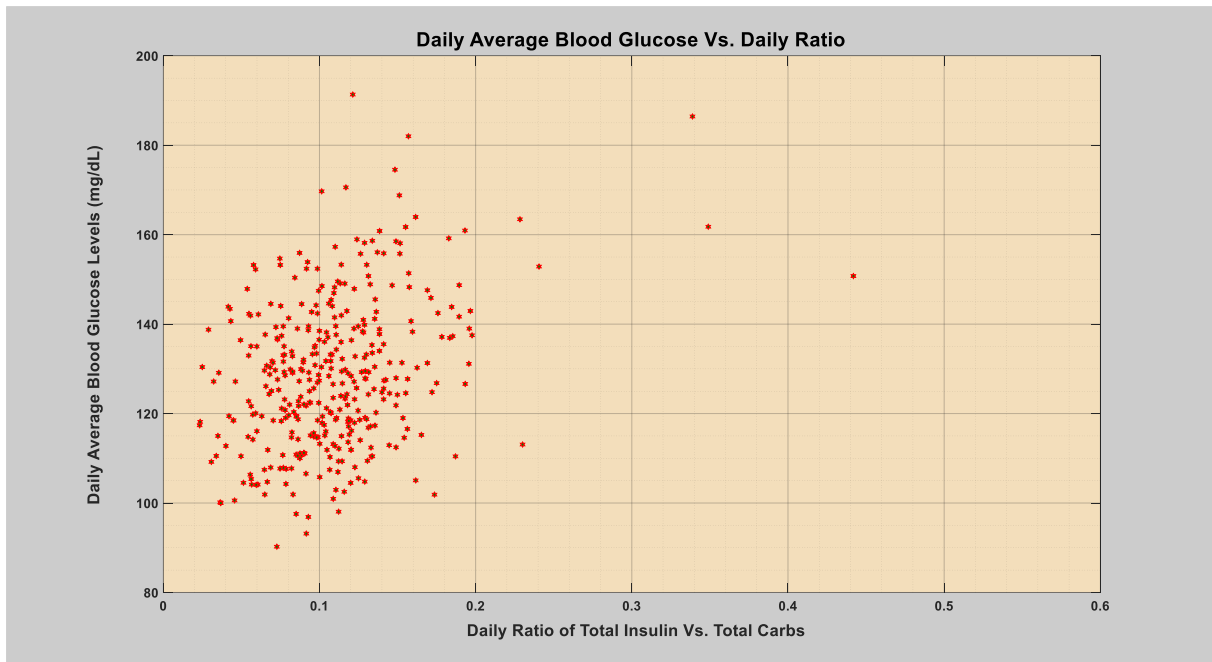


a) A daily plot of typical patient-year with an event of infection influenza (flu), source [258], Figure 3.

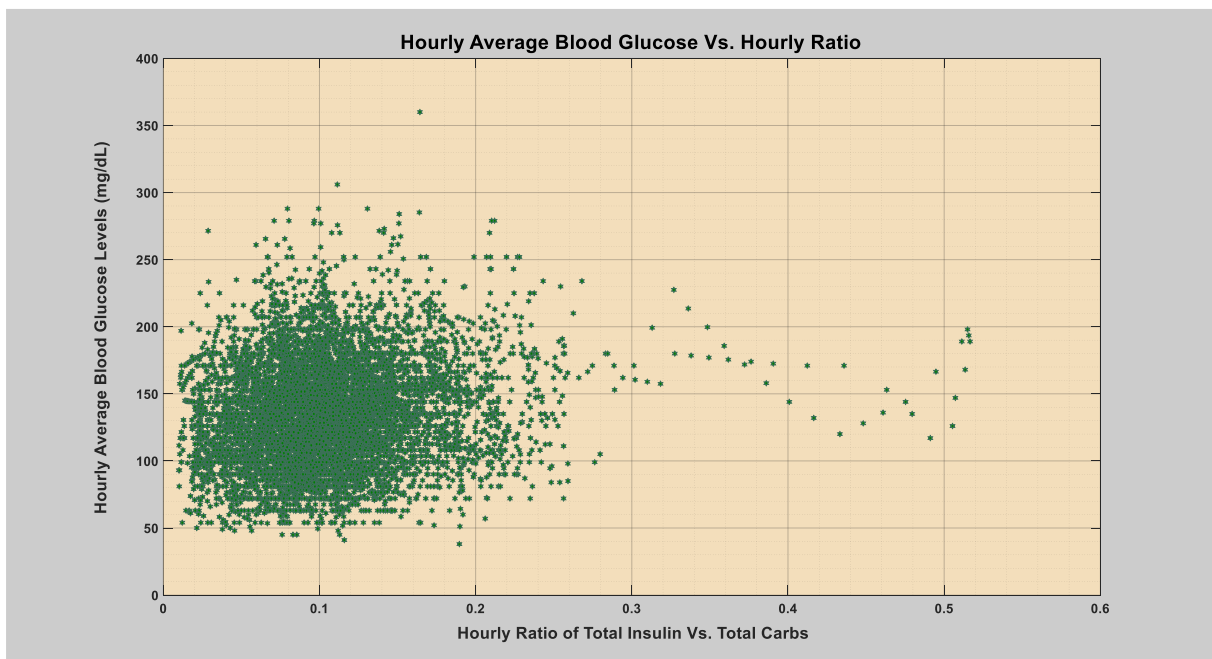


b) An hourly plot of typical patient-year with an event of infection influenza (flu), source [258], Figure 4.

**Figure 22:** A daily and hourly plot of typical smoothed patient-year with an event of infection influenza (flu), and depicting the scatter plot of average blood glucose levels vs. the ratio of insulin to carbohydrate.



a) A daily plot of typical patient-year with an event of infection influenza (flu), source [258] Appendix 2, Figure 8.

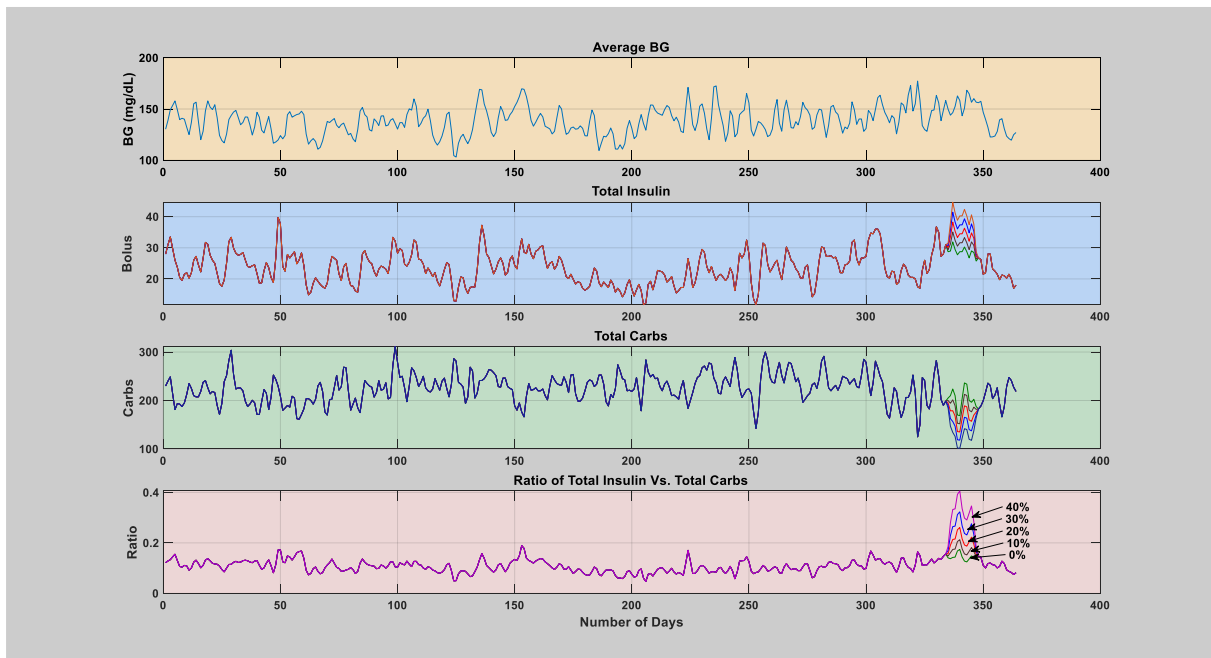


b) An hourly plot of typical patient-year with an event of infection influenza (flu) source [258], Appendix 2, Figure 8.

**Figure 23:** A daily and hourly plot of typical smoothed patient-year with an event of infection influenza (flu), and depicting the scatter plot of average blood glucose levels vs. the ratio of insulin to carbohydrate.

### 6.2.1.2 Data containing simulated events of infections

The degree of infection triggered deviations on blood glucose dynamics depends on several factors and is mainly related to the type of pathogens involved and the individual immunity [259]. Per the findings and description provided in *chapter five*, and considering pathogen-specific deviations, infection states of different sizes and shapes were simulated and injected into the daily regular patient-years. The simulated infection states were for a duration of two-weeks and incorporate 10%, 20%, 30%, and 40% of simultaneous deviation, i.e. higher insulin and lower carbohydrate by the same factor, from the aggregated insulin and carbohydrate profile of an individual, as shown in **Figure 24**. Those deviations were added to the five regular patient-years. The simulated infection states were used to assess and evaluate the model's performance in response to the different degrees of infection-induced changes from small to large changes in the individual blood glucose dynamics. This directly corresponds to the fact that different pathogens trigger a different degree of deviation in blood glucose dynamics.



**Figure 24:** Characteristics of the input feature, i.e. insulin to carbohydrate ratio, with simulated infection states of varying degree and shape ( $\alpha = 0\%$ ,  $10\%$ ,  $20\%$ ,  $30\%$ , and  $40\%$ ).

## 6.3 Practical experiments and results

Two categories of approaches were evaluated and compared; one-class classifiers and unsupervised methods. A detailed description of the models given in **Table 11** can be found in *Appendix 1 of Paper 3* [258]. For the one-class classifiers, during training, the regular period of the patient year was designated as a target class, and the period incorporating the infection episode as a non-target class. The exact location of the anomalies (novel) days was determined based on the individual's self-reported time-window of the infection episode and through the analysis conducted in *chapter five*. As demonstrated and presented in *chapter five*, the number of days with abnormal values of the insulin-to-carbohydrate ratio is regarded as non-target (anomalies) days. The one-class classifiers were trained using the target sample objects and

tested with a dataset containing both target and non-target sample objects. The unsupervised models were evaluated by presenting the entire dataset at once. The computational time required to execute each model was also estimated. The classifiers were compared based on their performance, required sample size to produce an acceptable description of the data, and computational time. Furthermore, the performance of those classifiers was compared with the unsupervised models.

**Table 11:** Models tested for the proposed personalized health model. BD = boundary and domain-based, DN= density-based, RE = reconstruction-based methods.

Models	One-class classifier	Unsupervised
Support vector data description (SVDD) [217; 218; 221]	√ BD	X
Incremental SVDD (IncSVDD) [218]	√ BD	X
One-class support vector machine (V-SVM) [198]	√ BD	X
Nearest neighbor (NN) [183; 216; 217; 220]	√ BD	X
Minimum spanning tree (MST) [116]	√ BD	X
Gaussian [183; 216; 218; 219]	√ DN	X
Minimum covariance Gaussian (MCG) [188; 216]	√ DN	X
Mixture of Gaussian (MOG) [216; 218]	√ DN	X
Parzen [174; 183; 216; 219]	√ DN	X
Naïve Parzen [216; 218]	√ DN	X
k-nearest neighbor (KNN) [183; 216; 219]	√ DN	X
Local outlier factor (LOF) [42; 218]	√ DN	√
Principal component analysis (PCA) [216-218]	√ RE	X
K-means [218]	√ RE	X
Self-organizing map (SOM) [217; 218]	√ RE	X
Auto-encoder network (AE) [216; 217]	√ RE	X
Connectivity-based outlier factor (COF) [215]	X	√

### 6.3.1 Model evaluation

The model evaluation was conducted based on *data type* - raw and smoothed data, *data granularities* - hourly and daily timeframes, and *data sample sizes* - one, two, three, and four months. For each data granularity, the ratio of total insulin to total carbohydrate was computed and used as either a univariate input or used along with average blood glucose levels as bivariate input. For the daily case, the models were evaluated based on raw data and its smoothed version with a two-days moving average filter. The model's performance was evaluated for each individual's dataset and reported using three performance metrics; *m* runs average and standard deviation of the *area under the receiver operating characteristic (AUC)*, *specificity*, and *F1-score*. Comparison of the overall models' performance among all the individual's dataset was carried based on these metrics, however, for the sake of clarity, the findings are depicted in terms of *F1-score*, given its practical implication, and the rest of the metrics values are given in *Appendix B of this dissertation and Appendix 4 of Paper 3* [258].

### 6.3.1.1 One-class classifiers

The hyper-parameters of most of the one-class classifier models, i.e. complexity parameter  $\gamma$ , were optimized based on the consistency approach [222] except for Parzen, and NN, which was optimized by using the leave-one-out error. Min-max was used for normalizing the dataset [118]. For MST, the complete MST is selected. For PCA, the fraction of variance retained from the training dataset was determined based on repeated experiments and set to be 0.6. Twenty times 5-fold & ten times 3-fold stratified cross-validation was used to evaluate the performance of the bivariate and univariate one-class classifiers respectively. In all the cases, a pre-specified threshold of outlier fraction in the training dataset was set to be  $\epsilon^t = 0.01$ , where one percent of the most dissimilar target data could be excluded from generating the data description. As described earlier in this chapter, the models were evaluated based on three important data features; *granularity* (hourly and daily), *data nature* (raw and smoothed data), and *sample object size* (1, 2, 3, and 4 months).

### 6.3.1.2 Unsupervised method

For comparison purposes, two unsupervised models; local outlier factor (LOF) and connectivity based outlier factor (COF), were also evaluated based on two important data features; *granularity* (hourly and daily), and *data nature* (raw and smoothed data). The whole patient-year, i.e., a sample size of 365 days for the daily case and 365\*24 hours for the hourly case was used during evaluation. The model's performance was measured based on the average of twenty runs. Both LOF and COF require a user-supplied parameter, i.e. number of neighbors ( $k$ ). In this dissertation, the value was determined based on a set of repeated experiments, and an optimal value with superior performance was selected. In this regard,  $k$  was set to be 30 days and 240 hrs for the daily and hourly scenario respectively. Performance evaluation was carried out by setting a detection threshold, and an optimal threshold was determined through repeated evaluation of different values.

## 6.3.2 One-class classifier performance

This section presents the comparison of the models' performance with respect to the type of input features; *data nature* (raw and smoothed), *data granularity* (daily and hourly), and *training sample sizes* (1,2,3,4 months). The median and average performance of each model was computed from their respective performance on the individual's infection states in relation to different sample sizes, data granularity, and nature of data, which can be seen in *Appendix B of this dissertation and Appendix 4 of Paper 3* [258]. Whenever necessary, the overall standard deviation of each model was computed as a pooled standard deviation. Generally, the evaluations demonstrated that the models' performance improved with increasing *sample size*, and *smoothing* the data enabled the models to generate better descriptions with smaller sample sizes. Despite the stated improvement, one of the drawbacks of smoothing is related to the delay incurred, which could result in late detection.

### 6.3.2.1 Bivariate input feature - Contextual anomalies

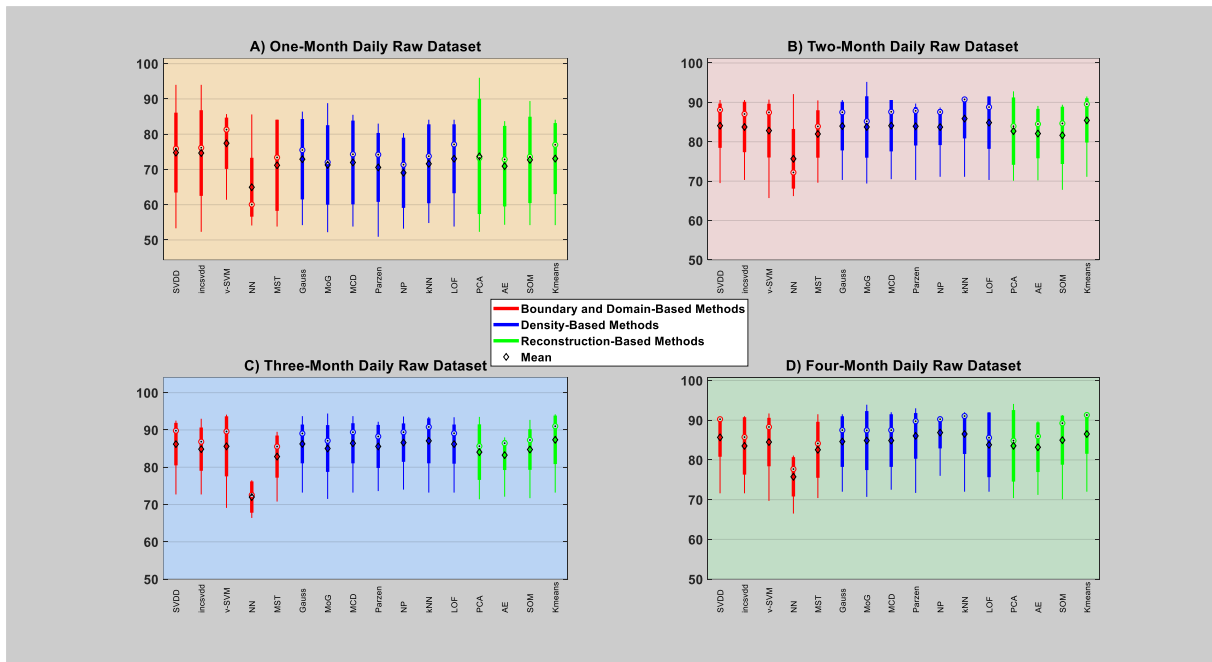
The bivariate input describes the use of both *blood glucose levels* and the *ratio of insulin-to-carbohydrate* as input parameters in modeling the detection algorithm. As a *contextual anomaly*, the blood glucose level is regarded as a behavior, which is evaluated by considering within the context of the insulin-to-carbohydrate ratio. Regarding the training sample size, the

models generate different descriptions, which mainly depend on the *data nature* and *granularity*, as shown in **Figure 25-27**. As a rule of thumb, a sample size of 3 months for the daily raw data, 1 month for the daily smoothed data, and 2 months for the hourly data could be sufficient to start with when the individual participants join the system. However, it is better to note that through time there will exist more data, and further model improvement (re-training) can be carried out. As expected, in contrast to the raw data, the smoothed version achieved excellent description. Among the three categories, the boundary and domain-based method produced a better performance with a 1-month sample size in both data granularities, specifically v-SVM. For a higher sample size, the density-based method generated a better description with the raw daily data, the boundary and domain-based method are better with the smoothed daily data, and the density and reconstruction-based methods are better with the hourly data. In particular, on average, v-SVM, K-means, and K-NN are the three top-performing models from their respective category demonstrating superior performance (excellent descriptions) as compared to the rest of the models. The performance (score) plot of the models can be found in *Appendix 3 of Paper 3* [258], depicting the capability of each model in detecting the infection episode from the regular period. These models were trained on a random block of *120 regular days (4 months)* of the patient year and tested on the entire patient-year.

#### **6.3.2.1.1 Daily raw dataset**

The raw data is pre-processed data without further smoothing and can contain short-term and fast-scale features that could affect the model's generalization. As expected, as shown in **Figure 25**, the models suffer in performance degradations, where the models' performance exhibits wider variations. Increasing the sample size has shown little improvement on the models' descriptions, specifically after the three-month sample size. As compared to the other methods, the boundary and domain-based method, specifically v-SVM, performed better with a 1-month sample size. With a two-month sample size, all the three methods improved, and density-based method, particularly K-NN, and reconstruction based method, particularly K-means, performed better. For higher sample sizes (three and four-month), generally, the density-based method outperforms the other methods. In this regard, the boundary and domain-based method (i.e. SVDD and v-SVM), density-based method (i.e. K-NN, Parzen, and Naïve Parzen), and reconstruction-based method (i.e. K-means and SOM) demonstrated better performance. Overall, a model such as v-SVM performed better with the 1-month sample size, and SVDD, K-NN, Naïve Parzen, and K-means performed better with higher sample sizes.

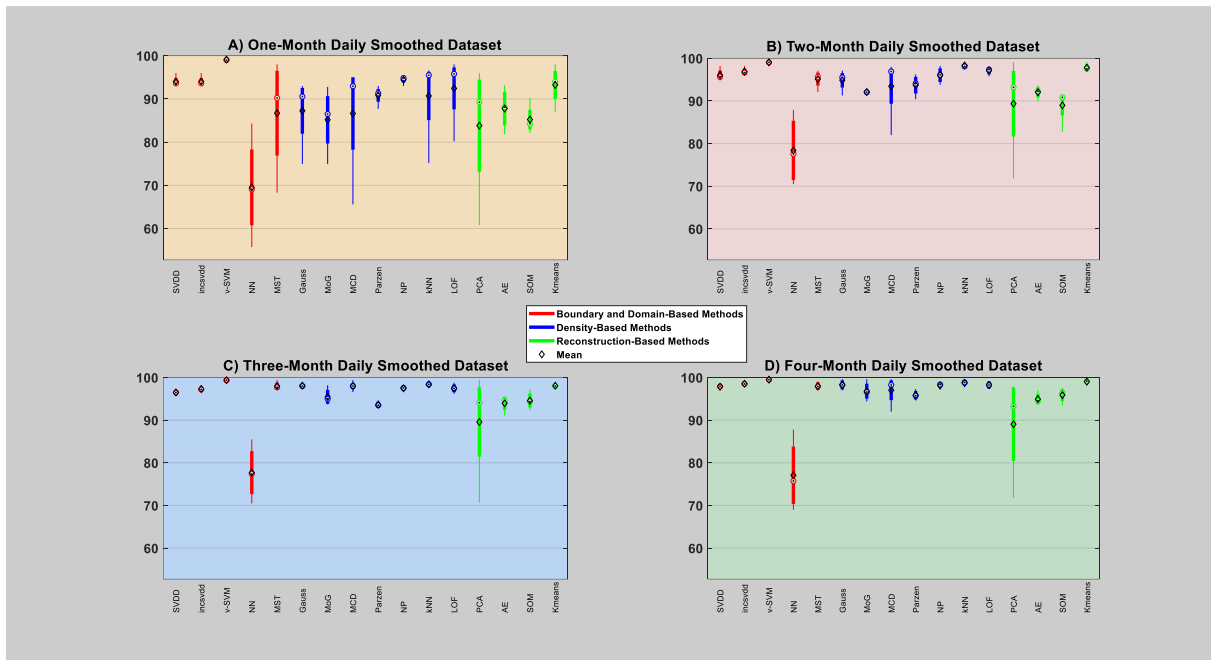




**Figure 25:** Bivariate input feature - The models’ median and average F1-score over the daily raw datasets, source partially from [258], Figure 8.

### 6.3.2.1.2 Daily smoothed dataset

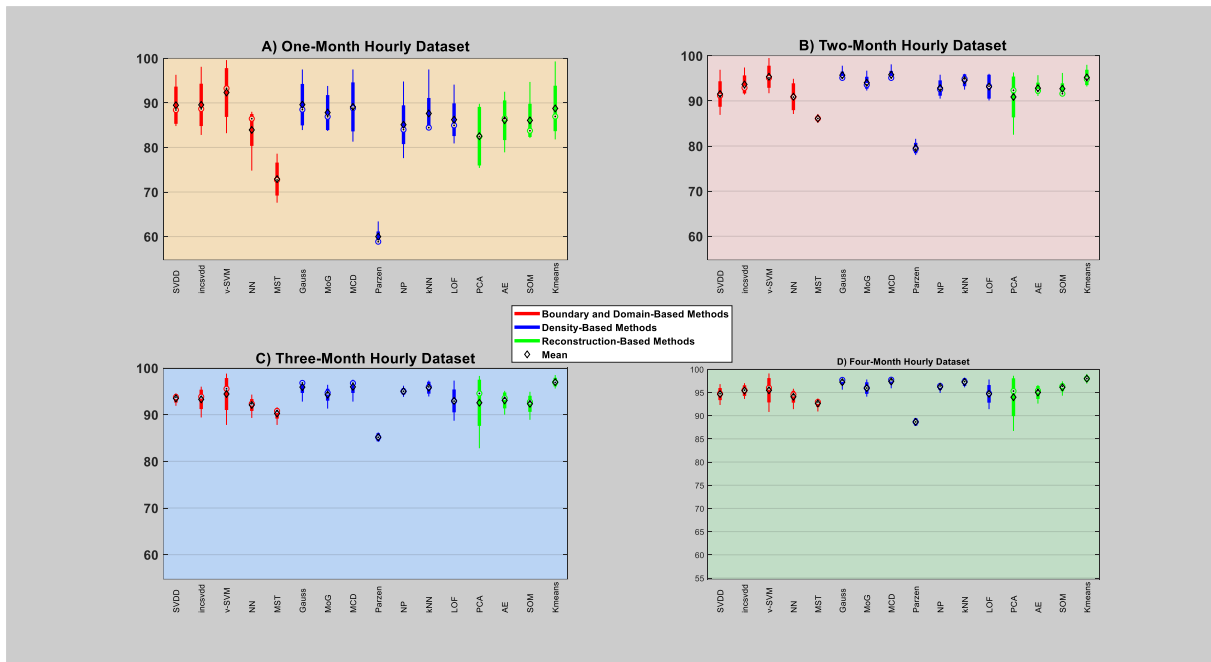
The smoothed dataset is a filtered version with fewer short-term and fast-scale features, and in this version, the models are generally expected to improve. As shown in **Figure 26**, as expected the models achieved significant improvements over the raw version of the data. Increasing the sample size has drastically improved the models’ description attaining enhanced performance, specifically can be seen from the two-month sample size and onwards. As compared to the other methods, the boundary and domain-based method, particularly v-SVM, achieved excellent descriptions with a 1-month sample size. With a two-month sample size, all the three methods improved, and boundary and domain-based method, particularly v-SVM, density-based method, particularly K-NN, and reconstruction based method, particularly K-means, performed better. For higher sample sizes (three and four-month), all the methods achieved comparable descriptions. Overall, v-SVM is the best model, which generates excellent descriptions in all sample sizes. From each respective method, on average, models such as v-SVM, K-NN, and K-means achieved superior descriptions for all the higher sample sizes and the worst-performing models being NN and PCA.



**Figure 26:** Bivariate input feature - The models’ median and average F1-score over the daily smoothed datasets, source partially from [258], Figure 8.

### 6.3.2.1.3 Hourly smoothed dataset

The hourly dataset depicts the average change in blood glucose levels in relation to the ratio of insulin-to-carbohydrate within each hour of the day. Unlike the daily scenario, the hourly dataset incorporates more training examples per day (24 sample objects). Increasing the data granularity could provide finer details and early detection, however, at the cost of unwanted features, which might become very significant as the level gets higher. As can be seen in **Figure 27**, despite smoothing the models still exhibit wider performance variations as compared to the daily smoothed dataset. As per the previous findings, generally, increasing the sample size and smoothing the data should enable the models to generalize well with little variations, however, the presence of unwanted features within the hourly data hampers the models’ generalization ability. In this dataset, in a comparison between methods, the boundary and domain-based method, particularly v-SVM, achieved better descriptions with a 1-month sample size. With a two-month sample size, all the three methods improved, and boundary and domain-based method (i.e. v-SVM), density-based method (i.e. Gaussian families and K-NN), and reconstruction based method (i.e. K-means) performed better. For higher sample sizes (three and four-month), generally, the density-based method is better. Overall, a model such as v-SVM performed better with the 1-month sample size, and Gaussian, MCD Gaussian, KNN, and K-means performed better with higher sample sizes and the worst-performing models being NN, Parzen, MST, and PCA.



**Figure 27:** Bivariate input feature - The models’ median and average F1-score over the hourly smoothed datasets, source partially from [258], Figure 8.

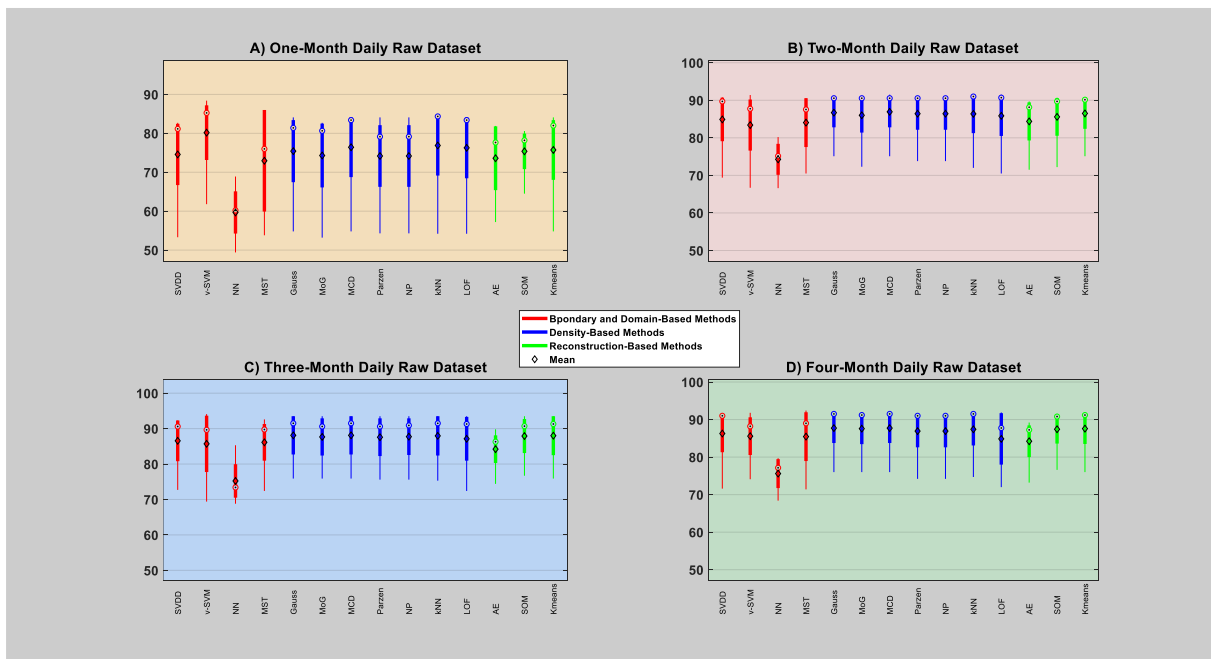
### 6.3.2.2 Univariate input feature - Point anomalies

The univariate input describes the use of the ratio of insulin-to-carbohydrate as the sole input parameters in modeling the detection algorithms. Per the findings presented in *chapter five*, the ratio of insulin-to-carbohydrate is demonstrated to be highly informative containing high discriminative power compared to the other key parameters of blood glucose dynamics. In this regard, this input feature is expected to generate improved performance compared to the bivariate input feature. This alternative approach presents an advantage when there is a lack of access to blood glucose measurements. As a *point anomaly*, each point in the time-series of the ratio is evaluated against a reference description generated by the models to determine its degree of normality. As per the findings, generally, the models exhibit performance variations with different sample sizes, as shown in **Figure 28-30**, and the degree of these variations is mainly dependent on the *data nature* and *granularity*. Regarding the optimal sample size, there is no much difference with the bivariate input feature. As a rule of thumb, a sample size of 2 months for the daily raw data, 1 month for the daily smoothed data, and 2 months for the hourly data could be sufficient to start with when the individual participants join the system. However, through time there will exist more data, and further model improvement can be carried out after the individual participants join the proposed system. Similar to the bivariate feature, smoothing allows the models to generate excellent description compared to the raw dataset. In general, among the three methods, the boundary and domain-based method produced a better performance with a 1-month sample size in both data granularities. For a higher sample size, all the three methods achieved comparable description with the raw and smoothed daily data, and the density and reconstruction-based methods are better with the hourly data. In particular, on average, v-SVM produced a better performance from 1-month sample size and all the models achieved comparable performance for the higher sample size except NN. The performance

(score) plot of the models can be found in *Appendix A of this dissertation*, which depicts the capability of each model in detecting the infection episode from the regular period. These models were trained on a random block of *120 regular days (4 months)* of the patient year and tested on the entire patient-year.

### 6.3.2.2.1 Daily raw dataset

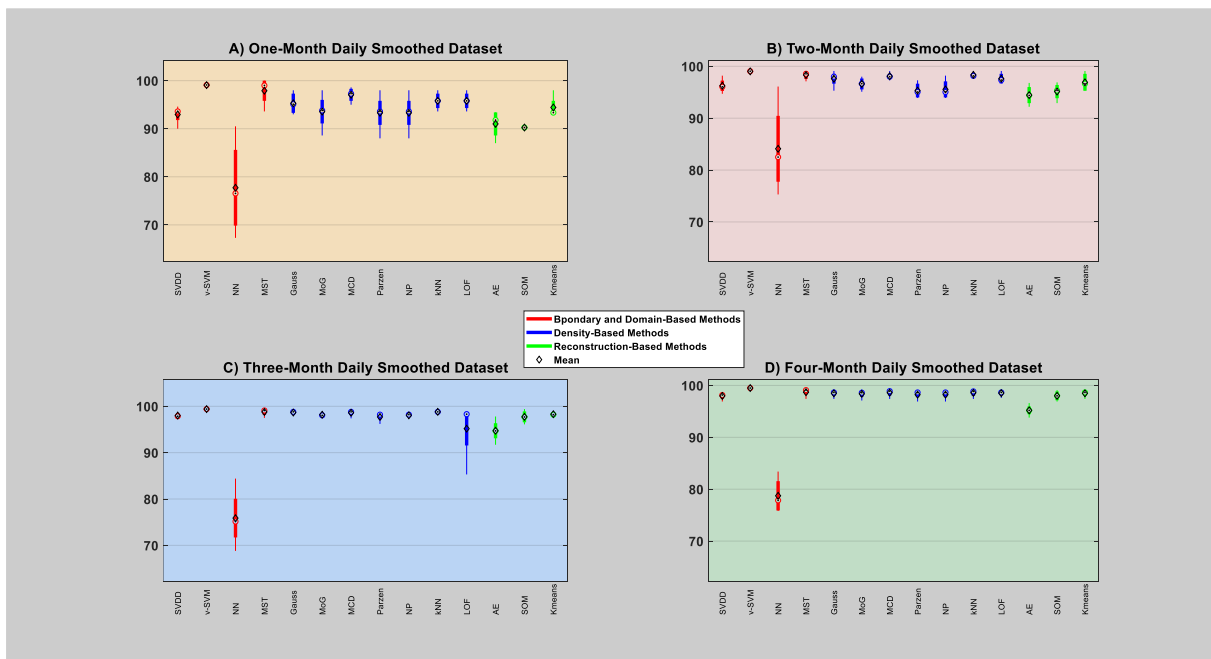
As described earlier in this chapter, this input dataset is without smoothing and can contain short and fast-scale features, which could affect the models' generalization. As compared to the bivariate feature, the models demonstrated better performance with the univariate input feature with smaller sample sizes, i.e. one and two months, and comparable performance with higher sample sizes (three and four months), as shown in **Figure 28**. Increasing the sample size has a better effect in improving the model's performance as compared to the bivariate input case. The performance improvement is mainly related to the characteristic of the insulin-to-carbohydrate ratio, which exhibit similar patterns of values during the regular days of the year. As compared to the other methods, the boundary and domain-based method, specifically v-SVM, performed better with 1-month sample size. With a two-month sample size, all the three methods demonstrated significant improvement, and all the models except NN generated comparable description. For higher sample sizes (three and four-month), there is no difference among all the three methods, and the worst model being NN. Overall, a model such as v-SVM performed better with the 1-month sample size, and except NN, all the other models achieved comparable performance. As the sample size increase, all the models except NN improved and produced comparable performances.



**Figure 28:** Univariate input feature - The models' median and average F1-score over the daily raw datasets.

### 6.3.2.2.2 Daily smoothed dataset

As described earlier in this chapter, this input dataset is a filtered (smoothed) version of the raw dataset, and generally, the models are expected to generate improved description compared to the raw version. As compared to the raw dataset, as expected, the models have achieved significant performance improvement, as shown in **Figure 29**. Further comparison with the bivariate input demonstrated the advantage of the univariate input, which enabled the models to achieve better description in all the sample sizes. In this specific dataset, increasing the sample size has little effect on performance improvement as the models have already achieved better description with lower sample sizes. As compared to the other methods, the boundary and domain-based method, specifically v-SVM, performed better with a 1-month sample size. With a two-month sample size, all the three methods improved, and boundary and domain-based method (i.e. v-SVM and MST), density-based method (i.e. Gaussian families, LOF and K-NN), and reconstruction based method (i.e. K-means) performed better. For higher sample sizes (three and four-month), all three methods generate comparable performances, and all the models produced similar descriptions except NN, which produces the worst description. Overall, on average, models such as v-SVM and MST achieved relatively greater performance in all the sample sizes, while all the other models except NN achieved a comparable description with two and more sample sizes. Generally, as the sample size increases, all the models except NN achieved comparable performance.

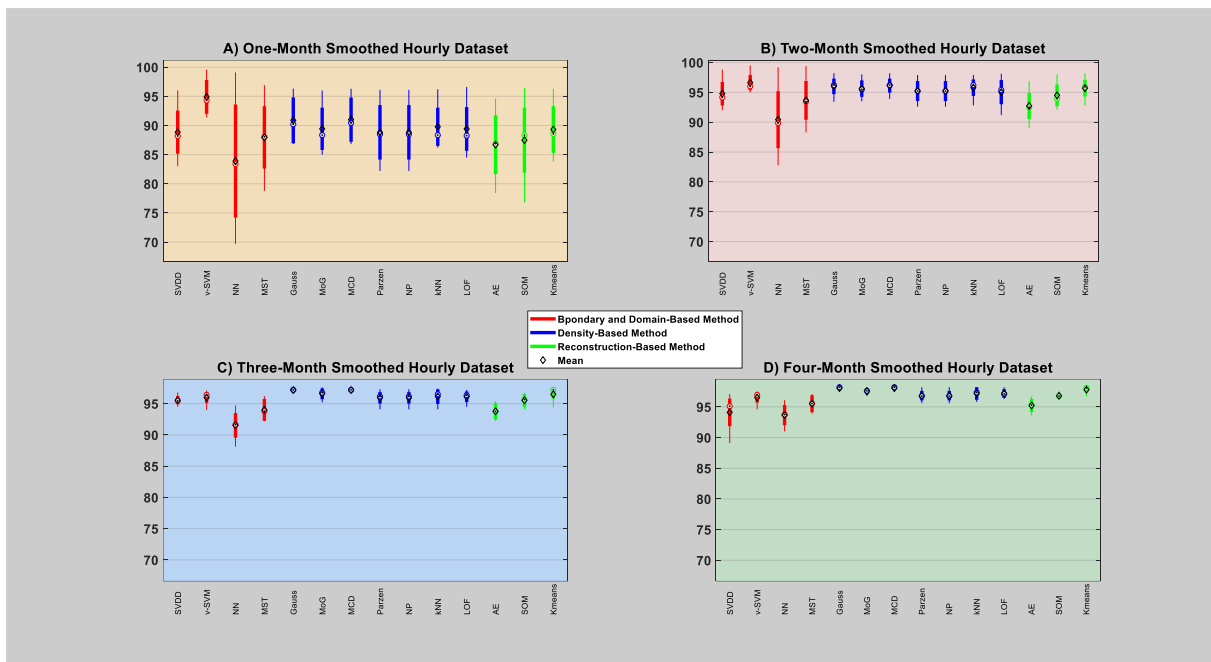


**Figure 29:** Univariate input feature - The models’ median and average F1-score over the daily smoothed datasets.

### 6.3.2.2.3 Hourly smoothed dataset

As described earlier, the hourly dataset is a filtered version depicting the relationship between the average blood glucose levels and the ratio. However, in the univariate sense, only the user’s estimated carb consumption and insulin requirements within each hour of the day are

considered. As described earlier, increasing the data granularity could improve the possibility of detecting the anomalies state at an earlier stage, however, at the cost of accuracy. Therefore, as expected, despite presenting a large sample size and smoothing the data, as can be seen from **Figure 30**, the models exhibit high variance as compared to the daily smoothed dataset. Further comparison with the bivariate input demonstrated the advantage of the univariate input only with smaller sample sizes, and with higher sample sizes both the input produced comparable descriptions. As compared to the other methods, the boundary and domain-based method, specifically v-SVM, performed better with 1-month sample size. With a two-month sample size, all the three methods improved, and boundary and domain-based method (i.e. v-SVM, and SVDD), density-based method (i.e. Gaussian families, LOF, Parzen, Naïve Parzen, and K-NN), and reconstruction based method (i.e. SOM and, K-means) performed better. For higher sample sizes (three and four-month), all three methods generate comparable performances, and the models achieved somewhat comparable descriptions except NN, which generated the worst description in all the sample sizes. Generally, increasing the sample size has helped the models to capture the data distribution better. Overall, models such as v-SVM achieved relatively greater performance with a 1-month sample size, while models such as Gaussian families and K-means achieved better descriptions with higher sample sizes.

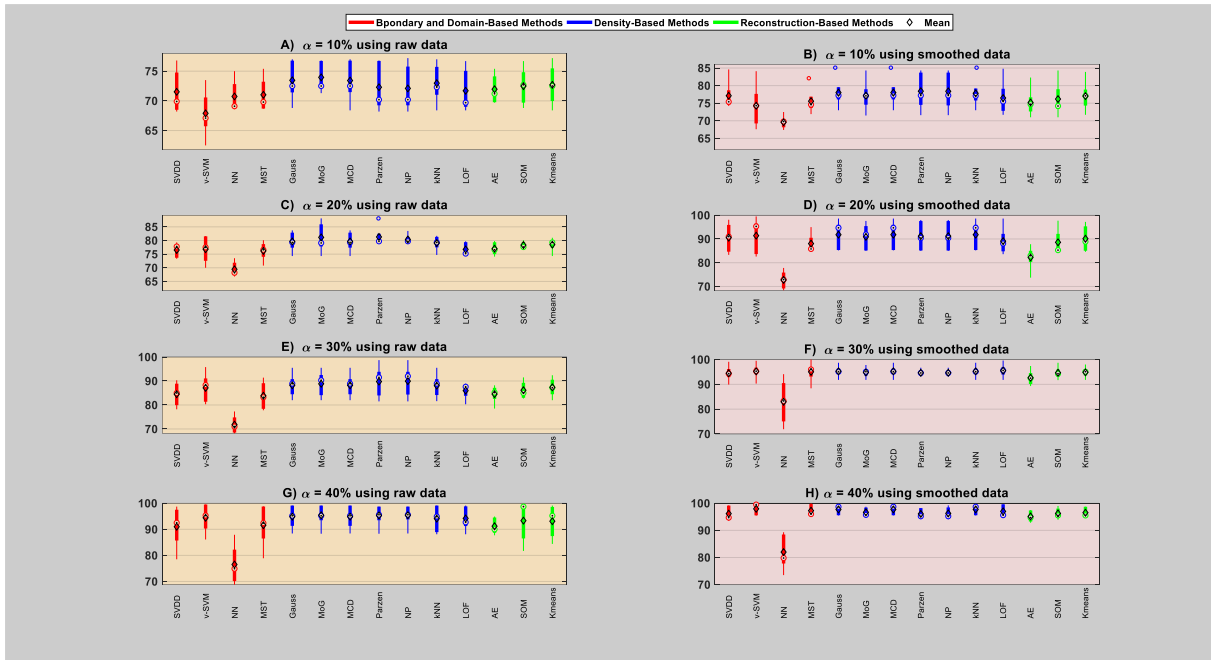


**Figure 30:** Univariate input feature - The models’ median and average F1-score over the hourly smoothed datasets.

#### 6.3.2.2.4 Models performance with a different level of deviations

This section presents evaluations of the model’s performance with data containing simulated deviations of different degrees. As described earlier in this chapter, a two-week-long simulated infection triggered deviations was injected into the five regular patient-years. Simultaneous deviations of  $\alpha = 10\%$ ,  $20\%$ ,  $30\%$ , and  $40\%$  were added to the daily aggregated insulin and carbohydrate values. Univariate input was used and the models were trained with 4-month

sample size, and their performance was evaluated and compared on the raw and smoothed version of each patient-years. As shown in **Figure 31**, with increasing deviation  $\alpha$ , the model's detection performance improves. In a comparison of the data types, the models achieved better detection performance with the smoothed dataset. Detecting infection states that induce very small deviations, i.e.  $\alpha < 10\%$  change, requires training the models with a suitable threshold that could reject outliers in the training dataset that exceeds the induced deviations (i.e.  $\alpha < 10\%$ ). However, this could in turn increase the false alarm rate and make the model less sensitive flagging regular days as an infection state. In this regard, for an application that involves detecting an infection state, it is necessary to favor the inclusion of some of the less significant outliers in the data description to avoid frequent false alarm, however, at the expense of missing infection state that induces small deviations (i.e.  $\alpha < 10\%$ ) on the blood glucose dynamics. Per the findings, with  $\alpha = 10\%$ , the density-based method performed a better detection task, and specifically, MOG achieved better description. In this regard, generally, the Gaussian family achieved better performance with the raw dataset, however, all the models achieved comparable descriptions except NN with the smoothed dataset. It is better to note that, despite the small deviation ( $\alpha = 10\%$ ), smoothing the data has helped the models to achieve good descriptions. For  $\alpha = 20\%$ , among the three methods, the density-based method (i.e. Gaussian families, Parzen and naïve Parzen) and reconstruction-based method (i.e. SOM and K-means) achieved better description with the raw data, and regarding the smoothed data, the boundary and domain-based method (i.e. SVDD, and v-SVM), density-based method (i.e. Gaussian families, and K-NN) and reconstruction based method (i.e. SOM, and K-means) achieved relatively better and comparable performance. For  $\alpha = 30\%$ , the density-based method performed better, and specifically, Parzen and naïve Parzen achieved slightly better performance with the raw data, and almost all the models except the nearest neighbor achieved comparable performance with the smoothed dataset. For  $\alpha = 40\%$ , all the three methods achieved comparable performance, and all the models except NN achieved comparable performances.



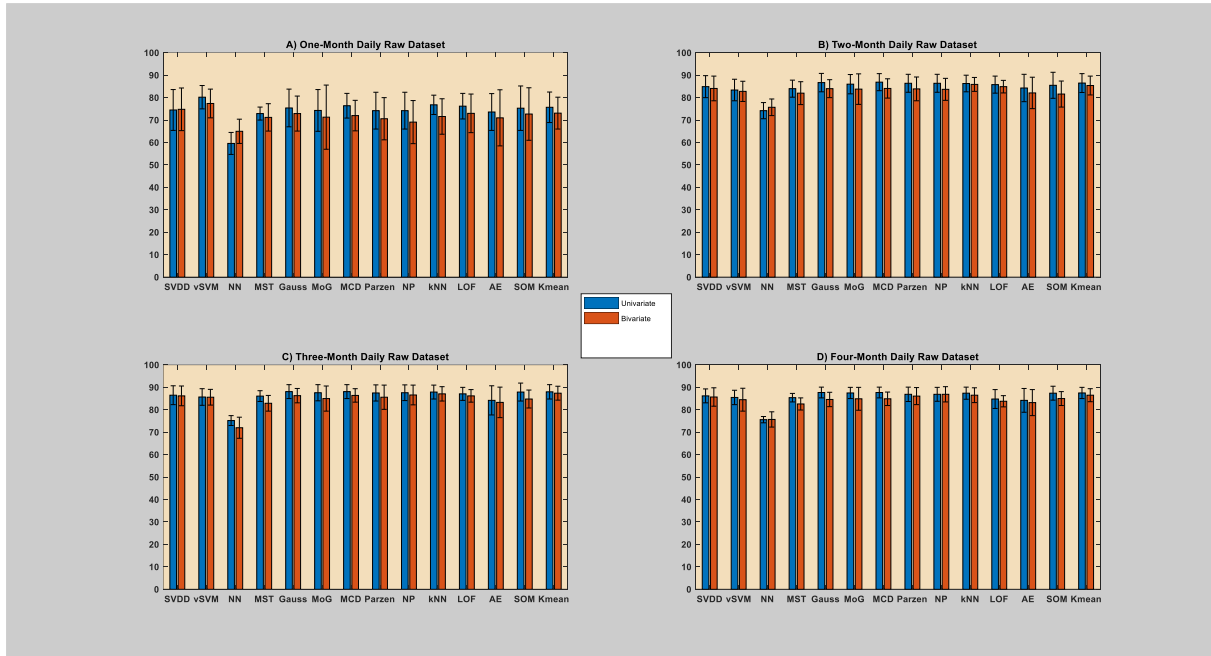
**Figure 31:** Univariate input feature - median and average performance (F1-score) of the models over the five patient years injected with different degree of deviations.

### 6.3.3 Comparison of input features

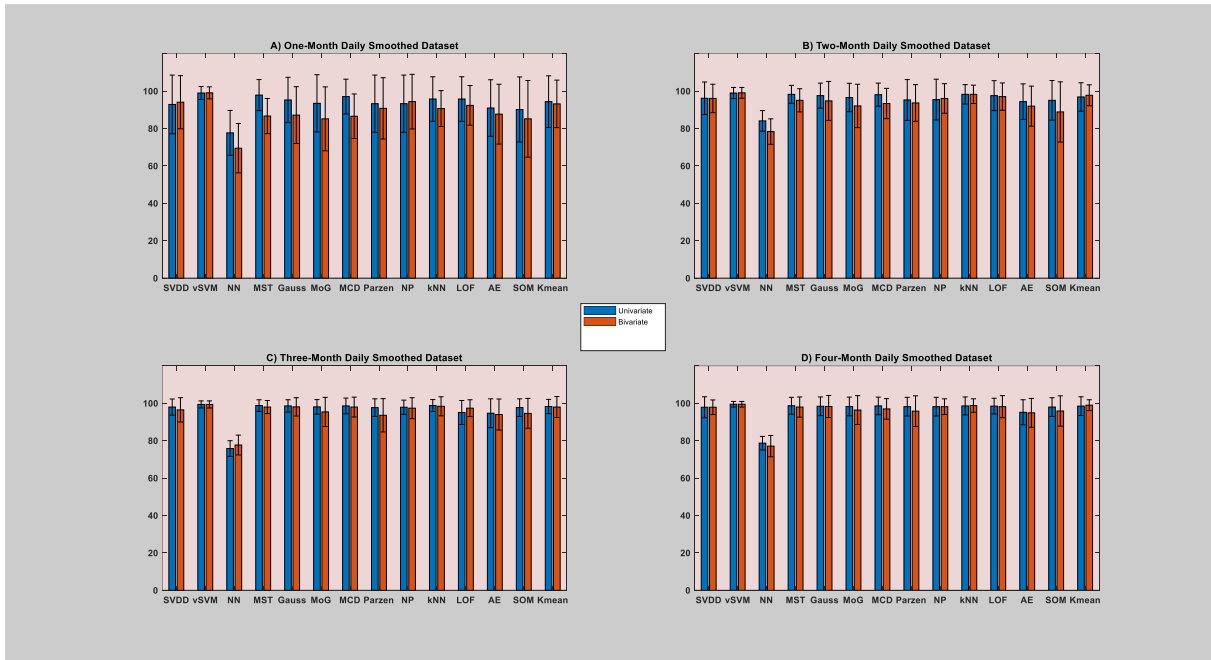
The performance of any model can be greatly affected by the input features selected for modeling [226]. In this regard, this section presents the comparative analysis of the performance achieved with a univariate input, i.e. insulin to carbohydrate ratio, and a bivariate input, i.e. blood glucose levels and insulin to carbohydrate ratio. As shown in **Figure 32-34**, except under certain circumstances, where both achieved comparable performance, the univariate models display superior performances. However, despite the improved performance one of the drawbacks of the univariate input-based models emanates from the fact that these models cannot differentiate points that lie within *quadrants 1 and 2* or *quadrants 3 and 4*, as shown in **Figure 19** [258]. In this regard, theoretically, a very large ratio that resides in quadrant 3 is considered as normal value as long as the individual blood glucose levels go to the hypoglycemia state responding to the high insulin injection and low carbohydrate intake [197; 258], and however, the univariate models consider such a situation as abnormal by just looking upon the ratio values as outliers. However, it should be noted that in practical settings such incidence might be fatal for the individuals and might end up being unconscious and sometimes dead [61; 151], and therefore, such ratio values might be almost non-existent in a practical sense. Therefore, it can be concluded that the univariate models could do the same task as the bivariate models when it comes to detecting infection incidences, i.e. large ratio values, despite lacking the capability to differentiate between quadrant 3 and 4. However, as compared to the bivariate model, the univariate model might sometimes generate a false alarm in rare situations that are very sparse, i.e. too small values not included in training the models. This rare situation is manifested in the individual's blood glucose management practice, for instance, if the patient on random days prefers to replace insulin requirement with physical exercise/activity sessions, this instance could end up with very small insulin to carbohydrate



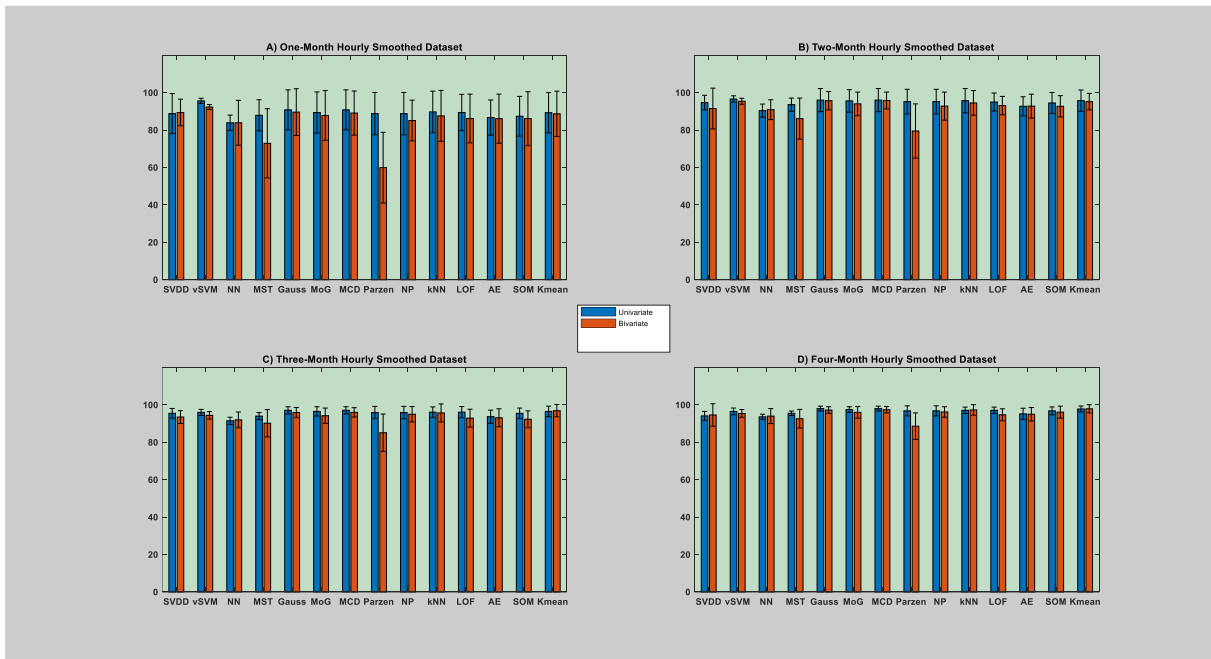
ratio and be flagged as a false alarm. In these circumstances, without the blood glucose level feature, it will remain difficult for these models to differentiate between the normality or abnormality of a small value of insulin to carbohydrate ratio. Hence, having blood glucose levels as an additional input feature could minimize such unnecessary alarms.



**Figure 32:** Daily raw dataset - performance comparison (F1-score) of models using bivariate input, i.e. blood glucose levels and insulin to carbohydrate ratio, and univariate input feature, i.e. insulin to carbohydrate ratio, based on the daily raw dataset. The error bars are given in terms of the overall mean and standard deviation of each model across all the patient-years and infection states.



**Figure 33:** Daily smoothed dataset - performance comparison (F1-score) of models using bivariate input, i.e. blood glucose levels and insulin to carbohydrate ratio, and univariate input feature, i.e. insulin to carbohydrate ratio, based on the daily smoothed dataset. The error bars are given in terms of the overall mean and standard deviation of each model across all the patient-years and infection states.



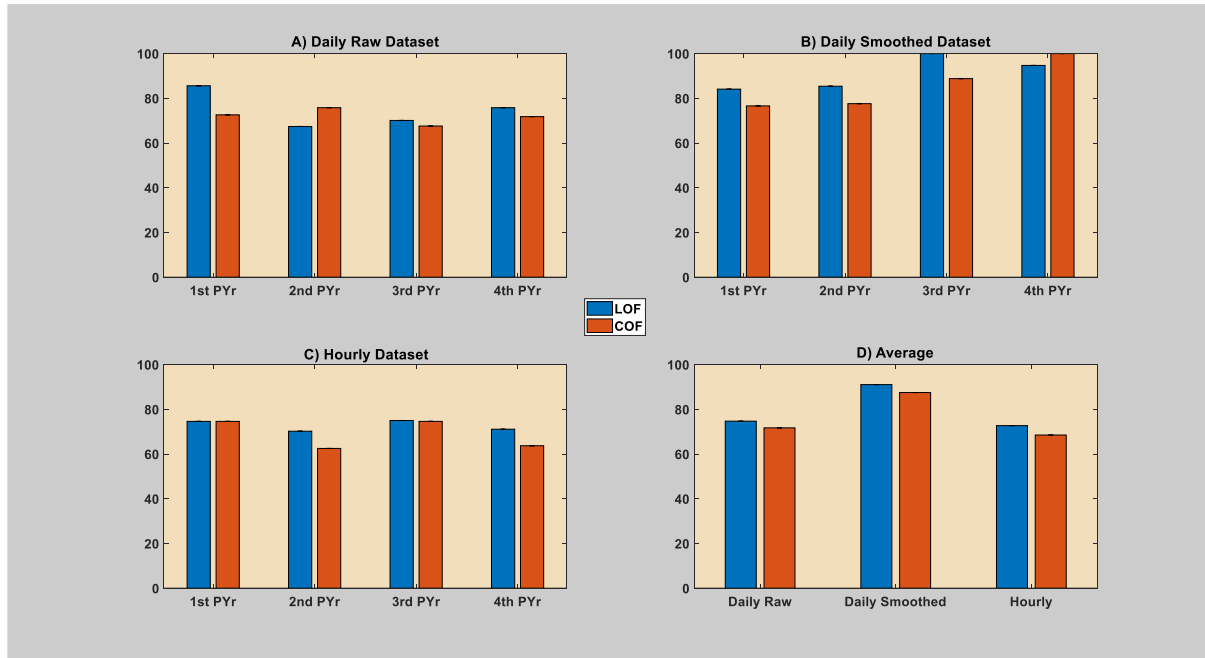
**Figure 34:** Hourly smoothed dataset - performance comparison (F1-score) of models using bivariate input, i.e. blood glucose levels and insulin to carbohydrate ratio, and univariate input feature, i.e. insulin to carbohydrate ratio, based on the hourly smoothed dataset. The error bars are given in terms of the overall mean and standard deviation of each model across all the patient-years and infection states.

### 6.3.4 Unsupervised method performance

For comparison purposes, this section presents the performance of two density-based unsupervised models, LOF and COF, as given in **Figure 35-37**. The performance of these models was compared based on the type of *input features*, *data nature* (raw and smoothed), and *data granularity* (daily and hourly). The average performance of each model was computed using their respective performance on the individual infection states in relation to different data granularity, and the nature of data can be seen in *Appendix B of this dissertation and Appendix 4 of Paper 3* [258]. The optimal threshold and number of neighbors were selected after performing repeated evaluations for different combinations of values. Generally, LOF has a slight edge over the COF in all the infection states, and this could be linked with the nature of the data distribution that agrees with the LOF assumption of spherical neighbor's distribution [42].

#### 6.3.4.1 Bivariate input feature - Contextual anomalies

The performance of the unsupervised models, LOF and COF, using a bivariate input, i.e. blood glucose levels and the insulin-to-carbohydrate ratio is given in **Figure 35**. The optimal threshold values used in the model evaluation are given in **Table 12**. As can be seen from the figure, smoothing the data enabled the models to generate better performance, and as a result, both the models were able to yield excellent performance with the daily dataset [258]. However, this is not the case for the hourly dataset, mainly related to the presence of highly detailed information containing unwanted short-term and fast-scale features. In general, both the models were able to produce a comparable performance in all the datasets, however, LOF has displayed better performance. The performance (score) plot of the models can be found in *Appendix 3 of Paper 3* [258], depicting the capability of each model in detecting the infection episode from the regular period. These models were tested on the entire patient-year.



**Figure 35:** Bivariate input features - performance comparison (F1-score) of the unsupervised models using bivariate input, i.e. blood glucose levels and insulin to carbohydrate ratio, source [258], Table 8.

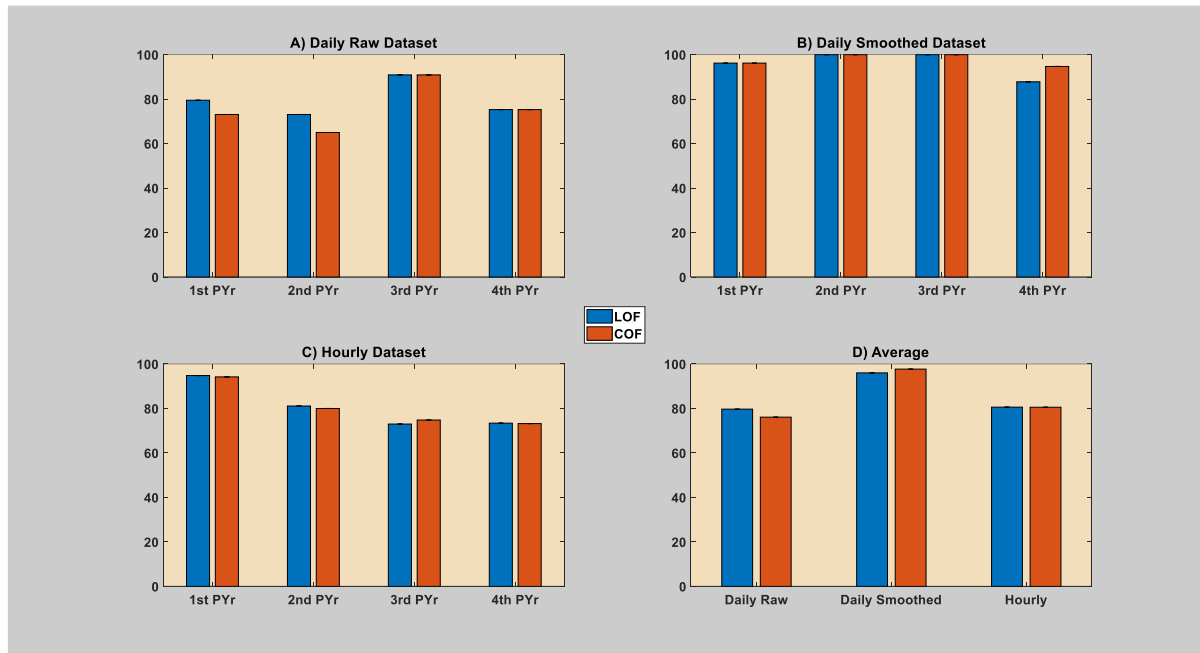
**Table 12:** Bivariate input features - optimal values of thresholds used in performance evaluation. The values given as  $T_h$  are the optimal threshold values used for each patient-year,  $h$  depicting that particular year, source [258], Table 8.

Granularity	Pre-pro.	Model (Threshold <sub>patient-year</sub> )
Daily	Without filter	LOF ( $T_1=2.4, T_2=1.2, T_3=1.45, T_4=1.8$ )
		COF ( $T_1=1.4, T_2=1.3, T_3=1.4, T_4=1.4$ )
	With filter	LOF ( $T_1=1.7, T_2=1.6, T_3=1.95, T_4=2.2$ )
		COF ( $T_1=1.3, T_2=1.3, T_3=1.8, T_4=1.8$ )
Hourly	With filter	LOF ( $T_1=1.4, T_2=1.3, T_3=1.35, T_4=1.5$ )
		COF ( $T_1=1.2, T_2=1.1, T_3=, T_4=1.1$ )

#### 6.3.4.2 Univariate input feature - Point anomalies

The performance of the unsupervised models, LOF and COF, using a univariate input, i.e. the insulin-to-carbohydrate ratio is given in **Figure 36**. The optimal threshold values used in the model evaluation are given in **Table 13**. As can be seen from the figure, like the bivariate case, smoothing the daily dataset significantly improved the models' performance. As described above, similar to the bivariate case, smoothing the hourly dataset yields inferior performance to the daily scenario, due to the presence of unwanted short-term and fast-scale features as a result of higher data granularity. In general, both these models achieved comparable performance in all the infection states. Regarding the raw dataset, LOF showed a slight edge over the COF. For the hourly and daily smoothed dataset, both the models achieved comparable performance. The performance (score) plot of the models can be found in *Appendix A of this*

*dissertation*, depicting the capability of each model in detecting the infection state from the regular period. These models were tested on the entire patient-year.



**Figure 36:** Univariate input features - performance comparison (F1-score) of the unsupervised models using univariate input, i.e. insulin to carbohydrate ratio.

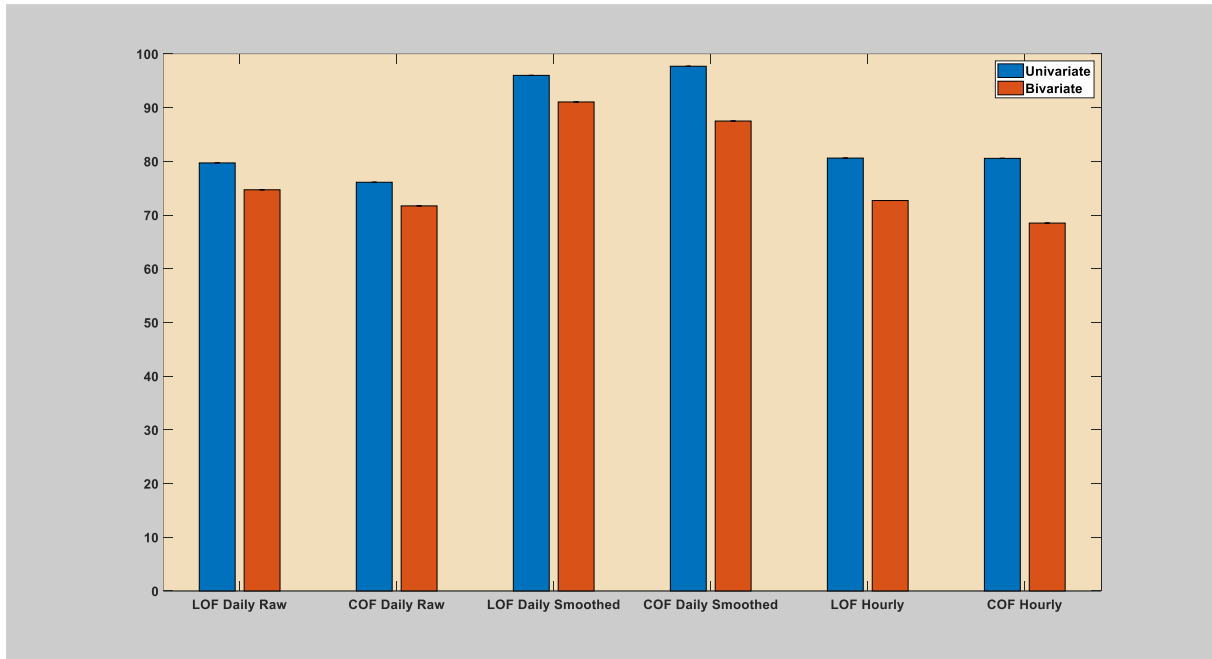
**Table 13:** Univariate input features - optimal values of thresholds used in performance evaluation. The values given as  $T_h$  are the optimal threshold values used for each patient-year,  $h$  depicting that particular year.

Granularity	Pre-pro.	Model (Threshold <sub>patient-year</sub> )
Daily	Without filter	LOF ( $T_1=2.7, T_2=1.5, T_3=2.95, T_4=2.2$ )
		COF ( $T_1=1.4, T_2=1.1, T_3=2.3, T_4=1.8$ )
	With filter	LOF ( $T_1=1.9, T_2=1.9, T_3=2.8, T_4=2.8$ )
		COF ( $T_1=1.6, T_2=1.5, T_3=2.8, T_4=3.1$ )
Hourly	With filter	LOF ( $T_1=1.9, T_2=1.6, T_3=1.2, T_4=1.7$ )
		COF ( $T_1=1.6, T_2=1.3, T_3=1.2, T_4=1.3$ )

### 6.3.4.3 Comparison of input features

This section presents the comparison of the average performance that can be gained by utilizing either of the bivariate or univariate input features. To this end, the performance of the unsupervised models, LOF and COF, using blood glucose levels and insulin-to-carbohydrate ratio versus the single insulin-to-carbohydrate ratio, is given in **Figure 37**. Per the findings, the models with a univariate input feature have achieved improved performance compared to the

bivariate input. This improvement could be linked with the discriminative power of the ratio compared to blood glucose levels.

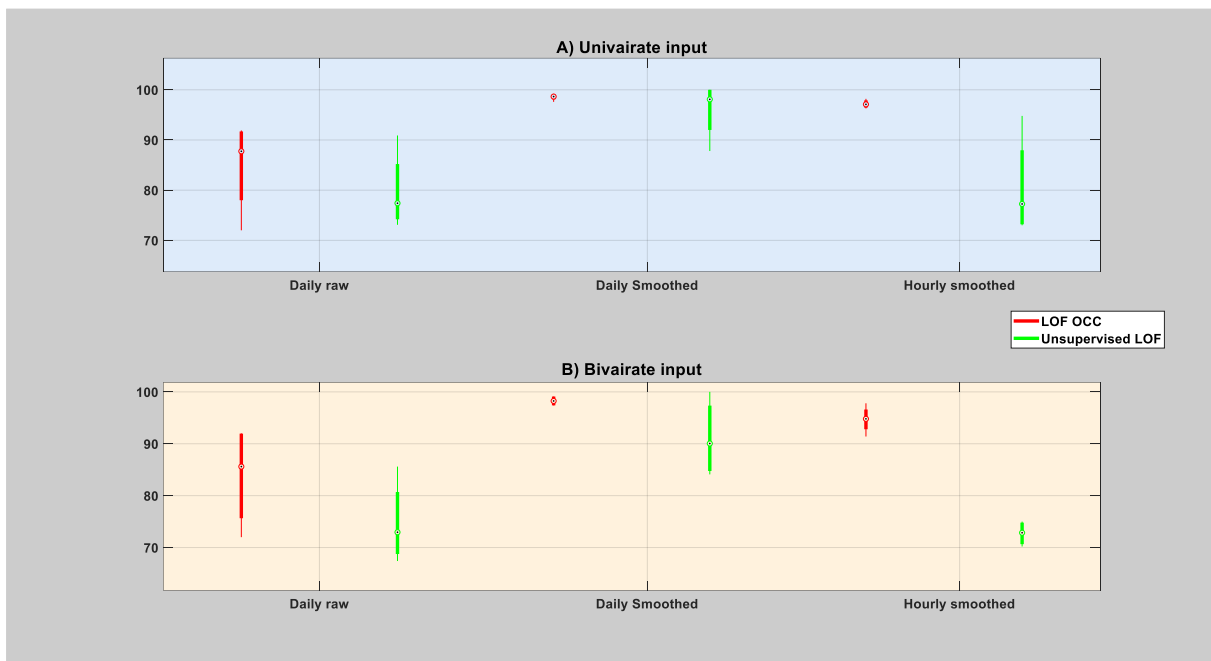


**Figure 37:** Performance comparison (F1-score) of the unsupervised models using a bivariate input, i.e. blood glucose levels and insulin to carbohydrate ratio versus a univariate input, i.e. insulin to carbohydrate ratio, source partially from [258], Table 8.

### 6.3.5 Comparison of unsupervised versus one-class classifier

This section presents the comparative performance of the one-class classifier and the unsupervised method. In this regard, as per the findings, compared to the semi-supervised method, i.e. one-class classifier, the unsupervised method fails to achieve comparable performance, especially with the bivariate input feature. One of the drawbacks of the unsupervised method is related to the fact that they require a fairly large sample size to at least produce comparable performance with the one-class classifiers [87]. This characteristic can be easily observed by looking at the performance of the unsupervised method based on the whole patient-year and the one-class classifier trained only with four months of the patient-year. To further illustrate the difference in performance, the comparison of the best performing unsupervised model, LOF, to its one-class classifier version, is given in **Figure 38**. As can be seen from the figure, under almost all the circumstances the one-class classifier model achieved superior performance. The characteristic of the data distribution, which contains a high and sparse density pattern, could be the reason behind the performance degradation of the unsupervised method [176] since these models require perfect demarcation between normal and abnormal values. As described in the input features, the state of blood glucose dynamics contains very rare events that contribute to the existence of sparse regions within the four quadrants. In this regard, a typical example could be a holiday season, where an individual happens to consume too many carbohydrates. Furthermore, the individual decision to switch to physical activity or exercise sessions to compensate for insulin requirements could result in a

similar pattern. These sparse data or rare events are a normal portion of the data, which needs to be treated as such by the detection methods. However, in these typical scenarios, the unsupervised method could end up considering these situations as abnormal resulting in a false alarm. In this regard, one of the main drawbacks of the unsupervised method is related to the fact that they determine anomalies from the data themselves and, there is no mechanism to let the model learn and accept certain sparse regions just like the one-class classifiers. Thus, the atypical nature of the underlying data distribution affects the performance of an unsupervised method. Yet, the one-class classifier method can handle such kind of situation if properly introduced with such an example during the learning phase [258].

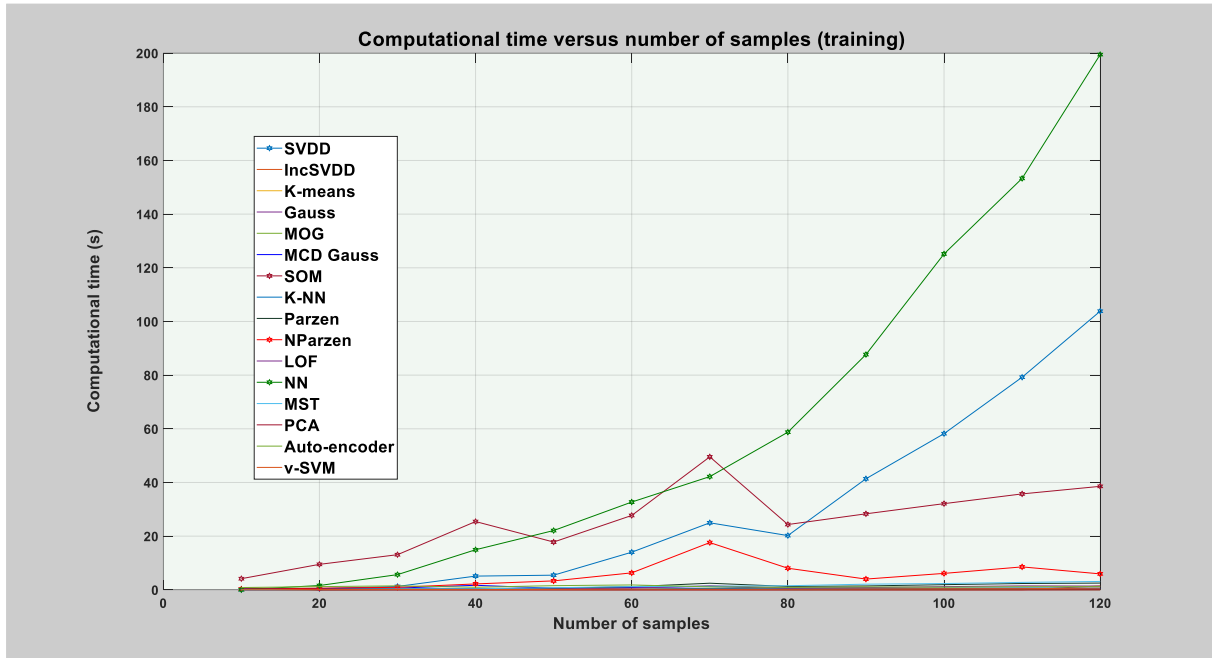


**Figure 38:** Performance comparison (F1-score) of unsupervised and one-class classifier version of the local outlier factor (LOF) model using both a bivariate input, i.e. blood glucose levels and insulin to carbohydrate ratio, and a univariate input, i.e. insulin to carbohydrate ratio.

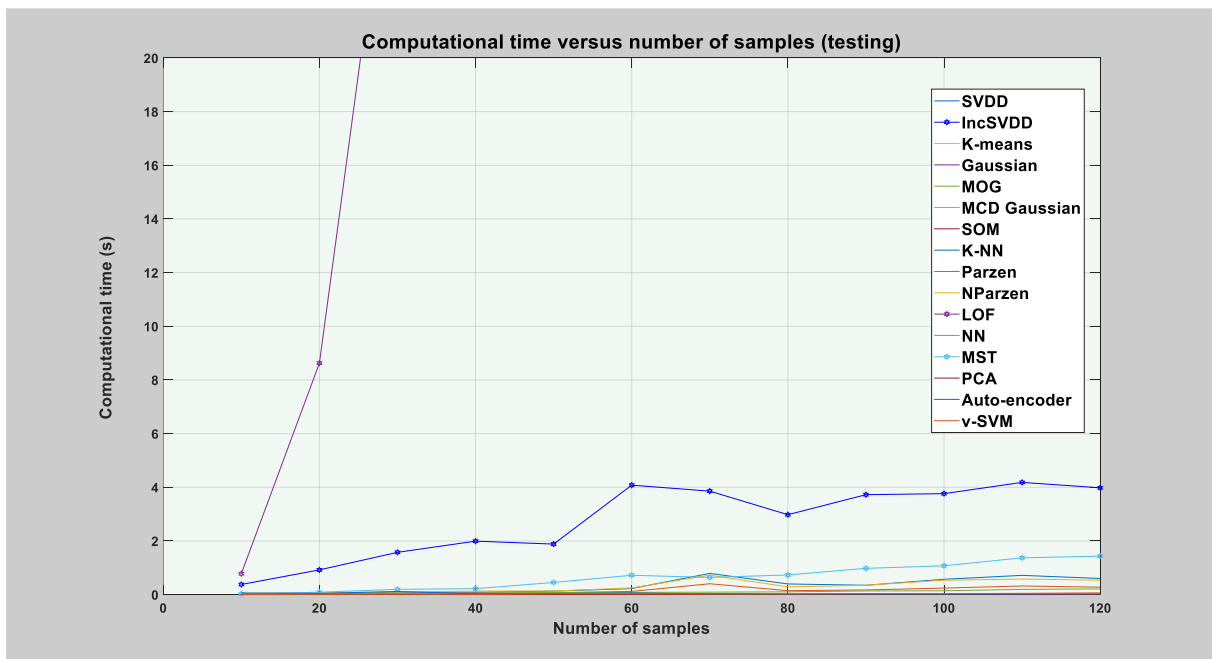
### 6.3.6 Computational Time

Models computational (running) time depicts the necessary time required by a model to learn and classify given sample objects [217; 258]. In real-world applications, it can become a bottleneck for a large scale implementation, i.e. real-time mass outbreak detection. In a certain application, however, the learning time can be compromised considering that the model can be trained offline [258]. However, the testing time is very crucial in almost all applications and is one of the determinants of the model's success in achieving the proposed real-time task. To this end, the model's average running time (learning and testing time) is estimated and compared. To quantify and compare the required computational time by each model, different sample sizes of bivariate training and testing data were used, i.e. 240, 480, 720, 960, 1200, 1440, 1680, 1920, 2160, 2400, 2640, 2880 objects. The resulting rough (average) computational time requirements of the models is given in **Figure 39 (a & b)**. Per the finding, generally, models

that took a significant amount of training time include NN, SVDD, and SOM. During the detection phase, almost all the models took considerably less time except certain models such as LOF and Incsvdd [258].



a) The computational time for the training phase, source [258], Figure 5.



b) The computational time for the detection phase, source [258], Figure 6.

**Figure 39:** Average running time required for the training and detection phase. Figure (a) depicts the models' time requirement during the training phase. Figure (b) depicts the models' time requirement during the detection phase. The labels in X-axis stands for a quotient of the sample size divided by twenty-four, i.e.10 (10\*24), 20 (20\*24), 30 (30\*24), 40 (40\*24), 50



(50\*24), 60 (60\*24), 70 (70\*24), 80 (80\*24), 90 (90\*24), 100 (100\*24), 110 (110\*24), 120 (120\*24) sample objects. The models' average computational time requirement is depicted on the Y-axis.

### 6.3.7 Practical significance

The practical significance of the presented personalized health model can be viewed within the scope of infectious disease outbreak detection, as demonstrated in this dissertation, and assisting the individuals during infection episodes, i.e. decision support and learning platform [258]. In this regard, the ability to detecting infection onset in real-time and under free-living conditions can be devised to provide crucial and supportive information for decision making [258; 259]. The following information can be vital to manage the crises;

- As discussed in *chapter five*, infection episodes induce elevated blood glucose levels despite administering proper insulin and consuming less carbohydrate [259]. In this regard, having real-time health status information can assist to provide information on blood glucose evolution during the course of infection, such as time in-range blood glucose levels [80], and glucose variability [50; 207; 251]. Such information coupled with insulin sensitivity change could be very important to the individual. This is mainly important in short term to mitigate the crises and also in the long term to learn how to cope with a similar situation in the future.
- The other big challenge for the individuals during infection onset is setting right the ratio of insulin to carbohydrate while taking a closer look at the blood glucose levels. As discussed in *chapter five*, the reason is the presence of insulin resistance and glucose production as the body's response mechanism to the infection-causing pathogens [259]. Thus providing information regarding the evolution of the ratio during the crises can help the individuals to estimate and set right the ratio during the crises. Further, this kind of information can enable the individual to be aware of the degree of insulin sensitivity change during infection episodes to better cope in the future.

Apart from this, the other important area could be a learning platform that can be used for educational purposes. In this regard, it could be essential to educating individuals about the effect of different infection pathogens and their associated deviations triggered on the key parameters of blood glucose dynamics [259]. However, to use such a service in practice require to analyze data containing different kind of pathogens and a large population to estimate the degree of deviations associated with each pathogen among different individuals. The idea here is to devise a simulation model, where a user can enter his body mass index, age, infection type, days since infection begins, and other similar information to see his/her estimated range of blood glucose levels, and change in the ratio (insulin sensitivity). In this regard, educating individuals with such simulated information about what to anticipate at each phase of the disease progression could be very essential [258; 259].

## 6.4 Limitation

Taking into account the complexity of blood glucose dynamics, the development of the proposed personalized health model requires considering a dataset from a large number of participants and incorporating all the key parameters of the blood glucose dynamics. In this regard, this study could be benefited from incorporating physical activity data and other information. Furthermore, utilizing data from a large number of participants could further validate the conclusion. However, it was very challenging and difficult for me to get a rich and precise dataset to further strengthen the conclusion. Actually, my inclusion criteria were stiff given the fact that I only considered datasets that are complete, accurate, and long enough (more than 3 months) as well as are required to contain at least one infection episode, which makes the data collection challenge far more difficult.

## 6.5 Knowledge Summary

In the context of the original research question, the following section presents the added knowledge within the scope of the presented results.

*What do we know about the topic already?*

- Semi-supervised (one-class classifier) and unsupervised methods have been used in various medical applications such as diagnosis and monitoring [57; 110; 164; 230].
- There is literature that exploits data from different sources such as google search, school, and works absenteeism, wearables data such as resting heart rate and sleep quality, pharmacy drug sales, and others to forecast infectious disease outbreaks [181; 194; 227]. However, none of this literature has used a personalized health model to screen the health status of an individual in near real-time and under free-living conditions based on their self-recorded health-related digital data for infectious disease outbreak detection purposes.

*What does this chapter add to our knowledge?*

- To the best of my knowledge, this is the first attempt towards realizing a personalized health model to screen health status and capture infection episodes among people with type 1 diabetes using self-recorded data.
- Approaches and design alternatives for realizing a personalized health model is presented. This model is the core and building block of the proposed personalized digital infectious disease detection system.
- A personalized health model for health status monitoring and detection of infection episodes among people with type 1 diabetes is realized. The experimental result indicates the success of the proposed approach achieving excellent performance in monitoring health status and detecting the infection episodes.
- Models with two alternative input features were evaluated and reported; bivariate input, i.e. blood glucose levels and insulin-to-carbohydrate ratio, and univariate input, i.e. insulin-to-carbohydrate ratio. The point is to consider the univariate input as an alternative when there is a challenge in getting blood glucose measurements. The

experimental result demonstrated the potential of these input features in the proposed tasks. Among these two input alternatives, the univariate input feature achieved slightly better performance.

- Two approaches for setting the frequency of monitoring are evaluated and reported. In this regard, the models were tested in two different data granularities, i.e. hourly and daily. The experimental results demonstrated the capability of these approaches achieving excellent performance in both data granularities. This capability allows the proposed system to be developed considering both frequencies of monitoring, where the personalized health model can run and perform computation either every hour or every day, or both.
- Within the scope of the realized personalized health model, the initial data requirement from an individual to join the proposed system is evaluated and reported. The assessment mainly considers the required sample size to generate acceptable model performance in real-world settings. In this regard, the experimental result indicates that every new participant joining such a system is required to fulfill minimum pre-collected data requirement, and per the findings, a sample size of three-month training sample size for the daily raw dataset (60 sample objects), a one-month (30 sample objects) training sample size for the daily smoothed dataset, and a two-month training sample sizes for the hourly dataset (60\*24 objects) can be regarded as a rule of thumb data size, to begin with. Of course, the presence of more data at the beginning will be favored, however, this is not the case in practical settings. The point here was to establish a rule of thumb data size for everybody to join the proposed system without compromising the system's accuracy.
- Among the models tested, on average, models such as v-SVM, K-NN, and K-means achieved relatively better performance with all evaluation criteria. The unsupervised models underperformed compared to the one-class classifiers, and mainly due to the nature of the data distribution.
- The degree and severity of infection episodes dictate its detectability through the developed personalized health models. In this regard, the experimental result demonstrated that infection episodes, which trigger large and medium deviations are captured with greater accuracy by the models. As per the findings, influenza (flu) episodes were detected with superior accuracy, and infection such as mild common cold without fever was also detected. However, infection such as light cold without fever was not detected.

## 6.6 Chapter summary

This chapter presented and discussed results related to the development of a personalized health model for detecting infection episodes among people with type 1 diabetes. A group of one-class classifiers and unsupervised methods were tested, evaluated and performance assessed. The one-class classifier models were evaluated in terms of *input features*, i.e. bivariate and univariate, *data granularity*, i.e. daily and hourly, and *required sample sizes*, i.e. one, two, three, and four months. The unsupervised models were evaluated with all these criteria but the sample size. Generally, the experimental results demonstrated the potential of the proposed approaches in achieving the intended task. In this regard, the proposed methods especially the one-class

classifiers have demonstrated superior performance in detecting deviations from the norm due to infection onset. In comparison to each particular one-class classifier model from their respective groups, v-SVM, KNN, and, K-means achieved relatively better performance on average with all the evaluation criteria. From the unsupervised models, LOF has demonstrated a slight edge over COF with all the evaluation criteria. Comparing the one-class classifier with the unsupervised method, the experimental result demonstrated the potential of the former to the latter in these typical datasets. As far as my knowledge is concerned, this is the first study that targeted a personalized health model to detect infection episodes among people with type 1 diabetes exploiting self-recorded data. In conclusion, these findings provide a novel approach to support the ongoing surveillance efforts, and however, additional large-scale studies might be needed to further strengthen the conclusion.

## 7 Concerns, Expectations, and Willingness Towards Sharing Self-Collected Health-Related Data

*Synopsis:* This chapter puts forward findings related to user concerns, expectations, and willingness for long-term and successful sharing of self-collected health-related data. The objective was to identify factors that enhance users' participation in the sharing of self-recorded health-related data to the proposed system. The chapter presents results from a quantitative survey conducted to identify these factors and closes with a concluding remark. This chapter provides answers to the fourth research question (Q4).

### 7.1 Introduction

As presented and discussed in *chapter four*, the fundamental requirements in the proposed EDMON system, are the availability of continuous, accurate, and precise self-recorded data from the individual participants, and the willingness to share these data with the center for further processing [259]. In fact, the individual participant usually self-records various parameters of the blood glucose dynamics as part of his/her self-management practice. In this regard, the idea behind the EDMON system is to use these data for secondary purposes, i.e. infectious disease outbreak detection. This objective especially calls for a user, who is motivated to continuously record accurate and precise health data and also willing to share the data with the intended system. However, from a practical perspective, peoples motivation to self-record accurate and precise data and willingness to share them vary considerably, and there are several factors, concerns, and barriers to realizing these requirements in practical settings [9; 24; 53; 60; 83; 93; 126; 189; 200; 201; 244; 245; 247; 249; 250; 252]. Data sharing concerns, barriers, and willingness could vary under different circumstances, however, factors such as lack of trust, incentives, and reciprocity, and concerns such as data ownership, access controls, and confidentiality are among those depicted in the literature [177; 229; 246; 247]. To be successful, a system like EDNON needs to consider different factors, concerns, barriers, and enablers during system design and implementation. Therefore, this chapter focus on assessing factors related to user concerns, expectations, and willingness for long-term and successful mass sharing of self-collected health-related data. Moreover, it also assesses the user mHealth apps feature preference and experience with health tracking technologies during data collection. The study was carried out in two steps using *exploratory sequential method* [75; 171; 239], where we primarily designed a qualitative interview guide incorporating five separate themes and performed data collection and analysis, and then used the findings from the qualitative study to further inform and rectify design of the quantitative survey questioners [105] and data collection. The qualitative study was developed based on a detailed concept and application scenarios of the proposed EDMON system, in general, and the patient unit, i.e. mHealth app, in particular [7; 130]. The initial qualitative exploration assesses factors within the scope of four main themes; sensors and wearables, data sharing, data integration, and social media and entertainment. Within these themes, face-to-face interviews were conducted assessing user knowledge, experience, and expectation [252]. The results from this qualitative exploration were used to develop the quantitative survey questionnaire [105]. To pinpoint factors that stand out for people with diabetes, the study performed a comparative analysis against other chronic patients, and healthy individuals [257]. In total 430 participants responded to the survey, of

which 61 individuals with diabetes, 82 individuals with other chronic diseases, and 285 healthy individuals. The respondents were from Norway (59), Switzerland (187), Germany (13), US/UK/Australia/Canada (77), France (26), and the other 46 respondents were from 35 different countries over the globe [257]. As can be seen from **Table 14**, there doesn't exist an age difference ( $p=.083$ ) or gender difference ( $p=.133$ ) among the study groups. The results presented in this chapter are part of the findings presented in *Paper 5* [257].

**Table 14:** Participants demographics, familiarity with wearable technologies, and data sharing experiences. Na depicts the number of respondents who declined to answer, source [257], Table 1.

Variable	Diabetes	Other chronic	Healthy individual
Age: <30 y; 30-50y; >50y	15; 17; 27 (Na: 2)	9; 32; 40 (Na: 1)	34; 96; 146 (Na: 6)
Gender: female; male; other	35; 25; 1	59; 20; 3	177; 99; 3
Wearable device: yes; no	59; 2	44; 36	143; 137
Sharing experience: yes; no	24; 27 (Na: 10)	22; 45 (Na: 14)	47; 160; (Na: 11)

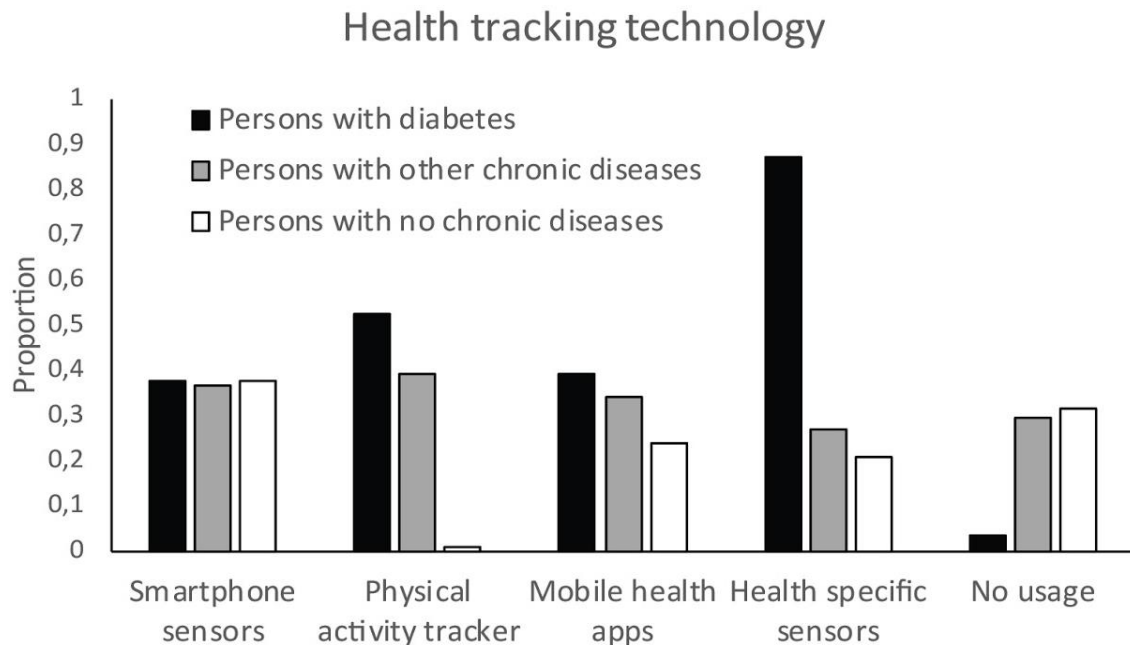
## 7.2 Data recording related experiences and preference

As described in *chapter four*, the proposed EDMON system considers mHealth app as a patient unit, which is expected to integrate readings from various diabetes-related sensors and other wearables technologies [259]. In this regard, this section explores the participant's experience and regular usage of health-tracking technologies and further spotted features of mHealth apps that the respondents rated important. For the EDMON system to be successful, it is crucial to take into account the users' preferences and experiences while designing and developing the patient unit [2; 77; 140; 252].

### 7.2.1 Experience and usage of health-tracking technologies

This section presents and discusses the result related to the participant's experience and regular usage of health tracking technologies. The respondents were presented with two questions assessing the individual's experiences and regular usage of health tracking technologies. The first question asked the respondents if they have experience with any wearable devices (Yes/No) and the second question presented the following options; sensors integrated into the smartphone, physical activity tracker, mobile health (mHealth) apps, health-specific measurements, and none [105], and asked if they regularly use any of these health-tracking technologies. As per the findings, as can be seen in **Table 14** above, many people with diabetes, close to 97%, in one way or another have used wearable device for self-management purposes through collecting physical activity and other health-related data, as compared to people with other chronic diseases (55%), and healthy individuals (51%),  $\chi^2(423) = 44.04$ ,  $p=.001$  [257]. Furthermore, as shown in **Figure 40**, people with diabetes tend to often regularly use health-specific measurement devices and physical activity as compared to the other study groups. In this regard, the result demonstrated that most people with diabetes (87%) regularly use health-

specific devices in combination with a physical activity tracker, and only 3% reports no use of sensor or wearable devices [257]. Among the other study groups including people with other chronic diseases and healthy individuals, smartphones integrated with sensors and mobile health apps are popular.

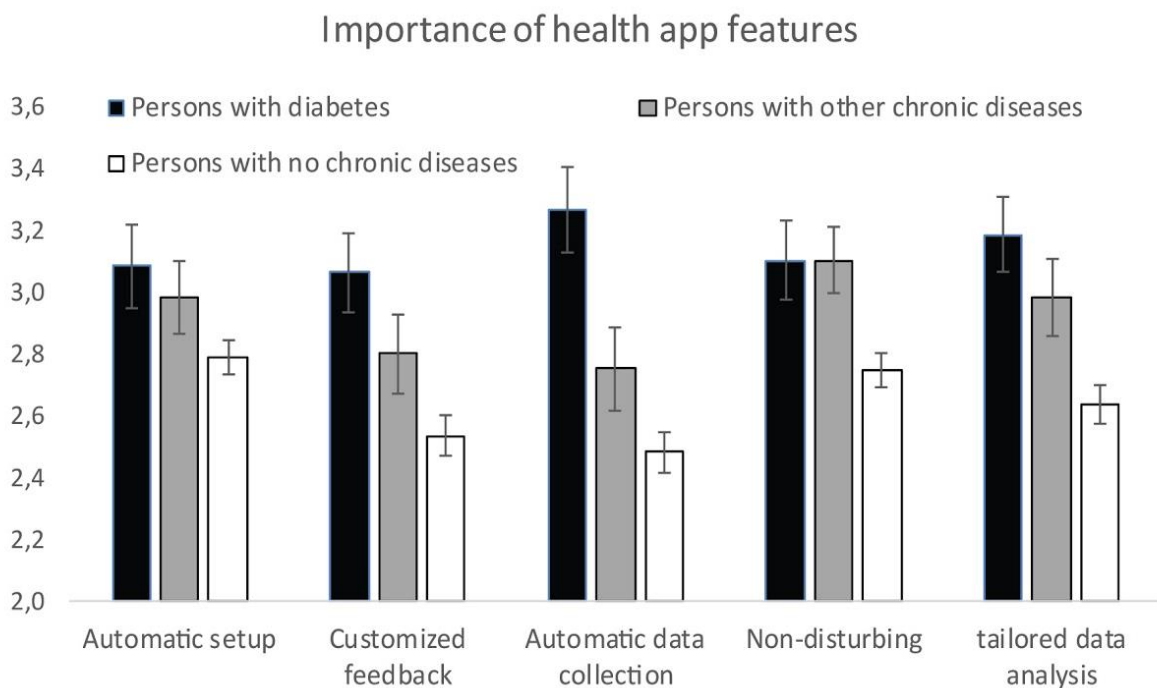


**Figure 40:** Health-tracking technologies usage and experience among the three groups, source [257], Figure 2 (b).

## 7.2.2 mHealth apps and features preference

This section presents and discusses the results related to the participant's preferences for mHealth app features. As described in *chapter four*, the proposed EDMON system relies on a mHealth app for collecting the necessary data from the individual subject [259], and, in this regard, it is crucial to understand what kind of features this group of patients prefer. Existing literature outlined the importance of mHealth app features and functionalities for its rating among the users, and are determinant factors that could have a positive impact on the user's long term engagement [111; 156; 252]. Taking this into account, the respondents were asked to rate these features; automatic setup and easy to understand/use, customizable feedback, automatic data collection, non-disturbing, and tailored data analysis (1=not important to 4=very important) [105]. The respondent's feature preferences are given in **Figure 41**. As per the findings, non-disturbing features that require fewer interactions and features that allow automatic setup, and are easy to understand (use) is rated as very important by most respondents, i.e. main effect of feature,  $F(3.83,1163.75)=3.389$ ,  $p=.01$ ,  $\eta^2=.011$  [257]. Generally, people with chronic disease rated all the features as equally important,  $F(2,304)=12.09$ ,  $p<.001$ ,  $\eta^2=.074$ . When it comes to people with diabetes, features that enable automatic data collection is rated as the most crucial feature, yielding a significant interaction

effect,  $F(7.66,1163.75)=2.104$ ,  $p=.035$ ,  $\eta^2=.014$ . In this regard, the most probable reason is that manual registration of data is more time consuming, and full of hassle. To this end, one respondent said, "What is missing in most of the diabetes app is an automatic integration of sensors that automatically collect data and different types of wearables for automatic collection of data. In the short term manual input is not a big problem but in the long term, it would be a lot of hassle to do that. Take a lot of time. Simple is a lot better " [252]. Furthermore, automatic setup and understandability, customized feedback, non-disturbing, and tailored analysis and functionalities were also favored by people with diabetes [257].



**Figure 41:** Preference of mobile health (mHealth) app features, source [257], Figure 1 (b).

### 7.3 Data sharing

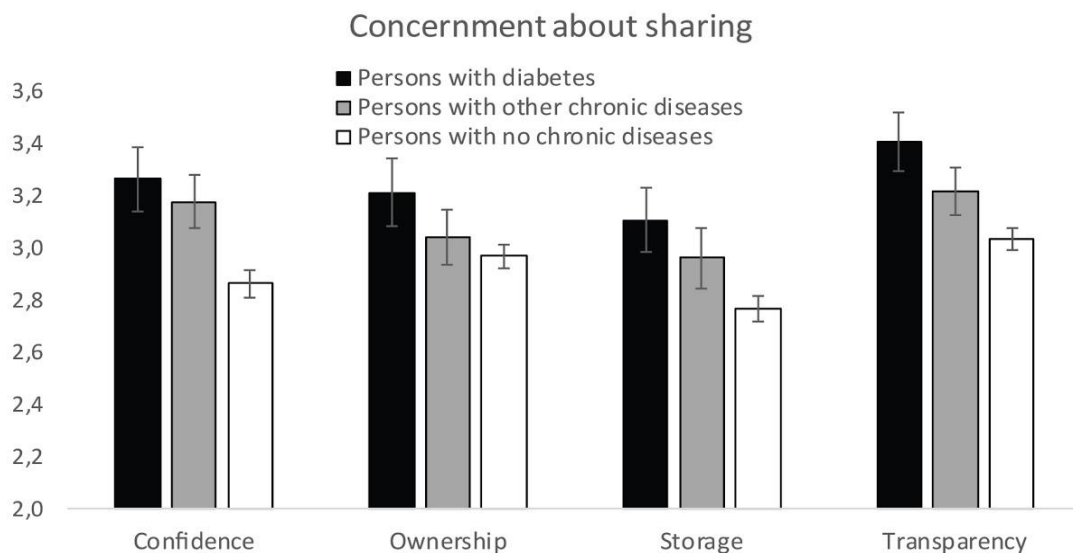
As discussed in *chapter four*, the proposed EDMON system expects the individual participant to continuously record accurate, and precise data and share these data with the center for further processing [259]. Taking into account the sensitivity of health-related data, sharing of such data faces significant challenges for several reasons [81; 117; 119; 126; 131; 225; 229]. To be successful, it is necessary to involve all the stakeholders throughout system design and implementation. In this regard, this section explores the participant's concerns, expectations, and willingness towards mass sharing of self-recorded health-related data to the intended system. Among the three groups, many people with diabetes reported data sharing experience (47%), as shown in **Table 14** above, compared to people with other chronic diseases (33%), and healthy individuals (23%),  $\chi^2=19.6$ ,  $p<.001$  [257].



### 7.3.1 Data sharing concerns and expectations

This section presents and discusses results related to the participant's concerns and expectations towards sharing data for the intended system. Understanding participant's concerns and expectations towards sharing self-collected health data are important steps for successful system design and development [15; 81; 117; 225; 257]. Taking this into account, the respondents were asked about their degree of concerns and what do they expect in return for sharing a given health data.

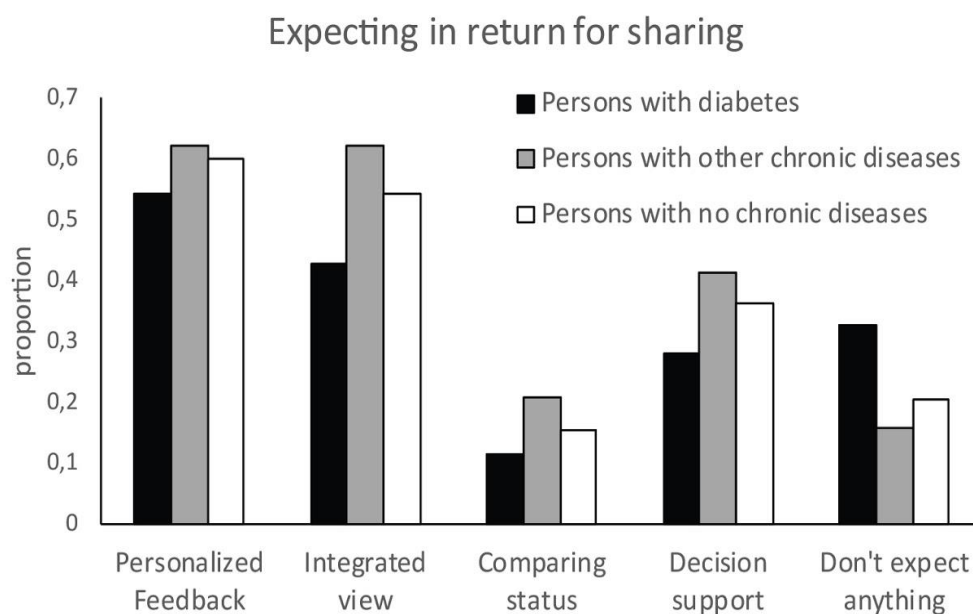
Several concerns could arise during data sharing [15; 117; 225], and in this regard, the respondents were asked to rate different concerns, which are expected to arise during data sharing, based on a 1 to 4 scale; trust-related issues (e.g. privacy/confidentiality/ security), data ownership - who owns the data, data storage - where is the data stored/service availability, and transparency of health data usage by 3<sup>rd</sup> parties [105]. As per the findings, as shown in **Figure 42**, for people with diabetes, transparency is the highly-rated concerns, followed by confidence related to data security and confidentiality (trust), ownership related to who owns the data irrespective of storage location, and storage related to where the data is stored. As can be seen from the figure, the group of healthy individuals is the least concerned with any of those stated issues, and this could be related to the type of data these groups usually collect, which is less health-specific [257]. All the respondents reported transparency as the most crucial concern and storage location as the last issue of concern.



**Figure 42:** Participants' concerns towards sharing health data, source [257], Figure 1 (a).

Understanding what kind of services the users might expect, i.e. incentives, in return for sharing data and addressing those expectations during system design and implementation could have a positive impact on a long-lasting engagement with the proposed EDMON system [225; 257].

To this end, the users' expectations were assessed by asking the respondents to rate the following features based on a 1 to 4 scale; personalized feedback and notification, integrated view (including data analysis, aggregated results, and trends), decision support, comparing status with others, and don't expect anything [105]. Per the findings, as shown in **Figure 43**, generally, 60% of the respondents (256 out of 429) rated personalized feedback as the most important feature, 53% of the respondents (233 out of 429) rated integrated view, 36% of the respondents (154 out of 429) rated decision support, and least rated feature was a status comparison, and only 16% of the respondents (68 out of 429) expected it in return [257]. Expectations among the three groups didn't differ, smallest  $p > .74$ . In this regard, providing such features as personalized feedback, integrated view, decision support, and comparing status with peer groups for people with diabetes could motivate this group for successful sharing of self-collected health data to the proposed EDMON system. Interestingly, a high number of people with diabetes also reported that they don't expect anything in return for sharing data (29%).



**Figure 43:** Participants' expectations in response to data sharing, source [257], Figure 2 (a).

### 7.3.2 Willingness towards anonymous sharing of data

This section presents and discusses results related to the participant's willingness to anonymously share different types of health-related data to the proposed EDMON system. User's willingness to share his/her health data depends on various factors such as the sensitiveness of the data type to be shared, views and attitude towards data sharing, perceived benefits and risks of data sharing, trust issues, fear of privacy, and data breaches, and a lot of other factors [117; 120; 257]. As described in *chapter four*, the proposed EDMON system, requires the individual participants to share mandatory data types (such as blood glucose levels,

insulin, diet, medication, and geographical location) and optional data types (such as physical activity, other physiological parameters, and others) [259]. Hence, the point was to assess the users' willingness to anonymously share these and other data types to the proposed EDMON system. To this end, the respondents were asked about their willingness to anonymously share health-related data to a central server. The following options were presented based on a multiple-choice format; lifestyle and dietary intake, medication intake and treatment, physiological indicators, physical activity and exercises data, geographical location (GPS), signs of infection, daily mood and feelings, weight, sleep duration, social environment (e.g. risks of infection), and none of these [105]. Per the findings, as shown in **Table 15**, generally, people with diabetes happen to be willing to share most parameters [257]. However, they are reserved for certain parameters such as geographical location (GPS), and signs of infection. In comparison between groups (corrected for multiple comparisons), generally, the respondents don't differ towards the sharing of certain parameters such as lifestyle and dietary intake, daily mood and feelings, sleep duration, geographical location, and social environment. In general, the respondents happen to be least willing (restrictive) towards sharing information about their geographical location (GPS). As described above, geographical location information is one of the mandatory data types that need to be shared with the proposed EDMON system for performing a clustering and outbreak detection analysis. In an ideal scenario, having fine-grained information such as where the individual work, live, school, and recently been can assist in tracing back contacts if an infectious outbreak is detected [170; 199; 259; 261]. In this regard, the practical implication is that the proposed EDMON system needs to put a substantial effort to provide a proper privacy-preserving mechanism that can persuade the users or use simply use a high levels address such as postal code and avoid the individual's location [39; 259; 262].

**Table 15:** Participants' willingness towards sharing specific health-related data, source [257], Table 2.

	Medication intake and treatment	Lifestyle/dietary intakes	Signs of infection	Physiological indicators	Daily mood	PA	Geographical location	Sleep duration	Weight	Social environment	None of these
N	213	219	195	202	193	248	87	240	221	173	89
Diabetes	53	41	31	46	29	43	20	43	43	30	5
Other chronic	44	45	45	40	46	54	16	50	49	35	16
No chronic	116	133	119	116	118	151	51	147	129	108	68
Chi-square	43.48	9.05	5.22	24.39	5.73	8.93	6.92	8.29	15.48	2.87	7.59
p-value	<.001	.011	.074	<.001	.057	.012	.031	.016	<.001	.238	.023

## 7.4 Practical Implication

As presented and discussed in *chapter four*, the proposed EDMON system requires a continuous recording of accurate and precise data and sharing of these data to the system. However, in practical settings, there are a number of challenges that arise during system design and implementation and require proper mitigations options to properly address them. As per the findings, the challenge can arise from 1) lack of motivation to stick to the EDMON app for continuous, accurate, and precise data recording, 2) failure to address users concerns that arise in response to data sharing 3) failure to fulfill user expectations out of the system, and others [252; 257]. In this regard, **Table 16** provides the main findings in this regard and puts forward mitigation options for the challenges, and also highlights promising findings that foster the system implementation.

**Table 16:** Topics and findings related to data recording and sharing, requirements of the proposed EDMON system, and mitigation options for challenges that arise.

Category	Topics	Findings	EDMON system requirements	Importance and mitigation options
<b>Data recording</b>	<b>Wearable device usage experience</b>	Most people with diabetes, close to 97%, have reported experience with wearable devices and sensors.	<i>Accurate and precise data recording in continuous manners</i>	This is very relevant to accomplish the stated requirements; <i>accurate and precise data</i> recordings. Having well-experienced users is an asset.
	<b>Regular usage</b>	Most people with diabetes (87%) regularly use health-specific devices and activity tracker, and only 3% reports no use of sensor or wearable devices.		This is a great advantage and to some extent can guarantee the <i>continuity</i> of the data to the proposed EDMON system. Regular usage and active engagement is a requirement.
	<b>mHealth App feature preferences</b>	Most people with diabetes prefer automatic data collection, automatic setup, understandability, customized feedback, non-disturbing, and tailored analysis.		Designing and developing a usable mHealth app for long term engagement requires involving the users in the process and incorporating users' preferences. In this regard, adding these <i>features to the EDMON patient unit app</i> can facilitate long-lasting engagements.
<b>Data sharing</b>	<b>Experience</b>	Almost half of the people with diabetes reported data sharing experience (47%).	<i>Continuous (real-time) sharing of recorded data</i>	This is a promising level of experience that can be further nurtured through properly addressing concerns, expectations, and also carefully looking at willingness and conditions for willingness to certain types of data.
	<b>Concerns</b>	People with diabetes raised concerns including transparency, data security, confidentiality (trust), data ownership, and data storage location.		Failure to resolve users' concerns in regard to data sharing could have a great impact on hampering their motivations to engage with the system. Hence, it is crucial to address these users' concerns.

	<b>Expectation</b>	Personalized feedback, integrated view, decision support, status comparison.	Providing some kind of incentives to better help manage their condition in return to data sharing could motivate the individuals to better stick to the system for a long time. Hence, it is necessary to include these functionalities in the proposed EDMON system.
	<b>Willingness</b>	People with diabetes happen to be willing to share most parameters but geographical location.	Regarding the geographical location, users are clearly against it. Hence, a substantial effort is needed to persuade the users by providing strong privacy, security, and confidentiality measures or simply rely on higher-level address such as postal code, which is less sensitive.

## 7.5 Limitation

One of the main limitations of the study could be the number of participants. The study could be benefited more if the number of participants is large enough to further strengthen the conclusion. The other possible limitation could be linked with the unbalanced number of participants among the group used during the comparison. However, it should be noted that the survey was distributed in different social media and forms for longer periods, and the questionnaires were also developed in three different languages to reach broader audiences.

## 7.6 Knowledge summary

In the context of the original research question, the following section presents the added knowledge within the scope of the presented results.

*What do we know about the topic already?*

- There are rapid advancements in health tracking technology, and these technologies are being integrated into our daily life, and consequently, a large amount of data are being generated each day.
- Continuous collection of quality health data (active and long term engagement) and willingness to share these data are mainly hampered by various factors.
- Several pieces of literature have previously addressed the issue under different contexts and objectives.
- Data sharing in particular are affected by the underlying objective. For instance, most people could refrain from sharing if there is an underlying commercial objective rather than public or patient benefits.

*What does this chapter add to our knowledge?*

- The result identified that the rate of adoption of health tracking devices is increasing, for instance, among the diabetes participants, almost 87% of the respondents reported regular use of health-specific devices and activity tracker, and only 3% reports no use of such devices. Furthermore, 97% of the respondents reported experience with some type of wearable devices and sensors. Besides, features such as automatic data

collection, easier interface, e.g. voice command, tailored and personalized feedback are reported to enhance usage time and long-term engagement.

- The result indicates that user's willingness to share health-related data mainly depends on the type of data under consideration and related concerns such as privacy, security, confidentiality, transparency, and ownership. For people with diabetes, transparency and confidentiality of the data are found to be very critical when sharing self-management data. The underlying reason can be linked to fear associated with information leakage or data usage by 3<sup>rd</sup> parties. In this regard, for example, blood glucose information containing experiences of repeated lows blood glucose level (hypoglycemia), if discovered by a licensing authority, might end up suspending one's driving license.
- Fulfillment of user's expectations and resolving individual concerns could highly facilitate sharing health-related data to a central repository. In this regard, the result indicates that participants expect some kind of immediate benefits or incentives from sharing, and this includes tailored and personalized data analysis, integrated view, customized feedback, and others. However, comparison of status among peers in the society is found to be less relevant. The respondents also highlight benefits to the general public health as an additional facilitator.
- The result identified that despite anonymization, respondents still refrain from sharing some kind of data. The respondent's willingness to share anonymous health-related data does differ for different data types. For example, most of the participants don't want to share their geographical location for various reasons. However, on the other hand, participants are willing to share health-related data including medication, physiological parameters, weight, and others.

## 7.7 Chapter summary

This chapter presented and discussed the individual's concerns, expectations, and willingness towards successful sharing of self-collected data to the proposed EDMON system along with mHealth features that could enhance active usage time and engagement. The study relies on an *exploratory sequential method*, where the findings from a qualitative interview were used to further inform and rectify the design of the quantitative survey questioners and data collection. The qualitative interview guide was developed based on *a detailed concept and application scenarios of the proposed EDMON system in general and the patient unit, i.e. mHealth app, in particular*. The result identified that among people with diabetes, the rate of adoption of health tracking and fitness devices is higher. Besides, factors such as automatic data collection, easier interface, e.g. voice command, tailored and personalized feedback are reported to enhance usage time and long-term engagement. When it comes to sharing self-collected health-related data, the findings demonstrated that people with diabetes are willing to share certain parameters such as medication intake and treatments, physiological indicators, lifestyle and dietary intake, physical activity data, sleep duration, and weight. However, they are reserved for certain parameters such as geographical location (GPS), and signs of infection. In return to data sharing, people with diabetes expects some kinds of benefits or incentives such as tailored and personalized data analysis, integrated view, customized feedback, and others. There are various concerns people with diabetes have regarding data sharing, including transparency, confidence

related to data security and confidentiality (trust), ownership related to who owns the data, and storage related to where the data is stored. Generally, the findings presented in this chapter can support the effort towards enhancing user engagement and facilitate successful data sharing, however, further large scale study might be needed to further strengthen the conclusion.





## 8 Conclusion and Future Works

Generally, there has been a consensus about the effect of infection onset in people with type 1 diabetes, and it has been known for a long time, however, there has been little to no attempt towards exploiting self-recorded data from these patient group to detect infection onset on an individual basis; thereby assisting the effort towards detecting infectious disease outbreaks. Moreover, in the bodies of literature, little to no studies have previously considered infection onset as a major parameter while designing and modeling either prediction or anomaly detection services for this group of patients. These knowledge gaps have stimulated towards defining the main research problem (MP) as follows: *What is the effect of infection incidence on key parameters of blood glucose dynamics amongst people with type 1 diabetes, and how can self-recorded data from this patient group assist in detecting an infectious disease outbreak?*

### 8.1 Contribution

This section will highlight and discuss the main contributions within the scope of the original research problem and the objective of the dissertation.

**Contribution 1 (C1) - Paper 1 related to Q1.** *A general framework of a personalized health model-based digital infectious disease detection system for realizing a system that can collect, analyze, detect, and notify the concerned bodies about the ongoing outbreak.*

This contribution is linked to the first research question (**Q1**) and aims to provide a general framework for the proposed personalized health model-based-digital infection detection system, i.e. EDMON system. In literature, there are various implemented systems designed for diabetes-related service, i.e. remote monitoring [14; 109; 185; 193; 268], and disease surveillance purposes [8; 54; 55; 91; 227]. Therefore, the point was to integrate concepts from those systems and come up with a novel framework that utilizes self-recorded data from people with type 1 diabetes, and assist in detecting infectious disease outbreaks. The framework is one of a kind introducing the individualization concept for surveillance purposes in practical settings. In this regard, the framework put forward a personalized health model-based-digital infectious disease detection system that detects infection onset at an individual level and uses such information for detecting infectious disease outbreaks among the public by identifying a cluster of infected individuals based on time and geographical locations. Moreover, it also highlights challenges that could emanate from practical implementation. Therefore, the contribution is the framework itself, from where the actual system can be implemented.

*The entire work related to this contribution is presented and discussed in Chapter 4.*

**Contribution 2 (C2) - Paper 1 related to Q2.** *A proof of concept towards using key parameters of blood glucose dynamics for infection detection.*

This contribution is linked to the second research question (**Q2**) and aims to provide a characterization of the effect of infection episodes on key parameters of blood glucose dynamics amongst people with type 1 diabetes. Among other things, the characterization

identifies optimal parameters for infection detection by quantifying the nature and degree of infection triggered deviations. The point was to select optimal parameters from the group of key parameters of blood glucose dynamics to realize the proposed personalized health model with a minimum false alarm rate as possible. In this regard, some studies in the literature have investigated the possibility of using self-recorded data from this patient group for outbreak detection and surveillance purposes, *see chapter five*. However, none of these studies investigated and characterized the effect of infection episodes on each key parameter of the blood glucose dynamics and pinpoint optimal parameters with high accuracy. To the best of my knowledge, this is the first study that empirically and numerically quantifies the effect of infection episodes on each key parameter of the blood glucose dynamics among people with type 1 diabetes exploiting self-recorded data. The study carried out a trend analysis and data distribution estimation to pinpoint the optimal parameters relying on a real dataset of ten patient-years. Therefore, the contribution is the characterization of the effect of infection episodes, and identification of optimal parameters for developing the proposed personalized health model.

*The entire work related to this contribution is presented and discussed in **Chapter 5**.*

**Contribution 3 (C3) - Paper 3 related Q4. *A personalized health model for detecting infection incidences in people with type 1 diabetes using blood glucose, insulin, and carbohydrate information.***

This contribution is linked to the fourth research question (**Q4**) and aims to provide a personalized health model that is capable of monitoring the individual's health status in free-living conditions. To the best of my knowledge, this is the first attempt towards realizing a personalized health model to capture infection episodes among people with type 1 diabetes using self-recorded data. The model accepts the user's self-recorded data continuously and automatically detects when the individual becomes sick. The realized model is a type of contextual anomaly detector exploiting blood glucose levels and insulin-to-carbohydrate ratio, where blood glucose is the behavior to be observed and the ratio of insulin-to-carbohydrate is the context. A group of one-class classifiers and unsupervised models were tested and evaluated. Therefore, the contribution is the personalized health model.

*The entire work related to this contribution is presented and discussed in the **first half of Chapter 6**.*

**Contribution 4 (C4) - Paper 4 related Q4. *An alternative personalized health model for detecting infection incidences in people with type 1 diabetes using only insulin and carbohydrate information.***

This contribution is linked to the fourth research question (**Q4**) and aims to provide an alternative approach towards realizing a personalized health model for monitoring the individual's health status in free-living conditions, and automatic detection of infection onset.

The point was to realize a model with less number of input parameters, i.e. using only insulin-to-carbohydrate ratio, and provide an alternative model. The realized model is a type of point anomaly detector exploiting the insulin-to-carbohydrate ratio, where each point of the time-series is observed. Almost the same group of one-class classifiers and unsupervised models were tested and evaluated. Therefore, the contribution is the alternative personalized health model.

*The entire work related to this contribution is presented and discussed in the **second half of Chapter 6.***

**Contribution 5 (C5) - Paper 5 related to Q5. Assessments of user concerns, expectations, and willingness towards sharing self-collected health-related data.**

This contribution is linked to the fifth research question (**Q5**) and aims to provide a list of user concerns, expectations, and willingness for long-term engagement and successful sharing of health-related data to the proposed EDMON system. *An exploratory sequential method was used*, where we primarily designed a qualitative interview guide incorporating five separate themes and performed data collection and analysis, and then used the findings from the qualitative study to further inform and rectify the design of the quantitative survey questionnaires and data collection. In this regard, initially, a qualitative exploration, i.e. face-to-face interview, was conducted based on a *detailed concept and application scenarios* of the proposed EDMON system in general, and the patient unit, i.e. mHealth app, in particular. Then, the findings from the qualitative interviews were used as input to inform, rectify, and refine the quantitative survey questionnaires. The contribution is a list of user concerns, expectations, and willingness for long-term engagement and successful sharing of self-collected health-related data to the proposed EDMON system. These factors are essential and need to be taken into account during the design and implementation of the proposed EDMON system.

*The entire work related to this contribution is presented and discussed in **Chapter 7.***

## **8.2 Main Conclusion**

To effectively answer the main research problem, it is necessary to briefly address the different sub-questions as they were derived to address part of the main research problem.

**Question 1 (Q1).** *How is a personalized health model-based digital infectious disease detection system that collects self-recorded data from participants, analyses the data, detect deviations on an individual basis, identify a cluster of individual, and notify the status of an outbreak to be designed and implemented?*

To answer this in a commendable way, it was quite necessary to identify and address different sub-components of the proposed EDMON system and realize a framework that depicts the relationship and task requirements of the components. In this regard, the realized framework contains different units performing a series of tasks; patient unit, data repository (database) unit, infection detection unit, clustering unit, information visualization unit, and a wireless communication platform. The *patient unit* is a standalone smartphone app (mHealth app) that

integrates and stores different sensor readings, which could be through either manual or automatic recordings. The individual's data-structure should at least contain information on blood glucose levels, insulin and carbohydrate registration, geographical location, and time of registration. Currently, most diabetes technologies enable Bluetooth connections to foster the integration of sensor readings. The data stored in the mobile app needs to be transmitted to a *database server*, where it is stored for further processing. However, given the sensitivity of health data, high emphasis needs to be given to data security, privacy, and confidentiality. In this regard, the transmission and storage of data in a server need to strictly follow major international guidelines, e.g. HIPPA compliance. The *infection detection unit* should have access to the individual's records from the database and execute the personalized health model, which is trained on the individual's historical data, to look for any abnormal deviations promptly. At any given time, the output from the personalized health model is the individual's health status coded as normal (0), suspicious (-1), and infected (1). The *clustering unit* should accept the timely health status from each individual participant along with their respective geographical location and time of data registration to perform a Spatio-temporal analysis to detect a group of infected individuals under a region of surveillance. The status of the region or city under surveillance can be visualized based on either a standalone smartphone app (mHealth app) or a web-based or both. The information can be useful for the patient himself/herself, relatives, societies, healthcare, and public health officials.

**Question 2 (Q2).** *What is the effect of infection incidence on key parameters of blood glucose dynamics among people with type 1 diabetes, and which parameters can effectively be used for detecting infection incidences in people with type 1 diabetes?*

To answer this question, background work was conducted assessing previous similar works and existing knowledge about the effect of infection onset amongst people with type 1 diabetes. In this regard, particular emphasis was given to empirical studies that numerically assess the nature, and degree of infection triggered deviations on different parameters of the individual's blood glucose dynamics as compared to the regular period. A knowledge gap was identified and empirical data analysis was conducted using data containing real infection episodes [253; 259]. The empirical analysis demonstrated that compared to the regular period, infection onset triggers significant deviations from the typical norm of blood glucose dynamics. Per the findings, during the entire course of infection, blood glucose levels remain elevated regardless of higher insulin injections and consumption of lesser carbohydrates. This event marks the onset of infection in people with type 1 diabetes. During the regular period, blood glucose usually drops to the hypoglycemia region with higher insulin injections and lesser carbohydrate consumption. Hence, the individual diabetes profiles including blood glucose level, insulin injection, carbohydrate ingestion, and insulin-to-carbohydrate ratio are significantly affected by infection episode and can be successfully be used for developing a personalized infection detection model. Among these parameters, blood glucose level and the ratio of insulin-to-carbohydrate happen to be more informative of infection onset.

**Question 3 (Q3).** *What is the status regarding an infection detection system using self-recorded data from people with type 1 diabetes and to what extent do the existing personalized decision support systems, and blood glucose alarm events applications consider infection incidence and its effect while developing blood glucose anomalies detection algorithms?*

As a roadmap to the realization of a personalized health model for detecting infection onset, it was necessary to perform a review that assesses the existence of any previously developed and implemented infection detection system exploiting self-recorded data from people with type 1 diabetes, and also other existing systems such as personalized decision support system and blood glucose alarm event applications. The point here was to look after similar systems and also to evaluate how and to what extent do the existing blood glucose anomaly detection algorithm uses information related to infection. In this regard, there are few attempts to realize an infection detection system, *see chapter two*. However, it turns out that there doesn't exist any developed and implemented infection detection system and almost all the studies don't consider information related to infection episode as a parameter while developing anomalies detection and prediction model [253]. For instance, an anomalies detection algorithm designed to detect hyperglycemia episodes (high blood glucose levels) in people with type 1 diabetes consider only detecting the presence or absence of hyperglycemia and ignores the underlying cause, which in some cases might be infection onset.

**Question 4 (Q4).** *How to design and develop a personalized health model that can continuously monitor individual health status and automatically detect infection onset using self-recorded data from people with type 1 diabetes?*

To answer this question, it was necessary to look for a model that can continuously monitor the individual's health status and automatically detect infection onset in this patient group while reducing false alarms. This can be achieved via either a *blood glucose prediction approach* or an *anomaly detection approach*, *see chapter four*. However, the choice is quite dependent on the state-of-the-art performance of these approaches under free-living conditions. In this regard, given the state-of-the-art performance of blood glucose prediction models, which rapidly degrades after 30 min prediction horizon, it is natural to look for other possible approaches that can operate in hourly and daily timeframes (data granularity). The optimal approach, in this case, will be to rely on semi-supervised (one-class classifiers) and unsupervised anomaly detection strategies. One-class classifier requires to learn the domain knowledge (reference description) that accurately describes what is normal to detect deviations from normality. However, the unsupervised approach doesn't require training and detect deviation from normality based on the entire dataset presented during testing.

**Question 5 (Q5).** *What are the user concerns and expectations towards sharing self-collected health-related data with the proposed system? and what type of data are they willing to anonymously share?*

To answer this in a commendable way, it was necessary to perform either a qualitative interview or quantitative survey, or both to identify a list of factors based on either a prototype system or a detailed concept and application scenarios of the proposed EDMON system. In this regard, an exploratory sequential method was used, where findings from a qualitative face-to-face interview were used to further inform, rectify, and refine the quantitative survey questionnaires. Per the findings, factors such as automatic data collection, easier interface, e.g. voice command, tailored and personalized feedback were reported to enhance usage time and long-term engagement. People with diabetes are found to be willing to anonymously share most of the parameters including medication intake and treatments, physiological indicators, lifestyle and dietary intake, physical activity data, sleep duration, and weight. However, they are reserved for certain parameters such as geographical location (GPS), and signs of infection. In return to data sharing, people with diabetes expects tailored and personalized data analysis, integrated view, feedback, and others. There are various concerns people with diabetes have in relation to data sharing including transparency, confidence related to data security and confidentiality (trust), ownership related to who owns the data, and storage related to where the data is stored.

### **8.2.1 Main Research Problem – MP**

**MP:** *What is the effect of infection incidence on key parameters of blood glucose dynamics amongst people with type 1 diabetes, and how can self-recorded data from this patient group assist in detecting an infectious disease outbreak?*

Blood glucose dynamics are affected by numerous factors and infection onset is one of them. This dissertation has shown that infection onset amongst people with type 1 diabetes triggers a substantial shift in the operating point of the blood glucose dynamics, which results in prolonged hyperglycemia regardless of higher insulin injection and fewer carbohydrate consumptions, which contradicts the typical norm of blood glucose dynamics. It is important to note that these circumstances during the regular day usually result in severe hypoglycemia. Generally, per the findings, infection episode triggers substantial deviations on the key parameters of blood glucose dynamics; and blood glucose level, amount of insulin injection, amount of carbohydrate consumption, and the ratio of insulin-to-carbohydrate and can be regarded as optimal parameters for infection detections. In this regard, the results demonstrated the potential of the individual's diabetes profile containing these self-recorded data for realizing a personalized health model for continuous monitoring of an individual's health status, and automatic detection of infection onset in free-living conditions; thereby using such information for detecting infectious disease outbreaks by identifying a cluster of infected individuals based on time and geographical locations. In regard to the personalized health model, the experimental results demonstrated the potential of the models designed with bivariate input (blood glucose levels and insulin-to-carbohydrate ratio), and univariate input (insulin-to-carbohydrate ratio). The proposed personalized health model with a one-class classifier and unsupervised methods have shown significant detection capability.

The results presented have practical significance for understanding the effect of infection episodes amongst people with type 1 diabetes, and the nature of infection triggered deviations incurred on blood glucose dynamics. Given the criticality of detection time in infectious disease surveillance, recent efforts have been focused on analyzing different data sources, which are mainly aggregated data at population levels. In this regard, little to no efforts have been made to realize individualized surveillance based on a personalized health model exploiting self-recorded data. In this regard, the result presented in this dissertation has practical significance for understanding the potential of personalized health model in outbreak detection settings. Actually, the advantage behind the personalized health model concept introduced in this dissertation lies in its usefulness beyond the surveillance purpose. In this regard, the presented personalized health model can also be used for other purposes such as to devise decision support tools and learning platforms for the patient to properly manage infection-induced crises. Generally, developing a personalized health model-based digital infectious disease detection system like EDMON, which aims for early detection, i.e. during the incubation period, requires considering various aspects, and the results presented in this dissertation construct evidence that supports the efforts towards building the next generation personalized health model-based-digital infectious disease surveillance systems and provoke further thoughts in this challenging field.

### 8.3 Future Works

This dissertation produced several topics that need to be realized in long term for the proposed EDMON system, along with other important additional topics to be considered for further improvement.

***Prototype development:*** This dissertation has presented and discussed the required framework along with user concerns, expectations, and willingness for data sharing, and the necessary personalized health model to realize the prototype system. During the Ph.D. period, an initial prototype for the backend server, and mHealth app has been developed [58]. These components need further improvements and currently, the integration of different sensors into the mHealth app is underway. Therefore, the effort to realize the prototype system is still ongoing. During the entire system design and development, it is necessary to involve all the stakeholders.

***Effect of different pathogens:*** This dissertation has presented findings from a real incidence of influenza and common cold without fever and demonstrated its effect on key parameters of blood glucose dynamics. However, a large scale study is required to further validate the findings with a large number of participants. Moreover, it is necessary to perform analysis and identify the effect of other different pathogens on blood glucose dynamics.

***Blood glucose prediction model:*** This dissertation has developed the infection detection unit with an anomaly detection method and demonstrated its applicability to the task. The choice was made based on the state-of-the-art performance of these two methods. However, it is necessary to test the prediction based method and compare the performance with the anomaly detection method presented.

**Physical activity data:** The models and empirical analysis presented in this dissertation exploits the three key parameters of blood glucose dynamics: blood glucose levels, insulin, carbohydrate, and insulin-to-carbohydrate ratio, and the findings demonstrated the potential of these parameters. However, it is necessary to investigate the added benefit of utilizing physical activity or exercise sessions data.

**Disease Characterization:** After detection of an infectious disease outbreak comes disease characterization, which is an integral part of outbreak reporting. Disease characterization is a process by which the type of causative agent or pathogens is determined [237]. In the existing disease surveillance systems, characterization is often carried out mainly by investigating the pathogen's mode of transmission, epidemics curve (incubation period), and its symptoms manifestations [1; 59; 237]. However, in the context of people with type 1 diabetes, it might be possible to characterize the pathogens through the individual's self-recorded data. In this regard, there are studies that report the associations between the degree of infection-triggered deviations on blood glucose dynamics and the type of pathogens involved, and hence, thoroughly analyzing large-scale datasets incorporating a different type of infections or pathogens could characterize them via quantifying the deviations each pathogen induces on blood glucose dynamics.

**Detection time:** Early detection is a crucial characteristic of any infectious disease surveillance system. In this regard, a surveillance system capable of detecting infectious disease outbreaks during the incubation period is highly sought. To this end, there is evidence suggesting that people with type 1 diabetes sometimes experience elevated blood glucose levels before symptom onset. Therefore, it might be possible to detect infection onset in this patient group before the individual experiences symptoms. In this regard, it is necessary to investigate and analyze CGM features within the context of other parameters of blood glucose dynamics, and along with different physiological parameters such as heart rate, body temperature, and other vital signs.

**Harnessing data from wearables:** The ubiquitous nature of smartphones and wearables equipped with a variety of physiological sensors have presented a huge potential and unbounded possibility to realize a personalized health model-based digital infectious disease detection system [181; 194]. A huge amount of data is being generated each day that can be geared toward detecting infectious disease outbreaks. As demonstrated in this dissertation, apart from its use in detecting infectious disease outbreaks, this type of early warning system can also address the need of having real-time health status information for different groups in the population, including chronic patients, aging populations, people using ambient assisted living, and healthy individuals. In literature, there is evidence suggesting the predictive potential of different physiological parameters in response to infection onset including heart rate, sleep quality, body temperature, and other similar variables [181]. Therefore, it might be possible to realize a personalized health model after performing a large scale study to further assess the predictive potential of these and other similar physiological parameters. The advantage of this approach lies in the fact that it can incorporate almost every individual in the population that uses wearable devices for their own purposes. From a surveillance perspective, having a large



population under surveillance is an advantage to produce statistically significant outbreak detection results.

## 9 References

- [1] Developing infectious disease surveillance systems, *Nature Communications* **11** (2020), 4962.
- [2] M.D. Adu, U.H. Malabu, A.E.O. Malau-Aduli, and B.S. Malau-Aduli, Users' preferences and design recommendations to promote engagements with mobile apps for diabetes self-management: Multi-national perspectives, *PLoS One* **13** (2018), e0208942.
- [3] M.D. Adu, U.H. Malabu, A.E.O. Malau-Aduli, and B.S. Malau-Aduli, The development of My Care Hub Mobile-Phone App to Support Self-Management in Australians with Type 1 or Type 2 Diabetes, *Scientific Reports* **10** (2020), 7.
- [4] S. Agrawal and J. Agrawal, Survey on Anomaly Detection using Data Mining Techniques, *Procedia Computer Science* **60** (2015), 708-713.
- [5] B.I. Ahmad, M.A. Ayu, I. Abdullahi, and Y. Yakubu, Remote patient monitoring system architecture for diabetes management, in: *2017 International Conference on Computing, Engineering, and Design (ICCED)*, 2017, pp. 1-6.
- [6] S. Ahmad, A. Lavin, S. Purdy, and Z. Agha, Unsupervised real-time anomaly detection for streaming data, *Neurocomputing* **262** (2017), 134-147.
- [7] A. Ahtinen, *Mobile Applications to Support Physical Exercise - Motivational Factors and Design Strategies*, Tampere University of Technology, 2015.
- [8] E.L. Aiken, S.F. McGough, M.S. Majumder, G. Wachtel, A.T. Nguyen, C. Viboud, and M. Santillana, Real-time estimation of disease activity in emerging outbreaks using internet search information, *PLoS Comput Biol* **16** (2020), e1008117.
- [9] M. Aitken, J. de St Jorre, C. Pagliari, R. Jepson, and S. Cunningham-Burley, Public responses to the sharing and linkage of health data for research purposes: a systematic review and thematic synthesis of qualitative studies, *BMC medical ethics* **17** (2016), 73-73.
- [10] I. Ajmera, M. Swat, C. Laibe, N. Le Novere, and V. Chelliah, The impact of mathematical modeling on the understanding of diabetes and related complications, *CPT Pharmacometrics Syst Pharmacol* **2** (2013), e54.
- [11] H.K. Akturk and S. Garg, Technological advances shaping diabetes care, *Curr Opin Endocrinol Diabetes Obes* **26** (2019), 84-89.
- [12] M. Al-khafajiy, T. Baker, C. Chalmers, M. Asim, H. Kolivand, M. Fahim, and A. Waraich, Remote health monitoring of elderly through wearable sensors, *Multimedia Tools and Applications* **78** (2019), 24681-24706.
- [13] M.A. Al-Tae, W. Al-Nuaimy, A. Al-Ataby, Z.J. Muhsin, and S.N. Abood, Mobile health platform for diabetes management based on the Internet-of-Things, in: *2015 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT)*, 2015, pp. 1-5.
- [14] G. Alfian, M. Syafrudin, M.F. Ijaz, M.A. Syaekhoni, N.L. Fitriyani, and J. Rhee, A Personalized Healthcare Monitoring System for Diabetic Patients by Utilizing BLE-Based Sensors and Real-Time Data Processing, *Sensors (Basel)* **18** (2018).
- [15] A. Alqhatani and H.R. Lipford, "There is nothing that I need to keep secret": Sharing Practices and Concerns of Wearable Fitness Data, in: *Fifteenth Symposium on Usable Privacy and Security* 2019.
- [16] E. Årsand, M. Muzny, M. Bradway, J. Muzik, and G. Hartvigsen, Performance of the First Combined Smartwatch and Smartphone Diabetes Diary Application Study, *Journal of Diabetes Science and Technology* **9** (2015), 556-563.
- [17] E. Årsand, O. Walseth, N. Andersson, R. Fernando, O. Granberg, J. Bellika, and G. Hartvigsen, Using Blood Glucose Data as an Indicator for Epidemic Disease Outbreaks, *Studies in Health Technology and Informatics* (2005), 217 - 222.
- [18] H. Azami, K. Mohammadi, and B. Bozorgtabar, An Improved Signal Segmentation Using Moving Average and Savitzky-Golay Filter, *Journal of Signal and Information Processing* **03** (2012), 39-44.
- [19] T.S. Bailey, J. Walsh, and J.Y. Stone, Emerging Technologies for Diabetes Care, *Diabetes Technol Ther* **20** (2018), S278-S284.
- [20] H. Banaee, M.U. Ahmed, and A. Loutfi, Data mining for wearable sensors in health monitoring systems: a review of recent trends and challenges, *Sensors (Basel)* **13** (2013), 17472-17500.

- [21] D. Baud, X. Qi, K. Nielsen-Saines, D. Musso, L. Pomar, and G. Favre, Real estimates of mortality following COVID-19 infection, *The Lancet Infectious Diseases* **20** (2020), 773.
- [22] N.A. Bazaev, A.N. Pletenev, and K.V. Pozhar, Classification of Factors Affecting Blood Glucose Concentration Dynamics, *Biomedical Engineering* **47** (2013), 100-103.
- [23] R.W. Beck, R.M. Bergenstal, L.M. Laffel, and J.C. Pickup, Advances in technology for management of type 1 diabetes, *The Lancet* **394** (2019), 1265-1273.
- [24] E.A. Bell, L. Ohno-Machado, and M.A. Grando, Sharing my health data: a survey of data sharing preferences of healthy individuals, *AMIA ... Annual Symposium proceedings. AMIA Symposium* **2014** (2014), 1699-1708.
- [25] S. Benjamin, A. Kirstin, M.F. Sarah, C. Anoushka, D. Stephan, P. Karena, R. Adam, M.H. Frederick, and E.M. Ashley, Feasibility of continuous fever monitoring using wearable devices, *Research Square* (2020).
- [26] W.C. Bevier, H.C. Zisser, L. Jovanovic, D.A. Finan, C.C. Palerm, D.E. Seborg, and F.J. Doyle, 3rd, Use of continuous glucose monitoring to estimate insulin requirements in patients with type 1 diabetes mellitus during a short course of prednisone, *J Diabetes Sci Technol* **2** (2008), 578-583.
- [27] L.B. Borghouts and H.A. Keizer, Exercise and insulin sensitivity: a review, *Int J Sports Med* **21** (2000), 1-12.
- [28] Z.I. Botev, J.F. Grotowski, and D.P. Kroese, Kernel density estimation via diffusion, *The annals of Statistics* **38** (2010), 2916-2957.
- [29] T. Botsis, J.G. Bellika, and G. Hartvigsen, Disease surveillance systems for sensitive population groups, in: *Advances in Disease Surveillance*, 2007.
- [30] T. Botsis, J.G. Bellika, and G. Hartvigsen, New Directions in Electronic Disease Surveillance: Detection of Infectious Diseases during the Incubation Period, in: *International Conference on eHealth, Telemedicine, and Social Medicine*, IEEE, Cancun, Mexico, 2009, pp. 176-183.
- [31] T. Botsis and G. Hartvigsen, Exploring new directions in disease surveillance for people with diabetes: Lessons learned and future plans, *Studies in Health Technology and Informatics* **160** (2010), 466 - 470.
- [32] T. Botsis, O. Hejlesen, J. Bellika, and G. Hartvigsen, Blood glucose levels as a censor for early detection of infection in type-1 diabetics, in, 2007.
- [33] T. Botsis, O. Hejlesen, J. Bellika, and G. Hartvigsen, Disease surveillance systems for diabetics, in: *American Telemedicine Association*, 2008.
- [34] T. Botsis, O. Hejlesen, J. Bellika, and G. Hartvigsen, Electronic Disease Surveillance for Sensitive Population Groups – The Diabetics Case Study, *Studies in Health Technology and Informatics* **136** (2008), 365-370.
- [35] T. Botsis, O. Hejlesen, J. Bellika, and G. Hartvigsen, Electronic infectious disease surveillance systems for diabetics, in: *International Conference on Advanced Technologies & Treatments for Diabetes*, Prague, Czech Republic, 2008.
- [36] T. Botsis, O. Hejlesen, J.G. Bellika, and G. Hartvigsen, Blood glucose levels as an indicator for the early detection of infections in type-1 diabetics, in: *Advances in Disease Surveillance*, 2007.
- [37] T. Botsis, O. Hejlesen, J.G. Bellika, and G. Hartvigsen, Electronic disease surveillance for sensitive population groups - the diabetics case study, *Stud Health Technol Inform* **136** (2008), 365-370.
- [38] T. Botsis, A.M. Lai, G. Hripesak, W. Palmas, J.B. Starren, and G. Hartvigsen, Proof of concept for the role of glycemic control in the early detection of infections in diabetics, *Health Informatics J* **18** (2012), 26-35.
- [39] M.N.K. Boulos, A.J. Curtis, and P. AbdelMalik, Musings on privacy issues in health research involving disaggregate geographic data about individuals, *International Journal of Health Geographics* **8** (2009), 46.
- [40] A.W. Bowman and A. Azzalini, *Applied smoothing techniques for data analysis: the kernel approach with S-Plus illustrations*, OUP Oxford, 1997.
- [41] D. Brealey and M. Singer, Hyperglycemia in critical illness: a review, *J Diabetes Sci Technol* **3** (2009), 1250-1260.
- [42] M.M. Breunig, H.-P. Kriegel, R.T. Ng, and J. Sander, LOF: identifying density-based local outliers, *SIGMOD Rec.* **29** (2000), 93-104.

- [43] A. Brown, Factors that affect blood glucose levels, in, diaTribe, 2018.
- [44] J.S. Brownstein, C.C. Freifeld, and L.C. Madoff, Digital disease detection--harnessing the Web for public health surveillance, *N Engl J Med* **360** (2009), 2153-2155, 2157.
- [45] J.S. Brownstein, C.C. Freifeld, B.Y. Reis, and K.D. Mandl, Surveillance Sans Frontieres: Internet-based emerging infectious disease intelligence and the HealthMap project, *PLoS Med* **5** (2008), e151.
- [46] D.L. Buckeridge, H. Burkom, M. Campbell, W.R. Hogan, and A.W. Moore, Algorithms for rapid outbreak detection: a research synthesis, *Journal of Biomedical Informatics* **38** (2005), 99-113.
- [47] S.O. Butler, I.F. Btaiche, and C. Alaniz, Relationship Between Hyperglycemia and Infection in Critically Ill Patients, *Pharmacotherapy* **25** (2005), 963-976.
- [48] G. Carole, R. Frances, H. Philip, and C. Fiona, The Leeds Insulin Pump Workbook for Children and Young People, in, 2012.
- [49] CDC, Event Vs. Indicator-based Surveillance, in, Centers for Disease Control and Prevention USA, 2019.
- [50] A. Ceriello and M.A. Ihnat, 'Glycaemic variability': a new therapeutic challenge in diabetes and the critical care setting, *Diabet Med* **27** (2010), 862-867.
- [51] V. Chandola, A. Banerjee, and V. Kumar, Anomaly detection: A survey, *ACM computing surveys (CSUR)* **41** (2009), 15.
- [52] H. Chen, D. Zeng, and P. Yan, *Infectious Disease Informatics: Syndromic Surveillance for Public Health and Bio-Defense*, Springer Publishing Company, Incorporated, 2009.
- [53] J. Chen, A. Bauman, and M. Allman-Farinelli, A Study to Determine the Most Popular Lifestyle Smartphone Applications and Willingness of the Public to Share Their Personal Data for Health Research, *Telemed J E Health* **22** (2016), 655-665.
- [54] B.C.K. Choi, The past, present, and future of public health surveillance, *Scientifica* **2012** (2012), 875253-875253.
- [55] J. Choi, Y. Cho, E. Shim, and H. Woo, Web-based infectious disease surveillance systems and public health perspectives: a systematic review, *BMC Public Health* **16** (2016), 1238-1238.
- [56] M. Clark, *What Is Diabetes?*, John Wiley & Sons, Ltd., England, 2004.
- [57] G. Cohen, M. Hilario, H. Sax, S. Hugonnet, C. Pellegrini, and A. Geissbühler, An Application of One-class Support Vector Machines to Nosocomial Infection Detection, in: *Medinfo*, 2004, pp. 716-720.
- [58] S. Coucheron, A.Z. Woldaregay, E. Årsand, T. Botsis, and G. Hartvigsen, EDMON - A System Architecture for Real-Time Infection Monitoring and Outbreak Detection Based on Self-Recorded Data from People with Type 1 Diabetes: System Design and Prototype Implementation, in: *SHI 2019. Proceedings of the 17th Scandinavian Conference on Health Informatics, November 12-13, 2019, Oslo, Norway*, Linköping University Electronic Press, Linköpings universitet, 2019, pp. 37-44.
- [59] N.R. Council, *BioWatch and Public Health Surveillance: Evaluating Systems for the Early Detection of Biological Threats: Abbreviated Version*, National Academies Press, 2011.
- [60] S. Courbier, R. Dimond, and V. Bros-Facer, Share and protect our health data: an evidence based approach to rare disease patients' perspectives on data sharing and data protection - quantitative survey and recommendations, *Orphanet journal of rare diseases* **14** (2019), 175-175.
- [61] P.E. Cryer, Hypoglycemia in type 1 diabetes mellitus, *Endocrinology and metabolism clinics of North America* **39** (2010), 641-654.
- [62] L. David, Infections in Diabetes, (2011), 139-155.
- [63] P.C. Davidson, H.R. Hebblewhite, R.D. Steed, and B.W. Bode, Analysis of guidelines for basal-bolus insulin dosing: basal insulin, correction factor, and carbohydrate-to-insulin ratio, *Endocr Pract* **14** (2008), 1095-1101.
- [64] L.A. DiMeglio, C. Evans-Molina, and R.A. Oram, Type 1 diabetes, *Lancet (London, England)* **391** (2018), 2449-2462.
- [65] G. Dritsakis, L. Murdin, D. Kikidis, G.H. Saunders, P. Katrakazas, D. Brdarić, K. Ploumidou, and D.E. Bamiou, Challenges and Strengths of Multidisciplinary Research in Audiology: The EVOTION Example, *Am J Audiol* **28** (2019), 1046-1051.

- [66] dsmi-lab-ntust, Anomaly Detection Toolbox, in, GitHub, GitHub, 2016.
- [67] K. Duangchaemkarn, V. Chaovatut, P. Wiwatanadate, and E. Boonchieng, Symptom-based data preprocessing for the detection of disease outbreak, in: *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2017, pp. 2614-2617.
- [68] R.P. Duin, P. Juszczak, P. Paclik, E. Pekalska, D. De Ridder, D.M. Tax, and S. Verzakov, Prtools4. 1, a matlab toolbox for pattern recognition, *Delft University of technology* **2600** (2007).
- [69] D.L. Duke, *Intelligent Diabetes Assistant: A Telemedicine System for Modeling and Managing Blood Glucose*, PhD, Carnegie Mellon University, 2010.
- [70] T. Dunning and E. Friedman, *Practical machine learning: a new look at anomaly detection*, "O'Reilly Media, Inc.", 2014.
- [71] T. Eckmanns, H. Fuller, and S.L. Roberts, Digital epidemiology and global health security; an interdisciplinary conversation, *Life sciences, society and policy* **15** (2019), 2.
- [72] J. Evans, A. Papadopoulos, C.T. Silvers, N. Charness, W.R. Boot, L. Schlachta-Fairchild, C. Crump, M. Martinez, and C.B. Ent, Remote Health Monitoring for Older Adults and Those with Heart Failure: Adherence and System Usability, *Telemedicine journal and e-health : the official journal of the American Telemedicine Association* **22** (2016), 480-488.
- [73] K.R. Evenson, M.M. Goto, and R.D. Furberg, Systematic review of the validity and reliability of consumer-wearable activity trackers, *International Journal of Behavioral Nutrition and Physical Activity* **12** (2015), 159.
- [74] T. Fawcett, Mining the Quantified Self: Personal Knowledge Discovery as a Challenge for Data Science, *Big Data* **3** (2015), 249-266.
- [75] M.D. Fetters, L.A. Curry, and J.W. Creswell, Achieving integration in mixed methods designs-principles and practices, *Health Serv Res* **48** (2013), 2134-2156.
- [76] D.A. Finan, H. Zisser, L. Jovanovic, W.C. Bevier, and D.E. Seborg, Automatic Detection of Stress States in Type 1 Diabetes Subjects in Ambulatory Conditions, *Ind Eng Chem Res* **49** (2010), 7843-7848.
- [77] G.A. Fleming, J.R. Petrie, R.M. Bergenstal, R.W. Holl, A.L. Peters, and L. Heinemann, Diabetes Digital App Technology: Benefits, Challenges, and Recommendations. A Consensus Report by the European Association for the Study of Diabetes (EASD) and the American Diabetes Association (ADA) Diabetes Technology Working Group, *Diabetes Care* **43** (2020), 250-260.
- [78] D.U. Forums, Flu and high blood sugar, in, Diabetes UK, UK.
- [79] C.C. Freifeld, K.D. Mandl, B.Y. Reis, and J.S. Brownstein, HealthMap: global infectious disease monitoring through automated classification and visualization of Internet media reports, *Journal of the American Medical Informatics Association : JAMIA* **15** (2008), 150-157.
- [80] M.A.L. Gabbay, M. Rodacki, L.E. Calliari, A.G.D. Vianna, M. Krakauer, M.S. Pinto, J.S. Reis, M. Puñales, L.G. Miranda, A.C. Ramalho, D.R. Franco, and H.P.C. Pedrosa, Time in range: a new parameter to evaluate blood glucose control in patients with diabetes, *Diabetology & Metabolic Syndrome* **12** (2020), 22.
- [81] L.D. Genevieve, A. Martani, M.C. Mallet, T. Wangmo, and B.S. Elger, Factors influencing harmonized health data collection, sharing and linkage in Denmark and Switzerland: A systematic review, *PLoS One* **14** (2019), e0226015.
- [82] L.D. Genevieve, A. Martani, T. Wangmo, D. Paolotti, C. Koppeschaar, C. Kjelso, C. Guerrisi, M. Hirsch, O. Woolley-Meza, P. Lukowicz, A. Flahault, and B.S. Elger, Participatory Disease Surveillance Systems: Ethical Framework, *J Med Internet Res* **21** (2019), e12273.
- [83] M. George, A.M. Chacko, S.K. Kurien, and N. Ali, Diabetes care in cloud - research challenges, in: *Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing*, ACM, Limassol, Cyprus, 2019, pp. 160-162.
- [84] S.J. Gillespie, K.D. Kulkarni, and A.E. Daly, Using Carbohydrate Counting in Diabetes Clinical Practice, *Journal of the American Dietetic Association* **98** (1998), 897-905.
- [85] H. Gimpel, M. Nißen, and R. Görlitz, Quantifying the quantified self: A study on the motivations of patients to track their own health, (2013).
- [86] J. Ginsberg, M.H. Mohebbi, R.S. Patel, L. Brammer, M.S. Smolinski, and L. Brilliant, Detecting influenza epidemics using search engine query data, *Nature* **457** (2009), 1012-1014.

- [87] M. Goldstein and S. Uchida, A Comparative Evaluation of Unsupervised Anomaly Detection Algorithms for Multivariate Data, *PLoS One* **11** (2016), e0152173.
- [88] A. Gramacki, Bandwidth Selectors for Kernel Density Estimation, in: *Nonparametric Kernel Density Estimation and Its Computational Aspects*, Springer International Publishing, Cham, 2018, pp. 63-83.
- [89] A. Gramacki, Kernel Density Estimation, in: *Nonparametric Kernel Density Estimation and Its Computational Aspects*, Springer International Publishing, Cham, 2018, pp. 25-62.
- [90] O. Granberg, J.G. Bellika, E. Årsand, and G. Hartvigsen, Automatic infection detection system, *Studies in Health Technology and Informatics* **129** (2007), 566-570.
- [91] S.L. Groseclose and D.L. Buckeridge, Public Health Surveillance Systems: Recent Advances in Their Use and Evaluation, *Annual Review of Public Health* **38** (2017), 57-79.
- [92] T.H. Grubestic and T.C. Matisziw, On the use of ZIP codes and ZIP code tabulation areas (ZCTAs) for the spatial analysis of epidemiological data, *Int J Health Geogr* **5** (2006), 58.
- [93] Q. Grundy, K. Chiu, F. Held, A. Continella, L. Bero, and R. Holz, Data sharing practices of medicines related apps and the mobile ecosystem: traffic, content, and network analysis, *BMJ* **364** (2019), 1920.
- [94] K. Guan, M. Shao, and S. Wu, A Remote Health Monitoring System for the Elderly Based on Smart Home Gateway, *Journal of Healthcare Engineering* **2017** (2017), 5843504.
- [95] C. Gurrin, A.F. Smeaton, and A.R. Doherty, LifeLogging: Personal Big Data, *Foundations and Trends® in Information Retrieval* **8** (2014), 1-125.
- [96] S. Hajizadeh, Z. Li, R.P.B.J. Dollevoet, and D.M.J. Tax, Evaluating Classification Performance with only Positive and Unlabeled Samples, in: *Structural, Syntactic, and Statistical Pattern Recognition*, P. Fränti, G. Brown, M. Loog, F. Escolano, and M. Pelillo, eds., Springer Berlin Heidelberg, Berlin, Heidelberg, 2014, pp. 233-242.
- [97] S. Hajizadeh, A. Núñez, and D.M.J. Tax, Semi-supervised Rail Defect Detection from Imbalanced Image Data, *IFAC-PapersOnLine* **49** (2016), 78-83.
- [98] G. Hartvigsen, E. Årsand, T. Botsis, K. van Vuurden, M. Johansen, and J.G. Bellika, Reusing patient data to enhance patient empowerment and electronic disease surveillance, *Journal on Information Technology in Healthcare* **7** (2009), 4-12.
- [99] A.C. Harvey and S.J. Koopman, Diagnostic Checking of Unobserved-Components Time Series Models, *Journal of Business & Economic Statistics* **10** (1992), 377-389.
- [100] J. Haslett and S.J. Haslett, The Three Basic Types of Residuals for a Linear Model, *International Statistical Review / Revue Internationale de Statistique* **75** (2007), 1-24.
- [101] N.-B. Heidenreich, A. Schindler, and S. Sperlich, Bandwidth selection for kernel density estimation: a review of fully automatic selectors, *AStA Advances in Statistical Analysis* **97** (2013), 403-433.
- [102] K.J. Henning, What is Syndromic Surveillance?, *Morbidity and Mortality Weekly Report* **53** (2004), 7-11.
- [103] A. Henriksen, M. Haugen Mikalsen, A.Z. Woldaregay, M. Muzny, G. Hartvigsen, L.A. Hopstock, and S. Grimsgaard, Using Fitness Trackers and Smartwatches to Measure Physical Activity in Research: Analysis of Consumer Wrist-Worn Wearables, *J Med Internet Res* **20** (2018), e110.
- [104] A. Henriksen, F. Svartdal, S. Grimsgaard, G. Hartvigsen, and L. Hopstock, Polar Vantage and Oura physical activity and sleep trackers: A validation study, *medRxiv* (2020).
- [105] A. Henriksen, A.Z. Woldaregay, D.-Z. Issom, G. Pfuhl, A. Richard, E. Årsand, K. Sato, G. Hartvigsen, and J. Rochat, Replication data for: User expectations and willingness to share self-collected health, in: U.-T.A.U.o. Norway, ed., DataverseNO, Tromsø, Norway, 2019.
- [106] J.L. Hicks, T. Althoff, R. Susic, P. Kuhar, B. Bostjancic, A.C. King, J. Leskovec, and S.L. Delp, Best practices for analyzing large-scale health data from wearables and smartphone apps, *npj Digital Medicine* **2** (2019), 45.
- [107] M. Hintze, Viewing the GDPR through a de-identification lens: a tool for compliance, clarification, and consistency, *International Data Privacy Law* **8** (2017), 86-101.
- [108] K. Hope, D.N. Durrheim, E.T. d'Espaignet, and C. Dalton, Syndromic Surveillance: is it a useful tool for local outbreak detection?, *Journal of epidemiology and community health* **60** (2006), 374-375.

- [109] G. Huzooree, K.K. Khedo, and N. Joonas, Wireless Body Area Network System Architecture for Real-Time Diabetes Monitoring, in, Springer International Publishing, Cham, 2017, pp. 262-271.
- [110] I. Irigoien, B. Sierra, and C. Arenas, Towards application of one-class classification methods to medical data, *ScientificWorldJournal* **2014** (2014), 730712.
- [111] S. Izahar, Q.Y. Lean, M.A. Hameed, M.K. Murugiah, R.P. Patel, Y.M. Al-Worafi, T.W. Wong, and L.C. Ming, Content Analysis of Mobile Health Applications on Diabetes Mellitus, *Front Endocrinol (Lausanne)* **8** (2017), 318.
- [112] S. James, *The efficiency of conformal predictors for anomaly detection*, Doctor of Philosophy, Royal Holloway, University of London, 2016.
- [113] N. Japkowicz, *Concept-Learning in the Absence of Counter-Examples: An Autoassociation-Based Approach to Classification*, PhD, Graduate School-New Brunswick Rutgers, The State University of New Jersey, 1999.
- [114] A.J. Jara, M.A. Zamora-Izquierdo, and A.F. Gomez-Skarmeta, An Ambient Assisted Living System for Telemedicine with Detection of Symptoms, in, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009, pp. 75-84.
- [115] N. Joshi, G.M. Caputo, M.R. Weitekamp, and A.W. Karchmer, Infections in patients with diabetes mellitus, *N Engl J Med* **341** (1999), 1906-1912.
- [116] P. Juszczak, D.M.J. Tax, E. Pełkalska, and R.P.W. Duin, Minimum spanning tree based one-class classifier, *Neurocomputing* **72** (2009), 1859-1869.
- [117] S. Kalkman, J. van Delden, A. Banerjee, B. Tyl, M. Mostert, and G. van Thiel, Patients' and public views and attitudes towards the sharing of health data for research: a narrative review of the empirical evidence, *Journal of Medical Ethics* (2019), medethics-2019-105651.
- [118] S. Kandanaarachchi, M.A. Munoz, R.J. Hyndman, and K. Smith-Miles, On normalization and algorithm selection for unsupervised outlier detection, in, 2018.
- [119] M. Karampela, S. Ouhbi, and M. Isomursu, Connected Health User Willingness to Share Personal Health Data: Questionnaire Study, *J Med Internet Res* **21** (2019), e14537.
- [120] M. Karampela, S. Ouhbi, and M. Isomursu, Exploring users' willingness to share their health and personal data under the prism of the new GDPR: implications in healthcare, in: *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019, pp. 6509-6512.
- [121] N. Katsilambros, E. Diakoumopoulou, I. Ioannidis, S. Liatis, K. Makrilakis, N. Tentolouris, and P. Tsapogas, Acute Illness in Diabetes, in: *Diabetes in Clinical Practice Questions and Answers from Case Studies*, M. Konstantinos, ed., John Wiley & Sons, Ltd, England, 2006, pp. 103-107.
- [122] J. Kauffmann, L. Ruff, G. Montavon, and K.-R. Müller, The Clever Hans effect in anomaly detection, *arXiv preprint arXiv:2006.10609* (2020).
- [123] S.S. Khan and M.G. Madden, A Survey of Recent Trends in One Class Classification, in, Springer Berlin Heidelberg, Berlin, Heidelberg, 2010, pp. 188-197.
- [124] S.S. Khan and M.G. Madden, One-class classification: taxonomy of study and review of techniques, *The Knowledge Engineering Review* **29** (2014), 345-374.
- [125] A.B. King, A. Kuroda, M. Matsuhisa, and T. Hobbs, A Review of Insulin-Dosing Formulas for Continuous Subcutaneous Insulin Infusion (CSII) for Adults with Type 1 Diabetes, *Curr Diab Rep* **16** (2016), 83.
- [126] J. Kinsella, C. Hawthorne, I. Piper, M. Shaw, and L. Moss, Sharing of Big Data in Healthcare: Public Opinion, Trust, and Privacy Considerations for Health Informatics Researchers, (2017), 463-468.
- [127] J.K. Kirk and J. Stegner, Self-monitoring of blood glucose: practical aspects, *Journal of Diabetes Science and Technology* **4** (2010), 435-439.
- [128] C. Klingler, D.S. Silva, C. Schuermann, A.A. Reis, A. Saxena, and D. Strech, Ethical issues in public health surveillance: a systematic qualitative review, *BMC Public Health* **17** (2017), 295.
- [129] D.C. Klonoff, Continuous Glucose Monitoring, *Roadmap for 21st century diabetes therapy* **28** (2005), 1231-1239.
- [130] I. Koskinen, J. Zimmerman, T. Binder, J. Redstrom, and S. Wensveen, *Design Research Through Practice: From the Lab, Field, and Showroom*, Morgan Kaufmann Publishers Inc., 2012.

- [131] P. Kostkova, Disease surveillance data sharing for public health: the next ethical frontiers, *Life sciences, society and policy* **14** (2018), 16.
- [132] A. Kuroda, T. Yasuda, M. Takahara, F. Sakamoto, R. Kasami, K. Miyashita, S. Yoshida, E. Kondo, K. Aihara, I. Endo, T.A. Matsuoka, H. Kaneto, T. Matsumoto, I. Shimomura, and M. Matsuhisa, Carbohydrate-to-insulin ratio is estimated from 300-400 divided by total daily insulin dose in type 1 diabetes patients who use the insulin pump, *Diabetes Technol Ther* **14** (2012), 1077-1080.
- [133] C.H. Lang, Mechanism of Insulin Resistance in Infection, in: *Pathophysiology of Shock, Sepsis, and Organ Failure*, G. Schlag and H. Redl, eds., Springer Berlin Heidelberg, Berlin, Heidelberg, 1993, pp. 609-625.
- [134] J. Lauritzen, J. E.Arsand, K. Van Vuurden, O. Hejlesen, and G. Hartvigsen, Exploring illness prediction in Type 1 Diabetes Mellitus pre-symptom onset, *Using administrative databases to identify cases of chronic kidney disease: a systematic review* (2011), 25.
- [135] J.N. Lauritzen, E. Arsand, K. Van Vuurden, J.G. Bellika, O.K. Hejlesen, and G. Hartvigsen, Towards a mobile solution for predicting illness in Type 1 Diabetes Mellitus: Development of a prediction model for detecting risk of illness in Type 1 Diabetes prior to symptom onset, (2011), 1-5.
- [136] R. Laxhammar, *Conformal anomaly detection : Detecting abnormal trajectories in surveillance applications*, PhD dissertation, University of Skövde, 2014.
- [137] R. Laxhammar and G. Falkman, Inductive conformal anomaly detection for sequential detection of anomalous sub-trajectories, *Annals of Mathematics and Artificial Intelligence* **74** (2015), 67-94.
- [138] D. Lazer, R. Kennedy, G. King, and A. Vespignani, Big data. The parable of Google Flu: traps in big data analysis, *Science* **343** (2014), 1203-1205.
- [139] L. Leelarathna, S.A. Roberts, A. Hindle, K. Markakis, T. Alam, A. Chapman, J. Morris, A. Urwin, P. Jinadev, and M.K. Rutter, Comparison of different insulin pump makes under routine care conditions in adults with Type 1 diabetes, *Diabet Med* **34** (2017), 1372-1379.
- [140] C. LeRouge and N. Wickramasinghe, A review of user-centered design for diabetes-related consumer health informatics technologies, in: *Journal of Diabetes Science and Technology*, 2013, pp. 1039-1056.
- [141] D. Lewis, Sick days solved with a DIY closed loop #OpenAPS, in, DIYPS.org, Diabetes, DIYPS, OpenAPS, 2016.
- [142] D. Lewis, Quantified sickness when you have #OpenAPS and the flu, in, DIYPS.org, Diabetes, OpenAPS, Research, 2018.
- [143] L.C. Madoff, ProMED-mail: an early warning system for emerging diseases, *Clin Infect Dis* **39** (2004), 227-232.
- [144] K. Makrilakis and N. Katsilambros, Diabetic Emergencies, Diagnosis and Clinical Management: Sick-Day Rules in Diabetes, Part 2.
- [145] K. Makrilakis and N. Katsilambros, Sick-day rules in diabetes, *Diabetic emergencies* (2011), 178.
- [146] S. Maldonado, C. Montecinos, K. Smith-Miles, and R. Weber, Robust classification of imbalanced data using one-class and two-class SVM-based multiclassifiers, *Intelligent Data Analysis* **18** (2014), 95-112.
- [147] M.L. Marcovecchio and F. Chiarelli, The effects of acute and chronic stress on diabetes control, *Sci Signal* **5** (2012), pt10.
- [148] A. Martinez-Millana, E. Jarones, C. Fernandez-Llatas, G. Hartvigsen, and V. Traver, App Features for Type 1 Diabetes Support and Patient Empowerment: Systematic Literature Review and Benchmark Comparison, *JMIR Mhealth Uhealth* **6** (2018), e12237.
- [149] A. Martinez, W. Ruba, A.B. Sánchez, M.T. Meneu, and V. Traver, Architecture for Lifestyle Monitoring Platform in Diabetes Management, in, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011, pp. 194-200.
- [150] O. Mazhelis, One-class classifiers: a review and analysis of suitability in the context of mobile-masquerader detection, *South African Computer Journal* **2006** (2006), 29-48.
- [151] R.J. McCrimmon and R.S. Sherwin, Hypoglycemia in type 1 diabetes, *Diabetes* **59** (2010), 2333-2339.



- [152] O.P. McGuinness, Defective glucose homeostasis during infection, *Annu Rev Nutr* **25** (2005), 9-35.
- [153] K.G. Mehrotra, C.K. Mohan, and H. Huang, *Anomaly detection principles and algorithms*, Springer, 2017.
- [154] S. Meißner, Effects of Quantified Self Beyond Self-Optimization, in: *Lifelogging: Digital self-tracking and Lifelogging - between disruptive technology and cultural transformation*, S. Selke, ed., Springer Fachmedien Wiesbaden, Wiesbaden, 2016, pp. 235-248.
- [155] M. Memon, S.R. Wagner, C.F. Pedersen, F.H.A. Beevi, and F.O. Hansen, Ambient assisted living healthcare frameworks, platforms, standards, and quality attributes, *Sensors (Basel, Switzerland)* **14** (2014), 4312-4341.
- [156] M.F. Mendiola, M. Kalnicki, and S. Lindenauer, Valuable features in mobile health apps for patients and consumers: content analysis of apps and user ratings, *JMIR Mhealth Uhealth* **3** (2015), e40.
- [157] D.K. Ming, S. Sangkaew, H.Q. Chanh, P.T.H. Nhat, S. Yacoub, P. Georgiou, and A.H. Holmes, Continuous physiological monitoring using wearable technology to inform individual management of infectious diseases, public health and outbreak responses, *International Journal of Infectious Diseases* **96** (2020), 648-654.
- [158] M. MiniMed, MiniMed® 530G: System User Guide, in: <http://www.medtronicdiabetes.com/patents>, USA, 2012.
- [159] B.A. Mizock, Alterations in carbohydrate metabolism during stress: A review of the literature, *The American Journal of Medicine* **98** (1995), 75-84.
- [160] B.A. Mizock, Alterations in fuel metabolism in critical illness: hyperglycaemia, *Best Pract Res Clin Endocrinol Metab* **15** (2001), 533-551.
- [161] M. Mobasser, M. Shirmohammadi, T. Amiri, N. Vahed, H. Hosseini Fard, and M. Ghojzadeh, Prevalence and incidence of type 1 diabetes in the world: a systematic review and meta-analysis, *Health promotion perspectives* **10** (2020), 98-115.
- [162] D. Morrow, M. Hasegawa-Johnson, T. Huang, W. Schuh, R.F.L. Azevedo, K. Gu, Y. Zhang, B. Roy, and R. Garcia-Retamero, A multidisciplinary approach to designing and evaluating Electronic Medical Record portal messages that support patient self-care, *Journal of Biomedical Informatics* **69** (2017), 63-74.
- [163] S.G. Mougiakakou, C.S. Bartsocas, E. Bozas, N. Chaniotakis, D. Iliopoulou, I. Kouris, S. Pavlopoulos, A. Prountzou, M. Skevofilakas, A. Tsoukalis, K. Varotsis, A. Vazeou, K. Zarkogianni, and K.S. Nikita, SMARTDIAB: A Communication and Information Technology Approach for the Intelligent Monitoring, Management and Follow-up of Type 1 Diabetes Patients, *IEEE Transactions on Information Technology in Biomedicine* **14** (2010), 622-633.
- [164] N. Mounika and P. Vaijyanthi, Analysis of algorithms for one class classification of heart disease identification, (2017), 907-912.
- [165] M. Munoz-Organero, Deep Physiological Model for Blood Glucose Prediction in T1DM Patients, *Sensors (Basel, Switzerland)* **20** (2020), 3896.
- [166] E. Mykhalovskiy and L. Weir, The Global Public Health Intelligence Network and early warning outbreak detection: a Canadian contribution to global public health, *Can J Public Health* **97** (2006), 42-44.
- [167] T. Nakamura, Y. Hirota, N. Hashimoto, T. Matsuda, M. Takabe, K. Sakaguchi, W. Ogawa, and S. Seino, Diurnal variation of carbohydrate insulin ratio in adult type 1 diabetic patients treated with continuous subcutaneous insulin infusion, *Journal of diabetes investigation* **5** (2014), 48-50.
- [168] G.H. Nguyen, A. Bouzerdoum, and S.L. Phung, Learning pattern classification tasks with imbalanced data sets, in: *Pattern Recognition*, IntechOpen, 2009.
- [169] H.V. Nguyen, T.T. Nguyen, and Q.U. Nguyen, Sequential ensemble method for unsupervised anomaly detection, in: *2017 9th International Conference on Knowledge and Systems Engineering (KSE)*, 2017, pp. 71-76.
- [170] J. O'Shea, Digital disease detection: A systematic review of event-based internet biosurveillance systems, *Int J Med Inform* **101** (2017), 15-22.
- [171] A.J. Onwuegbuzie, R.M. Bustamante, and J.A. Nelson, Mixed Research as a Tool for Developing Quantitative Instruments, *Journal of Mixed Methods Research* **4** (2009), 56-78.

- [172] K. Ostherr, S. Borodina, R.C. Bracken, C. Lotterman, E. Storer, and B. Williams, Trust and privacy in the context of user-generated health data, *Big Data & Society* **4** (2017), 2053951717704673.
- [173] S. Oviedo, J. Vehi, R. Calm, and J. Armengol, A review of personalized blood glucose prediction strategies for T1DM patients, *Int J Numer Method Biomed Eng* **33** (2017).
- [174] E. Parzen, On Estimation of a Probability Density Function and Mode, *The Annals of Mathematical Statistics* **33** (1962), 1065-1076.
- [175] M.W. Percival, W.C. Bevier, Y. Wang, E. Dassau, H.C. Zisser, L. Jovanovic, and F.J. Doyle, 3rd, Modeling the effects of subcutaneous insulin administration and carbohydrate consumption on blood glucose, *J Diabetes Sci Technol* **4** (2010), 1214-1228.
- [176] M.I. Petrovskiy, Outlier Detection Algorithms in Data Mining Systems, *Programming and Computer Software* **29** (2003), 228-237.
- [177] K.T. Pickard and M. Swan, Big desire to share big health data: A shift in consumer attitudes toward personal health information, in: *2014 AAAI Spring Symposium Series*, 2014.
- [178] M.A.F. Pimentel, D.A. Clifton, L. Clifton, and L. Tarassenko, A review of novelty detection, *Signal Processing* **99** (2014), 215-249.
- [179] C.G. Pretty, A.J. Le Compte, J.G. Chase, G.M. Shaw, J.C. Preiser, S. Penning, and T. Desai, Variability of insulin sensitivity during the first 4 days of critical illness: implications for tight glycemic control, *Ann Intensive Care* **2** (2012), 17.
- [180] G. Quer, J.M. Radin, M. Gadaleta, K. Baca-Motes, L. Ariniello, E. Ramos, V. Kheterpal, E.J. Topol, and S.R. Steinhubl, Wearable sensor data and self-reported symptoms for COVID-19 detection, *Nature Medicine* (2020).
- [181] J.M. Radin, N.E. Wineinger, E.J. Topol, and S.R. Steinhubl, Harnessing wearable device data to improve state-level real-time surveillance of influenza-like illness in the USA: a population-based study, *The Lancet Digital Health* **2** (2020), e85-e93.
- [182] V. Rasoulzadeh, E.C. Erkus, T.A. Yogurt, I. Ulusoy, and S.A. Zergeroğlu, A comparative stationarity analysis of EEG signals, *Annals of Operations Research* **258** (2017), 133-157.
- [183] D.d. Ridder, D.M.J. Tax, and R.P.W. Duin, An experimental comparison of one-class classification methods, in: *Proceedings of the 4th Annual Conference of the Advanced School for Computing and Imaging, Delft*, 1998.
- [184] P.V. Röder, B. Wu, Y. Liu, and W. Han, Pancreatic regulation of glucose homeostasis, *Experimental & molecular medicine* **48** (2016), e219-e219.
- [185] I. Rodriguez-Rodriguez, J.V. Rodriguez, and M.A. Zamora-Izquierdo, Variables to Be Monitored via Biomedical Sensors for Complete Type 1 Diabetes Mellitus Management: An Extension of the "On-Board" Concept, *J Diabetes Res* **2018** (2018), 4826984.
- [186] E. Rodriguez Mega, COVID has killed more than one million people. How many more will die?, *Nature* (2020).
- [187] A.J. Rossini, "Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-Plus Illustrations" by Adrian W. Bowman and Adelchi Azzalini, *Computational Statistics* **15** (2000), 301-302.
- [188] P.J. Rousseeuw and K.V. Driessen, A Fast Algorithm for the Minimum Covariance Determinant Estimator, *Technometrics* **41** (1999), 212-223.
- [189] A. Rowhani-Farid, M. Allen, and A.G. Barnett, What incentives increase data sharing in health and medical research? A systematic review, *Research Integrity and Peer Review* **2** (2017), 4.
- [190] E. Ruiz-Velazquez, A.Y. Alanis, R. Femat, and G. Quiroz, Neural modeling of the blood glucose level for Type 1 Diabetes Mellitus patients, (2011), 696-701.
- [191] C. Saint-Pierre, V. Herskovic, and M. Sepúlveda, Multidisciplinary collaboration in primary care: a systematic review, *Family Practice* **35** (2017), 132-141.
- [192] T. Saito and M. Rehmsmeier, The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets, *PLoS One* **10** (2015), e0118432.
- [193] S. Salehi, A. Olyaeemanesh, M. Mobinizadeh, E. Nasli-Esfahani, and H. Riazi, Assessment of remote patient monitoring (RPM) systems for patients with type 2 diabetes: a systematic review and meta-analysis, *Journal of Diabetes & Metabolic Disorders* **19** (2020), 115-127.
- [194] S. Samerski, Individuals on alert: digital epidemiology and the individualization of surveillance, *Life sciences, society and policy* **14** (2018), 13.

- [195] A.H. Sapci and H.A. Sapci, The Effectiveness of Hands-on Health Informatics Skills Exercises in the Multidisciplinary Smart Home Healthcare and Health Informatics Training Laboratories, *Applied clinical informatics* **8** (2017), 1184-1196.
- [196] A.H. Sapci and H.A. Sapci, Innovative Assisted Living Tools, Remote Monitoring Technologies, Artificial Intelligence-Driven Solutions, and Robotic Systems for Aging Societies: Systematic Review, *JMIR Aging* **2** (2019), e15429.
- [197] M. Schiavon, C. Dalla Man, and C. Cobelli, Insulin Sensitivity Index-Based Optimization of Insulin to Carbohydrate Ratio: In Silico Study Shows Efficacious Protection Against Hypoglycemic Events Caused by Suboptimal Therapy, *Diabetes Technol Ther* **20** (2018), 98-105.
- [198] B. Schölkopf, R.C. Williamson, A.J. Smola, J. Shawe-Taylor, and J.C. Platt, Support Vector Method for Novelty Detection, in: *Advances in neural information processing systems*, 1999, pp. 582-588.
- [199] P. Sebastiani and K. Mandl, Biosurveillance and outbreak detection, *Data mining: next generation challenges and future directions* (2004), 185-198.
- [200] A. Seifert, M. Christen, and M. Martin, Willingness of Older Adults to Share Mobile Health Data with Researchers, *GeroPsych* **31** (2018), 41-49.
- [201] E. Seltzer, J. Goldshear, S.C. Guntuku, D. Grande, D.A. Asch, E.V. Klinger, and R.M. Merchant, Patients' willingness to share digital health and non-health data for research: a cross-sectional study, *BMC Medical Informatics and Decision Making* **19** (2019), 157.
- [202] D.R. Seshadri, E.V. Davies, E.R. Harlow, J.J. Hsu, S.C. Knighton, T.A. Walker, J.E. Voos, and C.K. Drummond, Wearable Sensors for COVID-19: A Call to Action to Harness Our Digital Infrastructure for Remote Patient Monitoring and Virtual Assessments, *Frontiers in Digital Health* **2** (2020).
- [203] D.R. Seshadri, J.R. Rowbottom, C. Drummond, J.E. Voos, and J. Craker, A review of wearable technology: Moving beyond the hype: From need through sensor implementation, in: *2016 8th Cairo International Biomedical Engineering Conference (CIBEC)*, 2016, pp. 52-55.
- [204] G. Shafer and V. Vovk, A tutorial on conformal prediction, *Journal of Machine Learning Research* **9** (2008), 371-421.
- [205] M. Shahidul Islam, M.T. Islam, A.F. Almutairi, G.K. Beng, N. Misran, and N. Amin, Monitoring of the Human Body Signal through the Internet of Things (IoT) Based LoRa Wireless Network System, *Applied Sciences* **9** (2019), 1884.
- [206] L. Shoemaker and L.O. Hall, Anomaly Detection Using Ensembles, in, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011, pp. 6-15.
- [207] S.E. Siegelaar, F. Holleman, J.B. Hoekstra, and J.H. DeVries, Glucose variability; does it matter?, *Endocr Rev* **31** (2010), 171-182.
- [208] E.J. Simon, J. Dickey, J.B. Reece, and R.A. Burton, *Campbell Essential Biology with Physiology*, Pearson Education, United States of America, 2015.
- [209] S.O. Skovseth, J.G. Bellika, and F. Godtliebsen, Causality in scale space as an approach to change detection, *PLoS One* **7** (2012), e52253.
- [210] C.E. Smart, F. Annan, L.P. Bruno, L.A. Higgins, C.L. Acerini, P. International Society for, and D. Adolescent, ISPAD Clinical Practice Consensus Guidelines 2014. Nutritional management in children and adolescents with diabetes, *Pediatr Diabetes* **15 Suppl 20** (2014), 135-153.
- [211] J. Smith, I. Nouretdinov, R. Craddock, C. Offer, and A. Gammerman, Anomaly Detection of Trajectories with Kernel Density Estimation by Conformal Prediction, in, Springer Berlin Heidelberg, Berlin, Heidelberg, 2014, pp. 271-280.
- [212] Spike, Spike App Features, in: *Spike app for diabetes self-management* github.com, SpikeApp /Spike 2018, pp. Spike app for diabetes self-management
- [213] N. Straiton, M. Alharbi, A. Bauman, L. Neubeck, J. Gullick, R. Bhindi, and R. Gallagher, The validity and reliability of consumer-grade activity trackers in older, community-dwelling adults: A systematic review, *Maturitas* **112** (2018), 85-93.
- [214] W. Sun, Z. Cai, Y. Li, F. Liu, S. Fang, and G. Wang, Security and Privacy in the Medical Internet of Things: A Review, *Security and Communication Networks* **2018** (2018), 1-9.

- [215] J. Tang, Z. Chen, A.W.-c. Fu, and D.W. Cheung, Enhancing Effectiveness of Outlier Detections for Low Density Patterns, in: *Advances in Knowledge Discovery and Data Mining*, M.-S. Chen, P.S. Yu, and B. Liu, eds., Springer Berlin Heidelberg, Berlin, Heidelberg, 2002, pp. 535-548.
- [216] D.M. Tax and R.P. Duin, Characterizing one-class datasets, in: *Proceedings of the sixteenth annual symposium of the pattern recognition association of South Africa*, 2006, pp. 21-26.
- [217] D.M.J. Tax, *One-class classification: Concept learning in the absence of counter-examples*, PhD, 2002.
- [218] D.M.J. Tax, DDTTools, the data description toolbox for MATLAB, version 2.1. 2, *Delft University of Technology, Delft, Netherlands* (2015).
- [219] D.M.J. Tax and R.P.W. Duin, Support vector domain description, *Pattern Recognition Letters* **20** (1999), 1191-1199.
- [220] D.M.J. Tax and R.P.W. Duin, Data description in subspaces, in: *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*, 2000, pp. 672-675 vol.672.
- [221] D.M.J. Tax and R.P.W. Duin, Support Vector Data Description, *Machine Learning* **54** (2004), 45-66.
- [222] D.M.J. Tax and K. Muller, A consistency-based model selection for one-class classification, in: *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, 2004, pp. 363-366 Vol.363.
- [223] İ. Tayfur and M.A. Afacan, Reliability of smartphone measurements of vital parameters: A prospective study using a reference method, *Am J Emerg Med* **37** (2019), 1527-1530.
- [224] A. Tharwat, Classification assessment methods, *Applied Computing and Informatics* (2018).
- [225] C. Thinking, Understanding public expectations of the use of health and care data, in: London: The Curved Thinking Partnership, 2019.
- [226] S. Thudumu, P. Branch, J. Jin, and J. Singh, A comprehensive survey of anomaly detection techniques for high dimensional big data, *Journal of Big Data* **7** (2020), 42.
- [227] F.C. Tsui, J.U. Espino, V.M. Dato, P.H. Gesteland, J. Hutman, and M.M. Wagner, Technical description of RODS: a real-time public health surveillance system, *Journal of the American Medical Informatics Association : JAMIA* **10** (2003), 399-408.
- [228] M. van der Hoogt, J.C. van Dyk, R.C. Dolman, and M. Pieters, Protein and fat meal content increase insulin requirement in children with type 1 diabetes - Role of duration of diabetes, *Journal of clinical & translational endocrinology* **10** (2017), 15-21.
- [229] W.G. van Panhuis, P. Paul, C. Emerson, J. Grefenstette, R. Wilder, A.J. Herbst, D. Heymann, and D.S. Burke, A systematic review of barriers to data sharing in public health, *BMC Public Health* **14** (2014), 1144.
- [230] A. Vasighizaker, A. Sharma, and A. Dehzangi, A novel one-class classification approach to accurately predict disease-gene association in acute myeloid leukemia cancer, *PLoS One* **14** (2019), e0226115.
- [231] E. Vayena, M. Salathe, L.C. Madoff, and J.S. Brownstein, Ethical challenges of big data in public health, *PLoS Comput Biol* **11** (2015), e1003904.
- [232] E. Velasco, Disease detection, epidemiology and outbreak response: the digital future of public health practice, *Life sciences, society and policy* **14** (2018), 7.
- [233] V. Vovk, The Basic Conformal Prediction Framework, (2014), 3-19.
- [234] V. Vovk, Beyond the Basic Conformal Prediction Framework, (2014), 21-46.
- [235] V. Vovk, A. Gammerman, and G. Shafer, Conformal prediction, in: *Algorithmic Learning in a Random World*, Springer US, Boston, MA, 2005, pp. 17-51.
- [236] M.M. Wagner, L.S. Gresham, and V. Dato, Case Detection, Outbreak Detection, and Outbreak Characterization, *Handbook of Biosurveillance* (2006), 27-50.
- [237] M.M. Wagner, A.W. Moore, and R.M. Aryel, *Handbook of biosurveillance*, Elsevier, 2011.
- [238] W.K. Waldhausl, P. Bratusch-Marrain, M. Komjati, F. Breitenecker, and I. Troch, Blood glucose response to stress hormone exposure in healthy man and insulin dependent diabetic patients: prediction by computer modeling, *IEEE Trans Biomed Eng* **39** (1992), 779-790.
- [239] S. Wallace, M. Clark, and J. White, 'It's on my iPhone': attitudes to the use of mobile computing devices in medical education, a mixed-methods study, *BMJ Open* **2** (2012).

- [240] J. Walsh, R. Roberts, and T. Bailey, Guidelines for insulin dosing in continuous subcutaneous insulin infusion using new formulas from a retrospective study of individuals with optimal glucose levels, *J Diabetes Sci Technol* **4** (2010), 1174-1181.
- [241] J. Walsh, R. Roberts, and T. Bailey, Guidelines for optimal bolus calculator settings in adults, *J Diabetes Sci Technol* **5** (2011), 129-135.
- [242] B. Wang and Z. Mao, One-class classifiers ensemble based anomaly detection scheme for process control systems, *Transactions of the Institute of Measurement and Control* **40** (2017), 3466-3476.
- [243] H. Wang, Y. Zhao, L. Yu, J. Liu, I.M. Zwetsloot, J. Cabrera, and K.-L. Tsui, A Personalized Health Monitoring System for Community-Dwelling Elderly People in Hong Kong: Design, Implementation, and Evaluation Study, *J Med Internet Res* **22** (2020), e19223.
- [244] G. Watts, Data sharing: keeping patients on board, *The Lancet Digital Health* **1** (2019), e332-e333.
- [245] E.R. Weitzman, B. Adida, S. Kelemen, and K.D. Mandl, Sharing data for public health research by members of an international online diabetes social network, *PLoS One* **6** (2011), e19256.
- [246] E.R. Weitzman, L. Kaci, and K.D. Mandl, Acceptability of a personally controlled health record in a community-based setting: implications for policy and design, *J Med Internet Res* **11** (2009), e14.
- [247] E.R. Weitzman, S. Kelemen, L. Kaci, and K.D. Mandl, Willingness to share personal health record data for care improvement and public health: a survey of experienced personal health record users, *BMC Medical Informatics and Decision Making* **12** (2012), 39-39.
- [248] L.J. Wheat, Infection and Diabetes Mellitus, *Diabetes Care* **3** (1980), 187-197.
- [249] D. Whicher, M. Ahmed, S. Siddiqi, I. Adams, C. Grossmann, and K. Carman, eds., *Health Data Sharing to Support Better Outcomes: Building a Foundation of Stakeholder Trust*, NAM Special Publication, Washington, DC: National Academy of Medicine, 2020.
- [250] R. Whiddett, I. Hunter, J. Engelbrecht, and J. Handy, Patients' attitudes towards sharing their health information, *International Journal of Medical Informatics* **75** (2006), 530-541.
- [251] M. Wiley, R. Bunescu, C. Marling, J. Shubrook, and F. Schwartz, Automatic Detection of Excessive Glycemic Variability for Diabetes Management, in: *2011 10th International Conference on Machine Learning and Applications and Workshops*, Honolulu, HI, USA, 2011, pp. 148-154.
- [252] A. Woldaregay, D. Issom, A. Henriksen, H. Marttila, M. Mikalsen, G. Pfuhl, K. Sato, C. Lovis, and G. Hartvigsen, Motivational Factors for User Engagement with mHealth Apps, *Studies in Health Technology and Informatics* **249** (2018), 151.
- [253] A.Z. Woldaregay, E. Årsand, T. Botsis, D. Albers, L. Mamykina, and G. Hartvigsen, Data-Driven Blood Glucose Pattern Classification and Anomalies Detection: Machine-Learning Applications in Type 1 Diabetes, *J Med Internet Res* **21** (2019), e11030.
- [254] A.Z. Woldaregay, E. Årsand, T. Botsis, and G. Hartvigsen, An Early Infectious Disease Outbreak Detection Mechanism Based on Self-Recorded Data from People with Diabetes, *Studies in Health Technology and Informatics* **245** (2017), 619-623.
- [255] A.Z. Woldaregay, E. Årsand, A. Giordanengo, D. Albers, L. Mamykina, T. Botsis, and G. Hartvigsen, EDMON-A Wireless Communication Platform for a Real-Time Infectious Disease Outbreak De-tection System Using Self-Recorded Data from People with Type 1 Diabetes, in: *Proceedings from The 15th Scandinavian Conference on Health Informatics 2017 Kristiansand, Norway, August 29–30, 2017*, Linköping University Electronic Press, 2018, pp. 14-20.
- [256] A.Z. Woldaregay, E. Årsand, S. Walderhaug, D. Albers, L. Mamykina, T. Botsis, and G. Hartvigsen, Data-driven modeling and prediction of blood glucose dynamics: Machine learning applications in type 1 diabetes, *Artificial Intelligence in Medicine* **98** (2019), 109-134.
- [257] A.Z. Woldaregay, A. Henriksen, D.Z. Issom, G. Pfuhl, K. Sato, A. Richard, C. Lovis, E. Årsand, J. Rochat, and G. Hartvigsen, User Expectations and Willingness to Share Self-Collected Health Data, *Stud Health Technol Inform* **270** (2020), 894-898.
- [258] A.Z. Woldaregay, I.K. Launonen, D. Albers, J. Igual, E. Årsand, and G. Hartvigsen, A Novel Approach for Continuous Health Status Monitoring and Automatic Detection of Infection Incidences in People With Type 1 Diabetes Using Machine Learning Algorithms (Part 2): A

- Personalized Digital Infectious Disease Detection Mechanism, *J Med Internet Res* **22** (2020), e18912.
- [259] A.Z. Woldaregay, I.K. Launonen, E. Årsand, D. Albers, A. Holubová, and G. Hartvigsen, Toward Detecting Infection Incidence in People With Type 1 Diabetes Using Self-Recorded Data (Part 1): A Novel Framework for a Personalized Digital Infectious Disease Detection System, *J Med Internet Res* **22** (2020), e18911.
- [260] P. World Health Organization. Regional Office for the Western, *A guide to establishing event-based surveillance*, Manila : WHO Regional Office for the Western Pacific, 2008.
- [261] P.K. Yeng, A.Z. Woldaregay, T. Solvoll, and G. Hartvigsen, A systematic review of cluster detection mechanisms in syndromic surveillance: Towards developing a framework of cluster detection mechanisms for EDMON system, in: *Proceedings from The 16th Scandinavian Conference on Health Informatics 2018, Aalborg, Denmark August 28–29, 2018*, Linköping University Electronic Press, 2018, pp. 62-69.
- [262] P.K. Yeng, A.Z. Woldaregay, T. Solvoll, and G. Hartvigsen, Cluster Detection Mechanisms for Syndromic Surveillance Systems: Systematic Review and Framework Development, *JMIR Public Health Surveill* **6** (2020), e11512.
- [263] C. Yi, Bivariant Kernel Density Estimation, in, 2013, pp. A tool for bivariant pdf, cdf and icdf estimation using Gaussian kernel function.
- [264] H. Yki-Jarvinen, K. Sammalkorpi, V.A. Koivisto, and E.A. Nikkila, Severity, duration, and mechanisms of insulin resistance during acute infections, *J Clin Endocrinol Metab* **69** (1989), 317-323.
- [265] E. Yu and P. Parekh, A Bayesian Ensemble for Unsupervised Anomaly Detection, *arXiv preprint arXiv:1610.07677* (2016).
- [266] Y. Yu, Y. Zhu, S. Li, and D. Wan, Time Series Outlier Detection Based on Sliding Window Prediction, *Mathematical Problems in Engineering* **2014** (2014), 14.
- [267] Z. Zhao, K.G. Mehrotra, and C.K. Mohan, Ensemble Algorithms for Unsupervised Anomaly Detection, in, Springer International Publishing, Cham, 2015, pp. 514-525.
- [268] S. Zulj, G. Seketa, D. Dzaja, F. Sklebar, S. Drobnjak, L. Celic, and R. Magjarevic, Supporting Diabetic Patients with a Remote Patient Monitoring Systems, in, Springer Singapore, Singapore, 2017, pp. 577-580.

# 10 Appendix A - Score plot of the models using the univariate input feature: Ratio of Insulin to Carbohydrate

The following appendix presents the performance score of the different models evaluated in the dissertation using univariate input – insulin-to-carbohydrate ratio. Both the one-class classifiers and unsupervised models were tested. For the one-class classifier models, a random block of a 4-month sample size from each patient-year was used to train the models and evaluated with the whole patient-year – containing both the regular and infection days. Unsupervised models were tested with the whole patient-year. The smoothed version of the dataset is used and the models were tested with both the hourly and daily datasets. The score plot of the models is given below, and, as can be seen from the figures, each model generated different scores and rejected varying portions of the infection episodes.

## 10.1. One-class classifier Method

### 10.1.1. Daily

#### 10.1.1.1. The First Infection Episode (Flu)

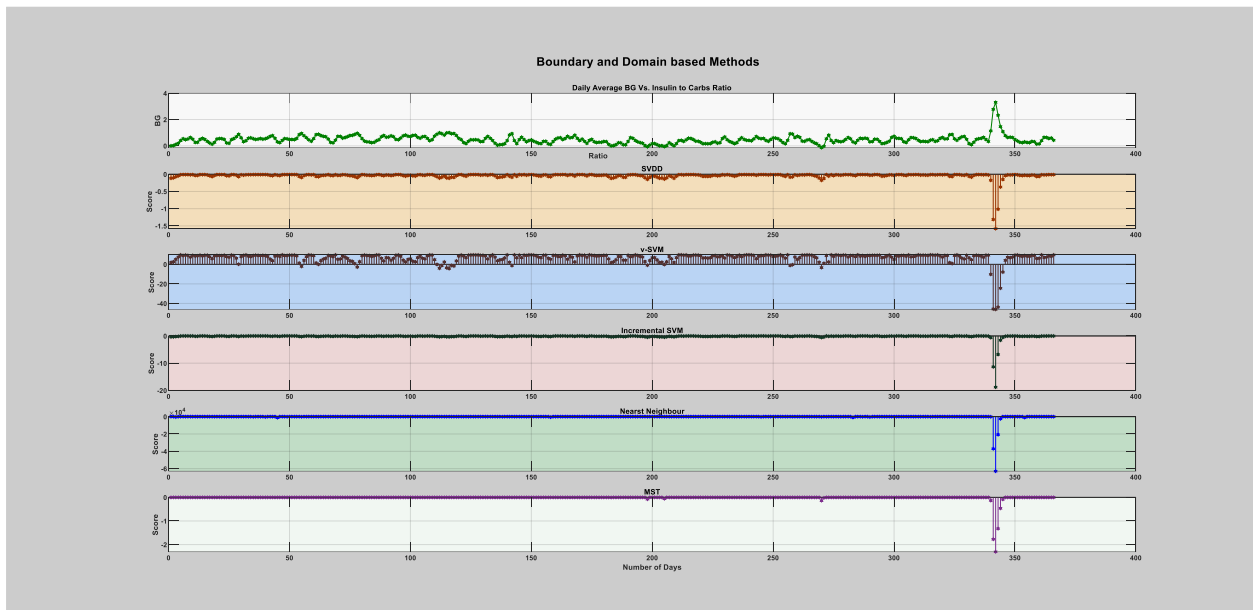
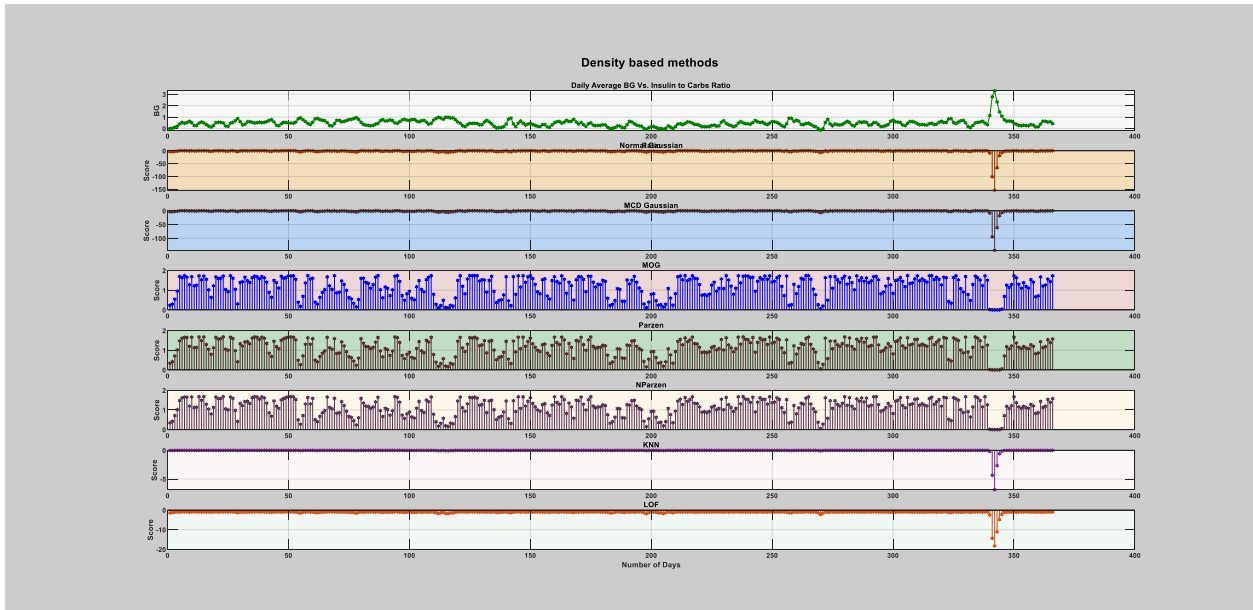
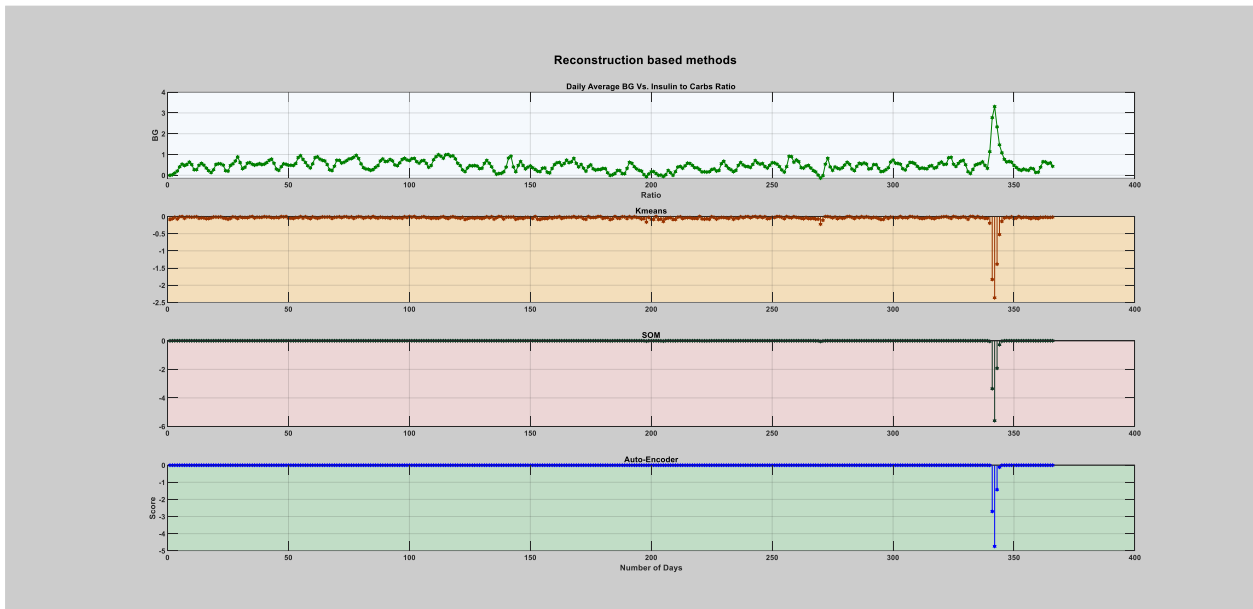


Figure 1: Boundary and domain-based method.



**Figure 2:** Density-based method.



**Figure 3:** Reconstruction-based method.



### 10.1.1.2. The Second Infection Episode (Flu)

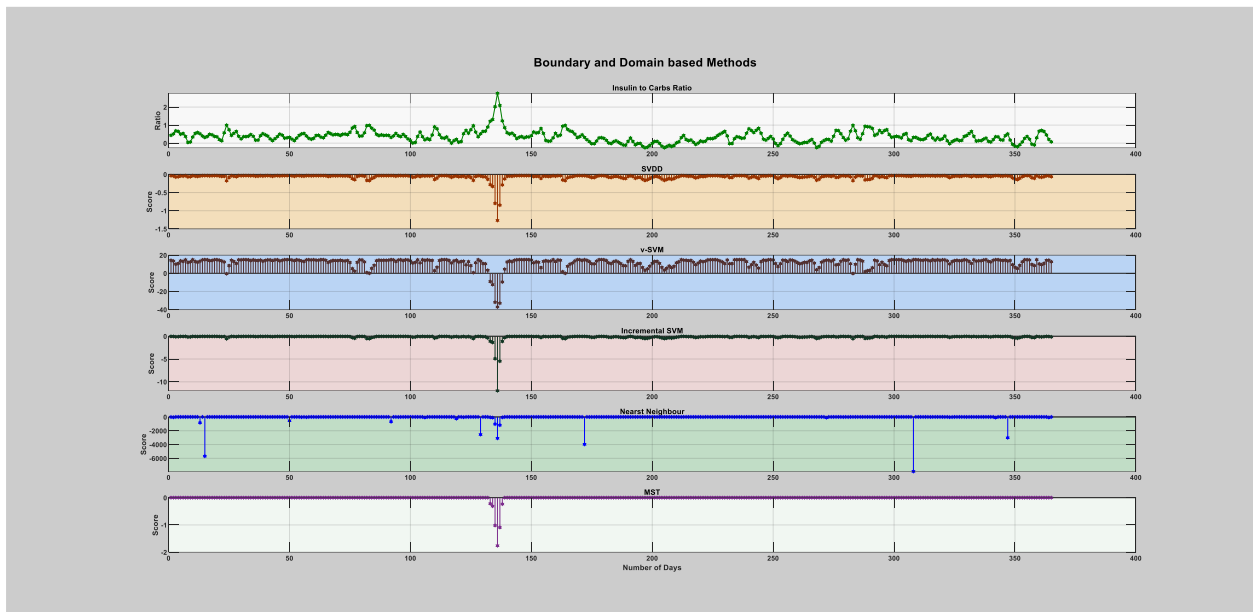


Figure 4: Boundary and domain-based method.

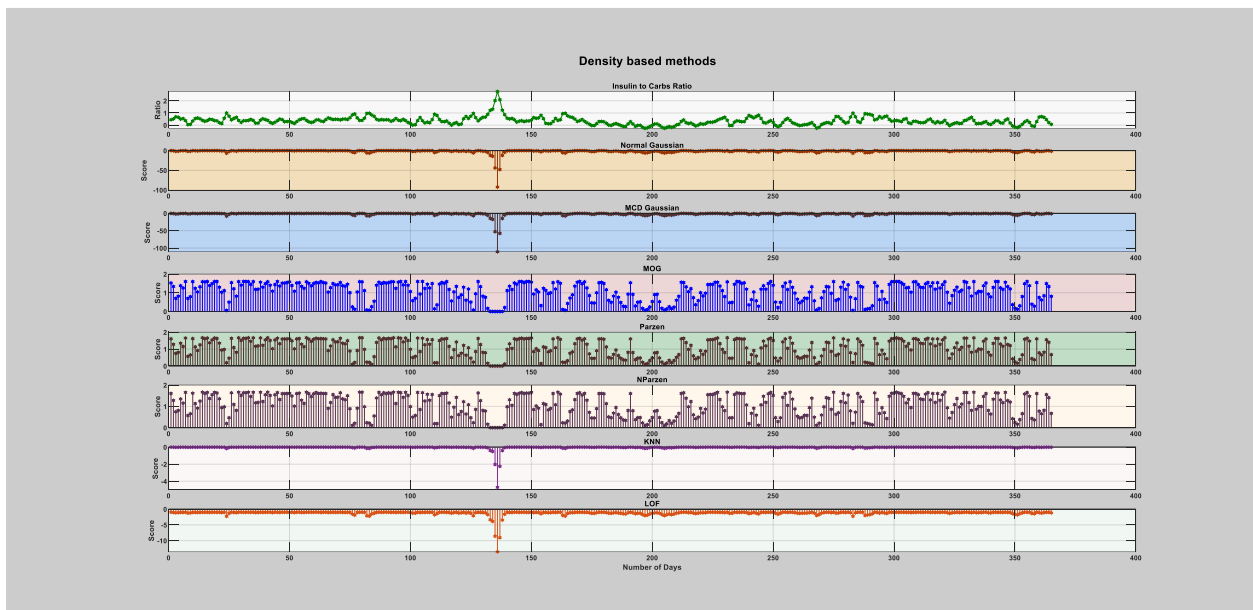
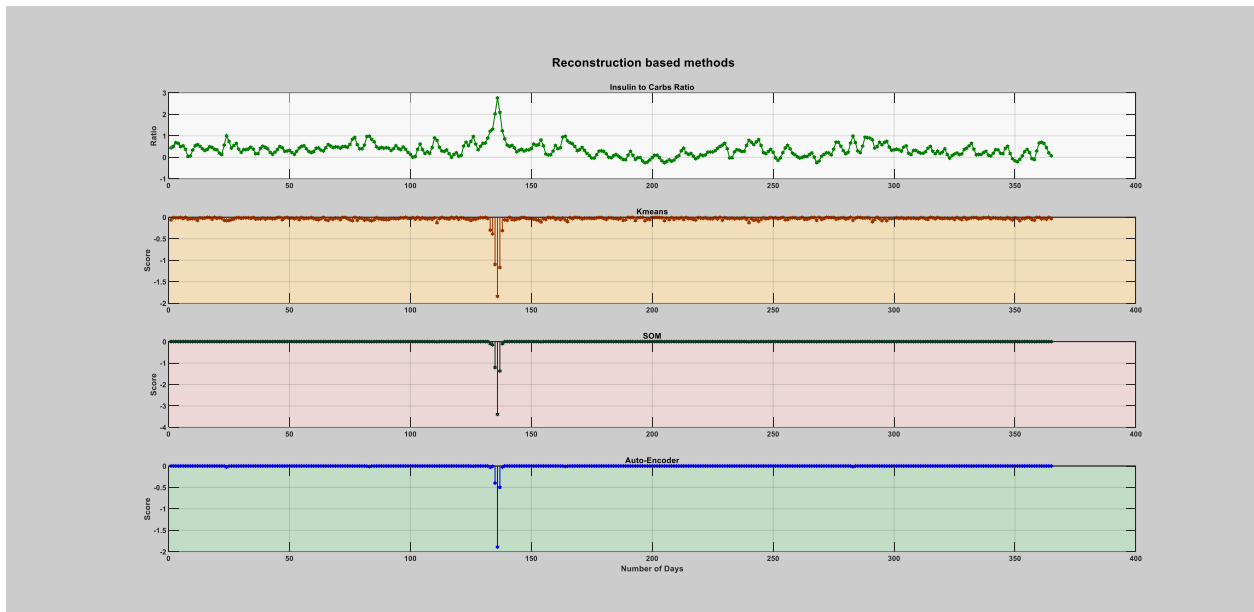
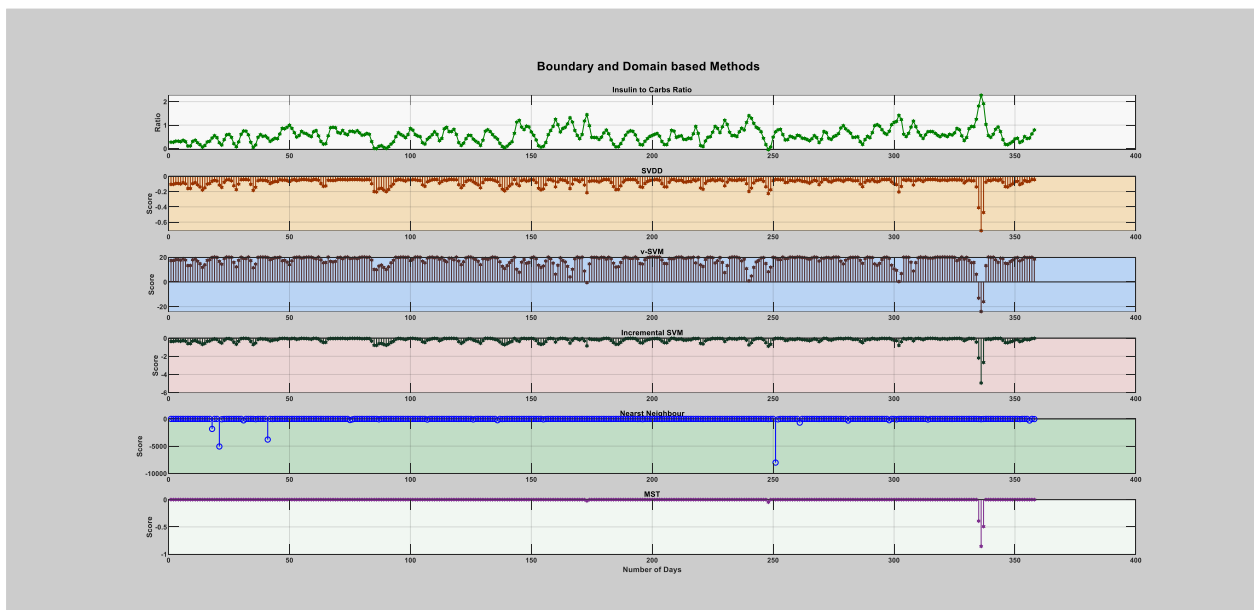


Figure 5: Density-based method.

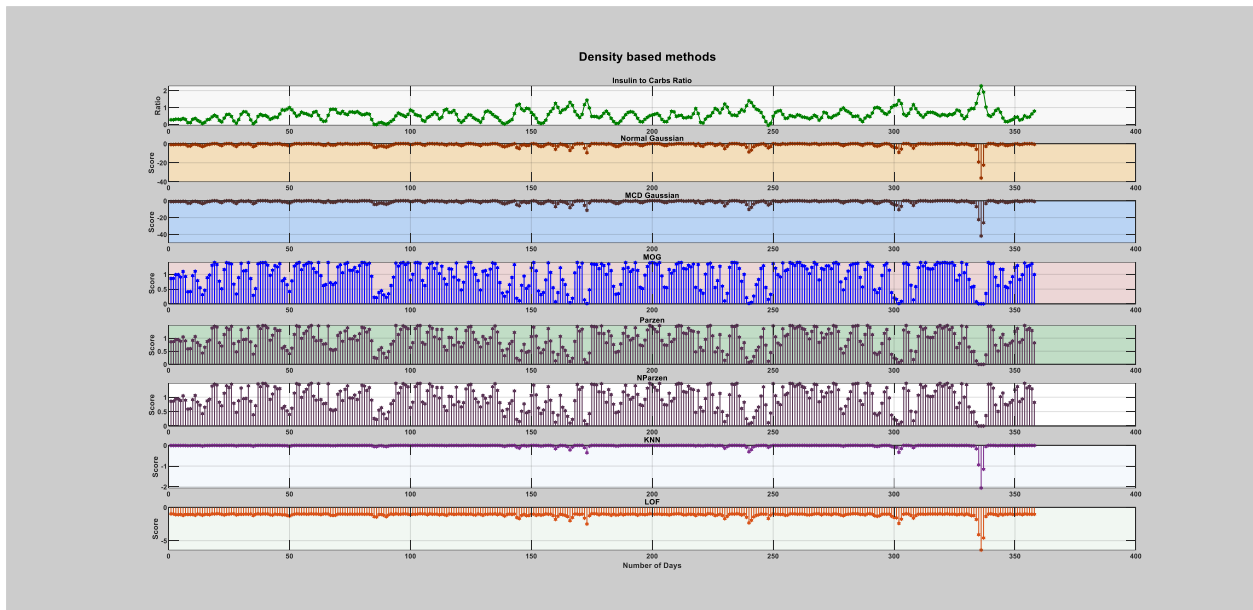


**Figure 6:** Reconstruction-based method.

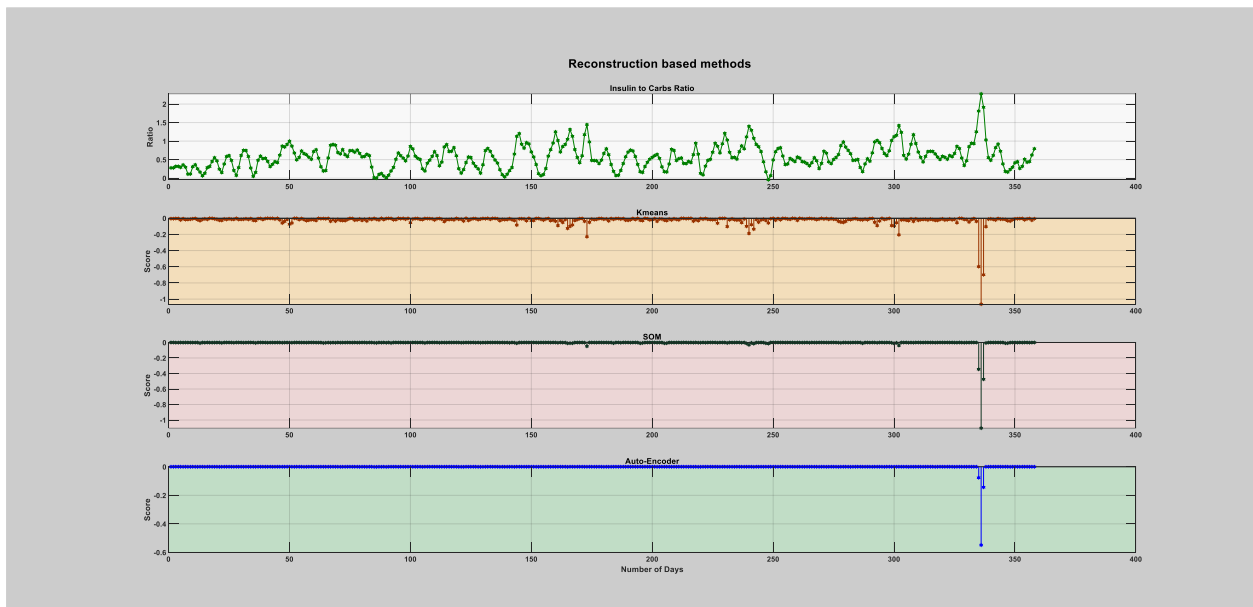
### 10.1.1.3. The Third Infection Episode (Flu)



**Figure 7:** Boundary and domain-based method.



**Figure 8:** Density-based method.



**Figure 9:** The score of the reconstruction-based method on the whole patient-year.

#### 10.1.1.4. The Fourth Infection Episode (Flu)

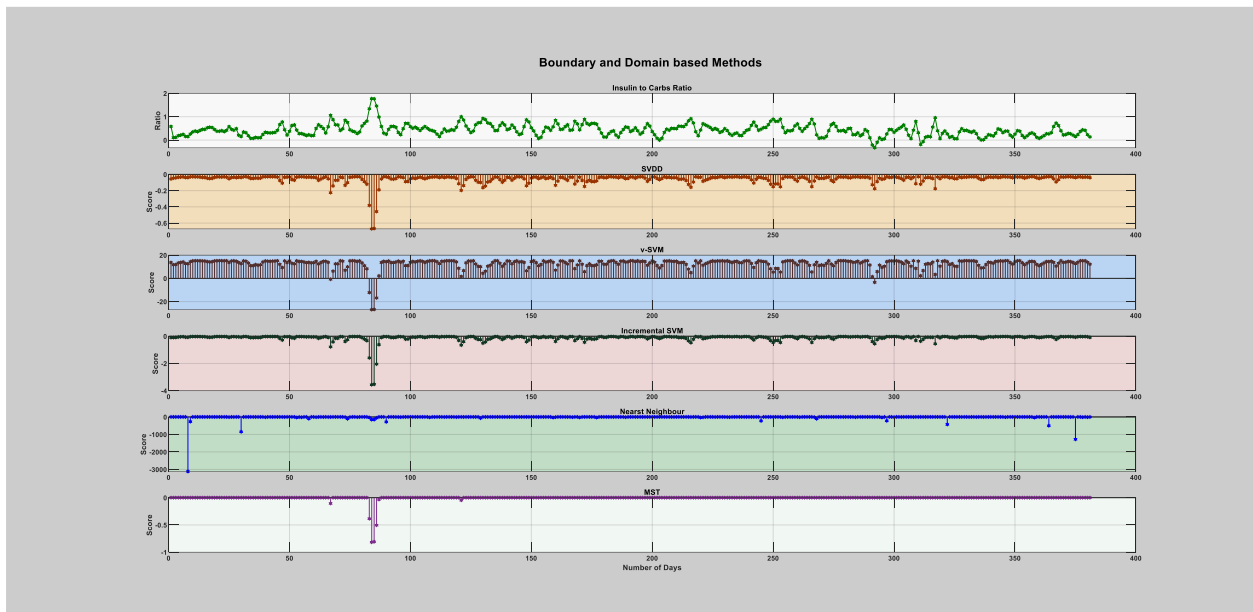


Figure 10: Boundary and domain-based.

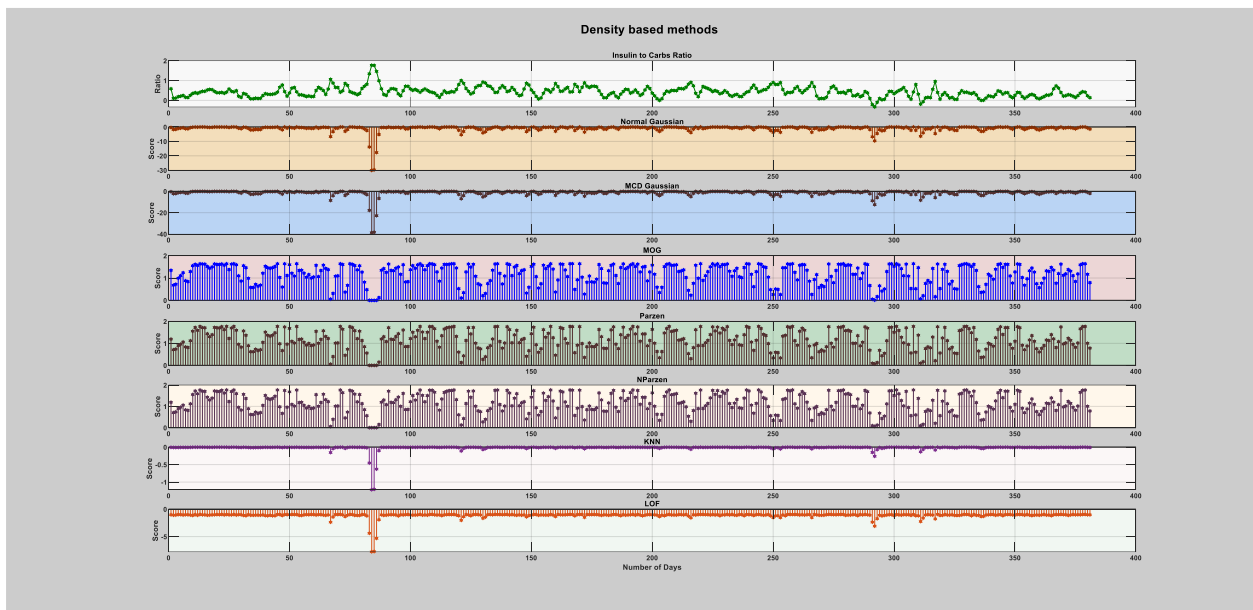
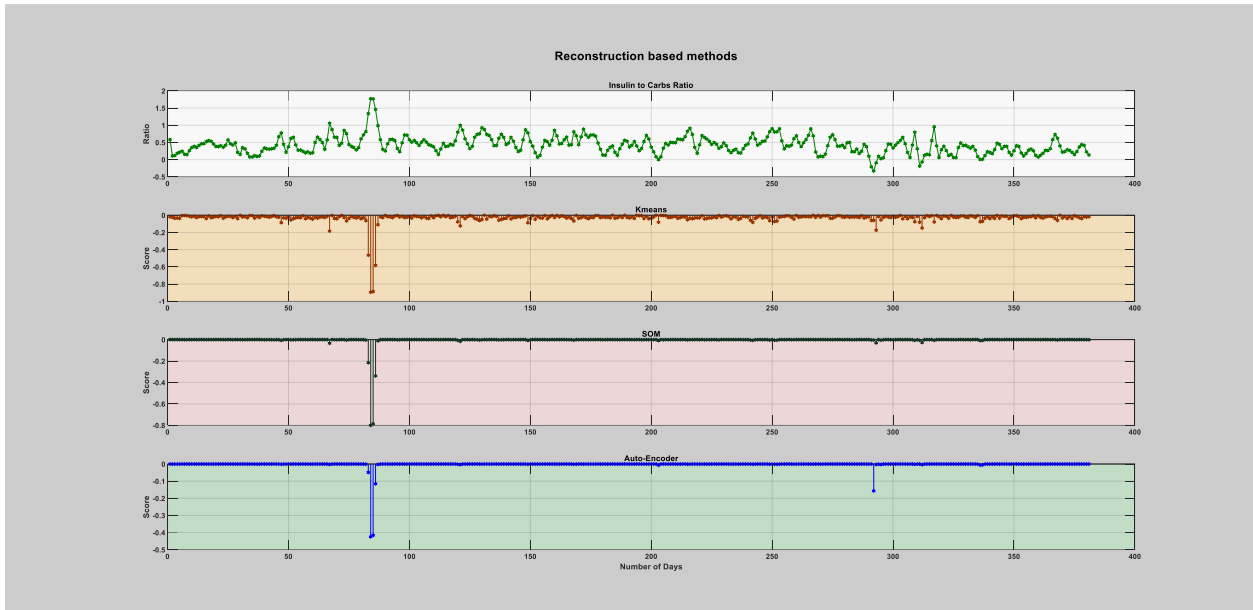


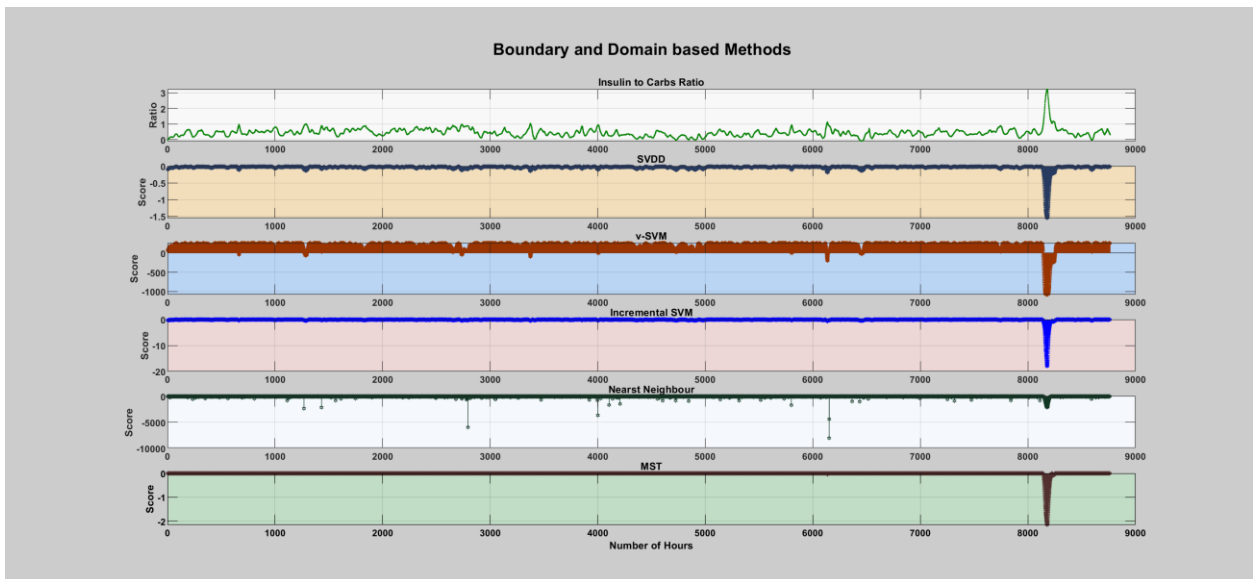
Figure 11: Density-based method.



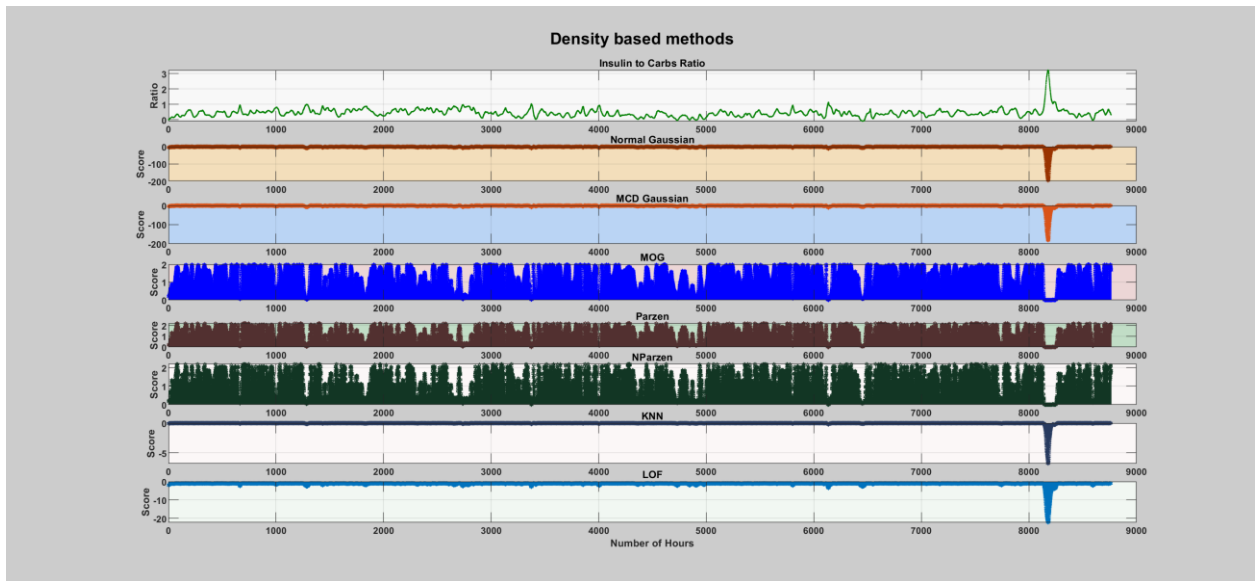
**Figure 12:** Reconstruction-based method.

## 10.1.2. Hourly

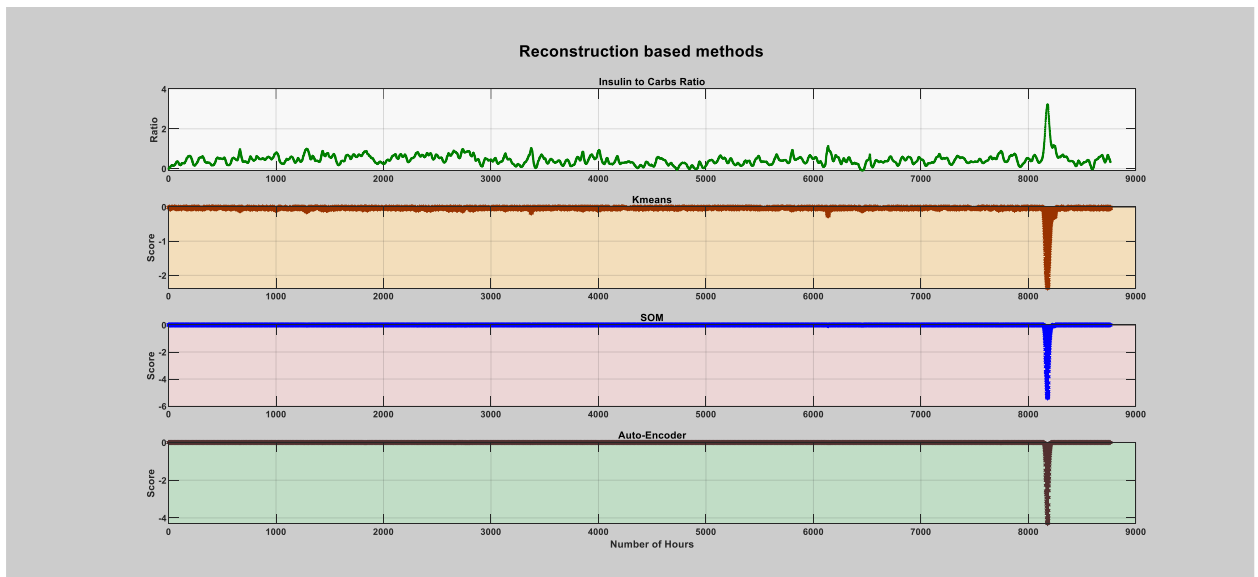
### 10.1.2.1. The First Infection Episode (Flu)



**Figure 13:** Boundary and domain-based method.



**Figure 14:** Density-based method.



**Figure 15:** Reconstruction-based method.

### 10.1.2.2. The Second Infection Episode (Flu)

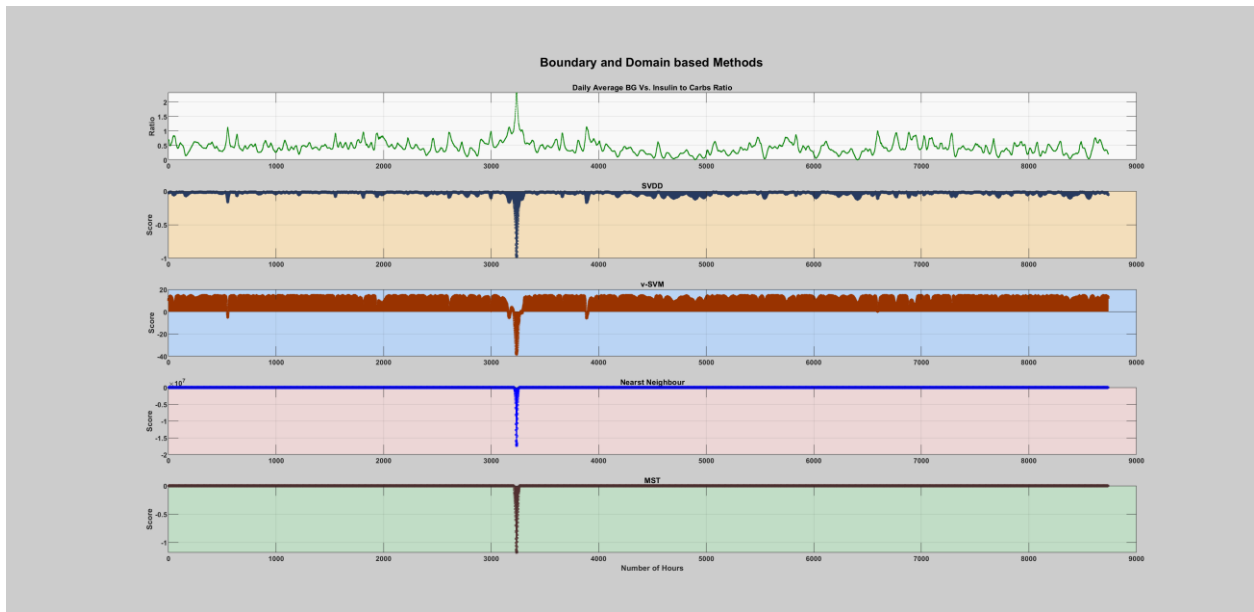


Figure 16: Boundary and domain-based.

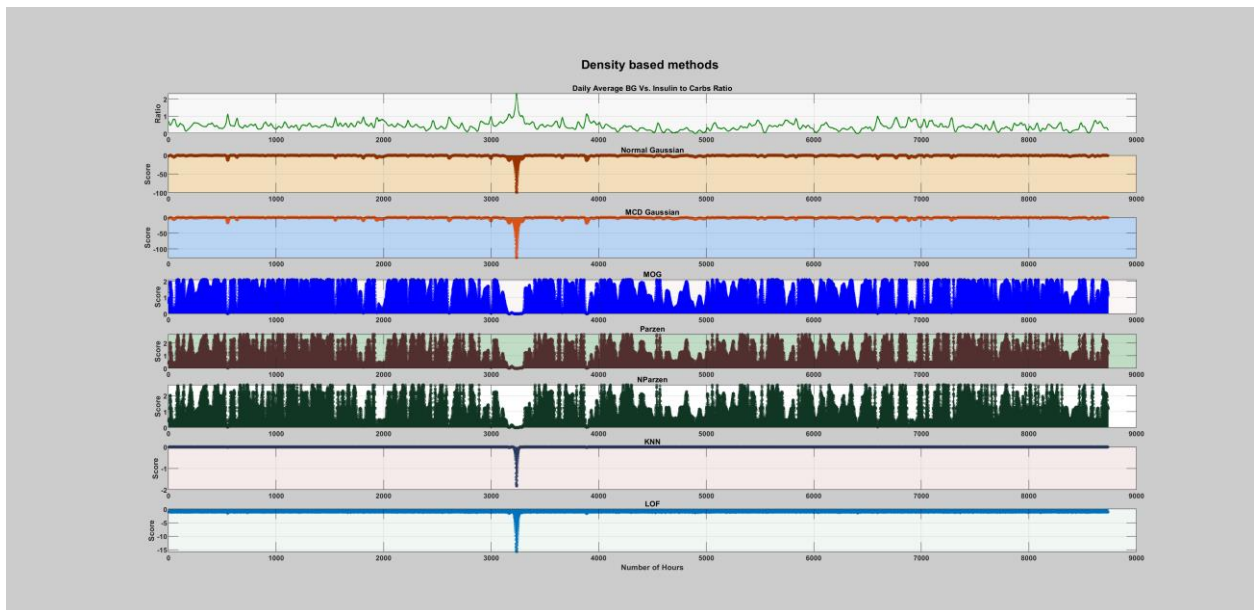
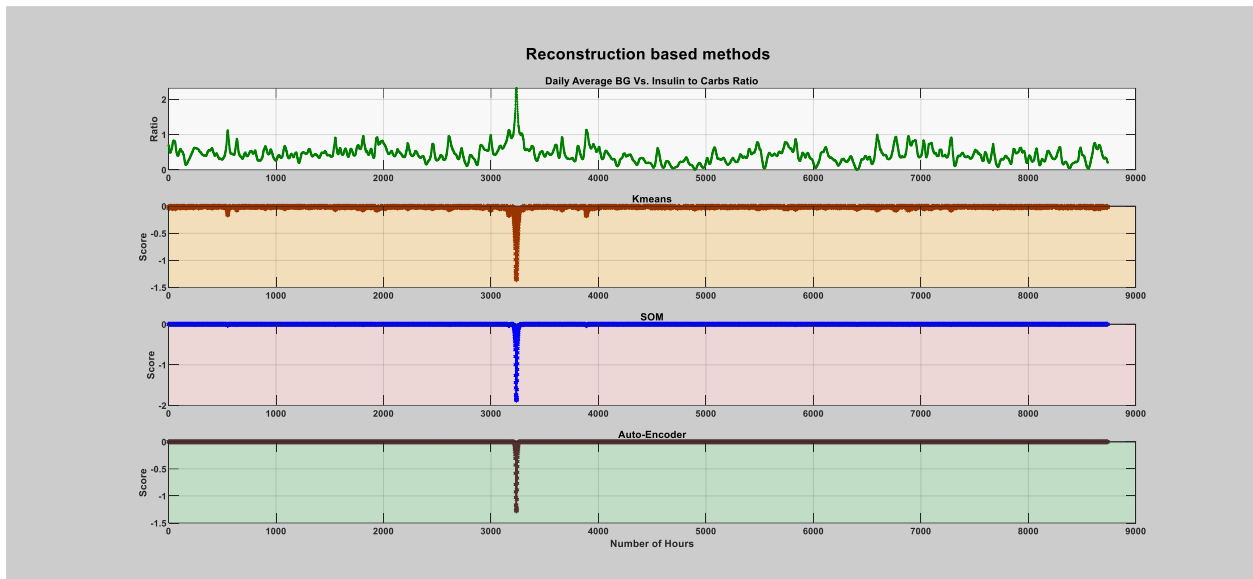


Figure 17: Density-based method.



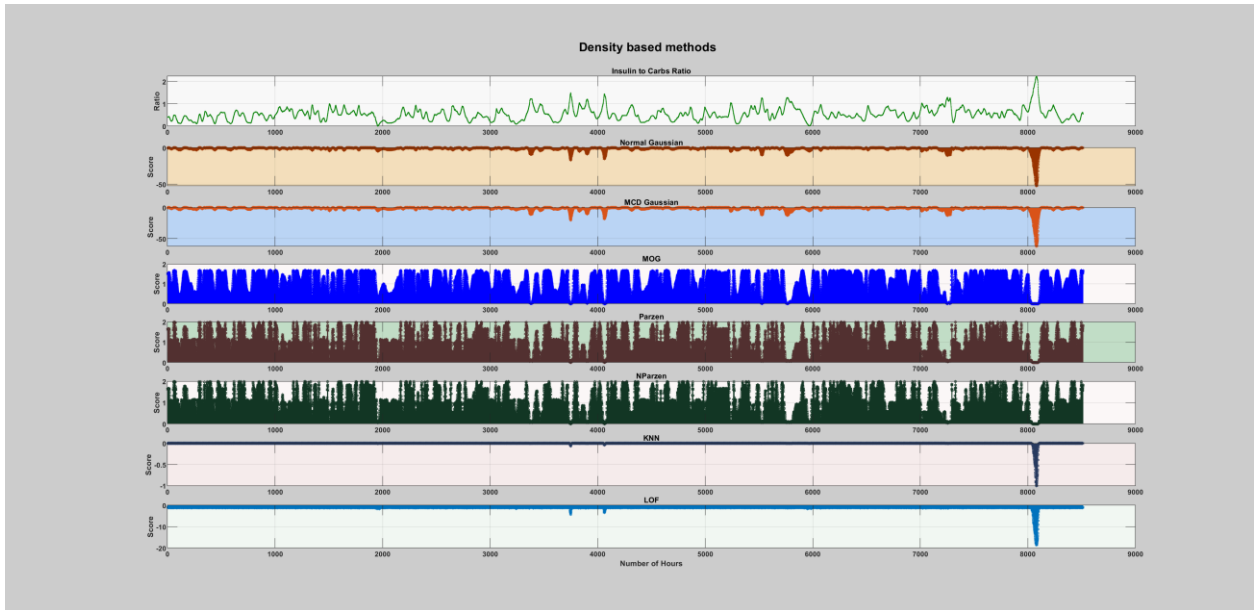
**Figure 18:** Reconstruction-based method.

### 10.1.2.3. The Third Infection Episode (Flu)

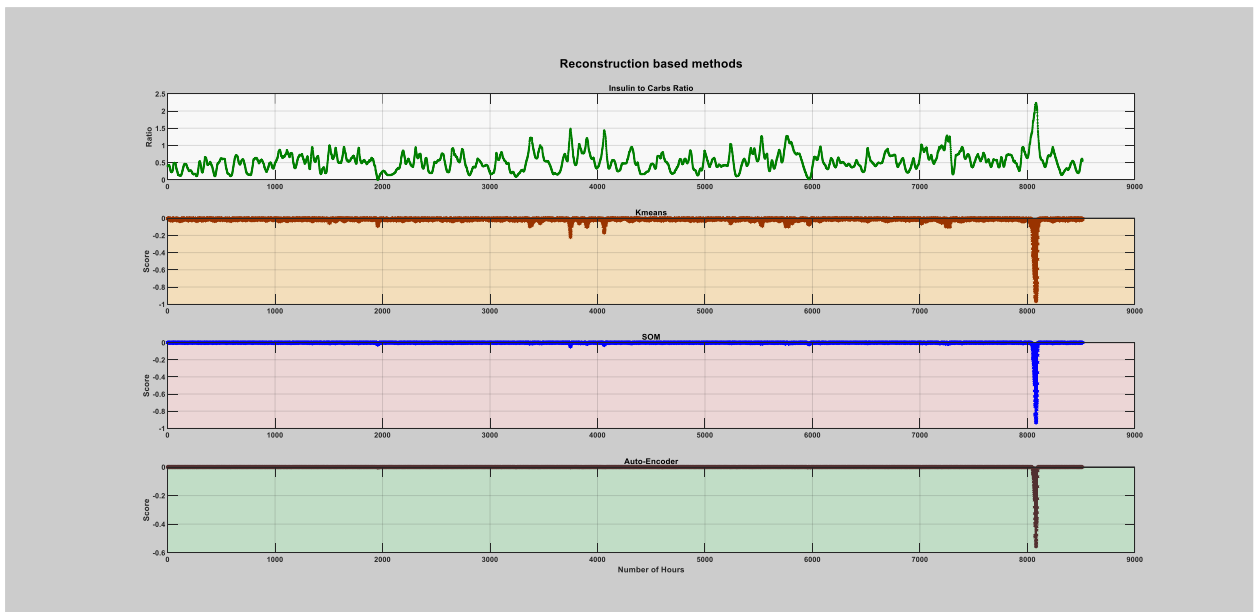


**Figure 19:** Boundary and domain-based method.





**Figure 20:** Density-based method.



**Figure 21:** Reconstruction-based method.

### 10.1.2.4. The Fourth Infection Episode (Flu)

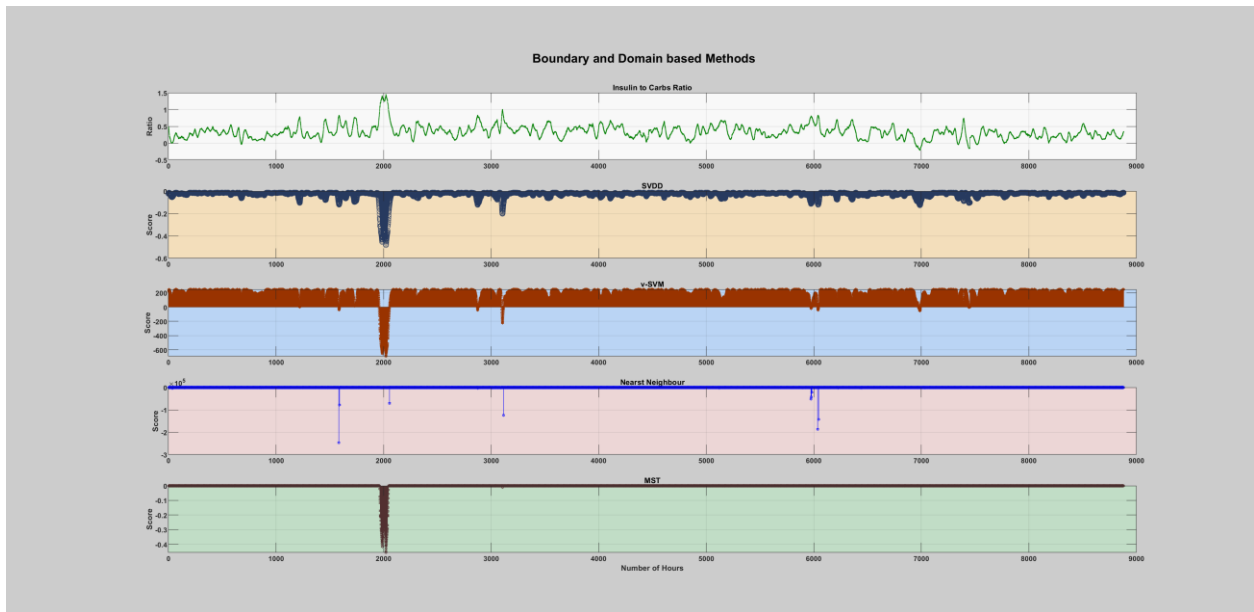


Figure 22: Boundary and domain-based method.

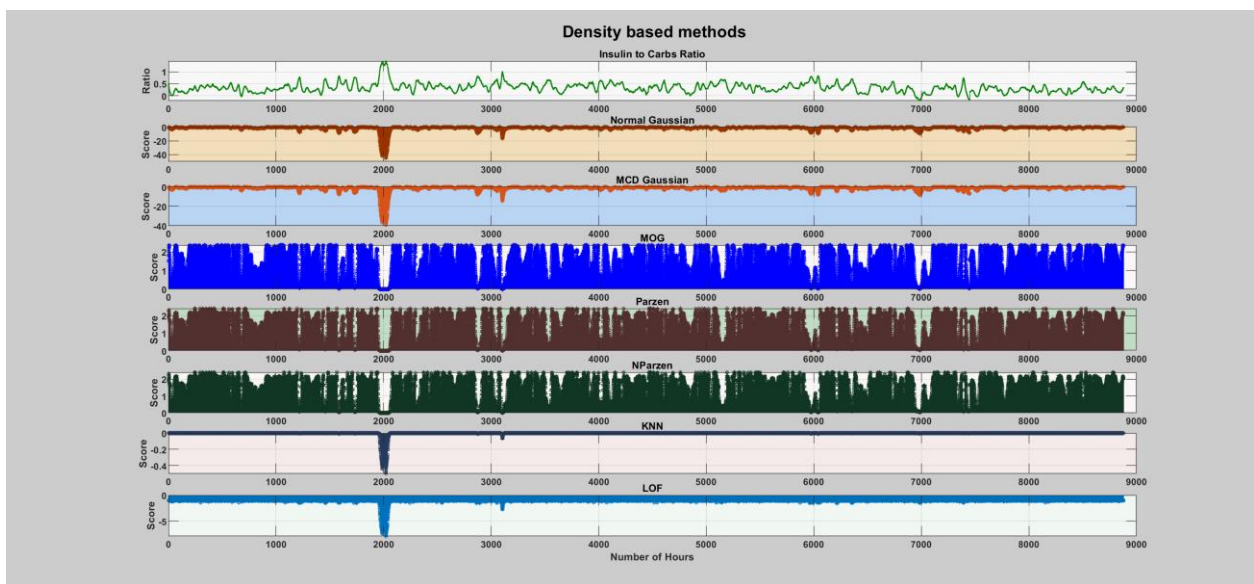


Figure 23: Density-based method.

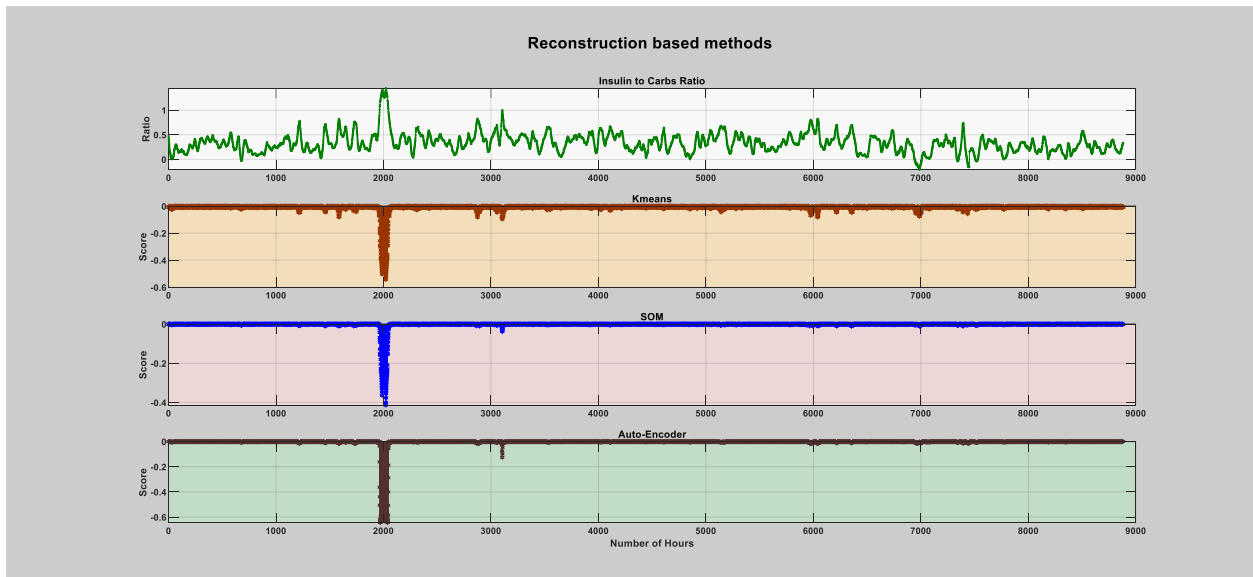


Figure 24: Reconstruction-based method.

## 10.2. Unsupervised method

### 10.2.1. Daily

#### 10.2.1.1. The First Infection Episode (Flu)

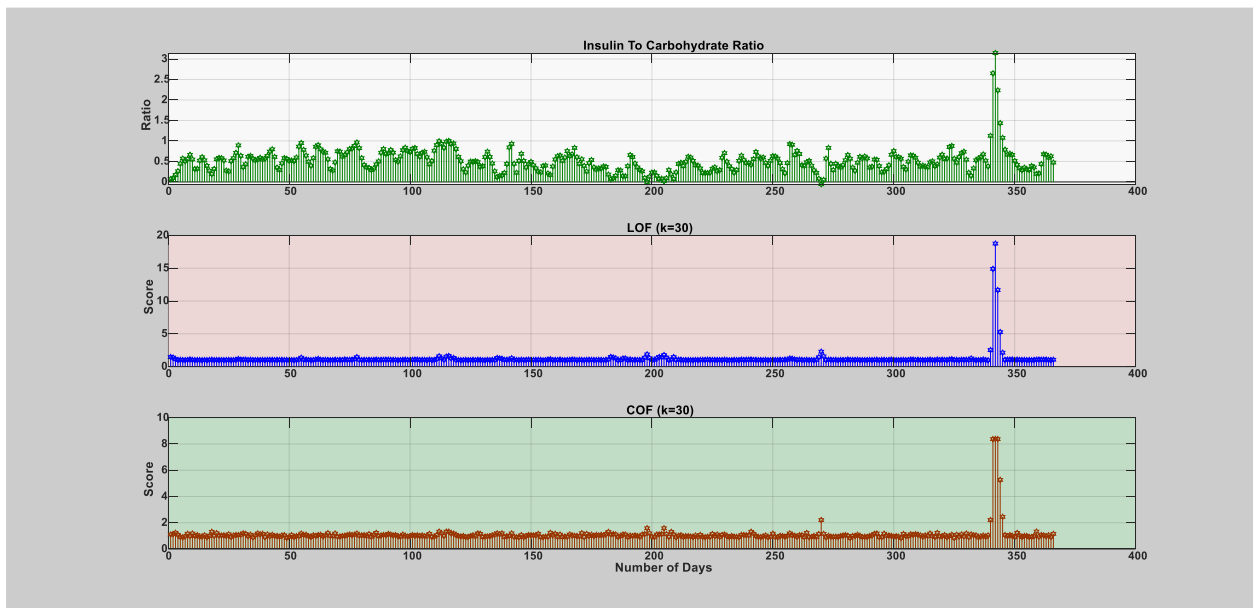


Figure 25: LOF and COF ( $k = 30$  data points).

### 10.2.1.2. The Second Infection Episode (Flu)

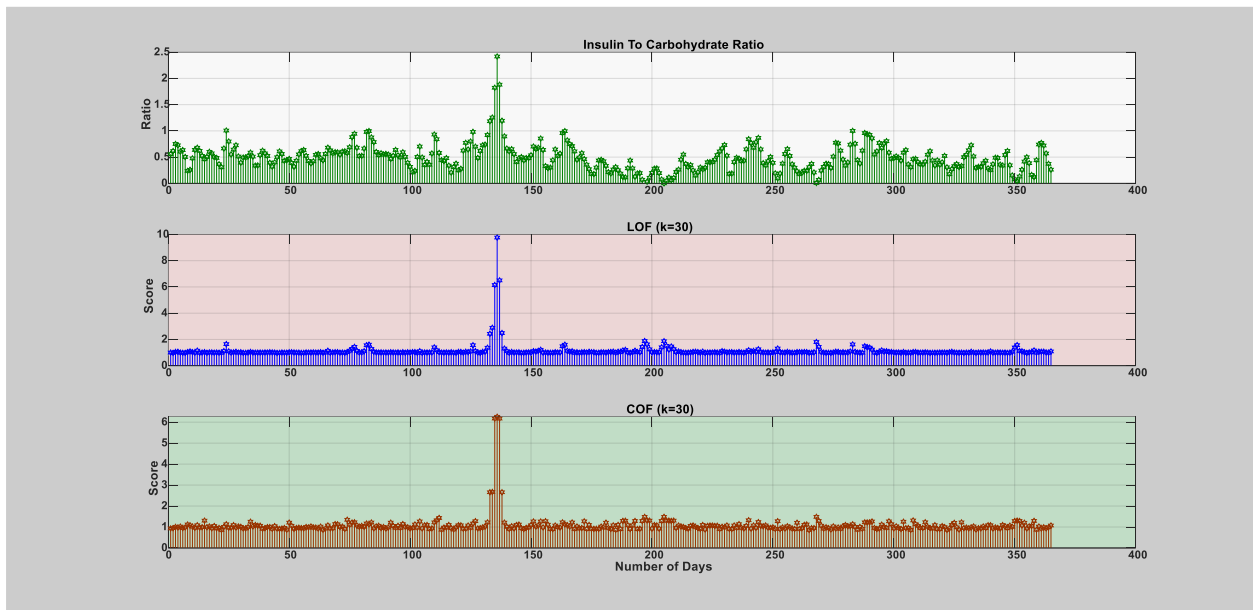


Figure 26: LOF and COF ( $k = 30$  data points).

### 10.2.1.3. The Third Infection Episode (Flu)

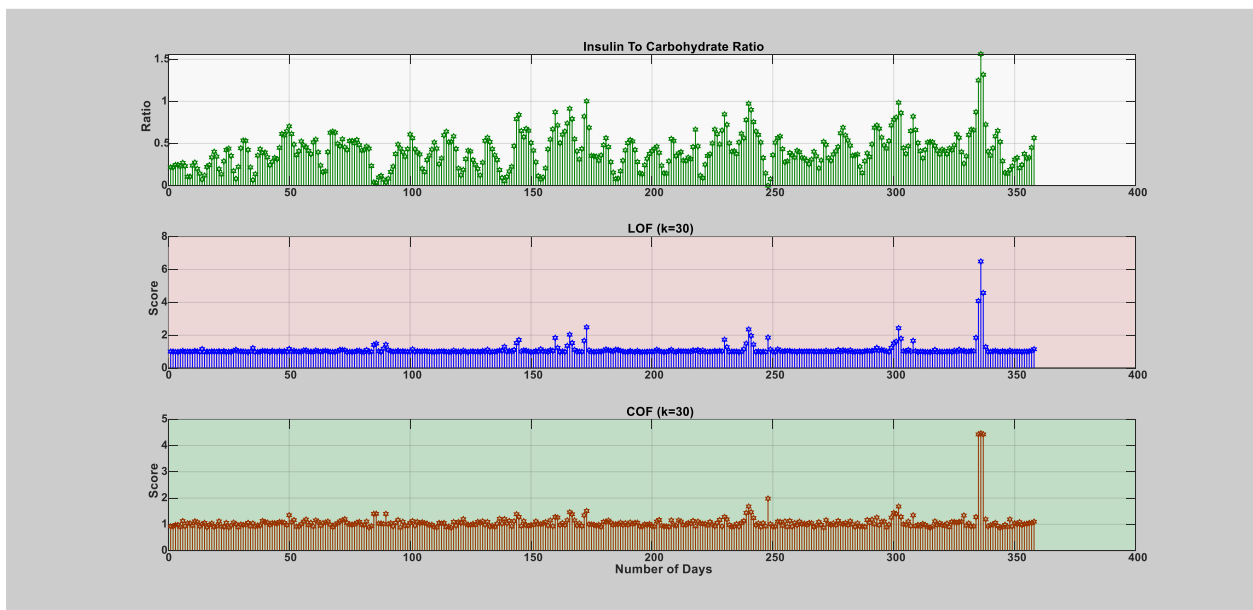


Figure 27: LOF and COF ( $k = 30$  data points).

### 10.2.1.4. The Fourth Infection Episode (Flu)

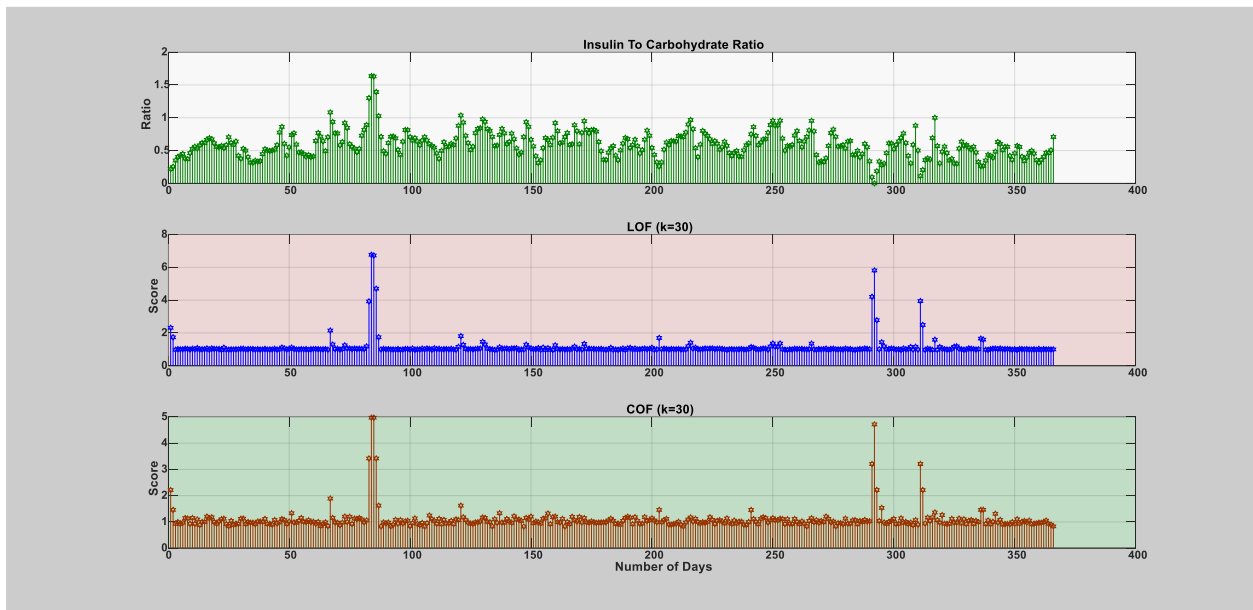


Figure 28: LOF and COF (k = 30 data points).

## 10.2.2. Hourly

### 10.2.2.1. The First Infection Episode (Flu)

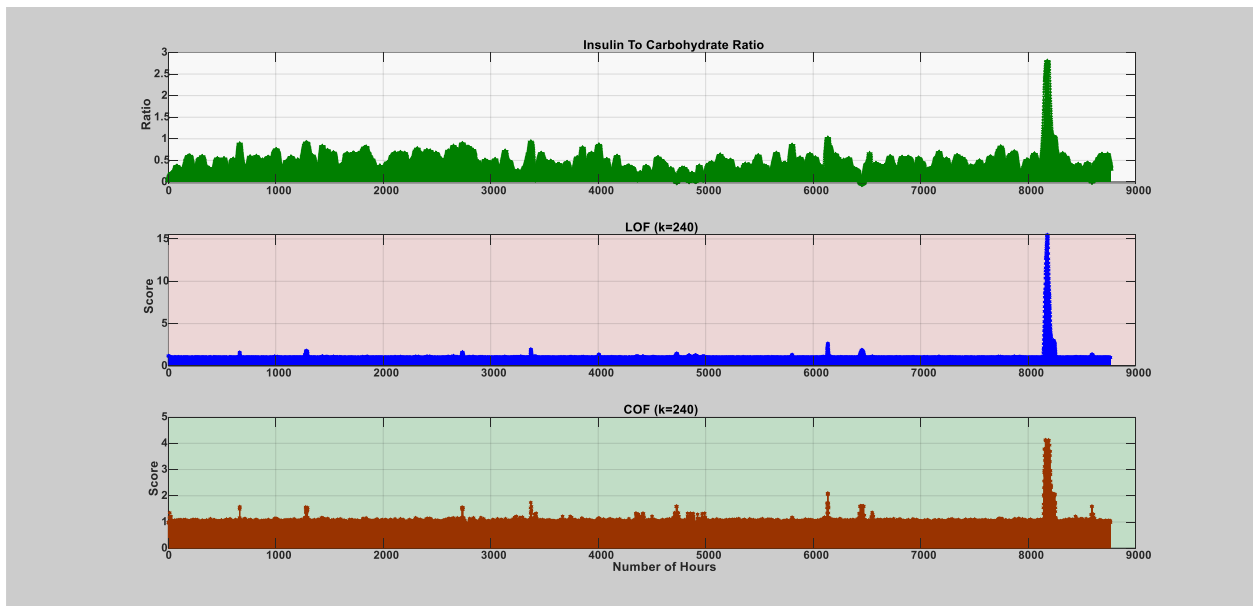


Figure 29: LOF and COF (k = 240 data points).

### 10.2.2.2. The Second Infection Episode (Flu)

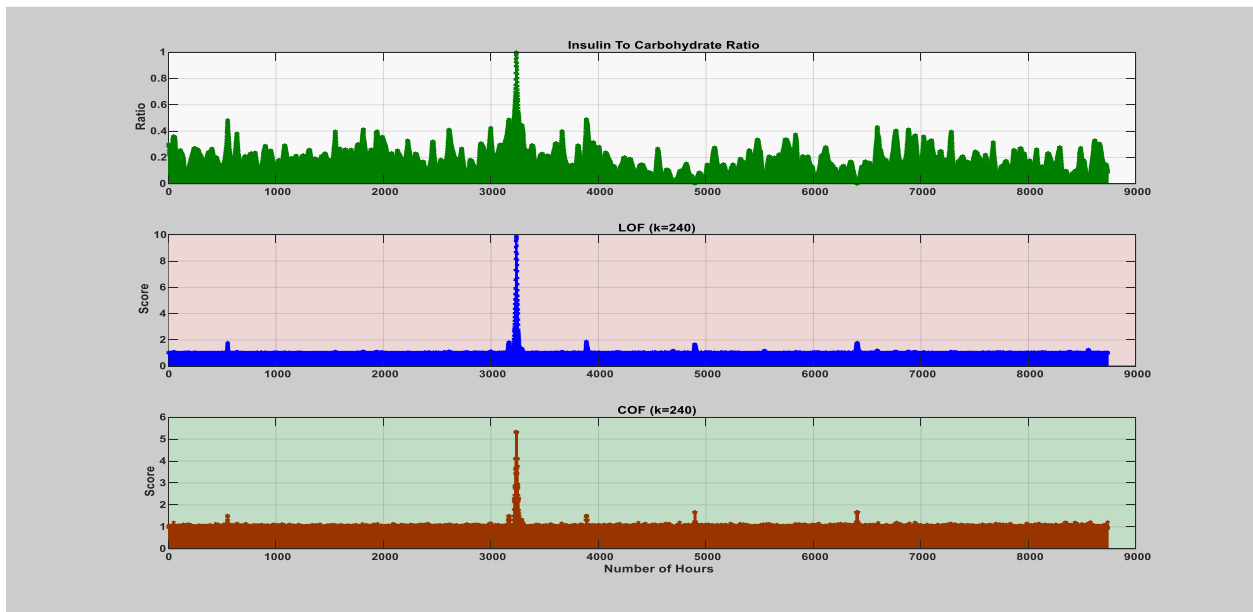


Figure 30: LOF and COF( $k = 240$  data points).

### 9.2.2.3. The Third Infection Episode (Flu)

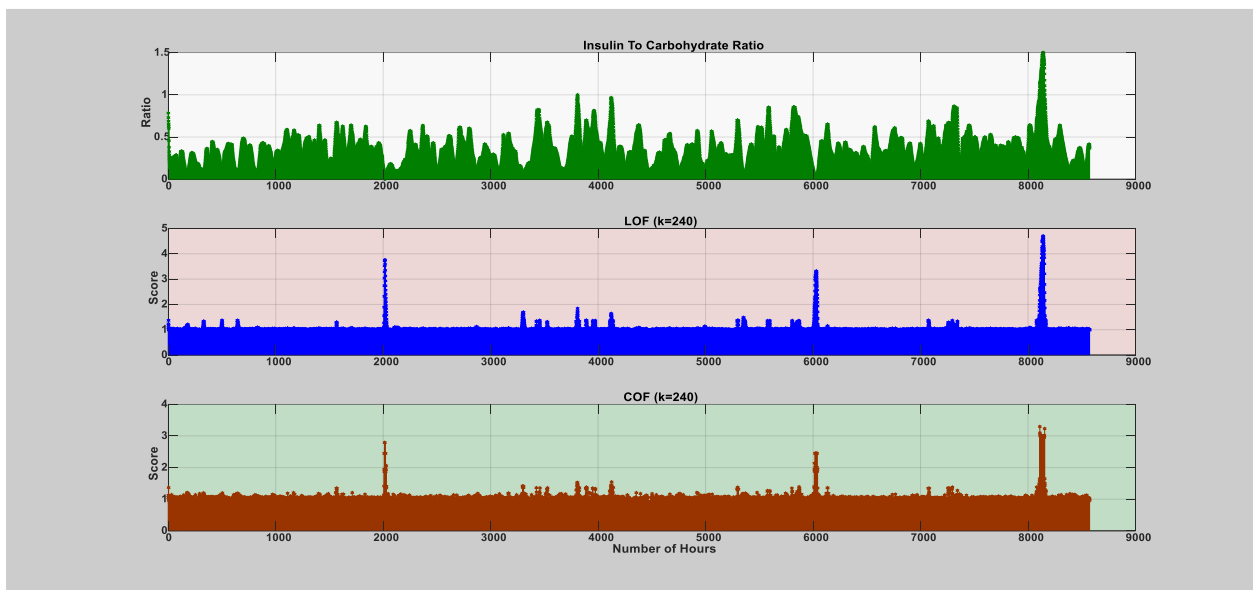


Figure 31: LOF and COF( $k = 240$  data points).

10.2.2.4. The Fourth Infection Episode (Flu)

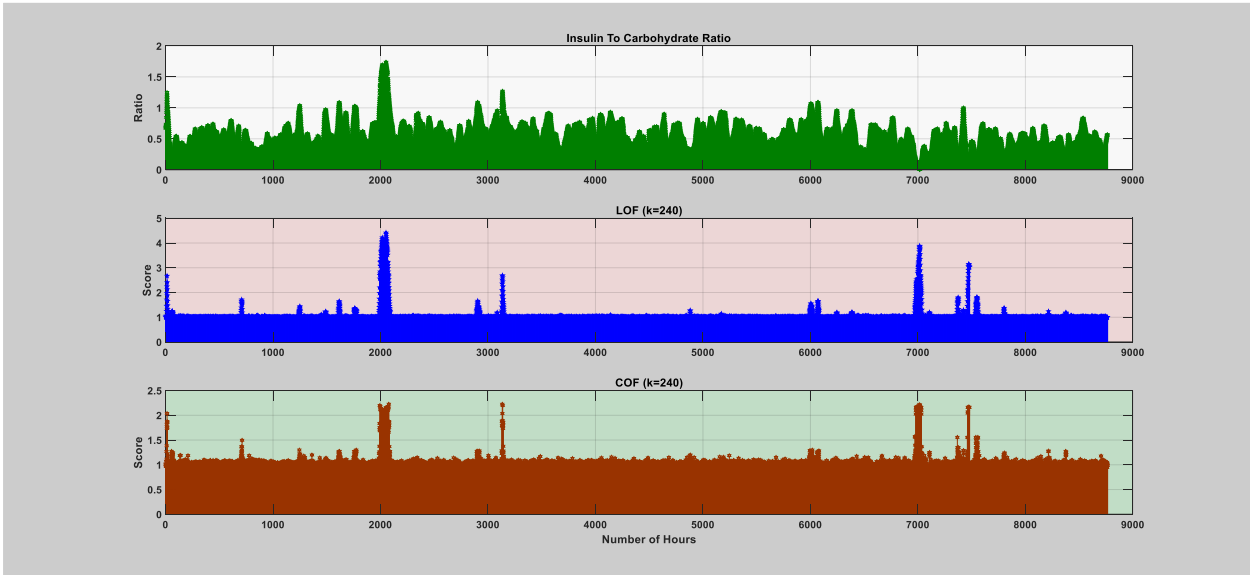


Figure 32: LOF and COF(k = 240 data points).

# 11 Appendix B – Models performance using the univariate input feature: Ratio of Insulin-to-Carbohydrate

This appendix presents the evaluation of the models in each patient-year. The performance is depicted in terms of the area under the ROC curve (AUC), specificity, and F1-score. The table depicts the performance of each model in regard to different evaluation criteria; 1) model performance across each infection episode and thereby depicting performance variations among individuals, 2) sample sizes depicts model performance with limited data sample sizes, 3) data granularity, i.e. daily and hourly, depicts the model's performance in response to variations in detail within the data, 4) data nature, i.e. raw and smoothed data (moving average window size = 2 days or 48 hrs), depicts the model performance improvement gained by removing short term and fast scale features from the data.

## 11.1. One-class classifier Method

### 11.1.1. Daily

#### 11.1.1.1. The 1<sup>st</sup> Infection Episode (Flu)

**Table 1:** Univariate Input Raw Data (different sample sizes) - Average and standard deviation (F1-score, AUC, specificity).

Fraction = 0.01												
Models	Boundary and Domain-Based Method											
	1 Month			2 Months			3 Months			4 Months		
	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1
SVDD	86.4 (6.2)	83.3 (0.0)	82.7 (6.3)	88.4 (4.1)	83.3 (0.0)	90.8 (3.2)	92.6 (2.2)	88.9 (0.0)	92.3 (5.0)	88.5 (3.1)	83.3 (0.0)	91.0 (2.9)
<i>v</i> -SVM	86.9 (4.7)	83.3 (0.0)	88.4 (3.6)	90.9 (3.4)	83.3 (0.0)	91.4 (2.8)	96.1 (1.6)	88.9 (0.0)	94.1 (2.2)	93.6 (2.7)	83.3 (0.0)	91.8 (1.4)
Nearest Neighbor	84.6 (6.8)	31.7 (13.9)	59.1 (4.1)	93.8 (5.7)	43.3 (8.2)	76.6 (2.9)	89.9 (1.9)	66.7 (0.0)	85.3 (1.8)	92.9 (5.5)	48.3 (5.0)	79.2 (1.6)
MST	90.3 (3.2)	50.0 (0.0)	66.0 (2.0)	90.9 (1.8)	65.0 (5.0)	84.6 (2.3)	94.2 (0.5)	78.9 (3.3)	90.0 (0.5)	91.5 (0.4)	68.3 (5.0)	86.5 (2.0)
Density-Based Method												
Gaussian OC	89.0 (6.4)	83.3 (0.0)	82.7 (6.3)	90.8 (3.5)	83.3 (0.0)	90.6 (3.7)	95.6 (2.1)	88.9 (0.0)	93.5 (3.7)	93.0 (3.3)	83.3 (0.0)	91.9 (1.3)
MoG	89.0 (6.4)	83.3 (0.0)	82.7 (6.3)	90.8 (3.5)	83.3 (0.0)	90.6 (3.7)	95.6 (2.1)	88.9 (0.0)	93.5 (3.7)	93.0 (3.3)	83.3 (0.0)	91.9 (1.3)
MCD Gaussian	91.1 (5.1)	83.3 (0.0)	84.1 (5.9)	94.6 (2.7)	83.3 (0.0)	90.6 (3.7)	96.4 (1.5)	88.9 (0.0)	93.5 (3.7)	94.5 (2.6)	83.3 (0.0)	91.9 (1.3)
Parzen	91.8 (5.4)	73.3 (8.2)	78.2 (4.9)	94.1 (3.3)	83.3 (0.0)	90.8 (3.2)	97.2 (1.5)	88.9 (0.0)	93.5 (3.7)	95.7 (2.7)	83.3 (0.0)	91.0 (2.9)
Naive Parzen	91.8 (5.4)	73.3 (8.2)	78.2 (4.9)	94.1 (3.3)	83.3 (0.0)	90.8 (3.2)	97.2 (1.5)	88.9 (0.0)	93.5 (3.7)	95.7 (2.7)	83.3 (0.0)	91.0 (2.9)
k-NN	89.0 (5.3)	83.3 (0.0)	84.6 (3.2)	93.7 (3.0)	83.3 (0.0)	91.5 (2.8)	97.1 (1.4)	88.9 (0.0)	93.5 (3.7)	95.7 (2.7)	83.3 (0.0)	91.9 (1.3)
LOF	85.6 (3.7)	83.3 (0.0)	84.1 (5.9)	94.6 (2.7)	81.7 (5.0)	90.9 (3.2)	96.8 (1.5)	87.8 (3.3)	93.1 (3.9)	94.4 (3.7)	63.3 (14.6)	84.0 (6.7)
Reconstruction-Based Method												
Auto – encoder	88.1 (10)	68.4 (17.7)	73.6 (11)	92.5 (7.0)	76.2 (14)	87.1 (6.7)	94.8 (4.8)	82.7 (9.7)	89.8 (6.0)	92.4 (7.2)	75.5 (11.1)	87.7 (5.4)
SOM	94.7 (5.9)	77.6 (11.8)	77.1 (10)	97.3 (2.9)	83.0 (2.9)	88.9 (6.0)	95.7 (2.2)	88.9 (0.0)	93.5 (2.6)	94.8 (2.1)	83.3 (0.0)	90.6 (3.0)
K-means	89.0 (6.4)	83.3 (0.0)	82.7 (6.3)	90.8 (3.5)	83.3 (0.0)	90.6 (3.7)	95.6 (2.1)	88.9 (0.0)	93.5 (3.7)	93.0 (3.3)	83.3 (0.0)	91.9 (1.3)



**Table 2:** Univariate Input Smoothed Data (different sample sizes) - Average and standard deviation (F1-score, AUC, specificity).

Fraction = 0.01												
Models	Boundary and Domain-Based Method											
	1 Month			2 Months			3 Months			4 Months		
	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1
SVDD	100 (0.0)	100 (0.0)	93.6 (15.2)	100 (0.0)	100 (0.0)	95.8 (10.1)	100 (0.0)	100 (0.0)	97.5 (5.0)	100 (0.0)	100 (0.0)	98.3 (3.5)
<i>v-SVM</i>	100 (0.0)	100 (0.0)	98.9 (3.4)	100 (0.0)	100 (0.0)	99.2 (2.5)	100 (0.0)	100 (0.0)	99.4 (2.0)	100 (0.0)	100 (0.0)	99.6 (1.2)
Nearest Neighbor	98.3 (3.6)	63.3 (10.0)	72.4 (5.7)	98.7 (2.2)	62 (19.8)	84.8 (5.8)	99.0 (1.7)	63.3 (13.2)	84.4 (5.2)	98.9 (1.5)	60 (15.3)	83.4 (4.2)
MST	100 (0.0)	98.3 (5.0)	93.6 (15.2)	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	97.5 (5.0)	100 (0.0)	100 (0.0)	99.2 (2.5)
Density-Based Method												
Gaussian OC	100 (0.0)	100 (0.0)	93.6 (15.2)	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	98.8 (3.8)	100 (0.0)	100 (0.0)	99.2 (2.5)
MoG	100 (0.0)	100 (0.0)	93.6 (15.2)	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	98.8 (3.8)	100 (0.0)	100 (0.0)	99.2 (2.5)
MCD Gaussian	100 (0.0)	100 (0.0)	96.6 (7.0)	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	98.8 (3.8)	100 (0.0)	100 (0.0)	99.2 (2.5)
Parzen	99.6 (1.3)	100 (0.0)	93.6 (15.2)	100 (0.0)	100 (0.0)	94.1 (15.0)	100 (0.0)	100 (0.0)	98.2 (4.0)	100 (0.0)	100 (0.0)	98.7 (2.7)
Naive Parzen	99.6 (1.3)	100 (0.0)	93.6 (15.2)	100 (0.0)	100 (0.0)	94.1 (15.0)	100 (0.0)	100 (0.0)	98.2 (4.0)	100 (0.0)	100 (0.0)	98.7 (2.7)
k-NN	100 (0.0)	100 (0.0)	93.6 (15.2)	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	98.8 (3.8)	100 (0.0)	100 (0.0)	99.2 (2.5)
LOF	100 (0.0)	100 (0.0)	93.6 (15.2)	100 (0.0)	100 (0.0)	96.7 (10.0)	100 (0.0)	100 (0.0)	98.8 (3.8)	100 (0.0)	100 (0.0)	99.2 (2.5)
Reconstruction-Based Method												
Auto – encoder	96.8 (7.8)	89.8 (16.3)	87.0 (16.7)	98.5 (4.0)	92.2 (14)	92.2 (12.4)	99.0 (3.3)	93.9 (12.1)	94.9 (6.4)	98.3 (4.1)	92.9 (12.0)	95.1 (6.0)
SOM	99.6 (1.3)	100 (0.0)	90.4 (17.8)	100 (0.0)	100 (0.0)	92.9 (15.2)	100 (0.0)	100 (0.0)	96.1 (6.4)	100 (0.0)	100 (0.0)	97.4 (3.6)
K-means	100 (0.0)	100 (0.0)	93.6 (15.2)	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	98.8 (3.8)	100 (0.0)	100 (0.0)	99.2 (2.5)

### 11.1.1.2. The 2<sup>nd</sup> Infection Episode (Flu)

**Table 3:** Univariate Input Raw Data (different sample sizes) - Average and standard deviation (F1-score, AUC, specificity).

Fraction = 0.01												
Models	Boundary and Domain-Based Method											
	1 Month			2 Months			3 Months			4 Months		
	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1
SVDD	77.2 (11.5)	20 (10.0)	53.3 (5.2)	82.2 (10.2)	23.3 (20.1)	69.4 (5.5)	86.1 (6.9)	28.9 (20.1)	72.7 (3.7)	80.4 (7.5)	23.3 (20.1)	71.6 (4.1)
<i>v-SVM</i>	86.1 (9.3)	18.8 (8.2)	61.8 (6.8)	94.4 (6.0)	21.7 (13.7)	66.7 (7.2)	96.8 (4.4)	27.3 (13.5)	69.4 (5.9)	97.3 (3.4)	38.9 (12.6)	74.1 (5.2)
Nearest Neighbor	56.1 (11.8)	0.0 (0.0)	49.4 (4.8)	52.3 (14.1)	1.7 (5.0)	66.6 (3.0)	59.1 (10.7)	18.9 (15.8)	68.8 (3.5)	53.8 (8.5)	11.7 (7.7)	68.4 (1.9)
MST	56.7 (4.2)	16.7 (0.0)	53.8 (5.1)	59.0 (4.9)	21.7 (15.0)	70.5 (1.6)	62.8 (6.6)	26.7 (13.4)	72.4 (0.8)	60.3 (7.3)	21.7 (15.0)	71.4 (1.9)
Density-Based Method												
Gaussian OC	86.7 (9.4)	20 (10.0)	54.8 (2.6)	94.2 (5.9)	38.3 (15.0)	75.1 (1.1)	96.7 (4.4)	38.9 (16.7)	75.9 (2.2)	97.1 (3.3)	38.3 (15.0)	76.0 (2.5)
MoG	85.8 (11.9)	20.2 (9.7)	53.2 (5.6)	93.1 (7.6)	31.6 (19.1)	72.3 (3.2)	96.6 (4.5)	38.9 (16.7)	75.9 (2.2)	97.3 (3.3)	38.3 (15.0)	76.0 (2.5)
MCD Gaussian	91.7 (8.7)	20 (10.0)	54.8 (2.6)	94.8 (6.3)	38.3 (15.0)	75.1 (1.1)	97.1 (4.5)	38.9 (16.7)	75.9 (2.2)	97.7 (3.5)	38.3 (15.0)	76.0 (2.5)
Parzen	90.0 (10.5)	23.3 (13.4)	54.3 (3.0)	94.2 (6.9)	33.3 (18.3)	73.8 (1.7)	96.8 (4.6)	37.8 (17.4)	75.6 (2.5)	97.6 (3.4)	31.7 (19.0)	74.2 (3.8)
Naive Parzen	90.0 (10.5)	23.3 (13.4)	54.3 (3.0)	94.2 (6.9)	33.3 (18.3)	73.8 (1.7)	96.8 (4.6)	36.7 (18.0)	75.3 (2.5)	97.6 (3.4)	31.7 (19.0)	74.2 (3.8)
k-NN	91.7 (8.7)	18.3 (5.0)	54.2 (3.9)	95.4 (6.5)	26.7 (20.1)	72.0 (1.8)	97.1 (4.5)	36.7 (18.0)	75.3 (2.5)	97.1 (3.3)	33.3 (18.3)	74.7 (3.4)
LOF	87.6 (7.4)	18.3 (5.0)	54.2 (3.9)	93.4 (7.0)	21.7 (15.0)	70.5 (1.6)	95.3 (4.6)	26.7 (13.4)	72.4 (0.8)	95.7 (3.6)	23.3 (20.1)	72.0 (3.8)
Reconstruction-Based Method												
Auto – encoder	72.9 (16.1)	31.3 (19.6)	57.2 (7.5)	71.6 (14.7)	28.4 (18.7)	71.5 (5.4)	74.2 (12.7)	34.2 (18.0)	74.4 (4.7)	74.6 (14)	29.6 (19.7)	73.2 (4.8)
SOM	77.1 (9.6)	62.9 (10.6)	64.5 (9.5)	78.4 (6.2)	42.8 (12.7)	72.2 (6.8)	82.6 (6.1)	49.7 (11.6)	76.7 (5.3)	84.6 (4.9)	43.1 (14.1)	76.6 (3.2)
K-means	86.7 (9.4)	20.0 (10.0)	54.8 (2.6)	94.2 (5.9)	38.3 (15.0)	75.1 (1.1)	96.7 (4.4)	38.9 (16.7)	75.9 (2.2)	97.1 (3.3)	38.3 (15.0)	76.0 (2.5)

**Table 4:** Univariate Input Smoothed Data (different sample sizes) - Average and standard deviation (F1-score, AUC, specificity).

Fraction = 0.01												
Models	Boundary and Domain-Based Method											
	1 Month			2 Months			3 Months			4 Months		
	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1
SVDD	100 (0.0)	100 (0.0)	90.0 (20.1)	100 (0.0)	100 (0.0)	96.3 (7.4)	100 (0.0)	100 (0.0)	98.2 (4.0)	100 (0.0)	100 (0.0)	96.9 (7.9)
<i>v-SVM</i>	100 (0.0)	100 (0.0)	99.5 (2.9)	100 (0.0)	100 (0.0)	99.1 (2.5)	100 (0.0)	100 (0.0)	99.5 (1.6)	100 (0.0)	100 (0.0)	99.6 (1.3)
Nearest Neighbor	100 (0.0)	91.7 (25.1)	90.5 (19.2)	99.3 (2.1)	90 (30.1)	96.1 (9.1)	91.7 (10.7)	12.2 (12.7)	68.8 (1.7)	95.4 (4.3)	35 (19.0)	76 (5.3)
MST	100 (0.0)	100 (0.0)	100 (0.0)	100 (0.0)	100 (0.0)	99.1 (2.7)	100 (0.0)	100 (0.0)	99.4 (1.8)	100 (0.0)	100 (0.0)	97.4 (7.9)
Density-Based Method												
Gaussian OC	100 (0.0)	100 (0.0)	93 (15.6)	100 (0.0)	100 (0.0)	95.3 (9.6)	100 (0.0)	100 (0.0)	98.8 (2.4)	100 (0.0)	100 (0.0)	97.4 (7.9)
MoG	100 (0.0)	100 (0.0)	88.6 (19.8)	100 (0.0)	100 (0.0)	95.1 (10.2)	100 (0.0)	100 (0.0)	97.8 (4.0)	100 (0.0)	100 (0.0)	97.1 (7.9)
MCD Gaussian	100 (0.0)	100 (0.0)	95 (15.0)	100 (0.0)	100 (0.0)	97.3 (8.2)	100 (0.0)	100 (0.0)	97.4 (6.1)	100 (0.0)	100 (0.0)	97.4 (7.9)
Parzen	100 (0.0)	100 (0.0)	88 (19.9)	100 (0.0)	100 (0.0)	93.9 (12.2)	100 (0.0)	100 (0.0)	96.2 (6.7)	100 (0.0)	100 (0.0)	96.9 (7.9)
Naive Parzen	100 (0.0)	100 (0.0)	88 (19.9)	100 (0.0)	100 (0.0)	93.9 (12.2)	100 (0.0)	100 (0.0)	97.6 (4.1)	100 (0.0)	100 (0.0)	96.9 (7.9)
k-NN	100 (0.0)	100 (0.0)	95 (15.0)	100 (0.0)	100 (0.0)	98.3 (5.0)	100 (0.0)	100 (0.0)	98.8 (2.4)	100 (0.0)	100 (0.0)	97.4 (7.9)
LOF	100 (0.0)	100 (0.0)	95(15.0)	100 (0.0)	100 (0.0)	96.7 (10.0)	98.4 (4.8)	67.8 (23.1)	85.3 (10.9)	100 (0.0)	100 (0.0)	97.6 (6.0)
Reconstruction-Based Method												
Auto – encoder	99.4 (3.6)	96.3 (11.9)	90.2 (15.8)	99.3 (3.5)	97.3 (9.7)	95.2 (8.6)	95.5 (9.0)	86.4 (20.5)	91.7 (10.7)	96.8 (7.3)	90 (16.7)	93.8 (8.3)
SOM	100 (0.0)	100 (0.0)	90.1 (17.8)	100 (0.0)	100 (0.0)	95.9 (7.4)	100 (0.0)	100 (0.0)	97.0 (4.9)	100 (0.0)	100 (0.0)	96.9 (8.0)
K-means	100 (0.0)	100 (0.0)	93.0 (15.6)	100 (0.0)	100 (0.0)	95.3 (9.6)	100 (0.0)	100 (0.0)	98.2 (4.0)	100 (0.0)	100 (0.0)	97.4 (7.9)

### 11.1.1.3. The 3<sup>rd</sup> Infection Episode (Flu)

**Table 5:** Univariate Input Raw Data (different sample sizes) - Average and standard deviation (F1-score, AUC, specificity).

Fraction = 0.01												
Models	Boundary and Domain-Based Method											
	1 Month			2 Months			3 Months			4 Months		
	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1
SVDD	68.9 (4.5)	66.7 (0.0)	80.1 (13.8)	75.4 (6.6)	66.7 (0.0)	88.8 (6.2)	74.3 (3.0)	66.7 (0.0)	88.9 (2.4)	73.5 (4.9)	66.7 (0.0)	91.0 (2.9)
<i>v-SVM</i>	69.4 (5.9)	66.7 (0.0)	84.5 (4.1)	69.1 (4.3)	66.7 (0.0)	86.5 (4.1)	67.8 (2.3)	66.7 (0.0)	86.1 (2.5)	71.0 (4.2)	66.7 (0.0)	87.0 (2.7)
Nearest Neighbor	93.6 (10.1)	6.7 (20.1)	68.9 (6.7)	74.7 (10.6)	0.0 (0.0)	80.2 (0.7)	77.1 (13.1)	0.0 (0.0)	74.6 (1.8)	74.6 (9.4)	0.0 (0.0)	79.7 (1.3)
MST	81.8 (3.4)	66.7 (0.0)	86.0 (1.0)	82.8 (1.7)	66.7 (0.0)	90.5 (5.9)	83.0 (0.7)	66.7 (0.0)	89.5 (1.8)	82.9 (0.9)	66.7 (0.0)	91.5 (2.7)
Density-Based Method												
Gaussian OC	68.9 (4.5)	66.7 (0.0)	80.1 (13.8)	68.8 (3.6)	66.7 (0.0)	90.5 (5.9)	67.4 (1.4)	66.7 (0.0)	89.5 (1.8)	68.9 (2.7)	66.7 (0.0)	91.5 (2.7)
MoG	70.1 (6.3)	66.7 (0.0)	79.0 (14.0)	72.8 (6.6)	66.7 (0.0)	90.5 (5.9)	73.5 (3.9)	66.7 (0.0)	88.9 (2.4)	78.8 (7.0)	66.7 (0.0)	91.5 (2.7)
MCD Gaussian	68.9 (4.5)	66.7 (0.0)	82.7 (6.3)	68.3 (2.4)	66.7 (0.0)	90.5 (5.9)	69.2 (3.7)	66.7 (0.0)	89.5 (1.8)	71.6 (4.1)	66.7 (0.0)	91.5 (2.7)
Parzen	67.8 (3.3)	66.7 (0.0)	80.1 (13.8)	71.5 (5.1)	66.7 (0.0)	90.5 (5.9)	70.3 (4.3)	66.7 (0.0)	88.9 (2.4)	74.7 (4.8)	66.7 (0.0)	91.5 (2.7)
Naive Parzen	67.8 (3.3)	66.7 (0.0)	80.1 (13.8)	71.5 (5.1)	66.7 (0.0)	90.5 (5.9)	70.3 (4.3)	66.7 (0.0)	89.5 (1.8)	74.7 (4.8)	66.7 (0.0)	91.5 (2.7)
k-NN	79.7 (9.8)	66.7 (0.0)	84.6 (3.2)	68.9 (2.7)	66.7 (0.0)	90.5 (5.9)	74.3 (4.5)	66.7 (0.0)	89.5 (1.8)	77.7 (3.1)	66.7 (0.0)	91.5 (2.7)
LOF	71.1 (7.4)	66.7 (0.0)	82.7 (6.3)	72.1 (5.4)	66.7 (0.0)	90.5 (5.9)	71.1 (4.3)	66.7 (0.0)	89.5 (1.8)	77.8 (6.3)	66.7 (0.0)	91.5 (2.7)
Reconstruction-Based Method												
Auto – encoder	80.8 (10.9)	65.8 (5.2)	81.9 (10.5)	88.5 (9.8)	65.8 (5.2)	89.7 (5.4)	84.7 (11.8)	58.3 (16.3)	86.4 (5.5)	88.5 (9.9)	58 (14.9)	89.2 (4.2)
SOM	93.9 (8.2)	67.7 (5.7)	80.6 (9.7)	91.2 (5.1)	66.7 (0.0)	90.5 (5.9)	88.3 (6.0)	66.7 (0.0)	89.5 (1.8)	89.6 (5.0)	66.7 (0.0)	91.5 (2.7)
K-means	83.2 (12.7)	66.7 (0.0)	81.3 (9.8)	85.6 (8.5)	66.7 (0.0)	89.7 (6.1)	86.8 (7.5)	66.7 (0.0)	89.1 (2.6)	88.5 (6.5)	66.7 (0.0)	91.4 (2.8)

**Table 6:** Univariate Input Smoothed Data (different sample sizes) - Average and standard deviation (F1-score, AUC, specificity).

Fraction = 0.01												
Models	Boundary and Domain-Based Method											
	1 Month			2 Months			3 Months			4 Months		
	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1
SVDD	100 (0.0)	100 (0.0)	93.6 (15.2)	100 (0.0)	100 (0.0)	94.7 (11.1)	100 (0.0)	100 (0.0)	97.5 (5.0)	100 (0.0)	100 (0.0)	98.0 (6.0)
$\nu$ -SVM	100 (0.0)	100 (0.0)	98.6 (4.2)	100 (0.0)	100 (0.0)	98.8 (3.7)	100 (0.0)	100 (0.0)	99.3 (2.0)	100 (0.0)	100 (0.0)	99.4 (1.8)
Nearest Neighbor	80.0 (24.3)	0.0 (0.0)	67.3 (1.8)	93 (10.1)	0.0 (0.0)	80.2 (0.7)	88.3 (8.5)	0.0 (0.0)	74.6 (1.8)	91.2 (5.8)	0.0 (0.0)	79.7 (0.8)
MST	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	97.1 (6.3)	100 (0.0)	100 (0.0)	98.8 (2.4)	100 (0.0)	100 (0.0)	99.1 (2.7)
Density-Based Method												
Gaussian OC	100 (0.0)	100 (0.0)	96.6 (7.0)	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	98.8 (3.8)	100 (0.0)	100 (0.0)	98.6 (4.3)
MoG	100 (0.0)	100 (0.0)	94.0 (15.1)	100 (0.0)	100 (0.0)	96.0 (8.0)	100 (0.0)	100 (0.0)	97.9 (4.5)	100 (0.0)	100 (0.0)	98.5 (4.3)
MCD Gaussian	100 (0.0)	100 (0.0)	98.6 (4.3)	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	98.8 (3.8)	100 (0.0)	100 (0.0)	99.1 (2.7)
Parzen	100 (0.0)	100 (0.0)	93.6 (15.2)	100 (0.0)	100 (0.0)	96.0 (8.0)	100 (0.0)	100 (0.0)	98.2 (4.0)	100 (0.0)	100 (0.0)	98.6 (4.3)
Naive Parzen	100 (0.0)	100 (0.0)	93.6 (15.2)	100 (0.0)	100 (0.0)	96.0 (8.0)	100 (0.0)	100 (0.0)	98.2 (4.0)	100 (0.0)	100 (0.0)	98.6 (4.3)
k-NN	100 (0.0)	100 (0.0)	96.6 (7.0)	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	98.8 (3.8)	100 (0.0)	100 (0.0)	98.6 (4.3)
LOF	100 (0.0)	100 (0.0)	96.6 (7.0)	100 (0.0)	100 (0.0)	98.0 (6.0)	99.3 (2.2)	96.7 (10.0)	97.8 (4.5)	100 (0.0)	100 (0.0)	98.6 (4.3)
Reconstruction-Based Method												
Auto – encoder	98.8 (7.2)	97.3 (11.7)	93.4 (14.6)	96.4 (9.4)	88.2 (23.1)	93.6 (8.9)	97.1 (7.2)	87.5 (24.0)	94.5 (7.5)	96.8 (8.1)	85.5 (23.5)	95.3 (6.1)
SOM	100 (0.0)	100 (0.0)	89.9 (15.6)	99.9 (0.9)	100 (0.0)	94.8 (10.8)	100 (0.0)	100 (0.0)	98.4 (4.1)	100 (0.0)	100 (0.0)	98.6 (4.3)
K-means	100 (0.0)	100 (0.0)	93.1 (14.6)	100 (0.0)	100 (0.0)	95.3 (9.4)	100 (0.0)	100 (0.0)	98.0 (4.3)	100 (0.0)	100 (0.0)	98.6 (4.3)

#### 11.1.1.4. The 4<sup>th</sup> Infection Episode (Flu)

**Table 7:** Univariate Input Raw Data (different sample sizes) - Average and standard deviation (F1-score, AUC, specificity).

Fraction = 0.01												
Models	Boundary and Domain-Based Method											
	1 Month			2 Months			3 Months			4 Months		
	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1
SVDD	95.0 (4.1)	75.0 (0.0)	82.2 (7.9)	98.3 (2.0)	75.0 (0.0)	90.6 (3.7)	92.7 (2.5)	85.7 (0.0)	92.3 (5.0)	87.2 (3.1)	75.0 (0.0)	91.5 (1.8)
$\nu$ -SVM	98.3 (3.4)	75.3 (2.5)	86.0 (5.0)	98.7 (1.9)	75.1 (1.8)	89.0 (3.5)	98.5 (1.8)	85.7 (0.0)	93.2 (2.4)	96.0 (2.7)	75.0 (0.0)	89.4 (2.0)
Nearest Neighbor	74.0 (21.2)	7.5 (22.6)	61.3 (2.7)	77.4 (8.5)	5.0 (15.0)	73.7 (5.7)	84.7 (10.4)	0.0 (0.0)	72.2 (0.6)	86.4 (6.6)	0.0 (0.0)	75.1 (0.4)
MST	86.7 (1.7)	75.0 (0.0)	86.0 (1.0)	87.1 (0.8)	75.0 (0.0)	90.6 (3.7)	92.5 (0.6)	82.9 (8.6)	92.6 (4.4)	87.2 (0.5)	75.0 (0.0)	92.4 (0.2)
Density-Based Method												
Gaussian OC	97.5 (3.8)	75.0 (0.0)	84.1 (5.9)	98.8 (1.9)	75.0 (0.0)	90.6 (3.7)	98.1 (2.0)	85.7 (0.0)	93.5 (3.7)	95.4 (2.6)	75.0 (0.0)	91.5 (2.7)
MoG	97.2 (4.1)	75.0 (0.0)	82.3 (7.8)	97.7 (3.5)	75.0 (0.0)	90.6 (3.7)	98.6 (1.8)	85.7 (0.0)	92.3 (5.0)	96.3 (2.6)	75.0 (0.0)	90.9 (3.0)
MCD Gaussian	98.3 (3.3)	75.0 (0.0)	84.1 (5.9)	98.8 (1.9)	75.0 (0.0)	91.5 (2.8)	98.4 (1.7)	85.7 (0.0)	93.5 (3.7)	96.7 (2.1)	75.0 (0.0)	91.5 (2.7)
Parzen	97.5 (3.8)	75.0 (0.0)	84.1 (5.9)	98.3 (2.0)	75.0 (0.0)	90.6 (3.7)	98.6 (1.8)	85.7 (0.0)	92.3 (5.0)	96.5 (2.5)	75.0 (0.0)	91.0 (2.9)
Naive Parzen	97.5 (3.8)	75.0 (0.0)	84.1 (5.9)	98.3 (2.0)	75.0 (0.0)	90.6 (3.7)	98.6 (1.8)	85.7 (0.0)	92.3 (5.0)	96.5 (2.5)	75.0 (0.0)	91.0 (2.9)
k-NN	98.3 (3.3)	75.0 (0.0)	84.1 (5.9)	98.8 (1.9)	75.0 (0.0)	91.5 (2.8)	98.6 (1.8)	85.7 (0.0)	93.5 (3.7)	96.5 (2.5)	75.0 (0.0)	91.5 (2.7)
LOF	98.3 (3.3)	75.0 (0.0)	84.1 (5.9)	97.5 (3.8)	75.0 (0.0)	91.5 (2.8)	98.6 (1.8)	85.7 (0.0)	93.5 (3.7)	96.3 (2.6)	75.0 (0.0)	91.9 (1.3)
Reconstruction-Based Method												
Auto – encoder	85.0 (8.9)	72.4 (10.5)	81.7 (8.7)	87.9 (9.6)	71.3 (12.0)	89.2 (6.1)	87.9 (12.5)	67.6 (24.9)	86.2 (8.7)	87.7 (12.7)	57.0 (24.4)	86.8 (6.2)
SOM	79.5 (5.5)	75.0 (0.0)	79.3 (9.5)	84.7 (8.6)	75.4 (3.0)	90.7 (3.6)	97.8 (4.3)	87.0 (4.8)	91.9 (4.8)	95.6 (5.9)	74.0 (7.0)	91.0 (3.2)
K-means	97.5 (3.8)	75.0 (0.0)	84.1 (5.9)	98.8 (1.9)	75.0 (0.0)	90.6 (3.7)	98.1 (2.0)	85.7 (0.0)	93.5 (3.7)	95.4 (2.6)	75.0 (0.0)	91.0 (2.9)

**Table 8:** Univariate Input Smoothed Data (different sample sizes) - Average and standard deviation (F1-score, AUC, specificity).

Fraction = 0.01												
Models	Boundary and Domain-Based Method											
	1 Month			2 Months			3 Months			4 Months		
	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1
SVDD	100 (0.0)	100 (0.0)	94.6 (8.4)	100 (0.0)	100 (0.0)	98.2 (3.6)	100 (0.0)	100 (0.0)	98.8 (2.4)	100 (0.0)	100 (0.0)	98.7 (2.9)
<i>v-SVM</i>	100 (0.0)	100 (0.0)	99.3 (3.1)	100 (0.0)	100 (0.0)	99.0 (2.9)	100 (0.0)	100 (0.0)	99.4 (1.9)	100 (0.0)	100 (0.0)	99.4 (1.6)
Nearest Neighbor	99.2 (2.5)	62.5 (34.1)	80.7 (12)	97.5 (5.3)	0.0 (0.0)	75.3 (0.8)	97.8 (4.5)	17.1 (26.3)	75.7 (5.9)	98.3 (2.6)	5.0 (15.0)	75.8 (2.0)
MST	100 (0.0)	100 (0.0)	100 (0.0)	100 (0.0)	100 (0.0)	99.1 (2.7)	100 (0.0)	100 (0.0)	99.4 (1.8)	100 (0.0)	100 (0.0)	99.1 (1.7)
Density-Based Method												
Gaussian OC	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	99.1 (2.7)	100 (0.0)	100 (0.0)	98.2 (2.7)	100 (0.0)	100 (0.0)	98.7 (2.9)
MoG	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	97.3 (4.2)	100 (0.0)	100 (0.0)	98.2 (2.7)	100 (0.0)	100 (0.0)	98.7 (2.9)
MCD Gaussian	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	99.1 (2.7)	100 (0.0)	100 (0.0)	99.4 (1.8)	100 (0.0)	100 (0.0)	98.7 (2.9)
Parzen	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	97.3 (4.2)	100 (0.0)	100 (0.0)	98.2 (2.7)	100 (0.0)	100 (0.0)	98.7 (2.9)
Naive Parzen	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	98.2 (3.6)	100 (0.0)	100 (0.0)	98.2 (2.7)	100 (0.0)	100 (0.0)	98.7 (2.9)
k-NN	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	99.1 (2.7)	100 (0.0)	100 (0.0)	98.8 (2.4)	100 (0.0)	100 (0.0)	99.1 (1.7)
LOF	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	99.1 (2.7)	100 (0.0)	100 (0.0)	98.8 (2.4)	100 (0.0)	100 (0.0)	98.7 (2.9)
Reconstruction-Based Method												
Auto – encoder	98.7 (5.6)	95.8 (14.6)	93.4 (11.2)	99.7 (2.3)	97.0 (12.9)	96.8 (6.2)	99.4 (3.5)	97.9 (8.0)	97.8 (3.9)	98.7 (4.9)	92.0 (17.9)	96.6 (5.4)
SOM	100 (0.0)	100 (0.0)	90.6 (16.8)	100 (0.0)	100 (0.0)	96.9 (4.9)	100 (0.0)	100 (0.0)	99.4 (1.8)	100 (0.0)	100 (0.0)	99.1 (1.7)
K-means	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	99.1 (2.7)	100 (0.0)	100 (0.0)	98.2 (2.7)	100 (0.0)	100 (0.0)	98.7 (2.9)

## 11.1.2. Hourly

### 11.1.2.1. The 1<sup>st</sup> Infection Episode (Flu)

**Table 9:** Univariate Input Smoothed Data (different sample sizes) - Average and standard deviation (F1-score, AUC, specificity).

Fraction = 0.01												
Models	Boundary and Domain-Based Method											
	1 Month			2 Months			3 Months			4 Months		
	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1
SVDD	97.8 (3.4)	78.7 (4.9)	83.0 (2.0)	98.2 (1.5)	81.7 (4.0)	92.0 (0.9)	98.1 (2.4)	85.4 (3.5)	94.5 (2.9)	99.0 (1.3)	91.3 (1.4)	97.1 (2.3)
<i>v-SVM</i>	99.3 (1.2)	86.2 (2.0)	91.4 (1.7)	99.7 (0.5)	93.1 (0.7)	96.3 (1.7)	99.8 (0.5)	94.2 (0.0)	97.1 (1.6)	99.8 (0.3)	94.2 (0.0)	97.1 (1.5)
Nearest Neighbor	83.8 (5.3)	51.1 (13.3)	69.7 (5.6)	85.7 (4.5)	54.6 (12.4)	82.8 (3.6)	86.2 (3.1)	54.8 (12.4)	88.1 (2.1)	86.6 (3.4)	56.8 (12.0)	91.0 (1.3)
MST	87.5 (2.6)	70.9 (7.5)	78.8 (2.5)	87.8 (2.7)	71.1 (7.5)	88.3 (1.5)	88.0 (2.7)	71.8 (7.2)	92.2 (0.9)	88.0 (2.7)	74.0 (6.5)	94.4 (0.7)
Density-Based Method												
Gaussian OC	99.3 (1.2)	83.8 (3.5)	86.8 (1.6)	99.7 (0.5)	92.1 (1.2)	96.0 (2.0)	99.8 (0.5)	94.8 (1.7)	97.2 (3.1)	99.8 (0.3)	94.4 (0.5)	98.2 (1.1)
MoG	98.8 (1.8)	81.6 (3.8)	85.0 (1.7)	99.6 (0.6)	90.2 (1.8)	95.0 (1.6)	99.8 (0.5)	94.7 (1.9)	97.3 (3.1)	99.7 (0.4)	94.5 (0.5)	97.9 (1.2)
MCD Gaussian	99.3 (1.2)	83.8 (3.5)	86.8 (1.6)	99.7 (0.5)	92.1 (1.2)	96.0 (2.0)	99.8 (0.5)	94.8 (1.7)	97.2 (3.1)	99.8 (0.3)	94.4 (0.5)	98.2 (1.1)
Parzen	98.6 (2.3)	80.2 (4.4)	82.2 (5.8)	99.4 (1.0)	89.7 (1.5)	94.5 (2.6)	99.5 (0.8)	95.9 (1.6)	96.4 (4.2)	99.6 (0.8)	94.8 (0.8)	97.0 (3.7)
Naive Parzen	98.6 (2.3)	80.2 (4.4)	82.2 (5.8)	99.4 (1.0)	89.7 (1.5)	94.5 (2.6)	99.5 (0.8)	95.9 (1.6)	96.4 (4.2)	99.6 (0.8)	94.8 (0.8)	97.0 (3.7)
k-NN	99.3 (1.2)	83.8 (3.5)	86.8 (1.6)	99.7 (0.5)	92.1 (1.2)	96.0 (2.0)	99.8 (0.5)	94.8 (1.7)	97.2 (3.1)	99.8 (0.3)	94.4 (0.5)	98.2 (1.1)
LOF	99.3 (1.1)	83.8 (3.5)	86.8 (1.6)	99.7 (0.5)	92.1 (1.2)	96.0 (2.0)	99.8 (0.5)	94.8 (1.7)	97.2 (3.1)	99.7 (0.3)	94.4 (0.5)	98.2 (1.1)
Reconstruction-Based Method												
Auto – encoder	88.0 (5.9)	71.3 (9.5)	78.4 (5.1)	91.0 (4.2)	74.6 (9.7)	89.0 (3.1)	91.0 (5.3)	76.3 (10.5)	92.2 (4.3)	92.0 (5.3)	77.3 (10.1)	93.6 (4.2)
SOM	92.3 (2.3)	68.5 (9.9)	76.8 (2.8)	96.1 (3.1)	84.2 (8.4)	92.1 (3.6)	97.9 (0.8)	90.7 (2.9)	96.4 (1.8)	98.1 (0.7)	92.7 (1.5)	97.0 (2.6)
K-means	99.3 (1.2)	83.8 (3.5)	86.8 (1.6)	99.7 (0.5)	92.1 (1.2)	96.0 (2.0)	99.8 (0.5)	94.8 (1.7)	97.2 (3.1)	99.8 (0.3)	94.4 (0.5)	98.2 (1.1)

### 11.1.2.2. The 2<sup>nd</sup> Infection Episode (Flu)

**Table 10:** Univariate Input Smoothed Data (different sample sizes) - Average and standard deviation (F1-score, AUC, specificity).

Fraction = 0.01												
Models	Boundary and Domain-Based Method											
	1 Month			2 Months			3 Months			4 Months		
	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1
SVDD	100 (0.0)	100 (0.0)	96.0 (10.4)	100 (0.0)	100 (0.0)	98.8 (2.6)	99.5 (0.6)	86.0 (4.3)	95.3 (0.7)	99.6 (0.6)	89.1 (6.1)	89.1 (0.9)
<i>v-SVM</i>	100 (0.0)	100 (0.0)	99.6 (1.1)	100 (0.0)	100 (0.0)	99.5 (1.3)	99.8 (0.4)	91.9 (1.0)	96.2 (1.5)	99.9 (0.3)	95.3 (0.5)	97.5 (1.6)
Nearest Neighbor	99.9 (0.1)	99.5 (1.5)	99.1 (1.1)	100 (0.0)	100 (0.0)	99.2 (0.5)	84.6 (1.4)	66.8 (8.5)	91.0 (1.9)	84.7 (1.2)	68.0 (7.8)	93.0 (1.4)
MST	100 (0.0)	100 (0.0)	96.9 (7.9)	100 (0.0)	100 (0.0)	99.4 (1.2)	86.8 (1.6)	71.6 (4.0)	92.3 (0.7)	86.7 (1.8)	71.9 (3.9)	93.9 (1.3)
Density-Based Method												
Gaussian OC	100 (0.0)	100 (0.0)	96.3 (11.2)	100 (0.0)	100 (0.0)	98.2 (4.9)	99.8 (0.4)	92.1 (2.7)	97.2 (0.9)	99.9 (0.3)	96.0 (1.8)	98.3 (1.6)
MoG	100 (0.0)	100 (0.0)	96.0 (10.7)	100 (0.0)	100 (0.0)	98.0 (4.8)	99.5 (0.6)	90.3 (3.9)	96.2 (1.4)	99.8 (0.4)	93.3 (3.6)	97.3 (1.7)
MCD Gaussian	100 (0.0)	100 (0.0)	96.3 (11.2)	100 (0.0)	100 (0.0)	98.2 (5.1)	99.8 (0.4)	92.1 (2.7)	97.2 (0.9)	99.9 (0.3)	96.0 (1.8)	98.3 (1.6)
Parzen	100 (0.0)	100 (0.0)	96.1 (10.7)	100 (0.0)	100 (0.0)	97.9 (5.2)	99.3 (0.7)	88.5 (2.0)	95.9 (1.5)	99.4 (0.8)	90.2 (2.8)	96.5 (2.0)
Naive Parzen	100 (0.0)	100 (0.0)	96.1 (10.7)	100 (0.0)	100 (0.0)	97.9 (5.2)	99.3 (0.7)	88.5 (2.0)	95.9 (1.5)	99.4 (0.8)	90.2 (2.8)	96.5 (2.0)
k-NN	100 (0.0)	100 (0.0)	96.2 (11.1)	100 (0.0)	100 (0.0)	97.9 (5.2)	99.2 (0.7)	88.3 (2.5)	95.8 (1.3)	99.6 (0.5)	89.2 (3.7)	96.5 (1.4)
LOF	100 (0.0)	100 (0.0)	96.6 (9.6)	100 (0.0)	100 (0.0)	98.1 (5.1)	91.6 (0.9)	86.6 (1.0)	96.0 (0.8)	91.9 (2.3)	87.1 (1.7)	96.6 (1.5)
Reconstruction-Based Method												
Auto – encoder	98.2 (4.1)	97.2 (5.1)	94.6 (9.3)	98.6 (2.7)	96.3 (6.6)	96.9 (4.6)	91.9 (6.7)	76.1 (11.1)	92.8 (2.7)	94.0 (5.3)	79.7 (10.8)	94.7 (2.7)
SOM	100 (0.0)	100 (0.0)	96.4 (10.0)	100 (0.0)	100 (0.0)	98.0 (4.1)	95.0 (4.5)	85.3 (5.4)	94.9 (1.8)	98.0 (2.1)	88.2 (3.1)	96.5 (1.9)
K-means	100 (0.0)	100 (0.0)	96.3 (11.2)	100 (0.0)	100 (0.0)	98.2 (4.9)	99.8 (0.4)	92.1 (2.7)	97.2 (0.9)	99.9 (0.3)	96.0 (1.8)	98.3 (1.6)

### 11.1.2.3. The 3<sup>rd</sup> Infection Episode (Flu)

**Table 11:** Univariate Input Smoothed Data (different sample sizes) - Average and standard deviation (F1-score, AUC, specificity).

Fraction = 0.01												
Models	Boundary and Domain-Based Method											
	1 Month			2 Months			3 Months			4 Months		
	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1
SVDD	97.2 (2.6)	85.9 (2.9)	87.3 (15.0)	97.0 (1.7)	86.0 (3.6)	93.6 (6.9)	97.7 (2.6)	84.9 (1.8)	95.6 (4.1)	98.0 (1.4)	86.5 (4.3)	95.6 (4.1)
<i>v-SVM</i>	97.7 (2.7)	85.9 (0.7)	92.6 (1.5)	99.2 (1.5)	88.9 (0.0)	95.0 (2.1)	98.6 (1.3)	85.3 (0.4)	94.0 (1.7)	98.9 (0.9)	86.8 (0.8)	94.6 (1.9)
Nearest Neighbor	87.8 (3.3)	55.8 (9.7)	78.7 (5.0)	90.2 (1.5)	67.8 (8.8)	88.5 (5.7)	89.1 (1.9)	64.1 (3.5)	92.3 (2.6)	89.0 (2.6)	67.3 (3.2)	94.5 (2.0)
MST	90.2 (0.3)	81.1 (3.0)	86.4 (10.6)	90.5 (1.2)	81.6 (4.1)	92.5 (6.6)	89.9 (1.5)	80.4 (0.4)	95.3 (3.5)	90.1 (0.8)	80.4 (0.4)	96.7 (1.9)
Density-Based Method												
Gaussian OC	97.5 (2.7)	85.8 (1.9)	87.1 (15.6)	99.2 (1.5)	89.8 (2.6)	93.4 (10.8)	98.5 (1.3)	85.7 (1.1)	96.7 (1.1)	98.9 (0.9)	87.2 (1.1)	97.5 (1.1)
MoG	97.2 (2.6)	85.1 (3.0)	86.6 (15.8)	99.0 (1.5)	89.6 (2.2)	93.5 (10.4)	98.2 (2.2)	82.2 (2.8)	95.2 (3.2)	98.7 (1.2)	85.7 (2.0)	96.8 (1.7)
MCD Gaussian	97.5 (2.7)	85.8 (1.9)	87.6 (14.2)	99.3 (1.4)	89.8 (2.6)	93.9 (9.3)	98.7 (1.3)	85.7 (1.1)	96.7 (1.1)	99.1 (0.9)	87.2 (1.1)	97.5 (1.1)
Parzen	95.9 (2.7)	84.9 (3.1)	86.1 (15.0)	98.8 (1.5)	88.1 (3.2)	92.6 (11.3)	97.6 (2.4)	82.3 (1.5)	94.1 (4.3)	98.3 (1.4)	83.5 (1.5)	95.6 (2.7)
Naive Parzen	95.9 (2.7)	84.9 (3.1)	86.1 (15.0)	98.8 (1.5)	88.1 (3.2)	92.6 (11.3)	97.6 (2.4)	82.3 (1.5)	94.1 (4.3)	98.3 (1.4)	83.5 (1.5)	95.6 (2.7)
k-NN	96.1 (2.4)	84.2 (3.9)	86.2 (15.1)	98.8 (1.4)	88.5 (2.6)	92.8 (11.4)	97.5 (2.6)	81.0 (1.6)	94.1 (4.2)	98.3 (1.4)	82.5 (1.5)	95.8 (2.5)
LOF	92.6 (1.7)	78.8 (5.2)	84.5 (12.4)	93.7 (1.3)	77.2 (6.4)	91.2 (6.5)	90.7 (1.5)	81.6 (0.9)	94.5 (4.2)	90.8 (0.8)	81.5 (0.6)	96.3 (2.1)
Reconstruction-Based Method												
Auto – encoder	90.1 (5.4)	79.1 (6.7)	84.9 (11.6)	93.5 (4.1)	81.0 (6.9)	92.1 (7.5)	93.2 (3.4)	80.2 (5.7)	94.7 (3.9)	93.8 (3.2)	80.9 (6.3)	95.9 (2.5)
SOM	95.2 (1.6)	85.6 (2.8)	87.0 (14.5)	96.5 (0.8)	87.5 (3.2)	93.3 (9.0)	93.1 (2.2)	81.7 (2.2)	94.1 (4.2)	94.3 (1.6)	82.1 (2.1)	96.2 (2.1)
K-means	91.9 (1.8)	79.2 (5.3)	83.8 (14.7)	95.1 (1.6)	86.2 (5.3)	92.8 (9.6)	93.8 (2.5)	79.0 (3.9)	94.4 (4.2)	95.7 (0.9)	84.1 (2.3)	96.7 (2.1)

### 11.1.2.4. The 4th Infection Episode (Flu)

**Table 12:** Univariate Input Smoothed Data (different sample sizes) - Average and standard deviation (F1-score, AUC, specificity).

Fraction = 0.01												
Models	Boundary and Domain-Based Method											
	1 Month			2 Months			3 Months			4 Months		
	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1
SVDD	99.0 (1.9)	93.1 (3.2)	89.1 (10.2)	97.9 (1.5)	86.6 (3.0)	94.7 (1.9)	98.5 (1.1)	87.0 (2.6)	96.8 (0.7)	99.1 (0.8)	93.4 (4.2)	94.6 (0.3)
<i>v-SVM</i>	99.7 (0.7)	93.4 (0.0)	96.0 (1.3)	99.6 (0.6)	91.0 (0.5)	95.7 (1.8)	99.8 (0.4)	92.7 (0.5)	96.6 (1.8)	99.8 (0.3)	93.4 (0.0)	96.8 (2.1)
Nearest Neighbor	94.5 (1.7)	81.2 (1.9)	88.1 (3.0)	89.8 (1.4)	72.3 (4.2)	91.2 (0.9)	90.8 (1.4)	76.5 (2.0)	94.7 (0.4)	91.2 (1.5)	77.5 (2.1)	96.1 (0.3)
MST	95.5 (1.0)	90.9 (2.3)	89.7 (9.0)	91.5 (1.1)	83.2 (2.3)	94.4 (0.8)	91.6 (1.1)	83.2 (2.3)	96.2 (0.5)	91.6 (1.2)	83.2 (2.3)	97.1 (0.4)
Density-Based Method												
Gaussian OC	99.6 (0.8)	94.1 (2.0)	93.3 (8.1)	99.6 (0.6)	90.4 (1.0)	96.4 (1.8)	99.7 (0.4)	92.7 (0.9)	97.6 (1.6)	99.8 (0.3)	93.8 (0.9)	98.2 (1.5)
MoG	99.1 (1.6)	93.4 (2.3)	90.1 (10.3)	99.4 (0.7)	90.3 (1.1)	96.0 (2.1)	99.7 (0.4)	92.7 (0.9)	97.6 (1.6)	99.8 (0.3)	93.9 (1.0)	98.2 (1.5)
MCD Gaussian	99.7 (0.7)	94.1 (2.0)	93.3 (8.1)	99.6 (0.6)	90.4 (1.0)	96.4 (1.8)	99.7 (0.4)	92.7 (0.9)	97.6 (1.6)	99.7 (0.3)	93.7 (0.7)	98.2 (1.4)
Parzen	99.1 (1.5)	94.9 (1.7)	90.9 (10.6)	99.6 (0.7)	91.0 (1.8)	95.9 (1.9)	99.7 (0.5)	92.6 (2.2)	97.3 (1.7)	99.8 (0.3)	95.2 (1.3)	98.2 (1.7)
Naive Parzen	99.1 (1.5)	94.9 (1.7)	90.9 (10.6)	99.6 (0.7)	91.0 (1.8)	95.9 (1.9)	99.7 (0.5)	92.6 (2.2)	97.3 (1.7)	99.8 (0.3)	95.2 (1.3)	98.2 (1.7)
k-NN	99.1 (1.7)	93.0 (2.4)	89.9 (10.8)	99.6 (0.6)	90.4 (1.8)	96.4 (1.5)	99.8 (0.4)	92.0 (1.6)	97.4 (1.5)	99.8 (0.3)	94.7 (1.3)	98.2 (1.6)
LOF	96.1 (1.5)	92.3 (2.4)	89.7 (10.7)	97.0 (1.0)	89.9 (1.5)	94.8 (3.9)	97.4 (0.6)	91.2 (1.3)	96.7 (2.7)	97.1 (1.5)	91.5 (0.7)	97.6 (1.9)
Reconstruction-Based Method												
Auto – encoder	95.1 (5.5)	89.5 (5.7)	88.9 (9.5)	94.7 (4.0)	80.6 (11.2)	92.9 (3.6)	95.7 (3.8)	82.8 (10.6)	95.4 (2.4)	96.3 (3.9)	85.1 (10.9)	96.7 (2.1)
SOM	97.3 (1.2)	92.9 (2.4)	89.7 (10.5)	96.7 (1.4)	88.5 (2.4)	94.6 (3.1)	97.5 (0.6)	90.8 (1.5)	96.7 (2.4)	97.4 (0.5)	91.1 (1.2)	97.5 (1.8)
K-means	96.5 (1.6)	93.2 (2.3)	90.3 (9.5)	96.2 (1.0)	89.5 (2.1)	95.8 (2.2)	97.1 (1.2)	91.4 (2.5)	97.2 (1.6)	98.2 (1.0)	93.2 (2.0)	98.0 (1.4)

## 11.2. Unsupervised Method

**Table 13:** Univariate Input Raw and Smoothed Data for both the daily and hourly data granularity - Average and standard deviation (F1-score, AUC, specificity). The parameters  $k_d$  and  $k_h$  represent the optimal number of nearest neighbors for the daily and hourly cases respectively.

Freq.		Density-Based Methods												
	Pre-pro.	Methods (Threshold)	1 <sup>st</sup> Infection Episode ( $k_d=30, k_h=240$ )			2 <sup>nd</sup> Infection Episode ( $k_d=30, k_h=240$ )			3 <sup>rd</sup> Infection Episode ( $k_d=30, k_h=240$ )			4 <sup>th</sup> Infection Episode ( $k_d=30, k_h=240$ )		
			AUC	Sensitivity	F1	AUC	Sensitivity	F1	AUC	Sensitivity	F1	AUC	Sensitivity	F1
Daily	Without filter	LOF ( $T_1=2.7, T_2=1.5, T_3=2.95, T_4=2.2$ )	90.7	83.3	79.5	89.4	83.3	73.1	90.1	66.7	90.9	86.7	75.0	75.3
		COF ( $T_1=1.4, T_2=1.1, T_3=2.3, T_4=1.8$ )	89.4	83.3	73.1	85.0	100	65.0	83.1	66.7	90.9	86.7	75.0	75.3
	With filter	LOF ( $T_1=1.9, T_2=1.9, T_3=2.8, T_4=2.8$ )	99.9	100	96.2	100	100	100	100	100	100	99.6	100	87.8
		COF ( $T_1=1.6, T_2=1.5, T_3=2.8, T_4=3.1$ )	99.9	100	96.2	100	100	100	100	100	100	99.9	100	94.7
Hourly		LOF ( $T_1=1.9, T_2=1.6, T_3=1.2, T_4=1.7$ )	95.0	90.1	94.8	94.2	89.7	81.1	93.9	91.4	73.0	86.8	75.3	73.4
		COF ( $T_1=1.6, T_2=1.3, T_3=1.2, T_4=1.3$ )	94.9	90.1	94.2	87.7	76.3	80.0	96.5	96.3	74.8	86.8	75.3	73.2

## **12 Included Papers**

# Paper 1

**Woldaregay, A. Z.,** Launonen, I. K., Årsand, E., Albers, D., Holubová, A., & Hartvigsen, G. (2020). *Toward Detecting Infection Incidence in People With Type 1 Diabetes Using Self-Recorded Data (Part 1): A Novel Framework for a Personalized Digital Infectious Disease Detection System.* *J Med Internet Res*, 22(8), e18911. doi:10.2196/18911



Original Paper

# Toward Detecting Infection Incidence in People With Type 1 Diabetes Using Self-Recorded Data (Part 1): A Novel Framework for a Personalized Digital Infectious Disease Detection System

Ashenafi Zebene Woldaregay<sup>1</sup>, MSc; Ilkka Kalervo Launonen<sup>2</sup>, PhD; Eirik Årsand<sup>1</sup>, PhD; David Albers<sup>3,4</sup>, PhD; Anna Holubová<sup>5,6</sup>, MSc; Gunnar Hartvigsen<sup>1</sup>, PhD

<sup>1</sup>Department of Computer Science, University of Tromsø – The Arctic University of Norway, Tromsø, Norway

<sup>2</sup>Department of Clinical Research, University Hospital of North Norway, Tromsø, Norway

<sup>3</sup>Department of Pediatrics, Informatics and Data Science, University of Colorado, Aurora, CO, United States

<sup>4</sup>Department of Biomedical Informatics, Columbia University, New York, NY, United States

<sup>5</sup>Department of ICT in Medicine, Faculty of Biomedical Engineering, Czech Technical University, Prague, Czech Republic

<sup>6</sup>Spin-off Company and Research Results Commercialization Center of the First Faculty of Medicine, Charles University, Prague, Czech Republic

**Corresponding Author:**

Ashenafi Zebene Woldaregay, MSc

Department of Computer Science

University of Tromsø – The Arctic University of Norway

Hansine Hansens veg 54, Science building Realfagbygget, office A124

Tromsø

Norway

Phone: 47 46359333

Email: [ashenafi.z.woldaregay@uit.no](mailto:ashenafi.z.woldaregay@uit.no)

## Abstract

**Background:** Type 1 diabetes is a chronic condition of blood glucose metabolic disorder caused by a lack of insulin secretion from pancreas cells. In people with type 1 diabetes, hyperglycemia often occurs upon infection incidences. Despite the fact that patients increasingly gather data about themselves, there are no solid findings that uncover the effect of infection incidences on key parameters of blood glucose dynamics to support the effort toward developing a digital infectious disease detection system.

**Objective:** The study aims to retrospectively analyze the effect of infection incidence and pinpoint optimal parameters that can effectively be used as input variables for developing an infection detection algorithm and to provide a general framework regarding how a digital infectious disease detection system can be designed and developed using self-recorded data from people with type 1 diabetes as a secondary source of information.

**Methods:** We retrospectively analyzed high precision self-recorded data of 10 patient-years captured within the longitudinal records of three people with type 1 diabetes. Obtaining such a rich and large data set from a large number of participants is extremely expensive and difficult to acquire, if not impossible. The data set incorporates blood glucose, insulin, carbohydrate, and self-reported events of infections. We investigated the temporal evolution and probability distribution of the key blood glucose parameters within a specified timeframe (weekly, daily, and hourly).

**Results:** Our analysis demonstrated that upon infection incidence, there is a dramatic shift in the operating point of the individual blood glucose dynamics in all the timeframes (weekly, daily, and hourly), which clearly violates the usual norm of blood glucose dynamics. During regular or normal situations, higher insulin and reduced carbohydrate intake usually results in lower blood glucose levels. However, in all infection cases as opposed to the regular or normal days, blood glucose levels were elevated for a prolonged period despite higher insulin and reduced carbohydrates intake. For instance, compared with the preinfection and postinfection weeks, on average, blood glucose levels were elevated by 6.1% and 16%, insulin (bolus) was increased by 42% and 39.3%, and carbohydrate consumption was reduced by 19% and 28.1%, respectively.

**Conclusions:** We presented the effect of infection incidence on key parameters of blood glucose dynamics along with the necessary framework to exploit the information for realizing a digital infectious disease detection system. The results demonstrated that compared with regular or normal days, infection incidence substantially alters the norm of blood glucose dynamics, which are quite significant changes that could possibly be detected through personalized modeling, for example, prediction models and

anomaly detection algorithms. Generally, we foresee that these findings can benefit the efforts toward building next generation digital infectious disease detection systems and provoke further thoughts in this challenging field.

(*J Med Internet Res* 2020;22(8):e18911) doi: [10.2196/18911](https://doi.org/10.2196/18911)

## KEYWORDS

type 1 diabetes; self-recorded health data; infection incidence; decision making; infectious disease outbreaks; public health surveillance

## Introduction

The incidence of infectious disease outbreaks can create panic in society and is a threat to local and global health security. Such outbreaks require immediate detection and appropriate response during the initial phase of the incidence to reduce fatality and save lives [1]. The timeliness of outbreak detection defines the success of the appropriate response by the concerned bodies. The state-of-the-art syndromic surveillance systems have been improved compared with the traditional surveillance system, which is generally passive and dependent on laboratory confirmation [2]. Syndromic surveillance makes use of features that come before diagnosis, including different activities triggered by the onset of symptoms, such as Google search, Twitter, school and work absenteeism, pharmacy drug sells, and other sources as a signal of change in individual and population health [2]. These signals are mainly acquired from the secondary source of information, typically built for other purposes. However, to keep up the pace with the rapidly changing social and biological dynamics, novel outbreak detection mechanisms are highly sought [2].

The advancement and omnipresence of smartphones, Internet of Things (IoT) devices, wearables, and sensors have enabled individuals to easily self-record health-related events often for self-tracking or self-managing their disease [3,4]. The recent movement known as quantified self and lifelogging is the result of such technological advancement, where people collect various kinds of health-related events and data for personal informatics purposes, that is, self-surveillance and self-management [5-8]. To this end, people with diabetes are not an exception, where they self-record detailed information as part of their self-management, including blood glucose levels, diet and insulin intake, physical activity, medication, and other information [4,9,10]. Consequently, a huge amount of self-recorded, personal health-related data is generated each day that have great potential to be used as a secondary source of information for other purposes such as digital epidemiology [11,12]. According to recent reports, personal health data or self-collected health-related data have provided an enormous opportunity to enhance the possibility of detecting infection incidence during the presymptomatic stage (improved sensitivity and timeliness), specifically during the incubation period, where most of the existing systems neglect from their process [13].

Type 1 diabetes is a chronic condition of blood glucose metabolic disorder caused by lack of insulin secretion from pancreas cells [14]. These patient groups are recommended to maintain their blood glucose levels within a specified range through self-management practice [14,15]. Blood glucose levels are controlled by balancing insulin and meal intake along with

other contexts such as physical activity, medications, and others. Blood glucose dynamics are affected by various factors that can be categorized as common, individual, and unpredictable factors [16]. These factors could be further categorized as patient-controllable and patient-uncontrollable parameters [17]. Patient-controllable parameters incorporate factors on which the patient has direct control and can roughly understand their immediate effect on blood glucose dynamics. However, patient-uncontrollable parameters include factors in which the patient does not have direct control and faces a challenge to understand their immediate effect on blood glucose levels. From the patient perspective, usually patient-controllable parameters induce reasonable deviations on blood glucose levels; however, patient-uncontrollable parameters induce unreasonable blood glucose deviations and usually differ from the usual norm of blood glucose dynamics [18]. The total number of people living with diabetes is increasing worldwide. According to recent reports [14], there were 415 million people between the ages of 20 and 79 years in 2015, and this value is projected to increase by 54% in 2040. From this figure, 5% are believed to have type 1 diabetes. In these patient groups, infection incidence often results in complications and difficulties in controlling blood glucose levels within the recommended range [19-21]. As a result, early detection of infection incidence among these patient groups could provide a way to assist the individual and at the same time can be used to realize a digital infectious disease detection system.

Currently, with the advancement of technology, the need to have a system that is able to detect infection incidence at the presymptomatic stage is highly sought [13]. In this regard, there are some previous investigations that have showcased the use of self-recorded data from people with diabetes as surveillance events (indicators) by uncovering the effect of infection incidence on blood glucose levels and glycemic control in real-life settings [18,22-36]. These studies reported the presence of prolonged hyperglycemia episodes as a result of infection incidence, thereby revealing the potential of self-recorded data as a secondary source of information for realizing a digital infectious disease detection system. For instance, Botsis et al [22] conducted a proof-of-concept study based on daily glycemic control data of 248 people with type 2 diabetes and concluded that blood glucose levels, insulin dosage, diet (carbohydrate consumption), physical activity, and other physiological parameters could be used as potential event indicators of infection incidence but calls for further investigations. Furthermore, Botsis et al [18] also reported elevated glycated hemoglobin (HbA<sub>1c</sub>) levels after infections regardless of tight blood glucose control, which only settled down to normal levels after the patient recovered. Moreover, other studies conducted in hospital settings also reported similar results in this direction

[37,38]. Despite reporting the potential of using self-recorded data as a surveillance event indicator, none of these studies demonstrated the extent to which each parameter is affected at an individual level as a result of infection incidence. Therefore, the purpose of this study was to retrospectively analyze the effect of infection incidence at an individual level and pinpoint optimal parameters that can effectively be used as input variables for developing an infection detection algorithm, thereby illustrating how these patient groups can assist in detecting infectious disease outbreaks. Moreover, this study provides a general framework regarding how a digital infectious disease detection system can be designed using self-recorded data from people with type 1 diabetes as a secondary source of information. Furthermore, this sheds light on the possibility of assisting the individual during such an incident. To this end, we analyzed temporal trends and probability distributions of different diabetes profile parameters (ie, blood glucose, insulin, carbohydrate, and others) to uncover the effect of infection incidence on the blood glucose dynamics, thereby identifying parameters that can effectively be used as potential events (indicators) of infection incidence. In addition, a framework is presented depicting the necessary structure to properly exploit self-recorded data from these patient groups to realize a real-time digital infectious disease detection system. This paper is structured as follows: the Methods section describes the materials and methods used to analyze the data sets. The Results section presents the results depicting the effect of acute infection incidence in comparison with regular or normal situations. The Discussion section presents the overall findings and proposes a framework for designing and developing a real-time digital infectious disease detection system using self-recorded data

from these patient groups. The final section of *Discussion* presents our concluding remarks.

## Methods

### Materials

High precision self-recorded data of 10 patient years collected from 3 real subjects (2 males and 1 female) with type 1 diabetes were used. The patients were free from any other chronic or other form of disease, except the self-reported acute infection incidence throughout the entire data collection period. The data sets consisted of blood glucose measurements (self-monitoring of blood glucose [SMBG] and continuous glucose monitoring [CGM]), injected insulin (basal and bolus), diet (carbohydrate in grams), and self-reported events of acute infection. The patients used different diabetes self-management technologies throughout the data collection period to gather these data sets including the Diabetes Diary app (Norwegian Centre for E-health Research) [39], the Spike app [40], the xDrip with app, Dexcom CGM, insulin pens, and insulin pumps, as shown in [Table 1](#). The data sets consist of both normal years, without any significant acute infection incidence, and years with at least one or more acute infection incidence. The normal (without infection) patient years were used as a baseline to compare the effect of all patient-controllable parameters and patient-uncontrollable parameters against the self-reported incidence of acute infection. The self-reported incidences of acute infections were a case of influenza (flu) and mild and light common cold without fever. All the experiments and analyses were conducted using MATLAB version 2018a (Mathworks).

**Table 1.** Equipment used in diabetes self-management.

Patients	Self-management		
	BG <sup>a</sup>	Insulin administration	Diet
Subject 1	SMBG <sup>b</sup> —finger pricks recorded in the Diabetes Diary mobile app and Dexcom CGM <sup>c</sup>	Insulin pen (multiple bolus and one-time basal in the morning) recorded in the Diabetes Diary mobile app	Carbohydrate in grams recorded in the Diabetes Diary mobile app
Subject 2	SMBG—finger pricks recorded in the Spike mobile app and Dexcom G4 CGM	Insulin pen (multiple bolus [Humalog] and one-time basal [Toujeo] before bed) recorded in the Spike mobile app	Carbohydrate in grams recorded in the Spike mobile app
Subject 3	Enlite (Medtronic) CGM and Dexcom G4 CGM	Medtronic MinMed G640 insulin pump (basal rates profile [Fiasp] and multiple bolus [Fiasp])	Carbohydrate in grams recorded in pump information

<sup>a</sup>BG: blood glucose.

<sup>b</sup>SMBG: self-monitoring of blood glucose.

<sup>c</sup>CGM: continuous glucose monitoring.

### Patient Characteristics

The participants were highly motivated individuals with type 1 diabetes who had advanced knowledge and understanding of several diabetes-related technologies. Hence, the self-recorded

data can be regarded as highly precise and accurate. All the participants had advanced knowledge of carbohydrate counting, which can be considered as level 3 (advanced) [41]. The long-term average HbA<sub>1c</sub> and characteristics of the participants are given in [Table 2](#).

**Table 2.** Participants characteristics.

Variables	Values
<b>Gender, n</b>	
Male	2
Female	1
Age (years), mean (SD)	34 (13.2)
<b>Body weight (kg)</b>	
Subject 1	83
Subject 2	77
Subject 3	70
<b>HbA<sub>1c</sub><sup>a</sup>(%)</b>	
Subject 1	6.0
Subject 2	7.3
Subject 3	6.2
Carbohydrate counting	Level 3 (advanced)

<sup>a</sup>HbA<sub>1c</sub>: glycated hemoglobin.

### Data Collection and Ethics

The study protocol has been submitted to the Norwegian Regional Committees for Medical Health Research Ethics Northern Norway (REK) for evaluation and was found exempted from regional ethics review because it resides outside of the scope of medical research (reference number: 108435). Written consent was obtained and the participants donated the data sets. All data from the participants were anonymized.

### Approaches

We retrospectively assessed and analyzed the diabetes profile (blood glucose, insulin, carbohydrate, and insulin-to-carbohydrate ratio) to uncover the nature, size, and shape of the infection-induced shift in the operating region of the blood glucose dynamics. A data size of 10 patient years incorporating blood glucose levels (SMBG and CGM), insulin (bolus and basal), diet (carbohydrate in grams), and self-reported events of acute infection was used. The analysis was performed based on specified timeframes (weekly, daily, and hourly) to reveal the effect of acute infection development on blood glucose dynamics. The data set incorporates 5 normal patient years without any infection incidence and 5 patient years each with at least one case of self-reported incidence of acute infection. Normal patient years were used as a baseline for comparison purposes. We analyzed the temporal evolution and probability distribution of blood glucose levels, injected insulin, carbohydrate intake (grams), and insulin-to-carbohydrate ratio within the stated timeframe. For the daily and hourly timeframes, a moving-average filter and nonparametric density estimation techniques, the kernel density estimator, were used to analyze the trend and data distribution before, during, and after the infection incidence. A moving-average filter with a window size of 2 days was employed to remove fast timescale features through smoothing. The window size includes  $N-1$  observations from the previous data points and the current data point, where  $N$  is the window size. Generally, the window size of a

moving-average filter is determined based on complementary issues of better smoothing and the cost of significant delay (shift) incurred [42,43]. A small window size often generates less delay (shift) but at the cost of more short-term features and having a larger window size will smoothen the data in a better manner but at the cost of significant delay in the timeliness of detecting the infection incidence. Therefore, the window size was determined based on these complementary issues, and more importance was given to minimize the inherent delay (shift) incurred due to the window size. To this end, window sizes of 1, 2, 3, and 4 days were applied and tested to choose the optimal size of the window, and as a result, a window size of 2 days was found to be satisfactory. The preinfection, infection, and postinfection week analyses were carried out on the raw data set based on the week's daily average and SD of blood glucose levels and daily sum and SD of insulin and carbohydrate. A statistical boxplot was used to depict the comparison during preinfection, infection, and postinfection weeks.

### Data Resampling, Imputation, and Preprocessing

The features of the self-collected data from individuals with type 1 diabetes are shown in Table 3. The raw data were resampled at a uniform rate by assigning each measurement into the nearest time-bin based on its time stamp. Generally, whenever there is more than one measurement within each time-bin, the measurements are combined into a single measurement by either summing or averaging the elements. For blood glucose levels (both CGM and SMBG), the measurements were averaged into their respective sampling time-bins. However, regarding carbohydrate consumption and insulin injections, the sum of the elements in their respective sampling time-bin was computed, as shown in Table 4. In each time-bin, the effect of total insulin and total carbohydrate on the average blood glucose level was considered. The resampled data were further preprocessed using a moving-average filter with a 2-day (48-hour) window size to capture only the important patterns—long-term variation, while filtering and smoothing

local and short-term variations. Moreover, for narrower time-bin resampling, for example, an hour, there are more frequent zeros of measurement, especially for carbohydrate and insulin measurements, which poses a significant challenge to compute the insulin-to-carbohydrate ratio as the ratio goes to infinity given that the carbohydrate amount is zero. Therefore, in such

cases of a narrower time-bin, the ratio was computed only after computing the moving-average value of insulin and carbohydrate based on a window size of 48 hours. Regarding the missing blood glucose values during the hourly computations, a cubic spline interpolation was used to estimate the missing values.

**Table 3.** Self-collected user data.

Variable names	Subject's record variables	
	Description	Units
$U^c_g$	Continuous glucose reading	mg/dL
$U^g$	Self-management blood glucose reading	mg/dL
$U^I_b$	Injected insulin (bolus)	Units
$U^I_s$	Injected insulin (basal)	Units
$U^C$	Ingested carbohydrate	Grams

**Table 4.** Data preprocessing.

Variable name	Preprocessed variables	
	Description	Units
$U^{cg}_{[t-\Delta t, t]}$	Average continuous glucose reading	mg/dL
$U^g_{[t-\Delta t, t]}$	Average self-management blood glucose reading	mg/dL
$\sum_{t-\Delta t}^t U_f^I$	Sum injected insulin (bolus)	Units
$\sum_{t-\Delta t}^t U_s^I$	Sum injected insulin (basal)	Units
$\sum_{t-\Delta t}^t U^C$	Sum ingested carbohydrate	Grams
$\frac{\sum_{t-\Delta t}^t U_f^I}{\sum_{t-\Delta t}^t U^C}$	Ratio of insulin (bolus) to carbohydrate	Units/grams
$\frac{\sum_{t-\Delta t}^t U_s^I}{\sum_{t-\Delta t}^t U^C}$	Ratio of insulin (basal) to carbohydrate	Units/grams

**Kernel Density Estimation**

Nonparametric density estimation is an alternative to the parametric approach, which involves specifying a model using a number of parameters that can be estimated through the likelihood principle [44,45]. In this study, we used kernel density estimation techniques [46-48] to estimate the probability distribution of the diabetes profile key parameters to uncover the deviation incurred by the acute infection incidence. In this regard, both univariate and bivariate kernel density estimators are used to assess and analyze the insulin-to-carbohydrate ratio

(univariate) and blood glucose levels along with the insulin-to-carbohydrate ratio (bivariate), respectively. An adaptive kernel density estimator with a Gaussian kernel was used in both cases. For the univariate kernel density estimator [49], bandwidth selection is based on the suggestion from Botev et al [44], which is a data-driven and plug-in bandwidth selector that does not use normal reference rules. For the bivariate estimator, a rule-of-thumb bandwidth selection suggested by Bowman et al [50,51] was used to determine the appropriate bandwidth [52]. These computations are carried out based on the procedures given in [Textboxes 1](#) and [2](#).

**Textbox 1.** One-dimensional adaptive kernel density estimation.

Approach: one-dimensional adaptive kernel density estimation

- Given: time series data sets of the insulin-to-carbohydrate ratio  $X \in D$  and an adaptive kernel density estimator  $M$  – one – dimensional
- Remove the reported days of infection from the time series data sets  $D$  and form a new data set  $X \in Q$
- Compute the one dimensional density based on the kernel density estimator  $M$  using  $D$  and  $Q$
- Compare the distribution from  $M$

**Textbox 2.** Two-dimensional adaptive kernel density estimation.

Approach: two-dimensional adaptive kernel density estimation

- Given: time series data sets of blood glucose level and the insulin-to-carbohydrate ratio  $X, Y \in D$  and an adaptive kernel density estimator  $N$  – two - dimensional
- Remove the reported days of infection from the time series data sets  $D$  and form a new data set  $X, Y \in Q$
- Compute the two-dimensional density based on the kernel density estimator  $N$  using  $D$  and  $Q$
- Compare the distribution from  $N$

## Results

### Overview

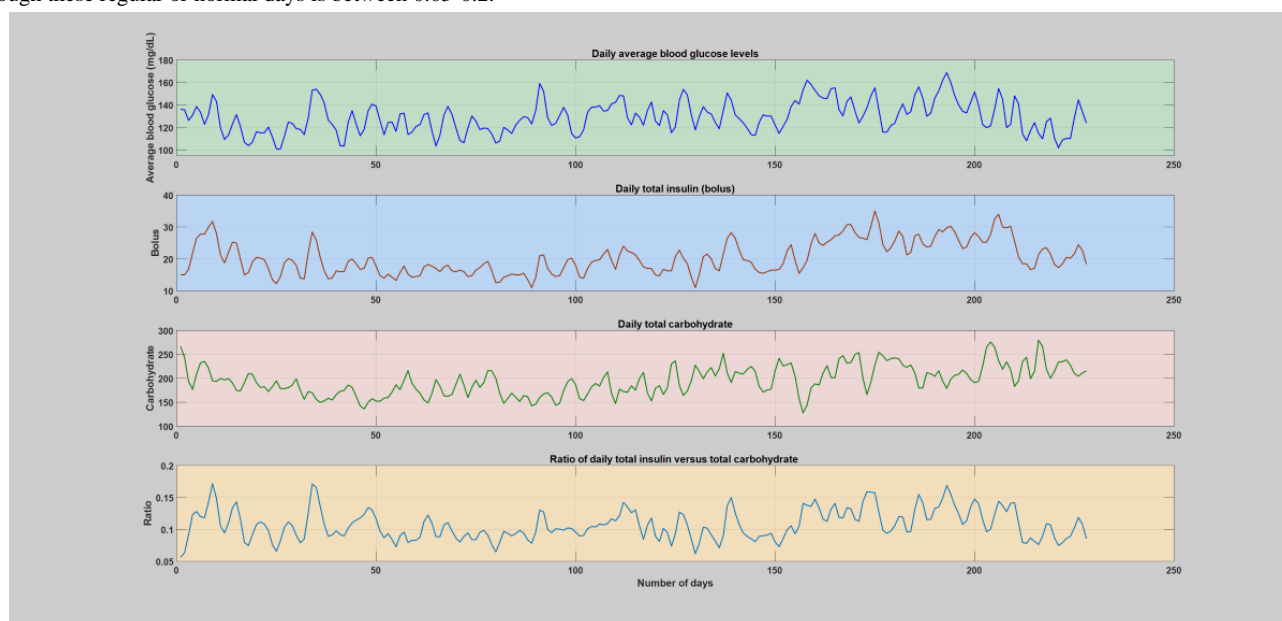
The analysis was conducted based on an hourly, daily, and weekly basis to reveal the deviations incurred due to the infection incidence. A total of 10 patient years were analyzed, and 5 of these years were found to include at least one incidence of acute infection lasting around 1-2 weeks. The proposed approach is designed to smooth out short-duration variations and include the 2 major patient-controllable factors, insulin and diet intake. Normal patient years were used to compare the effect of all patient-controllable parameters and patient-uncontrollable parameters against the self-reported incidence of acute infection. The trend analysis for both the normal patient years and patient years with acute infections using the proposed approach is presented below along with the nonparametric probability distribution. The weekly mean deviations of key diabetes parameters (blood glucose, insulin, and diet) during the preinfection, infection, and postinfection weeks are given in [Multimedia Appendix 1](#).

### Trend Analysis

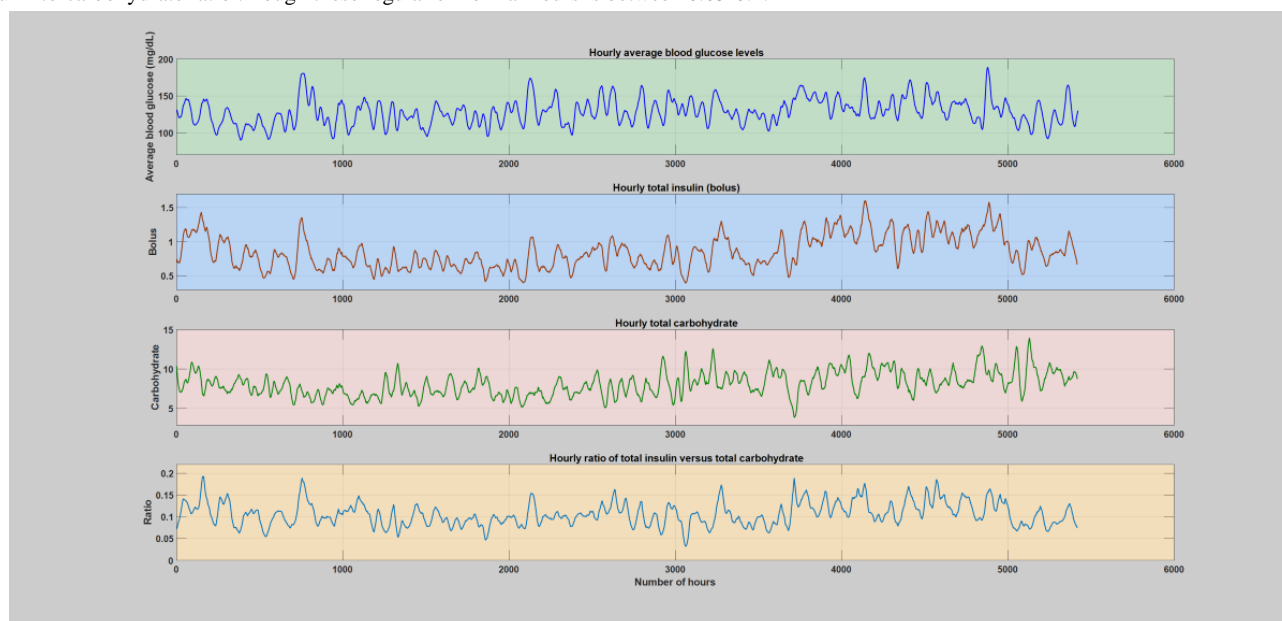
#### Trend Comparison for Normal Patient Years

During normal years when patients do not have any significant illness or infections ([Multimedia Appendix 2](#)), the insulin-to-carbohydrate ratio follows a similar trend in all the subjects, where the insulin-to-carbohydrate ratio lies between 0.05 and 0.2. An elaborate analytical plot of a typical patient year without infection incidence showing the phenomena is depicted in [Figures 1 and 2](#). A detailed analytical plot of the 5 patient years depicting the same phenomena can be found in [Multimedia Appendix 2](#). The insulin-to-carbohydrate ratio conveys interesting information about the usual operating point of the patient, depicting the necessary amount of insulin (bolus) required for every gram of carbohydrate consumed to maintain the blood glucose levels within a healthy range (typically recommended to be between 70 and 180 mg/dL). As can be seen from the yearlong trend analysis of the regular or normal patient years ([Multimedia Appendix 2](#)), despite the presence of various factors that are known to disturb blood glucose dynamics, both patient-controllable parameters and patient-uncontrollable parameters except infection incidence, the insulin-to-carbohydrate ratio remains to be relatively stable.

**Figure 1.** The first patient year, where there is no incidence of acute infections. The figure depicts the daily variation of average blood glucose levels, total insulin (bolus), total carbohydrate, and total insulin-to-total carbohydrate ratio. The operating point of the patient's insulin-to-carbohydrate ratio through these regular or normal days is between 0.05-0.2.



**Figure 2.** The first patient year, where there is no incidence of acute infections. The figure depicts variation of average blood glucose levels, total insulin (bolus), total carbohydrate, and total insulin-to-carbohydrate ratio during each hours of the day. The operating point of the patient's insulin-to-carbohydrate ratio through these regular or normal hours is between 0.05-0.2.



### ***Trend Comparison of Patient Years With Acute Infection***

The trend analysis of the key diabetes parameters, blood glucose, insulin, and carbohydrate, during acute infection suggests that there is a dramatic shift in the evolution of blood glucose, insulin, and carbohydrate (for detailed information, see [Multimedia Appendices 1 and 3](#)). Infection incidence brought about a dramatic increase in blood glucose levels, insulin intake, and reduction in carbohydrate consumption. The detailed analysis and the shift incurred on a weekly, daily, and hourly basis are presented in the following section.

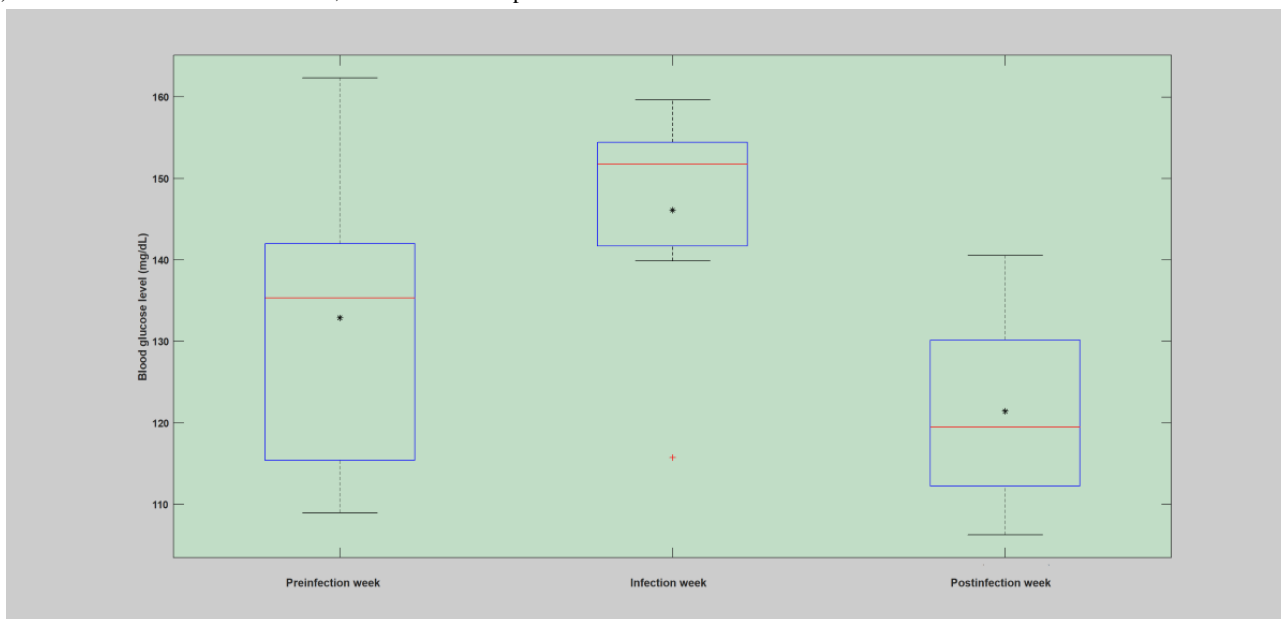
### ***Weekly Analysis***

The weekly analysis of the patient years was conducted by analyzing the deviation incurred on the key parameters of the blood glucose dynamics during the infection week in comparison with before and after the infection incidence. The raw data were used to estimate the deviations incurred due to infection incidence. The mean and SD of blood glucose levels, total insulin (bolus), and total carbohydrate were computed and used for comparison of the infection-induced deviations. As shown

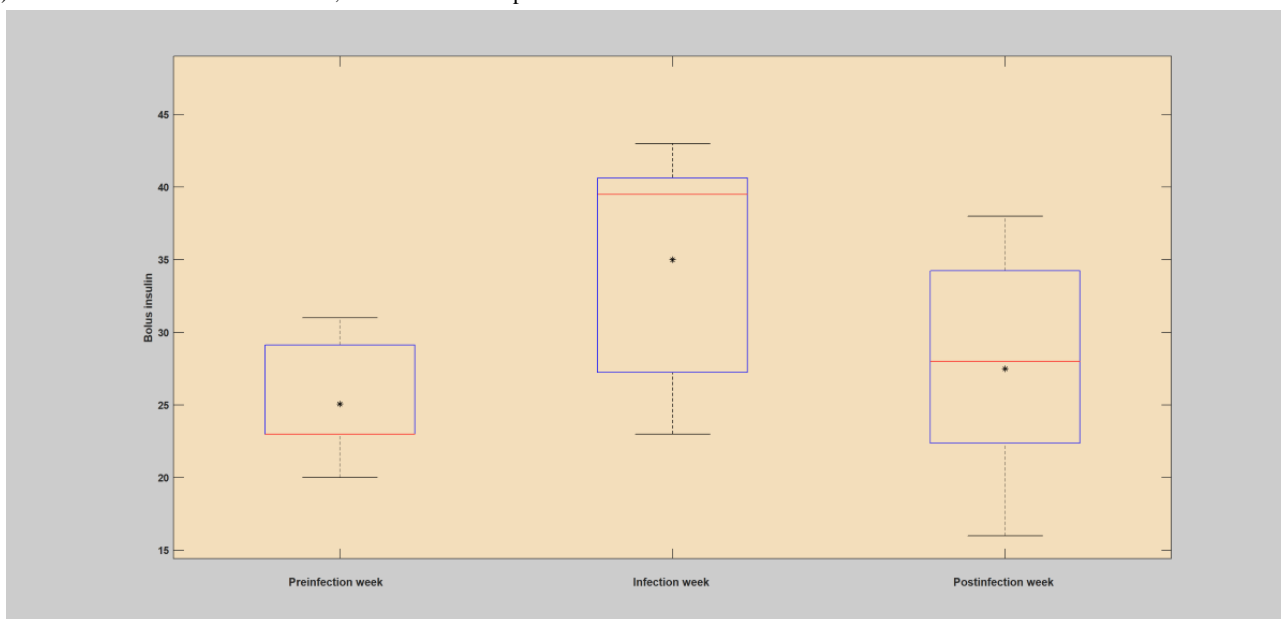
in [Figures 3-5](#) and [Table 5](#), in all the infection cases, the weekly analysis demonstrated that blood glucose levels were elevated despite higher insulin injection and reduced carbohydrate consumption. In all of these cases, it is clear that the incidence of infection has brought unreasonable deviation, with respect to the patient-controllable parameters, in the operation of the overall blood glucose dynamics as compared with the usual norm of the blood glucose dynamics. The presence of elevated blood glucose levels in the infection week, regardless of the high amount of insulin injections and lower carbohydrate consumption, clearly violated the norm of the blood glucose dynamics, where during normal situations the blood glucose levels are expected to drop with high insulin and reduced carbohydrate consumption. The fact that the blood glucose remains elevated during the infection incidence despite higher insulin injections and low carbohydrate consumption is highly associated with the infection phenomenon, which enhances the production of glucose and increased insulin resistance within the body to deliver more energy for the body to fight the pathogens. A more detailed description of the weekly analysis can be found in [Multimedia Appendix 1](#).



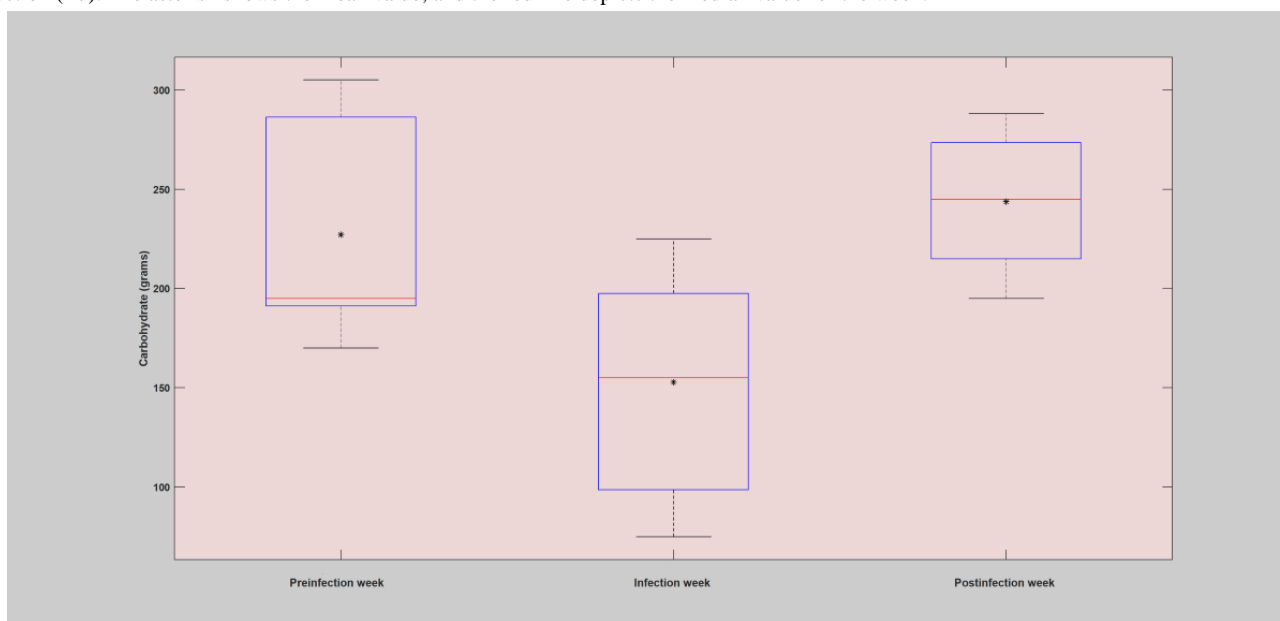
**Figure 3.** Analysis of blood glucose levels during the preinfection week, infection week, and postinfection week based on the first case of infection (flu). The asterisk shows the mean value, and the red line depicts the median value for the week.



**Figure 4.** Analysis of total insulin (bolus) intake during preinfection week, infection week, and postinfection week based on the first case of infection (flu). The asterisk shows the mean value, and the red line depicts the median value for the week.



**Figure 5.** Analysis of total carbohydrate (grams) intake during preinfection week, infection week, and postinfection week based on the first case of infection (flu). The asterisk shows the mean value, and the red line depicts the median value for the week.



**Table 5.** Mean and standard deviation of blood glucose levels, total insulin (bolus), and total carbohydrate during the preinfection week, infection week, and postinfection week.

Parameters	Preinfection week, mean (SD)	Infection week, mean (SD)	Postinfection week, mean (SD)
<b>The first case of infection (flu)</b>			
BG <sup>a</sup> (mg/dL)	130.74 (16.89)	141.95 (14.37)	119.16 (7.39)
Total insulin (bolus)	23.39 (4.91)	35.30 (6.11)	21.32 (4.61)
Carbohydrate (grams)	241.11 (57.27)	178.80 (65.69)	241.18 (37.63)
<b>The second case of infection (flu)</b>			
BG (mg/dL)	143.01 (19.53)	155.36 (21.99)	126.17 (11.70)
Total insulin (bolus)	28.07 (8.85)	41.07 (9.44)	25.36 (6.93)
Carbohydrate (grams)	190.14 (43.93)	161.14 (58.43)	214.57 (34.66)
<b>The third case of infection (flu)</b>			
BG (mg/dL)	136.93 (18.58)	144.12 (20.30)	134.18 (11.96)
Total insulin (bolus)	20.08 (5.44)	31.50 (10.84)	22.83 (3.86)
Carbohydrate (grams)	178.0 (45.87)	144.83 (37.63)	195.83 (42.59)
<b>The fourth case of infection (flu)</b>			
BG (mg/dL)	157.74 (31.12)	161.34 (19.88)	138.57 (19.83)
Total insulin (bolus)	24.43 (5.26)	32.14 (7.01)	29.29 (5.22)
Carbohydrate (grams)	199.06 (53.45)	167.04 (44.94)	226.07 (18.23)
<b>The fifth case of infection (flu)</b>			
BG (mg/dL)	135.21 (14.58)	139.88 (15.54)	122.87 (14.49)
Insulin (bolus)	32.80 (4.59)	40.37 (8.31)	33.36 (7.94)
Insulin (basal)	19.20 (1.21)	20.42 (2.06)	18.68 (1.56)
Total insulin	52.33 (5.14)	61.21 (8.26)	52.46 (8.47)

<sup>a</sup>BG: blood glucose.

## Blood Glucose Levels

In all these infection incidences, the individual blood glucose levels remain elevated for a prolonged period of time despite low carbohydrate consumption and increased insulin injections as compared with the regular or normal days. Blood glucose levels were elevated during the infection week as compared with the preinfection and postinfection weeks.

- During the first case of infection, the overall mean percentage increase in the infection week's blood glucose levels was 8.57% over the preinfection week and 19.12% over the postinfection week, as shown in [Table 5](#).
- During the second case of infection, the overall mean percentage increase in the infection week's blood glucose levels was 8.63% over the preinfection week and 23.13% over the postinfection week, as shown in [Table 5](#).
- During the third case of infection, the overall mean percentage increase in the infection week's blood glucose levels was 7.26% over the preinfection week and 7.41% over the postinfection week, as shown in [Table 5](#).
- During the fourth case of infection, the overall mean percentage increase in the infection week's blood glucose levels was 2.28% over the preinfection week and 16.43% over the postinfection week, as shown in [Table 5](#).
- During the fifth case of infection, the overall mean percentage increase in the infection week's blood glucose levels was 3.45% over the preinfection week and 13.84% over the postinfection week, as shown in [Table 5](#).

## Insulin Intake

The comparison of infection week insulin injections with preinfection and postinfection weeks revealed that there was a dramatic increase in the amount of insulin intake during the infection period.

- During the first case of infection, the overall mean percentage increase in the infection week's insulin (bolus) injection was 50.93% over the preinfection week and 65.59% over the postinfection week, as shown in [Table 5](#).
- During the second case of infection, the overall mean percentage increase in the infection week's insulin (bolus) injection was 46.31% over the preinfection week and 61.94% over the postinfection week, as shown in [Table 5](#).
- During the third case of infection, the overall mean percentage increase in the infection week's insulin (bolus) injection was 56.87% over the preinfection week and 37.98% over the postinfection week, as shown in [Table 5](#).
- During the fourth case of infection, the overall mean percentage increase in the infection week's insulin (bolus) injection was 31.56% over the preinfection week and 9.7% over the postinfection week, as shown in [Table 5](#).
- During the fifth case of infection, the overall mean percentage increase in the infection week's insulin (bolus) injection was 23.08% over the preinfection week and 21.01% over the postinfection week, as shown in [Table 5](#).

## Carbohydrate Consumption

Comparison of the amount of carbohydrate consumption during the infection week with the preinfection and postinfection weeks

revealed that there was a significant reduction during the infection period.

- During the first case of infection, the overall mean percentage reduction in the infection week's carbohydrate consumption was 25.84% below the preinfection week and 25.87% below the postinfection week, as shown in [Table 5](#).
- During the second case of infection, the overall mean percentage reduction in the infection week's carbohydrate consumption was 15.25% below the preinfection week and 24.90% below the postinfection week, as shown in [Table 5](#).
- During the third case of infection, the overall mean percentage increase in the infection week's carbohydrate consumption was 18.63% below the preinfection week and 26.04% below the postinfection week, as shown in [Table 5](#).
- During the fourth case of infection, the overall mean percentage increase in the infection week's carbohydrate consumption was 16.09% below the preinfection week and 35.34% below the postinfection week, as shown in [Table 5](#).

## Insulin-to-Carbohydrate Ratio

The insulin-to-carbohydrate ratio defines the amount of insulin a patient needs to take for every gram of carbohydrate consumed. The value of the insulin-to-carbohydrate ratio usually lies between 0.05 and 0.2 on normal occasions. However, it has dramatically increased upon the incidence of infection.

- During the first case of infection, the overall mean percentage increase in the infection week's insulin-to-carbohydrate ratio was around 125.84% above the normal operating point of the patient, as shown in [Table 5](#).
- During the second case of infection, the overall mean percentage increase in the infection week's insulin-to-carbohydrate ratio was approximately 144.43% above the normal operating point of the patient, as shown in [Table 5](#).
- During the first case of infection, the overall mean percentage increase in the infection week's insulin-to-carbohydrate ratio was around 93.75% above the normal operating point of the patient, as shown in [Table 5](#).
- During the fourth case of infection, the overall mean percentage increase in the infection week's insulin-to-carbohydrate ratio was approximately 70.84% above the normal operating point of the patient, as shown in [Table 5](#).

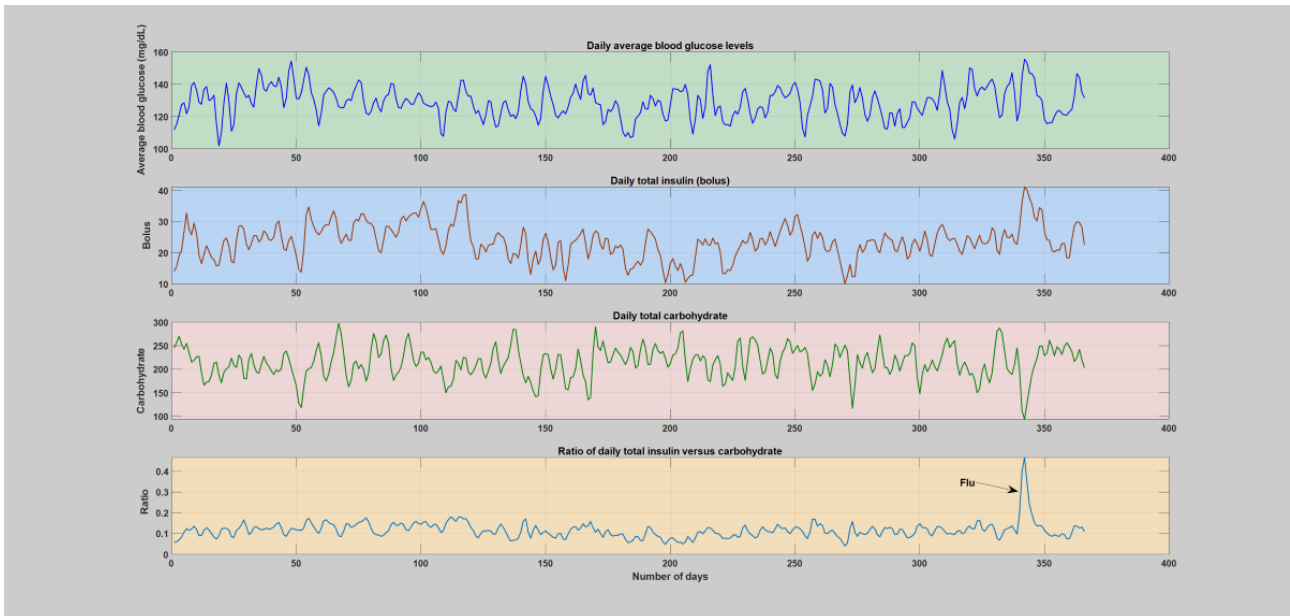
## Daily and Hourly Analysis

Hourly and daily analyses were conducted by analyzing the deviations incurred on the key diabetes parameters, blood glucose levels, insulin, carbohydrate, and the insulin-to-carbohydrate ratio as a result of infection incidence in contrast to the whole patient year. The comparison was carried out based on the smoothed version of the data, that is, 2 days window moving-average filter. Similar to the weekly analysis, the infection-induced shift of the blood glucose dynamics, that

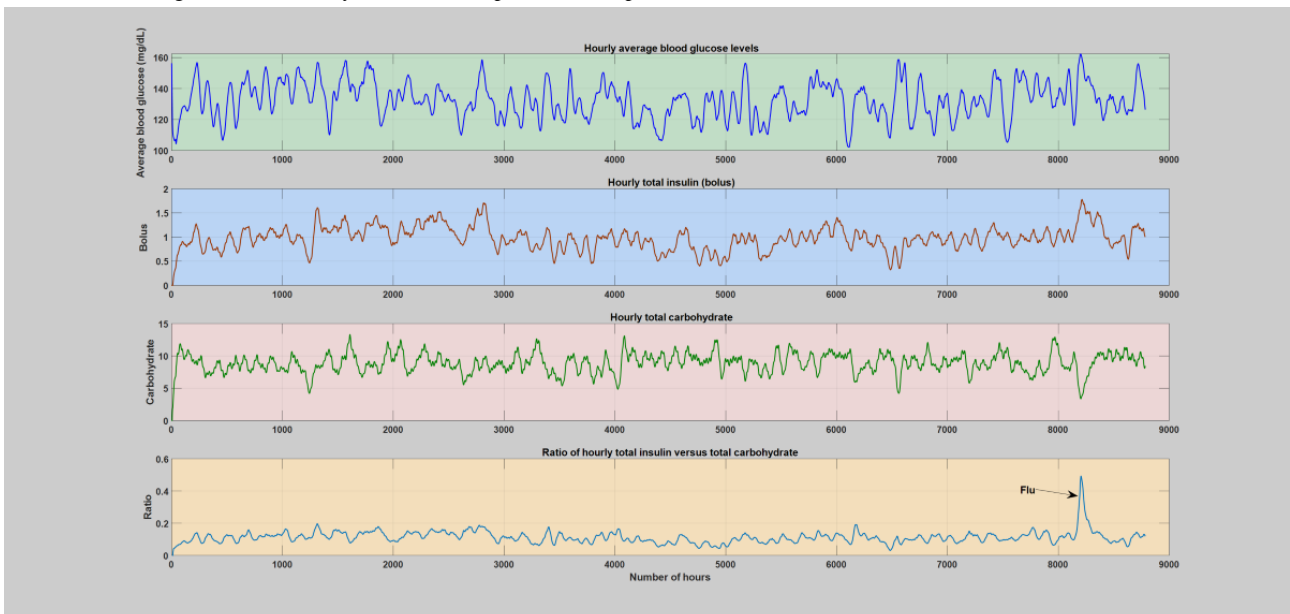
is, higher glucose production and increased insulin resistance, is clearly shown in both the daily and hourly analyses. As can be seen in Figures 6-11, the insulin-to-carbohydrate ratio of the patient has drastically shifted to a higher value to account for the effect of increased glucose production and insulin resistance (see Multimedia Appendix 3 for a detailed plot of the hourly

analysis in all the infection cases). In all of these cases, the insulin-to-carbohydrate ratio increases from the usual values of 0.05 to 0.2 during the normal period to higher values reaching 0.6, depending on the degree of severity of the infection incidence, type of pathogens involved, and the individual immunity.

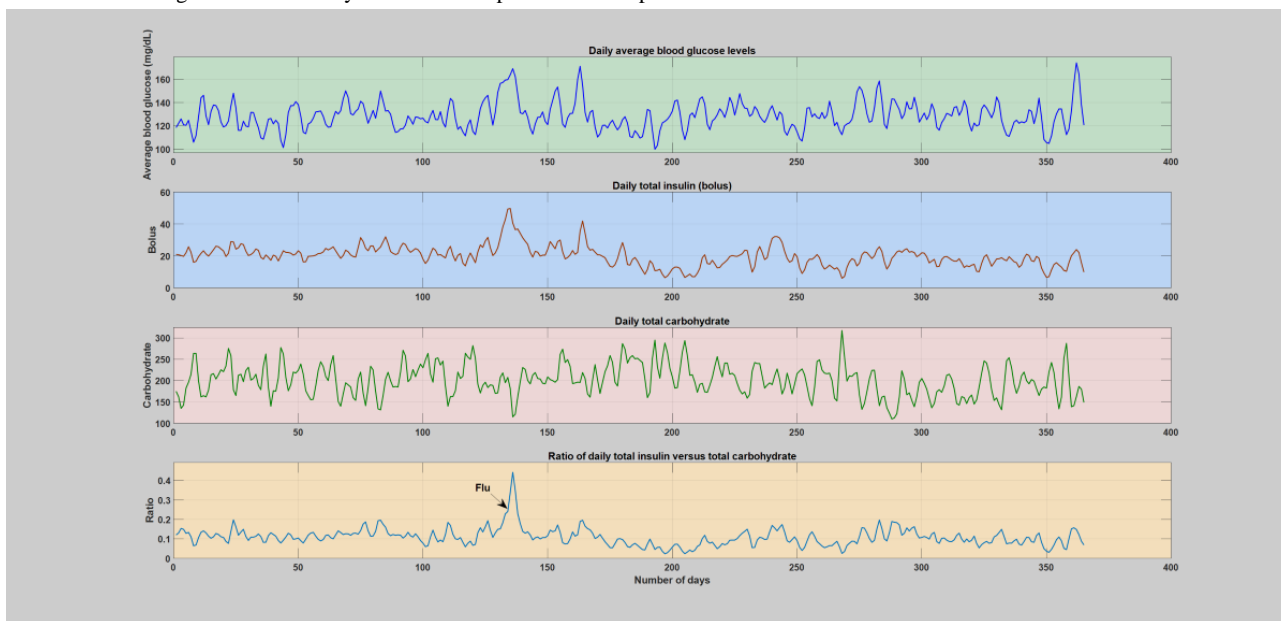
**Figure 6.** Daily analysis of the first infection case (flu). The figure depicts variation of average blood glucose levels, total insulin (bolus), total carbohydrate, and total insulin-to-total carbohydrate ratio. The operating point of the patient’s insulin-to-carbohydrate ratio had dramatically shifted and raised above the regular or normal days and reach a top around 0.5 upon midinfection week.



**Figure 7.** Hourly analysis of the first infection case (flu). The figure depicts variation of average blood glucose levels, total insulin (bolus), total carbohydrate, and total insulin-to-total carbohydrate ratio. The operating point of the patient’s insulin-to-carbohydrate ratio had dramatically shifted and raised above the regular or normal days and reach a top around 0.5 upon midinfection week.



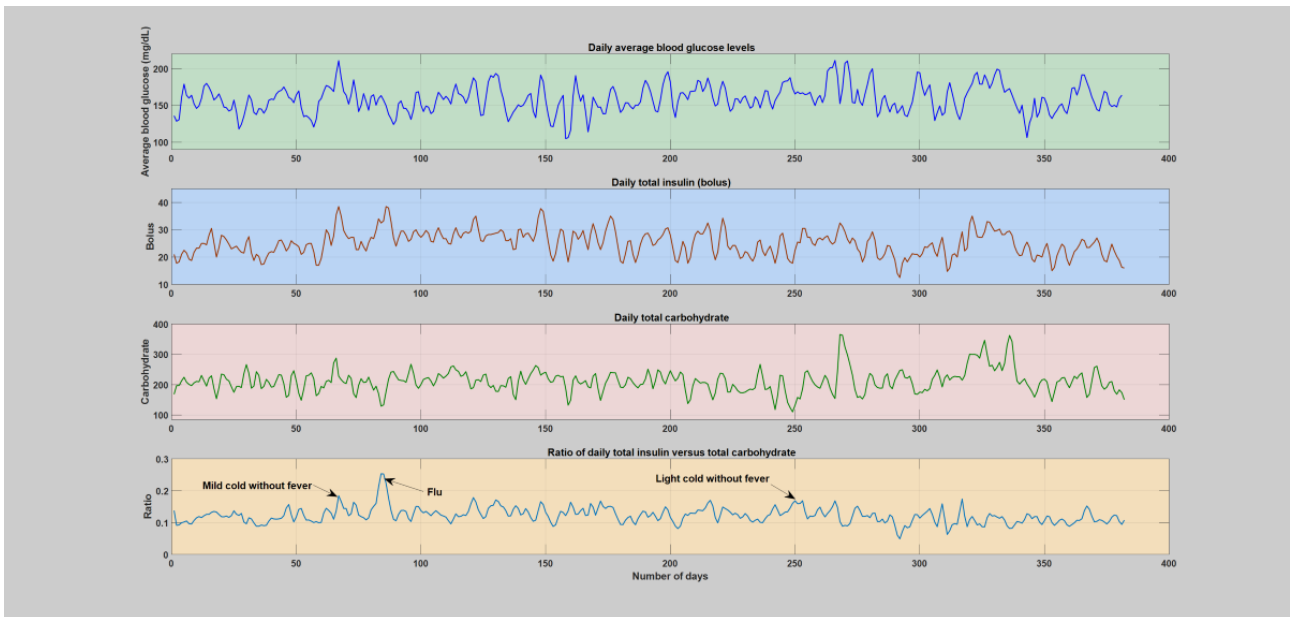
**Figure 8.** Daily analysis of the second infection case (flu). The figure depicts variation of average blood glucose levels, total insulin (bolus), total carbohydrate, and total insulin-to-total carbohydrate ratio. The operating point of the patient’s insulin-to-carbohydrate ratio had dramatically shifted and raised above the regular or normal days and reach a top around 0.45 upon midinfection week.



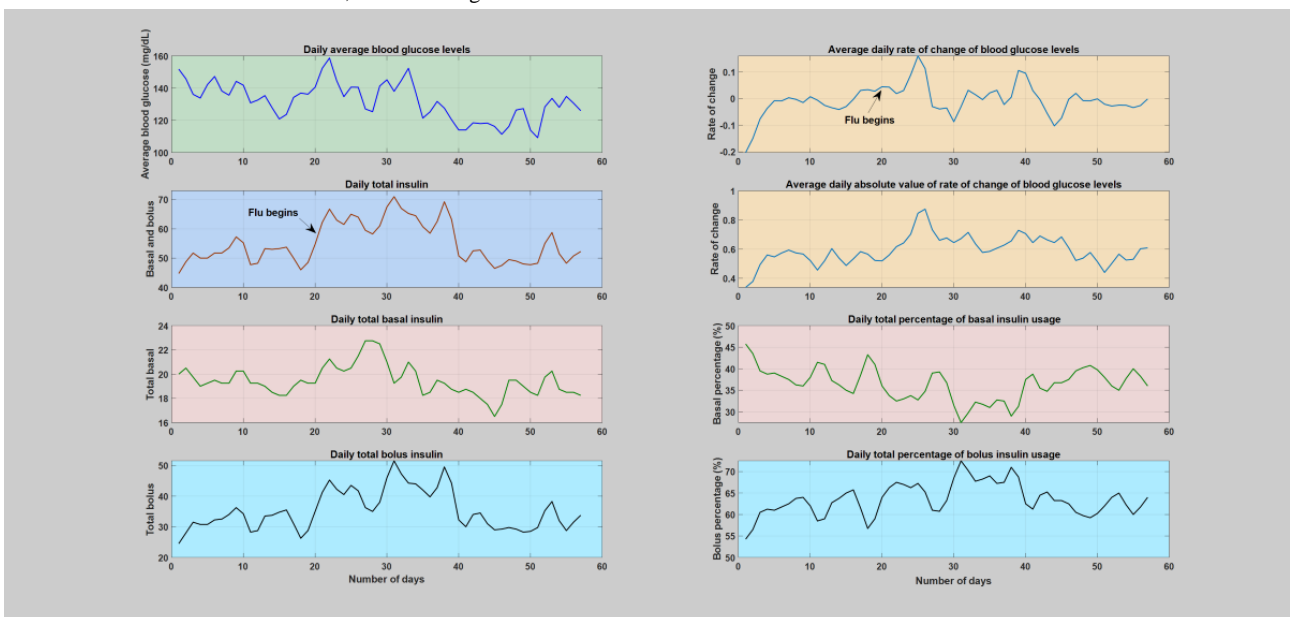
**Figure 9.** Daily analysis of the third infection case (flu). The figure depicts variation of average blood glucose levels, total insulin (bolus), total carbohydrate, and total insulin-to-total carbohydrate ratio. The operating point of the patient’s insulin-to-carbohydrate ratio had dramatically shifted and raised above the regular or normal days and topped around 0.4 upon midinfection week.



**Figure 10.** Daily analysis of the fourth infection case (mild common cold without fever, light common cold without fever, and flu). The figure depicts variation of average blood glucose levels, total insulin (bolus), total carbohydrate, and total insulin-to-total carbohydrate ratio. The operating point of the patient’s insulin-to-carbohydrate ratio had dramatically shifted and raised above the regular or normal days and reach a top around 0.28 upon midinfection week. A light common cold without fever seems to not significantly affect the operating point.



**Figure 11.** Daily analysis of the fifth infection case (flu). The figure depicts variation of average blood glucose levels, total insulin including both bolus and basal insulin, daily average rate of change of CGM and, absolute value of rate of change of CGM (computed based on CGM direction from the pump information), percentage of basal and bolus per total insulin units. As can be seen, as a result of the ongoing infection incidence, there is clear and dramatic rise in the amount of insulin, while blood glucose levels remain elevated.

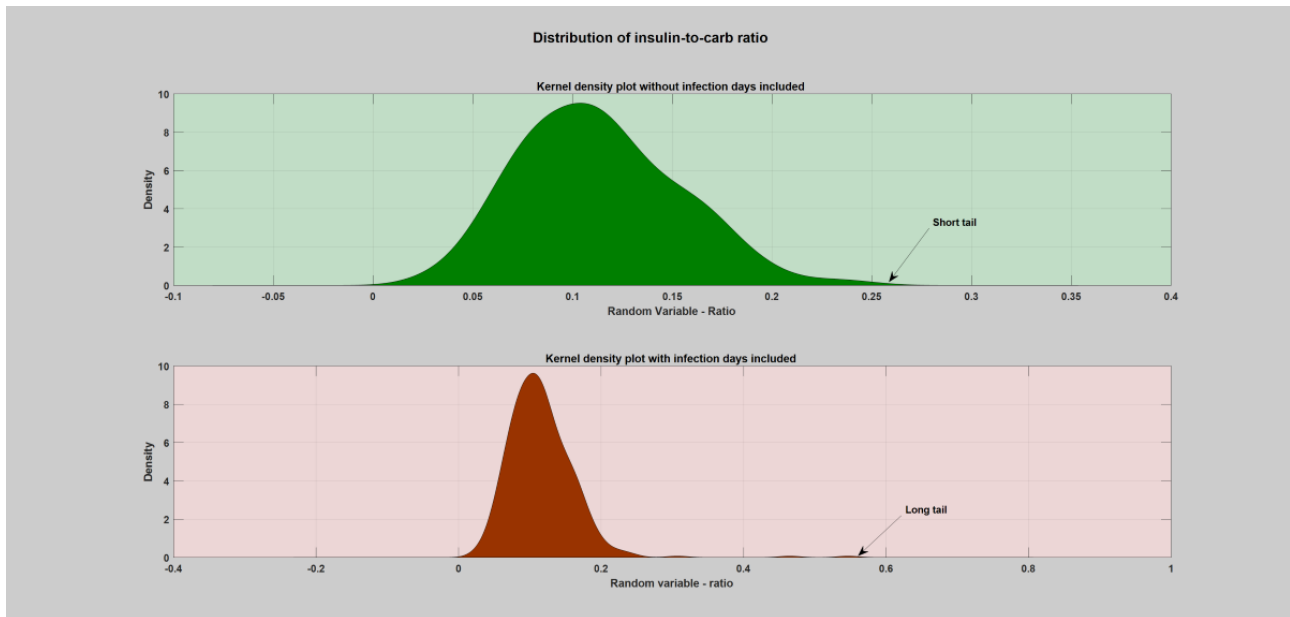


**Kernel Density Estimation–Probability Distribution**

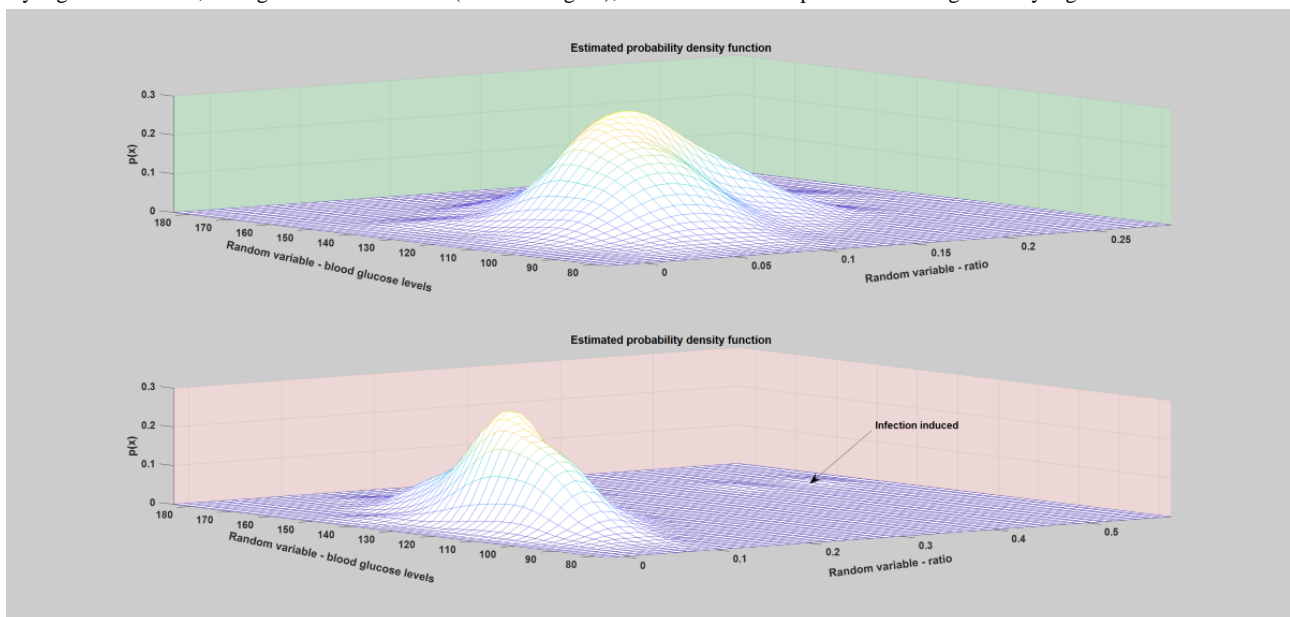
Kernel density was estimated to study and characterize the nature, shape, and degree of severity of the deviations incurred due to infection incidence by analyzing the probability distribution of the individual key parameters of the blood glucose dynamics. A univariate and bivariate kernel density estimation based on the insulin-to-carbohydrate ratio and blood

glucose levels was carried out on the yearlong data, as shown in Figures 12 and 13 (a detailed plot for all the infection cases, both hourly and daily, can be found in Multimedia Appendix 1). As can be seen from the figures, the infection incidence has brought a significant change in the probability distribution. However, the nature, shape, and degree of outlieriness depend on the type of pathogen involved, severity of infection, and individual immunity.

**Figure 12.** Univariate kernel density estimation of a patient year using the daily insulin-to-carbohydrate ratio. As can be seen from the tail of the distribution, during regular or normal days (the green shaded region), the yearly distribution of the patient’s insulin-to-carbohydrate ratio lies within the values of 0.005 and 0.2. However, during infection incidence (the red shaded region), there is a clear deviation in the tail of the distribution, where the values reaches around 0.58.



**Figure 13.** Bivariate kernel density estimation of a patient year using both the daily average blood glucose levels and insulin-to-carbohydrate ratio. As can be seen from the bivariate distribution, during regular or normal days (the top light green figure), the distributions are concentrated around the high density regions. However, during infection incidence (the lower figure), there is a clear bump far from the high density regions.



## Discussion

### Principal Findings

Presently, in relation to people’s mobility and travel, there is a growing concern regarding an infectious disease outbreak. Such an incident can be a menace to our global health security, which calls for early detection and immediate response. Thus, there is a growing need for new approaches and technologies to upgrade the existing surveillance system for early detection of emerging infectious diseases [1]. Existing disease surveillance systems detect the incidence of outbreaks long after the incidence of the first symptoms. Therefore, the purpose of this study was to

demonstrate how people with type 1 diabetes can assist in outbreak detection and further to shed light upon the possibility of assisting the individual during such an incident.

The advancement and omnipresence of smartphones, IoT devices, wearables, and sensors have enabled individuals to easily self-record health-related events often for self-tracking or self-managing their disease [5,6,53]. People with diabetes self-record detailed information including blood glucose levels, diet and insulin intake, physical activity, medication, and other parameters. The presence of such large self-recorded health data presents an opportunity to be used as a secondary source of information for other purposes such as digital epidemiology

and decision support applications. According to recent reports, the use of personal health information or self-collected data could mitigate the possibility of detecting infection incidence during the presymptomatic stage (improved sensitivity and timeliness), specifically during the incubation period, of which most of the current systems neglect from their process [13]. Our findings demonstrated that upon infection incidence, there is a dramatic shift in the operating point of the individual's blood glucose dynamics, which clearly violates the usual norm of blood glucose dynamics. During regular or normal days, blood glucose levels usually decrease when there is a significant increase in insulin injection and reduction in carbohydrate consumption. However, in all of the infection cases we analyzed, compared with the preinfection and postinfection weeks, the following were noticed:

- Blood glucose levels were elevated by an average of 6.1% and 16% over the preinfection and postinfection weeks, respectively.
- Insulin injection (bolus) increased by 42% and 39.3% over preinfection and postinfection weeks, respectively.
- Carbohydrate consumption was reduced by 19% and 28.1% compared with preinfection and postinfection weeks, respectively.
- The insulin-to-carbohydrate ratio increased by 108.7% on average in all cases.

In general, all of these findings confirm that during infection incidence, blood glucose levels are elevated despite injecting higher amounts of insulin and reduced carbohydrate consumption. The identified changes are quite significant anomalies compared with the regular or normal days and could potentially be detected with a dedicated personalized (individualized) computational health model. Various algorithms that span from prediction models to anomaly detection algorithms can be investigated to detect such infection-induced changes in blood glucose dynamics. Apart from the potential use of these findings in personalized digital infectious disease detection systems, it could also be used for decision support in self-management during infection and illness. As presented earlier, during the course of infection, individuals with diabetes usually struggle with severe hyperglycemia. Managing blood glucose levels during infection incidence is not an easy task, given the fact that it is caused by a mixed effect of both patient-controllable and patient-uncontrollable parameters. The patient can only estimate the disturbance caused by the amount of carbohydrate consumption, insulin injection, and physical activity load, which is not the case during infection incidence. Apart from these known major factors, that is, patient-controllable parameters, there is an underlying and unknown disturbance caused by the patient's uncontrollable parameters, such as counterregulatory hormones (CRHs), as a result of infection incidence. This unknown disturbance mainly increases glucose production from the liver and reduces insulin sensitivity. To this end, people with type 1 diabetes face a very difficult challenge to estimate the necessary amount of insulin for a given amount of carbohydrate consumption. In this regard, providing real-time decision support could reduce the burden during such a crisis. One possible approach could be characterizing the effect of different pathogens on blood glucose

dynamics, mainly on insulin resistance and its sensitivity change over the course of infection. However, a large set of infection-related self-recorded data need to be analyzed for investigating how each pathogen affects the key parameters of blood glucose dynamics during the entire course of infection. This requires collecting and analyzing infection-related data, and estimating the overall changes each pathogen could bring on insulin sensitivity during the course of infection. To this end, the presented result reflects a promising result that can be geared toward decision support during infection or illness. For example, the change in insulin-to-carbohydrate ratio can be used to provide general information related to each pathogen on what to expect, such as the percentage of insulin resistance during the first days, in the middle, and at the final days of the infection.

### Infection-Induced Shift of Operating Point in Blood Glucose Dynamics

During infection incidence, people with diabetes usually struggle with severe hyperglycemia and critical hypoglycemia if not properly managed. However, during regular or normal days, the patient can manage the incidence of hyperglycemia, which is mostly diet-induced, by properly controlling the patient-controllable parameters, for example, amount of carbohydrate consumption, insulin injection, and performing balanced physical activity or exercise. Yet, during infection incidence, it turns out to be very difficult to manage the hyperglycemia incidence due to the fact that it is caused by a mixed effect of both patient-controllable and patient-uncontrollable parameters. The patient's uncontrollable parameters define the action of hormonal effects such as CRH induced by either physiological stress or emotional stress. The hormonal effect is two-sided, which is a higher glucose production from the liver and inhibiting insulin production and reducing sensitivity [54,55]. A detailed study conducted by Waldhausl et al [56] demonstrated the significant effect of stress hormones on the production of glucose and insulin resistance. The study was conducted by infusing different stress hormones to investigate the effect of exposure to these hormones on blood glucose response [55]. The extent and degree of hyperglycemia events and insulin resistance during infection incidence directly correlate with the type of pathogen, the type of hormone involved and the severity of the infection [37,38,55]. Generally, the phenomenal effect of infection incidence on blood glucose dynamics in people with diabetes can be simply described using the following relationships:

$$BG_t = CH_t + CRH_t - \phi IN_t - PA_t$$

Where  $\phi$  is an insulin sensitivity factor, BG is the blood glucose level, CH is the amount of carbohydrate consumption, IN is the amount of insulin injection, PA is the amount of physical activity session or exercise load, and CRH is the effect of CRHs. The equation depicts the phenomena that occur during infection incidence, where blood glucose levels are raised by the action of both patient-controllable parameters (CH) and patient-uncontrollable parameters (CRHs, such as cortisol and adrenalin). Thus, consumption of any regular diet in an individual can induce severe hyperglycemia due to the added effect of glucose production from the liver as a result of the



CRH effect [55]. For this reason, the patient is expected to reduce the amount of carbohydrate intake to a certain extent to optimally manage the hyperglycemia crises and at the same time avoiding any critical hypoglycemia incidences (for more information, see [Multimedia Appendix 1](#)). By the same token, blood glucose levels can be lowered to euglycemia by the patient-controllable parameters (insulin [IN] and physical activity session or exercise load [PA]). However, due to the change in insulin sensitivity, the action of insulin is reduced ( $\phi$  is affected by infection incidence), and the patient is expected to deliver more insulin injections to counterbalance the effect of insulin resistance [57]. According to our results, all these scenarios are reflected in the individual's blood glucose dynamic infected with flu (influenza), where a dynamic shift occurred from the usual operating point of the blood glucose dynamics. There are elevated blood glucose levels, despite injecting a higher amount of insulin and consuming less carbohydrate than the regular or normal days. These characteristics are clearly demonstrated on the shift incurred on the individual's insulin-to-carbohydrate ratio as compared with the regular or normal days. Therefore, blood glucose, amount of injected insulin, diet intake, and insulin-to-carbohydrate ratio and other supporting physiological parameters such as body temperature and blood pressure can be exploited to develop a personalized health model for detecting infection incidence among people with type 1 diabetes. Given the similarity, this result can also be translated to other types of diabetes, such as people with type 2 diabetes. It is worth mentioning that apart from infection incidence, other factors such as emotional stress could also result in similar variable episodes of elevated blood glucose levels [17]. This can obviously impact the detection performance of the model. However, our results based on yearlong patients' data demonstrated that the use of carbohydrate consumption, insulin injections, and insulin-to-carbohydrate ratio along with the blood glucose could solve this confounding nature. Moreover, acute emotional stress, other than the chronic ones, might have less influence on one's meal appetite compared with infection incidence to skew the insulin-to-carbohydrate ratio [17].

### Relevance of the Data

The informational values of the data, availability of the data, and cost of the data are the 3 key metrics necessary to evaluate the relevance of new surveillance data for a digital infectious disease detection system [58]. The informational value of the data assesses how informative the data are to facilitate the detection or characterization of infectious disease outbreaks. In this regard, the surveillance data must clearly indicate the absence or presence of infections either on an individual or population level or both in a timely manner. Furthermore, the rate of false alarms derived from the data is an important factor that dictates the acceptability of the surveillance data, which is in turn governed by the signal-to-noise ratio defining the signal's strength depicting the infection period as compared to the regular or normal period (baseline data) [58]. In this regard, our results demonstrated that the infection-induced signal exhibits high discriminative power from the baseline (normal or regular) patterns. The availability of surveillance data is another crucial indicator for screening potential types of data, which needs to

be addressed [58]. In this regard, given the widespread and ubiquitous nature of mobile apps, and different sensors, people with type 1 diabetes collect far more data than ever. For example, many people with type 1 diabetes use continuous glucose monitors and insulin pumps, which are predicted to grow further in terms of both quality and quantity of data in the coming years. The most crucial challenge in this direction includes issues related to security, privacy, and confidentiality of user data if there is a necessity to collect user data into a central server than deploying the detection algorithm on the user's own mobile device. The cost of data delineates the associated cost in relation to acquiring the data in question, including the cost incurred for realizing the data collection system [58]. In this regard, the individual's self-recorded data are solely collected for their own use and used as a secondary source of information for disease surveillance purposes. Providing tailored and valuable feedback to the individual patient might further motivate them to participate on a large scale (for further details, see the section Ethical and Motivational Challenges).

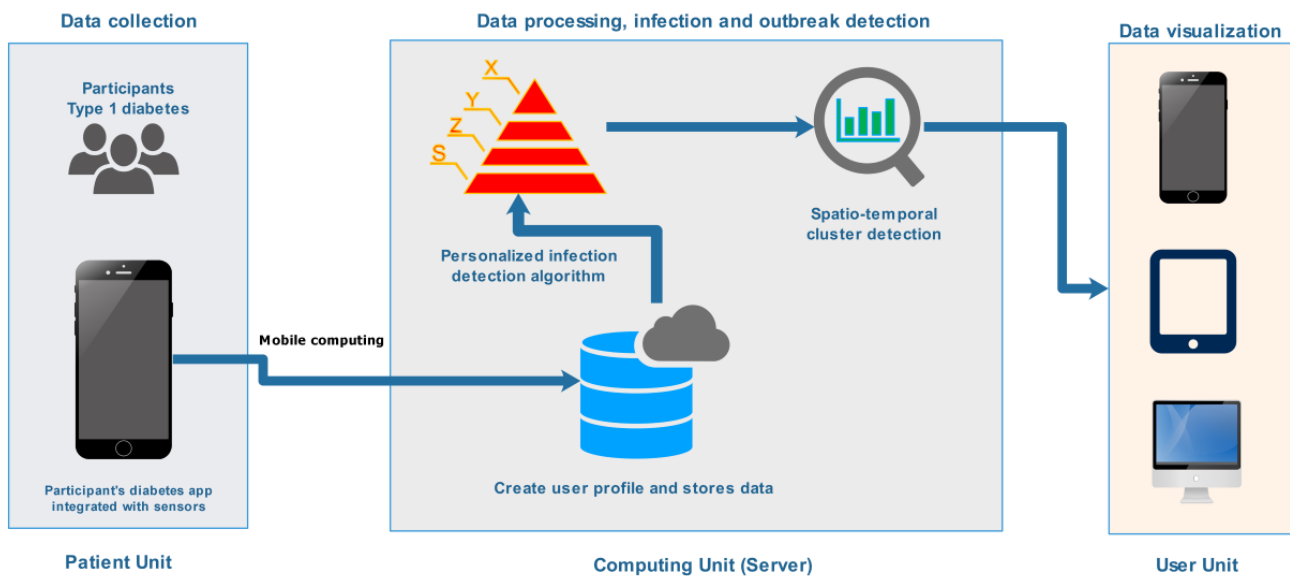
### Framework of a Personalized Digital Infectious Disease Detection System

Epidemic intelligence encompasses activities directed toward early detections, verification, and assessment of potential public health threats to notify and recommend necessary measures for the concerned bodies regarding the ongoing situation [56]. Early detection systems such as Google Flu Trends and other existing systems have certain limitations because they do not have the mechanisms to identify or track individual cases through diagnosis or screening based on a personalized health model. This limitation has a major impact and certainly introduces bias in disease outbreak prediction. Currently, a personalized health model, which resembles the way clinicians and epidemiologists classify an individual as normal, suspected, or confirmed case, for screening and case detection doesn't exist [58]. Having a personalized health model can provide information for both individual health-related decision support purposes and at the same time can be used for tracking infectious disease outbreaks among the public. The results of this study demonstrated that commencement of infection in people with type 1 diabetes significantly alters the individual blood glucose dynamics, and such a change can potentially be detected through modeling of the individual blood glucose dynamics. Moreover, incorporating various physiological parameters, for example, heart rate and body temperature, to a personalized health model will further enable the capture of infection incidence as early as possible, that is, incubation period. Therefore, the development of a personalized health model-based digital infectious disease detection system is vital for the success of next-generation public health surveillance systems. The data sources and signal exploited, outbreak detection algorithms employed, clustering approaches, and visualization techniques used to play a central role in any digital infectious disease detection systems by determining its accuracy (sensitivity) and timeliness (lead time) [56]. On the basis of the kind of data sources and signals exploited, infectious disease surveillance systems can be generally grouped into an indicator-based and event-based system [56,59,60]. Event-based systems mainly rely on

unstructured data collected through formal or informal sources and is characterized by quick detection, reporting, and assessments of public health events, including clusters of disease [56,60]. On the other hand, indicator-based systems mainly use structured data, which are collected following a standard case definition and is characterized by routine reporting of disease cases [56,60]. The proposed system [26,61], as shown in Figure 14, is categorized under event-based digital infectious disease detection systems, where the events are grouped under

microevents and macroevents [56]. Under the umbrella of these events and the proposed system in general, a framework of several components such as infection detection algorithms (how to develop an algorithm to detect infection incidence at the individual level); clustering algorithms (how to group the infected individuals to form a cluster); visualization techniques (how to report and display the detected outbreak incidence) and further ethical and motivation challenges are briefly discussed below.

Figure 14. The Proposed System Architecture.

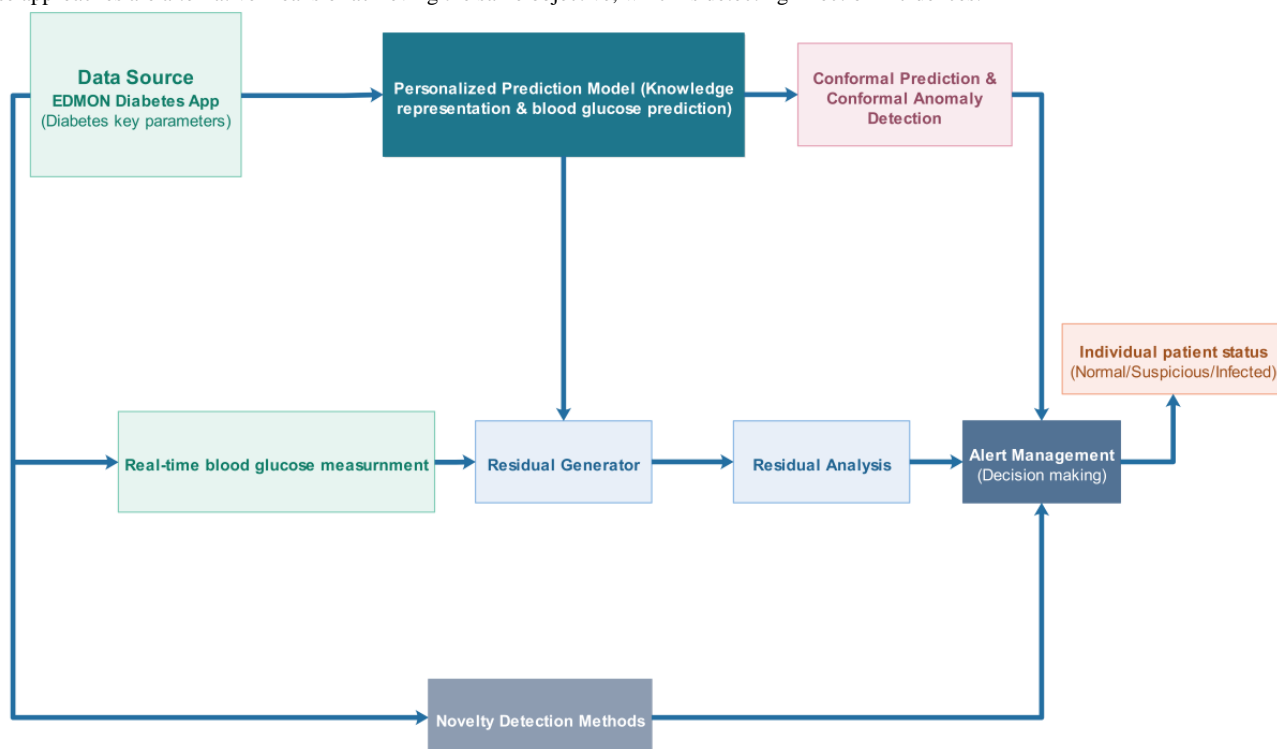


**Microevent: Individual-Level Detection of Infection Incidence**

The detection of microevents as the name suggests is carried out at an individual level by tracking the individual’s diabetes profile including blood glucose levels, amount of insulin injections, carbohydrate consumption, physical activity or exercise sessions, and others. The presence of elevated blood glucose levels despite injecting higher amounts of insulin and consumption of less carbohydrates is regarded as a marker of an event of infection incidence and hence can be defined as a microevent for the event-based digital infectious disease

detection system. Detecting the incidence of these kinds of deviation from the usual norm of blood glucose dynamics requires a proper personalized health model, which can learn from past history of the patient and judge whether the information conforms with the usual trend. Hence, the proposed personalized health model for detecting these types of microevents incorporates 3 components: a data source, personalized infection detection algorithm, and alarm management module, as shown in Figure 15. As can be seen from the figure, the personalized infection detection algorithm can be modeled using either a prediction model-based approach or a novelty or anomaly detection-based approach.

**Figure 15.** The proposed personalized infection detection algorithms for detecting microevents (incidences of infections) in people with type 1 diabetes. These approaches are alternative means of achieving the same objective, which is detecting infection incidences.



**Data Sources and Input**

The patient unit is a mobile health app, as shown in Figure 14, which integrates data from different sensors and wearables that record key diabetes parameters, such as blood glucose levels, insulin dosage, diet, physical activity, and other optional physiological parameters including body temperature, heart rate, blood pressure, and others [26,61]. The app is also expected to record the geographical location of the individual along with the time of data registration. For example, one way of estimating user location can be carried out based on global positioning system (GPS) information from the mobile phone during data registration [61]. The geographical location of the user can be the geographical coordinates of longitude and latitude [62], postal code address [63], or any local reference coordinates.

**Personalized Infection Detection Algorithm**

Detection of microevents can be carried out using individual self-recorded historical data based on a personalized health model, that is, either a prediction model [64,65] or novelty detection algorithms [66-68], as shown in Figure 15. The prediction model-based algorithm requires learning the individual blood glucose dynamics for accurate prediction, and for the purpose of detecting the microevents, it can be implemented as either a residual-based [69-72] or a conformal prediction-based approach [73-80]. In a similar fashion, novelty detection-based algorithms can be other alternatives for detecting novel microevents relying on either supervised, semisupervised, and unsupervised approaches [4,66,67,81]. Different categories of novelty detection approaches could be exploited for detecting infection-induced deviations in blood glucose dynamics, including approaches based on statistical techniques [68], prediction, density [82-85], distance [67,86],

classification or domain [4,62,63,87-92], clustering [62,93], and ensemble [67,68,80-82,85,86,92-95].

**Alarm Management (Decision Making)**

The alarm management module accepts the score computed by the personalized infection detection algorithm as input and evaluates the degree of severity of the infection incidence. The severity is evaluated based on the degree of abnormalities of the anomalies score, and a label could be assigned to the individual patient status as normal (0), suspicious (-1), and infected (1). For example, a rule-based fuzzy logic with membership functions of infected, normal, and suspicious can be used to assign the label indicating the severity of the infection incidence using the anomaly score. The output from the alarm management will be directly fed to the cluster detection analysis, which is used to detect a group of patients based on geography (space) and time so as to reveal if there is any ongoing infectious disease outbreak.

**Macroevents: Population-Level Infectious Disease Outbreak Detection**

**Cluster Detection Mechanism**

Cluster detection is defined as the process of identifying a group of infected individuals with similar spatial, temporal, or spatio-temporal attributes [96]. A spatial cluster analysis only considers a patient’s geographical location, and a temporal cluster analysis considers only the time aspect of the events. However, a spatio-temporal cluster analysis is conducted to look for aberrant patterns and detect a cluster of infected people within a specified geographical region and predefined timeframe [96,97]. The analysis of space time clusters is carried out based on a couple of steps: geocoding and identification, which transforms the patient address into meaningful coordinates and

detecting the clusters based on the transformed location and time. A space time cluster analysis is the most favored approach when it comes to early detection of an infectious disease outbreak. A space time cluster analysis can be designed by performing a spatial analysis first and then superimposing the temporal aspect [97]. Regarding the proposed system, the input to the space time cluster detection analysis consists of the individual patient status from alarm management, user location, and time of data registration [26]. The status of the individual patient at any time can be normal (0), infected (1), or suspicious (-1), which comes from alarm management. The user's geographical location can be geographical coordinates of longitude and latitude [98], postal code address [99], or any local reference coordinates. Estimation of user location can be carried out using GPS information from the user's mobile phone, which can be accessed during each data registration. The time aspects depend on the requirement of detection frequency and can be set to either an hourly or daily window. One optimal approach could be tracking the individual during each hour of the day for any statistically significant deviations and performing a concluding analysis at the end of each day based on the daily analysis. Various algorithms have been implemented in the literature, including the density-based clustering algorithm, Bayesian spatial scan statistics, K-NN with Haversian distance (K-nearness), cumulative summation, space time scan statistics, space time permutation scan statistic, and space scan statistics [96,97,100], which can be further tested and adopted. The most important challenge is the sparsity of the data set considering the small proportion of people with type 1 diabetes that can be under surveillance over a large region. Therefore, it is necessary to adopt these cluster detection techniques to overcome data sparsity and produce acceptable detection accuracy. In the proposed system, the detected clusters, if there is any, can be displayed and viewed based on real-time and interactive data visualization tools.

### Data Visualization

Data visualization is a mechanism by which detected clusters of disease outbreaks, if there is any, are presented to the responsible bodies for quicker public health actions and responses. Generally, such a visualization tool could report outbreaks of epidemic cases for investigation and follow-up, and it could also report the duration of the epidemic (timing), degree of severity of the epidemic, and the region under threat. In the literature, there are various implemented visualization tools and visual displays with regard to disease outbreak detection systems, including ArcGIS, Google map API, TwiInfo, OpenStreetMap, and JFreeChart, and display mechanisms such as maps, time series, graphs, and color indicators [96]. These visualization tools and display mechanisms can be further tested and adopted in the proposed system. The real-time health status of an individual from the ongoing tracking could be accessible to the end user and can be displayed in a stand-alone software app based on smartphones, tablets, and computers or a dedicated website [26]. Generally, both the data providers (participants) and the general population could benefit from the system in the sense that they can take actions needed to avoid being infected. Moreover, the individual patient could also receive analysis and feedback from the system to learn the situation, such as the

degree and severity of deviation of different parameters, including blood glucose, insulin, diet, and insulin-to-carbohydrate ratio, along with their trend as compared with the noninfection period.

### Ethical and Motivational Challenges

The implementation of a digital infectious disease detection system based on self-recorded data poses serious challenges that require special attention, such as user privacy and security, data confidentiality, user acceptance, and motivations [26,101], especially during data collection, transmission, and data storage [102,103]. Personal health-related data are sensitive, and the data collection, transmission, and data storage procedure need to follow the standards and regulations provided by the major governing bodies, such as General Data Protection Regulation (GDPR) and Health Insurance Portability and Accountability Act (HIPAA) [104,105]. This includes privacy-preserving mechanisms such as pseudonymization and anonymization to meet the necessary data compliance requirement along with user informed consent [102,103]. According to GDPR, the deidentification procedure is one of the recommended anonymization standards to preserve data confidentiality [104,105]. Moreover, from the technology perspective, it is necessary to look for a robust mechanism to ensure that user privacy and security are respected during data collection, transmission, and storage, as this is highly critical for successful acceptance of the proposed system [26,106]. One such alternative is to look for the possibility of deploying the infection detection algorithm (app logic) on the user (client) mobile device terminal to avoid transmission of patient data to a central server, where only the timely computed infection status of the patient will be sent to the central server for further cluster detection processing. However, this choice requires further feasibility studies to determine the cost, especially in terms of power constraints related to the mobile device terminal, since the detection algorithms need to continuously run in the background to compute the individual's infection status, at the most each hour of the day [26]. In addition, users might also lack willingness to adopt a new technology or system for various reasons ranging from lack of trust, lack of motivation, lack of perceived usefulness, and ease of use [26,101]. However, these challenges can be mitigated by properly buying user trust by developing state-of-the-art technology for preserving privacy, security, and confidentiality of the user and addressing factors that enhance user motivation, including usability knowledge, simplicity and ease of use, reduced time and frequency of interaction with the system, incentives, and others [101].

### Conclusions

The relationship between infection incidents and elevated blood glucose levels has been known for a long time. People with type 1 diabetes often experience prolonged episodes of elevated blood glucose levels as a result of infection incidence. Despite the fact that patients increasingly gather data about themselves, there are no solid findings on how to use such self-recorded data as a secondary source of information for other purposes, such as self-management-related decision support during infection incidence and digital infectious disease detection systems. We presented the effect of infection incidence on key

parameters of the blood glucose dynamics along with the necessary framework to exploit the information for realizing a digital infectious disease detection system and further shed light on the possibility of assisting individuals during infection-related blood glucose management crises. The results demonstrated that despite tight blood glucose control, blood glucose level is still elevated during infection incidence. The analysis shows that infection incidences have a significant impact on blood glucose dynamics as compared with the other patient-uncontrollable factors. All of these findings indicate that blood glucose levels were elevated despite a higher amount of insulin injection and reduced carbohydrate consumption, which are quite significant changes that could possibly be detected through personalized modeling that spans from prediction models to anomaly detection algorithms. However, further large-scale studies are required to strengthen the findings. Moreover, future research should investigate the possibility of improving detection time and disease characterization. Early detection, that is, during the incubation period, is a critical

component of any outbreak detection system and therefore needs to be improved by analyzing how various features of CGM can be used in context with other parameters, such as diet, insulin, and physical activity data. For instance, different individuals with type 1 diabetes often reported the experiences of an elevated episode of blood glucose levels before the onset of the first symptoms. Disease characterization involves determining the type and nature of pathogens that cause the infection, which is an important component of outbreak reporting. The extent and degree of the impact of infection incidence on blood glucose dynamics are highly correlated with the disease pathogens involved. In this regard, carefully analyzing a large-scale self-recorded data set containing several infection incidences (different pathogens) could characterize them based on their effect on blood glucose dynamics. Generally, we foresee that these findings can benefit the efforts toward building next-generation digital infectious disease surveillance systems and provoke further thoughts in this challenging field.

---

### Acknowledgments

The work presented in this paper is part of the project “Electronic Disease Surveillance Monitoring Network (EDMON) system,” which is funded by the University of Tromsø—The Arctic University of Norway and National Library of Medicine “Mechanistic machine learning” (grant number: LM012734) and is the PhD program of the first author, AW. The authors would like to extend their sincere gratitude to all the participants of the study.

---

### Authors' Contributions

The first author, AW, conceived the study, designed and performed the experiments, and wrote the manuscript. IK, EÅ, AH, DA, and GH provided successive inputs and revised the manuscript. All authors approved the final manuscript.

---

### Conflicts of Interest

None declared.

---

### Multimedia Appendix 1

Comparative analysis of parameters of blood glucose dynamics with and without infection incidences.

[\[DOCX File , 9296 KB-Multimedia Appendix 1\]](#)

---

### Multimedia Appendix 2

Analytical plot of the normal/regular patient years.

[\[DOCX File , 1346 KB-Multimedia Appendix 2\]](#)

---

### Multimedia Appendix 3

Analytical plot of the patient years with acute infection.

[\[DOCX File , 1662 KB-Multimedia Appendix 3\]](#)

---

### References

1. Choi J, Cho Y, Shim E, Woo H. Web-based infectious disease surveillance systems and public health perspectives: a systematic review. *BMC Public Health* 2016 Dec 8;16(1):1238 [\[FREE Full text\]](#) [doi: [10.1186/s12889-016-3893-0](https://doi.org/10.1186/s12889-016-3893-0)] [Medline: [27931204](https://pubmed.ncbi.nlm.nih.gov/27931204/)]
2. Hope K, Durrheim DN, d'Espaignet ET, Dalton C. Syndromic Surveillance: is it a useful tool for local outbreak detection? *J Epidemiol Community Health* 2006 May;60(5):374-375 [\[FREE Full text\]](#) [doi: [10.1136/jech.2005.035337](https://doi.org/10.1136/jech.2005.035337)] [Medline: [16680907](https://pubmed.ncbi.nlm.nih.gov/16680907/)]
3. Banaee H, Ahmed MU, Loutfi A. Data mining for wearable sensors in health monitoring systems: a review of recent trends and challenges. *Sensors (Basel)* 2013 Dec 17;13(12):17472-17500 [\[FREE Full text\]](#) [doi: [10.3390/s131217472](https://doi.org/10.3390/s131217472)] [Medline: [24351646](https://pubmed.ncbi.nlm.nih.gov/24351646/)]

4. Woldaregay AZ, Årsand E, Botsis T, Albers D, Mamykina L, Hartvigsen G. Data-driven blood glucose pattern classification and anomalies detection: machine-learning applications in type 1 diabetes. *J Med Internet Res* 2019 May 1;21(5):e11030 [FREE Full text] [doi: [10.2196/11030](https://doi.org/10.2196/11030)] [Medline: [31042157](https://pubmed.ncbi.nlm.nih.gov/31042157/)]
5. Fawcett T. Mining the quantified self: personal knowledge discovery as a challenge for data science. *Big Data* 2015 Dec;3(4):249-266. [doi: [10.1089/big.2015.0049](https://doi.org/10.1089/big.2015.0049)] [Medline: [27441406](https://pubmed.ncbi.nlm.nih.gov/27441406/)]
6. Gurrin C, Smeaton AF, Doherty AR. LifeLogging: personal big data. *FNT Inf Ret* 2014;8(1):1-125 [FREE Full text] [doi: [10.1561/1500000033](https://doi.org/10.1561/1500000033)]
7. Prince JD. The quantified self: operationalizing the quotidien. *J Electron Resour Med Libr* 2014 May 30;11(2):91-99 [FREE Full text] [doi: [10.1080/15424065.2014.909145](https://doi.org/10.1080/15424065.2014.909145)]
8. Rawassizadeh R, Momeni E, Dobbins C, Mirza-Babaei P, Rahnamoun R. Lesson learned from collecting quantified self information via mobile and wearable devices. *J Sens Actuator Netw* 2015 Nov 5;4(4):315-335 [FREE Full text] [doi: [10.3390/jsan4040315](https://doi.org/10.3390/jsan4040315)]
9. Bellazzi R, Abu-Hanna A. Data mining technologies for blood glucose and diabetes management. *J Diabetes Sci Technol* 2009 May 1;3(3):603-612 [FREE Full text] [doi: [10.1177/193229680900300326](https://doi.org/10.1177/193229680900300326)] [Medline: [20144300](https://pubmed.ncbi.nlm.nih.gov/20144300/)]
10. Rodríguez-Rodríguez I, Rodríguez JV, Zamora-Izquierdo M. Variables to be monitored via biomedical sensors for complete type 1 diabetes mellitus management: an extension of the 'on-board' concept. *J Diabetes Res* 2018;2018:4826984 [FREE Full text] [doi: [10.1155/2018/4826984](https://doi.org/10.1155/2018/4826984)] [Medline: [30363935](https://pubmed.ncbi.nlm.nih.gov/30363935/)]
11. Vayena E, Salathé M, Madoff LC, Brownstein JS. Ethical challenges of big data in public health. *PLoS Comput Biol* 2015 Feb;11(2):e1003904 [FREE Full text] [doi: [10.1371/journal.pcbi.1003904](https://doi.org/10.1371/journal.pcbi.1003904)] [Medline: [25664461](https://pubmed.ncbi.nlm.nih.gov/25664461/)]
12. Denecke K. An ethical assessment model for digital disease detection technologies. *Life Sci Soc Policy* 2017 Sep 20;13(1):16 [FREE Full text] [doi: [10.1186/s40504-017-0062-x](https://doi.org/10.1186/s40504-017-0062-x)] [Medline: [28929347](https://pubmed.ncbi.nlm.nih.gov/28929347/)]
13. Tsui F, Espino JU, Dato VM, Gesteland PH, Hutman J, Wagner MM. Technical description of RODS: a real-time public health surveillance system. *J Am Med Inform Assoc* 2003;10(5):399-408 [FREE Full text] [doi: [10.1197/jamia.M1345](https://doi.org/10.1197/jamia.M1345)] [Medline: [12807803](https://pubmed.ncbi.nlm.nih.gov/12807803/)]
14. Ogurtsova K, da Rocha Fernandes J, Huang Y, Linnenkamp U, Guariguata L, Cho N, et al. IDF diabetes atlas: global estimates for the prevalence of diabetes for 2015 and 2040. *Diabetes Res Clin Pract* 2017 Jun;128:40-50. [doi: [10.1016/j.diabres.2017.03.024](https://doi.org/10.1016/j.diabres.2017.03.024)] [Medline: [28437734](https://pubmed.ncbi.nlm.nih.gov/28437734/)]
15. Clark M. What is diabetes? In: Ogden J, editor. *Understanding Diabetes*. New Jersey, United States: John Wiley & Sons Ltd; 2004.
16. Bazaev NA, Pletenev AN, Pozhar KV. Classification of factors affecting blood glucose concentration dynamics. *Biomed Eng* 2013 Jul 18;47(2):100-103. [doi: [10.1007/s10527-013-9344-7](https://doi.org/10.1007/s10527-013-9344-7)] [Medline: [23789157](https://pubmed.ncbi.nlm.nih.gov/23789157/)]
17. Finan DA, Zisser H, Jovanović L, Bevier WC, Seborg DE. Automatic detection of stress states in type 1 diabetes subjects in ambulatory conditions. *Ind Eng Chem Res* 2010 Sep 1;49(17):7843-7848 [FREE Full text] [doi: [10.1021/ie901891c](https://doi.org/10.1021/ie901891c)] [Medline: [20953334](https://pubmed.ncbi.nlm.nih.gov/20953334/)]
18. Botsis T, Hejlesen O, Bellika JG, Hartvigsen G. Electronic disease surveillance for sensitive population groups - the diabetics case study. *Stud Health Technol Inform* 2008;136:365-370. [Medline: [18487758](https://pubmed.ncbi.nlm.nih.gov/18487758/)]
19. Rayfield EJ, Ault MJ, Keusch GT, Brothers MJ, Nechemias C, Smith H. Infection and diabetes: the case for glucose control. *Am J Med* 1982 Mar;72(3):439-450. [doi: [10.1016/0002-9343\(82\)90511-3](https://doi.org/10.1016/0002-9343(82)90511-3)] [Medline: [7036735](https://pubmed.ncbi.nlm.nih.gov/7036735/)]
20. Simonsen JR, Harjutsalo V, Järvinen A, Kirveskari J, Forsblom C, Groop P, FinnDiane Study Group. Bacterial infections in patients with type 1 diabetes: a 14-year follow-up study. *BMJ Open Diabetes Res Care* 2015;3(1):e000067 [FREE Full text] [doi: [10.1136/bmjdr-2014-000067](https://doi.org/10.1136/bmjdr-2014-000067)] [Medline: [25767718](https://pubmed.ncbi.nlm.nih.gov/25767718/)]
21. Yki-Järvinen H, Sammalkorpi K, Koivisto VA, Nikkilä EA. Severity, duration, and mechanisms of insulin resistance during acute infections. *J Clin Endocrinol Metab* 1989 Aug;69(2):317-323. [doi: [10.1210/jcem-69-2-317](https://doi.org/10.1210/jcem-69-2-317)] [Medline: [2666428](https://pubmed.ncbi.nlm.nih.gov/2666428/)]
22. Botsis T, Lai AM, Hripsak G, Palmas W, Starren JB, Hartvigsen G. Proof of concept for the role of glycemic control in the early detection of infections in diabetics. *Health Informatics J* 2012 Mar;18(1):26-35. [doi: [10.1177/1460458211428427](https://doi.org/10.1177/1460458211428427)] [Medline: [22447875](https://pubmed.ncbi.nlm.nih.gov/22447875/)]
23. Botsis T, Hartvigsen G. Exploring new directions in disease surveillance for people with diabetes: lessons learned and future plans. *Stud Health Technol Inform* 2010;160(Pt 1):466-470. [Medline: [20841730](https://pubmed.ncbi.nlm.nih.gov/20841730/)]
24. Årsand E, Walseth OA, Andersson N, Fernando R, Granberg O, Bellika JG, et al. Using blood glucose data as an indicator for epidemic disease outbreaks. *Stud Health Technol Inform* 2005;116:217-222. [Medline: [16160262](https://pubmed.ncbi.nlm.nih.gov/16160262/)]
25. Granberg O, Bellika JG, Årsand E, Hartvigsen G. Automatic infection detection system. *Stud Health Technol Inform* 2007;129(Pt 1):566-570. [Medline: [17911780](https://pubmed.ncbi.nlm.nih.gov/17911780/)]
26. Woldaregay AZ, Årsand E, Giordanengo A, Albers D, Mamykina L, Botsis T, et al. EDMON-A Wireless Communication Platform for a Real-Time Infectious Disease Outbreak Detection System Using Self-Recorded Data from People with Type 1 Diabetes. In: *The 15th Scandinavian Conference on Health Informatics*. 2017 Presented at: CHI'17; August 29, 2017; Kristiansand, Norway URL: [https://ep.liu.se/konferensartikel.aspx?series=eCP&issue=145&Article\\_No=3](https://ep.liu.se/konferensartikel.aspx?series=eCP&issue=145&Article_No=3)
27. Lauritzen JN, Årsand E, Vuurden KV, Bellika JG, Hejlesen OK, Hartvigsen G. Towards a Mobile Solution for Predicting Illness in Type 1 Diabetes Mellitus: Development of a Prediction Model for Detecting Risk of Illness in Type 1 Diabetes Prior to Symptom Onset. In: *2nd International Conference on Wireless Communication, Vehicular Technology, Information*

- Theory and Aerospace & Electronic Systems Technology. 2011 Presented at: Wireless VITAE; February 28-March 3, 2011; Chennai, India. [doi: [10.1109/wirelessvitae.2011.5940877](https://doi.org/10.1109/wirelessvitae.2011.5940877)]
28. Skrøvseth SO, Bellika JG, Godtliebsen F. Causality in scale space as an approach to change detection. *PLoS One* 2012;7(12):e52253 [FREE Full text] [doi: [10.1371/journal.pone.0052253](https://doi.org/10.1371/journal.pone.0052253)] [Medline: [23300626](https://pubmed.ncbi.nlm.nih.gov/23300626/)]
  29. Botsis T, Bellika JG, Hartvigsen G. Disease Surveillance Systems for Sensitive Population Groups. Faculty & Staff Insider - University of Washington. 2007. URL: <https://faculty.washington.edu/lober/www.isdsjournal.org/htdocs/articles/2026.pdf> [accessed 2020-07-23]
  30. Botsis T, Hejlesen O, Bellika JG, Hartvigsen G. Blood Glucose Levels as an Indicator for the Early Detection of Infections In Type-1 Diabetics. Faculty & Staff Insider - University of Washington. 2007. URL: <http://faculty.washington.edu/lober/www.isdsjournal.org/htdocs/articles/2025.pdf> [accessed 2020-07-23]
  31. Lauritzen J, Årsand E, Van Vuurden K, Hejlesen O, Hartvigsen G. Exploring illness prediction in Type 1 Diabetes Mellitus Pre-Symptom Onset. *E-health Research*. 2010. URL: <https://ehealthresearch.no/en/publications/exploring-illness-prediction-in-type-1-diabetes-mellitus-pre-symptom-onset-1> [accessed 2020-07-23]
  32. Hartvigsen G, Årsand E, Botsis T, van Vuurden K, Johansen M, Bellika JG. Reusing patient data to enhance patient empowerment and electronic disease surveillance. *J Inf Technol* 2009;7(1):4-12.
  33. Botsis T, Bellika JG, Hartvigsen G. New Directions in Electronic Disease Surveillance: Detection of Infectious Diseases during the Incubation Period. In: International Conference on eHealth. 2009 Presented at: EH'09; February 1-7, 2009; Cancun, Mexico p. 176-183. [doi: [10.1109/etelemed.2009.9](https://doi.org/10.1109/etelemed.2009.9)]
  34. Botsis T, Hejlesen O, Bellika JG, Hartvigsen G. Blood Glucose Levels as a Sensor for Early Detection of Infection in Type-1 Diabetics. Faculty & Staff Insider - University of Washington. 2007. URL: <http://faculty.washington.edu/lober/www.isdsjournal.org/htdocs/articles/2025.pdf> [accessed 2020-07-23]
  35. Botsis T, Hejlesen O, Bellika JG, Hartvigsen G. Disease surveillance systems for diabetics. In: Telemedicine and e-Health. 2008 Mar Presented at: The American Telemedicine Association Annual International Meeting and Exposition; 2008; Seattle p. 108-108. [doi: [10.1089/tmj.2008.9983.suppl](https://doi.org/10.1089/tmj.2008.9983.suppl)]
  36. Botsis T, Hejlesen O, Bellika JG, Hartvigsen G. Electronic infectious disease surveillance systems for diabetics. 2008 Presented at: International Conference on Advanced Technologies & Treatments for Diabetes; Prague, Czech Republic.
  37. Brealey D, Singer M. Hyperglycemia in critical illness: a review. *J Diabetes Sci Technol* 2009 Nov 1;3(6):1250-1260 [FREE Full text] [doi: [10.1177/1932296809000300604](https://doi.org/10.1177/1932296809000300604)] [Medline: [20144378](https://pubmed.ncbi.nlm.nih.gov/20144378/)]
  38. Marik PE, Bellomo R. Stress hyperglycemia: an essential survival response!. *Crit Care* 2013 Mar 6;17(2):305 [FREE Full text] [doi: [10.1186/cc12514](https://doi.org/10.1186/cc12514)] [Medline: [23470218](https://pubmed.ncbi.nlm.nih.gov/23470218/)]
  39. Årsand E, Muzny M, Bradway M, Muzik J, Hartvigsen G. Performance of the first combined smartwatch and smartphone diabetes diary application study. *J Diabetes Sci Technol* 2015 May;9(3):556-563 [FREE Full text] [doi: [10.1177/1932296814567708](https://doi.org/10.1177/1932296814567708)] [Medline: [25591859](https://pubmed.ncbi.nlm.nih.gov/25591859/)]
  40. Features. Spike App. URL: <https://spike-app.com/#features1> [accessed 2019-09-20]
  41. Gillespie SJ, Kulkarni KD, Daly AE. Using carbohydrate counting in diabetes clinical practice. *J Am Diet Assoc* 1998 Aug;98(8):897-905. [doi: [10.1016/S0002-8223\(98\)00206-5](https://doi.org/10.1016/S0002-8223(98)00206-5)] [Medline: [9710660](https://pubmed.ncbi.nlm.nih.gov/9710660/)]
  42. Rasoulzadeh V, Erkus EC, Yogurt TA, Ulusoy I, Zergeroğlu SA. A comparative stationarity analysis of EEG signals. *Ann Oper Res* 2016 Apr 26;258(1):133-157 [FREE Full text] [doi: [10.1007/s10479-016-2187-3](https://doi.org/10.1007/s10479-016-2187-3)]
  43. Azami H, Mohammadi K, Bozorgtabar B. An improved signal segmentation using moving average and Savitzky-Golay filter. *J Signal Inf Process* 2012;03(01):39-44. [doi: [10.4236/jsip.2012.31006](https://doi.org/10.4236/jsip.2012.31006)]
  44. Botev ZI, Grotowski JF, Kroese DP. Kernel density estimation via diffusion. *Ann Statist* 2010 Oct;38(5):2916-2957. [doi: [10.1214/10-aos799](https://doi.org/10.1214/10-aos799)]
  45. Gramacki A. Nonparametric density estimation. In: Nonparametric Kernel Density Estimation and Its Computational Aspects. Cham, Switzerland: Springer International Publishing; 2018:7-24.
  46. Gramacki A. Bandwidth selectors for kernel density estimation. In: Nonparametric Kernel Density Estimation and Its Computational Aspects. Cham, Switzerland: Springer International Publishing; 2018:63-83.
  47. Heidenreich N, Schindler A, Sperlich S. Bandwidth selection for kernel density estimation: a review of fully automatic selectors. *AStA Adv Stat Anal* 2013 Jun 2;97(4):403-433 [FREE Full text] [doi: [10.1007/s10182-013-0216-y](https://doi.org/10.1007/s10182-013-0216-y)]
  48. Gramacki A. Kernel density estimation. In: Nonparametric Kernel Density Estimation and Its Computational Aspects. Cham, Switzerland: Springer International Publishing; 2018:25-62.
  49. Zdravko B. Kernel Density Estimator. MathWorks. 2015. URL: <https://www.mathworks.com/matlabcentral/fileexchange/14034-kernel-density-estimator> [accessed 2019-06-22]
  50. Rossini AJ. 'Applied smoothing techniques for data analysis: the kernel approach with s-plus illustrations' by Adrian W Bowman and Adelchi Azzalini. *Comput Stat* 2000 Sep 11;15(2):301-302 [FREE Full text] [doi: [10.1007/s001800000033](https://doi.org/10.1007/s001800000033)]
  51. Bowman AW, Azzalini A. Applied Smoothing Techniques for Data Analysis: The Kernel Approach With S-Plus Illustrations. Oxford, UK: OUP Oxford; 1997.
  52. Yi C. Bivariate Kernel Density Estimation (V2.1). MathWorks. 2013. URL: [https://la.mathworks.com/matlabcentral/fileexchange/19280-bivariate-kernel-density-estimation-v2-1?s\\_tid=prof\\_contriblnk](https://la.mathworks.com/matlabcentral/fileexchange/19280-bivariate-kernel-density-estimation-v2-1?s_tid=prof_contriblnk) [accessed 2019-06-22]

53. Yamaguchi M, Kaseda C, Yamazaki K, Kobayashi M. Prediction of blood glucose level of type 1 diabetics using response surface methodology and data mining. *Med Biol Eng Comput* 2006 Jun;44(6):451-457. [doi: [10.1007/s11517-006-0049-x](https://doi.org/10.1007/s11517-006-0049-x)] [Medline: [16937196](https://pubmed.ncbi.nlm.nih.gov/16937196/)]
54. van Herpe T, Pluymers B, Espinoza M, van den Berghe G, de Moor BA. A minimal model for glycemia control in critically ill patients. *Conf Proc IEEE Eng Med Biol Soc* 2006;2006:5432-5435 [FREE Full text] [doi: [10.1109/IEMBS.2006.260613](https://doi.org/10.1109/IEMBS.2006.260613)] [Medline: [17946700](https://pubmed.ncbi.nlm.nih.gov/17946700/)]
55. Waldhäusl WK, Bratusch-Marrain P, Komjati M, Breitenecker F, Troch I. Blood glucose response to stress hormone exposure in healthy man and insulin dependent diabetic patients: prediction by computer modeling. *IEEE Trans Biomed Eng* 1992 Aug;39(8):779-790 [FREE Full text] [doi: [10.1109/10.148386](https://doi.org/10.1109/10.148386)] [Medline: [1354649](https://pubmed.ncbi.nlm.nih.gov/1354649/)]
56. O'Shea J. Digital disease detection: a systematic review of event-based internet biosurveillance systems. *Int J Med Inform* 2017 May;101:15-22 [FREE Full text] [doi: [10.1016/j.ijmedinf.2017.01.019](https://doi.org/10.1016/j.ijmedinf.2017.01.019)] [Medline: [28347443](https://pubmed.ncbi.nlm.nih.gov/28347443/)]
57. Mizock BA. Alterations in carbohydrate metabolism during stress: a review of the literature. *Am J Med* 1995 Jan;98(1):75-84 [FREE Full text] [doi: [10.1016/S0002-9343\(99\)80083-7](https://doi.org/10.1016/S0002-9343(99)80083-7)] [Medline: [7825623](https://pubmed.ncbi.nlm.nih.gov/7825623/)]
58. Wagner MM, Moore AW, Aryel RM, editors. *Handbook of Biosurveillance*. New York, USA: Elsevier; 2006.
59. Event-Based Surveillance. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/globalhealth/healthprotection/gddopscenter/how.html> [accessed 2019-06-26]
60. World Health Organization. *A Guide to Establishing Event-Based Surveillance*. Geneva, Switzerland: World Health Organization; 2008.
61. Coucheron S, Woldaregay AZ, Årsand E, Botsis T, Hartvigsen G. EDMON - A System Architecture for Real-Time Infection Monitoring and Outbreak Detection Based on Self-Recorded Data from People with Type 1 Diabetes: System Design and Prototype Implementation. In: *The 17th Scandinavian Conference on Health Informatics*. 2019 Presented at: CHI'19; November 12, 2019; Oslo, Norway p. 37-44 URL: <https://ep.liu.se/ecp/161/007/ecp19161007.pdf>
62. Tax DM. *One-Class Classification: Concept Learning in the Absence of Counter-Examples*. South Holland, Netherlands: Technische Universiteit Delft; 2002.
63. Mazhelis O. One-class classifiers: a review and analysis of suitability in the context of mobile-masquerader detection. *S Afr Comput J* 2006;36:29-48 [FREE Full text]
64. Oviedo S, Vehí J, Calm R, Armengol J. A review of personalized blood glucose prediction strategies for T1DM patients. *Int J Numer Method Biomed Eng* 2017 Jun;33(6):---. [doi: [10.1002/cnm.2833](https://doi.org/10.1002/cnm.2833)] [Medline: [27644067](https://pubmed.ncbi.nlm.nih.gov/27644067/)]
65. Woldaregay AZ, Årsand E, Walderhaug S, Albers D, Mamykina L, Botsis T, et al. Data-driven modeling and prediction of blood glucose dynamics: machine learning applications in type 1 diabetes. *Artif Intell Med* 2019 Jul;98:109-134 [FREE Full text] [doi: [10.1016/j.artmed.2019.07.007](https://doi.org/10.1016/j.artmed.2019.07.007)] [Medline: [31383477](https://pubmed.ncbi.nlm.nih.gov/31383477/)]
66. Dunning T, Friedman E. In: Loukides M, editor. *Practical Machine Learning: A New Look at Anomaly Detection*. New York, USA: O'Reilly Media Inc; 2018.
67. Chandola V, Banerjee A, Kumar V. Anomaly detection: A survey. *ACM Comput Surv* 2009 Jul;41(3):1-58 [FREE Full text] [doi: [10.1145/1541880.1541882](https://doi.org/10.1145/1541880.1541882)]
68. Pimentel MA, Clifton DA, Clifton L, Tarassenko L. A review of novelty detection. *Sig Process* 2014 Jun;99:215-249 [FREE Full text] [doi: [10.1016/j.sigpro.2013.12.026](https://doi.org/10.1016/j.sigpro.2013.12.026)]
69. Haslett J, Haslett SJ. The three basic types of residuals for a linear model. *Int Statistical Rev* 2007 Apr;75(1):1-24 [FREE Full text] [doi: [10.1111/j.1751-5823.2006.00001.x](https://doi.org/10.1111/j.1751-5823.2006.00001.x)]
70. Harvey AC, Koopman SJ. Diagnostic checking of unobserved-components time series models. *J Bus Econ Stat* 1992 Oct;10(4):377-389. [doi: [10.1080/07350015.1992.10509913](https://doi.org/10.1080/07350015.1992.10509913)]
71. Yu Y, Zhu Y, Li S, Wan D. Time series outlier detection based on sliding window prediction. *Math Probl Eng* 2014;2014:1-14 [FREE Full text] [doi: [10.1155/2014/879736](https://doi.org/10.1155/2014/879736)]
72. Woldaregay AZ, Årsand E, Botsis T, Hartvigsen G. An early infectious disease outbreak detection mechanism based on self-recorded data from people with diabetes. *Stud Health Technol Inform* 2017;245:619-623. [Medline: [29295170](https://pubmed.ncbi.nlm.nih.gov/29295170/)]
73. Vovk V, Gammernan A, Shafer G. *Conformal prediction*. In: *Algorithmic Learning in a Random World*. Boston, MA: Springer US; 2005:17-51.
74. Vovk V. The basic conformal prediction framework. In: *Conformal Prediction for Reliable Machine Learning*. New York, USA: Morgan Kaufmann; 2014:3-19.
75. Vovk V. Beyond the basic conformal prediction framework. In: *Conformal Prediction for Reliable Machine Learning*. New York, USA: Morgan Kaufmann; 2014:21-46.
76. James S. *The Efficiency of Conformal Predictors for Anomaly Detection*. Royal Holloway, University of London: Research. 2016. URL: [https://pure.royalholloway.ac.uk/portal/en/publications/the-efficiency-of-conformal-predictors-for-anomaly-detection\(d68cdce7-7d84-414b-9498-0c8539bb57b8\)/export.html](https://pure.royalholloway.ac.uk/portal/en/publications/the-efficiency-of-conformal-predictors-for-anomaly-detection(d68cdce7-7d84-414b-9498-0c8539bb57b8)/export.html) [accessed 2020-07-23]
77. Laxhammar R, Falkman G. Inductive conformal anomaly detection for sequential detection of anomalous sub-trajectories. *Ann Math Artif Intell* 2013 Sep 20;74(1-2):67-94 [FREE Full text] [doi: [10.1007/s10472-013-9381-7](https://doi.org/10.1007/s10472-013-9381-7)]
78. Shafer G, Vovk V. A tutorial on conformal prediction. *J Mach Learn Res* 2008 Jun 1;9(-):371-421 [FREE Full text] [doi: [10.5555/1390681.1390693](https://doi.org/10.5555/1390681.1390693)]



79. Smith J, Nouretdinov I, Craddock R, Offer C, Gammerman A. Anomaly Detection of Trajectories with Kernel Density Estimation by Conformal Prediction. In: IFIP International Conference on Artificial Intelligence Applications and Innovations. 2014 Presented at: AIAI'14; September 19-21, 2014; Rhodos, Greece URL: [https://doi.org/10.1007/978-3-662-44722-2\\_29](https://doi.org/10.1007/978-3-662-44722-2_29) [doi: [10.1007/978-3-662-44722-2\\_29](https://doi.org/10.1007/978-3-662-44722-2_29)]
80. Laxhammar R. Conformal Anomaly Detection: Detecting Abnormal Trajectories in Surveillance Applications. Skövde: School of Informatics, University of Skövde; 2014.
81. Mehrotra K, Mohan CK, Huang H. Anomaly Detection Principles and Algorithms. New York, USA: Springer; 2017.
82. Zhao Z, Mehrotra KG, Mohan CK. Ensemble Algorithms for Unsupervised Anomaly Detection. In: International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems. 2015 Presented at: IEA/AIE'15; June 10-12, 2015; Seoul, Korea (Republic of) URL: <https://link.springer.com/content/pdf/10.1007%2F978-3-319-19066-2.pdf> [doi: [10.1007/978-3-319-19066-2\\_50](https://doi.org/10.1007/978-3-319-19066-2_50)]
83. Breunig MM, Kriegel HP, Ng RT, Sanders J. LOF: Identifying Density-Based Local Outliers. In: Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data. 2000 Presented at: SIGMOD'00; May 16-18, 2000; Dallas, Texas, USA. [doi: [10.1145/342009.335388](https://doi.org/10.1145/342009.335388)]
84. Tang J, Chen Z, Fu AW, Cheung DW. Enhancing Effectiveness of Outlier Detections for Low Density Patterns. In: Pacific-Asia Conference on Knowledge Discovery and Data Mining. 2002 Presented at: PAKDD'02; May 6-8, 2002; Taipei, Taiwan. [doi: [10.1007/3-540-47887-6\\_53](https://doi.org/10.1007/3-540-47887-6_53)]
85. Nguyen HV, Nguyen TT, Nguyen QU. Sequential Ensemble Method for Unsupervised Anomaly Detection. In: 9th International Conference on Knowledge and Systems Engineering (KSE). 2017 Presented at: KSE'17; October 19-21, 2017; Hue, Vietnam. [doi: [10.1109/kse.2017.8119437](https://doi.org/10.1109/kse.2017.8119437)]
86. Goldstein M, Uchida S. A comparative evaluation of unsupervised anomaly detection algorithms for multivariate data. PLoS One 2016;11(4):e0152173 [FREE Full text] [doi: [10.1371/journal.pone.0152173](https://doi.org/10.1371/journal.pone.0152173)] [Medline: [27093601](https://pubmed.ncbi.nlm.nih.gov/27093601/)]
87. Irigoien I, Sierra B, Arenas C. Towards application of one-class classification methods to medical data. ScientificWorldJournal 2014;2014:730712 [FREE Full text] [doi: [10.1155/2014/730712](https://doi.org/10.1155/2014/730712)] [Medline: [24778600](https://pubmed.ncbi.nlm.nih.gov/24778600/)]
88. Tax DM, Duin RP. Support vector data description. Mach Learn 2004 Jan;54(1):45-66 [FREE Full text] [doi: [10.1023/b:mach.0000008084.60811.49](https://doi.org/10.1023/b:mach.0000008084.60811.49)]
89. Schölkopf B, Williamson RC, Smola AJ, Shawe-Taylor J, Platt JC. Support Vector Method for Novelty Detection. In: The 12th International Conference on Neural Information Processing. 1999 Presented at: NIPS'99; November 29-December 4, 1999; Denver, Colorado, USA URL: <https://dl.acm.org/doi/10.5555/3009657.3009740> [doi: [10.5555/3009657.3009740](https://doi.org/10.5555/3009657.3009740)]
90. Juszczak P, Tax DM, Pe kalska E, Duin RP. Minimum spanning tree based one-class classifier. Neurocomputing 2009 Mar;72(7-9):1859-1869 [FREE Full text] [doi: [10.1016/j.neucom.2008.05.003](https://doi.org/10.1016/j.neucom.2008.05.003)]
91. Khan SS, Madden MG. One-class classification: taxonomy of study and review of techniques. Knowl Eng Rev 2014 Jan 24;29(3):345-374 [FREE Full text] [doi: [10.1017/s026988891300043x](https://doi.org/10.1017/s026988891300043x)]
92. Wang B, Mao Z. One-class classifiers ensemble based anomaly detection scheme for process control systems. T I Meas Control 2017 Sep 21;40(12):3466-3476 [FREE Full text] [doi: [10.1177/0142331217724508](https://doi.org/10.1177/0142331217724508)]
93. Ahmad S, Lavin A, Purdy S, Agha Z. Unsupervised real-time anomaly detection for streaming data. Neurocomputing 2017 Nov;262:134-147 [FREE Full text] [doi: [10.1016/j.neucom.2017.04.070](https://doi.org/10.1016/j.neucom.2017.04.070)]
94. Yu E, Parekh P. A bayesian ensemble for unsupervised anomaly detection. arXiv 2016:- preprint [FREE Full text]
95. Shoemaker L, Hall LO. Anomaly Detection Using Ensembles. In: International Workshop on Multiple Classifier Systems. 2011 Presented at: MCS'11; June 15-17, 2011; Naples, Italy URL: [https://doi.org/10.1007/978-3-642-21557-5\\_3](https://doi.org/10.1007/978-3-642-21557-5_3) [doi: [10.1007/978-3-642-21557-5\\_3](https://doi.org/10.1007/978-3-642-21557-5_3)]
96. Yeng PK, Woldaregay AZ, Solvoll T, Hartvigsen G. Cluster detection mechanisms for syndromic surveillance systems: systematic review and framework development. JMIR Public Health Surveill 2020 May 26;6(2):e11512 [FREE Full text] [doi: [10.2196/11512](https://doi.org/10.2196/11512)] [Medline: [32357126](https://pubmed.ncbi.nlm.nih.gov/32357126/)]
97. Shi Z, Pun-Cheng L. Spatiotemporal data clustering: a survey of methods. Int J Geogr Inf Sci 2019 Feb 28;8(3):112 [FREE Full text] [doi: [10.3390/ijgi8030112](https://doi.org/10.3390/ijgi8030112)]
98. Duangchaemkarn K, Chaovatur V, Wiwatanadate P, Boonchieng E. Symptom-based data preprocessing for the detection of disease outbreak. Conf Proc IEEE Eng Med Biol Soc 2017 Jul;2017:2614-2617 [FREE Full text] [doi: [10.1109/EMBC.2017.8037393](https://doi.org/10.1109/EMBC.2017.8037393)] [Medline: [29060435](https://pubmed.ncbi.nlm.nih.gov/29060435/)]
99. Grubestic TH, Matisziw TC. On the use of ZIP codes and ZIP code tabulation areas (ZCTAs) for the spatial analysis of epidemiological data. Int J Health Geogr 2006 Dec 13;5:58 [FREE Full text] [doi: [10.1186/1476-072X-5-58](https://doi.org/10.1186/1476-072X-5-58)] [Medline: [17166283](https://pubmed.ncbi.nlm.nih.gov/17166283/)]
100. Glatman-Freedman A, Kaufman Z, Kopel E, Bassal R, Taran D, Valinsky L, et al. Near real-time space-time cluster analysis for detection of enteric disease outbreaks in a community setting. J Infect 2016 Aug;73(2):99-106. [doi: [10.1016/j.jinf.2016.04.038](https://doi.org/10.1016/j.jinf.2016.04.038)] [Medline: [27311747](https://pubmed.ncbi.nlm.nih.gov/27311747/)]
101. Woldaregay AZ, Issom D, Henriksen A, Marttila H, Mikalsen M, Pfuhl G, et al. Motivational factors for user engagement with mhealth apps. Stud Health Technol Inform 2018;249:151-157. [Medline: [29866972](https://pubmed.ncbi.nlm.nih.gov/29866972/)]

102. Klingler C, Silva DS, Schuermann C, Reis AA, Saxena A, Strech D. Ethical issues in public health surveillance: a systematic qualitative review. *BMC Public Health* 2017 Apr 4;17(1):295 [FREE Full text] [doi: [10.1186/s12889-017-4200-4](https://doi.org/10.1186/s12889-017-4200-4)] [Medline: [28376752](https://pubmed.ncbi.nlm.nih.gov/28376752/)]
103. Geneviève LD, Martani A, Wangmo T, Paolotti D, Koppeschaar C, Kjelsø C, et al. Participatory disease surveillance systems: ethical framework. *J Med Internet Res* 2019 May 23;21(5):e12273 [FREE Full text] [doi: [10.2196/12273](https://doi.org/10.2196/12273)] [Medline: [31124466](https://pubmed.ncbi.nlm.nih.gov/31124466/)]
104. Rumbold JM, Pierscionek B. The effect of the general data protection regulation on medical research. *J Med Internet Res* 2017 Feb 24;19(2):e47 [FREE Full text] [doi: [10.2196/jmir.7108](https://doi.org/10.2196/jmir.7108)] [Medline: [28235748](https://pubmed.ncbi.nlm.nih.gov/28235748/)]
105. Hintze M. Viewing the GDPR through a de-identification lens: a tool for compliance, clarification, and consistency. *Int Data Priv Law* 2017;8(1):86-101 [FREE Full text] [doi: [10.1093/idpl/ix020](https://doi.org/10.1093/idpl/ix020)]
106. Sun W, Cai Z, Li Y, Liu F, Fang S, Wang G. Security and privacy in the medical internet of things: a review. *Secur Commun Netw* 2018;2018:1-9 [FREE Full text] [doi: [10.1155/2018/5978636](https://doi.org/10.1155/2018/5978636)]

## Abbreviations

**CGM:** continuous glucose monitoring  
**CRH:** counterregulatory hormone  
**GDPR:** General Data Protection Regulation  
**GPS:** global positioning system  
**IoT:** Internet of Things  
**SMBG:** self-monitoring of blood glucose

*Edited by G Eysenbach; submitted 26.03.20; peer-reviewed by M Nomali, S Sarbadhikari; comments to author 15.04.20; revised version received 06.06.20; accepted 11.06.20; published 12.08.20*

*Please cite as:*

Woldaregay AZ, Launonen IK, Årsand E, Albers D, Holubová A, Hartvigsen G

*Toward Detecting Infection Incidence in People With Type 1 Diabetes Using Self-Recorded Data (Part 1): A Novel Framework for a Personalized Digital Infectious Disease Detection System*

*J Med Internet Res* 2020;22(8):e18911

URL: <https://www.jmir.org/2020/8/e18911>

doi: [10.2196/18911](https://doi.org/10.2196/18911)

PMID:

©Ashenafi Zebene Woldaregay, Ilkka Kalervo Launonen, Eirik Årsand, David Albers, Anna Holubová, Gunnar Hartvigsen. Originally published in the *Journal of Medical Internet Research* (<http://www.jmir.org>), 12.08.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the *Journal of Medical Internet Research*, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

# Appendix 1: Comparative Analysis of Parameters of Blood Glucose Dynamics with and without Infection Incidences

## I. Pre-Infection, Infection, and Post-Infection Week Analysis

The analysis is done based on the daily average blood glucose levels, total insulin including both bolus and basal whenever possible and total carbohydrate consumptions. As shown in the **Table 1** and **Figure 1-5**, the analysis has demonstrated that during all the infection incidences,

- Blood glucose is too resistant to drop and remains elevated in all the cases.
- There is high injection of insulin compared to the pre and post infection week.
- Carbohydrate is significantly reduced as compared to the pre and post infection week.

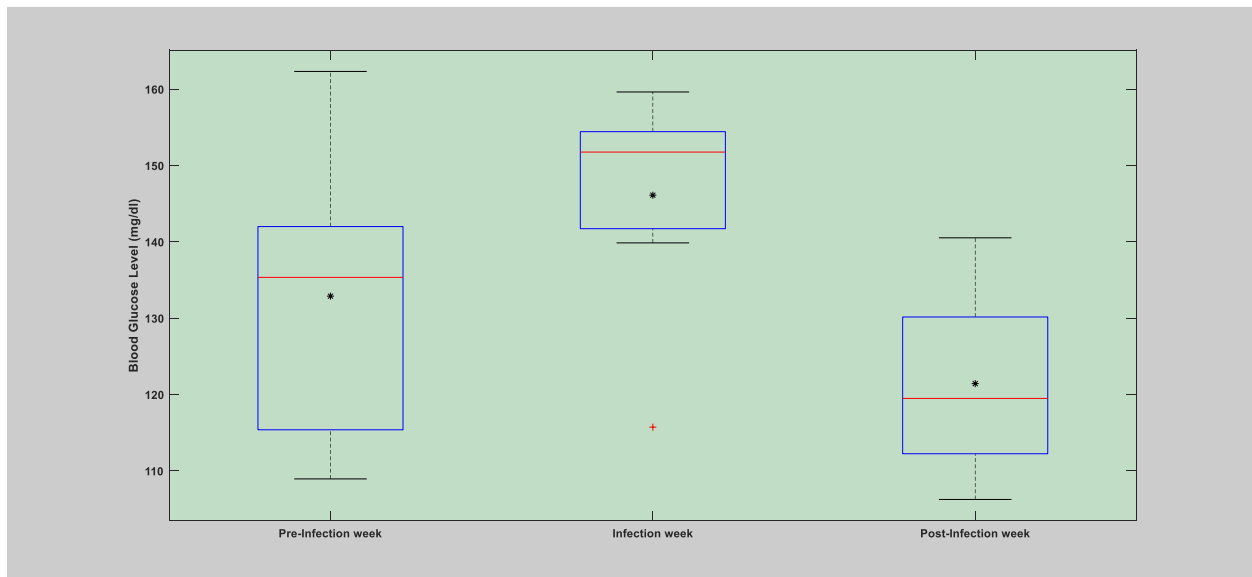
Detailed demonstration is presented in the following section.

**Table 1:** Mean and standard deviation of BG levels, total insulin (bolus), and total carbohydrate during the pre-infection week, infection week and post-infection week.

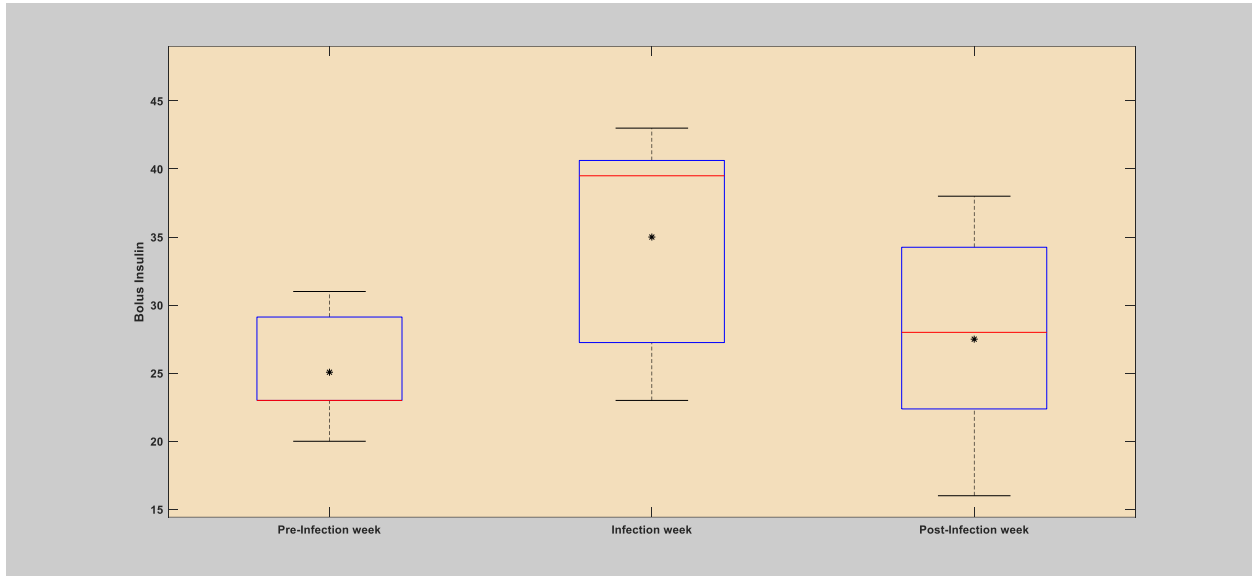
The first case of infection (flu)			
Parameter	Pre-infection week ( Mean (SD))	Infection week ( Mean (SD))	Post-infection week ( Mean (SD))
Daily			
Blood Glucose (mg/dl)	130.74 (16.89)	141.95 (14.37)	119.16 (7.39)
Total insulin (bolus)	23.39 (4.91)	35.30 (6.11)	21.32 (4.61)
Carbohydrate (grams)	241.11 (57.27)	178.80 (65.69)	241.18 (37.63)
Hourly			
Blood Glucose (mg/dl)	134.24 (32.67)	147.38 (36.89)	129.77 (39.13)
Total insulin (bolus)	0.99 (1.87)	1.51 (2.32)	0.89 (1.77)
Carbohydrate (grams)	10.16 (16.50)	7.32 (13.76)	10.20 (15.84)
The second case of infection (flu)			
Daily			
Blood Glucose (mg/dl)	143.01 (19.53)	155.36 (21.99)	126.17 (11.70)
Total insulin (bolus)	28.07 ( 8.85)	41.07 (9.44)	25.36 (6.93)
Carbohydrate (grams)	190.14 (43.93)	161.14 (58.43)	214.57 (34.66)
Hourly			
Blood Glucose (mg/dl)	147.51 (40.75)	162.32 (39.49)	130.85 (36.49)
Total insulin (bolus)	1.14 (1.92)	1.68 (2.48)	0.98 (1.94)
Carbohydrate (grams)	8.3517 (14.68)	6.88 (14.25)	8.39 (14.56)
The third case of infection (flu)			
Daily			
Blood Glucose (mg/dl)	136.93 (18.58)	144.12 (20.30)	134.18 (11.96)
Total insulin (bolus)	20.08 ( 5.44)	31.50 (10.84)	22.83 (3.86)

Carbohydrate (grams)	178.0 (45.87)	144.83 (37.63)	195.83 (42.59)
	Hourly		
Blood Glucose (mg/dl)	143.30 (40.51)	149.84 (32.90)	139.77 (39.61)
Total insulin (bolus)	1.00 (1.77)	1.50 (2.17)	0.93 (1.62)
Carbohydrate (grams)	7.71 (14.31)	6.48 (12.82)	7.28 (13.42)
<b>The fourth case of infection (flu)</b>			
	Daily		
Blood Glucose (mg/dl)	157.74 (31.12)	161.34 (19.88)	138.57 (19.83)
Total insulin (bolus)	24.43 (5.26)	32.14 (7.01)	29.29 (5.22)
Carbohydrate (grams)	199.06 (53.45)	167.04 (44.94)	226.07 (18.23)
	Hourly		
Blood Glucose (mg/dl)	151.57 (51.15)	157.68 (52.43)	142.16 (47.79)
Total insulin (bolus)	0.96 (2.60)	1.39 (2.97)	1.19 (3.24)
Carbohydrate (grams)	8.38 (20.25)	6.78 (18.52)	9.78 (22.57)
<b>The fifth case of infection (flu)</b>			
	Daily		
Blood Glucose (mg/dl)	135.21 (14.58)	139.88 (15.54)	122.87 (14.49)
Insulin (bolus)	32.80 (4.59)	40.37 (8.31)	33.36 (7.94)
Insulin (basal)	19.20 (1.21)	20.42 (2.06)	18.68 (1.56)
Total Insulin	52.33 (5.14)	61.21 (8.26)	52.46 (8.47)
	Hourly		
Blood Glucose (mg/dl)	134.23 (34.16)	144.09 (44.25)	122.12 (35.99)
Insulin (bolus)	1.36 (2.58)	1.76 (2.82)	1.45 (2.63)

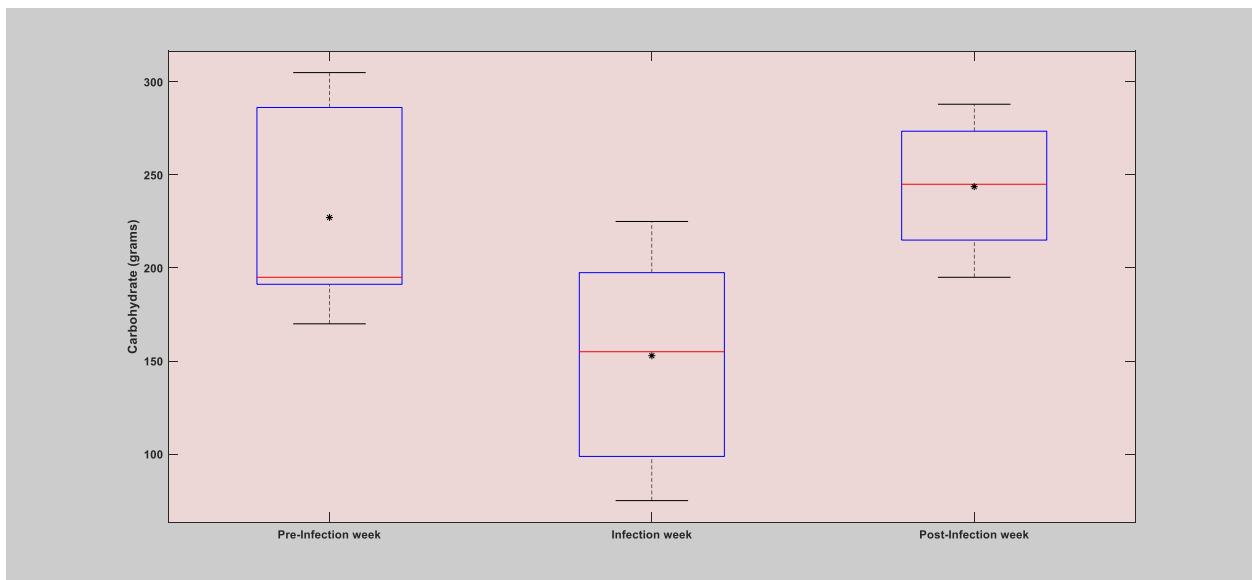
## 1. The First Case of infection (flu)



- a) Comparison of Blood Glucose levels during pre-infection week, infection week and post-infection week. As can be seen, the blood glucose is elevated during the infection week as compared to the pre and post-infection week.



- b) Comparison of Insulin (bolus) intake during pre-infection week, infection week and post-infection week. As can be seen, the amount of insulin (bolus) intake is elevated during the infection week as compared to the pre and post-infection week.

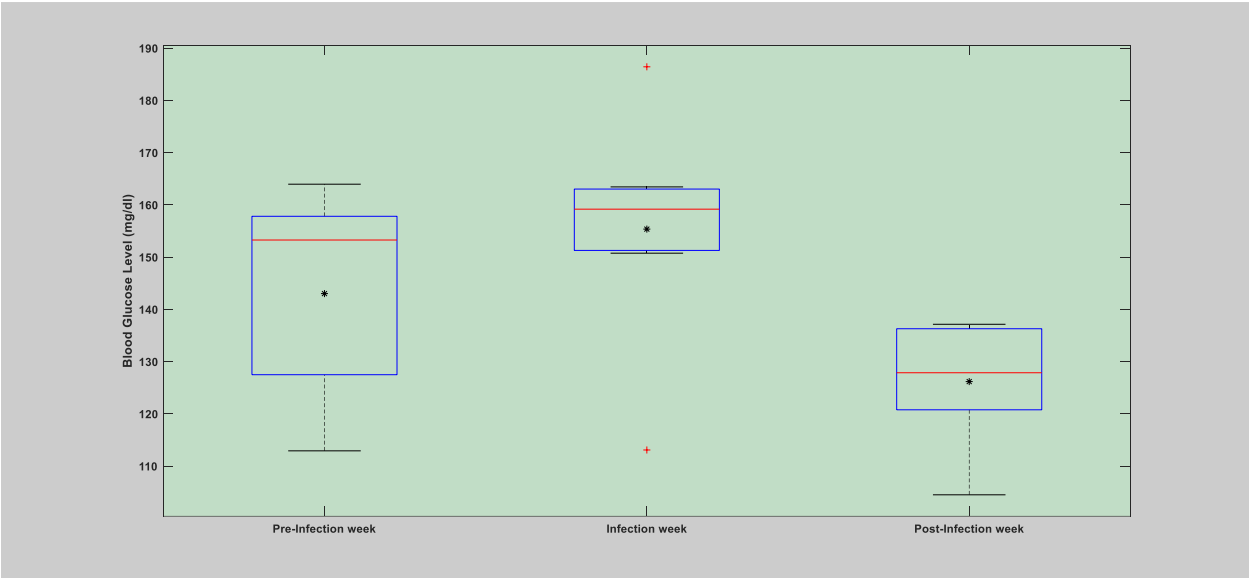


- c) Comparison of Carbohydrate (grams) intake during pre-infection week, infection week and post-infection week. As can be seen, the amount of Carbohydrate (grams) intake is significantly reduced during the infection week as compared to the pre and post-infection week.
- d) **Table 2:** Mean Percentage Change between pre-infection week versus infection week and post-infection week and infection week. For further reference, see **Table 1** Above.

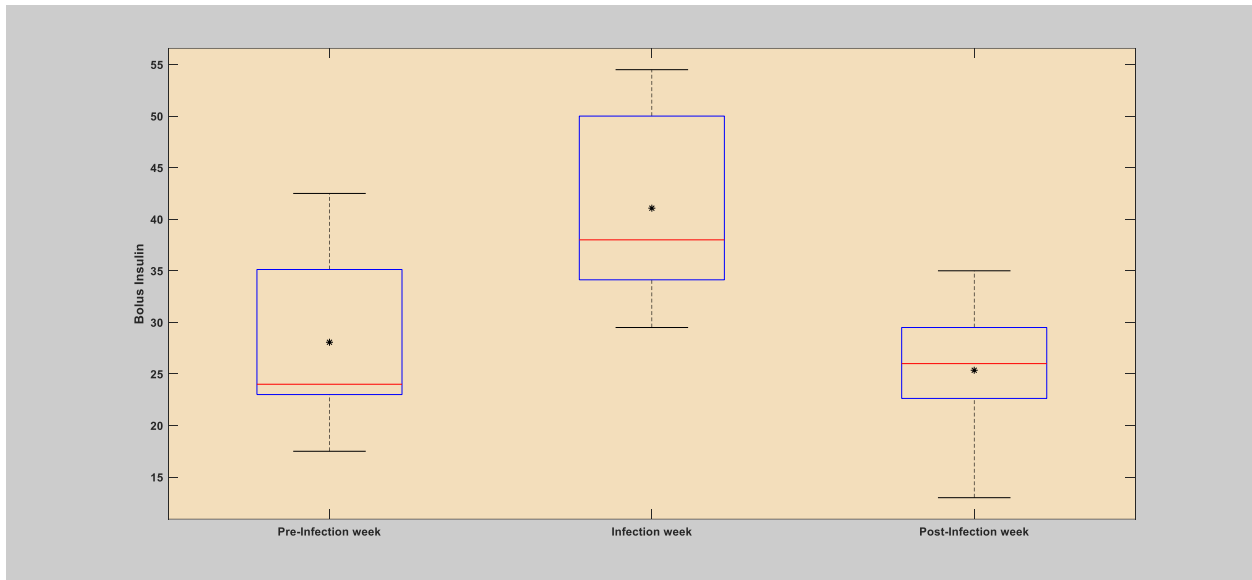
Parameters	Pre-infection Vs. Infection week	Post-infection week Vs. Infection week
<b>Blood Glucose</b>	<b>8.57%</b>	<b>19.12%</b>
<b>Insulin (bolus)</b>	<b>50.93%</b>	<b>65.59%</b>
<b>Carbohydrate</b>	<b>-25.84%</b>	<b>-25.87%</b>
<b>Hourly</b>		
<b>Blood Glucose</b>	<b>9.79%</b>	<b>13.57%</b>
<b>Insulin (bolus)</b>	<b>52.53%</b>	<b>70.00%</b>
<b>Carbohydrate</b>	<b>-27.95%</b>	<b>-28.24%</b>

Figure 1: Analysis of pre-infection week, infection week, and post-infection week based on the first patient year. Figure (a) depicts the blood glucose levels during these weeks. Figure (b) depicts the amount of insulin (bolus) injected during these weeks. Figure (c) depicts the amount of carbohydrate consumed in grams during these weeks. Table 2 shows the mean percentage change between these weeks. In all the figures, the asterisk shows the mean value and the red line depicts the median value for the week.

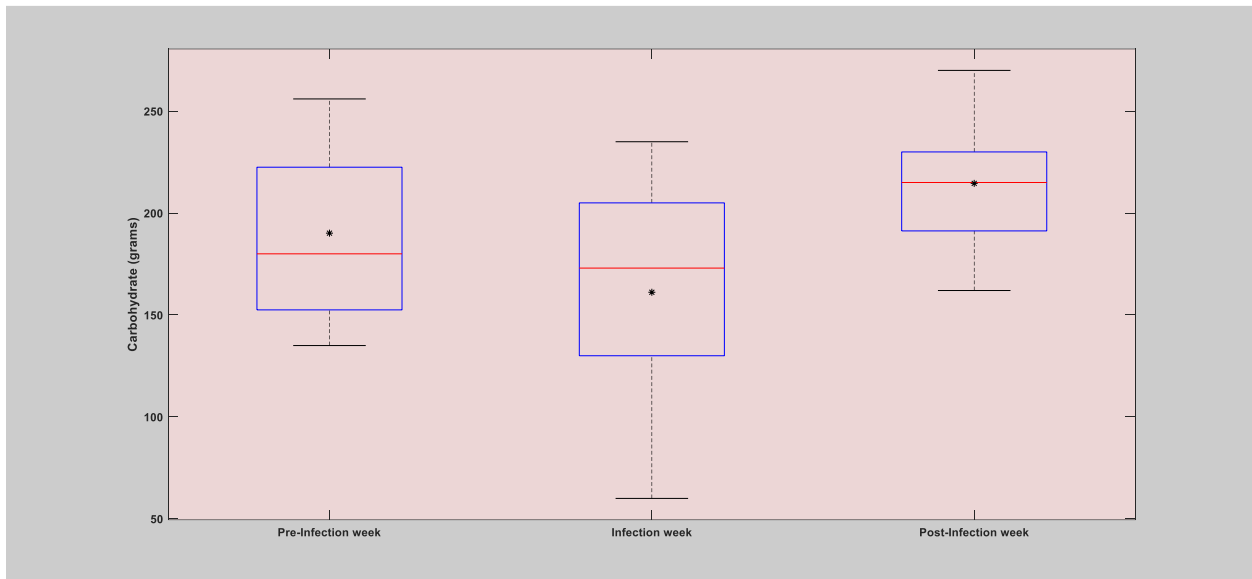
## 2. The Second Case of Infection (flu)



- a) Comparison of Blood Glucose levels during pre-infection week, infection week and post-infection week. As can be seen, the blood glucose is elevated during the infection week as compared to the pre and post-infection week.



b) Comparison of Insulin (bolus) intake during pre-infection week, infection week and post-infection week. As can be seen, the amount of insulin (bolus) intake is elevated during the infection week as compared to the pre and post-infection week.



c) Comparison of Carbohydrate (grams) intake during pre-infection week, infection week and post-infection week. As can be seen, the amount of Carbohydrate (grams) intake is significantly reduced during the infection week as compared to the pre and post-infection week.

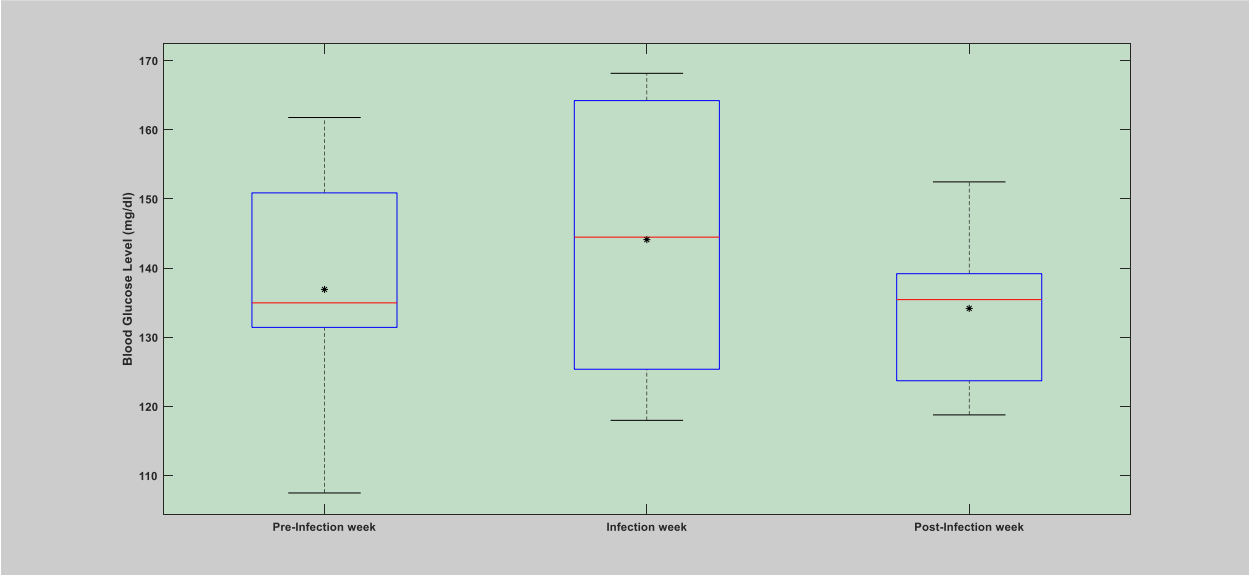
e) Table 3: Mean Percentage Change between pre-infection week versus infection week and post-infection week and infection week. For further reference, see Table 1 Above.

Parameters	Pre-infection Vs. Infection week	Post-infection week Vs. Infection week
Blood Glucose (mg/dl)	8.63%	23.13%
Insulin (bolus)	46.31%	61.94%

<b>Carbohydrate (grams)</b>	<b>-15.25%</b>	<b>-24.90%</b>
<b>Hourly</b>		
<b>Blood Glucose (mg/dl)</b>	<b>10.04%</b>	<b>24.05%</b>
<b>Insulin (bolus)</b>	<b>47.13%</b>	<b>70.87%</b>
<b>Carbohydrate (grams)</b>	<b>-17.59%</b>	<b>-18.0%</b>

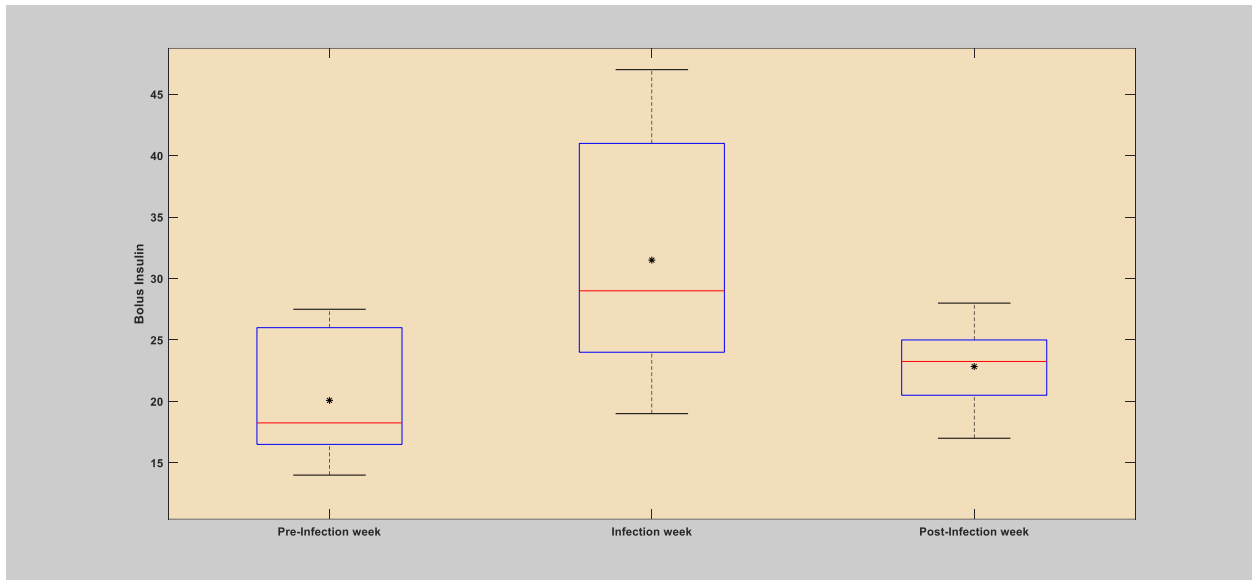
**Figure 2:** Analysis of pre-infection week, infection week, and post-infection week based on the second patient year. Figure (a) depicts the blood glucose levels during these weeks. Figure (b) depicts the amount of insulin (bolus) injected during these weeks. Figure (c) depicts the amount of carbohydrate consumed in grams during these weeks. Table 3 shows the mean percentage change between these weeks. In all the figures, the asterisk shows the mean value and the red line depicts the median value for the week.

### 3. The Third Case of Infection (flu)

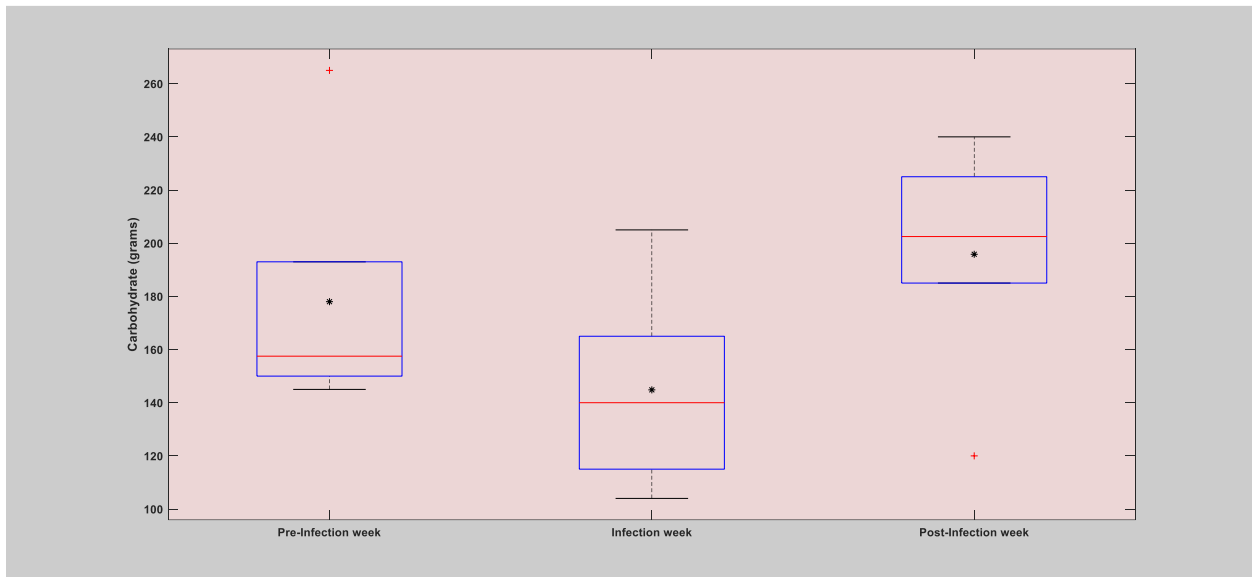


a) Comparison of Blood Glucose levels during pre-infection week, infection week and post-infection week. As can be seen, the blood glucose is elevated during the infection week as compared to the pre and post-infection week.





b) Comparison of Insulin (bolus) intake during pre-infection week, infection week and post-infection week. As can be seen, the amount of insulin (bolus) intake is elevated during the infection week as compared to the pre and post-infection week.



c) Comparison of Carbohydrate (grams) intake during pre-infection week, infection week and post-infection week. As can be seen, the amount of Carbohydrate (grams) intake is significantly reduced during the infection week as compared to the pre and post-infection week.

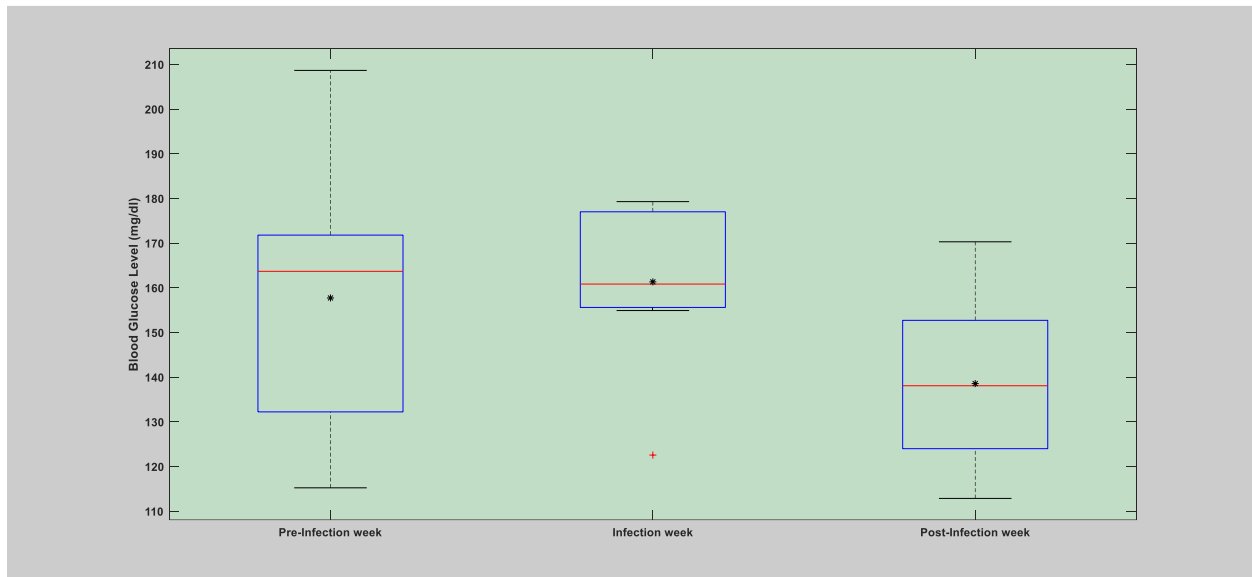
f) **Table 4:** Mean Percentage Change between pre-infection week versus infection week and post-infection week and infection week. For further reference, see **Table 1** Above.

Parameters	Pre-infection Vs. Infection week	Post-infection week Vs. Infection week
Blood Glucose (mg/dl)	7.26%	7.41%
Insulin (bolus)	56.87%	37.98%

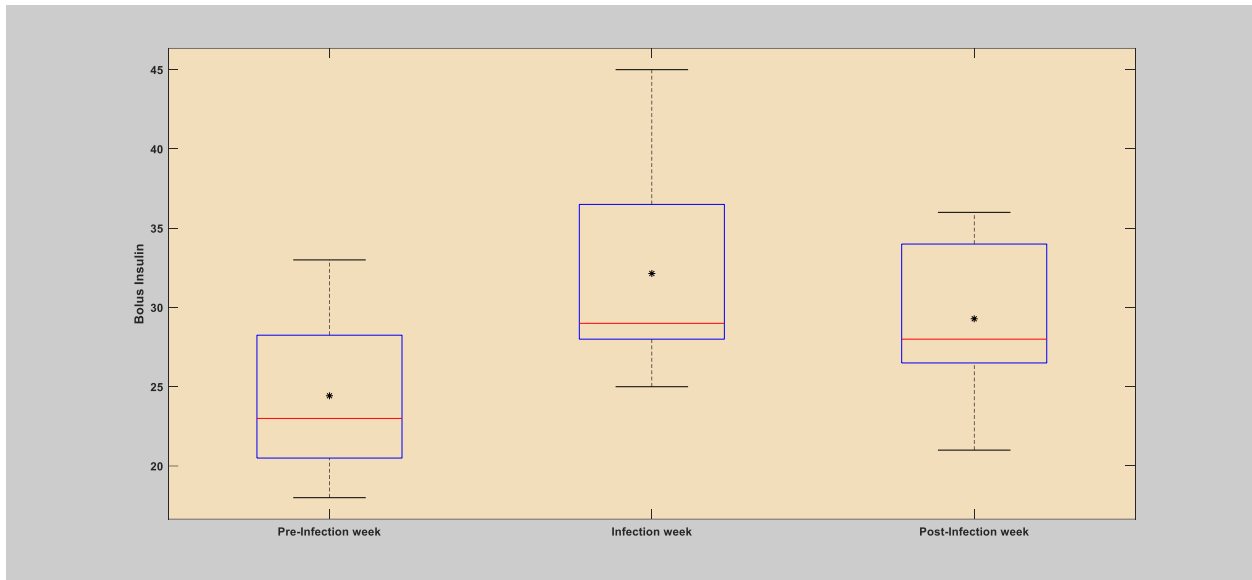
<b>Carbohydrate (grams)</b>	<b>-18.63%</b>	<b>-26.04%</b>
<b>Hourly</b>		
<b>Blood Glucose (mg/dl)</b>	<b>4.56%</b>	<b>7.21%</b>
<b>Insulin (bolus)</b>	<b>50.0%</b>	<b>61.29%</b>
<b>Carbohydrate (grams)</b>	<b>-15.95%</b>	<b>-10.99%</b>

Figure 3: Analysis of pre-infection week, infection week, and post-infection week based on the third patient year. Figure (a) depicts the blood glucose levels during these weeks. Figure (b) depicts the amount of insulin (bolus) injected during these weeks. Figure (c) depicts the amount of carbohydrate consumed in grams during these weeks. Table 4 shows the mean percentage change between these weeks. In all the figures, the asterisk shows the mean value and the red line depicts the median value for the week.

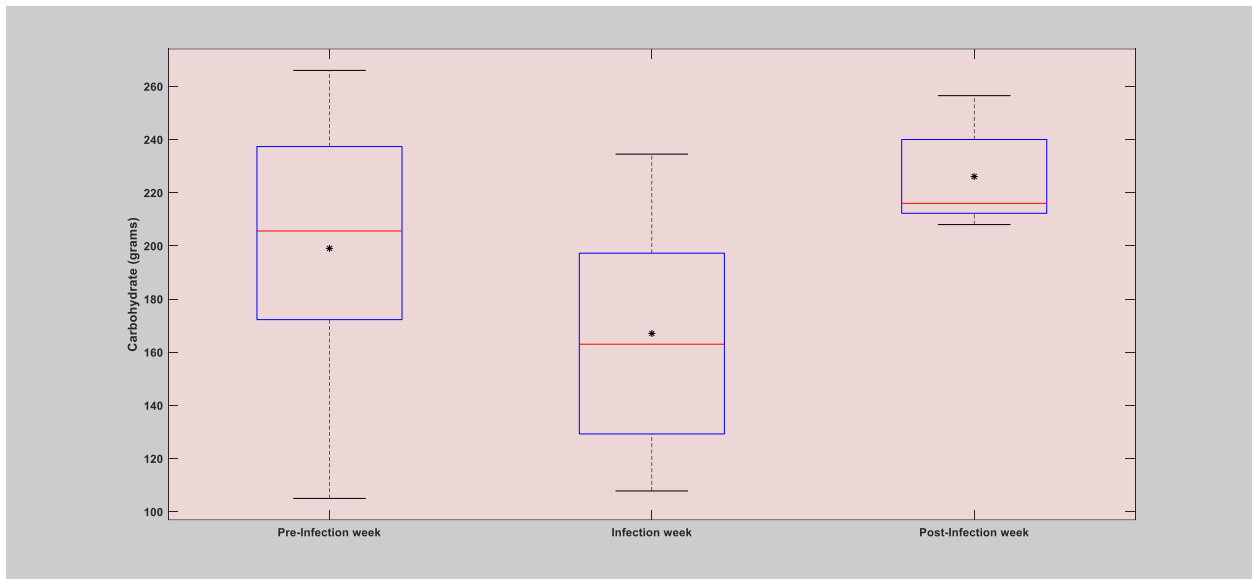
#### 4. The Fourth Case of Infection (flu)



- a) Comparison of Blood Glucose levels during pre-infection week, infection week and post-infection week. As can be seen, the blood glucose is elevated during the infection week as compared to the pre and post-infection week.



b) Comparison of Insulin (bolus) intake during pre-infection week, infection week and post-infection week. As can be seen, the amount of insulin (bolus) intake is elevated during the infection week as compared to the pre and post-infection week.



c) Comparison of Carbohydrate (grams) intake during pre-infection week, infection week and post-infection week. As can be seen, the amount of Carbohydrate (grams) intake is significantly reduced during the infection week as compared to the pre and post-infection week.

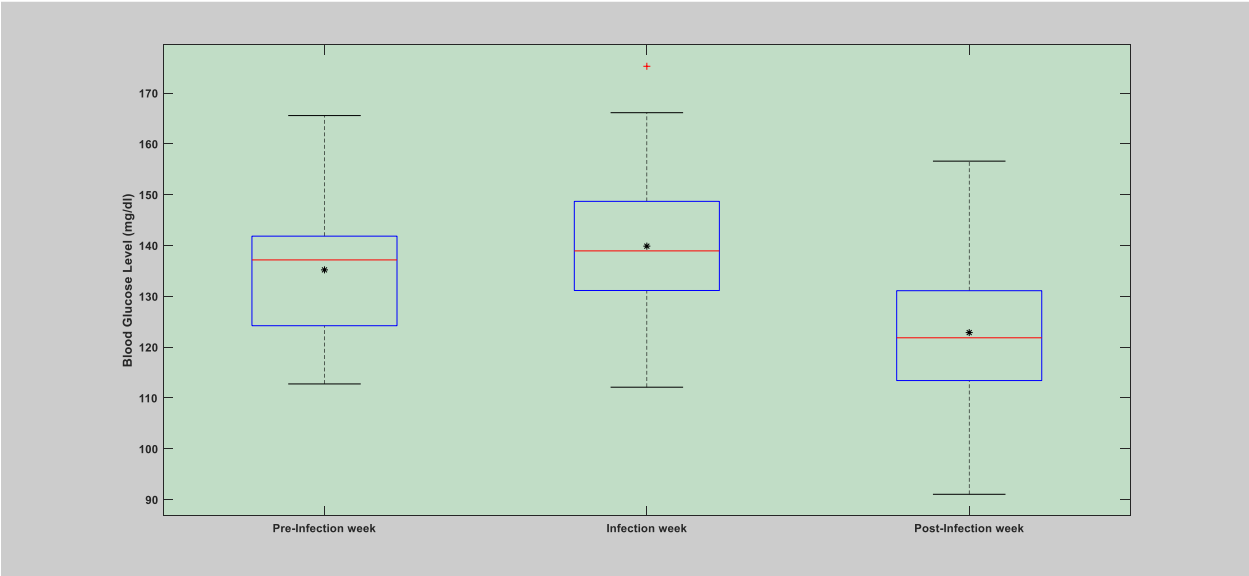
g) Table 5: Mean Percentage Change between pre-infection week versus infection week and post-infection week and infection week. For further reference, see Table 1 Above.

Parameters	Pre-infection Vs. Infection week	Post-infection week Vs. Infection week
Blood Glucose (mg/dl)	2.28%	16.43%
Insulin (bolus)	31.56%	9.73%

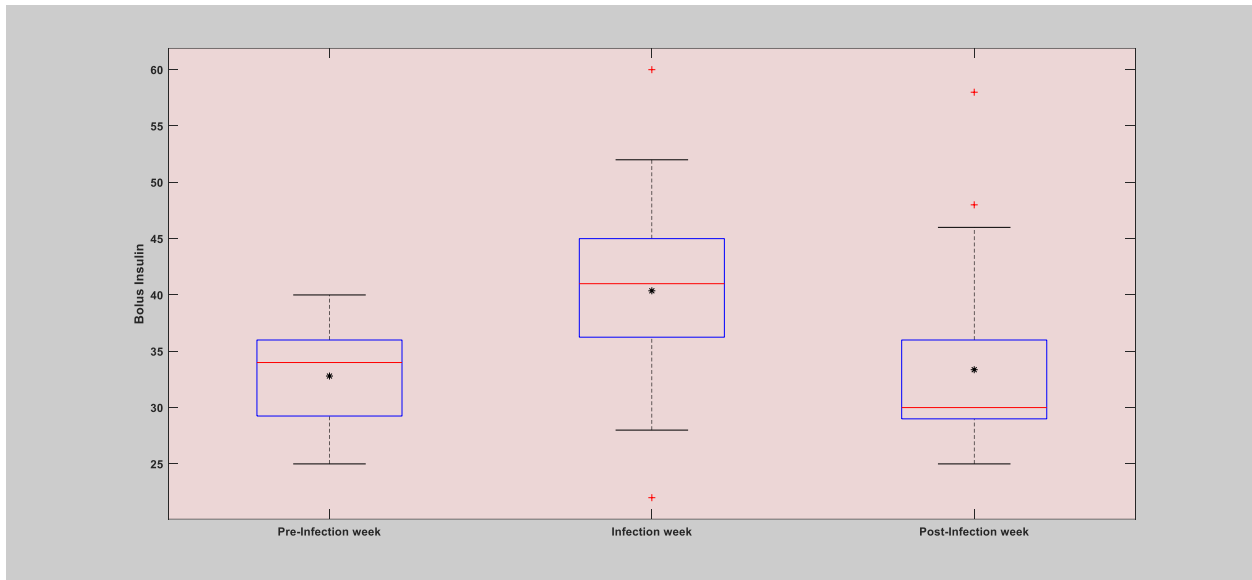
<b>Carbohydrate (grams)</b>	<b>-16.09%</b>	<b>-35.34%</b>
<b>Hourly</b>		
<b>Blood Glucose (mg/dl)</b>	<b>4.03%</b>	<b>10.92%</b>
<b>Insulin (bolus)</b>	<b>44.79%</b>	<b>16.81%</b>
<b>Carbohydrate (grams)</b>	<b>-19.09%</b>	<b>-30.68%</b>

**Figure 4:** Analysis of pre-infection week, infection week, and post-infection week based on the fourth patient year. Figure (a) depicts the blood glucose levels during these weeks. Figure (b) depicts the amount of insulin (bolus) injected during these weeks. Figure (c) depicts the amount of carbohydrate consumed in grams during these weeks. Table 5 shows the mean percentage change between these weeks. In all the figures, the asterisk shows the mean value and the red line depicts the median value for the week.

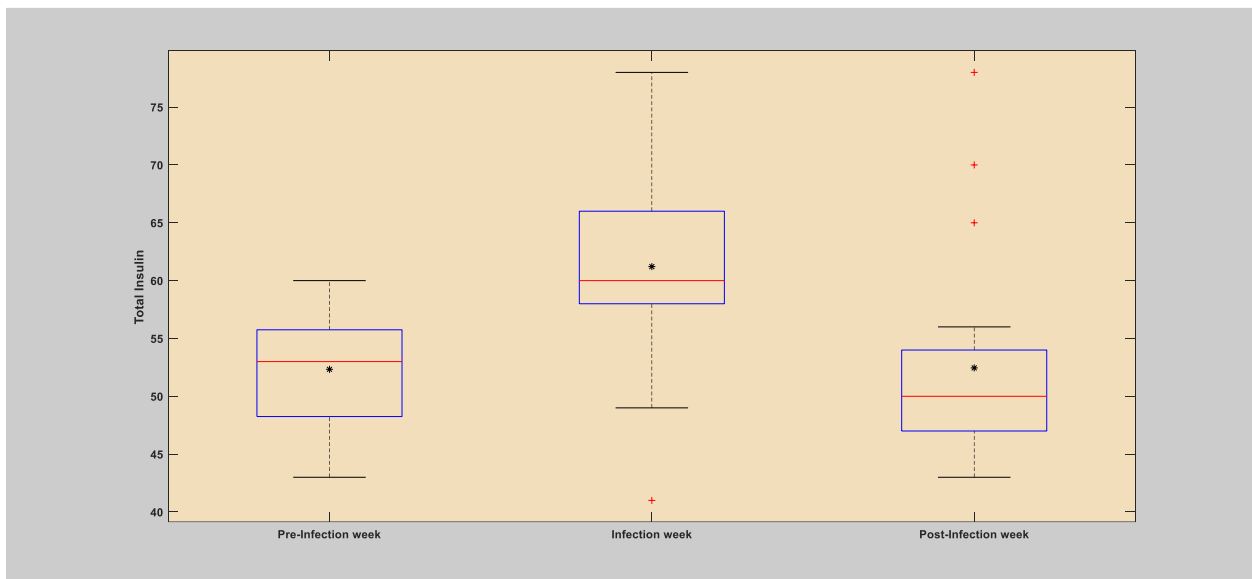
### 5. The Fifth Case of Infection (flu)



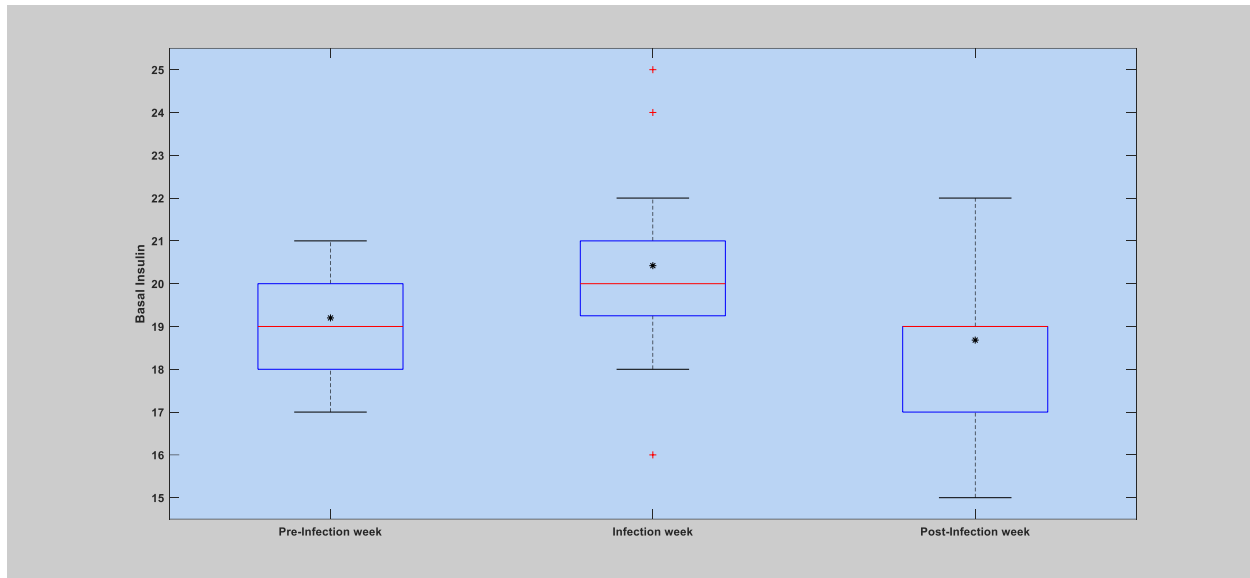
- a) Comparison of Blood Glucose levels during pre-infection week, infection week and post-infection week. As can be seen, the blood glucose is elevated during the infection week as compared to the pre and post-infection week.



- b) Comparison of Insulin (bolus) intake during pre-infection week, infection week and post-infection week. As can be seen, the amount of insulin (bolus) intake is elevated during the infection week as compared to the pre and post-infection week.



- c) Comparison of total Insulin (bolus and basal) intake during pre-infection week, infection week and post-infection week. As can be seen, the amount of total insulin intake is elevated during the infection week as compared to the pre and post-infection week.



- d) Comparison of Insulin (basal) intake during pre-infection week, infection week and post-infection week. As can be seen, the amount of insulin (basal) intake is elevated during the infection week as compared to the pre and post-infection week.
- h) Table 6: Mean Percentage Change between pre-infection week versus infection week and post-infection week and infection week. For further reference, see Table 1 Above.

Parameters	Pre-infection Vs. Infection week	Post-infection week Vs. Infection week
<b>Blood Glucose (mg/dl)</b>	<b>3.45%</b>	<b>13.84%</b>
<b>Insulin (bolus)</b>	<b>23.08%</b>	<b>21.01%</b>
<b>Insulin (basal)</b>	<b>6.35%</b>	<b>9.32%</b>
<b>Total Insulin</b>	<b>16.97%</b>	<b>16.68%</b>
<b>Hourly</b>		
<b>Blood Glucose (mg/dl)</b>	<b>7.346%</b>	<b>17.99%</b>
<b>Insulin (bolus)</b>	<b>29.42%</b>	<b>21.38%</b>

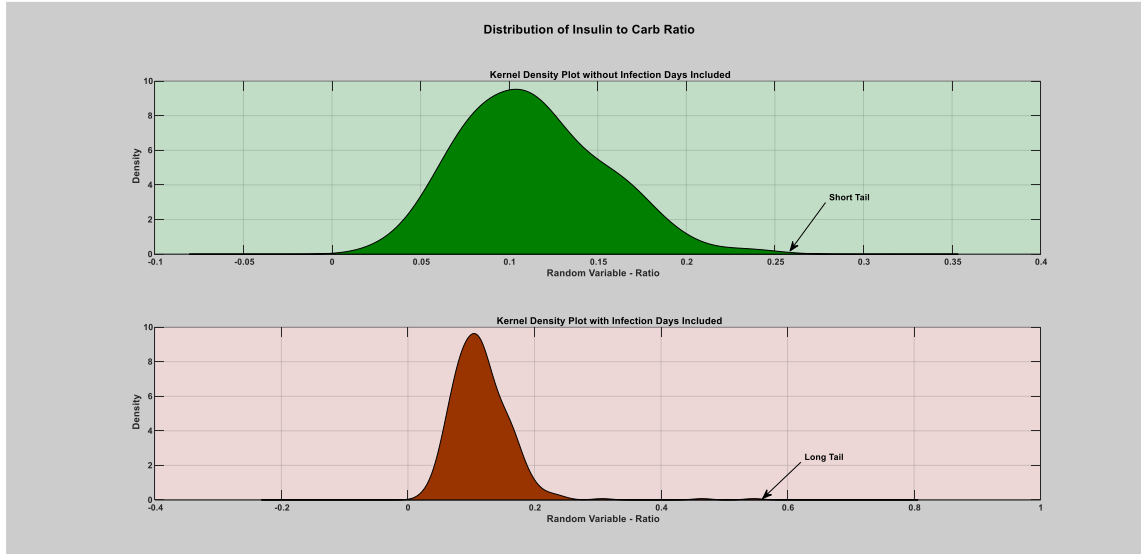
Figure 5: Analysis of pre-infection week, infection week, and post-infection week based on the fifth patient year. Figure (a) depicts the blood glucose levels during these weeks. Figure (b) depicts the amount of insulin (bolus) injected during these weeks. Figure (c) depicts the amount of total insulin (bolus + basal) injected during these weeks. Figure (d) depicts the amount of insulin (basal) injected during these weeks. Table 6 shows the mean percentage change between these weeks. In all the figures, the asterisk shows the mean value and the red line depicts the median value for the week.

## II. Kernel Density Estimation

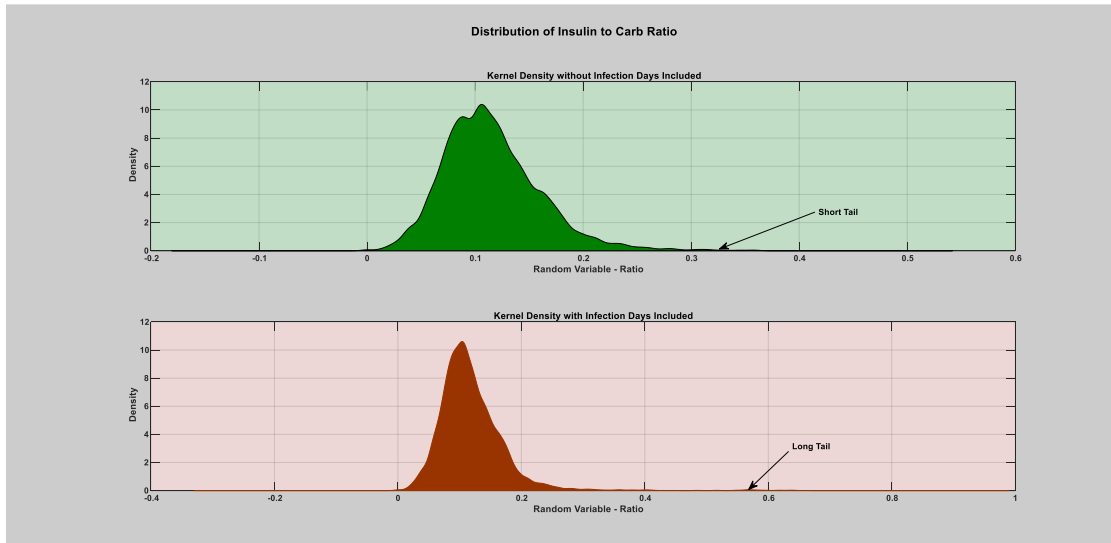
The kernel density was estimated relying on two procedures; by removing the infection period from the yearly data and computing the distribution and computing the kernel density for the whole year including the infection period.

This is carried out so as to identify the effect of the infection period on the distribution of the data (please see the manuscript for further explanation).

### 1. The First Case of infection (flu)



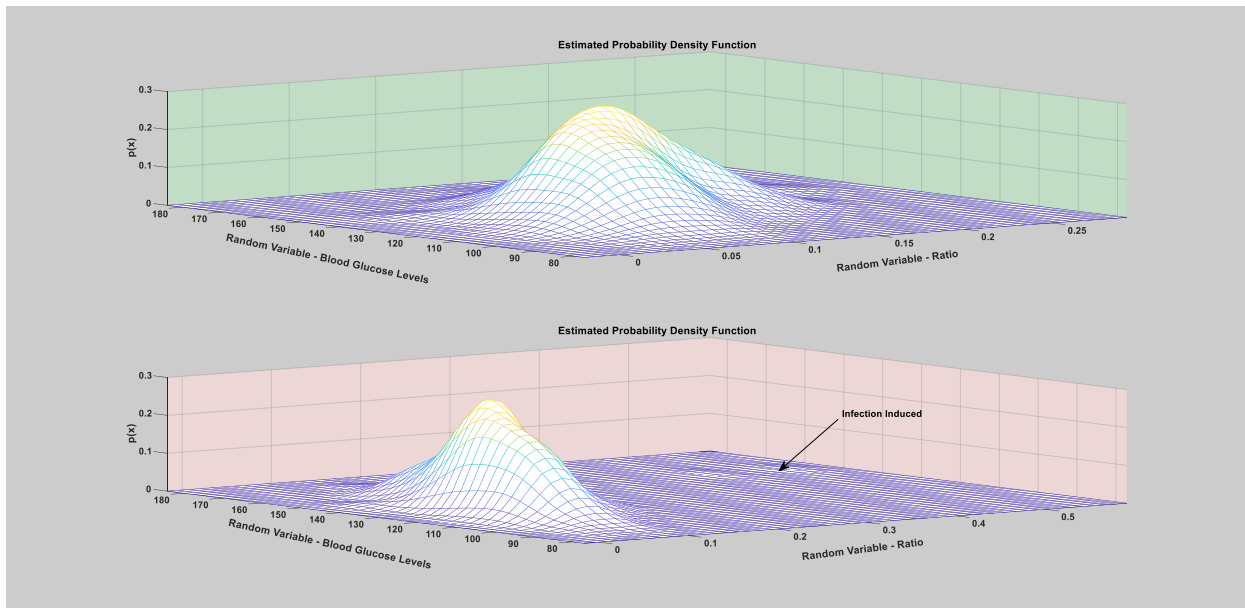
a) Kernel density estimation of daily total insulin (bolus) to carbohydrate ratio.



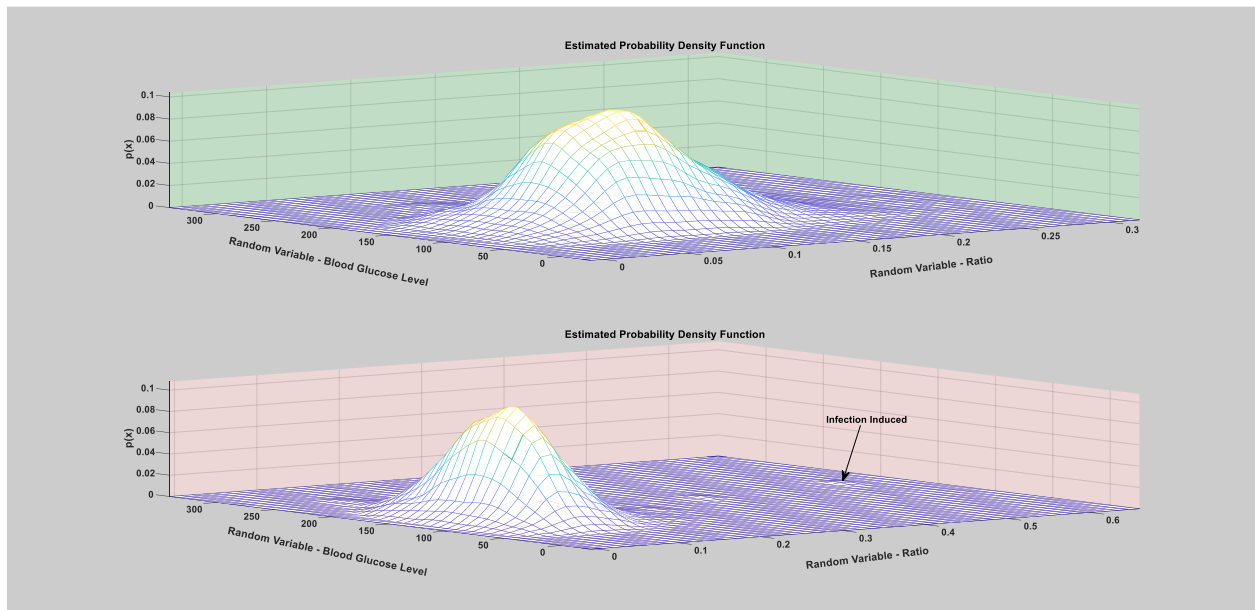
b) Kernel density estimation of hourly total insulin (bolus) to carbohydrate ratio.

**Figure 6:** Univariate kernel density estimation of a patient year using the daily insulin (bolus) to carbohydrate ratio. Figure (a) depicts the univariate kernel estimation of the daily total insulin (bolus) to carbohydrate ratio. Figure (b) depicts the univariate kernel estimation of the hourly total insulin (bolus) to carb ratio. As can be seen from the tail of the distribution, during normal days (the green shaded region) almost most of the yearly distribution of the patient insulin (bolus) to carbohydrate ratio lies within the values of 0.005 and 0.2. However,

during infection incidence (the red shaded region) there is a clear deviation in the tail of the distribution, where the values reaches around 0.58.



a) Kernel density estimation of daily average blood glucose levels vs. total insulin (bolus) to carbohydrate ratio.



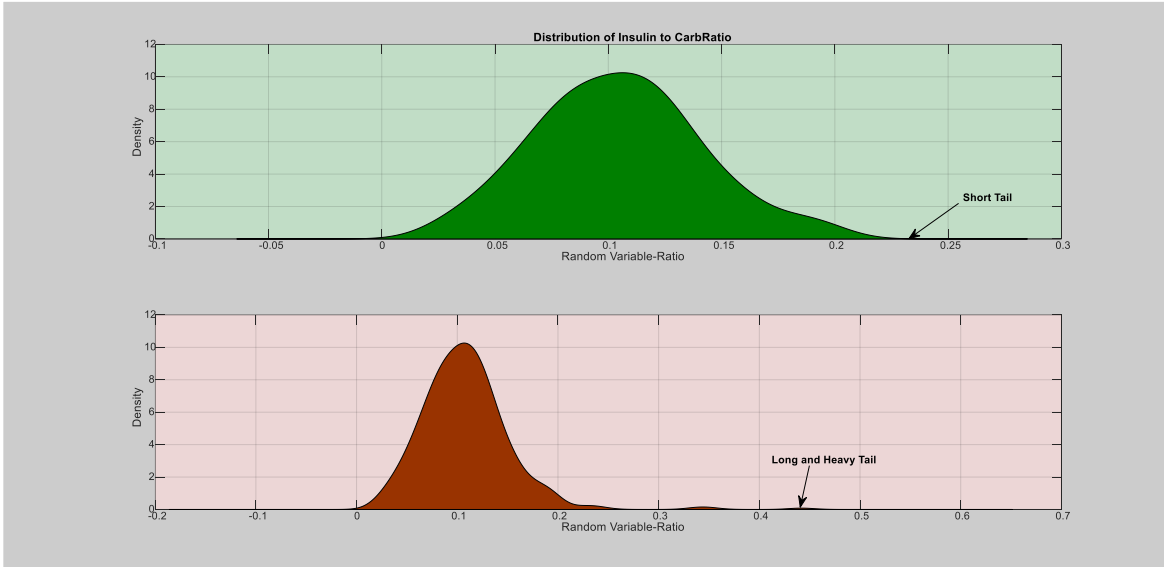
b) Kernel density estimation of hourly average blood glucose levels vs. total insulin (bolus) to carbohydrate ratio.

**Figure 7:** Bivariate Kernel Density estimation of a patient year using both the average BG levels and insulin (bolus) to carbohydrate ratio. Figure (a) depicts the bivariate kernel estimation of the daily average BG vs. total insulin (bolus) to carbohydrate ratio. Figure (b) depicts the bivariate kernel estimation of the hourly average BG vs. total insulin (bolus) to carbohydrate ratio. As can be seen from the bivariate distribution, during regular/normal

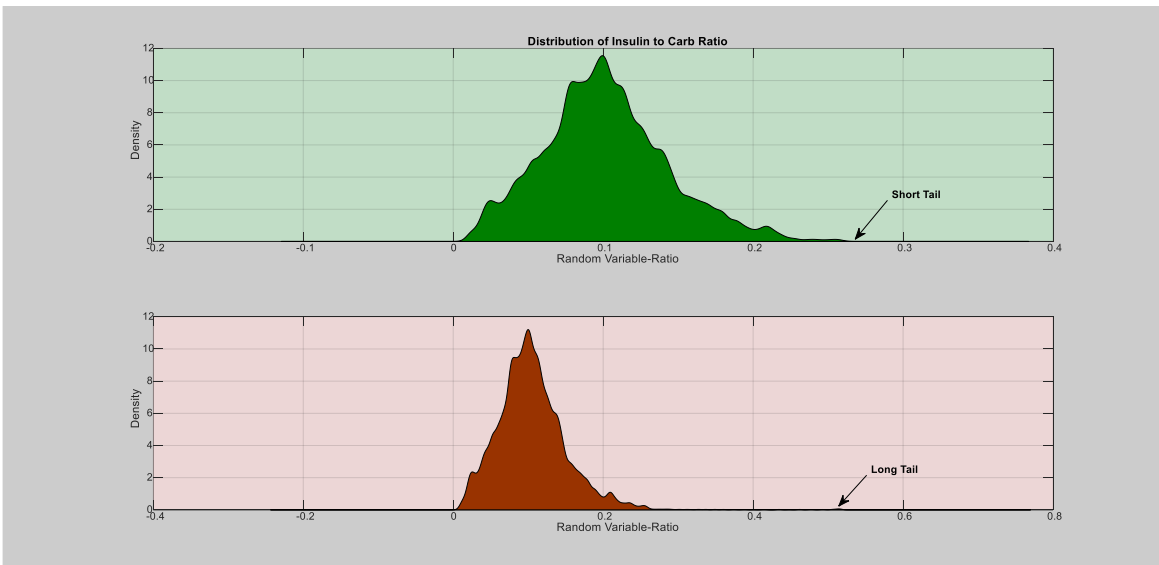


days (the top light green figure), the distributions are concentrated around the high density regions. However, during infection incidence (the lower figure), there is a clear bump far from the high density regions.

## 2. The Second Case of infection (flu)



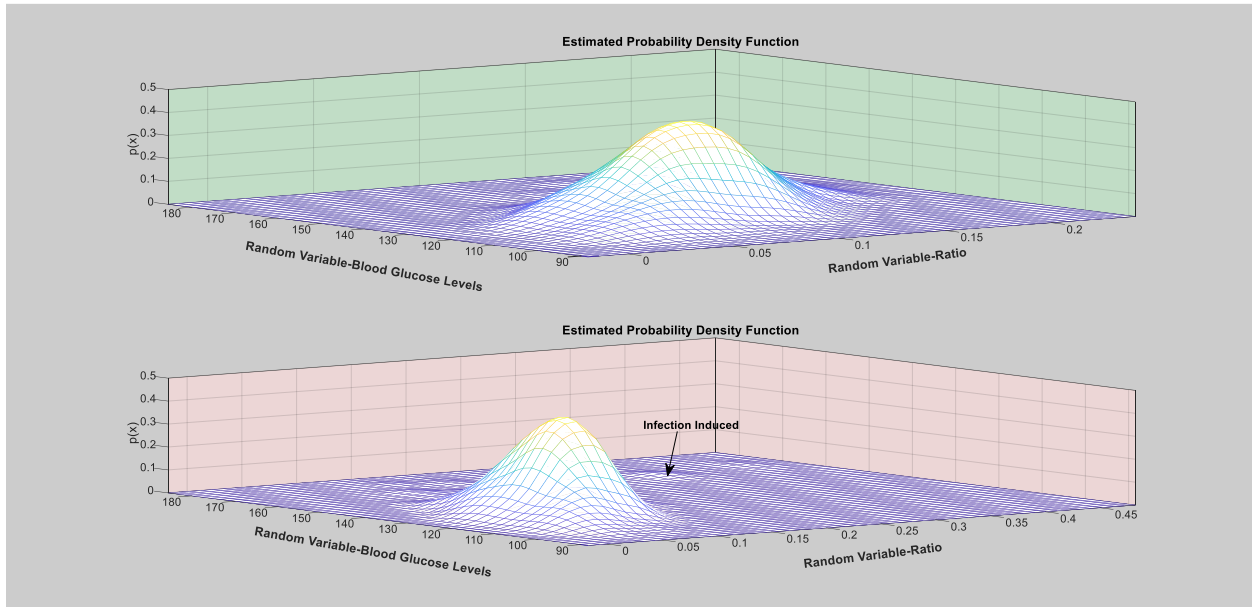
a) Kernel density estimation of daily total insulin (bolus) to carbohydrate ratio.



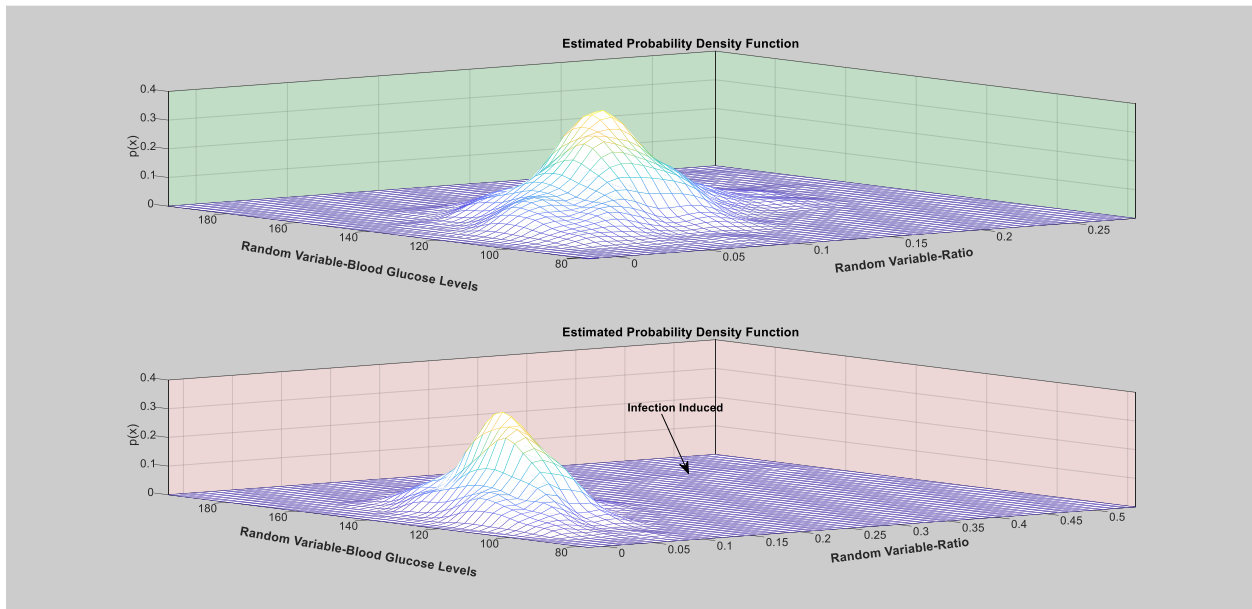
b) Kernel density estimation of hourly total insulin to carbohydrate ratio.

**Figure 8:** Univariate Kernel Density estimation of a patient year using the daily insulin (bolus) to carbohydrate ratio. Figure (a) depicts the univariate kernel estimation of the daily total insulin (bolus) to carbohydrate ratio. Figure (b) depicts the univariate kernel estimation of the hourly total insulin (bolus) to carb ratio. As can be seen from the tail of the distribution, during regular/normal days (the green shaded region) almost most of the yearly distribution of the patient insulin to carbohydrate ratio lies within the values of 0.005 and 0.25. However, during

infection incidence (the red shaded region) there is a clear deviation in the tail of the distribution, where the values reaches around 0.7.



a) Kernel density estimation of daily average blood glucose levels and total insulin (bolus) to carbohydrate ratio.

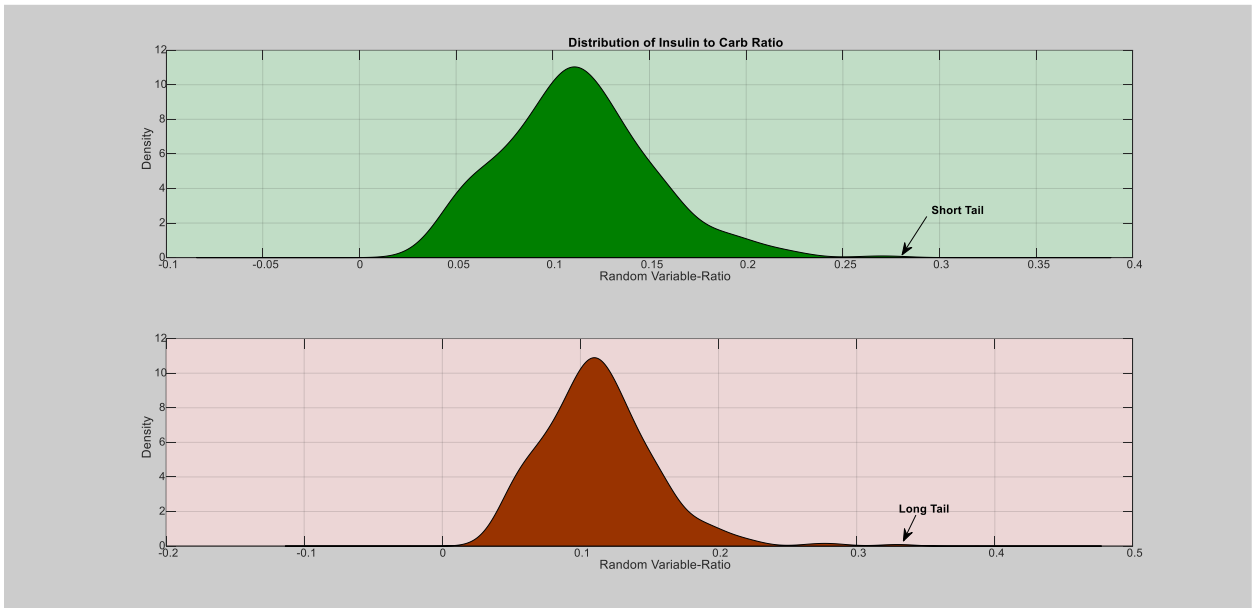


b) Kernel density estimation of hourly average blood glucose levels and total insulin (bolus) to carbohydrate ratio.

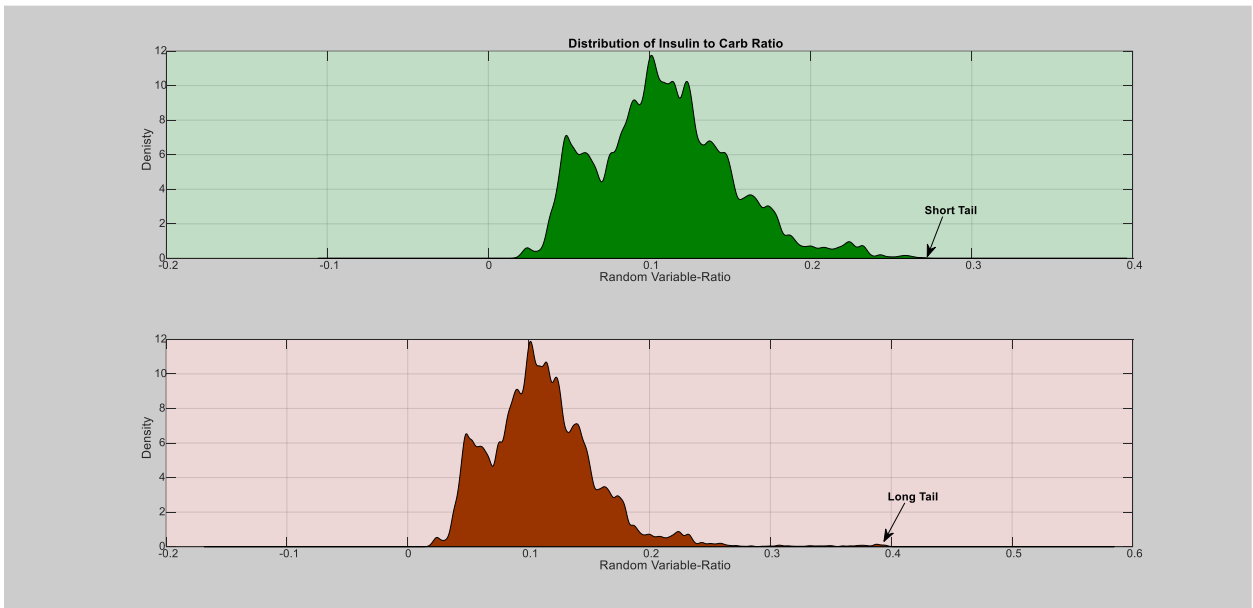
**Figure 9:** Bivariate Kernel Density estimation of a patient year using both the average blood glucose levels and insulin (bolus) to carbohydrate ratio. Figure (a) depicts the bivariate kernel estimation of the daily average BG vs.

total insulin (bolus) to carbohydrate ratio. Figure (b) depicts the bivariate kernel estimation of the hourly average BG vs. total insulin (bolus) to carbohydrate ratio. As can be seen from the bivariate distribution, during regular/normal days (the top light green figure), the distributions are concentrated around the high density regions. However, during infection incidence (the lower figure), there is a clear bump far from the high density regions.

### 3. The Third Case of infection (flu)

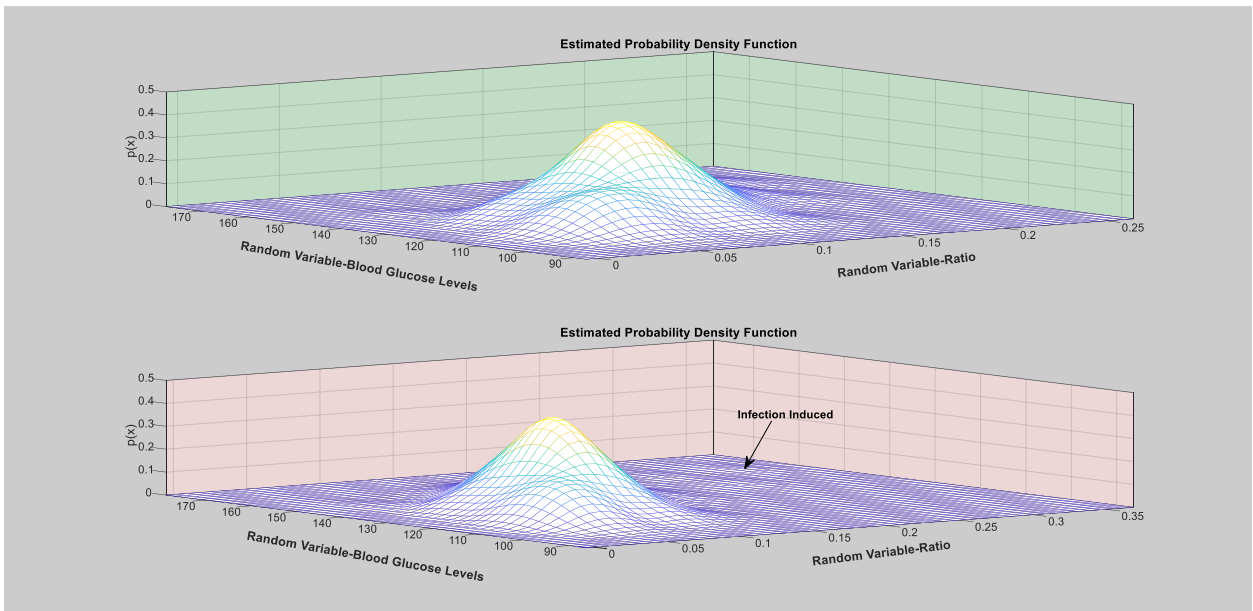


a) Kernel density estimation of daily total insulin (bolus) to carbohydrate ratio.

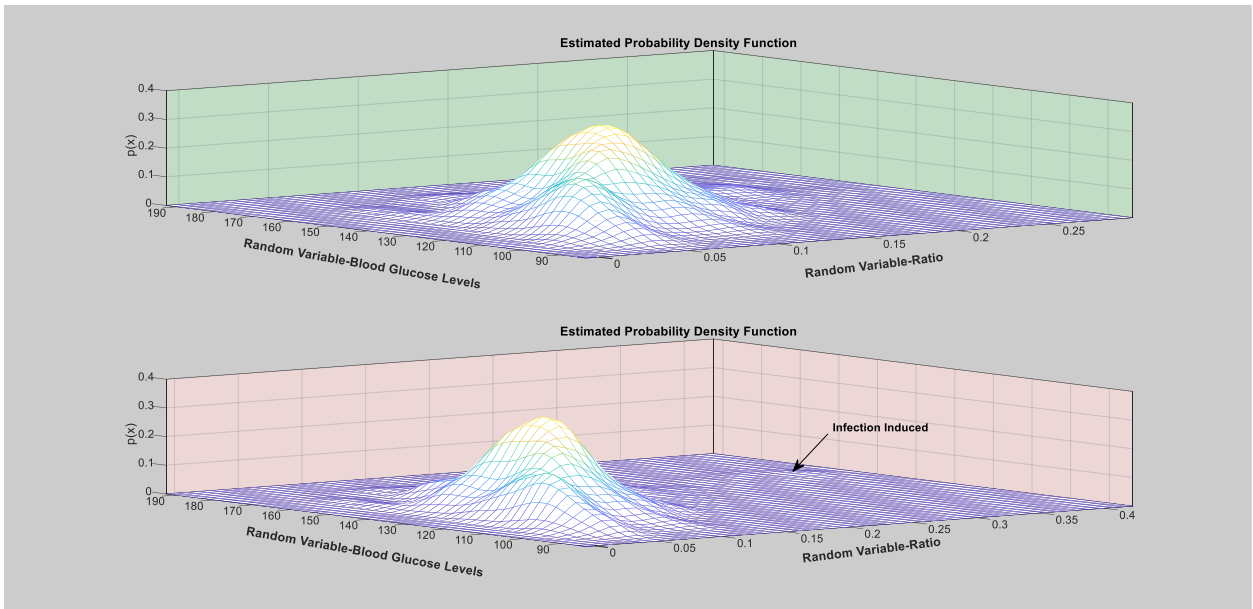


b) Kernel density estimation of hourly total insulin (bolus) to carbohydrate ratio.

**Figure 10:** Univariate Kernel Density estimation of a patient year using the daily insulin (bolus) to carbohydrate ratio. Figure (a) depicts the univariate kernel estimation of the daily total insulin (bolus) to carbohydrate ratio. Figure (b) depicts the univariate kernel estimation of the hourly total insulin (bolus) to carb ratio. As can be seen from the tail of the distribution, during regular/normal days (the green shaded region) almost most of the yearly distribution of the patient insulin (bolus) to carbohydrate ratio lies within the values of 0.005 and 0.26. However, during infection incidence (the red shaded region) there is a clear deviation in the tail of the distribution, where the values reaches around 0.5.



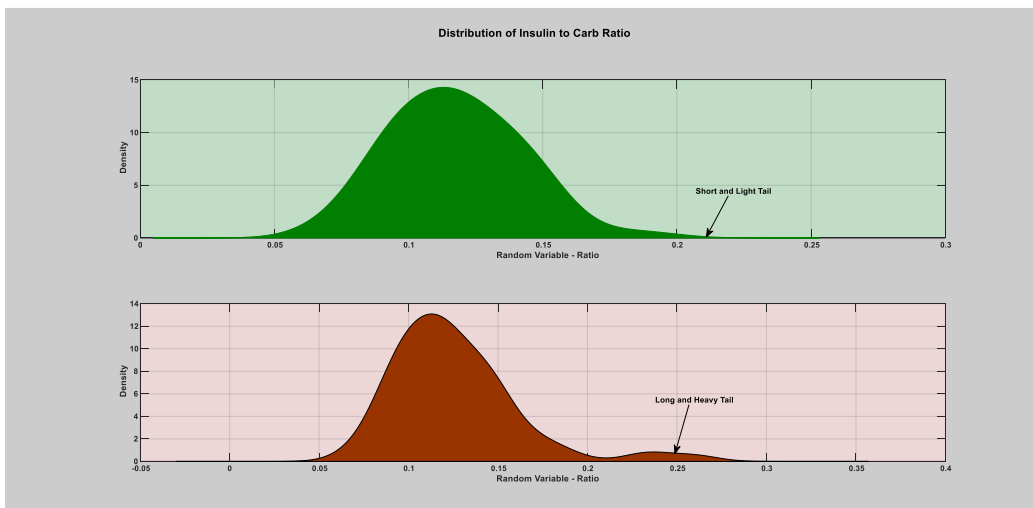
a) Kernel density estimation of daily average blood glucose levels and total insulin (bolus) to carbohydrate ratio.



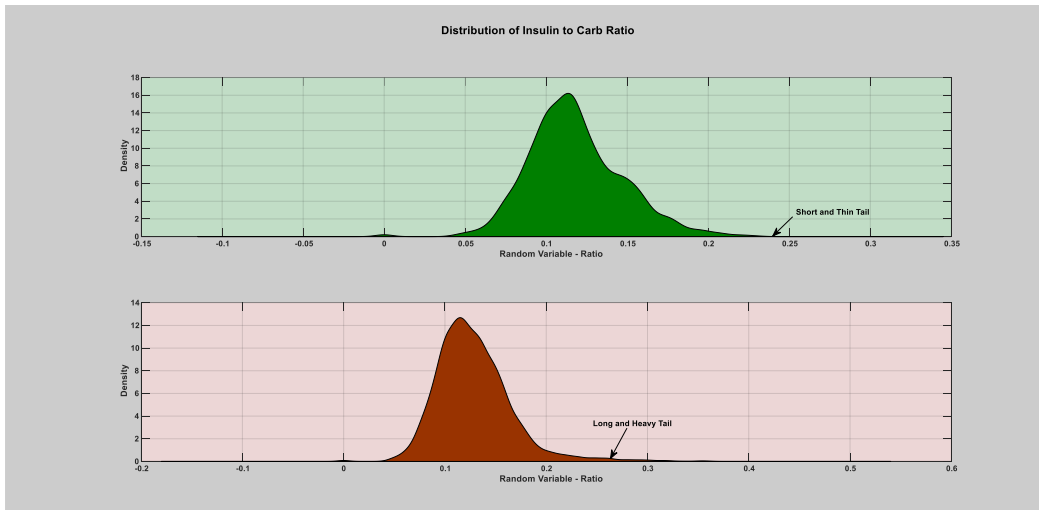
b) Kernel density estimation of hourly average blood glucose levels and total insulin (bolus) to carbohydrate ratio.

**Figure 11:** Bivariate Kernel Density estimation of a patient year using both the average blood glucose levels and insulin (bolus) to carbohydrate ratio. Figure (a) depicts the bivariate kernel estimation of the daily average BG vs. total insulin (bolus) to carbohydrate ratio. Figure (b) depicts the bivariate kernel estimation of the hourly average BG vs. total insulin (bolus) to carbohydrate ratio. As can be seen from the bivariate distribution, during regular/normal days (the top light green figure), the distributions are concentrated around the high density regions. However, during infection incidence (the lower figure), there is a clear bump far from the high density regions.

#### 4. The Fourth Case of infection (flu)

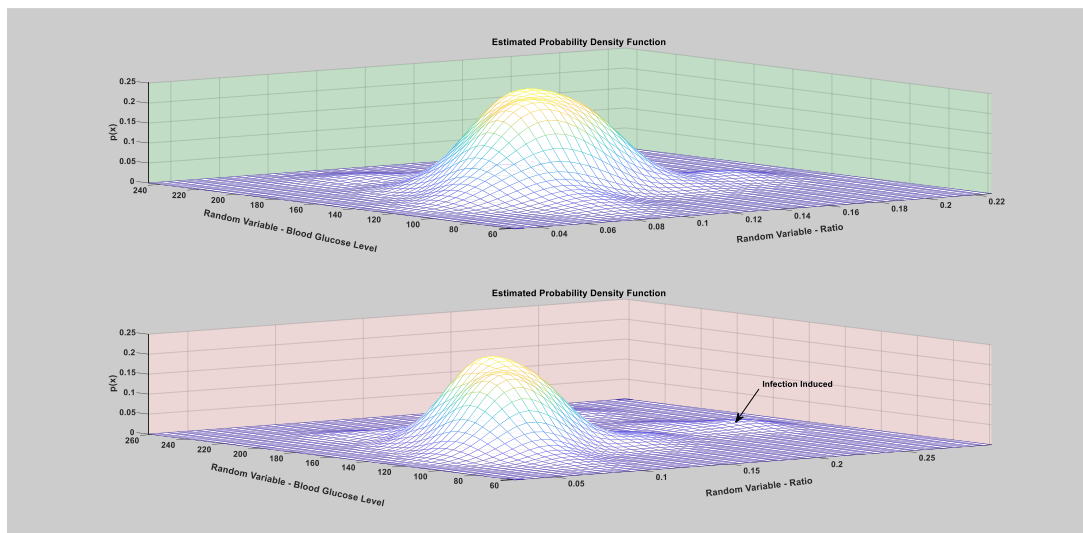


a) Kernel density estimation of daily total insulin (bolus) to carbohydrate ratio.

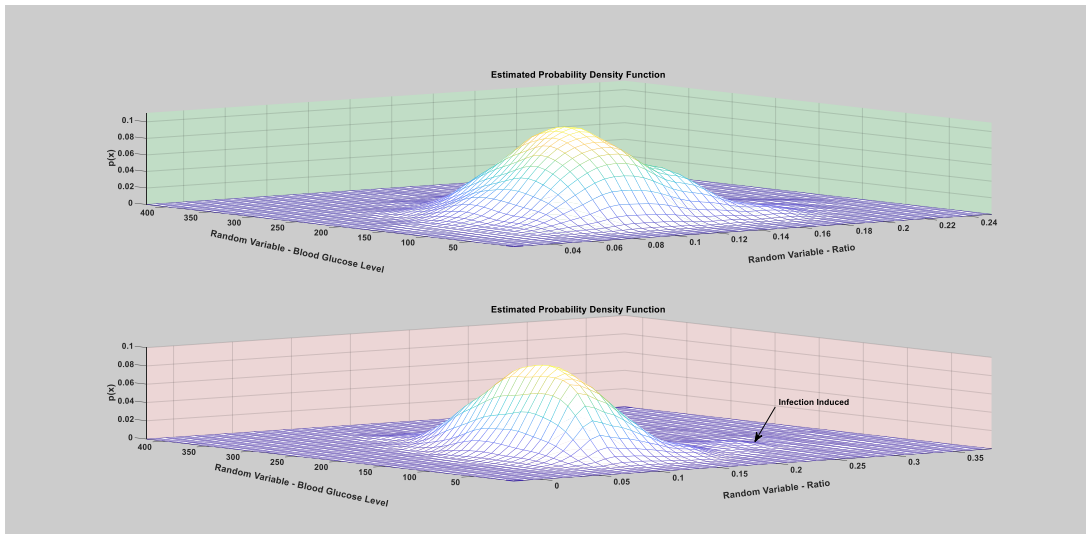


b) Kernel density estimation of hourly total insulin (bolus) to carbohydrate ratio.

**Figure 12:** Univariate Kernel Density estimation of a patient year using the daily insulin (bolus) to carbohydrate ratio. Figure (a) depicts the univariate kernel estimation of the daily total insulin (bolus) to carbohydrate ratio. Figure (b) depicts the univariate kernel estimation of the hourly total insulin (bolus) to carb ratio. As can be seen from the tail of the distribution, during regular/normal days (the green shaded region) almost most of the yearly distribution of the patient insulin to carbohydrate ratio lies within the values of 0.005 and 0.2. However, during infection incidence (the red shaded region) there is a clear deviation in the tail of the distribution, where the values reaches around 0.45.



a) Kernel density estimation of daily average blood glucose levels and total insulin (bolus) to carbohydrate ratio.



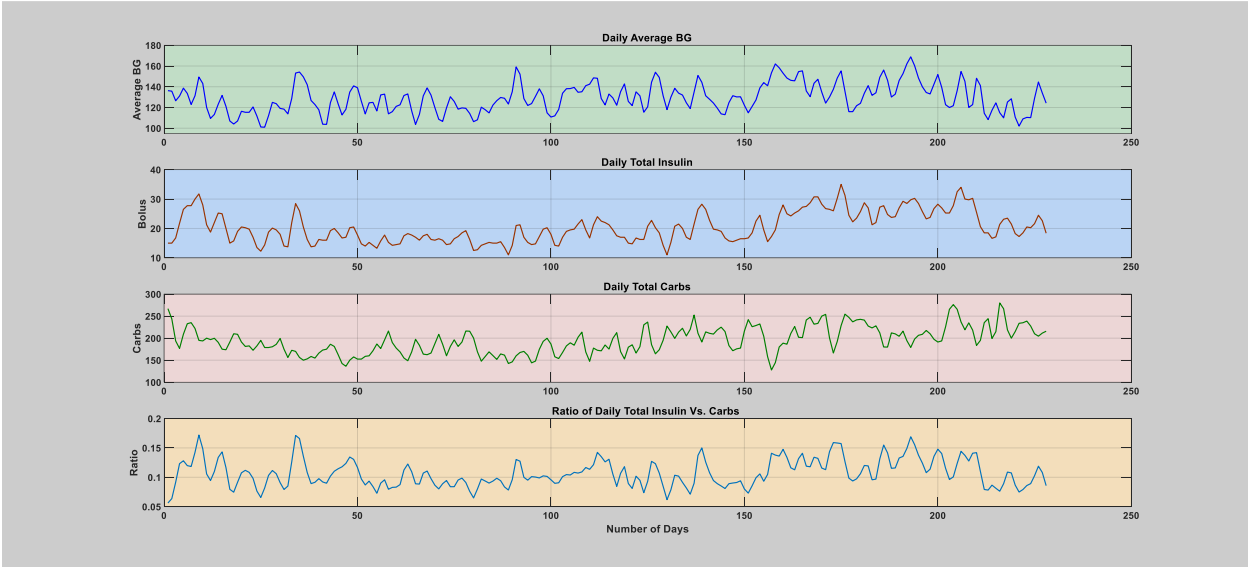
b) Kernel density estimation of hourly average blood glucose levels and total insulin (bolus) to carbohydrate ratio.

**Figure 13:** Bivariate Kernel Density estimation of a patient year using both the average blood glucose levels and insulin (bolus) to carbohydrate ratio. Figure (a) depicts the bivariate kernel estimation of the daily average BG vs. total insulin (bolus) to carbohydrate ratio. Figure (b) depicts the bivariate kernel estimation of the hourly average BG vs. total insulin (bolus) to carbohydrate ratio. As can be seen from the bivariate distribution, during regular/normal days (the top light green figure), the distributions are concentrated around the high density regions. However, during infection incidence (the lower figure), there is a clear bump far from the high density regions.

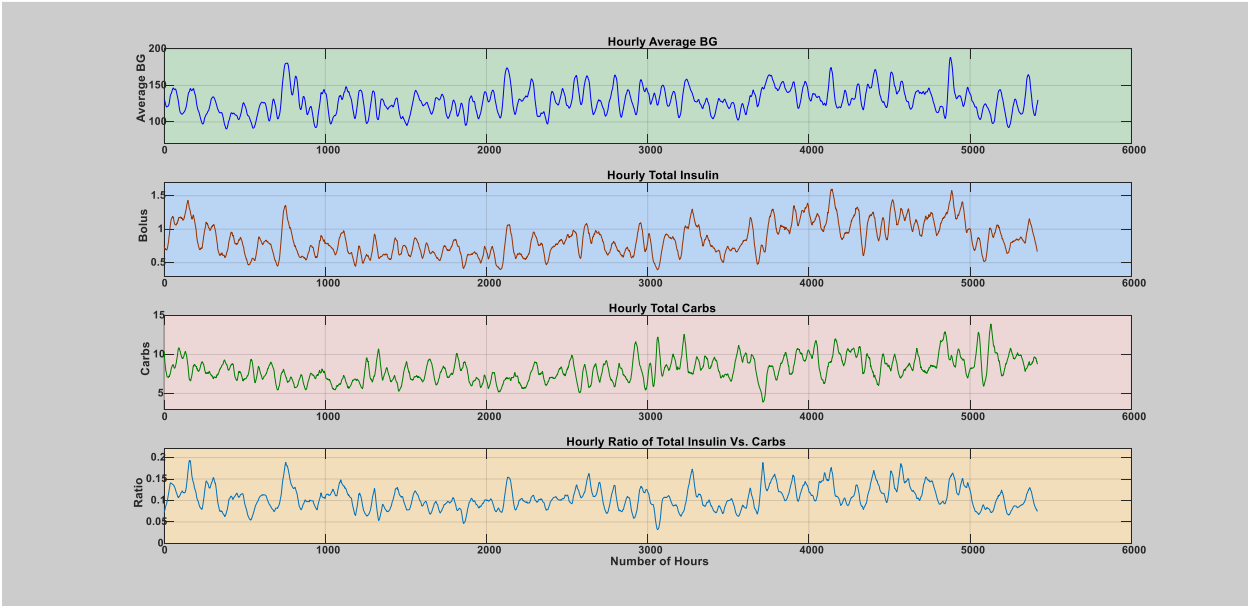
# Appendix 2: Analytical plot of the Normal/Regular Patient Years

The normal/regular patient year depicts the absence of significant infection incidences on the individual patient. These data are used to validate the change associated with presence of infections in an individual. The data were analyzed after computing the daily or hourly average BG, total insulin and carbohydrate consumption and smoothing with a 2-days window size moving average filter. The data was filtered to remove short term noise. The analytical plot as shown in the figure below demonstrated that under normal conditions the insulin to carbohydrate ratio remains between 0.05 and 0.2 in all the normal patient years.

## 1. The First Patient Year



a) Daily average BG levels, total insulin (bolus), total carbohydrate, and insulin to carbohydrate ratio.

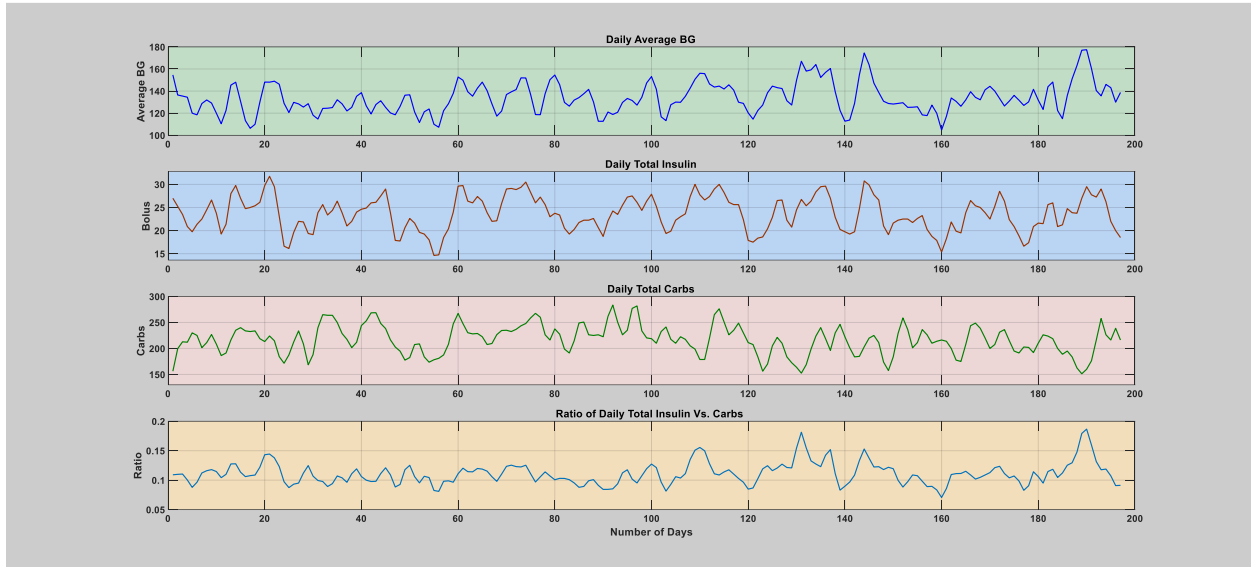




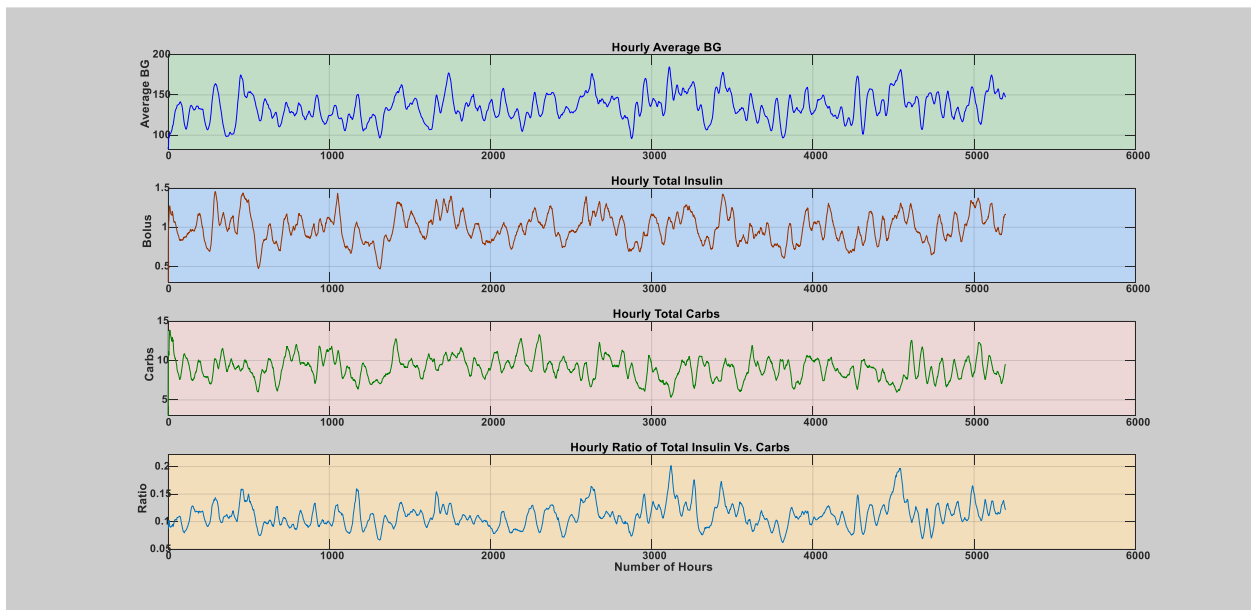
b) Hourly average BG levels, total insulin (bolus), total carbohydrate, and insulin to carbohydrate ratio.

**Figure 1:** The first patient year, where there is no incidence of acute infections. Figure (a) depicts the daily variation of BG, total insulin (bolus), carbohydrate, and insulin to carbohydrate ratio. Figure (b) depicts variation of the same variable during each hours of the day. The operating point of the patient's insulin to carbohydrate ratio through these normal days is between 0.05 to 0.2.

## 2. The Second Patient Year



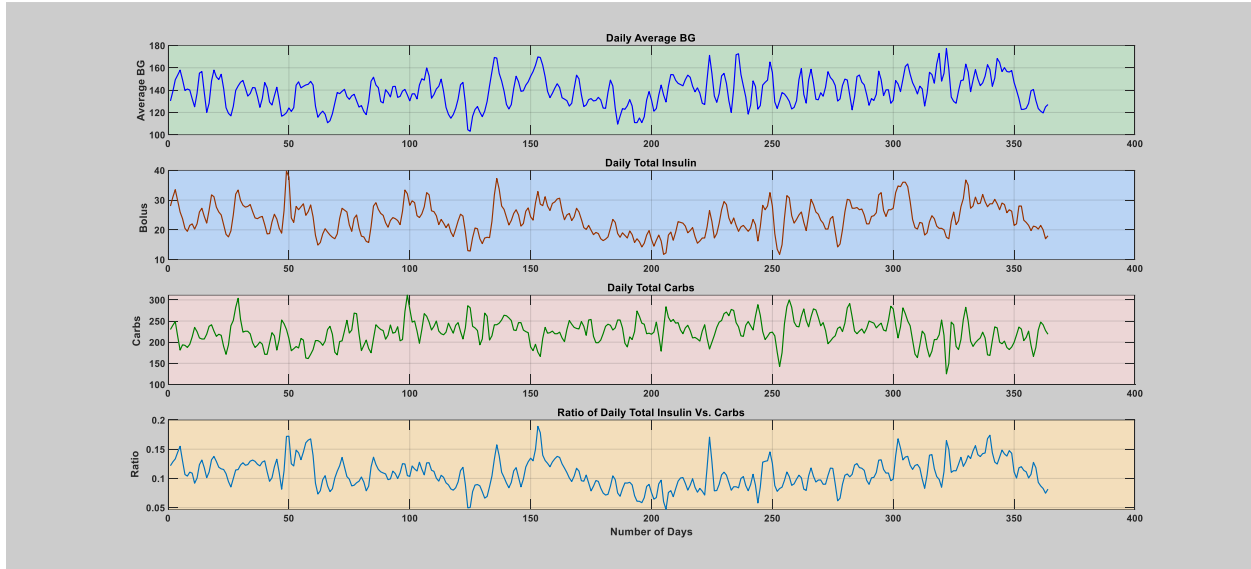
a) Daily average BG levels, total insulin (bolus), total carbohydrate, and insulin to carbohydrate ratio.



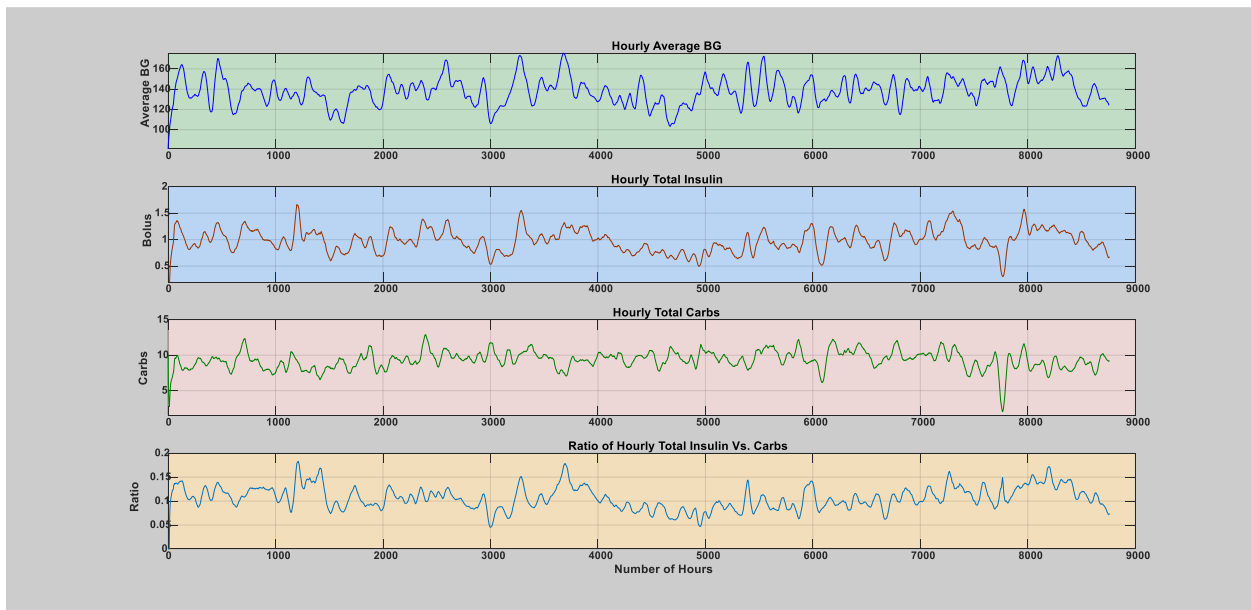
b) Hourly average BG levels, total insulin (bolus), total carbohydrate, and insulin to carbohydrate ratio.

**Figure 2:** The second patient year, where there is no incidence of acute infections. Figure (a) depicts the daily variation of BG, total insulin (bolus), carbohydrate, and insulin to carbohydrate ratio. Figure (b) depicts variation of the same variable during each hours of the day. The operating point of the patient's insulin to carbohydrate ratio through these normal days is between 0.05 to 0.2.

### 3. The Third Patient Year



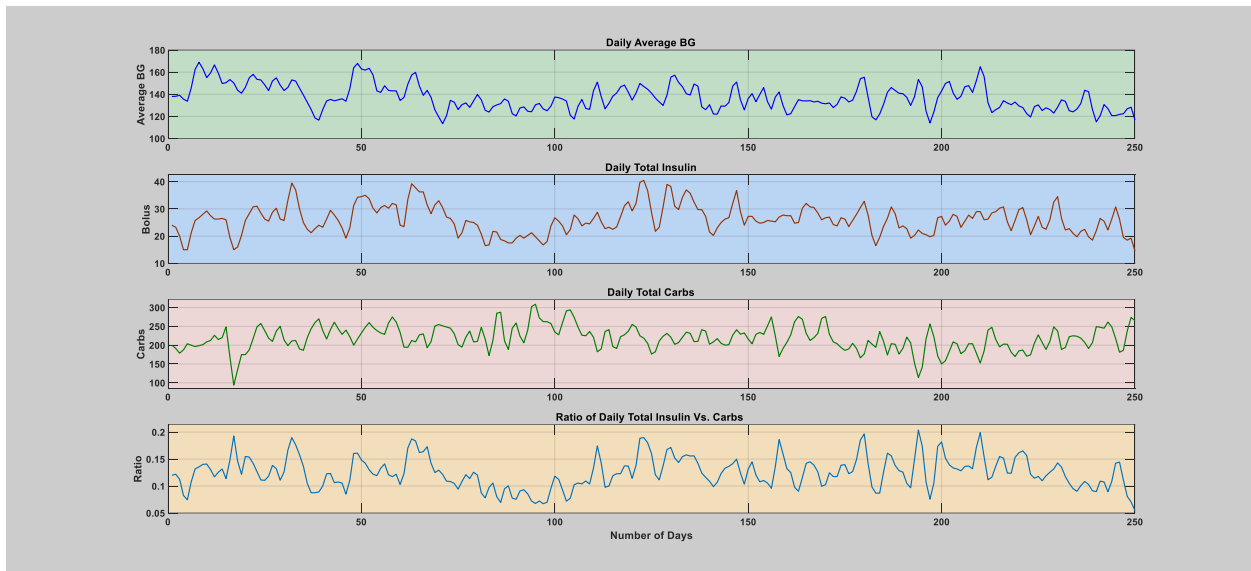
a) Daily average BG levels, total insulin (bolus), total carbohydrate, and insulin to carbohydrate ratio.



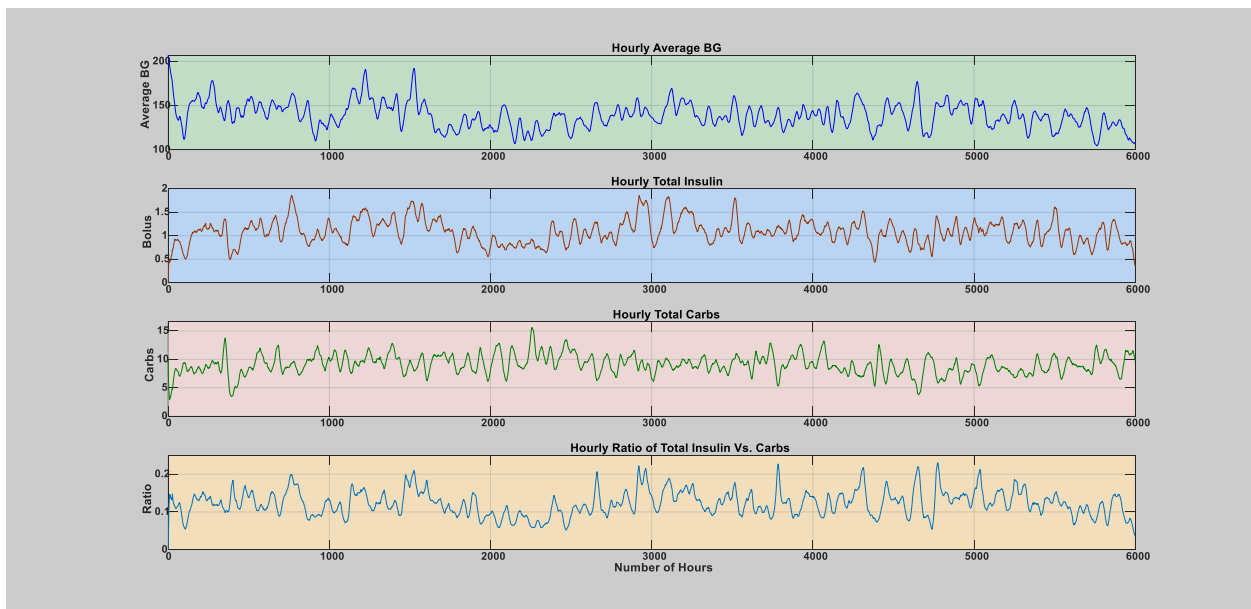
b) Hourly average BG levels, total insulin (bolus), total carbohydrate, and insulin to carbohydrate ratio.

Figure 3: The third patient year, where there is no incidence of acute infections. Figure (a) depicts the daily variation of BG, total insulin (bolus), carbohydrate, and insulin to carbohydrate ratio. Figure (b) depicts variation of the same variable during each hours of the day. The operating point of the patient's insulin to carbohydrate ratio through these normal days is between 0.05 to 0.2.

#### 4. The Fourth Patient Year



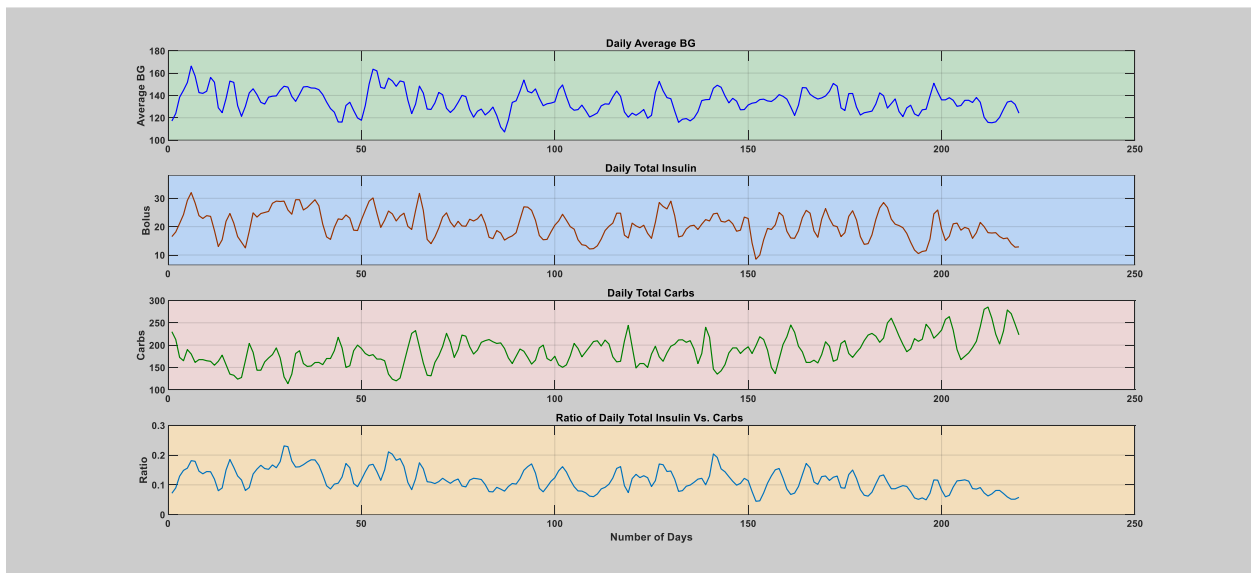
a) Daily average BG levels, total insulin (bolus), total carbohydrate, and insulin to carbohydrate ratio.



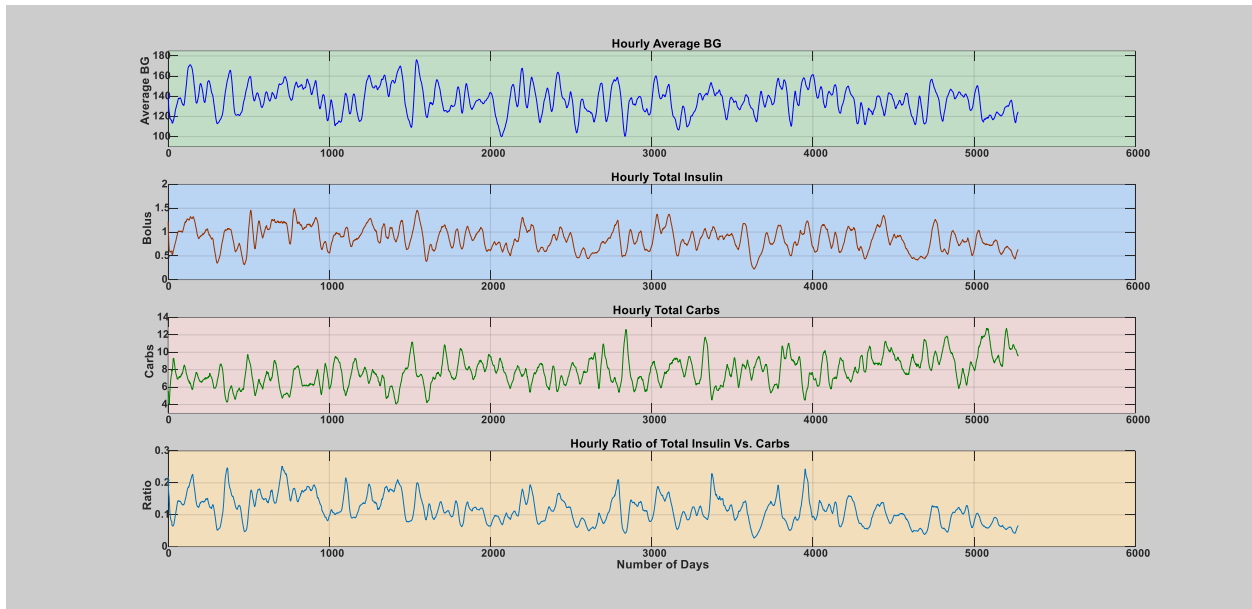
b) Hourly average BG levels, total insulin (bolus), total carbohydrate, and insulin to carbohydrate ratio.

**Figure 4:** The fourth patient year, where there is no incidence of acute infections. Figure (a) depicts the daily variation of BG, total insulin (bolus), carbohydrate, and insulin to carbohydrate ratio. Figure (b) depicts variation of the same variable during each hours of the day. The operating point of the patient's insulin to carbohydrate ratio through these normal days is between 0.05 to 0.21.

## 5. The Fifth Patient Year



a) Daily average BG levels, total insulin (bolus), total carbohydrate, and insulin to carbohydrate ratio.



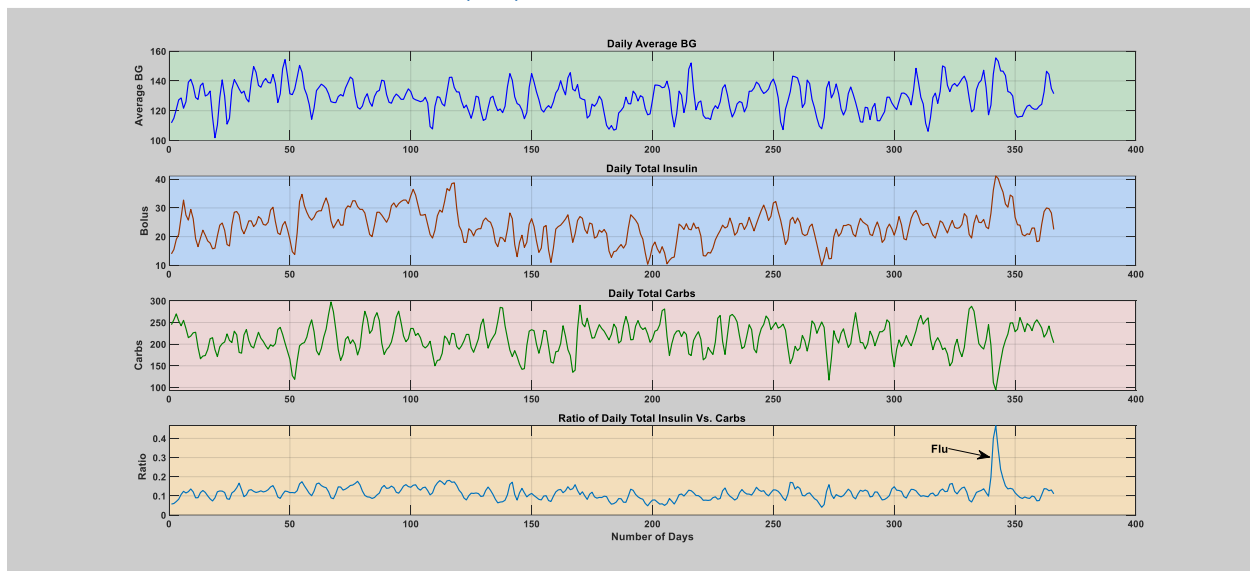
b) Hourly average BG levels, total insulin (bolus), total carbohydrate, and insulin to carbohydrate ratio.

**Figure 5:** The fifth patient year, where there is no incidence of acute infections. Figure (a) depicts the daily variation of BG, total insulin (bolus), carbohydrate, and insulin to carbohydrate ratio. Figure (b) depicts variation of the same variable during each hours of the day. The operating point of the patient's insulin to carbohydrate ratio through these normal days is between 0.05 to 0.22.

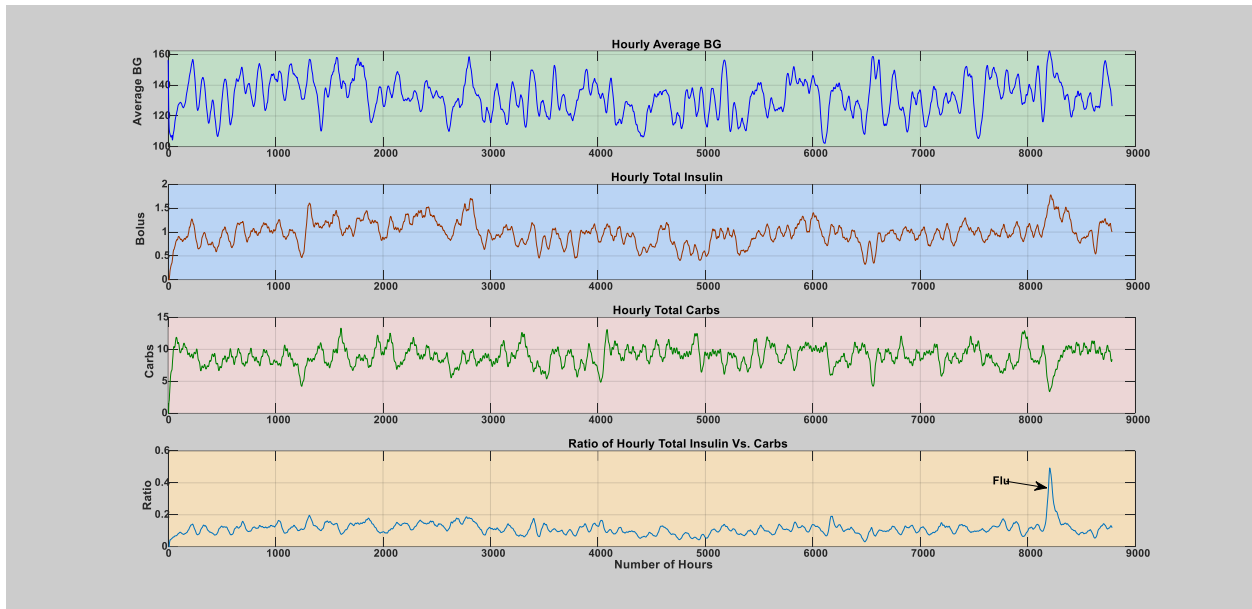
## Appendix 3: Analytical plot of the Patient Years with acute infection

The patient year with acute infection depicts a patient year containing at least with one or more infection incidences. These data are used to compare and evaluate the effect of infection incidences on the key parameters of the BG dynamics on an individual basis including BG levels, insulin intake, carbohydrate consumption, and ratio of insulin to carbohydrate. The data were analyzed after computing the daily or hourly average BG, total insulin and carbohydrate consumption and smoothing with a 2-days window size moving average filter. The data was filtered to remove short term noise. The analytical plot as shown in the figure below demonstrated that during infection incidences the key parameters of the BG dynamics are highly affected. The patient experiences elevated BG levels while the insulin to carbohydrate ratio is dramatically shifted to a higher value depicting higher insulin intake with low carbohydrate ingestions. During normal conditions the insulin to carbohydrate ratio remains between 0.05 and 0.2.

### 1. The first case of infection (Flu)



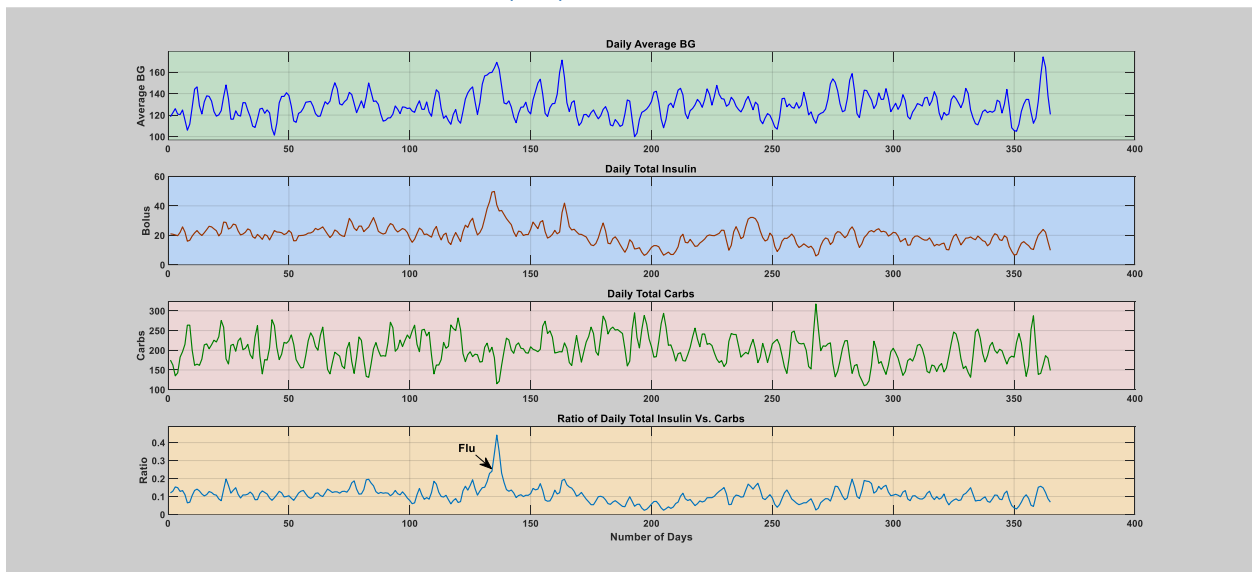
- a) Daily average BG levels, total insulin (bolus), total carbohydrate, and insulin to carbohydrate ratio based on a window size of two days.



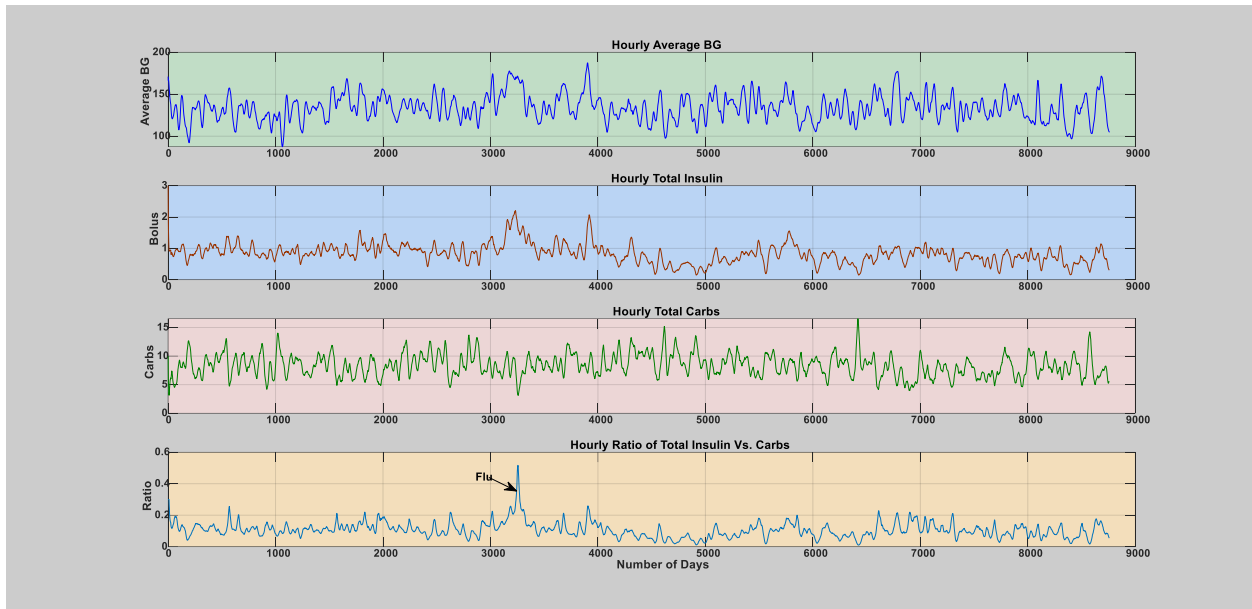
b) Hourly average BG levels, total insulin (bolus), total carbohydrate, and insulin to carbohydrate ratio based on a window size of 48 hours.

**Figure 1:** The sixth patient year, where the patient was infected with influenza (flu) starting from the first week of December. Figure (a) depicts the daily variation of BG, total insulin (bolus), carbohydrate, and insulin to carbohydrate ratio. Figure (b) depicts variation of the same variable during each hours of the day. The operating point of the patient's insulin to carbohydrate ratio had dramatically shifted and raised above the regular/normal days and reach a top around 0.5 upon mid infection week.

## 2. The second case of infection (Flu)



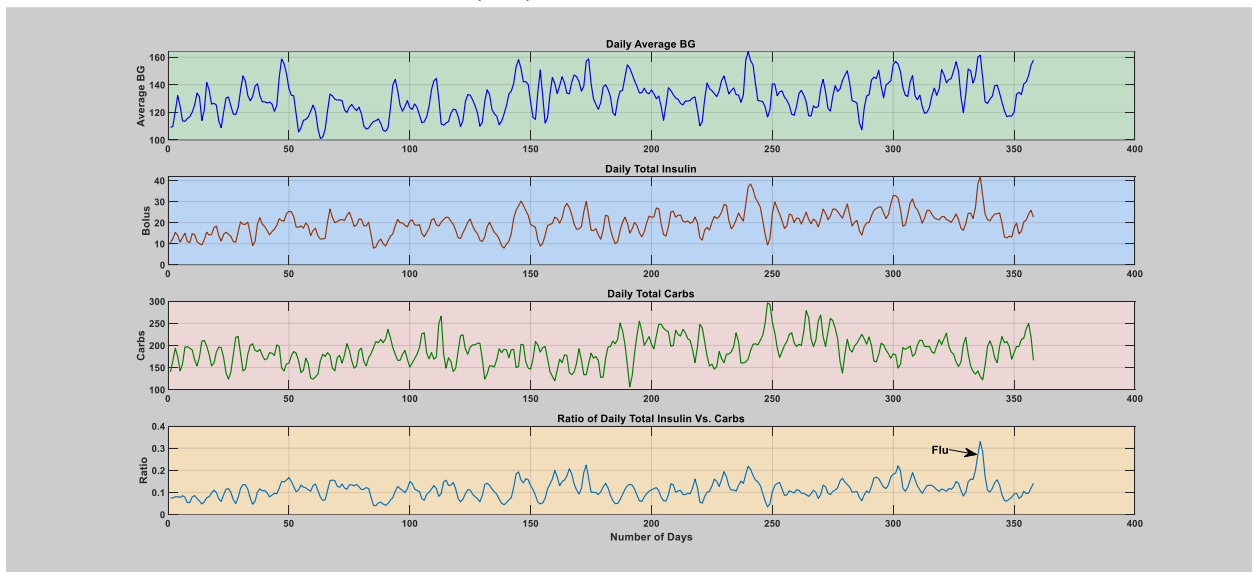
a) Daily average BG levels, total insulin (bolus), total carbohydrate, and insulin to carbohydrate ratio based on a window size of two days.



b) Hourly average BG levels, total insulin (bolus), total carbohydrate, and insulin to carbohydrate ratio based on a window size of 48 hours.

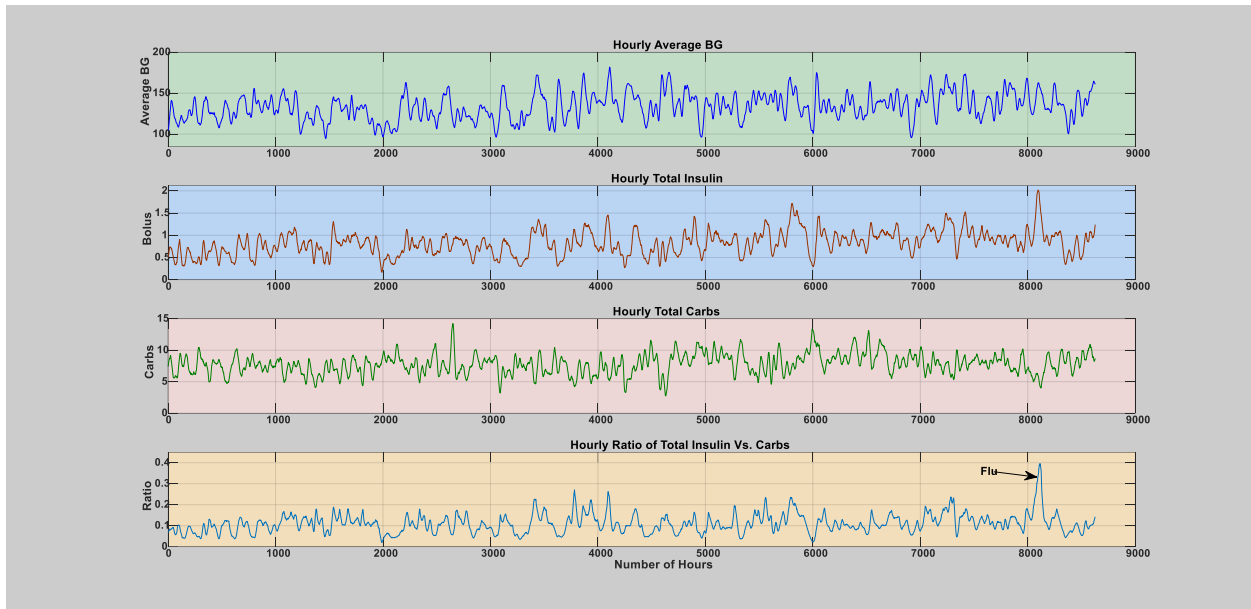
**Figure 2:** The seventh patient year, where the patient was infected with influenza (flu) starting from the first week of April. Figure (a) depicts the daily variation of BG, total insulin (bolus), carbohydrate, and insulin to carbohydrate ratio. Figure (b) depicts variation of the same variable during each hours of the day. The operating point of the patient's insulin to carbohydrate ratio had dramatically shifted and raised above the regular/normal days and reach a top around 0.45 upon mid infection week.

### 3. The third case of infection (Flu)



a) Daily average BG levels, total insulin (bolus), total carbohydrate, and insulin to carbohydrate ratio based on a window size of two days.

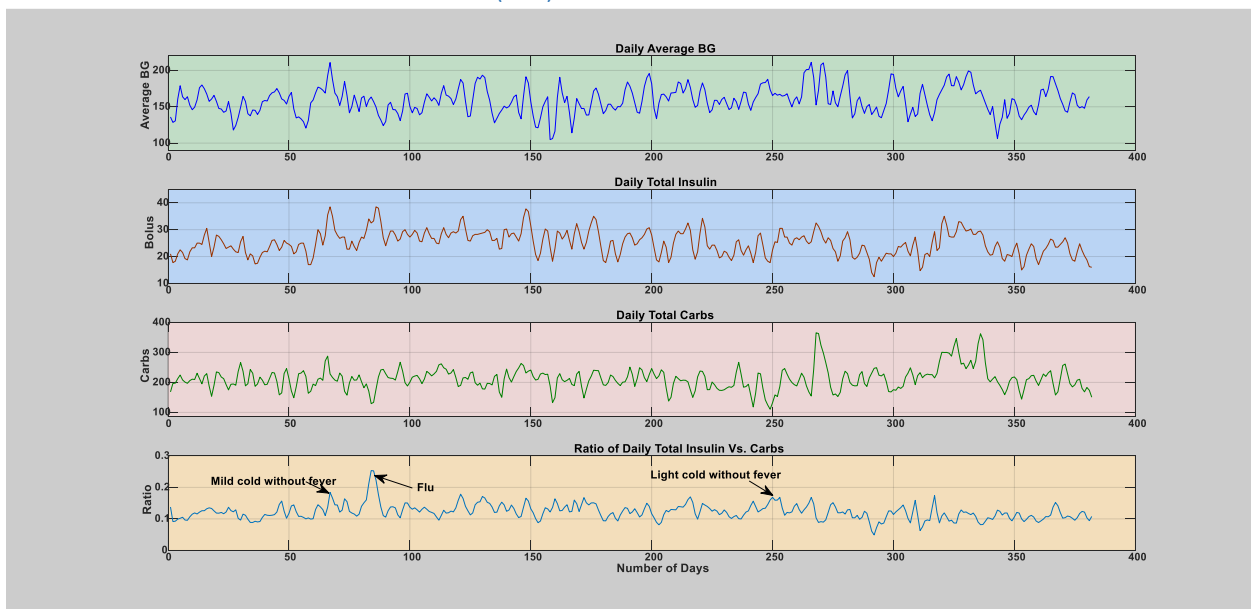




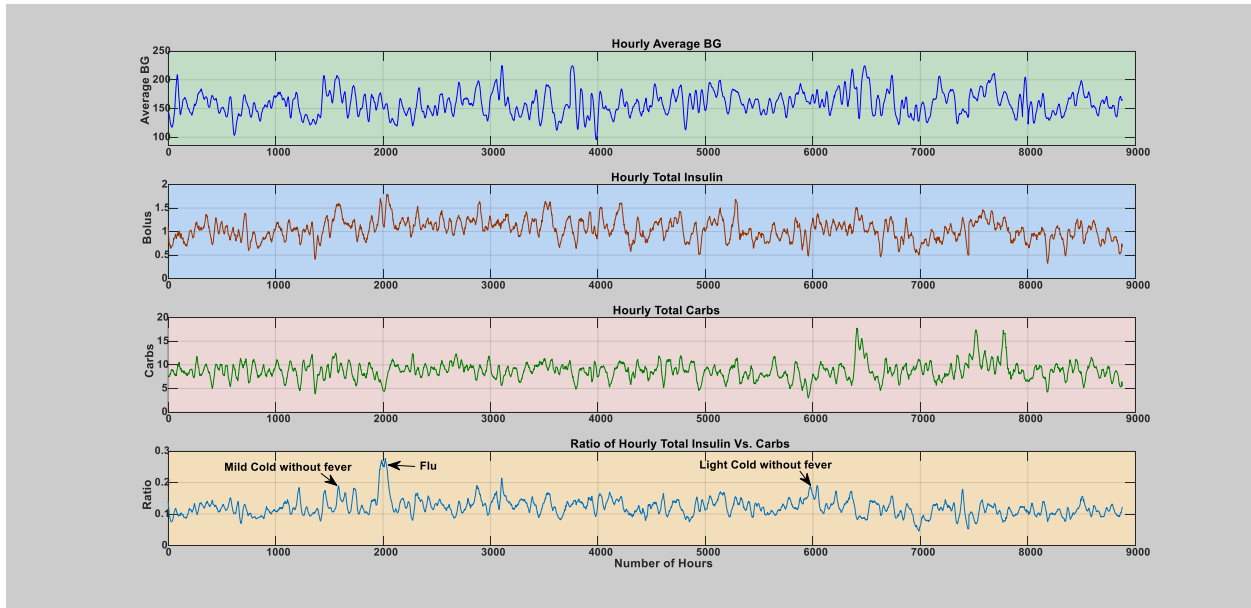
b) Hourly average BG levels, total insulin (bolus), total carbohydrate, and insulin to carbohydrate ratio based on a window size of 48 hours.

**Figure 3:** The eighth patient year, where the patient was infected with influenza (flu) starting from the last week of November. Figure (a) depicts the daily variation of BG, total insulin (bolus), carbohydrate, and insulin to carbohydrate ratio. Figure (b) depicts variation of the same variable during each hours of the day. The operating point of the patient's insulin to carbohydrate ratio had dramatically shifted and raised above the normal days and topped around 0.4 upon mid infection week.

#### 4. The fourth case of infection (Flu)



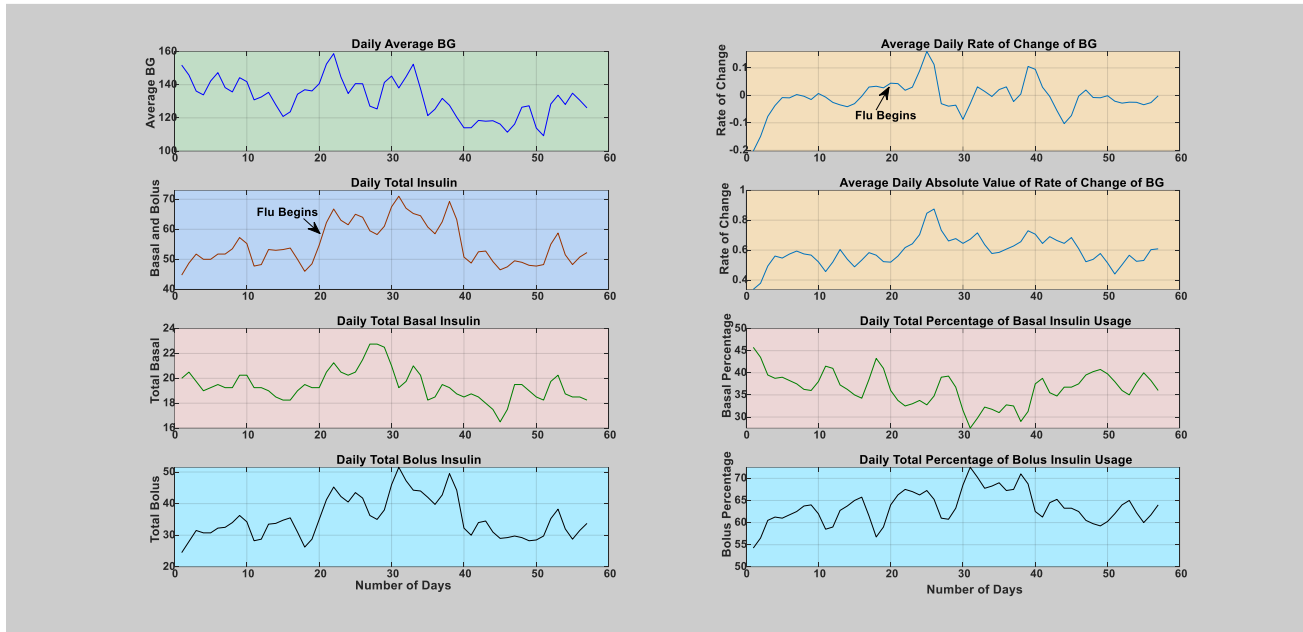
a) Daily average BG levels, total insulin (bolus), total carbohydrate, and ratio of insulin to carbohydrate based on a window size of two days.



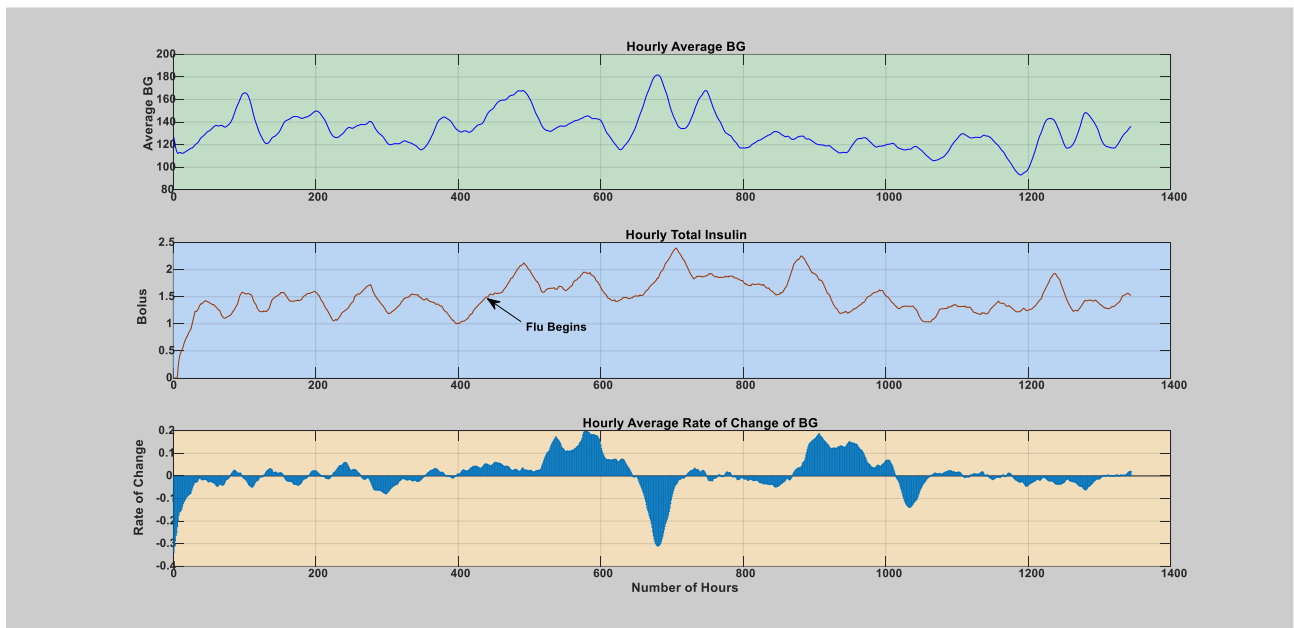
b) Hourly average BG levels, total insulin (bolus), total carbohydrate, and insulin to carbohydrate ratio based on a window size of 48 hours.

**Figure 4:** The ninth patient year, where the patient was infected with three series of acute infection (mild common cold without fever starting on the first week of august and light common cold without fever starting of mid-February, influenza (flu) starting from mid-August). Figure (a) depicts the daily variation of BG, total insulin (bolus), carbohydrate, and insulin to carbohydrate ratio. Figure (b) depicts variation of the same variable during each hours of the day. The operating point of the patient’s insulin to carbohydrate ratio had dramatically shifted and raised above the normal days and reach a top around 0.28 upon mid infection week. A light common cold without fever seems to not significantly affect the operating point.

## 5. The fifth case of infection (Flu)



- a) Daily average BG levels, total insulin including both bolus and basal insulin, daily average rate of change of CGM and absolute value of rate of change of CGM (computed based on CGM direction from the pump information), percentage of basal and bolus per total insulin units based on a window size of two days.



- b) Hourly average BG levels, total insulin (bolus), hourly average rate of change of CGM (computed based on CGM direction from the pump information) based on a window size of 48 hours.

**Figure 5:** The tenth patient year, where the patient was infected with influenza (flu) starting from mid-January. The first Figure (a) depicts the daily variation and Figure (b) depicts variation during each hours

of the day. As can be seen, there is clear and dramatic rise in insulin amount, while the BG levels remain elevated due to the ongoing infection incidence.

## Paper 2

**Woldaregay, A. Z.**, Arsand, E., Botsis, T., Albers, D., Mamykina, L., & Hartvigsen, G. (2019). *Data-Driven Blood Glucose Pattern Classification and Anomalies Detection: Machine-Learning Applications in Type 1 Diabetes*. *J Med Internet Res*, 21(5), e11030. doi:10.2196/11030

Review

# Data-Driven Blood Glucose Pattern Classification and Anomalies Detection: Machine-Learning Applications in Type 1 Diabetes

Ashenafi Zebene Woldaregay<sup>1</sup>, MSc; Eirik Årsand<sup>2</sup>, PhD; Taxiarchis Botsis<sup>3</sup>, PhD; David Albers<sup>4</sup>, PhD; Lena Mamykina<sup>4</sup>, PhD; Gunnar Hartvigsen<sup>1</sup>, PhD

<sup>1</sup>Department of Computer Science, University of Tromsø – The Arctic University of Norway, Tromsø, Norway

<sup>2</sup>Norwegian Centre for E-health Research, University Hospital of North Norway, Tromsø, Norway

<sup>3</sup>The Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University School of Medicine, Baltimore, MD, United States

<sup>4</sup>Department of Biomedical Informatics, Columbia University, New York, NY, United States

**Corresponding Author:**

Ashenafi Zebene Woldaregay, MSc  
Department of Computer Science  
University of Tromsø – The Arctic University of Norway  
Realfagbygget, Hansine Hansens vei 54  
Tromsø,  
Norway  
Phone: 47 77646444  
Email: [ashenafi.z.woldaregay@uit.no](mailto:ashenafi.z.woldaregay@uit.no)

## Abstract

**Background:** Diabetes mellitus is a chronic metabolic disorder that results in abnormal blood glucose (BG) regulations. The BG level is preferably maintained close to normality through self-management practices, which involves actively tracking BG levels and taking proper actions including adjusting diet and insulin medications. BG anomalies could be defined as any undesirable reading because of either a precisely known reason (normal cause variation) or an unknown reason (special cause variation) to the patient. Recently, machine-learning applications have been widely introduced within diabetes research in general and BG anomaly detection in particular. However, irrespective of their expanding and increasing popularity, there is a lack of up-to-date reviews that materialize the current trends in modeling options and strategies for BG anomaly classification and detection in people with diabetes.

**Objective:** This review aimed to identify, assess, and analyze the state-of-the-art machine-learning strategies and their hybrid systems focusing on BG anomaly classification and detection including glycemic variability (GV), hyperglycemia, and hypoglycemia in type 1 diabetes within the context of personalized decision support systems and BG alarm events applications, which are important constituents for optimal diabetes self-management.

**Methods:** A rigorous literature search was conducted between September 1 and October 1, 2017, and October 15 and November 5, 2018, through various Web-based databases. Peer-reviewed journals and articles were considered. Information from the selected literature was extracted based on predefined categories, which were based on previous research and further elaborated through brainstorming.

**Results:** The initial results were vetted using the title, abstract, and keywords and retrieved 496 papers. After a thorough assessment and screening, 47 articles remained, which were critically analyzed. The interrater agreement was measured using a Cohen kappa test, and disagreements were resolved through discussion. The state-of-the-art classes of machine learning have been developed and tested up to the task and achieved promising performance including artificial neural network, support vector machine, decision tree, genetic algorithm, Gaussian process regression, Bayesian neural network, deep belief network, and others.

**Conclusions:** Despite the complexity of BG dynamics, there are many attempts to capture hypoglycemia and hyperglycemia incidences and the extent of an individual's GV using different approaches. Recently, the advancement of diabetes technologies and continuous accumulation of self-collected health data have paved the way for popularity of machine learning in these tasks. According to the review, most of the identified studies used a theoretical threshold, which suffers from inter- and inpatient variation. Therefore, future studies should consider the difference among patients and also track its temporal change over time. Moreover, studies should also give more emphasis on the types of inputs used and their associated time lag. Generally, we foresee that these developments might encourage researchers to further develop and test these systems on a large-scale basis.

**KEYWORDS**

type 1 diabetes; blood glucose dynamics; anomalies detection; machine learning

## Introduction

### Background

Diabetes mellitus is a chronic metabolic disorder that results in abnormal blood glucose (BG) regulation. The BG level is maintained close to normality through self-management practices, which involves actively tracking BG levels and taking proper actions including diet and insulin medications. The estimated number of people with diabetes aged between 20 and 79 years was 415 million (uncertainty interval: 340-536 million) in 2015 and is expected to reach 642 million (uncertainty interval: 521-829 million) by 2040 [1]. The global economic burden of diabetes in adults aged between 20 and 79 years was estimated to be US \$1.31 trillion (95% CI 1.28-1.36) in 2015 [2]. The total number of deaths attributed to diabetes is estimated to be 5 million in people with diabetes aged between 20 and 79 years [1]. People with diabetes have a higher risk of getting infections as compared with the normal population, which potentially increases their morbidity and mortality [3]. The greater and frequent risk of infections is mainly correlated with a hyperglycemia environment [3,4]. Moreover, studies suggest a hypoglycemia episode could result in a higher hospitalization and mortality rate [5].

The individual's BG dynamic is affected by various factors, which are mainly categorized as common, individual, and unpredictable factors [6]. The common factors include amount of food intake, insulin intake, previous level of BG, pregnancy, drug and vitamin intake, smoking, and alcohol intake. The individual factors include dawn phenomena, physical exercise load, and menstruation. The unpredictable factors include stress, concomitant diseases, and infections [6]. Swings in BG dynamics, that is, hypoglycemia and hyperglycemia, could be generally categorized under a normal cause variation and special cause variation. The normal cause variation is regarded as caused by those common and individual factors, whereas the special cause variation is caused by those unpredictable factors. The underlying reason of the special cause variations is difficult to understand and remains a challenge for the patient during the incidences. For instance, during stress and infections, the patient usually struggles with hyperglycemia and injects frequent insulin to lower his or her BG levels.

BG anomalies could be defined as any undesirable reading because of either a precisely known reason (normal cause variation) or an unknown reason (special cause variation) to the patient [7]. Even if the advancement in self-management applications and diabetes monitoring technologies has made things easier, the challenge of BG anomalies remains to be managed by the patient themselves. There are some technological developments in the direction of personalized decision systems and BG event alarms to provide an alert and decision support to the patient in the time of these challenges. Techniques such as classification and detection of glycemic variability (GV), hypoglycemia, and hyperglycemia, in

particular, and BG anomalies, in general, are central to the development of these diabetes technologies. The ubiquitous nature and widespread use of mobile health (mHealth) apps, sensors and wearables, and other point-of-care (POC) devices for self-monitoring and management purposes have made possible the generation of automated and continuous diabetes-related data, which brought an opportunity for the introduction of machine learning and its application for intelligent and improved systems, which is capable of solving complex tasks within a dynamic knowledge and dynamic environment. In this regard, there are some reviews conducted toward the applications of artificial intelligence in diabetes-related tasks. For instance, Contreras et al [8] conducted literature reviews on the applications of artificial intelligence in the context of critical diabetes management issues such as BG prediction and strategies for BG control, adverse glycemic events detection, bolus calculators and advisory system, patient personalization (tailored features), and others [8]. Moreover, Rigla et al [9] also conducted a review to provide a general overview and popularity of artificial intelligence applications to diabetes problems. Generally, both Contreras et al [8] and Rigla et al [9] tried to demonstrate the potential of artificial intelligence with regard to all groups of people with diabetes focusing on general self-management issues. As far as our knowledge is concerned, there are almost no reviews conducted toward techniques of BG anomaly classification and detection focusing on various approaches, in general, and machine-learning applications, in particular. However, there were some reviews conducted to evaluate the significant effect of pattern management based on self-monitoring BG (SMBG) with regard to clinical practices [10]. Therefore, we suggest that there is a lack of reviews focusing on BG anomaly classification and detection. The objective of this review was to identify, assess, and analyze the state-of-the-art machine-learning strategies in BG anomaly classification and detection including GV, hyperglycemia, and hypoglycemia in people with type 1 diabetes. Moreover, it has presented the current modeling options of machine-learning applications and their hybrid systems. The review covers machine-learning approaches pertinent to personalized decision support systems and BG alarm events applications in type 1 diabetes.

### Machine Learning Tasks in Type 1 Diabetes

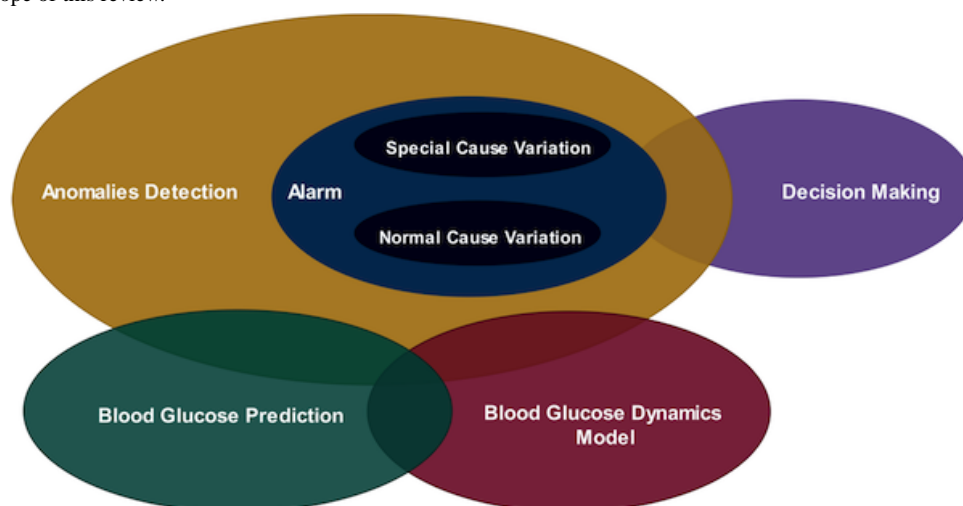
Machine-learning approaches (tasks) are generally categorized as regression, prediction, classification, detection, and clustering, which are grouped either in supervised, semisupervised, unsupervised, or reinforcement learning based on the type of learning employed. Generally, reinforcement learning is out of the scope of this review, where we mainly focus on the other 3 categories. Machine learning-based data mining tasks could be categorized as descriptive or unsupervised (ie, clustering, association, and summarization), semisupervised (ie, classification and detection), and predictive or supervised learning (ie, classification and regression) [11]. In this regard,

the most widely used machine learning–based data mining tasks in the literature are BG anomalies detection, BG prediction, modeling of BG dynamics, and decision making or education, as shown in Figure 1. In this review, we will focus on the typical applications of classification and detection tasks in diabetes research, specifically in BG anomaly detection within the context of a personalized decision support system and BG alarm events applications. The review considers various classes of machine learning algorithms: artificial neural network (ANN), decision trees (DTs), support vector machine (SVM), evolutionary algorithms (EAs), and others.

An ANN is a biologically inspired computational model consisting of a set of interconnected neurons and a scaled connection between them that is called weights [12]. On the basis of network topology, an ANN is mainly categorized as a feedforward ANN (single-layer perceptron (SLP), multi-layer perceptron (MLP), and radial basis function [RBF]) and feedback ANN (recurrent neural network [RNN], Elman net, Kohonen’s self-organizing map (SOM), and Hopfield networks) [12]. The SVM works based on the theory of structural risk

minimization principle [13]. Learning in the SVM occurs through finding an optimal hyperplane that can maximize the margin between the classes. The SVM has been widely exploited in numerous applications such as regression and prediction, pattern identification and recognition, categorization, and classification [13]. An EA is a biologically inspired approach to problem solving [14]. The 2 most used variants of EA are genetic programming (GP) and genetic algorithm (GA). Random forest (RF) or DTs are a kind of an ensemble approach of learning for different classification and regression applications, which mainly learns by constructing a multitude of DTs generating the mode of the class or mean of prediction. The hidden Markov model (HMM) is a variant of the statistical Markov model, where the system being modeled is assumed to follow a Markov property with unobserved states [15]. There are various versions of HMMs; however, in this review, we considered only those trained with a framework close to machine learning families. Hybridization is the process of combining 2 or more different approaches in parallel or serious connection, either at the preprocessing stage, feature extraction, or learning stage, when looking for an improved performance [16].

**Figure 1.** Most widely used machine learning–based data mining tasks based on self-recorded data in people with type 1 diabetes. The yellow shaded ellipse depicts the scope of this review.



**Blood Glucose Anomaly Classification and Detection**

Hawkins defined anomalies as “observations that deviate much from the other observations so as to arouse suspicions that it could be generated by a different process” [7,17,18]. There are terms that are often used interchangeably with anomalies, such as outliers, deviations, exceptions, rare instances, and irregularities. The problem of identifying and capturing anomalies in data can be supervised, semisupervised, and unsupervised tasks [19,20]. These strategies can roughly be categorized as classifier- or model-based (detection) approach. The semisupervised is better when anomalous instances are not easily available, whereas supervised techniques are more suitable when there are sufficient labeled instances of both normal and anomalous instances. The unsupervised approach does not require any reference data labels, where normal behaviors have to be determined dynamically, and the detections are mainly performed with regard to the entire datasets. The model-based strategies can be considered as a diagnosis of the system’s behavior during abnormal situations through modeling

and adequately characterizing the system’s behavior during normal situations [19,21]. It uses a system’s model to either estimate or predict the underlying system (process) dynamics to capture anomalies in the data. The most important design requirement in using a model includes discovering and characterizing what is to be considered a normal pattern of behaviors [22]. Unlike the classifier-based strategies, the model-based strategies do not require rigorous knowledge of the underlying expected anomalies, that is, to fully understand and characterize the shape and nature of the expected anomalies [22]. By simply defining what is the expected normal pattern the system should exhibit, the model-based anomaly detection is capable of detecting abnormal behavior, which is not considered as the normal behavior of the system. Defining and discovering what is *normal* is a challenging task especially for dynamic and complex systems, for example, BG dynamics. However, this is often tackled in a dynamic and complex system by relying on either a machine learning model trained on a large enough dataset or using an explicit mathematical model, for



example, physiological model of BG dynamics, of the system if it exists already.

BG readings are time series data, and anomalies in BG levels could be regarded as any undesirable readings, as shown in [Figure 1](#), because of either a predictable cause (normal cause variation) or an unpredictable cause (special cause variation). A normal cause variation could be defined as any hypoglycemia or hyperglycemia incidences with the underlying cause known to the patient herself or himself and also referred as predictable (patient controllable) factors such as insulin injection, diet intake, physical activity, and others. However, special cause variation refers to any hypoglycemia or hyperglycemia incidences with the underlying cause unknown to the patient and also called unpredictable (patient uncontrollable) factors such as stress, infections, insulin set failure, and others. The classifier, semisupervised (model)- and unsupervised-based approach could be used to solve the challenge of capturing BG anomalies caused by both the predictable factors (normal cause variation) and unpredictable cause (special cause variation). However, regarding the unpredictable factors (special cause variation), the classifier-based approach remains to be very challenging with limited feasibility as the classifier-based strategies require a thorough understanding and characterization of the nature, size, and shape of the anomalies, along with its inter- and intravariability among the patients. With the same token, the unsupervised approach could face the same challenge as it does not have any mechanisms for differentiating the one with special cause from the normal cause variations. However, the model-based (semisupervised) approach happens to be more appropriate given that it only requires to characterize what is considered to be normal so as to detect what is believed to be abnormal. For example, infection (stress)-related hyperglycemia and a diet-induced hyperglycemia are treated differently according to the model-based (semisupervised) anomaly detection strategies. In this regard, diet-induced hyperglycemia is treated as normal, as the model could describe the underlying cause (certain meal), but infection-related hyperglycemia is considered as an anomaly because the model cannot describe the underlying cause based on patient controllable variables (eg, meals and insulin).

GV measures the degree or the rate at which the patient's BG fluctuates between high and low levels [23]. GV is useful to provide all-inclusive information on one's self-management practices concerning postprandial spikes in BG, as well as episodes of hypoglycemic and hyperglycemic events [23,24], which are the main factors that contribute for a higher risk of cardiovascular events in people with diabetes. The evaluation of GV helps to comprehend and assess the effect of the patient's timely actions on the hypoglycemia and hyperglycemia incidence by associating out-of-target BG levels with patient-specific factors, such as insulin dosage, other medication, meals, activity, stress, and illness [23]. However, there is no gold standard approach for assessing GV, and despite its importance, it remains to be challenging.

### **Blood Glucose Prediction**

BG prediction is about forecasting an individual's future BG levels using current and past information and is also an

important constituent of BG anomaly classification and detection approaches. It mainly aims to provide crucial alarms for patients in advance with sufficient lead time so as to avoid further complications from hypoglycemia or hyperglycemia incidences. According to Oviedo et al [25], BG prediction models could be categorized into 3 main groups: physiological models, data-driven models, and hybrid models [25]. These categories are solely demarcated based on the necessity of extensive knowledge of the underlying BG dynamics: black box approach (data-driven model), intermediate knowledge (hybrid model), and extensive knowledge (physiological model). The data-driven model, which is mainly referred to as black box model, uses the patient's continuous glucose monitoring (CGM), insulin, dietary, and other relevant information to develop a prediction model, for example, machine learning and time series approaches. There are a variety of data-driven models developed and tested in the literature including machine learning (neural network, support vector regression, jump neural network, RNN, and others) and time series models (autoregressive [AR] with exogenous input, AR moving average with exogenous input, AR moving average, and others) [25]. Hybrid models make use of the advantages from the data-driven and physiological models [25]. Most of the hybrid models rely on the physiological model to compute meal and insulin information as input for the data-driven models [25]. Physiological models mainly rely on 3 sets of mathematical (differential) equations to describe the underlying dynamics: BG dynamics, insulin dynamics, and meal absorption dynamics. Physiological models are roughly grouped into lumped and comprehensive models based on the way the model treats each organ and tissue so as to develop the differential equations [26]. There are a variety of physiological models developed in the literature such as Berger, Hovorka, Cobelli, Lehmann and Deutsch model, and others [26]. Generally, there are plenty of models implemented in the literature on the prediction of BG levels [25,26]. However, BG prediction is not under the scope of this review, and we mainly focus on the data-driven BG pattern classification and anomaly detection approaches under the umbrella of machine learning.

## **Methods**

### **Search Strategy**

The objective of this review was to identify, assess, and analyze the state-of-the-art machine learning strategies and their hybrid system focusing on BG anomaly classification and detection including GV, hyperglycemia, and hypoglycemia in people with type 1 diabetes. The review covers machine learning approaches pertinent to personalized decision support systems and BG alarm events applications. Therefore, for the purpose of the study, a rigorous literature search was conducted between September 1 and October 1, 2017, through various Web-based databases including Google scholar, IEEE Xplore, DBLP Computer Science Bibliography, ScienceDirect, PubMed or Medline, Journal of Diabetes Science and Technology, and Diabetes Technology & Therapeutics. Additional search was also conducted between October 15 and November 5, 2018, on those databases to refine and update the records. Furthermore, the reference list of the selected articles was used to extract additional articles to get a complete overview of the field.

Peer-reviewed journals and articles published between 2000 and 2018 were considered. The inclusion and exclusion criteria were setup through rigorous discussion and brainstorming among the authors. Different combinations of terms such as *diabetes*, *intelligent system*, *hybrid system*, *machine learning*, *BG event indicators (hypo- and hyperglycemia prediction)*, *BG event alarm*, *BG personalized decision system*, *clinical*, *closed-loop system*, *hyperglycemia*, *hypoglycemia*, *GV*, and *personalized profile* were used during the search. The terms were combined using *AND/OR* for a better search strategy. Relevant articles were first identified by reviewing the title, keywords, and abstracts for a preliminary filter with our selection criteria, and then we reviewed full text articles that seemed relevant. Information from the selected literature was extracted based on some predefined categories, which were based on previous research, and further elaborated through brainstorming.

### **Inclusion and Exclusion Criteria**

To be included in the review, the studies should have developed, implemented, tested, and discussed machine learning and any of its hybrid approaches in type 1 diabetes focusing on one or more of the following application areas:

- BG anomaly detection
- Hypoglycemia prediction, classification, or detection
- Hyperglycemia prediction, classification, or detection
- Glycemic or BG variability classification or detection

Therefore, the studies that reside outside of these stated scopes were excluded from the review including all articles written in other languages but English.

### **Data Categorization and Data Collection**

Information was extracted from the selected studies based on predefined parameters (variables) and categories. The categories were defined based on rigorous brainstorming and discussion among the authors. These categories were demarcated solely to collect the relevant data and to assess, analyze, and evaluate the model's characteristics and its experimental setup.

### **Application Scenario**

This category defines the type of applications where the machine learning algorithm is being exploited. It can be hypoglycemia and hyperglycemia prediction, classification and detection, or GV classification and detection.

### **Type of Input**

This category was defined to assess, analyze, and evaluate the type of inputs used to develop the algorithm. This includes the key diabetes parameters and other physiological parameters relevant for BG anomaly classification and detections: BG, heart rate variability, and others.

### **Data Format, Type, Size, and Data Source**

This category was defined to assess, analyze, and evaluate the type of data format used as input to the algorithm. This depends on the basis of the type of diabetes technologies, mobile apps, and POC devices used for data collection and algorithm development. It includes different data formats such as from

CGM devices, mHealth apps (ie, diabetes diary), heart rate monitoring devices, and others.

### **Input Preprocessing**

This category defines the kind of preprocessing algorithm the system implements so as to avoid missing, sparse, and corrupted input data.

### **Class of Machine Learning**

This category defines the class of machine learning algorithm used to train and test the BG anomaly classification and detection algorithm. It includes different classes of machine learning algorithm: ANN, SVM, Bayesian network, DT, and others.

### **Training or Learning Method and Algorithm**

This category defines the class of learning algorithms used to train the model. It includes different training algorithms such as the backpropagation algorithm, kernel, optimization techniques, and others.

### **Performance Metrics or Evaluation Criteria**

This category defines the type of evaluation metrics used to assess the accuracy of the classification and detection algorithm implemented. It includes different performance metrics such as specificity, sensitivity, receiver operating characteristic (ROC) curves, and others.

### **Literature Evaluation**

The included literature was analyzed and evaluated based on the above defined categories and variables to uncover the state-of-the-art machine learning applications in hyperglycemia or hypoglycemia prediction, classification and detection, and GV classification and detection. It also tries to pinpoint their characteristics along with the experimental setup used to implement and test the algorithms. The first evaluation and analysis was carried out based on the type of input used to develop the algorithms to uncover the state-of-the-art inputs used in these circumstances. The second evaluation and analysis was carried out based on the various classes of machine learning used to develop these algorithms to uncover the rate of adoption and their suitability to the task. The third evaluation and analysis was carried out based on the performance metrics used to evaluate the performance of these algorithms.

## **Results**

### **Relevant Literature**

The initial hit was vetted using the title, abstract, and keywords and retrieved a total of 496 papers (DBLP Computer Science (20), Diabetes Technology & Therapeutics (23), Google Scholar (160), IEEE (215), Journal of Diabetes Science and Technology (22), PubMed Medlin (27), and ScienceDirect (29); see [Figure 2](#)). After removing duplicates from the list, 410 records remained. Then, we did an independent assessment of the articles and screening based on the inclusion and exclusion criteria, which eliminated another 215 papers, leaving 195 relevant papers. After a full-text assessment, 47 articles were left (hyperglycemia=5, glycemic variabilities=3, and hypoglycemia=39), which were critically analyzed as shown in

Figures 2 and 3. The interrater agreement was measured using a Cohen kappa test, and disagreements were resolved through discussion.

Figure 2. Flow diagram of the review process.

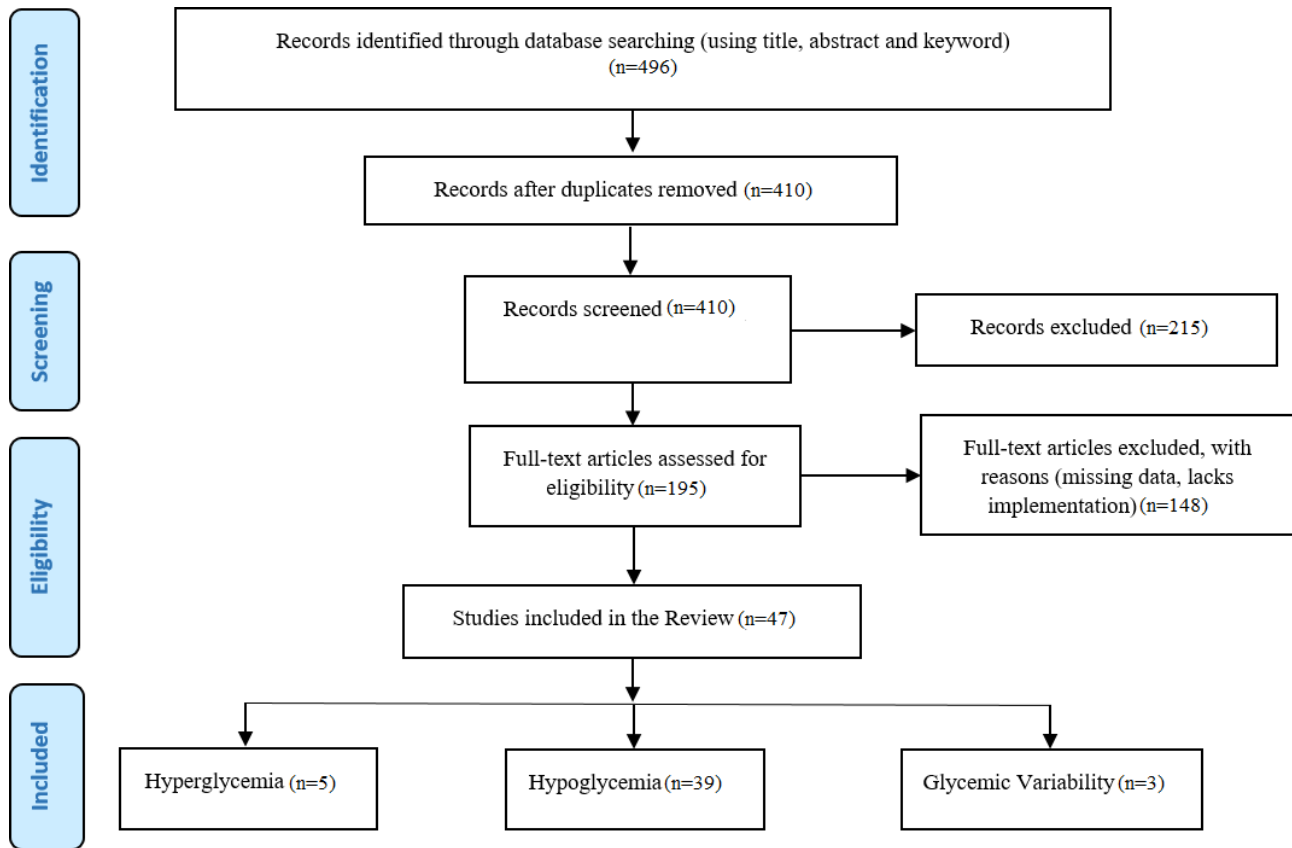
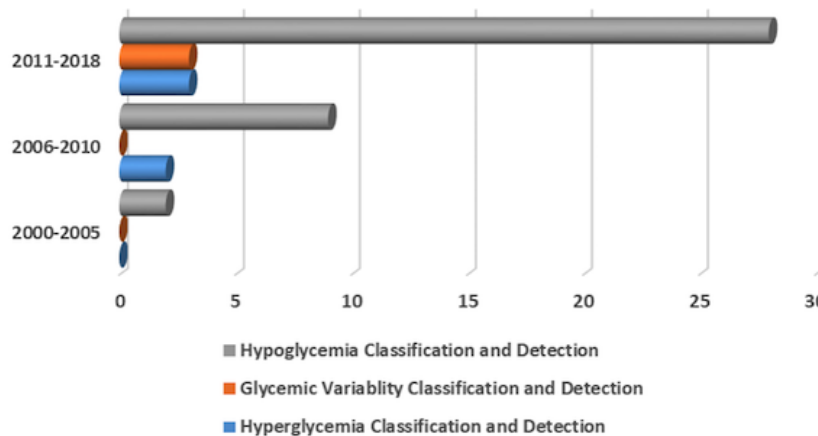


Figure 3. The number of articles published per year of publication.



**Evaluation and Analysis of the Literature**

The literature, as described previously, was evaluated based on the type of machine learning used to develop the algorithm, the type of input used to train the system, and the performance metrics used to evaluate the algorithm performance based on the tables in Multimedia Appendices 1 and 2.

**Data Characteristics and Input Parameters**

**Input Parameters**

Selecting the proper types of input parameters is one of the crucial design strategies for successful classification and

detection algorithm development. In this regard, the outer bigger ring, the middle ring, and the inner ring in Multimedia Appendix 1 depict the types of input used in hypoglycemia, hyperglycemia, and GV classification and detection algorithm, respectively. According to hypoglycemia classification and detection algorithm, BG, heart rate, and QT interval are the most used types of input parameters (25/39, 64%). BG alone is the second most used type of input parameter (4/39, 10%). BG and insulin are the third most used types of input parameters along with BG, insulin, diet, physical activity, and others (3/39, 8%). BG and diet alone, along with BG, insulin, and diet, and BG, heart rate, skin impedance, and BG, insulin, diet, heart rate, galvanic

response, skin impedance are the fifth most used types of input parameters (1/39, 3%). According to hyperglycemia classification and detection algorithm, BG alone, and BG and insulin represent the most used types of input parameters (2/5, 40%). BG, heart rate, and QT interval represent the second most

used types of input parameters (1/5, 20%). According to GV classification and detection algorithm, BG alone (3/6, 50%), and BG and insulin (3/6, 50%) are equally ranked as the most used types of input parameters, as shown in Figure 4.

**Figure 4.** Reported input features, machine learning class, and accuracy. ANN: artificial neural network; BBNN: block-based neural network; BG: blood glucose; BNN: Bayesian Neural Network; DBN: deep belief network; DT: decision tree; ELM: extreme learning machine; GA: genetic algorithm; GP: genetic programming; HMM: hidden Markov model; NAR: nonlinear autoregressive network; NARX: nonlinear autoregressive network with exogenous inputs; NBC: Naive Bayes classifier; RNN: recurrent neural network; SVM: support vector machine; VTWNN: variable translation wavelet neural network.

Study	Features								Type of Machine Learning																Performance							
	BG	Insulin	Phy. Activity	Diet	HR	QT interval	Skin Impedance	Gabranic response	DT	SVM	RNN	ANN	NARX & NAR	Fuzzy	NBC	Rule Based	ELM	HMM	DBN	Gaussian Pro.	BBNN	VTWNN	BNN	CNN	GP	GA	Hybrid	Sensitivity	Specificity	Accuracy		
[25]																													76%	58%	--	
[26]& [27]																														79%	52%	--
[28]																														86%	96%	96%
[29]																														80%	95%	98%
[30]																														--	--	--
[31]																														--	--	88%
[32]																														--	--	--
[33] & [34]																														78%	96%	--
[35] & [36]																														81%	93%	--
[37]																														80%	80%	80%
[38]																														--	--	68%
[39]																														60%	100%	85%
[40] & [41] & [42]																														75%	50%	--
[43]																														86%	80%	--
[44]																														77%	53%	--
[45]																														79%	54%	--
[46]																														78%	60%	--
[47]																														--	--	--
[48]																														80%	73%	--
[49]																														--	95%	--
[50]																														81%	41%	--
[51] & [52] & [53]																														83%	64%	--
[54] & [55]																														82%	63%	--
[56]																														75%	60%	--
[57]																														82%	60%	--
[58] & [59]																														74%	59%	63%
[60] & [61] [62]																														75%	83%	--
[63]																														77%	52%	--
[63] & [64]																														79%	50%	--
[65] & [66]																														77%	51%	--
[67]																														79%	52%	--
[68]																														77%	55%	--
[69]																														84%	52%	--
[70]																														80%	50%	--
[71]																														78%	60%	--
[72]																														--	90%	86%
[73]																														75%	55%	--
[74]																														--	--	100%
[75]																														--	--	100%
[15]																														--	--	--
[76]																														71%	65%	--
[77]																														--	--	--
[78]																														--	--	85%
[79]																														81%	74%	90%
[80]																														87%	97%	94%
[81]																														72.23%	92%	90%
[82]																														73%	60%	--

## Data Characteristics

### Data Sources

Different kinds of data sources ranging from BG monitors, physical activity, electrocardiogram (ECG), and heart rate sensors have been used in the reviewed articles for hyperglycemia, hypoglycemia, and GV classification and detection algorithms. The reviewed articles relied on different

kinds of data formats including SMBG (finger sticks), CGM, and ECG signals, as shown in [Table 1](#). Generally, ECG signal is the most used type of data format (51%), followed by CGM (39%) and SMBG (10%). Specifically, hypoglycemia classification and detection involve (CGM (n=11), ECG (n=24), and SMBG (n=5)). Regarding, hyperglycemia classification and detection (CGM (n=5) and ECG (n=1)) and GV classification and detection (CGM (n=3)).

**Table 1.** Types of data formats used in the studies (N=49).

Data type/format	Count, n (%)
Continuous glucose monitoring	19 (39)
Self-monitoring blood glucose	5 (10)
Electrocardiogram signal	25 (51)

With regard to BG monitoring, different devices and brands have been exploited for developing hypo-/hyperglycemia and GV classification and detection algorithms, as shown in [Table 2](#). Generally, Yellow Spring Instruments is the most used device (50%) followed by Guardian Real Time (MinMed CGM; 28%). GlycoMark (7%) is the third most used device followed by HemoCue Glucose 201 (5%) and Self-Monitored BG (5%). Specifically, the most used devices for hypoglycemia classification and detection are Guardian Real Time (MinMed, CGM; n=7), Yellow Spring Instruments (n=21), HemoCue

Glucose 201 (n=2), Dexcom CGM system (n=1), Self-Monitored BG (SMBG; n=2), Medtronic Enlite CGM sensors (n=1), Medtronic insulin pump (n=4), SensorWear armband (physical activity; n=2), and Basis Peak fitness band (n=1). and Basis Peak fitness band (n=1). As for hyperglycemia classification and detection, Guardian Real Time (MinMed CGM; n=2) and Medtronic insulin pump (n=3) had been used. With regard to GV classification and detection, GlycoMark (n=3), Guardian Real Time (MinMed CGM; n=3), and Medtronic insulin pump (n=3) had been used.

**Table 2.** Types of devices used for the monitoring of blood glucose levels (N=42).

Devices	Count, n (%)
Guardian Real Time (MinMed, CGM <sup>a</sup> )	12 (28)
HemoCue Glucose 201 (HemoCue)	2 (5)
Yellow Spring Instruments	21 (50)
Dexcom CGM system	1 (3)
Medtronic Enlite CGM sensors	1 (3)
GlycoMark	3 (7)
Self-Monitored Blood Glucose-unknown device	2 (5)

<sup>a</sup>CGM: continuous glucose monitoring.

Various brands of physiological monitoring (heart rate and ECG signals) devices have been exploited in the reviewed articles. Generally, as shown in [Table 3](#), Compumedics system is the most used system (52%) followed by a customized device such as a battery-powered chest belt-worn device (22%). HypoMon is the third most used device (13%) followed by Basis Peak fitness band (9%) and a self-designed portable apparatus (4%). Specifically, for hypoglycemia classification and detection

purposes, various devices have been used such as HypoMon (n=3), Basis Peak (n=2), Compumedics system (n=11), a battery-powered chest belt-worn (n=5), and self-designed portable apparatus (n=1). With regard to hyperglycemia classification and detection, only 1 article has used the Compumedics system (n=1), which indicates that heart rate and ECG signals have a limited use in this case.

**Table 3.** Types of devices used for the monitoring of physiological parameters (heart rate and electrocardiogram signals; N=23).

Devices	Count, n (%)
HypoMon	3 (13)
Basis peak fitness band	2 (9)
Compumedics system	12 (52)
A battery-powered chest belt-worn (customized)	5 (22)
Self-designed portable apparatus (customized)	1 (4)

### Data Preprocessing

Data preprocessing is an important stage of any machine learning strategy. In this regard, there were various kinds of data preprocessing strategies used in the reviewed articles. The reviewed articles had relied on both BG and other physiological (heart rate, ECG, skin impedance, and others) data, which of course involves different preprocessing strategies depending on the data type under consideration. Regarding the BG data, various preprocessing approaches had been used including differencing (derivative) BG values [27,28], CGM data reconstruction, or smoothing using different methods such as spline interpolation [29-33], a rough feature elimination, such as fast *separability* and correlation analysis algorithm [28,29], representing BG temporal change information [34], feature selection and feature ranking [35], filtering using Pearson's correlation coefficient (PCC) and the *t* test, and the wrapper approach using greedy backward elimination [33]. The other physiological parameters (heart rate, ECG, skin impedance, and others) had been preprocessed using different methods such as normalization [36-38], feature extraction and selection [39,40], feature extraction using fast Fourier transform (FFT) [41], unsupervised restricted Boltzmann machine-based feature representation [42], filtering techniques such as Infinite impulse response high pass filter [41,43], correlation analysis [44-46], and transformation of frequency domain into time domain (FFT) [47].

### Class of Machine Learning

#### Hypoglycemia Classification and Detection

Different classes of machine learning techniques have been adopted in hypoglycemia prediction, classification, and detection algorithms to predict, classify, and detect the incoming hypoglycemia incident in people with type 1 diabetes, as shown in Figure 4. Conventional feedforward ANN is the most adopted class of machine learning, which is used in 26% (17/65) of the studies, as shown in Multimedia Appendix 1. Hybridization of machine learning techniques with other approaches such as time series, fuzzy logic, and others are the second most adopted approach (12/65, 18%). The SVM ranked the third most adopted class of machine learning (9/65, 14%). DT ranked the fourth most adopted technique (4/65, 6%). GA, time delay ANN and time sensitive ANN, block-based neural network (BBNN), and adaptive neural fuzzy inference system (ANFIS) are the fifth most used classes of machine learning (3/65, 5%). Nonlinear autoregressive network with exogenous inputs (NARX) and nonlinear autoregressive network (NAR) along with Gaussian process regression, combinational neural logic network, and Bayesian neural network (BNN) ranked as the sixth most used

classes of machine learning (2/65, 3%). Deep belief network (DBN), radial basis function neural network (RBFNN), and variable translation wavelet neural network (VTWNN) are the seventh most used classes of machine learning (1/65, 2%).

#### Hyperglycemia Classification and Detection

Hyperglycemia classification, prediction, and detection has been practiced less when compared with hypoglycemia, which might be linked because of its less severe short-term complications as opposed to hypoglycemia incidences. However, irrespective of this limitation, different types of machine learning techniques have been adopted, as shown in Figure 4. For example, ANN is the most used machine learning technique in 34% (3/9) of the studies (feedforward (1/9) and feedback RNN (2/9)), as shown in Multimedia Appendix 1 along with EA (3/9,34%) (GA (1/9) and GP (2/9)). The HMM (2/9, 22%) is the third most used followed by a hybrid approach (1/9, 11%).

#### Glycemic Variability Classification and Detection

GV detection is a recent development, which has great importance in quantifying factors associated with hypo-/hyperglycemia incidence. In this regard, there is some research and development involving machine learning techniques, as shown in Figure 4. For example, feedforward ANN is the most used class of machine learning (3/8, 37%), as shown in Multimedia Appendix 1. Naive Bayes classifier (NBC) and SVM are the second most adopted techniques of machine learning (2/8, 25%). DT is the third most used class of machine learning (1/8, 13%).

#### Performance Metrics

The performance metrics used in the evaluation of hypoglycemia, hyperglycemia, and GV classification and detection algorithms are depicted in the outer ring, the middle ring, and inner ring, respectively, as shown in Multimedia Appendix 1. According to hypoglycemia classification and detection, sensitivity, and specificity are the most used performance metrics (37/58, 64%). Accuracy and precision are the second most used performance metrics (9/58, 15%). Root mean square error and mean square error are the third most used performance metrics (4/58, 7%). Geometric mean is the fourth most used performance metric (3/58, 5%). Correlation coefficient is the fifth most used performance metric (2/58, 3%). Time lag (TL), recall, and ROC curve are the sixth most used performance metrics (1/58, 2%). According to hyperglycemia classification and detection, accuracy and precision, root mean square error and mean square error, time lag (TL), correlation coefficient, recall, and false positive rate are the most used performance metrics (2/15, 13%). ROC curve, geometric mean, sensitivity, and specificity are the third most used performance

metrics (1/15, 7%). According to GV classification and detection, accuracy, and precision are the most used performance metrics (3/5, 60%). Sensitivity and specificity are the second most used performance metrics (2/5, 40%).

## Discussion

### Principal Findings

The objective of this review was to identify, assess, and analyze the state-of-the-art machine applications in BG pattern classifications and anomaly detection: hyperglycemia, hypoglycemia, and GV classification and detection. According to the reviewed literature, the anomaly classification and detection approach could be roughly categorized as either a classifier-based or a model-based approach [19,21]. The classifier-based approach mainly relies on using either a specified threshold or some kinds of rules to classify the BG levels as either normal or abnormal. The difference is that unlike the model-based approach, the classifier-based approach requires rigorous and deeper knowledge regarding the nature, size, and shape of the underlying anomalies under consideration so as to develop the necessary threshold or rule to capture them. However, the model-based approach only requires to demarcate the boundary of what is known to be normal so as to capture what is believed to be abnormal [21]. The model-based approach does not require rigorous knowledge of the underlying expected anomalies, that is, to fully understand and characterize the shape and nature of the expected anomalies [22]. By simply defining what is the expected normal pattern that the system should exhibit, a model-based approach is capable of detecting abnormal behavior, which is not considered as the normal behavior of the system. Defining and discovering what is *normal* is a challenging task especially for dynamic and complex systems, for example, BG dynamics. However, this is often tackled in a dynamic and complex system by either relying on a machine learning model trained on large enough datasets or using an explicit mathematical model of the system such as a physiological or compartmental BG dynamics model [21].

Various classes of machine learning algorithms have been adopted for the task. Regarding hypoglycemia classification and detection, feedforward ANN, hybrid systems, SVM, DT, GA, ANFIS, NARX, NAR, Gaussian process regression, DBN, and BNN have been developed and tested. These techniques have explored various kinds of input parameters notably BG, heart rate, QT interval, insulin, diet, physical activity, galvanic response, and skin impedance. Concerning hyperglycemia classification and detection, RNN, GP, HMM, feedforward ANN, GA, and hybrid systems have been developed and tested, exploring various types of input parameters including BG, insulin, heart rate, and QT interval. GV detection is a recent development, which has great importance in quantifying factors associated with hypoglycemia and hyperglycemia incidence. In this regard, there is some research and development involving machine learning techniques. For example, feedforward ANN, NBC, SVM, and DT have been tested up to the task using BG and insulin delivery profiles.

Generally, all of the studies have relied on either indirect indicator variables such as heart rate, QT interval, and others

or a subset of input parameters that affects BG dynamics. The patient's contextual information, for example, meals, physical activity, insulin, and sleep, have a significant effect on BG dynamics, and a proper anomaly classification and detection algorithm should consider the effects of these parameters. In this regard, however, the individual patient is expected to record meal, insulin, and physical activity data. One of the main limitations is meal modeling, where most of the algorithm depends on the individual estimation of carbohydrate, which is prone to errors and further aggravates the degradation to detection performance. With regard to physical activity, there are various wearables and sensors that can record the individual's physical activity load and durations. However, there is the lack of a uniform approach among the studies with certain limitations on the way these signals are employed in the classification and detection algorithms. For example, there are some studies that consider levels of activity as low, moderate, and high and others consider descriptive features by summarizing the number, intensity, steps, exercise durations, and others to better quantify the effect of physical activities. Moreover, recording insulin dosage has its inherent limitations, which might affect the detection performance. For example, blockage of insulin flow from the insulin pump because of the infusion set failure and error incurred during manual registrations might pose a significant challenge in the performance of the detection system. Furthermore, CGM is becoming one of the most important components in these classification and detection algorithms. However, even if CGM advancement has enabled patients to have continuous estimation of their subcutaneous glucose levels, it has limitations when used in a personalized detection system (an alarm). In this regard, recent studies have showed that autocorrelation of the CGM reading vanishes after 30 min, making the detection performance to degrade afterward. These findings suggest that any classification and detection algorithms aiming for a better lead time should consider other patient's contextual information and various features of the CGM itself. There are some studies that develop a model by assessing several features of the CGM signal so as to compensate for its inaccuracy. Moreover, CGM is found to be inaccurate during hypoglycemia episodes, that is, insulin-induced hypoglycemia versus spontaneous hypoglycemia. In this regard, insulin-induced hypoglycemia is found to be difficult to detect as compared with spontaneous hypoglycemia. Fast occurring hypoglycemia is difficult to detect because of the blood-interstitial delay, which makes them important features to be detected by a given model. Furthermore, CGM calibration frequency and timing also affects the performance of the classification and detection algorithm.

The reviewed studies are limited to and could be roughly categorized by age groups (children, young, adult, and old), time of the day (diurnal vs nocturnal) and configurations (online vs offline). For example, most of the studies consider nocturnal hypoglycemia detection, considering the fact that most of hypoglycemia crises occurred during nighttime and also the crises during this time have a bad consequence as compared with the diurnal period. Moreover, it is a fact that nocturnal detection is simpler as compared with the diurnal considering the dynamics of the patients. However, irrespective of these challenges, there are also studies that consider the diurnal period.

However, there are limited studies that attempt to develop an algorithm that could detect anomalies in both of those contexts. With regard to the age group, most surveys reported that age group has a great effect on BG dynamics, which is typically related with the dynamics and active lifestyle adopted by each group. Therefore, it is deemed a necessary approach to consider a personalized algorithm for each age group. With regard to the configuration, there are fewer attempts of online (real-time) algorithms, where almost all of the algorithms were tested and implemented in the offline mode. In this regard, the most crucial issues concerning machine learning strategies could be the necessity of frequent retraining when subjected to a real-time and dynamic task.

In addition, the most important component in classification and detection algorithms is the threshold used to differentiate the normal from the abnormal. In this regard, almost most of the studies have used a static threshold based on suggestions either from the literature or physicians and other concerned bodies such as the American Diabetes Association. However, the critical issues in this approach are that the threshold might vary from patient to patient and also some patients might not feel any symptoms at the specified threshold (when using indirect indicators such as heart rate, QT interval, and others). However, there are some studies that employed a fuzzy logic-based approach by having a continuous decision space.

In principle, any future BG anomaly classification and detection algorithm should be expected to detect any upcoming anomalies as soon as possible (lead time—giving more response time), avoid any false alarm at any cost, perform in real-time (in an online fashion), adapt with the dynamics of BG evolution (learn continuously), automatically tune its parameters without user intervention, be able to perform throughout the day in a free living condition (diurnal and nocturnal periods), and incorporate as many input variables to better capture the dynamics. In this regard, for example, the most crucial issues concerning a real-time (online) machine learning algorithm could be the necessity of frequent retraining when subject to a real-time and dynamic task. Moreover, developing a model that considers a real-time and adaption-to-free-living condition needs to incorporate a wide range of parameters that affect BG dynamics. Furthermore, it should properly consider and address the inherent technological limitation that affects the performance of the detection algorithm. Almost all of the studies need a proper clinical validation to be integrated into a smartphone and CGM for a real-time application. This can be better described by looking at the number of samples used and their validation strategies (see [Multimedia Appendix 2](#)). Therefore, future studies should give more emphasis on clinical validation by taking a sufficient number of subjects in the development and testing phase so as to better quantify the inter- and intravariability among patients. In addition, the most crucial concept of justifying and reporting the underlying cause, as because of either patient controllable or patient uncontrollable parameters, for the detected anomalies is not addressed in any of the reviewed literature. For example, the underlying cause of hyperglycemia incidences could be patient controllable parameters such as diet or patient uncontrollable parameters such as stress and infections. Therefore, in this regard, a proper

hyperglycemia classification and detection system might be expected to be able to identify and report the underlying cause, which has a greater significance to the patient especially during infection crises.

## Summary of Existing Efforts: Machine Learning Techniques

### *Artificial Neural Network*

There are various types of ANNs used in solving BG classification and anomaly detection tasks: hypoglycemia, hyperglycemia, and GV classification and detection. Regarding hypoglycemia classification and detection, for instance, Eljil et al [48], had proposed a special type of ANN known as the time-sensitive ANN and compared the result with a time delay neural network, NARX, distributed time delay neural network, and NAR. San et al [37,49] proposed an evolvable BBNN and compared the result with feedforward ANNs and multiple regression. Moreover, San et al [42] proposed a DBN and compared the result with a wavelet neural network, a feedforward ANN, and multiple regression models. Some of the studies have also investigated the advantage of having a separate feature extraction and classification unit. In this regard, for example, both Laione et al [47] and Nguyen et al [41,50] have proposed an ANN using FFT for data extraction. Nguyen et al [41] have further trained the network through a 2-step process that combines the advantage of GA and the Levenberg Marquardt algorithm. Chan et al and Yan et al [51,52] also proposed a neural network-based rule discovery system that consisted of a neural network-based classification unit and rule-based extraction unit. There are some studies that optimized the ANN parameters through a particle swarm optimization technique. For example, Ling et al [53], Phyo et al [36,54,55], and San et al [56] proposed a new hybrid rough neural network, a VTWNN, a normalized RBFNN, and a combinational neural logic network with the neural logic AND, OR, and NOT gates, respectively, where the design parameters of the network were optimized through a hybrid particle swarm optimization with wavelet mutation operation. Moreover, Nguyen et al [43,57] also proposed an ANN that is optimized through a standard particle swarm optimization strategy. Furthermore, some studies have investigated extreme learning machines (ELMs). For instance, Ling et al [58] and San et al [59] proposed a feedforward ANN trained through an ELM and compared the result with a feedforward ANN optimized through particle swarm optimization, multiple regression-based fuzzy inference system, fuzzy inference system, and linear multiple regression. Mo et al [60] have also used ELMs and regularized the ELMs on CGM data. In addition, Nguyen et al [61–63] and Ngo et al [64] had proposed an optimal BNN algorithm using feedforward ANN architecture. There are some studies that tried to integrate a physiological model with ANN. For instance, Bertachi et al [65] integrated the physiological model of an individual diabetes patient with an ANN to predict nocturnal hypoglycemia events. Regarding, hyperglycemia classification and detection, there is only 1 study by Nguyen et al [38] that uses a feedforward multilayer ANN trained using different training algorithms, that is, gradient descent, gradient descent with momentum, scaled conjugate gradient, and resilient back propagation. Regarding



GV classification and detection, the reviewed studies had been performed either for detection purposes or for automated metrics purposes. For the detection purpose, for example, Wiley et al [33] proposed Naive Bayes (NB), multilayer perceptron (MLP) ANN, and SVM models to detect excessive GV on CGM data and compared the accuracy of the result with the other 2 diabetes experts. Concerning the automated metrics, Marling et al [66] had developed an NBC (probabilistic reasoning), an MLP ANN, and a logistic model tree (DT built using logistic regression), which could be used to monitor CGM data. Moreover, Marling et al [32] also proposed an MLP ANN and support vector regression to develop a consensus perceived GV metric.

### ***Support Vector Machines, Kernel Function, and Gaussian Process Regression***

SVM, kernel function (KF), and Gaussian process regression have been exploited for hypoglycemia classification and detection purposes in the reviewed literature. For example, Georga et al [67] developed a support vector regression for hypoglycemia prediction and compared the performance with a feedforward MLP ANN and Gaussian process regression. Georga et al [68] also proposed support vector regression and Gaussian process regression for BG prediction so as to indicate the daily incidences of hyperglycemia and hypoglycemia to the patients as well as provision of decision support to physicians in making the decision about treatment and risk of complications. Moreover, Jensen et al [29,30] developed an automatic pattern recognition system so as to detect hypoglycemia incidences retrospectively using CGM data and thereby to foster a thorough evaluation of past events and discussion with their caregivers. Jensen et al [28,69] also proposed a real-time pattern classification model by using several features from the CGM data so as to detect hypoglycemia incidences in real-time. Furthermore, Marling et al [70] proposed a hypoglycemia detection algorithm that incorporates noninvasive sensor data from fitness bands and also compared different kernels for the task: linear, Gaussian, and quadratic kernels. Nuryani et al [71] also proposed a swarm-based SVM algorithm using the repolarization variabilities as input so as to detect hypoglycemia incidences.

### ***Genetic Programming and Genetic Algorithm***

There is little visibility of GP and GA usage in their nonhybrid form for BG classification and anomaly detection tasks: hypoglycemia, hyperglycemia, and GV classification and detection. However, there are some studies that use these techniques in their hybrid form. For example, Ling et al [44,72,73] developed a hypoglycemia detection algorithm using a GA-based multiple regression coupled with a fuzzy inference system. The study exploited the GA so as to optimize the fuzzy rules, membership function of the fuzzy inference system, and also model parameters of the regression.

### ***Random Forest***

RF and DT have been mostly used in the context of hypoglycemia classification and detection tasks. For example, Eljil et al [27] proposed DTs using different techniques, namely, C4.5, J4.8, REPTree, bagging, and the cost-sensitive version of J4.8. Jung et al [74] also proposed DTs using new predictor

variables using CGM data. Moreover, Jung et al [34] proposed a DT- and SVM-based prediction model using self-monitored BG. Zhang et al [35] also proposed a new approach using the classification tree to predict the occurrences of acute hypoglycemia during intravenous insulin infusion before the actual hypoglycemic events take place.

### ***Hidden Markov Model***

Generally, HMM is used to model an environment that could better describe the evolution of the individual BG dynamics. In this regard, there are some studies that use HMM to develop model-based BG anomaly classification and detection algorithms. For example, Zhu et al [15,75] studied an approach for automatic detection of anomalies in individual BG data, using a model trained with historical data containing daily normal measurements. The trained Markovian world tries to analyze the incoming BG data and flags if it deviates from what is known by the model.

### ***Hybrid and Ensemble Models***

Hybridization approaches have been extensively used when looking for performance improvement by exploiting the advantage from 2 or more different approaches [16]. In this regard, there are some attempts in the reviewed articles which tried merging different approaches for enhanced performance in hypoglycemia classification and detection. For example, hybridization of an ANN with other techniques is demonstrated in some of these studies. Chan et al [76] developed a hybrid system that consisted of an ANN and a GA and also compared the performance with MLP ANN and classical statistical algorithms. Ghevondian et al [77] proposed a novel hybrid system of a fuzzy neural network ANN estimator to predict the BG profile and hypoglycemia incidences. San et al [78] proposed a hybrid system using an ANFIS and compared the performance with the wavelet neural network, feedforward ANN, and multiple regression. There is also some literature that tries to hybridize the SVM with other techniques. For example, Nuryani et al [39,79] proposed a hybrid fuzzy SVM and investigated the applicability of 3 KFs: radial basis, exponential radial basis, and polynomial function for the task. Moreover, Nuryani et al [40,80] also further developed a novel strategy using a hybrid particle swarm-based fuzzy SVM technique. Fuzzy reasoning models are also tested in some of the studies. For example, Ling et al [81] developed a hybrid particle swarm-optimization-based fuzzy reasoning model, where the fuzzy rules and the fuzzy-membership functions are optimized through a hybrid particle swarm optimization with wavelet mutation. The model is also compared with feedforward ANN and multiple-regression models. Mathews et al [46] developed a hybrid model using a fuzzy inference system with multiple regression, where the fuzzy rules are optimized through a GA. The study also compares the performance of the developed system with an ANN whose parameters are optimized through particle swarm optimization. In addition, San et al [82] proposed a hybrid system based on rough sets concepts and neural computing. The study has compared various hybrid approaches trained through hybrid particle swarm optimization with wavelet mutation including the rough BBNN, BBNN, rough feedforward ANN, wavelet neural network, SVM with an RBF, and

conventional feedforward ANN. Ling et al [45] also proposed an alarm system based on the hybrid neural logic network with multiple regression. Lai et al [83] developed a fuzzy inference system for hypoglycemia detection, where the system parameters are optimized through an intelligent optimizer.

Owing to the complexity of BG dynamics, it remains difficult to achieve an accurate result in every circumstance. One model can have better accuracy in some circumstances and the other model can achieve better accuracy where the first model fails to achieve a comparable result. Therefore, it is natural to look for possibilities to exploit the strengths from these different models to achieve better accuracy in most of the circumstances, which lead to ensemble approaches [16]. An ensemble approach is generally favored when one is interested to merge 2 or more different models for improved performance. In this regard, there are some studies that try to combine 2 different models looking for performance improvement in the overall system. In this regard, Daskalaki et al [84] proposed an early warning system, for both hyperglycemia and hypoglycemia, using RNN and AR with output correction module models. Moreover, the study investigated the performance improvement from the combined use of both RNN and AR with an output correction module. Moreover, Botwey et al [31] proposed combining an AR model with output correction and an RNN based on different data fusion schemes including the Dempster-Shafer evidential theory, GAs, and GP.

## Conclusions

Despite the complexity of BG dynamics, there are many attempts to capture hypoglycemia and hyperglycemia incidences and the extent of an individual GV using different approaches. Recently, because of the ubiquitous nature of self-management mHealth apps, sensors and wearables have paved the way for the continuous accumulation of self-collected health data, which in turn contributed for the widespread research of machine learning applications in these tasks. In the reviewed articles, generally, the anomaly classification and detection approaches could be categorized as either model (process)-based or

classifier (rule)-based approaches. Hypoglycemia classification and detection has been given more attention than hyperglycemia and GV detection, which might be because of its serious complication and the comparable complexity involved. The state-of-the-art indicates that various classes of machine learning have been developed and tested in these tasks. Regarding hypoglycemia classification and detection, feedforward ANNs, hybrid systems, SVM, DT, GA, adaptive neural fuzzy inference system, NARX, and NAR, Gaussian process regression, DBN, and BNN have been developed and tested. These techniques have explored various kinds of input parameters, notably BG, heart rate, QT interval, insulin, diet, physical activity, galvanic response, and skin impedance. Concerning hyperglycemia classification and detection, RNN, GP, HMM, feedforward ANN, GA, and hybrid systems have been developed and tested, exploring various types of input parameters including BG, insulin, heart rate, and QT interval. GV detection is a recent development, which has great importance in quantifying factors associated with hypoglycemia and hyperglycemia incidence. In this regard, there is some research and development involving machine learning techniques, for example, the feedforward ANN, NBC, and SVM.

Most of these studies have used a theoretical threshold suggested either by the literature or physicians and various concerned bodies such as the American Diabetes Association. However, the problem here is that some patients might feel no symptoms at the specified threshold, and it may vary from patient to patient. Therefore, a model should consider such differences among the patients (intra- and intervariability) and also track its temporal change over time. Moreover, the studies should give more emphasis on the TL and various types of inputs used. Furthermore, researchers should give proper emphasis to develop anomaly classification and detection models, which are capable of justifying and reporting the underlying cause, as either due to patient controllable or patient uncontrollable parameters. Generally, we foresee that these developments might encourage researchers to further develop and test these systems on a large-scale basis.

---

## Acknowledgments

The project is funded by the University of Tromsø – The Arctic University of Norway and National Library of Medicine “Mechanistic machine learning” (grant number: LM012734).

---

## Conflicts of Interest

None declared.

---

## Multimedia Appendix 1

Analysis of reported parameters, data characteristics, class of machine learning, and performance metrics.

[\[PDF File \(Adobe PDF File\), 201KB-Multimedia Appendix 1\]](#)

---

## Multimedia Appendix 2

Detail on reported accuracy, inputs and performance metrics used, and machine learning categorization.

[\[PDF File \(Adobe PDF File\), 281KB-Multimedia Appendix 2\]](#)

---

## References

1. Ogurtsova K, da Rocha Fernandes JD, Huang Y, Linnenkamp U, Guariguata L, Cho NH, et al. IDF Diabetes Atlas: Global estimates for the prevalence of diabetes for 2015 and 2040. *Diabetes Res Clin Pract* 2017 Jun;128:40-50. [doi: [10.1016/j.diabres.2017.03.024](https://doi.org/10.1016/j.diabres.2017.03.024)] [Medline: [28437734](https://pubmed.ncbi.nlm.nih.gov/28437734/)]
2. Bommer C, Heesemann E, Sagalova V, Manne-Goehler J, Atun R, Bärnighausen T, et al. The global economic burden of diabetes in adults aged 20-79 years: a cost-of-illness study. *Lancet Diabetes Endocrinol* 2017 Dec;5(6):423-430. [doi: [10.1016/S2213-8587\(17\)30097-9](https://doi.org/10.1016/S2213-8587(17)30097-9)] [Medline: [28456416](https://pubmed.ncbi.nlm.nih.gov/28456416/)]
3. Casqueiro J, Casqueiro J, Alves C. Infections in patients with diabetes mellitus: a review of pathogenesis. *Indian J Endocrinol Metab* 2012 Mar;16(Suppl 1):S27-S36 [FREE Full text] [doi: [10.4103/2230-8210.94253](https://doi.org/10.4103/2230-8210.94253)] [Medline: [22701840](https://pubmed.ncbi.nlm.nih.gov/22701840/)]
4. Knapp S. Diabetes and infection: is there a link?--A mini-review. *Gerontology* 2013;59(2):99-104 [FREE Full text] [doi: [10.1159/000345107](https://doi.org/10.1159/000345107)] [Medline: [23182884](https://pubmed.ncbi.nlm.nih.gov/23182884/)]
5. McCoy R, van Houten HK, Ziegenfuss JY, Shah ND, Wermers RA, Smith SA. Increased mortality of patients with diabetes reporting severe hypoglycemia. *Diabetes Care* 2012 Sep;35(9):1897-1901 [FREE Full text] [doi: [10.2337/dc11-2054](https://doi.org/10.2337/dc11-2054)] [Medline: [22699297](https://pubmed.ncbi.nlm.nih.gov/22699297/)]
6. Bazaev N, Pletenev AN, Pozhar KV. Classification of factors affecting blood glucose concentration dynamics. *Biomed Eng* 2013 Jul 18;47(2):100-103. [doi: [10.1007/s10527-013-9344-7](https://doi.org/10.1007/s10527-013-9344-7)]
7. Ahmad S, Purdy S. arXiv. 2016. Real-Time Anomaly Detection for Streaming Analytics URL:<https://arxiv.org/abs/1607.02480> [accessed 2019-04-02] [WebCite Cache ID 77Kc7RB7g]
8. Contreras I, Vehi J. Artificial intelligence for diabetes management and decision support: literature review. *J Med Internet Res* 2018 May 30;20(5):e10775 [FREE Full text] [doi: [10.2196/10775](https://doi.org/10.2196/10775)] [Medline: [29848472](https://pubmed.ncbi.nlm.nih.gov/29848472/)]
9. Rigla M, García-Sáez G, Pons B, Hernando ME. Artificial intelligence methodologies and their application to diabetes. *J Diabetes Sci Technol* 2018 Dec;12(2):303-310 [FREE Full text] [doi: [10.1177/1932296817710475](https://doi.org/10.1177/1932296817710475)] [Medline: [28539087](https://pubmed.ncbi.nlm.nih.gov/28539087/)]
10. Choudhary J, Genovese S, Reach G. Blood glucose pattern management in diabetes: creating order from disorder. *J Diabetes Sci Technol* 2013 Nov 1;7(6):1575-1584 [FREE Full text] [doi: [10.1177/193229681300700618](https://doi.org/10.1177/193229681300700618)] [Medline: [24351184](https://pubmed.ncbi.nlm.nih.gov/24351184/)]
11. Banaee H, Ahmed MU, Loufi A. Data mining for wearable sensors in health monitoring systems: a review of recent trends and challenges. *Sensors (Basel)* 2013 Dec 17;13(12):17472-17500 [FREE Full text] [doi: [10.3390/s131217472](https://doi.org/10.3390/s131217472)] [Medline: [24351646](https://pubmed.ncbi.nlm.nih.gov/24351646/)]
12. Shanmuganathan S, Samarasinghe S, editors. Artificial neural network modelling. In: *Studies in Computational Intelligence Vol 628*. Switzerland: Springer International Publishing; 2016.
13. Campbell C, Ying Y. Learning with Support Vector Machines: Synthesis Lectures on Artificial Intelligence and Machine Learning. San Rafael, California, USA: Morgan & Claypool Publishers; Feb 11, 2011:1-95.
14. Wong ML, Leung KS. Data mining using grammar based genetic programming and applications. In: Koza J, editor. *Genetic Programming Series*. New York: Kluwer Academic Publishers; 2002.
15. Zhu Y. Automatic detection of anomalies in blood glucose using a machine learning approach. *J Commun Netw* 2011 Apr;13(2):125-131. [doi: [10.1109/JCN.2011.6157411](https://doi.org/10.1109/JCN.2011.6157411)]
16. Georga EI, Fotiadis DI, Tigas SK. 6 - Nonlinear models of glucose concentration. In: *Personalized Predictive Modeling in Type 1 Diabetes*. London: Academic Press; 2018.
17. Hawkins DM. Identification of outliers. In: *Monographs on Statistics and Applied Probability*. London: Springer; 1987.
18. Lavin A, Ahmad S. Evaluating Real-Time Anomaly Detection Algorithms -- The Numenta Anomaly Benchmark. 2015 Presented at: 2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA); December 9-11, 2015; Miami, FL, USA p. 38-44. [doi: [10.1109/icmla.2015.141](https://doi.org/10.1109/icmla.2015.141)]
19. Mehrotra KG, Mohan CK, Huang H. Anomaly detection principles and algorithms. In: Subrahmanian VS, editor. *Terrorism, Security, and Computation*. Cham, Switzerland: Springer, Cham; 2017.
20. Chandola V, Banerjee A, Kumar V. Anomaly detection. *ACM Comput Surv* 2009 Jul 1;41(3):1-58. [doi: [10.1145/1541880.1541882](https://doi.org/10.1145/1541880.1541882)]
21. Klerx T, Anderka M, Büning HK, Priesterjahn S. Model-Based Anomaly Detection for Discrete Event Systems. 2014 Presented at: 2014 IEEE 26th International Conference on Tools with Artificial Intelligence; November 10-12, 2014; Limassol, Cyprus p. 665-672. [doi: [10.1109/ICTAI.2014.105](https://doi.org/10.1109/ICTAI.2014.105)]
22. Dunning T, Friedman E. In: Loukides M, editor. *Practical Machine Learning: A New Look at Anomaly Detection*. Sebastopol, CA: O'Reilly Media, Inc; 2014.
23. Frandes M, Timar B, Timar R, Lungeanu D. Chaotic time series prediction for glucose dynamics in type 1 diabetes mellitus using regime-switching models. *Sci Rep* 2017 Dec 24;7(1):6232 [FREE Full text] [doi: [10.1038/s41598-017-06478-4](https://doi.org/10.1038/s41598-017-06478-4)] [Medline: [28740090](https://pubmed.ncbi.nlm.nih.gov/28740090/)]
24. Suh S, Kim JH. Glycemic variability: how do we measure it and why is it important? *Diabetes Metab J* 2015 Aug;39(4):273-282 [FREE Full text] [doi: [10.4093/dmj.2015.39.4.273](https://doi.org/10.4093/dmj.2015.39.4.273)] [Medline: [26301188](https://pubmed.ncbi.nlm.nih.gov/26301188/)]
25. Oviedo S, Vehí J, Calm R, Armengol J. A review of personalized blood glucose prediction strategies for T1DM patients. *Int J Numer Method Biomed Eng* 2017 Dec;33(6):1-21. [doi: [10.1002/cnm.2833](https://doi.org/10.1002/cnm.2833)] [Medline: [27644067](https://pubmed.ncbi.nlm.nih.gov/27644067/)]
26. Balakrishnan N, Rangaiah GP, Samavedham L. Review and analysis of blood glucose (BG) models for type 1 diabetic patients. *Ind Eng Chem Res* 2011 Nov 2;50(21):12041-12066. [doi: [10.1021/ie2004779](https://doi.org/10.1021/ie2004779)]

27. Eljil SK, Qadah G, Pasquier M. Predicting hypoglycemia in diabetic patients using data mining techniques. 2013 Presented at: 2013 9th International Conference on Innovations in Information Technology (IIT); March 17-19, 2013; Abu Dhabi, United Arab Emirates. [doi: [10.1109/Innovations.2013.6544406](https://doi.org/10.1109/Innovations.2013.6544406)]
28. Jensen M, Mahmoudi Z, Christensen TF, Tarnow L, Seto E, Johansen MD, et al. Evaluation of an algorithm for retrospective hypoglycemia detection using professional continuous glucose monitoring data. *J Diabetes Sci Technol* 2014 Jan;8(1):117-122 [FREE Full text] [doi: [10.1177/1932296813511744](https://doi.org/10.1177/1932296813511744)] [Medline: [24876547](https://pubmed.ncbi.nlm.nih.gov/24876547/)]
29. Jensen M, Christensen TF, Tarnow L, Mahmoudi Z, Johansen MD, Hejlesen OK. Professional continuous glucose monitoring in subjects with type 1 diabetes: retrospective hypoglycemia detection. *J Diabetes Sci Technol* 2013 Jan 1;7(1):135-143 [FREE Full text] [doi: [10.1177/193229681300700116](https://doi.org/10.1177/193229681300700116)] [Medline: [23439169](https://pubmed.ncbi.nlm.nih.gov/23439169/)]
30. Jensen M, Christensen TF, Tarnow L, Johansen MD, Hejlesen OK. An information and communication technology system to detect hypoglycemia in people with type 1 diabetes. *Stud Health Technol Inform* 2013;192:38-41. [Medline: [23920511](https://pubmed.ncbi.nlm.nih.gov/23920511/)]
31. Botwey R, Daskalaki E, Diem P, Mougiakakou SG. Multi-model data fusion to improve an early warning system for hypo-/hyperglycemic events. *Conf Proc IEEE Eng Med Biol Soc* 2014;2014:4843-4846. [doi: [10.1109/EMBC.2014.6944708](https://doi.org/10.1109/EMBC.2014.6944708)] [Medline: [25571076](https://pubmed.ncbi.nlm.nih.gov/25571076/)]
32. Marling C, Struble NW, Bunescu RC, Shubrook JH, Schwartz FL. A consensus perceived glycemic variability metric. *J Diabetes Sci Technol* 2013 Jul 1;7(4):871-879 [FREE Full text] [doi: [10.1177/193229681300700409](https://doi.org/10.1177/193229681300700409)] [Medline: [23911168](https://pubmed.ncbi.nlm.nih.gov/23911168/)]
33. Wiley M, Bunescu R, Marling C, Shubrook J, Schwartz F. Automatic Detection of Excessive Glycemic Variability for Diabetes Management. 2011 Presented at: 10th International Conference on Machine Learning Applications Workshops; December 18-21, 2011; Honolulu, HI, USA p. 148-154. [doi: [10.1109/ICMLA.2011.39](https://doi.org/10.1109/ICMLA.2011.39)]
34. Jung M. Toward Designing Mobile Software to Predict Hypoglycemia for Patients with Diabetes. 2016 Presented at: 2016 IEEE/ACM International Conference on Mobile Software Engineering Systems (MOBILESoft); May 16-17, 2016; Austin, TX, USA. [doi: [10.1109/MobileSoft.2016.024](https://doi.org/10.1109/MobileSoft.2016.024)]
35. Zhang Y. Predicting occurrences of acute hypoglycemia during insulin therapy in the intensive care unit. *Conf Proc IEEE Eng Med Biol Soc* 2008;2008:3297-3300. [doi: [10.1109/EMBS.2008.4649909](https://doi.org/10.1109/EMBS.2008.4649909)] [Medline: [19163412](https://pubmed.ncbi.nlm.nih.gov/19163412/)]
36. San PP, Ling SH, Nguyen HT. Optimized variable translation wavelet neural network and its application in hypoglycemia detection system. 2012 Presented at: 7th IEEE Conference on Industrial Electronics and Applications (ICIEA); July 18-20, 2012; Singapore, Singapore p. 547-551. [doi: [10.1109/ICIEA.2012.6360788](https://doi.org/10.1109/ICIEA.2012.6360788)]
37. San P, Ling SH, Nguyen HT. Block based neural network for hypoglycemia detection. *Conf Proc IEEE Eng Med Biol Soc* 2011;2011:5666-5669. [doi: [10.1109/EMBS.2011.6091371](https://doi.org/10.1109/EMBS.2011.6091371)] [Medline: [22255625](https://pubmed.ncbi.nlm.nih.gov/22255625/)]
38. Nguyen L, Su S, Nguyen HT. Neural network approach for non-invasive detection of hyperglycemia using electrocardiographic signals. *Conf Proc IEEE Eng Med Biol Soc* 2014;2014:4475-4478. [doi: [10.1109/EMBC.2014.6944617](https://doi.org/10.1109/EMBC.2014.6944617)] [Medline: [25570985](https://pubmed.ncbi.nlm.nih.gov/25570985/)]
39. Nuryani N, Ling SH, Nguyen HT. Hypoglycaemia detection for type 1 diabetic patients based on ECG parameters using Fuzzy Support Vector Machine. 2010 Presented at: The International Joint Conference on Neural Networks (IJCNN); July 18-23, 2010; Barcelona, Spain p. 1-7. [doi: [10.1109/IJCNN.2010.5596916](https://doi.org/10.1109/IJCNN.2010.5596916)]
40. Nuryani N, Ling SH, Nguyen HT. Hybrid particle swarm-based fuzzy support vector machine for hypoglycemia detection. 2012 Presented at: IEEE International Conference on Fuzzy Systems; June 10-15, 2012; Brisbane, QLD, Australia p. 1-6. [doi: [10.1109/FUZZ-IEEE.2012.6250828](https://doi.org/10.1109/FUZZ-IEEE.2012.6250828)]
41. Nguyen L, Nguyen AV, Ling SH, Nguyen HT. Combining genetic algorithm and Levenberg-Marquardt algorithm in training neural network for hypoglycemia detection using EEG signals. 2013 Presented at: 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); July 3-7, 2013; Osaka, Japan p. 5386-5389. [doi: [10.1109/EMBC.2013.6610766](https://doi.org/10.1109/EMBC.2013.6610766)]
42. San PP, Ling SH, Nguyen HT. Deep learning framework for detection of hypoglycemic episodes in children with type 1 diabetes. *Conf Proc IEEE Eng Med Biol Soc* 2016 Dec;2016:3503-3506. [doi: [10.1109/EMBC.2016.7591483](https://doi.org/10.1109/EMBC.2016.7591483)] [Medline: [28269053](https://pubmed.ncbi.nlm.nih.gov/28269053/)]
43. Nguyen LB, Nguyen AV, Ling SH, Nguyen HT. A particle swarm optimization-based neural network for detecting nocturnal hypoglycemia using electroencephalography signals. 2012 Presented at: The International Joint Conference on Neural Networks (IJCNN); June 10-15, 2012; Brisbane, QLD, Australia. [doi: [10.1109/IJCNN.2012.6252745](https://doi.org/10.1109/IJCNN.2012.6252745)]
44. Ling SH, Nguyen H, Chan KY. Genetic algorithm based fuzzy multiple regression for the nocturnal Hypoglycaemia detection. 2010 Presented at: IEEE Congress on Evolutionary Computation; July 18-23, 2010; Barcelona, Spain. [doi: [10.1109/CEC.2010.5586315](https://doi.org/10.1109/CEC.2010.5586315)]
45. Ling S, San PP, Lam HK, Nguyen HT. Hypoglycemia detection: multiple regression-based combinational neural logic approach. *Soft Comput* 2015 Jul 25;21(2):543-553. [doi: [10.1007/s00500-015-1809-z](https://doi.org/10.1007/s00500-015-1809-z)]
46. Mathews S. Fuzzy inference system and multiple regression for the detection of hypoglycemia. *Int J Comp Sci Appl* 2012 Apr 30;2(2):37-50. [doi: [10.5121/ijcsa.2012.2204](https://doi.org/10.5121/ijcsa.2012.2204)]
47. Iaione F, Marques JL. Methodology for hypoglycaemia detection based on the processing, analysis and classification of the electroencephalogram. *Med Biol Eng Comput* 2005 Jul;43(4):501-507. [doi: [10.1007/BF02344732](https://doi.org/10.1007/BF02344732)] [Medline: [16255433](https://pubmed.ncbi.nlm.nih.gov/16255433/)]
48. Eljil SK, Qadah G, Pasquier M. Predicting hypoglycemia in diabetic patients using time-sensitive artificial neural networks. *Int J Healthc Inf Syst Inform* 2016;11(4):70-88. [doi: [10.4018/ijhisi.2016100104](https://doi.org/10.4018/ijhisi.2016100104)]

49. San P, Ling SH, Nguyen HT. Industrial application of evolvable block-based neural network to hypoglycemia monitoring system. *IEEE Trans Ind Electron* 2013 Dec;60(12):5892-5901. [doi: [10.1109/TIE.2012.2228143](https://doi.org/10.1109/TIE.2012.2228143)]
50. Nguyen H, Ghevondian N, Jones TW. Neural-network detection of hypoglycemic episodes in children with type 1 diabetes using physiological parameters. 2006 Presented at: International Conference of the IEEE Engineering in Medicine and Biology Society; January 17-18, 2006; New York, NY, USA p. 6053-6056. [doi: [10.1109/IEMBS.2006.259482](https://doi.org/10.1109/IEMBS.2006.259482)]
51. Chan K, Ling S, Dillon T, Nguyen H. Diagnosis of hypoglycemic episodes using a neural network based rule discovery system. *Expert Syst Appl* 2011 Aug;38(8):9799-9808. [doi: [10.1016/j.eswa.2011.02.020](https://doi.org/10.1016/j.eswa.2011.02.020)]
52. Chan KY, Ling SH, Nguyen HT, Jiang F. A hypoglycemic episode diagnosis system based on neural networks for Type 1 diabetes mellitus. 2012 Presented at: 2012 IEEE Congress on Evolutionary Computation; June 10-15, 2012; Brisbane, QLD, Australia. [doi: [10.1109/CEC.2012.6256604](https://doi.org/10.1109/CEC.2012.6256604)]
53. Ling SH, San PP, Lam HK, Nguyen HT. Non-invasive detection of hypoglycemic episodes in Type 1 diabetes using intelligent hybrid rough neural system. 2014 Presented at: 2014 IEEE Congress on Evolutionary Computation (CEC); July 6-11, 2014; Beijing, China. [doi: [10.1109/CEC.2014.6900229](https://doi.org/10.1109/CEC.2014.6900229)]
54. San PP, Ling SH, Nguyen HT. Hybrid particle swarm optimization based normalized radial basis function neural network for hypoglycemia detection. 2012 Presented at: The International Joint Conference on Neural Networks (IJCNN); June 10-15, 2012; Brisbane, QLD, Australia p. 1-6. [doi: [10.1109/IJCNN.2012.6252743](https://doi.org/10.1109/IJCNN.2012.6252743)]
55. San P, Ling SH, Nguyen HT. Hybrid PSO-based variable translation wavelet neural network and its application to hypoglycemia detection system. *Neural Comput & Applic* 2012 Sep 27;23(7-8):2177-2184. [doi: [10.1007/s00521-012-1168-y](https://doi.org/10.1007/s00521-012-1168-y)]
56. San PP, Ling SH, Nguyen HT. Combinational neural logic systems industrial application on hypoglycemia monitoring system. 2013 Presented at: IEEE 8th Conference on Industrial Electronics and Applications (ICIEA); June 19-21, 2013; Melbourne, VIC, Australia p. 947-952. [doi: [10.1109/ICIEA.2013.6566503](https://doi.org/10.1109/ICIEA.2013.6566503)]
57. Nguyen L, Nguyen AV, Ling SH, Nguyen HT. An adaptive strategy of classification for detecting hypoglycemia using only two EEG channels. 2012 Presented at: Annual International Conference of the IEEE Engineering in Medicine Biology Society; August 28-September 1, 2012; San Diego, CA, USA p. 3515-3518. [doi: [10.1109/EMBC.2012.6346724](https://doi.org/10.1109/EMBC.2012.6346724)]
58. Ling S, San PP, Nguyen HT. Non-invasive hypoglycemia monitoring system using extreme learning machine for Type 1 diabetes. *ISA Trans* 2016 Sep;64:440-446. [doi: [10.1016/j.isatra.2016.05.008](https://doi.org/10.1016/j.isatra.2016.05.008)] [Medline: [27311357](https://pubmed.ncbi.nlm.nih.gov/27311357/)]
59. San PP, Ling SH, Soe NN, Nguyen HT. A novel extreme learning machine for hypoglycemia detection. *Conf Proc IEEE Eng Med Biol Soc* 2014;2014:302-305. [doi: [10.1109/EMBC.2014.6943589](https://doi.org/10.1109/EMBC.2014.6943589)] [Medline: [25569957](https://pubmed.ncbi.nlm.nih.gov/25569957/)]
60. Mo X, Wang Y, Wu X. Hypoglycemia prediction using extreme learning machine (ELM) and regularized ELM. 2013 Presented at: 25th Chinese Control and Decision Conference (CCDC); May 25-27, 2013; Guiyang, China. [doi: [10.1109/CCDC.2013.6561727](https://doi.org/10.1109/CCDC.2013.6561727)]
61. Nguyen H, Ghevondian N, Jones TW. Detection of nocturnal hypoglycemic episodes (natural occurrence) in children with Type 1 diabetes using an optimal Bayesian neural network algorithm. 2008 Presented at: 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; August 20-25, 2008; Vancouver, BC, Canada p. 1311-1314. [doi: [10.1109/IEMBS.2008.4649405](https://doi.org/10.1109/IEMBS.2008.4649405)]
62. Nguyen H, Ghevondian N, Nguyen ST, Jones TW. Detection of hypoglycemic episodes in children with type 1 diabetes using an optimal Bayesian neural network algorithm. 2007 Presented at: 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; August 22-26, 2007; Lyon, France p. 3140-3143. [doi: [10.1109/IEMBS.2007.4352995](https://doi.org/10.1109/IEMBS.2007.4352995)]
63. Nguyen H, Jones TW. Detection of nocturnal hypoglycemic episodes using EEG signals. 2010 Presented at: Annual International Conference of the IEEE Engineering in Medicine and Biology; August 31-September 4, 2010; Buenos Aires, Argentina p. 4930-4933. [doi: [10.1109/IEMBS.2010.5627233](https://doi.org/10.1109/IEMBS.2010.5627233)]
64. Ngo CQ, Truong BC, Jones TW, Nguyen HT. Occipital EEG Activity for the Detection of Nocturnal Hypoglycemia. 2018 Presented at: 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); July 18-21, 2018; Honolulu, HI, USA p. 3862-3865.
65. Bertachi A, Biagi L, Contreras I, Luo N, Vehí J. Prediction of Blood Glucose Levels And Nocturnal Hypoglycemia Using Physiological Models and Artificial Neural Networks. 2018 Presented at: 3rd International Workshop on Knowledge Discovery in Healthcare Data co-located with the 27th International Joint Conference on Artificial Intelligence and the 23rd European Conference on Artificial Intelligence (IJCAI-ECAI 2018); July 13, 2018; Stockholm, Sweden p. 85-90.
66. Marling C, Shubrook JH, Vernier SJ, Wiley MT, Schwartz FL. Characterizing blood glucose variability using new metrics with continuous glucose monitoring data. *J Diabetes Sci Technol* 2011 Jul 1;5(4):871-878 [FREE Full text] [doi: [10.1177/193229681100500408](https://doi.org/10.1177/193229681100500408)] [Medline: [21880228](https://pubmed.ncbi.nlm.nih.gov/21880228/)]
67. Georga E, Protopappas VS, Ardigò D, Polyzos D, Fotiadis DI. A glucose model based on support vector regression for the prediction of hypoglycemic events under free-living conditions. *Diabetes Technol Ther* 2013 Aug;15(8):634-643. [doi: [10.1089/dia.2012.0285](https://doi.org/10.1089/dia.2012.0285)] [Medline: [23848178](https://pubmed.ncbi.nlm.nih.gov/23848178/)]
68. Georga EI, Protopappas VC, Mougiakakou SG, Fotiadis DI. Short-term vs long-term analysis of diabetes data: application of machine learning and data mining techniques. 2013 Presented at: 13th IEEE International Conference on BioInformatics and BioEngineering; November 10-13, 2013; Chania, Greece. [doi: [10.1109/BIBE.2013.6701622](https://doi.org/10.1109/BIBE.2013.6701622)]

69. Jensen M, Christensen TK, Tarnow L, Seto E, Johansen MD, Hejlesen OK. Real-time hypoglycemia detection from continuous glucose monitoring data of subjects with type 1 diabetes. *Diabetes Technol Ther* 2013 Jul;15(7):538-543. [doi: [10.1089/dia.2013.0069](https://doi.org/10.1089/dia.2013.0069)] [Medline: [23631608](https://pubmed.ncbi.nlm.nih.gov/23631608/)]
70. Marling C, Xia L, Bunesco R, Schwartz F. Machine learning experiments with noninvasive sensors for hypoglycemia detection. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc; 2016 Presented at: IJCAI Workshop on Knowledge Discovery in Healthcare Data; July 10, 2016; New York, NY.
71. Ling S, Nguyen HT. Ventricular repolarization variability for hypoglycemia detection. 2011 Presented at: Annual International Conference of the IEEE Engineering in Medicine Biology Society; August 30-September 3, 2011; Boston, MA, USA p. 7961-7964. [doi: [10.1109/IEMBS.2011.6091963](https://doi.org/10.1109/IEMBS.2011.6091963)]
72. Ling S, Nguyen HT. Genetic-algorithm-based multiple regression with fuzzy inference system for detection of nocturnal hypoglycemic episodes. *IEEE Trans Inf Technol Biomed* 2011 Mar;15(2):308-315. [doi: [10.1109/TITB.2010.2103953](https://doi.org/10.1109/TITB.2010.2103953)] [Medline: [21349796](https://pubmed.ncbi.nlm.nih.gov/21349796/)]
73. Ling SH, Nguyen HT, Leung FH. Hypoglycemia detection using fuzzy inference system with genetic algorithm. 2011 Presented at: IEEE International Conference on Fuzzy Systems (FUZZ-IEEE ); June 27-30, 2011; Taipei, Taiwan. [doi: [10.1109/FUZZY.2011.6007319](https://doi.org/10.1109/FUZZY.2011.6007319)]
74. Jung M, Lee YB, Jin SM, Park SM. arXiv. 2017. Prediction of Daytime Hypoglycemic Events Using Continuous Glucose Monitoring Data and Classification Technique URL:<https://arxiv.org/abs/1704.08769> [accessed 2019-04-02] [[WebCite Cache ID 77KnFAH7R](https://arxiv.org/abs/1704.08769)]
75. Zhu Y. Automatic detection of anomalies in blood glucose using a machine learning approach. 2010 Presented at: IEEE International Conference on Information Reuse & Integration; August 4-6, 2010; Las Vegas, NV. USA p. 92-97. [doi: [10.1109/IRI.2010.5558959](https://doi.org/10.1109/IRI.2010.5558959)]
76. Chan K, Ling SH, Dillon TS, Nguyen H. Classification of hypoglycemic episodes for Type 1 diabetes mellitus based on neural networks. 2010 Presented at: IEEE Congress on Evolutionary Computation; July 18-23, 2010; Barcelona, Spain. [doi: [10.1109/CEC.2010.5586320](https://doi.org/10.1109/CEC.2010.5586320)]
77. Ghevondian N, Nguyen HT, Colagiuri S. A novel fuzzy neural network estimator for predicting hypoglycaemia in insulin-induced subjects. 2001 Presented at: 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society; October 25-28, 2001; Istanbul, Turkey. [doi: [10.1109/IEMBS.2001.1020533](https://doi.org/10.1109/IEMBS.2001.1020533)]
78. San P, Ling SH, Nguyen HT. Intelligent detection of hypoglycemic episodes in children with type 1 diabetes using adaptive neural-fuzzy inference system. *Conf Proc IEEE Eng Med Biol Soc* 2012;2012:6325-6328. [doi: [10.1109/EMBC.2012.6347440](https://doi.org/10.1109/EMBC.2012.6347440)] [Medline: [23367375](https://pubmed.ncbi.nlm.nih.gov/23367375/)]
79. Nuryani SL, Nguyen HT. Electrocardiographic T-wave peak-to-end interval for hypoglycaemia detection. *Conf Proc IEEE Eng Med Biol Soc* 2010;2010:618-621. [doi: [10.1109/IEMBS.2010.5627430](https://doi.org/10.1109/IEMBS.2010.5627430)] [Medline: [21096769](https://pubmed.ncbi.nlm.nih.gov/21096769/)]
80. Nuryani N, Ling SS, Nguyen HT. Electrocardiographic signals and swarm-based support vector machine for hypoglycemia detection. *Ann Biomed Eng* 2012 Apr;40(4):934-945. [doi: [10.1007/s10439-011-0446-7](https://doi.org/10.1007/s10439-011-0446-7)] [Medline: [22012087](https://pubmed.ncbi.nlm.nih.gov/22012087/)]
81. Ling S, Nguyen HT. Natural occurrence of nocturnal hypoglycemia detection using hybrid particle swarm optimized fuzzy reasoning model. *Artif Intell Med* 2012 Jul;55(3):177-184. [doi: [10.1016/j.artmed.2012.04.003](https://doi.org/10.1016/j.artmed.2012.04.003)] [Medline: [22698854](https://pubmed.ncbi.nlm.nih.gov/22698854/)]
82. San PP, Ling SH, Nuryani N, Nguyen H. Evolvable rough-block-based neural network and its biomedical application to hypoglycemia detection system. *IEEE Trans Cybern* 2014 Aug;44(8):1338-1349. [doi: [10.1109/TCYB.2013.2283296](https://doi.org/10.1109/TCYB.2013.2283296)] [Medline: [24122616](https://pubmed.ncbi.nlm.nih.gov/24122616/)]
83. Lai J, Leung F, Ling S. Hypoglycaemia detection using fuzzy inference system with intelligent optimiser. *Appl Soft Comput* 2014 Jul;20:54-65. [doi: [10.1016/j.asoc.2013.12.015](https://doi.org/10.1016/j.asoc.2013.12.015)]
84. Daskalaki E, Nørgaard K, Züger T, Proutzou A, Diem P, Mougiakakou S. An early warning system for hypoglycemic/hyperglycemic events based on fusion of adaptive prediction models. *J Diabetes Sci Technol* 2013 May 1;7(3):689-698 [[FREE Full text](https://doi.org/10.1177/193229681300700314)] [doi: [10.1177/193229681300700314](https://doi.org/10.1177/193229681300700314)] [Medline: [23759402](https://pubmed.ncbi.nlm.nih.gov/23759402/)]

## Abbreviations

- ANFIS:** adaptive neural fuzzy inference system
- ANN:** artificial neural network
- AR:** autoregressive
- BBNN:** block-based neural network
- BG:** blood glucose
- BNN:** Bayesian Neural Network
- CGM:** continuous glucose monitoring.
- DBN:** deep belief network
- DT:** decision tree
- EA:** evolutionary algorithm
- ECG:** electrocardiogram
- ELM:** extreme learning machine

**GA:** genetic algorithm  
**GP:** genetic programming  
**GV:** glycemic variability  
**HMM:** hidden Markov model  
**MLP:** multilayer perceptron  
**NAR:** nonlinear autoregressive network  
**NARX:** nonlinear autoregressive network with exogenous inputs  
**NBC:** Naive Bayes classifier  
**POC:** point-of-care  
**RF:** random forest  
**RNN:** recurrent neural network  
**ROC:** receiver operating characteristic  
**SLP:** single-layer perceptron  
**SMBG:** self-monitoring blood glucose  
**SVM:** support vector machine  
**VTWNN:** variable translation wavelet neural network

*Edited by G Eysenbach; submitted 11.05.18; peer-reviewed by U Sartipy, J Vehi; comments to author 11.10.18; revised version received 27.11.18; accepted 30.01.19; published 01.05.19*

*Please cite as:*

Woldaregay AZ, Årsand E, Botsis T, Albers D, Mamykina L, Hartvigsen G  
*Data-Driven Blood Glucose Pattern Classification and Anomalies Detection: Machine-Learning Applications in Type 1 Diabetes*  
*J Med Internet Res* 2019;21(5):e11030  
URL: <https://www.jmir.org/2019/5/e11030/>  
doi: [10.2196/11030](https://doi.org/10.2196/11030)  
PMID: [31042157](https://pubmed.ncbi.nlm.nih.gov/31042157/)

©Ashenafi Zebene Woldaregay, Eirik Årsand, Taxiarchis Botsis, David Albers, Lena Mamykina, Gunnar Hartvigsen. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 01.05.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Multimedia Appendix 1: Analysis of reported parameters, data characteristics, machine learning class, and performance metrics

Table 1: Reported input features, machine learning class and accuracy.

Study	Features							Type of Machine Learning																Performance								
	BG	Insulin	Phy. Activity	Diet	HR	QT interval	Skin Impedance	Galvanic response	DT	SVM	RNN	ANN	NARX & NAR	Fuzzy	NBC	Rule Based	ELM	HMM	DBN	Gaussian Pro.	BBNN	VTWNN	BNN	CNLN	GP	GA	Hybrid	Sensitivity	Specificity	Accuracy		
[25]																													76%	58%	--	
[26]& [27]																														79%	52%	--
[28]																													86%	96%	96%	
[29]																													80%	98%	98%	
[30]																													84%	--	--	
[31]																													--	--	88%	
[32]																													--	--	--	
[33] & [34]																													78%	96%	--	
[35] & [36]																													81%	93%	--	
[37]																													80%	80%	80%	
[38]																													--	--	68%	
[39]																													60%	100%	85%	
[40] & [41] & [42]																													75%	50%	--	
[43]																													86%	80%	--	
[44]																													77%	53%	--	
[45]																													79%	54%	--	
[46]																													78%	60%	--	
[47]																													--	--	--	
[48]																													80%	73%	--	
[49]																													--	95%	--	
[50]																													85%	41%	--	
[51] & [52] & [53]																													83%	64%	--	
[54] & [55]																													82%	63%	--	
[56]																													75%	60%	--	
[57]																													82%	60%	--	
[58] & [59]																													74%	59%	63%	
[60] & [61]																													75%	83%	--	
[62]																													77%	52%	--	
[63] & [64]																													79%	50%	--	
[65] & [66]																													77%	51%	--	
[67]																													79%	52%	--	
[68]																													77%	55%	--	





Class of machine learning

Hypoglycemia classification & detection

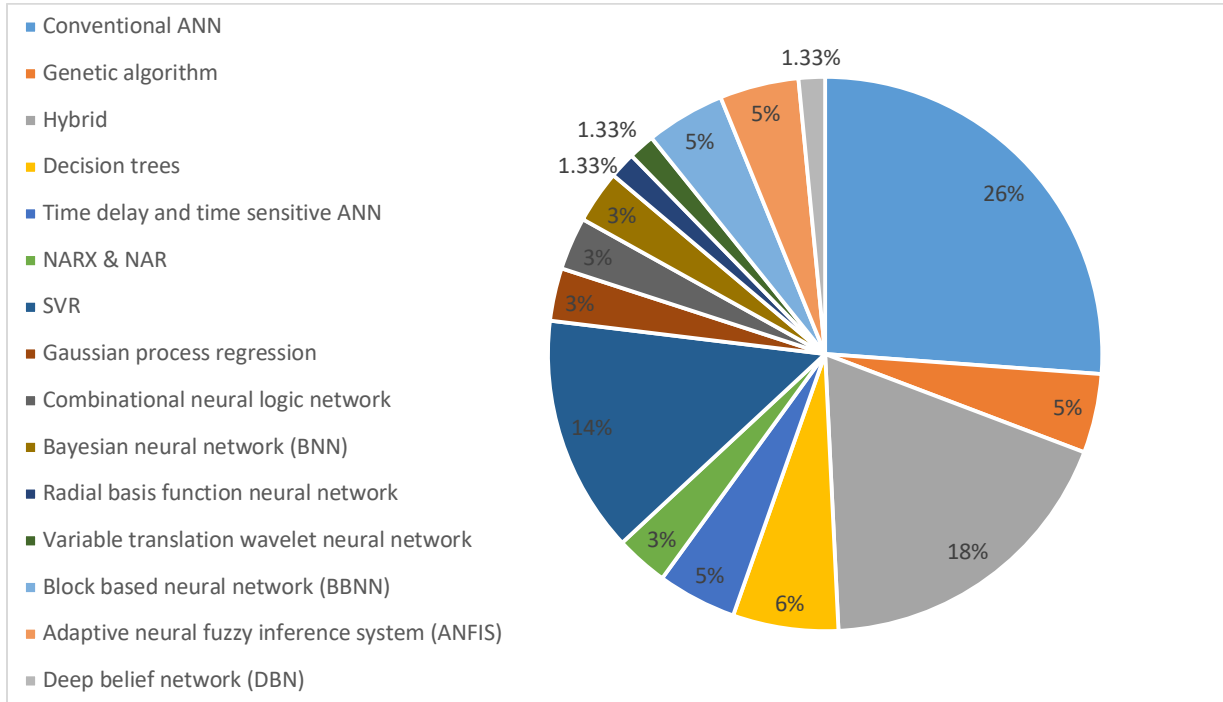


Figure 6: Classes of machine learning used in hypoglycaemia classification and detection.

Hyperglycemia classification & detection

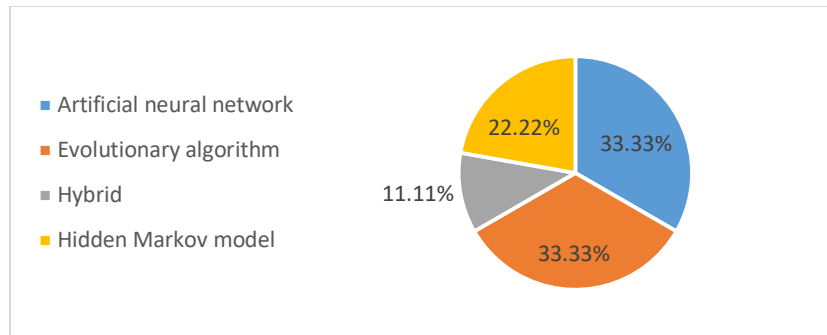


Figure 7: Classes of machine learning used in hyperglycaemias classification and detection.

## Glycemic variability classification & detection

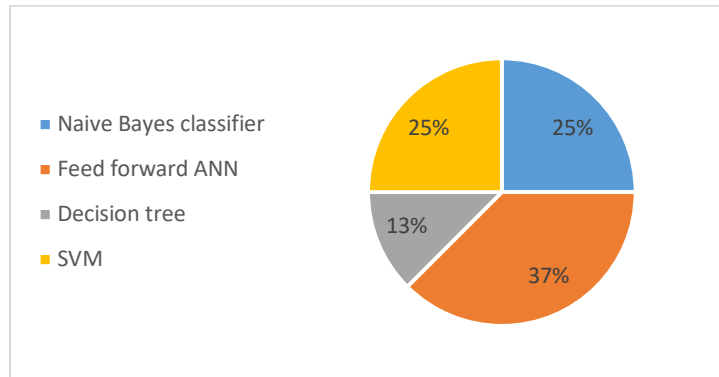


Figure 8: Classes of machine learning used in glycaemic variability classification and detection.

## Performance metrics

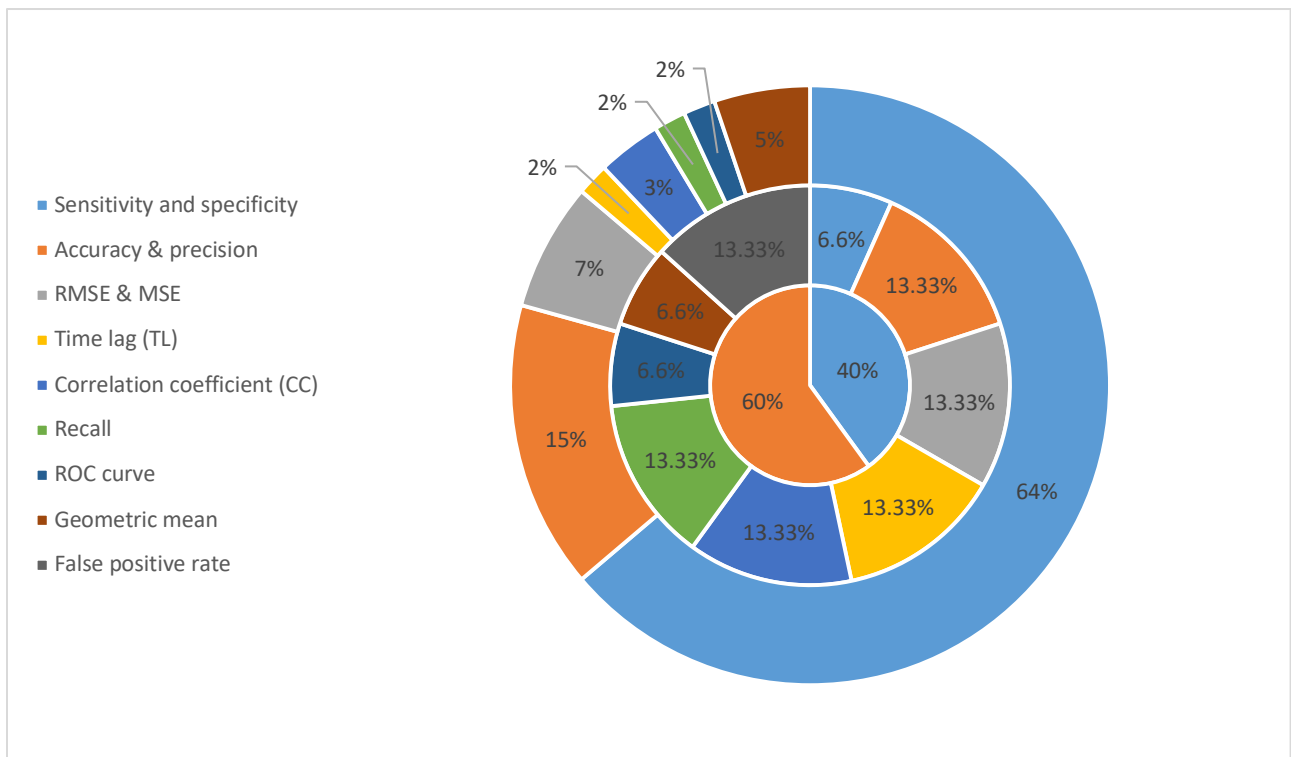


Figure 9: Depicts the type of performance metrics used in the studies. The outer ring, middle ring and inner rings depicts the type of performance metrics used in hypoglycaemia, hyperglycaemias and glycaemic variability classification and detection algorithms.

# Multimedia Appendix 2: Detail on reported accuracy, inputs and performance metrics used, and machine learning categorization.

Table 1: Data extracted from the literatures included in the study.

Ref.	Subject	Type of input	Data Format/Data source	Input Pre-processing	Class of Machine Learning	Performance Metrics
[1]	16 Real (14.6±1.5 years of age)	BG, rate of change of heart rate, Corrected QT interval of electrocardiogram signal, Rate of change of corrected QT interval	Department of Health, Government of Western Australia, Yellow Spring Instrument for BG	N/A	Hybrid-(feed-forward neural network and genetic algorithm)	Sensitivity and specificity
[2] & [3]	16 Real (14.6±1.5 years of age)	BG, rate of change of heart rate, Corrected QT interval of electrocardiogram signal, Rate of change of corrected QT interval	HypoMon (Hypoglycemia Monitor from AIMedics Pty, Ltd.) to measure the required physiological parameters, Yellow Spring Instrument to measure BG.	N/A	Hybrid-(feed-forward neural network, rule discovery, and genetic algorithm)	Sensitivity and specificity
[4]	10 Real (2 male & 8 female, 11-70 years of age)	BG (past and present), Rate of change of BG	American Hospital in Dubai-Medtronic CGM & Insulin Pump	Differencing BG Values	Decision trees	Accuracy, sensitivity and specificity
[5]	10 Real (2 male & 8 female, 11-70 years of age)	Subcutaneous glucose measurements, insulin, and carbohydrate intake	American Hospital in Dubai-Medtronic CGM & Insulin Pump	N/A	Artificial Neural Networks (ANN), Time- Sensitive ANN (TS-ANN), Time Delay Neural Network (TDNN), Nonlinear Autoregressive Network with exogenous inputs (NARX), Distributed Time Delay Neural Network (DTDNN), and Nonlinear Autoregressive Network (NAR.)	Root mean squared error, sensitivity, specificity, and Accuracy
[6]	15 Real (3 women and 12 men whose ages ranged from 19 to 65 years (average, 40.3 – 13.5 years))	Glucose profile, meals, insulin intake, and physical activities (exercise), sleep	Guardian Real-time CGM system (Medtronic Minimed Inc., Northridge, CA), SenseWear armband (BodyMedia Inc., Pittsburgh, PA) physical activity monitor	N/A	Support vector for regression (SVR), feed-forward multilayer perceptron (MLP) and Gaussian processes (GP) regression	Sensitivity, precision, and time lag
[7]	15 Real	Glucose, rate of meal glucose appearance, plasma insulin concentration, meal-derived glucose inserted in plasma, energy expenditure, hour of day	Guardian Real-time CGM system (Medtronic Minimed Inc., Northridge, CA), SenseWear armband (BodyMedia Inc., Pittsburgh, PA) physical activity monitor	N/A	Support vector regression (SVR) or Gaussian processes (GP) regression	Standard deviation of the RMSE and correlation coefficient, sensitivity
[8]	12 real (6 normal & 6 T1DM)-6 males & 6 females aged 26 ± 3 years)	Measured BGLs, skin impedances, heart rates	N/A	Normalization	A novel fuzzy neural network estimator algorithm (FNNE)- a parallel combination of fuzzy inference mechanism (FIM) and a multi-layered neural network	Mean square error, correlation coefficient
[9] & [10]	10 Real (all male, age 44 ± 15 years)	Current CGM reading, first derivative of the current CGM reading and the reading before, Time since last insulin injection, Linear regression, Skewness, Kurtosis of the CGM readings in multiple intervals.	CGM device (Guardian RT®, Minimed Inc., USA), emoCue Glucose 201+ glucose analyzer (HemoCue®, Ängelholm, Sweden)	Reconstruction of CGM data using spline interpolation and a rough feature elimination, using fast SEPCOR algorithm.	Support vector machine	Sensitivity, and specificity
[11] & [12]	10 Real (all male, age 44 ± 15 years)	Current CGM reading, first derivative of the current CGM reading and the reading before, Time since last insulin injection, Linear regression, Skewness, Kurtosis of the CGM readings in multiple intervals.	CGM device (Guardian RT®, Minimed Inc., USA), emoCue Glucose 201+ glucose analyzer (HemoCue®, Ängelholm, Sweden)	Reconstruction of CGM data using spline interpolation and a rough feature elimination, using fast SEPCOR algorithm.	Support vector machine	Sensitivity, and specificity
[13]	21 Real	BG, Meal, Rate of decrease from a peak and absolute level of the BG at the decision point	Diagnostic (professional) CGM devices	N/A	Decision trees	Accuracy, sensitivity, and specificity
[14]	1 Real (Male)	Glucose levels right before meals (G1), Glucose levels after more than 5 hours (G2), Time interval (T), Average Fasting glucose level (AG1), The rate of decrease in [Glu], Ratio of current level to average	Self-Monitored Blood Glucose (SMBG)	Extraction and representing temporal change information,	Decision tree and SVM	Accuracy
[15]	8 Real (five male, three female aged 35 + 13.5 years (mean +SD))	ECG Signal information	Portable apparatus was developed to record the EEG- To record the EEG signal, a digitalization module was developed that was linked to a palmtop PC.	Amplitude Normalization and Transformation (FFT).	Artificial neural networks (MLP-ANNs)	Accuracy, sensitivity, and specificity
[16] & [17] & [18]	16 Real (children-14.6±1.5 years)	BG, heart rate (HR), corrected QT interval of the ECG signal, change of HR, and the change of corrected QT interval	skin surface biosensor electrodes for the measurement of physiological parameters, Yellow Spring Instruments (YSI) used for collection of reference BG	Normalization and linear correction analysis..	Genetic algorithm (GA)-based multiple regression with fuzzy inference system (FIS)	Sensitivity and specificity
[19]	16 Real (children-14.6±1.5 years)	BG, heart rate and the corrected QT interval of the electrocardiogram (ECG) signal	skin surface biosensor electrodes for the measurement of physiological parameters, Yellow Spring Instruments (YSI) used for collection of reference BG	Normalization	Hybrid particle-swarm-optimization-based fuzzy-reasoning model, Feed-Forward Neural Network (FFNN)	Sensitivity and specificity
[20]	15 Real (children-14.6±1.5 years)	BG, heart rate and the corrected QT interval of the electrocardiogram (ECG) signal	skin surface biosensor electrodes for the measurement of physiological parameters, Yellow Spring Instruments (YSI) used for collection of reference BG	Normalization and Partitioning of the input.	Hybrid rough set based neural network (FRNN)	Sensitivity and specificity
[21]	15 Real (children-14.6±1.5 years)	BG, electrocardiogram signal (heart rate (HR) & corrected QT interval (QTc), change of heart rate (HR) and corrected QT interval (QTc))	skin surface biosensor electrodes for the measurement of physiological parameters, Yellow Spring Instruments (YSI) used for collection of reference BG	Correlation analysis	Multiple regression (MR)-based combinational neural logic approach - Combinational neural logic network (NLN-feedforward neural network (FFNN) & rule based logic)	Sensitivity and specificity
[22]	16 Real (children-14.6±1.5 years)	BG, heart rate (HR), corrected QT interval of ECG (QTc), change of heart rate (ΔHR), and change of corrected QT interval of ECG (ΔQTc).	skin surface biosensor electrodes for the measurement of physiological parameters, Yellow Spring Instruments (YSI) used for collection of reference BG	Normalization	Single-hidden Layer Feedforward neural Network (SLFN) with L hidden nodes (Extreme learning machine(ELM)-based neural network)	Sensitivity and specificity
[23]	1 Real (middle-aged male)	BG, Insulin, meals, heart rate (HR), galvanic skin response (GSR), and skin and air temperatures (ST and AT).	Medtronic insulin pump & Dexcom CGM system, smart phone (reported meals, sleep & exercise), Fitness band, Basis Peak (heart rate (HR), galvanic skin response (GSR), & skin and air temperatures (ST & AT))	Feature extraction and selection- using greedy feature selection	Support vector machines (SVM)	Sensitivity, specificity, precision and recall
[24]	5 real	BG, heart rate(HR),corrected QT interval, change of HR and change of corrected QT interval	N/A	Linear correction analysis	Multiple regression with fuzzy inference system(FIS), Neural network. Parameters optimized through genetic algorithm (GA), particle-swarm optimization respectively.	Sensitivity and specificity
[25]	N/A	Blood Glucose (CGM)	N/A	N/A	Single-hidden Layer Feedforward neural Network (SLFN) with L hidden nodes (extreme learning machines (ELM) and regularized ELM (RELM))	Root mean square error (RMSE), sensitivity, specificity, ROC
[26]	21 Real (children, 14.4±1.6 years)	BG, heart rate, corrected QT interval of the ECG signal and skin impedance	HypoMon, blood glucose (BG) levels were collected as reference using Yellow Spring Instruments	Normalization	Feedforward multi-layer neural network	Sensitivity and specificity
[27] & [28] & [29]	16 Real (children, 14.4±1.6 years)	BG, heart rate, corrected QT interval of the ECG signal and skin impedance	HypoMon, actual blood glucose (BG) levels were collected as reference using Yellow Spring Instruments	Normalization	Feed-forward multi-layer neural network (Bayesian neural network-Bayesian learning)	Sensitivity and specificity
[30] & [31]	5 real (adolescent patients between the ages of 12 and 18 year old)	BG, EEG responses-the centroid theta frequency and the centroid alpha frequency from each channel	Compumedics system, electromyogram (EMG) signals, electrooculogram (EOG) signals, BGLs were acquired using Yellow Spring Instruments	Filtering (IIR highpass filter and A notch filter at 50Hz)	Feed-forward multi-layer neural network (standard particle swarm optimization strategy is applied to optimize the parameters)	Sensitivity and specificity, ROC Curve
[32]	5 real (adolescent patients between the ages of 12 and 18 year old)	BG, EEG responses-the centroid theta frequency and the centroid alpha frequency from each channel	Compumedics system, electromyogram (EMG) signals, electrooculogram (EOG) signals, BGLs were acquired using Yellow Spring Instruments	Filtering (IIR highpass filter), Feature extraction (Fast Fourier Transform (FFT))	Feed-forward multi-layer neural network (genetic algorithm and Levenberg Marquardt algorithm)	Sensitivity and specificity, ROC Curve
[33]	5 real	BG, Repolarization variabilities- QTcVI, TpTeVI, ToTeVI and RTpVI	Compumedics system, BGLs were acquired using Yellow Spring Instruments	N/A	Swarm-based Support vector machine (SVM) (Radial basis function (RBF) standard particle swarm optimization strategy is applied to optimize the parameters)	Sensitivity, specificity, and geometric mean
[34] & [35]	5 real (with age of 16±0.7 years)	BG, heart rate, corrected QT (QT c) interval and corrected TpTe (TpTec) interval	Compumedics system, BGLs were acquired using Yellow Spring Instruments	Feature extraction	Fuzzy Support Vector Machine - kernel functions (radial basis function (RBF), exponential radial basis function (ERBF) and polynomial function)	Sensitivity, specificity and accuracy
[36] & [37]	5 real	BG, HR, RTpTe, QTc, TpTec, ToTec and QTpe	Princess Margaret Hospital in Perth, Australia, with approval from Women's and Children's Health Service,	Normalization and feature extraction	Hybrid particle swarm - based fuzzy support vector machine (SFisSvm) technique- kernel functions (radial basis function	Sensitivity, specificity and geometric mean

			Department of Health, Government of Western Australia		(RBF), exponential radial basis function (ERBF) and polynomial function)-hybrid particle swarm optimization	
[38]	15 Real (children)	BG, heart rate (HR), corrected QT (QTc), change in the heart rate (ΔHR) and change in the QTc interval (ΔQTc)	Compumedics system, BGLs were acquired using Yellow Spring Instruments	Normalization	Hybrid particle swarm optimization based normalized radial basis function neural network (NRBFNN)- hybrid particle swarm optimization with wavelet mutation (HPSOWM)	Sensitivity and specificity
[39] & [40]	15 Real (children)	BG, heart rate (HR) and corrected QT interval (QTc)	Compumedics system, BGLs were acquired using Yellow Spring Instruments	Normalization	Variable translation wavelet neural network (VTWNN)-hybrid particle swarm optimization with wavelet mutation (HPSOWM)	Sensitivity and specificity
[41] & [42]	15 Real (children)	BG, heart rate (HR) and the corrected QT interval (QTc)	Compumedics system, BGLs were acquired using Yellow Spring Instruments	Normalization	Evolvable block based neural network (BBNN)- hybrid particle swarm optimization with wavelet mutation (HPSOWM)	Sensitivity, specificity, ROC Curve, and geometric mean value
[43]	15 Real (children)	BG, heart rate (HR) and corrected QT (QTc)	Compumedics system, BGLs were acquired using Yellow Spring Instruments	N/A	Adaptive neural fuzzy inference system (ANFIS)- hybrid particle swarm optimization with wavelet mutation (HPSOWM)	Sensitivity and specificity
[44]	15 Real (children with ages 14.6 ± 1.5 years)	BG, heart rate (HR) and the corrected QT interval (QTc)	N/A	N/A	Combinational neural logic network (NLN) - hybrid particle swarm optimization with wavelet mutation (HPSOWM)	Sensitivity and specificity
[45]	15 Real (children with ages 14.6 ± 1.5 years)	BG, HR, QTc, change in HR and change in QTc	Compumedics system, BGLs were acquired using Yellow Spring Instruments	Rough Set based Pre-processing	Hybrid rough-block-based neural network (R-BBNN)-hybrid particle swarm optimization with wavelet mutation (HPSOWM)	Sensitivity and specificity
[46]	15 Real (children with ages 14.6 ± 1.5 years)	BG, heart rate (HR) and corrected QT (QTc)	Compumedics system, BGLs were acquired using Yellow Spring Instruments	Feature representation using unsupervised restricted Boltzmann machines (RBM)	Deep belief network (DBN) block based neural network (BBNN)- greedy layer-wise manner training-backpropagation of error derivatives.	Sensitivity and specificity
[47]	16 Real (children with ages 14.6 ± 1.5 years)	BG, HR, QTc, Change in HR and Change in QTc	N/A	N/A	ELM trained feed-forward neural network (ELM-FFNN)-single hidden layer feedforward neural network (FFNN)-Extreme Learning Machine (ELM)	Sensitivity and specificity
[48]	N/A	BG levels, Slope of changes in BG levels, dose response to IV insulin, Insulin titration promptness, Cumulated, administered insulin	MIMIC II (Multiparameter Intelligent Monitoring in Intensive Care database II) from the ICUs at Beth Israel Deaconess Medical Center, Boston, MA.	Feature selection and ranking	Decision tree (Classification tree-C5.0)	Sensitivity and specificity
[49]	16 Real (children with ages 14.6 ± 1.5 years)	BG, heart rate (HR), corrected QT interval of the electrocardiogram (ECG) signal (QTc), change of HR, and change of QTc	Princess Hospital for Children in Perth, Western Australia, Australia. The actual BG levels were collected as reference using the Yellow Spring Instruments	N/A	Fuzzy inference system (FIS)- multi-objective optimization approach- wavelet mutated differential evolution optimizers	Sensitivity and specificity
[50]	23 real (17 to 70 years of age)	Glucose and insulin	Medtronic insulin pumps (Medtronic MiniMed Inc., Northridge, CA, USA) combined with a real-time CGM system under normal daily living conditions	Smoothing and Filtering	Adaptive data-driven models (autoregressive with output correction - cARX, & a recurrent neural network - RNN)- Data fusion techniques ( Dempster-Slafer Evidential Theory (DST), Genetic Algorithms (GA), & Genetic Programming (GP) - teacher-forced, real-time, recurrent learning algorithm	RMSE, time lag (TL), correlation coefficient, and receiver operating characteristic (ROC) curve
[51]	23 real (17 to 70 years of age)	Glucose and insulin	Medtronic insulin pumps (Medtronic MiniMed Inc., Northridge, CA, USA) combined with a real-time CGM system under normal daily living conditions	N/A	Hybrid autoregressive with an output correction module/recurrent neural network (ARN)-based EWS, recurrent neural network (RNN)- teacher-forced, real-time, recurrent learning algorithm.	Root mean square error, TL, and correlation coefficient
[52]	N/A	BG	N/A	N/A	Hidden Markov model (HMM)- Baum-Welch algorithm that belongs to the family of Expectation Maximization algorithms, Forward-Backward algorithm	Precision, false positives rates, Recall
[53]	10 real (adolescent)	BG, HR, PR, QTc, RTC, TPTEC, time domain (Mean RR interval (MeanRR), standard deviation of the RR interval index (SDNN), root mean square of successive RR interval differences(RMSSD), Percentage of consecutive RR intervals that differ by more than 50 ms (pNN50), HRV triangular index (HRVi), Baseline width of the RR interval histogram evaluated through triangular interpolation (TINN)) & frequency domain, Total spectral power (TotalPw), Ratio between LF and HF components (LF/HF)	During the study period, ECG signals were continuously recorded by a medical device called Compumedics with the sampling rate of 512 Hz, while actual blood glucose was collected as reference using Yellow Springs Instruments. Kubios HRV Analysis Software package	Normalization & interpolation	Multilayer feed-forward neural network- Levenberg-Marquardt (LM) algorithm-error back propagation learning method	Sensitivity & specificity, geometric mean.
[54]	N/A	BG	N/A	N/A	Hidden Markov model (HMM)- Baum-Welch algorithm that belongs to the family of Expectation Maximization algorithms, Forward-Backward algorithm	Precision, false positives rates, recall
[55]	11 patients (9 female and 2 male patients ranged in age from 26 to 67 years)	Mean amplitude of glycemic excursion (MAGE), i) excursion frequency, ii) distance traveled and along with physicians' variability classification for that day (26 to 67 years)	Medtronic Paradigm® insulin pumps with Real-Time continuous glucose monitors. GlycoMark™. The CGM data were extracted from the Medtronic CareLink® database into clinical diabetes research database.	N/A	Naive Bayes classifier (probabilistic reasoning), a multilayer perceptron (ANN), and a logistic model tree (decision tree built using logistic regression)	Accuracy (classifier vs. physician glycemic variability classifications of daily CGM charts)
[56]	19 patients (14 female and 5 male patients, ranging in age from 17 to 71 (mean 47) years)	MAGE, Excursion Frequency (EF), Distance Traveled (DT), SD, Area Under the Curve, Central Image Moments, Eccentricity, Discrete Fourier Transform, Roundness Ratio, Bending Energy, Direction Codes, Maximum Slope	Medtronic Paradigm® insulin pumps with Real-Time continuous glucose monitors. GlycoMark™. The CGM data were extracted from the Medtronic CareLink® database into our clinical diabetes research database.	Smoothing and feature selection (greedy forward selection and greedy backward elimination)	Multilayer perceptrons (MPs) and support vector machines for regression (SVR)- Gaussian kernel, Back propagation	Accuracy, sensitivity, and specificity
[57]	11 patients (9 female and 2 male patients ranged in age from 26 to 67 years)	Minimum Input-Direction Codes, Excursion Frequency, Standard Deviation, and Distance Traveled	Medtronic Paradigm® insulin pumps with Real-Time continuous glucose monitors. GlycoMark™. The CGM data were extracted from the Medtronic CareLink® database into clinical diabetes research database.	Smoothing using cubic splines & feature selection (Pearson's Correlation Coefficient & t-test, wrapper approach using greedy backward elimination)	Naive Bayes (NB), Multilayer Perceptron (MP), and Support Vector Machine (SVM)- Gaussian kernel, Back propagation, grid search optimization	Accuracy, sensitivity, and specificity
[58]	6 real patients (Two male, and four female)	SMBG (finger sticks), CGM, Insulin, Physical activity, Diet, illness and other life events	OhioT1DM Dataset (Medtronic 530G insulin pumps, Medtronic Elite CGM sensors, and Reported life-event data via a custom smartphone app and provided physiological data from a Basis Peak fitness band.)	N/A	Artificial Neural Network (ANN)	Accuracy, sensitivity, and specificity
[59]	8 real patients (12 and 18 years age)	Frequency features of occipital lobe (Centroid frequency and spectral entropy)	Yellow Spring Instruments for BG and Compumedics System for EEG	Fast Fourier Transform	Bayesian regularized neural network	Sensitivity and specificity

Table 2: Reported accuracy from the literatures.

Ref.	Reported System performance (Accuracy, Sensitivity, Specificity & Time horizon)	Comment
[1]	Neural network trained with Genetic algorithm (NN-GA) (Sensitivity (75.57%), Specificity (57.68%))	Compared with Neural network trained with a Levenberg Marquardt (LM) algorithm, Statistical Regression (SR), Fuzzy Regression (FR), Genetic Programming (GP), and Genetic Programming based Fuzzy Regression (GP-FR) for detecting hypoglycemia incidences.
[2] & [3]	GA-NN based Rule discovery ( Specificity (79.11%) and Sensitivity (52.01%))	Compared with Fuzzy regression (FR), Genetic programming (GP), fuzzy regression based Genetic programming (FR-GP), neural network trained with back propagation ( NN-BP) and Neural network trained genetic algorithm (NN-GA) for hypoglycemia detection.
[4]	Decision tree - 30 min Hypoglycemia prediction-(Sensitivity (86.47%), Specificity (96.22), Accuracy (95.97%))	Compared various decision tree approach J4.8, REPTree, Bagging, J4.8 and cost sensitive version of J4.8 using CGM data.
[5]	Time-Sensitive ANN (TS-ANN) - 30 min hypoglycemia prediction-(average specificity (98.2%), average accuracy (97.6%) and average sensitivity (80.19%) with a maximum value reaching 93%).	Hypoglycemia were detected by BG prediction using neural network NARX (Nonlinear Autoregressive network with Exogenous Inputs) trained with Bayesian Regularization back propagation training Algorithm.
[6]	Support vector for regression - Free-living conditions (Nocturnal hypoglycemic - Sensitivity (30-min - 94% & 60-min - 94%) with time lags of 5.43 min and 4.57 min, respectively. Diurnal - without physical activities sensitivity (30-min - 92% and 60-min	Compared Support vector for regression, Multilayer perceptron, and Gaussian processes for prediction of nocturnal and diurnal hypoglycemic events. The study also investigated the effect of Physical activity information.

	-96%), with both time lags being less than 5 min. Diurnal – with physical activities decreases the sensitivity by 8% and 3%, respectively. Both nocturnal and diurnal predictions show a high (> 90%) precision.	
[7]	SVR - (30min- Average prediction Accuracy Hypo (87%) & Hyper (96%), 60min-Hypo (83%) & Hyper (94%)). GP (30min- Average prediction Accuracy Hypo (88%) & Hyper (95%), 60min- Hypo (85%) & Hyper (88%))	Compared support vector regression (SVR) and Gaussian process (GP).
[8]	Fuzzy neural network estimator algorithm (FNNE) predicted the onset of hypoglycemia episodes with a mean error of 0.071 (p < 0.03)	The FNNE algorithm was developed as a parallel combination of fuzzy inference mechanism (FIM) and a multi-layered neural network architecture.
[9] & [10]	Support vector regression (SVR) - with an event-based sensitivity of 100%, the algorithm produced only one false hypoglycemia detection. The sample-based sensitivity and specificity levels were 78% and 96%, respectively	Developed an android based system to detect hypoglycemia incidence using CGM and other information.
[11] & [12]	SVR with CGM- sample based Sensitivity 81%, and Specificity 93%	Compared CGM with and without SVR algorithm for hypoglycemia detection.
[13]	Classification and Regression Tree (CART) - Average accuracy (79.8%), average sensitivity (80.05%), overall specificity (79.9%)- The model was able to detect almost 80% of hypoglycemic events 15 min in advance	Investigated Classification and Regression Tree (CART) for hypoglycemia detection.
[14]	Decision tree (Accuracy (65.2%)) & Linear SVM (Accuracy (68.4%))	Investigated DT and SVM for hypoglycemia prediction.
[15]	ANN real time - Accuracy (85.2%), sensitivity (60%) and specificity (100%)	Invested real time and offline hypoglycemia detection using ECG signal.
[16] & [17] & [18]	Genetic algorithm based multiple regression with fuzzy inference - Sensitivity (75%) and specificity (over 50%)	Genetic algorithm is used to optimize regression and fuzzy rules. Compared various order multiple Regression Fuzzy Inference System and Linear multiple regression with various number of inputs.
[19]	Hybrid particle-swarm-optimization-based fuzzy-reasoning - Advanced hypoglycemic episodes (sensitivity (85.71%) & specificity (79.84%)) and hypoglycemic episodes (sensitivity (80.00%) & specificity (55.14%))	Investigated the applicability of PSO to optimize fuzzy rules and membership function of FRM. Compared with neural network and a regression method.
[20]	Hybrid rough set based neural network (RNN) - sensitivity (76.74%) and specificity (52.73%) whereas conventional FWNN with no rough approximation gives sensitivity (69.77%) and specificity (49.09%)	Hybrid particle swarm optimization with wavelet mutation (HPSOWM) is used to optimize RNN. Compared the result with a feedforward neural network (FWNN).
[21]	Combinational neural logic network with multiple regression - Sensitivity (79.07%) and Specificity (53.64%)	Hybrid particle swarm optimization with wavelet mutation (HPSOWM) is used to optimize the model parameters. Compared the result with neural logic network (NLN), wavelet neural network (WNN), feedforward neural network (FFNN), and multiple regression (MR)
[22]	Extreme learning machine (ELM)-based neural network -Sensitivity (78.00%) and Specificity (60.00%)	Compared the result with Particle swarm optimization based neural network (PSO-NN), Second order multiple regression fuzzy inference system (MR- FIS), Fuzzy inference system (FIS) and Linear multiple regression (LMR).
[24]	Fuzzy inference system with multiple regression - Sensitivity (80%) and Specificity (72.5%)	The fuzzy membership functions and rules are optimized using genetic algorithm. Compared the result with particle swarm optimization neural network.
[25]	Extreme learning machines (ELM)-the mean Specificity (95.4%) and the standard deviation (1.13)	Proposed and compared extreme learning machines (ELM) and regularized ELM (RELM) to predict hypoglycemia incidences using CGM readings.
[26]	Feedforward multi-layer neural network - Sensitivity (95.16%) and specificity (41.42%)	Proposed a neural network based hypoglycemia detection algorithm using ECG signal and skin impedance.
[27] & [28] & [29]	Bayesian neural network - Sensitivity (83.46%) and specificity (63.88%)	Investigated the applicability of Bayesian neural network to detect hypoglycemia from real time physiological parameters.
[30] & [31]	Particle Swarm Optimization-based Neural Network - Sensitivity (82%) and Specificity (63%)	Neural network parameters are optimized through PSO.
[32]	Neural network - Sensitivity (75%) and specificity (60%)	Investigated the possibility of combining Genetic Algorithm and Levenberg-Marquardt for neural network training in hypoglycemia detection algorithm.
[33]	Swarm-based support vector machine (SVM) - Sensitivity (82.14% ) and Specificity (60.19%)	Investigated SVM-RBF for hypoglycemia detections and optimized the parameters through PSO.
[34] & [35]	Fuzzy Support Vector Machine (FSVM-RBF) - (Sensitivity (74.19%), Specificity (58.54%), Accuracy (63.20%))	Compared FSVM and SVM along with three different kernel functions (radial basis function (RBF), exponential radial basis function (ERBF) and polynomial function) for the classification purpose.
[36] & [37]	Hybrid particle swarm - based fuzzy support vector machine (SFisSvm) - Sensitivity (75.19%), Specificity (83.71%) and Geometric mean (79.33%)	The FIS and SVM parameters are optimized using a hybrid particle swarm optimization with wavelet mutation algorithm. The swarm based SVM uses RBF kernel (SSvmR), sigmoid kernel (SSvmS) and linear kernel function (SSvmL).
[38]	Normalized radial basis function neural network (NRBFNN) - Sensitivity (76.74% ) and Specificity (51.82%)	The parameters of NRBFNN are optimized through hybrid particle swarm optimization with wavelet mutation (HPSOWM). Compared the result with radial basis function network, feedforward neural network, and multi regression.
[39] & [40]	Optimized variable translation wavelet neural network (VTWNN) - Sensitivity (79.07 %) and Specificity (50.00 %)	The parameters of VTWNN are optimized using a hybrid particle swarm optimization with wavelet mutation. Compared the result with wavelet neural network (WNN), feedforward neural network(FWNN2) and multi regression(MR)
[41] & [42]	Block Based Neural Network (BBNN) - Sensitivity (76.74%) and specificity (50.91%)	The BBNN parameters are optimized through a hybrid particle swarm optimization with wavelet mutation. Compared the result with feedforward neural networks and multiple regression.
[43]	Adaptive neural fuzzy inference system (ANFIS) – Sensitivity (79.09%) and specificity (51.82%)	The membership function and network parameters are optimized using swarm optimization with wavelet mutation (HPSOWM). Compared the result with fuzzy inference system (FIS), wavelet neural network (WNN), feedforward neural network (FWNN) and multiple regression (MR)
[44]	Combinational neural logic network (NLN) - Sensitivity (76.74%) and specificity (54.55%)	The NLN parameters are trained by hybrid particle swarm optimization with wavelet mutation (HPSOWM). Compared the result with neural logic network (NLN), wavelet neural network (WNN), feedforward neural network (FFNN) and multi regression (MR).
[45]	Hybrid rough-block-based neural network (R-BBNN) - Sensitivity (83.72%) and specificity (51.91%)	The R-BBNN parameters are optimized through a hybrid particle swarm optimization with wavelet mutation. Compared the result with BBNN, rough feedforward neural network (R-FWNN), wavelet neural network (WNN), SVM with a radial basis function and conventional feedforward neural network (FWNN).
[46]	Deep belief network (DBN) – Sensitivity (80.00%) and specificity (50.00%)	Compared the result with Block based neural network (BBNN), wavelet neural network (WNN), feedforward neural network (FFNN), and multiple regression (MR) models.
[47]	Extreme learning machine based feed-forward neural network (ELM-FFNN) - Sensitivity (78%) and specificity (60%)	Compared the result with multiple regression fuzzy inference system (MRFIS), Feed-forward neural network trained with particle swarm optimization (FFNN-PSO), Fuzzy inference system, and Linear multiple regression.

[48]	Classification tree - Predicted 82.12% of acute hypoglycemic events (specificity: 89.87%; positive predictive value: 88.72%; accuracy: 86.00%) and 76.99% of severe acute hypoglycemic events (80.53%, 74.31%, and 78.76% respectively).	Investigated towards predicting hypoglycemia incidence during intravenous (IV) insulin infusion for ICU patients.
[49]	Fuzzy inference system (FIS) - Sensitivity (75%) and Specificity (55%)	FIS parameters are tuned by an intelligent optimizer with two wavelet-mutated differential evolutions (WM-DE) engines. Compared the result with Neural network based rule discovery, linear multiple regression, evolved multiple regressions, feed-forward neural network (FFNN), and evolved fuzzy inference system.
[50]	Hypoglycemia (EWS (correct alarms=100%, detection time=16.7min, daily false alarms=0.08), EWS-DST (CA=100,DT=18.4min, DFA=1.0), EWS-GA(CA=100,DT=13min, DFA=0.17), EWS-GP(CA=100,DT=12.3min,DFA=0.17)), Hyperglycemia (EWS (CA=100,DT=14.7,DFA=0.8), EWS-DST (CA=100,DT=11.6min,DFA=0.73), EWS-GA (CA=100,DT=12.1min,DFA=0.73), EWS-GP(CA=100,DT=12min, DFA=0.33))	Investigated into advanced data fusion schemes for merging output of different hypo/hyperglycemia predictors such as Dempster-Schafer Evidential Theory and Evolutionary Methods (Genetic Algorithms, Genetic Programming). Compared the results with cARX and RNN models, and a linear fusion of the two.
[51]	ARX-based system - hypoglycemic (hyperglycemic) event prediction (accuracy of 100.0% (100.0%), detection time of 10.0 (8.0) min, and daily false alarms of 0.7 (0.5)). cARX-based system - Accuracy 100.0% (100.0%), DT 17.5 (14.8) min, & DFA 1.5 (1.3) and, RNN-based system Accuracy 100.0% (92.0%), DT 8.4 (7.0) min, and DFA 0.1 (0.2). The hybrid cARN-based EWS - 100.0% (100.0%) prediction accuracy, detection 16.7 (14.7) min in advance, and 0.8 (0.8) daily false alarms.	Investigated the performance improvement using a hybrid autoregressive with an output correction module/recurrent neural network (cARN). Compared performance of ARX, cARX, and RNN models.
[53]	Feed forward multi-layer neural network - Sensitivity (70.59%), specificity (65.38%) and geometric mean (67.94%)	Compared the ANN model with Linear Discriminant Analysis (LDA) and K-Nearest Neighbors (KNN) on hyperglycemia detection.
[52, 54]	Hidden Markov model (HMM) - The simulation result show that the proposed model is capable of detecting anomalies (i.e., no false positives) from the CGM readings based on historical data (in the presence of reasonable changes in the patient's daily routine).	Investigated the applicability of Hidden Markov model (HMM) in anomalies detection from the change in the patient's daily lifestyle.
[55]	Naïve Bayes classifier - matched the physicians' classifications 85% of the time that they were internally consistent and in agreement with each other.	Investigated into the applicability of characterizing blood glucose variability using new metrics with CGM data using Naïve Bayes classifier.
[56]	SVR models - When applied to 262 different CGM plots as a screen for excessive GV (accuracy (90.1%), sensitivity (97.0%), and specificity (74.1%).	Investigated the applicability of developing a perceived glycemic variability metric using SVM model. Compared the result with mean amplitude of glycemic excursion, standard deviation, distance travelled, and excursion frequency.
[57]	Multilayer Perceptron (MP)- (Accuracy 93.8%, Sensitivity 86.6%, Specificity 96.6%), Support Vector Machine (SVM)- (Accuracy 91.4%, Sensitivity 80.0%, Specificity 96.0%), Naive Bayes (NB) - (Accuracy 91.9%, Sensitivity 88.3%, Specificity 93.3%)	Investigated on an automatic glycemic variability detection and compared Naive Bayes (NB), Multilayer Perceptron (MP), and Support Vector Machine (SVM) models using CGM data.
[58]	Artificial Neural network – Average Accuracy (90%), Average sensitivity (72.23%) and Average specificity (92%)	Developed Artificial Neural Network integrated with physiological model for both blood glucose prediction and classification of hypoglycemia and further compared the result with existing models.
[59]	Bayesian regularized neural network - Sensitivity (73%) and specificity (60%)	Investigated and tested a feed-forward neural network trained with Bayesian regularization algorithm.

## References

1. Chan, K.Y., S.H. Ling, T.S. Dillon, and H. Nguyen. *Classification of hypoglycemic episodes for Type 1 diabetes mellitus based on neural networks*. in *IEEE Congress on Evolutionary Computation*. 2010. Barcelona, Spain.
2. Chan, K.Y., S.H. Ling, T.S. Dillon, and H.T. Nguyen, *Diagnosis of hypoglycemic episodes using a neural network based rule discovery system*. *Expert Systems with Applications*, 2011. **38**(8): p. 9799-9808.
3. Kit Yan, C., L. Sai Ho, H.T. Nguyen, and F. Jiang. *A hypoglycemic episode diagnosis system based on neural networks for Type 1 diabetes mellitus*. in *2012 IEEE Congress on Evolutionary Computation*. 2012. Brisbane, QLD, Australia.
4. Eljil, K.S., G. Qadah, and M. Pasquier. *Predicting hypoglycemia in diabetic patients using data mining techniques*. in *2013 9th International Conference on Innovations in Information Technology (IIT)*. 2013. Abu Dhabi, United Arab Emirates.
5. Eljil, K.S., G. Qadah, and M. Pasquier, *Predicting Hypoglycemia in Diabetic Patients Using Time-Sensitive Artificial Neural Networks*. *International Journal of Healthcare Information Systems and Informatics*, 2016. **11**(4): p. 70-88.
6. Georga, E.I., V.C. Protopappas, D. Ardigo, D. Polyzos, and D.I. Fotiadis, *A glucose model based on support vector regression for the prediction of hypoglycemic events under free-living conditions*. *Diabetes Technol Ther*, 2013. **15**(8): p. 634-43.
7. Georga, E.I., V.C. Protopappas, S.G. Mouggiakakou, and D.I. Fotiadis. *Short-term vs. long-term analysis of diabetes data: Application of machine learning and data mining techniques*. in *13th IEEE International Conference on Bioinformatics and BioEngineering*. 2013. Chania, Greece.

8. Ghevondian, N., H.T. Nguyen, and S. Colagiuri. *A novel fuzzy neural network estimator for predicting hypoglycaemia in insulin-induced subjects*. in *2001 Conference Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 2001. Istanbul, Turkey, Turkey.
9. Jensen, M.H., T.F. Christensen, L. Tarnow, Z. Mahmoudi, M.D. Johansen, and O.K. Hejlesen, *Professional Continuous Glucose Monitoring in Subjects with Type 1 Diabetes: Retrospective Hypoglycemia Detection*. *Journal of Diabetes Science and Technology*, 2013. **7**(1): p. 135-143.
10. Jensen, M.H., T.F. Christensen, L. Tarnow, M.D. Johansen, and O.K. Hejlesen, *An information and communication technology system to detect hypoglycemia in people with type 1 diabetes*. *Studies in health technology and informatics*, 2012. **192**: p. 38-41.
11. Jensen, M.H., T.F. Christensen, L. Tarnow, E. Seto, M. Dencker Johansen, and O.K. Hejlesen, *Real-time hypoglycemia detection from continuous glucose monitoring data of subjects with type 1 diabetes*. *Diabetes Technol Ther*, 2013. **15**(7): p. 538-43.
12. Jensen, M.H., Z. Mahmoudi, T.F. Christensen, L. Tarnow, E. Seto, M.D. Johansen, and O.K. Hejlesen, *Evaluation of an Algorithm for Retrospective Hypoglycemia Detection Using Professional Continuous Glucose Monitoring Data*. *J Diabetes Sci Technol*, 2014. **8**(1): p. 117-122.
13. Jung, M., Y.-B. Lee, S.-M. Jin, and S.-M. Park, *Prediction of Daytime Hypoglycemic Events Using Continuous Glucose Monitoring Data and Classification Technique*. arXiv preprint arXiv:1704.08769, 2017.
14. Jung, M. *Toward Designing Mobile Software to Predict Hypoglycemia for Patients with Diabetes*. in *2016 IEEE/ACM International Conference on Mobile Software Engineering and Systems (MOBILESoft)*. 2016. Austin, TX, USA.
15. Laione, F. and J. Marques, *Methodology for hypoglycaemia detection based on the processing, analysis and classification of the electroencephalogram*. *Medical and Biological Engineering and Computing*, 2005. **43**(4): p. 501-507.
16. Ling, S.H., H. Nguyen, and K.Y. Chan. *Genetic algorithm based fuzzy multiple regression for the nocturnal Hypoglycaemia detection*. in *IEEE Congress on Evolutionary Computation*. 2010. Barcelona, Spain.
17. Ling, S.S.H. and H.T. Nguyen, *Genetic-Algorithm-Based Multiple Regression With Fuzzy Inference System for Detection of Nocturnal Hypoglycemic Episodes*. *IEEE Transactions on Information Technology in Biomedicine*, 2011. **15**(2): p. 308-315.
18. Ling, S.H., H.T. Nguyen, and F.H.F. Leung. *Hypoglycemia detection using fuzzy inference system with genetic algorithm*. in *2011 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2011)*. 2011. Taipei, Taiwan.
19. Ling, S.H. and H.T. Nguyen, *Natural occurrence of nocturnal hypoglycemia detection using hybrid particle swarm optimized fuzzy reasoning model*. *Artif Intell Med*, 2012. **55**(3): p. 177-84.
20. Ling, S.H., P.P. San, H.K. Lam, and H.T. Nguyen. *Non-invasive detection of hypoglycemic episodes in Type 1 diabetes using intelligent hybrid rough neural system*. in *2014 IEEE Congress on Evolutionary Computation (CEC)*. 2014. Beijing, China.
21. Ling, S.H., P.P. San, H.K. Lam, and H.T. Nguyen, *Hypoglycemia detection: multiple regression-based combinational neural logic approach*. *Soft Computing*, 2015. **21**(2): p. 543-553.
22. Ling, S.H., P.P. San, and H.T. Nguyen, *Non-invasive hypoglycemia monitoring system using extreme learning machine for Type 1 diabetes*. *ISA Trans*, 2016. **64**: p. 440-446.
23. Marling, C., L. Xia, R. Bunescu, and F. Schwartz. *Machine learning experiments with noninvasive sensors for hypoglycemia detection*. in *Proceedings of IJCAI 2016 Workshop on Knowledge Discovery in Healthcare Data*. 2016. New York, NY.
24. Mathews, S., *Fuzzy Inference System And Multiple Regression For Detection Of Hypoglycemia*. *International Journal on Computational Science & Applications*, 2012. **2**(2): p. 37-50.



25. Mo, X., Y. Wang, and X. Wu. *Hypoglycemia prediction using extreme learning machine (ELM) and regularized ELM*. in *2013 25th Chinese Control and Decision Conference (CCDC)*. 2013. Guiyang, China.
26. Nguyen, H.T., N. Ghevondian, and T.W. Jones. *Neural-Network Detection of Hypoglycemic Episodes in Children with Type 1 Diabetes using Physiological Parameters*. in *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*. 2006. New York, NY, USA.
27. Nguyen, H.T., N. Ghevondian, and T.W. Jones. *Detection of nocturnal hypoglycemic episodes (natural occurrence) in children with Type 1 diabetes using an optimal Bayesian neural network algorithm*. in *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 2008. Vancouver, BC, Canada.
28. Nguyen, H.T., N. Ghevondian, S.T. Nguyen, and T.W. Jones. *Detection of Hypoglycemic Episodes in Children with Type 1 Diabetes using an Optimal Bayesian Neural Network Algorithm*. in *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 2007. Lyon, France.
29. Nguyen, H.T. and T.W. Jones. *Detection of nocturnal hypoglycemic episodes using EEG signals*. in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*. 2010. Buenos Aires, Argentina.
30. Nguyen, L.B., A.V. Nguyen, S.H. Ling, and H.T. Nguyen. *A particle swarm optimization-based neural network for detecting nocturnal hypoglycemia using electroencephalography signals*. in *The 2012 International Joint Conference on Neural Networks (IJCNN)*. 2012. Brisbane, QLD, Australia.
31. Nguyen, L.B., A.V. Nguyen, S.H. Ling, and H.T. Nguyen. *An adaptive strategy of classification for detecting hypoglycemia using only two EEG channels*. in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 2012. San Diego, CA, USA
32. Nguyen, L.B., A.V. Nguyen, S.H. Ling, and H.T. Nguyen. *Combining genetic algorithm and Levenberg-Marquardt algorithm in training neural network for hypoglycemia detection using EEG signals*. in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2013. Osaka, Japan.
33. Nuryani, S. Ling, and H.T. Nguyen. *Ventricular repolarization variability for hypoglycemia detection*. in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 2011. Boston, MA, USA.
34. Nuryani, S.H. Ling, and H.T. Nguyen. *Hypoglycaemia detection for type 1 diabetic patients based on ECG parameters using Fuzzy Support Vector Machine*. in *The 2010 International Joint Conference on Neural Networks (IJCNN)*. 2010. Barcelona, Spain.
35. Nuryani, S.L. and H.T. Nguyen. *Electrocardiographic T-wave peak-to-end interval for hypoglycaemia detection*. in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*. 2010. Buenos Aires, Argentina.
36. Nuryani, N., S.H. Ling, and H.T. Nguyen. *Hybrid particle swarm - based fuzzy support vector machine for hypoglycemia detection*. in *2012 IEEE International Conference on Fuzzy Systems*. 2012. Brisbane, QLD, Australia.
37. Nuryani, N., S.S. Ling, and H.T. Nguyen, *Electrocardiographic signals and swarm-based support vector machine for hypoglycemia detection*. *Ann Biomed Eng*, 2012. **40**(4): p. 934-45.
38. Phyo Phyo, S., S.H. Ling, and H.T. Nguyen. *Hybrid particle swarm optimization based normalized radial basis function neural network for hypoglycemia detection*. in *The 2012 International Joint Conference on Neural Networks (IJCNN)*. 2012. Brisbane, QLD, Australia.
39. Phyo Phyo, S., S.H. Ling, and H.T. Nguyen. *Optimized variable translation wavelet neural network and its application in hypoglycemia detection system*. in *2012 7th IEEE Conference on Industrial Electronics and Applications (ICIEA)*. 2012. Singapore, Singapore.

40. San, P.P., S.H. Ling, and H.T. Nguyen, *Hybrid PSO-based variable translation wavelet neural network and its application to hypoglycemia detection system*. Neural Computing and Applications, 2012. **23**(7-8): p. 2177-2184.
41. San, P.P., S.H. Ling, and H.T. Nguyen. *Block based neural network for hypoglycemia detection*. in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 2011. Boston, MA, USA.
42. San, P.P., S.H. Ling, and H.T. Nguyen, *Industrial Application of Evolvable Block-Based Neural Network to Hypoglycemia Monitoring System*. IEEE Transactions on Industrial Electronics, 2013. **60**(12): p. 5892-5901.
43. San, P.P., S.H. Ling, and H.T. Nguyen. *Intelligent detection of hypoglycemic episodes in children with type 1 diabetes using adaptive neural-fuzzy inference system*. in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 2012. San Diego, CA, USA.
44. San, P.P., S.H. Ling, and H.T. Nguyen. *Combinational neural logic system and its industrial application on hypoglycemia monitoring system*. in *2013 IEEE 8th Conference on Industrial Electronics and Applications (ICIEA)*. 2013. Melbourne, VIC, Australia
45. San, P.P., S.H. Ling, Nuryani, and H. Nguyen, *Evolvable Rough-Block-Based Neural Network and its Biomedical Application to Hypoglycemia Detection System*. IEEE Transactions on Cybernetics, 2014. **44**(8): p. 1338-1349.
46. San, P.P., S.H. Ling, and H.T. Nguyen. *Deep learning framework for detection of hypoglycemic episodes in children with type 1 diabetes*. in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2016. Orlando, FL, USA.
47. San, P.P., S.H. Ling, N.N. Soe, and H.T. Nguyen. *A novel extreme learning machine for hypoglycemia detection*. in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 2014. Chicago, IL, USA
48. Zhang, Y. *Predicting occurrences of acute hypoglycemia during insulin therapy in the intensive care unit*. in *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 2008. Vancouver, BC, Canada.
49. Lai, J.C.Y., F.H.F. Leung, and S.H. Ling, *Hypoglycaemia detection using fuzzy inference system with intelligent optimiser*. Applied Soft Computing, 2014. **20**: p. 54-65.
50. Botwey, R.H., E. Daskalaki, P. Diem, and S.G. Mougiakakou. *Multi-model data fusion to improve an early warning system for hypo-/hyperglycemic events*. in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 2014. Chicago, IL, USA
51. Daskalaki, E., K. Nørgaard, T. Züger, A. Proutzou, P. Diem, and S. Mougiakakou, *An Early Warning System for Hypoglycemic/Hyperglycemic Events Based on Fusion of Adaptive Prediction Models*. Journal of Diabetes Science and Technology, 2013. **7**(3): p. 689-698.
52. Zhu, Y., *Automatic detection of anomalies in blood glucose using a machine learning approach*. Journal of Communications and Networks, 2011. **13**(2): p. 125-131.
53. Nguyen, L.L., S. Su, and H.T. Nguyen. *Neural network approach for non-invasive detection of hyperglycemia using electrocardiographic signals*. in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 2014. Chicago, IL, USA.
54. Zhu, Y. *Automatic detection of anomalies in blood glucose using a machine learning approach*. in *2010 IEEE International Conference on Information Reuse & Integration*. 2010. Las Vegas, NV, USA
55. Marling, C.R., J.H. Shubrook, S.J. Vernier, M.T. Wiley, and F.L. Schwartz, *Characterizing Blood Glucose Variability Using New Metrics with Continuous Glucose Monitoring Data*. Journal of Diabetes Science and Technology, 2011. **5**(4): p. 871-878.
56. Marling, C.R., N.W. Struble, R.C. Bunescu, J.H. Shubrook, and F.L. Schwartz, *A Consensus Perceived Glycemic Variability Metric*. Journal of Diabetes Science and Technology, 2013. **7**(4): p. 871-879.

57. Wiley, M., R. Bunescu, C. Marling, J. Shubrook, and F. Schwartz. *Automatic Detection of Excessive Glycemic Variability for Diabetes Management*. in *2011 10th International Conference on Machine Learning and Applications and Workshops*. 2011. Honolulu, HI, USA.
58. Bertachi, A., L. Biagi, I. Contreras, N. Luo, and J. Vehí, *Prediction of Blood Glucose Levels And Nocturnal Hypoglycemia Using Physiological Models and Artificial Neural Networks*. 2018. p. 85-90.
59. Ngo, C.Q., B.C.Q. Truong, T.W. Jones, and H.T. Nguyen, *Occipital EEG Activity for the Detection of Nocturnal Hypoglycemia*. *Conf Proc IEEE Eng Med Biol Soc*, 2018. **2018**: p. 3862-3865.

## Paper 3

**Woldaregay, A. Z.,** Launonen, I. K., Albers, D., Igual, J., Årsand, E., & Hartvigsen, G. (2020). *A Novel Approach for Continuous Health Status Monitoring and Automatic Detection of Infection Incidences in People With Type 1 Diabetes Using Machine Learning Algorithms (Part 2): A Personalized Digital Infectious Disease Detection Mechanism. J Med Internet Res, 22(8), e18912. doi:10.2196/18912*

Original Paper

# A Novel Approach for Continuous Health Status Monitoring and Automatic Detection of Infection Incidences in People With Type 1 Diabetes Using Machine Learning Algorithms (Part 2): A Personalized Digital Infectious Disease Detection Mechanism

Ashenafi Zebene Woldaregay<sup>1</sup>, MSc; Ilkka Kalervo Launonen<sup>2</sup>, PhD; David Albers<sup>3,4</sup>, PhD; Jorge Iguual<sup>5</sup>, PhD; Eirik Årsand<sup>1</sup>, PhD; Gunnar Hartvigsen<sup>1</sup>, PhD

<sup>1</sup>Department of Computer Science, University of Tromsø – The Arctic University of Norway, Tromsø, Norway

<sup>2</sup>Department of Clinical Research, University Hospital of North Norway, Tromsø, Norway

<sup>3</sup>Department of Pediatrics, Informatics and Data Science, University of Colorado, Aurora, CO, United States

<sup>4</sup>Department of Biomedical Informatics, Columbia University, New York, NY, United States

<sup>5</sup>Universidad Politecnica Valencia, Valencia, Spain

**Corresponding Author:**

Ashenafi Zebene Woldaregay, MSc

Department of Computer Science

University of Tromsø – The Arctic University of Norway

Hansine Hansens veg 54, Science building Realfagbygget, office A124

Tromsø

Norway

Phone: 47 46359333

Email: [ashenafi.z.woldaregay@uit.no](mailto:ashenafi.z.woldaregay@uit.no)

## Abstract

**Background:** Semisupervised and unsupervised anomaly detection methods have been widely used in various applications to detect anomalous objects from a given data set. Specifically, these methods are popular in the medical domain because of their suitability for applications where there is a lack of a sufficient data set for the other classes. Infection incidence often brings prolonged hyperglycemia and frequent insulin injections in people with type 1 diabetes, which are significant anomalies. Despite these potentials, there have been very few studies that focused on detecting infection incidences in individuals with type 1 diabetes using a dedicated personalized health model.

**Objective:** This study aims to develop a personalized health model that can automatically detect the incidence of infection in people with type 1 diabetes using blood glucose levels and insulin-to-carbohydrate ratio as input variables. The model is expected to detect deviations from the norm because of infection incidences considering elevated blood glucose levels coupled with unusual changes in the insulin-to-carbohydrate ratio.

**Methods:** Three groups of one-class classifiers were trained on target data sets (regular days) and tested on a data set containing both the target and the nontarget (infection days). For comparison, two unsupervised models were also tested. The data set consists of high-precision self-recorded data collected from three real subjects with type 1 diabetes incorporating blood glucose, insulin, diet, and events of infection. The models were evaluated on two groups of data: raw and filtered data and compared based on their performance, computational time, and number of samples required.

**Results:** The one-class classifiers achieved excellent performance. In comparison, the unsupervised models suffered from performance degradation mainly because of the atypical nature of the data. Among the one-class classifiers, the boundary and domain-based method produced a better description of the data. Regarding the computational time, nearest neighbor, support vector data description, and self-organizing map took considerable training time, which typically increased as the sample size increased, and only local outlier factor and connectivity-based outlier factor took considerable testing time.

**Conclusions:** We demonstrated the applicability of one-class classifiers and unsupervised models for the detection of infection incidence in people with type 1 diabetes. In this patient group, detecting infection can provide an opportunity to devise tailored services and also to detect potential public health threats. The proposed approaches achieved excellent performance; in particular,

the boundary and domain-based method performed better. Among the respective groups, particular models such as one-class support vector machine, K-nearest neighbor, and K-means achieved excellent performance in all the sample sizes and infection cases. Overall, we foresee that the results could encourage researchers to examine beyond the presented features into other additional features of the self-recorded data, for example, continuous glucose monitoring features and physical activity data, on a large scale.

(*J Med Internet Res* 2020;22(8):e18912) doi: [10.2196/18912](https://doi.org/10.2196/18912)

## KEYWORDS

type 1 diabetes; self-recorded health data; infection detection; decision support techniques; outbreak detection system; syndromic surveillance

## Introduction

Anomaly or novelty detection problem involves identifying the anomalous or novel instances, which exhibit different characteristics, from the rest of the data set and has been widely used in various applications including machine fault and sensor failure detection, prevention of credit card or identity fraud, health and medical diagnostics and monitoring, cyber-intrusion detection, and others [1-7]. The term anomaly was precisely coined by Hawkins [8] as “observations that deviate much from the other observations so as to arouse suspicions that it could be generated by a different process.” Anomalousness is usually described as point, contextual, and collective, depending on how the degree of anomaly is computed [1,7,9]. On the basis of the necessity of having labeled data instances for the respective class, the anomaly detection problem can be approached as supervised, semisupervised, and unsupervised [3,7,9-11]. Supervised anomaly detection, for example, multiclass classification, requires labeled data instances for both the target and the nontarget (anomaly) classes. This characteristic makes it impractical for tasks where there is difficulty in either finding enough samples for the anomaly class, that is, poorly sampled and unbalanced data, or demarcating boundaries of the anomaly class [7,10,12]. Moreover, anomalies could also evolve over time, and what is known today might not be valid through time, making the characterization of anomalies class more challenging. In this case, semisupervised anomaly detection, that is, one-class classification, is preferred given that it only requires characterizing what is believed to be normal (target data instances) to detect the abnormal (nontarget data instances) [7]. Under certain circumstances, for example, medical domain, obtaining and demarcating the anomalous (nontarget) data instances can become very difficult, expensive, and time consuming, if not impossible [7,13]. For instance, assume a health diagnostic and monitoring system that detects health changes in an individual by tracking the individual’s physiological parameters, where the current health status is examined based on a set of parameters, and raises a notification alarm when the individual health deteriorates [12]. In such a system, it becomes feasible to rely on a method that can be trained using only the regular or normal day measurements (target days) so as to detect deviation from normality [12,14]. This is because demarcating the exact boundaries between normal and abnormal health conditions is very challenging given that each pathogen has a different effect on the individual physiology. The one-class classifiers–based anomaly detection

methods can be roughly grouped into 3 main groups: boundary and domain-based, density-based, and reconstruction-based methods based on how their internal function is defined and the approach used for minimization [3,10,12,13,15,16]. These models take into account different characteristics of the data set, and depending on the data set under consideration, these models could achieve different generalization performance, overfitting, and bias [12]. Unlike supervised and semisupervised anomaly detection methods, unsupervised methods do not require labeled instances to detect the anomaly (nontarget) instances because they rely on the entire data set to determine the anomalies and can be another possible alternative to semisupervised anomaly detection methods [7,10,12]. One of the drawbacks of unsupervised methods is that they require significant amount of data to achieve comparable performance. Both semisupervised and unsupervised methods have been used in various applications to detect anomalous instances [1,7,10,16]. In particular, these methods have been popular in the medical domain owing to their suitability for such applications, where there is lack of a sufficient data set for the other classes [13]. Accordingly, considering the difficulty and expense of obtaining enough sample data sets for the infection days from people with type 1 diabetes, a one-class classifier and unsupervised models are proposed for detecting infection incidence in people with type 1 diabetes.

Type 1 diabetes, also known as insulin-dependent diabetes, is a chronic disease of blood glucose regulation (hemostasis), and is caused by the lack of insulin secretion from pancreatic cells [17,18]. In people with type 1 diabetes, the incidence of infection often results in hyperglycemia and frequent insulin injection [19-26]. Infection-induced anomalies are characterized by violation of the norm of blood glucose dynamics, where blood glucose remains elevated despite taking a higher amount of insulin injection with less carbohydrate consumption [19]. Despite these potentials, there have been very few studies that focused on detecting infection incidence in individuals with type 1 diabetes using a dedicated personalized health model. Therefore, the objective of this study was to develop an algorithm, that is, a personalized health model that can automatically detect the incidence of infection in people with type 1 diabetes using blood glucose levels and insulin-to-carbohydrate ratio as input variables. For this, a one-class classifier and unsupervised models are proposed. The model is expected to detect any deviations from the norm because of infection incidences considering elevated blood glucose level (hyperglycemia incidences) coupled with unusual changes in the insulin-to-carbohydrate ratio, that is, frequent

insulin injections and unusual reduction in the amount of carbohydrate intake [19]. Three groups of one-class classifiers and two unsupervised density-based models were explored. A detailed theoretical description of the proposed models is given in [Multimedia Appendix 1](#) [1,7-16,27-37]. The anomaly detection problem studied in this paper can be regarded as a contextual anomaly, where the ratio of insulin-to-carbohydrate is the context and the average blood glucose level is the behavioral attribute. This is mainly because of the fact that elevated blood glucose levels do not always signify being anomalies without looking at the context of the ratio of insulin-to-carbohydrate in this case. Throughout the paper, the term object is used to describe a feature vector incorporating the number of parameters under consideration. For example, an object  $X$  can define a specific event of an individual blood glucose dynamics at a specified time index  $k$  and is represented by a feature vector  $X_k=(x_{k,1}, x_{k,2})$ , where  $x_{k,1}$  represents the ratio of total insulin-to-total carbohydrate and  $x_{k,2}$  represents the average blood glucose level in a specific time-bin (interval) around  $k$ .

## Methods

A group of one-class classifiers and unsupervised models were tested and compared. The one-class classifier incorporates 3 groups: boundary and domain-based, density-based, and reconstruction-based methods. The boundary and domain-based method contains support vector data description (SVDD) [27], one-class support vector machine (V-SVM) [28], incremental support vector machine [29], nearest neighbor (NN) [12], and minimum spanning tree (MST) [15]. Density-based method includes normal Gaussian [32], minimum covariance Gaussian [38], mixture of Gaussian (MOG) [32], Parzen [39], naïve Parzen [32], K-nearest neighbor (KNN) [12,30], and local outlier factor (LOF) [31]. The reconstruction-based method includes principal component analysis (PCA) [12,32], K-means [32], self-organizing maps (SOM) [12,32], and auto-encoder networks [12]. In addition, the unsupervised models were also tested, including the LOF [31,33] and the connectivity-based outlier factor (COF) [33,34]. The input variables, average blood glucose levels and ratio of total insulin (bolus) to total carbohydrate, used in training and testing of the models were selected in accordance with the description provided by Woldaregay et al

[19], and the ratio was calculated by dividing the total insulin with the total carbohydrate within a specified time-bin. The data set consists of high-precision self-recorded data collected from 3 real subjects (2 males and 1 female; average age 34 [SD 13.2] years) with type 1 diabetes. It incorporates blood glucose levels, insulin, carbohydrate information, and self-reported infections cases of influenza (flu) and, mild and light common cold without fever, as shown in [Table 1](#). Exemplar data depicting the model's input features for 2 specific patient years with and without infection are shown in [Figures 1-4](#), and a more detailed description of the input features for 10-patient years with and without infection incidences can be found in [Multimedia Appendix 2](#) [12,19]. The data were resampled and imputed in accordance with the description provided by Woldaregay et al [19], and the preprocessed data were smoothed using a moving average filter of 2 days' (48 hours) window size to remove short-term and small-scale features [19,40,41]. Feature scaling was carried out using min-max scaling [42] to normalize the data between 0 and 1, which is important to ensure that larger parameters do not dominate the smaller ones. The data sets are labeled as target and nontarget data sets, where the target data sets include all the self-recorded normal period of the year and the nontarget data set includes only the self-reported infection periods when the individual was sick. Accordingly, the one-class classifiers were trained using only the target data sets containing the regular or normal period of the year and tested using both the target and the nontarget (infection period) data sets. For the unsupervised models, all the data sets containing both the target and the nontarget data sets were presented during testing. The hyperparameters of most of the one-class classifiers were optimized using a consistency approach [43]. Models such as naïve Parzen and Parzen were optimized using the leave-one-out method. For MST, the entire MST was used. For PCA, the fraction of variance retained from the training data set was set to be 0.67. The models were evaluated based on different characteristics including data nature (with and without filter), data granularity (hourly and daily), data sample size, and required computational time. All the experiments were conducted using MATLAB 2018b (Mathworks, Inc). Most of the models were implemented using *ddtools*, *prtools*, and *anomaly detection toolbox*, which are MATLAB toolboxes [32,33,35].

**Table 1.** Equipments used in the self-management of diabetes.

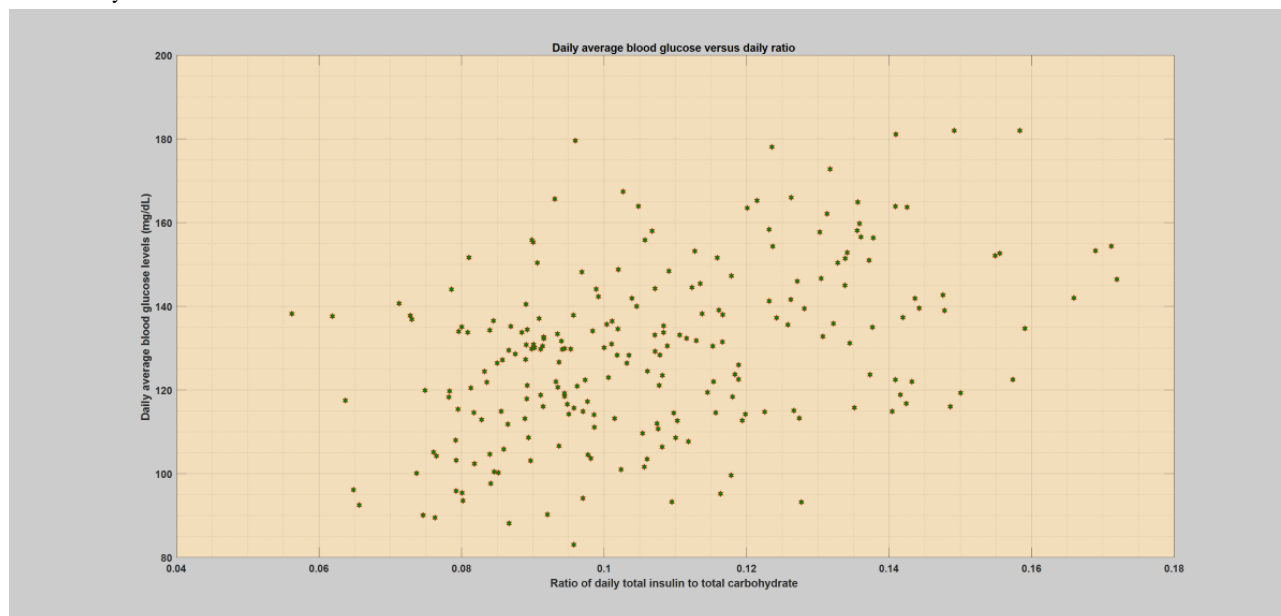
Patients	Self-management				
	BG <sup>a</sup>	Insulin administration	Diet	Body weight (kg)	HbA <sub>1c</sub> <sup>b</sup> (%)
Subject 1	Finger pricks recorded in the Diabetes Diary mobile app and Dexcom CGM <sup>c</sup>	Insulin Pen (multiple bolus and 1-time basal in the morning) recorded in the Diabetes Diary mobile app	Carbohydrate in grams recorded in the Diabetes Diary mobile app; level 3 (advanced carb counting)	83	6.0
Subject 2	Finger pricks recorded in the Spike mobile app and Dexcom G4 CGM <sup>c</sup>	Insulin Pen (multiple bolus [Humalog] and 1-time basal [Toujeo] before bed) recorded in the Spike mobile app	Carbohydrate in grams recorded in the Spike mobile app; level 3 (advanced carb counting)	77	7.3
Subject 3	Enlite (Medtronic) CGM <sup>c</sup> and Dexcom G4	Medtronic MinMed G640 insulin pump (basal rates profile [Fiasp] and multiple bolus [Fiasp])	Carbohydrate in grams recorded in pump information; level 3 (advanced carb counting)	70	6.2

<sup>a</sup>BG: blood glucose.

<sup>b</sup>HbA<sub>1c</sub>: hemoglobin A<sub>1c</sub>.

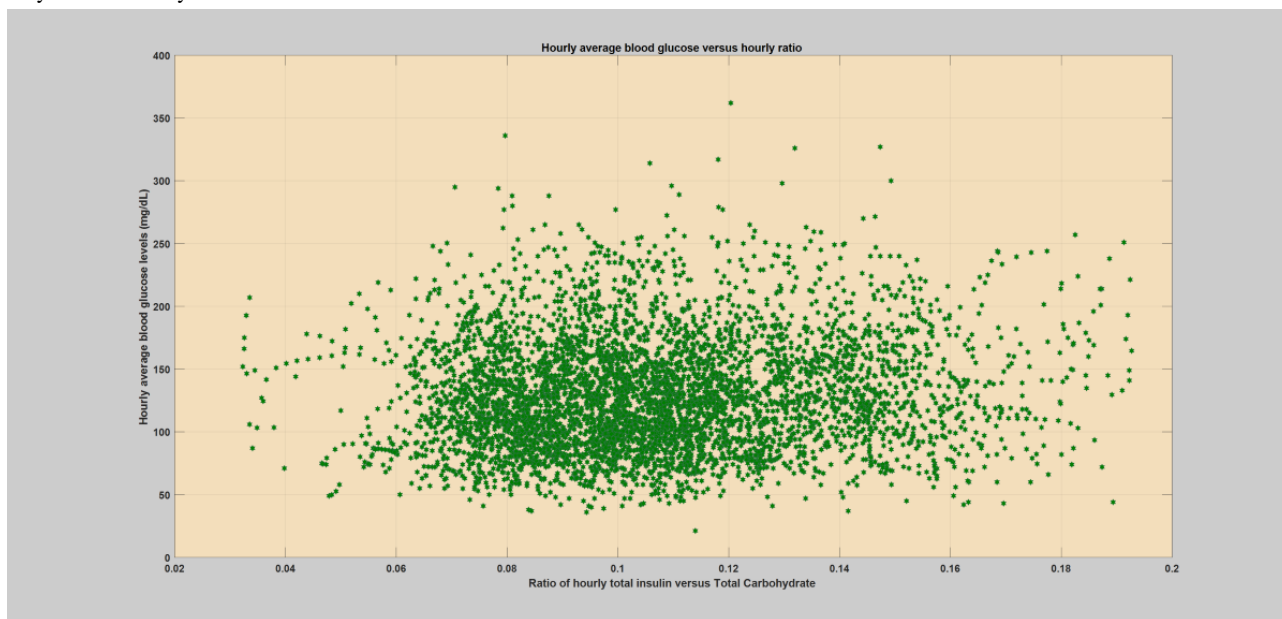
<sup>c</sup>CGM: continuous glucose monitoring.

**Figure 1.** Daily scatter plot of average blood glucose levels versus total insulin (bolus) to total carbohydrate ratio for a specific regular or normal patient year without any infection incidences.

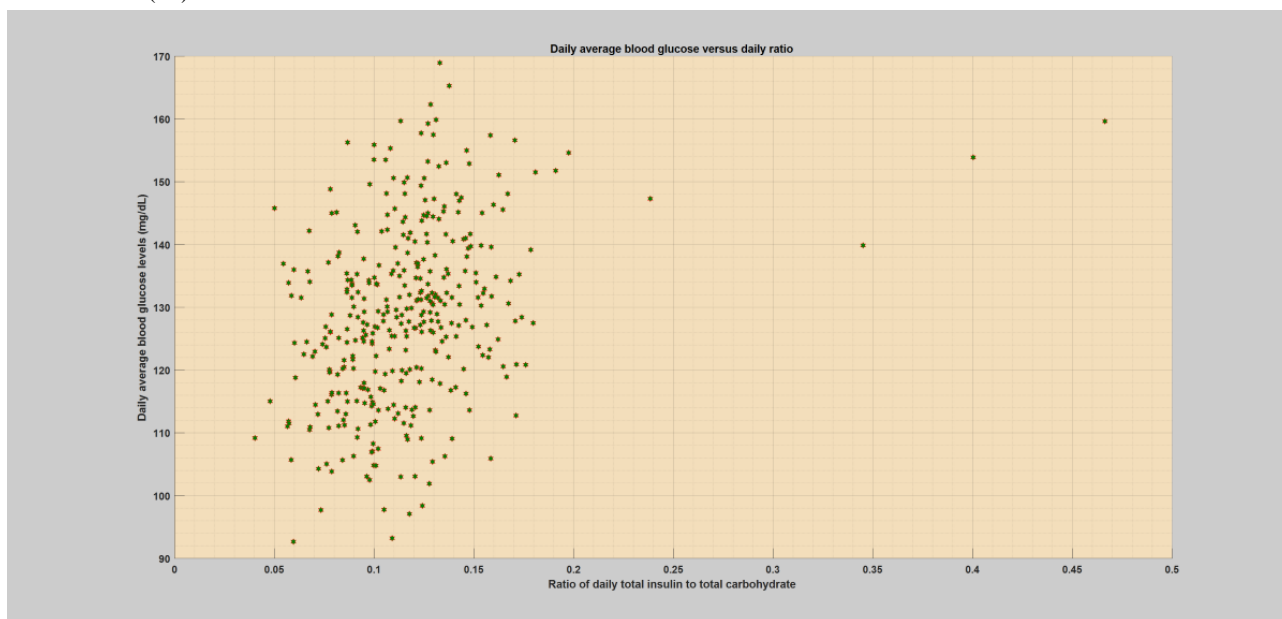




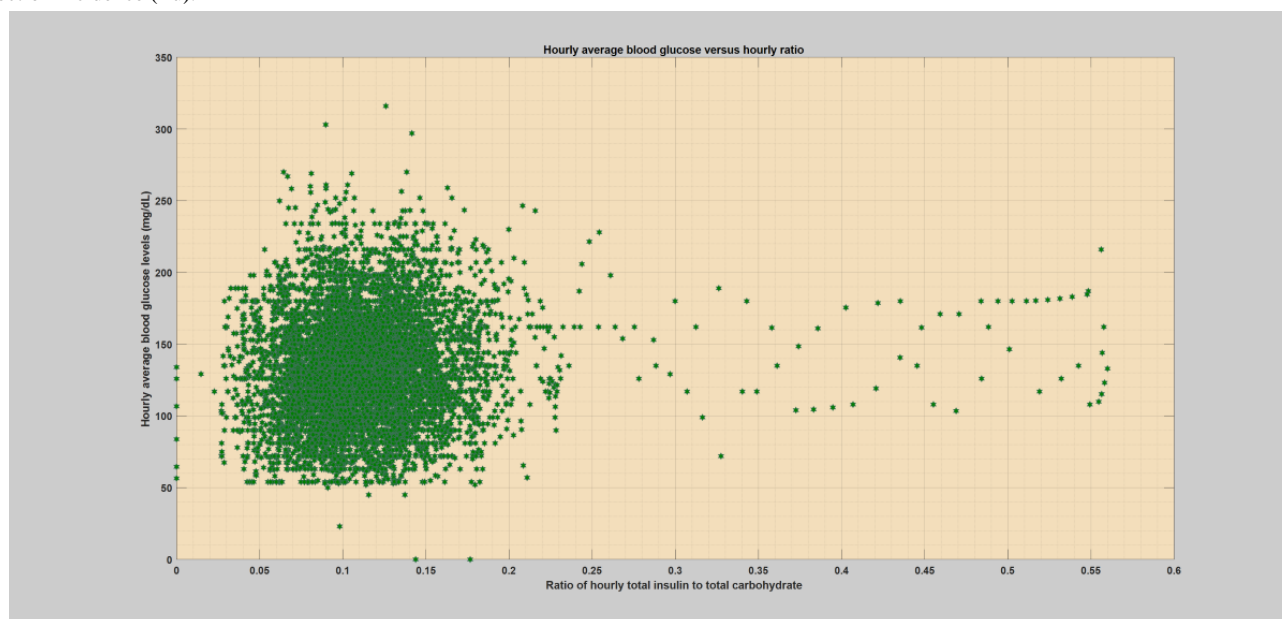
**Figure 2.** Hourly scatter plot of average blood glucose levels versus total insulin (bolus) to total carbohydrate ratio for a specific regular or normal patient year without any infection incidences.



**Figure 3.** Daily scatter plot of average blood glucose levels versus total insulin (bolus) to total carbohydrate ratio for a specific patient year with an infection incidence (flu).



**Figure 4.** Hourly scatter plot of average blood glucose levels versus total insulin (bolus) to total carbohydrate ratio for a specific patient year with an infection incidence (flu).



## Model Evaluation

The performance of the one-class classifiers was evaluated using 20 times 5-fold stratified cross-validation. For both daily and hourly cases, the user-specified outlier fraction threshold  $\beta$  was set to 0.01 such that 1% of the training target data are allowed to be classified as outlier or get rejected [12]. Class imbalance was mitigated by oversampling of the nontarget data sets through random sampling [44]. Performance was measured using the area under the receiver operating characteristic (ROC) curve (AUC), specificity, and F1-score [45-48]. The AUC, specificity, and F1-score were reported as the average (SD) of twenty times five-fold stratified cross-validation rounds. AUC is the result of integration (summation) of the ROC curve over a range of possible classification thresholds [49]. It is regarded as robust (insensitive) when it comes to the presence of data imbalance; however, it is impractical for real-world implementation because it is independent of a single threshold [48]. Specificity measures the ratio of correctly classified negative samples from the total number of available negative samples [50]. Thus, it depicts the proportion of infection days (nontarget samples) that are correctly classified as such to the total number of infection days (period). It is only used to examine how the model performs in regard to the nontarget class (infection days). F1-score is the harmonic mean of precision and recall, where the value ranges from 0 to 1, and high F1 scores depict high classification performance [45]. F1-score is considered appropriate when evaluating model performance with regard to one target class and in the presence of unbalanced data sets [10,46-48]. The models were further compared based on various criteria, which can contribute to the implementation of the models in real-world settings, including computation time, sample size, number of user-defined parameters, and sensitivity to outliers in the training data sets:

- **Computation time:** this characteristic defines the amount of time taken to train and test the model. Regarding personal use, response time is crucial for acceptance of the services

by a wide range of users. Furthermore, with regard to the outbreak detection settings, this is an important parameter given that a system that uses data from many participants needs to have an acceptable response time. However, in real-world applications, the training phase can be performed in an offline mode, which makes the testing response time very crucial.

- **Sample size:** this characteristic specifies the minimum amount of training data required to generate an acceptable performance. This is an important factor given that the system relies on self-recorded data; it is difficult to accumulate a large set of data for an individual initially.
- **Number of user-defined parameters:** this characteristic defines the complexity of the model. It is simpler and less data are required to estimate a model with fewer parameters. This is an important factor because it is easier for an individual to implement the simple model compared with the complex model.
- **Sensitivity to outliers in the training data sets:** this characteristic defines how the model estimation is affected by outliers in the training set. This is a crucial characteristic because the model training depends on self-reported data, which are highly dependent on the accuracy of the user data registration. It is possible that the user might forget to report some infection incidence and hence might be considered as target data sets and be used as a training data set. Furthermore, errors incurred during manual registration of data can also affect model generalization.

## Data Collection and Ethical Declaration

The study protocol has been submitted to the Norwegian Regional Committees for Medical Health Research Northern Norway for evaluation and was found exempted from regional ethics review because it is outside the scope of medical research (reference number: 108435). Written consent was obtained, and the participants donated the data sets. All data from the participants were anonymized.

## Results

The models were evaluated based on two different versions of the same data set: raw and filtered. The input variables to the models were the average blood glucose levels and the ratio of total insulin (bolus)-to-total carbohydrate. The necessary computational time for both training and testing of the models was also estimated. A comparison of the classifiers was carried out taking into account their performance, necessary sample size for producing acceptable performance, and computational time. These models were further compared based on their theoretical guarantee provided for robustness to outliers in the target data set and based on their complexity. In addition, these classifiers were compared with the unsupervised version of some selected models.

### Model Evaluation

Model training and evaluations were carried out on an individual basis taking into account different characteristics of the data, specified time window or resolution (hourly and daily), and nature of the data (raw data and its smoothed version). For daily evaluation, we compared the performance of the models on raw data and its smoothed version with a 2-day moving average filter. For hourly evaluation, we compared the performance of the model on a smoothed version of the data set. The purpose of the comparison was to study the performance gain achieved by removing short-time noises from the data set through smoothing. The average and SD of AUC, specificity, and F1-score are computed and reported for each model. The top performing models from each category are highlighted in italics within each tables.

### Semisupervised Models

The regular or normal days were labeled as the target class data set and the infection period as the nontarget class data set. Three groups of one-class classifiers were trained on the target class and tested on a data set containing both the target and the nontarget classes. In addition to the data characteristics stated above, resolution and data nature, the one-class classifier performance was also assessed taking into account the required sample object size to produce acceptable data description. In this direction, we consider four groups of sample size: 1 month, 2 months, 3 months, and 4 months data sets. In the model evaluation, the data set containing the infection period was presented during testing. The evaluation was carried out based on 20 times 5-fold stratified cross-validation. The performance of the model was reported as the average and SD of AUC, specificity, and F1-score of the rounds. A score plot of each model for both the hourly and the daily scenarios using the smoothed version of the data can be found in [Multimedia Appendix 3](#), where the models were trained on random 120 regular or normal days of the patient year and tested over the whole year.

### Daily

As can be seen in [Tables 2](#) and [3](#) below (see also [Multimedia Appendix 4](#)), the performance of the models generally improves as the size of the sample increases. The models performed well with respect to the raw data sets; however, the performance significantly improved with the smoothed version of the data. The results indicate that the sample size greatly affects the model performance and that there is a larger variation in performance when the training data set is small. Generally, it can be seen that the models generalize well with the 3-month data set (90 sample objects) and further improve after 3 months. In general, on average, with both the raw and smoothed data sets, the boundary and domain-based method performed better with a small sample size. As the sample size increased, all the three groups produced comparable descriptions of the data. From each respective category, models such as V-SVM, K-NN, and K-means performed well across all the sample sizes.

### First Case of Infection (Flu)

The boundary and domain-based method achieved a better description of the data with a small sample size when compared with the other two groups. However, as the sample size increased, all the three groups achieved relatively comparable descriptions of the data. Specific models such as V-SVM, K-NN, and K-means performed better from their respective group. Regarding the raw data, as seen in [Table 2](#), all the models failed to generalize from the 1-month data set as compared with the large sample objects, that is, 3 months, which was expected:

1. From the boundary and domain-based method, V-SVM performed better in all the sample sizes and achieved comparable performance even with 60 objects and improved significantly afterward. SVDD produced a comparable description with higher sample sizes, that is, 3 months and later.
2. From the density-based method, K-NN performed better in all the sample sizes and achieved better performance even with 60 objects. Naïve Parzen produced comparable performance with higher sample sizes, that is, 3 months and later.
3. From the reconstruction-based method, K-means achieved better performance for all sample sizes.

Smoothing the data, as shown in [Table 3](#), improved the model performance even with 30 sample objects:

1. From the boundary and domain-based method, V-SVM achieved better performance in all sample sizes.
2. From the density-based method, K-NN achieved better performance for all sample sizes, minimum covariance determinant (MCD) Gaussian produced a comparable description with 30 and 60 sample objects, and naïve Parzen achieved comparable description of the data with 4-month sample objects.
3. Regarding the reconstruction-based method, PCA achieved good performance with 30 and 60 sample objects, whereas K-means performed better with larger sample objects.

**Table 2.** Average (SD) of area under the receiver operating characteristic curve, specificity, F1-score for the raw data set (without smoothing), and different sample size. Fraction=0.01.

Models	1 month			2 months			3 months			4 months		
	AUC <sup>a</sup> , mean (SD)	Specificity, mean (SD)	F1, mean (SD)	AUC, mean (SD)	Specificity, mean (SD)	F1, mean (SD)	AUC, mean (SD)	Specificity, mean (SD)	F1, mean (SD)	AUC, mean (SD)	Specificity, mean (SD)	F1, mean (SD)
<b>Boundary and domain-based method</b>												
SVDD <sup>b</sup>	90.7 (8.8)	71.7 (7.7)	73.6 (5.5)	93.4 (6.2)	81.7 (5.0)	87.4 (8.1)	96.4 (2.9)	87.8 (3.3)	91.3 (6.0)	94.6 (3.7)	81.7 (5.0)	90.0 (4.6)
IncSVDD <sup>c</sup>	90.4 (8.9)	66.7 (7.5)	72.7 (4.9)	91.8 (5.9)	66.7 (7.5)	84.4 (3.2)	95.8 (2.9)	70.0 (7.1)	85.4 (1.2)	93.7 (3.6)	55 (10.7)	81.0 (2.7)
V-SVM <sup>d</sup>	93.1 (6.0)	63 (10.6)	78.9 (6.2) <sup>e</sup>	96.5 (2.3)	81.9 (4.7)	90.7 (3.4)	97.9 (1.5)	88.9 (0.0)	94.1 (2.0)	96.2 (2.3)	83.3 (0.0)	91.7 (1.4)
NN <sup>f</sup>	74.2 (9.3)	38.3 (7.7)	61.0 (4.7)	89.5 (9.3)	20.0 (6.7)	70.0 (4.6)	90.1 (6.6)	11.1 (18)	69.2 (3.8)	92.8 (3.3)	33.3 (0.0)	75.1 (0.4)
MST <sup>g</sup>	89.4 (8.1)	50.0 (0.0)	62.7 (6.6)	95.4 (5.6)	61.7 (7.7)	82.3 (5.9)	96.6 (2.7)	68.9 (4.5)	83.6 (4.7)	94.1 (2.8)	55.0 (7.7)	80.6 (2.3)
<b>Density-based method</b>												
Gaussian	90.6 (7.1)	60.0 (8.2)	68.8 (8.4)	95.4 (4.6)	70.0 (6.7)	85.3 (4.6)	97.3 (2.5)	80.0 (4.5)	89.2 (3.3)	95.5 (3.2)	66.7 (0.0)	84.5 (2.0)
MOG <sup>h</sup>	88.1 (9.9)	80.1 (17.3)	67.8 (16.4)	93.1 (7.1)	75.8 (14.8)	82.5 (10.1)	95.6 (3.4)	80.2 (7.5)	86.0 (6.7)	93.7 (3.9)	68.7 (11.6)	84.2 (5.7)
MCD <sup>i</sup> Gaussian	89.0 (8.5)	55.0 (7.7)	66.4 (9.0)	94.0 (4.6)	68.3 (5.0)	84.6 (6.3)	97.0 (2.7)	80.0 (4.5)	89.9 (2.4)	94.5 (3.2)	65.0 (5.0)	84.0 (3.2)
Parzen	89.0 (9.2)	70.0 (6.7)	70.7 (5.9)	94.6 (4.9)	83.3 (0.0)	87.9 (6.3)	97.2 (2.4)	88.9 (0.0)	90.5 (5.9)	95.2 (2.9)	83.3 (0.0)	88.9 (3.3)
Naïve Parzen	90.1 (7.6)	55 (10.7)	65.0 (5.0)	95.7 (3.9)	76.7 (8.2)	87.2 (3.5)	98.3 (1.4)	88.9 (0.0)	93.6 (2.4)	96.8 (2.1)	83.3 (0.0)	90.7 (2.0)
K-NN <sup>j</sup>	91.8 (6.9)	50.0 (0.0)	66.0 (2.0)	95.6 (3.1)	81.7 (5.0)	90.9 (3.2)	97.9 (1.6)	88.9 (0.0)	93.5 (3.7)	97.0 (2.2)	83.3 (0.0)	92.0 (1.0)
LOF <sup>k</sup>	88.5 (6.1)	66.7 (7.5)	72.7 (4.9)	97.0 (1.9)	71.7 (7.7)	86.1 (2.4)	96.8 (2.8)	78.9 (3.3)	88.7 (2.8)	92.6 (4.8)	50.0 (0.0)	79.3 (2.6)
<b>Reconstruction-based method</b>												
PCA <sup>l</sup>	87.8 (11.9)	50.0 (7.5)	62.4 (8.5)	93.5 (6.2)	51.7 (5.0)	78.2 (4.1)	93.6 (4.7)	60 (10.2)	81.8 (4.4)	91.3 (5.2)	46.7 (6.7)	78.7 (2.3)
Auto-encoder	82.2 (12.0)	57.9 (15.3)	64.7 (12.0)	88.2 (9.5)	61.6 (14.0)	81.4 (7.1)	93.4 (5.7)	74.4 (11)	86.4 (5.9)	88.4 (8.8)	61.3 (14.3)	82.7 (5.7)
SOM <sup>m</sup>	86.9 (9.4)	78.3 (13.3)	66.7 (16.9)	92.8 (7.3)	64.2 (12.4)	80.9 (7.0)	95.8 (3.7)	80.1 (6.3)	86.9 (5.5)	92.2 (4.1)	76.5 (9.0)	87.5 (4.5)
K-means	91.8 (6.9)	65.0 (9.0)	71.8 (5.1)	96.0 (2.4)	83.3 (0.0)	91.5 (2.8)	97.6 (1.6)	88.9 (0.0)	93.5 (3.7)	96.2 (2.2)	83.3 (0.0)	91.5 (1.6)

<sup>a</sup>AUC: area under the receiver operating characteristic curve.

<sup>b</sup>SVDD: support vector data description.

<sup>c</sup>IncSVDD: incremental support vector data description.

<sup>d</sup>V-SVM: one-class support vector machine.

<sup>e</sup>Italicized values indicates the top performing models.

<sup>f</sup>NN: nearest neighbor.

<sup>g</sup>MST: minimum spanning tree.

<sup>h</sup>MOG: mixture of Gaussian.

<sup>i</sup>MCD: minimum covariance determinant.

<sup>j</sup>K-NN: K-nearest neighbor.

<sup>k</sup>LOF: local outlier factor.

<sup>l</sup>PCA: principal component analysis.

<sup>m</sup>SOM: self-organizing maps.

**Table 3.** Average of area under the receiver operating characteristic curve, specificity, and F1-score for smoothed version of the data with a 2-day moving average filter and different sample size. Fraction=0.01.

Models	1 month			2 months			3 months			4 months		
	AUC <sup>a</sup> , mean (SD)	Specificity	F1	AUC <sup>a</sup> , mean (SD)	Specificity	F1	AUC <sup>a</sup> , mean (SD)	Specificity	F1	AUC <sup>a</sup> , mean (SD)	Specificity	F1
<b>Boundary and domain-based method</b>												
SVDD <sup>b</sup>	99.6 (1.3)	100 (0.0)	93.6 (15.2)	100 (0.0)	100 (0.0)	94.8 (10.1)	100 (0.0)	100 (0.0)	97.0 (4.1)	100 (0.0)	100 (0.0)	96.9 (4.0)
IncSVDD <sup>c</sup>	99.6 (1.3)	100 (0.0)	93.6 (15.2)	100 (0.0)	100 (0.0)	97.1 (6.3)	100 (0.0)	100 (0.0)	97.6 (4.1)	100 (0.0)	100 (0.0)	98.3 (2.8)
V-SVM <sup>d</sup>	100 (0.0)	99.5 (2.9)	98.9 (3.2) <sup>e</sup>	100 (0.0)	100 (0.0)	99.1 (2.6)	100 (0.0)	100 (0.0)	99.4 (1.7)	100 (0.0)	100 (0.0)	99.6 (1.2)
NN <sup>f</sup>	98.1 (3.9)	58.3 (15.4)	72.3 (9.9)	86.9 (12.5)	16.7 (22.4)	70.5 (5.3)	88.1 (6.5)	54.4 (22.5)	80.0 (8.6)	92.4 (5.3)	8.3 (17.1)	69.0 (4.8)
MST <sup>g</sup>	98.5 (2.4)	85.0 (5.0)	85.5 (2.1)	99.7 (0.8)	100 (0.0)	97.1 (6.3)	99.9 (0.4)	97.8 (4.5)	97.2 (4.0)	99.7 (0.8)	100 (0.0)	97.0 (7.9)
<b>Density-based method</b>												
Gaussian	100 (0.0)	98.3 (5.0)	92.1 (15.2)	100 (0.0)	100 (0.0)	97.1 (6.3)	99.8 (0.7)	100 (0.0)	97.6 (4.1)	99.4 (1.7)	100 (0.0)	97.0 (7.9)
MOG <sup>h</sup>	98.6 (3.2)	99.8 (1.7)	88.5 (16.8)	99.6 (1.2)	100 (0.0)	92.2 (11.1)	99.7 (0.7)	99.8 (1.4)	94 (10.3)	99.3 (2.0)	99.9 (1.2)	94.4 (11.8)
MCD <sup>i</sup> Gaussian	98.9 (2.2)	91.7 (8.4)	90.9 (7.7)	100 (0.0)	100 (0.0)	98.0 (6.0)	99.5 (1.1)	96.7 (5.1)	96.6 (5.9)	99.4 (1.7)	88.3 (7.7)	92.0 (6.8)
Parzen	99.6 (1.3)	100 (0.0)	87.7 (17.0)	100 (0.0)	100 (0.0)	95.1 (8.0)	100 (0.0)	100 (0.0)	94.6 (9.8)	99.9 (0.4)	100 (0.0)	94.6 (12.3)
Naïve Parzen	99.2 (2.5)	100 (0.0)	94.7 (11.1)	100 (0.0)	100 (0.0)	93.8 (11.0)	99.6 (1.1)	100 (0.0)	97.5 (5.0)	100 (0.0)	100 (0.0)	98.7 (2.7)
K-NN <sup>j</sup>	98.1 (3.9)	68.3 (5.0)	75.2 (4.3)	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	98.8 (3.8)	100 (0.0)	100 (0.0)	97.7 (4.7)
LOF <sup>k</sup>	98.6 (2.9)	75.0 (13.5)	80.2 (10.8)	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	96.9 (5.0)	99.7 (0.8)	100 (0.0)	97.4 (7.9)
<b>Reconstruction-based method</b>												
PCA <sup>l</sup>	98.9 (2.2)	85.0 (5.0)	85.5 (2.1)	99.2 (1.3)	85.0 (5.0)	91.4 (2.7)	98.6 (1.9)	88.9 (0.0)	92.2 (6.0)	97.8 (2.2)	83.3 (0.0)	89.1 (9.7)
Auto-en- coder	97.4 (6.0)	89.1 (13.0)	86.0 (14.2)	98.5 (3.2)	94.5 (9.6)	91.8 (9.4)	99.2 (2.4)	93.7 (10.2)	93.7 (8.3)	98.6 (3.8)	94.4 (9.5)	93.7 (9.7)
SOM <sup>m</sup>	99.3 (1.9)	99.9 (1.2)	84.7 (19.8)	99.8 (0.7)	100 (0.0)	91.4 (9.6)	99.9 (0.3)	100 (0.0)	95.2 (7.9)	99.6 (1.3)	100 (0.0)	93.4 (12.1)
K-means	99.2 (2.5)	85.0 (11.7)	87.0 (10.4)	100 (0.0)	100 (0.0)	97.1 (6.3)	100 (0.0)	100 (0.0)	98.8 (3.8)	100 (0.0)	100 (0.0)	99.2 (2.5)

<sup>a</sup>AUC: area under the receiver operating characteristic curve.

<sup>b</sup>SVDD: support vector data description.

<sup>c</sup>IncSVDD: incremental support vector data description.

<sup>d</sup>V-SVM: one-class support vector machine.

<sup>e</sup>Italicized values indicates the top performing models.

<sup>f</sup>NN: nearest neighbor.

<sup>g</sup>MST: minimum spanning tree.

<sup>h</sup>MOG: mixture of Gaussian.

<sup>i</sup>MCD: minimum covariance determinant.

<sup>j</sup>K-NN: K-nearest neighbor.

<sup>k</sup>LOF: local outlier factor.

<sup>l</sup>PCA: principal component analysis.

<sup>m</sup>SOM: self-organizing maps.

### Second Case of Infection (Flu)

The boundary and domain-based method achieved better performance with a small sample size compared with the density and reconstruction-based methods. However, as the sample size increased, all the three groups achieved comparable performance. The detailed numerical values of comparison are given in [Multimedia Appendix 4](#). Specific models such as V-SVM, K-NN, and K-means performed better from their respective group. Regarding the raw data, all the models failed to generalize from the 1-month data set as compared with the higher sample objects, that is, 3 months ([Multimedia Appendix 4](#)):

1. From the boundary and domain-based method, SVDD, MST, and incremental support vector data description (incSVDD) performed better with a larger sample object, and V-SVM achieved better description with 30 sample objects.
2. From the density-based method, all the models exhibited similar performance. Naïve Parzen and K-NN, with only 60 sample objects, achieved comparable performance with the higher sample objects.
3. From the reconstruction-based method, K-means achieved better performance for all sample sizes.

Smoothing the data significantly improved the performance of the model even with 30 objects, compared with the raw data ([Multimedia Appendix 4](#)):

1. From the boundary and domain-based method, the V-SVM achieved higher performance in all the sample sizes.
2. From the density-based method, LOF achieved better description with small sample objects, and K-NN produced better description with all the sample sizes. Gaussian families achieved improved and comparable performance with increased sample objects. Among them, K-NN with only 60 objects achieved comparable performance with larger sample objects.
3. Regarding the reconstruction-based method, K-means and SOM achieved better performance, whereas K-means performed better in all the sample sizes.

### Third Case of Infection (Flu)

The boundary and domain-based method achieved better performance with a small sample size compared with the density and reconstruction-based methods. However, as the sample size increased, all the three groups produced comparable descriptions. The detailed numerical values of comparison are given in [Multimedia Appendix 4](#). Specific models such as V-SVM, MST, LOF, and PCA performed better from their respective group. Regarding the raw data, surprisingly, in contrast to the previous two infection cases, all the models achieved higher generalization from the 1-month data set ([Multimedia Appendix 4](#)):

1. From the boundary and domain-based method, SVDD, V-SVM, MST, and incSVDD performed better in all the cases, with MST achieving better performance.

2. From the density-based method, normal and MCD Gaussian achieved better description of the data with 1-month sample objects. K-NN and LOF performed better with sample sizes larger than 1-month sample objects, and LOF outperformed all sample sizes. The LOF with only 60 objects achieved comparable performance with the higher sample objects.
3. From the reconstruction-based method, PCA produced better description for all sample sizes, whereas K-means and SOM achieved comparable performance with sample size larger than 1-month sample objects.

Smoothing the data allowed the models to generalize well and significantly improved the performance of the model even with 30 objects, compared with the raw data ([Multimedia Appendix 4](#)):

1. From the boundary and domain-based method, the V-SVM and MST achieved higher performance in all the sample sizes, whereas V-SVM outperformed all the models.
2. From the density-based method, the Gaussian families, LOF, and K-NN achieved better performance, whereas LOF achieved better performance in all sample sizes.
3. Regarding the reconstruction-based method, K-means and PCA achieved better performance, whereas PCA performed better in all the sample sizes.

### Fourth Case of Infection (Flu)

The boundary and domain-based method achieved better performance with small sample sizes compared with the density and reconstruction-based methods. All the three groups improved with increasing sample size. The detailed numerical values of comparison are given in [Multimedia Appendix 4](#). Specific models such as V-SVM, LOF, and K-means performed better from their respective group. Regarding the raw data, surprisingly, in contrast to all the previous three infection cases, all the models achieved higher generalization from the 1-month data set ([Multimedia Appendix 4](#)):

1. From the boundary and domain-based method, SVDD, V-SVM, and incSVDD performed better for all the sample sizes.
2. From the density-based method, MCD Gaussian performed better with a 1-month sample size, and all the models produced comparable descriptions as the sample size increased, whereas the LOF performed better for all the sample sizes.
3. From the reconstruction-based method, PCA performed relatively better for all the sample sizes, and K-means and SOM achieved comparable performance with a larger sample size.

Smoothing the data significantly improved the model performance even with 30 objects compared with the raw data ([Multimedia Appendix 4](#)):

1. From the boundary and domain-based method, the V-SVM achieved higher performance in all the sample sizes. As the sample size increased, the incSVDD and MST achieved comparable performance.

2. From the density-based method, K-NN and LOF produced better descriptions with a 1-month sample size. K-NN performed better in almost all sample sizes.
3. From the reconstruction-based method, K-means achieved better performance for all sample sizes.

### Hourly

As can be seen in [Table 4](#) (see also [Multimedia Appendix 4](#)), the performance of the model generally improved as more training sample data were presented. The models produced comparable performance even with the 1-month data set compared with the daily scenario. This is mainly because of the presence of more samples per day (24 samples per day), which enables the models to reach a better generalization. Generally, the results indicate that the models generalize well after 2 months. Both the boundary and domain-based method and reconstruction-based method achieved better performance even with a 1-month sample size. However, the density-based method suffers from large variation with 1-month training samples. In general, the boundary and domain-based method performed better in all the infection cases compared with the other two

methods. In addition, specific models such as V-SVM, K-NN, and K-means performed well from their respective groups.

### *First Case of Infection (Flu)*

The boundary and domain-based method achieved better performance compared with the density and reconstruction-based methods. As can be seen in [Table 4](#), the boundary and domain-based method achieved better generalization from the 1-month data set. Specific models such as V-SVM, K-NN, and K-means performed better from their respective group:

1. From the boundary and domain-based method, V-SVM achieved better description in all sample sizes, whereas SVDD, incSVDD, and V-SVM achieved comparable performance with a larger sample size.
2. From the density-based method, Gaussian families and naïve Parzen performed better at large sample sizes, whereas K-NN and LOF achieved better performance in all the sample sizes. K-NN outperformed all the models.
3. From the reconstruction-based method, K-means performed better in all the sample sizes, and all the other models performed better with larger sample sizes.



**Table 4.** Average (SD) of area under the receiver operating characteristic curve, specificity, F1-score for the smoothed version of the data with a 48-hour moving average filter and different sample size. Fraction=0.01.

Models	1 month			2 months			3 months			4 months		
	AUC <sup>a</sup> , mean (SD)	Specificity	F1	AUC <sup>a</sup> , mean (SD)	Specificity	F1	AUC <sup>a</sup> , mean (SD)	Specificity	F1	AUC <sup>a</sup> , mean (SD)	Specificity	F1
<b>Boundary and domain-based method</b>												
SVDD <sup>b</sup>	97.6 (1.9)	83.2 (3.4)	85.8 (1.7)	97.8 (1.2)	85.7 (5.0)	90.5 (9.6)	97.7 (1.2)	90.4 (5.1)	94.2 (2.9)	98.1 (0.9)	91.0 (3.7)	96.8 (0.9)
IncSVDD <sup>c</sup>	97.4 (1.9)	84.5 (2.8)	86.8 (1.9)	97.7 (1.2)	86.7 (2.0)	93.9 (1.0)	97.5 (1.2)	88.5 (1.5)	96.0 (1.1)	97.9 (0.9)	88.9 (1.2)	97.1 (0.7)
V-SVM <sup>d</sup>	98.1 (2.1)	84.5 (1.1)	90.5 (1.1) <sup>e</sup>	99.0 (1.1)	92.6 (0.0)	96.1 (1.3)	99.5 (0.6)	93.8 (0.5)	96.9 (1.4)	99.4 (0.4)	94.2 (0.0)	97.1 (1.3)
NN <sup>f</sup>	84.8 (6.0)	75.9 (4.5)	74.8 (6.0)	89.3 (2.2)	76.5 (4.1)	87.1 (3.3)	89.0 (4.0)	77.5 (3.9)	89.3 (4.4)	90.2 (4.7)	77.5 (3.8)	91.4 (6.4)
MST <sup>g</sup>	90.5 (3.1)	85.4 (3.9)	67.6 (14.5)	94.4 (2.0)	85.7 (4.0)	85.1 (7.0)	94.7 (2.4)	88.8 (3.5)	87.8 (8.5)	95.8 (2.2)	88.8 (3.0)	90.9 (5.9)
<b>Density-based method</b>												
Gaussian	98.1 (2.2)	79.8 (4.9)	83.9 (2.7)	99.5 (0.9)	90.1 (1.7)	95.2 (1.8)	99.6 (0.7)	92.9 (1.3)	97.1 (2.5)	99.5 (0.5)	92.2 (1.0)	97.7 (1.1)
MOG <sup>h</sup>	95.8 (3.6)	82.7 (4.3)	83.7 (5.0)	98.3 (1.5)	86.2 (2.7)	92.3 (2.7)	98.7 (1.4)	88.7 (4.6)	94.7 (3.5)	98.6 (1.6)	88.2 (3.1)	95.3 (3.2)
MCD <sup>i</sup> Gaussian	98.6 (2.1)	75.3 (6.9)	81.3 (2.5)	99.6 (0.9)	89.6 (1.9)	95.0 (1.8)	99.6 (0.7)	92.5 (1.8)	97.0 (2.3)	99.6 (0.4)	92.0 (1.2)	97.7 (1.1)
Parzen	91.9 (2.9)	93.6 (2.0)	63.4 (16.5)	96.2 (2.3)	94.4 (2.0)	81.6 (10.2)	96.6 (2.6)	94.8 (1.7)	84.2 (9.5)	97.4 (2.2)	95.6 (1.2)	87.9 (7.1)
Naïve Parzen	94.8 (3.7)	76.4 (5.6)	77.6 (7.9)	98.7 (1.2)	85.2 (3.3)	91.8 (2.9)	99.1 (1.1)	89.1 (3.8)	94.8 (2.5)	98.9 (0.9)	89.7 (2.4)	96.2 (1.6)
K-NN <sup>j</sup>	97.1 (3.4)	78.8 (2.0)	84.2 (2.1)	99.1 (1.0)	92.9 (0.7)	96.0 (1.8)	99.6 (0.4)	93.8 (0.7)	97.3 (1.9)	99.5 (0.3)	94.0 (0.6)	98.2 (0.9)
LOF <sup>k</sup>	96.9 (3.5)	78.3 (3.0)	84.2 (2.4)	99.2 (1.1)	91.9 (0.9)	96.0 (1.8)	99.6 (0.5)	93.7 (0.8)	97.3 (2.1)	99.5 (0.4)	93.1 (0.4)	97.8 (1.2)
<b>Reconstruction-based method</b>												
PCA <sup>l</sup>	97.1 (3.4)	63.9 (8.8)	75.4 (0.3)	99.4 (1.2)	76.4 (6.6)	90.2 (1.1)	99.1 (1.3)	75.1 (6.8)	92.4 (1.1)	98.9 (1.2)	69.1 (4.1)	93.1 (0.8)
Auto-encoder	92.0 (4.8)	79.5 (7.6)	78.9 (8.3)	96.2 (2.6)	83.1 (7.2)	91.1 (3.9)	96.3 (3.2)	84.3 (7.7)	92.7 (5.0)	96.7 (3.0)	84.0 (8.0)	94.6 (4.4)
SOM <sup>m</sup>	94.1 (2.3)	82.2 (3.3)	82.6 (4.9)	95.6 (1.1)	82.9 (3.1)	91.6 (1.9)	94.8 (2.3)	83.4 (5.8)	92.3 (4.1)	95.5 (1.9)	84.1 (3.8)	94.3 (3.8)
K-means	97.3 (3.2)	80.9 (2.5)	85.5 (2.5)	98.9 (1.1)	92.6 (0.7)	95.8 (1.8)	99.3 (0.6)	92.9 (0.7)	97.3 (1.4)	99.4 (0.4)	94.1 (0.2)	98.1 (1.1)

<sup>a</sup>AUC: area under the receiver operating characteristic curve.

<sup>b</sup>SVDD: support vector data description.

<sup>c</sup>IncSVDD: incremental support vector data description.

<sup>d</sup>V-SVM: one-class support vector machine.

<sup>e</sup>Italicized values indicates the top performing models.

<sup>f</sup>NN: nearest neighbor.

<sup>g</sup>MST: minimum spanning tree.

<sup>h</sup>MOG: mixture of Gaussian.

<sup>i</sup>MCD: minimum covariance determinant.

<sup>j</sup>K-NN: K-nearest neighbor.

<sup>k</sup>LOF: local outlier factor.

<sup>l</sup>PCA: principal component analysis.

<sup>m</sup>SOM: self-organizing maps.

### **Second Case of Infection (Flu)**

The boundary and domain-based method and reconstruction-based method achieved better performance for all sample sizes compared with the density-based method. Specifically, the boundary and domain-based method achieved better generalization from the 1-month data set. The detailed numerical values of comparison are given in [Multimedia Appendix 4](#). Specific models such as V-SVM, K-NN, and K-means performed better from their respective group:

1. From the boundary and domain-based method, V-SVM achieved better description for all the sample sizes, and SVDD, NN, and incSVDD improved with larger training sample size; however, V-SVM outperformed all the models for all the sample sizes.
2. From the density-based method, normal and MCD Gaussian performed better with the 1- and 2-month sample sizes, and models such as K-NN performed better on all the sample sizes, whereas naïve Parzen outperformed all the models with the 3- and 4-month data sets.
3. From the reconstruction-based method, K-means produced better description for all the sample sizes and the auto-encoder and SOM performed better with larger sample sizes.

### **Third Case of Infection (Flu)**

Generally, in comparison, all the groups performed better at large training sample sizes; however, the boundary and domain-based method achieved better performance with small training sample sizes. It achieved comparable generalization from the 1-month data set. The detailed numerical values of comparison are given in [Multimedia Appendix 4](#). Specific models such as V-SVM, families that utilize nearest neighbor distance (K-NN and LOF), and PCA performed better from their respective group:

1. From the boundary and domain-based method, SVDD, NN, MST, incSVDD, and V-SVM achieved better performance at larger training sample sizes, whereas V-SVM outperformed all the models for all the sample sizes.
2. From the density-based method, the Gaussian families, K-NN, LOF, and naïve Parzen achieved better performance at larger training sample sizes, whereas K-NN and LOF outperformed all the models for all the sample sizes.
3. From the reconstruction-based method, K-means, PCA, auto-encoder, and SOM achieved better performance at larger training sample sizes, whereas PCA performed better for all sample sizes.

### **Fourth Case of Infection (Flu)**

Generally, in comparison, all the group performed better at large training sample size; however, the boundary and domain-based method achieved better performance with small training sample sizes, for example, 1-month data set. It achieved comparable generalization from the 1-month data set. The detailed numerical values of comparison are given in [Multimedia Appendix 4](#). Specific models such as V-SVM, Gaussian families (Gaussian, MOG, and MCD Gaussian), and PCA performed better from their respective groups:

1. From the boundary and domain-based method, NN, incSVDD, and V-SVM achieved better performance at larger training sample sizes, whereas V-SVM outperformed all the models for all the sample sizes.
2. From the density-based method, Gaussian families, K-NN, LOF, and naïve Parzen achieved better performance at larger training sample sizes, whereas Gaussian families outperformed all the models for all the sample sizes.
3. From the reconstruction-based method, K-means, SOM, auto-encoder, and PCA achieved better performance at larger training sample sizes, whereas PCA performed better for all sample sizes.

### **Average Performance Across all the Infection Cases**

The average performances of the models across all the infection cases for different sample sizes, levels of data granularity (hourly and daily), and nature of data (raw and smoothed) are shown in [Tables 5-7](#). In general, the boundary and domain-based method performed better than the other two groups in both daily and hourly smoothed data sets; however, all the groups achieved comparable performance with respect to the daily raw data set. Specific models such as V-SVM, K-NN, and K-means performed better in all these circumstances.

### **Daily Raw Data Set**

Regarding the daily raw data set, as shown in [Table 5](#), specific models such as V-SVM, MCD Gaussian, K-NN, and K-means produced relatively better descriptions of the 1-month data. For the 2-month sample size, models such as incSVDD, K-NN, LOF, and K-means achieved better performance. For the 3-month sample size, SVDD, incSVDD, V-SVM, Gaussian, MCD Gaussian, K-NN, LOF, and K-means produced comparable descriptions. As expected, SVDD and most of the density-based method improved with larger training sizes. For the 4-month sample size, almost all the models produced much improved performance. In the group comparison, all three groups produced comparable descriptions in all the sample sizes.

**Table 5.** Average performance of each model across all the infection cases for the daily raw data set (without smoothing) and different sample sizes. Fraction=0.01.

Models	1 month			2 months			3 months			4 months		
	AUC <sup>a</sup> , mean (SD)	Specificity	F1	AUC <sup>a</sup> , mean (SD)	Specificity	F1	AUC <sup>a</sup> , mean (SD)	Specificity	F1	AUC <sup>a</sup> , mean (SD)	Specificity	F1
<b>Boundary and domain-based method</b>												
SVDD <sup>b</sup>	87.1 (11)	66.0 (13.5)	74.8 (9.5)	91.7 (7.3)	61.7 (10.6)	84.1 (5.5)	93.3 (4.6)	67.3 (10.5)	86.2 (4.4)	91.4 (4.3)	61.7 (10.6)	85.7 (4.1) <sup>c</sup>
IncSVDD <sup>d</sup>	85.2 (11)	63.0 (4.6)	74.7 (10.4)	90.5 (8.5)	57.9 (11)	83.8 (3.6)	92.8 (5.1)	62.8 (10.9)	84.9 (3.2)	90.8 (4.4)	55.0 (11.7)	83.5 (3.7)
V-SVM <sup>e</sup>	91.5 (8.0)	55.7 (7.0)	77.4 (6.4)	92.2 (5.1)	60.6 (5.0)	82.8 (4.5)	94.2 (3.8)	66.9 (6.1)	86.6 (3.5)	93.8 (4.1)	63.1 (11.9)	84.5 (5.1)
NN <sup>f</sup>	73.4 (12)	31.3 (6.5)	65.0 (5.4)	72.1 (11.9)	25.0 (9.6)	75.7 (3.7)	70.8 (11.2)	8.6 (17.6)	72.0 (4.7)	70.0 (9.0)	16.0 (14.4)	75.7 (3.4)
MST <sup>g</sup>	82.4 (8.7)	52.1 (0.0)	71.2 (6.1)	82.6 (9.1)	50.4 (9.0)	82.0 (5.1)	84.0 (6.3)	56.2 (9.3)	82.9 (3.5)	84.2 (6.6)	50.0 (11.4)	82.6 (2.7)
<b>Density-based method</b>												
Gaussian	91.5 (9.9)	56.9 (7.7)	72.9 (7.8)	93.6 (6.1)	58.8 (10.9)	84.0 (4.0)	95.1 (4.3)	65.3 (10.6)	86.3 (3.2)	95.0 (3.5)	57.9 (10.3)	84.6 (3.2)
MOG <sup>h</sup>	89.9 (12)	69.2 (11.9)	71.3 (14.3)	91.7 (6.1)	64.1 (14.0)	83.8 (6.8)	94.0 (4.4)	67.0 (11.4)	85.0 (5.6)	94.5 (3.7)	61.6 (12.6)	84.9 (5.1)
MCD <sup>i</sup> Gaussian	90.8 (9.1)	54.0 (5.5)	72.0 (6.8)	93.1 (6.0)	58.0 (8.1)	84.1 (4.3)	95.3 (4.2)	65.3 (10.6)	86.4 (3.0)	94.8 (3.5)	57.9 (10.6)	84.9 (3.0)
Parzen	89.7 (10)	59.6 (8.3)	70.6 (9.4)	91.7 (6.5)	62.1 (10.3)	83.9 (5.3)	93.9 (5.0)	68.7 (11.2)	85.6 (5.4)	94.3 (3.8)	66.1 (12.7)	86.1 (3.8)
Naïve Parzen	88.1 (8.7)	54.2 (6.5)	69.1 (9.6)	90.2 (7.1)	60.4 (11.2)	83.7 (4.9)	91.9 (5.5)	66.5 (12.8)	86.6 (4.4)	92.8 (4.7)	64.6 (10.0)	86.9 (3.4)
K-NN <sup>j</sup>	91.1 (7.8)	52.9 (5.1)	71.6 (7.9)	91.6 (5.0)	61.1 (11.3)	85.9 (3.1)	94.8 (4.8)	66.9 (11.2)	87.1 (3.2)	95.0 (3.8)	62.1 (10.3)	86.5 (3.3) )
LOF <sup>k</sup>	89.2 (8.9)	56.3 (3.9)	73.0 (8.6)	92.4 (6.0)	59.2 (11.1)	84.9 (2.8)	94.0 (4.8)	64.4 (11.4)	86.2 (2.8)	93.7 (4.3)	53.8 (10.3)	83.8 (2.5)
<b>Reconstruction-based method</b>												
PCA <sup>l</sup>	87.6 (8.8)	58.8 (4.6)	73.7 (8.3)	90.2 (6.4)	55.0 (6.8)	82.7 (4.5)	91.4 (4.9)	59.7 (6.2)	84.1 (3.2)	90.5 (4.5)	53.8 (7.2)	83.6 (2.9)
Auto-en- coder	83.6 (14)	58.3 (17.7)	71.0 (12.5)	84.6 (12.5)	53.1 (20.0)	82.1 (7.0)	88.4 (10.0)	57.7 (21.5)	83.3 (6.8)	88.5 (10.6)	52.3 (21.0)	83.2 (5.8)
SOM <sup>m</sup>	85.6 (12)	63.4 (10.3)	72.7 (11.7)	87.6 (7.2)	57.1 (10.2)	81.6 (5.8)	93.5 (5.4)	64.4 (8.5)	84.8 (4.0)	94.7 (4.0)	59.0 (5.8)	85.0 (3.1)
K-means	94.2 (7.6)	57.2 (7.6)	73.1 (7.1)	93.7 (6.2)	62.2 (10.5)	85.4 (4.2)	96.0 (4.4)	67.6 (10.3)	87.4 (3.1)	95.8 (3.9)	62.1 (10.3)	86.5 (2.9)

<sup>a</sup>AUC: area under the receiver operating characteristic curve.

<sup>b</sup>SVDD: support vector data description.

<sup>c</sup>Italicized values indicates the top performing models.

<sup>d</sup>IncSVDD: incremental support vector data description.

<sup>e</sup>V-SVM: one-class support vector machine.

<sup>f</sup>NN: nearest neighbor.

<sup>g</sup>MST: minimum spanning tree.

<sup>h</sup>MOG: mixture of Gaussian.

<sup>i</sup>MCD: minimum covariance determinant.

<sup>j</sup>K-NN: K-nearest neighbor.

<sup>k</sup>LOF: local outlier factor.

<sup>l</sup>PCA: principal component analysis.

<sup>m</sup>SOM: self-organizing maps.

### ***Daily Smoothed Data Set***

Regarding the daily smoothed data set, as shown in [Table 6](#), almost all models achieved excellent performance and much improved data description compared with the daily raw data set. As shown in [Table 6](#), specific models such as V-SVM,

K-NN, and K-means produced excellent descriptions of the data for all the sample sizes; however, V-SVM achieved superior performance compared with these models. In the group comparison, the boundary and domain-based method produced excellent description of the data for all sample sizes.

**Table 6.** Average performance of each model across all the infection cases for the daily smoothed data set (with filter) and different sample size. Fraction=0.01.

Models	1 month			2 months			3 months			4 months		
	AUC <sup>a</sup> , mean (SD)	Specificity	F1	AUC <sup>a</sup> , mean (SD)	Specificity	F1	AUC <sup>a</sup> , mean (SD)	Specificity	F1	AUC <sup>a</sup> , mean (SD)	Specificity	F1
<b>Boundary and domain-based method</b>												
SVDD <sup>b</sup>	99.9 (0.7)	100 (0.0)	94.1 (14.2)	100 (0.0)	100 (0.0)	96.1 (7.6)	100 (0.0)	100 (0.0)	96.5 (6.5)	100 (0.0)	100 (0.0)	97.9 (3.9)
IncSVDD <sup>c</sup>	99.9 (0.7)	100 (0.0)	94.1 (14.2)	100 (0.0)	100 (0.0)	96.9 (6.5)	100 (0.0)	100 (0.0)	97.3 (5.9)	100 (0.0)	100 (0.0)	98.6 (2.9)
V-SVM <sup>d</sup>	100 (0.0)	100 (0.0)	99.1 (3.2) <sup>e</sup>	100 (0.0)	100 (0.0)	99.1 (2.9)	100 (0.0)	100 (0.0)	99.4 (1.9)	100 (0.0)	100 (0.0)	99.5 (1.5)
NN <sup>f</sup>	90.1 (14.5)	40.0 (30.5)	69.5 (13.2)	88.9 (9.9)	33.1 (22.6)	78.4 (6.8)	89.2 (7.9)	33.6 (14.6)	77.7 (5.3)	90.5 (6.8)	23.5 (18.6)	77.1 (5.7)
MST <sup>g</sup>	98.9 (3.6)	85 (6.1)	86.7 (9.4)	99.8 (0.7)	96.7 (3.4)	95.1 (6.2)	99.9 (0.2)	98.9 (4.1)	98.0 (3.5)	99.9 (0.5)	100 (0.0)	98.0 (5.4)
<b>Density-based method</b>												
Gaussian	99.2 (5.1)	92.6 (9.0)	87.2 (15.2)	99.5 (2.5)	96.7 (7.5)	94.8 (10.4)	99.9 (0.4)	100 (0.0)	98.1 (4.9)	99.8 (0.8)	100 (0.0)	98.3 (5.9)
MOG <sup>h</sup>	98.8 (5.4)	92.9 (8.6)	85.2 (17.1)	99.4 (2.6)	97.0 (5.4)	92.1 (11.6)	99.9 (0.4)	99.9 (0.7)	95.4 (7.8)	99.8 (1.0)	99.9 (0.6)	96.4 (7.7)
MCD <sup>i</sup> Gaussian	98.4 (5.6)	86.6 (8.8)	86.6 (11.9)	99.3 (2.7)	90.0 (8.7)	93.4 (8.1)	99.8 (0.5)	99.2 (2.6)	98.0 (5.3)	99.8 (0.9)	97.1 (3.9)	97.0 (5.5)
Parzen	99.2 (3.5)	100 (0.0)	90.8 (16.4)	99.9 (0.4)	100 (0.0)	93.7 (9.8)	100 (0.0)	100 (0.0)	93.6 (8.9)	99.9 (0.3)	100 (0.0)	95.8 (8.2)
Naïve Parzen	99.8 (1.2)	100 (0.0)	94.4 (14.6)	100 (0.0)	100 (0.0)	96.1 (7.9)	99.9 (0.5)	100 (0.0)	97.4 (5.6)	100 (0.0)	100 (0.0)	98.2 (4.2)
K-NN <sup>j</sup>	99.5 (2.0)	91.6 (3.6)	90.7 (9.6)	99.9 (0.4)	100 (0.0)	98.3 (4.9)	100 (0.0)	100 (0.0)	98.4 (5.1)	100 (0.0)	100 (0.0)	98.8 (3.6)
LOF <sup>k</sup>	99.6 (1.5)	93.3 (7.3)	92.4 (10.6)	99.9 (0.5)	99.2 (3.4)	97.1 (7.3)	99.9 (0.2)	98.6 (2.8)	97.4 (4.5)	99.9 (0.4)	100 (0.0)	98.2 (5.9)
<b>Reconstruction-based method</b>												
PCA <sup>l</sup>	93.8 (6.7)	82.0 (7.3)	83.8 (10.4)	91.3 (4.3)	77.9 (7.3)	89.3 (8.7)	88.7 (5.9)	76.3 (8.6)	89.5 (5.3)	90.7 (3.6)	76.2 (8.6)	89.0 (6.9)
Auto-encoder	97.0 (8.1)	91.6 (14.6)	87.7 (16.0)	98.1 (5.4)	92.6 (15.3)	92.0 (10.7)	98.6 (4.6)	92.8 (14.8)	94.0 (8.3)	98.7 (4.0)	92.7 (15.8)	94.9 (7.7)
SOM <sup>m</sup>	99.1 (3.2)	99.9 (0.6)	85.2 (20.5)	99.8 (0.7)	100 (0.0)	88.9 (16.1)	99.9 (0.2)	100 (0.0)	94.6 (8.0)	99.8 (0.6)	100 (0.0)	95.9 (8.1)
K-means	99.8 (1.2)	96.2 (6.0)	93.2 (12.7)	100 (0.0)	100 (0.0)	97.8 (5.6)	100 (0.0)	100 (0.0)	98.0 (5.6)	100 (0.0)	100 (0.0)	99.0 (2.9)

<sup>a</sup>AUC: area under the receiver operating characteristic curve.

<sup>b</sup>SVDD: support vector data description.

<sup>c</sup>IncSVDD: incremental support vector data description.

<sup>d</sup>V-SVM: one-class support vector machine.

<sup>e</sup>Italicized values indicates the top performing models.

<sup>f</sup>NN: nearest neighbor.

<sup>g</sup>MST: minimum spanning tree.

<sup>h</sup>MOG: mixture of Gaussian.

<sup>i</sup>MCD: minimum covariance determinant.

<sup>j</sup>K-NN: K-nearest neighbor.

<sup>k</sup>LOF: local outlier factor.

<sup>l</sup>PCA: principal component analysis.

<sup>m</sup>SOM: self-organizing maps.

### ***Hourly Smoothed Data Set***

Regarding the hourly smoothed data set, as shown in [Table 7](#), almost all the models failed to produce acceptable data description from the 1-month sample size except V-SVM, which achieved the best description. The high variability between the performance of the models with the 1-month hourly data set could be associated with the high data granularity, and, in fact, the models require more data sets to capture the high variability

among the data objects. Models such as V-SVM, MCD Gaussian, and K-means achieved superior performance from their respective groups. In general, V-SVM outperformed in all the sample sizes. The density and reconstruction-based models improved with larger sample size. In the group comparison, the boundary and domain-based method produced better description in all the sample sizes, and the density and reconstruction-based method achieved equivalent performance with larger sample sizes.

**Table 7.** Average performance of each model across all the infection cases for the hourly data set with smoothing and different sample size. Fraction=0.01.

Models	1 month			2 months			3 months			4 months		
	AUC <sup>a</sup> , mean (SD)	Specificity	F1	AUC <sup>a</sup> , mean (SD)	Specificity	F1	AUC <sup>a</sup> , mean (SD)	Specificity	F1	AUC <sup>a</sup> , mean (SD)	Specificity	F1
<b>Boundary and domain-based method</b>												
SVDD <sup>b</sup>	97.4 (2.9)	89.0 (3.4)	89.4 (7.1)	97.4 (1.8)	86.7 (4.4)	91.5 (10.9)	97.2 (2.6)	80.1 (5.5)	93.5 (3.4)	97.6 (1.7)	81.8 (5.3)	94.6 (6.0)
IncSVDD <sup>c</sup>	97.1 (2.9)	87.7 (2.7)	89.5 (5.9)	97.2 (1.8)	86.4 (2.8)	93.6 (4.8)	97.0 (2.7)	76.2 (6.3)	93.2 (2.6)	97.4 (1.7)	79.0 (4.8)	95.4 (1.9) <sup>d</sup>
V-SVM <sup>e</sup>	98.1 (2.0)	85.5 (0.6)	92.3 (1.3)	98.9 (1.4)	89.8 (0.2)	95.4 (1.6)	98.7 (1.4)	86.4 (0.4)	94.4 (2.0)	99.0 (0.9)	89.2 (0.3)	95.4 (2.1)
NN <sup>f</sup>	93.2 (7.8)	92.0 (2.4)	83.9 (12.0)	94.4 (2.5)	88.4 (3.4)	90.9 (5.3)	93.3 (2.8)	83.0 (3.7)	92.0 (4.2)	94.0 (2.8)	82.9 (3.6)	94.0 (4.0)
MST <sup>g</sup>	96.1 (2.6)	94.4 (2.2)	72.9 (18.5)	97.3 (1.4)	94.2 (2.1)	86.1 (11.0)	96.1 (2.1)	93.5 (1.9)	90.2 (7.3)	97.0 (1.4)	93.6 (1.7)	92.6 (5.0)
<b>Density-based method</b>												
Gaussian	98.4 (1.6)	91.2 (2.6)	89.6 (12.5)	99.3 (0.9)	92.3 (1.7)	95.7 (4.9)	98.8 (1.3)	88.1 (4.0)	95.9 (2.7)	99.2 (0.7)	89.8 (3.1)	97.2 (1.8)
MOG <sup>h</sup>	97.5 (3.0)	91.7 (3.2)	87.8 (13.3)	98.9 (1.2)	90.9 (2.7)	94.0 (6.3)	98.2 (2.0)	85.4 (6.6)	94.2 (4.1)	98.5 (1.5)	88.0 (4.9)	96.0 (3.1)
MCD <sup>i</sup> Gaussian	98.5 (1.5)	89.9 (3.7)	89.1 (11.8)	99.5 (0.9)	92.2 (92.2)	95.8 (4.5)	98.9 (1.1)	87.9 (3.3)	96.0 (2.5)	99.2 (0.7)	90.4 (3.4)	97.4 (1.7)
Parzen	96.4 (2.6)	97.8 (1.1)	59.9 (18.9)	98.0 (1.6)	97.7 (1.1)	79.5 (14.5)	97.2 (2.3)	96.4 (1.2)	85.1 (10)	98.1 (1.6)	96.7 (1.1)	88.6 (7.1)
Naïve Parzen	96.4 (3.0)	87.5 (3.5)	85.1 (10.9)	98.7 (1.5)	89.2 (2.8)	92.8 (7.5)	96.0 (2.3)	90.8 (2.6)	95.0 (4.1)	98.2 (1.6)	90.0 (1.8)	96.2 (2.8)
K-NN <sup>j</sup>	97.6 (2.9)	91.1 (1.6)	87.6 (13.6)	99.0 (1.4)	92.4 (2.4)	94.5 (6.6)	98.4 (1.4)	92.6 (1.4)	95.7 (4.8)	98.7 (1.1)	93.3 (1.3)	97.3 (2.8)
LOF <sup>k</sup>	96.9 (2.9)	91.2 (1.6)	86.2 (13.0)	97.4 (1.8)	89.8 (4.8)	93.1 (4.9)	95.0 (3.0)	85.2 (4.6)	92.9 (4.8)	95.8 (1.7)	85.3 (4.7)	94.7 (3.2)
<b>Reconstruction-based method</b>												
PCA <sup>l</sup>	97.4 (3.2)	78.2 (6.1)	82.5 (10.9)	94.8 (3.8)	77.6 (4.5)	90.9 (3.6)	92.6 (4.2)	72.4 (3.8)	92.5 (1.9)	93.4 (3.2)	71.1 (2.5)	93.9 (1.1)
Auto-en- coder	95.4 (5.3)	88.7 (9.5)	86.1 (13.1)	96.9 (3.2)	87.1 (9.9)	92.8 (6.4)	95.0 (5.3)	79.3 (14.5)	93.1 (4.8)	95.9 (4.3)	80.3 (14.4)	95.0 (3.6)
SOM <sup>m</sup>	95.9 (2.9)	91.6 (2.6)	86.1 (14.4)	95.7 (1.7)	87.6 (4.1)	92.7 (5.7)	93.9 (3.5)	79.1 (10.9)	92.3 (4.5)	96.0 (2.5)	87.5 (7.0)	96.1 (3.2)
K-means	97.1 (3.9)	89.7 (6.7)	88.7 (12.1)	98.6 (1.7)	91.1 (4.2)	95.2 (4.4)	98.5 (1.5)	92.3 (2.9)	96.9 (3.3)	98.9 (1.0)	93.9 (1.3)	97.9 (2.2)

<sup>a</sup>AUC: area under the receiver operating characteristic curve.

<sup>b</sup>SVDD: support vector data description.

<sup>c</sup>IncSVDD: incremental support vector data description.

<sup>d</sup>Italicized values indicates the top performing models.

<sup>e</sup>V-SVM: one-class support vector machine.

<sup>f</sup>NN: nearest neighbor.

<sup>g</sup>MST: minimum spanning tree.

<sup>h</sup>MOG: mixture of Gaussian.

<sup>i</sup>MCD: minimum covariance determinant.

<sup>j</sup>K-NN: K-nearest neighbor.

<sup>k</sup>LOF: local outlier factor.

<sup>l</sup>PCA: principal component analysis.

<sup>m</sup>SOM: self-organizing maps.

### Unsupervised Methods

Two density-based unsupervised models were tested and evaluated on the same set of data as used in the one-class classifiers: LOF and COF. The average AUC, specificity, and F1-score were computed after 20 runs. The best performing thresholds for all the infection cases along with the optimal value of  $k$  (number of neighbors) are given in [Table 8](#). As can be seen from the table, both the LOF and the COF achieved better performance on the smoothed data set as compared with its raw version. In all the infection cases, LOF performed better than COF. This is mainly because of the characteristics of the data sets, which fulfill the LOF spherical assumption of neighbor distribution. Considering the average F1-score across all the infection cases, LOF achieved 74.7% on the raw daily data, 91.1% on the smoothed daily data, and 72.7% on the hourly

data, whereas COF achieved 71.9% on the raw daily data, 85.8% on the smoothed daily data, and 68.9% on the hourly data. However, compared with the one-class classifier, it suffers from performance degradation mainly because the data are not distributed uniformly, where some regions may contain high density and others might be sparse. However, the region of sparse density does not always signify anomalies (infection incidence). For example, an individual patient on certain days might prefer to take little insulin compared with most of the days and perform heavy physical activity to replace their insulin needs. This scenario could generate an outlier, a small ratio of insulin-to-carbohydrate, which will be considered and detected as outliers by unsupervised models. A detailed score plot of each model for the different infection cases can be found in [Multimedia Appendix 3](#).



**Table 8.** Average area under the receiver operating characteristic curve, specificity, and F1-score for both with and without smoothed versions of the data. The parameters  $k_d$  and  $k_h$  represent the optimal number of nearest neighbors for the daily and hourly cases, respectively.

Frequencies, density-based methods													
Pre-pro	Models (threshold)	1st case of infection ( $k_d=30, k_h=240$ )			2nd case of infection ( $k_d=30, k_h=240$ )			3rd case of infection ( $k_d=30, k_h=240$ )			4th case of infection ( $k_d=30, k_h=240$ )		
		AUC <sup>a</sup>	Specific	F1	AUC <sup>a</sup>	Specific	F1	AUC <sup>a</sup>	Specific	F1	AUC <sup>a</sup>	Specific	F1
<b>Daily</b>													
Without filter	LOF <sup>b</sup> ( $T_1=2.4, T_2=1.2, T_3=1.45, T_4=1.8$ ) <sup>c</sup>	75.0	50.0	85.6	90.0	100	67.4	92.1	66.7	70.1	98.2	100	75.8
	COF <sup>d</sup> ( $T_1=1.4, T_2=1.3, T_3=1.4, T_4=1.4$ )	82.1	66.7	72.6	97.4	100	75.8	75.2	66.7	67.6	96.7	100	71.8
With filter	LOF <sup>b</sup> ( $T_1=1.7, T_2=1.6, T_3=1.95, T_4=2.2$ )	99.0	100	84.1	99.2	100	85.4	100	100	100	99.9	100	94.7
	COF <sup>d</sup> ( $T_1=1.3, T_2=1.3, T_3=1.8, T_4=1.8$ )	97.6	100	76.6	97.9	100	77.6	99.5	100	88.8	100	100	100
<b>Hourly</b>													
	LOF <sup>b</sup> ( $T_1=1.4, T_2=1.3, T_3=1.35, T_4=1.5$ )	98.0	86.0	74.6	95.5	100	70.2	94.3	91.4	75.0	85.2	72.6	71.1
	COF <sup>d</sup> ( $T_1=1.2, T_2=1.1, T_3=, T_4=1.1$ )	92.4	88.4	74.6	77.0	66.0	62.5	90.3	82.7	74.6	82.6	82.2	63.7

<sup>a</sup>AUC: area under the receiver operating characteristic curve.

<sup>b</sup>LOF: local outlier factor.

<sup>c</sup> $T_k$ : threshold for the kth month.

<sup>d</sup>COF: connectivity-based outlier factor.

### Computational Time

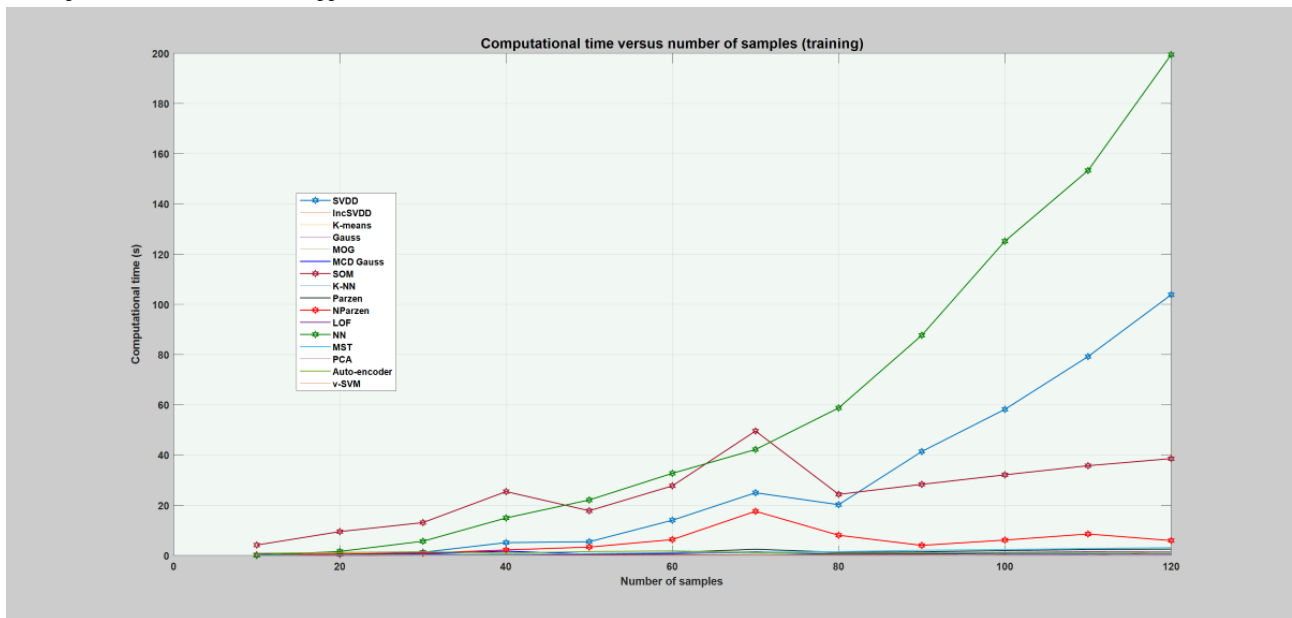
Computational time is the amount of time a particular model needs to learn and execute a given task [12]. It can be regarded as one of the best performance indicators for real-time systems. For a real-time application, an optimal model is the one that achieves superior detection performance with small training and testing time. Depending on the application, sometimes models can be trained offline, which makes the training time less important [12]. In this regard, the computational times of all the models were estimated and compared with each other. The computational time was measured for different sample sizes

of the training and testing data sets. The sample size of the training and testing data includes 240, 480, 720, 960, 1200, 1440, 1680, 1920, 2160, 2400, 2640, and 2880 sample objects (data points) each. The required computational time for both training and testing each model is depicted in Figures 5 and 6. The figures demonstrate a rough estimation of the computational time, where each model learns the data set and classifies the sample objects. During the training phase, NN, SVDD, and SOM took considerable time. For a training sample size of 2880 objects, NN requires 296 times, SVDD requires 206 times, and SOM requires 42 times the time taken by K-NN on the same sample size. Generally, as the number of sample objects

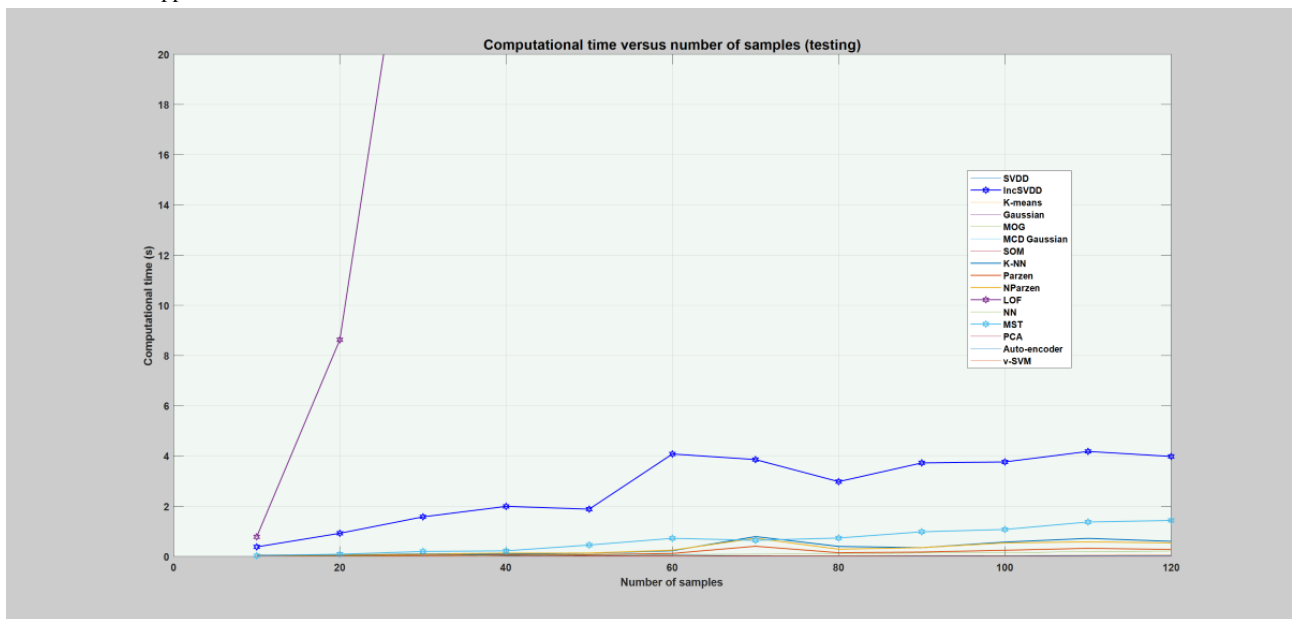
increases, these models require much more time. However, K-means, Gaussian families, LOF, MST, K-NN, V-SVM, PCA, auto-encoder, and incSVDD took less time. These models took almost constant time even when the number of samples

increased. During the testing phase, only the LOF took considerable time compared with the other models, as can be seen in Figure 6.

**Figure 5.** Plot of models’ average computational time for the training phase. The x-axis depicts the sample size, and each label stands for total sample size divided by 24. The y-axis depicts the computational time required by each model. Gauss: Gaussian; IncSVDD: incremental support vector data description; K-NN: K-nearest neighbor; LOF: local outlier factor; MCD: minimum covariance determinant; MOG: mixture of Gaussian; MST: minimum spanning tree; NN: nearest neighbor; NParzen: naïve Parzen; PCA: principal component analysis; SOM: self-organizing maps; SVDD: support vector data description; V-SVM: one-class support vector machine.



**Figure 6.** Plot of models’ average computational time for the testing phase. The x-axis depicts the sample size, and each label stands for total sample size divided by 24. The y-axis depicts the computational time required by each model. Gauss: Gaussian; IncSVDD: incremental support vector data description; K-NN: K-nearest neighbor; LOF: local outlier factor; MCD Gauss: Gaussian; SOM: self-organizing maps; MOG: mixture of Gaussian; MST: minimum spanning tree; NN: nearest neighbor; NParzen: naïve Parzen; PCA: principal component analysis; SVDD: support vector data description; V-SVM: one-class support vector machine.



## Discussion

### Principal Findings

Anomaly or novelty detection problem has been widely used in various applications including machine fault and sensor

failure detection, prevention of credit card or identity fraud, health and medical diagnostics and monitoring, cyber-intrusion detection, and others [1-3]. In applications related to health and medical diagnostics and monitoring, the anomaly detection problem has been used to detect and identify the abnormal health

state of an individual, for example, detecting abnormal patterns of heartbeat recorded using an electrocardiogram [1,51-54]. The omnipresence of various physiological sensors has facilitated circumstances for individuals to easily self-record health-related events and data for the purpose of self-informatics and management [55]. Currently, people are generating huge amounts of data on a daily basis that can contribute to both individual and public health purposes [54]. To this end, people with diabetes are not an exception, generating rich data in both quality and quantity, which is expected to further improve with advances in diabetes technologies. These data can provide valuable information if processed with the right tools and methodology, and in this regard, particular instance includes detecting novel or anomalous data points for various purposes. The availability of labeled data constrains the choice of methods in the anomaly detection problem [3,9-11]. Supervised anomaly detection methods are impractical for applications such as detecting infection incidences in people with type 1 diabetes for a number of reasons [10,12]. Blood glucose dynamics are affected by various other factors apart from infection incidences [19,56,57], and characterization of infection-induced anomalies (abnormal class) from the normal class [13] is a challenging task because of the following reasons:

1. There are no well-defined boundaries regarding how different pathogens affect various key parameters of blood glucose dynamics, including blood glucose levels, insulin injections, carbohydrate ingestions, physical activity or exercise load, and others. This results in poor boundary demarcation between the normal and abnormal classes.
2. Class boundaries defined for a single pathogen might not work for the other pathogens because the effect of different pathogens on the blood glucose dynamics could be different.
3. It is expensive and time consuming to collect infection-related data to explore and characterize pathogen-specific class boundaries. This results in ill-defined class boundaries even for an infection related to a single pathogen.
4. The degree of effect of the same pathogens on the blood glucose dynamics could differ between different individuals because of the difference in individual immunity, which further complicates the characterization task.
5. Lack of sufficient sample size for both the abnormal and the normal classes results in poor training and testing data sample size or imbalanced class problems.

Given these challenges, the best possible approach is to identify methods that can learn from the normal health state of an individual and classify abnormalities relying on the boundaries learnt from the normal health state, which is a one-class classifier approach. This definitely reduces the challenge because it only requires the characterization of what is believed to be a normal health state. For instance, assume a health diagnostic and monitoring system that detects health changes in an individual by tracking the individual's physiological parameters, where the current health status is examined based on set of parameters, and raises a notification alarm when the individual health deteriorates [12]. In such a system, it becomes feasible to rely on a method that can be trained using only the regular or normal day measurements (target days) so as to detect

deviation from normality [12,14]. Another possible alternative approach is to identify a method that does not require any characterization and labeling of classes, which is unsupervised methods [7]. Accordingly, considering the previously mentioned challenges, one-class classifiers and unsupervised models were proposed for detecting infection incidence in people with type 1 diabetes. The objective was to develop a personalized health model that can automatically detect the incidence of infection in people with type 1 diabetes using blood glucose levels and insulin-to-carbohydrate ratio as input variables. The model is expected to detect any deviations from the norm as a result of infection incidences considering blood glucose level (hyperglycemia incidences) coupled with unusual changes in the insulin-to-carbohydrate ratio, that is, frequent insulin injections and unusual reduction in the amount of carbohydrate intake [19]. A personalized health model based on one-class classifiers and unsupervised methods was tested using blood glucose levels and the insulin-to-carbohydrate ratio as a bivariate input. The result demonstrated the potential of the proposed approach, which achieved excellent performance in describing the data set, that is, detecting infection days from the regular or normal days, and, in particular, the boundary and domain-based method performed better. Among the respective group, particular models such as V-SVM, K-NN, and K-means achieved excellent performance in all the sample sizes and infection cases. However, the unsupervised approaches suffer performance degradation compared with the one-class classifier mainly because of the atypical nature of the data, which are not distributed uniformly, where some regions may contain high density and others might be sparse (Multimedia Appendix 2). There are rare events (sparse region) of blood glucose dynamics that are a normal response; however, the unsupervised methods can still detect and flag false alarms including the following:

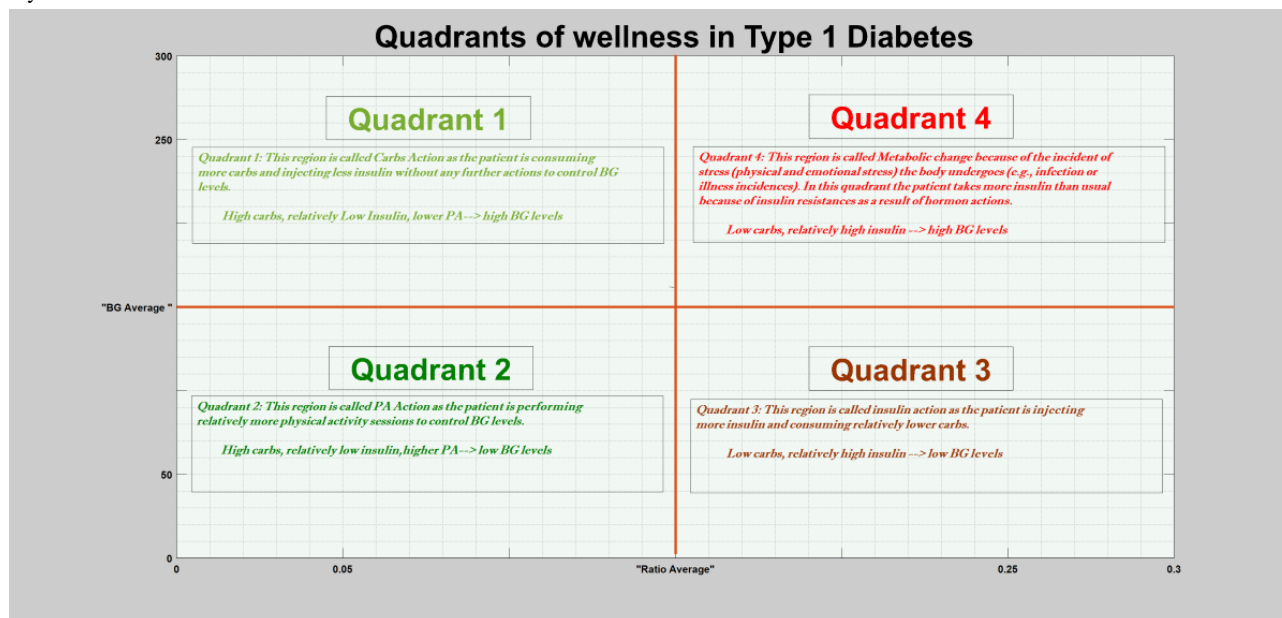
1. Carbohydrate action: a situation in which the ratio of insulin-to-carbohydrate is small and the blood glucose levels are high (hyperglycemia), *Carb Action-Quadrant 1* in Figure 7. This is a normal response to blood glucose dynamics as consumption of more carbohydrates and less insulin intake can derive blood glucose dynamics into the hyperglycemia region (high blood glucose levels) if there is no physical activity session. A typical example of this particular situation is holiday seasons, where people consume too many carbohydrates.
2. Physical activity action: despite a small ratio of insulin-to-carbohydrate, the blood glucose levels still drop to low levels (hypoglycemia), *PA Action-Quadrant 2* in Figure 7. Normally, a small ratio of insulin-to-carbohydrate signifies that the patient consumed more carbohydrates and injected less insulin, which normally derives the blood glucose dynamics into the hyperglycemia region. However, despite taking more carbohydrates and less insulin, a rigorous physical exercise can still derive the blood glucose dynamics into the hypoglycemia region. Therefore, this is a normal response of blood glucose dynamics as the action of physical activity or exercise can derive the patient into hypoglycemic regions even if the patient consumes more carbohydrates. For example, an individual patient on certain days might prefer to take little insulin as compared with most of the days and perform heavy physical activity to

replace their insulin needs. This scenario could generate an outlier, a small ratio of insulin-to-carbohydrate, which will be considered and detected as anomalies by the unsupervised models. However, this could be mitigated by incorporating physical activity data as an input variable.

- Insulin action: the ratio of insulin-to-carbohydrate is large, that is, high insulin intake and low carbohydrate

consumption, and blood glucose levels are low (hypoglycemia), *Insulin Action-Quadrant 3* in Figure 7. This is a normal response to blood glucose dynamics as administration of high insulin with little carbohydrate consumption can derive the blood glucose dynamics into the hypoglycemic region.

**Figure 7.** Quadrants of wellness in people with type 1 diabetes. The figure depicts the 4 possible scenarios of different parameters: carbohydrate action, insulin action, physical activity action, and abnormality because of metabolic change such as infection and stress. BG: blood glucose; PA: physical activity.



The drawback of unsupervised methods is that they do not have any mechanism to handle rare events even if the events are normal. This is mainly because unsupervised methods define an anomaly on the basis of the entire data set. However, the one-class classifier can learn and handle such scenarios appropriately if presented during the training phase. This is mainly because one-class classifiers produce a reference description based on the available normal (target) data set, including the rare events. With regard to the one-class classifiers, the boundary and domain-based method achieved a better description of the data set compared with the density and reconstruction-based methods, mainly because of the ability of such models to handle the atypical nature of the data [12]. Detectability of the infection incidence is directly related to the extent and degree of the effect it induces on the blood glucose dynamics. The type of pathogen, individual's immunity, and hormones involved could play a role in determining the degree of severity in this regard [19,24,58-62]. To this end, the results demonstrated that the models were capable of detecting all the infection incidences that can significantly alter the blood glucose dynamics, such as influenza. Moreover, infection incidence that had a moderate effect on the blood glucose dynamics, such as mild common cold without fever, was also detected. However, as expected, infection incidences that had almost little effect on the blood glucose dynamics, such as light common cold without fever, as reported by the individual patient, were not detected. Regarding the computational time, NN, SVDD, and SOM took considerable training time, which typically increased as the number of sample objects increased. Moreover, compared with

the other models, only LOF and COF took considerable testing time.

### Comparative Analysis of the Methods

Selecting the proper model for implementation in a real-world setting requires considering different characteristics of the model. This includes typical model characteristics such as performance in limited training sample size, robustness to outliers in the training data, required training and testing time, and complexity of the model (in terms of the number of model parameters).

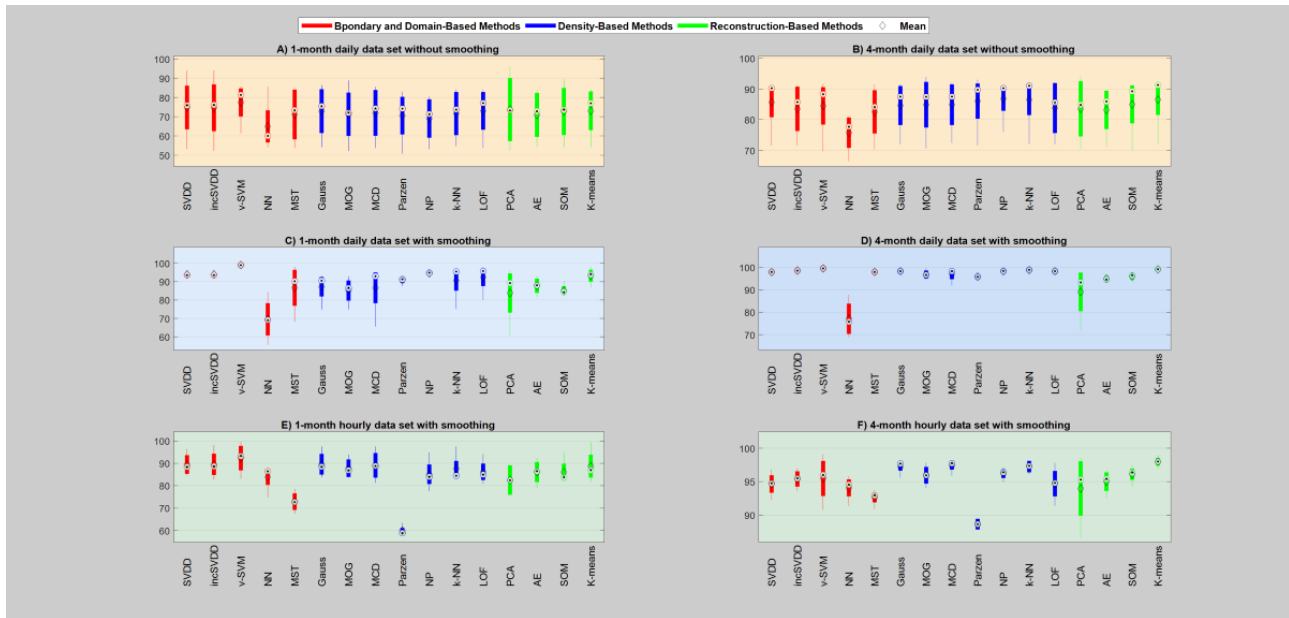
### Performance and Sample Size

The sample size,  $N$ , is the number of sample objects used during the training phase and highly affects the generalization power of the model [12,13]. Models trained with small sample sizes often fail to produce satisfactory descriptions mainly associated with the presence of large variance in the sample objects [3,12,13,63]. To this end, the results indicate that most of the models fail to make good descriptions with a 1-month (30 objects) data set, mainly with the daily raw data set, as shown in Figure 8. The figure depicts the average performance of each model across all the infection cases over the 1- and 4-month sample sizes. Specifically, MST, Gaussian families, SOM, and auto-encoders require a considerable amount of training sample objects to better describe the data. There is some exception, for instance V-SVM, which produces a satisfactory description of the 1-month data sets in all the infection cases and data granularity. Models such as NN and PCA produced the worst

description in most cases. As the number of training sample objects increased, all the models improved and produced a comparable description of the data. As a rule of thumb, for the daily scenario, a 3-month training sample (90 sample objects) produces a good description of the data, which can be considered for real-world applications. Moreover, if smoothing is

considered, a 1-month sample size produces better description than the 4-month sample size without smoothing, as shown in Figure 8. However, for the hourly scenario, a 1-month training sample object produces a comparable description and anything more than this size will be enough.

**Figure 8.** Average performance (F1-score) of each model across all the infection cases. AE: auto-encoder; Gauss: Gaussian; IncSVDD: incremental support vector data description; K-NN: K-nearest neighbor; LOF: local outlier factor; MCD: minimum covariance determinant; MOG: mixture of Gaussian; MST: minimum spanning tree; NN: nearest neighbor; NP: naïve Parzen; PCA: principal component analysis; SOM: self-organizing maps; SVDD: support vector data description; V-SVM: one-class support vector machine.



### Computational Time

For real-time applications, the time a model takes to learn and classify the sample object is essential in model selection. Table 9 depicts the rough estimation of average training and testing time required by different classifiers, both the one-class classifiers and the unsupervised models, based on 2880 training and testing sample objects each. Most of the models, as shown in Figures 5 and 6 and Table 9, require reasonable training and testing time, except NN, SVDD, and SOM, which took a

considerably longer time. However, it is possible that in some cases models can be trained offline, which makes the training time less important. With regard to the testing time, most of the models executed the classification task in a reasonable time except COF and one class classifier version of LOF, which consume considerable time to classify the 2880 objects. The computational time in these particular models grows exponentially as the sample size increases, which makes them resource demanding in a big data setting.

**Table 9.** Rough estimation of average training and testing time required by the different classifiers.

Methods	Training time, mean (SD)	Testing time, mean (SD)
<b>One-class classifiers</b>		
SVDD <sup>a</sup>	105.2 (2.03)	0.008 (0.002)
IncSVDD <sup>b</sup>	0.05 (0.16)	2.41 (0.83)
K-means	0.0047 (0.0014)	0.0032 (0.0010)
Gaussian	0.0055 (0.0032)	0.0032 (0.0012)
MOG <sup>c</sup>	0.076 (0.018)	0.0036 (0.0011)
MCD <sup>d</sup> Gaussian	0.27 (0.075)	0.0034 (0.0015)
SOM <sup>e</sup>	21.62 (5.91)	0.0033 (0.00087)
K-NN <sup>f</sup>	0.51 (0.11)	0.52 (0.12)
Parzen	2.02 (0.41)	0.21 (0.052)
Naïve Parzen	4.02 (0.82)	0.40 (0.10)
LOF <sup>g</sup>	1.15 (0.28)	1198.05 (323.07)
NN <sup>h</sup>	151.34 (22.52)	0.18 (0.024)
MST <sup>i</sup>	2.39 (0.31)	1.24 (0.19)
PCA <sup>j</sup>	0.046 (0.20)	0.0031 (0.00086)
Auto-encoder	0.65 (0.094)	0.017 (0.0034)
V-SVM <sup>k</sup>	0.32 (0.024)	0.035 (0.0066)
<b>Unsupervised</b>		
LOF <sup>l</sup>	N/A <sup>m</sup>	0.2 (0.0)
COF <sup>n</sup>	N/A	82.8 (1.5)

<sup>a</sup>SVDD: support vector data description.

<sup>b</sup>IncSVDD: incremental support vector data description.

<sup>c</sup>MOG: mixture of Gaussian.

<sup>d</sup>MCD: minimum covariance determinant.

<sup>e</sup>SOM: self-organizing maps.

<sup>f</sup>K-NN: K-nearest neighbor.

<sup>g</sup>LOF: local outlier factor.

<sup>h</sup>NN: nearest neighbor.

<sup>i</sup>MST: minimum spanning tree.

<sup>j</sup>PCA: principal component analysis.

<sup>k</sup>V-SVM: one-class support vector machine.

<sup>l</sup>LOF: local outlier factor.

<sup>m</sup>N/A: not applicable.

<sup>n</sup>COF: connectivity-based outlier factor.

### Robustness to Outliers in the Training Data Set

The presence of outliers in the training data set could significantly affect the model's generalization ability. Outlier objects are samples that exhibit different characteristics compared with the rest of the objects in the data set [8,63]. For instance, an individual might forget a previous infection incident and could label these days as a regular or normal period during self-reporting, which could end up being used as target data sets for training. Another important example could be error recorded

during data registration, that is, carbohydrate, blood glucose levels, and insulin registration. Such errors could occur during the manual registration of carbohydrates, associated with infusion set failures and other similar situations. In this scenario, an individual could record lower or higher values incorrectly affecting the input features, for example, ratio of insulin-to-carbohydrate and blood glucose levels, resulting in an outlier that could greatly affect the model's generalization ability. In this type of situation, a model's sensitivity to outliers in the training data is crucial to curb the influence of outliers

on the accuracy of the description generated. To some extent, a user-specified empirical rejection rate is incorporated in the models to reduce the effect of outliers in the training data by rejecting the most dissimilar objects from the description generated. For example, a rejection rate of 1% on training data sets implies that 1% of outliers in the training data set are rejected. Nevertheless, the sensitivity of models to outliers in the training data sets differs greatly between models. Among the models, NN is regarded as the most sensitive model to outliers in the training data set [12]. The presence of outliers in the training data changes the shape of the description generated by the model, forcing a larger portion of the feature space to be accepted as the target class [10,12]. Furthermore, models that rely on an estimation of the covariance matrix, for example, Gaussian families, also suffer from the presence of outliers in the training data sets [12,36]. However, when equipped with regularization, Gaussian models can withstand such outliers. Local density estimators such as Parzen can withstand outliers, considering the fact that only the local density is affected [12]. Models that rely on prototype estimation, such as SOM and K-means, are highly affected by the presence of outliers in the training data set, which could force the estimated prototype to be placed near or at the nontarget data set [2,12,13]. Nevertheless, boundary and domain-based method such as SVDD and V-SVM and reconstruction-based method such as auto-encoders are more or less insensitive to outliers and can generate acceptable solutions [3,12,64].

### **Model Parameters and Associated Complexity**

The parameters of a model can be either free or user defined. These two parameters, free and user defined, provide insight into how flexible the model is, how sensitive the model is to overtraining, and how easy the model is to configure (simplicity) [12,16]. Considering the number of these parameters, there exist large variations among the models. For instance, NN does not possess any free parameters; therefore, its performance completely relies on the training data set [12]. This constraint has limitations, mainly because training data that contain outliers could ruin the model's performance [12,15,16]. A model that possess large number of free and user defined parameters is too flexible and complex [12]. Regarding the user-defined parameters, also known as hyper-parameters, a model equipped with small number of parameters and preferably with intuitive meaning are easy to configure. Setting up the user defined parameters incorrectly can degrade the model's performance and selecting the proper values (optimization) becomes complex and vague as the number of model parameters become too large. One of the simplest models is Parzen density and NN, which do not require the user to specify any parameters [3,12,13]. Some models, such as support vector families, require the user to specify parameters that have intuitive meaning, for example, the ratio of training objects to be rejected by the description [12,65]. There are also models that are complex enough given that the user is expected to specify many parameters, which are not intuitive and require careful choice. Examples of such models include SOM and auto-encoders, where the user is expected to supply the number of neuron, hidden units, and learning rate [10,12,37,66].

### **Practical Illustration and Area of Applications**

For a real-world application, apart from the performance of the model, it is important to consider two important aspects of the data set, the time window of detection (data granularity) and the required sample size. The time window or data granularity, that is, hourly and daily, defines the frequency (continuity) of computation one needs to conduct throughout the day to screen the health status of the individual with type 1 diabetes. In an hourly time window, one is expected to carry out the computation at the end of each hour throughout the day. However, in the daily time window, one needs to carry out one aggregate computation at the end of the day. Decreasing the time window (increasing the granularity of the data) enhances early detections; however, at the coast of accuracy, for example, more unwanted features (noise) in the data. The results demonstrated that almost all the models produced fairly comparable detection performances in both time windows. Moreover, the required sample size determines the necessary amount of data an individual with type 1 diabetes needs to collect in advance before joining such an infection detection system. Models that could generalize well with small sample sizes could be preferred in a real-world application to enable more people to join the system with ease. Generally, the results demonstrated that the models require at least a sample size of 3-month data for the daily case and 2-month data for hourly case to perform better. Automating the detection of infection incidences among people with type 1 diabetes can deliver a means to provide personalized decision support and learning platforms for the individuals and, at the same time, can be used to detect infectious disease outbreaks on a large scale through spatio-temporal cluster detection [19,67,68]. Detailed descriptions of these instances are given below:

1. A personalized decision support system and learning platform relies on an individual's self-recorded data to provide relevant information in relation to decision making to assist the individuals during crises [19,67,68]. Moreover, it can also provide a learning platform concerning the extent to which infection incidence affects the key parameters of the blood glucose dynamics. Information regarding what to expect at each stage of the course of infection could be very important to the individuals [19]. During infection incidences, various kinds of information could be vital for an individual to properly manage blood glucose levels, including time in range (blood glucose), to what extent is the evolution of blood glucose affected during the course of infection, to what extent does insulin sensitivity change, and how much does the insulin-to-carbohydrate ratio shift, that is, changes in insulin requirements for each gram of carbohydrate intake.
2. A population-based early outbreak detection system relies on self-recorded information from an individual with type 1 diabetes to detect individuals' infection cases and, thereby, detect a group of infected individuals on a spatio-temporal basis. Such a system should collect individuals' self-recorded data to a central server, analyze individuals' data on a timely basis, identify and locate a cluster of people based on space and time, and notify the responsible bodies if there is an ongoing outbreak [19,67-71].

## Conclusions

Anomaly or novelty detection problem has been widely used in various applications including machine fault and sensor failure detection, prevention of credit card or identity fraud, health and medical diagnostics and monitoring, cyber-intrusion detection, and others. In this study, we demonstrated the applicability of one-class classifiers and unsupervised anomaly detection methods for the purpose of detecting infection incidences in people with type 1 diabetes. In general, the proposed methods produced excellent performance in describing the data set, and particularly the boundary and domain-based method performed better. In contrast to the specific models,

V-SVM, K-NN, and K-means achieved better generalization in describing the data set in all infection cases. Detecting the incidence of infection in people with type 1 diabetes can provide an opportunity to devise tailored services, that is, personalized decision support and a learning platform for the individuals, and can simultaneously be used for detecting potential public health threats, that is, infectious disease outbreaks, on a large-scale basis through a spatio-temporal cluster detection. Generally, we foresee that the results presented could encourage researchers to further examine the presented features along with other additional features of self-recorded data, for example, various CGM features and physical activity data, on a large-scale basis.

## Acknowledgments

The work presented in this paper is part of the project *Electronic Disease Surveillance Monitoring Network (EDMON) system*, which is funded by the University of Tromsø–The Arctic University of Norway and National Library of Medicine *Mechanistic machine learning* (grant number: LM012734) and is the PhD program of the first author, AW. The authors would like to extend their sincere gratitude to all the participants of the study.

## Authors' Contributions

The first author, AW, conceived the study, designed and performed the experiments, and wrote the manuscript. IK, EÅ, JI, DA, and GH provided successive inputs and revised the manuscript. All authors approved the final manuscript.

## Conflicts of Interest

None declared.

## Multimedia Appendix 1

Theoretical background of the methods.

[\[DOCX File , 73 KB-Multimedia Appendix 1\]](#)

## Multimedia Appendix 2

Detailed description of the models input features.

[\[DOCX File , 12076 KB-Multimedia Appendix 2\]](#)

## Multimedia Appendix 3

Score plot of the models for each patient year.

[\[DOCX File , 12308 KB-Multimedia Appendix 3\]](#)

## Multimedia Appendix 4

Model evaluations – performance of the models for each patient year.

[\[DOCX File , 62 KB-Multimedia Appendix 4\]](#)

## References

1. Dunning T, Friedman E. In: Loukides M, editor. *Practical Machine Learning: A New Look at Anomaly Detection*. New York, USA: O'Reilly Media Inc; 2014.
2. Agrawal S, Agrawal J. Survey on anomaly detection using data mining techniques. *Procedia Comput Sci* 2015;60:708-713 [[FREE Full text](#)] [doi: [10.1016/j.procs.2015.08.220](https://doi.org/10.1016/j.procs.2015.08.220)]
3. Pimentel MA, Clifton DA, Clifton L, Tarassenko L. A review of novelty detection. *Sig Process* 2014 Jun;99:215-249 [[FREE Full text](#)] [doi: [10.1016/j.sigpro.2013.12.026](https://doi.org/10.1016/j.sigpro.2013.12.026)]
4. Cohen G, Hilario M, Sax H, Hugonnet S, Pellegrini C, Geissbuhler A. An application of one-class support vector machine to nosocomial infection detection. *Stud Health Technol Inform* 2004;107(Pt 1):716-720. [Medline: [15360906](https://pubmed.ncbi.nlm.nih.gov/15360906/)]
5. Cohen G, Sax H, Geissbuhler A. Novelty detection using one-class Parzen density estimator. An application to surveillance of nosocomial infections. *Stud Health Technol Inform* 2008;136:21-26. [Medline: [18487702](https://pubmed.ncbi.nlm.nih.gov/18487702/)]



6. Cohen G, Hilario M, Sax H, Hugonnet S, Geissbuhler A. Learning from imbalanced data in surveillance of nosocomial infection. *Artif Intell Med* 2006 May;37(1):7-18. [doi: [10.1016/j.artmed.2005.03.002](https://doi.org/10.1016/j.artmed.2005.03.002)] [Medline: [16233974](https://pubmed.ncbi.nlm.nih.gov/16233974/)]
7. Chandola V, Banerjee A, Kumar V. Anomaly detection: a survey. *ACM Comput Surv* 2009 Jul;41(3):1-58 [FREE Full text] [doi: [10.1145/1541880.1541882](https://doi.org/10.1145/1541880.1541882)]
8. Hawkins DM. Identification of Outliers. Netherlands: Springer; 1980.
9. Mehrotra KG, Mohan CK, Huang H. In: Subrahmanian VS, editor. Anomaly Detection Principles and Algorithms. Cham, Switzerland: Springer International Publishing; 2017.
10. Khan SS, Madden MG. One-class classification: taxonomy of study and review of techniques. *Knowl Eng Rev* 2014 Jan 24;29(3):345-374 [FREE Full text] [doi: [10.1017/s026988891300043x](https://doi.org/10.1017/s026988891300043x)]
11. Ding X, Li Y, Belatreche A, Maguire LP. An experimental evaluation of novelty detection methods. *Neurocomputing* 2014 Jul;135:313-327 [FREE Full text] [doi: [10.1016/j.neucom.2013.12.002](https://doi.org/10.1016/j.neucom.2013.12.002)]
12. Tax DM. One-Class Classification: Concept Learning in the Absence of Counter-Examples. South Holland: Technische Universiteit Delft; 2002.
13. Irigoien I, Sierra B, Arenas C. Towards application of one-class classification methods to medical data. *ScientificWorldJournal* 2014;2014:730712 [FREE Full text] [doi: [10.1155/2014/730712](https://doi.org/10.1155/2014/730712)] [Medline: [24778600](https://pubmed.ncbi.nlm.nih.gov/24778600/)]
14. Japkowicz N. Concept Learning in the Absence of Counterexamples: an Autoassociation-Based Approach to Classification. Semantic Scholar. 1999. URL: <https://www.semanticscholar.org/paper/Concept-learning-in-the-absence-of-counterexamples%3A-Japkowicz-Hanson/03ed0a73d1f7a7b16505d6cb9c8bfbeef7b19bb> [accessed 2020-07-23]
15. Juszczak P, Tax DM, Pe, kalska E, Duin RP. Minimum spanning tree based one-class classifier. *Neurocomputing* 2009 Mar;72(7-9):1859-1869 [FREE Full text] [doi: [10.1016/j.neucom.2008.05.003](https://doi.org/10.1016/j.neucom.2008.05.003)]
16. Mazhelis O. One-class classifiers: a review and analysis of suitability in the context of mobile-masquerader detection. *S Afr Comput J* 2006;2006(36):29-48 [FREE Full text]
17. Clark M. What is diabetes? In: Ogden J, editor. Understanding Diabetes. New Jersey, United States: John Wiley & Sons, Ltd; Jan 20, 2004.
18. Ogurtsova K, da Rocha FJ, Huang Y, Linnenkamp U, Guariguata L, Cho N, et al. IDF diabetes atlas: global estimates for the prevalence of diabetes for 2015 and 2040. *Diabetes Res Clin Pract* 2017 Jun;128:40-50. [doi: [10.1016/j.diabres.2017.03.024](https://doi.org/10.1016/j.diabres.2017.03.024)] [Medline: [28437734](https://pubmed.ncbi.nlm.nih.gov/28437734/)]
19. Woldaregay A, Årsand E, Albers D, Launonen I, Hartvigsen G. Towards detecting infection incidences in people with type 1 diabetes using self-recorded data: a novel framework for a digital infectious disease detection mechanism. *JMIR preprints* 2020:- preprint [FREE Full text] [doi: [10.2196/preprints.18911](https://doi.org/10.2196/preprints.18911)]
20. Marcovecchio ML, Chiarelli F. The effects of acute and chronic stress on diabetes control. *Sci Signal* 2012 Oct 23;5(247):pt10. [doi: [10.1126/scisignal.2003508](https://doi.org/10.1126/scisignal.2003508)] [Medline: [23092890](https://pubmed.ncbi.nlm.nih.gov/23092890/)]
21. Rayfield EJ, Ault MJ, Keusch GT, Brothers MJ, Nechemias C, Smith H. Infection and diabetes: the case for glucose control. *Am J Med* 1982 Mar;72(3):439-450. [doi: [10.1016/0002-9343\(82\)90511-3](https://doi.org/10.1016/0002-9343(82)90511-3)] [Medline: [7036735](https://pubmed.ncbi.nlm.nih.gov/7036735/)]
22. Botsis T, Lai AM, Hripisak G, Palmas W, Starren JB, Hartvigsen G. Proof of concept for the role of glycemic control in the early detection of infections in diabetics. *Health Informatics J* 2012 Mar;18(1):26-35. [doi: [10.1177/1460458211428427](https://doi.org/10.1177/1460458211428427)] [Medline: [22447875](https://pubmed.ncbi.nlm.nih.gov/22447875/)]
23. Botsis T, Hejlesen O, Bellika JG, Hartvigsen G. Blood Glucose Levels as an Indicator for the Early Detection of Infections In Type-1 Diabetics. Faculty & Staff Insider - University of Washington. 2007. URL: <http://faculty.washington.edu/lober/www.isdsjournal.org/htdocs/articles/2025.pdf> [accessed 2020-07-23]
24. Mizock BA. Alterations in carbohydrate metabolism during stress: a review of the literature. *Am J Med* 1995 Jan;98(1):75-84. [doi: [10.1016/S0002-9343\(99\)80083-7](https://doi.org/10.1016/S0002-9343(99)80083-7)] [Medline: [7825623](https://pubmed.ncbi.nlm.nih.gov/7825623/)]
25. Bosarge PL, Kerby JD. Stress-induced hyperglycemia: is it harmful following trauma? *Adv Surg* 2013;47:287-297. [doi: [10.1016/j.yasu.2013.03.002](https://doi.org/10.1016/j.yasu.2013.03.002)] [Medline: [24298857](https://pubmed.ncbi.nlm.nih.gov/24298857/)]
26. Kajbaf F, Mojtahedzadeh M, Abdollahi M. Mechanisms underlying stress-induced hyperglycemia in critically ill patients. *Therapy* 2007 Jan;4(1):97-106 [FREE Full text] [doi: [10.2217/14750708.4.1.97](https://doi.org/10.2217/14750708.4.1.97)]
27. Tax DM, Duin RP. Support vector data description. *Mach Learn* 2004 Jan;54(1):45-66 [FREE Full text] [doi: [10.1023/b:mach.0000008084.60811.49](https://doi.org/10.1023/b:mach.0000008084.60811.49)]
28. Schölkopf B, Williamson RC, Smola AJ, Shawe-Taylor J, Platt JC. Support Vector Method for Novelty Detection. NIPS Proceedings. 1999. URL: <https://papers.nips.cc/paper/1723-support-vector-method-for-novelty-detection.pdf> [accessed 2020-07-23]
29. Tax DM, Duin RP. Support vector domain description. *Pattern Recognit Lett* 1999 Nov;20(11-13):1191-1199 [FREE Full text] [doi: [10.1016/s0167-8655\(99\)00087-2](https://doi.org/10.1016/s0167-8655(99)00087-2)]
30. Ridder DD, Tax DM, Duin, RP. An Experimental Comparison of One-Class Classification Methods. In: Proceedings of the 4th Annual Conference of the Advanced School for Computing and Imaging. 1998 Presented at: Annual Conference of the Advacned School for Computing and Imaging; September 29-October 4, 1998; Delft, Netherlands URL: <https://www.researchgate.net/publication/282673032> An experimental comparison of classification algorithm performances for highly imbalanced datasets

31. Breunig MM, Kriegel H, Ng RT, Sander J. LOF: identifying density-based local outliers. *SIGMOD Rec* 2000 Jun 1;29(2):93-104 [FREE Full text] [doi: [10.1145/335191.335388](https://doi.org/10.1145/335191.335388)]
32. Tax D. Software. DDTools. Delft: Delft University of Technology; 2015. URL: <https://www.tudelft.nl/ewi/over-de-faculteit/afdelingen/intelligent-systems/pattern-recognition-bioinformatics/pattern-recognition-bioinformatics/data-and-software/dd-tools> [accessed 2019-02-10]
33. A Collection of Algorithms for Anomaly/Outlier Detection. Anomaly Detection Toolbox. 2016. URL: <http://dsmi-lab-ntust.github.io/AnomalyDetectionToolbox/> [accessed 2019-03-25]
34. Tang J, Chen Z, Fu AW, Cheung DW. Enhancing effectiveness of outlier detections for low density patterns. In: *Advances in Knowledge Discovery and Data Mining*. New York, USA: Springer; 2002.
35. Duin R, Juszczak P, Paclik P, Pekalska E, De Ridder D, Tax DM. Software. Delft University of Technology. 2007. URL: <http://prtools.tudelft.nl/Guide/37Pages/software.html> [accessed 2019-02-25]
36. Goldstein M, Uchida S. A comparative evaluation of unsupervised anomaly detection algorithms for multivariate data. *PLoS One* 2016;11(4):e0152173 [FREE Full text] [doi: [10.1371/journal.pone.0152173](https://doi.org/10.1371/journal.pone.0152173)] [Medline: [27093601](https://pubmed.ncbi.nlm.nih.gov/27093601/)]
37. Swersky L, Marques H, Sander J, Campello RJ, Zimek A. On the Evaluation of Outlier Detection and One-Class Classification Methods. In: *IEEE International Conference on Data Science and Advanced Analytics (DSAA)*. 2016 Presented at: DSAA'16; October 17-19, 2016; Montreal, QC, Canada URL: <https://doi.org/10.1109/DSAA.2016.8> [doi: [10.1109/dsaa.2016.8](https://doi.org/10.1109/dsaa.2016.8)]
38. Rousseeuw PJ, Driessen KV. A fast algorithm for the minimum covariance determinant estimator. *Technometrics* 1999 Aug;41(3):212-223 [FREE Full text] [doi: [10.1080/00401706.1999.10485670](https://doi.org/10.1080/00401706.1999.10485670)]
39. Parzen E. On estimation of a probability density function and mode. *Ann Math Statist* 1962 Sep;33(3):1065-1076 [FREE Full text] [doi: [10.1214/aoms/1177704472](https://doi.org/10.1214/aoms/1177704472)]
40. Rasoulzadeh V, Erkus EC, Yogurt TA, Ulusoy I, Zergeroğlu SA. A comparative stationarity analysis of EEG signals. *Ann Oper Res* 2016 Apr 26;258(1):133-157 [FREE Full text] [doi: [10.1007/s10479-016-2187-3](https://doi.org/10.1007/s10479-016-2187-3)]
41. Azami H, Mohammadi K, Bozorgtabar B. An improved signal segmentation using moving average and Savitzky-Golay filter. *J Signal Inf Process* 2012;03(01):39-44. [doi: [10.4236/jsip.2012.31006](https://doi.org/10.4236/jsip.2012.31006)]
42. Kandanaarachchi S, Muñoz MA, Hyndman RJ, Smith-Miles K. On normalization and algorithm selection for unsupervised outlier detection. *Data Min Knowl Disc* 2019 Nov 21;34(2):309-354 [FREE Full text] [doi: [10.1007/s10618-019-00661-z](https://doi.org/10.1007/s10618-019-00661-z)]
43. Tax DM, Muller KA. A Consistency-based Model Selection for One-Class Classification. In: *Proceedings of the 17th International Conference on Pattern Recognition*. 2004 Presented at: ICPR'04; August 26, 2004; Cambridge, UK. [doi: [10.1109/icpr.2004.1334542](https://doi.org/10.1109/icpr.2004.1334542)]
44. Maldonado S, Montecinos C. Robust classification of imbalanced data using one-class and two-class SVM-based multiclassifiers. *Intell Data Anal* 2014 Jan 1;18(1):95-112. [doi: [10.3233/ida-130630](https://doi.org/10.3233/ida-130630)]
45. Tharwat A. Classification assessment methods. *Appl Comput Inf* 2018 Aug:- epub ahead of print [FREE Full text] [doi: [10.1016/j.aci.2018.08.003](https://doi.org/10.1016/j.aci.2018.08.003)]
46. Nguyen GH, Bouzerdoum A, Phung SL. Learning pattern classification tasks with imbalanced data sets. *Pattern Recog* 2009:- [FREE Full text] [doi: [10.5772/7544](https://doi.org/10.5772/7544)]
47. Hajizadeh S, Li Z, Dollevoet RP, Tax DM. Evaluating classification performance with only positive and unlabeled samples. In: Fränti P, Brown G, Loog M, Escolano F, Pelillo M, editors. *Structural, Syntactic, and Statistical Pattern Recognition: Joint IAPR International Workshop*. Berlin, Heidelberg: Springer; 2014.
48. Hajizadeh S, Núñez A, Tax DM. Semi-supervised rail defect detection from imbalanced image data. *IFAC-PapersOnLine* 2016;49(3):78-83 [FREE Full text] [doi: [10.1016/j.ifacol.2016.07.014](https://doi.org/10.1016/j.ifacol.2016.07.014)]
49. Bradley AP. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recog* 1997 Jul;30(7):1145-1159 [FREE Full text] [doi: [10.1016/s0031-3203\(96\)00142-2](https://doi.org/10.1016/s0031-3203(96)00142-2)]
50. McNamara LA, Martin S. Principles of epidemiology and public health. In: Long SS, Prober CG, Fischer M, editors. *Principles and Practice of Pediatric Infectious Diseases*. Fifth Edition. New York, USA: Elsevier; 2018.
51. Zhu Y. Automatic detection of anomalies in blood glucose using a machine learning approach. *J Commun Netw* 2011 Apr;13(2):125-131 [FREE Full text] [doi: [10.1109/jcn.2011.6157411](https://doi.org/10.1109/jcn.2011.6157411)]
52. Zhu Y. Automatic Detection of Anomalies in Blood Glucose Using a Machine Learning Approach. In: *International Conference on Information Reuse & Integration*. 2010 Presented at: IER'10; August 4-6, 2010; Las Vegas, NV, USA URL: <https://ieeexplore.ieee.org/document/5558959> [doi: [10.1109/jcn.2011.6157411](https://doi.org/10.1109/jcn.2011.6157411)]
53. Spinosa EJ, Carvalho AC. Support vector machines for novel class detection in Bioinformatics. *Genet Mol Res* 2005 Sep 30;4(3):608-615. [Medline: [16342046](https://pubmed.ncbi.nlm.nih.gov/16342046/)]
54. Lotze TH. Anomaly Detection in Time Series: Theoretical and Practical Improvements for Disease Outbreak Detection. Digital Repository at the University of Maryland. 2009. URL: <https://drum.lib.umd.edu/handle/1903/9857> [accessed 2020-07-23]
55. Tsui F, Espino JU, Dato VM, Gesteland PH, Hutman J, Wagner MM. Technical description of RODS: a real-time public health surveillance system. *J Am Med Inform Assoc* 2003;10(5):399-408 [FREE Full text] [doi: [10.1197/jamia.M1345](https://doi.org/10.1197/jamia.M1345)] [Medline: [12807803](https://pubmed.ncbi.nlm.nih.gov/12807803/)]

56. Woldaregay AZ, Årsand E, Botsis T, Albers D, Mamykina L, Hartvigsen G. Data-driven blood glucose pattern classification and anomalies detection: machine-learning applications in type 1 diabetes. *J Med Internet Res* 2019 May 1;21(5):e11030 [FREE Full text] [doi: [10.2196/11030](https://doi.org/10.2196/11030)] [Medline: [31042157](https://pubmed.ncbi.nlm.nih.gov/31042157/)]
57. Oviedo S, Vehí J, Calm R, Armengol J. A review of personalized blood glucose prediction strategies for T1DM patients. *Int J Numer Method Biomed Eng* 2017 Jun;33(6):---. [doi: [10.1002/cnm.2833](https://doi.org/10.1002/cnm.2833)] [Medline: [27644067](https://pubmed.ncbi.nlm.nih.gov/27644067/)]
58. Yki-Järvinen H, Sammalkorpi K, Koivisto VA, Nikkilä EA. Severity, duration, and mechanisms of insulin resistance during acute infections. *J Clin Endocrinol Metab* 1989 Aug;69(2):317-323. [doi: [10.1210/jcem-69-2-317](https://doi.org/10.1210/jcem-69-2-317)] [Medline: [2666428](https://pubmed.ncbi.nlm.nih.gov/2666428/)]
59. Rayfield EJ, Curnow RT, George DT, Beisel WR. Impaired carbohydrate metabolism during a mild viral illness. *N Engl J Med* 1973 Sep 20;289(12):618-621. [doi: [10.1056/NEJM197309202891207](https://doi.org/10.1056/NEJM197309202891207)] [Medline: [4198822](https://pubmed.ncbi.nlm.nih.gov/4198822/)]
60. McGuinness OP. Defective glucose homeostasis during infection. *Annu Rev Nutr* 2005;25:9-35. [doi: [10.1146/annurev.nutr.24.012003.132159](https://doi.org/10.1146/annurev.nutr.24.012003.132159)] [Medline: [16011457](https://pubmed.ncbi.nlm.nih.gov/16011457/)]
61. Brealey D, Singer M. Hyperglycemia in critical illness: a review. *J Diabetes Sci Technol* 2009 Nov 1;3(6):1250-1260 [FREE Full text] [doi: [10.1177/1932296809000300604](https://doi.org/10.1177/1932296809000300604)] [Medline: [20144378](https://pubmed.ncbi.nlm.nih.gov/20144378/)]
62. Mizock BA. Alterations in fuel metabolism in critical illness: hyperglycaemia. *Best Pract Res Clin Endocrinol Metab* 2001 Dec;15(4):533-551. [doi: [10.1053/beem.2001.0168](https://doi.org/10.1053/beem.2001.0168)] [Medline: [11800522](https://pubmed.ncbi.nlm.nih.gov/11800522/)]
63. Tax DM, Duin RP. Characterizing One-Class Datasets. CiteSeerX. 2005. URL: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.460.8322&rep=rep1&type=pdf> [accessed 2019-07-10]
64. Wang D, Yeung DS, Tsang EC. Structured one-class classification. *IEEE Trans Syst Man Cybern B Cybern* 2006 Dec;36(6):1283-1295. [doi: [10.1109/tsmcb.2006.876189](https://doi.org/10.1109/tsmcb.2006.876189)] [Medline: [17186805](https://pubmed.ncbi.nlm.nih.gov/17186805/)]
65. Janssens JH. Outlier Selection and One-Class Classification. Netherlands: Embedded Systems Institute, Tilburg University; 2013.
66. Wang B, Mao Z. One-class classifiers ensemble based anomaly detection scheme for process control systems. *T I Meas Control* 2017 Sep 21;40(12):3466-3476 [FREE Full text] [doi: [10.1177/0142331217724508](https://doi.org/10.1177/0142331217724508)]
67. Samerski S. Individuals on alert: digital epidemiology and the individualization of surveillance. *Life Sci Soc Policy* 2018 Jun 14;14(1):13 [FREE Full text] [doi: [10.1186/s40504-018-0076-z](https://doi.org/10.1186/s40504-018-0076-z)] [Medline: [29900518](https://pubmed.ncbi.nlm.nih.gov/29900518/)]
68. Radin JM, Wineinger NE, Topol EJ, Steinhubl SR. Harnessing wearable device data to improve state-level real-time surveillance of influenza-like illness in the USA: a population-based study. *Lancet Digit Health* 2020 Feb;2(2):e85-e93 [FREE Full text] [doi: [10.1016/s2589-7500\(19\)30222-5](https://doi.org/10.1016/s2589-7500(19)30222-5)]
69. Woldaregay A, Årsand E, Giordanengo A, Albers D, Mamykina L, Botsis T, et al. EDMON-A Wireless Communication Platform for a Real-time Infectious Disease Outbreak Detection System Using Self-recorded Data From People With Type 1 Diabetes. In: *The 15th Scandinavian Conference on Health Informatics*. 2017 Presented at: SHI'17; August 29, 2017; Kristiansand, Norway URL: [https://ep.liu.se/konferensartikel.aspx?series=ecp&issue=145&Article\\_No=3](https://ep.liu.se/konferensartikel.aspx?series=ecp&issue=145&Article_No=3)
70. Coucheron S, Woldaregay AZ, Årsand E, Botsis T, Hartvigsen G. EDMON - A System Architecture for Real-Time Infection Monitoring and Outbreak Detection Based on Self-Recorded Data from People with Type 1 Diabetes: System Design and Prototype Implementation. In: *The 17th Scandinavian Conference on Health Informatics*. 2019 Presented at: CHI'19; November 12-13, 2019; Oslo, Norway URL: <https://ep.liu.se/ecp/161/007/ecp19161007.pdf>
71. Yeng PK, Woldaregay AZ, Solvoll T, Hartvigsen G. Cluster detection mechanisms for syndromic surveillance systems: systematic review and framework development. *JMIR Public Health Surveill* 2020 May 26;6(2):e11512 [FREE Full text] [doi: [10.2196/11512](https://doi.org/10.2196/11512)] [Medline: [32357126](https://pubmed.ncbi.nlm.nih.gov/32357126/)]

## Abbreviations

- AUC:** area under the receiver operating characteristic curve
- COF:** connectivity-based outlier factor
- IncSVDD:** incremental support vector data description
- K-NN:** K-nearest neighbor
- LOF:** local outlier factor
- MCD:** minimum covariance determinant
- MOG:** mixture of Gaussian
- MST:** minimum spanning tree
- NN:** nearest neighbor
- PCA:** principal component analysis
- SOM:** self-organizing maps
- SVDD:** support vector data description
- ROC:** receiver operating characteristic curve
- V-SVM:** one-class support vector machine

*Edited by G Eysenbach; submitted 26.03.20; peer-reviewed by S Sarbadhikari, M Nomali; comments to author 13.04.20; revised version received 06.06.20; accepted 11.06.20; published 12.08.20*

*Please cite as:*

*Woldaregay AZ, Launonen IK, Albers D, Igual J, Årsand E, Hartvigsen G*

*A Novel Approach for Continuous Health Status Monitoring and Automatic Detection of Infection Incidences in People With Type 1 Diabetes Using Machine Learning Algorithms (Part 2): A Personalized Digital Infectious Disease Detection Mechanism*

*J Med Internet Res 2020;22(8):e18912*

*URL: <https://www.jmir.org/2020/8/e18912>*

*doi: [10.2196/18912](https://doi.org/10.2196/18912)*

*PMID:*

©Ashenafi Zebene Woldaregay, Ilkka Kalervo Launonen, David Albers, Jorge Igual, Eirik Årsand, Gunnar Hartvigsen. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 12.08.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

# Appendix 1: Theoretical Background of the Methods

## Background

### 1.1.1. Notion of Object

An object is described with a feature vector encompassing the number of parameters under consideration. For example, an object  $k$  can define a specific event of an individual BG dynamics at a specified time index  $k$ , and is represented by a feature vector  $\mathcal{X}_k = (x_{1,1}, x_{1,2})$ , where  $x_{1,1}$  represent the insulin to carb ratio at the time index  $k$  and  $x_{1,2}$  represent the average BG level in the specified time bin around  $k$ .

### 1.1.2. One-class Classifier

One-class classification problem can be regarded as a special type of two-class classification problem incorporating a target and non-target class [7, 9]. The target class is where a model is trained on, and is expected to incorporate well sampled representative object of the target that reflect the region of the data in the feature space. The non-target (outlier) class is sparsely represented and sometimes can be totally absent. The task of one-classification is governed mainly by distinct two elements; function (model) and threshold. Function (model) measures the resemblance of an object  $x$  depending on the distance  $d(x)$  or probability  $\mathcal{P}(x)$  to the target class described by the training dataset  $\mathcal{X}_{tr}$ . A threshold ( $\beta$ ) is used to decide the belongingness of a test object  $x$  to either of the classes, i.e. non-target or the target class, depending on the definition of the function's internal structure. For example, test objects are rejected when computed distance by the function (model) is greater than some specified threshold  $\beta$ ,

$$\psi(x) = \begin{cases} target & I(d(x) \leq \beta) \\ non - target & I(d(x) > \beta) \end{cases} \quad (1)$$

Or when computed resemblance by the function (model) is less than some specified threshold  $\beta$ ,

$$\psi(x) = \begin{cases} target & I(\mathcal{P}(x) \geq \beta) \\ non - target & I(\mathcal{P}(x) < \beta) \end{cases} \quad (2)$$

Where  $\psi(x)$  is the class label and  $I$  is an indicator function.

Depending on the type of internal function (model) used, one-class classifier can be broadly categorized into three main groups; boundary and domain-based, density-based, and reconstruction-based method [5, 7, 10-12]. The main difference between these methods is the way they define of the function (model), and minimization approaches, thereby achieving different generalization, bias and overfitting as they consider different data characteristics [7].

#### 1.1.2.1. Boundary and Domain-Based Method

Boundary-based method estimate a boundary, e.g. hyperplane/hypersphere, around majority of the training (target) dataset, where a predefined small percentage (fraction) of the target data are allowed to lie outside the specified boundary [11, 12]. A test object is regarded as outlier if it falls outside of the defined boundary. The resemblance of the test object is determined by computing the distance from the test object to the boundary estimated around the

training objects. The distance computation is conducted by taking into account both I) the inter-distance between the new object and the training datasets and II) intra-distance between the objects in the training datasets. Different boundary and domain-based method exist including support vector data description (SVDD), one-class support vector machine (v-SVM), incremental support vector machine, nearest neighbors (NN), and minimum spanning tree (MST) [7, 10, 11].

**Support Vector, and Incremental Support Vector Data Description Vs. One-class Support Vector Classifier:** These methods describe the target class by fitting the training (target) dataset into either a hypersphere or a hyperplane respectively [7, 12, 13]. SVDD defines a hypersphere that encompass the entire target dataset with a volume as minimum as possible [14, 15]. SVDD carries out the minimization task through quadratic programming problem [9]. The incremental version is applicable to a problem that involves online and sequential data [16]. During the training phase, the parameter  $\alpha_i$  are estimated by minimizing

$$L = \varepsilon_{SVDD} = \sum_i \alpha_i (x_i \cdot x_j) - \sum_{i,j} \alpha_i \alpha_j (x_i \cdot x_j) \quad (3)$$

Subject to the constraints  $\sum_i \alpha_i = 1$  and  $0 \leq \alpha_i \leq \mathfrak{C}$ , where  $\mathfrak{C}$  specifies the number of vectors that will not be covered by the description. The minimization is solved via quadratic programming problem by relying on kernel function, e.g. Gaussian, to replace the inner product so as to transform the vectors into a higher dimensional feature space for a more accurate description. A new test object  $x$  is evaluated by computing its distance to the hypersphere's center and comparing against the hypersphere's radius:

$$\psi(x) = \begin{cases} target, & \text{if } \|x - a\|^2 \leq \mathbb{R}^2 \\ non - target, & \text{otherwise} \end{cases} \quad (4)$$

Where,  $a$  is the hypersphere's center, and is computed as  $\sum_i \alpha_i x_i$ . The hypersphere's radius is calculated as:

$$\mathbb{R} = (x_k \cdot x_k) - 2 \sum_i \alpha_i (x_i \cdot x_k) + \sum_{i,j} \alpha_i \alpha_j (x_i \cdot x_j) \quad (5)$$

Where  $x_k$  are the vectors, which have  $\alpha_i < \mathfrak{C}$ .

The one-class support vector machine, i.e. v-SVM, considers a hyperplane that separate the target datasets from the origin with a maximum margin [13]. The following minimization problem is carried out to reach at a solution for the parameter  $\alpha_i$ :

$$\min_{\alpha_{i,j}} \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j K(x_i, x_j) \quad (6)$$

$$\text{Subject to } 0 \leq \alpha_i \leq \frac{1}{Nv}, \sum_i \alpha_i = 1,$$

Where  $v$  is similar to  $\mathfrak{C}$  in SVDD, and plays the role of regularization term. A new test object  $x$  is evaluated by calculating the distance from the test object to the origin as follows:

$$\psi(x) = \begin{cases} target, & \text{if } \sum_i \alpha_i K(x_i, x) - \zeta \geq 0, \\ non - target, & \text{otherwise} \end{cases} \quad (7)$$

$\zeta$  is computed as  $\sum_j \alpha_j(x_i, x_i)$ , where  $x_i$  is an object vector, for which  $\alpha_i$  is not at the lower or upper bound.

**Nearest-Nearest Neighbor Data Description:** Uses the distances to the first nearest neighbor to approximate the local density of the target class [7, 17, 18]. A new test object  $x_i$  is evaluated by measuring the distance to the first nearest neighbor,  $NN(x)$ , in the target dataset. The estimated distance is then normalized by its nearest neighbor distance as given below:

$$\rho_{NN}(x) = \frac{\|x - NN^{tr}(x)\|}{\|NN^{tr}(x) - NN^{tr}(NN^{tr}(x))\|} \quad (8)$$

Where,  $NN^{tr}(x)$  is the nearest neighbor of object  $x$ . The classifier is defined based on a threshold  $\beta$  as:

$$\psi(x) = \begin{cases} target & \text{if } \rho_{NN}(x) \leq \beta \\ non - target & \text{if } \rho_{NN}(x) > \beta \end{cases} \quad (9)$$

**Minimum Spanning Tree Data Description:** Exploits the structure of minimum spanning tree to describe the target class [10]. A new test object is evaluated by calculating the distance from the object to the closest edge of the tree [9, 10]. The classifier is defined based on a specified threshold as given in equation 9.

#### 1.1.2.2. Density-Based Method

Density-based method estimate the probability density distribution of the target object, where a test object that lies in the high density region is regarded as normal and anomaly if it lies in a low density region [6, 11]. Different variants of density-based methods exist such as Gaussian, minimum covariance Gaussian, mixture of Gaussian, Parzen, Naïve Parzen, local outlier factor, and k-nearest neighbor [7].

**Gaussian, Minimum Covariance Gaussian, and Mixture of Gaussian Data Description:** Gaussian data descriptions describe the target data by assuming that the data is either normally distributed or mixture of a number of normal distributions [6, 7, 9, 12]. Gaussian and MCD Gaussian data description models uses Mahalanobis distance estimate as resemblance measure instead of density estimate [9]. Normal Gaussian data description defines the Mahalanobis distance from a new test object  $x$  to the training set  $X$  based on the mean and covariance matrix of the training set [6, 9]:

$$\rho_{Maha}(x) = (x - \mu^{tr})^T \Sigma^{-1} (x - \mu^{tr}),$$

where  $\mu^{tr}$  is training sample mean

$$\Sigma = \frac{1}{|X|} \sum_{x^k \in X} (x^k - \mu^{tr})(x^k - \mu^{tr}), \quad (10)$$

where  $|X|$  is the number of objects in the training dataset.

The classifier is defined as:

$$\psi(x) = \begin{cases} target & \text{if } \rho_{Maha}(x) \leq \beta \\ non - target & \text{if } \rho_{Maha}(x) > \beta \end{cases} \quad (11)$$

The threshold  $\beta$  is estimated by taking the user specified target error into consideration.

Minimum Covariance Gaussian data description is similar to Gaussian data description, except that the mean and covariance matrix is estimated using only a fraction of the target datasets that minimize the determinant of the covariance matrix [9, 19]. Mixture of Gaussian data description model define the target class using a linear combination of  $k$  Gaussian [9, 11, 20]. A new test object  $x$  is evaluated as follows:

$$\rho_{MOG}(x) = \sum_{i=1}^K P_i \exp(-(x - \mu_i)^T \Sigma_i^{-1} (x - \mu_i)) \quad (12)$$

$P_i$  and  $\Sigma_i$  are optimized using expectation minimization (EM) algorithm. The classifier is defined based on a specified threshold  $\beta$  as:

$$\psi(x) = \begin{cases} target & \text{if } \rho_{MOG}(x) \geq \beta \\ non - target & \text{if } \rho_{MOG}(x) < \beta \end{cases} \quad (13)$$

**Parzen and Naïve Parzen Data Description:** These models are non-parametric density estimators, which don't take into account any assumption about the underlying data distribution [7, 11]. The density is directly estimated from the training datasets using a mixture of kernels, most often a Gaussian kernel, centered on each individual training dataset, with diagonal covariance matrix  $\Sigma_i = hI$ . The smoothing parameter  $h$  characterizes the density estimate, where large values results in overestimate and small values results in noisy estimation. The optimal value of the smoothing parameter is computed based on the maximum likelihood on the training data using leave-one-out approach [6, 7, 12]. A new test object is evaluated as follows:

$$\rho_{Parzen}(x) = \sum_{i=1}^N \exp(-(x - x_i)^T h^{-2} (x - x_i)) \quad (14)$$

The classifier is defined as given in equation 13.

**K-Nearest Neighbor Data Description:** KNN describes the target class by approximating the local density of the training (target) datasets [12]. The distance to the dataset can be computed based on the distance to the  $k^{\text{th}}$  nearest neighbor, distance to the average of the  $k$ -nn's, or average squared distance to the  $k$ -nn's [9]. For example, considering the distance to the  $k^{\text{th}}$  nearest neighbor, an object is evaluated as a function of the score, which is the ratio of the distance from the object to its  $k^{\text{th}}$  nearest neighbors and the distance between the  $k^{\text{th}}$  nearest neighbor and its  $k^{\text{th}}$  nearest neighbors [12].

$$\rho_{Knn}(x_i, k) = \frac{d(x_i, NN_k(x_i))}{d(NN_k(x_i), NN_k(NN_k(x_i)))} \quad (15)$$

The classifier is defined based on a specified threshold as given in equation 9.

**Local Outlier Factor Data Description:** Like KNN, LOF considers the local density of an object to its respective neighbors, however, the distance is replaced by reachability distance [21, 22].

$$reach - dist_k(x_i \leftarrow x_j) = \max \left\{ d(x_j, NN_k(x_j)), d(x_i, x_j) \right\} \quad (16)$$



For an object  $x_i$ , the local reachability density is computed by taking the inverse average reachability distance from the set of  $x_i$ 's neighbors, which are located within the  $k$ -nearest neighbor distance around  $x_i$ :

$$lrd_k = \frac{1}{\sum_{x_j \in kNN(x_i)} reach-dist_k(x_i \leftarrow x_j)} / |kNN(x_i)| \quad (17)$$

For an object  $x_i$ , the degree of outlierness score, called LOF score, is evaluated by comparing its reachability density ( $lrd$ ) with its neighbors:

$$LOF_k(x_i) = \frac{\sum_{x_j \in kNN(x_i)} lrd_k(x_j)}{|kNN(x_i)|} \quad (18)$$

The classifier is defined based on a specified threshold as given in equation 9.

### 1.1.2.3. *Reconstruction-Based Method*

Reconstruction-based method make assumptions about the underlying data characteristics, which involves modelling of the data generating process by estimating the parameters during the training phase using the target objects [6, 17]. It is characterized by a set of prototypes/subspaces with minimal reconstruction errors. A test object is determined as either normal or anomaly based on the reconstruction error, which indicates how the test object fits to the model. Normal test objects usually generate minimum reconstruction error (closer fit) and anomalies generate high reconstruction errors. Reconstruction-based method includes different models; principal component analysis (PCA), self-organizing map (SOM), auto-encoder, and K-means, which mainly differ in their prototype/subspace definition, optimization principle, and the way the reconstruction error is used [7].

**Principal Component Analysis Data Description (PCA):** Computes the internal variance and external covariance structures of the target data in terms of set of principal components, which are a linear combinations of the original variables, to describe the data on a linear subspace [7, 9, 12]. The eigenvectors of the data covariance matrix  $\Sigma$  are used to define the corresponding subspace. Different PCA optimization techniques exist and yet eigenvalue decomposition is the simplest procedure to compute the eigenvectors of the target covariance matrix,  $\Sigma$  [7]. The number of basis vectors are computed depending on the fraction of variance the user intends to retain in the description. The projection is carried out as:

$$x_{projected} = W(W^T W)^{-1} W^T x, \quad (19)$$

where  $x_{projected}$ , and  $x$  are the new projected data and the original data respectively. Whereas  $W$  is a  $d \times k$  matrix containing  $k$  eigenvectors, and  $d$  represent the original feature space dimensionality. A new test object is evaluated based on the reconstruction error, which is the difference between the original test object  $x$  and its projection onto the space,  $x_{projected}$  [7].

$$\epsilon_{reconstruction} = \|x - x_{projected}\|^2 \quad (20)$$

The classifier is defined based on a threshold  $\beta$  as:

$$\psi(x) = \begin{cases} target & \text{if } \varepsilon_{reconstruction} \leq \beta \\ non - target & \text{if } \varepsilon_{reconstruction} > \beta \end{cases} \quad (21)$$

**Auto-Encoder Data Description:** Neural network contains of a series of interconnected neurons at one or several layers, where training updates the weights connecting each respective neuron [12]. Neural network is capable of learning a complex functional mapping between the input and output features. Auto-encoder is a special type of neural network that learns the internal structure of the data to reconstruct the input features at the output [12]. Therefore, in one-class classification, the difference between the input and output features is taken into consideration to characterized the target class [7]. The reconstruction error is computed as:

$$\varepsilon_{reconstruction} = \|x - f_{auto}(x)\|^2 \quad (22)$$

The classifier is similar to equation 21.

**K-Means and Self-Organizing Maps (SOM) Data Description:** Both are type of clustering methods, which relies on the assumption that the data can be clustered and described by a set of prototypes or codebook vectors  $\mu_i$  [6, 7, 12]. Often the nearest prototypes, measured in terms of Euclidian distance, are used to represent the target object. K-means data description describes the training (target) data by k number of clusters, where the average distance to the cluster center is minimized [6, 7]. The standard k-means clustering procedure is used to place the center of the clusters ( $\mu_i$ ) as follows:

$$\varepsilon_{im} = \sum_k (\min_i \|x_k - \mu_i\|^2) \quad (23)$$

Self-organizing map (SOM) is an unsupervised clustering method, where objects in the feature space are mapped into a space while retaining their distance and neighborhood relationships [7]. SOM performs a competitive learning so as to locate the position of the prototype vectors [12]. An update is carried out not only on the nearest prototype but also prototypes in the neighborhoods of the nearest prototype, which is specified by a predefined topology. However, the magnitude of update decreases as the distance increases and distant porotypes get smaller updates. A new test object is evaluated based on a reconstruction error, which is the difference between the test object and its closest cluster center (neuron) in either the K-means or SOM [7, 12]:

$$\varepsilon_{reconstruction} = \min_i \|x - \mu_i\|^2 \quad (24)$$

The classifier is similar to equation 21.

### 1.1.3. Unsupervised approach

Unsupervised approaches take unlabeled datasets as input and determine whether each objects in the datasets are normal or abnormal with respect to the entire dataset [21, 23]. There exist variety of unsupervised approaches in literatures, which can be categorized into statistical, nearest-neighbor based techniques, and cluster based approaches [21, 23]. In this paper, we have tested two nearest-neighbor based techniques (local density based methods), local outlier factor (LOF) and connectivity based outlier factor (COF), which compare the density of an object with its neighbors than the entire dataset [24]. The connectivity-based outlier factor and local outlier factor only differ in the

way the density is estimated for a given object [23, 25]. LOF exploits the Euclidian distance measure to select the k-nearest neighbors, which is valid only if the data is distributed spherically around the object. This method often fails in some condition, for example, when objects in the dataset have a direct linear correlation. In this regard, COF improves this drawback by computing the local density of the neighborhood based on shortest-path approach, also known as chaining distance, which is the minimum of the sum of all distances connecting all k-neighbors and the object [23, 25].

## Reference

1. Hawkins, D.M., *Introduction*, in *Identification of Outliers*. 1980, Springer Netherlands: Dordrecht. p. 1-12.
2. Chandola, V., Banerjee, A., and Kumar, V., *Anomaly detection: A survey*. ACM computing surveys (CSUR), 2009. 41(3): p. 15.
3. Dunning, T. and Friedman, E., *Practical machine learning: a new look at anomaly detection*. 2014: "O'Reilly Media, Inc."
4. Mehrotra, K.G., Mohan, C.K., and Huang, H., *Anomaly detection principles and algorithms*. 2017: Springer.
5. Khan, S.S. and Madden, M.G., *One-class classification: taxonomy of study and review of techniques*. The Knowledge Engineering Review, 2014. 29(3): p. 345-374.
6. Ding, X., Li, Y., Belatreche, A., and Maguire, L.P., *An experimental evaluation of novelty detection methods*. Neurocomputing, 2014. 135: p. 313-327.
7. Tax, D.M.J., *One-class classification: Concept learning in the absence of counter-examples*, in *Technische Universiteit Delft*. 2002.
8. Japkowicz, N., *Concept-Learning in the Absence of Counter-Examples: An Autoassociation-Based Approach to Classification*, in *Computer Science*. 1999, Graduate School-New Brunswick Rutgers, The State University of New Jersey: New Brunswick, New Jersey.
9. Tax, D.M.J., *DDTools, the data description toolbox for MATLAB, version 2.1*. 2. Delft University of Technology, Delft, Netherlands, 2015.
10. Juszczak, P., Tax, D.M.J., Pe, kalska, E., and Duin, R.P.W., *Minimum spanning tree based one-class classifier*. Neurocomputing, 2009. 72(7): p. 1859-1869.
11. Irigoiien, I., Sierra, B., and Arenas, C., *Towards application of one-class classification methods to medical data*. ScientificWorldJournal, 2014. 2014: p. 730712.
12. Mazhelis, O., *One-class classifiers: a review and analysis of suitability in the context of mobile-masquerader detection*. South African Computer Journal, 2006. 2006(36): p. 29-48.

13. Schölkopf, B., Williamson, R.C., Smola, A.J., Shawe-Taylor, J., and Platt, J.C., *Support Vector Method for Novelty Detection*. 1999. p. 582-588.
14. Tax, D.M.J. and Duin, R.P.W., *Support vector domain description*. Pattern Recognition Letters, 1999. 20(11): p. 1191-1199.
15. Tax, D.M.J. and Duin, R.P.W., *Support Vector Data Description*. Machine Learning, 2004. 54(1): p. 45-66.
16. Tax, D.M.J. and Laskov, P. *Online SVM learning: from classification to data description and back*. in *2003 IEEE XIII Workshop on Neural Networks for Signal Processing (IEEE Cat. No.03TH8718)*. 2003.
17. Ridder, D.d., Tax, D.M.J., and Duin, R.P.W. *An experimental comparison of one-class classification methods*. in *Proceedings of the 4th Annual Conference of the Advanced School for Computing and Imaging, Delft*. 1998.
18. Tax, D.M.J. and Duin, R.P.W. *Data description in subspaces*. in *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*. 2000.
19. Duin, R.P., Juszczak, P., Paclik, P., Pekalska, E., De Ridder, D., Tax, D.M., and Verzakov, S., *Prtools4. 1, a matlab toolbox for pattern recognition*. Delft University of technology, 2007. 2600.
20. Irigoien, I., Sierra, B., Arenas, C., and #xf3, *Towards Application of One-Class Classification Methods to Medical Data*. The Scientific World Journal, 2014. 2014: p. 7.
21. Swersky, L., Marques, H.O., Sander, J., Campello, R.J.G.B., and Zimek, A. *On the Evaluation of Outlier Detection and One-Class Classification Methods*. in *2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*. 2016.
22. Breunig, M.M., Kriegel, H.-P., Ng, R.T., and Sander, J., *LOF: identifying density-based local outliers*. SIGMOD Rec., 2000. 29(2): p. 93-104.
23. Goldstein, M. and Uchida, S., *A Comparative Evaluation of Unsupervised Anomaly Detection Algorithms for Multivariate Data*. PLoS One, 2016. 11(4): p. e0152173.
24. *Anomaly Detection Toolbox*. 2016 [cited 2020 3/25]; Available from: <http://dsmi-lab-ntust.github.io/AnomalyDetectionToolbox/>.
25. Tang, J., Chen, Z., Fu, A.W.-c., and Cheung, D.W. *Enhancing Effectiveness of Outlier Detections for Low Density Patterns*. in *Advances in Knowledge Discovery and Data Mining*. 2002. Berlin, Heidelberg: Springer Berlin Heidelberg.

## Appendix 2 - Detailed Description of the Models Input Features

### 1.1. Quadrants of wellness in people with type 1 diabetes

The four quadrants of wellness in people with type 1 diabetes, as shown in the **Figure 1**, tries to defines the state of BG dynamics (blood glucose levels) at any time  $t$  using carbohydrate, insulin and physical activity parameters. The first quadrant is called *carbohydrate action*, where the ratio of insulin to carbohydrate is small and the blood glucose levels are elevated (hyperglycemia). This is a normal response of blood glucose dynamics, since consumption of more carbohydrate can elevate blood glucose levels. The second quadrant is called *physical activity action*, where the ratio of insulin to carbohydrate is small but blood glucose levels drops (hypoglycemia). This is a normal response of blood glucose dynamics, since the action of physical activity can derive the patient into hypoglycemia regions even if the patient consumes. The third quadrant is called *insulin action*, where the ratio of insulin to carbohydrate is large (high insulin and low carbohydrate consumption and blood glucose levels drops (hypoglycemia). This is a normal response of blood glucose dynamics, since administration of high insulin with little carbohydrate consumption can derive the patient into hypoglycemia region. The fourth quadrant is called *effect of metabolic change*, where the ratio of insulin to carbohydrate is large (high insulin and low carbohydrate consumption but blood glucose levels are elevated (hyperglycemia). The patient experiences hyperglycemia despite injecting higher amount of insulin and consuming less carbohydrate. This quadrant is an abnormal response due to the effect of metabolic change incurred due to the incidence of infection and stress.

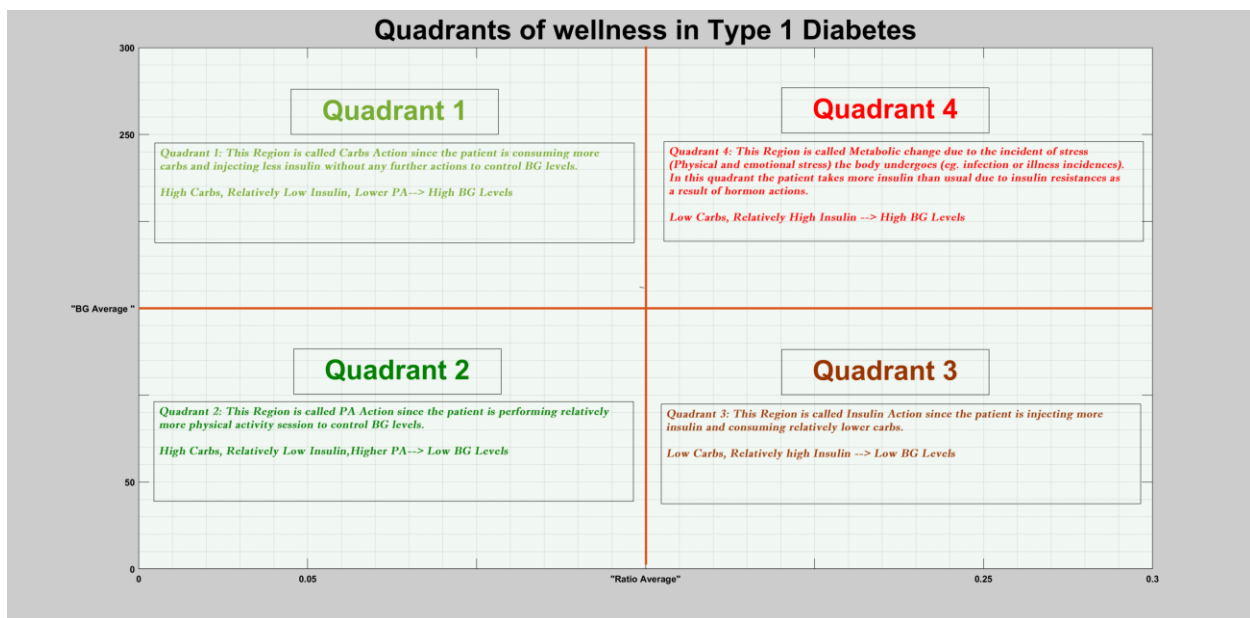


Figure 1: Quadrants of wellness in people with type 1 diabetes. The figure depicts the four possible scenarios of

different parameters: -carbohydrate action, insulin action, physical activity action, and abnormality due to metabolic change such as infection and stress.

## 1.2. Atypical Data

Atypical data signify when the data is not uniformly distributed and some region contain high density and other region contain sparse density<sup>1</sup>. In this kind of data, only the boundary between the target and outlier data is the same. However, the exact target density within the boundary could differ. As shown in the **Figure 2-11** below, the scatter plot of the input features, i.e. blood glucose levels vs. insulin to carbohydrate ratio, depicts similar characteristics to the nature of atypical dataset. In this kind of datasets, boundary and domain based method are more preferable than the others. Typically, density methods such as parzen density could suffer in performance degradation to this kind of dataset.

## 1.3. Description of Input Features

The input features, i.e. average blood glucose levels vs. insulin (bolus) to carbohydrate ratio, used for model training and testing are selected in accordance with the description provided in Woldaregay et al.<sup>2</sup>. The **Figures 2-11** depicts the scatter plot of the input features. The data are smoothed using a moving average filter to remove short term noises. For both daily and hourly analysis a moving average filter window size of two days or forty-eight hours were used respectively. Understanding of the data characteristics is essential to select the optimal anomaly detection model to better capture the data distribution during normal situations. The scatter plots of the input features presented in this section incorporates the data of ten different patient years under free living conditions. Five patient years depicting regular years without any significant infection incidences and five patient years with at least one or more incidences of infections<sup>1</sup>.

### 1.3.1. Description of Input Features During the Normal Patient Years

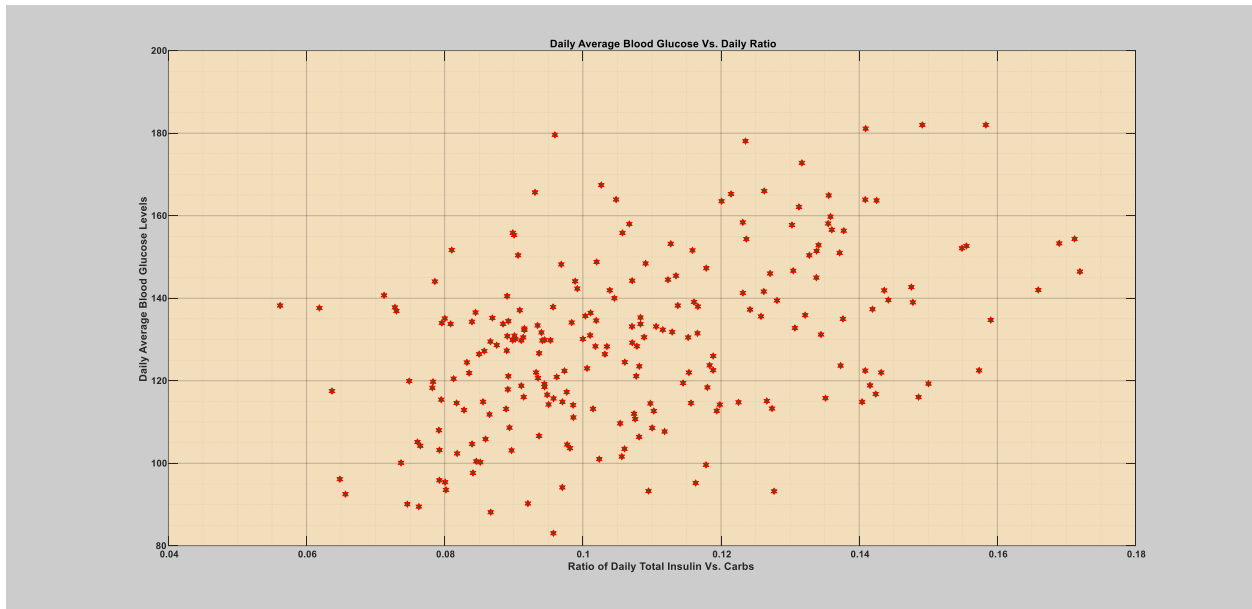
During the normal patients, the input features are characterized to be bounded with similar values of insulin to carbohydrate ratio<sup>1</sup>. However, from the scatter plot it appears to be a typical in distribution containing regions with high density and low density. The challenge with such kind of data is varying density and rare events that are still normal.

- **The First Patient Year (Normal year)**

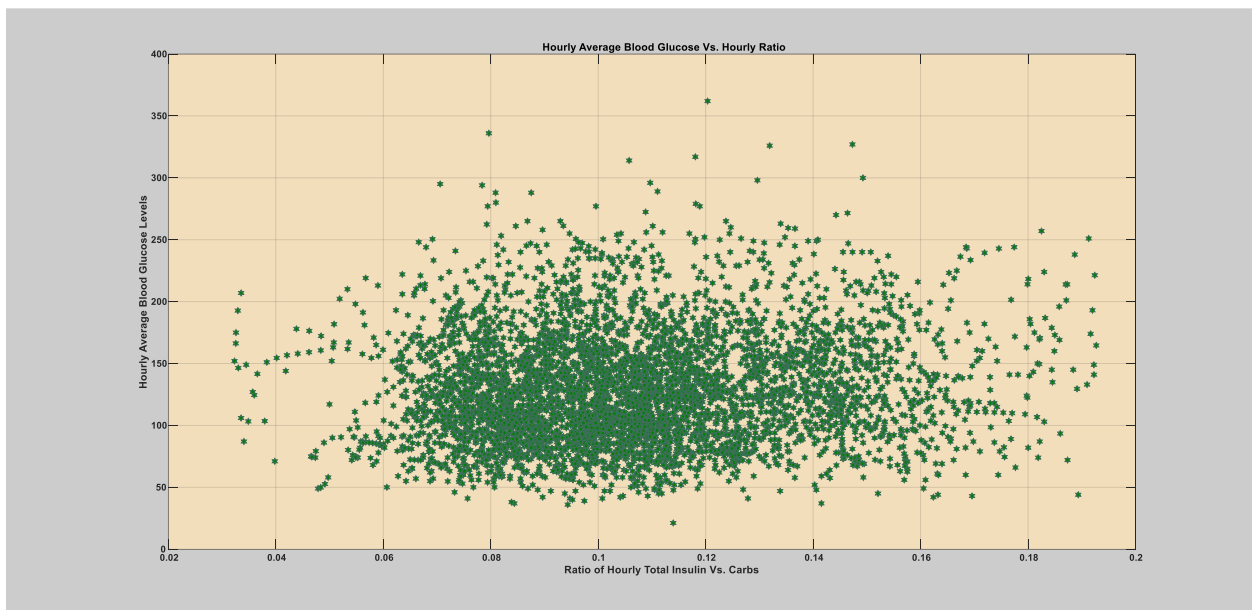
---

<sup>1</sup> Tax, D.M.J., *One-class classification: Concept learning in the absence of counter-examples*, in *Technische Universiteit Delft*. 2002.

<sup>2</sup> Woldaregay, A.Z., Årsand, E., Albers, D., Launonen, I., Holubová, A., and Hartvigsen, G., *Towards Detecting Infection Incidences in People with Type 1 Diabetes Using Self-Recorded Data: A Novel Framework for a Digital Infectious Disease Detection Mechanism*. JMIR Preprints, 2020. 26/03/2020:18911.



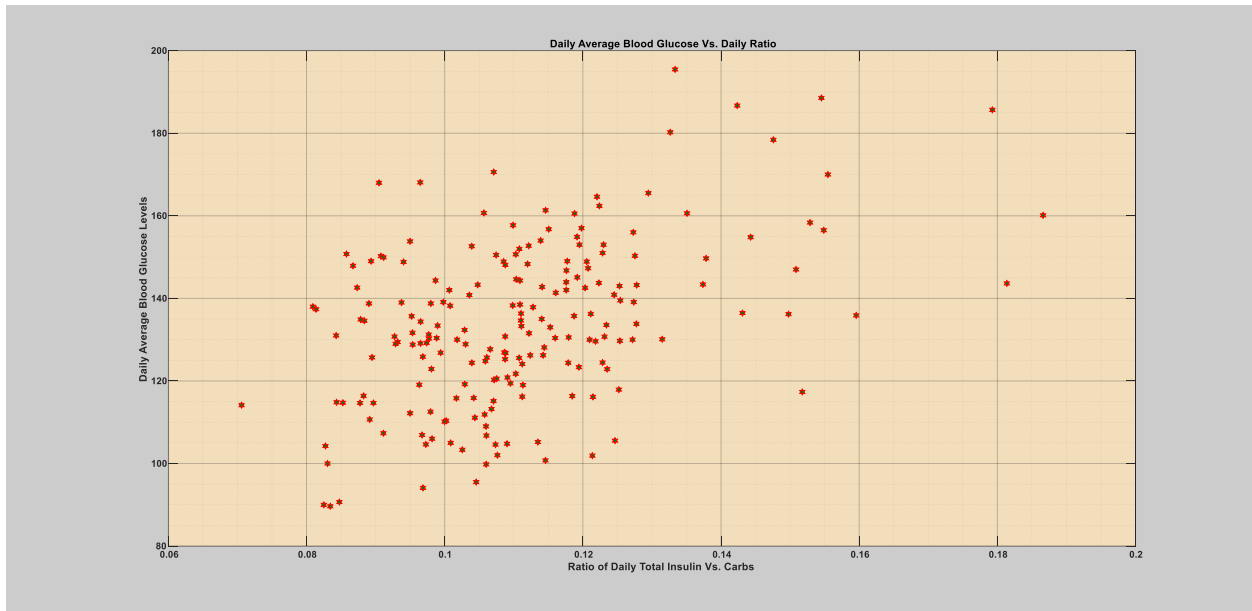
a) Daily average blood glucose levels vs. total insulin (bolus) to carbohydrate ratio.



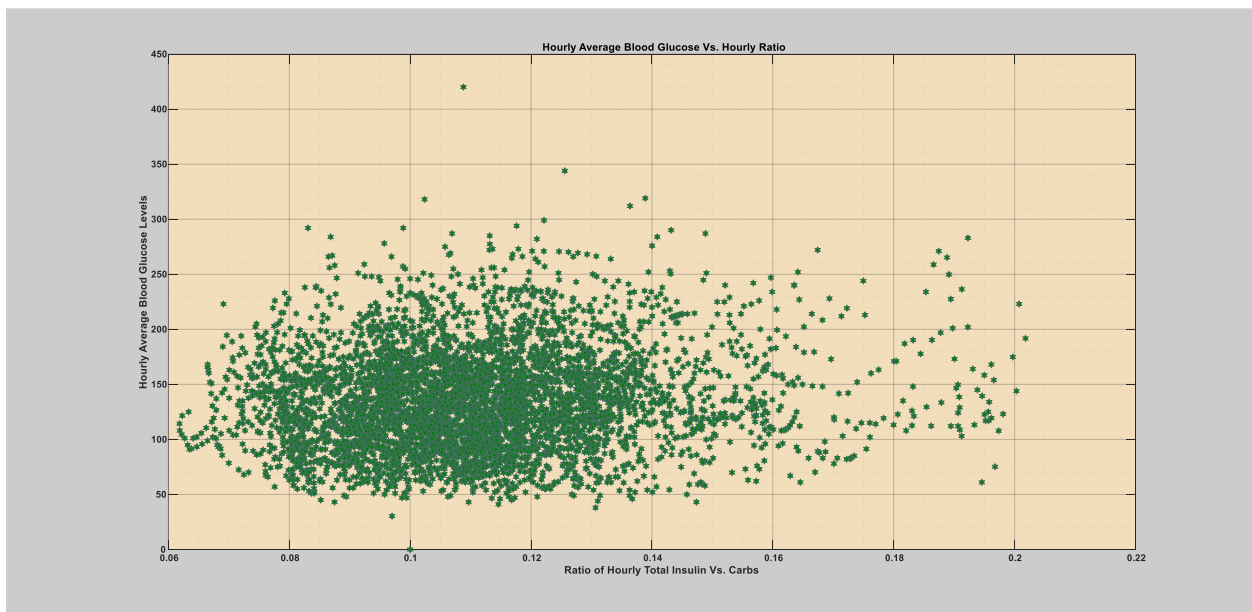
b) Hourly average blood glucose levels vs. total insulin (bolus) to carbohydrate ratio.

**Figure 2:** The first patient year, where there is no incidence of acute infections. Figure (a & b) depicts the daily and hourly scatter plot of the input features.

- **The Second Patient Year (Normal year)**



a) Daily average blood glucose levels vs. total insulin (bolus) to carbohydrate ratio.

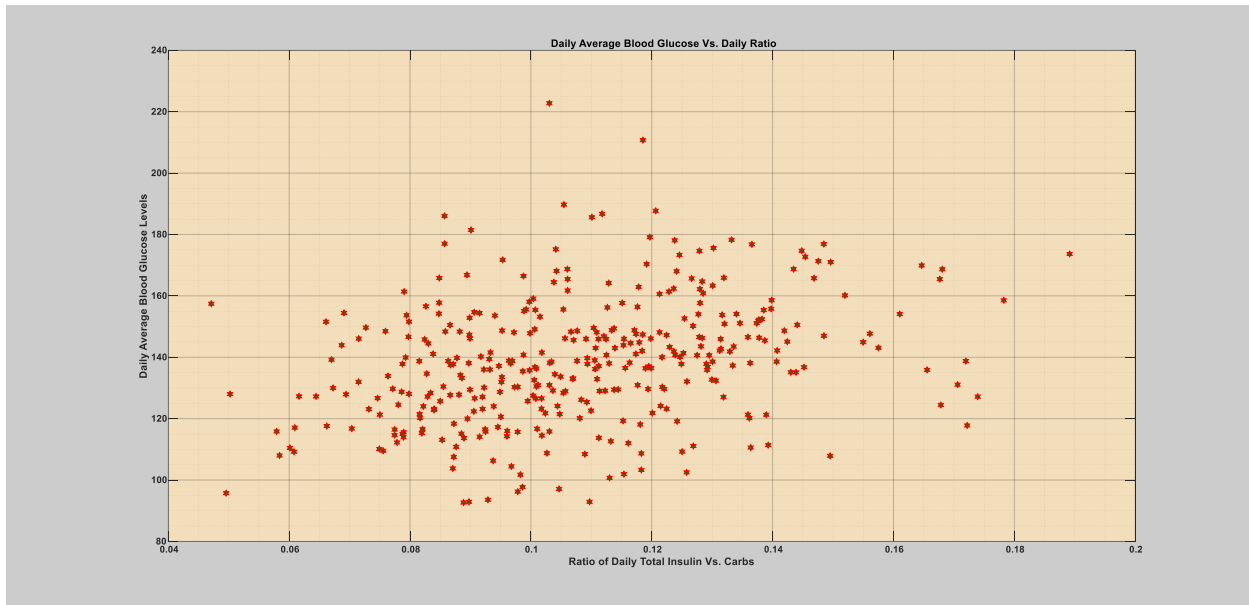


b) Hourly average blood glucose levels vs. total insulin (bolus) to carbohydrate ratio.

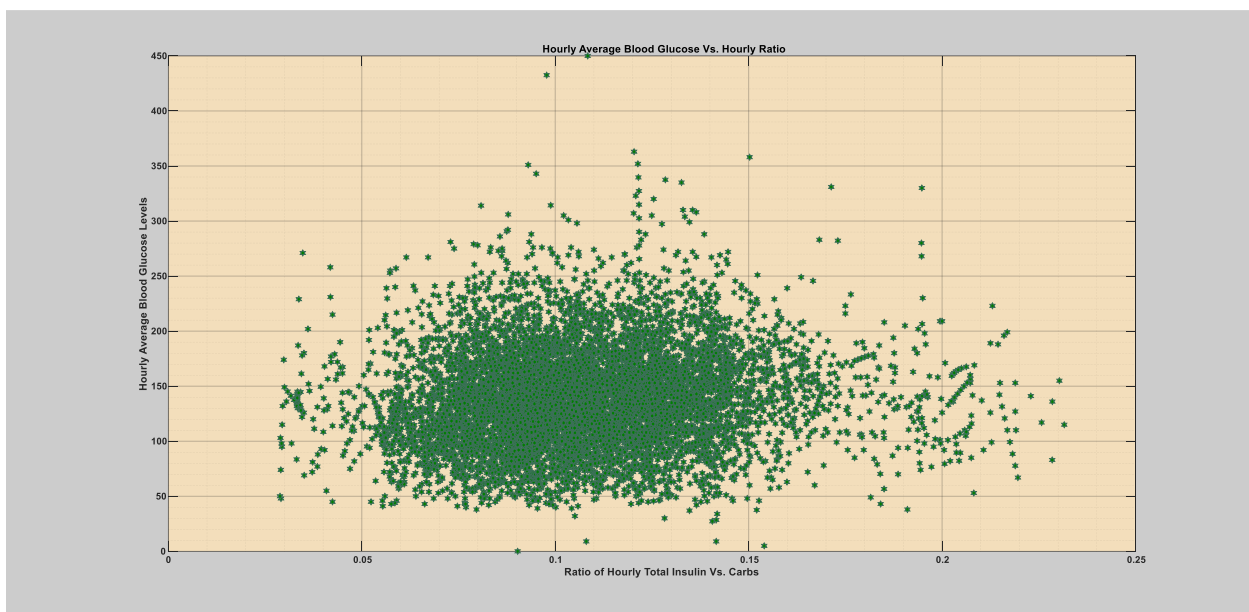
**Figure 3:** The second patient year, where there is no incidence of acute infections. Figure (a & b) depicts the daily and hourly scatter plot of the input features.

- **The Third Patient Year (Normal year)**





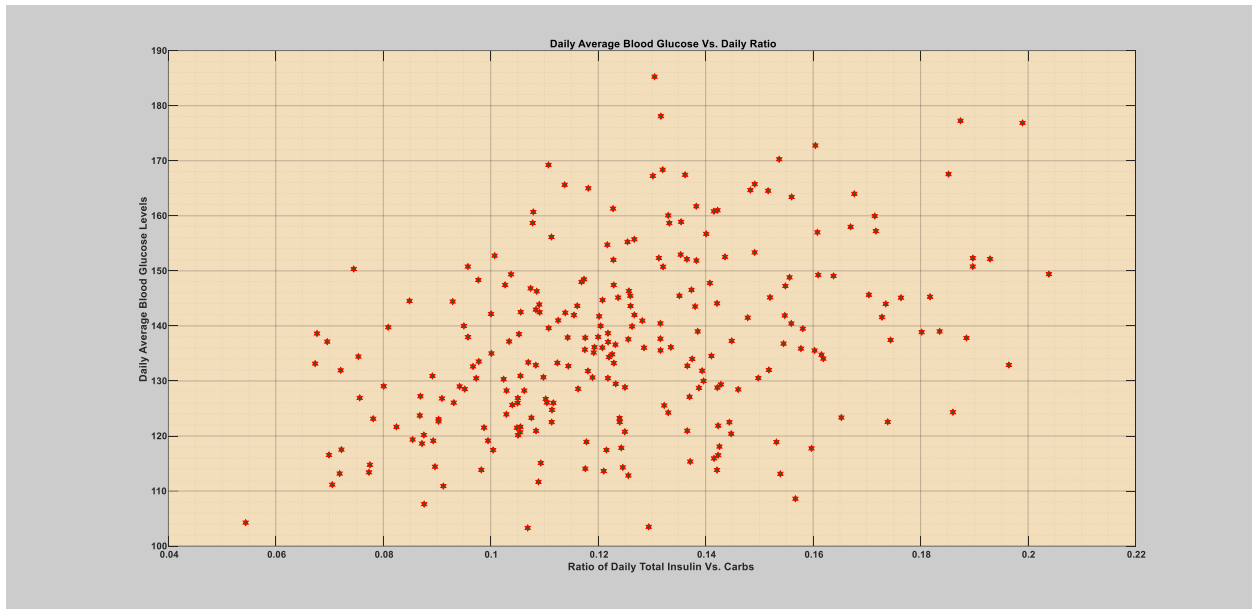
a) Daily average blood glucose levels vs. total insulin (bolus) to carbohydrate ratio.



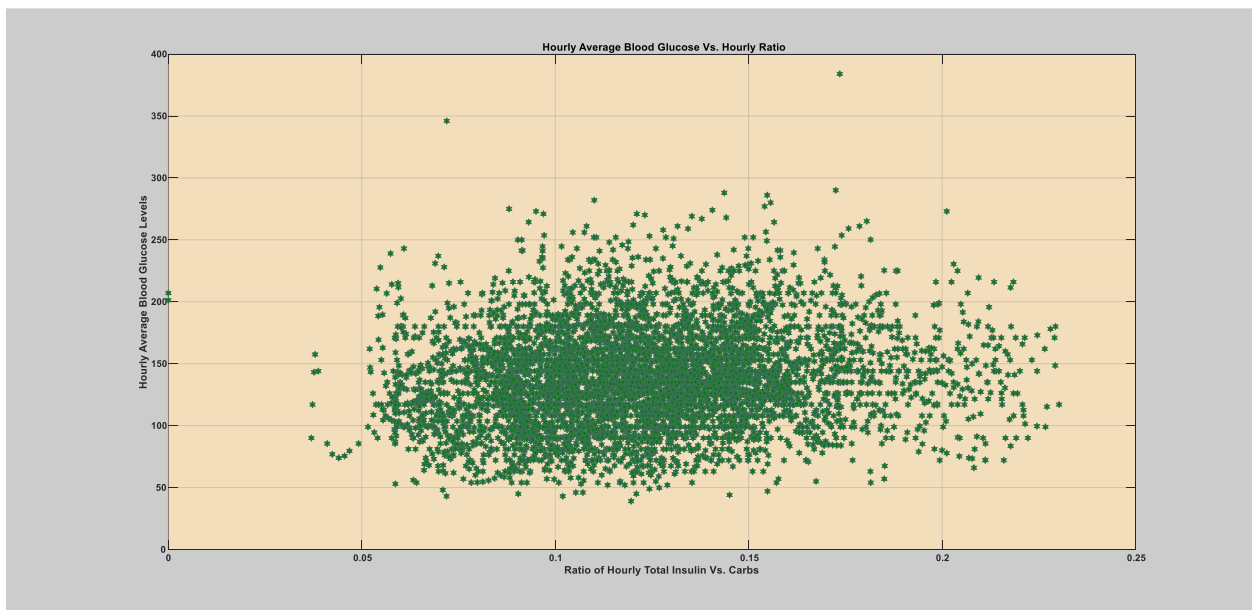
b) Hourly average blood glucose levels vs. total insulin (bolus) to carbohydrate ratio.

Figure 4: The third patient year, where there is no incidence of acute infections. Figure (a & b) depicts the daily and hourly scatter plot of the input features.

- **The Fourth Patient Year (Normal year)**



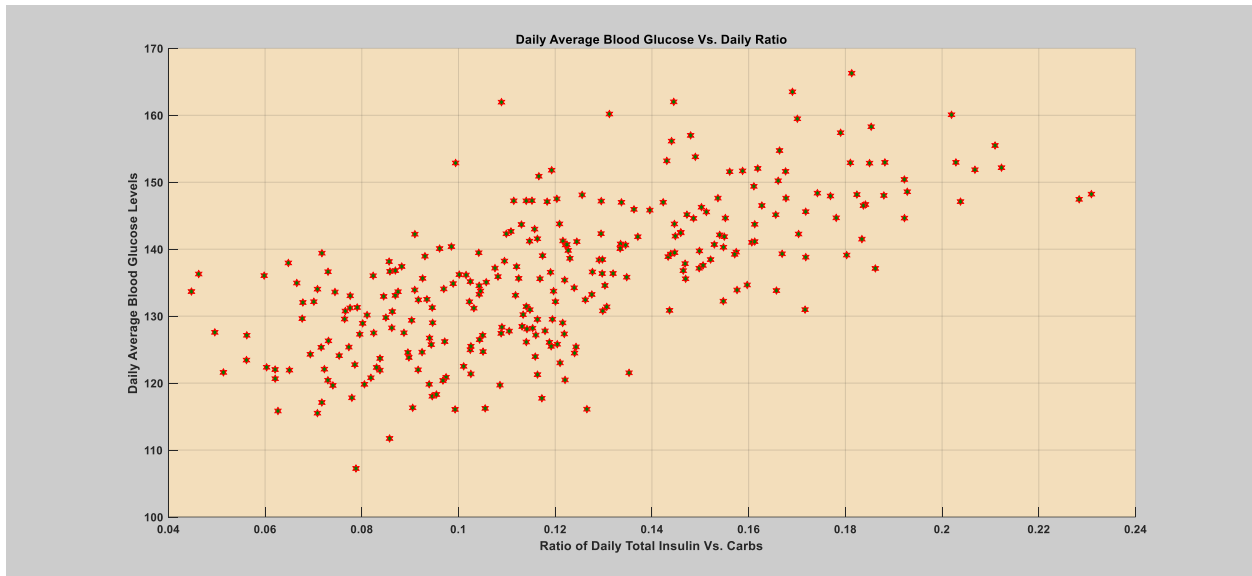
a) Daily average blood glucose levels vs. total insulin (bolus) to carbohydrate ratio.



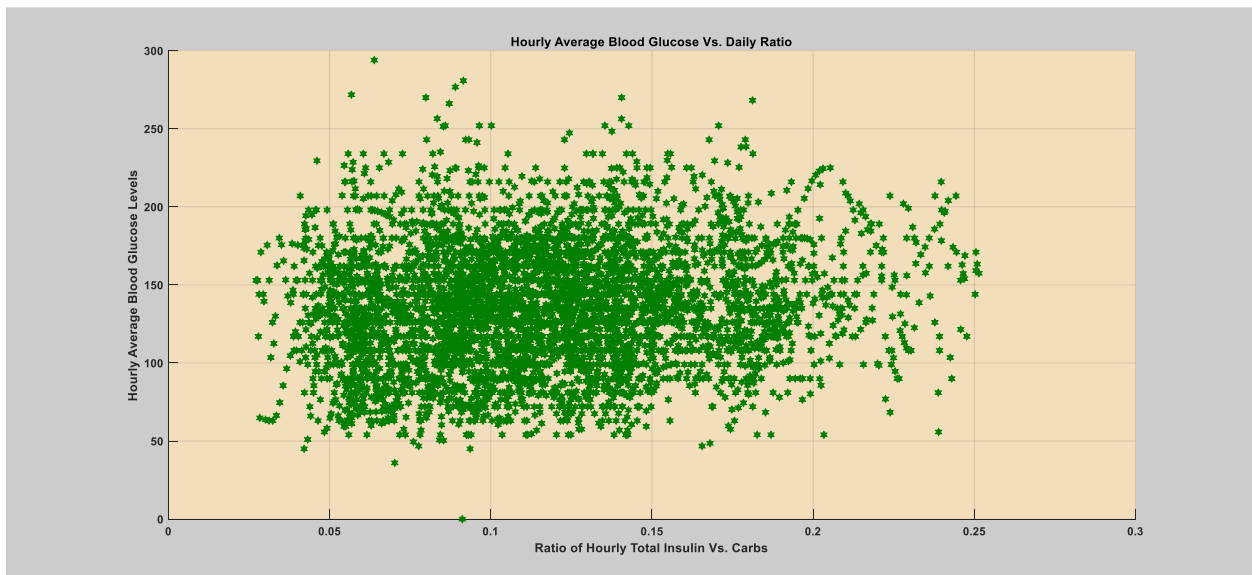
b) Hourly average blood glucose levels vs. total insulin (bolus) to carbohydrate ratio.

**Figure 5:** The fourth patient year, where there is no incidence of acute infections. Figure (a & b) depicts the daily and hourly scatter plot of the input features.

- **The Fifth Patient Year (Normal year)**



a) Daily average blood glucose levels vs. total insulin (bolus) to carbohydrate ratio.



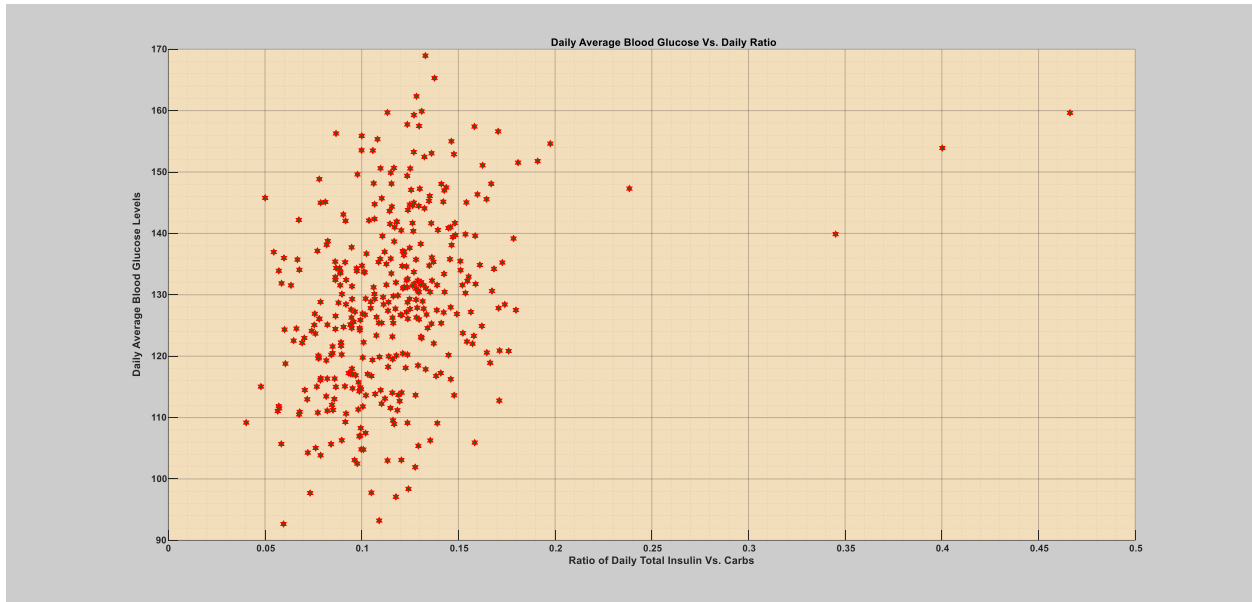
b) Hourly average blood glucose levels vs. total insulin (bolus) to carbohydrate ratio.

**Figure 6:** The fifth patient year, where there is no incidence of acute infections. Figure (a & b) depicts the daily and hourly scatter plot of the input features.

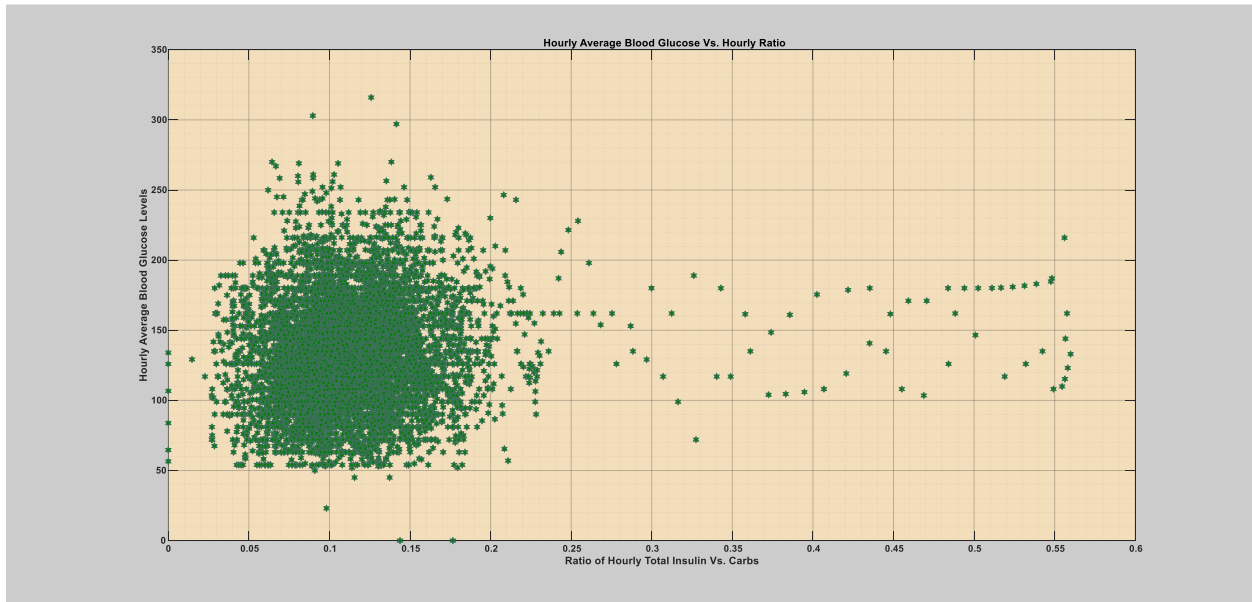
### 1.1.1. Description of Input Features During the Patient Years with Infection incidences

As can be seen from the scatter plot of the input features, the anomalies from the patient years are visible in both the daily and hourly cases. As described for the normal patient years above, the challenge is mostly modelling the normal portion of the data while minimizing false alarms. There are rare events that are normal response of the blood glucose dynamics and the optimal model is the one that captures those rare events along with the entire normal days.

- **The Six Patient Year (flu)**



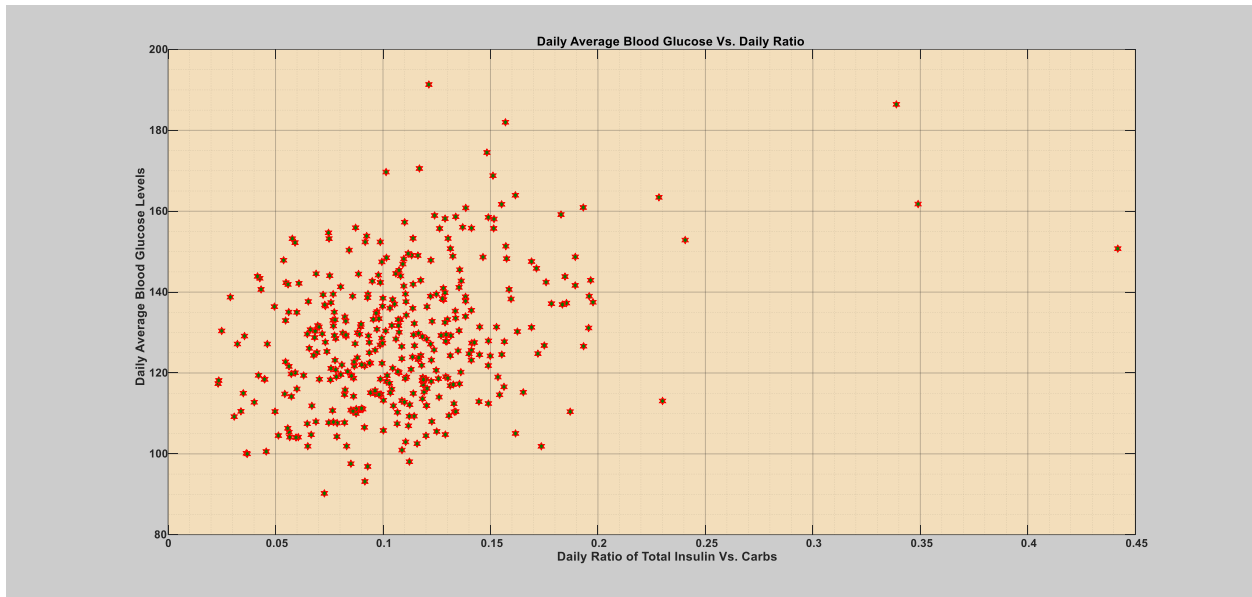
a) Daily average blood glucose levels vs. total insulin (bolus) to carbohydrate ratio.



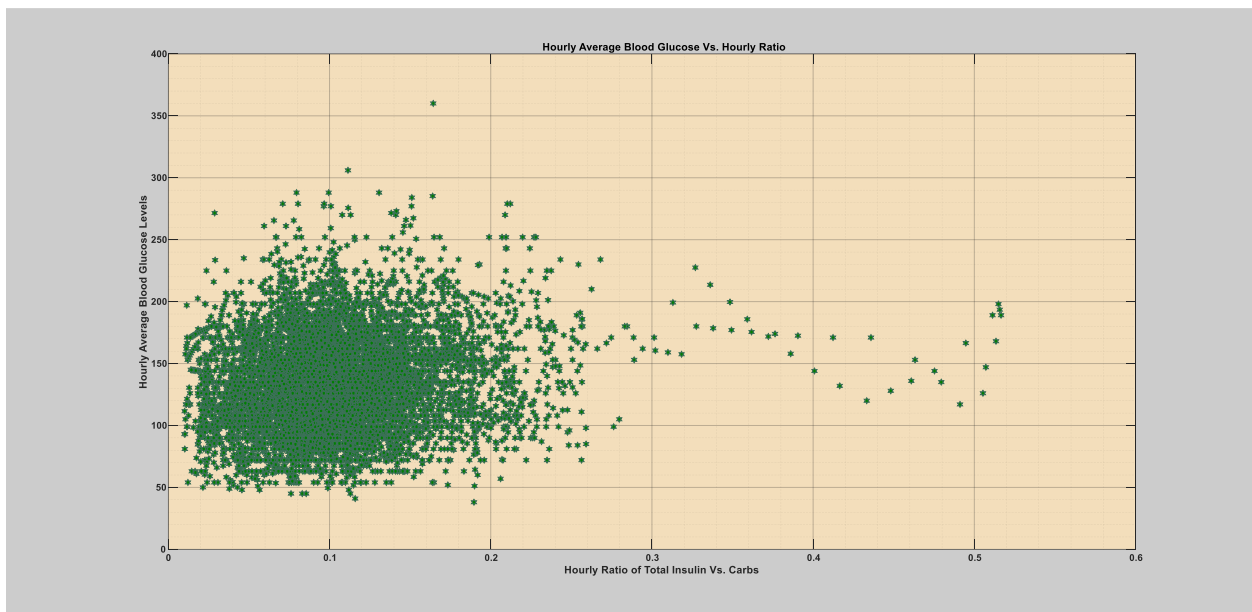
b) Hourly average blood glucose levels vs. total insulin (bolus) to carbohydrate ratio.

Figure 7: The six patient year, where the patient was infected with influenza (flu). Figure (a & b) depicts the daily and hourly scatter plot of the input features.

- **The Seventh Patient Year (flu)**



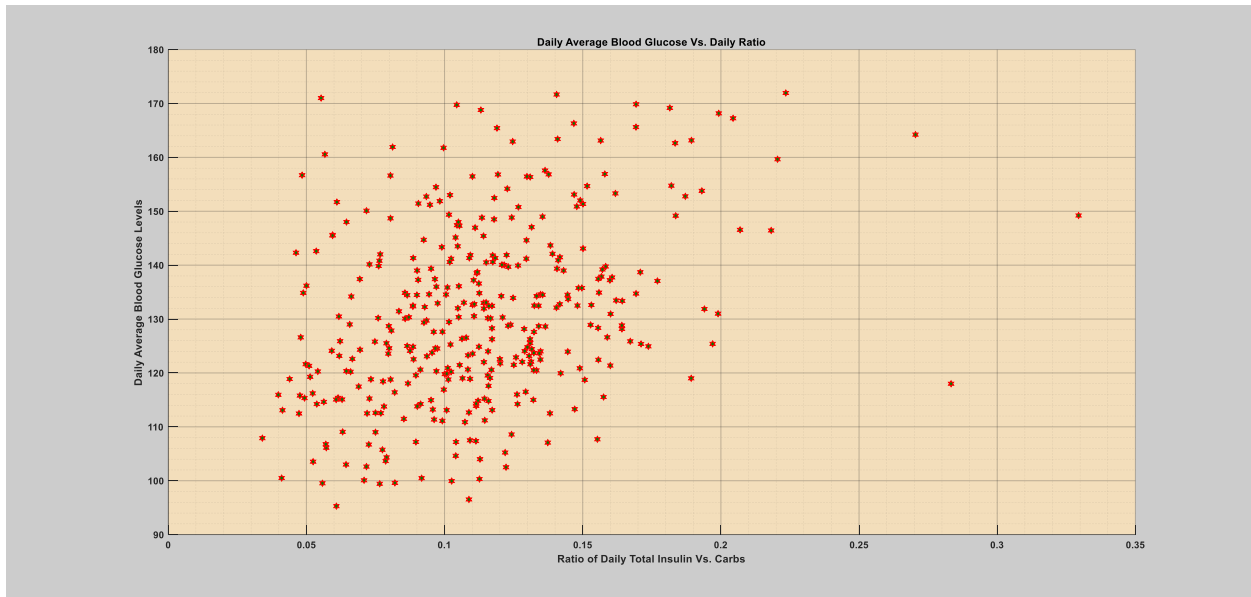
a) Daily average blood glucose levels vs. total insulin (bolus) to carbohydrate ratio.



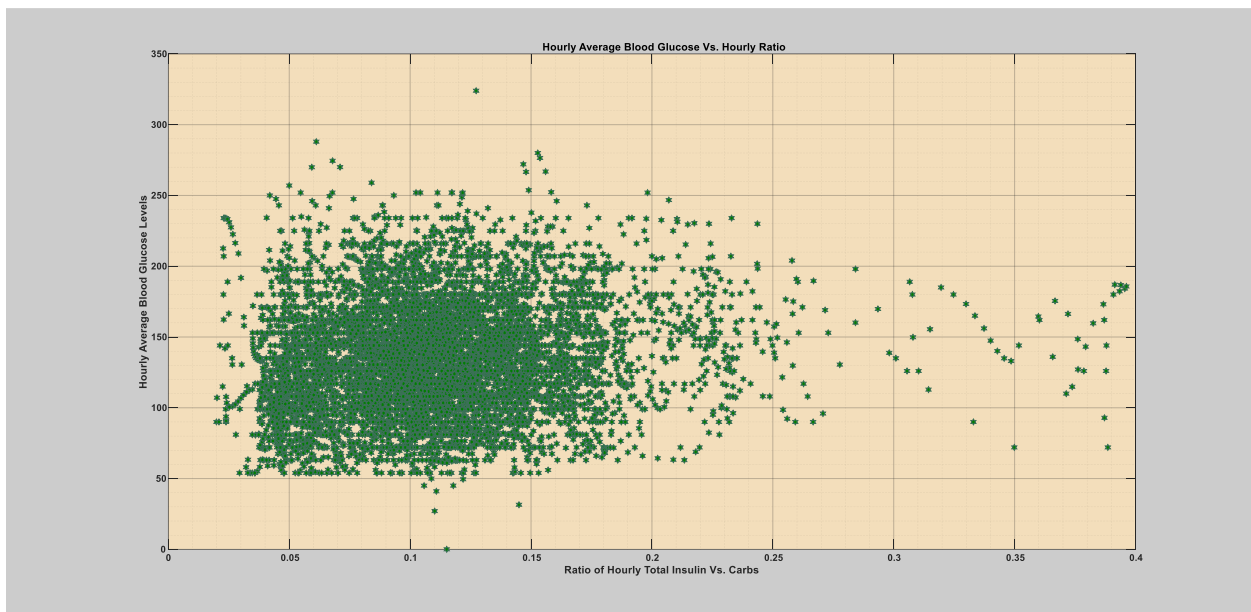
b) Hourly average blood glucose levels vs. total insulin (bolus) to carbohydrate ratio.

Figure 8: The seventh patient year, where the patient was infected with influenza (flu). Figure (a & b) depicts the daily and hourly scatter plot of the input features.

- The Eighth Patient Year (flu)



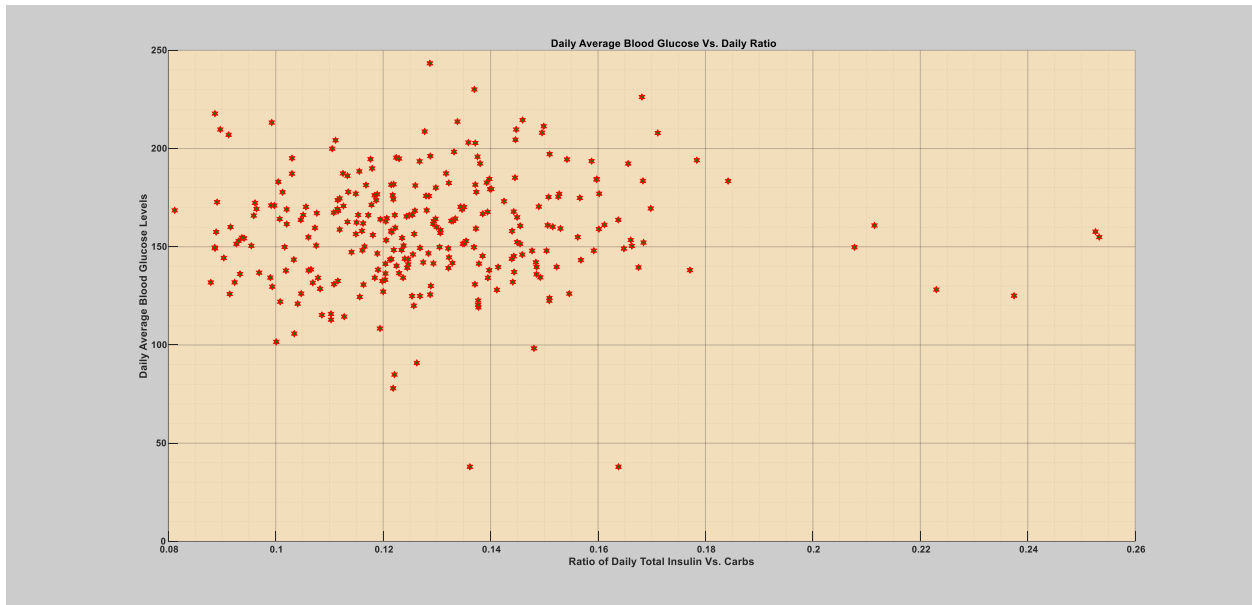
a) Daily average blood glucose levels vs. total insulin (bolus) to carbohydrate ratio.



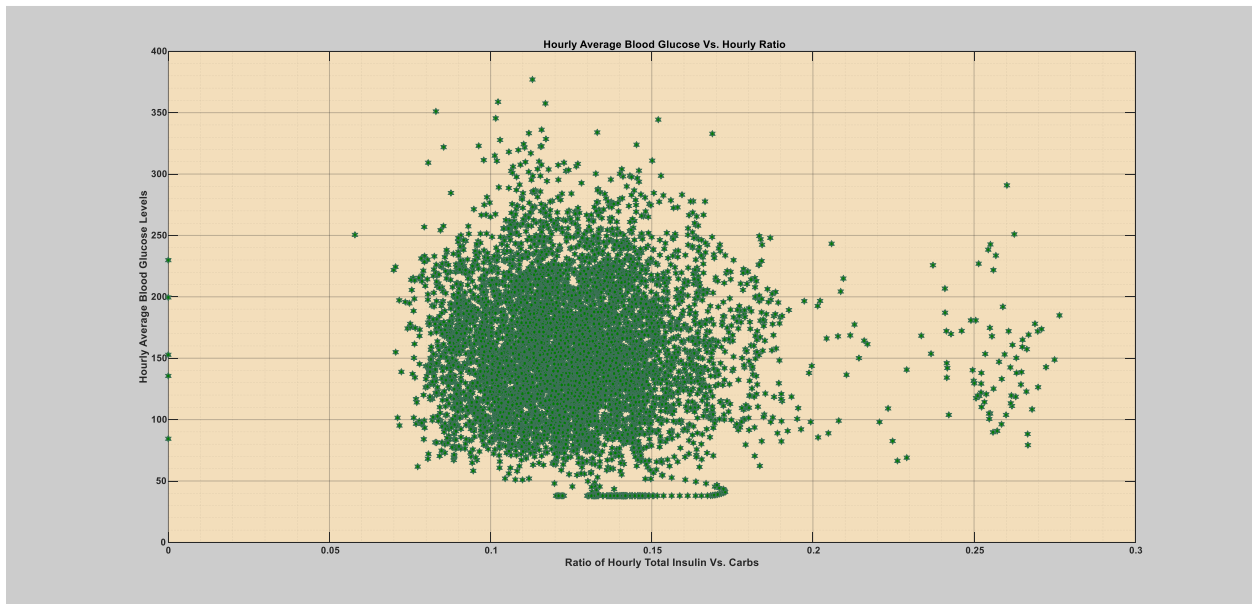
b) Hourly average blood glucose levels vs. total insulin (bolus) to carbohydrate ratio.

Figure 9: The eighth patient year, where the patient was infected with influenza (flu). Figure (a & b) depicts the daily and hourly scatter plot of the input features.

- **The Ninth Patient Year (flu)**



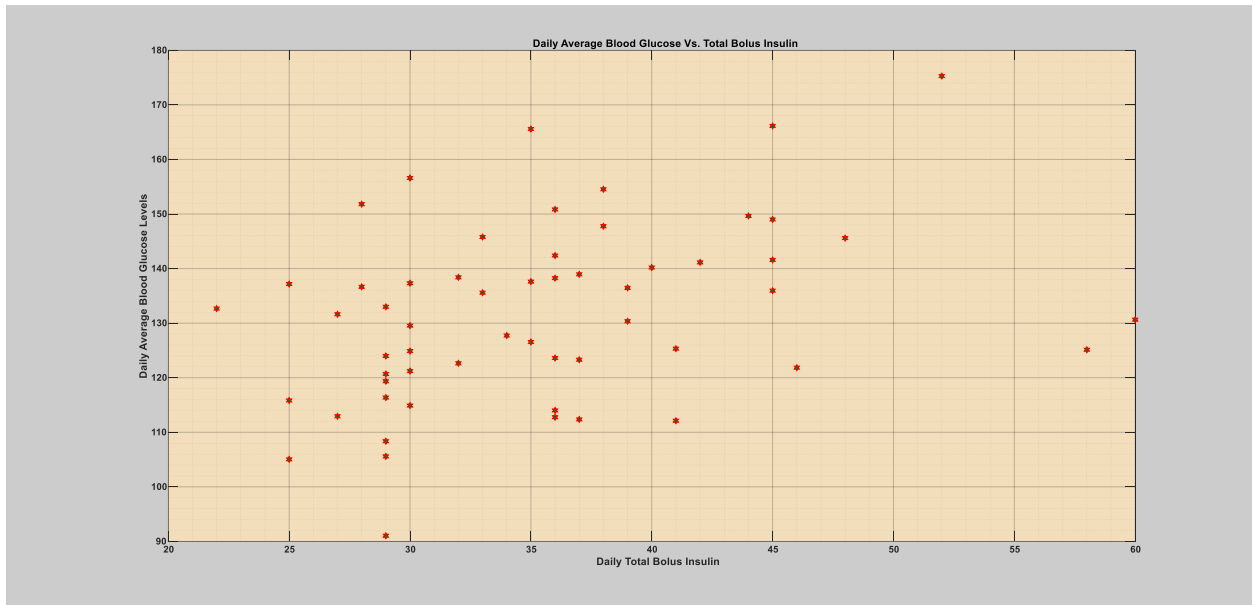
a) Daily average blood glucose levels vs. total insulin (bolus) to carbohydrate ratio.



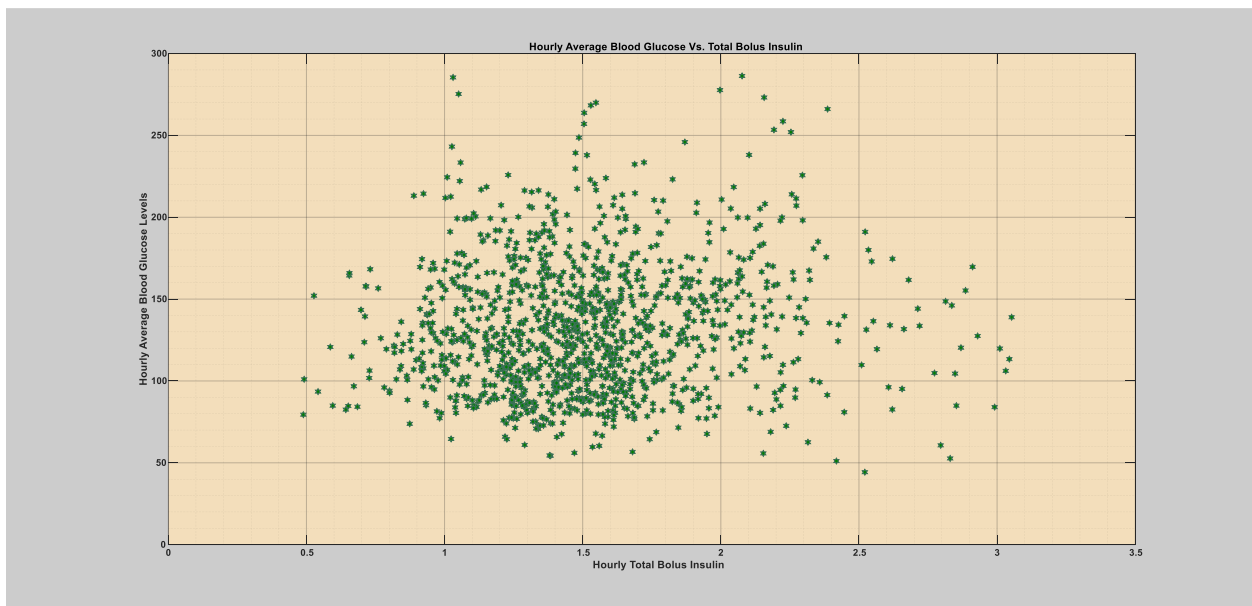
b) Hourly average blood glucose levels vs. total insulin (bolus) to carbohydrate ratio.

**Figure 10:** The ninth patient year, where the patient was infected with influenza (flu) and light and mild common cold without fever. Figure (a & b) depicts the daily and hourly scatter plot of the input features.

- **The Tenth Patient Year (Flu)**



a) Daily average blood glucose levels vs. total insulin (bolus).



b) Hourly average blood glucose levels vs. total insulin (bolus).

**Figure 11:** The tenth patient year, where the patient was infected with influenza (flu). Figure (a & b) depicts the daily and hourly scatter plot of average blood glucose levels vs. total insulin (bolus).





## Appendix 3-Score Plot of the Models for Each Patient Year

The evaluations of the models were carried out based on each patient year depending on two specified time-window; hourly and daily. The data were smoothed using a moving average with a window size of two days (forty-eight hours). The models were trained using a target dataset (regular/normal days) and tested on the whole patient year containing both the target (regular/normal days) and non-target (infection period). A training sample size of 120 days, which are randomly selected from the patient year excluding the infection period, were used to train the models and the score of each model for the whole patient year were plotted. Regarding the unsupervised methods, the whole patient year was presented during testing. The portion rejected by the models in all the figures represent the infection period.

### A. Semi-supervised (One-class classifiers)

#### 1. Daily

##### 1.1. The First Case of Infection (flu)

##### 1.1.1. Boundary and Domain-Based Method

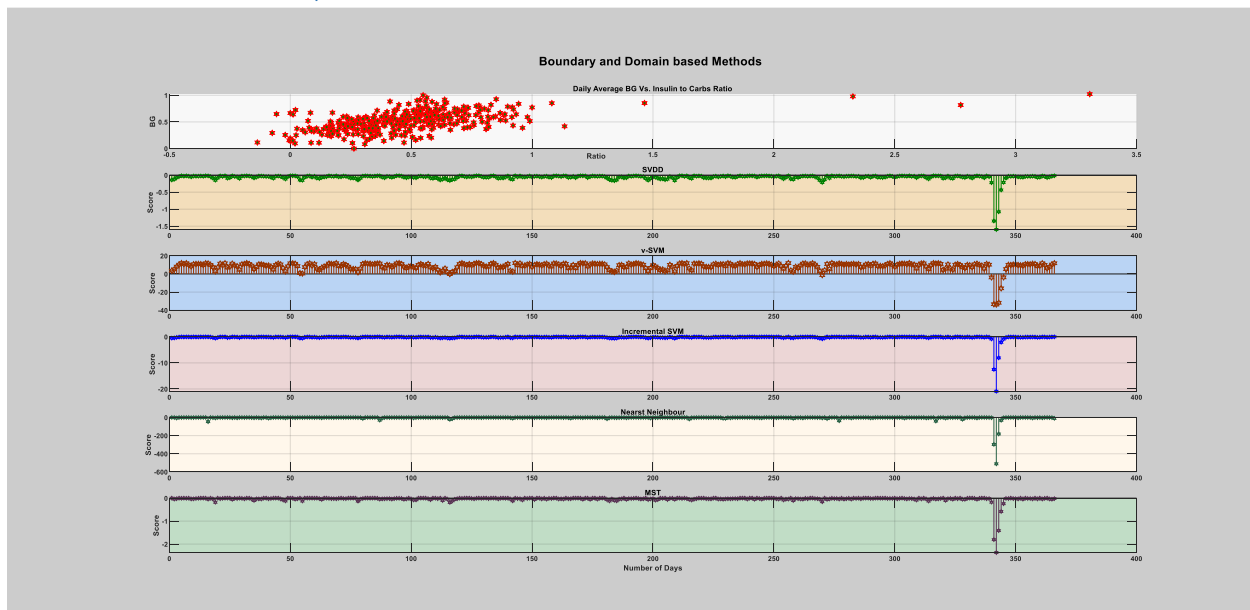


Figure 1: The score of the boundary and domain-based method on the whole patient year.

1.1.2. Density-Based Method

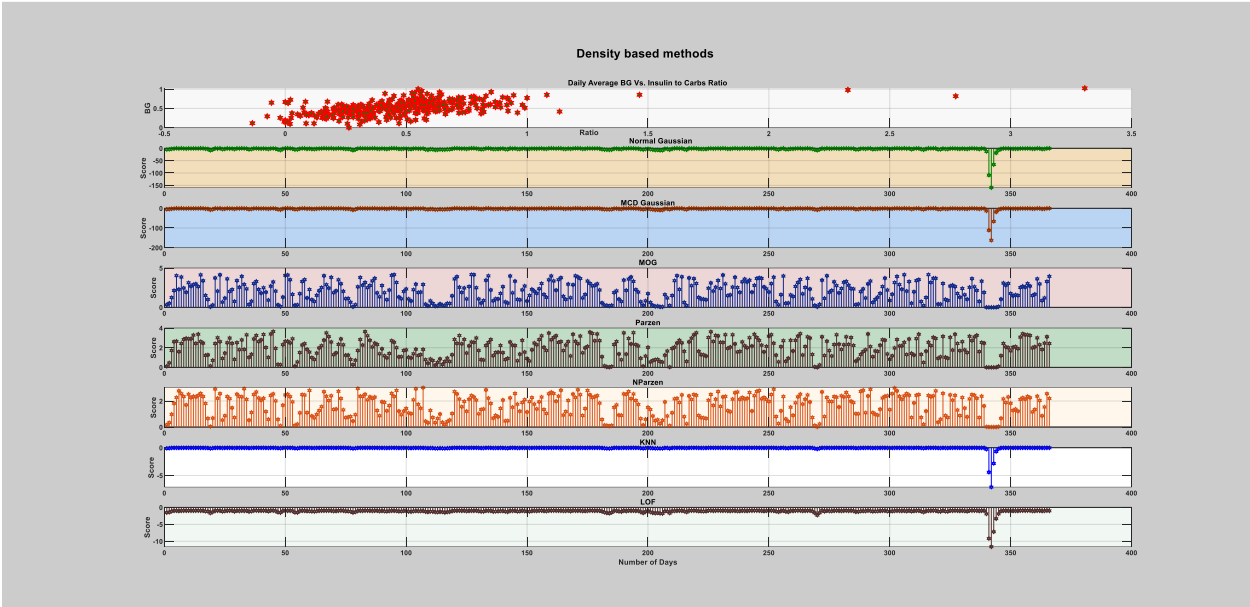


Figure 2: The score of the density-based method on the on the whole patient year.

1.1.3. Reconstruction-Based Method

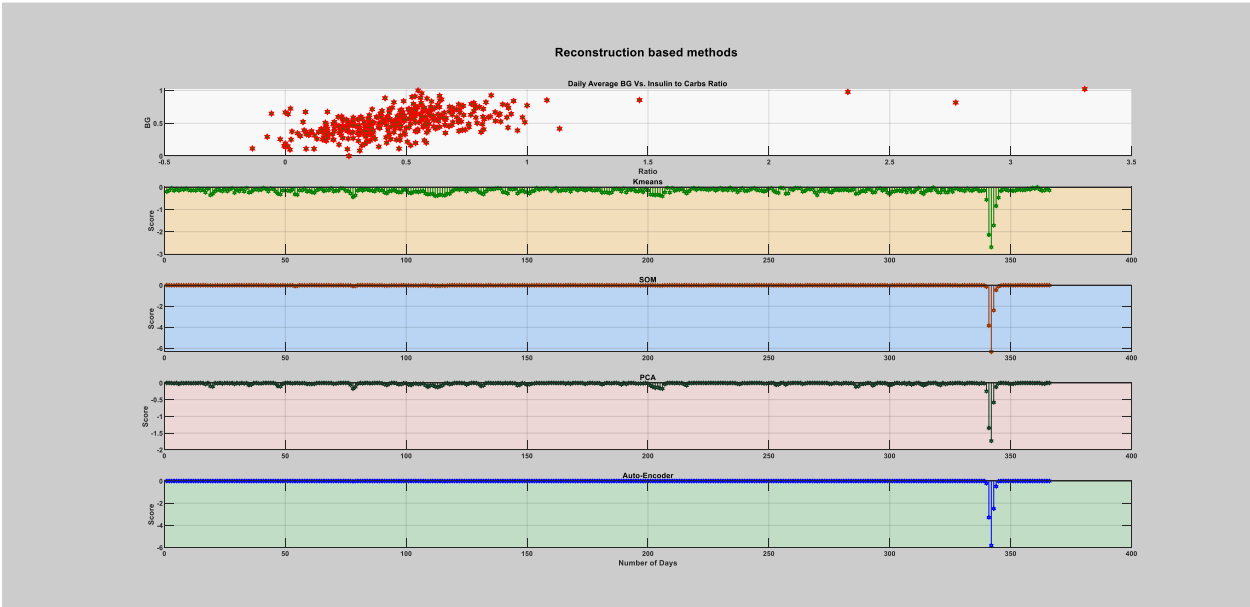


Figure 3: The score of the reconstruction-based method on the whole patient year.

## 1.2. The Second Case of Infection (flu)

### 1.2.1. Boundary and Domain-Based Method

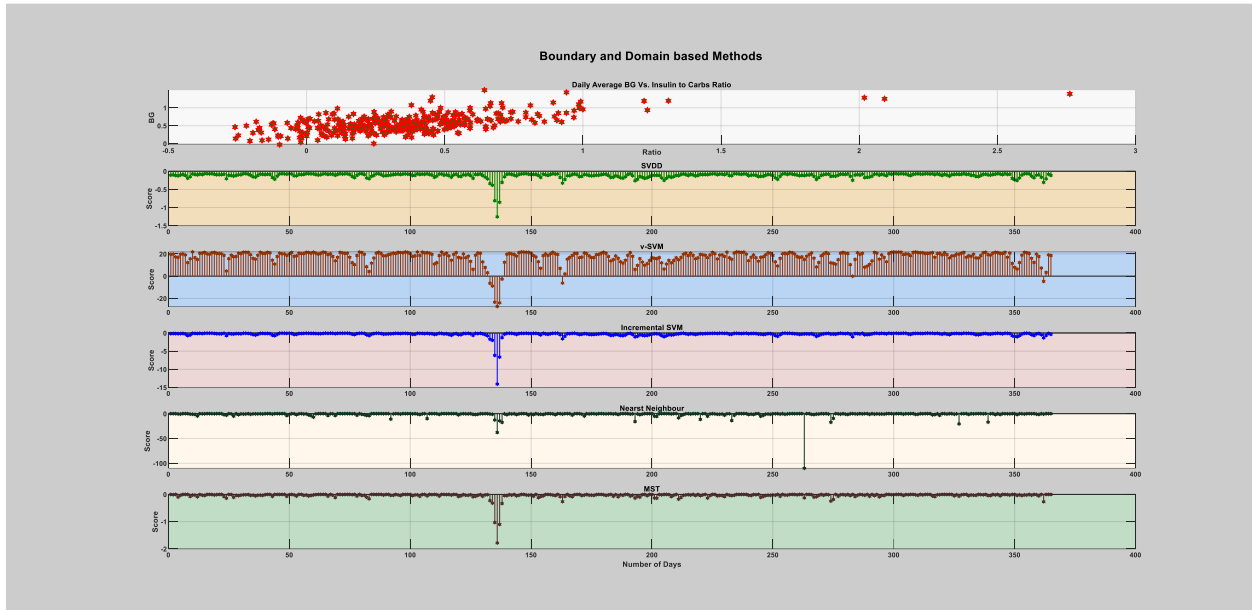


Figure 4: The score of the boundary and domain-based method on the whole patient year.

### 1.2.2. Density-Based Method

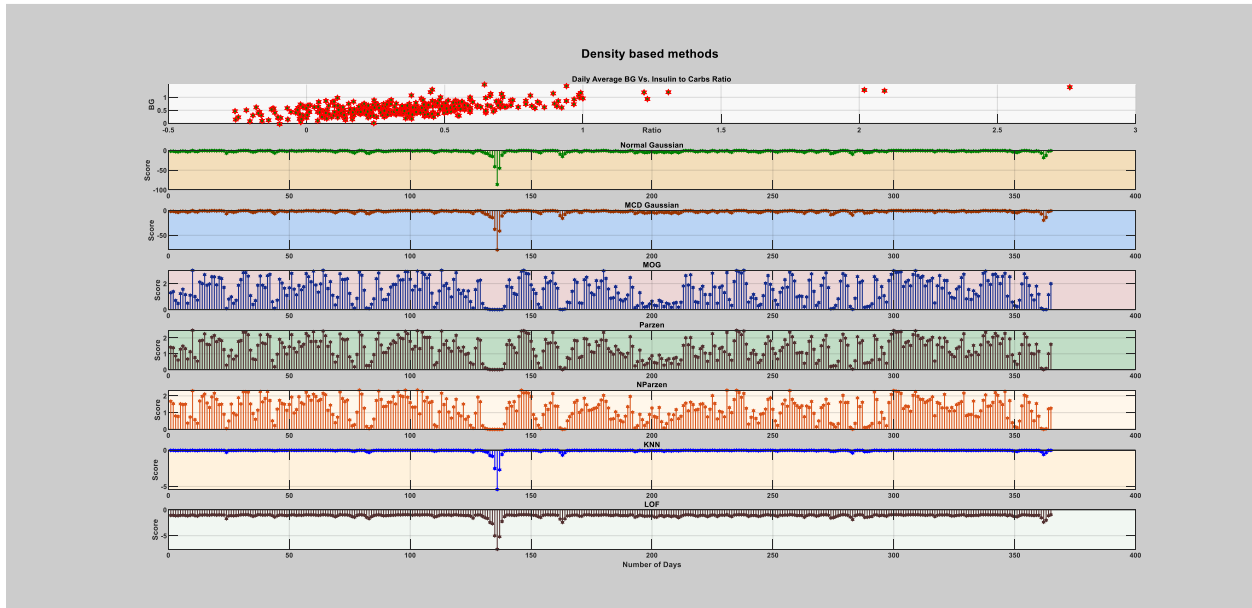


Figure 5: The score of the density-based method on the on the whole patient year.

1.2.3. Reconstruction-Based Method

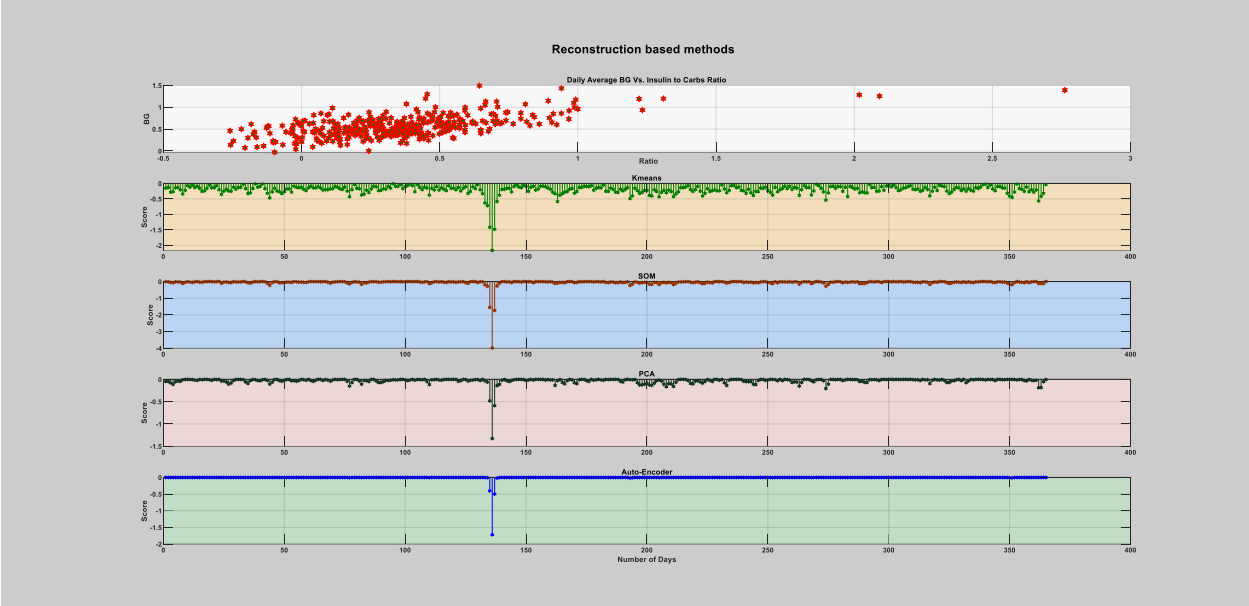


Figure 6: The score of the reconstruction-based method on the whole patient year.

1.3. The Third Case of Infection (flu)

1.3.1. Boundary and Domain-Based Method

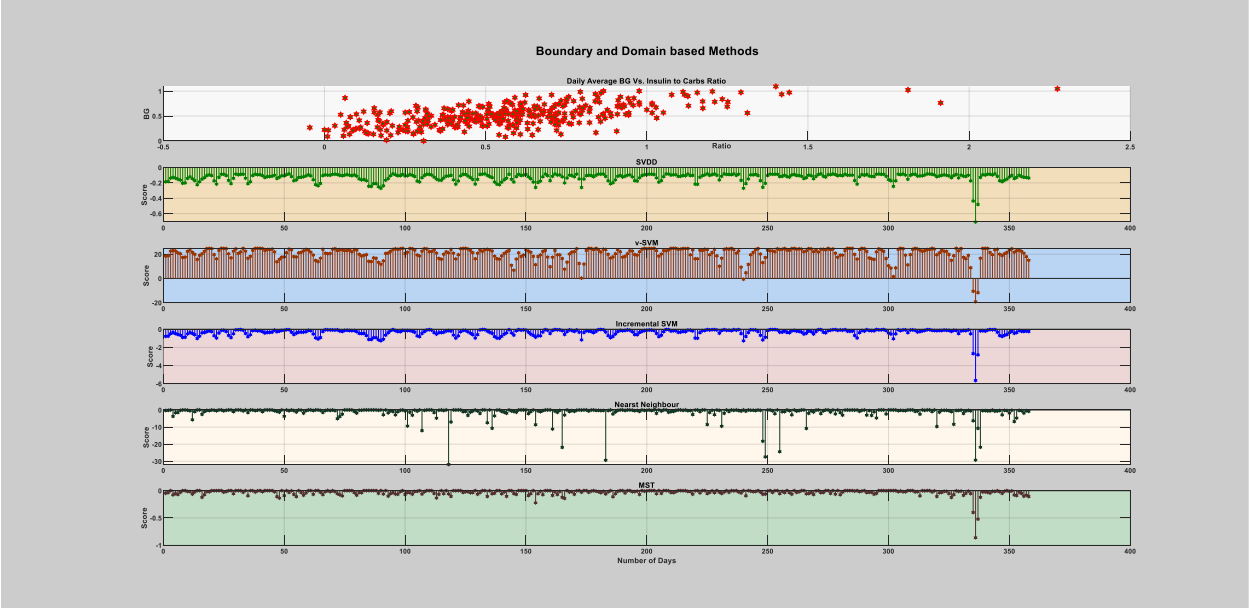


Figure 7: The score of the boundary and domain-based method on the whole patient year.

### 1.3.2. Density-Based Method

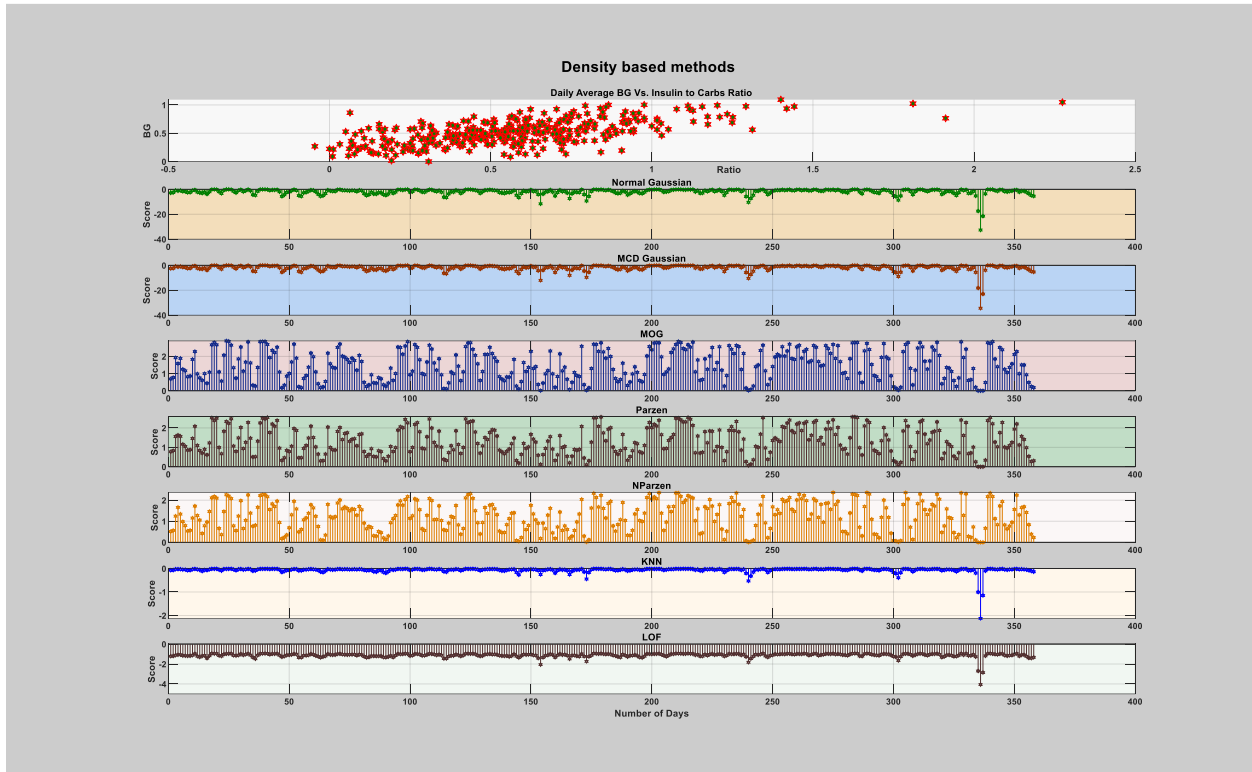


Figure 8: The score of the density-based method on the whole patient year.

### 1.3.3. Reconstruction-Based Method

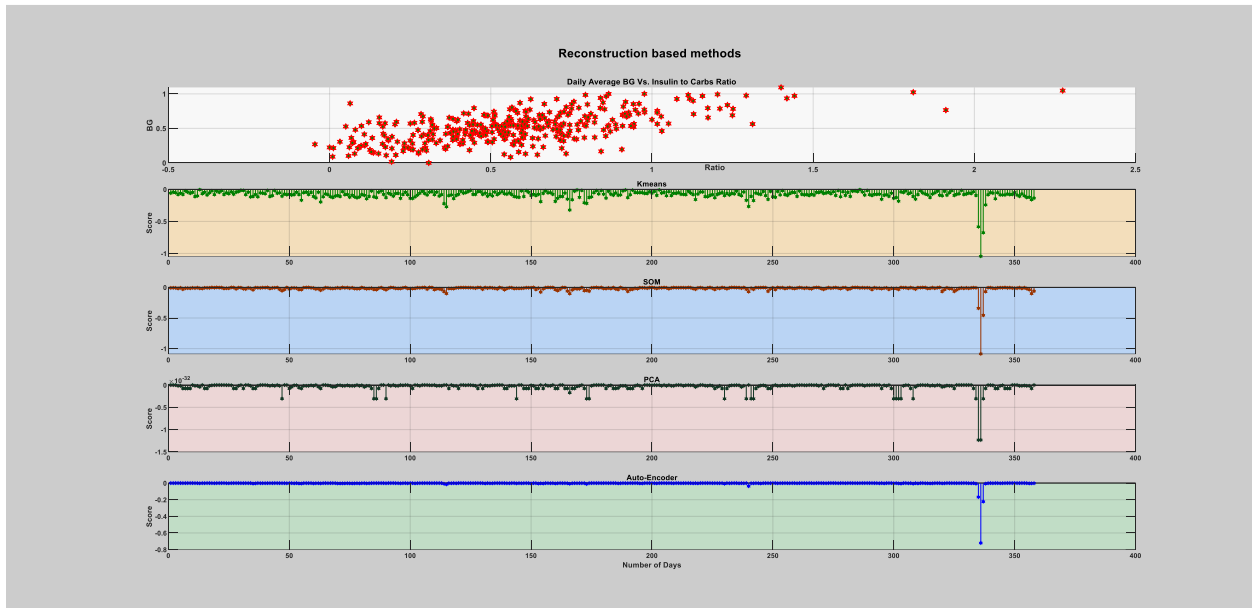


Figure 9: The score of the reconstruction-based method on the whole patient year.

## 1.4. The Fourth Case of Infection (flu)

### 1.4.1. Boundary and Domain-Based Method

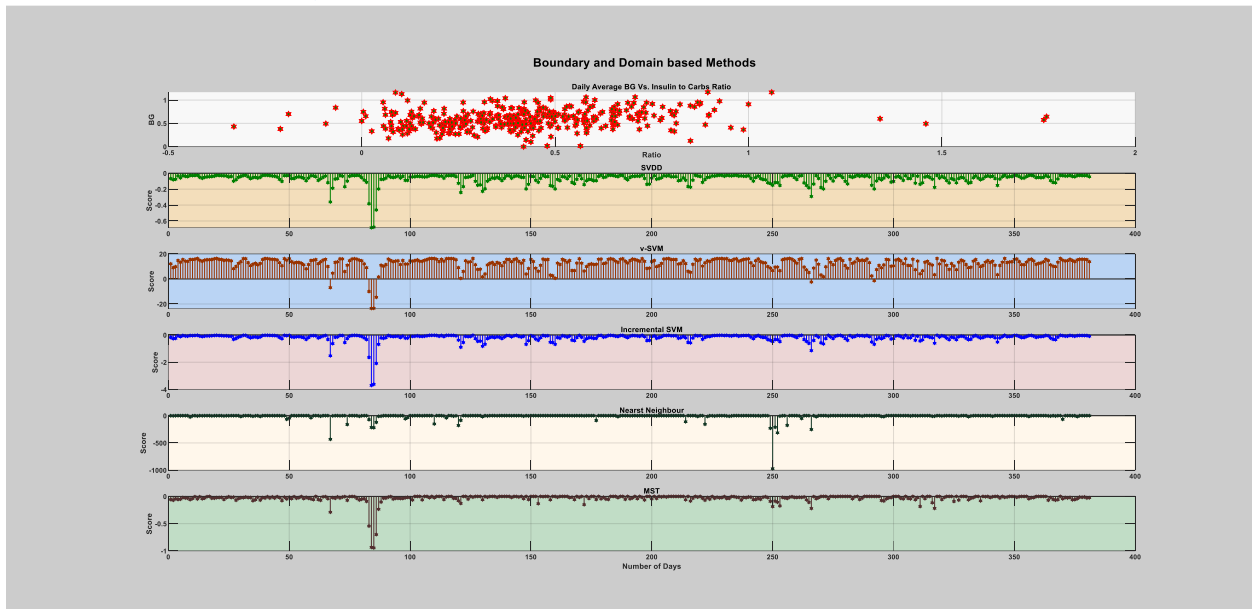


Figure 10: The score of the boundary and domain-based method on the whole patient year.

### 1.4.2. Density-Based Method

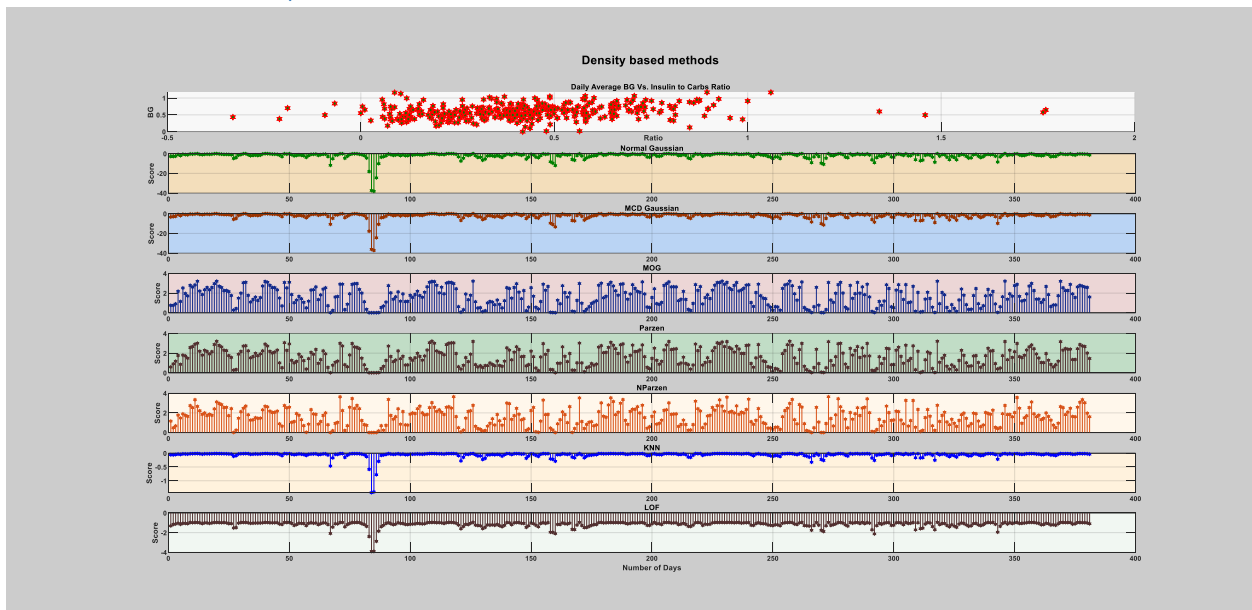


Figure 11: The score of the density-based method on the whole patient year.

1.4.3. Reconstruction-Based Method

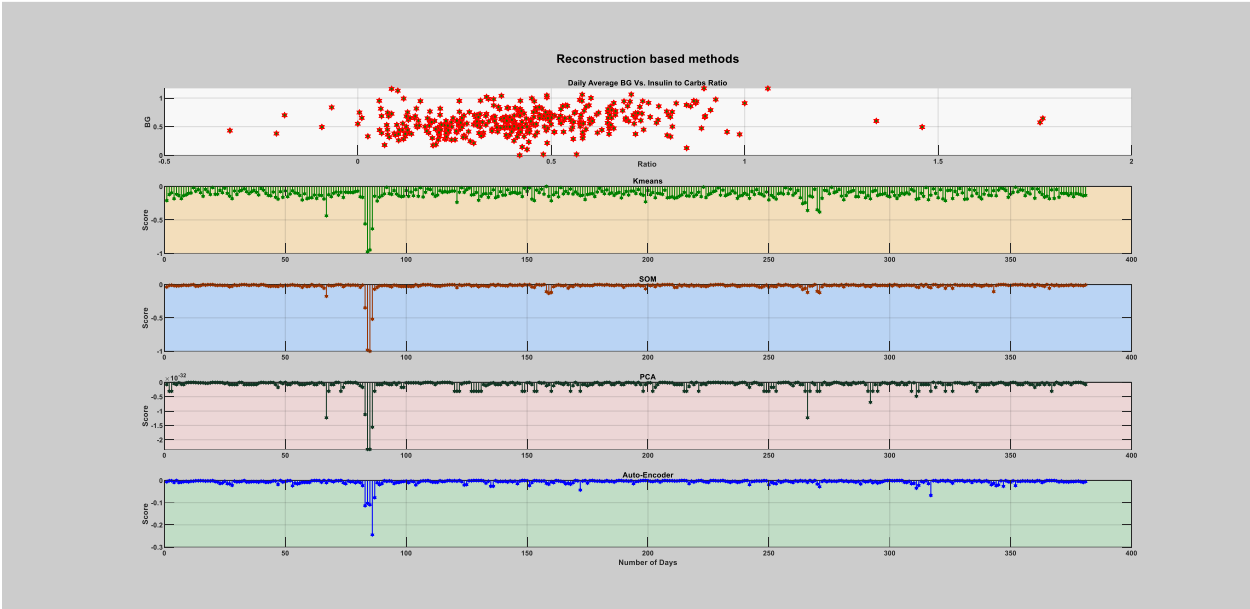


Figure 12: The score of the reconstruction-based method on the whole patient year.

2. Hourly

2.1. The First Case of Infection (flu)

2.1.1. Boundary and Domain-Based Method

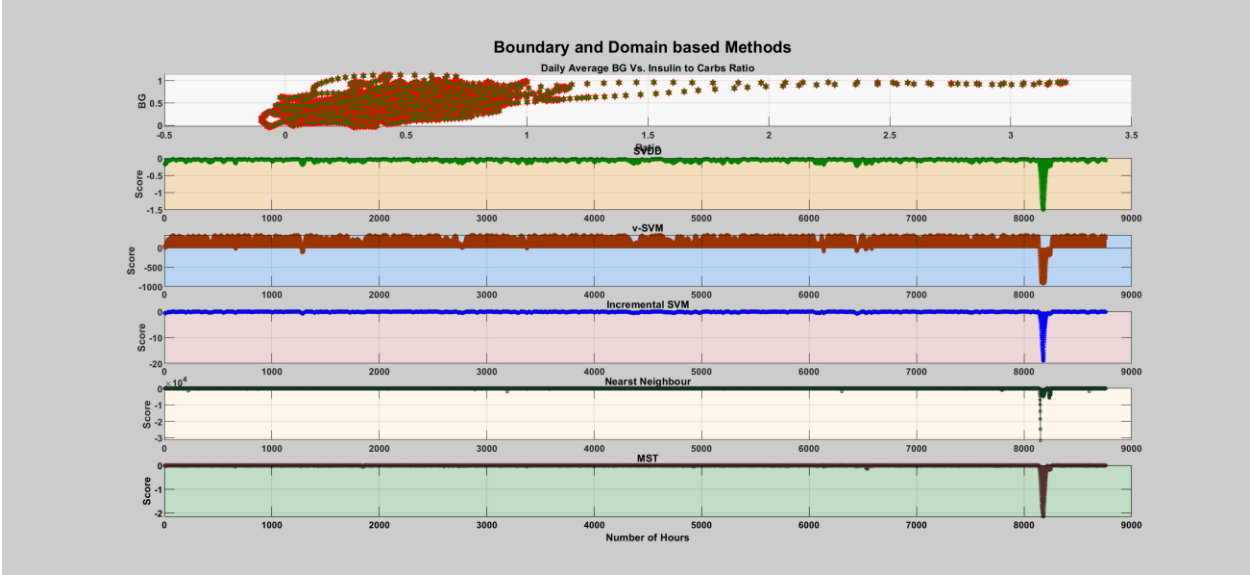


Figure 13: The score of the boundary and domain-based method on the whole patient year.



### 2.1.2. Density-Based Method

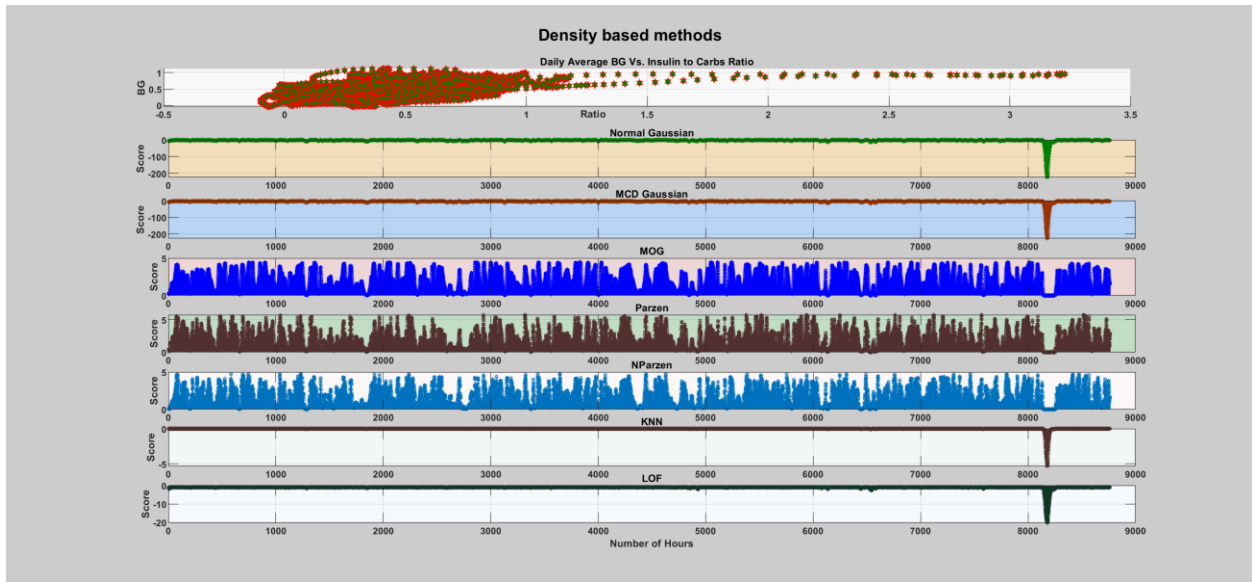


Figure 14: The score of the density-based method on the whole patient year.

### 2.1.3. Reconstruction-Based Method

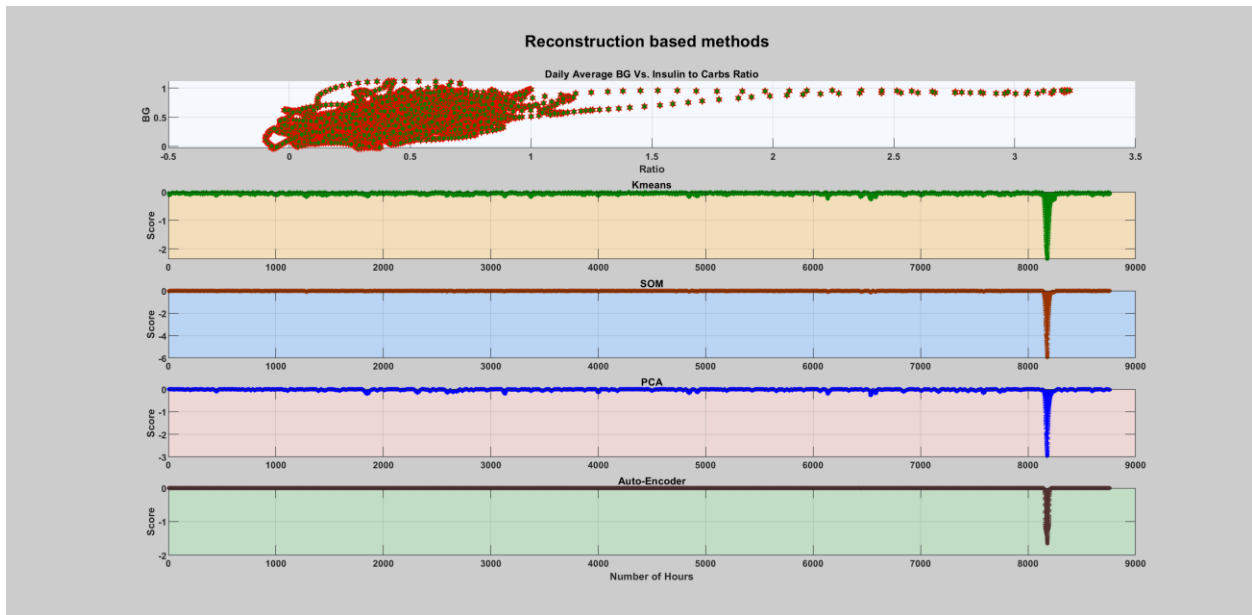


Figure 15: The score of the reconstruction-based method on the whole patient year.

## 2.2. The Second Case of Infection (flu)

### 2.2.1. Boundary and Domain-Based Method

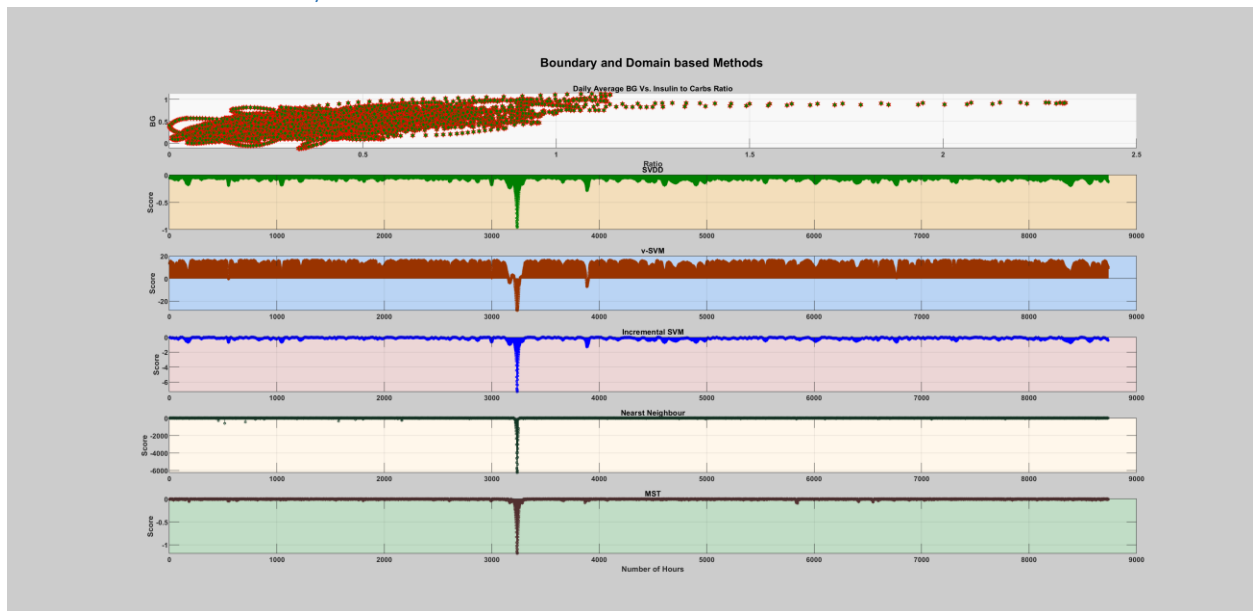


Figure 16: The score of the boundary and domain-based method on the whole patient year.

### 2.2.2. Density-Based Method

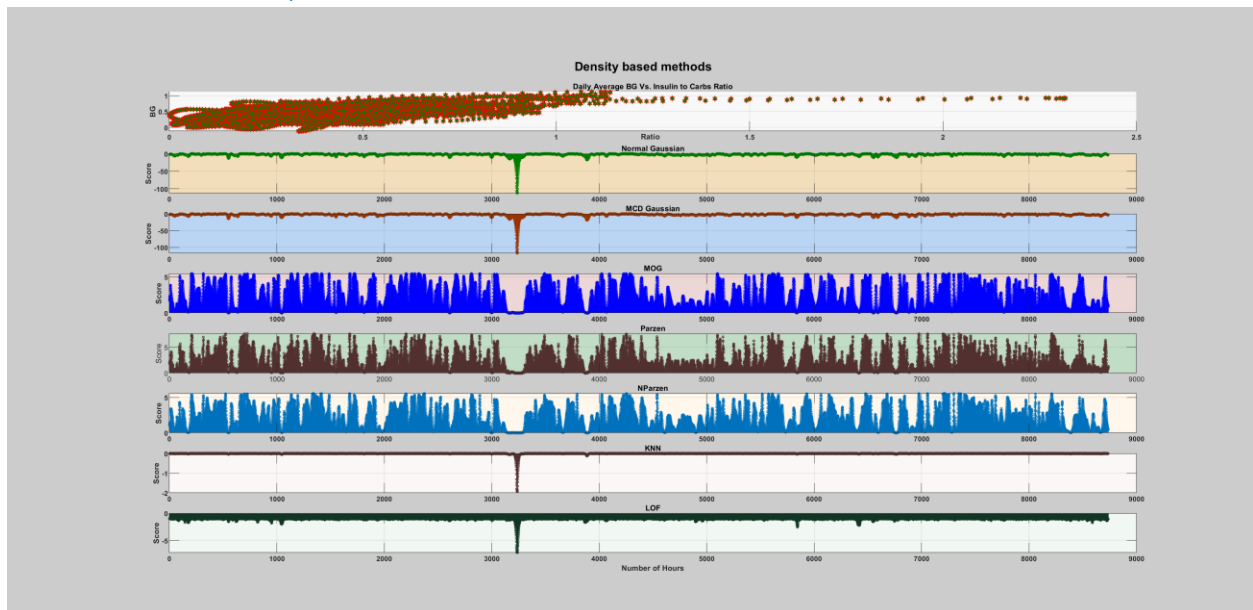


Figure 17: The score of the density-based method on the whole patient year.

### 2.2.3. Reconstruction-Based Method

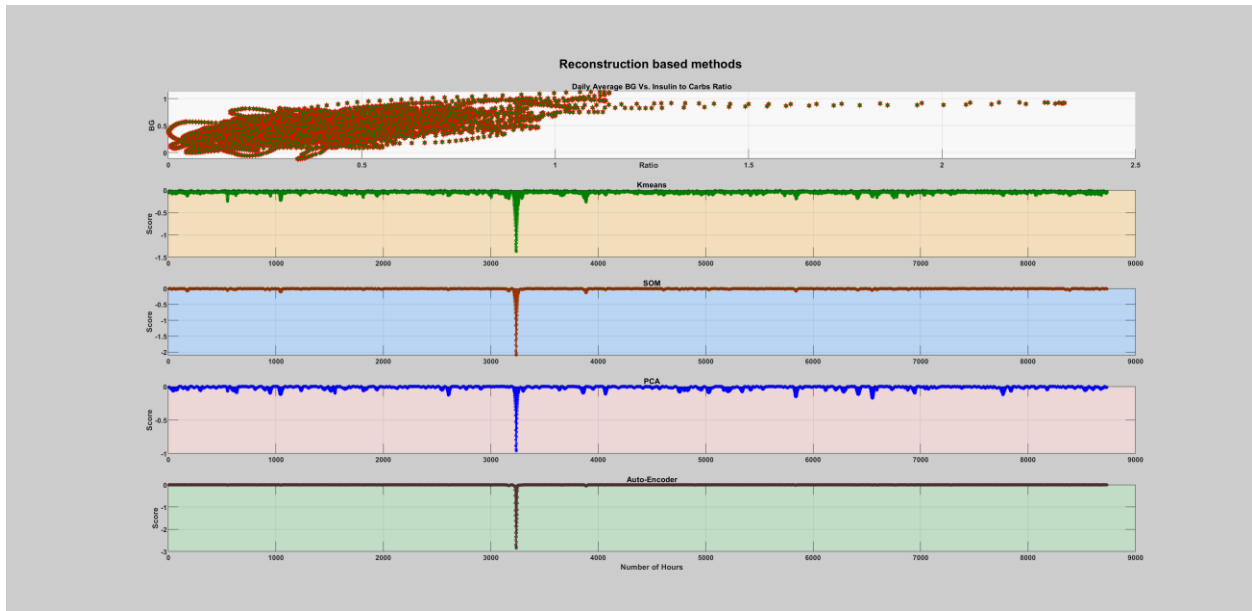


Figure 18: The score of the reconstruction-based method on the whole patient year.

## 2.3. The Third Case of Infection (flu)

### 2.3.1. Boundary and Domain-Based Method

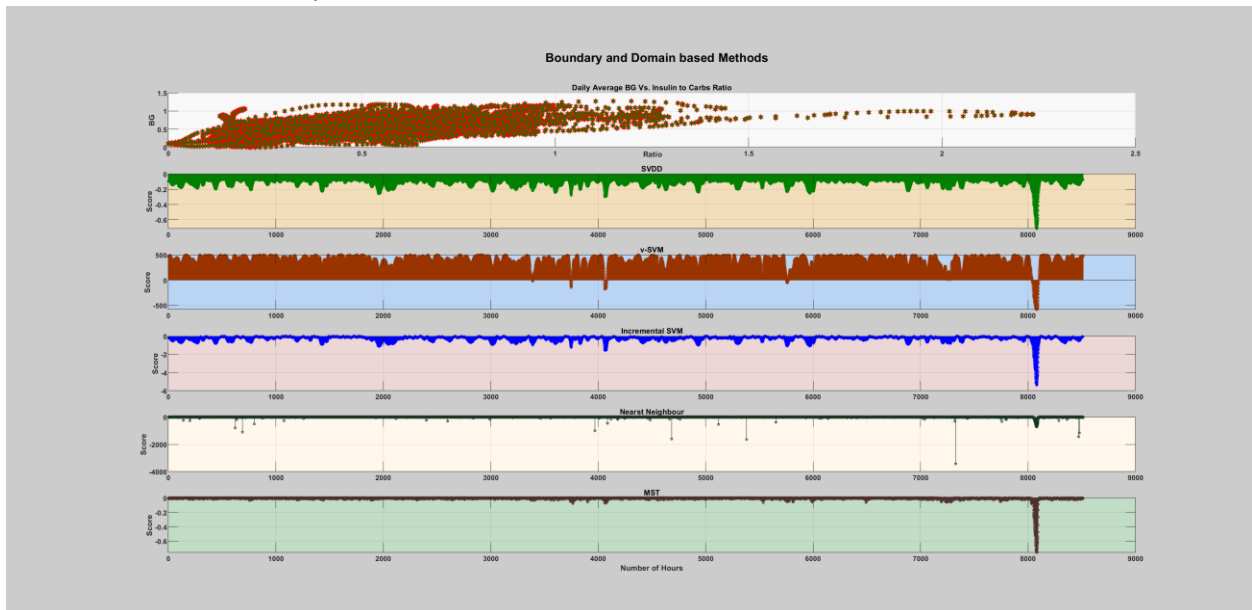


Figure 19: The score of the boundary and domain-based method on the whole patient year.

### 2.3.2. Density-Based Method

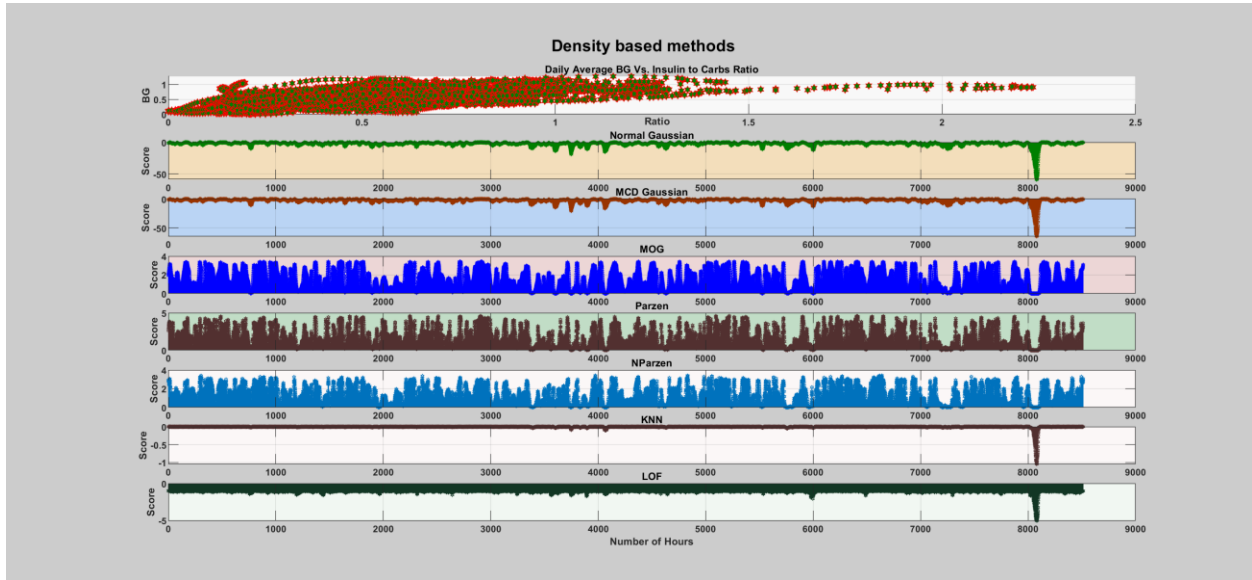


Figure 20: The score of the density-based method on the whole patient year.

### 2.3.3. Reconstruction-Based Method

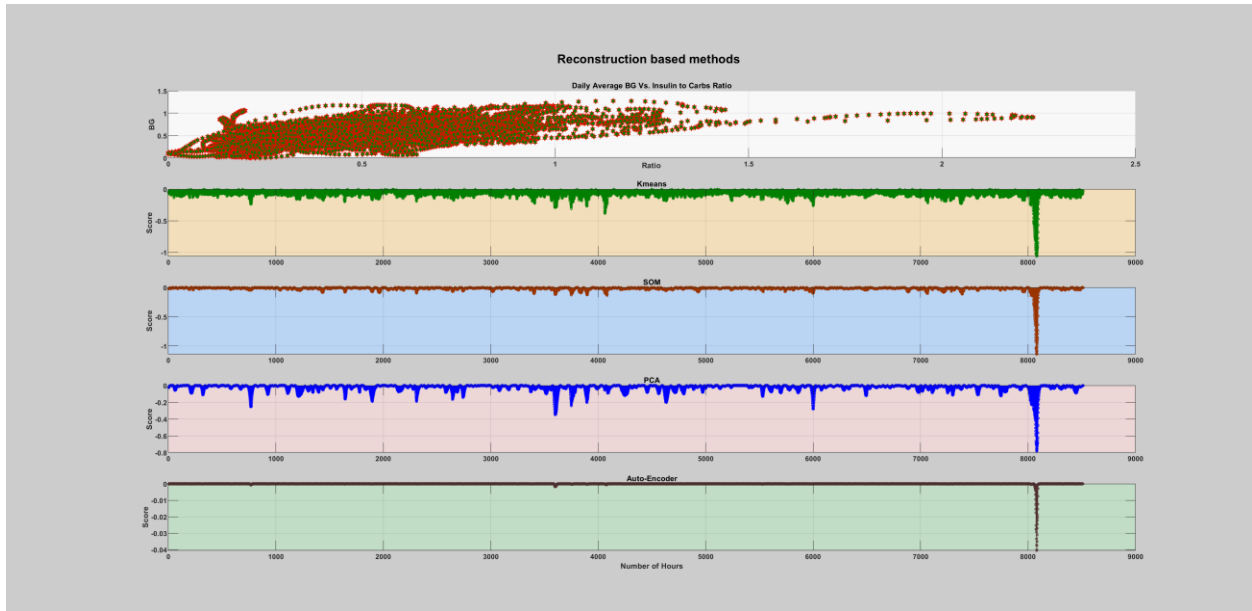


Figure 21: The score of the reconstruction-based method on the whole patient year.

## 2.4. The Fourth Case of Infection (flu)

### 2.4.1. Boundary and Domain-Based Method

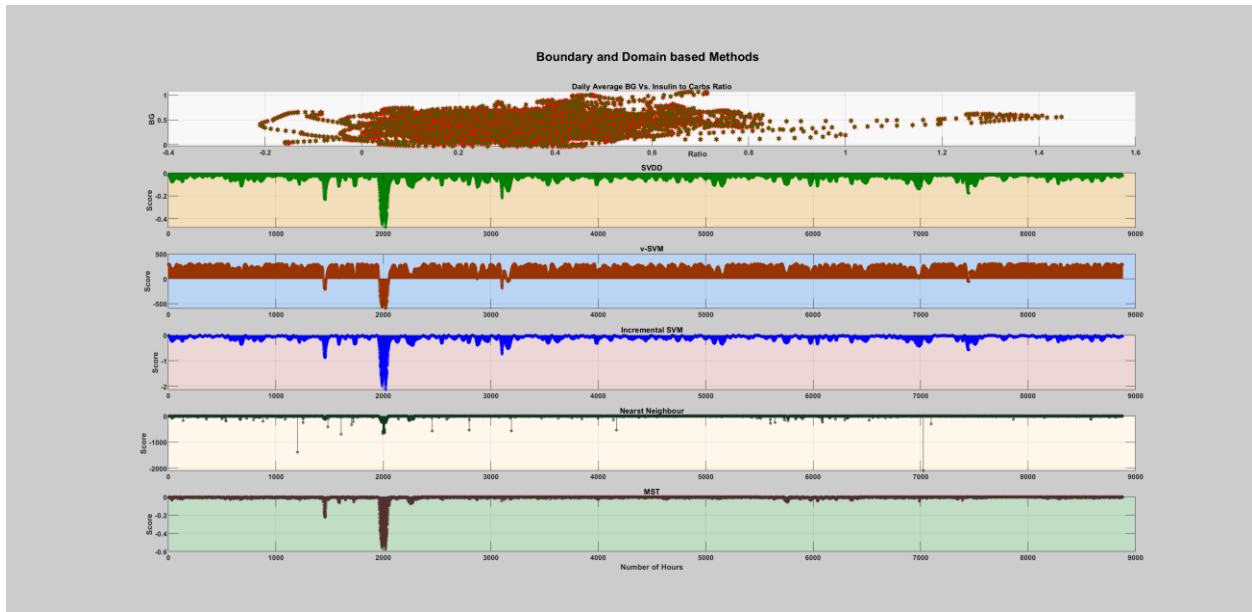


Figure 22: The score of the boundary and domain-based method on the whole patient year.

### 2.4.2. Density-Based Method

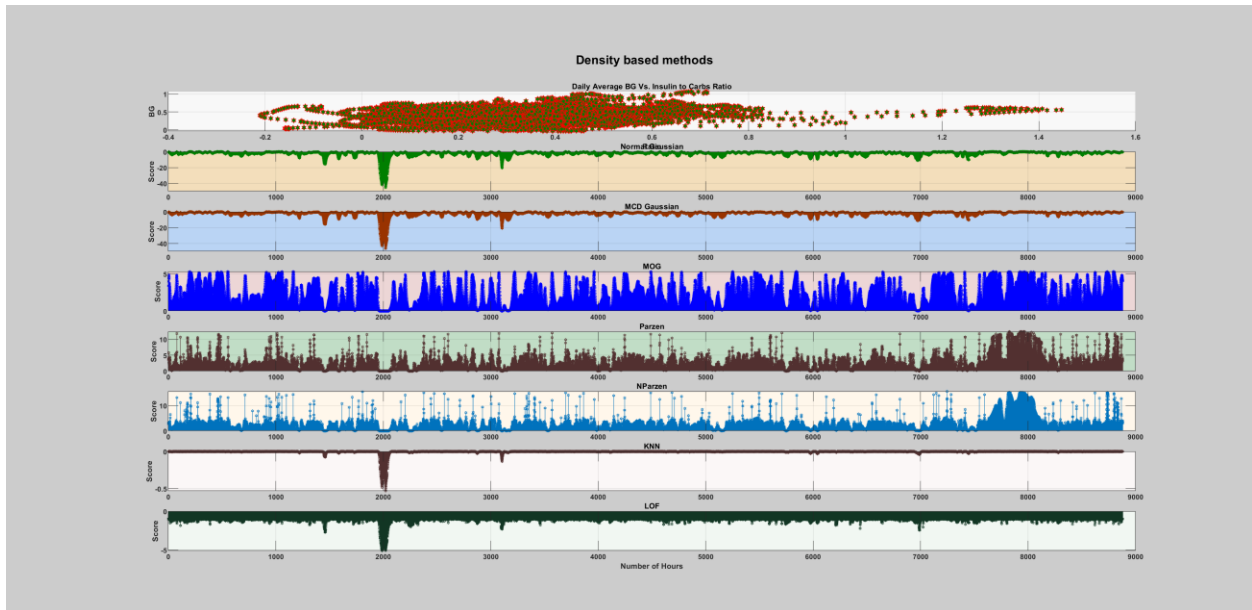


Figure 23: The score of the density-based method on the whole patient year.

### 2.4.3. Reconstruction-Based Method

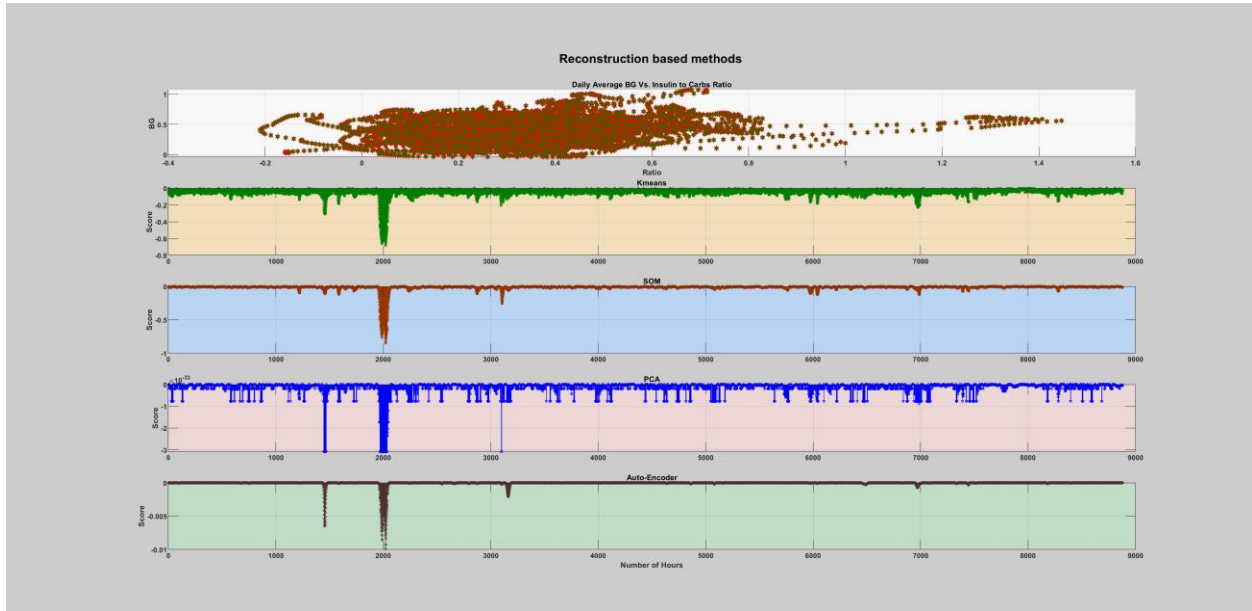


Figure 24: The score of the reconstruction-based method on the whole patient year.

## B. Unsupervised Approach

### 1. Daily

#### 1.1. The First Case of Infection (flu)

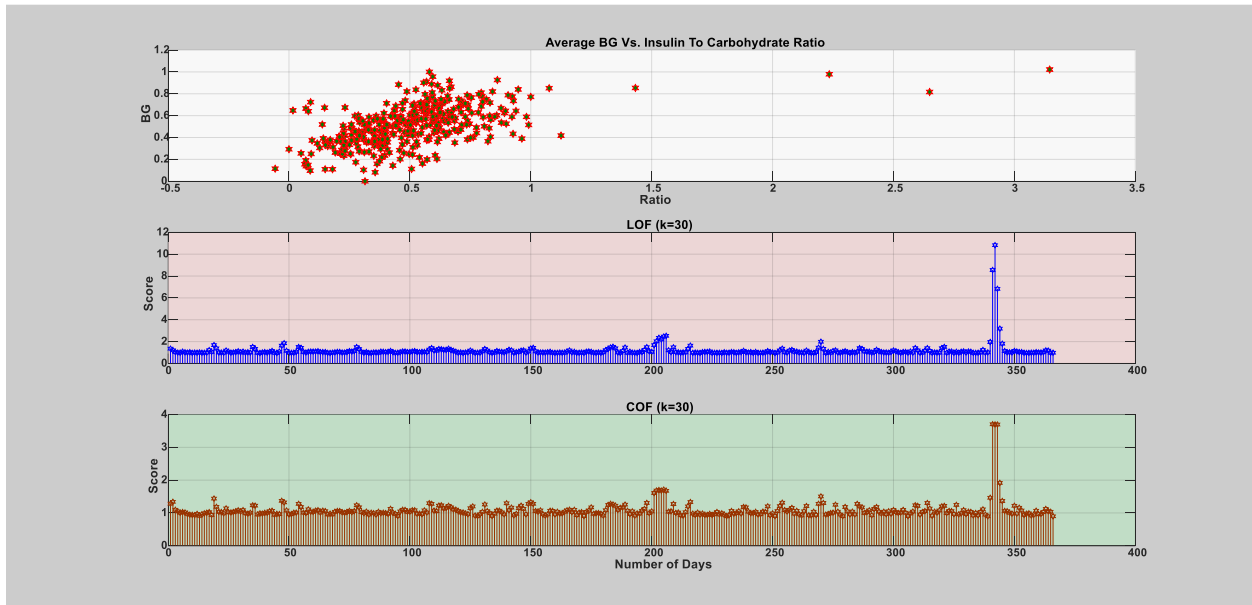


Figure 25: The score of unsupervised methods, LOF and COF, on the whole patient year. The value of  $k$  is set to be 30 data points.

### 1.2. The Second Case of Infection (flu)

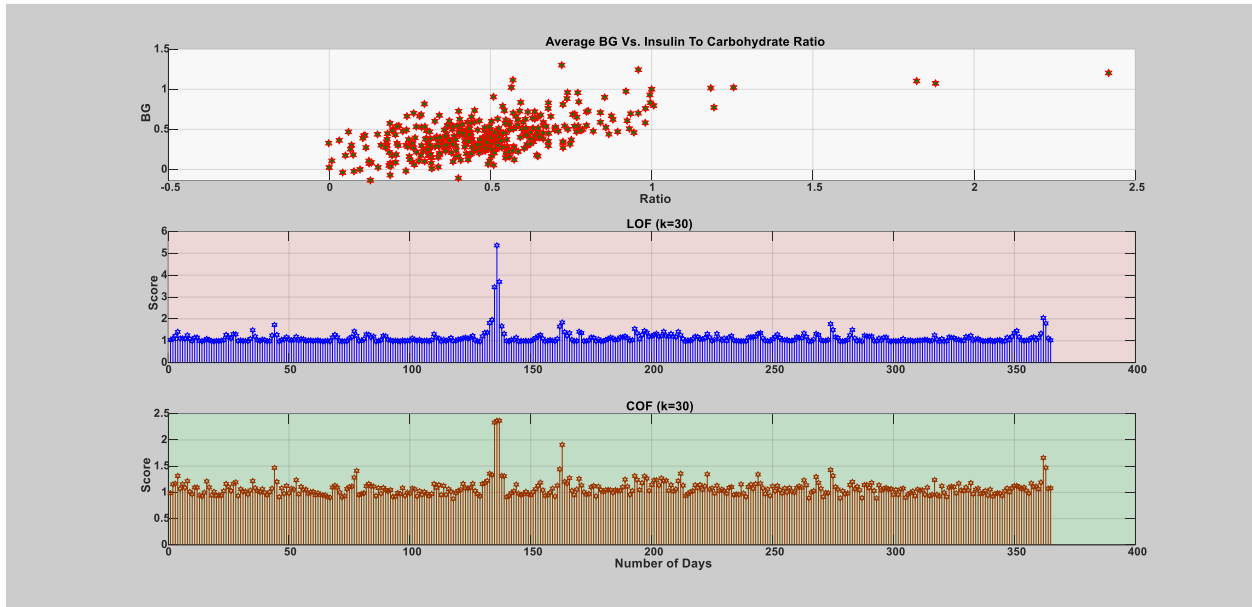


Figure 26: The score of unsupervised methods, LOF and COF, on the whole patient year. The value of k is set to be 30 data points.

### 1.3. The Third Case of Infection (flu)

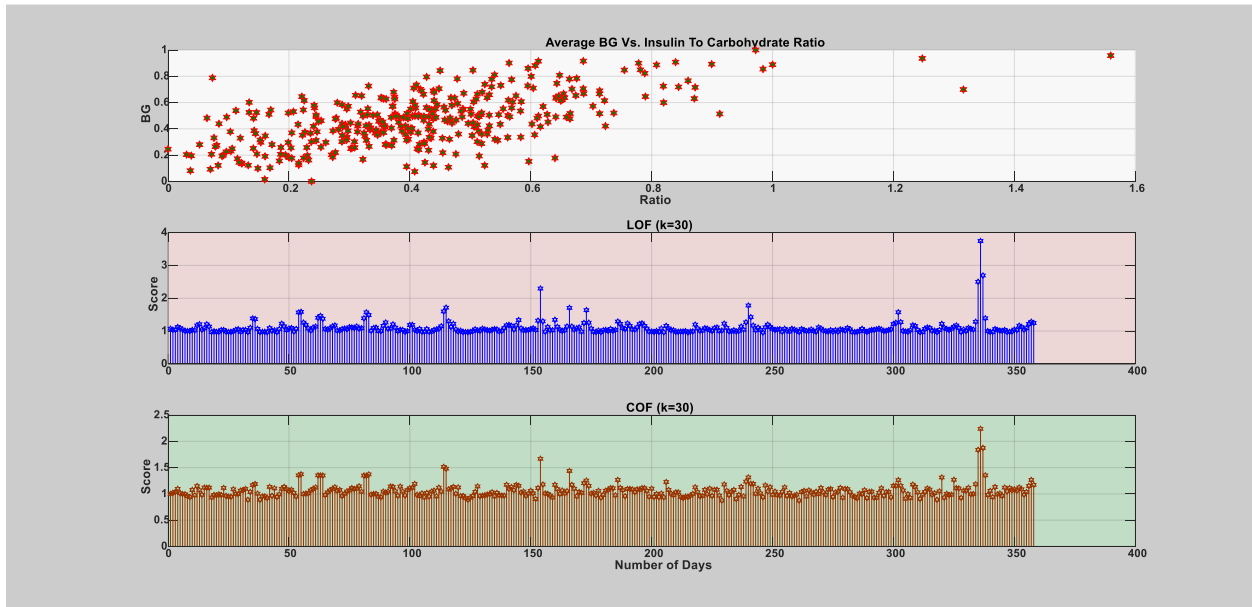


Figure 27: The score of unsupervised methods, LOF and COF, on the whole patient year. The value of k is set to be 30 data points.

### 1.4. The Fourth Case of Infection (flu)

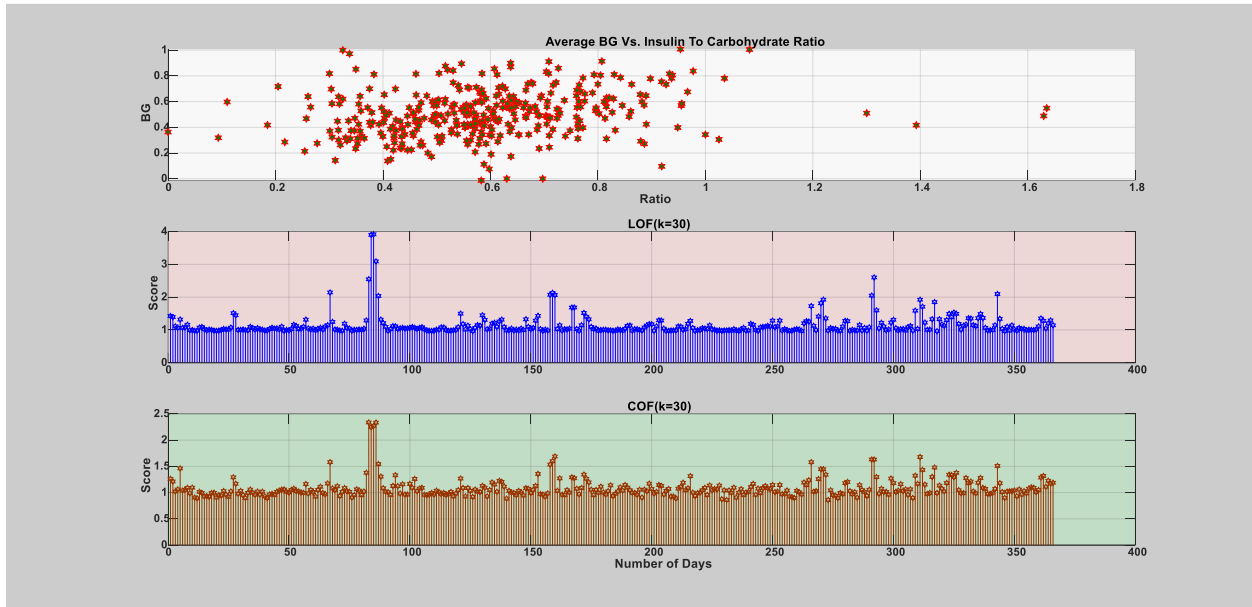


Figure 28: The score of unsupervised methods, LOF and COF, on the whole patient year. The value of  $k$  is set to be 30 data points.

## 2. Hourly

### 2.1. The First Case of Infection (flu)

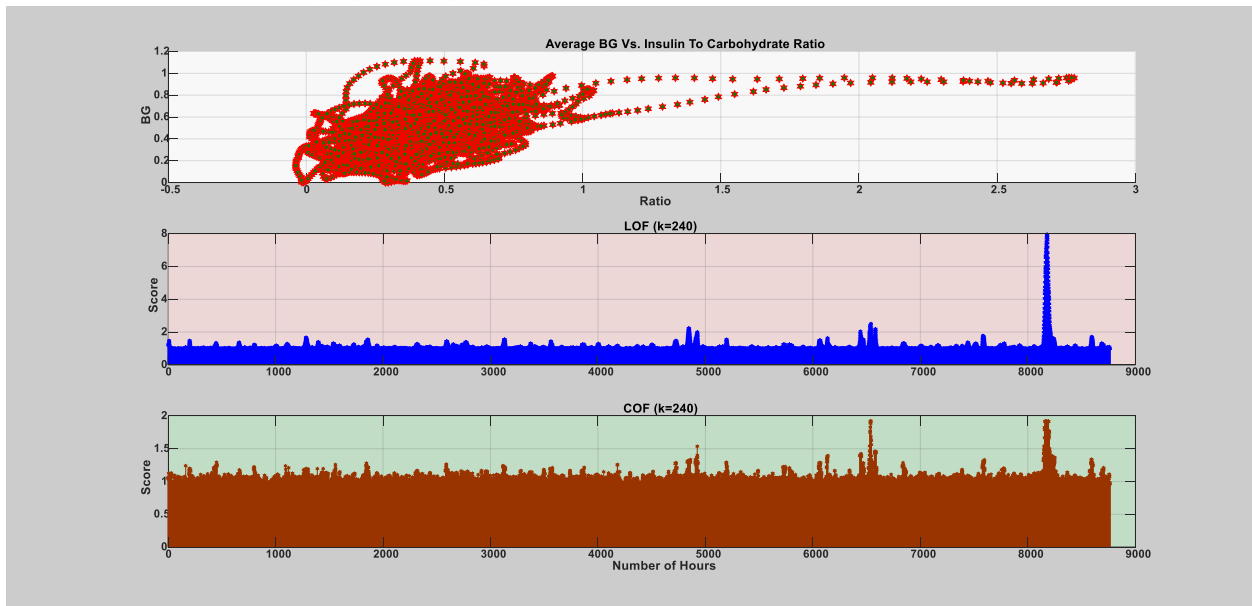


Figure 29: The score of unsupervised methods, LOF and COF, on the whole patient year. The value of  $k$  is set to be 240 data points.



### 2.2. The Second Case of Infection (flu)

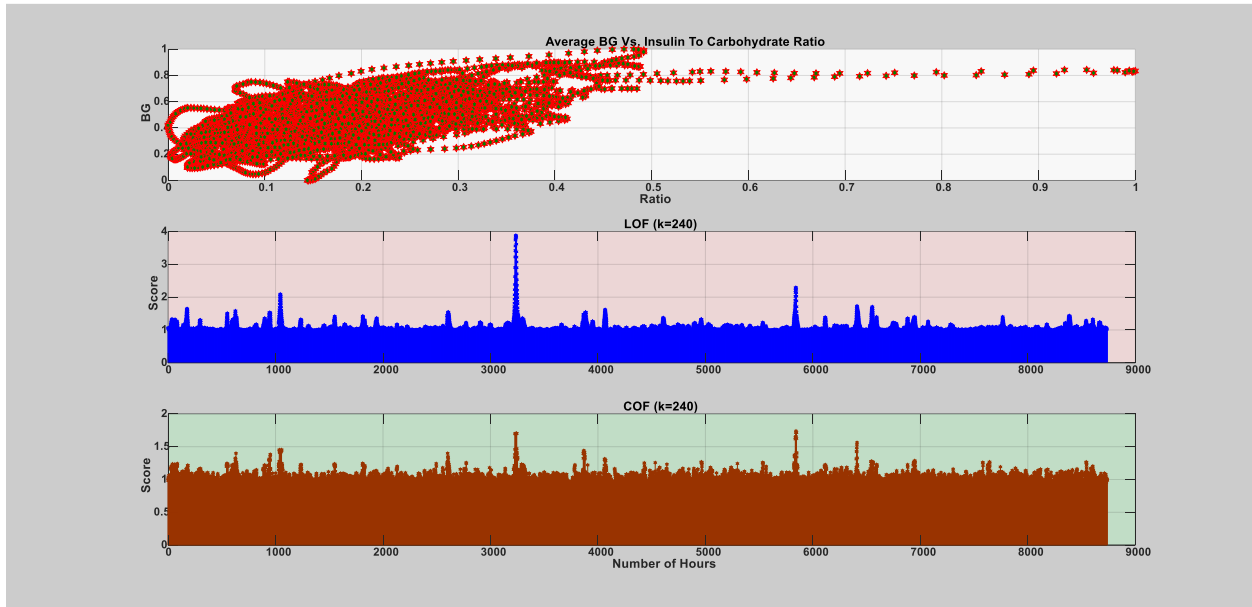


Figure 30: The score of unsupervised methods, LOF and COF, on the whole patient year. The value of k is set to be 240 data points.

### 2.3. The Third Case of Infection (flu)

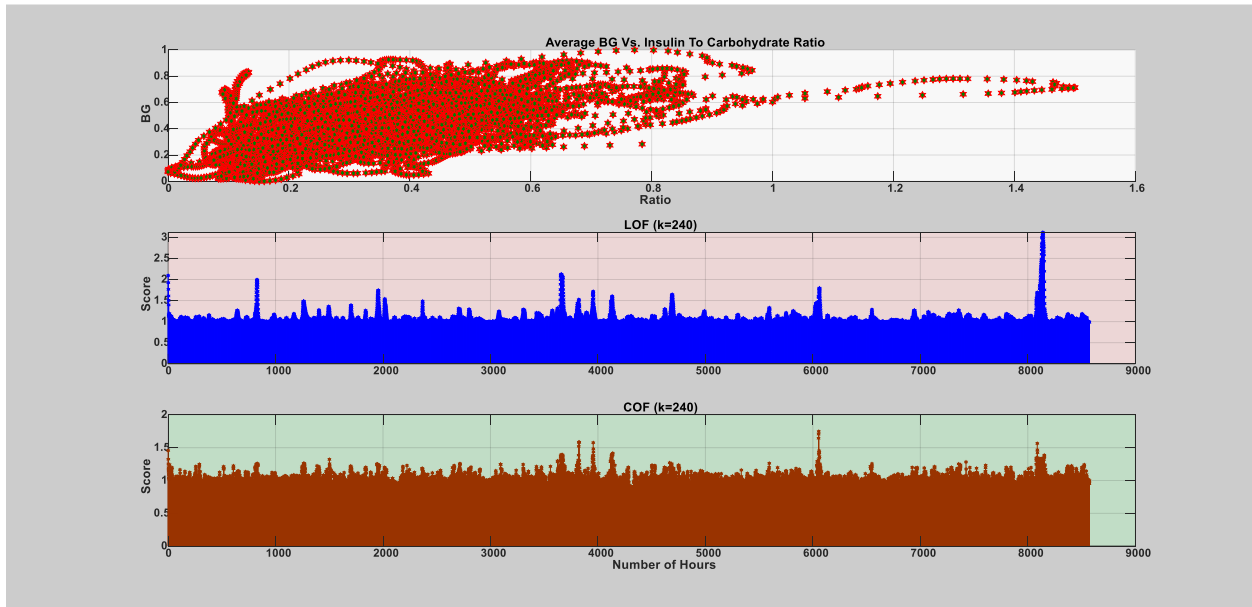


Figure 31: The score of unsupervised methods, LOF and COF, on the whole patient year. The value of k is set to be 240 data points.

## 2.4. The Fourth Case of Infection (flu)

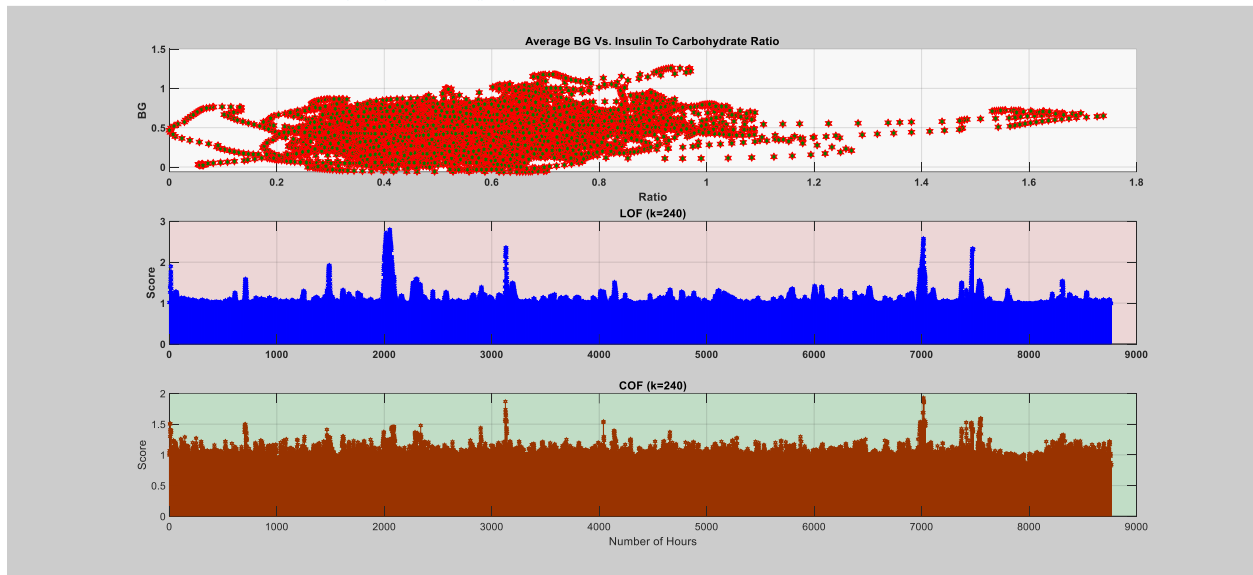


Figure 32: The score of unsupervised methods, LOF and COF, on the whole patient year. The value of k is set to be 240 data points.

## Appendix 4 - Model Evaluations – Performance of the Models for Each Patient Year

The one-class classifier models presented in this section consists of three categories; boundary and domain based, density based, and reconstruction based. The model evaluation is carried on a dataset that represent a daily and hourly scope. The models were evaluated using twenty times fivefold stratified cross-validation. During training phase, only the regular/normal day measurements were used to train the models. During testing phase, a dataset containing both regular/normal and infection day measurements were used. The performances are reported as average and standard deviation of the twenty rounds. Performance metrics like area under the ROC curve (AUC), specificity, and F1-score were used to evaluate the model performances. Two version of the same data was used to assess the performance; raw data and smoothed data. The smoothed data is the filtered version of the raw data using a moving average filter of 2-days window size. The models were compared considering required sample sizes to obtain satisfactory performance. The performance of the models is given in the Table 1-7 below for different individuals and infection years.

### I. Daily

#### 1. The First Case of Infection (Flu)

**Table 1 (a):** Average and standard deviation of AUC, specificity, F1-score for the raw dataset (without smoothing) and different sample size.

<b>Fraction = 0.01</b>												
<b>Models</b>	<b>Boundary and Domain-Based Method</b>											
	<b>1 Month</b>			<b>2 Months</b>			<b>3 Months</b>			<b>4 Months</b>		
	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1
SVDD	90.7 (8.8)	71.7 (7.7)	<b>73.6 (5.5)</b>	93.4 (6.2)	81.7 (5.0)	87.4 (8.1)	96.4 (2.9)	87.8 (3.3)	91.3 (6.0)	94.6 (3.7)	81.7 (5.0)	90.0 (4.6)
incsvdd	90.4 (8.9)	66.7 (7.5)	72.7 (4.9)	91.8 (5.9)	66.7 (7.5)	84.4 (3.2)	95.8 (2.9)	70.0 (7.1)	85.4 (1.2)	93.7 (3.6)	55 (10.7)	81.0 (2.7)
v-SVM	93.1 (6.0)	63 (10.6)	78.9 (6.2)	96.5 (2.3)	81.9 (4.7)	<b>90.7 (3.4)</b>	97.9 (1.5)	88.9 (0.0)	<b>94.1 (2.0)</b>	96.2 (2.3)	83.3 (0.0)	<b>91.7 (1.4)</b>
Nearest Neighbour	74.2 (9.3)	38.3 (7.7)	61.0 (4.7)	89.5 (9.3)	20.0 (6.7)	70.0 (4.6)	90.1 (6.6)	11.1 (18)	69.2 (3.8)	92.8 (3.3)	33.3 (0.0)	75.1 (0.4)
MST	89.4 (8.1)	50.0 (0.0)	62.7 (6.6)	95.4 (5.6)	61.7 (7.7)	82.3 (5.9)	96.6 (2.7)	68.9 (4.5)	83.6 (4.7)	94.1 (2.8)	55.0 (7.7)	80.6 (2.3)
<b>Density-Based Method</b>												
Gaussian	90.6 (7.1)	60.0 (8.2)	68.8 (8.4)	95.4 (4.6)	70.0 (6.7)	85.3 (4.6)	97.3 (2.5)	80.0 (4.5)	89.2 (3.3)	95.5 (3.2)	66.7 (0.0)	84.5 (2.0)
MoG	88.1 (9.9)	80.1 (17.3)	67.8 (16.4)	93.1 (7.1)	75.8 (14.8)	82.5 (10.1)	95.6 (3.4)	80.2 (7.5)	86.0 (6.7)	93.7 (3.9)	68.7 (11.6)	84.2 (5.7)
MCD Gaussian	89.0 (8.5)	55.0 (7.7)	66.4 (9.0)	94.0 (4.6)	68.3 (5.0)	84.6 (6.3)	97.0 (2.7)	80.0 (4.5)	89.9 (2.4)	94.5 (3.2)	65.0 (5.0)	84.0 (3.2)
Parzen	89.0 (9.2)	70.0 (6.7)	70.7 (5.9)	94.6 (4.9)	83.3 (0.0)	87.9 (6.3)	97.2 (2.4)	88.9 (0.0)	90.5 (5.9)	95.2 (2.9)	83.3 (0.0)	88.9 (3.3)
Naive Parzen	90.1 (7.6)	55 (10.7)	65.0 (5.0)	95.7 (3.9)	76.7 (8.2)	87.2 (3.5)	98.3 (1.4)	88.9 (0.0)	<b>93.6 (2.4)</b>	96.8 (2.1)	83.3 (0.0)	90.7 (2.0)
k-NN	91.8 (6.9)	50.0 (0.0)	66.0 (2.0)	95.6 (3.1)	81.7 (5.0)	<b>90.9 (3.2)</b>	97.9 (1.6)	88.9 (0.0)	93.5 (3.7)	97.0 (2.2)	83.3 (0.0)	<b>92.0 (1.0)</b>
LOF	88.5 (6.1)	66.7 (7.5)	<b>72.7 (4.9)</b>	97.0 (1.9)	71.7 (7.7)	86.1 (2.4)	96.8 (2.8)	78.9 (3.3)	88.7 (2.8)	92.6 (4.8)	50.0 (0.0)	79.3 (2.6)

	<b>Reconstruction-Based Method</b>											
PCA	87.8 (11.9)	50.0 (7.5)	62.4 (8.5)	93.5 (6.2)	51.7 (5.0)	78.2 (4.1)	93.6 (4.7)	60 (10.2)	81.8 (4.4)	91.3 (5.2)	46.7 (6.7)	78.7 (2.3)
Auto-encoder	82.2 (12.0)	57.9 (15.3)	64.7 (12.0)	88.2 (9.5)	61.6 (14.0)	81.4 (7.1)	93.4 (5.7)	74.4 (11)	86.4 (5.9)	88.4 (8.8)	61.3 (14.3)	82.7 (5.7)
SOM	86.9 (9.4)	78.3 (13.3)	66.7 (16.9)	92.8 (7.3)	64.2 (12.4)	80.9 (7.0)	95.8 (3.7)	80.1 (6.3)	86.9 (5.5)	92.2 (4.1)	76.5 (9.0)	87.5 (4.5)
K-means	91.8 (6.9)	65.0 (9.0)	<b>71.8 (5.1)</b>	96.0 (2.4)	83.3 (0.0)	<b>91.5 (2.8)</b>	97.6 (1.6)	88.9 (0.0)	<b>93.5 (3.7)</b>	96.2 (2.2)	83.3 (0.0)	<b>91.5 (1.6)</b>

Table 1 (b): Average and standard deviation of AUC, specificity, F1-score for smoothed version of the data with a two-days moving average filter and different sample size.

<b>Fraction = 0.01</b>												
<b>Models</b>	<b>Boundary and Domain-Based Method</b>											
	<b>1 Month</b>			<b>2 Months</b>			<b>3 Months</b>			<b>4 Months</b>		
	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1
SVDD	99.6 (1.3)	100 (0.0)	93.6 (15.2)	100 (0.0)	100 (0.0)	94.8 (10.1)	100 (0.0)	100 (0.0)	97.0 (4.1)	100 (0.0)	100 (0.0)	96.9 (4.0)
incsvdd	99.6 (1.3)	100 (0.0)	93.6 (15.2)	100 (0.0)	100 (0.0)	97.1 (6.3)	100 (0.0)	100 (0.0)	97.6 (4.1)	100 (0.0)	100 (0.0)	98.3 (2.8)
<i>v-SVM</i>	100 (0.0)	99.5 (2.9)	<b>98.9 (3.2)</b>	100 (0.0)	100 (0.0)	<b>99.1 (2.6)</b>	100 (0.0)	100 (0.0)	<b>99.4 (1.7)</b>	100 (0.0)	100 (0.0)	<b>99.6 (1.2)</b>
Nearest Neighbour	98.1 (3.9)	58.3 (15.4)	72.3 (9.9)	86.9 (12.5)	16.7 (22.4)	70.5 (5.3)	88.1 (6.5)	54.4 (22.5)	80.0 (8.6)	92.4 (5.3)	8.3 (17.1)	69.0 (4.8)
MST	98.5 (2.4)	85.0 (5.0)	85.5 (2.1)	99.7 (0.8)	100 (0.0)	97.1 (6.3)	99.9 (0.4)	97.8 (4.5)	97.2 (4.0)	99.7 (0.8)	100 (0.0)	97.0 (7.9)
<b>Density-Based Method</b>												
Gaussian	100 (0.0)	98.3 (5.0)	92.1 (15.2)	100 (0.0)	100 (0.0)	97.1 (6.3)	99.8 (0.7)	100 (0.0)	97.6 (4.1)	99.4 (1.7)	100 (0.0)	97.0 (7.9)
MoG	98.6 (3.2)	99.8 (1.7)	88.5 (16.8)	99.6 (1.2)	100 (0.0)	92.2 (11.1)	99.7 (0.7)	99.8 (1.4)	94 (10.3)	99.3 (2.0)	99.9 (1.2)	94.4 (11.8)
MCD Gaussian	98.9 (2.2)	91.7 (8.4)	<b>90.9 (7.7)</b>	100 (0.0)	100 (0.0)	<b>98.0 (6.0)</b>	99.5 (1.1)	96.7 (5.1)	96.6 (5.9)	99.4 (1.7)	88.3 (7.7)	92.0 (6.8)
Parzen	99.6 (1.3)	100 (0.0)	87.7 (17.0)	100 (0.0)	100 (0.0)	95.1 (8.0)	100 (0.0)	100 (0.0)	94.6 (9.8)	99.9 (0.4)	100 (0.0)	94.6 (12.3)
Naive Parzen	99.2 (2.5)	100 (0.0)	94.7 (11.1)	100 (0.0)	100 (0.0)	93.8 (11.0)	99.6 (1.1)	100 (0.0)	97.5 (5.0)	100 (0.0)	100 (0.0)	<b>98.7 (2.7)</b>
k-NN	98.1 (3.9)	68.3 (5.0)	75.2 (4.3)	100 (0.0)	100 (0.0)	<b>98.0 (6.0)</b>	100 (0.0)	100 (0.0)	<b>98.8 (3.8)</b>	100 (0.0)	100 (0.0)	97.7 (4.7)
LOF	98.6 (2.9)	75.0 (13.5)	80.2 (10.8)	100 (0.0)	100 (0.0)	<b>98.0 (6.0)</b>	100 (0.0)	100 (0.0)	96.9 (5.0)	99.7 (0.8)	100 (0.0)	97.4 (7.9)
<b>Reconstruction-Based Method</b>												
PCA	98.9 (2.2)	85.0 (5.0)	<b>85.5 (2.1)</b>	99.2 (1.3)	85.0 (5.0)	<b>91.4 (2.7)</b>	98.6 (1.9)	88.9 (0.0)	92.2 (6.0)	97.8 (2.2)	83.3 (0.0)	89.1 (9.7)
Auto-encoder	97.4 (6.0)	89.1 (13.0)	86.0 (14.2)	98.5 (3.2)	94.5 (9.6)	91.8 (9.4)	99.2 (2.4)	93.7 (10.2)	93.7 (8.3)	98.6 (3.8)	94.4 (9.5)	93.7 (9.7)
SOM	99.3 (1.9)	99.9 (1.2)	84.7 (19.8)	99.8 (0.7)	100 (0.0)	91.4 (9.6)	99.9 (0.3)	100 (0.0)	95.2 (7.9)	99.6 (1.3)	100 (0.0)	93.4 (12.1)
K-means	99.2 (2.5)	85.0 (11.7)	87.0 (10.4)	100 (0.0)	100 (0.0)	97.1 (6.3)	100 (0.0)	100 (0.0)	<b>98.8 (3.8)</b>	100 (0.0)	100 (0.0)	<b>99.2 (2.5)</b>

## 2. The Second Case of Infection (Flu)

Table 2 (a): Average and standard deviation of AUC, specificity, F1-score for the raw dataset (without smoothing) and different sample size.

<b>Fraction = 0.01</b>												
<b>Models</b>	<b>Boundary and Domain-Based Method</b>											
	<b>1 Month</b>			<b>2 Months</b>			<b>3 Months</b>			<b>4 Months</b>		
	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1
SVDD	78.1 (19.2)	25 (25.1)	53.3 (5.2)	90 (11.4)	23.3 (20.1)	69.5 (3.7)	93.5 (7.0)	28.9 (20.1)	<b>72.7 (3.7)</b>	87.0 (5.6)	23.3 (20.1)	<b>71.6 (4.1)</b>
incsvdd	67.1 (18.0)	18.3 (5.0)	52.3 (9.4)	87.5 (14.2)	23.3 (20.1)	<b>70.3 (2.8)</b>	91.9 (8.4)	28.9 (20.1)	<b>72.7 (3.7)</b>	86.1 (5.8)	23.3 (20.1)	<b>71.6 (4.1)</b>
<i>v-SVM</i>	91.7 (11.4)	18.2 (8.6)	<b>61.4 (7.5)</b>	93.0 (6.8)	19.0 (8.2)	65.7 (5.7)	96.5 (5.1)	26.4 (11.9)	69.1 (5.7)	96.5 (3.8)	27.6 (23.2)	69.7 (9.5)
Nearest Neighbour	39.4 (11.8)	16.7 (0.0)	54.1 (1.4)	32.8 (14.9)	6.7 (11.1)	66.2 (3.2)	38.3 (9.5)	8.6 (9.4)	66.4 (0.7)	29.7 (9.5)	4.5 (13.7)	66.5 (0.5)
MST	45.3 (12.7)	16.7 (0.0)	53.8 (5.1)	42.1 (16.2)	18.3 (5.0)	69.6 (4.2)	52.1 (10.6)	23.3 (3.3)	70.8 (3.0)	52 (11.7)	18.3 (5.0)	71.4 (1.3)

<b>Density-Based Method</b>												
<b>Gaussian</b>	90.0 (15.9)	18.3 (5.0)	54.2 (3.9)	93.9 (8.2)	23.3 (20.1)	70.3 (2.8)	96.2 (5.6)	28.9(20.1)	73.2 (3.1)	97.4 (3.9)	23.3 (20.1)	72.0 (3.8)
<b>MoG</b>	81.6 (19.7)	30.7 (10.6)	52.2 (10.1)	89.0 (7.9)	23.2 (19.4)	69.4 (3.8)	93.0 (5.0)	28.9 (19.7)	71.5 (4.4)	94.8 (3.7)	23.6 (17.6)	70.7 (3.7)
<b>MCD Gaussian</b>	89.4 (13.0)	16.7 (0.0)	<b>53.8 (5.1)</b>	93.1 (8.0)	21.7 (15.0)	70.5 (1.6)	97.1 (5.0)	28.9 (20.1)	73.2 (3.1)	97.5 (4.0)	25.0 (20.1)	72.5 (3.9)
<b>Parzen</b>	87.2 (16.3)	21.7 (10.7)	<b>50.9 (6.8)</b>	90.9 (8.1)	23.3 (20.1)	70.3 (2.8)	94.2 (5.8)	32.2 (21.4)	73.6 (4.9)	95.5 (4.3)	26.7 (21.4)	71.7 (4.2)
<b>Naive Parzen</b>	88.9 (11.9)	20.0 (6.7)	<b>53.2 (4.9)</b>	91.2 (8.2)	23.3 (20.1)	<b>71.1 (1.1)</b>	95.4 (5.7)	33.3 (21.1)	<b>74.0 (3.4)</b>	95.9 (4.3)	38.3 (16.8)	<b>76.0 (3.4)</b>
<b>k-NN</b>	92.2 (10.6)	20.0 (10.0)	<b>54.8 (2.6)</b>	93.9 (6.7)	23.3 (20.1)	<b>71.1 (1.1)</b>	97.1 (5.3)	28.9 (20.1)	73.2 (3.1)	97.4 (3.9)	23.3 (20.1)	72.0 (3.8)
<b>LOF</b>	91.1 (10.9)	16.7 (0.0)	53.8 (5.1)	94.5 (8.4)	23.3 (20.1)	70.3 (2.8)	95.9 (5.5)	28.9 (20.1)	73.2 (3.1)	97.2 (4.0)	23.3 (20.1)	72.0 (3.8)
<b>Reconstruction-Based Method</b>												
<b>PCA</b>	71.0 (9.6)	18.3 (5.0)	52.3 (9.4)	74.2 (9.5)	16.7 (0.0)	70.1 (2.7)	79.0 (7.7)	23.3 (3.3)	71.4 (1.9)	77.8 (6.0)	16.7 (0.0)	70.4 (1.3)
<b>Auto-encoder</b>	70.2 (19.2)	25.3 (10.8)	54.3 (7.8)	73.4 (16.3)	21.8 (14.5)	70.2 (3.8)	81.0 (13.0)	28.3 (15.0)	72.1 (3.3)	84 (13.0)	22.3 (13.9)	71.2 (3.1)
<b>SOM</b>	61.5 (19.9)	24.0 (8.3)	54.2 (6.6)	67.3 (10.4)	21.8 (13.2)	67.8 (4.5)	86.2 (8.3)	27.3 (13.1)	71.7 (2.2)	94.1 (4.8)	17.7 (6.7)	70.1 (2.1)
<b>K-means</b>	93.3 (9.6)	18.3 (5.0)	<b>54.2 (3.9)</b>	92.5 (6.3)	23.3 (20.1)	<b>71.1 (1.1)</b>	96.2 (5.5)	28.9 (20.1)	<b>73.2 (3.1)</b>	96.0 (4.1)	23.3 (20.1)	<b>72.0 (3.8)</b>

Table 2 (b): Average and standard deviation of AUC, specificity, F1-score for smoothed version of the data with a two-days moving average filter and different sample size.

<b>Fraction = 0.01</b>												
<b>Models</b>	<b>Boundary and Domain-Based Method</b>											
	<b>1 Month</b>			<b>2 Months</b>			<b>3 Months</b>			<b>4 Months</b>		
	<b>AUC</b>	<b>Specificity</b>	<b>F1</b>	<b>AUC</b>	<b>Specificity</b>	<b>F1</b>	<b>AUC</b>	<b>Specificity</b>	<b>F1</b>	<b>AUC</b>	<b>Specificity</b>	<b>F1</b>
<b>SVDD</b>	100 (0.0)	100 (0.0)	96.0 (8.0)	100 (0.0)	100 (0.0)	98.2 (3.6)	100 (0.0)	100 (0.0)	96.0 (8.5)	100 (0.0)	100 (0.0)	98.7 (2.0)
<b>incsvdd</b>	100 (0.0)	100 (0.0)	96.0 (8.0)	100 (0.0)	100 (0.0)	98.2 (3.6)	100 (0.0)	100 (0.0)	96.6 (8.6)	100 (0.0)	100 (0.0)	98.7 (2.0)
<b>v-SVM</b>	100 (0.0)	100 (0.0)	<b>99.1 (3.1)</b>	100 (0.0)	100 (0.0)	<b>99.2 (2.3)</b>	100 (0.0)	100 (0.0)	<b>99.5 (1.6)</b>	100 (0.0)	100 (0.0)	<b>99.5 (1.3)</b>
<b>Nearest Neighbour</b>	76.7 (21.3)	13.3 (25.7)	55.7 (10.6)	79 (12.0)	25.0 (26.2)	72.4 (9.8)	77 (11.4)	17 (13.4)	70.5 (4.3)	79.7 (8.5)	20 (27.8)	71.7 (8.2)
<b>MST</b>	97.2 (6.7)	55 (10.7)	68.3 (8.5)	99.4 (1.1)	86.7 (6.7)	92.1 (0.7)	100 (0.0)	97.8 (6.7)	97.2 (4.6)	99.7 (0.6)	100 (0.0)	98.7 (2.9)
<b>Density-Based Method</b>												
<b>Gaussian</b>	96.7 (10.0)	72 (16.8)	74.9 (12.4)	98.1 (5.0)	87 (14.6)	91.3 (9.2)	100 (0.0)	100 (0.0)	98.8 (2.4)	100 (0.0)	100 (0.0)	<b>99.6 (1.3)</b>
<b>MoG</b>	96.7 (10.0)	72 (16.8)	74.9 (12.4)	98.1 (5.0)	88 (10.7)	91.9 (8.6)	100 (0.0)	100 (0.0)	98.2 (4.0)	100 (0.0)	100 (0.0)	<b>99.6 (1.3)</b>
<b>MCD Gaussian</b>	95.0 (10.7)	55 (15.0)	65.6 (5.5)	97.2 (5.3)	60 (17.0)	82.0 (8.5)	100 (0.0)	100 (0.0)	<b>99.4 (1.8)</b>	100 (0.0)	100 (0.0)	<b>99.6 (1.3)</b>
<b>Parzen</b>	97.2 (6.7)	100 (0.0)	93.0 (15.6)	99.7 (0.8)	100 (0.0)	96.2 (6.5)	100 (0.0)	100 (0.0)	94.1 (8.6)	99.9 (0.4)	100 (0.0)	96.4 (4.5)
<b>Naive Parzen</b>	100.0 (0.0)	100 (0.0)	95.0 (15.0)	100 (0.0)	100 (0.0)	98.2 (3.6)	100 (0.0)	100 (0.0)	97.6 (4.1)	100 (0.0)	100 (0.0)	98.1 (4.3)
<b>k-NN</b>	100.0 (0.0)	98.3 (5.0)	<b>96.6 (7.0)</b>	99.7 (0.8)	100 (0.0)	<b>99.1 (2.7)</b>	100 (0.0)	100 (0.0)	98.8 (3.8)	100 (0.0)	100 (0.0)	<b>99.6 (1.3)</b>
<b>LOF</b>	100.0 (0.0)	98.3 (5.0)	<b>96.6 (7.0)</b>	99.4 (1.1)	96.7 (6.7)	96.7 (5.6)	99.9 (0.4)	94.4 (5.6)	96.2 (3.5)	100 (0.0)	100 (0.0)	99.1 (1.7)
<b>Reconstruction-Based Method</b>												
<b>PCA</b>	76.5 (13.0)	43.3 (13.4)	60.8 (10.0)	66.3 (8.3)	26.7 (13.4)	71.9 (7.1)	56.9 (11.2)	22.2 (12.2)	70.7 (5.1)	65.4 (6.8)	25 (13.5)	71.8 (4.6)
<b>Auto-encoder</b>	93.6 (11.6)	83.2 (19.6)	81.8 (17.5)	96.5 (7.2)	86.2 (19.3)	89.9 (10.4)	97.4 (5.9)	86.7 (19.4)	91.1 (9.5)	98.2 (4.3)	89.4 (16.6)	93.7 (7.1)
<b>SOM</b>	97.1 (6.1)	100 (0.0)	82.1 (19.6)	99.7 (1.2)	100 (0.0)	90.4 (13.2)	100 (0.0)	100 (0.0)	97.3 (4.1)	99.9 (0.4)	100 (0.0)	97.4 (4.6)
<b>K-means</b>	100 (0.0)	100 (0.0)	<b>98.0 (6.0)</b>	100 (0.0)	100 (0.0)	<b>99.1 (2.7)</b>	100 (0.0)	100 (0.0)	<b>98.0 (6.0)</b>	100 (0.0)	100 (0.0)	<b>99.6 (1.3)</b>

### 3. The Third Case of Infection (Flu)

Table 3 (a): Average and standard deviation of AUC, specificity, F1-score for the raw dataset (without smoothing) and different sample size.

Fraction = 0.01												
Models	<u>Boundary and Domain-Based Method</u>											
	1 Month			2 Months			3 Months			4 Months		
	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1
SVDD	79.7 (6.8)	66.7 (0.0)	78.2 (14.2)	86.4 (4.7)	66.7 (0.0)	90.6 (3.7)	84.6 (4.4)	66.7 (0.0)	88.3 (3.9)	85.4 (4.7)	66.7 (0.0)	90.5 (4.3)
incsvdd	83.1 (9.2)	66.7 (0.0)	79.6 (14.5)	85.9 (4.7)	66.7 (0.0)	90.6 (3.7)	84.9 (4.0)	66.7 (0.0)	88.3 (3.9)	85.6 (4.9)	66.7 (0.0)	90.5 (4.3)
<i>v-SVM</i>	83.1 (7.8)	66.7 (0.0)	83.7 (5.8)	82.9 (5.1)	66.7 (0.0)	86.3 (4.2)	84.1 (4.8)	66.7 (0.0)	86.0 (2.6)	84.7 (5.9)	66.7 (0.0)	87.1 (2.4)
Nearest Neighbour	89.4 (11.8)	70.0 (10.0)	85.6 (7.6)	96.7 (5.1)	73.3 (13.4)	<b>92.1 (3.8)</b>	81.2 (13.7)	3.3 (10.0)	75.9 (2.1)	79.3 (9.6)	3.3 (10.0)	80.3 (0.6)
MST	95.6 (7.4)	66.7 (0.0)	<b>84.1 (5.9)</b>	94.0 (3.7)	66.7 (0.0)	90.5 (5.9)	89.0 (4.7)	66.7 (0.0)	<b>89.5 (1.8)</b>	92.0 (4.1)	66.7 (0.0)	<b>91.5 (1.8)</b>
<u>Density-Based Method</u>												
Gaussian	87.2 (7.9)	66.7 (0.0)	<b>82.2 (7.9)</b>	87.5 (6.7)	66.7 (0.0)	89.7 (4.2)	87.7 (5.7)	66.7 (0.0)	88.9 (3.7)	88.2 (4.3)	66.7 (0.0)	90.5 (4.3)
MoG	90.6 (8.8)	68.7 (7.9)	76.3 (14.8)	85.6 (5.1)	66.7 (0.0)	87.9 (5.9)	88.6 (5.5)	66.7 (0.0)	88.2 (4.0)	90.5 (3.8)	66.7 (0.0)	90.7 (3.9)
MCD Gaussian	87.2 (7.9)	66.7 (0.0)	<b>82.2 (7.9)</b>	87.5 (6.7)	66.7 (0.0)	90.6 (3.7)	88.0 (5.9)	66.7 (0.0)	88.9 (3.7)	88.2 (4.3)	66.7 (0.0)	91.0 (2.9)
Parzen	86.4 (8.7)	66.7 (0.0)	77.7 (14.8)	85.4 (6.4)	66.7 (0.0)	87.8 (6.5)	85.8 (7.1)	66.7 (0.0)	86.0 (4.0)	87.4 (5.0)	66.7 (0.0)	90.6 (3.0)
Naive Parzen	79.7 (6.8)	66.7 (0.0)	77.7 (14.8)	80.3 (3.6)	66.7 (0.0)	87.9 (6.3)	77.7 (6.3)	66.7 (0.0)	88.9 (2.4)	82.2 (6.2)	66.7 (0.0)	91.0 (4.2)
k-NN	81.9 (7.6)	66.7 (0.0)	81.5 (13.9)	82.1 (4.4)	66.7 (0.0)	90.6 (3.7)	86.2 (7.0)	66.7 (0.0)	88.9 (2.4)	87.9 (5.0)	66.7 (0.0)	91.0 (4.2)
LOF	81.1 (11.2)	66.7 (0.0)	81.5 (13.9)	82.1 (6.1)	66.7 (0.0)	<b>91.5 (2.8)</b>	84.7 (6.7)	66.7 (0.0)	<b>89.5 (1.8)</b>	86.3 (4.9)	66.7 (0.0)	<b>91.9 (1.3)</b>
<u>Reconstruction-Based Method</u>												
PCA	92.5 (7.1)	66.7 (0.0)	<b>84.1 (5.9)</b>	94.1 (4.8)	66.7 (0.0)	<b>89.7 (4.2)</b>	93.4 (2.6)	66.7 (0.0)	<b>89.5 (1.8)</b>	93.4 (3.5)	66.7 (0.0)	<b>91.0 (2.9)</b>
Auto-encoder	86.4 (13.3)	68.8 (13.8)	81.0 (13.3)	84.2 (10.0)	65.2 (8.4)	89.1 (6.2)	83.3 (10.7)	60 (16.0)	86.5 (5.3)	85.4 (9.6)	59.7 (16.2)	89.7 (4.4)
SOM	94.7 (7.8)	66.8 (2.4)	80.5 (11.9)	93.6 (4.8)	67.5 (5.2)	88.4 (6.3)	93.4 (4.8)	66.7 (0.0)	87.7 (3.7)	94.2 (3.9)	66.7 (0.0)	<b>91.3 (2.6)</b>
K-means	94.6 (8.1)	70.5 (10.7)	82.2 (10.7)	90.4 (8.8)	67.0 (3.3)	88.5 (6.7)	92.2 (5.8)	66.7 (0.0)	88.5 (3.0)	94.4 (5.0)	66.7 (0.0)	<b>91.1 (2.8)</b>

Table 3 (b): Average and standard deviation of AUC, specificity, F1-score for smoothed version of the data with a two-days moving average filter and different sample size.

Fraction = 0.01												
Models	<u>Boundary and Domain-Based Method</u>											
	1 Month			2 Months			3 Months			4 Months		
	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1
SVDD	100 (0.0)	100 (0.0)	93.6 (15.2)	100 (0.0)	100 (0.0)	96.2 (6.5)	100 (0.0)	100 (0.0)	96.8 (6.1)	100 (0.0)	100 (0.0)	98.1 (4.3)
incsvdd	100 (0.0)	100 (0.0)	93.6 (15.2)	100 (0.0)	100 (0.0)	96.2 (6.5)	100 (0.0)	100 (0.0)	96.8 (6.1)	100 (0.0)	100 (0.0)	98.1 (4.3)
<i>v-SVM</i>	100 (0.0)	100 (0.0)	<b>99.6 (2.5)</b>	100 (0.0)	100 (0.0)	<b>98.9 (3.5)</b>	100 (0.0)	100 (0.0)	<b>99.3 (2.2)</b>	100 (0.0)	100 (0.0)	<b>99.4 (1.9)</b>
Nearest Neighbour	95.6 (13.4)	73.3 (41.7)	84.3 (18.7)	96.2 (3.5)	43.3 (21.4)	87.9 (4.6)	93.7 (7.8)	3.3 (10.0)	74.8 (3.3)	92.5 (8.3)	3.3 (10.0)	79.8 (3.3)
MST	100 (0.0)	100 (0.0)	95.0 (15.0)	100 (0.0)	100 (0.0)	95.1 (8.0)	100 (0.0)	100 (0.0)	98.2 (2.7)	100 (0.0)	100 (0.0)	<b>99.1 (1.7)</b>
<u>Density-Based Method</u>												
Gaussian	100 (0.0)	100 (0.0)	93.0 (15.6)	100 (0.0)	100 (0.0)	96.0 (8.0)	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	<b>99.1 (2.7)</b>
MoG	100 (0.0)	100 (0.0)	<b>92.8 (12.9)</b>	100 (0.0)	100 (0.0)	92.5 (13.0)	100 (0.0)	100 (0.0)	96.0 (6.6)	100 (0.0)	100 (0.0)	97.5 (4.4)
MCD Gaussian	100 (0.0)	100 (0.0)	95.0 (15.0)	100 (0.0)	100 (0.0)	97.1 (6.3)	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	<b>99.1 (2.7)</b>
Parzen	100 (0.0)	100 (0.0)	91.6 (15.5)	100 (0.0)	100 (0.0)	93.1 (9.0)	100 (0.0)	100 (0.0)	92.9 (8.5)	100 (0.0)	100 (0.0)	97.3 (4.3)
Naive Parzen	100 (0.0)	100 (0.0)	95.0 (15.0)	100 (0.0)	100 (0.0)	97.1 (6.3)	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	98.6 (4.3)
k-NN	100 (0.0)	100 (0.0)	95.0 (15.0)	100 (0.0)	100 (0.0)	<b>98.0 (6.0)</b>	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	<b>99.1 (2.7)</b>
LOF	100 (0.0)	100 (0.0)	95.0 (15.0)	100 (0.0)	100 (0.0)	<b>98.0 (6.0)</b>	100 (0.0)	100 (0.0)	<b>98.8 (2.4)</b>	100 (0.0)	100 (0.0)	<b>99.1 (2.7)</b>
<u>Reconstruction-Based Method</u>												
PCA	100 (0.0)	100 (0.0)	<b>96.0 (8.0)</b>	100 (0.0)	100 (0.0)	<b>99.1 (2.7)</b>	100 (0.0)	100 (0.0)	<b>99.4 (1.8)</b>	100 (0.0)	96.7 (10.0)	97.9 (2.7)

<b>Auto-encoder</b>	99.2 (4.9)	99.5 (4.1)	93.2 (14.5)	98.5 (5.3)	94.3 (15.4)	93.7 (9.9)	98.8 (5.2)	94.3 (15.7)	95.5 (7.3)	99.1 (3.8)	93.7 (16.8)	97.0 (4.8)
<b>SOM</b>	100 (0.0)	100 (0.0)	90.3 (17.8)	100 (0.0)	100 (0.0)	91.3 (13.5)	100 (0.0)	100 (0.0)	93.5 (9.3)	100 (0.0)	100 (0.0)	97.3 (3.8)
<b>K-means</b>	100 (0.0)	100 (0.0)	95.0 (15.0)	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	<b>99.1 (2.7)</b>

#### 4. The Fourth Case of Infection (Flu)

**Table 4 (a):** Average and standard deviation of AUC, specificity, F1-score for the raw dataset (without smoothing) and different sample size.

<b>Fraction = 0.01</b>												
<b>Models</b>	<b><u>Boundary and Domain-Based Method</u></b>											
	<b>1 Month</b>			<b>2 Months</b>			<b>3 Months</b>			<b>4 Months</b>		
	<b>AUC</b>	<b>Specificity</b>	<b>F1</b>	<b>AUC</b>	<b>Specificity</b>	<b>F1</b>	<b>AUC</b>	<b>Specificity</b>	<b>F1</b>	<b>AUC</b>	<b>Specificity</b>	<b>F1</b>
<b>SVDD</b>	100 (0.0)	100 (0.0)	<b>94.0 (9.2)</b>	97.1 (3.3)	75.0 (0.0)	88.8 (4.5)	98.7 (1.4)	85.7 (0.0)	92.5 (2.8)	98.4 (1.8)	75.0 (0.0)	90.6 (2.9)
<b>incsvdd</b>	100 (0.0)	100 (0.0)	<b>94.0 (9.2)</b>	96.7 (4.1)	75.0 (0.0)	<b>89.7 (4.2)</b>	98.6 (1.3)	85.7 (0.0)	<b>93.0 (2.7)</b>	97.9 (2.1)	75.0 (0.0)	<b>91.0 (2.9)</b>
<b>v-SVM</b>	97.9 (3.6)	75.0 (0.0)	85.7 (5.4)	96.5 (4.7)	74.8 (2.5)	88.6 (3.8)	98.4 (2.0)	85.7 (0.0)	<b>93.2 (2.1)</b>	97.6 (3.0)	75.0 (0.0)	89.5 (1.8)
<b>Nearest Neighbour</b>	90.6 (13)	5.9 (13.3)	59.1 (5.3)	69.5 (14.1)	8.5 (10.4)	74.4 (2.7)	73.5 (12.4)	20.0 (27.3)	76.5 (8.1)	78.1 (10.7)	27.5 (26.2)	81.1 (6.6)
<b>MST</b>	99.2 (2.5)	75.0 (0.0)	84.1 (5.9)	98.8 (2.7)	55.0 (15.0)	85.5 (3.3)	98.2 (3.3)	65.7 (17.2)	87.5 (3.5)	98.8 (2.1)	60.0 (20.1)	87.7 (4.1)
<b><u>Density-Based Method</u></b>												
<b>Gaussian</b>	98.3 (3.3)	82.5 (11.5)	86.4 (9.0)	97.5 (2.8)	75.0 (0.0)	90.6 (3.7)	99.1 (1.0)	85.7 (0.0)	<b>93.7 (2.1)</b>	98.8 (1.7)	75.0 (0.0)	91.5 (1.6)
<b>MoG</b>	99.1 (3.4)	97.1 (8.0)	88.8 (13.4)	99.1 (2.1)	90.5 (12.2)	95.2 (4.8)	98.7 (2.5)	92.1 (7.1)	94.4 (6.1)	98.8 (2.7)	87.4 (12.5)	93.9 (6.0)
<b>MCD Gaussian</b>	97.7 (3.6)	77.5 (7.5)	<b>85.5 (2.1)</b>	97.9 (2.8)	75.0 (0.0)	90.6 (3.7)	99.1 (1.0)	85.7 (0.0)	<b>93.7 (2.1)</b>	99.0 (1.7)	75.0 (0.0)	<b>92.0 (1.0)</b>
<b>Parzen</b>	96.0 (4.0)	80.0 (10.0)	83.0 (6.0)	95.8 (5.6)	75.0 (0.0)	89.7 (4.2)	98.2 (2.6)	87.1 (4.3)	92.2 (6.0)	99.0 (1.6)	87.5 (12.5)	93.0 (4.0)
<b>Naive Parzen</b>	93.5 (6.2)	75.0 (0.0)	80.3 (9.0)	93.3 (9.7)	75.0 (0.0)	88.7 (6.4)	96.1 (6.3)	77.1 (13.1)	89.9 (7.0)	96.3 (4.6)	70.0 (10.0)	89.8 (3.3)
<b>k-NN</b>	98.3 (3.3)	75.0 (0.0)	84.1 (5.9)	94.6 (4.6)	72.5 (7.5)	90.9 (3.2)	98.0 (2.7)	82.9 (8.6)	92.7 (3.3)	97.5 (3.1)	75.0 (0.0)	91.1 (2.8)
<b>LOF</b>	96.0 (4.0)	75.0 (0.0)	84.1 (5.9)	95.8 (4.9)	75.0 (0.0)	<b>91.5 (2.8)</b>	98.5 (2.1)	82.9 (8.6)	93.4 (2.9)	98.6 (2.4)	75.0 (0.0)	<b>92.0 (1.0)</b>
<b><u>Reconstruction-Based Method</u></b>												
<b>PCA</b>	99.2 (2.5)	100 (0.0)	<b>96.0 (8.0)</b>	98.8 (1.9)	85.0 (12.3)	<b>92.8 (5.8)</b>	99.4 (1.0)	88.6 (5.7)	93.5 (3.5)	99.4 (1.3)	85.0 (12.3)	<b>94.1 (4.0)</b>
<b>Auto-encoder</b>	95.7 (9.1)	81.3 (25.4)	83.7 (14.5)	92.6 (11.7)	63.9 (32.3)	87.6 (9.0)	95.9 (8.0)	68.1 (33.9)	88 (10.1)	96.2 (9.2)	66.0 (31.7)	89.2 (8.1)
<b>SOM</b>	99.2 (2.5)	84.4 (12.1)	89.4 (6.9)	96.8 (3.8)	74.9 (6.4)	89.4 (4.3)	98.6 (1.8)	83.3 (8.0)	92.7 (3.1)	98.4 (2.6)	75.0 (0.0)	91.0 (2.3)
<b>K-means</b>	96.9 (3.9)	75.0 (0.0)	84.1 (5.9)	95.8 (4.9)	75.0 (0.0)	90.6 (3.7)	98.0 (2.3)	85.7 (0.0)	<b>94.2 (1.8)</b>	96.7 (3.5)	75.0 (0.0)	91.5 (2.7)

**Table 4 (b):** Average and standard deviation of AUC, specificity, F1-score for smoothed version of the data with a two-days moving average filter and different sample size.

<b>Fraction = 0.01</b>												
<b>Models</b>	<b><u>Boundary and Domain-Based Method</u></b>											
	<b>1 Month</b>			<b>2 Months</b>			<b>3 Months</b>			<b>4 Months</b>		
	<b>AUC</b>	<b>Specificity</b>	<b>F1</b>	<b>AUC</b>	<b>Specificity</b>	<b>F1</b>	<b>AUC</b>	<b>Specificity</b>	<b>F1</b>	<b>AUC</b>	<b>Specificity</b>	<b>F1</b>
<b>SVDD</b>	100 (0.0)	100 (0.0)	93 (15.6)	100 (0.0)	100 (0.0)	95.1 (8.0)	100 (0.0)	100 (0.0)	96.2 (6.0)	100 (0.0)	100 (0.0)	97.7 (4.4)
<b>incsvdd</b>	100 (0.0)	100 (0.0)	93.0 (15.6)	100 (0.0)	100 (0.0)	96.0 (8.0)	100 (0.0)	100 (0.0)	98.2 (2.7)	100 (0.0)	100 (0.0)	<b>99.1 (1.7)</b>
<b>v-SVM</b>	100 (0.0)	100 (0.0)	<b>98.9 (3.7)</b>	100 (0.0)	100 (0.0)	<b>99.0 (3.0)</b>	100 (0.0)	100 (0.0)	<b>99.3 (2.1)</b>	100 (0.0)	100 (0.0)	<b>99.5 (1.7)</b>
<b>Nearest Neighbour</b>	90.0 (12.3)	15 (30.1)	65.8 (10.2)	93.3 (7.5)	47.5 (17.5)	82.8 (5.6)	97.8 (2.6)	60.0 (5.7)	85.5 (2.3)	97.5 (3.1)	62.5 (12.5)	87.8 (4.7)
<b>MST</b>	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	96.2 (6.5)	100 (0.0)	100 (0.0)	<b>99.4 (1.8)</b>	100 (0.0)	100 (0.0)	97.1 (6.0)
<b><u>Density-Based Method</u></b>												
<b>Gaussian</b>	100 (0.0)	100 (0.0)	89.0 (15.8)	100 (0.0)	100 (0.0)	95.0 (15.0)	100 (0.0)	100 (0.0)	<b>98.0 (6.0)</b>	100 (0.0)	100 (0.0)	97.4 (7.9)

MoG	100 (0.0)	100 (0.0)	84.5 (22.6)	100 (0.0)	100 (0.0)	91.8 (12.3)	99.9 (0.6)	100 (0.0)	93.7 (8.2)	100 (0.3)	100 (0.0)	95.7 (8.1)
MCD Gaussian	100 (0.0)	100 (0.0)	95.0 (15.0)	100 (0.0)	100 (0.0)	96.7 (10.0)	100 (0.0)	100 (0.0)	<b>98.0 (6.0)</b>	100 (0.0)	100 (0.0)	97.4 (7.9)
Parzen	100 (0.0)	100 (0.0)	91.0 (15.8)	100 (0.0)	100 (0.0)	90.4 (13.4)	100 (0.0)	100 (0.0)	92.9 (8.0)	100 (0.0)	100 (0.0)	95.1 (8.0)
Naive Parzen	100 (0.0)	100 (0.0)	93.0 (15.6)	100 (0.0)	100 (0.0)	95.1 (8.0)	100 (0.0)	100 (0.0)	96.8 (6.7)	100 (0.0)	100 (0.0)	97.2 (4.8)
k-NN	100 (0.0)	100 (0.0)	96.0 (8.0)	100 (0.0)	100 (0.0)	<b>98.2 (3.6)</b>	100 (0.0)	100 (0.0)	<b>98.0 (6.0)</b>	100 (0.0)	100 (0.0)	<b>98.6 (4.3)</b>
LOF	100 (0.0)	100 (0.0)	<b>98.0 (6.0)</b>	100 (0.0)	100 (0.0)	95.8 (10.1)	100 (0.0)	100 (0.0)	<b>98.0 (6.0)</b>	100 (0.0)	100 (0.0)	97.4 (7.9)
<b>Reconstruction-Based Method</b>												
PCA	100 (0.0)	100 (0.0)	<b>93.0 (15.6)</b>	100 (0.0)	100 (0.0)	95.0 (15.0)	99.4 (1.9)	94.3 (11.5)	96.0 (6.6)	99.8 (0.6)	100 (0.0)	97.4 (7.9)
Auto-encoder	97.9 (7.5)	94.8 (15.6)	90.1 (15.9)	98.9 (4.8)	95.4 (14.0)	92.8 (12.1)	99.0 (3.6)	96.5 (10.6)	95.5 (7.3)	99.0 (3.8)	93.5 (17.6)	95.4 (7.7)
SOM	100 (0.0)	100 (0.0)	83.8 (22.5)	100 (0.0)	100 (0.0)	82.8 (23.3)	100 (0.0)	100 (0.0)	92.4 (9.1)	100 (0.0)	100 (0.0)	95.5 (8.3)
K-means	100 (0.0)	100 (0.0)	<b>93.0 (15.6)</b>	100 (0.0)	100 (0.0)	<b>97.1 (6.3)</b>	100 (0.0)	100 (0.0)	<b>97.5 (6.0)</b>	100 (0.0)	100 (0.0)	<b>98.2 (4.3)</b>

## II. Hourly

### 1. The First Case of Infection (Flu)

Table 5: Average and standard deviation of AUC, specificity, F1-score for smoothed version of the data with 48 hours moving average filter and different sample size.

<b>Fraction = 0.01</b>												
Models	<b>Boundary and Domain-Based Method</b>											
	1 Month			2 Months			3 Months			4 Months		
	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1
SVDD	97.6 (1.9)	83.2 (3.4)	85.8 (1.7)	97.8 (1.2)	85.7 (5.0)	90.5 (9.6)	97.7 (1.2)	90.4 (5.1)	94.2 (2.9)	98.1 (0.9)	91.0 (3.7)	<b>96.8 (0.9)</b>
incsvdd	97.4 (1.9)	84.5 (2.8)	86.8 (1.9)	97.7 (1.2)	86.7 (2.0)	93.9 (1.0)	97.5 (1.2)	88.5 (1.5)	<b>96.0 (1.1)</b>	97.9 (0.9)	88.9 (1.2)	<b>97.0 (0.7)</b>
<i>v-SVM</i>	98.1 (2.1)	84.5 (1.1)	<b>90.5 (1.1)</b>	99.0 (1.1)	92.6 (0.0)	<b>96.1 (1.3)</b>	99.5 (0.6)	93.8 (0.5)	<b>96.9 (1.4)</b>	99.4 (0.4)	94.2 (0.0)	<b>97.1 (1.3)</b>
Nearest Neighbour	84.8 (6.0)	75.9 (4.5)	74.8 (6.0)	89.3 (2.2)	76.5 (4.1)	87.1 (3.3)	89.0 (4.0)	77.5 (3.9)	89.3 (4.4)	90.2 (4.7)	77.5 (3.8)	91.4 (6.4)
MST	90.5 (3.1)	85.4 (3.9)	67.6 (14.5)	94.4 (2.0)	85.7 (4.0)	85.1 (7.0)	94.7 (2.4)	88.8 (3.5)	87.8 (8.5)	95.8 (2.2)	88.8 (3.0)	90.9 (5.9)
<b>Density-Based Method</b>												
Gaussian	98.1 (2.2)	79.8 (4.9)	83.9 (2.7)	99.5 (0.9)	90.1 (1.7)	95.2 (1.8)	99.6 (0.7)	92.9 (1.3)	<b>97.1 (2.5)</b>	99.5 (0.5)	92.2 (1.0)	97.7 (1.1)
MoG	95.8 (3.6)	82.7 (4.3)	83.7 (5.0)	98.3 (1.5)	86.2 (2.7)	92.3 (2.7)	98.7 (1.4)	88.7 (4.6)	94.7 (3.5)	98.6 (1.6)	88.2 (3.1)	95.3 (3.2)
MCD Gaussian	98.6 (2.1)	75.3 (6.9)	81.3 (2.5)	99.6 (0.9)	89.6 (1.9)	95.0 (1.8)	99.6 (0.7)	92.5 (1.8)	<b>97.0 (2.3)</b>	99.6 (0.4)	92.0 (1.2)	97.7 (1.1)
Parzen	91.9 (2.9)	93.6 (2.0)	63.4 (16.5)	96.2 (2.3)	94.4 (2.0)	81.6 (10.2)	96.6 (2.6)	94.8 (1.7)	84.2 (9.5)	97.4 (2.2)	95.6 (1.2)	87.9 (7.1)
Naive Parzen	94.8 (3.7)	76.4 (5.6)	77.6 (7.9)	98.7 (1.2)	85.2 (3.3)	91.8 (2.9)	99.1 (1.1)	89.1 (3.8)	94.8 (2.5)	98.9 (0.9)	89.7 (2.4)	96.2 (1.6)
k-NN	97.1 (3.4)	78.8 (2.0)	<b>84.2 (2.1)</b>	99.1 (1.0)	92.9 (0.7)	<b>96.0 (1.8)</b>	99.6 (0.4)	93.8 (0.7)	<b>97.3 (1.9)</b>	99.5 (0.3)	94.0 (0.6)	<b>98.2 (0.9)</b>
LOF	96.9 (3.5)	78.3 (3.0)	<b>84.2 (2.4)</b>	99.2 (1.1)	91.9 (0.9)	<b>96.0 (1.8)</b>	99.6 (0.5)	93.7 (0.8)	<b>97.3 (2.1)</b>	99.5 (0.4)	93.1 (0.4)	97.8 (1.2)
<b>Reconstruction-Based Method</b>												
PCA	97.1 (3.4)	63.9 (8.8)	75.4 (0.3)	99.4 (1.2)	76.4 (6.6)	90.2 (1.1)	99.1 (1.3)	75.1 (6.8)	92.4 (1.1)	98.9 (1.2)	69.1 (4.1)	93.1 (0.8)
Auto-encoder	92.0 (4.8)	79.5 (7.6)	78.9 (8.3)	96.2 (2.6)	83.1 (7.2)	91.1 (3.9)	96.3 (3.2)	84.3 (7.7)	92.7 (5.0)	96.7 (3.0)	84.0 (8.0)	94.6 (4.4)
SOM	94.1 (2.3)	82.2 (3.3)	82.6 (4.9)	95.6 (1.1)	82.9 (3.1)	91.6 (1.9)	94.8 (2.3)	83.4 (5.8)	92.3 (4.1)	95.5 (1.9)	84.1 (3.8)	94.3 (3.8)
K-means	97.3 (3.2)	80.9 (2.5)	<b>85.5 (2.5)</b>	98.9 (1.1)	92.6 (0.7)	<b>95.8 (1.8)</b>	99.3 (0.6)	92.9 (0.7)	<b>97.3 (1.4)</b>	99.4 (0.4)	94.1 (0.2)	<b>98.1 (1.1)</b>

### 2. The Second Case of Infection (Flu)

Table 6: Average and standard deviation of AUC, specificity, F1-score for smoothed version of the data with 48 hours moving average filter and different sample size.



Fraction = 0.01

Models	<u>Boundary and Domain-Based Method</u>											
	1 Month			2 Months			3 Months			4 Months		
	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1
SVDD	100 (0.0)	100 (0.0)	96.3 (7.7)	100 (0.0)	100 (0.0)	96.9 (5.9)	98.5 (3.0)	76.9 (7.8)	91.9 (2.6)	98.8 (2.2)	81.0 (7.2)	94.4 (2.4)
incsvdd	100 (0.0)	100 (0.0)	98.1 (4.4)	100 (0.0)	100 (0.0)	97.4 (5.3)	98.6 (3.3)	65.7 (11.6)	89.4 (2.1)	99.0 (2.2)	74.3 (8.6)	93.6 (1.7)
<i>v-SVM</i>	100 (0.0)	100 (0.0)	<b>99.6 (0.9)</b>	100 (0.0)	100 (0.0)	<b>99.5 (1.6)</b>	99.7 (0.8)	98.7 (0.4)	<b>98.8 (1.9)</b>	99.8 (0.6)	99.6 (0.4)	<b>99.1 (2.0)</b>
Nearest Neighbour	97.6 (4.9)	100 (0.0)	85.9 (14.1)	99.4 (0.8)	100 (0.0)	94.9 (4.1)	97.6 (1.1)	92.5 (3.4)	94.3 (2.2)	98.1 (0.9)	92.1 (3.2)	95.8 (2.5)
MST	100 (0.0)	100 (0.0)	78.6 (16.1)	100 (0.0)	100 (0.0)	87.1 (8.6)	97.8 (1.5)	99.3 (0.2)	91.6 (6.1)	98.4 (1.3)	99.3 (0.2)	93.0 (4.2)
<u>Density-Based Method</u>												
Gaussian	100 (0.0)	100 (0.0)	<b>97.5 (7.6)</b>	100 (0.0)	100 (0.0)	97.8 (4.7)	99.0 (1.6)	79.3 (7.6)	92.8 (1.8)	99.6 (0.7)	84.9 (5.8)	95.6 (1.9)
MoG	100 (0.0)	99.4 (1.7)	93.8 (10.9)	100 (0.0)	100 (0.0)	96.7 (6.1)	98.0 (2.8)	78.2 (10.5)	91.3 (3.2)	98.7 (1.9)	82.0 (8.7)	94.1 (3.2)
MCD Gaussian	100 (0.0)	100 (0.0)	<b>97.5 (7.6)</b>	100 (0.0)	100 (0.0)	<b>98.1 (4.1)</b>	99.2 (1.3)	78.2 (6.1)	92.8 (1.9)	99.6 (0.6)	86.1 (6.5)	95.9 (1.8)
Parzen	100 (0.0)	100 (0.0)	58.9 (16.4)	100 (0.0)	100 (0.0)	79.8 (13.1)	98.3 (1.7)	99.3 (0.2)	84.4 (10.2)	98.9 (1.3)	99.3 (0.2)	87.8 (6.6)
Naive Parzen	100 (0.0)	100 (0.0)	94.8 (10.2)	100 (0.0)	100 (0.0)	95.8 (8.0)	99.0 (2.4)	95.0 (3.1)	<b>96.2 (4.0)</b>	98.9 (1.9)	98.0 (1.5)	<b>97.1 (3.2)</b>
k-NN	100 (0.0)	100 (0.0)	<b>97.5 (7.6)</b>	100 (0.0)	100 (0.0)	95.8 (7.6)	98.7 (1.4)	95.5 (2.3)	95.1 (4.0)	99.2 (1.1)	94.6 (2.3)	96.1 (3.5)
LOF	100 (0.0)	100 (0.0)	94.1 (11.2)	100 (0.0)	100 (0.0)	95.6 (3.8)	90.0 (4.9)	77.9 (6.9)	88.7 (5.6)	91.8 (2.2)	79.1 (6.3)	91.4 (3.4)
<u>Reconstruction-Based Method</u>												
PCA	95.9 (4.4)	65.5 (6.4)	76.5 (1.3)	91.7 (7.3)	54.2 (4.7)	82.5 (4.5)	74.4 (8.1)	33.7 (2.9)	82.8 (2.8)	77.1 (6.1)	32.1 (2.2)	86.7 (1.0)
Auto-encoder	99.5 (1.8)	97.1 (9.2)	92.5 (10.8)	99.4 (2.3)	97.7 (7.0)	95.7 (6.3)	93.4 (7.9)	69.4 (19.7)	90.0 (5.2)	95.1 (5.9)	71 (20.9)	92.6 (4.1)
SOM	100 (0.0)	100 (0.0)	94.7 (9.0)	100 (0.0)	100 (0.0)	96.2 (5.4)	93.8 (5.5)	69.0 (18.9)	88.9 (5.9)	99.1 (2.3)	95.5 (10.9)	97.4 (3.7)
K-means	100 (0.0)	100 (0.0)	<b>99.3 (1.8)</b>	100 (0.0)	100 (0.0)	<b>98.0 (4.3)</b>	99.7 (0.8)	98.4 (0.6)	<b>98.5 (3.2)</b>	99.8 (0.6)	99.3 (0.3)	<b>99.0 (2.4)</b>

### 3. The Third Case of Infection (Flu)

Table 7: Average and standard deviation of AUC, specificity, F1-score for smoothed version of the data with 48 hours moving average filter and different sample size.

Fraction = 0.01												
Models	<u>Boundary and Domain-Based Method</u>											
	1 Month			2 Months			3 Months			4 Months		
	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1
SVDD	93.5 (5.3)	79.8 (5.3)	84.8 (8.2)	96.3 (2.4)	77.7 (4.2)	91.8 (4.6)	95.9 (3.6)	69.8 (4.3)	<b>93.3 (3.0)</b>	96.3 (2.3)	68.9 (3.0)	<b>95.1 (1.3)</b>
incsvdd	92.8 (5.2)	74.4 (3.6)	82.8 (7.3)	95.7 (2.4)	76.7 (4.1)	91.4 (4.5)	95.2 (3.4)	67.4 (3.1)	<b>93.0 (2.3)</b>	95.5 (2.2)	69.6 (2.4)	<b>94.9 (1.5)</b>
<i>v-SVM</i>	94.7 (3.3)	64.1 (0.4)	<b>83.2 (1.7)</b>	97.7 (2.1)	80.2 (0.3)	<b>91.7 (1.6)</b>	96.5 (2.4)	67.4 (0.6)	87.8 (2.3)	97.4 (1.5)	75.3 (0.2)	90.8 (1.9)
Nearest Neighbour	91.8 (12.7)	92.1 (1.4)	87.0 (8.8)	96.2 (1.3)	92.1 (0.8)	92.9 (5.8)	92.9 (3.1)	81.6 (3.8)	92.4 (5.6)	93.6 (2.3)	81.5 (3.4)	<b>94.9 (2.8)</b>
MST	94.3 (4.1)	92.2 (1.9)	74.6 (22.3)	97.0 (1.8)	94.2 (0.8)	85.9 (14.8)	94.0 (3.0)	89.4 (1.1)	91.1 (6.5)	95.6 (1.3)	89.4 (1.1)	<b>93.7 (4.2)</b>
<u>Density-Based Method</u>												
Gaussian	96.0 (2.1)	88.3 (1.1)	86.0 (18.6)	98.3 (1.5)	89.4 (1.5)	<b>94.8 (6.5)</b>	97.1 (1.7)	87.0 (0.6)	<b>96.5 (1.8)</b>	97.9 (1.2)	88.0 (0.6)	<b>97.7 (1.0)</b>
MoG	95.0 (4.5)	85.1 (4.3)	84.1 (16.2)	98.1 (1.6)	88.2 (2.5)	93.1 (7.9)	96.6 (2.5)	82.1 (5.5)	94.7 (4.1)	97.2 (1.7)	87.1 (1.6)	96.6 (2.5)
MCD Gaussian	95.9 (2.1)	88.1 (1.3)	85.9 (18.6)	98.7 (1.4)	89.4 (1.5)	<b>94.9 (6.2)</b>	97.0 (1.6)	86.9 (0.6)	<b>96.5 (1.7)</b>	98.0 (1.2)	88.1 (0.6)	<b>97.8 (0.9)</b>
Parzen	94.1 (4.2)	97.8 (0.5)	58.8 (24.5)	97.2 (1.9)	96.5 (0.5)	78.0 (20.0)	95.1 (3.2)	91.6 (1.7)	86.2 (9.3)	96.7 (1.8)	92.0 (1.7)	89.4 (7.1)
Naive Parzen	93.4 (4.2)	80.5 (3.4)	83.9 (10.9)	97.3 (2.0)	81.2 (4.2)	90.5 (8.9)	95.2 (3.4)	80.1 (0.4)	93.8 (4.2)	95.9 (2.3)	80.2 (0.0)	94.9 (2.5)
k-NN	94.0 (4.5)	85.7 (2.4)	84.1 (17.7)	97.2 (2.4)	88.0 (2.7)	92.5 (7.7)	95.7 (2.4)	85.4 (0.5)	93.9 (6.4)	96.5 (1.8)	85.6 (0.6)	96.7 (2.6)
LOF	91.9 (4.3)	86.7 (1.2)	<b>80.9 (11.2)</b>	93.5 (2.8)	84.0 (7.1)	90.9 (4.9)	92.1 (3.0)	84.4 (2.3)	92.3 (5.4)	93.8 (2.2)	83.7 (4.5)	95.4 (3.3)

	<b>Reconstruction-Based Method</b>											
PCA	97.5 (2.6)	87.5 (1.4)	<b>88.4 (14.5)</b>	98.6 (1.2)	89.1 (1.2)	<b>94.5 (5.2)</b>	97.2 (1.4)	85.1 (0.9)	<b>96.7 (1.5)</b>	97.9 (1.1)	86.7 (0.5)	<b>97.5 (1.2)</b>
Auto-encoder	91.6 (8.2)	81.7 (12.4)	84.4 (14.0)	95.4 (3.5)	84.5 (8.9)	92.1 (8.0)	93.6 (4.0)	79.4 (10.9)	94.4 (4.1)	94.4 (4.0)	80.0 (12.9)	96.2 (2.4)
SOM	90.7 (4.8)	84.3 (4.0)	82.1 (18.1)	93.9 (2.4)	83.6 (5.1)	91.6 (7.6)	93.1 (2.7)	79.6 (5.7)	94.9 (2.8)	94.3 (2.2)	82.4 (4.3)	96.5 (1.8)
K-means	91.9 (6.9)	78.7 (12.8)	81.8 (14.7)	97.1 (2.7)	86.3 (6.4)	93.8 (5.8)	96.0 (2.6)	84.3 (4.1)	95.6 (3.4)	96.7 (1.8)	86.4 (0.9)	<b>96.9 (1.9)</b>

#### 4. The Fourth Case of Infection (Flu)

**Table 8:** Average and standard deviation of AUC, specificity, F1-score for smoothed version of the data with 48 hours moving average filter and different sample size.

<b>Fraction = 0.01</b>												
<b>Models</b>	<b>Boundary and Domain-Based Method</b>											
	<b>1 Month</b>			<b>2 Months</b>			<b>3 Months</b>			<b>4 Months</b>		
	<b>AUC</b>	<b>Specificity</b>	<b>F1</b>	<b>AUC</b>	<b>Specificity</b>	<b>F1</b>	<b>AUC</b>	<b>Specificity</b>	<b>F1</b>	<b>AUC</b>	<b>Specificity</b>	<b>F1</b>
SVDD	98.8 (1.3)	93.1 (2.4)	91.0 (8.0)	95.7 (2.4)	83.7 (5.7)	86.9 (17.4)	96.9 (2.0)	83.5 (3.3)	94.6 (4.5)	97.4 (1.1)	86.3 (5.7)	92.3 (11.4)
incsvdd	98.4 (1.6)	92.0 (2.7)	90.5 (7.6)	95.5 (2.4)	82.3 (3.0)	91.9 (6.3)	96.6 (2.2)	83.2 (2.3)	94.7 (4.0)	97.3 (1.0)	83.3 (2.6)	<b>96.1 (2.9)</b>
<i>v-SVM</i>	99.6 (0.6)	93.4 (0.0)	<b>96.0 (1.4)</b>	99.1 (1.5)	86.7 (0.3)	<b>94.1 (2.0)</b>	99.2 (1.2)	85.7 (0.1)	<b>94.2 (2.3)</b>	99.4 (0.8)	87.5 (0.5)	94.9 (2.8)
Nearest Neighbour	98.7 (3.3)	100 (0.0)	88.0 (15.5)	92.8 (4.1)	84.9 (5.2)	88.8 (6.9)	93.8 (1.9)	80.5 (3.6)	92.3 (3.6)	94.3 (1.7)	80.8 (3.7)	94.2 (2.5)
MST	99.6 (0.8)	100 (0.0)	70.8 (18.2)	97.8 (0.8)	97.0 (1.0)	86.3 (11.1)	98.1 (0.8)	96.8 (1.2)	90.5 (7.3)	98.4 (0.5)	97.0 (1.0)	92.9 (5.1)
<b>Density-Based Method</b>												
Gaussian	99.5 (1.0)	96.8 (1.5)	91.0 (13.7)	99.5 (0.8)	89.9 (2.6)	95.0 (5.1)	99.7 (0.7)	93.2 (1.6)	<b>97.2 (4.0)</b>	99.8 (0.4)	94.4 (1.4)	<b>97.9 (2.6)</b>
MoG	99.5 (1.1)	99.9 (0.6)	89.7 (16.5)	99.3 (1.1)	89.3 (3.9)	94.0 (6.7)	99.5 (0.8)	92.8 (2.8)	96.4 (5.0)	99.7 (0.5)	94.7 (2.5)	<b>97.8 (3.2)</b>
MCD Gaussian	99.5 (0.9)	96.3 (1.9)	<b>91.7 (11.2)</b>	99.7 (0.8)	89.8 (2.3)	<b>95.2 (4.3)</b>	99.7 (0.6)	94.1 (1.4)	<b>97.5 (3.7)</b>	99.8 (0.4)	95.6 (1.4)	<b>98.1 (2.6)</b>
Parzen	99.6 (0.8)	100 (0.0)	58.8 (14.8)	98.6 (0.9)	100 (0.0)	78.8 (11.4)	98.9 (0.9)	100 (0.0)	85.9 (10.0)	99.3 (0.6)	100.0 (0.0)	89.5 (7.0)
Naive Parzen	97.6 (1.9)	93.3 (2.4)	84.1 (13.1)	98.9 (1.9)	90.5 (1.4)	93.3 (8.0)	98.9 (1.4)	91.3 (1.3)	95.2 (5.1)	99.1 (1.0)	92.2 (2.1)	96.6 (3.4)
k-NN	99.3 (1.5)	100 (0.0)	84.7 (18.3)	99.8 (0.6)	88.9 (3.8)	94.0 (6.8)	99.7 (0.5)	95.9 (1.4)	96.8 (5.4)	99.8 (0.4)	99.1 (0.4)	<b>98.0 (3.4)</b>
LOF	99.1 (1.4)	100 (0.0)	85.7 (19.6)	96.8 (1.8)	83.5 (6.1)	90.1 (7.1)	98.3 (1.6)	84.9 (5.2)	93.5 (4.8)	98.3 (1.3)	85.4 (5.0)	94.2 (3.9)
<b>Reconstruction-Based Method</b>												
PCA	99.3 (1.1)	96.2 (4.5)	<b>89.8 (15.5)</b>	99.7 (0.6)	90.7 (3.2)	<b>96.3 (1.5)</b>	99.8 (0.4)	95.9 (1.4)	<b>98.3 (1.6)</b>	99.9 (0.3)	96.3 (1.5)	<b>98.6 (1.4)</b>
Auto-encoder	98.6 (3.6)	96.6 (7.0)	88.6 (16.3)	96.6 (4.1)	83 (14.1)	92.4 (6.1)	96.8 (4.6)	84.3 (15.5)	95.1 (4.4)	97.5 (3.2)	86.1 (11.3)	96.6 (2.8)
SOM	99.1 (1.9)	100 (0.2)	84.9 (18.9)	93.5 (2.2)	84.1 (5.4)	91.4 (5.9)	93.9 (2.4)	84.6 (5.3)	93.2 (4.4)	95.2 (3.4)	88.3 (5.9)	96.2 (3.1)
K-means	99.3 (1.4)	99.4 (1.7)	88.4 (18.2)	98.7 (1.9)	85.6 (5.2)	<b>93.2 (4.6)</b>	99.3 (1.1)	93.7 (3.8)	96.5 (4.3)	99.8 (0.4)	96.1 (2.4)	<b>97.9 (2.9)</b>

## Paper 4

**Woldaregay, A. Z.** *Automatic Detection of Infection State in Individuals with Type 1 Diabetes Under Free-Living Conditions Using Self-Recorded Insulin and Carbohydrate Information (Part 3)*. EARLY DRAFT MANUSCRIPT. This Manuscript contains overlapping texts and results from the dissertation.

# Automatic Detection of Infection State in Individuals with Type 1 Diabetes Under Free-Living Conditions Using Self-Recorded Insulin and Carbohydrate Information (Part 3)

**Ashenafi Zebene Woldaregay**

*Department of Computer Science, University of Tromsø – The Arctic University of Norway, Tromsø, Norway*

## Abstract

**Background:** Infection incidences are often associated with problematic BG management and people with type 1 diabetes usually face challenges to control BG in range throughout the period. A recent study has characterized infection states in people with type 1 diabetes and put forward diabetes profiles such as blood glucose levels, insulin injection, carbohydrate consumption, and the ratio of insulin-to-carbohydrate as optimal parameters to devise an infection detection model. Further, another study has devised a bivariate infection detection model and demonstrated the potential of blood glucose levels and insulin-to-carbohydrate ratio. Inspired by these previous studies, this study aims to investigate the potential of insulin and meals profile as event indicator variables by using the insulin-to-carbohydrate ratio as a sole input parameter.

**Objective:** The study aims to realize an infection detection model exploiting insulin and carbs profiles from people with type 1 diabetes under free-living conditions, where the ratio of insulin-to-carbohydrate is computed on a time bin basis to form an input feature.

**Method:** One-class classifier and unsupervised methods were tested and evaluated. The models were evaluated based on three attributes; data granularity, data nature, and sample sizes. Three different metrics; area under the ROC curve (AUC), specificity, and F1-score were used to measure (quantify) the models performances. A self-management dataset incorporating nine patient-years captured within the longitudinal records of 3 people with type 1 diabetes were used. The dataset consists of different parameters, however, only insulin and carbohydrate profiles were used as input features to the models. All experiments were performed using MATLAB®2018b (Mathworks, Inc, Natwick, MA). Further, the results are compared with previously published results.

**Result:** The models generally demonstrated excellent performance in describing the data as well as detecting the infection episodes. Generally, for a small sample size, the boundary and domain-based method, especially the v-SVM, performed better. However, with larger sample size, all three categories produced a comparable description of the data. Unsupervised models tested on the same dataset produced comparable performance, however, one of the drawbacks of the unsupervised models is that they require a pretty large data size to produce comparable performance.

**Conclusion:** Apart from its significance in outbreak detection, detecting infection episodes among people with type 1 diabetes can be useful for the individuals. In this regard, this study has demonstrated the potential of insulin and carbs profile in detecting infection episodes among this group of peoples. Generally, the proposed approaches have demonstrated superior performance in detecting deviations from the norm due to infection onset. In comparison to each particular model from their respective groups, v-SVM, K-means, KNN, and Gaussian families achieved better performance on average in all the infection states. In general, we foresee that the presented results could further encourage researchers to examine additional features on a large scale basis on top of the presented features.

**Keywords:** *Type 1 Diabetes mellitus, Self-recorded data, Detecting Infection incidence, Decision support system, Outbreak detection system, Digital infectious disease detection system.*

## 1. Introduction

Individualized surveillance refers to tracking the individual health status for detecting infectious disease outbreak among the public, and is believed to have a promising potential to revitalize the surveillance systems [26; 30; 49; 50]. In this regard, individualizing the surveillance effort can add double benefits; can monitor and notify any potential health changes to the individual and at the same time can deliver the information necessary for detecting infectious disease outbreaks [49]. In this kind of surveillance, the case detection needs to be realized by a personalized health model that can continuously examine the individual health status and detect when the individual becomes infected. The main drivers behind the conception of such a model are the rapid progress in information and communication technology, and the widespread availability of different smartphones and wearables equipped with a variety of physiological sensors, which created a suitable platform to easily self-track health data [17; 26; 28; 33; 35]. These technologies are increasingly been integrated into our daily life for a variety of reasons ranging from fitness tracking to managing diseases [3; 13; 14; 16; 17; 21; 34; 48]. As a result, a huge amount of data are being recorded on a daily basis that grows at an unprecedented rate [12; 36; 47]. The existence of these data is the cornerstone in the effort towards individualizing the surveillance systems. In this regard, for instance, recent studies have shown the feasibility of smartwatches and wearable technology in monitoring, detecting, and predicting illness [5; 22; 25; 26; 33; 45].

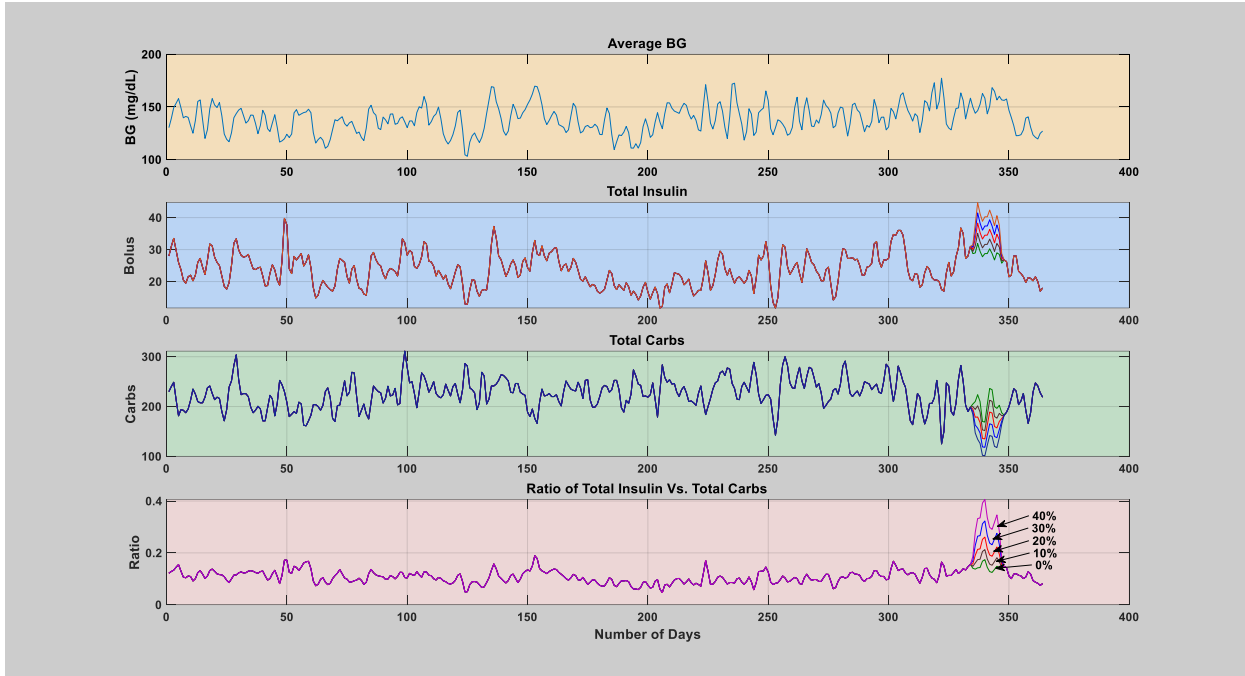
Further, self-recorded health-related data from people with type 1 diabetes could be one potential choice, which is becoming rich in both quality and quantity as a result of advancement in diabetes technologies [1; 2; 4]. Type 1 diabetes is a chronic disease that results in a lack of blood glucose control as a result of insulin deficiency within the body [7; 9]. This group of patients is expected to follow complex treatment regimens to control their blood glucose levels within the recommended targets including tracking blood glucose levels throughout the day, administrating balanced insulin while considering factors like blood glucose levels, meal intake, physical activity, and other possible factors. Recently, Woldaregay et.al. [50] has carried out a retrospective numerical analysis to study the effect of infection episodes (flu) on the key parameters of the blood glucose dynamics as an effort to pinpoint parameters for realizing a personalized health model in

individuals with type 1 diabetes. The study characterized infection states in this group of people by the presence of elevated BG levels, alteration in carbohydrate intake, and actions of the patient to lower it down e.g. enhanced and frequent insulin injection as compared to the regular days [50]. Further, Woldaregay et.al. [49] have developed a personalized health model using blood glucose levels and insulin-to-carbohydrate ratio as input features to the algorithm and demonstrated the success of the models in detecting infection state from the regular days. To this end, the aim of this study is twofold; to demonstrate the potential of using the insulin-to-carbohydrate ratio as an input feature to the algorithm, and further, compare the performance with the results presented in [49] using both blood glucose levels and insulin-to-carbohydrate ratio as input feature. Therefore, the purpose is to develop a personalized computational model for continuous and automatic detection of infection states in people with type 1 diabetes under free-living conditions using insulin and meals profile as event indicator variables, in which the ratio of insulin-to-carbohydrate is computed on a time-bin basis to form the input feature. This univariate input feature presents an advantage when there is a lack of access to blood glucose measurements. As a *point anomaly*, each point in the time-series of the ratio is evaluated against a reference description generated by the models to determine its degree of normality.

## 2. Materials and Methods

### 2.1. Materials

Both real and simulated infection episodes were used to evaluate the performance of the models. The dataset consists four patient-years with real influenza episodes and five regular patient-years with simulated infection episodes, detailed description of the dataset can be found from [49; 50], the participants characteristic is given in **Table 1**. The dataset consists of different BG dynamics parameters including blood glucose (BG), insulin, and carbohydrate information, however, the study has made use of only insulin and carbohydrate information. Moreover, based on the description provided in [50] and considering pathogen-specific deviations, infection states of different sizes and shapes were simulated and injected into the regular patient-years for performance evaluations. The simulated infection states were 10%, 20%, 30%, and 40% simultaneous deviation, i.e. higher insulin and lower carbohydrate by the same factor, from the total insulin and carbohydrate profile of individuals, as shown in **Figure 1**. The simulated infection states were used to assess the model's performance to a different degree of infection-induced changes from small to large changes in the individual BG dynamics. This directly corresponds to the fact that different pathogens induce a various degree of deviation on BG dynamics. The study protocol has been reviewed by the Norwegian Regional Committees for Medical Health Research Ethics Northern Norway (REK) (Reference number: 108435). Written consents have been obtained and the participants have donated the datasets. All the data from the participants are anonymized. All experiments were performed using MATLAB®2018b (Mathworks, Inc, Natwick, MA).



**Figure 1:** Characteristics of the input feature, i.e. insulin to carbohydrate ratio, with simulated infection states of varying degree and shape ( $\alpha = 0\%$ ,  $10\%$ ,  $20\%$ ,  $30\%$ , and  $40\%$ ).

**Table 1:** Participants characteristics [50].

Variables	
Gender	2 males and 1 female
Age	$34 \pm 13.2$ years
Bodyweight	Subject 1 (83 kg), Subject 2 (77kg), Subject 3 (70kg)
HbA1c	Subject 1 (6.0%), Subject 2 (7.3%), Subject 3 (6.2%)
Carbohydrate counting	Level 3 (advanced)

## 2.2. Models, Optimizations, Evaluations, and Performance Metrics

Almost the same group of models used in Woldaregay et.al. [49], i.e. one-class classifier and unsupervised methods, is used in this study but the models are tested and evaluated using only the insulin-to-carbohydrate ratio. Description of the models are given in **Table 2**, and further details can be found in [49]. The ratio was labeled as a set of target and non-target data. All the data, which are a regular period of the year was set as a target. The period containing the infection episode was set as a non-target. The one-class classifier models were trained on the target and tested using a dataset containing both the target and non-target data. 10 times 3-fold stratified cross-validation was used to evaluate the performance of the one-class classifier. Regarding the unsupervised method, no data labeling is required, and hence the entire patient-year was presented at once [49]. The one-class classifiers were evaluated and compared based on three features; data granularity, i.e. hourly and daily, data nature, i.e. raw and smoothed (2 days moving average), and different training sample sizes, and the unsupervised methods were evaluated with the same features but sample sizes. Further, the model's performance is compared with the results reported in Woldaregay et.al. [49].

The hyper-parameters of most of the one-class classifier models, i.e. complexity parameter  $\gamma$ , were optimized based on the consistency approach [44; 49] except for Parzen, and NN, which was optimized by using the leave-one-out error. Min-max was used for normalizing the dataset [19]. For MST, the complete MST is selected. In all the cases, a pre-specified threshold of outlier fraction in the training dataset was set to be  $\varepsilon^\dagger = 0.01$ , where one percent of the most dissimilar target data could be excluded from generating the data description. The number of neighbors ( $K$ ) in both local outlier factor (LOF) and connectivity-based outlier factor (COF) is determined based on repeated experiments. The model’s performance was evaluated for each individual’s dataset and reported using three performance metrics;  $m$  runs average and standard deviation of the *area under the receiver operating characteristic (AUC)*, *specificity*, and *F1-score*. Comparison of the overall models' performance among all the individual's dataset was carried based on these metrics, however, for the sake of clarity, the findings are depicted in terms of *F1-score*, given its practical implication, and the rest of the metrics values are given in *Appendix A*. The models were implemented using *MATLAB toolbox*, *ddtools*, *prtools*, and *Anomaly detection toolbox* [10; 11; 40].

**Table 2:** The models tested. BD = boundary and domain-based, DN= density-based, RE = reconstruction-based methods [49].

Models	One-class classifier	Unsupervised
Support vector data description (SVDD) [39; 40; 43]	√ <i>BD</i>	$\mathcal{X}$
One-class support vector machine (v-SVM) [32]	√ <i>BD</i>	$\mathcal{X}$
Nearest neighbor (NN) [27; 38; 39; 42]	√ <i>BD</i>	$\mathcal{X}$
Minimum spanning tree (MST) [18]	√ <i>BD</i>	$\mathcal{X}$
Gaussian [27; 38; 40; 41]	√ <i>DN</i>	$\mathcal{X}$
Minimum covariance Gaussian (MCG) [29; 38]	√ <i>DN</i>	$\mathcal{X}$
Mixture of Gaussian (MOG) [38; 40]	√ <i>DN</i>	$\mathcal{X}$
Parzen [23; 27; 38; 41]	√ <i>DN</i>	$\mathcal{X}$
Naïve Parzen [38; 40]	√ <i>DN</i>	$\mathcal{X}$
k-nearest neighbor (KNN) [27; 38; 41]	√ <i>DN</i>	$\mathcal{X}$
Local outlier factor (LOF) [6; 40]	√ <i>DN</i>	√
K-means [40]	√ <i>RE</i>	$\mathcal{X}$
Self-organizing map (SOM) [39; 40]	√ <i>RE</i>	$\mathcal{X}$
Auto-encoder network (AE) [38; 39]	√ <i>RE</i>	$\mathcal{X}$
Connectivity-based outlier factor (COF) [37]	$\mathcal{X}$	√

### 3. Results

The models were tested and evaluated using the ratio of insulin-to-carbohydrate as the sole input parameters. As a *point anomaly*, each point in the time-series of the ratio is evaluated against a reference description generated by the models to determine its degree of normality. As per the

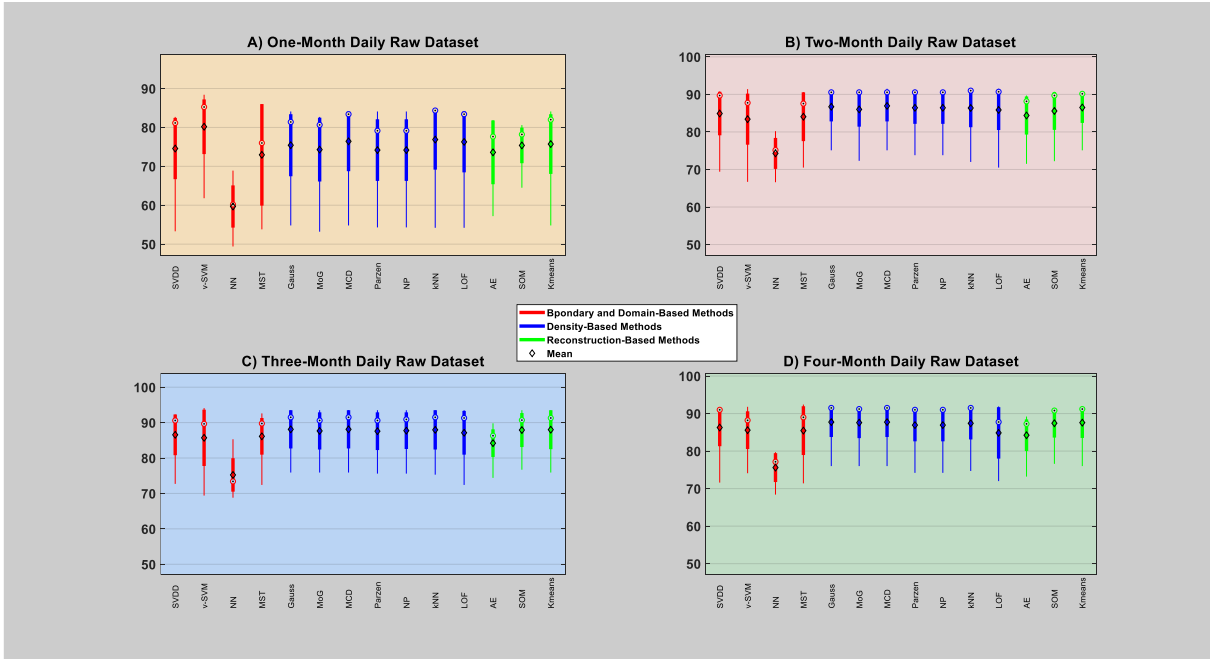


findings, generally, the models exhibit performance variations with different sample sizes, as shown in **Figure 2-4**, and the degree of these variations is mainly dependent on the *data nature* and *granularity*. As a rule of thumb, a sample size of 2 months for the daily raw data, 1 month for the daily smoothed data, and 2 months for the hourly data could be sufficient to start with when the individual participants join the system. Smoothing allows the models to generate excellent description compared to the raw dataset. In general, among the three methods, the boundary and domain-based method produced a better performance with a 1-month sample size in both data granularities. For a higher sample size, all the three methods achieved comparable description with the raw and smoothed daily data, and the density and reconstruction-based methods are better with the hourly data. In particular, on average, v-SVM produced a better performance from 1-month sample size and all the models achieved comparable performance for the higher sample size except NN. The performance (score) plot of the models can be found in *Appendix B*, which depicts the capability of each model in detecting the infection episode from the regular period. These models were trained on a random block of *120 regular days (4 months)* of the patient year and tested on the entire patient-year.

### 3.1. One class classifier

#### 3.1.1. Daily raw dataset

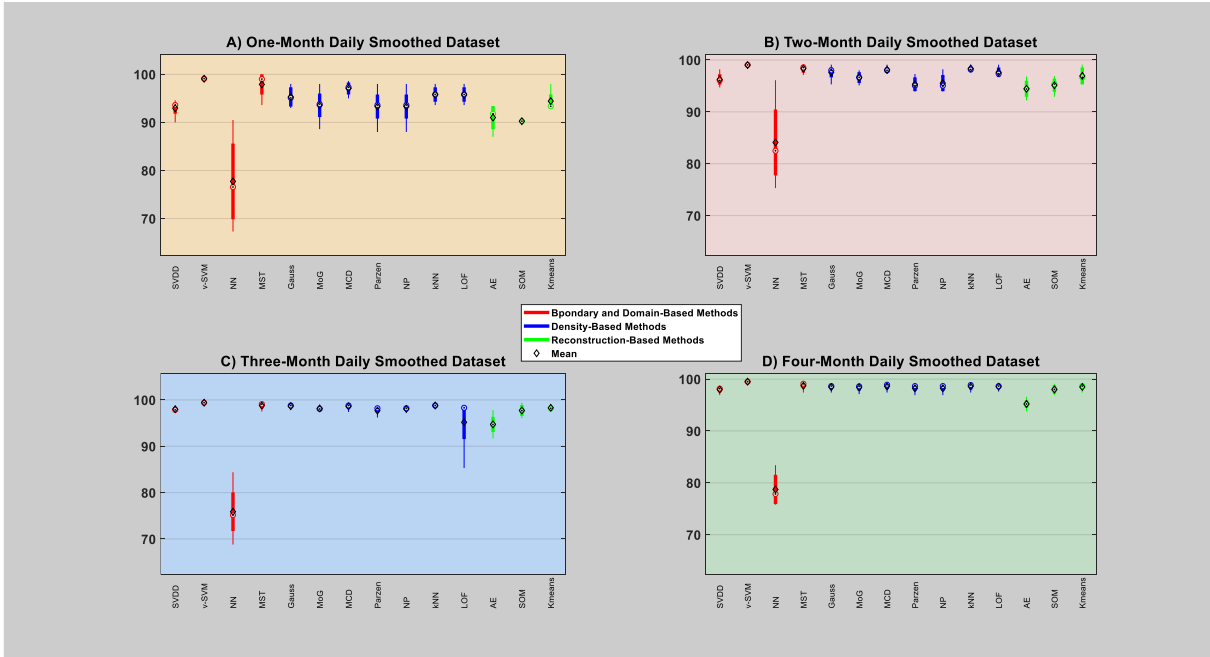
This input feature depicts the original pre-processed data without smoothing and can contain short and fast-scale features, which could affect the model's generalization. As expected, as shown in **Figure 2**, the models suffer in performance degradations, where the models' performance exhibits wider variations. Increasing the sample size has shown some improvement on the models' descriptions, specifically after the three-month sample size. As compared to the other group of methods, the boundary and domain-based method, specifically v-SVM, performed better with a 1-month sample size. With a two-month sample size, all the three methods demonstrated significant improvement, and all the models except NN generated comparable description. For higher sample sizes (three and four-month), there is no difference among all the three groups of methods, and the worst model being NN. Overall, a model such as v-SVM performed better with the 1-month sample size, and except NN, all the other models achieved comparable performance. As the sample size increase, all the models except NN improved and produced comparable performances.



**Figure 2:** Univariate input feature - The models’ median and average F1-score over the daily raw datasets.

### 3.1.2. Daily smoothed dataset

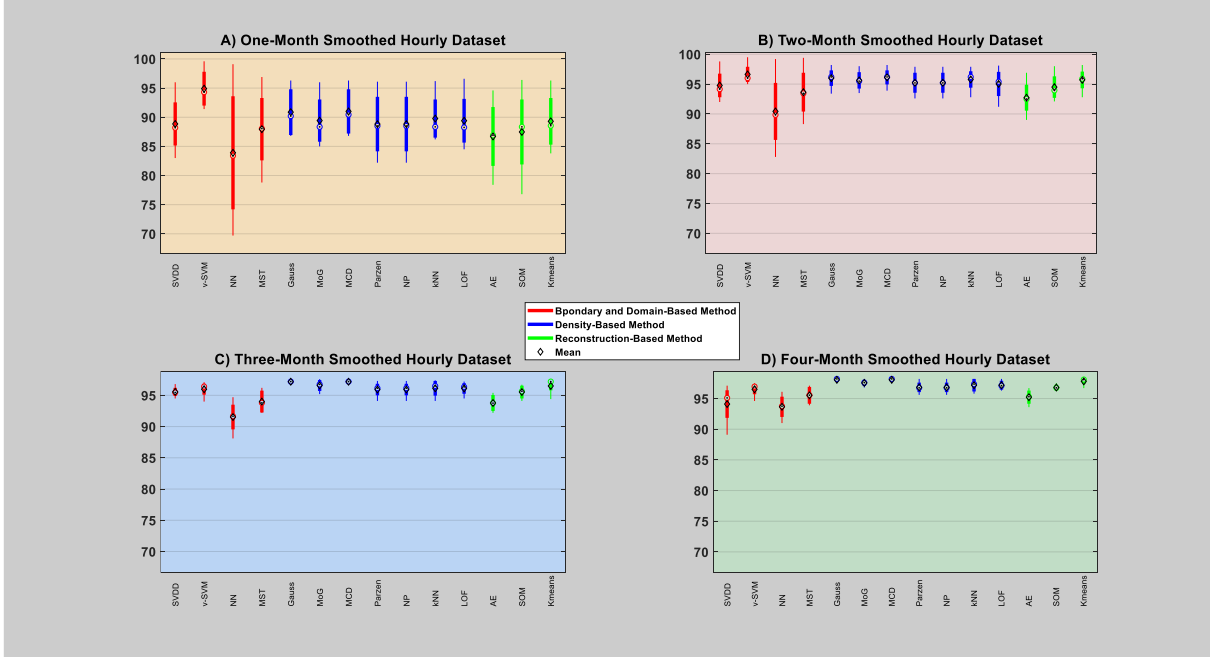
This input dataset is a filtered (smoothed) version of the raw dataset with fewer short-term and fast-scale features, and generally, the models are expected to generate improved description compared to the raw version. As compared to the raw dataset, as expected, the models have achieved significant performance improvement, as shown in **Figure 3**. In this specific dataset, increasing the sample size has little effect on performance improvement as the models have already achieved better description with lower sample sizes. As compared to the other group of methods, the boundary and domain-based method, specifically v-SVM, performed better with a 1-month sample size. With a two-month sample size, all the three methods improved, and boundary and domain-based method (i.e. v-SVM and MST), density-based method (i.e. Gaussian families, LOF and K-NN), and reconstruction based method (i.e. K-means) performed better. For higher sample sizes (three and four-month), all three methods generate comparable performances, and all the models produced similar descriptions except NN, which produces the worst description. Overall, on average, models such as v-SVM and MST achieved relatively greater performance in all the sample sizes, while all the other models except NN achieved a comparable description with two and more sample sizes. Generally, as the sample size increases, all the models except NN achieved comparable performance.



**Figure 3:** Univariate input feature - The models’ median and average F1-score over the daily smoothed datasets.

### 3.1.3. Hourly smoothed dataset

The hourly dataset is a filtered version depicting the relationship between the average blood glucose levels and the ratio. However, in the univariate sense, only the user’s estimated carb consumption and insulin requirements within each hour of the day are considered. It is obvious that increasing the data granularity could provide finer details and therefore early detection, however, at the cost of unwanted features, which might become very significant as the level gets higher. Therefore, as expected, despite presenting a large sample size and smoothing the data, as can be seen from **Figure 4**, the models exhibit high variance as compared to the daily smoothed dataset. As compared to the other group of methods, the boundary and domain-based method, specifically v-SVM, performed better with a 1-month sample size. With a two-month sample size, all the three methods improved, and boundary and domain-based method (i.e. v-SVM, and SVDD), density-based method (i.e. Gaussian families, LOF, Parzen, Naïve Parzen, and K-NN), and reconstruction based method (i.e. SOM and, K-means) performed better. For higher sample sizes (three and four-month), all three methods generate comparable performances, and the models achieved somewhat comparable descriptions except NN, which generated the worst description in all the sample sizes. Generally, increasing the sample size has helped the models to capture the data distribution better. Overall, models such as v-SVM achieved relatively greater performance with a 1-month sample size, while models such as Gaussian families and K-means achieved better descriptions with higher sample sizes.

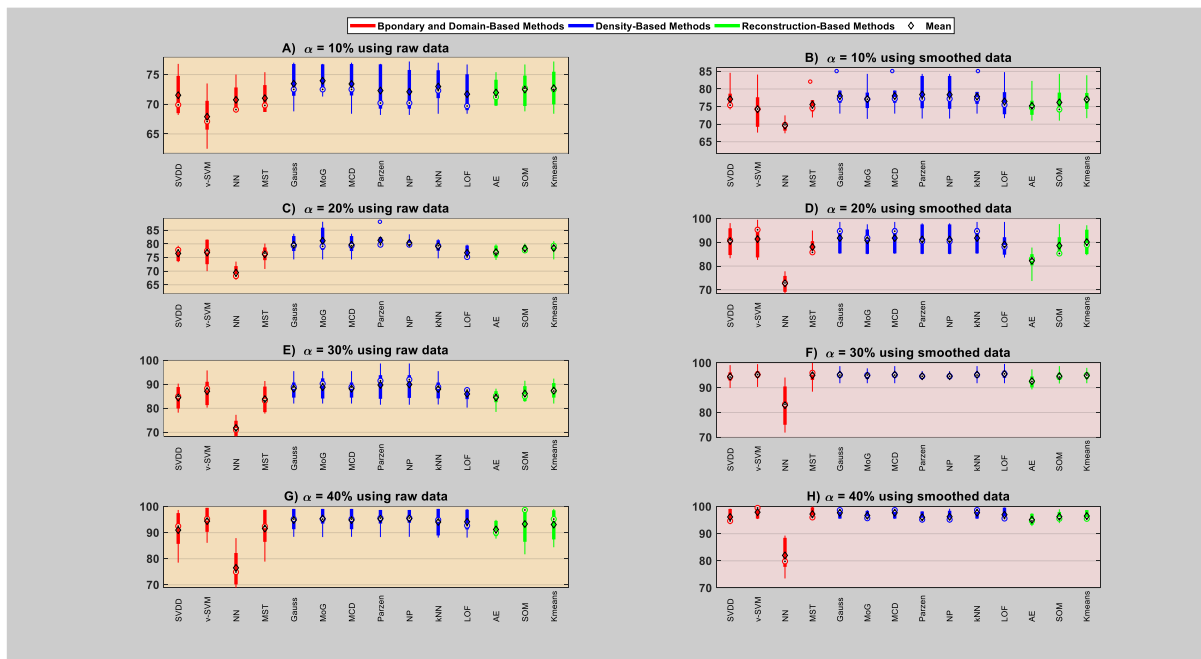


**Figure 4:** Univariate input feature - The models’ median and average F1-score over the hourly smoothed datasets.

### 3.1.4. Models performance with a different level of deviations

As described in the method section, the simulated data depicts a two-week-long simulated infection triggered deviations that were injected into each of the five regular patient-years. Simultaneous deviations of  $\alpha = 10\%$ ,  $20\%$ ,  $30\%$ , and  $40\%$  were added to the daily aggregated insulin and carbohydrate values. The models were trained with 4-month sample size, and their performance was evaluated and compared on the raw and smoothed version of each patient-years. For the sake of clarity, the median and average performance (F1-score) are presented. As shown in **Figure 5**, with increasing deviation  $\alpha$ , the model’s detection performance improves. In a comparison of the data types, the models achieved better detection performance with the smoothed dataset. Detecting infection states that induce very small deviations, i.e.  $\alpha < 10\%$  change, requires training the models with a suitable threshold that could reject outliers in the training dataset that exceeds the induced deviations (i.e.  $\alpha < 10\%$ ). However, this could in turn increases the false alarm rate and make the model less sensitive flagging regular days as an infection state. In this regard, for an application that involves detecting an infection state, it is necessary to favor the inclusion of some of the less significant outliers in the data description to avoid frequent false alarm, however, at the expense of missing infection state that induces small deviations (i.e.  $\alpha < 10\%$ ) on the blood glucose dynamics. Per the findings, with  $\alpha = 10\%$ , the density-based method performed a better detection task, and specifically, MOG achieved better description. In this regard, generally, the Gaussian family achieved better performance with the raw dataset, however, all the models achieved comparable descriptions except NN with the smoothed dataset. It is better to note that, despite the small deviation ( $\alpha = 10\%$ ), smoothing the data has helped the models to achieve good descriptions. For  $\alpha = 20\%$ , among the three methods, the density-based method (i.e. Gaussian families, Parzen and naïve Parzen) and reconstruction-based method (i.e. SOM and K-means)

achieved better description with the raw data, and regarding the smoothed data, the boundary and domain-based method (i.e. SVDD, and v-SVM), density-based method (i.e. Gaussian families, and K-NN) and reconstruction based method (i.e. SOM, and K-means) achieved relatively better and comparable performance. For  $\alpha = 30\%$ , the density-based method performed better, and specifically, Parzen and naïve Parzen achieved slightly better performance with the raw data, and almost all the models except the nearest neighbor achieved comparable performance with the smoothed dataset. For  $\alpha = 40\%$ , all the three methods achieved comparable performance, and all the models except NN achieved comparable performances.

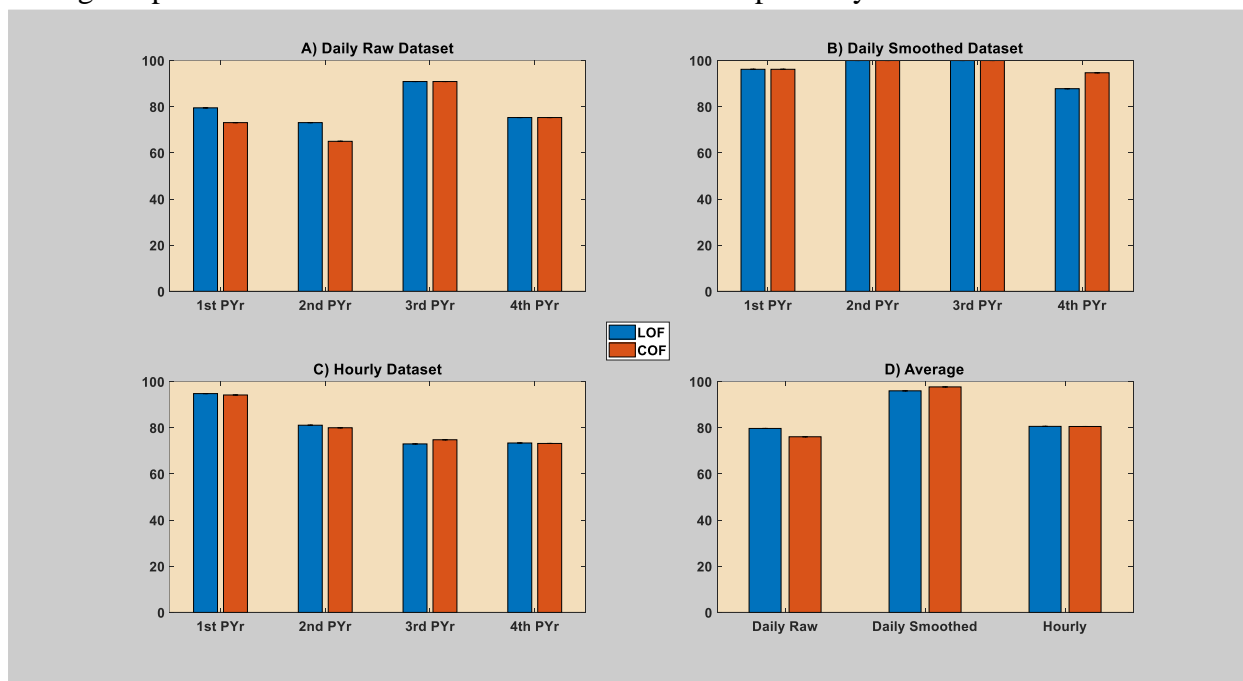


**Figure 5:** Univariate input feature - median and average performance (F1-score) of the models over the five patient years injected with different degree of deviations.

### 3.2. Unsupervised method

For comparison purposes, this section presents the performance of two density-based unsupervised models, LOF and COF, using a univariate input, i.e. the insulin-to-carbohydrate ratio, as given in **Figure 6**. The performance of these models was compared based on the *data nature* (raw and smoothed), and *data granularity* (daily and hourly). For the sake of clarity, the average performance of each model was computed using their respective performance on the individual infection states with different data granularity, and the details can be found in *Appendix A*. The optimal threshold and number of neighbors were selected after performing repeated evaluations for different combinations of values. The optimal threshold values used in the model evaluation are given in **Table 3**. As can be seen from the figure, smoothing the daily dataset significantly improved the models' performance. As expected, despite the smoothing the hourly dataset yields inferior performance to the daily scenario, due to the presence of unwanted short-term and fast-scale features as a result of higher data granularity. In general, both these

models achieved comparable performance in all the infection states. Regarding the raw dataset, LOF showed a slight edge over the COF. For the hourly and daily smoothed dataset, both the models achieved comparable performance. The performance (score) plot of the models can be found in *Appendix B*, depicting the capability of each model in detecting the infection state from the regular period. These models were tested on the entire patient-year.



**Figure 6:** Univariate input features - performance comparison (F1-score) of the unsupervised models using univariate input, i.e. insulin to carbohydrate ratio.

**Table 3:** Univariate input features - optimal values of thresholds used in performance evaluation. The values given as  $T_h$  are the optimal threshold values used for each patient-year,  $h$  depicting that particular year.

Granularity	Pre-pro.	Model (Threshold <sub>patient-year</sub> )
Daily	Without filter	LOF ( $T_1=2.7, T_2=1.5, T_3=2.95, T_4=2.2$ )
		COF ( $T_1=1.4, T_2=1.1, T_3=2.3, T_4=1.8$ )
	With filter	LOF ( $T_1=1.9, T_2=1.9, T_3=2.8, T_4=2.8$ )
		COF ( $T_1=1.6, T_2=1.5, T_3=2.8, T_4=3.1$ )
Hourly	With filter	LOF ( $T=1.9, T_2=1.6, T_3=1.2, T_4=1.7$ )
		COF ( $T_1=1.6, T_2=1.3, T_3=1.2, T_4=1.3$ )

## 4. Discussion

### 4.1. Principal finding

Recently, Woldaregay et.al. [50] have characterized infection states in people with type 1 diabetes by the presence of elevated BG episodes, alteration in carbohydrate intake, and actions of the

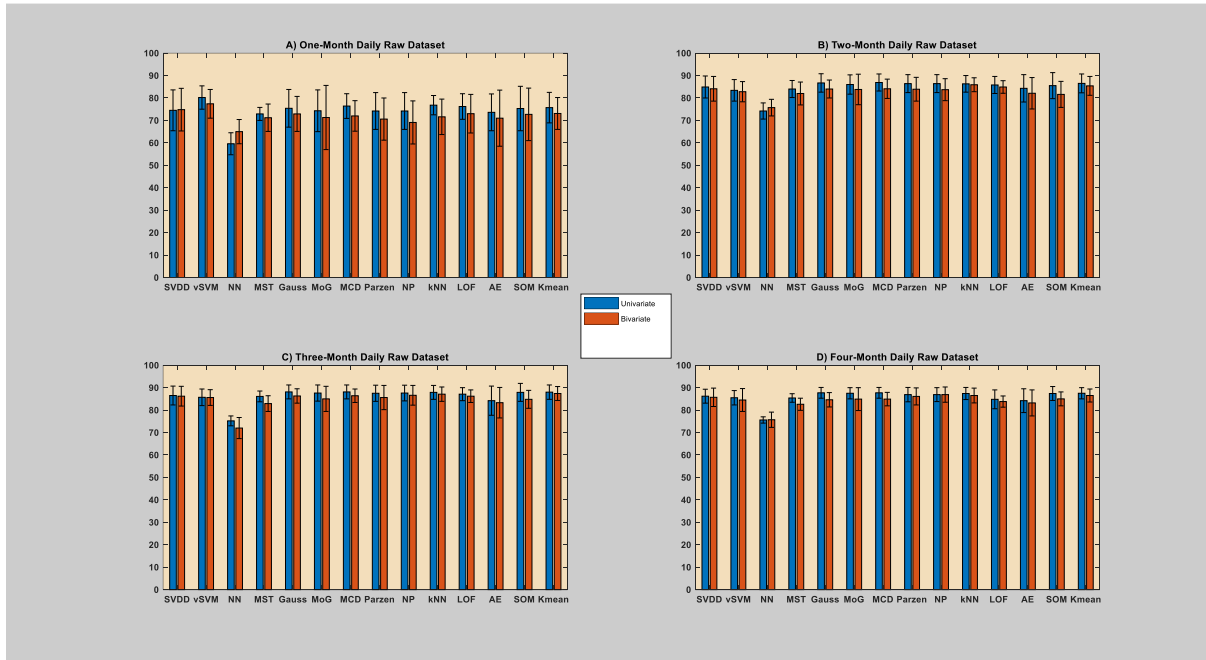
patient to lower it down, e.g. enhanced and frequent insulin injection as compared to the regular days. The study has put forward diabetes profiles such as blood glucose levels, insulin injection, carbohydrate consumption, and the ratio of insulin-to-carbohydrate as optimal parameters to devise an infection detection model. Further, Woldaregay et.al. [49] have devised a bivariate infection detection model and demonstrated the potential of blood glucose levels and insulin-to-carbohydrate ratio. Inspired by these previous studies, this study aims to demonstrate the potential of insulin and meals profile as event indicator variables by using the insulin-to-carbohydrate ratio as a sole input parameter to the model, where the ratio of insulin-to-carbohydrate is computed on a time bin basis to form an input feature. The performance evaluation of the models demonstrated the potential of the insulin and meals profile as event indicator variables for detecting the infection states from the regular days. Generally, for a small sample size, the boundary and domain-based method, especially v-SVM, performed better. However, for a larger sample size, all three categories produced a comparable description of the data. In comparison to each particular model from their respective groups, v-SVM, K-means, KNN, and Gaussian families achieved better performance on average in all the infection states.

## 4.2. Comparison of input features

### 4.2.1. One class classifier

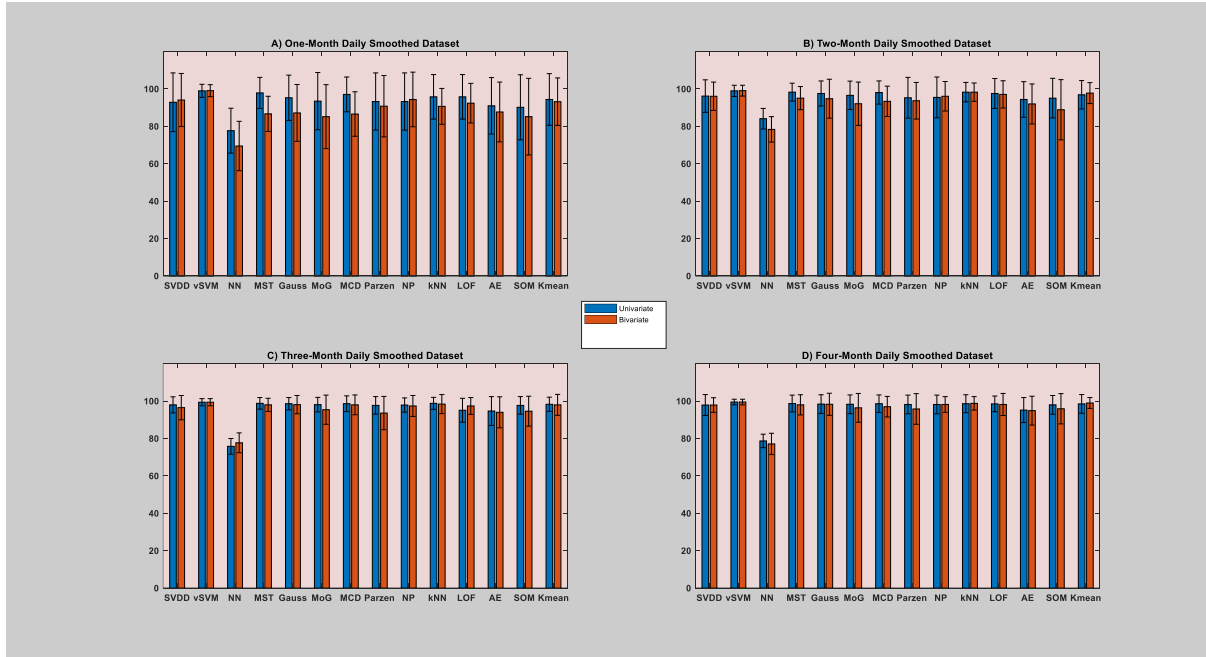
The performance of any model can be greatly affected by the input features selected for modeling [46]. In this regard, comparative analysis of the performance achieved in this study with a univariate input, i.e. insulin to carbohydrate ratio, and a bivariate input, i.e. blood glucose levels and insulin to carbohydrate ratio presented in Woldaregay et. al. [49] are performed. As shown in **Figure 7-9**, except under certain circumstances, where both achieved comparable performance, the univariate input feature-based models displayed superior performances. However, despite the improved performance one of the drawbacks of the univariate input feature-based models emanates from the fact that these models cannot differentiate between high ratio values that arise due to infection episode or just a regular day. In this regard, it worth mentioning that theoretically, a very large ratio is considered as normal value as long as the individual blood glucose levels go to the hypoglycemia state responding to the high insulin injection and low carbohydrate intake [31; 49; 50], and however, the univariate input feature-based models consider such a situation as abnormal by just looking upon the ratio values as outliers. However, it should be noted that in practical settings such incidence might be fatal for the individuals and might end up being unconscious and sometimes dead [8; 20], and therefore, such ratio values might be almost non-existent in a practical sense. Therefore, it can be concluded that the univariate input could do the same task as the bivariate input when it comes to detecting infection incidences, i.e. large ratio values, despite lacking the capability to differentiate between these ratio values. As compared to the bivariate input feature-based model, the other drawbacks of the univariate input feature-based model are related to the fact that this model might sometimes generate a false alarm in rare situations that are very sparse, i.e. too small values not included in training the models. This rare situation can be manifested in the individual's blood glucose management practice, for instance, if the patient on random days prefers to replace insulin requirement with physical exercise/activity

sessions, this instance could end up with a very small insulin-to-carbohydrate ratio and be flagged as a false alarm. In these circumstances, without the blood glucose level feature, it will remain difficult for the models to differentiate between the normality or abnormality of a small value of the insulin-to-carbohydrate ratio. Hence, having blood glucose levels as an additional input feature could minimize such unnecessary alarms.

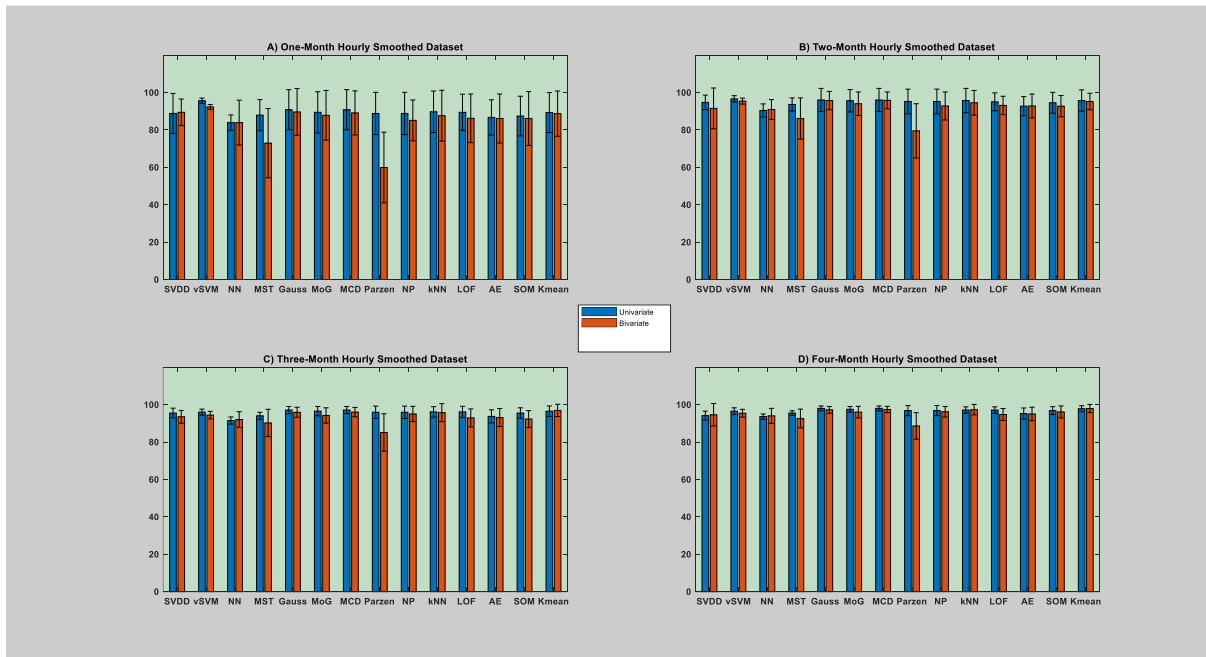


**Figure 7:** Daily raw dataset - performance comparison (F1-score) of models using bivariate input, i.e. blood glucose levels and insulin to carbohydrate ratio, and univariate input feature, i.e. insulin to carbohydrate ratio, based on the daily raw dataset. The error bars are given in terms of the overall mean and standard deviation of each model across all the patient-years and infection states.





**Figure 8:** Daily smoothed dataset - performance comparison (F1-score) of models using bivariate input, i.e. blood glucose levels and insulin to carbohydrate ratio, and univariate input feature, i.e. insulin to carbohydrate ratio, based on the daily smoothed dataset. The error bars are given in terms of the overall mean and standard deviation of each model across all the patient-years and infection states.

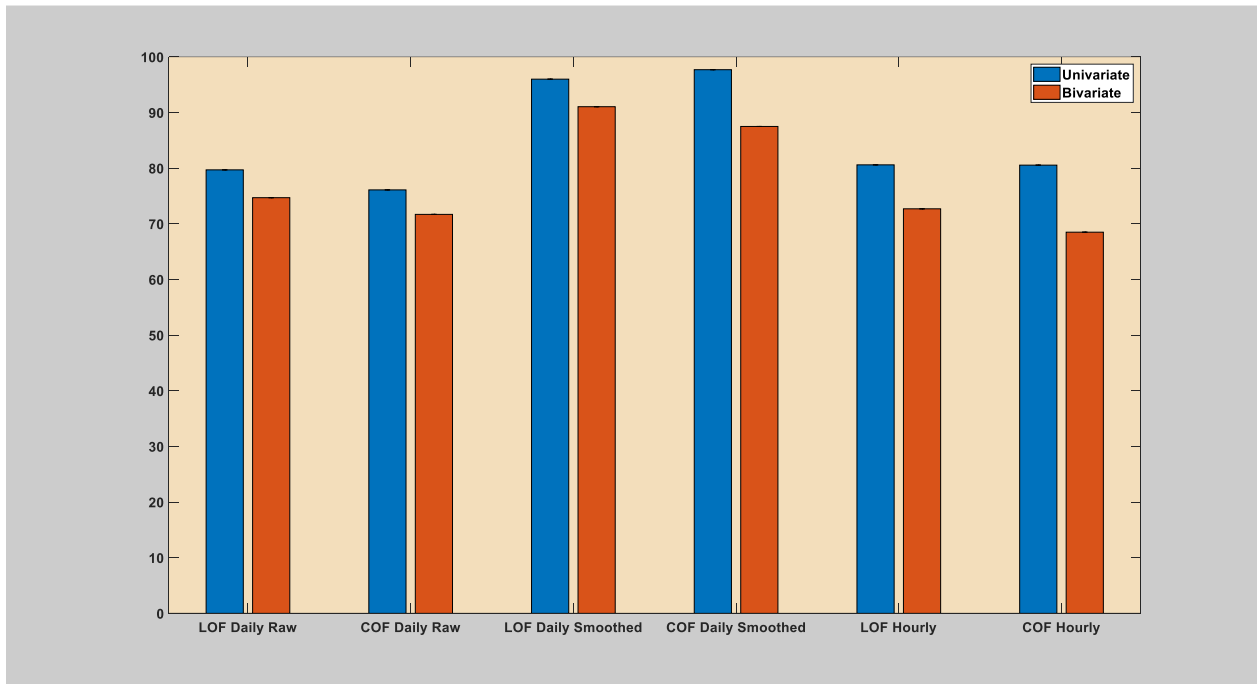


**Figure 9:** Hourly smoothed dataset - performance comparison (F1-score) of models using bivariate input, i.e. blood glucose levels and insulin to carbohydrate ratio, and univariate input feature, i.e. insulin to carbohydrate ratio, based on the hourly smoothed dataset. The error bars are given in

terms of the overall mean and standard deviation of each model across all the patient-years and infection states.

#### 4.2.2. Unsupervised Method

The comparative analysis of the model's performance in regard to the bivariate, i.e. blood glucose levels and insulin-to-carbohydrate ratio, and univariate input features, i.e. insulin-to-carbohydrate ratio, is shown in **Figure 10**. The analysis presents a comparison of the average performance that can be gained by utilizing either of the bivariate or univariate input features. Per the findings, the models with a univariate input feature have achieved improved performance compared to the bivariate input. This improvement could be linked with the discriminative power of the ratio compared to blood glucose levels.

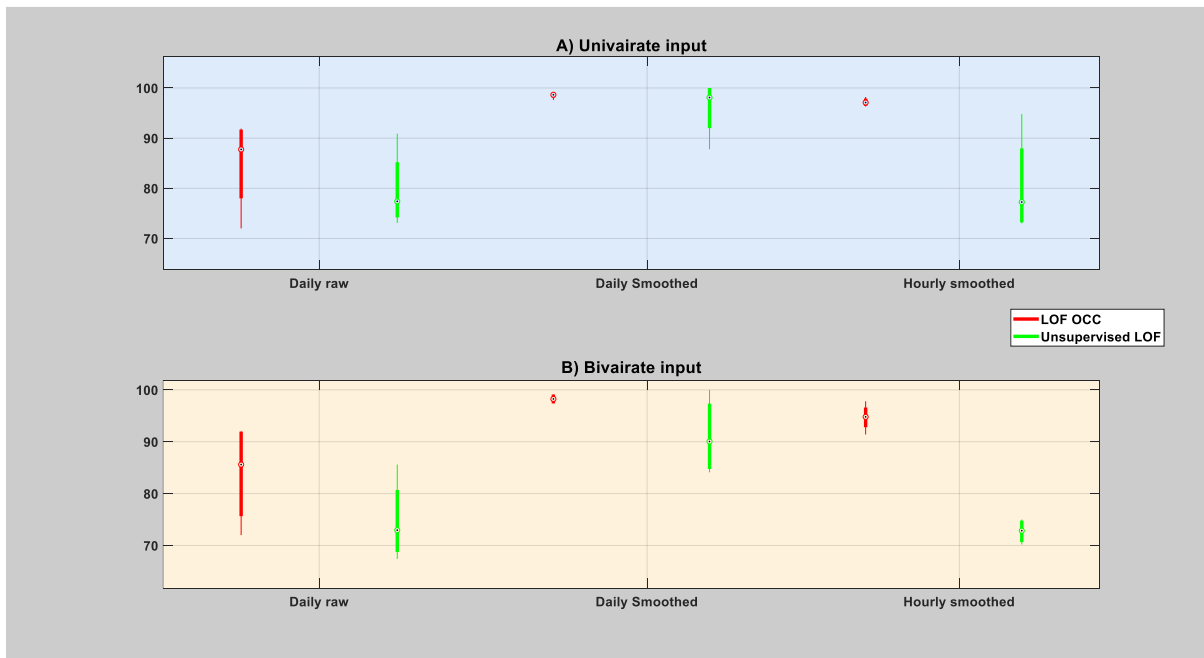


**Figure 10:** Performance comparison (F1-score) of the unsupervised models using a bivariate input, i.e. blood glucose levels and insulin to carbohydrate ratio versus a univariate input, i.e. insulin to carbohydrate ratio, source partially from [49], Table 8.

#### 4.3. Comparison of unsupervised versus one-class classifier methods

Generally, the comparative analysis of the performance achieved with the one-class classifier and unsupervised methods depicts that the unsupervised method fails to achieve comparable performance, especially with the bivariate input feature. One of the drawbacks of the unsupervised method is related to the fact that they require a fairly large sample size to at least produce comparable performance with the one-class classifiers [15; 49]. This characteristic can be easily observed by looking at the performance of the unsupervised method based on the whole patient-year and the one-class classifier trained only with four months of the patient-year. To further illustrate the difference in performance, the comparison of the best performing unsupervised

model, LOF, to its one-class classifier version, is given in **Figure 11**. As can be seen from the figure, under almost all the circumstances the one-class classifier model achieved superior performance. The characteristic of the data distribution, which contains a high and sparse density pattern, could be the reason behind the performance degradation of the unsupervised method [24] since these models expect clear demarcation between normal and abnormal values. Generally, the state of blood glucose dynamics contains very rare events that contribute to the existence of sparse regions within the distribution. In this regard, a typical example could be a holiday season, where an individual happens to consume too many carbohydrates. Furthermore, the individual decision to switch to physical activity or exercise sessions to compensate for insulin requirements could result in a similar pattern. These sparse data or rare events are a normal portion of the data, which needs to be treated as such by the detection methods. However, in these and other typical scenarios, the unsupervised method could end up considering these situations as abnormal resulting in a false alarm. In this regard, one of the main drawbacks of the unsupervised method is related to the fact that they determine anomalies from the data themselves and, there is no mechanism to let the model learn and accept certain sparse regions just like the one-class classifiers. Thus, the atypical nature of the underlying data distribution affects the performance of an unsupervised method. Yet, the one-class classifier method can handle such kind of situation if properly introduced with such an example during the learning phase [49].



**Figure 11:** Performance comparison (F1-score) of unsupervised and one-class classifier version of the local outlier factor (LOF) model using both a bivariate input, i.e. blood glucose levels and insulin to carbohydrate ratio, and a univariate input, i.e. insulin to carbohydrate ratio, source partially from [49].

## Conclusion

Infection incidences are often associated with problematic BG management and people with type 1 diabetes usually face challenges to control BG in range throughout the period. Apart from its significance in outbreak detection, detecting infection episodes among people with type 1 diabetes can be useful for the individuals. In this regard, this study has demonstrated the potential of insulin and carbs profile in detecting infection episodes among this group of peoples. Generally, the proposed approaches have demonstrated superior performance in detecting deviations from the norm due to infection onset. In comparison to each particular model from their respective groups, v-SVM, K-means, KNN, and Gaussian families achieved better performance on average in all the infection states. In general, we foresee that the presented approach could further encourage researchers to examine additional features on a large scale basis on top of the presented features.

## Reference

- [1] H.K. Akturk and S. Garg, Technological advances shaping diabetes care, *Curr Opin Endocrinol Diabetes Obes* **26** (2019), 84-89.
- [2] T.S. Bailey, J. Walsh, and J.Y. Stone, Emerging Technologies for Diabetes Care, *Diabetes Technol Ther* **20** (2018), S278-S284.
- [3] H. Banaee, M.U. Ahmed, and A. Loutfi, Data mining for wearable sensors in health monitoring systems: a review of recent trends and challenges, *Sensors (Basel)* **13** (2013), 17472-17500.
- [4] R.W. Beck, R.M. Bergenstal, L.M. Laffel, and J.C. Pickup, Advances in technology for management of type 1 diabetes, *The Lancet* **394** (2019), 1265-1273.
- [5] S. Benjamin, A. Kirstin, M.F. Sarah, C. Anoushka, D. Stephan, P. Karena, R. Adam, M.H. Frederick, and E.M. Ashley, Feasibility of continuous fever monitoring using wearable devices, *Research Square* (2020).
- [6] M.M. Breunig, H.-P. Kriegel, R.T. Ng, and J. Sander, LOF: identifying density-based local outliers, *SIGMOD Rec.* **29** (2000), 93-104.
- [7] M. Clark, *What Is Diabetes?*, John Wiley & Sons, Ltd., England, 2004.
- [8] P.E. Cryer, Hypoglycemia in type 1 diabetes mellitus, *Endocrinology and metabolism clinics of North America* **39** (2010), 641-654.
- [9] L.A. DiMeglio, C. Evans-Molina, and R.A. Oram, Type 1 diabetes, *Lancet (London, England)* **391** (2018), 2449-2462.
- [10] dsmi-lab-ntust, Anomaly Detection Toolbox, in, GitHub, GitHub, 2016.
- [11] R.P. Duin, P. Juszczak, P. Paclik, E. Pekalska, D. De Ridder, D.M. Tax, and S. Verzakov, Prtools4. 1, a matlab toolbox for pattern recognition, *Delft University of technology* **2600** (2007).
- [12] T. Eckmanns, H. Fuller, and S.L. Roberts, Digital epidemiology and global health security; an interdisciplinary conversation, *Life sciences, society and policy* **15** (2019), 2.
- [13] T. Fawcett, Mining the Quantified Self: Personal Knowledge Discovery as a Challenge for Data Science, *Big Data* **3** (2015), 249-266.
- [14] H. Gimpel, M. Nißen, and R. Görlitz, Quantifying the quantified self: A study on the motivations of patients to track their own health, (2013).
- [15] M. Goldstein and S. Uchida, A Comparative Evaluation of Unsupervised Anomaly Detection Algorithms for Multivariate Data, *PLoS One* **11** (2016), e0152173.

- [16] C. Gurrin, A.F. Smeaton, and A.R. Doherty, LifeLogging: Personal Big Data, *Foundations and Trends® in Information Retrieval* **8** (2014), 1-125.
- [17] J.L. Hicks, T. Althoff, R. Susic, P. Kuhar, B. Bostjancic, A.C. King, J. Leskovec, and S.L. Delp, Best practices for analyzing large-scale health data from wearables and smartphone apps, *npj Digital Medicine* **2** (2019), 45.
- [18] P. Juszczak, D.M.J. Tax, E. Pekalska, and R.P.W. Duin, Minimum spanning tree based one-class classifier, *Neurocomputing* **72** (2009), 1859-1869.
- [19] S. Kandanaarachchi, M.A. Munoz, R.J. Hyndman, and K. Smith-Miles, On normalization and algorithm selection for unsupervised outlier detection, in, 2018.
- [20] R.J. McCrimmon and R.S. Sherwin, Hypoglycemia in type 1 diabetes, *Diabetes* **59** (2010), 2333-2339.
- [21] S. Meißner, Effects of Quantified Self Beyond Self-Optimization, in: *Lifelogging: Digital self-tracking and Lifelogging - between disruptive technology and cultural transformation*, S. Selke, ed., Springer Fachmedien Wiesbaden, Wiesbaden, 2016, pp. 235-248.
- [22] D.K. Ming, S. Sangkaew, H.Q. Chanh, P.T.H. Nhat, S. Yacoub, P. Georgiou, and A.H. Holmes, Continuous physiological monitoring using wearable technology to inform individual management of infectious diseases, public health and outbreak responses, *International Journal of Infectious Diseases* **96** (2020), 648-654.
- [23] E. Parzen, On Estimation of a Probability Density Function and Mode, *The Annals of Mathematical Statistics* **33** (1962), 1065-1076.
- [24] M.I. Petrovskiy, Outlier Detection Algorithms in Data Mining Systems, *Programming and Computer Software* **29** (2003), 228-237.
- [25] G. Quer, J.M. Radin, M. Gadaleta, K. Baca-Motes, L. Ariniello, E. Ramos, V. Kheterpal, E.J. Topol, and S.R. Steinhubl, Wearable sensor data and self-reported symptoms for COVID-19 detection, *Nature Medicine* (2020).
- [26] J.M. Radin, N.E. Wineinger, E.J. Topol, and S.R. Steinhubl, Harnessing wearable device data to improve state-level real-time surveillance of influenza-like illness in the USA: a population-based study, *The Lancet Digital Health* **2** (2020), e85-e93.
- [27] D.d. Ridder, D.M.J. Tax, and R.P.W. Duin, An experimental comparison of one-class classification methods, in: *Proceedings of the 4th Annual Conference of the Advanced School for Computing and Imaging, Delft*, 1998.
- [28] I. Rodriguez-Rodriguez, J.V. Rodriguez, and M.A. Zamora-Izquierdo, Variables to Be Monitored via Biomedical Sensors for Complete Type 1 Diabetes Mellitus Management: An Extension of the "On-Board" Concept, *J Diabetes Res* **2018** (2018), 4826984.
- [29] P.J. Rousseeuw and K.V. Driessen, A Fast Algorithm for the Minimum Covariance Determinant Estimator, *Technometrics* **41** (1999), 212-223.
- [30] S. Samerski, Individuals on alert: digital epidemiology and the individualization of surveillance, *Life sciences, society and policy* **14** (2018), 13.
- [31] M. Schiavon, C. Dalla Man, and C. Cobelli, Insulin Sensitivity Index-Based Optimization of Insulin to Carbohydrate Ratio: In Silico Study Shows Efficacious Protection Against Hypoglycemic Events Caused by Suboptimal Therapy, *Diabetes Technol Ther* **20** (2018), 98-105.
- [32] B. Schölkopf, R.C. Williamson, A.J. Smola, J. Shawe-Taylor, and J.C. Platt, Support Vector Method for Novelty Detection, in: *Advances in neural information processing systems*, 1999, pp. 582-588.

- [33] D.R. Seshadri, E.V. Davies, E.R. Harlow, J.J. Hsu, S.C. Knighton, T.A. Walker, J.E. Voos, and C.K. Drummond, Wearable Sensors for COVID-19: A Call to Action to Harness Our Digital Infrastructure for Remote Patient Monitoring and Virtual Assessments, *Frontiers in Digital Health* **2** (2020).
- [34] D.R. Seshadri, J.R. Rowbottom, C. Drummond, J.E. Voos, and J. Craker, A review of wearable technology: Moving beyond the hype: From need through sensor implementation, in: *2016 8th Cairo International Biomedical Engineering Conference (CIBEC)*, 2016, pp. 52-55.
- [35] M. Shahidul Islam, M.T. Islam, A.F. Almutairi, G.K. Beng, N. Misran, and N. Amin, Monitoring of the Human Body Signal through the Internet of Things (IoT) Based LoRa Wireless Network System, *Applied Sciences* **9** (2019), 1884.
- [36] W. Sun, Z. Cai, Y. Li, F. Liu, S. Fang, and G. Wang, Security and Privacy in the Medical Internet of Things: A Review, *Security and Communication Networks* **2018** (2018), 1-9.
- [37] J. Tang, Z. Chen, A.W.-c. Fu, and D.W. Cheung, Enhancing Effectiveness of Outlier Detections for Low Density Patterns, in: *Advances in Knowledge Discovery and Data Mining*, M.-S. Chen, P.S. Yu, and B. Liu, eds., Springer Berlin Heidelberg, Berlin, Heidelberg, 2002, pp. 535-548.
- [38] D.M. Tax and R.P. Duin, Characterizing one-class datasets, in: *Proceedings of the sixteenth annual symposium of the pattern recognition association of South Africa*, 2006, pp. 21-26.
- [39] D.M.J. Tax, *One-class classification: Concept learning in the absence of counter-examples*, PhD, 2002.
- [40] D.M.J. Tax, DDTTools, the data description toolbox for MATLAB, version 2.1. 2, *Delft University of Technology, Delft, Netherlands* (2015).
- [41] D.M.J. Tax and R.P.W. Duin, Support vector domain description, *Pattern Recognition Letters* **20** (1999), 1191-1199.
- [42] D.M.J. Tax and R.P.W. Duin, Data description in subspaces, in: *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*, 2000, pp. 672-675 vol.672.
- [43] D.M.J. Tax and R.P.W. Duin, Support Vector Data Description, *Machine Learning* **54** (2004), 45-66.
- [44] D.M.J. Tax and K. Muller, A consistency-based model selection for one-class classification, in: *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, 2004, pp. 363-366 Vol.363.
- [45] İ. Tayfur and M.A. Afacan, Reliability of smartphone measurements of vital parameters: A prospective study using a reference method, *Am J Emerg Med* **37** (2019), 1527-1530.
- [46] S. Thudumu, P. Branch, J. Jin, and J. Singh, A comprehensive survey of anomaly detection techniques for high dimensional big data, *Journal of Big Data* **7** (2020), 42.
- [47] E. Vayena, M. Salathe, L.C. Madoff, and J.S. Brownstein, Ethical challenges of big data in public health, *PLoS Comput Biol* **11** (2015), e1003904.
- [48] H. Wang, Y. Zhao, L. Yu, J. Liu, I.M. Zwetsloot, J. Cabrera, and K.-L. Tsui, A Personalized Health Monitoring System for Community-Dwelling Elderly People in Hong Kong: Design, Implementation, and Evaluation Study, *J Med Internet Res* **22** (2020), e19223.
- [49] A.Z. Woldaregay, I.K. Launonen, D. Albers, J. Igual, E. Årsand, and G. Hartvigsen, A Novel Approach for Continuous Health Status Monitoring and Automatic Detection of Infection Incidences in People With Type 1 Diabetes Using Machine Learning Algorithms (Part 2): A Personalized Digital Infectious Disease Detection Mechanism, *J Med Internet Res* **22** (2020), e18912.

- [50] A.Z. Woldaregay, I.K. Launonen, E. Årsand, D. Albers, A. Holubová, and G. Hartvigsen, Toward Detecting Infection Incidence in People With Type 1 Diabetes Using Self-Recorded Data (Part 1): A Novel Framework for a Personalized Digital Infectious Disease Detection System, *J Med Internet Res* **22** (2020), e18911.

## Appendix A – Models performance using the univariate input feature: Ratio of Insulin-to- Carbohydrate

This appendix presents the evaluation of the models in each patient-year. The performance is depicted in terms of the area under the ROC curve (AUC), specificity, and F1-score. The table depicts the performance of each model in regard to different evaluation criteria; 1) model performance across each infection episode and thereby depicting performance variations among individuals, 2) sample sizes depicts model performance with limited data sample sizes, 3) data granularity, i.e. daily and hourly, depicts the model's performance in response to variations in detail within the data, 4) data nature, i.e. raw and smoothed data (moving average window size = 2 days or 48 hrs), depicts the model performance improvement gained by removing short term and fast scale features from the data.

### 1. One-class classifier Method

#### 1.1. Daily

##### 1.1.1. The 1<sup>st</sup> Infection Episode (Flu)

**Table 1:** Univariate Input Raw Data (different sample sizes) - Average and standard deviation (F1-score, AUC, specificity).

Fraction = 0.01												
Models	Boundary and Domain-Based Method											
	1 Month			2 Months			3 Months			4 Months		
	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1
SVDD	86.4 (6.2)	83.3 (0.0)	82.7 (6.3)	88.4 (4.1)	83.3 (0.0)	90.8 (3.2)	92.6 (2.2)	88.9 (0.0)	92.3 (5.0)	88.5 (3.1)	83.3 (0.0)	91.0 (2.9)
$\nu$ -SVM	86.9 (4.7)	83.3 (0.0)	88.4 (3.6)	90.9 (3.4)	83.3 (0.0)	91.4 (2.8)	96.1 (1.6)	88.9 (0.0)	94.1 (2.2)	93.6 (2.7)	83.3 (0.0)	91.8 (1.4)
Nearest Neighbor	84.6 (6.8)	31.7 (13.9)	59.1 (4.1)	93.8 (5.7)	43.3 (8.2)	76.6 (2.9)	89.9 (1.9)	66.7 (0.0)	85.3 (1.8)	92.9 (5.5)	48.3 (5.0)	79.2 (1.6)
MST	90.3 (3.2)	50.0 (0.0)	66.0 (2.0)	90.9 (1.8)	65.0 (5.0)	84.6 (2.3)	94.2 (0.5)	78.9 (3.3)	90.0 (0.5)	91.5 (0.4)	68.3 (5.0)	86.5 (2.0)
Density-Based Method												
Gaussian OC	89.0 (6.4)	83.3 (0.0)	82.7 (6.3)	90.8 (3.5)	83.3 (0.0)	90.6 (3.7)	95.6 (2.1)	88.9 (0.0)	93.5 (3.7)	93.0 (3.3)	83.3 (0.0)	91.9 (1.3)
MoG	89.0 (6.4)	83.3 (0.0)	82.7 (6.3)	90.8 (3.5)	83.3 (0.0)	90.6 (3.7)	95.6 (2.1)	88.9 (0.0)	93.5 (3.7)	93.0 (3.3)	83.3 (0.0)	91.9 (1.3)
MCD Gaussian	91.1 (5.1)	83.3 (0.0)	84.1 (5.9)	94.6 (2.7)	83.3 (0.0)	90.6 (3.7)	96.4 (1.5)	88.9 (0.0)	93.5 (3.7)	94.5 (2.6)	83.3 (0.0)	91.9 (1.3)
Parzen	91.8 (5.4)	73.3 (8.2)	78.2 (4.9)	94.1 (3.3)	83.3 (0.0)	90.8 (3.2)	97.2 (1.5)	88.9 (0.0)	93.5 (3.7)	95.7 (2.7)	83.3 (0.0)	91.0 (2.9)
Naive Parzen	91.8 (5.4)	73.3 (8.2)	78.2 (4.9)	94.1 (3.3)	83.3 (0.0)	90.8 (3.2)	97.2 (1.5)	88.9 (0.0)	93.5 (3.7)	95.7 (2.7)	83.3 (0.0)	91.0 (2.9)
k-NN	89.0 (5.3)	83.3 (0.0)	84.6 (3.2)	93.7 (3.0)	83.3 (0.0)	91.5 (2.8)	97.1 (1.4)	88.9 (0.0)	93.5 (3.7)	95.7 (2.7)	83.3 (0.0)	91.9 (1.3)
LOF	85.6 (3.7)	83.3 (0.0)	84.1 (5.9)	94.6 (2.7)	81.7 (5.0)	90.9 (3.2)	96.8 (1.5)	87.8 (3.3)	93.1 (3.9)	94.4 (3.7)	63.3 (14.6)	84.0 (6.7)
Reconstruction-Based Method												
Auto – encoder	88.1 (10)	68.4 (17.7)	73.6 (11)	92.5 (7.0)	76.2 (14)	87.1 (6.7)	94.8 (4.8)	82.7 (9.7)	89.8 (6.0)	92.4 (7.2)	75.5 (11.1)	87.7 (5.4)
SOM	94.7 (5.9)	77.6 (11.8)	77.1 (10)	97.3 (2.9)	83.0 (2.9)	88.9 (6.0)	95.7 (2.2)	88.9 (0.0)	93.5 (2.6)	94.8 (2.1)	83.3 (0.0)	90.6 (3.0)
K-means	89.0 (6.4)	83.3 (0.0)	82.7 (6.3)	90.8 (3.5)	83.3 (0.0)	90.6 (3.7)	95.6 (2.1)	88.9 (0.0)	93.5 (3.7)	93.0 (3.3)	83.3 (0.0)	91.9 (1.3)



**Table 2:** Univariate Input Smoothed Data (different sample sizes) - Average and standard deviation (F1-score, AUC, specificity).

Fraction = 0.01												
Models	Boundary and Domain-Based Method											
	1 Month			2 Months			3 Months			4 Months		
	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1
SVDD	100 (0.0)	100 (0.0)	93.6 (15.2)	100 (0.0)	100 (0.0)	95.8 (10.1)	100 (0.0)	100 (0.0)	97.5 (5.0)	100 (0.0)	100 (0.0)	98.3 (3.5)
$\nu$ -SVM	100 (0.0)	100 (0.0)	98.9 (3.4)	100 (0.0)	100 (0.0)	99.2 (2.5)	100 (0.0)	100 (0.0)	99.4 (2.0)	100 (0.0)	100 (0.0)	99.6 (1.2)
Nearest Neighbor	98.3 (3.6)	63.3 (10.0)	72.4 (5.7)	98.7 (2.2)	62 (19.8)	84.8 (5.8)	99.0 (1.7)	63.3 (13.2)	84.4 (5.2)	98.9 (1.5)	60 (15.3)	83.4 (4.2)
MST	100 (0.0)	98.3 (5.0)	93.6 (15.2)	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	97.5 (5.0)	100 (0.0)	100 (0.0)	99.2 (2.5)
Density-Based Method												
Gaussian OC	100 (0.0)	100 (0.0)	93.6 (15.2)	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	98.8 (3.8)	100 (0.0)	100 (0.0)	99.2 (2.5)
MoG	100 (0.0)	100 (0.0)	93.6 (15.2)	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	98.8 (3.8)	100 (0.0)	100 (0.0)	99.2 (2.5)
MCD Gaussian	100 (0.0)	100 (0.0)	96.6 (7.0)	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	98.8 (3.8)	100 (0.0)	100 (0.0)	99.2 (2.5)
Parzen	99.6 (1.3)	100 (0.0)	93.6 (15.2)	100 (0.0)	100 (0.0)	94.1 (15.0)	100 (0.0)	100 (0.0)	98.2 (4.0)	100 (0.0)	100 (0.0)	98.7 (2.7)
Naive Parzen	99.6 (1.3)	100 (0.0)	93.6 (15.2)	100 (0.0)	100 (0.0)	94.1 (15.0)	100 (0.0)	100 (0.0)	98.2 (4.0)	100 (0.0)	100 (0.0)	98.7 (2.7)
k-NN	100 (0.0)	100 (0.0)	93.6 (15.2)	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	98.8 (3.8)	100 (0.0)	100 (0.0)	99.2 (2.5)
LOF	100 (0.0)	100 (0.0)	93.6 (15.2)	100 (0.0)	100 (0.0)	96.7 (10.0)	100 (0.0)	100 (0.0)	98.8 (3.8)	100 (0.0)	100 (0.0)	99.2 (2.5)
Reconstruction-Based Method												
Auto – encoder	96.8 (7.8)	89.8 (16.3)	87.0 (16.7)	98.5 (4.0)	92.2 (14)	92.2 (12.4)	99.0 (3.3)	93.9 (12.1)	94.9 (6.4)	98.3 (4.1)	92.9 (12.0)	95.1 (6.0)
SOM	99.6 (1.3)	100 (0.0)	90.4 (17.8)	100 (0.0)	100 (0.0)	92.9 (15.2)	100 (0.0)	100 (0.0)	96.1 (6.4)	100 (0.0)	100 (0.0)	97.4 (3.6)
K-means	100 (0.0)	100 (0.0)	93.6 (15.2)	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	98.8 (3.8)	100 (0.0)	100 (0.0)	99.2 (2.5)

### 1.1.2. The 2<sup>nd</sup> Infection Episode (Flu)

**Table 3:** Univariate Input Raw Data (different sample sizes) - Average and standard deviation (F1-score, AUC, specificity).

Fraction = 0.01												
Models	Boundary and Domain-Based Method											
	1 Month			2 Months			3 Months			4 Months		
	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1
SVDD	77.2 (11.5)	20 (10.0)	53.3 (5.2)	82.2 (10.2)	23.3 (20.1)	69.4 (5.5)	86.1 (6.9)	28.9 (20.1)	72.7 (3.7)	80.4 (7.5)	23.3 (20.1)	71.6 (4.1)
$\nu$ -SVM	86.1 (9.3)	18.8 (8.2)	61.8 (6.8)	94.4 (6.0)	21.7 (13.7)	66.7 (7.2)	96.8 (4.4)	27.3 (13.5)	69.4 (5.9)	97.3 (3.4)	38.9 (12.6)	74.1 (5.2)
Nearest Neighbor	56.1 (11.8)	0.0 (0.0)	49.4 (4.8)	52.3 (14.1)	1.7 (5.0)	66.6 (3.0)	59.1 (10.7)	18.9 (15.8)	68.8 (3.5)	53.8 (8.5)	11.7 (7.7)	68.4 (1.9)
MST	56.7 (4.2)	16.7 (0.0)	53.8 (5.1)	59.0 (4.9)	21.7 (15.0)	70.5 (1.6)	62.8 (6.6)	26.7 (13.4)	72.4 (0.8)	60.3 (7.3)	21.7 (15.0)	71.4 (1.9)
Density-Based Method												
Gaussian OC	86.7 (9.4)	20 (10.0)	54.8 (2.6)	94.2 (5.9)	38.3 (15.0)	75.1 (1.1)	96.7 (4.4)	38.9 (16.7)	75.9 (2.2)	97.1 (3.3)	38.3 (15.0)	76.0 (2.5)
MoG	85.8 (11.9)	20.2 (9.7)	53.2 (5.6)	93.1 (7.6)	31.6 (19.1)	72.3 (3.2)	96.6 (4.5)	38.9 (16.7)	75.9 (2.2)	97.3 (3.3)	38.3 (15.0)	76.0 (2.5)
MCD Gaussian	91.7 (8.7)	20 (10.0)	54.8 (2.6)	94.8 (6.3)	38.3 (15.0)	75.1 (1.1)	97.1 (4.5)	38.9 (16.7)	75.9 (2.2)	97.7 (3.5)	38.3 (15.0)	76.0 (2.5)
Parzen	90.0 (10.5)	23.3 (13.4)	54.3 (3.0)	94.2 (6.9)	33.3 (18.3)	73.8 (1.7)	96.8 (4.6)	37.8 (17.4)	75.6 (2.5)	97.6 (3.4)	31.7 (19.0)	74.2 (3.8)
Naive Parzen	90.0 (10.5)	23.3 (13.4)	54.3 (3.0)	94.2 (6.9)	33.3 (18.3)	73.8 (1.7)	96.8 (4.6)	36.7 (18.0)	75.3 (2.5)	97.6 (3.4)	31.7 (19.0)	74.2 (3.8)
k-NN	91.7 (8.7)	18.3 (5.0)	54.2 (3.9)	95.4 (6.5)	26.7 (20.1)	72.0 (1.8)	97.1 (4.5)	36.7 (18.0)	75.3 (2.5)	97.1 (3.3)	33.3 (18.3)	74.7 (3.4)
LOF	87.6 (7.4)	18.3 (5.0)	54.2 (3.9)	93.4 (7.0)	21.7 (15.0)	70.5 (1.6)	95.3 (4.6)	26.7 (13.4)	72.4 (0.8)	95.7 (3.6)	23.3 (20.1)	72.0 (3.8)
Reconstruction-Based Method												
Auto – encoder	72.9 (16.1)	31.3 (19.6)	57.2 (7.5)	71.6 (14.7)	28.4 (18.7)	71.5 (5.4)	74.2 (12.7)	34.2 (18.0)	74.4 (4.7)	74.6 (14)	29.6 (19.7)	73.2 (4.8)
SOM	77.1 (9.6)	62.9 (10.6)	64.5 (9.5)	78.4 (6.2)	42.8 (12.7)	72.2 (6.8)	82.6 (6.1)	49.7 (11.6)	76.7 (5.3)	84.6 (4.9)	43.1 (14.1)	76.6 (3.2)
K-means	86.7 (9.4)	20.0 (10.0)	54.8 (2.6)	94.2 (5.9)	38.3 (15.0)	75.1 (1.1)	96.7 (4.4)	38.9 (16.7)	75.9 (2.2)	97.1 (3.3)	38.3 (15.0)	76.0 (2.5)

**Table 4:** Univariate Input Smoothed Data (different sample sizes) - Average and standard deviation (F1-score, AUC, specificity).

Fraction = 0.01												
Models	Boundary and Domain-Based Method											
	1 Month			2 Months			3 Months			4 Months		
	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1
SVDD	100 (0.0)	100 (0.0)	90.0 (20.1)	100 (0.0)	100 (0.0)	96.3 (7.4)	100 (0.0)	100 (0.0)	98.2 (4.0)	100 (0.0)	100 (0.0)	96.9 (7.9)
$\nu$ -SVM	100 (0.0)	100 (0.0)	99.5 (2.9)	100 (0.0)	100 (0.0)	99.1 (2.5)	100 (0.0)	100 (0.0)	99.5 (1.6)	100 (0.0)	100 (0.0)	99.6 (1.3)
Nearest Neighbor	100 (0.0)	91.7 (25.1)	90.5 (19.2)	99.3 (2.1)	90 (30.1)	96.1 (9.1)	91.7 (10.7)	12.2 (12.7)	68.8 (1.7)	95.4 (4.3)	35 (19.0)	76 (5.3)
MST	100 (0.0)	100 (0.0)	100 (0.0)	100 (0.0)	100 (0.0)	99.1 (2.7)	100 (0.0)	100 (0.0)	99.4 (1.8)	100 (0.0)	100 (0.0)	97.4 (7.9)
Density-Based Method												
Gaussian OC	100 (0.0)	100 (0.0)	93 (15.6)	100 (0.0)	100 (0.0)	95.3 (9.6)	100 (0.0)	100 (0.0)	98.8 (2.4)	100 (0.0)	100 (0.0)	97.4 (7.9)
MoG	100 (0.0)	100 (0.0)	88.6 (19.8)	100 (0.0)	100 (0.0)	95.1 (10.2)	100 (0.0)	100 (0.0)	97.8 (4.0)	100 (0.0)	100 (0.0)	97.1 (7.9)
MCD Gaussian	100 (0.0)	100 (0.0)	95 (15.0)	100 (0.0)	100 (0.0)	97.3 (8.2)	100 (0.0)	100 (0.0)	97.4 (6.1)	100 (0.0)	100 (0.0)	97.4 (7.9)
Parzen	100 (0.0)	100 (0.0)	88 (19.9)	100 (0.0)	100 (0.0)	93.9 (12.2)	100 (0.0)	100 (0.0)	96.2 (6.7)	100 (0.0)	100 (0.0)	96.9 (7.9)
Naive Parzen	100 (0.0)	100 (0.0)	88 (19.9)	100 (0.0)	100 (0.0)	93.9 (12.2)	100 (0.0)	100 (0.0)	97.6 (4.1)	100 (0.0)	100 (0.0)	96.9 (7.9)
k-NN	100 (0.0)	100 (0.0)	95 (15.0)	100 (0.0)	100 (0.0)	98.3 (5.0)	100 (0.0)	100 (0.0)	98.8 (2.4)	100 (0.0)	100 (0.0)	97.4 (7.9)
LOF	100 (0.0)	100 (0.0)	95(15.0)	100 (0.0)	100 (0.0)	96.7 (10.0)	98.4 (4.8)	67.8 (23.1)	85.3 (10.9)	100 (0.0)	100 (0.0)	97.6 (6.0)
Reconstruction-Based Method												
Auto – encoder	99.4 (3.6)	96.3 (11.9)	90.2 (15.8)	99.3 (3.5)	97.3 (9.7)	95.2 (8.6)	95.5 (9.0)	86.4 (20.5)	91.7 (10.7)	96.8 (7.3)	90 (16.7)	93.8 (8.3)
SOM	100 (0.0)	100 (0.0)	90.1 (17.8)	100 (0.0)	100 (0.0)	95.9 (7.4)	100 (0.0)	100 (0.0)	97.0 (4.9)	100 (0.0)	100 (0.0)	96.9 (8.0)
K-means	100 (0.0)	100 (0.0)	93.0 (15.6)	100 (0.0)	100 (0.0)	95.3 (9.6)	100 (0.0)	100 (0.0)	98.2 (4.0)	100 (0.0)	100 (0.0)	97.4 (7.9)

**1.1.3. The 3<sup>rd</sup> Infection Episode (Flu)**

**Table 5:** Univariate Input Raw Data (different sample sizes) - Average and standard deviation (F1-score, AUC, specificity).

Fraction = 0.01												
Models	Boundary and Domain-Based Method											
	1 Month			2 Months			3 Months			4 Months		
	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1
SVDD	68.9 (4.5)	66.7 (0.0)	80.1 (13.8)	75.4 (6.6)	66.7 (0.0)	88.8 (6.2)	74.3 (3.0)	66.7 (0.0)	88.9 (2.4)	73.5 (4.9)	66.7 (0.0)	91.0 (2.9)
$\nu$ -SVM	69.4 (5.9)	66.7 (0.0)	84.5 (4.1)	69.1 (4.3)	66.7 (0.0)	86.5 (4.1)	67.8 (2.3)	66.7 (0.0)	86.1 (2.5)	71.0 (4.2)	66.7 (0.0)	87.0 (2.7)
Nearest Neighbor	93.6 (10.1)	6.7 (20.1)	68.9 (6.7)	74.7 (10.6)	0.0 (0.0)	80.2 (0.7)	77.1 (13.1)	0.0 (0.0)	74.6 (1.8)	74.6 (9.4)	0.0 (0.0)	79.7 (1.3)
MST	81.8 (3.4)	66.7 (0.0)	86.0 (1.0)	82.8 (1.7)	66.7 (0.0)	90.5 (5.9)	83.0 (0.7)	66.7 (0.0)	89.5 (1.8)	82.9 (0.9)	66.7 (0.0)	91.5 (2.7)
Density-Based Method												
Gaussian OC	68.9 (4.5)	66.7 (0.0)	80.1 (13.8)	68.8 (3.6)	66.7 (0.0)	90.5 (5.9)	67.4 (1.4)	66.7 (0.0)	89.5 (1.8)	68.9 (2.7)	66.7 (0.0)	91.5 (2.7)
MoG	70.1 (6.3)	66.7 (0.0)	79.0 (14.0)	72.8 (6.6)	66.7 (0.0)	90.5 (5.9)	73.5 (3.9)	66.7 (0.0)	88.9 (2.4)	78.8 (7.0)	66.7 (0.0)	91.5 (2.7)
MCD Gaussian	68.9 (4.5)	66.7 (0.0)	82.7 (6.3)	68.3 (2.4)	66.7 (0.0)	90.5 (5.9)	69.2 (3.7)	66.7 (0.0)	89.5 (1.8)	71.6 (4.1)	66.7 (0.0)	91.5 (2.7)
Parzen	67.8 (3.3)	66.7 (0.0)	80.1 (13.8)	71.5 (5.1)	66.7 (0.0)	90.5 (5.9)	70.3 (4.3)	66.7 (0.0)	88.9 (2.4)	74.7 (4.8)	66.7 (0.0)	91.5 (2.7)
Naive Parzen	67.8 (3.3)	66.7 (0.0)	80.1 (13.8)	71.5 (5.1)	66.7 (0.0)	90.5 (5.9)	70.3 (4.3)	66.7 (0.0)	89.5 (1.8)	74.7 (4.8)	66.7 (0.0)	91.5 (2.7)
k-NN	79.7 (9.8)	66.7 (0.0)	84.6 (3.2)	68.9 (2.7)	66.7 (0.0)	90.5 (5.9)	74.3 (4.5)	66.7 (0.0)	89.5 (1.8)	77.7 (3.1)	66.7 (0.0)	91.5 (2.7)
LOF	71.1 (7.4)	66.7 (0.0)	82.7 (6.3)	72.1 (5.4)	66.7 (0.0)	90.5 (5.9)	71.1 (4.3)	66.7 (0.0)	89.5 (1.8)	77.8 (6.3)	66.7 (0.0)	91.5 (2.7)
Reconstruction-Based Method												
Auto – encoder	80.8 (10.9)	65.8 (5.2)	81.9 (10.5)	88.5 (9.8)	65.8 (5.2)	89.7 (5.4)	84.7 (11.8)	58.3 (16.3)	86.4 (5.5)	88.5 (9.9)	58 (14.9)	89.2 (4.2)
SOM	93.9 (8.2)	67.7 (5.7)	80.6 (9.7)	91.2 (5.1)	66.7 (0.0)	90.5 (5.9)	88.3 (6.0)	66.7 (0.0)	89.5 (1.8)	89.6 (5.0)	66.7 (0.0)	91.5 (2.7)
K-means	83.2 (12.7)	66.7 (0.0)	81.3 (9.8)	85.6 (8.5)	66.7 (0.0)	89.7 (6.1)	86.8 (7.5)	66.7 (0.0)	89.1 (2.6)	88.5 (6.5)	66.7 (0.0)	91.4 (2.8)

**Table 6:** Univariate Input Smoothed Data (different sample sizes) - Average and standard deviation (F1-score, AUC, specificity).

Fraction = 0.01												
Models	Boundary and Domain-Based Method											
	1 Month			2 Months			3 Months			4 Months		
	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1
SVDD	100 (0.0)	100 (0.0)	93.6 (15.2)	100 (0.0)	100 (0.0)	94.7 (11.1)	100 (0.0)	100 (0.0)	97.5 (5.0)	100 (0.0)	100 (0.0)	98.0 (6.0)
$\nu$ -SVM	100 (0.0)	100 (0.0)	98.6 (4.2)	100 (0.0)	100 (0.0)	98.8 (3.7)	100 (0.0)	100 (0.0)	99.3 (2.0)	100 (0.0)	100 (0.0)	99.4 (1.8)
Nearest Neighbor	80.0 (24.3)	0.0 (0.0)	67.3 (1.8)	93 (10.1)	0.0 (0.0)	80.2 (0.7)	88.3 (8.5)	0.0 (0.0)	74.6 (1.8)	91.2 (5.8)	0.0 (0.0)	79.7 (0.8)
MST	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	97.1 (6.3)	100 (0.0)	100 (0.0)	98.8 (2.4)	100 (0.0)	100 (0.0)	99.1 (2.7)
Density-Based Method												
Gaussian OC	100 (0.0)	100 (0.0)	96.6 (7.0)	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	98.8 (3.8)	100 (0.0)	100 (0.0)	98.6 (4.3)
MoG	100 (0.0)	100 (0.0)	94.0 (15.1)	100 (0.0)	100 (0.0)	96.0 (8.0)	100 (0.0)	100 (0.0)	97.9 (4.5)	100 (0.0)	100 (0.0)	98.5 (4.3)
MCD Gaussian	100 (0.0)	100 (0.0)	98.6 (4.3)	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	98.8 (3.8)	100 (0.0)	100 (0.0)	99.1 (2.7)
Parzen	100 (0.0)	100 (0.0)	93.6 (15.2)	100 (0.0)	100 (0.0)	96.0 (8.0)	100 (0.0)	100 (0.0)	98.2 (4.0)	100 (0.0)	100 (0.0)	98.6 (4.3)
Naive Parzen	100 (0.0)	100 (0.0)	93.6 (15.2)	100 (0.0)	100 (0.0)	96.0 (8.0)	100 (0.0)	100 (0.0)	98.2 (4.0)	100 (0.0)	100 (0.0)	98.6 (4.3)
k-NN	100 (0.0)	100 (0.0)	96.6 (7.0)	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	98.8 (3.8)	100 (0.0)	100 (0.0)	98.6 (4.3)
LOF	100 (0.0)	100 (0.0)	96.6 (7.0)	100 (0.0)	100 (0.0)	98.0 (6.0)	99.3 (2.2)	96.7 (10.0)	97.8 (4.5)	100 (0.0)	100 (0.0)	98.6 (4.3)
Reconstruction-Based Method												
Auto – encoder	98.8 (7.2)	97.3 (11.7)	93.4 (14.6)	96.4 (9.4)	88.2 (23.1)	93.6 (8.9)	97.1 (7.2)	87.5 (24.0)	94.5 (7.5)	96.8 (8.1)	85.5 (23.5)	95.3 (6.1)
SOM	100 (0.0)	100 (0.0)	89.9 (15.6)	99.9 (0.9)	100 (0.0)	94.8 (10.8)	100 (0.0)	100 (0.0)	98.4 (4.1)	100 (0.0)	100 (0.0)	98.6 (4.3)
K-means	100 (0.0)	100 (0.0)	93.1 (14.6)	100 (0.0)	100 (0.0)	95.3 (9.4)	100 (0.0)	100 (0.0)	98.0 (4.3)	100 (0.0)	100 (0.0)	98.6 (4.3)

### 1.1.4. The 4<sup>th</sup> Infection Episode (Flu)

**Table 7:** Univariate Input Raw Data (different sample sizes) - Average and standard deviation (F1-score, AUC, specificity).

Fraction = 0.01												
Models	Boundary and Domain-Based Method											
	1 Month			2 Months			3 Months			4 Months		
	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1
SVDD	95.0 (4.1)	75.0 (0.0)	82.2 (7.9)	98.3 (2.0)	75.0 (0.0)	90.6 (3.7)	92.7 (2.5)	85.7 (0.0)	92.3 (5.0)	87.2 (3.1)	75.0 (0.0)	91.5 (1.8)
$\nu$ -SVM	98.3 (3.4)	75.3 (2.5)	86.0 (5.0)	98.7 (1.9)	75.1 (1.8)	89.0 (3.5)	98.5 (1.8)	85.7 (0.0)	93.2 (2.4)	96.0 (2.7)	75.0 (0.0)	89.4 (2.0)
Nearest Neighbor	74.0 (21.2)	7.5 (22.6)	61.3 (2.7)	77.4 (8.5)	5.0 (15.0)	73.7 (5.7)	84.7 (10.4)	0.0 (0.0)	72.2 (0.6)	86.4 (6.6)	0.0 (0.0)	75.1 (0.4)
MST	86.7 (1.7)	75.0 (0.0)	86.0 (1.0)	87.1 (0.8)	75.0 (0.0)	90.6 (3.7)	92.5 (0.6)	82.9 (8.6)	92.6 (4.4)	87.2 (0.5)	75.0 (0.0)	92.4 (0.2)
Density-Based Method												
Gaussian OC	97.5 (3.8)	75.0 (0.0)	84.1 (5.9)	98.8 (1.9)	75.0 (0.0)	90.6 (3.7)	98.1 (2.0)	85.7 (0.0)	93.5 (3.7)	95.4 (2.6)	75.0 (0.0)	91.5 (2.7)
MoG	97.2 (4.1)	75.0 (0.0)	82.3 (7.8)	97.7 (3.5)	75.0 (0.0)	90.6 (3.7)	98.6 (1.8)	85.7 (0.0)	92.3 (5.0)	96.3 (2.6)	75.0 (0.0)	90.9 (3.0)
MCD Gaussian	98.3 (3.3)	75.0 (0.0)	84.1 (5.9)	98.8 (1.9)	75.0 (0.0)	91.5 (2.8)	98.4 (1.7)	85.7 (0.0)	93.5 (3.7)	96.7 (2.1)	75.0 (0.0)	91.5 (2.7)
Parzen	97.5 (3.8)	75.0 (0.0)	84.1 (5.9)	98.3 (2.0)	75.0 (0.0)	90.6 (3.7)	98.6 (1.8)	85.7 (0.0)	92.3 (5.0)	96.5 (2.5)	75.0 (0.0)	91.0 (2.9)
Naive Parzen	97.5 (3.8)	75.0 (0.0)	84.1 (5.9)	98.3 (2.0)	75.0 (0.0)	90.6 (3.7)	98.6 (1.8)	85.7 (0.0)	92.3 (5.0)	96.5 (2.5)	75.0 (0.0)	91.0 (2.9)
k-NN	98.3 (3.3)	75.0 (0.0)	84.1 (5.9)	98.8 (1.9)	75.0 (0.0)	91.5 (2.8)	98.6 (1.8)	85.7 (0.0)	93.5 (3.7)	96.5 (2.5)	75.0 (0.0)	91.5 (2.7)
LOF	98.3 (3.3)	75.0 (0.0)	84.1 (5.9)	97.5 (3.8)	75.0 (0.0)	91.5 (2.8)	98.6 (1.8)	85.7 (0.0)	93.5 (3.7)	96.3 (2.6)	75.0 (0.0)	91.9 (1.3)
Reconstruction-Based Method												
Auto – encoder	85.0 (8.9)	72.4 (10.5)	81.7 (8.7)	87.9 (9.6)	71.3 (12.0)	89.2 (6.1)	87.9 (12.5)	67.6 (24.9)	86.2 (8.7)	87.7 (12.7)	57.0 (24.4)	86.8 (6.2)
SOM	79.5 (5.5)	75.0 (0.0)	79.3 (9.5)	84.7 (8.6)	75.4 (3.0)	90.7 (3.6)	97.8 (4.3)	87.0 (4.8)	91.9 (4.8)	95.6 (5.9)	74.0 (7.0)	91.0 (3.2)
K-means	97.5 (3.8)	75.0 (0.0)	84.1 (5.9)	98.8 (1.9)	75.0 (0.0)	90.6 (3.7)	98.1 (2.0)	85.7 (0.0)	93.5 (3.7)	95.4 (2.6)	75.0 (0.0)	91.0 (2.9)

**Table 8:** Univariate Input Smoothed Data (different sample sizes) - Average and standard deviation (F1-score, AUC, specificity).

Fraction = 0.01												
Models	Boundary and Domain-Based Method											
	1 Month			2 Months			3 Months			4 Months		
	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1
SVDD	100 (0.0)	100 (0.0)	94.6 (8.4)	100 (0.0)	100 (0.0)	98.2 (3.6)	100 (0.0)	100 (0.0)	98.8 (2.4)	100 (0.0)	100 (0.0)	98.7 (2.9)
$\nu$ -SVM	100 (0.0)	100 (0.0)	99.3 (3.1)	100 (0.0)	100 (0.0)	99.0 (2.9)	100 (0.0)	100 (0.0)	99.4 (1.9)	100 (0.0)	100 (0.0)	99.4 (1.6)
Nearest Neighbor	99.2 (2.5)	62.5 (34.1)	80.7 (12)	97.5 (5.3)	0.0 (0.0)	75.3 (0.8)	97.8 (4.5)	17.1 (26.3)	75.7 (5.9)	98.3 (2.6)	5.0 (15.0)	75.8 (2.0)
MST	100 (0.0)	100 (0.0)	100 (0.0)	100 (0.0)	100 (0.0)	99.1 (2.7)	100 (0.0)	100 (0.0)	99.4 (1.8)	100 (0.0)	100 (0.0)	99.1 (1.7)
Density-Based Method												
Gaussian OC	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	99.1 (2.7)	100 (0.0)	100 (0.0)	98.2 (2.7)	100 (0.0)	100 (0.0)	98.7 (2.9)
MoG	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	97.3 (4.2)	100 (0.0)	100 (0.0)	98.2 (2.7)	100 (0.0)	100 (0.0)	98.7 (2.9)
MCD Gaussian	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	99.1 (2.7)	100 (0.0)	100 (0.0)	99.4 (1.8)	100 (0.0)	100 (0.0)	98.7 (2.9)
Parzen	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	97.3 (4.2)	100 (0.0)	100 (0.0)	98.2 (2.7)	100 (0.0)	100 (0.0)	98.7 (2.9)
Naive Parzen	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	98.2 (3.6)	100 (0.0)	100 (0.0)	98.2 (2.7)	100 (0.0)	100 (0.0)	98.7 (2.9)
k-NN	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	99.1 (2.7)	100 (0.0)	100 (0.0)	98.8 (2.4)	100 (0.0)	100 (0.0)	99.1 (1.7)
LOF	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	99.1 (2.7)	100 (0.0)	100 (0.0)	98.8 (2.4)	100 (0.0)	100 (0.0)	98.7 (2.9)
Reconstruction-Based Method												
Auto – encoder	98.7 (5.6)	95.8 (14.6)	93.4 (11.2)	99.7 (2.3)	97.0 (12.9)	96.8 (6.2)	99.4 (3.5)	97.9 (8.0)	97.8 (3.9)	98.7 (4.9)	92.0 (17.9)	96.6 (5.4)
SOM	100 (0.0)	100 (0.0)	90.6 (16.8)	100 (0.0)	100 (0.0)	96.9 (4.9)	100 (0.0)	100 (0.0)	99.4 (1.8)	100 (0.0)	100 (0.0)	99.1 (1.7)
K-means	100 (0.0)	100 (0.0)	98.0 (6.0)	100 (0.0)	100 (0.0)	99.1 (2.7)	100 (0.0)	100 (0.0)	98.2 (2.7)	100 (0.0)	100 (0.0)	98.7 (2.9)

## 1.2. Hourly

### 1.2.1. The 1<sup>st</sup> Infection Episode (Flu)

**Table 9:** Univariate Input Smoothed Data (different sample sizes) - Average and standard deviation (F1-score, AUC, specificity).

Fraction = 0.01												
Models	Boundary and Domain-Based Method											
	1 Month			2 Months			3 Months			4 Months		
	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1
SVDD	97.8 (3.4)	78.7 (4.9)	83.0 (2.0)	98.2 (1.5)	81.7 (4.0)	92.0 (0.9)	98.1 (2.4)	85.4 (3.5)	94.5 (2.9)	99.0 (1.3)	91.3 (1.4)	97.1 (2.3)
$\nu$ -SVM	99.3 (1.2)	86.2 (2.0)	91.4 (1.7)	99.7 (0.5)	93.1 (0.7)	96.3 (1.7)	99.8 (0.5)	94.2 (0.0)	97.1 (1.6)	99.8 (0.3)	94.2 (0.0)	97.1 (1.5)
Nearest Neighbor	83.8 (5.3)	51.1 (13.3)	69.7 (5.6)	85.7 (4.5)	54.6 (12.4)	82.8 (3.6)	86.2 (3.1)	54.8 (12.4)	88.1 (2.1)	86.6 (3.4)	56.8 (12.0)	91.0 (1.3)
MST	87.5 (2.6)	70.9 (7.5)	78.8 (2.5)	87.8 (2.7)	71.1 (7.5)	88.3 (1.5)	88.0 (2.7)	71.8 (7.2)	92.2 (0.9)	88.0 (2.7)	74.0 (6.5)	94.4 (0.7)
Density-Based Method												
Gaussian OC	99.3 (1.2)	83.8 (3.5)	86.8 (1.6)	99.7 (0.5)	92.1 (1.2)	96.0 (2.0)	99.8 (0.5)	94.8 (1.7)	97.2 (3.1)	99.8 (0.3)	94.4 (0.5)	98.2 (1.1)
MoG	98.8 (1.8)	81.6 (3.8)	85.0 (1.7)	99.6 (0.6)	90.2 (1.8)	95.0 (1.6)	99.8 (0.5)	94.7 (1.9)	97.3 (3.1)	99.7 (0.4)	94.5 (0.5)	97.9 (1.2)
MCD Gaussian	99.3 (1.2)	83.8 (3.5)	86.8 (1.6)	99.7 (0.5)	92.1 (1.2)	96.0 (2.0)	99.8 (0.5)	94.8 (1.7)	97.2 (3.1)	99.8 (0.3)	94.4 (0.5)	98.2 (1.1)
Parzen	98.6 (2.3)	80.2 (4.4)	82.2 (5.8)	99.4 (1.0)	89.7 (1.5)	94.5 (2.6)	99.5 (0.8)	95.9 (1.6)	96.4 (4.2)	99.6 (0.8)	94.8 (0.8)	97.0 (3.7)
Naive Parzen	98.6 (2.3)	80.2 (4.4)	82.2 (5.8)	99.4 (1.0)	89.7 (1.5)	94.5 (2.6)	99.5 (0.8)	95.9 (1.6)	96.4 (4.2)	99.6 (0.8)	94.8 (0.8)	97.0 (3.7)
k-NN	99.3 (1.2)	83.8 (3.5)	86.8 (1.6)	99.7 (0.5)	92.1 (1.2)	96.0 (2.0)	99.8 (0.5)	94.8 (1.7)	97.2 (3.1)	99.8 (0.3)	94.4 (0.5)	98.2 (1.1)
LOF	99.3 (1.1)	83.8 (3.5)	86.8 (1.6)	99.7 (0.5)	92.1 (1.2)	96.0 (2.0)	99.8 (0.5)	94.8 (1.7)	97.2 (3.1)	99.7 (0.3)	94.4 (0.5)	98.2 (1.1)
Reconstruction-Based Method												
Auto – encoder	88.0 (5.9)	71.3 (9.5)	78.4 (5.1)	91.0 (4.2)	74.6 (9.7)	89.0 (3.1)	91.0 (5.3)	76.3 (10.5)	92.2 (4.3)	92.0 (5.3)	77.3 (10.1)	93.6 (4.2)
SOM	92.3 (2.3)	68.5 (9.9)	76.8 (2.8)	96.1 (3.1)	84.2 (8.4)	92.1 (3.6)	97.9 (0.8)	90.7 (2.9)	96.4 (1.8)	98.1 (0.7)	92.7 (1.5)	97.0 (2.6)
K-means	99.3 (1.2)	83.8 (3.5)	86.8 (1.6)	99.7 (0.5)	92.1 (1.2)	96.0 (2.0)	99.8 (0.5)	94.8 (1.7)	97.2 (3.1)	99.8 (0.3)	94.4 (0.5)	98.2 (1.1)

### 1.2.2. The 2<sup>nd</sup> Infection Episode (Flu)

**Table 10:** Univariate Input Smoothed Data (different sample sizes) - Average and standard deviation (F1-score, AUC, specificity).

Fraction = 0.01												
Models	Boundary and Domain-Based Method											
	1 Month			2 Months			3 Months			4 Months		
	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1
SVDD	100 (0.0)	100 (0.0)	96.0 (10.4)	100 (0.0)	100 (0.0)	98.8 (2.6)	99.5 (0.6)	86.0 (4.3)	95.3 (0.7)	99.6 (0.6)	89.1 (6.1)	89.1 (0.9)
$\nu$ -SVM	100 (0.0)	100 (0.0)	99.6 (1.1)	100 (0.0)	100 (0.0)	99.5 (1.3)	99.8 (0.4)	91.9 (1.0)	96.2 (1.5)	99.9 (0.3)	95.3 (0.5)	97.5 (1.6)
Nearest Neighbor	99.9 (0.1)	99.5 (1.5)	99.1 (1.1)	100 (0.0)	100 (0.0)	99.2 (0.5)	84.6 (1.4)	66.8 (8.5)	91.0 (1.9)	84.7 (1.2)	68.0 (7.8)	93.0 (1.4)
MST	100 (0.0)	100 (0.0)	96.9 (7.9)	100 (0.0)	100 (0.0)	99.4 (1.2)	86.8 (1.6)	71.6 (4.0)	92.3 (0.7)	86.7 (1.8)	71.9 (3.9)	93.9 (1.3)
Density-Based Method												
Gaussian OC	100 (0.0)	100 (0.0)	96.3 (11.2)	100 (0.0)	100 (0.0)	98.2 (4.9)	99.8 (0.4)	92.1 (2.7)	97.2 (0.9)	99.9 (0.3)	96.0 (1.8)	98.3 (1.6)
MoG	100 (0.0)	100 (0.0)	96.0 (10.7)	100 (0.0)	100 (0.0)	98.0 (4.8)	99.5 (0.6)	90.3 (3.9)	96.2 (1.4)	99.8 (0.4)	93.3 (3.6)	97.3 (1.7)
MCD Gaussian	100 (0.0)	100 (0.0)	96.3 (11.2)	100 (0.0)	100 (0.0)	98.2 (5.1)	99.8 (0.4)	92.1 (2.7)	97.2 (0.9)	99.9 (0.3)	96.0 (1.8)	98.3 (1.6)
Parzen	100 (0.0)	100 (0.0)	96.1 (10.7)	100 (0.0)	100 (0.0)	97.9 (5.2)	99.3 (0.7)	88.5 (2.0)	95.9 (1.5)	99.4 (0.8)	90.2 (2.8)	96.5 (2.0)
Naive Parzen	100 (0.0)	100 (0.0)	96.1 (10.7)	100 (0.0)	100 (0.0)	97.9 (5.2)	99.3 (0.7)	88.5 (2.0)	95.9 (1.5)	99.4 (0.8)	90.2 (2.8)	96.5 (2.0)
k-NN	100 (0.0)	100 (0.0)	96.2 (11.1)	100 (0.0)	100 (0.0)	97.9 (5.2)	99.2 (0.7)	88.3 (2.5)	95.8 (1.3)	99.6 (0.5)	89.2 (3.7)	96.5 (1.4)
LOF	100 (0.0)	100 (0.0)	96.6 (9.6)	100 (0.0)	100 (0.0)	98.1 (5.1)	91.6 (0.9)	86.6 (1.0)	96.0 (0.8)	91.9 (2.3)	87.1 (1.7)	96.6 (1.5)
Reconstruction-Based Method												
Auto – encoder	98.2 (4.1)	97.2 (5.1)	94.6 (9.3)	98.6 (2.7)	96.3 (6.6)	96.9 (4.6)	91.9 (6.7)	76.1 (11.1)	92.8 (2.7)	94.0 (5.3)	79.7 (10.8)	94.7 (2.7)
SOM	100 (0.0)	100 (0.0)	96.4 (10.0)	100 (0.0)	100 (0.0)	98.0 (4.1)	95.0 (4.5)	85.3 (5.4)	94.9 (1.8)	98.0 (2.1)	88.2 (3.1)	96.5 (1.9)
K-means	100 (0.0)	100 (0.0)	96.3 (11.2)	100 (0.0)	100 (0.0)	98.2 (4.9)	99.8 (0.4)	92.1 (2.7)	97.2 (0.9)	99.9 (0.3)	96.0 (1.8)	98.3 (1.6)

### 1.2.3. The 3<sup>rd</sup> Infection Episode (Flu)

**Table 11:** Univariate Input Smoothed Data (different sample sizes) - Average and standard deviation (F1-score, AUC, specificity).

Fraction = 0.01												
Models	Boundary and Domain-Based Method											
	1 Month			2 Months			3 Months			4 Months		
	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1
SVDD	97.2 (2.6)	85.9 (2.9)	87.3 (15.0)	97.0 (1.7)	86.0 (3.6)	93.6 (6.9)	97.7 (2.6)	84.9 (1.8)	95.6 (4.1)	98.0 (1.4)	86.5 (4.3)	95.6 (4.1)
$\nu$ -SVM	97.7 (2.7)	85.9 (0.7)	92.6 (1.5)	99.2 (1.5)	88.9 (0.0)	95.0 (2.1)	98.6 (1.3)	85.3 (0.4)	94.0 (1.7)	98.9 (0.9)	86.8 (0.8)	94.6 (1.9)
Nearest Neighbor	87.8 (3.3)	55.8 (9.7)	78.7 (5.0)	90.2 (1.5)	67.8 (8.8)	88.5 (5.7)	89.1 (1.9)	64.1 (3.5)	92.3 (2.6)	89.0 (2.6)	67.3 (3.2)	94.5 (2.0)
MST	90.2 (0.3)	81.1 (3.0)	86.4 (10.6)	90.5 (1.2)	81.6 (4.1)	92.5 (6.6)	89.9 (1.5)	80.4 (0.4)	95.3 (3.5)	90.1 (0.8)	80.4 (0.4)	96.7 (1.9)
Density-Based Method												
Gaussian OC	97.5 (2.7)	85.8 (1.9)	87.1 (15.6)	99.2 (1.5)	89.8 (2.6)	93.4 (10.8)	98.5 (1.3)	85.7 (1.1)	96.7 (1.1)	98.9 (0.9)	87.2 (1.1)	97.5 (1.1)
MoG	97.2 (2.6)	85.1 (3.0)	86.6 (15.8)	99.0 (1.5)	89.6 (2.2)	93.5 (10.4)	98.2 (2.2)	82.2 (2.8)	95.2 (3.2)	98.7 (1.2)	85.7 (2.0)	96.8 (1.7)
MCD Gaussian	97.5 (2.7)	85.8 (1.9)	87.6 (14.2)	99.3 (1.4)	89.8 (2.6)	93.9 (9.3)	98.7 (1.3)	85.7 (1.1)	96.7 (1.1)	99.1 (0.9)	87.2 (1.1)	97.5 (1.1)
Parzen	95.9 (2.7)	84.9 (3.1)	86.1 (15.0)	98.8 (1.5)	88.1 (3.2)	92.6 (11.3)	97.6 (2.4)	82.3 (1.5)	94.1 (4.3)	98.3 (1.4)	83.5 (1.5)	95.6 (2.7)
Naive Parzen	95.9 (2.7)	84.9 (3.1)	86.1 (15.0)	98.8 (1.5)	88.1 (3.2)	92.6 (11.3)	97.6 (2.4)	82.3 (1.5)	94.1 (4.3)	98.3 (1.4)	83.5 (1.5)	95.6 (2.7)
k-NN	96.1 (2.4)	84.2 (3.9)	86.2 (15.1)	98.8 (1.4)	88.5 (2.6)	92.8 (11.4)	97.5 (2.6)	81.0 (1.6)	94.1 (4.2)	98.3 (1.4)	82.5 (1.5)	95.8 (2.5)
LOF	92.6 (1.7)	78.8 (5.2)	84.5 (12.4)	93.7 (1.3)	77.2 (6.4)	91.2 (6.5)	90.7 (1.5)	81.6 (0.9)	94.5 (4.2)	90.8 (0.8)	81.5 (0.6)	96.3 (2.1)
Reconstruction-Based Method												
Auto – encoder	90.1 (5.4)	79.1 (6.7)	84.9 (11.6)	93.5 (4.1)	81.0 (6.9)	92.1 (7.5)	93.2 (3.4)	80.2 (5.7)	94.7 (3.9)	93.8 (3.2)	80.9 (6.3)	95.9 (2.5)
SOM	95.2 (1.6)	85.6 (2.8)	87.0 (14.5)	96.5 (0.8)	87.5 (3.2)	93.3 (9.0)	93.1 (2.2)	81.7 (2.2)	94.1 (4.2)	94.3 (1.6)	82.1 (2.1)	96.2 (2.1)
K-means	91.9 (1.8)	79.2 (5.3)	83.8 (14.7)	95.1 (1.6)	86.2 (5.3)	92.8 (9.6)	93.8 (2.5)	79.0 (3.9)	94.4 (4.2)	95.7 (0.9)	84.1 (2.3)	96.7 (2.1)

### 1.2.4. The 4th Infection Episode (Flu)

**Table 12:** Univariate Input Smoothed Data (different sample sizes) - Average and standard deviation (F1-score, AUC, specificity).

Fraction = 0.01												
Models	Boundary and Domain-Based Method											
	1 Month			2 Months			3 Months			4 Months		
	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1	AUC	Specificity	F1
SVDD	99.0 (1.9)	93.1 (3.2)	89.1 (10.2)	97.9 (1.5)	86.6 (3.0)	94.7 (1.9)	98.5 (1.1)	87.0 (2.6)	96.8 (0.7)	99.1 (0.8)	93.4 (4.2)	94.6 (0.3)
$\nu$ -SVM	99.7 (0.7)	93.4 (0.0)	96.0 (1.3)	99.6 (0.6)	91.0 (0.5)	95.7 (1.8)	99.8 (0.4)	92.7 (0.5)	96.6 (1.8)	99.8 (0.3)	93.4 (0.0)	96.8 (2.1)
Nearest Neighbor	94.5 (1.7)	81.2 (1.9)	88.1 (3.0)	89.8 (1.4)	72.3 (4.2)	91.2 (0.9)	90.8 (1.4)	76.5 (2.0)	94.7 (0.4)	91.2 (1.5)	77.5 (2.1)	96.1 (0.3)
MST	95.5 (1.0)	90.9 (2.3)	89.7 (9.0)	91.5 (1.1)	83.2 (2.3)	94.4 (0.8)	91.6 (1.1)	83.2 (2.3)	96.2 (0.5)	91.6 (1.2)	83.2 (2.3)	97.1 (0.4)
Density-Based Method												
Gaussian OC	99.6 (0.8)	94.1 (2.0)	93.3 (8.1)	99.6 (0.6)	90.4 (1.0)	96.4 (1.8)	99.7 (0.4)	92.7 (0.9)	97.6 (1.6)	99.8 (0.3)	93.8 (0.9)	98.2 (1.5)
MoG	99.1 (1.6)	93.4 (2.3)	90.1 (10.3)	99.4 (0.7)	90.3 (1.1)	96.0 (2.1)	99.7 (0.4)	92.7 (0.9)	97.6 (1.6)	99.8 (0.3)	93.9 (1.0)	98.2 (1.5)
MCD Gaussian	99.7 (0.7)	94.1 (2.0)	93.3 (8.1)	99.6 (0.6)	90.4 (1.0)	96.4 (1.8)	99.7 (0.4)	92.7 (0.9)	97.6 (1.6)	99.7 (0.3)	93.7 (0.7)	98.2 (1.4)
Parzen	99.1 (1.5)	94.9 (1.7)	90.9 (10.6)	99.6 (0.7)	91.0 (1.8)	95.9 (1.9)	99.7 (0.5)	92.6 (2.2)	97.3 (1.7)	99.8 (0.3)	95.2 (1.3)	98.2 (1.7)
Naive Parzen	99.1 (1.5)	94.9 (1.7)	90.9 (10.6)	99.6 (0.7)	91.0 (1.8)	95.9 (1.9)	99.7 (0.5)	92.6 (2.2)	97.3 (1.7)	99.8 (0.3)	95.2 (1.3)	98.2 (1.7)
k-NN	99.1 (1.7)	93.0 (2.4)	89.9 (10.8)	99.6 (0.6)	90.4 (1.8)	96.4 (1.5)	99.8 (0.4)	92.0 (1.6)	97.4 (1.5)	99.8 (0.3)	94.7 (1.3)	98.2 (1.6)
LOF	96.1 (1.5)	92.3 (2.4)	89.7 (10.7)	97.0 (1.0)	89.9 (1.5)	94.8 (3.9)	97.4 (0.6)	91.2 (1.3)	96.7 (2.7)	97.1 (1.5)	91.5 (0.7)	97.6 (1.9)
Reconstruction-Based Method												
Auto – encoder	95.1 (5.5)	89.5 (5.7)	88.9 (9.5)	94.7 (4.0)	80.6 (11.2)	92.9 (3.6)	95.7 (3.8)	82.8 (10.6)	95.4 (2.4)	96.3 (3.9)	85.1 (10.9)	96.7 (2.1)
SOM	97.3 (1.2)	92.9 (2.4)	89.7 (10.5)	96.7 (1.4)	88.5 (2.4)	94.6 (3.1)	97.5 (0.6)	90.8 (1.5)	96.7 (2.4)	97.4 (0.5)	91.1 (1.2)	97.5 (1.8)
K-means	96.5 (1.6)	93.2 (2.3)	90.3 (9.5)	96.2 (1.0)	89.5 (2.1)	95.8 (2.2)	97.1 (1.2)	91.4 (2.5)	97.2 (1.6)	98.2 (1.0)	93.2 (2.0)	98.0 (1.4)

## 2. Unsupervised Method

**Table 13:** Univariate Input Raw and Smoothed Data for both the daily and hourly data granularity - Average and standard deviation (F1-score, AUC, specificity). The parameters  $k_d$ , and  $k_h$  represent the optimal number of nearest neighbors for the daily and hourly cases respectively.

Freq.		Density-Based Methods												
	Pre-pro.	Methods (Threshold)	1 <sup>st</sup> Infection Episode ( $k_d=30, k_h=240$ )			2 <sup>nd</sup> Infection Episode ( $k_d=30, k_h=240$ )			3 <sup>rd</sup> Infection Episode ( $k_d=30, k_h=240$ )			4 <sup>th</sup> Infection Episode ( $k_d=30, k_h=240$ )		
			AUC	Sensitivity	F1	AUC	Sensitivity	F1	AUC	Sensitivity	F1	AUC	Sensitivity	F1
Daily	Without filter	LOF ( $T_1=2.7, T_2=1.5, T_3=2.95, T_4=2.2$ )	90.7	83.3	79.5	89.4	83.3	73.1	90.1	66.7	90.9	86.7	75.0	75.3
		COF ( $T_1=1.4, T_2=1.1, T_3=2.3, T_4=1.8$ )	89.4	83.3	73.1	85.0	100	65.0	83.1	66.7	90.9	86.7	75.0	75.3
	With filter	LOF ( $T_1=1.9, T_2=1.9, T_3=2.8, T_4=2.8$ )	99.9	100	96.2	100	100	100	100	100	100	99.6	100	87.8
		COF ( $T_1=1.6, T_2=1.5, T_3=2.8, T_4=3.1$ )	99.9	100	96.2	100	100	100	100	100	100	99.9	100	94.7
Hourly		LOF ( $T_1=1.9, T_2=1.6, T_3=1.2, T_4=1.7$ )	95.0	90.1	94.8	94.2	89.7	81.1	93.9	91.4	73.0	86.8	75.3	73.4
		COF ( $T_1=1.6, T_2=1.3, T_3=1.2, T_4=1.3$ )	94.9	90.1	94.2	87.7	76.3	80.0	96.5	96.3	74.8	86.8	75.3	73.2

## Appendix B - Score plot of the models using the univariate input feature: Ratio of Insulin to Carbohydrate

The following appendix presents the performance score of the different models evaluated in the dissertation using univariate input – insulin-to-carbohydrate ratio. Both the one-class classifiers and unsupervised models were tested. For the one-class classifier models, a random block of a 4-month sample size from each patient-year was used to train the models and evaluated with the whole patient-year – containing both the regular and infection days. Unsupervised models were tested with the whole patient-year. The smoothed version of the dataset is used and the models were tested with both the hourly and daily datasets. The score plot of the models is given below, and, as can be seen from the figures, each model generated different scores and rejected varying portions of the infection episodes.

### 1. One-class classifier Method

#### 1.1. Daily

##### 1.1.1. *The First Infection Episode (Flu)*

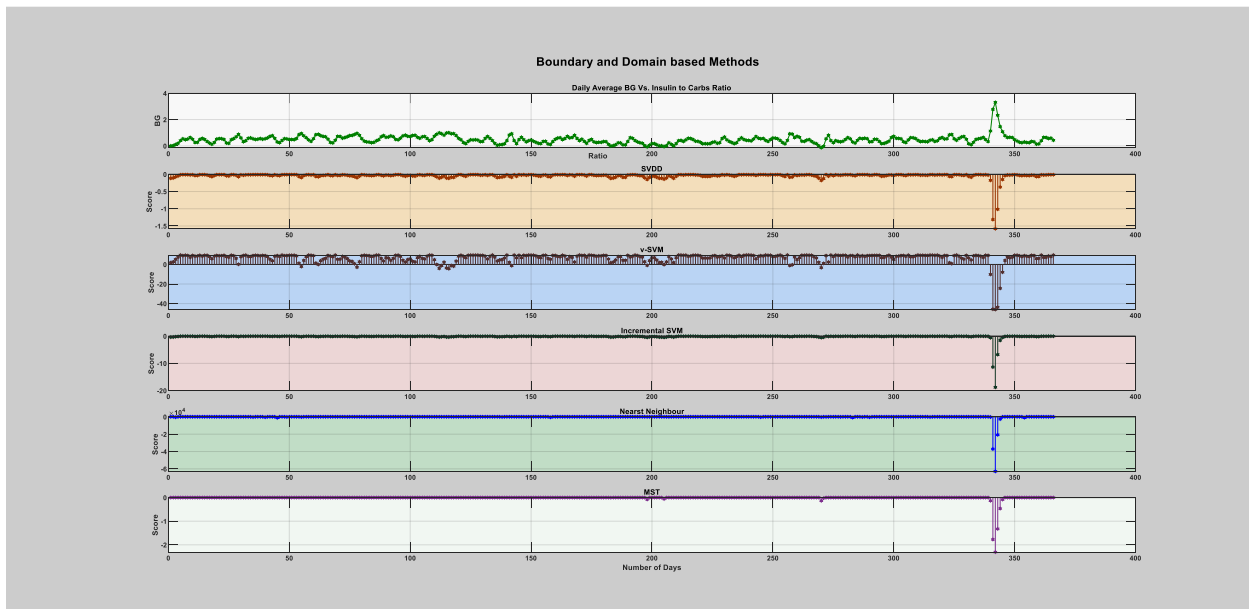
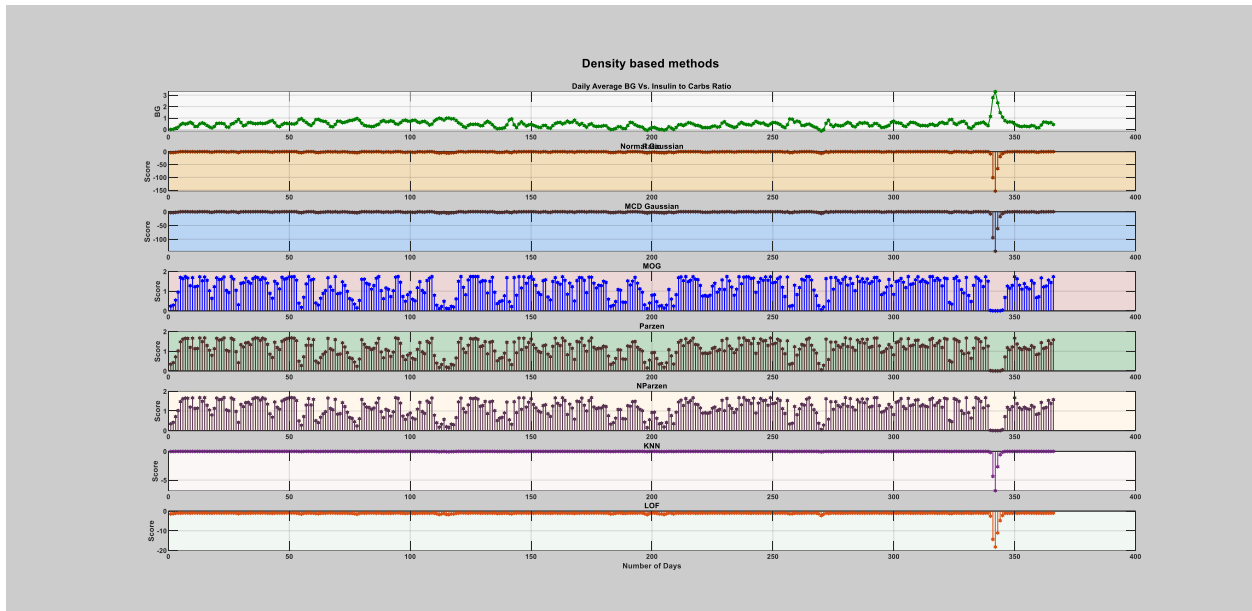
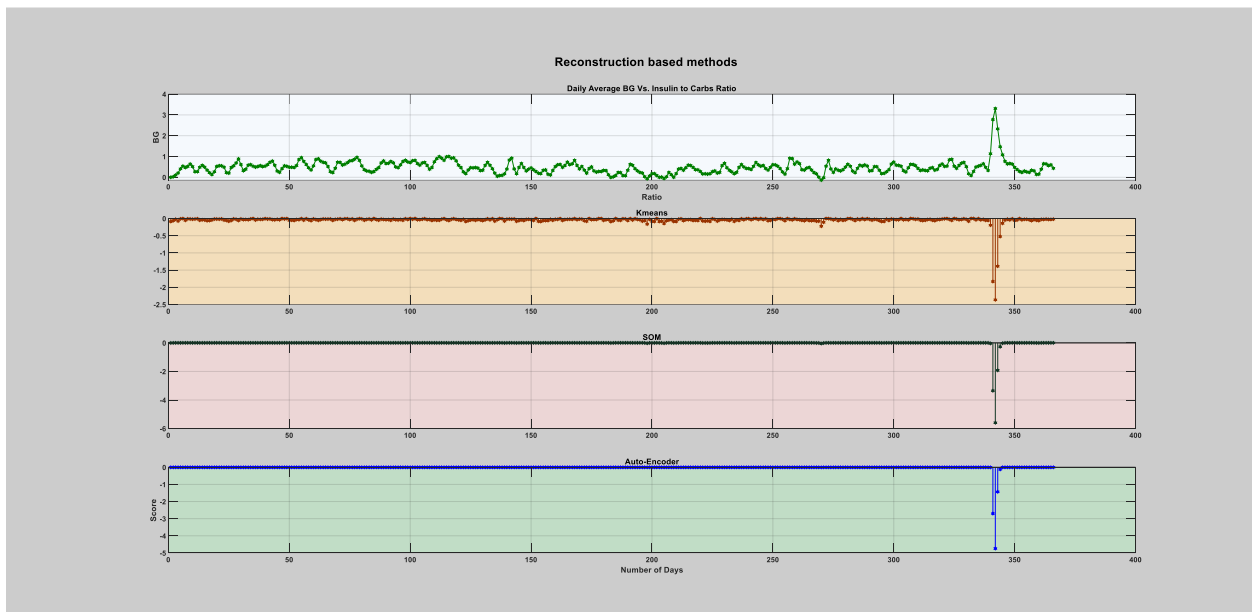


Figure 1: Boundary and domain-based method.



**Figure 2:** Density-based method.



**Figure 3:** Reconstruction-based method.



### 1.1.2. The Second Infection Episode (Flu)

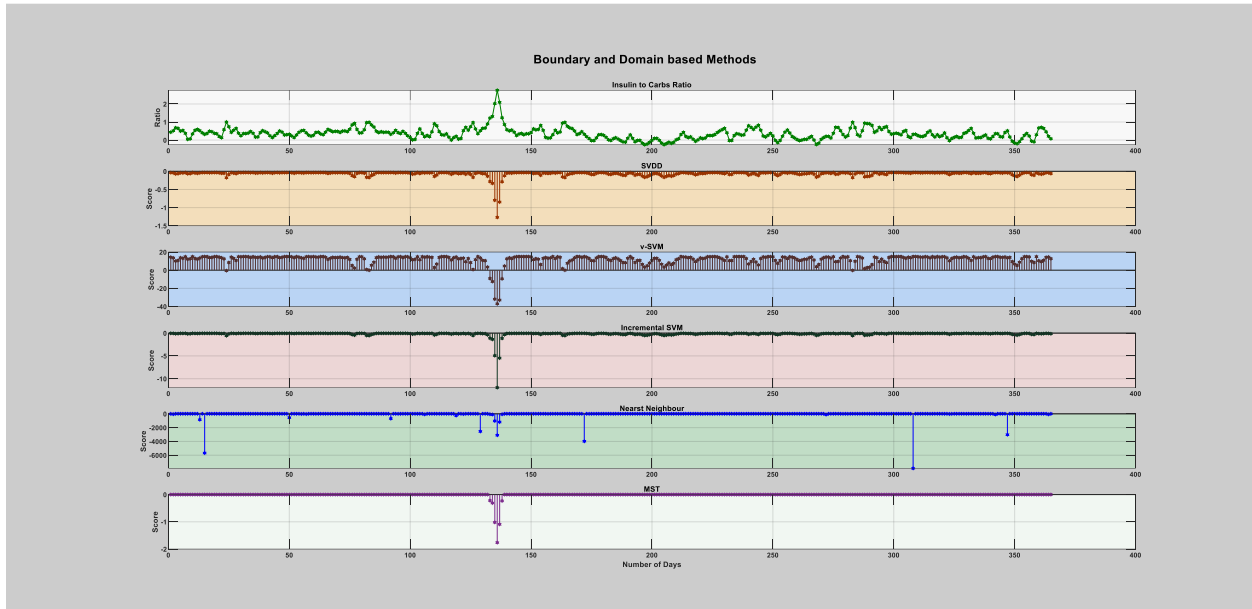


Figure 4: Boundary and domain-based method.

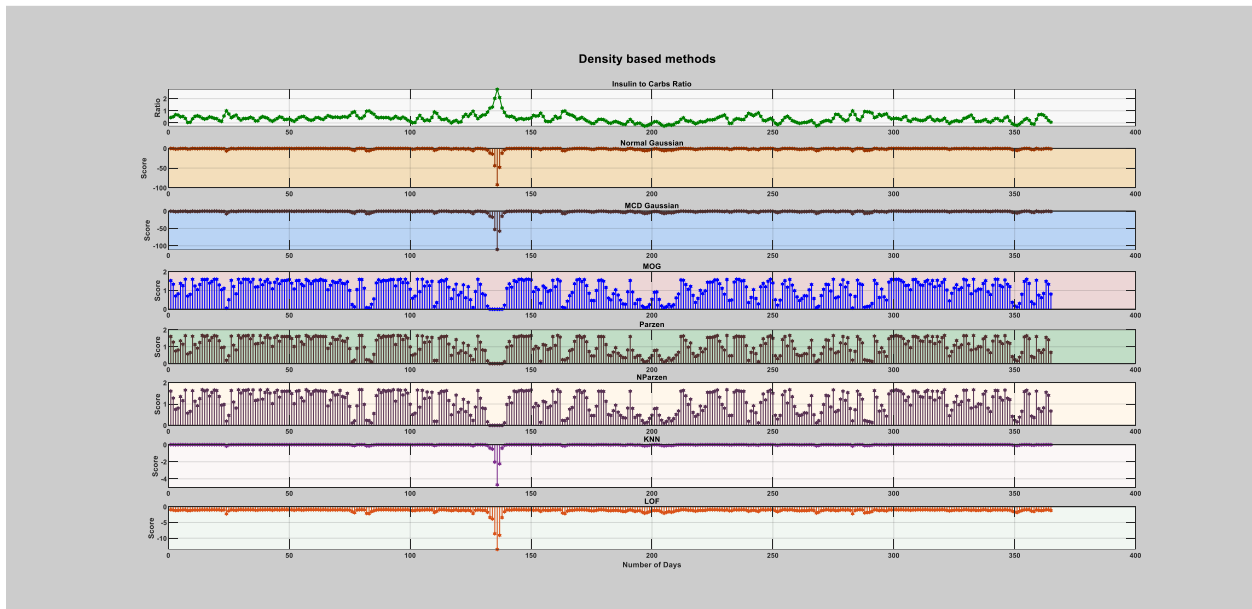
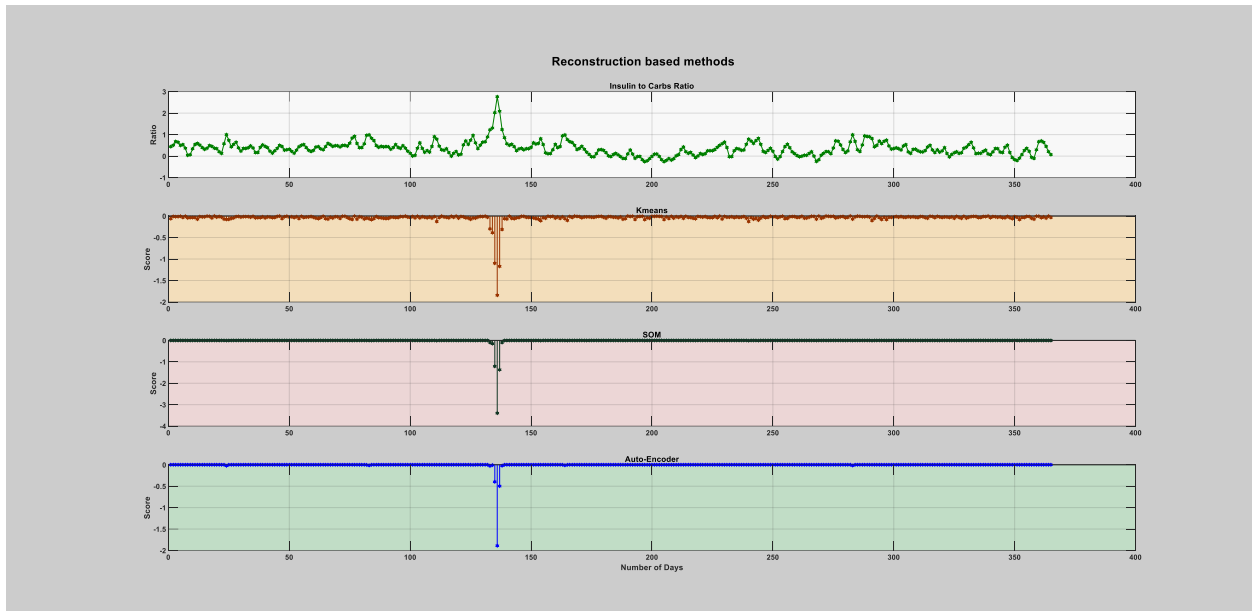
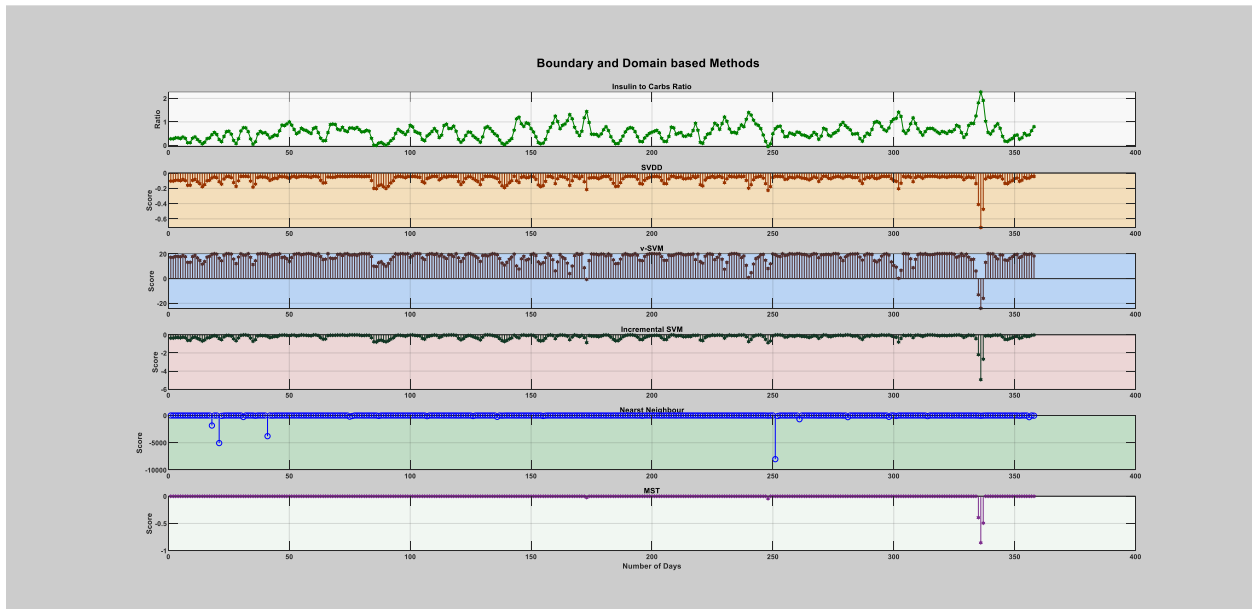


Figure 5: Density-based method.



**Figure 6:** Reconstruction-based method.

### 1.1.3. *The Third Infection Episode (Flu)*



**Figure 7:** Boundary and domain-based method.

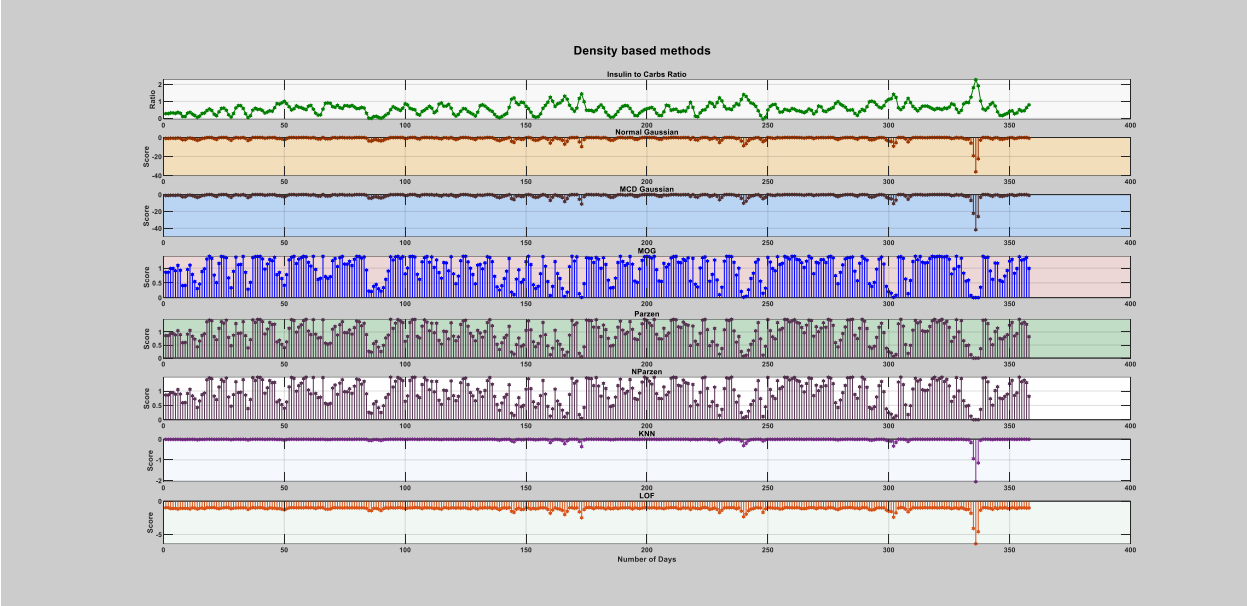


Figure 8: Density-based method.

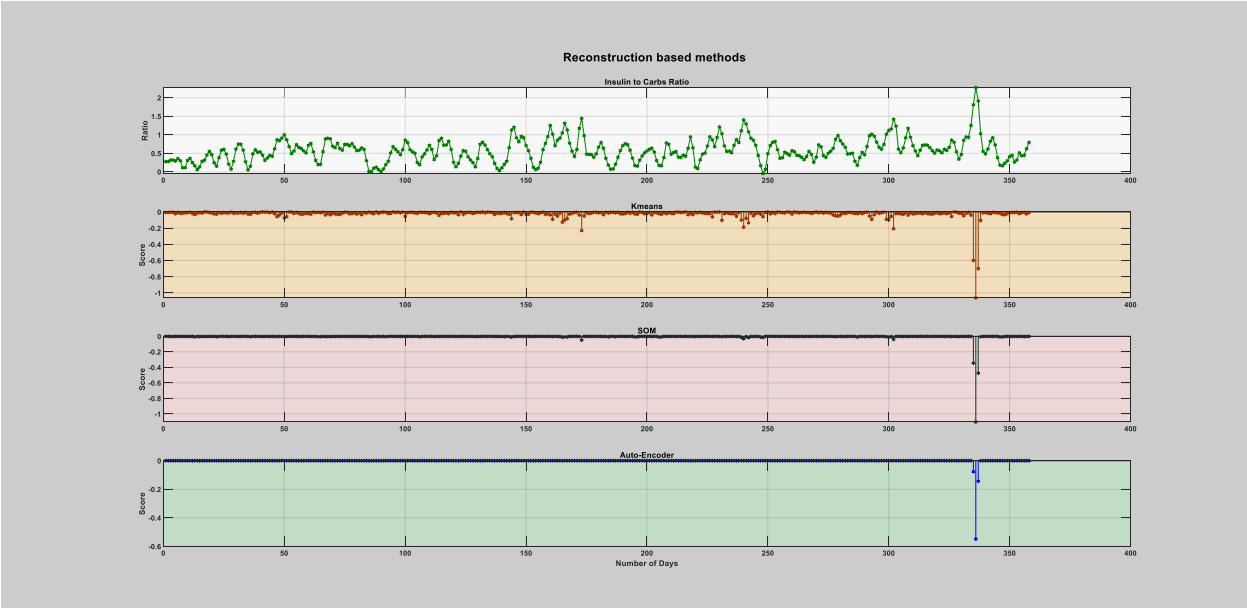


Figure 9: The score of the reconstruction-based method on the whole patient-year.

### 1.1.4. The Fourth Infection Episode (Flu)

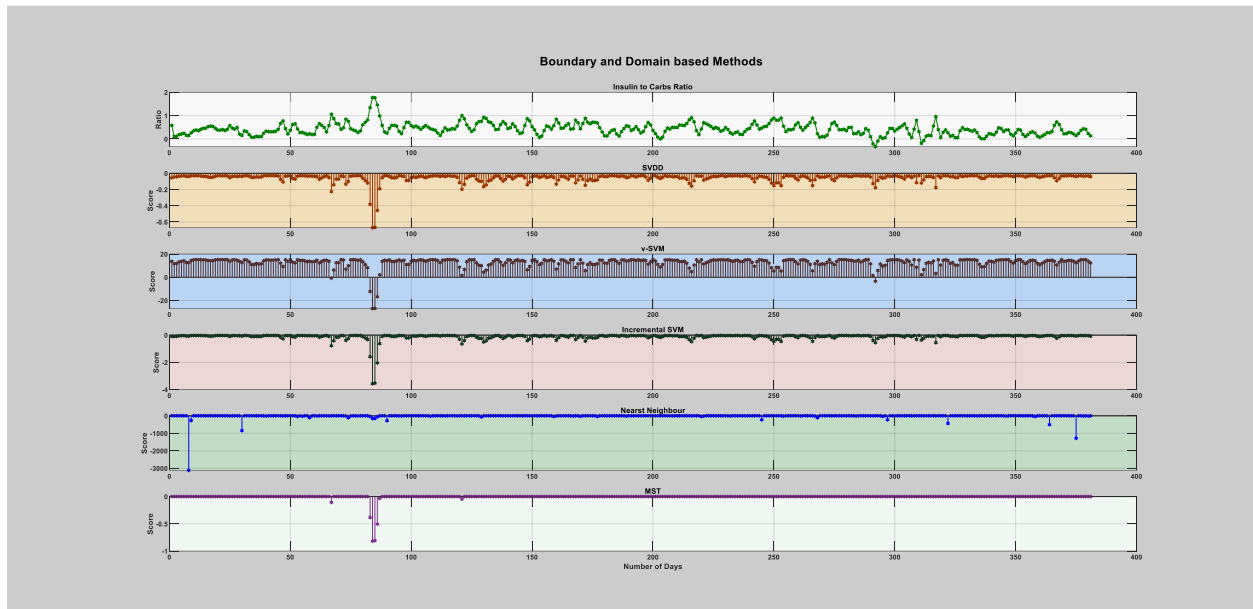


Figure 10: Boundary and domain-based.

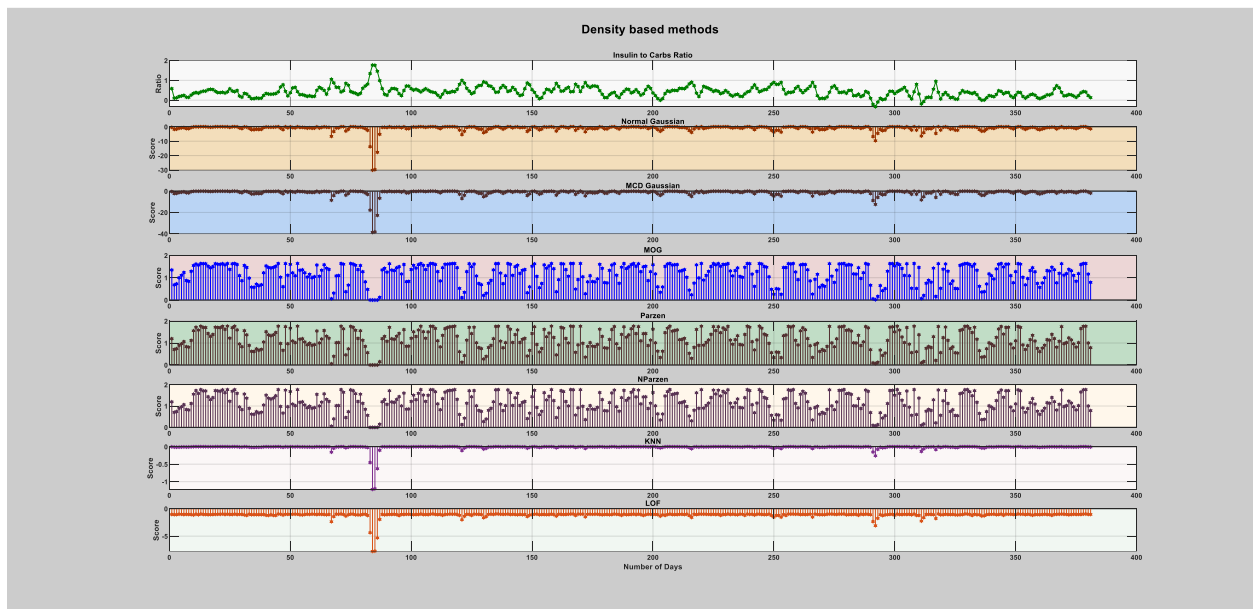
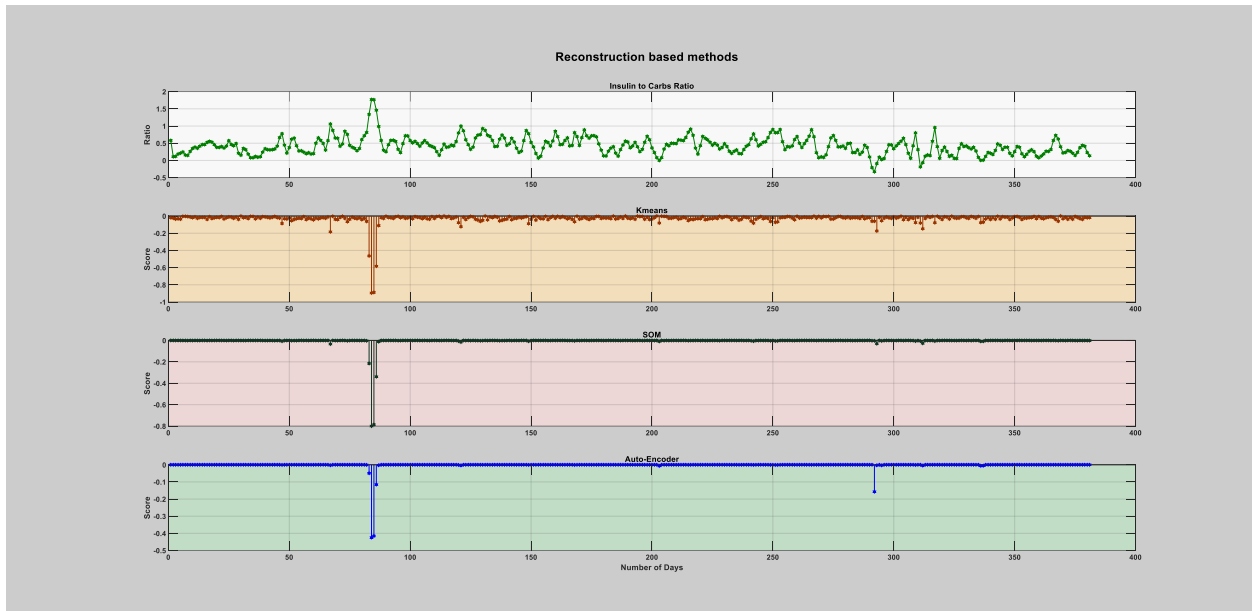


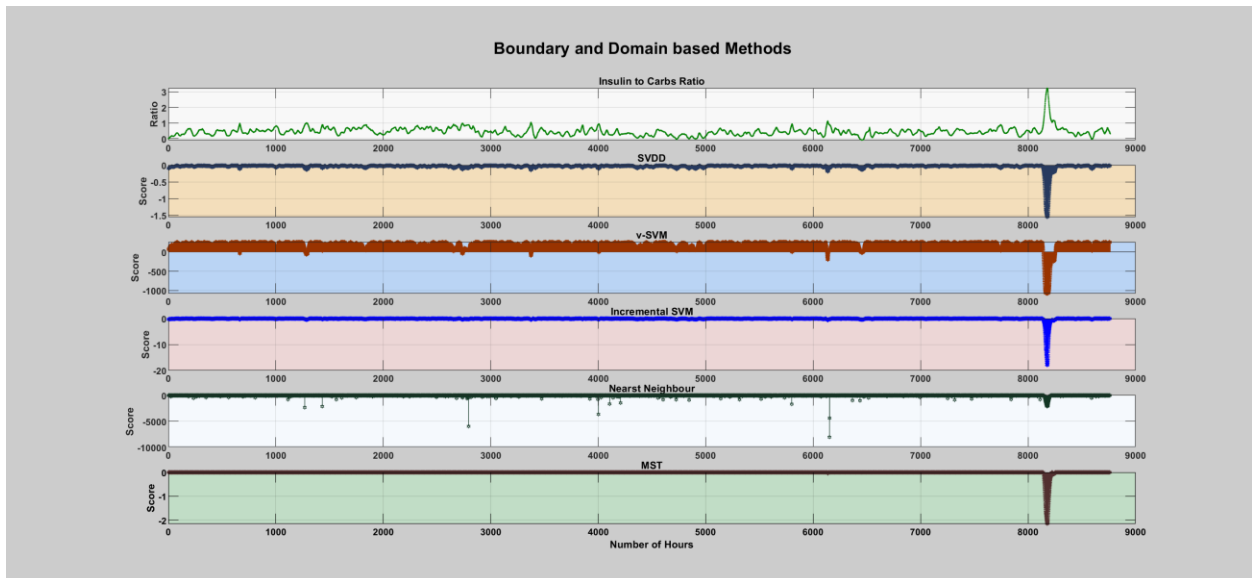
Figure 11: Density-based method.



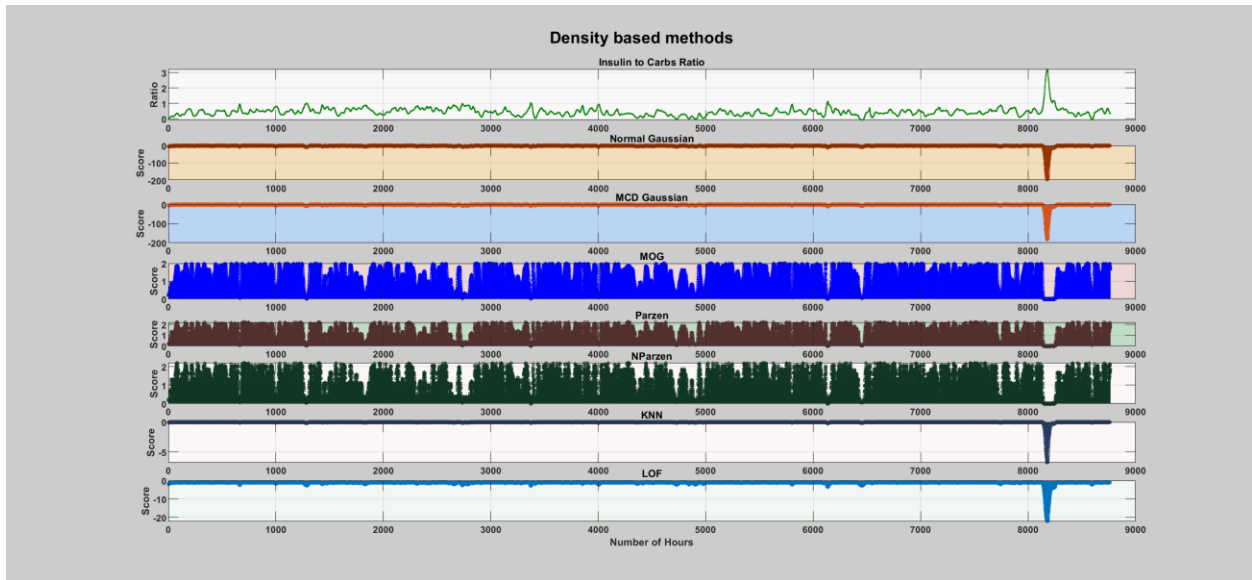
**Figure 12:** Reconstruction-based method.

## 1.2. Hourly

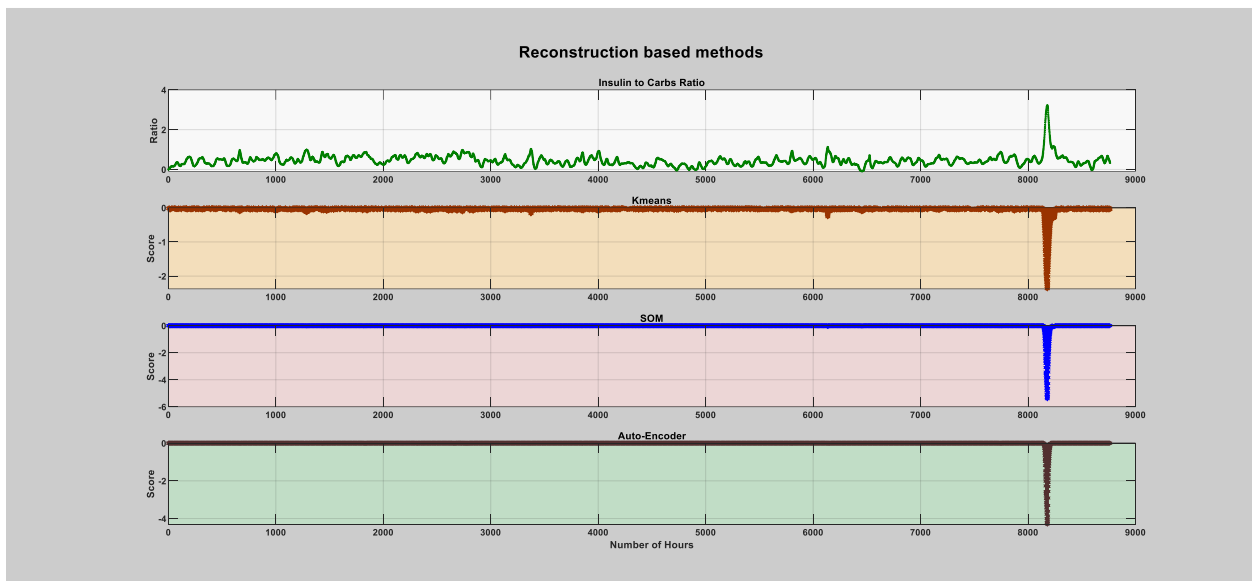
### 1.2.1. *The First Infection Episode (Flu)*



**Figure 13:** Boundary and domain-based method.



**Figure 14:** Density-based method.



**Figure 15:** Reconstruction-based method.

### 1.2.2. The Second Infection Episode (Flu)

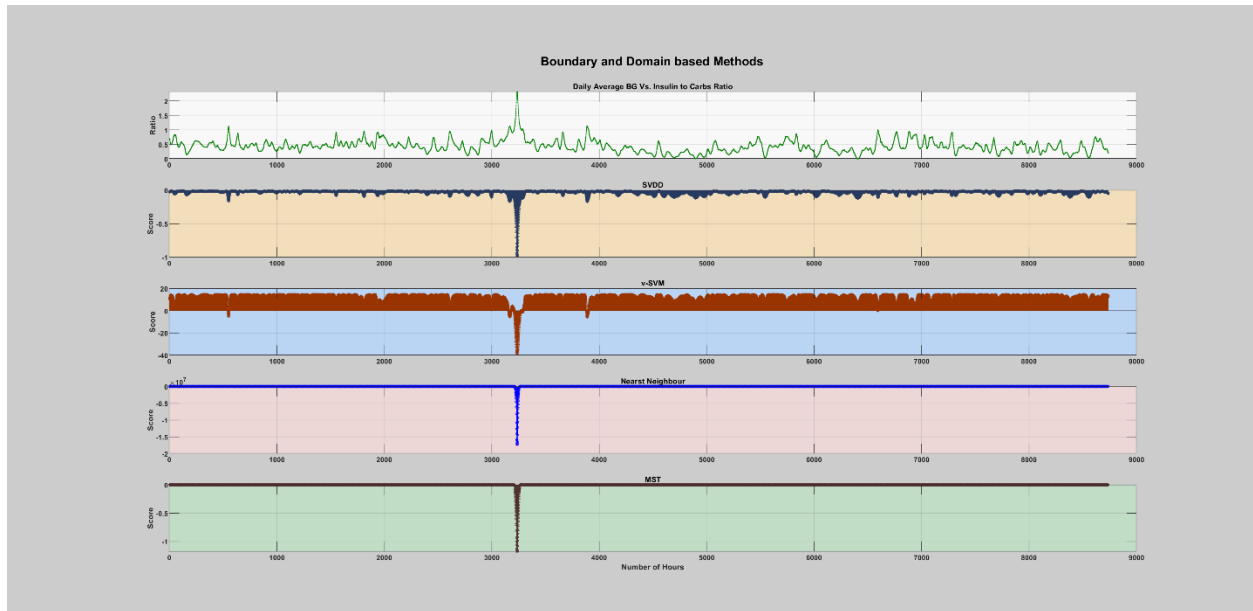


Figure 16: Boundary and domain-based.

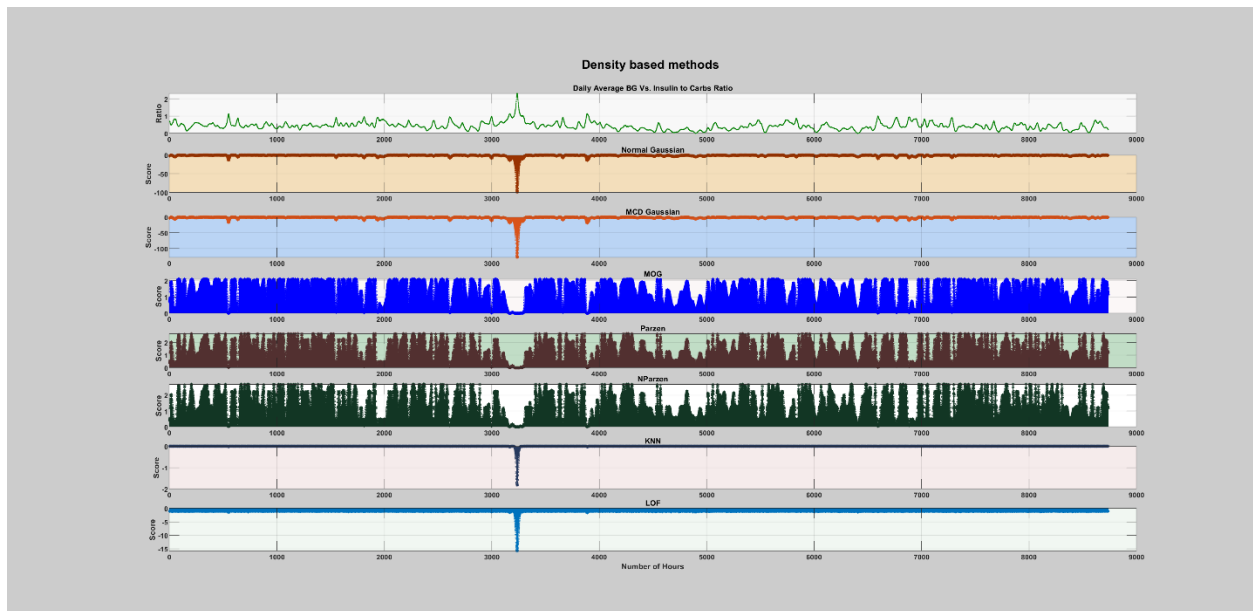
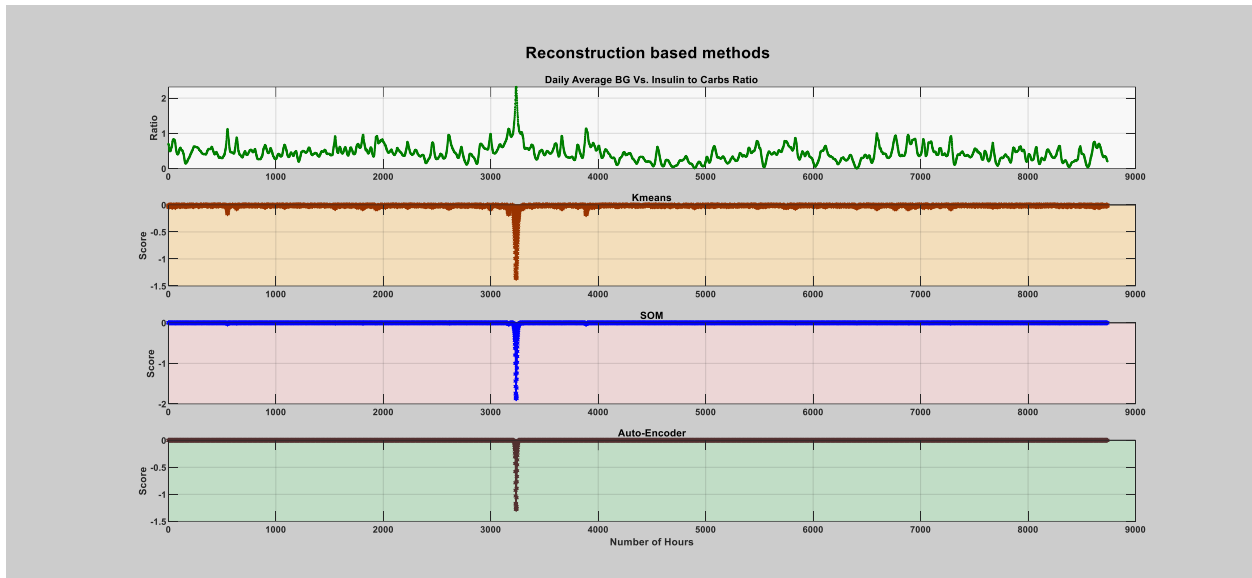
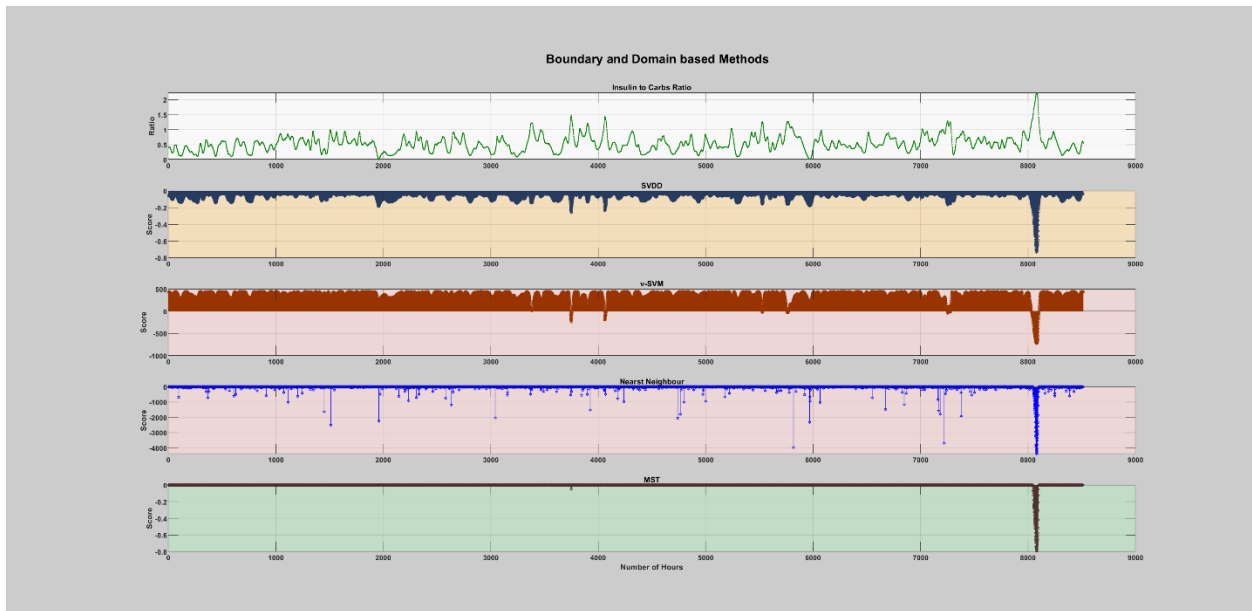


Figure 17: Density-based method.



**Figure 18:** Reconstruction-based method.

### 1.2.3. *The Third Infection Episode (Flu)*



**Figure 19:** Boundary and domain-based method.



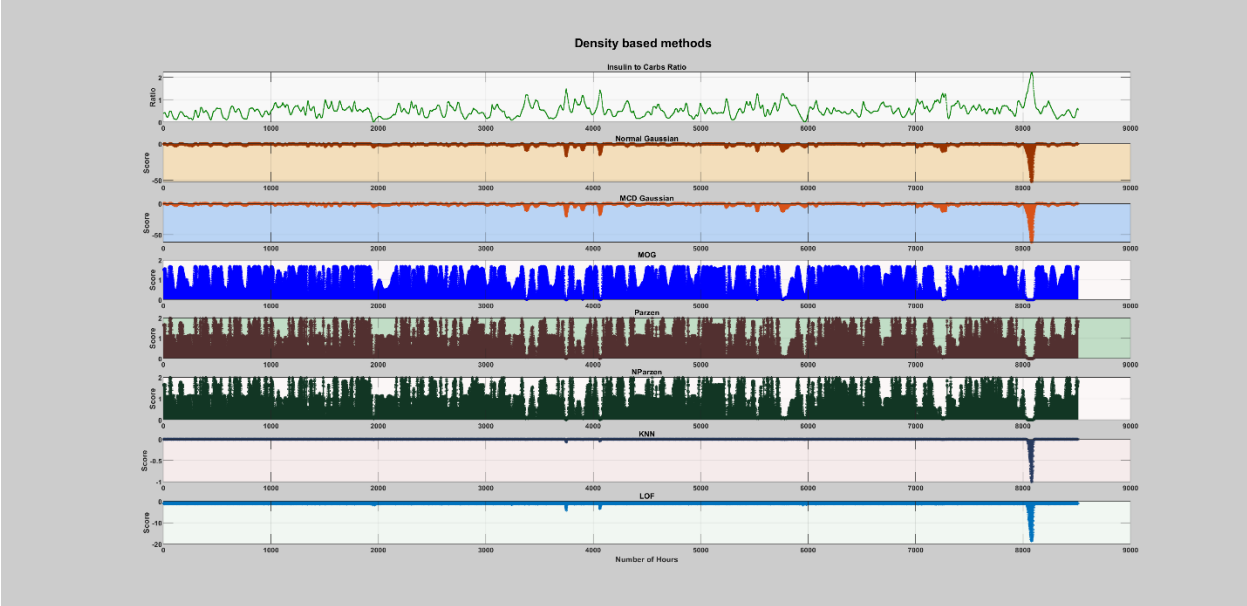


Figure 20: Density-based method.

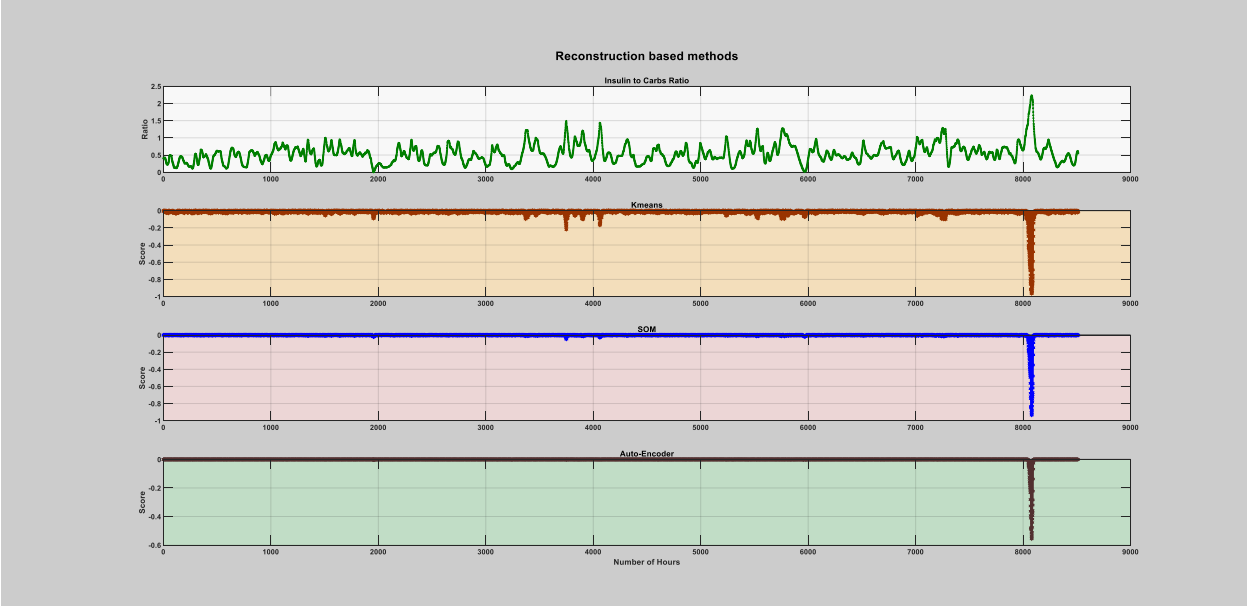


Figure 21: Reconstruction-based method.

### 1.2.4. The Fourth Infection Episode (Flu)

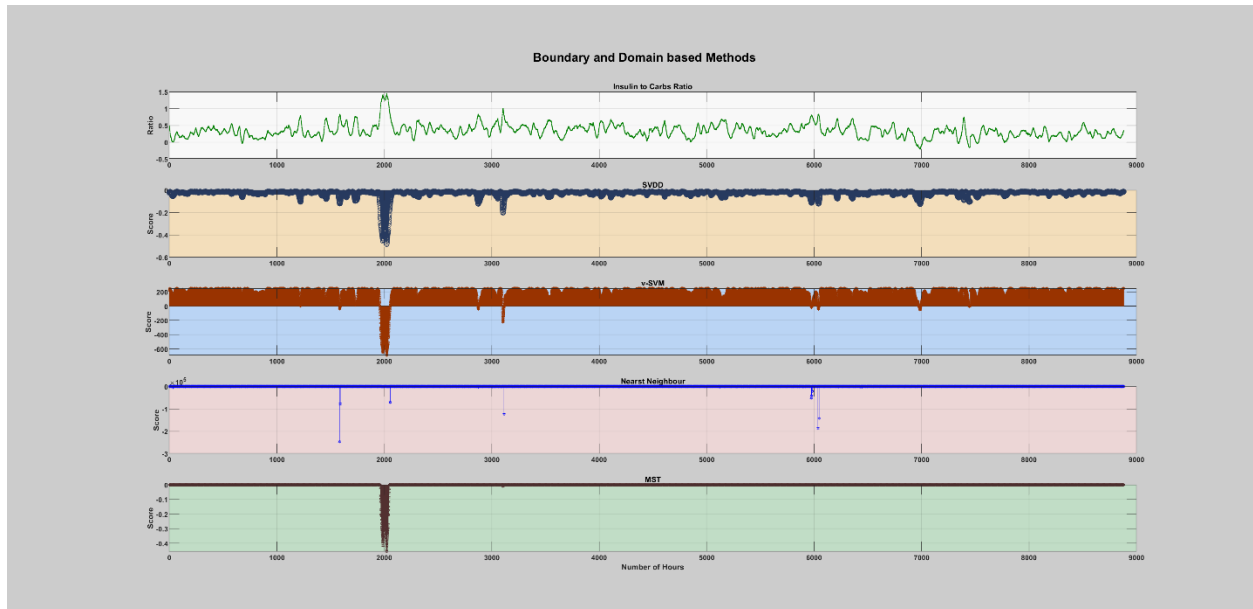


Figure 22: Boundary and domain-based method.

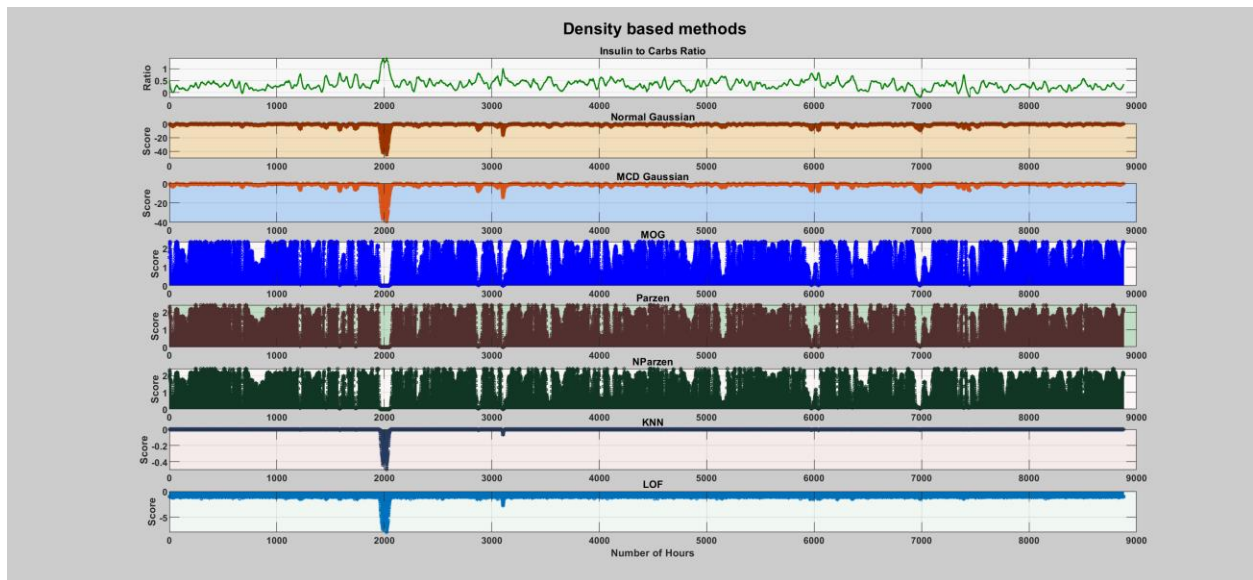


Figure 23: Density-based method.

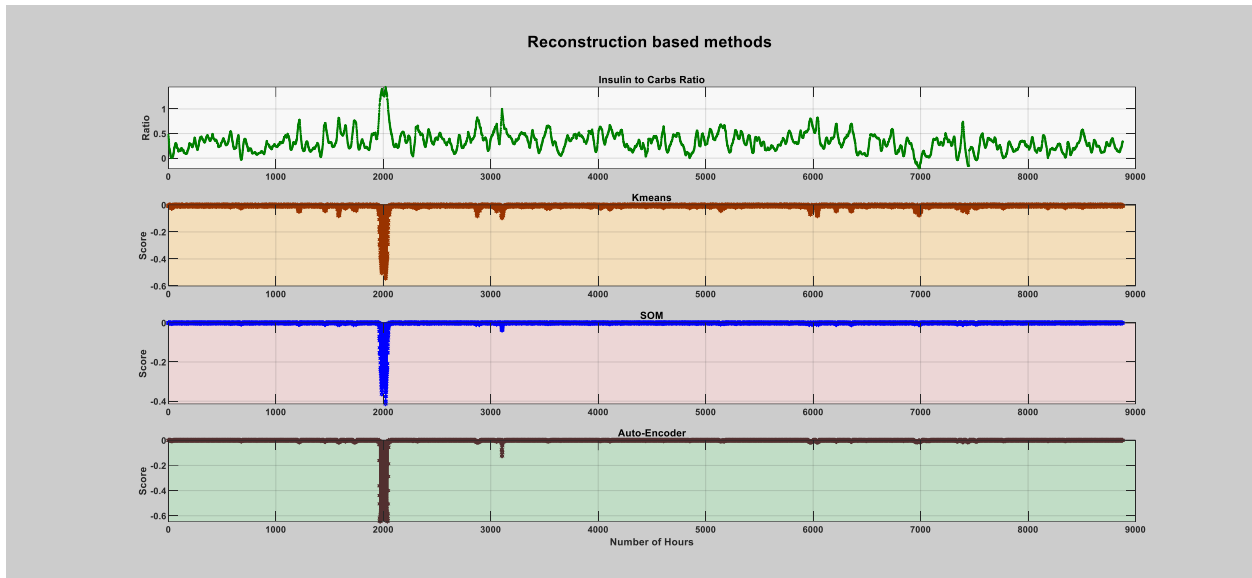


Figure 24: Reconstruction-based method.

## 2. Unsupervised method

### 2.1. Daily

#### 2.1.1. The First Infection Episode (Flu)

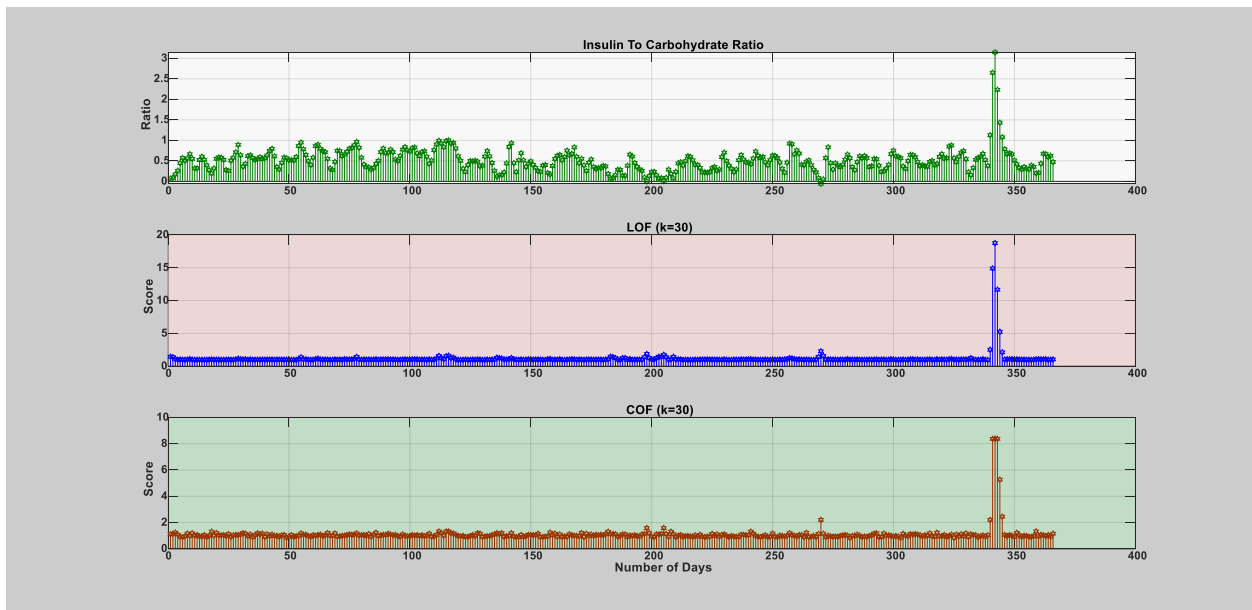


Figure 25: LOF and COF ( $k = 30$  data points).

### 2.1.2. The Second Infection Episode (Flu)

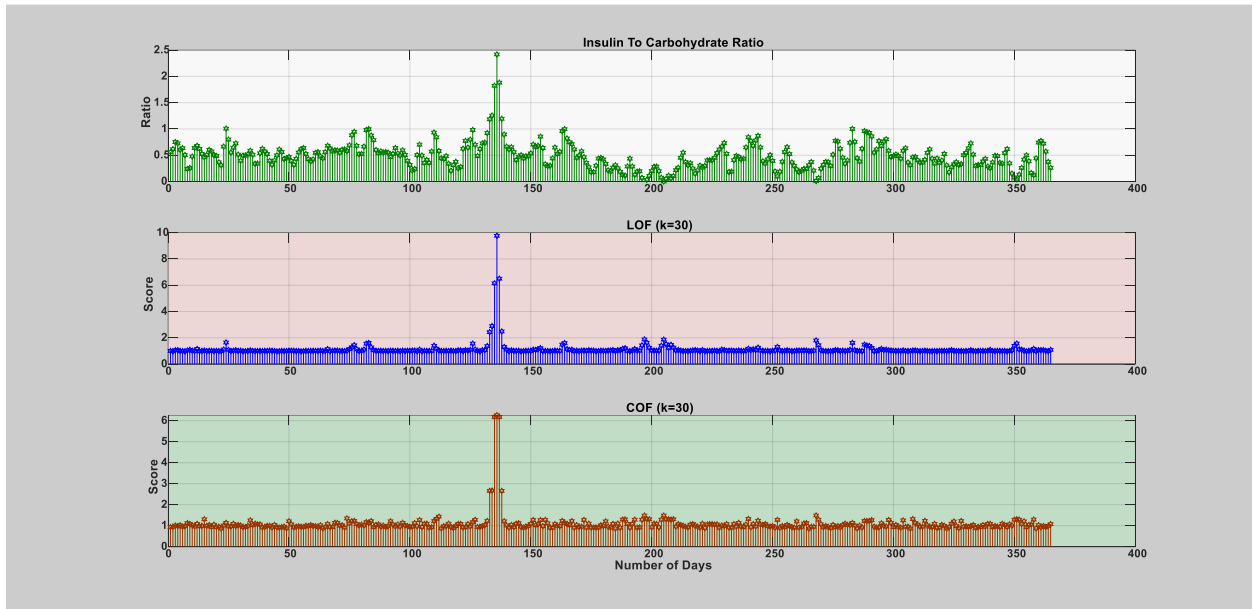


Figure 26: LOF and COF ( $k = 30$  data points).

### 2.1.3. The Third Infection Episode (Flu)

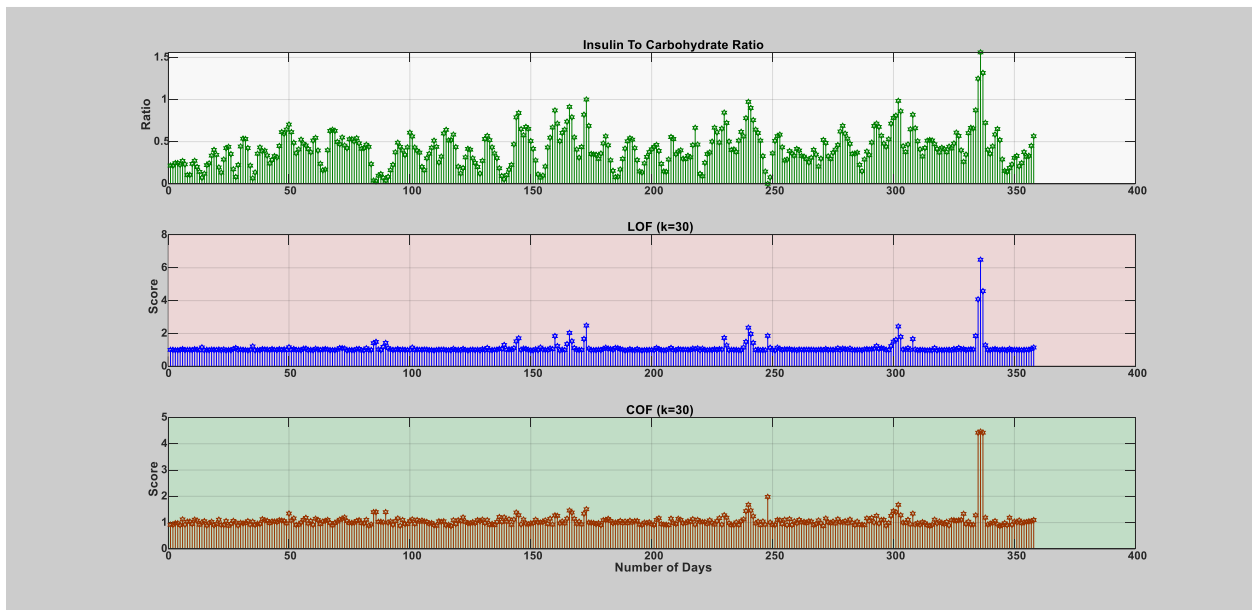


Figure 27: LOF and COF ( $k = 30$  data points).

### 2.1.4. The Fourth Infection Episode (Flu)

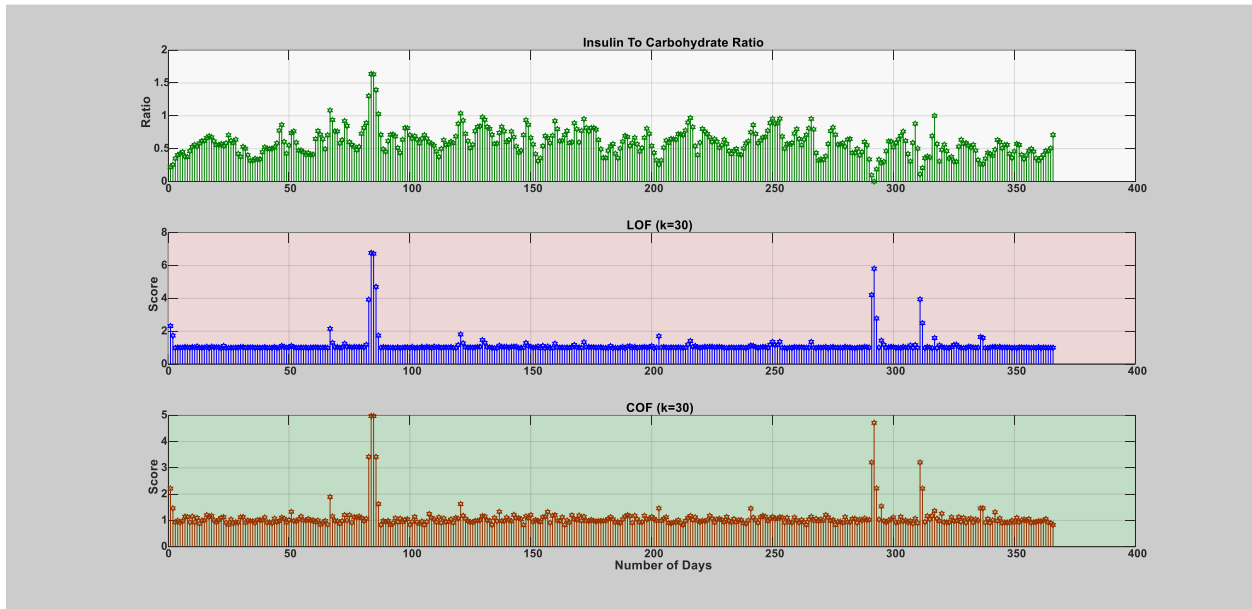


Figure 28: LOF and COF (k = 30 data points).

## 2.2. Hourly

### 2.2.1. The First Infection Episode (Flu)

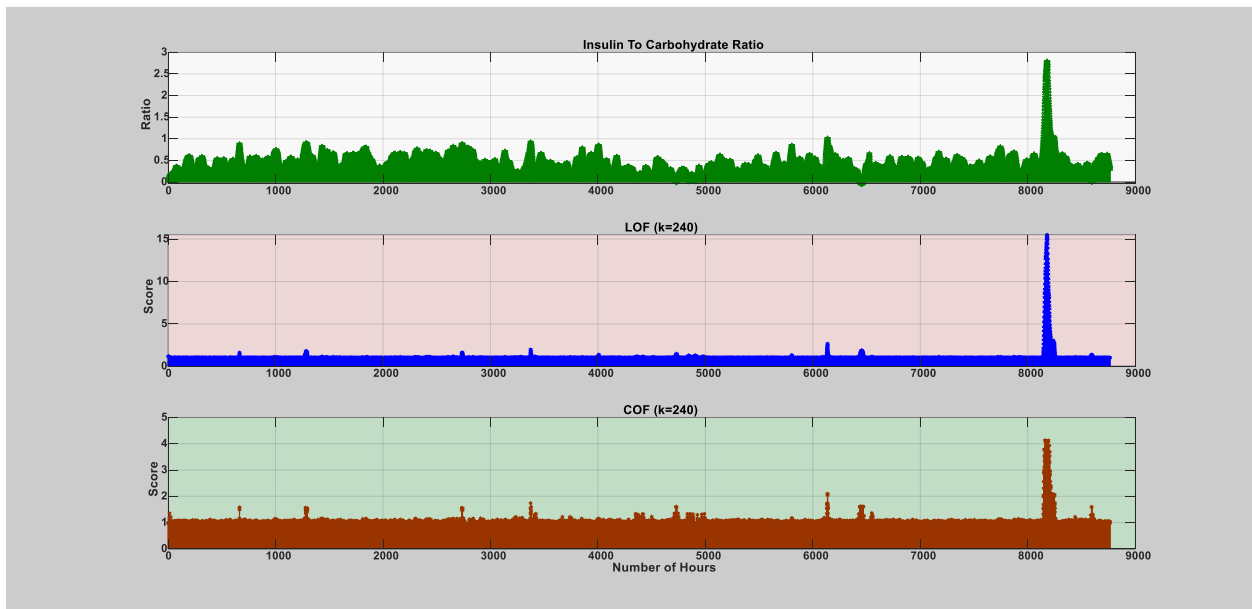


Figure 29: LOF and COF (k = 240 data points).

### 2.2.2. The Second Infection Episode (Flu)

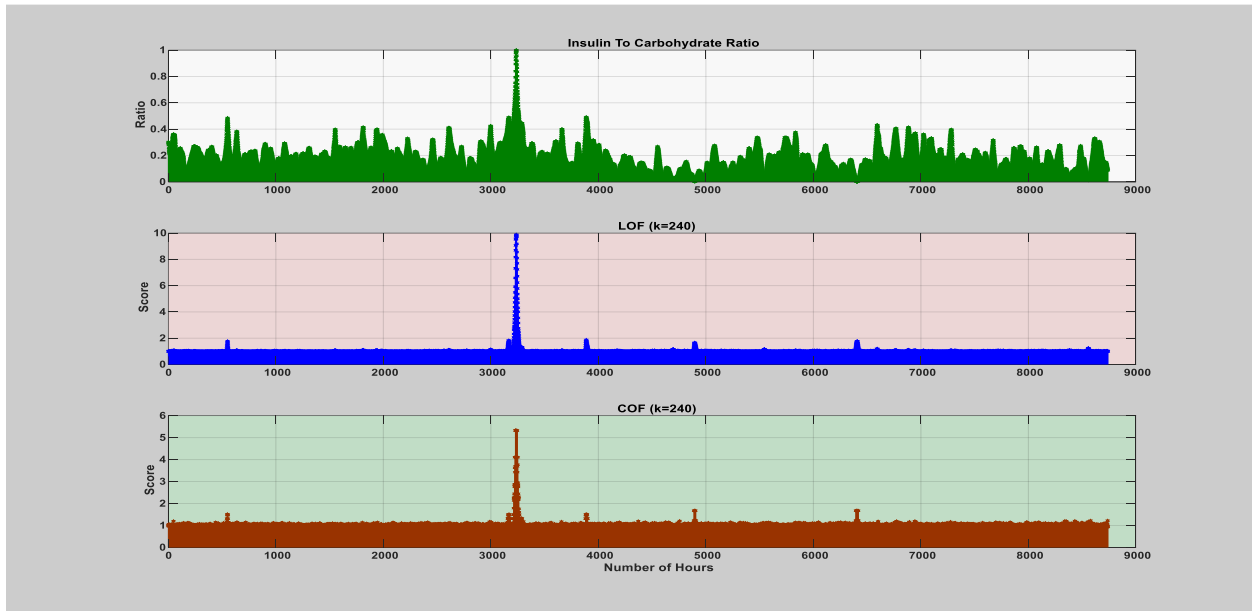


Figure 30: LOF and COF(k = 240 data points).

### 2.2.3. The Third Infection Episode (Flu)

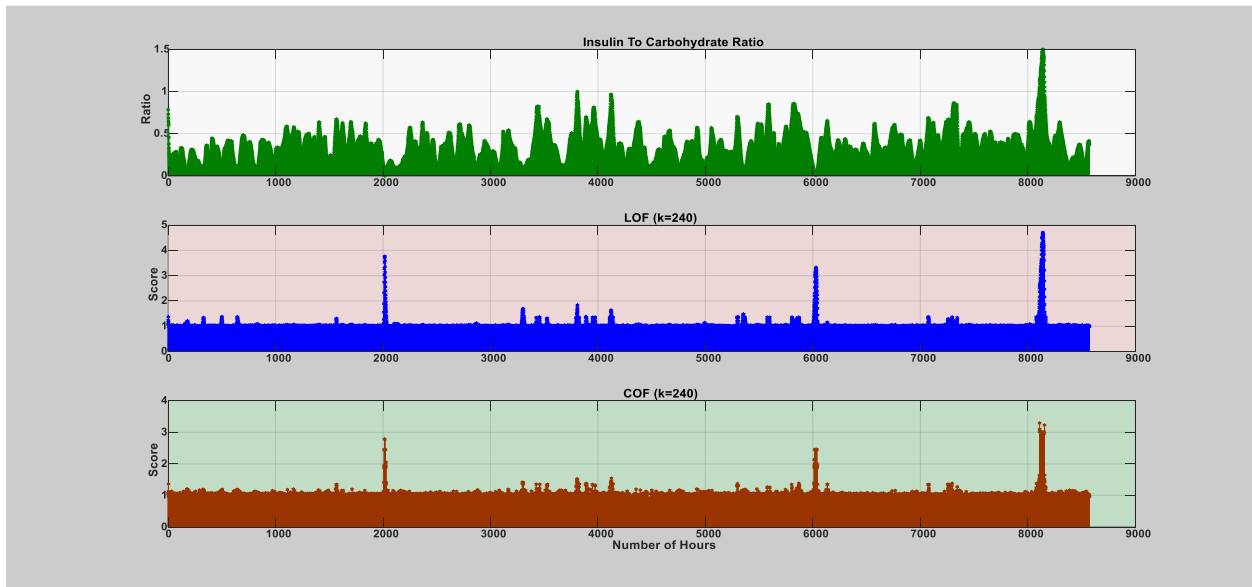
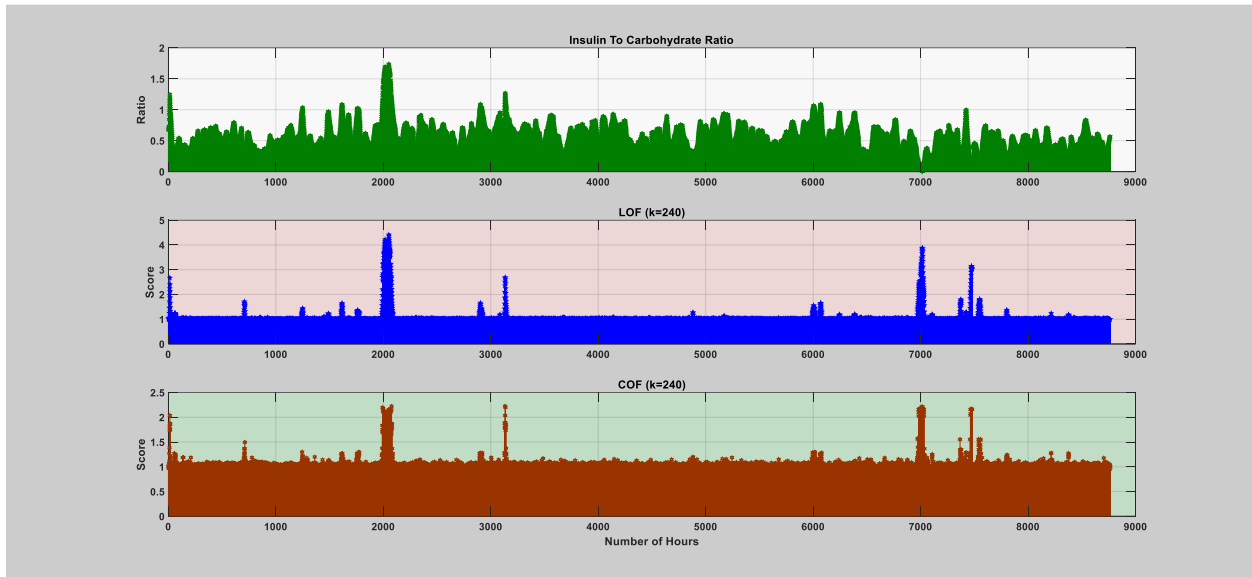


Figure 31: LOF and COF(k = 240 data points).

#### 2.2.4. The Fourth Infection Episode (Flu)



**Figure 32:** LOF and COF( $k = 240$  data points).

## Paper 5

**Woldaregay, A. Z.,** Henriksen, A., Issom, D. Z., Pfuhl, G., Sato, K., Richard, A., Lovis, C.,  
Arsand, E., Rochat, J., & Hartvigsen, G. (2020). *User Expectations and Willingness to  
Share Self-Collected Health Data*. *Stud Health Technol Inform*, 270, 894-898.  
doi:10.3233/SHTI200290



# User Expectations and Willingness to Share Self-Collected Health Data

Ashenafi Zebene WOLDAREGAY <sup>a,\*</sup>, André HENRIKSEN <sup>a,1,\*</sup>, David-Zacharie ISSOM <sup>b</sup>, Gerit PFUHL <sup>a</sup>, Keiichi SATO <sup>c</sup>, Aude RICHARD <sup>b</sup>, Christian LOVIS <sup>b</sup>, Eirik ÅRSAND <sup>d</sup>, Jessica ROCHAT <sup>b</sup>, and Gunnar HARTVIGSEN <sup>a</sup>

<sup>a</sup> UiT The Arctic University of Norway, Tromsø, Norway.

<sup>b</sup> University of Geneva, Geneva, Switzerland.

<sup>c</sup> Illinois Institute of Technology, Illinois, USA.

<sup>d</sup> Norwegian Centre for E-health Research, Tromsø, Norway.

**Abstract.** The rapid improvement in mobile health technologies revolutionized what and how people can self-record and manage data. This massive amount of information accumulated by these technologies has potentially many applications beyond personal need, i.e. for public health. A challenge with collecting this data is to motivate people to share this data for the benefit of all. The purpose of this study is to survey and examine factors that may motivate sharing this data. We asked 447 participants four questions related to health data sharing and motivation. Participants with a chronic disease were concerned about data sharing but also willing to share health data if personalized feedback is provided. Functionality, ease of use, and privacy are regarded as crucial features of health apps.

**Keywords.** Public health, mHealth, motivation, health surveys, health data sharing.

## 1. Introduction

The ubiquitous nature of smartphones, wearables, and sensors have revolutionized the way people collect health-related data. An increasing number of people collects large amounts of data for disease-management, fitness, and self-surveillance. In a recent study, more than 60% of American participants tracked various health parameters, including diet, weight, and physical exercise, where 21% relied on fitness tracker technology [1]. People with chronic diseases use different technologies, e.g. mHealth apps and continuous glucose monitors for diabetes management, collecting and processing health data for their self-management. This data can potentially be used as a secondary source of information for public health, including tracking of disease trends, behavioural patterns over time, chronic diseases status, research, and policy work [2, 3, 4, 5]. The impact of these self-collected data highly depends on people's willingness to share their data for the intended purpose. Considering the potential of these data to inform about individual and population health, understanding the users' expectations and willingness towards mass data sharing is an important area of research. Various factors could affect people's motivations to engage in mass data sharing, e.g. lack of trust, which is mainly

<sup>1</sup> Corresponding author, E-mail: [andre.henriksen@uit.no](mailto:andre.henriksen@uit.no) ; \* Authors contributed equally

subject to data security, privacy, and confidentiality issues [6]. When sharing such data, expectation and willingness often differ in different patient groups, as well as in healthy people [7, 8], e.g. people with type 1 diabetes have a high need to record health data several times per day. Here, we examine factors related to people's knowledge and expectations toward raising motivation for sharing of health-related data and comparing these factors across different groups.

## 2. Method

We created an online survey with questions related to motivational factors around mHealth apps and data sharing. Questions were derived from 16 in-person interviews [6]. The survey was conducted among English-speaking internet users in a Swiss cohort of healthy people and also in English and Norwegian online diabetes groups. We collected data between 11/2018 and 08/2019. The questions related to data sharing are: 1) How concerned are participants about sharing health data, 2) what do participants expect in return when sharing health data, 3) which data types are participants willing to share anonymously for research, and 4) how important are different criteria in order to agree to install an application that collects and shares data from their wearable device. Options were on a 4-point Likert scale from "not important at all" to "very important" (including "I don't know"), or multiple-choice. We stratified responses into three groups: 1) People with diabetes (PWD), 2) people with other chronic disease and 3) people without a chronic disease. Details about the questionnaire are available at [DataverseNO](#) [9]. We report descriptive statistics on age, gender, experience using wearable devices and mobile health apps and wearables for sharing health or activity data.

## 3. Results

Four hundred forty-seven (447) participants finished the survey, of which eight did not answer whether they have a disease, and nine selected "Do not want to answer". Further analyses are based on the remaining  $N=430$ . Sixty-one (61) participants had diabetes, 82 participants had another chronic disease, and 285 had no chronic disease. The majority of participants came from Switzerland (187), Norway (59), US/UK/Australia/Canada (77), France (26), and Germany (13). Remaining 46 came from 35 countries covering all continents. Table 1 gives their demographics, familiarity with mHealth apps, and sharing experience. There was no age ( $p=.083$ ) or gender ( $p=.133$ ) difference between the groups. However, 97% in the diabetes group use a wearable device for collecting activity or other health data, compared to only 51% in the "no disease" group, and 55% in the "other chronic disease" group,  $\chi^2(423)=44.04$ ,  $p=.001$ . Many PWD has experience in sharing data, less so people without a chronic disease,  $\chi^2=19.6$ ,  $p < .001$ .

**Table 1.** Demographics and familiarity with mHealth apps. Na= no answer.

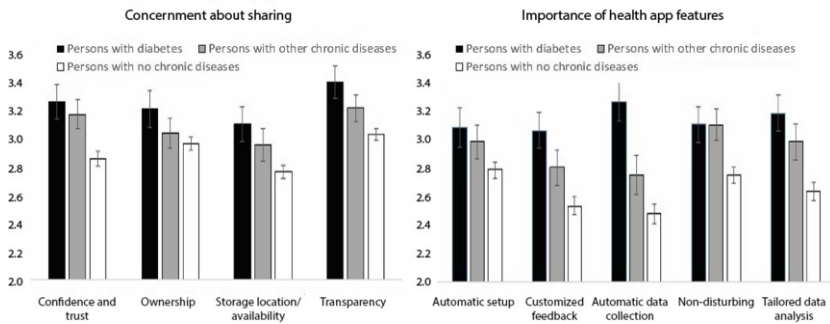
	With diabetes	With other chronic diseases	Without chronic disease
Age: <30 y; 30-50y; >50y	15; 17; 27 (Na: 2)	9; 32; 40 (Na: 1)	34; 96; 146 (Na: 6)
Gender: female; male; other	35; 25; 1	59; 20; 3	177; 99; 3
Wearable device: yes; no	59; 2	44; 36	143; 137
Sharing experience: yes; no	24; 27 (Na: 10)	22; 45 (Na: 14)	47; 160; (Na: 11)

### 3.1. Technologies use for health tracking

87% of those with diabetes use a health-specific device, in addition to often using a Physical Activity (PA) tracker. 3% use no sensor or wearable device. Among those with other- or no chronic diseases, the use of sensors integrated in the smartphone and PA trackers are most common. 25-30% use mobile health apps. Generally, participants preferred to discuss health issues with health providers. PWD would discuss it with others PWD (17%), but rarely with their family (6%). This is in contrast to persons with other- or no chronic disease, where 19% and 20%, respectively, would discuss health issues with their family and friends, see Figure 2b. This difference with whom to share was statistically significant,  $\chi^2=34.67$ ,  $p<.001$ .

### 3.2. Concerns about data sharing

Regarding what people are most concerned about sharing health data, persons with no chronic disease are in general least concerned, and all three groups rate storage as least and transparency as most concerning. Figure 1a shows how each group rate concerns about confidence and trust, data ownership, storage location/availability, and transparency of third party usage. Figure 1b shows how important certain features are for each group in order to agree to install an application that collects and share health data from their wearable device.



**Figure 1.** a: Concerns about different aspects of data sharing. b: Importance of features, answer option 1 to 4.

### 3.3. Participants expectations for sharing health data

Regarding what participants expected in return for sharing their health data, personalized feedback was chosen by 60% of participants, integrated view (i.e. aggregated results), was chosen by 53%, decision support by 36%, and least chosen (16%) was comparing

status with others. There was no difference between the groups in their expectations, smallest  $p > .74$ , but comparing with others was rated as least important (Figure 2a).

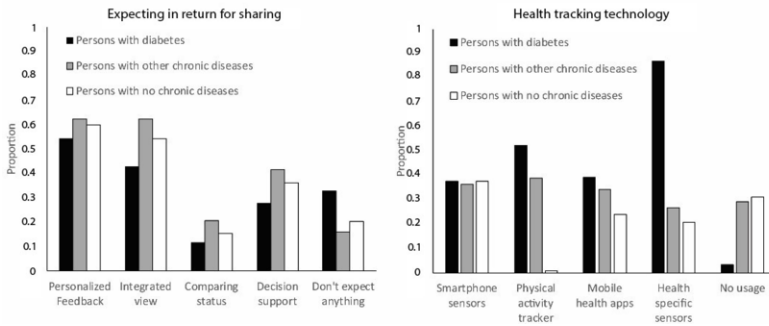


Figure 2. a: Expectations for sharing data. b: Group distribution of health tracking technology.

### 3.4. Willingness to anonymously share given data types in a research project

Groups do not differ in sharing lifestyle/dietary information, signs of infection, daily mood, geographical location, sleep duration, social environment (corrected for multiple comparisons). Participants are least willing to share their geographical location. Persons with chronic disease are more willing to share medication intake, physiological indicators and their weight. An overview is given in Table 2.

Table 2. Overview of willingness to share specific health data by group. Chr. = chronic.

	Medication intake and treatment	Lifestyle/dietary intakes	Signs of infection	Physiological indicators	Daily mood	Physical activity	Geographical location	Sleep duration	Weight	Social environment	None of these
N	213	219	195	202	193	248	87	240	221	173	89
Diabetes	53	41	31	46	29	43	20	43	43	30	5
Other chr.	44	45	45	40	46	54	16	50	49	35	16
No chr.	116	133	119	116	118	151	51	147	129	108	68
X <sup>2</sup>	43.48	9.05	5.22	24.39	5.73	8.93	6.92	8.29	15.48	2.87	7.59
p-value	<.001	.011	.074	<.001	.057	.012	.031	.016	<.001	.238	.023

### 3.5. Features' importance in mHealth app

Regarding which features of a health-app is important to participants (1=not important to 4=very important), non-disturbed tracking and automatic setup is rated by most as important, i.e. the main effect of feature,  $F(3.83, 1163.75)=3.389$ ,  $p=.01$ ,  $\eta^2=.011$ . All features were more important to people with a chronic disease,  $F(2, 304)=12.09$ ,  $p<.001$ ,  $\eta^2=.074$ . PWD rate automatic data collection as most important, yielding a significant interaction effect,  $F(7.66, 1163.75)=2.104$ ,  $p=.035$ ,  $\eta^2=.014$ .

## 4. Discussion

Mass sharing of health data could provide vital information for individual health management and public health. Continuous collection of quality health-related data and willingness to share these are limited by the user's motivation and expectations. The rate of acceptance of health tracking devices among clinicians is increasing [10], but

retention is decreasing [11]. Compared to PWD who use wearable sensors, participants with no- and other chronic disease reported higher adoption of sensors integrated in the smartphone and PA trackers. Automatic data collection, easier interface, e.g. voice command, tailored and personalized feedback is likely to increase usage and long-term engagement in such devices.

Concern about health data sharing is dependent on the type of data, and related concerns on issues such as privacy, security, confidentiality, transparency, and ownership. Participants with no chronic disease are less concerned compared to people with a chronic disease. Indeed, potential consequences of data leakage like repeated low blood glucose level (hypoglycemia) might result in the suspension of one's driving license or disqualification of health insurance enrollment in some countries. Some privacy and security shall be kept, i.e. many participants do not want to share their geographical location.

This study on how to motivate health data sharing is a collaboration between several projects where different systems for health data collection are under development. Results from the present study, and upcoming publications on related topics, will be used to direct the implementations of these systems for maximum acceptance. Future works include data collection, data quality and accuracy analysis, and detecting health patterns at the population level and in people with specific chronic diseases. Fulfilment of participant's expectations and resolving individual concerns could motivate sharing health-related data. Results indicate that participants expect some kind of immediate benefits from sharing their data, including tailored and personalized data analysis, integrated view, feedback and others. Comparison of status among peers was found to be less relevant.

## References

- [1] J. Vitak, Y. Liao, P. Kumar, M. Zimmer, K. Kritikos (eds.), *Privacy Attitudes and Data Valuation Among Fitness Tracker Users*, Cham: Springer International Publishing, 2018.
- [2] J. Chen, A. Bauman, M. Allman-Farinelli, A Study to Determine the Most Popular Lifestyle Smartphone Applications and Willingness of the Public to Share Their Personal Data for Health Research, *Telemed J E Health* **22(8)** (2016), 655-65.
- [3] K.T. Pickard, M. Swan, Big Desire to Share Big Health Data: A Shift in Consumer Attitudes toward Personal Health Information, *2014 AAAI Spring Symposium Series*, 2014.
- [4] K. Ostherr, S. Borodina, R.C. Bracken, C. Lotterman, E. Storer, B. Williams, Trust and privacy in the context of user-generated health data, *Big Data & Society* **4(1)** (2017).
- [5] A. Seifert, M. Christen, M. Martin, Willingness of Older Adults to Share Mobile Health Data with Researchers, *GeroPsych* **31(1)** (2018), 41-9.
- [6] A.Z. Woldaregay, D.Z. Issom, A. Henriksen, H. Marttila, M. Mikalsen, et al., Motivational Factors for User Engagement with mHealth Apps, *Studies in health technology and informatics* **249** (2018), 151-7.
- [7] K.K. Kim, J.G. Joseph, L. Ohno-Machado, Comparison of consumers' views on electronic data sharing for healthcare and research, *J Am Med Inform Assoc.* **22(4)** (2015), 821-30.
- [8] D.J. Willison, V. Steeves, C. Charles, L. Schwartz, J. Ranford, et al., Consent for use of personal information for health research: do people with potentially stigmatizing health conditions and the general public differ in their opinions?, *BMC Med Ethics* **10** (2009), 10.
- [9] A. Henriksen, A.Z. Woldaregay, D-Z. Issom, G. Pfuhl, A. Richard, et al., *Questionnaire for motivation in mobile health*, DataverseNO V1, 2019. <https://doi.org/10.18710/28SRMJ>
- [10] M. Bradway, L. Ribu, G. Hartvigsen, Årsand E, The evolution of clinicians' preparedness for mHealth use (2013-2017) and current barriers, *Proceedings from the 16th Scandinavian Conference on Health Informatics 2018*, Linköping University Electronic Press, Denmark, 2018.
- [11] C. Attig, T. Franke, Abandonment of personal quantification: A review and empirical study investigating reasons for wearable activity tracking attrition, *Computers in Human Behavior* **102** (2020), 223-37.

