UNIVERSITÀ POLITECNICA DELLE MARCHE

Facoltà di Ingegneria
Dipartimento di Ingegneria dell'Informazione

Study and analysis of insulin and glycaemic data for the development of user calibration optimisation of t:slim x2 insulin pump in children and adolescents

Author:
Ismaela AVELLINO

Supervisor:
Dr. Lorenzo PALMA

Co-supervisor:
Dr. Luisiana SABBATINI

Co-supervisor:
Dr. Valentino CHERUBINI

Academic year: 2020-2021

Here's to the fools who dream,

running away from any perverse drive towards conformity.

Contents

List of Figures

List of Tables

List of Equations

Chapter 1

Introduction

Diabetes mellitus is a metabolic disorder linked to total insulin deficiency, leading to carbohydrate and lipid metabolism's anomalies [1]. It is also known as juvenile as it usually attacks children and young adults.

It requires life-long monitoring and, even if insulin analogues, pump systems and vanguardist glucose monitoring devices have improved the efficiency of the control over this illness, current therapies cannot repeat the healthy-state feedback control that allows minute-to-minute regulations [2]. That means that full metabolic normalisation is not possible yet, which can cause short- and long-term complications. The latter can lead to a plethora of issues, both physical and psychological.

These are the main reasons why researchers are deeply interested in finding new creative solutions that can imitate a healthy pancreas.

1.1 Diabetes mellitus from a clinical point of view

Diabetes mellitus (coming from the Greek $\delta\iota\alpha\beta\eta\tau\eta\zeta$, translatable as 'passing through', coming from the presence of glucose in patients urine, which is, thus, as made 'of honey') is a multi-factorial and chronic metabolic disease that, according to a study by Cho et al. [3], afflicted 8,4% of the world population belonging to the age range of 18-99, in 2017. This value could reach 9.9% within 2045.

It is a pathology characterised by a glucose concentration increase in the blood (hyperglycemia) due to a struggle in production and proper insulin function. The latter is secreted by the pancreas and is fundamentally important in simple and complex sugars metabolism. After being assumed through eating, they represent the very first source of energy for the human body.

The alimentary tract is linked to metabolism [4] and is ruled by a complex system of tissues, organs and hormones. The latter are needed to deposit excess nutrients and release them in a fasting or energy demand situation. From this point of view, glucose metabolism is fundamental: the two main sources of plasma glucose are food and, consequently, the liver. Carbohydrates assumed during meals get hydrolysed in the intestine into simple sugars that are then absorbed, leading to an increase of postprandial glycaemia that reaches the peak at 90-120 minutes and whose amplitude depends on factors like the quantity of ingested food, the type (simple or complex) of assumed carbohydrates and the presence or lack of fats or fibres that slow the absorbing process. In fact, around 3-4 hours after the meal, the

liver releases 60-70% of glucose absorbed by organs that do not have deposits, while the rest gets employed for the immediate energy demand. Moreover, kidneys freely filter glucose and, in standard concentrations, the latter is totally reabsorbed by the proximal tubules until a quantity of 180 mg/dl is reached; once this threshold exceeds, glucose gets eliminated through micturition. As an answer to this increase in glycaemic values in the blood, the insulin hormone works in two modalities:

1. Basal (or long-acting): it keeps glycaemia low and constant throughout the day and between meals.

2. Postprandial (or short-acting): also known as bolus; it activates through food intake due to glycaemic peaks.

When insufficient production of insulin or when subjects tissues do not react correctly to its action, their metabolism gets altered. That leads to a rise in blood glucose levels that promotes the occurrence of diabetes.

This illness divides into two main types, plus a third one known as gestational that exclusively affects pregnant women and usually disappears after birth [5]:

- Type 1 (also known as *insulin-dependent* or *Juvenile*): it makes up to 10% of all diabetes cases and it usually onsets during childhood or adolescence [5]. If it occurs in an adult subject, it is *Latent Autoimmune Diabetes in Adults* (LADA). It is caused by a suppressed or highly reduced insulin production by the β cells of the pancreatic islets of Langerhans. This faulty production is due to the destruction, whose velocity varies from case to case, of these cells from the immune system of the patient. Thus, the lifelong injection of the hormone is necessary. Despite the remarkable and rapid evolution of scientific and laboratory knowledge and expertise, the triggering events are still unknown [6] even if some contributing factors [7] (genetic, environmental and immunologic) have been pinned down; in diabetics' blood, there are antibodies directed against antigens that are at the level of β cells, called ICA, GAD, IA-2 and IA-2β and the illness is classified as *autoimmune*, thus linked to the immune system's reaction against the organism, triggered by a combination of genetic and environmental factors. Symptoms, along with the possible complications, are various and variegated [6] and the main ones concern abundant and frequent micturition (polyuria), excessive thirst (polydipsia) and hunger, higher frequency of infections and inexplicable weight loss.

- Type 2: the most common one; it usually affects people who are at least 30-40 years old and it mainly targets overweight, obese and sedentary subjects [5]. In this case, insulin resistance plays a synergistic role with reduced secretion of insulin by β cells [7]: this hormone still gets produced, even if in a reduced amount, but the organism cannot efficiently use it. Along with these problems, other pathophysiological abnormalities contribute to the impairment of glucose metabolism regulation [8]. Generally, the onset of Type 2 diabetes mellitus can go undetected for many years as hyperglycaemia develops gradually and involves less evident symptoms compared to those of Type 1 diabetes mellitus. These symptoms are fatigue (asthenia), blurry vision and slow healing of wounds. Usually, the diagnosis happens casually or concomitantly to a situation of physical stress like infections and surgeries [5].

- Gestational: it can have acute adverse effects on fetal and neonatal outcomes [9], along with health issues in said offspring, including, but not limited to, long-term obesity

and glucose intolerance. Only around 10% of diagnosed subjects still have gestational diabetes shortly after delivering [10].

This thesis will focus on Type 1 Diabetes Mellitus (DMT1).

1.2 The diagnosis

The early diagnosis of the disease allows the clinicians to promptly intervene, reducing the possibility of future complications' occurrence. In fact, the more diabetes is left untreated, the worst will be the health of the patient [5].

Subjects that are then diagnosed with T1DM can present themselves with increased thirst and urination, fatigue, lack of energy, inexplicable bacterial and fungal infections, delayed wound healing, blurred vision and hands or foot's numbness or tingling.

They can be affected by modest hyperglycaemia that can evolve into a severe one or infection- or stress-related ketoacidosis (which could also lead to coma). So, as soon as the above-mentioned symptoms start, a diagnosis must be requested and pursued [11].

The diagnostic criteria are based on the following measurements [12]:

I. Fasting glycaemia (at least 8 hours) $\geq$ 200 mg/dl,

II. Random glycaemia (independent from the moment of the day) $\geq$ 126 mg/dl,

III. Glycaemia during load curve (OGTT), with an administration of 75g of glucose, $\geq$ 200 mg/dl.

1.3 The therapy

Diabetic patients need a systematic and continue therapy planned by a professional medical team [13]. The therapy's [4] main objective is to maintain glycaemia inside a range of standard values and its secondary objective is to implement an adequate lifestyle made up of correct food habits and physical activity. So, at the base of diabetes' care are therapeutic education, diet, sport and medications. It is also fundamental to abstain from smoking as it is a risk factor for the chronic complications of diabetes, but it also makes it complex to control glycaemia.

1.3.1 The pharmacological therapy

The main drug, as already mentioned, for a type 1 diabetic is insulin. Its dosage changes from person to person based on their daily activities, but the patient can never go without it. There are different kinds of insulin based on their function [14]:

- Basal and postprandial: while the first one, administered once every 24 hours, has a long action period and is used to maintain the glycaemia between one meal and another as constant as possible, the second one, which is also known as *short-acting insulin*, because it is employed in response to food intake, through an injection that comes before, during or after a meal to mimic the hormonal postprandial response.

- Premixed: it unites, in prefixed ratios, basal and postprandial insulin and is used by subjects that refuse to do more than two injections in one day, but it is not flexible, thus it is not recommended.

- Concentrated: it is a special type of basal insulin that is concentrated 2 to 5 times more than the common one and is used on patients that have high insulin resistance or that are particularly overweight, demanding more than 200 units per day.

The therapeutic education consists of the teaching of knowledge, that goes beyond the mere insulin dosage, that is helpful in the management of every aspect of the illness in the most efficient way possible. It is taught by doctors and nurses, along with other professional figures like dietitians and psychologists, to the patients and their families.

The training of young diabetic patients and the development of their awareness and know-how is pivotal for their adherence to treatment regimen [15]. So, improving this education is the necessary key to boost their constancy as, despite the wide availability of efficient therapies, adolescents affected by T1DM still have poorer adherence compared to other paediatric age groups [16] and non-adherence is directly bound to below-the-standard glycaemic control and increased morbidity and risk of premature death.

1.3.2 The non-pharmacological therapy

Non-pharmacological treatments are necessary for efficient monitoring of every aspect of the illness and, mainly, for the prevention of possible complications that could worsen the general health of the patient.

The most direct benefits of an optimal food habit are, of course, more notable for Type 2 diabetes mellitus patients, as the main causes of the onset of this illness are traceable to a sedentary lifestyle and an incorrect diet that leads to obesity. Nevertheless, even for those affected by T1DM, nutrition is an integral part of the illness monitoring, but also its therapy.

For many years, the diet of the patient has been limited, demonising certain foods, especially those rich in carbohydrates; today, fortunately, the consumable food is no longer imposed and the diabetic subject can freely eat, just paying, of course, attention to the quantities and types of carbohydrates, proteins and fats for the correct choice of insulin dosage and by trying to follow a reasonably healthy diet, just as it is recommended to the general population. Obviously, uncontrolled or non-appropriate feeding, like the one coming from a fast-food diet, can contribute to the rise in body weight, which is a serious risk factor that must be avoided.

Various studies have demonstrated that our Mediterranean diet [17] [18], for example, represents a reliable choice to get to a healthy lifestyle. It is based on a large intake of vegetables, fruits, legumes, cereals and seeds, as well as olive oil as the primary source of fats. Moreover, the consumption of fish, dairy products and white meats is limited. Thus, this kind of diet has a dramatically positive influence on the body weight, blood pressure and lipids blood concentration control, without being excessively restrictive and representing an easily accessible and sustainable option.

Physical activity [19] is obviously necessary for the diabetic patient as it increases the insulin response of the organism, increases glucose consumption by the muscles and contributes to the improvement of the overall health by increasing HDL (the so-called *good cholesterol*) and reducing the arterial pressure.

Sport must be regulated along with the insulin dosage and the meals to avoid metabolic disorders, hypo- and hyperglycaemias [4] and should never be enforced. In the diabetic subject, moderate physical activity causes:

I. in the first 5-10 minutes, muscles use the glucose that is deposited inside of them

II. if sport prolongs over 20-30 minutes, the muscle glycogen runs out and muscles must start to use the glucose injected in the blood by the liver

III. if the activity further prolongs, the muscle uses more and more fats coming from the adipose tissue as the source of energy.

During physical activity, the hepatic work gets stimulated and the counter-regulatory hormones, like adrenaline and cortisol (stress hormones that stimulate the body's insensitivity to hypoglycaemia) production is raised. After sport, the passage of glucose in the blood, muscle (and their sensitivity to insulin action) and liver to reconstruct the glycogen deposits raises.

Finally, one must focus on any inter-current disease, including the flu, and any kind of stress that could rise glycaemia and/or worsen the metabolic compensation.

1.4 The glyceamic monitoring and subsequent treatment

Glucose must be regularly monitored through tests carried out by the same patient and the analysis of glycated haemoglobin to prevent any occurrence or worsening of complications. The glycaemic monitoring is composed of various steps: the computation of plasma glucose through glucometer sticks, but also the interpretation of the output result and the subsequent therapeutic intervention, like the insulin injection, as a response. Self-monitoring is necessary to reach an adequate metabolic control and to reduce the risks of hypo- and hyperglycaemias: in fact, based on these checks, patients adjust their meals, daily physical activity and insulin dosage.

Among the most advanced technologies, today the interstitial constant glucose monitoring systems (CGMs) stand out. They represent the modern alternative to the standard glucometer. They are composed by:

- a glycaemic sensor usually implanted in the most adipose subcutaneous area of the abdomen or the arms,
- a monitor that allows to read and show the values of glucose in the interstitial fluid in real-time (CGM real-time) or postponed (CGM offline),
- a transmitter that allows the communication between the first two parts (through a wire or Bluetooth® technology), sending data on the plasma glucose concentration[1].

These devices must get calibrated using glucose measurements in the capillaries: the sensor performs measurements every 5/10 seconds and computes an average value every 5 minutes that it then sends to the monitor. In the past, the accuracy of this type of sensor

[1] `https://www.dexcom.com/g6-cgm-system`

was a major problem as it had to be calibrated various times by the patient through standard glycaemic controls.

CGMs can join the "open-loop" or "closed-loop" systems:

- Open-loop: coupled with multiple daily insulin injections or a subcutaneous insulin pump with the patient's input.

- Closed-loop: coupled to a subcutaneous insulin pump for the completely autonomous supply of basal and postprandial insulin, without human intervention.

The second system is also defined as the artificial pancreas or known as continuous subcutaneous insulin infusion (CSII) system; despite its name, it does not replace the natural organ inside the patient's body, but it mimics its action from the exterior. Different kinds exist, each with a different self-management level:

- Insulin infusion suspension due to hypoglycaemia: instead of sending an alarm to the patients and/or their families, it automatically suspends insulin any time the glycaemic level lowers under a determined threshold.

- Predictive insulin infusion suspension due to hypoglycaemia: it predicts the glycaemic lowering before it happens by studying the trend of blood glucose levels.

- Hypo- and hyperglycaemias' minimizer: it acts as a double control and works inside a range of values under which hypoglycaemia would occur and above which hyperglycaemia would occur. This system blocks the infusion before getting outside the first one of these two values of threshold and, on the contrary, it activates it before reaching the second one.

- Hybrid: it performs very well during the night, but not as much during the day. The basal control is done by the pump, but the postprandial injection is in the hands of the patient.

- Double hormone: the insulin pump is able not only to inject insulin but glucagon too. The latter is the main hormone for the response to hypoglycaemia as it stimulates hepatic glycogenolysis (which is to say the release of glucose by the liver) [4].

Until today, no final cure for T1DM has been found; there are only some therapies and the conventional one consists of the injection of insulin to simulate the action of β cells, but succeeding in finding the correct dosage of this hormone, especially for a tender age, can be very complex [13] especially due to the stress directly deriving from the illness, to physical changes that come with puberty and to the consequent hormonal fluctuations that influence the need and response of the body to insulin.

The injection is essential for subjects affected by T1DM and a wrong dosage combined with improper food habits and uncontrolled physical activity could cause severe hypo- and hyperglycaemias. To define the correct insulin dose to be injected, patients are required to manually write down the activities carried out during the day (sport and meal intake) in a glycaemic diary that, possibly, enables to adapt the insulin therapy in case of hypo- or hyperglycaemic events.

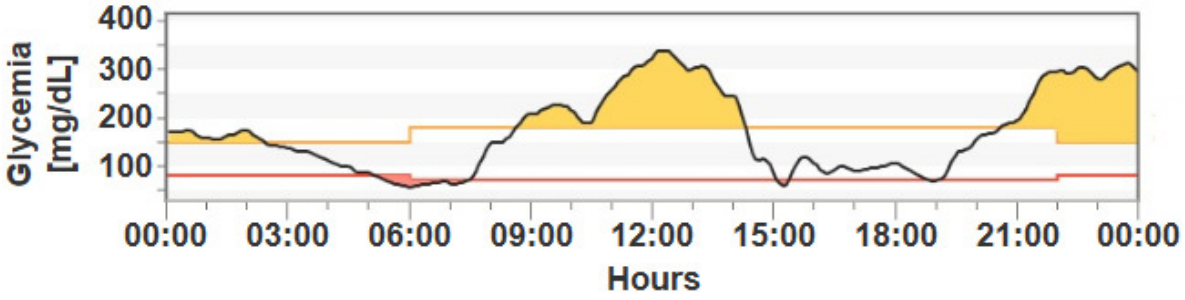


Figure 1.1: Example of a graph of the glycaemic trend recorded by the DexCom G6® sensor in 24 hours.

In fact, glucose curves computed by the DexCom G6® (Fig: 1.1) system help in keeping track of the glycaemic trends, but they are not clearly related to the causes of the peaks, which are, in turn, recorded in the diaries.

The rise of glucose in the blood is mainly caused by food intake, while the drop can be due to physical activity.

1.5 Diabetes' psycho-physical impact on the paediatric patient

The onset of a chronic disease like diabetes can be complex to face, especially at a tender age, and the patients will have to develop control strategies for T1DM and their general psycho-physical well-being. The main obstacle to overcome is the newly acquired independence that occurs during the passage from childhood to adult life: the family and the clinicians' team pass their role onto the subject, who will now have a dramatic influence in the future development of the pathology [20]. That potentially leads to conflicts and refusal of self-management [21].

The control of the disease will change over and over again through the years, requiring complete and continuous attention towards the patients and their issues linked to the illness itself and the rest of their daily lives [22]. Thus, routine check-ups are fundamental for evaluating the state of health of the young adult affected by T1DM.

1.5.1 The physical complications

T1DM urgently requires an improvement of the metabolic control and intensive treatment of neurological and cardiovascular risk factors to be able to substantially reduce the illness' complications [23].

This pathology can lead to the occurrence of acute complications directly deriving from improper monitoring and, consequently, inefficient therapy. The patient could face ketoacidosis coma due to the build-up of ketones that cause loss of consciousness, dehydration and severe haematic alterations. Complications are, on the contrary, chronic when they are caused by the general degradation of the overall state of health of the body and can attack various organs and tissues, like eyes, heart, blood vessels and nerves.

It has also been proved that severe hypoglycemia can impair some medial temporal-mediated cognitive skills, like delayed declarative memory, which is to say the ability to explicitly recall past events. In a study conducted by Hershey et al. [24] on T1DM-affected children, subjects used to drastically low blood glucose levels performed less accurately on a spatial declarative performance task and they were slower at a pattern recognition task, compared to the other diabetic patients not affected by severe hypoglycaemias and to control subjects. Moreover, both groups of diabetic children had significant impairment on a motor speed task when compared to their non-diabetic peers.

The most visible and impacting consequences in the short term are, of course, hypo- and hyperglycaemia, caused by many factors like meals irregularity, wrong insulin dose and alcohol and substance abuse [4]. In particular, hypoglycaemia can, in some cases, lead the diabetic individual to ask for others' assistance, occasionally leading to hospitalisation. A severe hypoglycaemic episode, mainly, but non exclusively, in a frail subject or in one who has concurrent diseases, can prove to be fatal. Hence, it is necessary to train the patients and their relatives to identify this event and promptly correct it.

The following are the most severe and widespread chronic problems [23]:

- Diabetic retinopathy; damages the small vessels that supply blood to the retina. It causes the loss of visual faculties and it is the most peculiar chronic issue caused by diabetes. It can lead to moderate lesions (micro-aneurysms, micro-haemorrhages, exudates) or more severe lesions (maculopathies, retinal ischemias, retinal detachment), with consequences that can also lead to blindness.

- Diabetic nephropathy; consists of the progressive reduction of the liver ability to filter the blood with the increment of creatinine and it can lead to renal impairment and the subsequent need for dialysis and/or the organ's transplant.

- Cardiovascular diseases (vasculopathies and cardiopathies); the risk of these pathologies in the diabetic subject is 2 to 4 times higher than the rest of the population.

- Diabetics neuropathy; it can cause loss of sensitivity, stomach, eyes and/or heart disorders, a pain of variable intensity and damage to the limbs, with the chance of the need for amputation in the most severe cases.

- Stroke and heart attack; T1DM shows up as one of the main causes.

- Diabetic foot; structural alterations of blood vessels and nerves can cause ulcers and issues mainly to the foot due to the loads it supports, making the amputation of the limb necessary. Dramatically, diabetes shows up as the primary cause of non-traumatic amputation. Correct education on feet hygiene is necessary as from apparently insignificant lesions, massive issues can arise. For this reason, even the most trivial of problems must never go underestimated and they must be checked on and cured by experienced clinicians.

In light of these complications, T1DM can be defined as a systemic pathology, meaning that it affects the whole organism. Since there is a chance to avoid or, at least, slow the progression of these chronic issues through strict control over risk factors (mainly attributable to wrong glucose and blood pressure levels), the disorder must never be neglected and overlooked. In fact, in the opposite case, patients could face disability and they could even die.

Prevention requires routine check-ups to evaluate the possibility of damage presence and to, consequently, engage the necessary diagnostic investigations and adequate treatments. Thus, it is fundamental to establish strict monitoring of target organs, such as eyes, liver and lower limbs. That is why people affected by T1DM must undergo regular check-ups, even if symptoms are absent.

Regarding the metabolic changes that the adolescent patient undergoes, it is important to talk about [23]:

- Insulin-resistance increase: it is caused by the rise in the production of the growth hormone (GH) that is typical of puberty, but which is even bigger in the diabetic subject. The persistent anomalies in the ratio between GH and insulin can interfere with the normal growth of the patient and worsen glucose homeostasis. This change in sensitivity is not constant, making the adjustment of the therapy even more complex, as it must be revised and reorganised even multiple times within the same month. GH is characterised by a quantitative pulsatile trend that reaches the peak during the night hours and, thus, it is responsible for the rise of glycaemia during the first hours of the morning: this phenomenon is known as the 'dawn phenomenon'. A modest rise in insulin during these hours could make up for it, but, in diabetic subjects, it can lead to morning hypoglycaemia.

- Reduction of the protein synthesis: their ingestion causes an increase of glycaemia, which does not occur in healthy subjects.

- Potential negative influences on growth and pubertal development in the case of poor metabolic control: even in well-managed patients a reduced pubertal growth spurt can be detected and it is more frequent in subjects who have been diagnosed in pre-school age.

- Increase in ketone bodies and subsequent ketoacidosis: due to insulin deficiency and the subsequent increase of counter-regulatory hormones (which is to say those that act in the opposite way of insulin).

- Menstrual irregularities: they are more frequent in diabetic girls rather than in their healthy peers.

- Weight gain: mainly caused by the rise in daily injected insulin dosage. It also leads to a thickening of the adipose basement membrane.

1.5.2 The psychological complications

Focusing on what paediatric subjects perceive as critical and threatening is necessary when designing therapeutic interventions and medical devices aimed at improving their health. That is the reason why exclusively meeting physical needs is not enough. The self-monitoring of T1DM is intense, constant, complex and surely does not go unnoticed, leading to feelings of stress and shame in a social context.

Among the so-called 'stressors' [25], the first one is school, followed by social life and diabetes. Specifically, the main factors that trigger stress divide into two macro-categories:

- General life: social adaptation, friendships, different chores' management, family life and pressure to succeed.

- Diabetes: being affected by the illness, manage emotions and monitor T1DM.

When at school and during their daily life, subjects find it difficult to keep up with their friends and their blood glucose levels affect their concentration, making them nervous and making them feel as if they cannot control themselves [21].

An adolescent affected by T1DM finds itself living a transition period already usually linked to mental health issues and that is additionally scarred by grudge and dissatisfaction that could lead to several self-harm mechanisms. The problematic behaviours typical of teenagers (linked to rebellion or experimentation) inevitably undermine the diabetic subject's adherence to therapy. Diabetic teenagers are exposed to material risks to develop psychiatric and/or eating disorders, along with substance abuse: this can result in the refusal of therapy with the subsequent degradation of metabolic control [26]. The frequency of eating disorders' cases among teenagers and young adults affected by T1DM is growing compared to their healthy peers [27]. These disorders concern, along with the most known illnesses like anorexia, bulimia nervosa and *binge eating disorder*, even general intrusive thoughts over body weight and food intake that lead to the deliberate lack of meal consumption and omission of insulin injections. The evolution of body image explains the attention that teenagers have over their physical appearance with a predisposition towards these disorders aimed at controlling body weight [25].

Paediatric subjects affected by T1DM often feel like a burden and find it difficult to express their needs related to their pathology.

Alcohol can cause severe hypoglycaemias and is related to a bigger risk of hospitalisation for diabetic ketoacidosis (DKA), while cannabis and other illicit substances can lead to the omission of insulin injection, wrong dosage or the lack of glycaemia monitoring [28] that have negative long-term effects on the latter factor itself [29].

Additionally, in the case of young patients that have been diagnosed at a tender age or of subjects who have had numerous episodes of hypo- and hyperglycaemia, a deficit in cognitive functions has been reported [26]. On the other hand, the benefit of psychological intervention in diabetic patients has been demonstrated [30], leading to optimal results in therapy's adherence and general life quality improvement.

Young patients affected by T1DM also present symptoms of anxiety, especially ones linked to the self-monitoring requirement, and they report that the pathology represents an additional layer of responsibility that requires an excessive amount of time and that limits their ability to integrally participate in all other aspects of their life [31]. In fact, sometimes, they skip or postpone their management of diabetes tasks.

The main cause of this struggling must be tracked down to the fact that adolescents affected by T1DM go through a transition from being completely dependent on their parents to have to self-manage in full autonomy: on one hand, this is a way to establish their freedom and independence, but on the other, it leads the patient to feel fully responsible for any mistake that is made [27].

Furthermore, some adolescents are seriously concerned about diabetes having an adverse impact on their future: they do not only worry about the possibility of suffering from bad health [32], becoming too ill and not getting to grow old but also about T1DM being a possible impediment to travelling, moving to other countries and getting into specific education programs and\or jobs [21]. This fear is grounded in their hyperglycaemia imposing them and their families to change their plans. Therefore, the illness is seen as a dreadful disturbance. That is also part of their worries about their own family's future life: they worry about the heredity of diabetes but also about the illness impact on their chances of finding a partner.

Some patients also express the fear that their diabetes principally affects their mothers (for the majority of paediatric subjects, their mother is the primary caretaker regarding diabetes management and conflict coming from the discrepancies in the perception of decision-making can arise [33]) and steals attention from their siblings [21].

There also are some psychological issues directly correlated to the use of specific diabetes equipment. Since subjects do not want to draw any attention to their illness, carrying it around in a non-discrete way becomes a crucial problem. Especially when using an insulin pump (which is a topic that will be fully explained later in this work), they feel as if they cannot wear the clothes they want, they get bruises and sometimes it detaches if they accidentally hit a piece of furniture or bump into another person and they cannot carry it with them while practising sports as they are cumbersome and could also irritate their skin during physical activity [21]. Regarding exercise, the main worry is their fear of hypoglycemia which limits their participation in sports and that obligates them to sit down and consume sugars.

Social support is fundamental, especially when coming from friends, while the parental one is oftentimes frustrating. When continuously asked about their blood glucose levels from their parents, subjects feel as if they are not trusted [34]. On the other hand, sometimes, family involvement, mainly talking about siblings [21], can be a great source of help. This dichotomy is typical of adolescents because they strive for autonomy, but also long for support in their safe transition towards new responsibilities [28] [35] [32].

Still, self-management is associated with effectively becoming an adult, which can represent an obstacle towards their adolescent identity and their socialisation: there is a discrepancy between diabetes self-care and life as a teenager [36].

Moreover, as already explained in Subsection 1.5.1, teenagers must deal with important physical changes during puberty, such as the development of secondary sexual characteristics, due to hormonal fluctuations, and the influence of the body need for and insulin sensitivity, making this phase even more challenging.

Adolescents are also subject to psycho-social development and the evolution of advanced reasoning skills and they must learn how to cope with stress, which is either general or specifically linked to diabetes and associated with glycaemic control, self-monitoring and quality of life [37].

Even the almost nonexistent knowledge of T1DM on the part of the patients' peers and friends can be seen as an obstacle: for example, their classmates' curiosity and enthusiasm can be sparked by a conversation on the topic. In fact, talking about it could be a chance to fully understand the illness as classmates usually show concern and empathy towards the diabetic subject, but they also find their routine check-ups to be strange, fear they could get infected and sick, in their turn, and are frightened by injections. That leads to the patient's isolation and alienation. Last but certainly not least, the peers of young diabetic patients usually label and ridicule them and, in some cases, they even get to bullying them [28].

Additionally, another of the main issues encountered by diabetic adolescents is the general worry of not being understood: other people, who are not affected by T1DM, cannot imagine how difficult life with this illness is like, so the subjects reluctantly accept their opinion on what they should or should not do. They also get frustrated due to public misconceptions about the pathology, leading to the choice of hesitating or avoiding telling strangers about their diabetes. That also comes from the desire of not being perceived as ill and vulnerable since they do not see themselves as such [21].

That is directly linked to the way diabetic adolescents perceive the care they receive:

they see it as being too target-driven and not enough person-centred [38] [32] [39]. They dislike health care professionals seeing them through their blood glucose values rather than as whole persons and the fact that they prefer to talk to their parents instead of them. On this topic, Starkman et al. [40] have studied how this can lead to depersonalising, which initiates a vicious circle as it can result in disengagement from therapy that causes poor glycaemic results and, subsequently, clinicians focusing even more on these.

Chapter 2

The state of the art

Throughout the history of diabetes' management, the control of this pathology has gone through many different stages. Multiple approaches, from the most intuitive to the most creative ones, have been planned and designed by teams of researchers. However, external control is yet to reach the ideal proper result given by the natural healthy insulin secretion pattern [41].

It is crucial to research and develop new and more efficient solutions. In fact, even if modern strategies and interventions on paediatric patients' internal and external support help minimise the demands placed upon them, non-optimal glycaemic control will still likely prevail for the majority of them until technological advancements will allow for automated insulin delivery in closed-loop systems [16].

Nevertheless, the main components of approaches going in this direction are CGM devices, insulin pumps, meaningful parameters and machine learning algorithms aimed at the improvement of closed-loop systems.

This work takes into consideration a CMG sensor by DexCom, an insulin pump by Tandem™ Diabetes Care and a machine learning approach to improve and perfect the choice of parameters made by the clinician to initialise the algorithm controlling the just quoted pump.

2.1 DexCom G6®

The glycaemic files used for the dataset analysed in this work contain data coming from recordings made by the G6 sensor[1] by DexCom.

It is an example of a CGM system, meaning a kind of wearable device of the medical type that can detect the blood glucose level of patients. This technology was designed mainly because self-monitoring of glycaemia plays a fundamental role in achieving the correct therapy, but only a few individuals affected by T1DM measure glucose levels as often as they should [42]. Instead, devices for the continuous glucose monitoring permit to assess the shifting interstitial glucose in an ongoing manner, also giving alarms to the subjects in case of emergency values, in a more user-friendly and easy way and facilitating them into making optimal decisions for their own treatments.

[1] `https://s3-us-west-2.amazonaws.com/dexcompdf/G6-CGM-Users-Guide.pdf`

Source: https://beyondtype1.org/comparing-the-dexcom-g6-to-the-g5/

Figure 2.1: DexCom G6® insertion device and sensor/transmitter.

Compared to the conventional glucose monitoring machines, CGM systems provide much greater insight into the glycaemic levels throughout the day, helping, through their data on trends, identifying and preventing hypo- and hyperglycaemias. The difference between an intermittent and continuous monitor for blood glucose levels is comparable to the one between a regular camera and a continuous security one [43]: the former takes discrete and clear snapshots which cannot predict the future and require a lot of effort to take, while the former takes multiple and poorly focused pictures that can be displayed in a sequential array, allowing trend predictions and operating automatically after being turned on. Thus, even if CGM are less accurate in the individual point analysis, they produce much more sensitive and profound data.

The G6 sensor is the one compatible with and accepted by the Tandem™ t:slim X2™ insulin pump. It is calibrated directly inside the factory, so it does not require any blood withdrawal from the fingertip. Fig. 2.1 shows the one-hand insertion device and the sensor (that also acts as the transmitter) with its adhesive patch.

Source: https://www.tandemdiabetes.com/products/infusion-sets

Figure 2.2: Tandem™'s t:slim X2™ insulin pump and one of its compatible infusion sets. Label 1 indicates the infusion set part that is in direct contact with the diabetic individuals and is composed of a tiny and fine soft tube (cannula) or steel needle that goes under their skin and an adhesive patch to secure it in place, label 2 indicates the thin and flexible tubing (of variable lengths) through which insulin flows and label 3 indicates the t:lock™ connector which is the physical part that connects and locks the tubing to the pump.

2.2 Tandem™ t:slim X2™

The t:slim X2™ insulin pump by Tandem™ is the one whose data is used in this work.

It is an example of a hybrid closed-loop system. The general idea of implementing a closed-loop glucose control has been a reality since the 1960s [44], but the actual approach was not feasible until much more recently: today, insulin pumps are smaller and more reliable and can be coupled with trustworthy glucose monitoring systems [16]. Moreover, being able to access safer wireless communication technologies, makes the development of wearable closed-loop insulin-delivery systems always more attainable.

Nevertheless, fully closed-loop systems, not requiring any input from users, are not a reality yet, while hybrid systems, requiring user-demanded boluses at meal-time and in case of need, are a functional and commercially widespread approach with proven efficacy at home-use for children, adolescents, adults and pregnant women [45] [46]. Thus, even if the prospects for fully closed-loop control of glycaemia have appreciably improved [47], hybrid ones are the best option in T1DM management, right now.

Tandem™'s t:slim X2™ insulin pump[2] (as shown in Fig. 2.2) is made up of the insulin pump, a 3 mL (which is the equivalent of 300 units of insulin) cartridge and a compatible infusion set. When combined with its Control-IQ™ technology and the compatible CGM system (the one on which Section 2.1 focuses), the whole set can be referred to as 'system' and the pump acts as a sensor receiver, getting data every 5 minutes.

[2] https://www.tandemdiabetes.com/docs/default-source/product-documents/t-slim-x2-insulin-pump/aw-1005628_c_user-guide-tslim-x2-control-iq-7-4-mgdl-artwork.pdf?sfvrsn=18a507d7_140

The pump has a touch-screen display to show all useful information to the user, but, in some world regions (not yet in Italy), a mobile app called t:connect™ enables the connection of a mobile device to the pump through Bluetooth® wireless communication to display data on the app itself.

The pump delivers the (at most) 300 units of insulin inside the cartridge (that must be replaced every 48-72 hours, depending on the subject's use) in either basal or bolus insulin mode.

The t:slim X2™ pump can be updated to carry an automated and personalised insulin dosing feature that enables the system to adjust insulin delivery based on the CGM sensor's measurements and prediction of glycaemic values 30 minutes into the future. However, since this is a hybrid closed-loop system, this feature cannot substitute the own patient's active diabetes management.

A Personal Profile is a group of settings that define basal and bolus delivery within specific time segments throughout 24 hours. Each profile can be personalised with a name and it allows the setting (inside specific ranges reported in the user guide) of basal rate (BR; in units/hour), a correction factor (also known as Insulin Sensitivity Factor (ISF)); it is the amount of blood glucose that is lowered by 1 unit of insulin), carb ratio (also known as insulin-to-carbohydrate ratio (ICR); it is the number of grams of carbohydrates that 1 unit of insulin covers), target BG (the target value of blood glucose, measured in mg/dL), insulin duration (the amount of time that insulin is active and available in the body after a bolus has been delivered) and carbohydrates setting (on/off). These settings are then used by the pump to calculate the right delivery of basal insulin and correction boluses. The user can create up to 6 different Personal Profiles (even though only one can be active at a time) and for each of them, up to 16 different time segments can be set. Setting multiple Personal Profiles allows for more flexibility and efficiency for patients' bodies and lifestyles, based on the different insulin delivery needs.

When using Control-IQ™ technology some of the previous settings cannot be personalised.

2.2.1 Control-IQ™ technology

After receiving FDA approval in December 2019[3], the Control-IQ™ algorithm became commercially available in the USA in 2020 from age 14 years upwards. This hybrid closed-loop system uses a model predictive control algorithm originally implemented on the DiAS platform (an ultra-portable artificial pancreas research platform for home studies of closed-loop control of blood glucose in patients affected by T1DM [48]), which was developed at the University of Virginia.

The system requires BR, ISF and ICR settings to appropriately modulate insulin delivery and it does not have any adaptive learning.

Studies and trials showed a reduction in hypo- and hyperglycaemias: these results demonstrate that Control-IQ™ is safe and efficient [49].

Control-IQ™ technology should not be used on subjects who use less than 10 units of insulin per day and that weigh less than 55 pounds (equivalent to 24,9 kilograms) as these are the minimum inputs required to initiate the algorithm and for it to operate safely. Moreover, it should not be used on children under the age of 6 years old.

[3] https://www.nsmedicaldevices.com/news/tandem-slim-x2-insulin-pump/

For safety reasons, this technology also limits the basal rate to 3 units/hour when the pump does not receive a CGM reading for 20 minutes and it shuts off when the sensor session ends.

While using Control-IQ™, the user should only inject insulin provided by and through the pump.

On the technology screen on the pump's display, the user can turn it on or off and see the weight and total daily insulin values that are set, manually entering them on the numerical keypad. The former should represent the value at the start of the system and should be updated during clinical check-ups.

Control-IQ™ is a feature that automatically adjusts insulin delivery rates and amounts in response to readings from the CGM, but the pump can be either used with or without it being enabled. When the algorithm is on, the subject should continue to take boluses to cover food intake or correct high glycaemic values.

The target CGM ranges used by the algorithm cannot be customised.

Control-IQ™ technology adjusts insulin delivery in several ways: it can decrease or suspend insulin injection when the predicted blood glucose levels are below target and increase it when they are above target and deliver up to 60% of correction bolus once per hour when needed. Each of these adjustments occurs in different ways depending on whether the user is sleeping, exercising or doing neither.

When the predicted CGM value is within the target range, the pump will deliver insulin at the rate determined by the active Personal Profile setting. When the predicted value is at or below the target range in 30 minutes in the future, the rate of delivered insulin will begin to gradually decrease to keep glycaemia inside the threshold values. Control-IQ™ can also reduce or completely suspend the basal delivery when it predicts a lowering of blood glucose levels. On the contrary, when the algorithm predicts an increase in glycaemia, it gradually increases the rate of delivered insulin. Once the maximum rate of insulin delivery has been reached, Control-IQ™ stops increasing the insulin delivery rate. The maximum insulin delivery rate is a calculated value depending on an individual's Correction Factor setting (found in the active Personal Profile), the Total Daily Insulin estimated by Control-IQ™ technology based on actual total daily insulin values and the current insulin on board (the insulin amount that is still active in the body after a bolus has been delivered). Additionally, When the algorithm predicts that the blood glucose level will be at or above 180 mg/dL 30 minutes in the future, and when the technology is either Increasing Insulin Delivery or delivering Maximum Insulin Delivery, the pump will automatically deliver correction boluses to attempt to achieve the target range by delivering 60% of the total correction bolus calculated based on the Personal Profile correction factor and predicted CGM reading. The target sensor glucose for the automatic correction bolus is 110 mg/dL. Automatic bolus delivery occurs at most once every 60 minutes and the percentage and duration between boluses is designed to avoid insulin stacking that may cause unsafe reductions in glucose values. The maximum amount of insulin that an automatic correction bolus will deliver is 6 units.

To use Control-IQ™ technology, these Personal Profile settings must be configured: basal rate, correction factor, carb ratio and target BG. Plus, carbohydrates settings must be turned on. Additionally, the weight and total daily insulin must be set by clinicians. The latter is set based on the doctor's own experience and on some general guidelines on continuous subcutaneous insulin demand linked to patients age coming from an article by Danne et al. [50] published in 2005.

Sleep mode should be scheduled through days, start and end time, but it can also be

activated manually. Another possible manual activation is the one of the Exercise mode.

2.3 Parameters and TIR

As already explained in Section 2.2 and Subsection 2.2.1, Control-IQ™ technology mainly uses 3 parameters per patient: the BR, the ISF and the ICR. These factors are inserted by the clinician and they influence the efficacy of the algorithm on the treatment of the single subject. For this reason, their choice and setting should be optimised, which is the main goal of this work. The principal value to focus on through this optimisation is the ISF as it does not affect the patients' quality of life.

The experience of clinicians is the main factor of influence on the efficient choice of settings along with the analysis of the subjects' ambulatory glucose profile [51], but how can we measure this efficiency? It is done through a parameter known as the time in range (TIR; also known as target glucose range).

The research towards such a parameter came from the improvements in CGM technologies but not in their performances. The need for agreed-upon glycaemic targets for both clinicians and diabetic patients was obvious. This is why, in February 2019, an international panel of physicians, researchers and individuals affected by T1DM with expertise in CGM met at the Advanced Technologies & Treatments for Diabetes (ATTD) Congress [52] to convene on consensus recommendations leading to the formulation of guidelines expressing the ranges of desired blood glucose levels which compose the time in ranges.

Until that moment, A1c [53] (which is to say glycosylated haemoglobin) was used to derive regression equations for the computation of the average blood glucose levels in the preceding 120 days. As such, it was the only metric used to evaluate the progression of T1DM, but it has been demonstrated to have accuracy limitations [54].

National and international medical organisations started developing and analysing glycaemic targets based on research on risks for acute and chronic complications and linked to the critical and insightful data coming from CGM systems. Some standards were born and core metrics were established in 2017 at the ATTD consensus conference [55], but it is often impractical to compute and use some of these metrics in daily clinical practice. This is why it was fundamental to choose how to interpret the data by identifying the time in ranges: target percentages of time in the different glycaemic ranges whose cut points can be adjusted to address the personal and specific needs of each patient. The metric is composed of 3 CGM measurements composed of a percentage of readings and time per day spent in these targets [52]:

- time within target glucose range (TIR),
- time below target glucose range (TBR),
- time above target glucose range (TAR).

The primary goal for safe and efficient glucose control is to increase the TIR and reduce TBR and TAR with a special focus on TBR. Thus, the consensus group agreed on choosing the threshold time percentage of CGM readings in the 24 hours that must be inside specific glycaemic ranges, but always with the clinician interest in the personalisation of these metrics to each individual's needs.

Generally, for a subject affected by T1DM [52]:

- TIR should be more than 70% of the daily readings and, during this time, blood glucose levels should be between 70 and 180 mg/dL,
- TBR should be less than 4% of the daily readings when blood glucose levels are between 54 and 70 mg/dL and less than 1% when they are below 54 mg/dL,
- TAR should be less than 25% of the daily readings when blood glucose levels are between 180 and 250 mg/dL and less than 5% when they are above 250 mg/dL.

Finally, it is crucial to point out that less glycaemic variability is linked to less cardiovascular risk [56], so it is in our interest to not only keep blood glucose levels inside TIR but to also avoid glycaemic fluctuations as much as possible.

2.4 Algorithms for the improvement of closed-loop systems

Closed-loop algorithms must be improved and optimised to reach the likelihood of producing devices that can actually act as an artificial pancreas. This is why several research groups have developed increasingly more complex algorithms to possibly achieve the goal of a fully automated glucose-responsive insulin delivery system. In the meantime, researchers have also studied improvements for the hybrid closed-loop systems to increase their performances.

For example, in 2017, Gao et al. [57] studied an approach towards an artificial pancreas made up of a CGM system, an insulin pump and an adaptive closed-loop control algorithm to try to find an efficient way to control blood glucose and insulin levels with their main goal being to create dynamic iteration and patient-oriented optimisation leading to good results in simulations. In the same year, Herrero et al. [58] focused on finding a new technique able to automatically adapt the pre-meal insulin bolus used by insulin pumps to compensate for delays in subcutaneous insulin absorption to avoid initial post-prandial hyperglycaemia, reaching a TIR of 77.5% in adolescents and 89.5% on adults through in silico tests. In 2018, Benhamou et al. [59] focused on the customisation of at home closed-loop insulin delivery in adult patients with T1DM assisted through structured remote monitoring. In 2020, Askari et al. [60] proposed an adaptive and predictive control framework to incorporate disturbance prediction and pattern learning based on the subject's historical data and subsequent forecasting. They reached a time in range of 84.4% and no hypo- or hyperglycaemic events. In the same year, Colmegna et al. [61] in silico tested a linear parameter-varying control law whose ultimate objective is reducing the user intervention to the minimum with a specific focus on moderate-intensity exercise.

While on the subject, there are multiple reviews on the machine learning and artificial intelligence approaches on this topic, like the one by Steil et al. [62] on the comparison between proportional-integral-derivative and model-predictive control and the one by Tyler er al. [63] on decision support systems for T1DM, but, at the date of writing of this work, there are no published papers or articles focused on the optimisation of the choice of the parameters required by Control-IQ™ technology implemented on Tandem™ t:slim X2™, which is the goal of this thesis.

Chapter 3

Materials and methods

This work's primary objective is to optimise the choice of the parameters set by the clinician team for the use of the Control-IQ™ algorithm, with the ultimate goal of improving the subjects' TIR.

As already explained in Section 2.3, the clinician sets the algorithm up by inserting BR, ICR and ISF settings. An inexperienced doctor will have difficulty in choosing the most efficient values, especially for the first two ones. Moreover, ISF does not affect the quality of life of the diabetic patient and, as such, it should be privileged in the optimisation.

The desired end result is an algorithm that automatically outputs the best values to be manually set by the clinicians to standardise the launch of Control-IQ™ and improve its efficacy. In order to do so, data coming from the insulin pump was used in machine learning approaches to gain insightful information on the relationship between the measurements and the settings to make an improvement feasible.

Thus, the experimental work was organised into 2 parts: the first one looking for the link and relationship between insulin and glycaemia and the second one looking for the connection between them and the parameters of the settings of Control-IQ™.

3.1 Dataset

Patients are required by the clinicians at Ancona's pediatric Ospedale Salesi to upload their DexCom G6® sensor and Tandem™'s t:slim X2™ insulin pump data on the Diasend[1] platform.

The hospital gave us data coming from the website after having removed any personal information (for privacy reasons) belonging to the subjects, except for their year of birth and sex, from their files.

At first, the most recent analysable data on the website belonged to a dataset of 48 patients: 28 females and 20 males (1 born in 1987, 1 born in 1990, 9 born between 2001 and 2003, 3 born in 2004, 16 born between 2006 and 2009, 12 born between 2010 and 2014 and 6 born between 2017 and 2019). For all of them, the main interest was on blood glucose level, insulin injections, carbohydrates intake and pump settings. For this reason, 3 subjects were immediately discarded as they had not uploaded insulin pump settings.

[1] `https://www.diasend.com/`

For the remaining subjects, data belonging to the same 91 days (3 months) span was downloaded and only 30 of them had data for at least 60 days (around 2 months). For them, additional days were downloaded, trying to choose the ones for which data was uploaded, while simultaneously acquiring the most continuous and the closest weeks and the least void days as possible.

Out of the remaining subjects, only the ones who had uploaded ingested carbohydrates quantities were chosen, thus discarding 10 of them. In the end, out of the initial 48 patients, only 20 of them were considered to be of interest.

Thus, 20 subjects compose the analysed dataset: 13 of them are females, while 7 are males. Moreover, 1 of them was born in 1990, 4 were born between 2001 and 2003, 10 were born between 2006 and 2009 and 5 were born between 2010 and 2014.

Table 3.1 displays the IDs of the chosen patients and their year of birth. The IDs are the combination of 1 letter (F or M, based on the subject's sex), one pair of numbers (representing the ascending order in which data was downloaded) and another pair of numbers (equal to the last 2 figures of the year of birth).

As already said, for each individual, data covering 3 months (91 days) was downloaded. Not all subjects' files have the same months or days as the main intent was to create a dataset with the most continuous weeks possible and not all of them had uploaded their data for the same periods. Nevertheless, they all have data regarding a 91-days-long span approximately coming from the same period of the year.

Eventually, two files were downloaded from the Diasend platform for each subject: one Excel file and a PDF one. The former is made up of 5 sheets:

- The first one reports the glycaemic values in mg/dL and the date and time at which they have been received by the pump, coming from the sensor.
- The second one reports the blood glucose values in mg/dL along with the date and time in which they were recorded by the DexCom G6® sensor (that measures values every 5 minutes).
- The third one reports insulin usage in U/h for the basal one and in U for the bolus, the bolus settings, the amount of ingested carbohydrates in grams and the linked date and time. Moreover, at the end of each day, it displays the overall amount of basal and bolus injections in U.
- The fourth one reports the pump settings; some of them concern alarms, while other insulin programs and some other ones are those required and used by Control-IQ™.
- The fifth one is just a report of the coded events daily recorded by the pump.

Before starting the preprocessing, the Excel files were reduced to two sheets: one reporting date, time and glycaemic values and the other one reporting date, time and basal, bolus and carbohydrates quantities.

Figure 3.1 displays subject M1006's blood glucose values (in mg/dL) detected by DexCom G6® sensor throughout 24 hours, taken from the Excel file: this plot makes it considerably evident how highly fluctuating glycaemia is in a normal day.

Figure 3.2 juxtaposes the single glycaemic values (in mg/dL and depicted by crosses) recorded in subject M1006's blood every 5 minutes by DexCom to the overall trend (in mg/dL) of Figure 3.1 throughout the same 24 hours.

Table 3.1: Subjects demographics. First column display their assigned ID, while the second column displays their year of birth.

ID	Year of birth
F0110	2010
F0207	2007
F1208	2008
F1408	2008
F1590	1990
F2403	2003
F2810	2010
F2910	2010
F3807	2007
F4301	2001
F4503	2003
F4609	2009
F4714	2014
M1006	2006
M2008	2008
M2711	2011
M3208	2008
M3309	2009
M4003	2003
M4409	2009

Figure 3.3 displays the amount (in grams) of ingested carbohydrates recorded by subject M1006 throughout the same 24 hours.

Figure 3.4 juxtaposes the single basal rate changes (in U/h and depicted by crosses) recorded for subject M1006 to the overall trend of basal injection velocity (in U/h) changes throughout the same 24 hours.

Figure 3.5 displays the single amounts of bolus injection (in U) for subject M1006 throughout the same 24 hours.

All just cited figures report data coming from 07/06/2021 recordings.

The PDF file, instead, reports the comparison of at most 6 of the most recently uploaded pump settings (the ones cited on the fourth point of the previous list and already defined in Sections 2.2 and 2.3 and Subsection 2.2.1) coming from the days in which users uploaded their data on the Diasend platform and highlighting the differences.

Since Personal Profiles are changed during the day-hospital check-up and, on that occasion, data is usually uploaded on the platform, the change in settings that can be downloaded should be all of them, but this is not certain. This represents an issue for our intent as it is impossible to keep track of all the changes in settings throughout the analysed period.

Nevertheless, a drop-down menu on Diasend gives the possibility to explore the settings for the days of data upload inside a specific time interval, making sure to get the most changes possible. By checking all these combinations for the subjects of interest (which are 10, as will be explained in Subsection 3.1.1), the hospital team gave us all possibly known changes inside their analysed time intervals.

3.1.1 Preprocessing

To prepare the data for being analysed, glycaemia and insulin pieces of information had to be comparable. In fact, as already explained in Section 3.1, the frequency at which blood glucose levels and insulin velocities and injections are recorded are different. Thus, the Excel files for glucose had several rows more.

Moreover, insulin and carbohydrate data were not only scarcer, but they were also often recorded at completely different times compared to the glycaemic one.

For this reason, a few steps were followed to homogenise data:

I. For those subjects who had, for almost all the rows, a discrepancy of a few minutes between glycaemia and insulin dates and times, a function was applied directly to their Excel files to either subtract or add the needed minutes to the insulin times.

II. A function called merge_asof[2] [3], from the Python's library Pandas, was used. This function was created to be applied to financial data. These time-series data points, which are attached to sequential time stamps, usually include measurements taken at very short periods.

   This function works on the columns specified in the "on" parameter and compares the value in the right data frame (B) with the one on the left (A) and if there is no perfect match, it takes the previous one and it merges them, by attaching the values of B

[2] https://pandas.pydata.org/pandas-docs/version/0.25.0/reference/api/pandas.merge_asof.html

[3] https://towardsdatascience.com/how-to-merge-not-matching-time-series-with-pandas-7993fcbce063

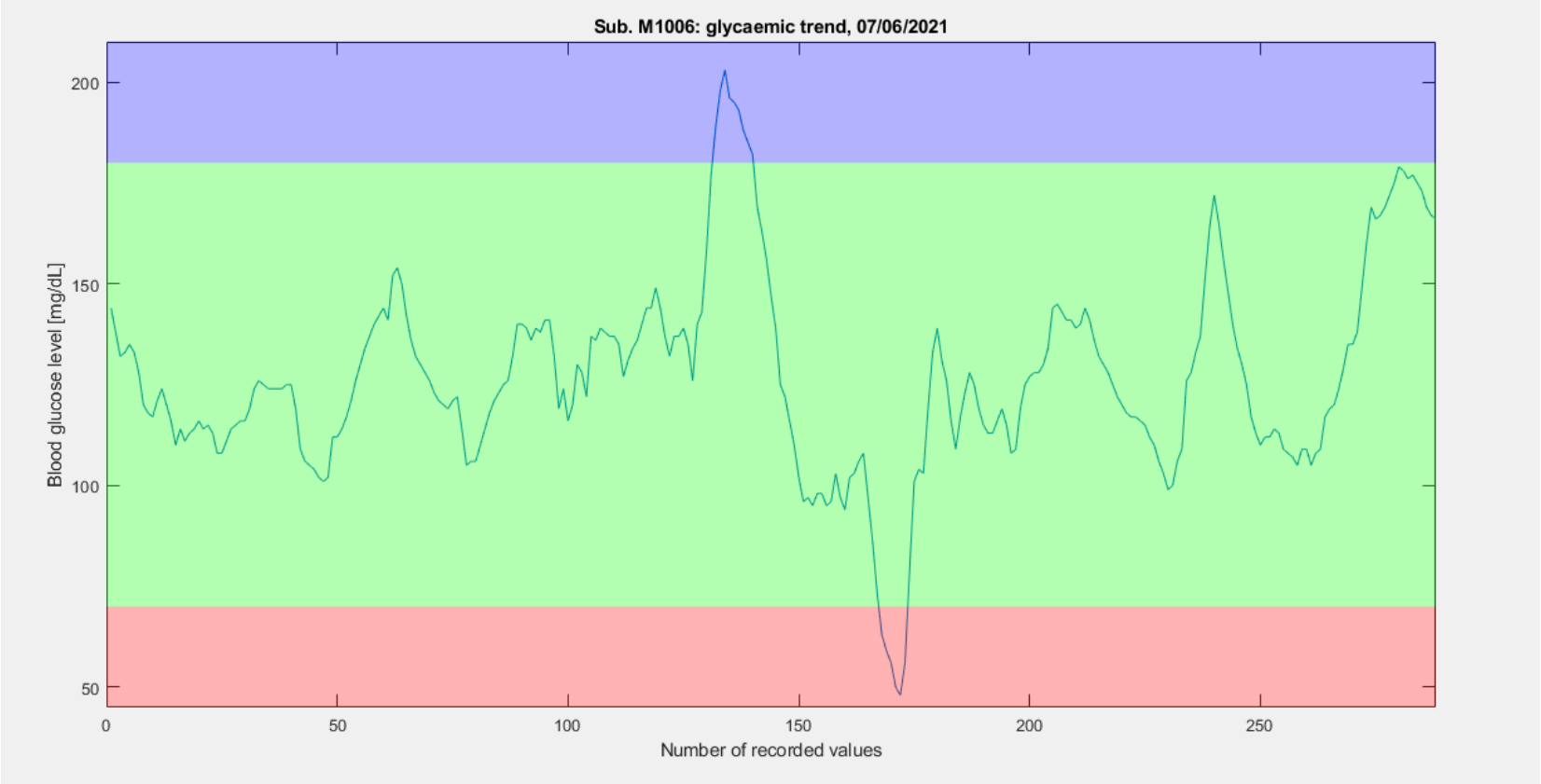


Figure 3.1: Plot of the glycaemic trend of subject M1006 on the 7^{th} of June 2021. The red, green and blue coloured areas indicate the TBR, TIR and TAR intervals (as defined in Section 2.3), respectively.

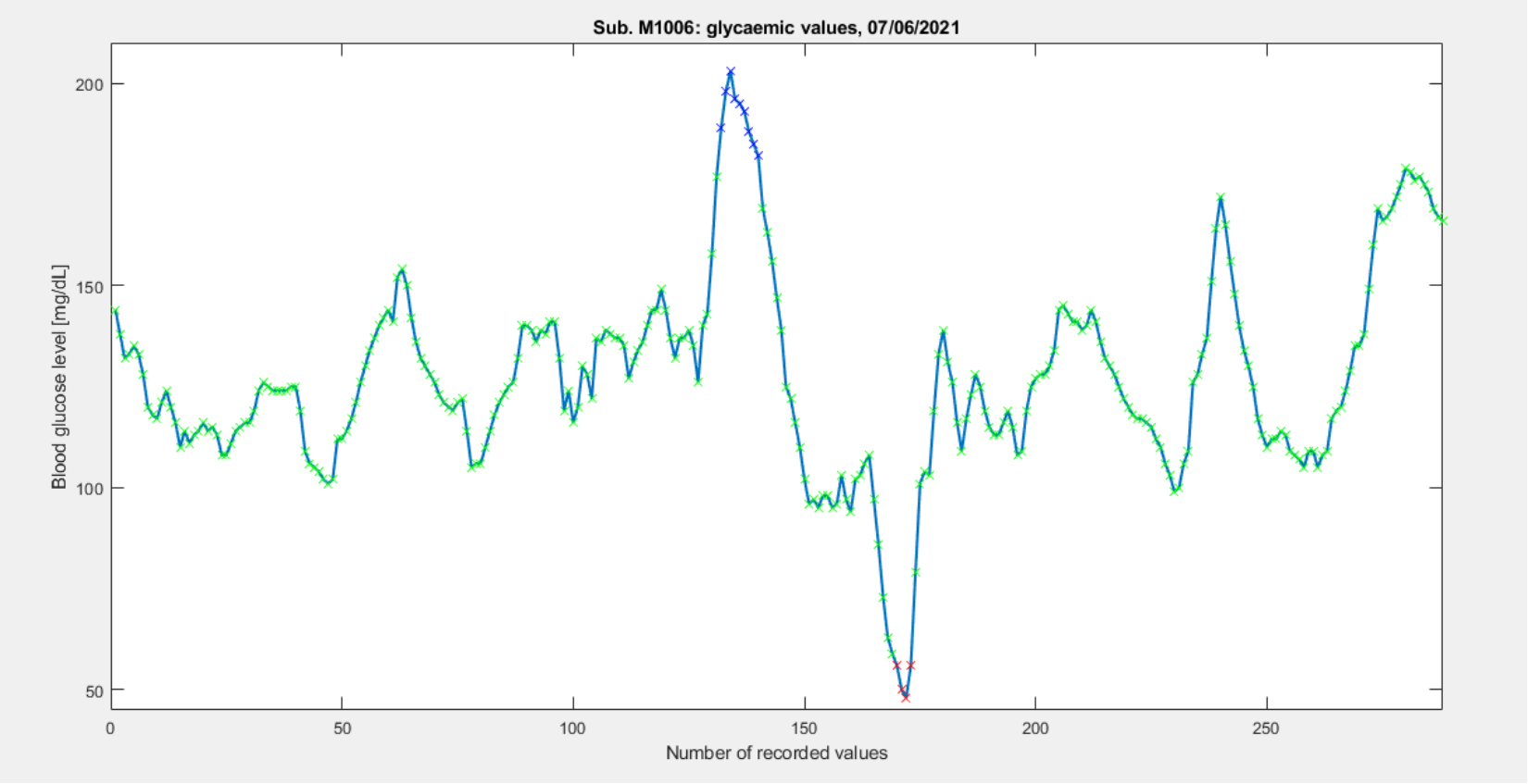


Figure 3.2: Plot of each individual glycaemic value of subject M1006 recorded on the 7^{th} of June 2021 justaposed to the overall trend. The red, green and blue coloured crosses indicate the TBR, TIR and TAR values, respectively.

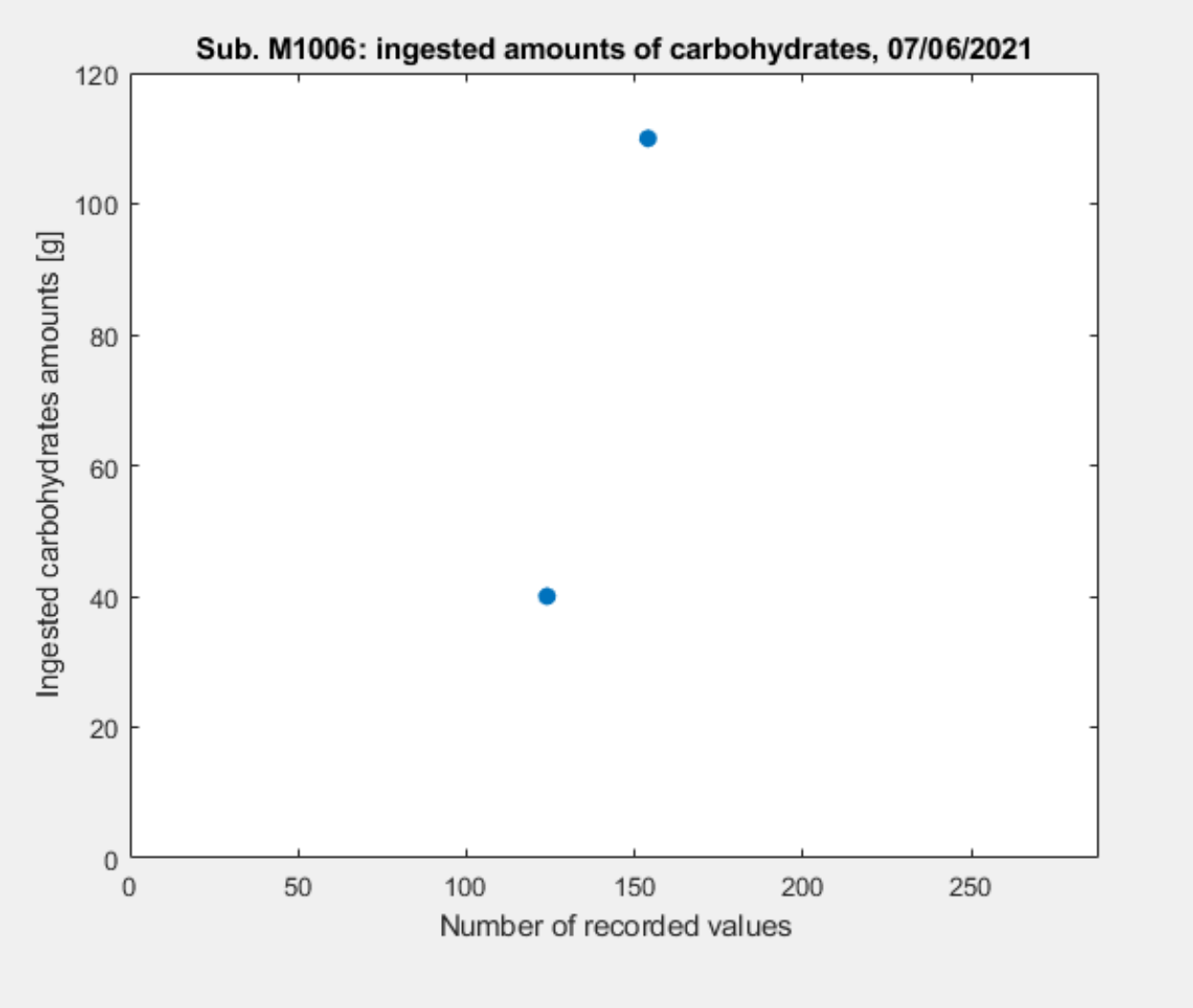


Figure 3.3: Plot of each individual carbohydrate quantity ingested and recorded by subject M1006 on the 7th of June 2021.

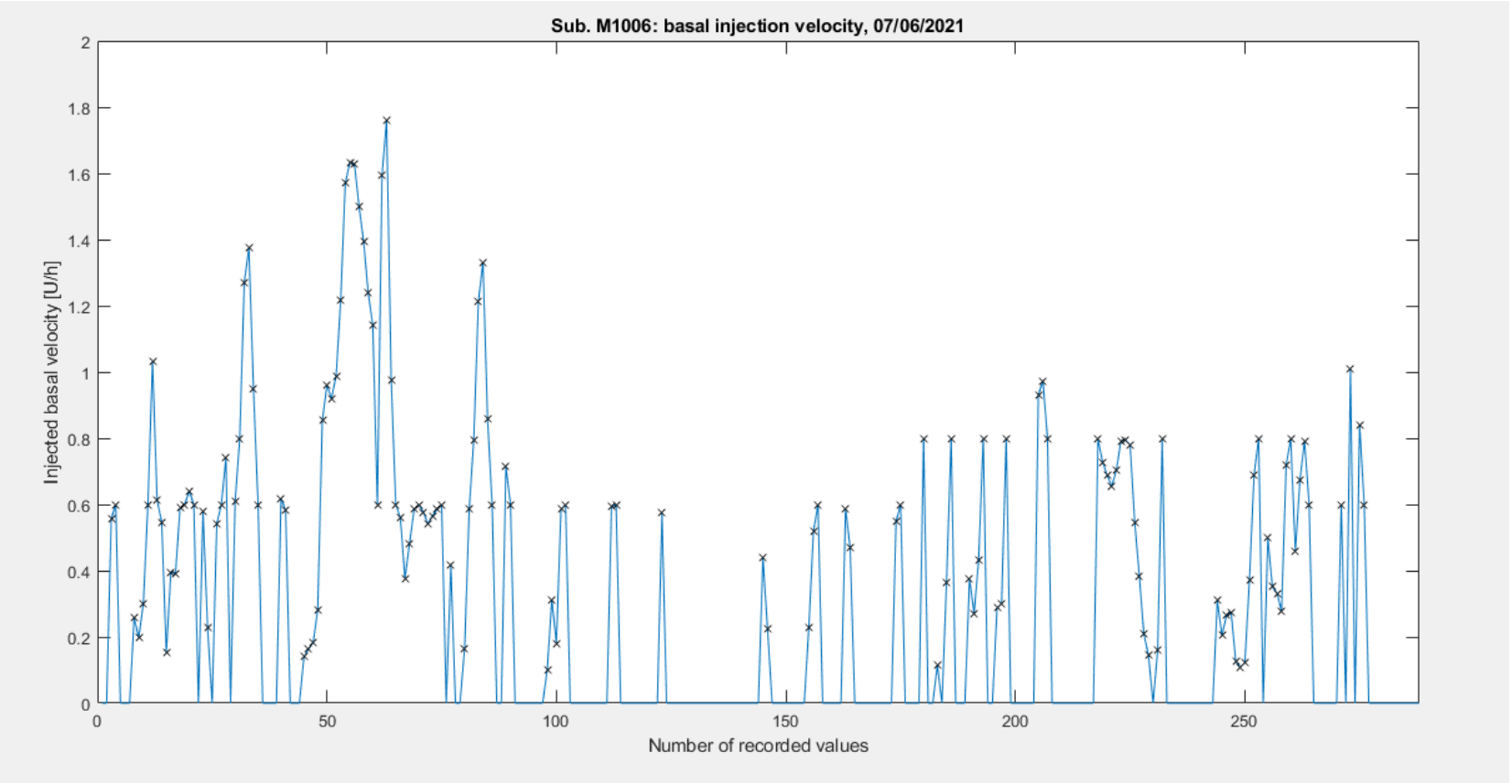


Figure 3.4: Plot of the basal rate changes (highlighted by crosses) and their overall trend for subject M1006 on the 7$^{\text{th}}$ of June 2021.

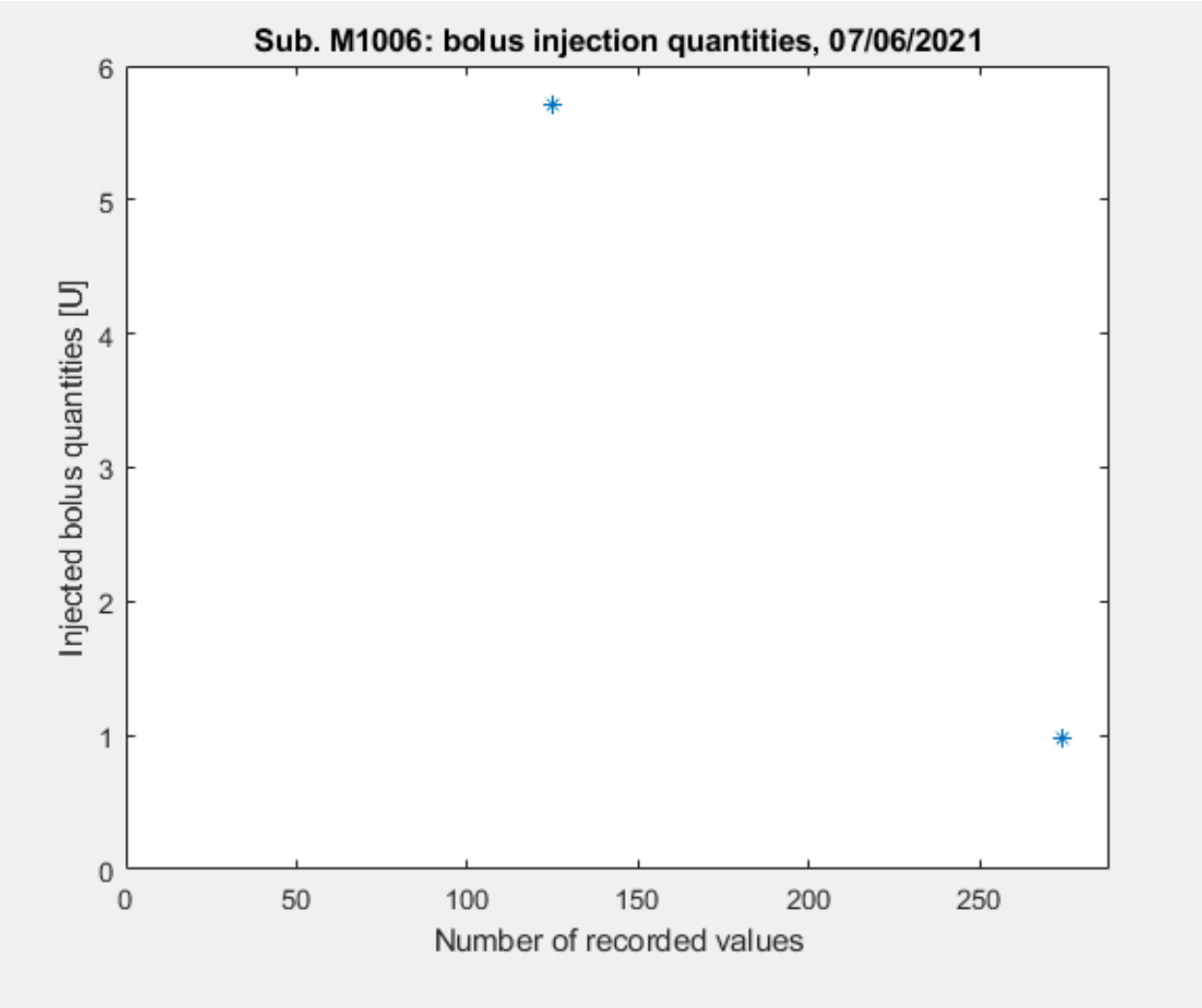


Figure 3.5: Plot of each individual bolus injection for subject M1006 on the 7^{th} of June 2021.

to the comparable times of A. It is also possible to choose a latency (in seconds) to be tolerated: in this case, since blood glucose levels are measured every 5 minutes by the DexCom G6® sensor, the time delta for the "tolerance" parameter was set at 240 seconds (4 minutes).

For the merge_asof function to be applied, 4 new excel files were created for each subject: the first one had one column for the time and dates coming from blood glucose levels' records and another one for the glycaemic values, while the second, third and fourth one had the times and dates coming from insulin and carbohydrates records on one column and, on the second one, they had the values of basal velocity, bolus injection and carbohydrates insulin, respectively.

So, the function was applied to each of these last 3 files for each subject for merging basal, bolus and carbohydrates data to the dates and times of glycaemia and the outcome was saved on an empty "back-up" Excel file to clearly divide each step and be able to easily pin down any possible error or loss of data.

III. To efficiently copy and paste these merge results, for each subject, the first column of the just created 2 insulin new files was changed with the one of the glycaemic one.

Subsequently, basal and bolus merged data were copied on the second and third column of the basal file, overwriting the preexisting data on the second column, to create a unique insulin one, while the carbohydrates merged data was copied on the third column of the glycaemia file.

IV. Lastly, it was time to fill the void rows of the merged data with zeros. Of course, the only columns with no 0 values were the ones of date and time and the blood glucose levels one.

In the end, each subject had 2 files: the first one with one column for dates and times, one for glycaemia and one for carbohydrates and the second one with one column for dates and times, one for basal and one for bolus.

In order to prepare the data for the two regressions and the random forest approach through which it was run, two paths were followed.

- Logistic regression and random forest were applied to basal insulin and glycaemic info, so the appropriate data had to be prepared.

  Of course, only the insulin file was considered. The column of date and time was deleted. A new column was created, made up of only 0s and 1s, based on either the presence or not of data on that specific row, 0 meaning no recorded value and 1 the opposite. These values were calculated through a simple Excel equation. Then, the column of basal velocities was deleted too and the new column was cut and pasted as the first and only one. This process was applied to all subjects, one at a time, ending up with 20 files, each only displaying the 0s and 1s column. Regarding the glycaemia file, the date and time column was deleted and the remaining columns were copied and pasted to cover the void one.

- Multivariate linear regression was applied to both basal insulin and glycaemic info, too.

First thing first, the basal file used at the beginning of the previous process was retrieved. Subsequently, for both the basal insulin and glycaemia file, the date and time column was deleted. In order to connect the first machine learning approach to the second one, the column of predicted labels coming from the forest approach (as it was the best one out of the previous models, as will be explained in Subsection 3.2.2) that consisted of only 0s and 1s, was copied and pasted on these two files. For the glycaemic one, 2 other new columns were created: one displaying the blood glucose level recorded in the previous row (5 minutes before) and the other one reporting the ones recorded 10 minutes before. The values inserted in the empty cells were the ones coming from the data of the previous day (taken from the original Excel files).

Then, for both Excel files, only the rows for which a basal value was recorded (based on the previous predictions) were kept, while the other ones were deleted through the help of an Excel filter and then, in both files, the column now only containing 1s was discarded. The remaining columns were cut and pasted to fill the voids. This process was applied to all subjects, one at a time, ending up with 20 files for basal insulin, each containing only the basal rows for which a velocity was recorded, and other 20 files for glycaemia, each containing the glycaemic and carbohydrates rows corresponding to the date and time for which a velocity was recorded.

Neither of these processes was performed on bolus data because recordings for it were scarce and because we cannot know from the files if the injection was manually started by the user or if it was an automatic correction made by the Control-IQ™ algorithm. So, working on it would have not been very insightful.

Finally, data was preprocessed to work on the pump settings (the ones of Personal Profiles; see Section 2.2). Work was done based on the settings' changes, on the basal velocities changes and insightful information computed from the glycaemic values:

- for what regards the settings, as already said in Section 3.1, their changes were not easy to download. The only available information on them was, unfortunately, limited,
- work done on the basal velocities changes was inferred from the initial preprocessing of the data, meaning the one coming from the IV point of the first list of this Section,
- glycaemic data was taken from its initial preprocessing, just like it was done for the basal one.

The work on the pump settings was done for only 10 subjects. This choice was done based on the balance of their basal data, the content of their settings' changes and the computation of their TIR.

To be able to determine the subjects on which the next approach was to be performed, the basal files (the ones just cited) of all of them were used inside a Python algorithm to compute the percentage of lines in which changes of basal velocities were recorded and the percentage of those in which they were not. This was a way to calculate how balanced the data was for each subject. At the end of this step, out of the 20 individuals composing the entire dataset, 9 of them were discarded as they had an imbalanced ratio of rows in favour of the ones for which a change was not recorded.

For the 11 subjects left, another Python algorithm was applied to calculate their TIR (based on the clinicians' guidelines [52]):

- F0110: TIR = 75,58%,
- F0207: TIR = 82,67%,
- F1208: TIR = 72,75%,
- F1408: TIR = 89,54%,
- F1590: TIR = 44,69%,
- F2910: TIR = 83,08%,
- F4609: TIR = 92,16%,
- M1006: TIR = 88,29%,
- M2008: TIR = 52,45%,
- M3208: TIR = 76,04%,
- M4003: TIR = 49,91%.

Subject F1408 was discarded as she had no retrievable changes in the Personal Profile settings in the analysed time interval.

Data analysis was done based on the settings' changes for each subject. All of them had changes in either BR or ICR over different date intervals and with a different number of parameters' settings. Only one subject (F0110) had ISF changes, too. So, even if it was the parameter on which we wanted to focus, it was not possible to do so.

The overall changes of the parameters created different date intervals regarding BR and/or ICR changes for each subject in which different time intervals of interest could be studied for the two approaches (A and B) that will be defined in Subsection 3.2.4:

- F0110
  - A: from 01/04/2021 to 08/04/2021, from 09/04/2021 to 21/05/2021, from 22/05/2021 to 05/06/2021 and 06/06/2021 to 30/06/2021 to study the lunch interval as no changes are reported for dinner time,
  - B:
    - from 01/04/2021 to 08/04/2021, from 09/04/2021 to 21/05/2021, from 22/05/2021 to 05/06/2021 and from 06/06/2021 to 30/06/2021 to study the effect of changes of BR and ICR in 2 time intervals (06:00-12:00 and 12:30-14:00),
    - from 01/04/2021 to 08/04/2021 and from 09/04/2021 to 30/06/2021 to study the effect of changes of ISF in 1 time interval (12:30-14:00),
- F0207
  - A:
    - from 01/04/2021 to 17/05/2021 and from 18/05/2021 to 30/06/2021 to study the lunch interval,

    - from 01/04/2021 to 12/05/2021, from 13/05/2021 to 17/05/2021 and from 18/05/2021 to 30/06/2021 to study the dinner interval,
  - B: from 01/04/2021 to 12/05/2021, from 13/05/2021 to 17/05/2021 and from 18/05/2021 to 30/06/2021 to study the effect of changes of BR in 1 time interval (15:00-20:00) as no changes are reported for ICR,

- F1208
  - A: from 01/04/2021 to 12/04/2021, from 13/04/2021 to 16/06/2021 and from 17/06/2021 to 30/06/2021 to study the lunch interval as no changes are reported for dinner time,
  - B: from 01/04/2021 to 16/05/2021 and from 16/05/2021 to 30/06/2021 to study the effect of changes of BR and ICR in 1 time interval (12:00-15:00),

- F1590
  - A: from 01/04/2021 to 12/04/2021, from 13/04/2021 to 06/05/2021, from 07/05/2021 to 12/05/2021, from 13/05/2021 to 16/05/2021 and from 17/05/2021 to 30/06/2021 to study the lunch interval as no changes are reported for dinner time,
  - B:
    - from 01/04/2021 to 12/04/2021, from 13/04/2021 to 06/05/2021 and from 07/05/2021 to 12/05/2021 to study the effect of changes of ICR in 3 time intervals (06:30-12:00, 12:00-15:00 and 18:30-23:00),
    - from 13/05/2021 to 16/05/2021 and from 17/05/2021 to 07/06/2021 to study the effect of changes of BR and ICR in 3 time intervals (00:00-04:00, 08:00-12:00 and 17:00-23:00),
    - from 01/04/2021 to 12/05/2021 and from 13/05/2021 to 30/06/2021 to study the effect of changes of the number of settings in the 24 hours in 4 time intervals (00:00-06:30, 06:30-12:00, 12:00-15:00 and 15:00-18:30),

- F2910
  - A: from 01/04/2021 to 26/04/2021 and from 27/04/2021 to 30/06/2021 to study the lunch interval as no changes are reported for dinner time,
  - B:
    - from 01/04/2021 to 26/04/2021, from 27/04/2021 to 02/05/2021 and from 03/05/2021 to 30/06/2021 to study the effect of changes of BR and ICR in 1 time interval (16:00-19:00),
    - from 01/04/2021 to 02/05/2021 and from 03/05/2021 to 30/06/2021 to study the effect of changes of the number of settings in the 24 hours in 1 time interval (03:00-08:00),

- F4609
  - A: from 01/04/2021 to 21/05/2021, from 22/05/2021 to 11/06/2021 and from 12/06/2021 to 30/06/2021 to study the lunch and dinner intervals,

  - B: from 01/04/2021 to 21/05/2021 and from 22/05/2021 to 30/06/2021 to study the effect of changes of BR and ICR in 6 time intervals (05:00-07:00, 09:00-11:00, 11:00-13:00, 15:00-16:00, 19:00-22:00 and 22:00-00:00),

- M1006
  - A: from 01/04/2021 to 19/04/2021, from 20/04/2021 to 26/04/2021 and from 27/04/2021 to 30/06/2021 to study the lunch and dinner intervals,
  - B: from 01/04/2021 to 19/04/2021 and from 20/04/2021 to 30/06/2021 to study the effect of changes of BR and ICR in 1 time interval (12:00-14:00),

- M2008
  - A: from 01/04/2021 to 22/04/2021 and from 23/04/2021 to 30/06/2021 to study the lunch and dinner intervals,
  - B: from 01/04/2021 to 19/04/2021 and from 20/04/2021 to 30/06/2021 to study the effect of changes of BR in 5 time intervals (00:00-03:00, 03:00-8:00, 8:00-15:00, 15:00-21:00 and 21:00-00:00) as no changes are reported for ICR,

- M3208
  - A: from 01/04/2021 to 15/04/2021 and from 16/04/2021 to 30/06/2021 to study the lunch interval as no changes are reported for dinner time,
  - B: from 01/04/2021 to 15/04/2021, from 16/04/2021 to 13/05/2021 and from 14/05/2021 to 30/06/2021 to study the effect of changes of BR and ICR in 1 time interval (07:00-15:00),

- M4003
  - A: from 01/04/2021 to 16/04/2021 and from 17/04/2021 to 30/06/2021 to study the lunch interval as no changes are reported for dinner time,
  - B: from 01/04/2021 to 16/04/2021 and from 17/04/2021 to 30/06/2021 to study the effect of changes of BR in 1 time interval (15:00-21:00) as no changes are reported for ICR.

For this reason, for each subject, for the two different approaches, the needed number of glycaemic files was made out of the initial one based on their date intervals of interest to study the blood glucose values at meals and the needed number of basal and glycaemic files was made out of their initial ones based on their date intervals that enabled the study of blood glucose values at the time intervals of interest.

Additional preprocessing to attempt to improve performances

To try to improve the performances of the first 2 machine learning approaches, the full glycaemic files (with all the original rows) were changed to portray, at first, 2 additional columns reporting the blood glucose levels recorded 5 and 10 minutes before and, then, the 6 columns with the levels recorded until 30 minutes before, meaning the 6 previous rows.

The process followed was the same as the one applied to the multivariate linear regression data.

While the addition of only 2 columns was done just to see if adding info could improve results as a previous approach had already been done on the starting data through MATLAB v. R2019b, the addition of the 6 columns was done based on strategy: Control-IQ™ predicts the glycaemic trend of the following 30 minutes, meaning that the basal velocity of each row where it is present is based on the blood glucose level recorded 30 minutes prior, which is to say 6 glycaemic measurements prior.

For each file for which glycaemic columns were added, another one was made in which the rows for which blood glucose levels were added were deleted (thus, the first 2 or 6 rows).

Also, the logistic regression files were copied and pasted to be normalised to have comparable orders of magnitude to see if results could be improved in this way, but the addition of columns was not done on them.

Finally, a few of the glycaemic files for the multivariate linear regression were changed to add the 6 columns with the levels recorded until 30 minutes before to test if that would improve the reached results.

3.2 Modelling

This section reports the definitions and explanation of the machine learning modelling approaches used in this work.

At first, the focus was on a model able to tell if the Control-IQ™ algorithm decided to either change the basal rate or not based on the analysis of the 0 and 1 labels where the former indicates a row for which there was no record of change in basal insulin velocities and the latter indicates one for which such a value was recorded. The logistic regression did not perform well enough, so a random forest model was implemented to look for better results.

Then, the outcome of the previous model (the random forest one), meaning the array of predicted labels, was used to preprocess (as said in Section 3.1.1) the data given as input to the multivariate linear regression. In fact, since this model's goal was to quantify the change in basal insulin delivery velocity, it was applied to files only containing the rows related to records of these changes.

Figure 3.6 shows a summarising flowchart of the thinking behind these approaches.

3.2.1 Logistic regression

The first machine learning approach was the logistic regression one. At first, as already anticipated in Subsection 3.1.1, it was done on MATLAB v. R2019b, but then the results were discarded and all the machine learning portion of this work has been done on Google's Colaboratory, working in Python[4].

Logistic regression uses classification to predict a binary outcome (in this case: 0 or 1, meaning no change in basal rate and change in basal velocity, respectively). Working on a set of independent variables, the logistic regression model classifies them and predicts how probable an event is to occur.

[4] `https://towardsdatascience.com/building-a-logistic-regression-in-python-step-by-step-becd4d56c9c8`

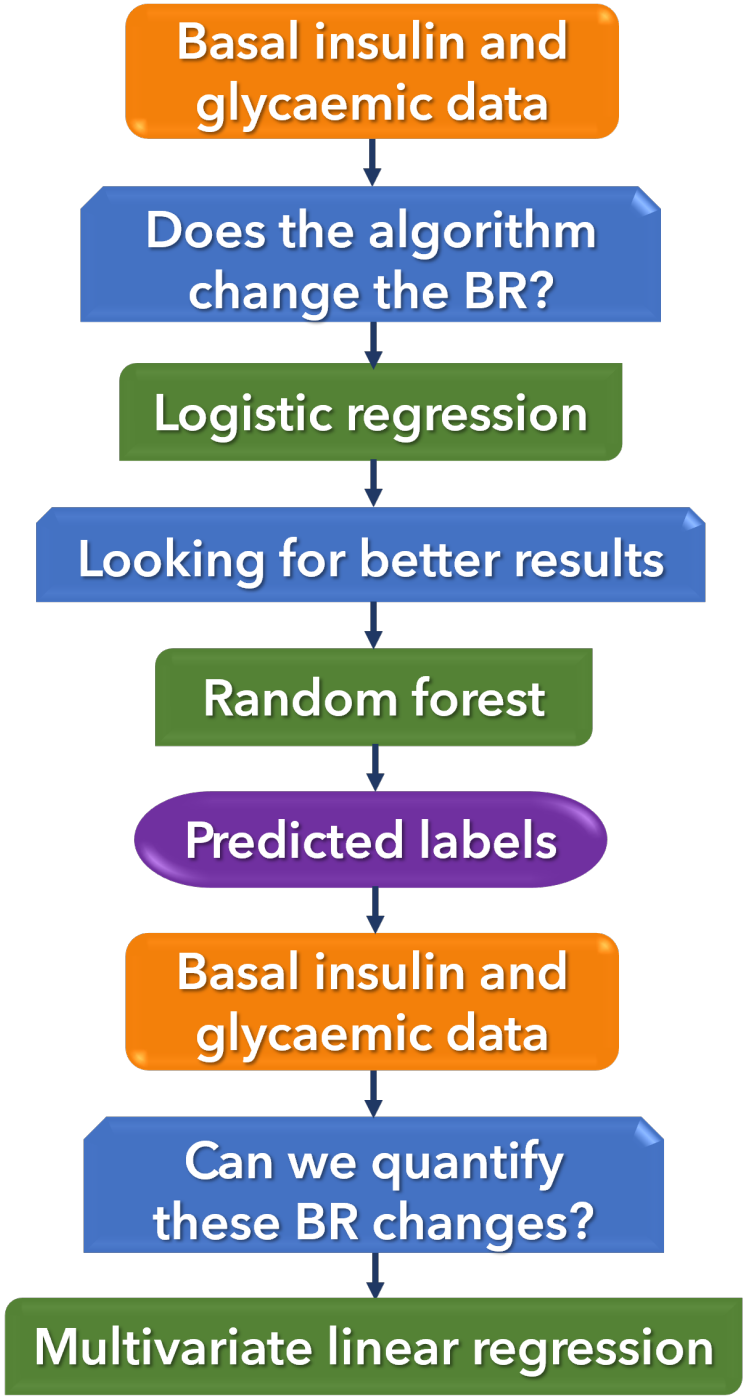


Figure 3.6: Flowchart of the modelling approach. The orange boxes report the data on which the ML algorithms were applied, the blue ones report the questions we tried to answer, the green ones report the applied ML approaches and the purple oval reports the output of the random forest approach that was used as an input for the preprocessing of the data on which the multivariate linear regression was used.

By using the module called statsmodels[5], a logistic regression model was built, on a Colab notebook, through the Logit(y,X) function where y is the column of 0 and 1s of the basal file and X is constituted by data coming from glycaemic and carbohydrate info.

In the example of logistic regression applied to the glycaemia file where the columns for the 2 previous basal levels were added, the logistic regression equation was:

$$Pr(y = 1|X) = \frac{exp(\beta_0 + \beta_1 Glycemia + \beta_2 Carbs + \beta_3 Glyc_{tminus1} + \beta_4 Glyc_{tminus2}}{1 + exp(\beta_0 + \beta_1 Glycemia + \beta_2 Carbs + \beta_3 Glyc_{tminus1} + \beta_4 Glyc_{tminus2})} \tag{3.1}$$

where $Glycemia$ indicates the glycaemic recorded values, $Carbs$ indicated the carbohydrates recorded intake, $Glyc_{tminus1}$ indicates the previous (5 minutes before) recorded values and $Glyc_{tminus2}$ indicates the second previous (10 minutes before) recorded values.

By using the open-source library Scikit-learn[6] and its functions, for each patient, the model was fitted and the script computed its accuracy (see Eq. 3.2), its predicted values, its confusion matrix, its precision ($Prec$, see Eq. 3.3), recall (Rec, see Eq. 3.4) and F1-score ($F1$, see Eq. 3.5) and its receiver operating characteristic (ROC) curve and area under it. Additionally, the percentage of rows reporting injection and no injection (represented by the recording of a change in insulin basal velocity) was computed to be used later on to choose the best individuals on which to work with the Control-IQ™ settings.

Here are reported the equations of the used metrics [64], where TN represents the true negatives, FP represents the false positives, FN represents the false negatives and TP represents the true positives:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3.2}$$

$$Prec = \frac{TP}{TP + FP} \tag{3.3}$$

$$Rec = \frac{TP}{TP + FN} \tag{3.4}$$

$$F1 = \frac{2 * Prec * Rec}{Prec + Rec} = \frac{2 * TP}{2 * TP + FP + FN} \tag{3.5}$$

The ultimate goal of this approach was to compute if Control-IQ™ decided to change the insulin rate or not based on info coming from glycaemic values and carbohydrate intake.

[5] `https://www.statsmodels.org/stable/index.html`

[6] `https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html`

3.2.2 Random forest

The random forest approach was used in an attempt to improve the performance (as will be explained in Chapter 5) of the algorithm applied to the basal insulin data linked to the presence of change in basal rate or not. Ir was implemented on Google's Colaboratory, working in Python[7].

It was used to solve both the regression and classification problems. The random forest algorithm picks N random records from the dataset and builds a decision tree based on them, repeatedly for the chosen number of trees wanted by us (in this case, 200 trees). In the case of a regression problem, each tree predicts a value for the output and the final value can then be calculated by taking the average of all previously calculated values. Instead, in the case of a classification problem, each tree predicts the category to which the new record belongs and then the latter is assigned to the category that wins the majority vote.

The random forest algorithm presents some interesting advantages: it is not biased because there are multiple trees and each one is trained on a subset of data, it is stable and it works well even when data has missing values or it has not been scaled well.

By using the open-source library Scikit-learn[8][9] and its functions, two random forest approaches were prepared, on a Colab notebook, through the RandomForestRegressor and RandomForestClassifier functions, after having scaled the data through Scikit-learn's StandardScaler[10] function.

Always through the same library as before, for each patient, the script was used to compute the mean absolute error (MAE), the mean squared error (MSE) and the root mean squared error (RMSE) for the regression and the $Prec$, Rec and $F1$, the accuracy and the confusion matrix for the classification.

Here are reported the equations of the used metrics [64]; the ones for the Accuracy (Eq. 3.2), $Prec$ (Eq. 3.3), Rec (Eq. 3.4) and $F1$ (Eq. 3.5) have been already defined.

$$MAE = \sum_{i=1}^{D} |x_i - y_i| \tag{3.6}$$

$$MSE = \sum_{i=1}^{D} (x_i - y_i)^2 \tag{3.7}$$

$$RMSE = \sqrt{\frac{1}{n} \Sigma_{i=1}^{n} \Big(\frac{d_i - f_i}{\sigma_i} \Big)^2} \tag{3.8}$$

[7] https://stackabuse.com/random-forest-algorithm-with-python-and-scikit-learn/

[8] https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html

[9] https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html

[10] https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html

3.2.3 Multivariate linear regression

Multivariate linear regression was used to try to understand the connection between the change of insulin rate and glycaemia. As already said on Subsection 3.2.1, this was done on Google's Colaboratory, working in Python[11].

Multivariate linear regression is used to understand if there is a relationship between multiple variables, how strong it is, which variable contributes the most, how accurately we can estimate the effect of each variable, how accurately we can predict the target and if the said relationship is linear.

By using the open-source library Scikit-learn[12] and its functions, a multivariate linear regression model was built, on a Colab notebook, through the LinearRegression(X,y) function where y is data coming from the basal file and X is constituted by data coming from glycaemic and carbohydrate info.

In the example of the multivariate linear regression applied to the glycaemia file where the columns for the 2 previous basal levels were added, the multivariate linear regression equation was:

$$Y = \beta_0 + \beta_1 Glycemia + \beta_2 Carbs + \beta_3 Glyc_{tminus1} + \beta_4 Glyc_{tminus2} \tag{3.9}$$

where $Glycemia$ indicates the glycaemic recorded values, $Carbs$ indicated the carbohydrates recorded intake, $Glyc_{tminus1}$ indicates the previous (5 minutes before) recorded values and $Glyc_{tminus2}$ indicates the second previous (10 minutes before) recorded values.

Always through the same library as before, for each patient, the model was fitted and the script computed the equation coefficients (the intercept and the regression coefficients), its predicted response, the p-value of each coefficient, the R-squared (R^2) value and F-statistic (F-stat) one. Moreover, the plot of the residuals was computed and saved.

Here are reported the equations of the used metrics [65] [64] [66]:

$$\text{p-value} = 1 - \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \tag{3.10}$$

$$R^2 = 1 - \frac{\text{sum squared regression}}{\text{total sum of squares}} = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y}_i)^2} \tag{3.11}$$

$$\text{F-stat} = \frac{\text{variance of } 1^{\text{st}} \text{ dataset}}{\text{variance of } 2^{\text{nd}} \text{ dataset}} = \frac{\sigma_1^2}{\sigma_2^2} \tag{3.12}$$

[11] https://towardsdatascience.com/the-complete-guide-to-linear-regression-in-python-3d3f8f06bf8

[12] https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html

3.2.4 Pump settings' analysis

To analyse how the pump settings' changes perform, work on TIR and basal velocities was to be done based on the Personal Profile's changes during the overall evaluated period (the 91 days of data used) and their time intervals and values of basal rate, insulin-to-carbohydrate ration and ISF.

Only for subject F0110, ISF changes were reported. Thus, even if it would have been preferable to work on that parameter for the whole dataset, only the BR and/or ICR could be considered for the others.

Data analysis was done on Google's Colaboratory, working in Python, based on the settings' changes for each subject working on the preprocessed files described in Subsection 3.1.1.

Two different approaches were applied, but both used a metric that will be called "tir" as it is based on TIR but calculated for different intervals.

(A) based on both the basal rate and insulin-to-carbohydrate ratio changes, the subjects' meal tirs (lunch: 11:30-15:00 and dinner 18:30 - 23:00) were computed for each of their different date intervals. This was done to analyse how the combination of changes of the parameters perform, meaning how well they control blood glucose levels keeping them inside the guidelines' thresholds [52]. The choice of these mealtime intervals was based on the metabolic regulation of food intake both in healthy and diabetic individuals [67] [68] [69] and on clinicians' experience on the matter.

It would have been interesting to work on another time interval in the middle of the night to analyse how good the settings perform while circadian rhythms influence the subjects' glucose metabolism. In fact, multiple researchers have found relationships between the two [70]. Circadian rhythms are fundamental biological processed that enable organisms to predict and prepare for changes in the surrounding environment due to the 24-hours Earth rotation on its own axis [71]. These rhythms are synchronised by the circadian clock that, based on light and darkness cycles, organises the chemical, molecular and psychological processes in time. A steady-state (bed-rest and fasting) experiment conducted by Trümper et al. [72] demonstrated that individuals affected by T1DM have less insulin sensitivity during the second half of the night, leading to higher insulin dosage, even in absence of external influences like physical activity and food intake.

Unfortunately, not all subjects have any changes in settings in the interval between 00:00 and 06:00, so this analysis was performed in approach B only on the eligible subjects.

(B) based solely on parameters' changes, the subjects' time intervals of interest were identified to calculate their tirs inside of them. For the time intervals for which BR changed, these values will be reported in the following list as they were further analysed:

- F0110
  - from 01/04/2021 to 08/04/2021: BR = 1,5 for time interval 06:00-12:00 and BR = 1,5 for time interval 12:30-14:00,
  - from 09/04/2021 to 21/05/2021: BR = 1,6 for time interval 06:00-12:00 and BR = 1,6 for time interval 12:30-14:00,

  - from 22/05/2021 to 05/06/2021: BR = 1,5 for time interval 06:00-12:00 and BR = 1,2 for time interval 12:30-14:00,
  - from 06/06/2021 to 30/06/2021: BR = 1,6 for time interval 06:00-12:00 and BR = 1,6 for time interval 12:30-14:00,
- F0207
  - from 01/04/2021 to 12/05/2021: BR = 1,2 for time interval 15:00-20:00,
  - from 13/05/2021 to 17/05/2021: BR = 0,9 for time interval 15:00-20:00,
  - from 18/05/2021 to 30/06/2021: BR = 0,5 for time interval 15:00-20:00,
- F1208
  - from 01/04/2021 to 16/05/2021: BR = 0,7 for time interval 12:00-15:00,
  - from 16/05/2021 to 30/06/2021: BR = 0,85 for time interval 12:00-15:00,
- F1590
  - from 13/05/2021 to 16/05/2021: BR = 0,6 for time interval 00:00-04:00, BR = 0,5 for time interval 08:00-12:00 and BR = 0,7 for time interval 17:00-23:00,
  - from 17/05/2021 to 07/06/2021: BR = 0,5 for time interval 00:00-04:00, BR = 0,4 for time interval 08:00-12:00 and BR = 0,6 for time interval 17:00-23:00,
- F2910
  - from 01/04/2021 to 26/04/2021: BR = 1,5 for time interval 16:00-19:00,
  - from 27/04/2021 to 02/05/2021: BR = 1 for time interval 16:00-19:00,
  - from 03/05/2021 to 30/06/2021: BR = 1,2 for time interval 16:00-19:00,
- F4609
  - from 01/04/2021 to 21/05/2021: BR = 0,6 for time interval 05:00-07:00, 0,6 for time interval 09:00-11:00, 0,6 for time interval 11:00-13:00, 0,8 for time interval 15:00-16:00, 0,8 for time interval 19:00-22:00 and 0,6 for time interval 22:00-00:00,
  - from 22/05/2021 to 30/06/2021: BR = 0,65 for time interval 05:00-07:00, 0,65 for time interval 09:00-11:00, 0,65 for time interval 11:00-13:00, 0,75 for time interval 15:00-16:00, 0,75 for time interval 19:00-22:00 and 0,75 for time interval 22:00-00:00,
- M1006
  - M1006: 12:00 - 14:00 basal same basal rate = 0,6 for both date intervals, but set at 13:00 for the first one and 12:30 for the second one) were computed for each of their 2 different date intervals. Before this, additional analysis on the magnitude and frequency of basal velocities changes (different from the basal rate set in the Personal Profile of the subject) was performed to be able to evaluate how the basal rate pump setting adheres to the real needs of the patient.
  - from 01/04/2021 to 19/04/2021: BR = 0,6 for time interval 12:00-14:00 (parameter set at 13:00),
  - from 20/04/2021 to 30/06/2021: BR = 0,6 for time interval 12:00-14:00 (parameter set at 12:30),

- M2008
  - from 01/04/2021 to 19/04/2021: BR = 1 for time interval 00:00-03:00, BR = 1,3 for time interval 03:00-8:00, BR = 1,1 for time interval 8:00-15:00, BR = 1,6 for time interval 15:00-21:00 and BR = 1,3 for time interval 21:00-00:00,
  - from 20/04/2021 to 30/06/2021: BR = 1,2 for time interval 00:00-03:00, BR = 1,4 for time interval 03:00-8:00, BR = 1,4 for time interval 8:00-15:00, BR = 1,4 for time interval 15:00-21:00 and BR = 1,4 for time interval 21:00-00:00,
- M3208
  - from 01/04/2021 to 15/04/2021: BR = 0,7 for time interval 07:00-15:00,
  - from 16/04/2021 to 13/05/2021: BR = 1 for time interval 07:00-15:00,
  - from 14/05/2021 to 30/06/2021: BR = 1,5 for time interval 07:00-15:00,
- M4003
  - from 01/04/2021 to 16/04/2021: BR = 1,2 for time interval 15:00-21:00,
  - from 17/04/2021 to 30/06/2021: BR = 1,5 for time interval 15:00-21:00.

By working on time masks coded in Python[13], computations were conducted, on a Colab notebook, only applied to the files linked to the date intervals of interest and on the data linked to the time intervals of interest. These time masks covered the different time intervals that have just been reported.

Approach A's script computed only the lunch and/or dinner tirs for the different date intervals for each subject.

Approach B's script computed the percentage of rows for which a change in basal velocity was recorded in the analysed time interval, the mean of these changes and their standard deviation (SD), their difference from the original Profile Personal setting and the tir for each time interval for the different date intervals for each subject. Moreover, it also computed just the tirs for the intervals of interest linked to changes of different parameters than BR.

The following is the equation [64] for the computation of SD:

$$SD = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2} \tag{3.13}$$

[13] `https://stackoverflow.com/questions/50250022/pandas-selecting-rows-in-a-specific-time-window/50250530`

Chapter 4

Results

This part reports, in each of its sections, the results achieved with the different machine learning approaches described in Chapter 3.

First, it reports the results coming from the logistic regression: confusion matrices, $Prec$, Rec and $F1$ values and ROC curves.

Then, it reports the results coming from the random forest: MAE, MSE, RMSE, confusion matrices and $Prec$, Rec and $F1$ values.

Then, it reports the results coming from the multivariate linear regression: the coefficients of the equation and their p-values and R^2 and F-stat values.

Finally, it reports the results coming from the analysis of the pump settings which is to say the tirs for the intervals of interest.

4.1 Logistic regression

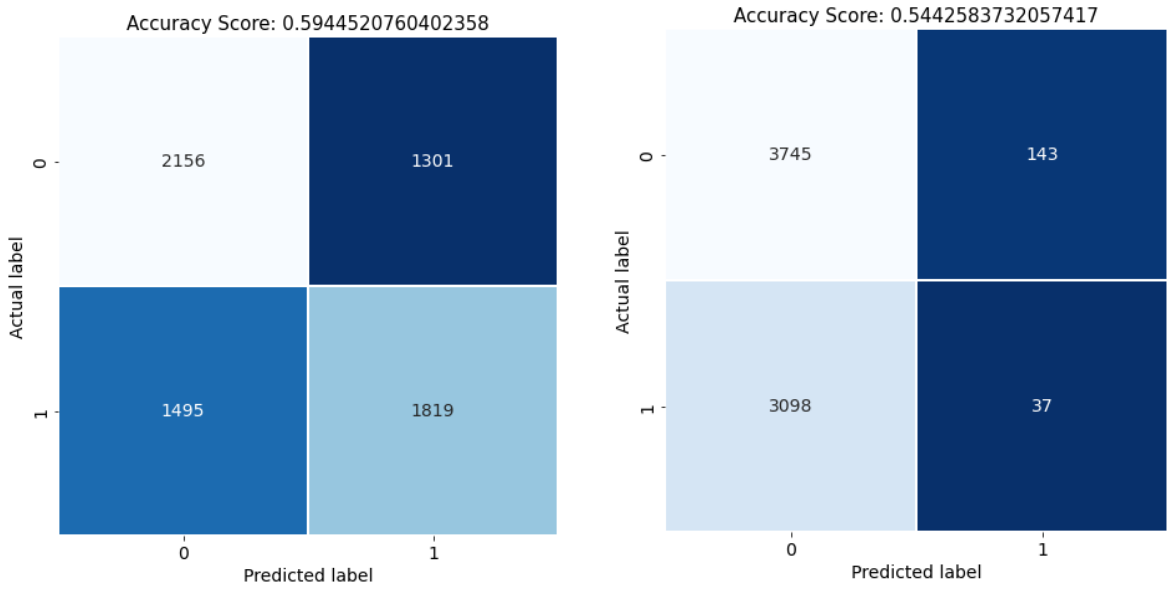
This section reports the results coming from the logistic regression procedure defined in Subsection 3.2.1.

Since, as will be explained in Section 5.1, some approaches where the analysed data was changed to try to improve performances did not give the hoped results, their results will not be reported here. For instance, the logistic regression in which the glycaemic columns reporting the blood glucose levels in the 2 previous rows were added resulted in a non-existent or, at least, a non-significant difference. Moreover, when the rows for which data had been added were deleted, performance changes were still not noteworthy and meaningful. Finally, normalising to try to reach comparable orders of magnitude resulted in minuscule and negligible variations.

Figures 4.1, 4.2, 4.3, 4.4 and 4.5 display the confusion matrix of each subject. The latter is reported in the subfigures' captions. The number on the first square (first row and column) represents the true negatives (TN), the second square (first row and second column) represents the false positives (FP), the third square (second row and first column) represents the false negatives (FN) and the fourth square represents the true positives (TP). Above each matrix, the accuracy score is reported.

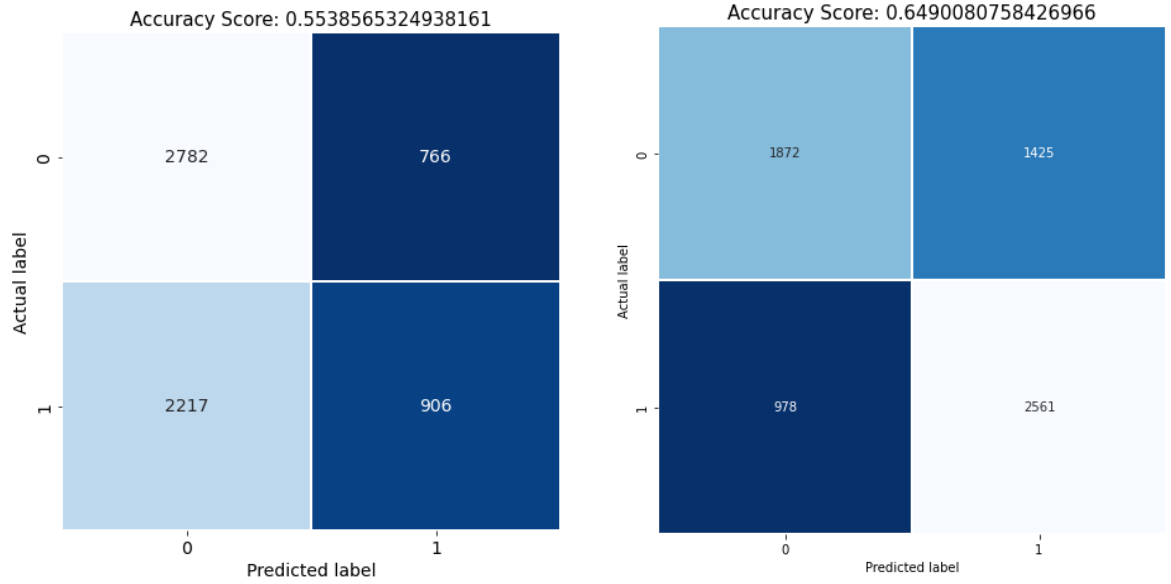
Table 4.1 displays the $Prec$, Rec and $F1$ values for the two labels 0 and 1 for each subject.

Regarding ROC curves, only the best one are displayed for space-saving reasons. Figures 4.6, 4.7, 4.8, 4.9 and 4.10 show the plots for subjects F0110, F1408, F1590, F2403 and M3309, respectively.



(a) Confusion matrix table resulting from the logistic regression on basal data for subject F0110.

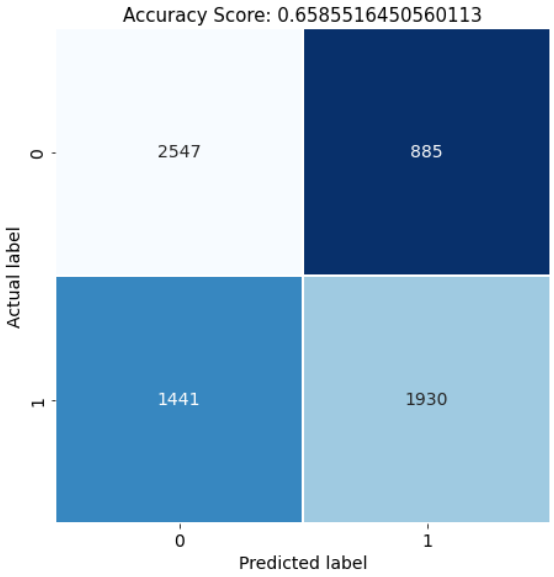
(b) Confusion matrix table resulting from the logistic regression on basal data for subject F0207.



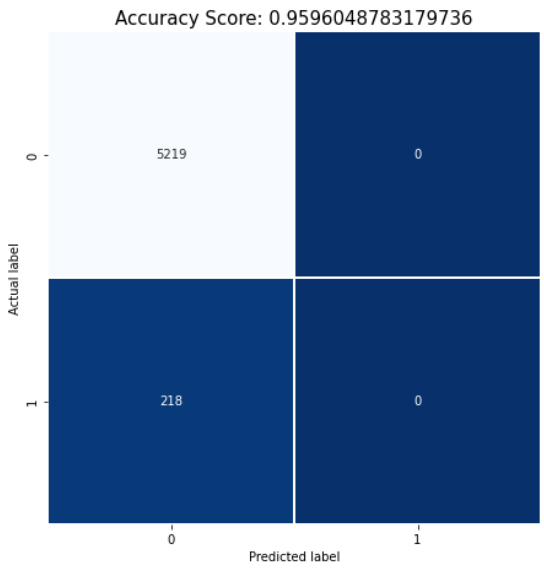
(c) Confusion matrix table resulting from the logistic regression on basal data for subject F1208.

(d) Confusion matrix table resulting from the logistic regression on basal data for subject F1408.

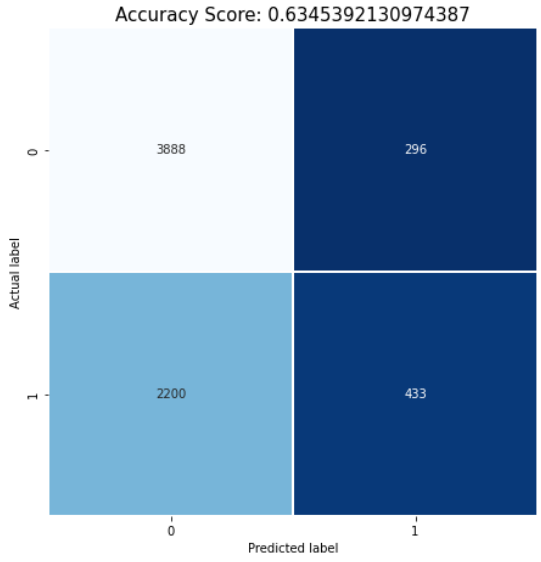
Figure 4.1: Confusion matrices resulting from the logistic regression on basal data for subjects F0110, F0207, F1208 and F1408.



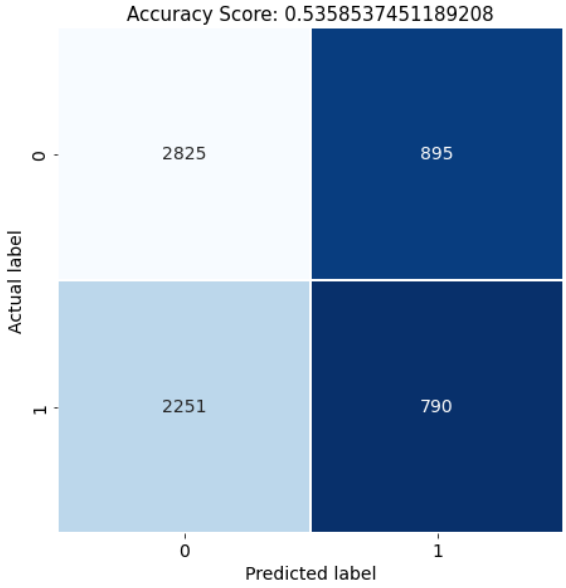
(a) Confusion matrix table resulting from the logistic regression on basal data for subject F1590.



(b) Confusion matrix table resulting from the logistic regression on basal data for subject F2403.

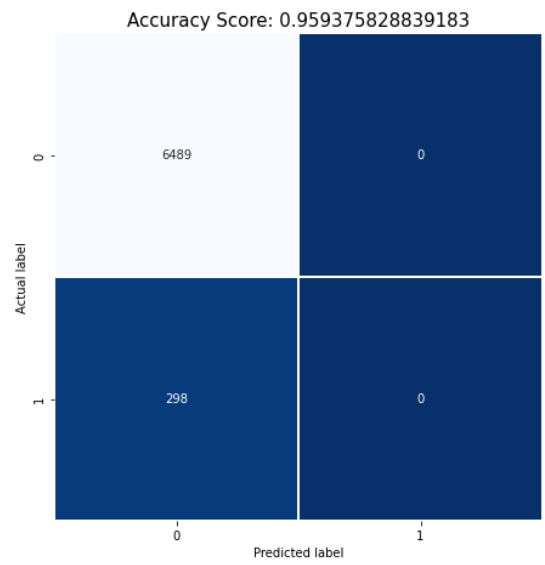


(c) Confusion matrix table resulting from the logistic regression on basal data for subject F2810.

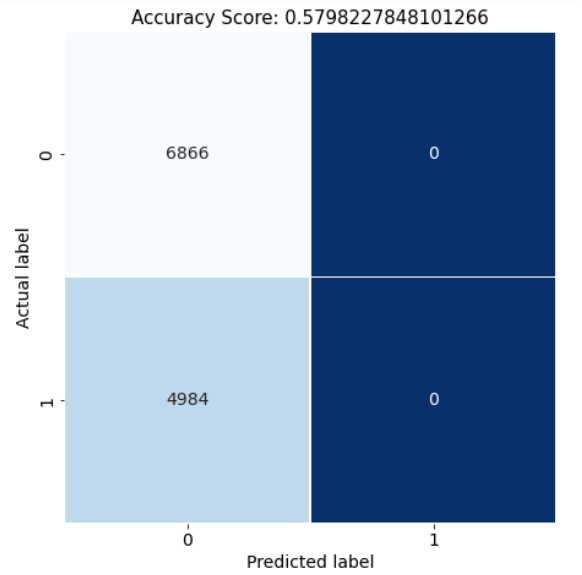


(d) Confusion matrix table resulting from the logistic regression on basal data for subject F2910.

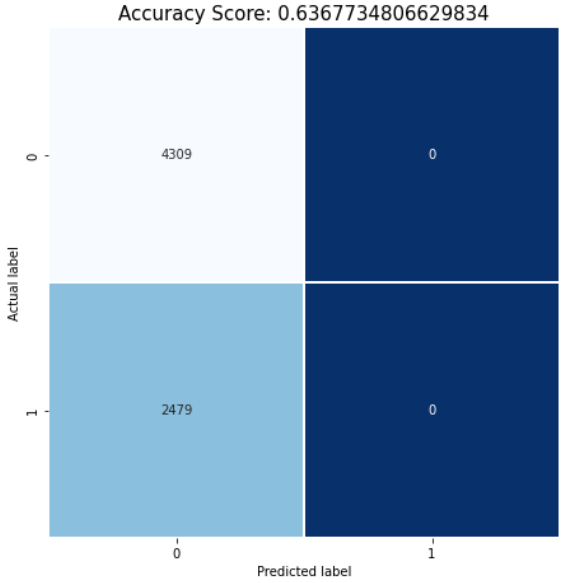
Figure 4.2: Confusion matrices resulting from the logistic regression on basal data for subjects F1590, F2403, F2810 and F2910.



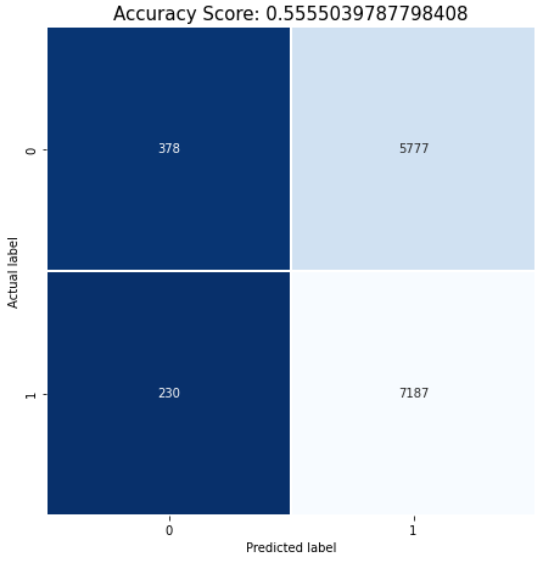
(a) Confusion matrix table resulting from the logistic regression on basal data for subject F3807.



(b) Confusion matrix table resulting from the logistic regression on basal data for subject F4301.

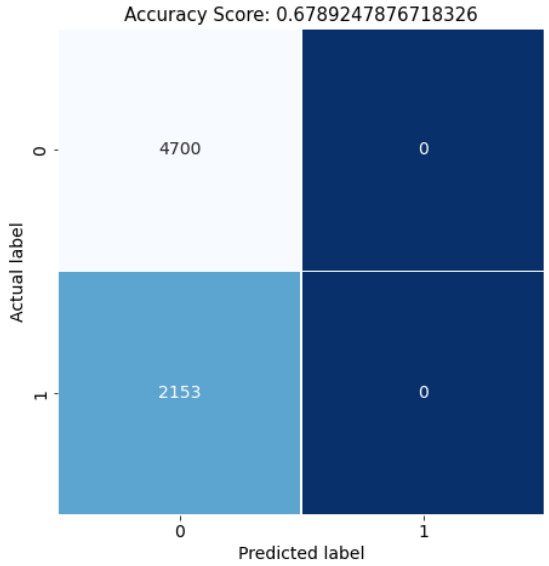


(c) Confusion matrix table resulting from the logistic regression on basal data for subject F4503.

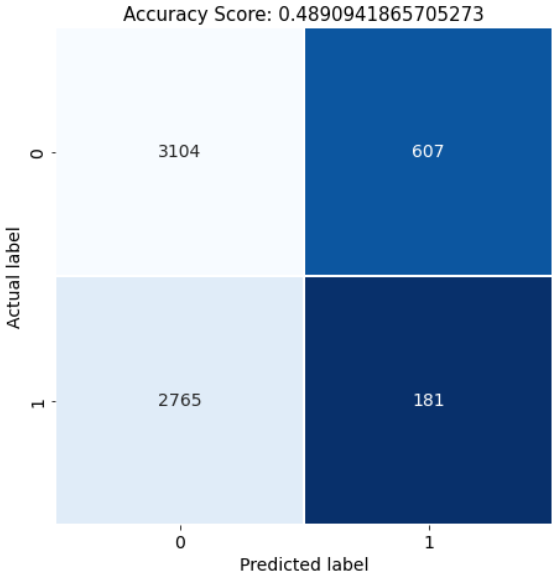


(d) Confusion matrix table resulting from the logistic regression on basal data for subject F4609.

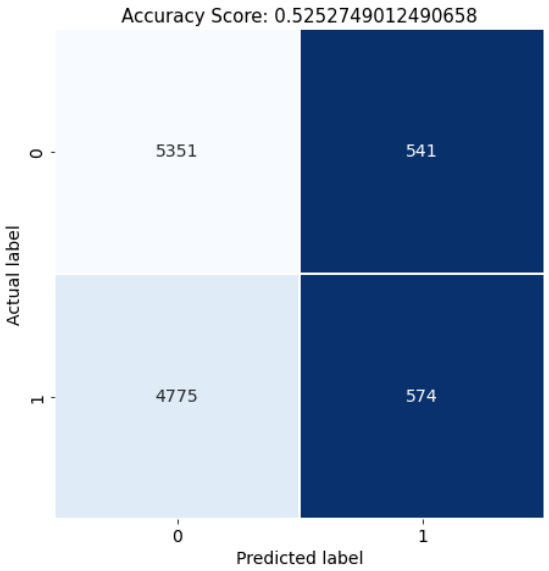
Figure 4.3: Confusion matrices resulting from the logistic regression on basal data for subjects F3807, F4301, F4503 and F4609.



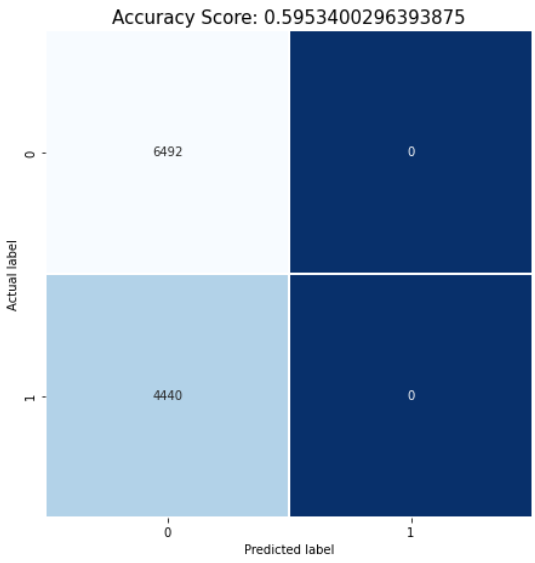
(a) Confusion matrix table resulting from the logistic regression on basal data for subject F4714.



(b) Confusion matrix table resulting from the logistic regression on basal data for subject M1006.

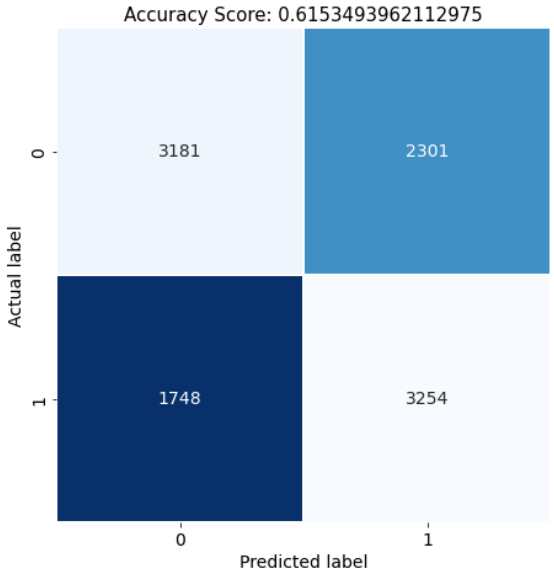


(c) Confusion matrix table resulting from the logistic regression on basal data for subject M2008.

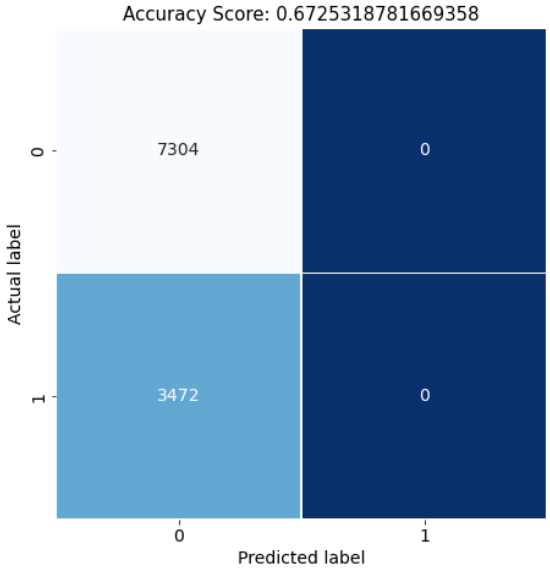


(d) Confusion matrix table resulting from the logistic regression on basal data for subject M2711.

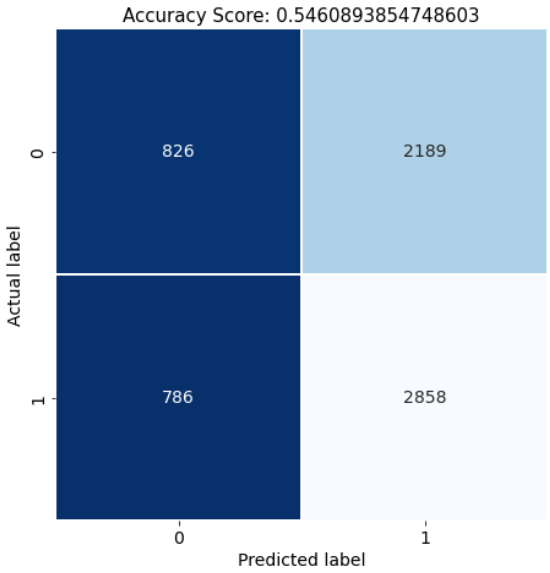
Figure 4.4: Confusion matrices resulting from the logistic regression on basal data for subjects F4714, M1006, M2008 and M2711.



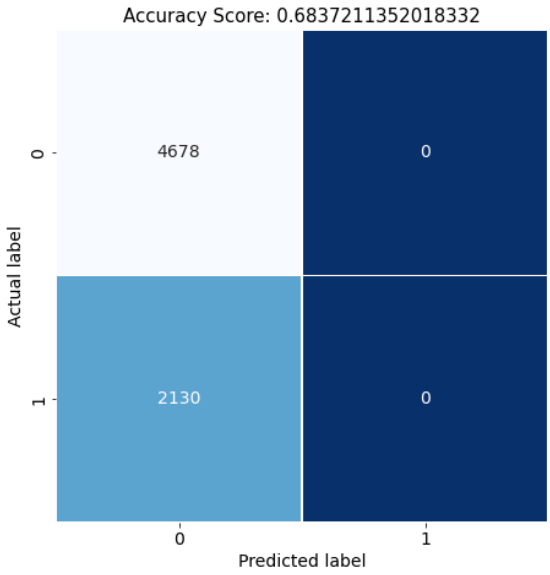
(a) Confusion matrix table resulting from the logistic regression on basal data for subject M3208.



(b) Confusion matrix table resulting from the logistic regression on basal data for subject M3309.



(c) Confusion matrix table resulting from the logistic regression on basal data for subject M4003.



(d) Confusion matrix table resulting from the logistic regression on basal data for subject M4409.

Figure 4.5: Confusion matrices resulting from the logistic regression on basal data for subjects M3208, M3309, M4003 and M4409.

Table 4.1: Metrics of the results coming from the logistic regression for all subjects. For each of them, the table reports the values for Precision ($Prec$), Recall (Rec) and F1-score ($F1$) for the two labels 0 and 1.

ID	Label	$Prec$	Rec	$F1$
F0110	0	0.59	0.62	0.61
	1	0.58	0.55	0.57
F0207	0	0.55	0.96	0.70
	1	0.21	0.01	0.02
F1208	0	0.56	0.78	0.65
	1	0.54	0.29	0.38
F1408	0	0.66	0.57	0.61
	1	0.64	0.72	0.68
F1590	0	0.64	0.74	0.69
	1	0.69	0.57	0.62
F2403	0	0.96	1.00	0.98
	1	0.00	0.00	0.00
F2810	0	0.64	0.93	0.76
	1	0.59	0.16	0.26
F2910	0	0.56	0.76	0.64
	1	0.47	0.26	0.33
F3807	0	0.96	1.00	0.98
	1	0.00	0.00	0.00
F4301	0	0.58	1.00	0.73
	1	0.00	0.00	0.00
F4503	0	0.63	1.00	0.78
	1	0.00	0.00	0.00
F4609	0	0.62	0.06	0.11
	1	0.55	0.97	0.71
F4714	0	0.69	1.00	0.81
	1	0.00	0.00	0.00
M1006	0	0.53	0.84	0.65
	1	0.23	0.06	0.10
M2008	0	0.53	0.91	0.67
	1	0.51	0.11	0.18
M2711	0	0.59	1.00	0.75
	1	0.00	0.00	0.00
M3208	0	0.65	0.58	0.61
	1	0.59	0.65	0.62
M3309	0	0.68	1.00	0.81
	1	0.00	0.00	0.00
M4003	0	0.51	0.27	0.36
	1	0.57	0.78	0.66
M4409	0	0.69	1.00	0.81
	1	0.00	0.00	0.00

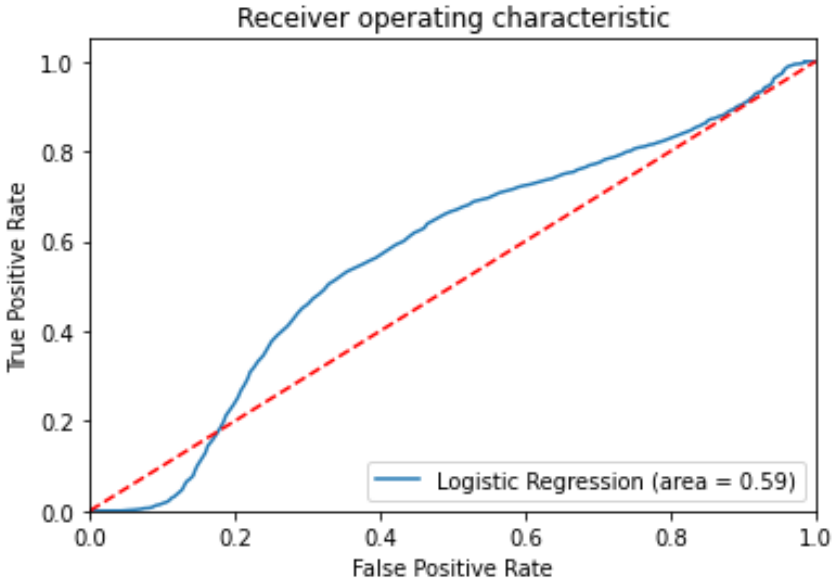


Figure 4.6: Plot for the ROC curve and area resulting from the logistic regression on basal data for subject F0110.

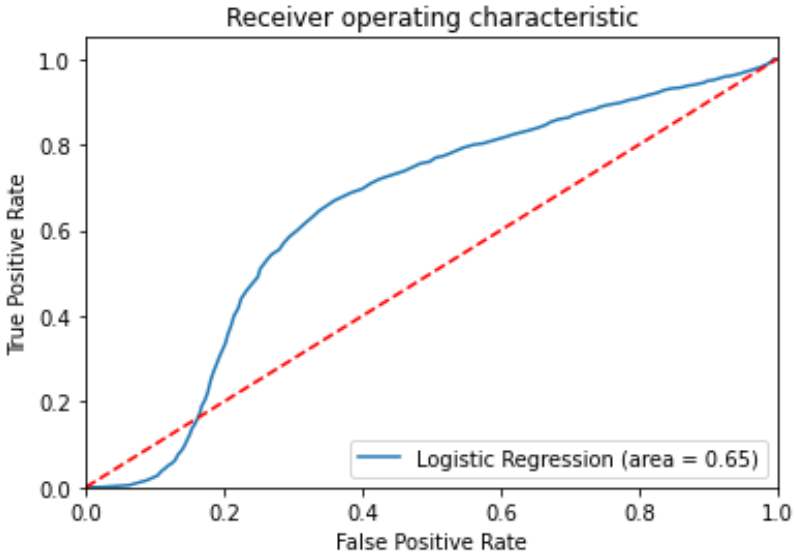


Figure 4.7: Plot for the ROC curve and area resulting from the logistic regression on basal data for subject F1408.

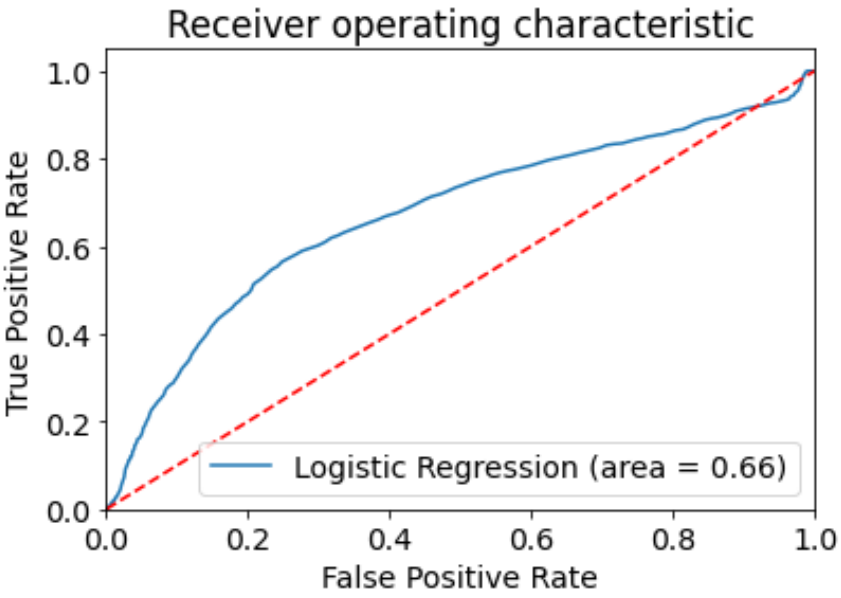


Figure 4.8: Plot for the ROC curve and area resulting from the logistic regression on basal data for subject F1590.

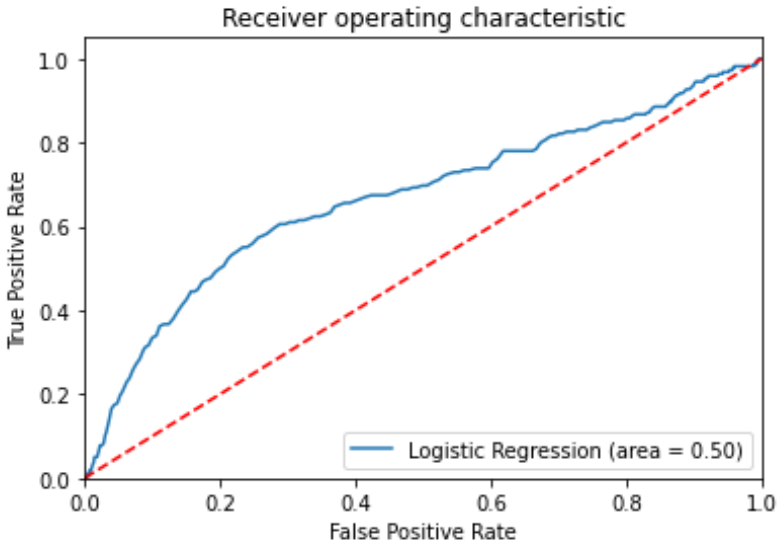


Figure 4.9: Plot for the ROC curve and area resulting from the logistic regression on basal data for subject F2403.

Figure 4.10: Plot for the ROC curve and area resulting from the logistic regression on basal data for subject M3309.

4.2 Random forest

This section reports the results coming from the random forest procedure defined in Subsection 3.2.2.

Since, as will be explained in Section 5.2, multiple approaches where the analysed data was changed to try to improve performances, not all of their results will be reported here. For instance, the random forest algorithm in which the used glycaemic files did not contain the added columns gave the least interesting results, while the ones for which the blood glucose levels of the 2 previous rows were added resulted in better values and the ones for which the added columns were 6 (until 30 minutes before) resulted in the best results. For this reason, only the results coming from the last approach are reported here.

The following list reports some of the metrics of the results coming from the random forest approach. It displays the MAE, MSE and RMSE values for each subject:

- F0110: MAE = 0.35, MSE = 0.19, RMSE = 0.44,
- F0207: MAE = 0.40, MSE = 0.21, RMSE = 0.46,
- F1208: MAE = 0.34, MSE = 0.18, RMSE = 0.43,
- F1408: MAE = 0.32, MSE = 0.17, RMSE = 0.42,
- F1590: MAE = 0.36, MSE = 0.20, RMSE = 0.44,
- F2403: MAE = 0.08, MSE = 0.04, RMSE = 0.20,
- F2810: MAE = 0.37, MSE = 0.17, RMSE = 0.41,

- F2910: MAE = 0.37, MSE = 0.20, RMSE = 0.45,
- F3807: MAE = 0.07, MSE = 0.04, RMSE = 0.19,
- F4301: MAE = 0.37, MSE = 0.20, RMSE = 0.44,
- F4503: MAE = 0.24, MSE = 0.13, RMSE = 0.36,
- F4609: MAE = 0.37, MSE = 0.21, RMSE = 0.46,
- F4714: MAE = 0.34, MSE = 0.18, RMSE = 0.42,
- M1006: MAE = 0.35, MSE = 0.18, RMSE = 0.43,
- M2008: MAE = 0.42, MSE = 0.23, RMSE = 0.48,
- M2711: MAE = 0.35, MSE = 0.19, RMSE = 0.44,
- M3208: MAE = 0.35, MSE = 0.19, RMSE = 0.45,
- M3309: MAE = 0.35, MSE = 0.19, RMSE = 0.43,
- M4003: MAE = 0.38, MSE = 0.20, RMSE = 0.45,
- M4409: MAE = 0.29, MSE = 0.15, RMSE = 0.40.

Figures 4.11, 4.12, 4.13, 4.14 and 4.15 display the confusion matrix of each subject. The latter is reported in the subfigures' captions. The number on the first square (first row and column) represents the true negatives (TN), the second square (first row and second column) represents the false positives (FP), the third square (second row and first column) represents the false negatives (FN) and the fourth square represents the true positives (TP). Above the matrix, the accuracy score is reported.

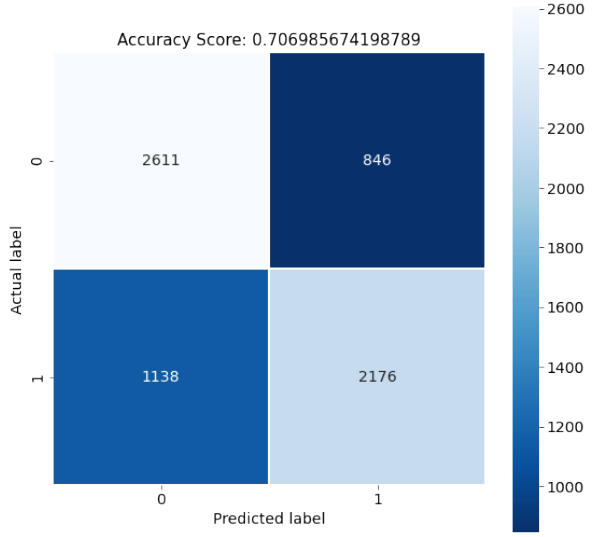
Table 4.2 displays the $Prec$, Rec and $F1$ values for the two labels 0 and 1 for each subject.

4.3 Multivariate linear regression

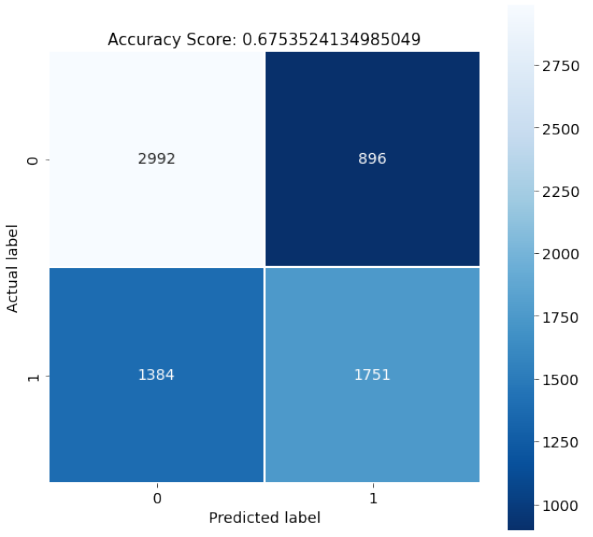
This section reports the results coming from the multivariate linear regression procedure defined in Subsection 3.2.3.

Since, as will be explained in Section 5.3, some approaches where the analysed data was changed to try to improve performances did not give the hoped results, the latter will not be reported here. For instance, the logistic regression in which the glycaemic columns reporting the blood glucose levels in the 6 previous rows were added resulted in non-existent or, at least, a non-significant difference for the residuals plots, for coefficients' p-values (for the ones already existing for the same approach applied to files reporting only 2 of the previous rows; of course, the new coefficients had new p-values) and R^2. The only thing that changed was F-stat: for example, it was 1769 (instead of 3867) for subject F0110 and 1276 (instead of 2942) for subject F0207.

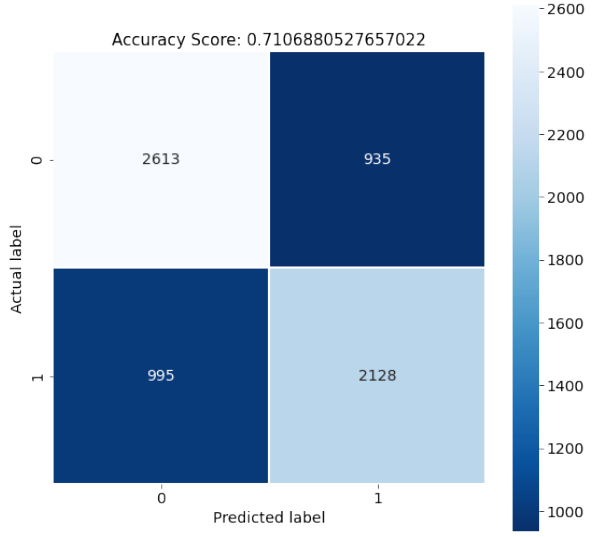
Tables 4.3, 4.4, 4.5 and 4.6 report the coefficients (the intercept and the slope coefficients) of Eq. 3.9 of the multivariate linear regression for each subject.



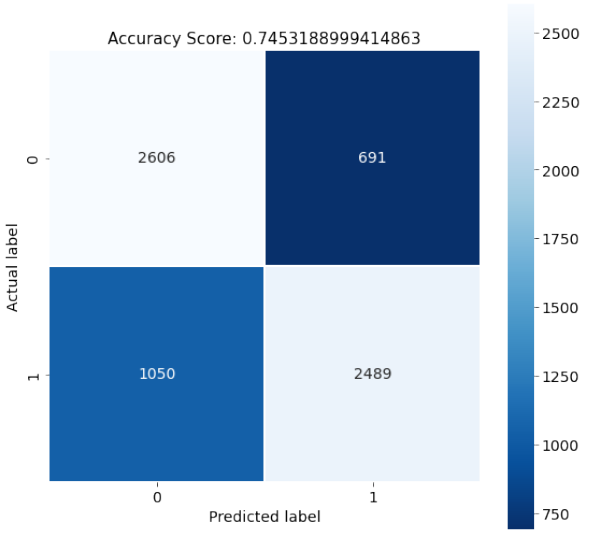
(a) Confusion matrix table resulting from the random forest approach on basal data for subject F0110.



(b) Confusion matrix table resulting from the random forest approach on basal data for subject F0207.

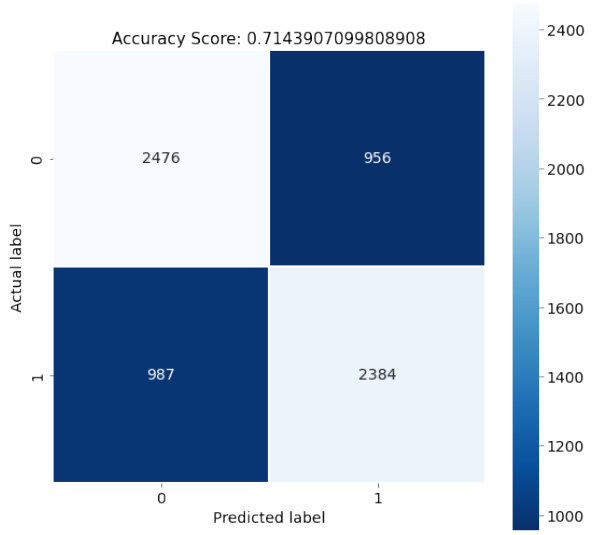


(c) Confusion matrix table resulting from the random forest approach on basal data for subject F1208.

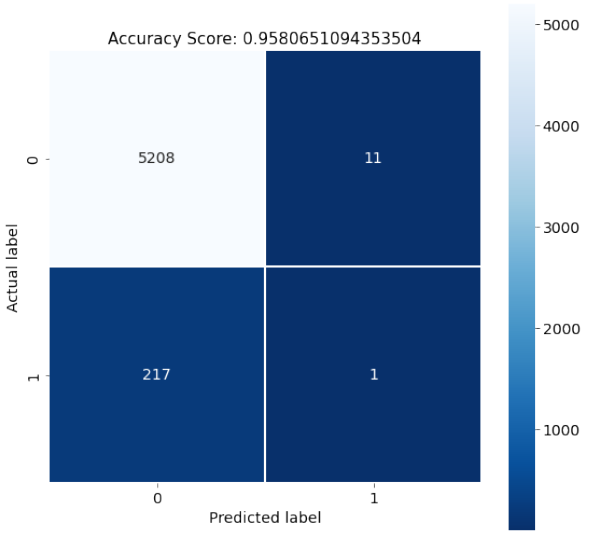


(d) Confusion matrix table resulting from the random forest approach on basal data for subject F1408.

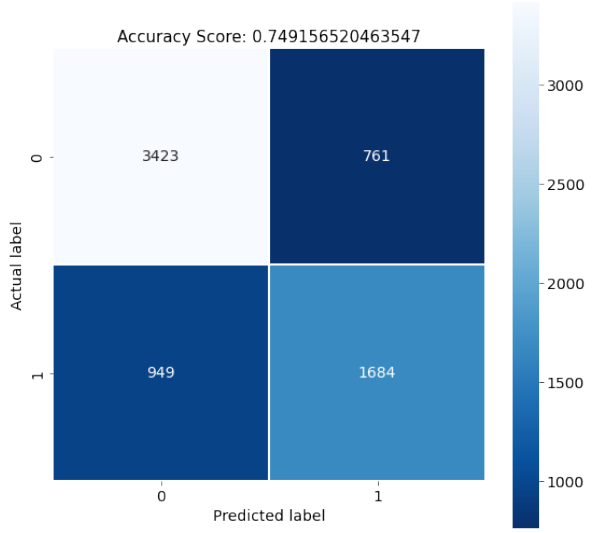
Figure 4.11: Confusion matrices resulting from the random forest approach on basal data for subjects F0110, F0207, F1208 and F1408.



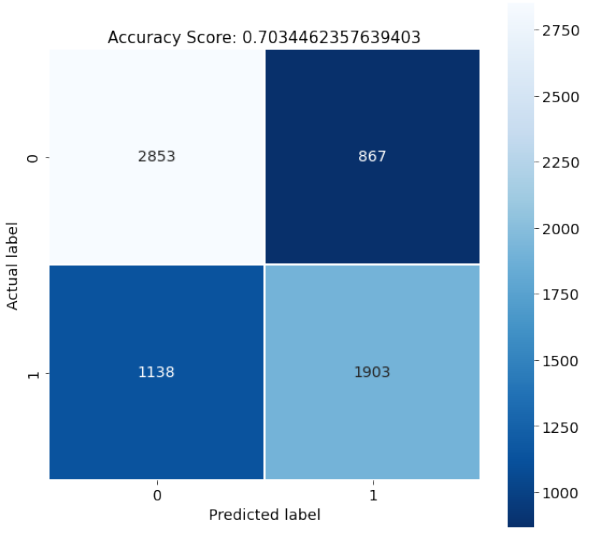
(a) Confusion matrix table resulting from the random forest approach on basal data for subject F1590.



(b) Confusion matrix table resulting from the random forest approach on basal data for subject F2403.

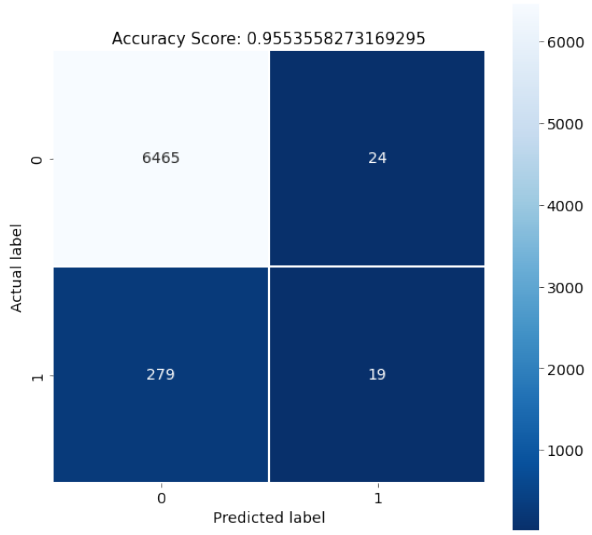


(c) Confusion matrix table resulting from the random forest approach on basal data for subject F2810.

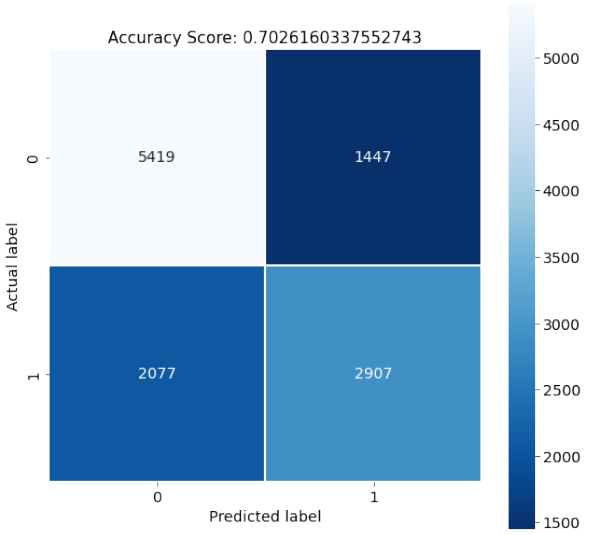


(d) Confusion matrix table resulting from the random forest approach on basal data for subject F2910.

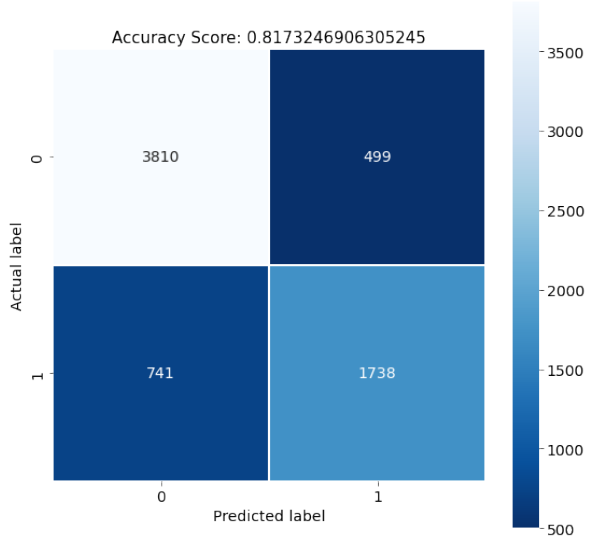
Figure 4.12: Confusion matrices resulting from the random forest approach on basal data for subjects F1590, F2403, F2810 and F2910.



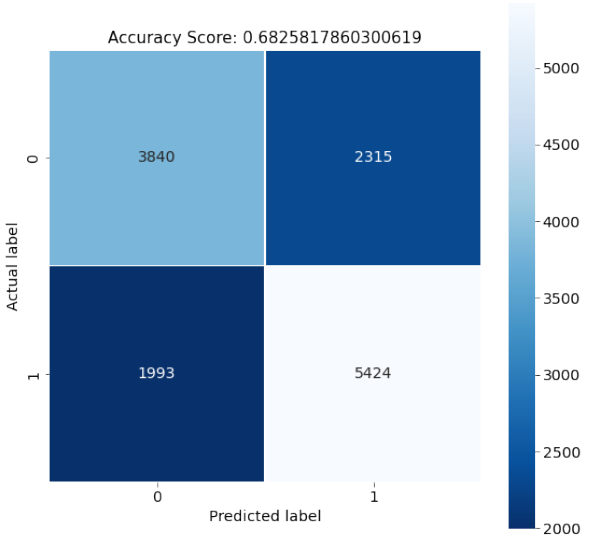
(a) Confusion matrix table resulting from the random forest approach on basal data for subject F3807.



(b) Confusion matrix table resulting from the random forest approach on basal data for subject F4301.

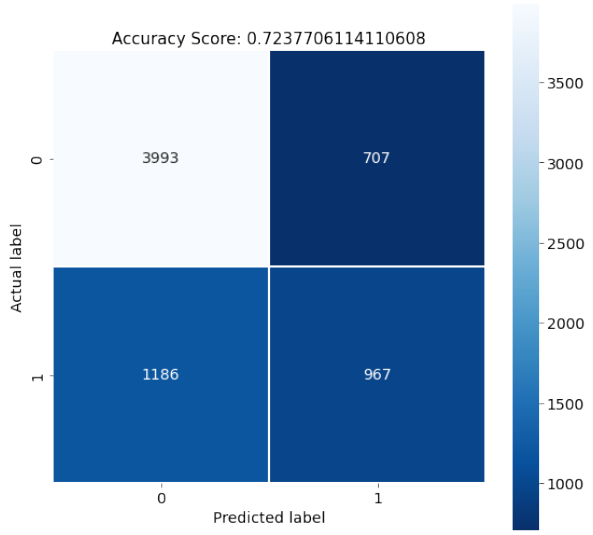


(c) Confusion matrix table resulting from the random forest approach on basal data for subject F4503.

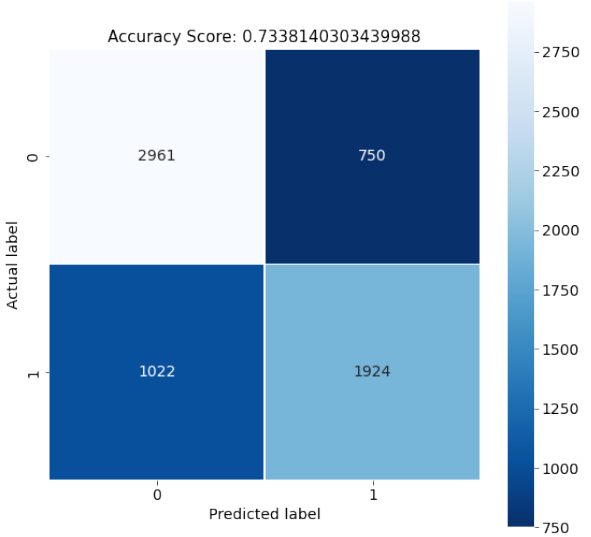


(d) Confusion matrix table resulting from the random forest approach on basal data for subject F4609.

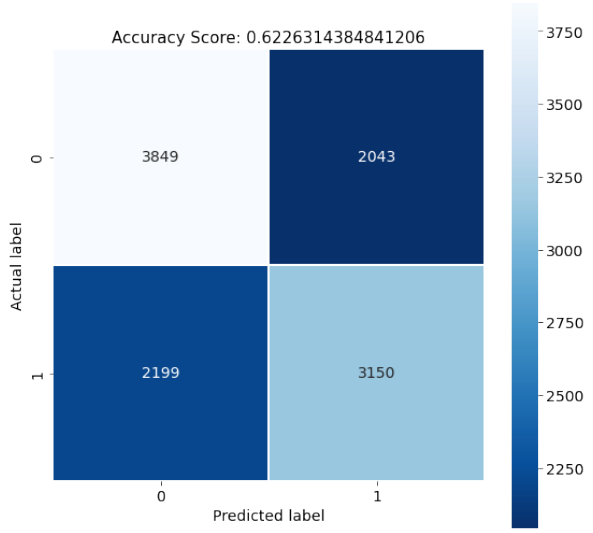
Figure 4.13: Confusion matrices resulting from the random forest approach on basal data for subjects F3807, F4301, F4503 and F4609.



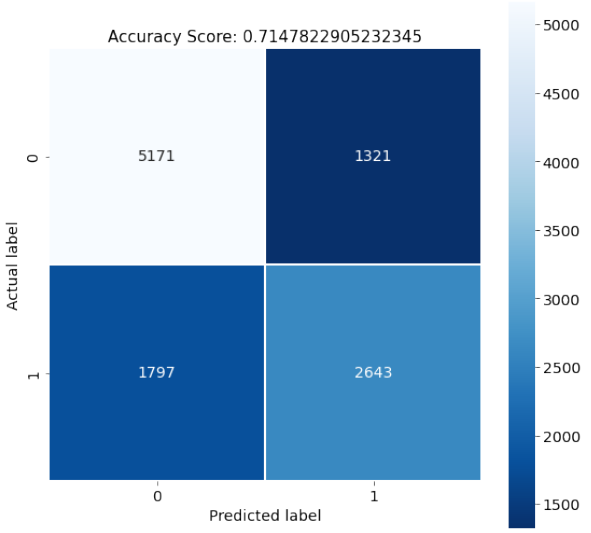
(a) Confusion matrix table resulting from the random forest approach on basal data for subject F4714.



(b) Confusion matrix table resulting from the random forest approach on basal data for subject M1006.

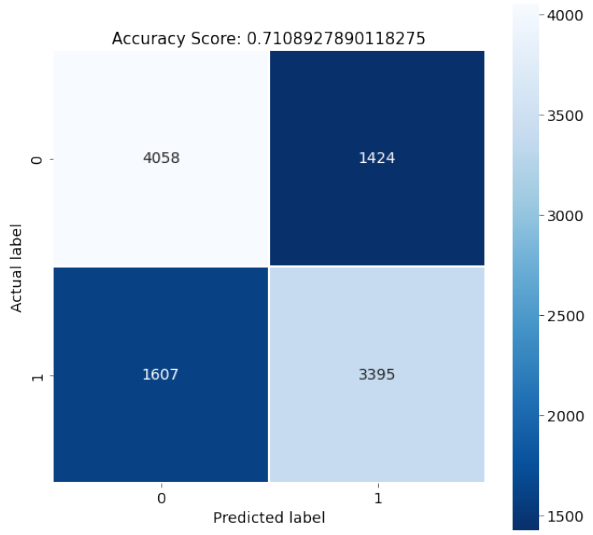


(c) Confusion matrix table resulting from the random forest approach on basal data for subject M2008.

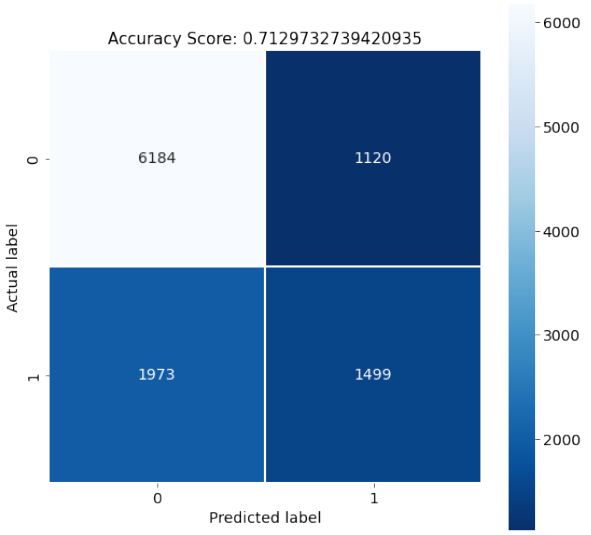


(d) Confusion matrix table resulting from the random forest approach on basal data for subject M2711.

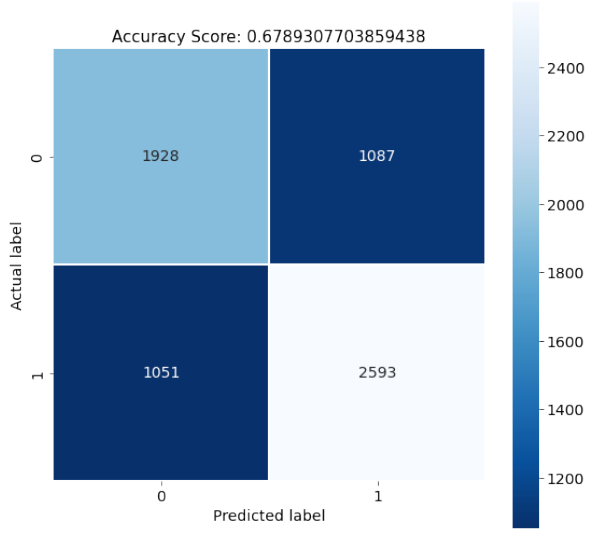
Figure 4.14: Confusion matrices resulting from the random forest approach on basal data for subjects F4714, M1006, M2008 and M2711.



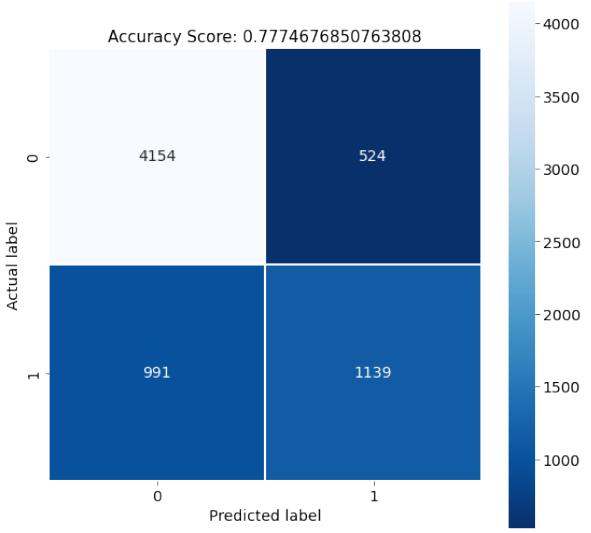
(a) Confusion matrix table resulting from the random forest approach on basal data for subject M3208.



(b) Confusion matrix table resulting from the random forest approach on basal data for subject M3309.



(c) Confusion matrix table resulting from the random forest approach on basal data for subject M4003.



(d) Confusion matrix table resulting from the random forest approach on basal data for subject M4409.

Figure 4.15: Confusion matrices resulting from the random forest approach on basal data for subjects M3208, M3309, M4003 and M4409.

Table 4.2: Metrics of the results coming from the random forest approach for all subjects. For each of them, the table reports the values for Precision ($Prec$), Recall (Rec) and F1-score ($F1$) for the two labels 0 and 1.

ID	Label	$Prec$	Rec	$F1$
F0110	0	0.70	0.76	0.72
	1	0.72	0.66	0.69
F0207	0	0.68	0.77	0.72
	1	0.66	0.56	0.61
F1208	0	00.72	0.74	0.73
	1	0.69	0.68	0.69
F1408	0	0.71	0.79	0.75
	1	0.78	0.70	0.74
F1590	0	0.71	0.72	0.72
	1	0.71	0.71	0.71
F2403	0	0.96	1.00	0.98
	1	0.08	0.00	0.01
F2810	0	0.78	0.82	0.80
	1	0.69	0.64	0.66
F2910	0	0.71	0.77	0.74
	1	0.69	0.63	0.65
F3807	0	0.96	1.00	0.98
	1	0.44	0.06	0.11
F4301	0	0.72	0.79	0.75
	1	0.67	0.58	0.62
F4503	0	0.84	0.88	0.86
	1	0.78	0.70	0.74
F4609	0	0.66	0.62	0.64
	1	0.70	0.73	0.72
F4714	0	0.77	0.85	0.81
	1	0.58	0.45	0.51
M1006	0	0.74	0.80	0.77
	1	0.72	0.65	0.68
M2008	0	0.64	0.65	0.64
	1	0.61	0.59	0.60
M2711	0	0.74	0.80	0.77
	1	0.67	0.60	0.63
M3208	0	0.72	0.74	0.73
	1	0.70	0.68	0.69
M3309	0	0.65	0.64	0.64
	1	0.70	0.71	0.71
M4003	0	0.51	0.27	0.36
	1	0.57	0.78	0.66
M4409	0	0.81	0.89	0.85
	1	0.68	0.53	0.60

Tables 4.7, 4.8, 4.9 and 4.10 report the p-value of each coefficient (the intercept and the slope coefficients) of Eq. 3.9 of the multivariate linear regression for each subject.

The following list reports the values of R^2 and F-stat coming from the multivariate linear regression. Both values are displayed for each subject:

- F0110: $R^2 = 0.515$, F-stat = 3867,
- F0207: $R^2 = 0.462$, F-stat = 2942,
- F1208: $R^2 = 0.515$, F-stat = 3711,
- F1408: $R^2 = 0.477$, F-stat = 3543,
- F1590: $R^2 = 0.321$, F-stat = 1737,
- F2403: $R^2 = 0.013$, F-stat = 3294,
- F2810: $R^2 = 0.463$, F-stat = 2515,
- F2910: $R^2 = 0.342$, F-stat = 1786,
- F3807: $R^2 = 0.025$, F-stat = 7979,
- F4301: $R^2 = 0.436$, F-stat = 4271,
- F4503: $R^2 = 0.568$, F-stat = 3604,
- F4609: $R^2 = 0.449$, F-stat = 6655,
- F4714: $R^2 = 0.343$, F-stat = 1276,
- M1006: $R^2 = 0.443$, F-stat = 2647,
- M2008: $R^2 = 0.333$, F-stat = 2978,
- M2711: $R^2 = 0.407$, F-stat = 3367,
- M3208: $R^2 = 0.515$, F-stat = 6030,
- M3309: $R^2 = 0.296$, F-stat = 1651,
- M4003: $R^2 = 0.448$, F-stat = 3238,
- M4409: $R^2 = 0.304$, F-stat = 1045.

Figures 4.16, 4.17, 4.18, 4.19, 4.20, 4.21, 4.22, 4.23, 4.24, 4.25, 4.26, 4.27, 4.28, 4.29, 4.30, 4.31, 4.32, 4.33, 4.34 and 4.35 display the plot of the residuals of each subject. The latter is reported in the figure's caption.

Table 4.3: Coefficients of the multivariate linear equation for each subject; part 1.

ID	**Coefficients**
F0110	-0.435 $1.944*e^{-2}$ $-1.735*e^{-18}$ $5.665*e^{-3}$ $-1.296*e^{-2}$
F0207	-0.917 $1.327*e^{-2}$ $-2.331*e^{-18}$ $3.946*e^{-3}$ $-4.092*e^{-3}$
F1208	-0.097 $9.347*e^{-3}$ $2.602*e^{-18}$ $2.034*e^{-3}$ $-4.993*e^{-3}$
F1408	-1.317 $2.608*e^{-2}$ $3.469*e^{-18}$ $9.346*e^{-3}$ $-1.313*e^{-2}$
F1590	0.116 $5.715*e^{-3}$ $-1.735*e^{-18}$ $1.467*e^{-3}$ $-3.271*e^{-3}$

Table 4.4: Coefficients of the multivariate linear equation for each subject; part 2.

ID	**Coefficients**
F2403	0.848 $3.014*e^{-4}$ $-5.421*e^{-20}$ $-4.704*e^{-5}$ $-1.094*e^{-4}$
F2810	-0.199 $3.892*e^{-3}$ 0.000 $1.027*e^{-3}$ $-9.284*e^{-4}$
F2910	-0.441 $1.613*e^{-2}$ $-1.301*e^{-18}$ $4.036*e^{-3}$ $-7.063*e^{-3}$
F3807	0.811 $1.181*e^{-4}$ $-3.388*e^{-20}$ $-2.598*e^{-6}$ $-8.217*e^{-5}$
F4301	-0.062 $9.781*e^{-3}$ $-1.464*e^{-17}$ $2.932*e^{-3}$ $-4.286*e^{-3}$

Table 4.5: Coefficients of the multivariate linear equation for each subject; part 3.

ID	**Coefficients**
F4503	-0.219 $8.925*e^{-3}$ $-2.927*e^{-18}$ $2.901*e^{-3}$ $-4.210*e^{-3}$
F4609	-0.571 $1.802*e^{-2}$ $-8.674*e^{-19}$ $6.089*e^{-4}$ $-8.200*e^{-3}$
F4714	0.023 $3.599*e^{-3}$ $-4.337*e^{-19}$ $5.285*e^{-4}$ $-1.493*e^{-3}$
M1006	-0.410 $1.278*e^{-2}$ $-4.337*e^{-19}$ $1.513*^{-3}$ $-5.625*e^{-3}$
M2008	0.333 $2.375*e^{-2}$ $-8.674*e^{-18}$ $1.154*e^{-3}$ $-1.259*e^{-2}$

Table 4.6: Coefficients of the multivariate linear equation for each subject; part 4.

ID	**Coefficients**
M2711	-0.310 $1.142*e^{-2}$ $4.337*e^{-19}$ $9.325*e^{-4}$ $-4.608*e^{-3}$
M3208	-0.163 $1.553*e^{-2}$ $4.770*e^{-18}$ $2.487*e^{-3}$ $-7.696*e^{-3}$
M3309	-0.075 $8.404*e^{-3}$ $-4.337*e^{-19}$ $7.477*e-^{-4}$ $-3.537*e^{-3}$
M4003	0.233 $9.495*e^{-3}$ $3.469*e^{-18}$ $2.491*e^{-3}$ $-5.598*e^{-3}$
M4409	-0.186 $5.690*e^{-3}$ $-5.421*e-^{-19}$ $1.486*e^{-3}$ $-9.209*e^{-4}$

Table 4.7: P-value of each coefficient of the multivariate linear equation for each subject; part 1.

ID	**Coefficient**	**p-value**
F0110	β_0	0.000
	β_1	0.000
	β_2	0.000
	β_3	0.000
	β_4	0.000
F0207	β_0	0.000
	β_1	0.000
	β_2	0.000
	β_3	0.000
	β_4	0.000
F1208	β_0	0.000
	β_1	0.000
	β_2	0.000
	β_3	0.000
	β_4	0.000
F1408	β_0	0.000
	β_1	0.000
	β_2	0.000
	β_3	0.000
	β_4	0.000
F1590	β_0	0.000
	β_1	0.000
	β_2	0.000
	β_3	0.000
	β_4	0.000

Table 4.8: P-value of each coefficient of the multivariate linear equation for each subject; part 2.

ID	**Coefficient**	**p-value**
F2403	β_0	0.000
	β_1	0.005
	β_2	0.002
	β_3	0.710
	β_4	0.309
F2810	β_0	0.000
	β_1	0.000
	β_2	0.000
	β_3	0.000
	β_4	0.000
F2910	β_0	0.000
	β_1	0.000
	β_2	0.000
	β_3	0.000
	β_4	0.000
F3807	β_0	0.000
	β_1	0.000
	β_2	0.000
	β_3	0.933
	β_4	0.005
F4301	β_0	0.000
	β_1	0.000
	β_2	0.002
	β_3	0.000
	β_4	0.000

Table 4.9: P-value of each coefficient of the multivariate linear equation for each subject; part 3.

ID	**Coefficient**	**p-value**
F4503	β_0	0.000
	β_1	0.000
	β_2	0.000
	β_3	0.000
	β_4	0.000
F4609	β_0	0.000
	β_1	0.000
	β_2	0.000
	β_3	0.188
	β_4	0.000
F4714	β_0	0.006
	β_1	0.000
	β_2	0.008
	β_3	0.001
	β_4	0.000
M1006	β_0	0.000
	β_1	0.000
	β_2	0.000
	β_3	0.000
	β_4	0.000
M2008	β_0	0.000
	β_1	0.000
	β_2	0.000
	β_3	0.288
	β_4	0.000

Table 4.10: P-value of each coefficient of the multivariate linear equation for each subject; part 4.

ID	**Coefficient**	**p-value**
M2711	β_0	0.000
	β_1	0.000
	β_2	0.025
	β_3	0.000
	β_4	0.000
M3208	β_0	0.000
	β_1	0.000
	β_2	0.000
	β_3	0.000
	β_4	0.000
M3309	β_0	0.000
	β_1	0.000
	β_2	0.000
	β_3	0.053
	β_4	0.000
M4003	β_0	0.000
	β_1	0.000
	β_2	0.000
	β_3	0.000
	β_4	0.000
M4409	β_0	0.000
	β_1	0.000
	β_2	0.000
	β_3	0.000
	β_4	0.002

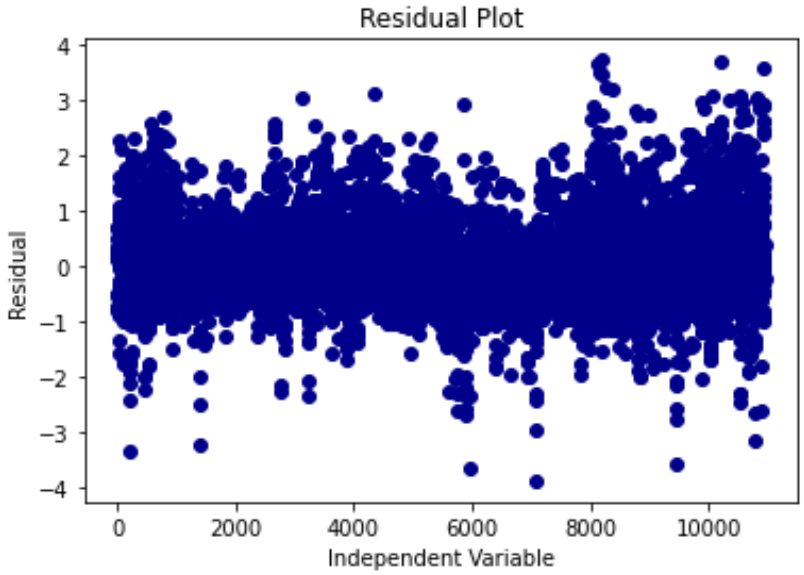


Figure 4.16: Plot of the residuals resulting from the multivariate linear regression for subject F0110.

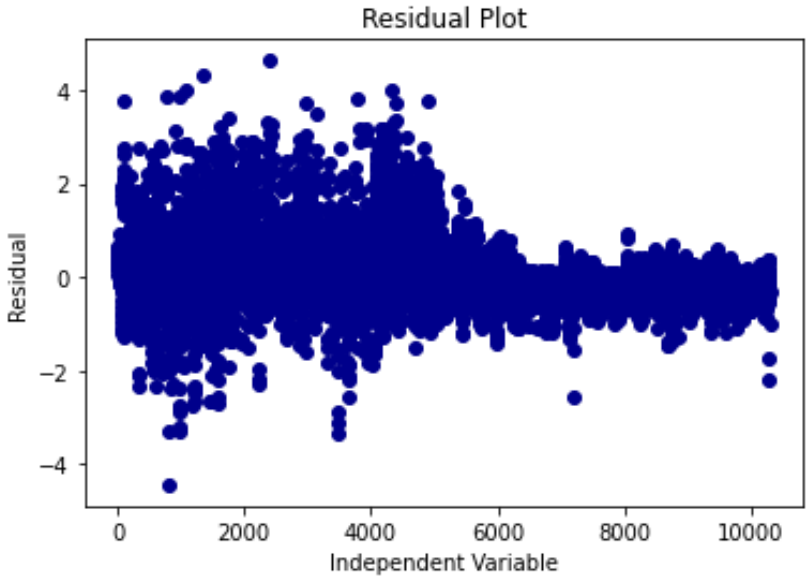


Figure 4.17: Plot of the residuals resulting from the multivariate linear regression for subject F0207.

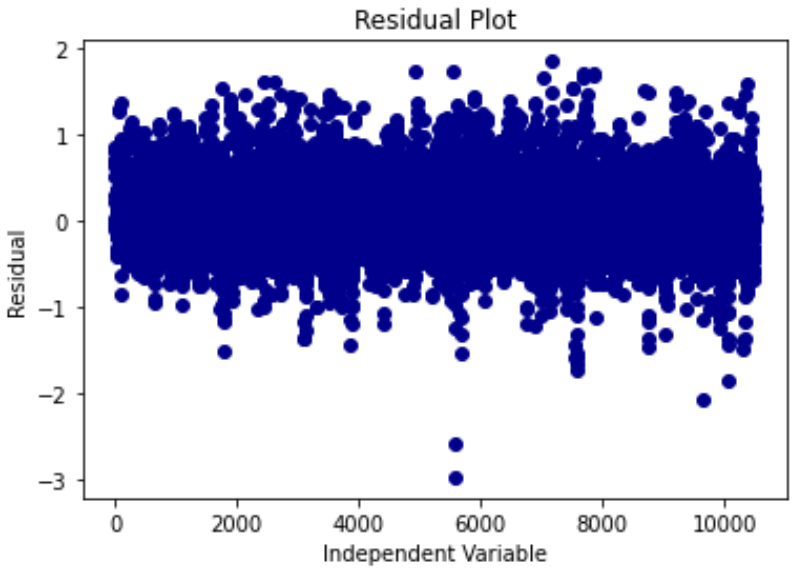


Figure 4.18: Plot of the residuals resulting from the multivariate linear regression for subject F1208.

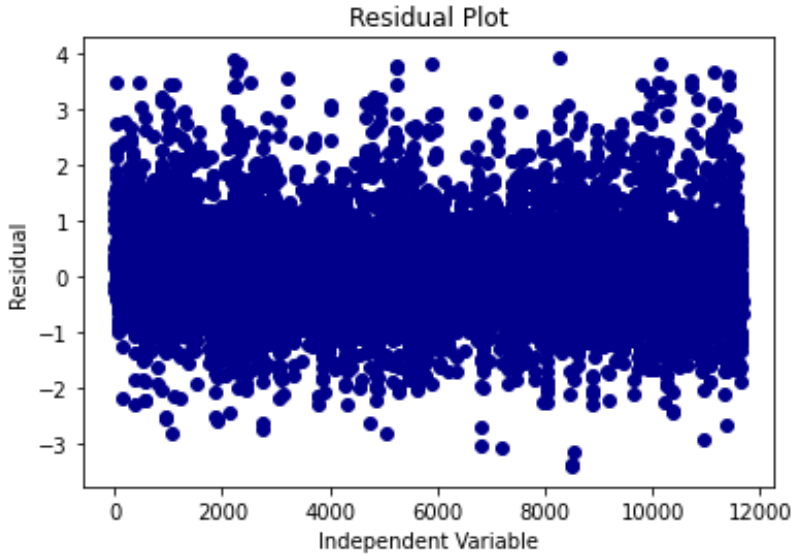


Figure 4.19: Plot of the residuals resulting from the multivariate linear regression for subject F1408.

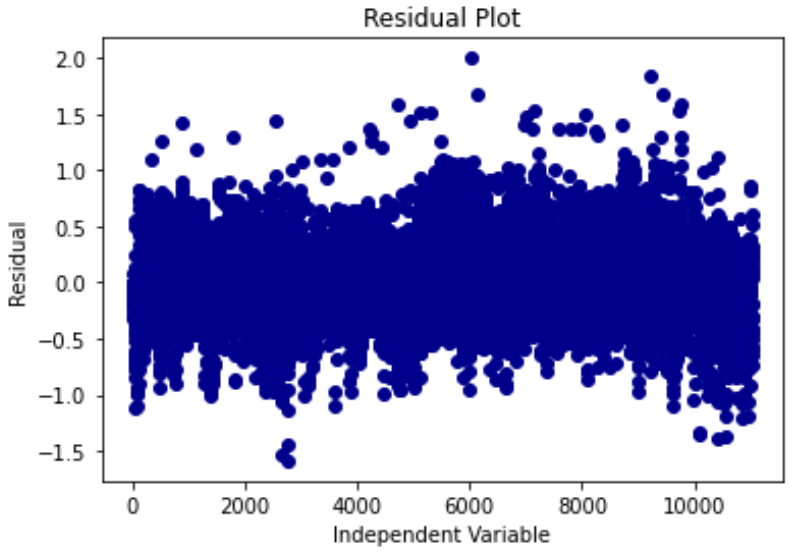


Figure 4.20: Plot of the residuals resulting from the multivariate linear regression for subject F1590.

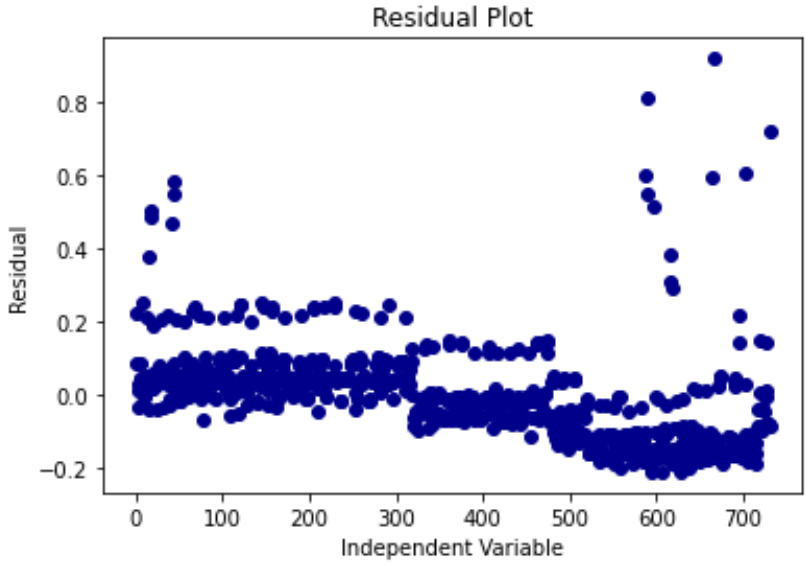


Figure 4.21: Plot of the residuals resulting from the multivariate linear regression for subject F2403.

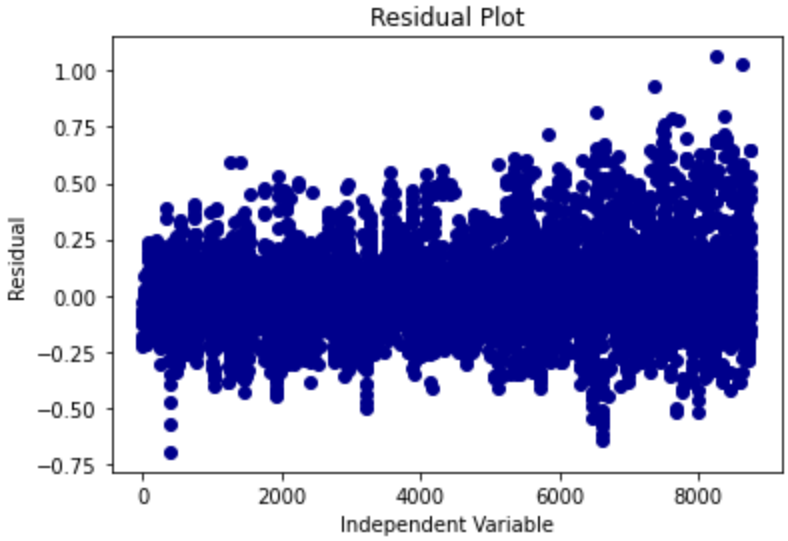


Figure 4.22: Plot of the residuals resulting from the multivariate linear regression for subject F2810.

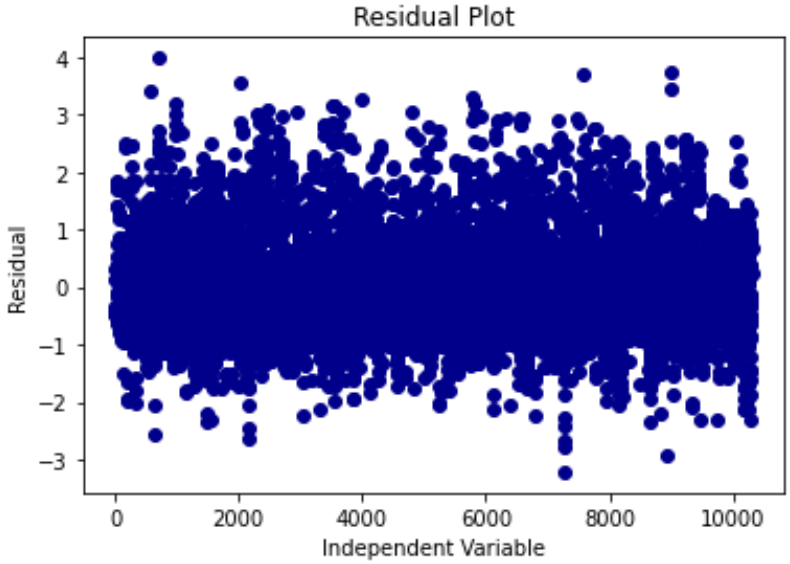


Figure 4.23: Plot of the residuals resulting from the multivariate linear regression for subject F2910.

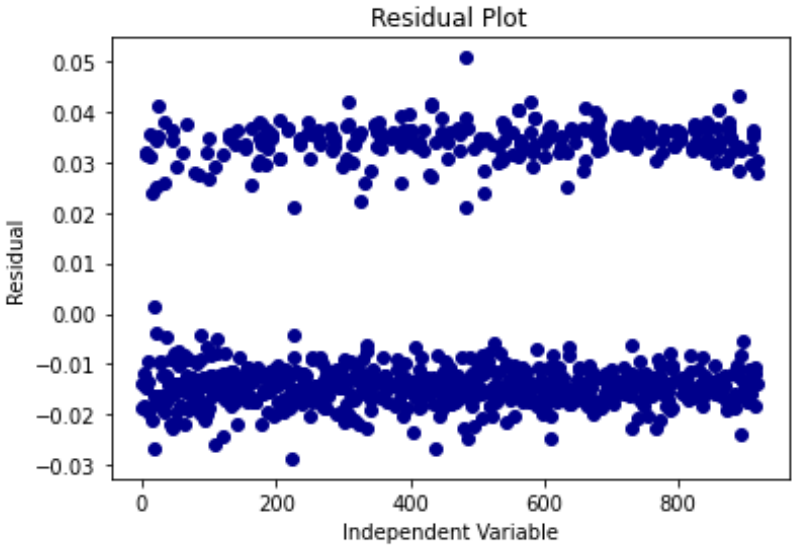


Figure 4.24: Plot of the residuals resulting from the multivariate linear regression for subject F3807.

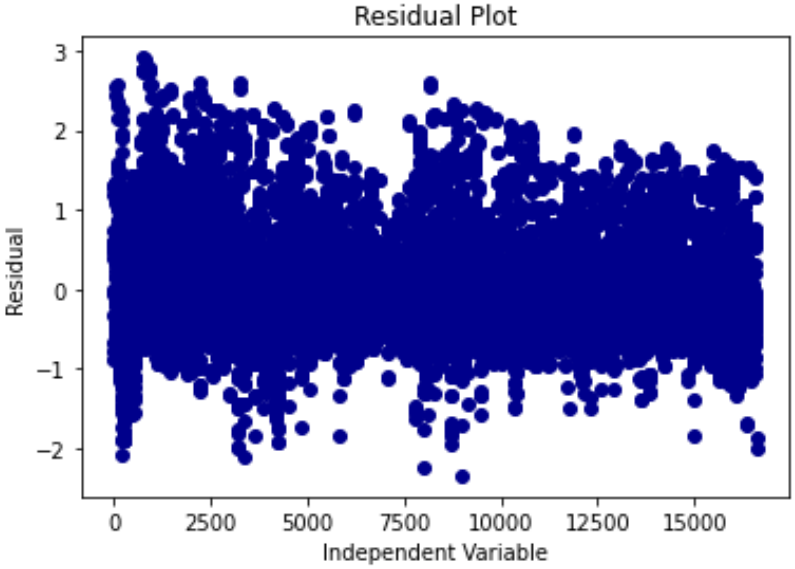


Figure 4.25: Plot of the residuals resulting from the multivariate linear regression for subject F4301.

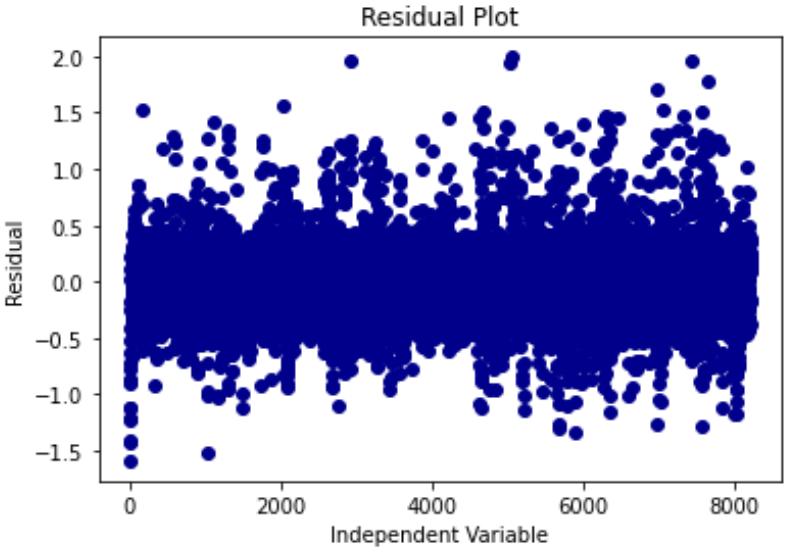


Figure 4.26: Plot of the residuals resulting from the multivariate linear regression for subject F4503.

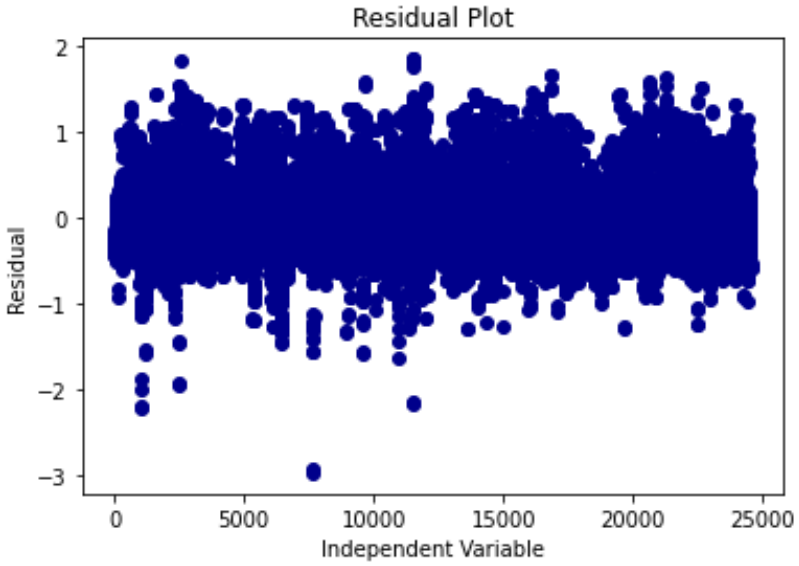


Figure 4.27: Plot of the residuals resulting from the multivariate linear regression for subject F4609.

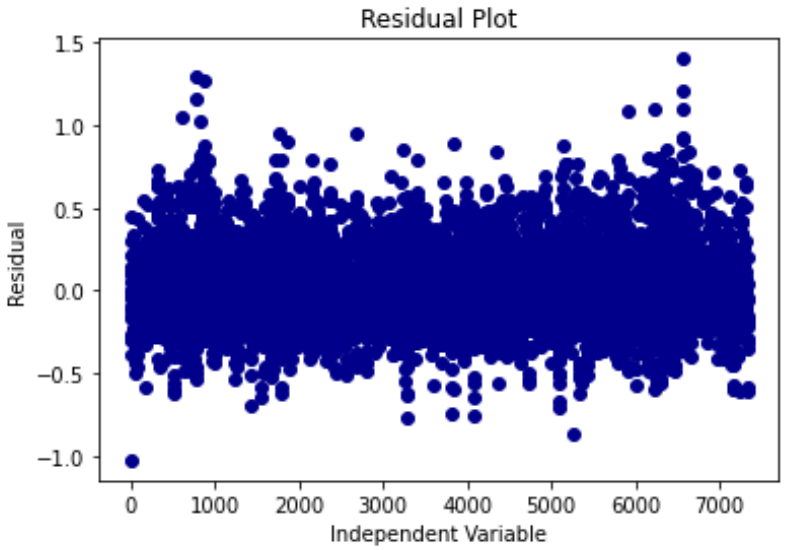


Figure 4.28: Plot of the residuals resulting from the multivariate linear regression for subject F4714.

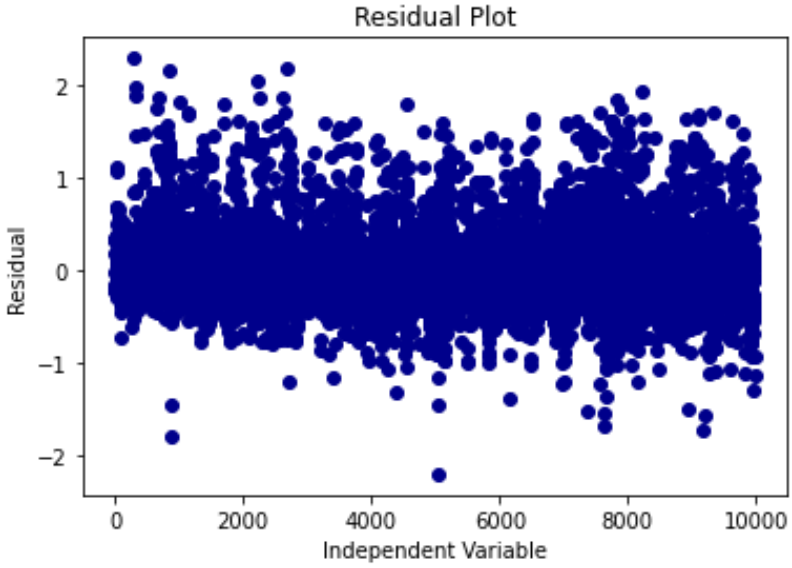


Figure 4.29: Plot of the residuals resulting from the multivariate linear regression for subject M1006.

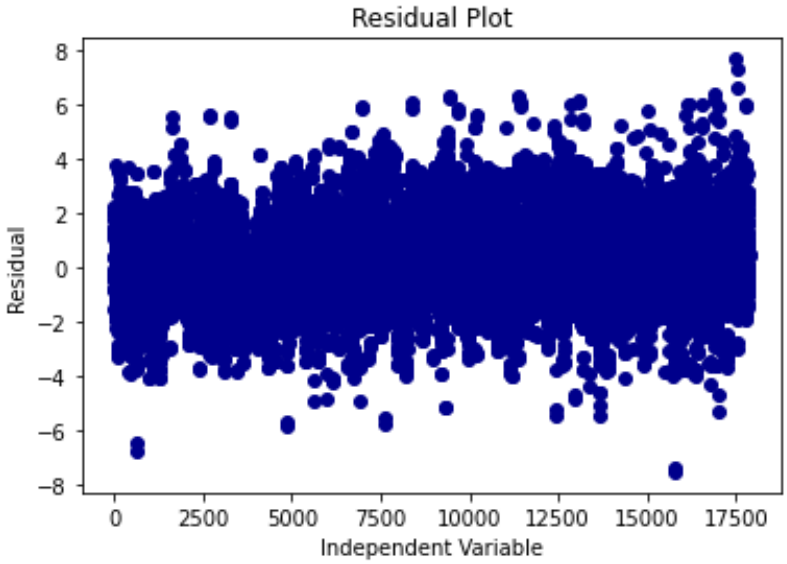


Figure 4.30: Plot of the residuals resulting from the multivariate linear regression for subject M2008.

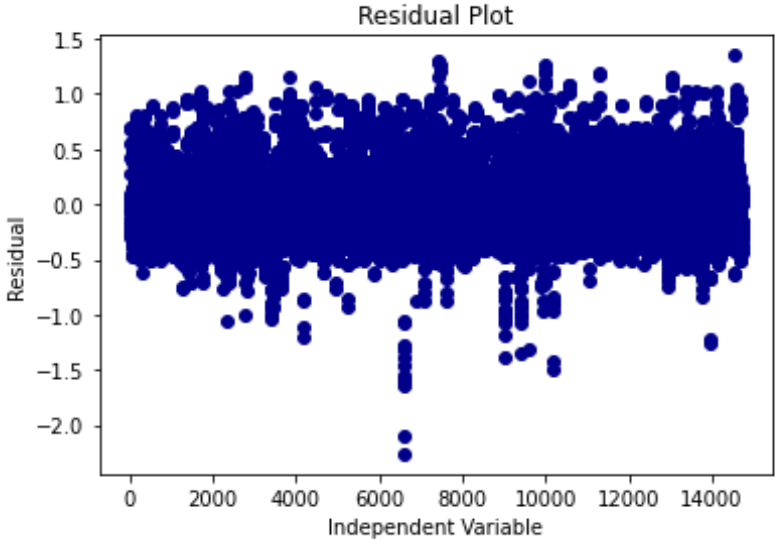


Figure 4.31: Plot of the residuals resulting from the multivariate linear regression for subject M2711.

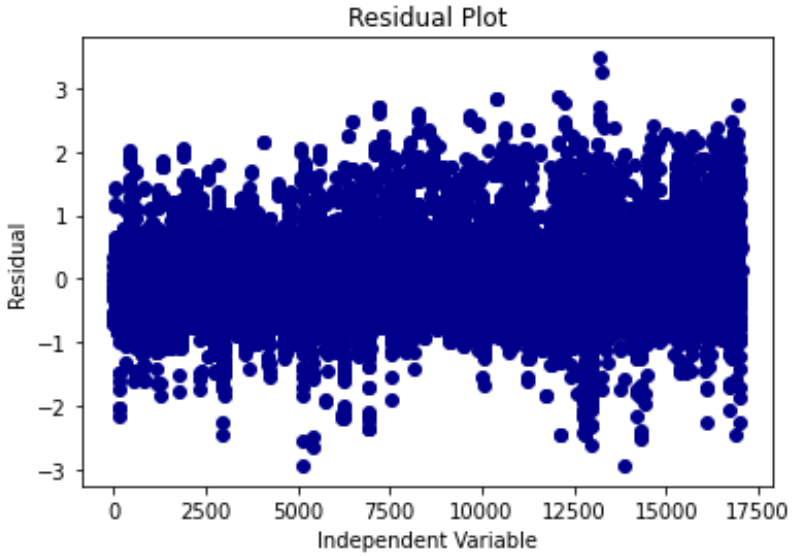


Figure 4.32: Plot of the residuals resulting from the multivariate linear regression for subject M3208.

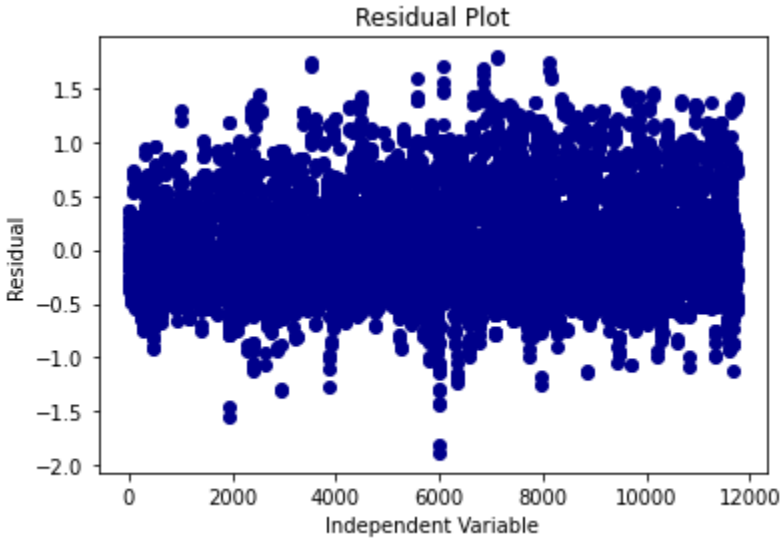


Figure 4.33: Plot of the residuals resulting from the multivariate linear regression for subject M3309.

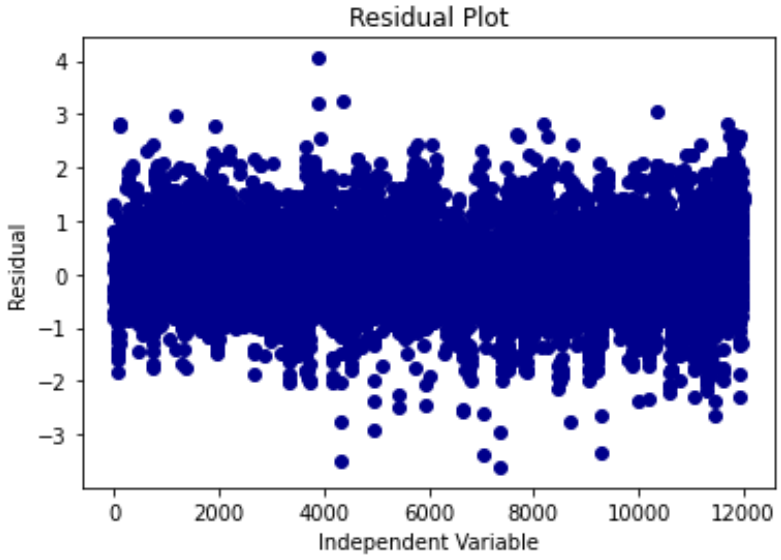


Figure 4.34: Plot of the residuals resulting from the multivariate linear regression for subject M4003.

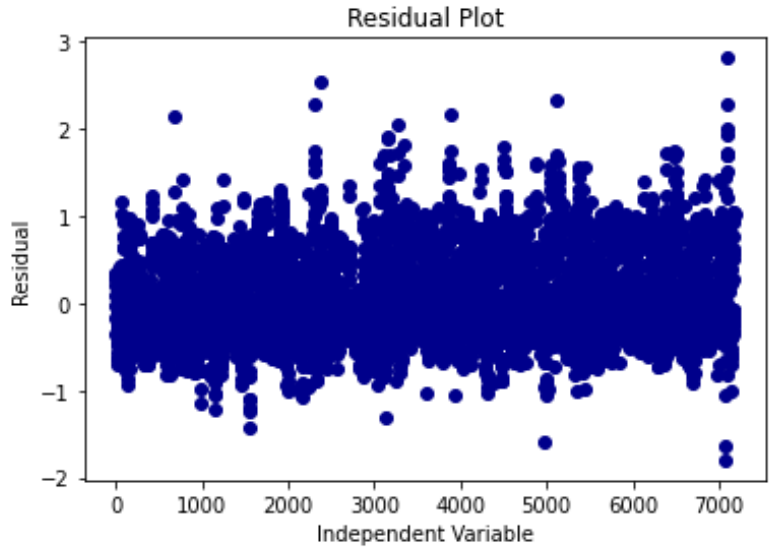


Figure 4.35: Plot of the residuals resulting from the multivariate linear regression for subject M4409.

4.4 Pump settings' analysis

This section reports the results coming from the pump settings' analysis approaches defined in Subsection 3.2.4.

First, are reported the results coming from Approach A, so the meal tirs. Some subjects only had insightful data changes for lunchtime, so their dinner tirs will not be reported.

Then, are reported the results coming from Approach B. For those subjects whose changes were linked to BR, the percentage of its changes, the means (Ms) of these changes and their standard deviations (SDs), the differences (diff) of said means from the basal rate settings in those intervals (SBRs) and the tirs for each date and time intervals are shown. For those subjects whose changes were not linked to BR, but rather to ISF or the increase of the quantity of setting time intervals, only the tirs are displayed.

4.4.1 Approach A

The meal tirs for each of the analysed 10 subjects are the following:

- F0110
  - first date interval; lunch 59,69%,
  - second date interval; lunch 77,65%,
  - third date interval; lunch 73,84%,
  - fourth date interval; lunch 59,40%,
- F0207
  - first date interval; lunch 63,11%,
  - second date interval; lunch 91,25%,
  - first date interval; dinner 48,77%,
  - second date interval; dinner 75,76%,
  - third date interval; dinner 94,89%,
- F1208
  - first date interval; lunch 75,29%,
  - second date interval; lunch 71,88%,
  - third date interval; lunch 50,00%,
- F1590
  - first date interval; lunch 44,29%,
  - second date interval; lunch 56,32%,
  - third date interval; lunch 11,88%,
  - fourth date interval; lunch 51,04%,
  - fifth date interval; lunch 52,86%,

- F2910
  - first date interval; lunch 81,36%,
  - second date interval; lunch 84,30%,
- F4609
  - first date interval; lunch 91,49% and dinner 86,96%,
  - second date interval; lunch 92,60% and dinner 93,39%,
  - third date interval; lunch 88,44% and dinner 98,66% ,
- M1006
  - first date interval; lunch 94,01% and dinner 91,01%,
  - second date interval; lunch 87,91% and dinner 81,71%,
  - third date interval; lunch 81,41% and dinner 77,07% ,
- M2008
  - first date interval; lunch 59,67% and dinner 20,87%,
  - second date interval; lunch 60,51% and dinner 23,64%,
- M3208
  - first date interval; lunch 74,62%,
  - second date interval; lunch 76,62%,
- M4003
  - first date interval; lunch 33,97% and dinner 58,14%,
  - second date interval; lunch 34,09% and dinner 40,82%.

4.4.2 Approach B

The results of Approach B for each of the analysed 10 subjects are the following:

- F0110
  - BR and ICR changes
    - 01/04/2021-08/04/2021 (06:00-12:00 SBR = 1,5; 12:30-14:00 SBR = 1,5)
      - first time interval; percentage of BR changes = 52,44%, M = 1,03, SD = 0,81, diff = 0,47 less and tir = 76,13%,
      - second time interval: percentage of BR changes = 43,86%, M = 1,16, SD = 0,94, diff = 0,34 less and tir = 78,12%,
    - 09/04/2021-21/05/2021 (06:00-12:00 SBR = 1,6; 12:30-14:00 SBR = 1,6)
      - first time interval; percentage of BR changes = 50,42%, M = 0,99, SD = 0,71, diff = 0,61 less and tir = 88,59%,

        - second time interval: percentage of BR changes = 45,24%, M = 1,05, SD = 0,84, diff = 0,55 less and tir = 84,31%,
    - 22/05/2021-05/06/2021 (06:00-12:00 SBR = 1,5; 12:30-14:00 SBR = 1,2)
        - first time interval; percentage of BR changes = 61,47%, M = 0,91, SD = 0,85, diff = 0,59 less and tir = 84,90%,
        - second time interval: percentage of BR changes = 41,38%, M = 1,38, SD = 1,21, diff = 0,18 more and tir = 74,57%,
    - 06/06/2021-30/06/2021 (06:00-12:00 SBR = 1,6; 12:30-14:00 SBR = 1,6)
        - first time interval; percentage of BR changes = 52,98%, M = 1,08, SD = 0,85, diff = 0,52 less and tir = 82,74%,
        - second time interval: percentage of BR changes = 42,86%, M = 1,31, SD = 0,93, diff = 0,29 less and tir = 67,06%,
  - ISF changes
    - 09/04/2021-30/06/2021; tir = 78,94%,

- F0207
  - 01/04/2021-12/05/2021 (15:00-20:00 SBR = 1,2); percentage of BR changes = 38,42%, M = 1,77, SD = 1,26, diff = 0,57 more and tir = 57,05%,
  - 13/05/2021-17/05/2021 (15:00-20:00 SBR = 0,90); percentage of BR changes = 39,00%, M = 0,70, SD = 0,32, diff = 0,20 less and tir = 96,34%,
  - 18/05/2021-30/06/2021 (15:00-20:00 SBR = 0,50); percentage of BR changes = 26,95%, M = 0,42, SD = 0,18, diff = 0,08 less and tir = 98,38%,

- F1208
  - 01/04/2021-16/06/2021 (12:00-15:00 SBR = 0,70); percentage of BR changes = 34,71%, M = 0,72, SD = 0,53, diff = 0,02 more and tir = 73,50%,
  - 17/06/2021-30/06/2021 (12:00-15:00 SBR = 0,85); percentage of BR changes = 38,89%, M = 1,16, SD = 0,63, diff = 0,22 less and tir = 56,60%,

- F1590
  - BR and ICR changes
    - 13/05/2021-16/05/2021 (00:00-04:00 SBR = 0,60, 08:00-12:00 SBR = 0,50 and 17:00-23:00 SBR = 0,70)
        - first time interval; percentage of BR changes = 59,67%, M = 0,89, SD = 0,33, diff = 0,29 more and tir = 55,19%,
        - second time interval: percentage of BR changes = 42,00%, M = 0,98, SD = 0,45, diff = 0,38 more and tir = 28,42%,
        - third time interval; percentage of BR changes = 49,92%, M = 1,06, SD = 0,42, diff = 0,46 more and tir = 38,16%,
    - 17/05/2021-07/06/2021 (00:00-04:00 SBR = 0,50, 08:00-12:00 SBR = 0,40 and 17:00-23:00 SBR = 0,60)

· first time interval; percentage of BR changes = 82,47%, M = 0,95, SD = 0,33, diff = 0,45 more and tir = 53,71%,
· second time interval: percentage of BR changes = 45,55%, M = 0,83, SD = 0,42, diff = 0,33 more and tir = 23,66%,
· third time interval; percentage of BR changes = 38,07%, M = 0,97, SD = 0,41, diff = 0,37 more and tir = 40,73%,

- ICR changes
  * 01/04/2021-12/04/2021 (06:30-12:00, 12:00-15:00 and 18:30-23:00)
    · first time interval; tir = 64,51%,
    · second time interval; tir = 45,58%,
    · third time interval; tir = 15,32%,
  * 13/04/2021-06/05/2021 (06:30-12:00, 12:00-15:00 and 18:30-23:00)
    · first time interval; tir = 58,29%,
    · second time interval; tir = 64,30%,
    · third time interval; tir = 23,50%,
  * 07/05/2021-12/05/2021 (06:30-12:00, 12:00-15:00 and 18:30-23:00)
    · first time interval; tir = 38,88%,
    · second time interval; tir = 14,48%,
    · third time interval; tir = 11,86%,
- Increase of quantity of setting time intervals
  * 01/04/2021-12/05/2021 (00:00-06:30, 06:30-12:00, 12:00-15:00 and 15:00-18:30)
    · first time interval; tir = 51,22%,
    · second time interval; tir = 57,35%,
    · third time interval; tir = 51,43%,
    · fourth time interval; tir = 17,60%,
  * 13/05/2021-30/06/2021 (00:00-06:30, 06:30-12:00, 12:00-15:00 and 15:00-18:30)
    · first time interval; tir = 64,83%,
    · second time interval; tir = 46,68%,
    · third time interval; tir = 56,79%,
    · fourth time interval; tir = 46,38%,

- F2910
  - BR and ICR changes
    * 01/04/2021-26/04/2021 (16:00-19:00 SBR = 1,50); percentage of BR changes = 37,63%, M = 0,97, SD = 0,80, diff = 0,53 less and tir = 78,29%,
    * 27/04/2021-02/05/2021 (16:00-19:00 SBR = 1); percentage of BR changes = 48,80%, M = 1,27, SD = 0,99, diff = 0,27 more and tir = 83,43%,
    * 03/05/2021-30/06/2021 (16:00-19:00 SBR = 1,20); percentage of BR changes = 37,57%, M = 0,95, SD = 0,81, diff = 0,25 less and tir = 79,31%,
  - Increase of quantity of setting time intervals

    * 01/04/2021-02/05/2021 (03:00-08:00); tir = 97,63%,
    * 03/05/2021-30/06/2021 (03:00-08:00); tir = 97,37%,

- F4609
  - 01/04/2021-21/05/2021 (05:00-07:00 SBR = 0,60, 09:00-11:00 SBR = 0,60, 11:00-13:00 SBR = 0,60, 15:00-16:00 SBR = 0,80, 19:00-22:00 SBR = 0,80 and 22:00-00:00 = 0,60)
    * first time interval; percentage of BR changes = 71,29%, M = 0,71, SD = 0,36, diff = 0,11 more and tir = 99,65%,
    * second time interval: percentage of BR changes = 48,79%, M = 0,63, SD = 0,49, diff = 0,03 more and tir = 90,96%,
    * third time interval; percentage of BR changes = 56,62%, M = 0,84, SD = 0,57, diff = 0,24 more and tir = 89,17%,
    * fourth time interval; percentage of BR changes = 49,58%, M = 0,90, SD = 0,59, diff = 0,10 more and tir = 87,05%,
    * fifth time interval: percentage of BR changes = 41,60%, M = 0,76, SD = 0,53, diff = 0,04 less and tir = 88,44%,
    * sixth time interval; percentage of BR changes = 59,71%, M = 0,97, SD = 0,54, diff = 0,37 more and tir = 83,18%,
  - 22/05/2021-30/06/2021 (05:00-07:00 SBR = 0,65, 09:00-11:00 SBR = 0,65, 11:00-13:00 SBR = 0,65, 15:00-16:00 SBR = 0,75, 19:00-22:00 SBR = 0,75 and 22:00-00:00 = 0,75)
    * first time interval; percentage of BR changes = 71,94%, M = 0,68, SD = 0,34, diff = 0,03 more and tir = 100,00%,
    * second time interval: percentage of BR changes = 45,73%, M = 0,67, SD = 0,52, diff = 0,02 more and tir = 93,03%,
    * third time interval; percentage of BR changes = 34,65%, M = 0,69, SD = 0,49, diff = 0,04 more and tir = 87,28%,
    * fourth time interval; percentage of BR changes = 47,46%, M = 0,74, SD = 0,44, diff = 0,01 less and tir = 87,76%,
    * fifth time interval: percentage of BR changes = 40,39%, M = 0,70, SD = 0,47, diff = 0,05 less and tir = 95,45%,
    * sixth time interval; percentage of BR changes = 63,42%, M = 0,84, SD = 0,51, diff = 0,14 more and tir = 93,97%,

- M1006
  - 01/04/2021-19/04/2021 (12:00-14:00 SBR = 0,60); percentage of BR changes = 28,50%, M = 0,42, SD = 0,23, diff = 0,18 less and tir = 90,45%,
  - 20/04/2021-30/06/2021 (12:00-14:00 SBR = 0,60); percentage of BR changes = 26,63%, M = 0,52, SD = 0,40, diff = 0,08 less and tir = 88,49%,

- M2008

- 01/04/2021-22/04/2021 (00:00-03:00 SBR = 1,00, 03:00-08:00 SBR = 1,30, 08:00-15:00 SBR = 1,10, 15:00-21:00 SBR = 1,60 and 21:00-00:00 SBR = 1,30)
  - first time interval; percentage of BR changes = 73,35%, M = 2,70, SD = 1,74, diff = 1,70 more and tir = 45,73%,
  - second time interval: percentage of BR changes = 65,07%, M = 2,33, SD = 1,41, diff = 1,33 more and tir = 85,79%,
  - third time interval; percentage of BR changes = 32,64%, M = 1,49, SD = 1,53, diff = 0,39 more and tir = 62,58%,
  - fourth time interval; percentage of BR changes = 46,88%, M = 3,07, SD = 1,58, diff = 1,47 more and tir = 18,31%,
  - fifth time interval: percentage of BR changes = 56,05%, M = 2,84, SD = 1,65, diff = 0,35 more and tir = 26,61%,
- 23/04/2021-30/06/2021 (00:00-03:00 SBR = 1,20, 03:00-08:00 SBR = 1,40, 08:00-15:00 SBR = 1,40, 15:00-21:00 SBR = 1,40 and 21:00-00:00 SBR = 1,40)
  - first time interval; percentage of BR changes = 70,98%, M = 3,14, SD = 1,72, diff = 1,94 more and tir = 52,63%,
  - second time interval: percentage of BR changes = 61,32%, M = 2,49, SD = 1,47, diff = 1,09 more and tir = 90,58%,
  - third time interval; percentage of BR changes = 32,46%, M = 1,91, SD = 1,81, diff = 0,51 more and tir = 54,46%,
  - fourth time interval; percentage of BR changes = 38,06%, M = 2,94, SD = 1,90, diff = 1,54 more and tir = 33,28%,
  - fifth time interval: percentage of BR changes = 43,43%, M = 3,33, SD = 1,94, diff = 0,54 more and tir = 25,54%,

- M3208
  - 01/04/2021-15/04/2021 (07:00-15:00 SBR = 0,70); percentage of BR changes = 34,30%, M = 0,88, SD = 0,61, diff = 0,18 more and tir = 78,43%,
  - 16/04/2021-13/05/2021 (07:00-15:00 SBR = 1,00); percentage of BR changes = 39,21%, M = 0,98, SD = 0,74, diff = 0,02 less and tir = 82,67%,
  - 14/05/2021-30/06/2021 (07:00-15:00 SBR = 1,50); percentage of BR changes = 43,49%, M = 1,48, SD = 1,05, diff = 0,02 less and tir = 67,43%,
- M4003
  - 01/04/2021-16/04/2021 (15:00-21:00 SBR = 1,20); percentage of BR changes = 42,44%, M = 2,01, SD = 0,92, diff = 0,81 more and tir = 38,68%,
  - 17/04/2021-30/06/2021 (15:00-21:00 SBR = 1,50); percentage of BR changes = 41,92%, M = 1,91, SD = 0,95, diff = 0,41 more and tir = 32,15%.

Chapter 5

Discussion

This part discusses, in each of its sections, the results reported in Chapter 4 and achieved with the different machine learning approaches described in Chapter 3.

5.1 Logistic regression

As already hinted at in Subsection 3.2.1, results for logistic regression did not improve when the preprocessing of files was changed: the ending results continued to be discouraging. Normalising the data or using information coming from the previous glycaemic recordings turned out to be useless.

The confusion matrices show how results were mediocre at best: for example, Figure 4.1a shows that 2156 cells were there was no injection (0; meaning that the basal velocity did not change) were correctly predicted and 1301 were wrongly labelled, while 1819 were there was injection (1; meaning that the basal velocity changed) were correctly predicted and 1495 were mislabelled. The ratios are off and the approach does not efficiently recognise how and why the change in basal rate is chosen by the Control-IQ™ technology. This reflects on the same subject's *Prec*, *Rec* and $F1$ values reported in Table 4.1. An analogous analysis can be made on subjects' F1408 (Figure 4.1d and Table 4.1), F1590 (Figure 4.2a and Table 4.1) results and M3208 (Figure 4.5a and Table 4.1) results.

The other subjects reported even worse results. Subject F0207 (Figure 4.1b and Table 4.1), F1208 (Figure 4.1c and Table 4.1), F2810 (Figure 4.2c and Table 4.1), F2910 (Figure 4.2d and Table 4.1), M1006 (Figure 4.4b and Table 4.1) and M2008 (Figure 4.4c and Table 4.1) showed very high predictions of 0s and very low predictions of 1 that even after having taken into consideration how the classes were imbalanced (way more 0s than 1s) show how this machine learning approach did not accurately predicted 0s and was extremely bad at recognising 1s.

Instead, one of the worst subject is F2403: as the confusion matrix (Figure 4.2b) and *Prec*, *Rec* and $F1$ values (Table 4.1) show, this machine learning approach never predicts 1s. Even if the Accuracy score is almost 0.96, it is only due to the highly imbalanced data (5219 rows of 0s compared to 218 rows of 1s): the probability of having to predict a 1 is so low, that even if the algorithm only predicts 0, it is still accurate. This means that the Accuracy score, in this case, is not a reliable metric. The low value of this approach on this specific subject is especially notable for *Prec*, *Rec* and $F1$ values which are all equal to 0.00. An analogous analysis can be made on subjects' F3807 (Figure 4.3a and Table 4.1),

Table 5.1: Confusion matrix table resulting from the random forest approach for subject F2403 when 2 rows of glycaemic info were added. The labels on the right are the actual ones, while the ones on the top are the predicted ones.

	0	1
0	5161	58
1	206	12

F4301 (Figure 4.3b and Table 4.1), F4503 (Figure 4.3c and Table 4.1), F4714 (Figure 4.4a and Table 4.1), M2711 (Figure 4.4d and Table 4.1) results, M3309 (Figure 4.5b and Table 4.1) and M4409 (Figure 4.5d and Table 4.1) results.

One of the only two subjects whose results have a different trend from any others is F4609 (Figure 4.3d and Table 4.1) as data is imbalanced in the opposite way: there are more 1s than 0s and the machine learning algorithm mostly predicts 1s even when they should be 0s. In fact, 378 0s are correctly predicted and 5777 are not, while 7187 1s are correctly labelled and only 230 get mislabelled. The other subject is M4003 (Figure 4.5c and Table 4.1).

The reported ROC curves (Figures 4.6, 4.7, 4.8, 4.9, 4.9 and 4.10), which were the best out of all the obtained ones, meaning that they had the most area under the curve, only endorse the previous discussion. The results achieved through the logistic regression were not satisfactory.

The main issue must probably be traced in the nature of the dataset: while every row has a glycaemic value record, basal insulin data is scarcer. This leads to a very imbalanced dataset which is also rather inhomogeneous, leading to different results based on the ratios of 0s and 1s for each patient. Even if some of them can be grouped by similarity, the group of subjects, analysed as a whole, gave very discording results.

5.2 Random forest

As already hinted, the random forest approach was implemented in an attempt to improve the results of the analysis of the choice of Control-IQ™ to either change or not the basal insulin rate as, as just said in Section 5.1, the ones coming from the logistic regression were not satisfactory.

The random forest algorithm did not really improve results when using the glycaemic files without the addition of the previous blood glucose levels. Instead, when using the files to which the 2 previous rows were added, results became better and they bettered even more when adding glycaemic information until half an hour before (which is to say for the 6 previous values). In fact, taking into consideration subject F2403, which was one of the worst ones in terms of results for the logistic regression (as said in Section 5.1), 1s were still never predicted, while, with the addition of the 2 previous rows, the confusion matrix for this individual was the one reported in Table 5.1:

Instead, the confusion matrix for subject F0110, which was already one of the best ones for the logistic regression (as said in Section 5.1), with the addition of the 2 previous rows, was the one reported in Table 5.2:

So, the results reported in Section 4.2 are the ones coming from the work on the glycaemic files to which the info on the glycaemic trend in the previous 30 minutes and are the best ones achieved. The simple comparison between Table 5.2 and Figure 4.11a or Table 5.1 and

Table 5.2: Confusion matrix table resulting from the random forest approach for subject F0110 when 2 rows of glycaemic info were added. The labels on the right are the actual ones, while the ones on the top are the predicted ones.

	0	1
0	2135	1322
1	1303	2011

Figure 4.12b, shows how results improved.

Moreover, it is interesting to compare the confusion matrices for the logistic regression and the random forest for each subject. The biggest improvement is notable in those subjects for which 1s were never predicted (the just cited F2403, but also F3807, F4301, F4503, F4714, M2711, M3309 and M4409) in the previous attempt. Now, some 1s get predicted, even if not perfectly and even if they are not a lot.

This also means that accuracy scores are generally a lot more valid and trustworthy and more homogeneous.

Table 4.2 reporting the $Prec$, Rec and $F1$ values for each subject is in full accordance with this analysis.

Finally, the lower value of MAE, MSE and RMSE implies higher accuracy of a regression model[1]. Thus, the results given by the random forest approach can be considered to be quite satisfactory.

5.3 Multivariate linear regression

The results reported in Section 4.3 are the ones coming from the work on the glycaemic files to which the info on the glycaemic trend regarding the 2 previous recordings was added. As already hinted at, in fact, adding the 6 previous rows did not generally improve results, but only F-stat, so it was discarded as the changes were not very telling. F-stat helps determine whether the relationship between your model and the response variable, whose strength is estimated by R^2, is statistically significant[2], but the practical interpretation of the equation that defines F-stat relating it to R^2 is that bigger values of the latter, lead to higher values of the former, giving strong evidence that at least some of the coefficients are non-zero.

Since the main goal of the multivariate linear regression approach was to understand the connection between the change of insulin rate and glycaemia, this can also be done through the analysis of R^2 values and the coefficients' p-values for each subject. In fact, just R^2 by itself is not enough to understand the relationships between the independent variables and dependent variable: it is irrelevant[3]. This is why it has to be taken into consideration along with the coefficients p-values reported in Tables 4.3, 4.4, 4.5 and 4.6. Even if in general a higher R^2 value means that the model fits the data better, when coupled with high p-values (>0.05) it means that the model is worthless as it explains a lot of the data variation, but it

[1] https://medium.com/analytics-vidhya/mae-mse-rmse-coefficient-of-determination-adjusted-r-squared-which-metric-is-better-cd0326a5697e

[2] https://blog.minitab.com/en/adventures-in-statistics-2/regression-analysis-how-do-i-interpret-r-squared-and-assess-the-goodness-of-fit

[3] https://statisticsbyjim.com/regression/how-high-r-squared/

is not significant. Instead, when coupled with low p-values ($\leq$0.05) it means that the model explains a lot of the variation of the data and it is significant. If, on the contrary, R^2 is low and p-values are high, the model is the worst one possible as it does not explain much of the data variation and it is not even significant, while when p-values are low, it means that the model is significant, but does not explain much of the variation of the data[4].

In this case, R^2 for subjects F0110, F1208, F4503 and M3208 are the only ones that can be considered moderate (not very high), while the others are all lower, especially the ones for subjects F2403 and F3807. The latter also have very high F-stat values: probably, since their data is very imbalanced, this regression cannot efficiently find a relationship between their independent and dependent variables. Nevertheless, as shown in Tables 4.7, 4.8, 4.9 and 4.10, the overall p-values are 0.000 or $<$0.050. The only exceptions can be seen for subjects F2403 and F3807, as expected, but also for subjects F4609, M2008 and M3309 even if these last p-values are lower.

The p-value for each coefficient test the null hypothesis that states that said coefficient is equal to zero, thus having no effect in the relationship[5]. A low p-value means that one can reject the said hypothesis and the predictor that coefficient is linked to is likely to be meaningful inside the model as its changes are related to changes in the response variable. On the contrary, a larger p-value suggests that that predictor is statistically insignificant. So, in this case:

- for the model of subject F2403, the columns of the 2 previous glycaemic values are not significant,
- for the model of subject F3807, the column of the previous glycaemic value is not significant,
- for the model of subject F4609, the column of the previous glycaemic value is probably not significant,
- for the model of subject M2008, the column of the previous glycaemic value is probably not significant,
- for the model of subject M3309, the column of the previous glycaemic value is probably not significant.

In order to extract meaningful info from the coefficients of Eq. 3.9 (reported in Tables 4.3, 4.4, 4.5 and 4.6) and try to see if there is any trend linked to their year of birth, Tables 5.3, 5.4, 5.5, 5.6, 5.7, 5.8, 5.9, 5.10, 5.11, 5.12, 5.13, 5.14, 5.15, 5.16 and 5.17 compare the first (β_0), second (β_1), third (β_2), fourth (β_3) and fifth (β_4) coefficient for birth groups 2001-2003, 2006-2009 and 2010-2014, respectively. Thus excluding only one subject (F1590) as she was the only one born in 1990.

These comparisons do not highlight any real or meaningful similarity in age groups. Their values are similar, but not enough, and even when numbers are close, their orders of magnitude are not.

[4] https://www.researchgate.net/post/What_is_the_relationship_between_R-squared_and_p-value_in_a_regression

[5] https://blog.minitab.com/en/adventures-in-statistics-2/how-to-interpret-regression-analysis-results-p-values-and-coefficients

Table 5.3: First coefficients of the multivariate linear regression; years of birth: 2001-2003.

F2403	0.848
F4301	-0.062
F4503	-0.219
M4003	0.233

Table 5.4: Second coefficients of the multivariate linear regression; years of birth: 2001-2003.

F2403	$3.014*e^{-4}$
F4301	$9.781*e^{-3}$
F4503	$8.925*e^{-3}$
M4003	$9.495*e^{-3}$

Moreover, something interesting to focus on is that β_2 (the third coefficient) is 0.000 for subject F2810, meaning that carbohydrates (the third parameter) are not significant in this specific model.

Finally, regarding the residuals' plots, we want them to display a cloud of points around 0 with constant variance and no trend[67]. This analysis on the plots is done retrospectively and, when it has a positive outcome, it demonstrates that the model is a good fit for the data.

In this case, Figures 4.16, 4.18, 4.19, 4.20, 4.23, 4.25, 4.26, 4.27, 4.28, 4.29, 4.30, 4.31, 4.32, 4.33, 4.34 and 4.35 that are the residuals' plots for subjects F0110, F1208, F1408, F1590, F2910, F4301, F4503, F4609, F4714, M1006, M2008, M2711, M3208, M3309, M4003 and M4409, respectively, show residuals that are independent and normally distributed.

On the contrary, Figures 4.17 of subject F0207, 4.21 of subject F2403, 4.22 of subject F2810 and 4.24 of subject F3807 show residuals that are not dependent and normally distributed. In fact, subjects F0207 and F2403's residual plots show a decreasing trend, subject F2810's residual plot show an increasing trend and subject F3807's residual plot has a categorical trend, meaning that there are 2 distinct clouds of points with a specific pattern.

[6] `https://towardsdatascience.com/how-to-use-residual-plots-for-regression-model-validation-c3c70e8ab378`

[7] `https://statisticsbyjim.com/regression/check-residual-plots-regression-analysis/`

Table 5.5: Third coefficients of the multivariate linear regression; years of birth: 2001-2003.

F2403	$-5.421*e^{-20}$
F4301	$-1.464*e^{-17}$
F4503	$-2.927*e^{-18}$
M4003	$3.469*e^{-18}$

Table 5.6: Fourth coefficients of the multivariate linear regression; years of birth: 2001-2003.

F2403	-4.701*e^{-5}
F4301	2.932*e^{-3}
F4503	2.901*e^{-3}
M4003	2.491*$^{-3}$

Table 5.7: Fifth coefficients of the multivariate linear regression; years of birth: 2001-2003.

F2403	-1.094*e^{-4}
F4301	-4.286*e^{-3}
F4503	-4.210*e^{-3}
M4003	-5.598*e^{-3}

Table 5.8: First coefficients of the multivariate linear regression; years of birth: 2006-2009.

F0207	-0.917
F1208	-0.097
F1408	-1.317
F3807	0.811
F4609	-0.571
M1006	-0.410
M2008	0.333
M3208	-0.163
M3309	-0.075
M4409	-0.186

Table 5.9: Second coefficients of the multivariate linear regression; years of birth: 2006-2009.

F0207	$1.327*e^{-2}$
F1208	$9.347*e^{-3}$
F1408	$2.608*e^{-2}$
F3807	$1.181*e^{-4}$
F4609	$1.802*e^{-2}$
M1006	$1.278*e^{-2}$
M2008	$2.375*e^{-2}$
M3208	$1.533*e^{-3}$
M3309	$8.404*e^{-3}$
M4409	$5.690*e^{-3}$

Table 5.10: Third coefficients of the multivariate linear regression; years of birth: 2006-2009.

F0207	$-2.331*e^{-18}$
F1208	$2.602*e^{-18}$
F1408	$3.469*e^{-18}$
F3807	$-3.388*e^{-20}$
F4609	$-8.674*e^{-19}$
M1006	$-4.337*e^{-19}$
M2008	$-8.674*e^{-18}$
M3208	$4.770*e^{-18}$
M3309	$-4.337*e^{-19}$
M4409	$-5.421*e^{-19}$

Table 5.11: Fourth coefficients of the multivariate linear regression; years of birth: 2006-2009.

F0207	$3.946*e^{-3}$
F1208	$2.034*e^{-3}$
F1408	$9.346*e^{-3}$
F3807	$-2.598*e^{-6}$
F4609	$6.089*e^{-4}$
M1006	$1.513*e^{-3}$
M2008	$1.154*e^{-3}$
M3208	$2.487*e^{-3}$
M3309	$7.477*e^{-4}$
M4409	$1.486*e^{-3}$

Table 5.12: Fifth coefficients of the multivariate linear regression; years of birth: 2006-2009.

F0207	$-4.092*e^{-3}$
F1208	$-4.993*e^{-3}$
F1408	$-1.313*e^{-2}$
F3807	$-8.217*e^{-5}$
F4609	$-8.200*e^{-3}$
M1006	$-5.625*e^{-3}$
M2008	$1.259*e^{-2}$
M3208	$-7.696*e^{-3}$
M3309	$-3.537*e^{-3}$
M4409	$-9.209*e^{-4}$

Table 5.13: First coefficients of the multivariate linear regression; years of birth: 2010-2014.

F0110	-0.435
F2810	-0.199
F2910	-0.441
F4714	0.023
M2711	-0.310

Table 5.14: Second coefficients of the multivariate linear regression; years of birth: 2010-2014.

F0110	$1.944*e^{-2}$
F2810	$3.892*e^{-3}$
F2910	$1.613*e^{-2}$
F4714	$3.599*e^{-3}$
M2711	$1.142\text{-}e^{-2}$

Table 5.15: Third coefficients of the multivariate linear regression; years of birth: 2010-2014.

F0110	$-1.735*e^{-18}$
F2810	0.000
F2910	$-1.301*e^{-18}$
F4714	$-4.337*e^{-19}$
M2711	$4.337*e^{-19}$

Table 5.16: Fourth coefficients of the multivariate linear regression; years of birth: 2010-2014.

F0110	$5.665*e^{-3}$
F2810	$1.027*e^{-3}$
F2910	$4.036*e^{-3}$
F4714	$5.285*e^{-4}$
M2711	$9.325*e^{-4}$

Table 5.17: Fifth coefficients of the multivariate linear regression; years of birth: 2010-2014.

F0110	$-1.296*e^{-2}$
F2810	$-9.824*e^{-4}$
F2910	$-7.063*e^{-3}$
F4714	$-1.493*e^{-3}$
M2711	$-4.608*e^{-3}$

5.4 Pump settings' analysis

The results reported in Section 4.4 for both approaches will be commented here.

At first, the discussion will be focused on approach A's results reported in Subsection 4.4.1.

Regarding subject F0110, lunch tirs improved then gradually worsened throughout the 91-days span of analysis. In the second date interval (09/04/2021-21/05/2021), at 12:30, the BR was raised from 1,5 to 1,6, ICR was lowered from 9 to 8 and ISF was raised from 60 to 70 leading to a good tir. In the third date interval (22/05/2021-05/06/2021) at the same hour, the first two settings became 1,2 and 10, while the third one remained the same, still keeping a good, even if worse, tir. In the last date interval (06/06/2021-30/06/2021), the settings (BR = 1,6, ICR = 9 and ISF = 70) were very similar to the ones of the second interval but still resulted in a worse tir. This could indicate that, for this patient, the ICR is the main influence on glycaemic control, but it would not explain the results of the third time interval that have an even higher setting of ICR. So, probably, there were other influencing factors in the last date interval.

It is interesting to highlight that this is the only subject for which ISF changes are reported in the analysed time period, but the same settings of this parameter have different outcomes, so it does not seem like ISF has a deep influence on F0110's glycaemic control during meals.

Regarding subject F0207, lunch and dinner tirs were calculated for different date intervals, but they both continued to improve in the analysed period. In this case, ICR and ISF never change, meaning that this subject is strongly and positively influenced by the BR.

Regarding subject F1208, only lunch tirs can be analysed and they show deterioration over time. Since the BR stayed the same from 13/05/2021 onward, while ICR changed another time, the results seem to suggest that the ICR changes were not the correct ones (until 16/06/2021 ICR = 8 and from 17/06/2021 ICR = 12).

Regarding subject F1590, lunch tirs improved at first, then had a drastic worsening in the third date interval (07/05/2021-12/05/2021). The small duration of this date interval highlights how bad glycaemic performances were detected and the settings were corrected: in fact, time settings were 5 until 12/05/2021 and became 9 from 13/05/2021 onward. Looking at the results, this looks like a good decision.

Regarding subject F2910, there is an improvement of lunch tirs concurrent to lower BR and ICR values, suggesting that this subject needs more modest settings' values.

Regarding subject F4609, dinner tirs continued to improve throughout the 91-days span of analysis, while the ones computed for lunchtime improved at first but then worsened. That probably indicates that the changes (increase of lunchtime tir and lunchtime ICR) done around the second half of the day were more efficient than the ones (decrease of dinnertime tir and increase of dinnertime ICR) done for the middle of the day as the date intervals are the same for both meal analysis.

Regarding subject M1006, both meal tirs worsened, indicating that these changes were not efficient. BR was not changed in value, but only in the time at which it was set (from 13:00 to 12:30 on the 20/04/2021), while ICR was raised throughout the period. Probably, this subject needed a change in BR and, maybe, the ICR rise was counterproductive.

Regarding subject M2008, both meal tirs result from the same date intervals and show an improvement, even if a slight one. In this case, ICR does not change and BR is set to 1,2 for the first setting time and to 1,4 for the remaining 8 setting times. Maybe, such a

repetition is not efficient and the values should show more heterogeneity throughout the day.

Regarding subject M3208, lunch tirs slightly improve in the second time interval when compared to the first one due to an increase of the BR value and a decrease of the ICR.

Regarding subject M4003, both meal tirs result from the same date intervals, but, while lunch tir slightly improves, the dinner one worsens. In this case, ICR does not change, so the results come from BR changes realised through raises. Maybe, for this individual, the lunch BR value needs to become even higher and the dinner one needs to be smaller.

Let's move on to the discussion of approach B's results reported in Subsection 4.4.2. The first evaluation metric is the percentage of basal rate changes as it represents an insight on how good the original setting when compared to the real needs of the subject: the most changes to the basal insulin velocities are made (based on blood glucose levels), the less correct is that setting for that time interval. Then, another useful evaluation metric is the difference of the mean of recorded basal rate changes compared to the original setting in that time interval: this could be a parameter helpful to understand how much smaller or bigger the magnitude of the basal rate setting should be for each interval of interest. Finally, the computation of tirs in said intervals is helpful to see how efficiently the recorded changes and the set velocity control glycaemia through a combined analysis of the tir value and the percentage of changes in the same interval.

Generally speaking, the standard deviations of the means of BR changes are quite high, highlighting the significant variability of this data.

Regarding subject F0110's BR and ICR changes, throughout the 4 date intervals, the first time interval shows a good tir (>75%), but almost half of the recorded basal rows show a change in BR values that highly differs from the original setting. Instead, the second time interval shows an increase in tir until 21/05/2021 and a worsening onward. The latter is concomitant with fewer BR corrections whose values are smaller suggesting that the BR setting is not efficient and bigger corrections would be needed. Subject F0110's ISF changes resulting tir can be compared with the one at lunchtime of the first time interval of Approach A: there is an improvement considering a smaller time period.

For a simpler comparison, figures 5.1, 5.2, 5.3 and 5.4 shows a visual summary of the Personal Profile settings' and tirs' changes throughout the 4 analysed date intervals for subject F0110.

Regarding subject F0207, the tirs improved and the percentage of corrections (38,42%, 39,00% and 26,95%) almost stayed the same and then decreased along with their difference (0,57 more, 0,20 less and 0,08 less) from the original setting. This suggests that the BR changes were efficient.

Regarding subject F1208, the tirs worsened and more (from 34,71% to 38,89%) and higher (from 0,02 more to 0,22 less) corrections were needed suggesting that the BR and ICR changes were not very efficient.

Regarding subject F1590's BR and ICR changes, throughout the 2 date intervals, all time intervals present an insufficient tir (<55%) and the best one out of them is linked to 82,47% of high BR changes. Thus, the changes chosen for this subject were not correct. To further explore this statement, subject F1590's ICR changes resulting tirs can be analysed: they show a general deterioration in the date intervals that as a whole are inside the first date interval for both BR and ICR changes, hinting at the fact that ICR does not have high influence on this specific individual. Regarding the increase of the quantity of setting time intervals, the results seem to suggest that, generally, more time intervals lead to better tirs.

Regarding subject F2910's BR and ICR changes, the 3 date intervals generally present

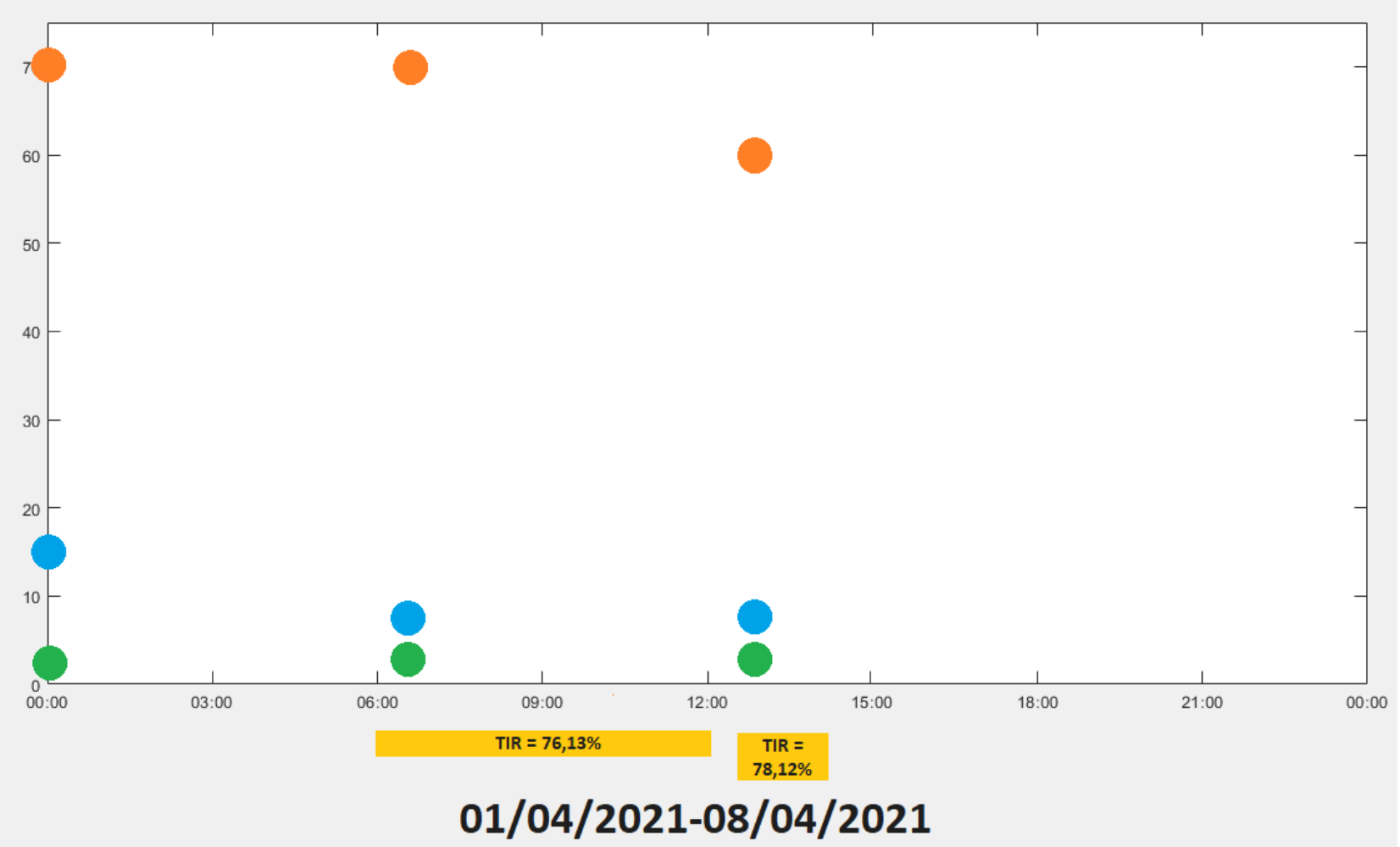


Figure 5.1: Plot of the Personal Profile settings for BR (represented by green dots), ICR (represented by blue dots) and ISF (represented by orange dots) throughout the time intervals of the 1st date interval of interest of subject F0110. Over the yellow background are reported the tirs for the 2 time intervals of interest over the same date interval.

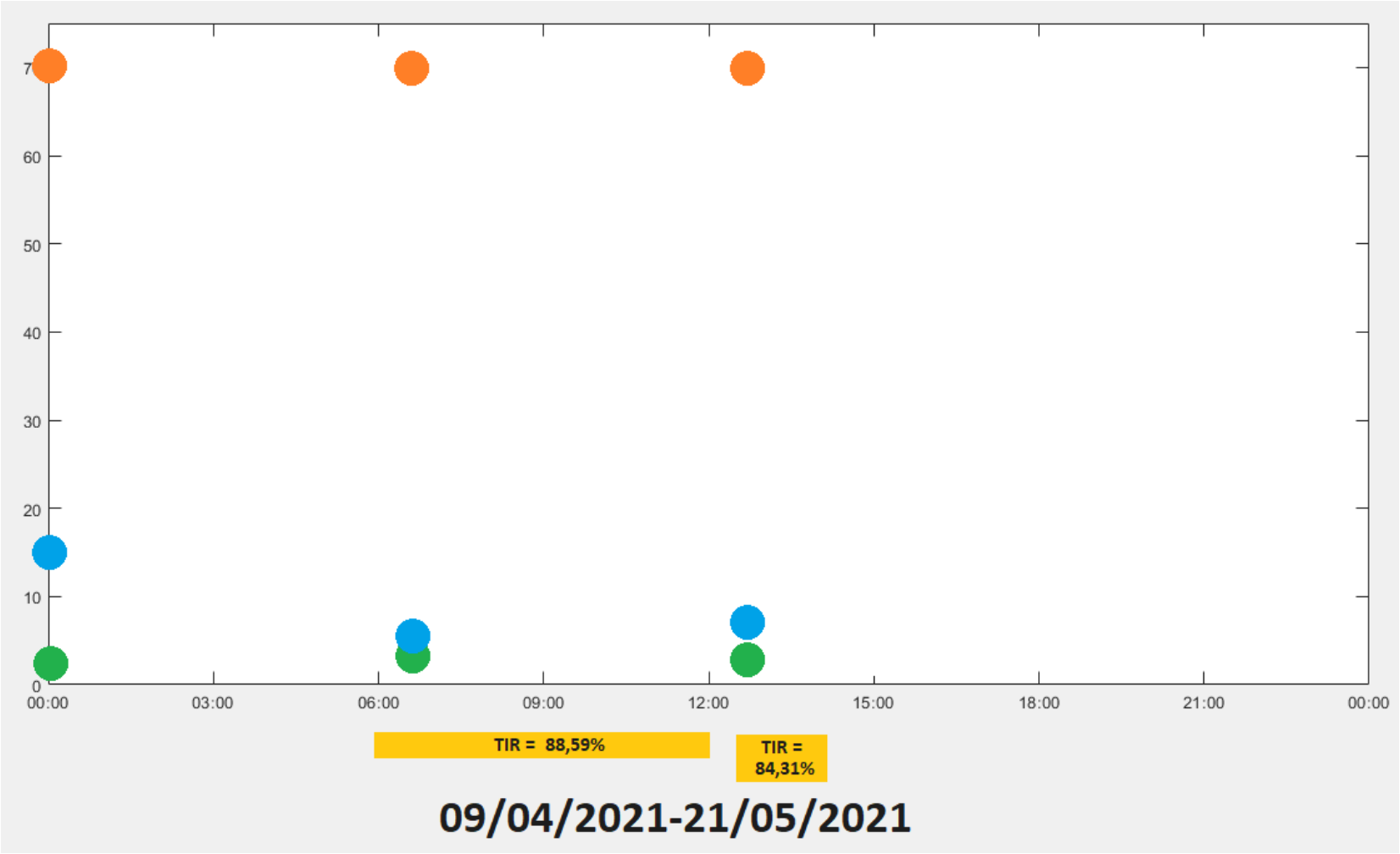


Figure 5.2: Plot of the Personal Profile settings for BR (represented by green dots), ICR (represented by blue dots) and ISF (represented by orange dots) throughout the time intervals of the 2nd date interval of interest of subject F0110. Over the yellow background are reported the tirs for the 2 time intervals of interest over the same date interval.

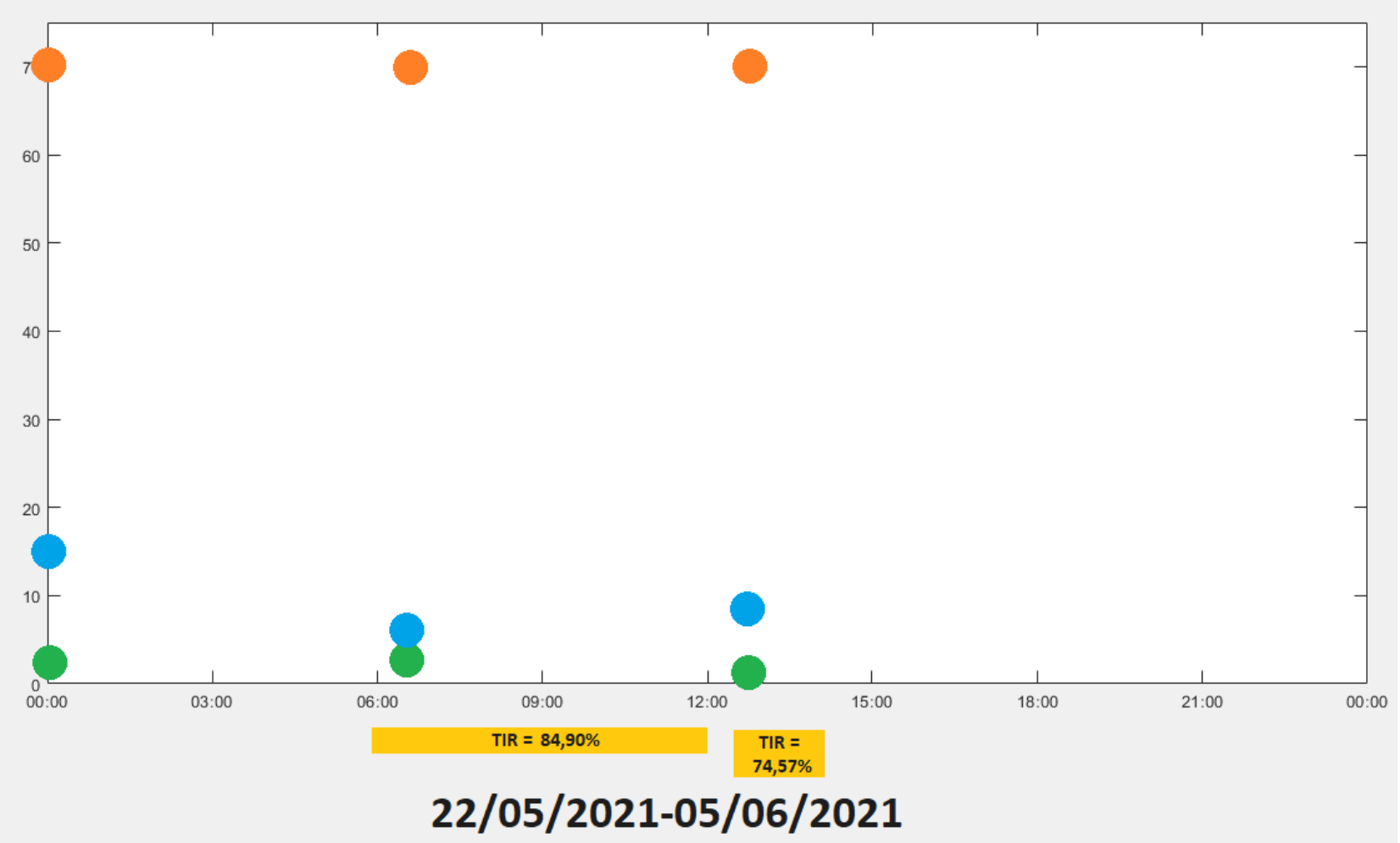


Figure 5.3: Plot of the Personal Profile settings for BR (represented by green dots), ICR (represented by blue dots) and ISF (represented by orange dots) throughout the time intervals of the 3rd date interval of interest of subject F0110. Over the yellow background are reported the tirs for the 2 time intervals of interest over the same date interval.

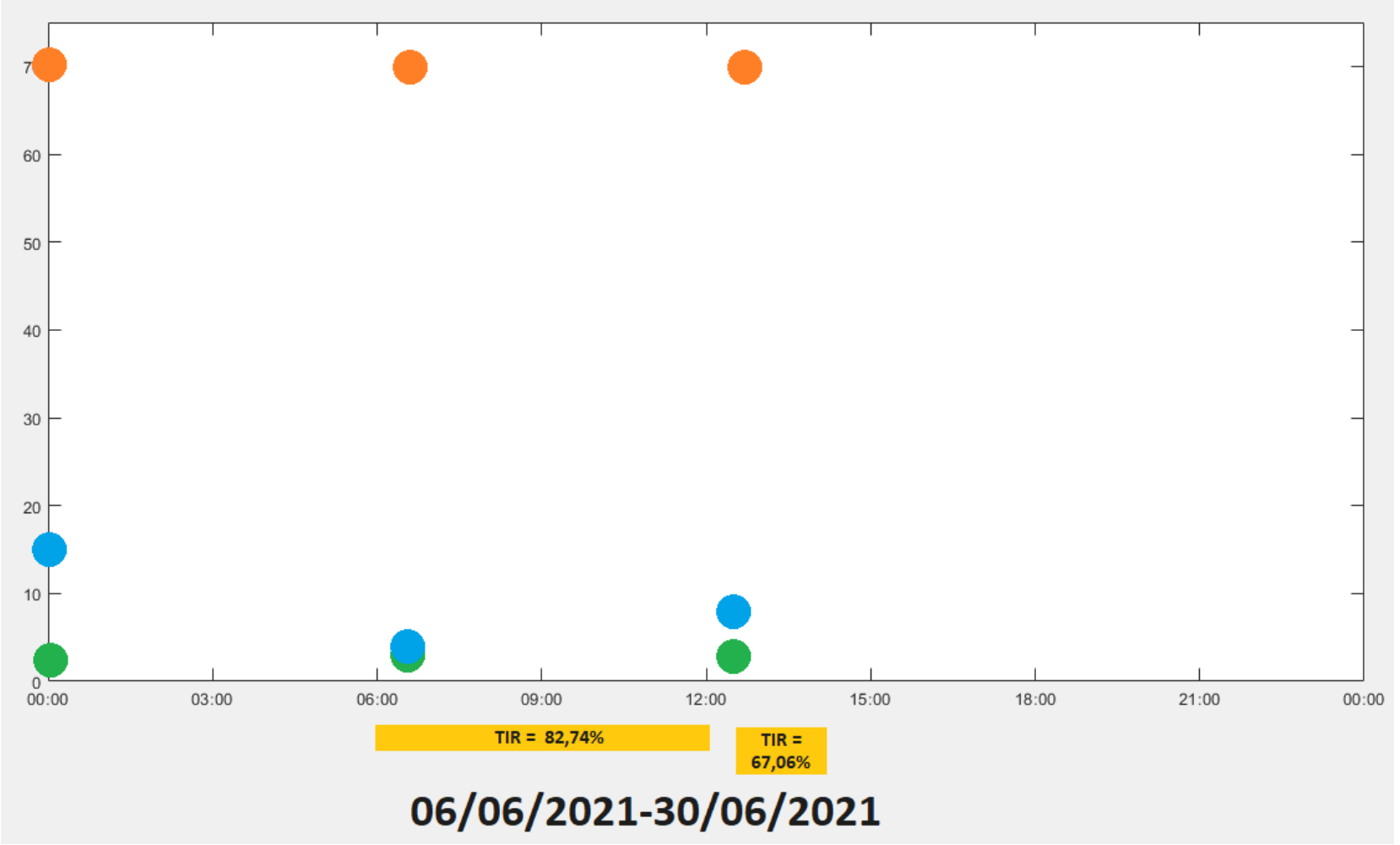


Figure 5.4: Plot of the Personal Profile settings for BR (represented by green dots), ICR (represented by blue dots) and ISF (represented by orange dots) throughout the time intervals of the 4^{th} date interval of interest of subject F0110. Over the yellow background are reported the tirs for the 2 time intervals of interest over the same date interval.

good tirs (>78%), but with moderate (both in value and in quantity) corrections. The increase of the quantity of the setting time intervals (from 8 to 9) does not highlight any specific improvement, suggesting that, probably, the addition of only 1 time interval (whose value is very close to the one that was present at the same time before the addition) is not very influential on this subject.

Regarding subject F4609, in the 2 date intervals, all time intervals' tirs show an improvement, except the third one which changes from 89,17% to 87,28% which is a small discrepancy. In the first date interval, the first, third and sixth time intervals are the ones for which the most corrections had to be done and the mean of the correction values differed from the set one only by a moderate amount (0,11 U/h more) for the first interval, but quite a lot (0,24 and 0,37 more, respectively) for the third and sixth ones. On the contrary, the second, fourth and fifth time intervals required fewer and smaller corrections (0,03 and 0,10 more and 0,04 less, respectively). Instead, in the second date interval, the corrections were generally less and, when more, still close to the previous ones, indicating an overall improvement of performance due to the new settings. Moreover, the differences were very low: at most 0,14 U/h more. This is another indicator of the efficiency of these parameters.

Regarding subject M1006, in the first and second time interval, the percentages, differences and tirs seem to suggest that fewer and smaller changes control blood glucose levels less efficiently. Still, the settings seem to be quite correct as the tirs are >88% and the percentage of corrections is <29%.

Regarding subject M2008, in the 2 date intervals, the first 2 time intervals and the fourth one show a tir improvement, while the other 2 show a tir deterioration. The percentage of changes and their magnitude remain pretty much the same throughout the 91-days span. This suggests that the BR setting (because this subject does not present ICR changes) of the second time interval (whose tir reaches 90,58%) is efficient, while the others should be corrected.

Regarding subject M3208, throughout the 3 date intervals of interest, the tirs improved at first and then worsened. This suggests that the BR change (from 0,70 to 1,00 and then 1,50) was too high in the third date interval and, so, not efficient for this individual. Additionally, the third date interval also shows a higher percentage of BR corrections, further proving the inefficiency of the set parameters.

Finally, regarding subject M4003, in the 2 date intervals, the tirs are very low (<39%), but, in comparison, the percentages of corrections are moderate (<43%). This suggests that the settings are so incorrect for this subject that the pump algorithm's corrections are not enough to counterbalance them.

Chapter 6

Conclusion

Hybrid closed-loop systems (like Tandem™'s t:slim X2™ insulin pump) enable better glycaemic control compared to the gold-standard insulin self-injection therapy, but they still cannot fully mimic the glucose metabolism of a healthy individual [49]. These systems have also been demonstrated to reduce the burden of diabetes management and improve the overall quality of life of patients who use them. So, even though they have the power to dramatically change the management of T1DM, further research is needed to improve their performances.

In this specific work, the focus on the improvement of the performances of this hybrid closed-loop system was on its technology called Control-IQ™ and its user-personalised parameters' settings.

A real-world study on the efficiency of the pump over one year of use by Breton et al. [73] demonstrated that t:slim X2™, along with its Control-IQ™ technology, keeps patients TIR at more than 70% and TBR at around 1%, proving its efficacy.

The main goal of this thesis was to gather the information that could lead towards the optimisation of the choice of the settings of the subjects' Personal Profiles performed by the clinicians, making this decision automatic and catered towards and personalised for each individual. Unfortunately, the available data was scarce and, for many subjects, pieces of information were missing. Moreover, the storage of said data is not adequate for satisfactory and complete studies and analysis. That means that the results of this thesis only represent an initial starting point towards what could be a definitive and reliable algorithm to be used by the clinical team. Still, they can be used to gather insight

- on how a useful dataset could be implemented and how the present one must be preprocessed,
- on the data itself, enabling the feasibility of comparison between different machine learning approaches and individuals,
- on the additional parameters that could be used to exhaustively complete the evaluations,
- on the meaning and differences of the pump's settings for each subject.

Now we know which one is the best preprocessing for the data that is currently available, but directly saving information differently would be preferable and more advantageous. First

thing first, having the pump itself assign the recorded basal rate to the closest time of recorded glucose value or repeat the closest recorded glucose value for the time of basal velocity change record would be time-saving. Definitely, subjects must be educated on the upload of their data and assured of how their effort could lead to a sharpened and straightforward glycaemic control, but if, at the same time, data was automatically made comparable it would be ideal. Also, having the possibility to download Excel files with only the columns of interest and having the time and date column already merged would be enormously convenient. Ultimately, the upload and download of the pump settings' changes should be optimised to report every change.

As could be expected, the machine learning approaches showed how algorithms that are oriented to imbalanced data work better on this dataset. They also highlighted how the change of predictors does not always have the same outcome: for example, while the addition of previous glycaemic values was excellent for the multivariate linear regression, it did not have the same results on the logistic regression. These differences could indicate that, based on the goal of the approach and on what we are trying to understand through it, for instance, the measurements taken 5 minutes before do not always have much influence on the model; so, it would be interesting to test the overall variability of closely recorded blood glucose levels for the subjects and compare that to these results. In addition, the output (the predicted labels) of the random forest used to preprocess the multivariate linear regression input files could be improved and this could, subsequently, lead to better results for the latter machine learning approach. Moreover, additional subject-focused work on the multivariate linear regression could result in fitter models and improved residual plots.

Regarding the pump settings' analysis, it is essential and crucial to work on more data belonging to more subjects. This study is too limited to be able to make any general assumption on its outcome and this is further demonstrated by the fact that observations made on some of the studied subjects are conflicting with the ones made on others. In fact, as expected, each subjects' Personal Profiles changes have different outcomes based on the individuals themselves. For example, sometimes the tirs were high because the Control-IQ™ corrections based on blood glucose levels were efficient, while some other times they were high because the set values were not that different from the subject's actual needs (as the corrections' values were quite close to the set ones). Other times, both the set parameter and the corrections were efficient in the glycaemic control. Thus, the analysis of the relationship between the number of changes and tirs values cannot be generalised and further research is needed.

Finally, generally using more data belonging to a longer time period of the same and of more subjects would be an excellent way to gain a deeper insight into the topics of interest. It would be also helpful to see if the year of birth and sex continue to not mark off any real difference in the results. Another predictor that would be interesting to analyse is the subject's weight which was a piece of information missing in the dataset used in this thesis: since it is part of the initial values set in t:slim X2™ insulin pump, maybe it could make more of an impact on the results.

Nevertheless, despite the limitations and issues of this work, this thesis presents an interesting closer look at the data recorded by Tandem™'s t:slim X2™ and a proper starting point towards deep and complete analysis to make an efficient and completely automated algorithm for the setting of the Personal Profiles a reality. The ultimate goal of this further research would be to take a step forward in the direction of a fully closed-loop system capable of realistically mimicking the healthy glucose metabolism, substituting the action

of the damaged pancreas. The development of such a system would be life-changing for all individuals affected by T1DM, but especially for children and adolescents (the focus of this thesis) who are left to face the psycho-physical repercussions while managing the disease and worrying about their future. Such a remarkable step in the advancement of diabetes care could make a true artificial pancreas a reality in the not-so-distant future.

Ringraziamenti

Eccomi arrivata al momento che sancisce la fine della stesura di questa tesi. Non potrei essere più fiera del mio percorso universitario e di questo elaborato che ne rappresenta la conclusione. Ho potuto esprimermi e dare sfogo alla mia creatività e alla mia passione per questo argomento e per la ricerca ingegneristica. Mi sento profondamente fortunata e voglio ringraziare tutte le persone che hanno reso possibile tutto questo.

Prima di tutto, un colossale grazie va al Dott. Palma, che ha accolto con entusiasmo la mia richiesta di lavorare sul diabete di tipo 1 e perchè ha quello stesso trascinante amore per l'ingegneria e lo sviluppo che ho anch'io. Grazie per ogni dubbio chiarito, per ogni domanda riposta e per ogni fiducia che ha riposto in me.

Un gigantesco grazie alla Dott.ssa Sabbatini che, nonostante stesse lavorando alla sua tesi di Dottorato, ha sempre trovato il tempo e il modo per aiutarmi nella parte sperimentale di questo lavoro e non potrei essergliene più grata.

Ultimo, ma non per importanza, un immenso grazie al Dott. Cherubini per ogni aiuto, ricerca e approfondimento che ha reso davvero completo questo lavoro. E per la passione che ha e ogni giorno dimostra per lo studio e la ricerca su questa malattia.

Un ringraziamento speciale va ai miei genitori. Se questo lavoro esiste, è anche grazie a voi. Perchè mi avete cresciuta curiosa e appassionata e avete stimolato e soddisfatto la mia fame e la mia sete di sapere. Perchè mi avete insegnato l'umiltà e, assieme a essa, l'ambizione. Perchè mi avete educata lasciandomi sempre l'indipendenza e la libertà di esplorare. E perchè è principalmente da voi che ho impararato a essere me stessa.

Un fortissimo grazie anche a mia sorella Clarissa e mio fratello Michelangelo. Perchè mi stimolano a essere ogni giorno una persona migliore e mi sono sempre accanto quando ne ho più bisogno, nonostante certe volte io risulti sfuggente, soprattutto, ma non solo, per gli impegni universitari. Grazie per la presenza, per le risate, per l'incoraggiamento e per scegliermi ogni giorno come vostra amica oltre che sorella.

Un enorme grazie a Guido per avermi supportata e sopportata durante gli innumerevoli progetti e tutto il processo della tesi. Mi hai offerto molto di più che un luogo di pace dove poter studiare e rilassarmi: mi hai dato la possibilità di scoprirmi, di evolvermi, di crescere. Te ne sarò per sempre grata.

Un gigantesco e collettivo grazie va a tutti i miei amici che da anni mi rendono ogni giorno più completa, che fanno parte di ogni mio traguardo e che li festeggiano al mio fianco.

A Camilla, la mia seconda sorella. Da 11 anni a questa parte continuiamo a sceglierci e non lo darò mai per scontato. In ogni mio passo, c'è un pezzo di te.

A Chiara, per avermi dischiuso una porticina che si affaccia sul tuo cuore e con l'augurio

di vivere insieme altre mille avventure da "scocchine".

A Serena perchè il nostro essere così diverse ci rende così unite ed è spesso una fonte di ispirazione, esattamente come ogni libro e ogni storia che condividiamo.

A Mabel, per le centinaia di videochiamate Leeds - Falconara Marittima e perchè ora che ci siamo ritrovate non ti lascio più.

A Giacomo, per tutte le nostre merende e per tutte le lacrime versate per il troppo ridere, ma anche per essere sempre aperto al confronto e al dialogo, su qualsiasi argomento.

A Michele, per la sua premura e per ogni chiacchierata sul cinema e su tutti gli altri piaceri della vita.

A Igor, per le infinite telefonate sul balcone, per le nostre inarrestabili passeggiate, per i nostri tour gastronomici, per i tuoi consigli e per ogni nostra battuta.

A Federica perchè mi hai accolta nella tua famiglia e a Montappone mi hai fatto sentire a casa.

Un caloroso ringraziamento va ai miei colleghi universitari. Grazie a Ilaria, la mia principale compagna di progetto, ma, anche e soprattutto, un'amica vera e sincera. Grazie per le giornate passate insieme, per le nostre chiacchierate e per tutti i paper e le presentazioni preparati insieme. Grazie a Maria Teresa, anche lei compagna di tanti progetti. Grazie per le telefonate durante il lockdown, per le confidenze e per ogni consiglio. Grazie a Lorenzo, ai nostri pranzi a mensa e alle tue improvvise domande filosofiche. Grazie anche a Delia e a Francesca per le pause caffè (anche per tutte quelle di cui siamo state derubate), per le chiacchiere al cambio lezione e per tutti i passi fatti assieme.

Infine, un generale grazie a tutti coloro che con un gesto, una parola o un'azione sono entrati a far parte della mia vita e mi hanno donato i litri di inchiostro con cui la sto scrivendo. *In case I don't see ya, good afternoon, good evening and goodnight.*

Bibliography

[1] D. Mellitus, "Diagnosis and classification of diabetes mellitus," *Diabetes care*, vol. 28, no. S37, pp. S5–S10, 2005.

[2] D. W. Cooke and L. Plotnick, "Type 1 diabetes mellitus in pediatrics," *pediatr Rev*, vol. 29, no. 11, pp. 374–84, 2008.

[3] N. Cho, J. Shaw, S. Karuranga, Y. Huang, J. da Rocha Fernandes, A. Ohlrogge, and B. Malanda, "IDF diabetes atlas: global estimates of diabetes prevalence for 2017 and projections for 2045," *Diabetes Research and Clinical Practice*, vol. 138, pp. 271–281, 2018.

[4] F. Dammacco, "Il trattamento integrato del diabete tipo 1 nel bambino e adolescente (i)," *Aprile 2000*, 2000.

[5] H. Baynes, "Classification, pathophysiology, diagnosis and management of diabetes mellitus," *Journal of Diabetes and Metabolism*, vol. 6, no. 5, pp. 1–9, 2015.

[6] R. Arcari, N. Grosso, A. Lezo, D. Boscolo, and P. Cavallo Perin, "Eziopatogenesi del diabete mellito di tipo 1," *Caleidoscopio Italiano*, vol. 126, pp. 3–64, 2007.

[7] S. A. Paschou, N. Papadopoulou-Marketou, G. P. Chrousos, and C. Kanaka-Gantenbein, "On type 1 diabetes mellitus pathogenesis," *Endocrine connections*, vol. 7, no. 1, pp. R38–R46, 2018.

[8] R. A. DeFronzo, E. Ferrannini, L. Groop, R. R. Henry, W. H. Herman, J. J. Holst, F. B. Hu, C. R. Kahn, I. Raz, G. I. Shulman *et al.*, "Type 2 diabetes mellitus," *Nature reviews Disease primers*, vol. 1, no. 1, pp. 1–22, 2015.

[9] T. A. Buchanan, A. H. Xiang *et al.*, "Gestational diabetes mellitus," *The Journal of clinical investigation*, vol. 115, no. 3, pp. 485–491, 2005.

[10] S. L. Kjos, T. A. Buchanan, J. S. Greenspoon, M. Montoro, G. S. Bernstein, and J. H. Mestman, "Gestational diabetes mellitus: the prevalence of glucose intolerance and diabetes mellitus in the first two months post partum," *American journal of obstetrics and gynecology*, vol. 163, no. 1, pp. 93–98, 1990.

[11] L. C. Martinez, D. Sherling, and A. Holley, "The screening and prevention of diabetes mellitus," *Primary Care: Clinics in Office Practice*, vol. 46, no. 1, pp. 41–52, 2019.

[12] M. Muggeo and R. C. Bonadonna, "La diagnosi di diabete," *Annali sanità pubblica*, 2002.

[13] A. D. Association *et al.*, "2. classification and diagnosis of diabetes: standards of medical care in diabetes—2018," *Diabetes care*, vol. 41, no. Supplement 1, pp. S13–S27, 2018.

[14] K. Ahmad, "Insulin sources and types: a review of insulin in terms of its mode on diabetes mellitus," *Journal of Traditional Chinese Medicine*, vol. 34, no. 2, pp. 234–237, 2014.

[15] D. Christie, "How do children and adolescents understand their diabetes?" *Practical Diabetes*, vol. 36, no. 4, pp. 117–120a, 2019.

[16] J. S. Borus and L. Laffel, "Adherence challenges in the management of type 1 diabetes in adolescents: prevention and intervention," *Current opinion in pediatrics*, vol. 22, no. 4, p. 405, 2010.

[17] M. A. Martínez-González, J. Salas-Salvadó, R. Estruch, D. Corella, M. Fitó, E. Ros, P. Investigators *et al.*, "Benefits of the mediterranean diet: insights from the predimed study," *Progress in cardiovascular diseases*, vol. 58, no. 1, pp. 50–60, 2015.

[18] J. Mattei, S. J. Bigornia, M. Sotos-Prieto, T. Scott, X. Gao, and K. L. Tucker, "The mediterranean diet and 2-year change in cognitive function by status of type 2 diabetes and glycemic control," *Diabetes care*, p. dc190130, 2019.

[19] C. Shugart, J. Jackson, and K. B. Fields, "Diabetes in sports," *Sports health*, vol. 2, no. 1, pp. 29–38, 2010.

[20] R. Streisand and M. Monaghan, "Young children with type 1 diabetes: Challenges, research, and future directions," *Current diabetes reports*, vol. 14, no. 9, p. 520, 2014.

[21] L. N. Jespersen, M. H. Vested, L. B. Johansen, and D. Grabowski, "Mirroring life of adolescents with type 1 diabetes—an outline of key aspects," *Diabetology*, vol. 2, no. 3, pp. 141–156, 2021.

[22] M. R. Frank, "Psychological issues in the care of children and adolescents with type 1 diabetes," *Paediatrics & child health*, vol. 10, no. 1, p. 18, 2005.

[23] G. Vespasiani, A. Nicolucci, and C. Giorda, "Epidemiologia del diabete," *Annali della sanità pubblica*, 2005.

[24] T. Hershey, N. Bhargava, M. Sadler, N. H. White, and S. Craft, "Conventional versus intensive diabetes therapy in children with type 1 diabetes: effects on memory and motor speed." *Diabetes care*, vol. 22, no. 8, pp. 1318–1324, 1999.

[25] A. M. Chao, K. E. Minges, C. Park, S. Dumser, K. M. Murphy, M. Grey, and R. Whittemore, "General life and diabetes-related stressors in early adolescents with type 1 diabetes," *Annual Meeting of the Society of Behavioral Medicine*, 2014.

[26] C. K. K. Kakleas, B. Kandyla and K. Karavanaki, "Psychosocial problems in adolescents with type 1 diabetes mellitus," *Diabetes amp; Metabolism*, vol. 1569, no. 5, pp. 339–424, 2009.

[27] E. O. Buschur, B. Glick, and M. K. Kamboj, "Transition of care for patients with type 1 diabetes mellitus from pediatric to adult health care systems," *Translational pediatrics*, vol. 6, no. 4, p. 373, 2017.

[28] M. Strand, A. Broström, and A. Haugstvedt, "Adolescents' perceptions of the transition process from parental management to self-management of type 1 diabetes," *Scandinavian Journal of Caring Sciences*, vol. 33, no. 1, pp. 128–135, 2019.

[29] K. Potter, H. Virtanen, P. Luca, D. Pacaud, A. Nettel-Aguirre, L. Kaminsky, and J. Ho, "Knowledge and practice of harm-reduction behaviours for alcohol and other illicit substance use in adolescents with type 1 diabetes," *Paediatrics & child health*, vol. 24, no. 1, pp. e51–e56, 2018.

[30] M. Edraki, M. Rambod, and Z. Molazem, "The effect of coping skills training on depression, anxiety, stress, and self-efficacy in adolescents with diabetes: a randomized controlled trial," *International journal of community based nursing and midwifery*, 2018.

[31] K. Rechenberg, M. Grey, and L. Sadler, "'anxiety and type 1 diabetes are like cousins': the experience of anxiety symptoms in youth with type 1 diabetes," *Research in Nursing Health*, vol. 41, no. 6, pp. 544–554, 2018.

[32] K. M. King, P. J. King, R. Nayar, and S. Wilkes, "Perceptions of adolescent patients of the "lived experience" of type 1 diabetes," *Diabetes Spectrum*, vol. 30, no. 1, pp. 23–35, 2017.

[33] V. A. Miller and D. Drotar, "Discrepancies between mother and adolescent perceptions of diabetes-related decision-making autonomy and their relationship to diabetes-related conflict and adherence to treatment," *Journal of pediatric psychology*, vol. 28, no. 4, pp. 265–274, 2003.

[34] K. P. Gürkan and Z. Bahar, "Living with diabetes: perceived barriers of adolescents," *Journal of Nursing Research*, vol. 28, no. 2, p. e73, 2020.

[35] J. Spencer, H. Cooper, and B. Milton, "The lived experiences of young people (13–16 years) with type 1 diabetes mellitus and their parents–a qualitative phenomenological study," *Diabetic Medicine*, vol. 30, no. 1, pp. e17–e24, 2013.

[36] D. Fonte, S. Colson, J. Côté, R. Reynaud, M.-C. Lagouanelle-Simeoni, and T. Apostolidis, "Representations and experiences of well-being among diabetic adolescents: Relational, normative, and identity tensions in diabetes self-management," *Journal of Health Psychology*, vol. 24, no. 14, pp. 1976–1992, 2019.

[37] K. Rechenberg, R. Whittemore, M. Holland, and M. Grey, "General and diabetes-specific stress in adolescents with type 1 diabetes," *Diabetes research and clinical practice*, 2017.

[38] T. Sanders, J. Elliott, P. Norman, B. Johnson, and S. Heller, "Disruptive illness contexts and liminality in the accounts of young people with type 1 diabetes," *Sociology of health & illness*, vol. 41, no. 7, pp. 1289–1304, 2019.

[39] E. Iversen, B.-C. H. Kolltveit, I. Hernar, J. Mårtensson, and A. Haugstvedt, "Transition from paediatric to adult care: a qualitative study of the experiences of young adults with type 1 diabetes," *Scandinavian journal of caring sciences*, vol. 33, no. 3, pp. 723–730, 2019.

[40] H. Starkman, K. Fisher, N. L. Pilek, G. Lopez-Henriquez, L. Lynch, and B. L. Bilkins-Morgis, "Listening to adolescents with uncontrolled diabetes, their parents and medical team." *Families, Systems, & Health*, vol. 37, no. 1, p. 30, 2019.

[41] S. J. Russell, F. H. El-Khatib, D. M. Nathan, K. L. Magyar, J. Jiang, and E. R. Damiano, "Blood glucose control in type 1 diabetes with a bihormonal bionic endocrine pancreas," *Diabetes care*, vol. 35, no. 11, pp. 2148–2155, 2012.

[42] J. D. R. F. C. G. M. S. Group, "Continuous glucose monitoring and intensive treatment of type 1 diabetes," *New England Journal of Medicine*, vol. 359, no. 14, pp. 1464–1476, 2008.

[43] D. C. Klonoff, "Continuous glucose monitoring: roadmap for 21st century diabetes therapy," *Diabetes care*, vol. 28, no. 5, pp. 1231–1239, 2005.

[44] C. Boughton and R. Hovorka, "Is an artificial pancreas (closed-loop system) for type 1 diabetes effective?" *Diabetic Medicine*, vol. 36, no. 3, pp. 279–286, 2019.

[45] A. Weisman, J.-W. Bai, M. Cardinez, C. K. Kramer, and B. A. Perkins, "Effect of artificial pancreas systems on glycaemic control in patients with type 1 diabetes: a systematic review and meta-analysis of outpatient randomised controlled trials," *The lancet Diabetes & endocrinology*, vol. 5, no. 7, pp. 501–512, 2017.

[46] E. Bekiari, K. Kitsios, H. Thabit, M. Tauschmann, E. Athanasiadou, T. Karagiannis, A.-B. Haidich, R. Hovorka, and A. Tsapas, "Artificial pancreas treatment for outpatients with type 1 diabetes: systematic review and meta-analysis," *bmj*, vol. 361, 2018.

[47] J. El Youssef, J. Castle, and W. K. Ward, "A review of closed-loop algorithms for glycemic control in the treatment of type 1 diabetes," *Algorithms*, vol. 2, no. 1, pp. 518–532, 2009.

[48] P. Keith-Hynes, B. Mize, A. Robert, and J. Place, "The diabetes assistant: a smartphone-based system for real-time control of blood glucose," *electronics*, vol. 3, no. 4, pp. 609–623, 2014.

[49] J. Fuchs and R. Hovorka, "Closed-loop control in insulin pumps for type-1 diabetes mellitus: safety and efficacy," *Expert Review of Medical Devices*, vol. 17, no. 7, pp. 707–720, 2020.

[50] T. Danne, T. Battelino, O. Kordonouri, R. Hanas, C. Klinkert, J. Ludvigsson, R. Barrio, C. Aebi, S. Gschwend, P.-E. Mullis *et al.*, "A cross-sectional international survey of continuous subcutaneous insulin infusion in 377 children and adolescents with type 1 diabetes mellitus from 10 countries," *Pediatric diabetes*, vol. 6, no. 4, pp. 193–198, 2005.

[51] R. M. Bergenstal, A. J. Ahmann, T. Bailey, R. W. Beck, J. Bissen, B. Buckingham, L. Deeb, R. H. Dolin, S. K. Garg, R. Goland *et al.*, "Recommendations for standardizing glucose reporting and analysis to optimize clinical decision making in diabetes: the ambulatory glucose profile (agp)," 2013.

[52] T. Battelino, T. Danne, R. M. Bergenstal, S. A. Amiel, R. Beck, T. Biester, E. Bosi, B. A. Buckingham, W. T. Cefalu, K. L. Close *et al.*, "Clinical targets for continuous glucose monitoring data interpretation: recommendations from the international consensus on time in range," *Diabetes care*, vol. 42, no. 8, pp. 1593–1603, 2019.

[53] E. J. Gallagher, D. Le Roith, and Z. Bloomgarden, "Review of hemoglobin a1c in the management of diabetes," *Journal of diabetes*, vol. 1, no. 1, pp. 9–17, 2009.

[54] J. G. Shepard, A. Airee, A. W. Dake, M. S. McFarland, and A. Vora, "Limitations of a1c interpretation," *South Med J*, vol. 108, no. 12, pp. 724–9, 2015.

[55] T. Danne, R. Nimri, T. Battelino, R. M. Bergenstal, K. L. Close, J. H. DeVries, S. Garg, L. Heinemann, I. Hirsch, S. A. Amiel *et al.*, "International consensus on use of continuous glucose monitoring," *Diabetes care*, vol. 40, no. 12, pp. 1631–1640, 2017.

[56] V. Klimontov, "Impact of glycemic variability on cardiovascular risk in diabetes," *Kardiologiia*, no. 10, pp. 80–87, 2018.

[57] J. Gao, P. Yi, Z. Chi, and T. Zhu, "A smart medical system for dynamic closed-loop blood glucose-insulin control," *Smart Health*, vol. 1, pp. 18–33, 2017.

[58] P. Herrero, J. Bondia, O. Adewuyi, P. Pesl, M. El-Sharkawy, M. Reddy, C. Toumazou, N. Oliver, and P. Georgiou, "Enhancing automatic closed-loop glucose control in type 1 diabetes with an adaptive meal bolus calculator–in silico evaluation under intra-day variability," *Computer methods and programs in biomedicine*, vol. 146, pp. 125–131, 2017.

[59] P. Y. Benhamou, E. Huneker, S. Franc, M. Doron, and G. Charpentier, "Customization of home closed-loop insulin delivery in adult patients with type 1 diabetes, assisted with structured remote monitoring: the pilot wp7 diabeloop study," *Acta diabetologica*, vol. 55, no. 6, pp. 549–556, 2018.

[60] M. R. Askari, I. Hajizadeh, M. Rashid, N. Hobbs, V. M. Zavala, and A. Cinar, "Adaptive-learning model predictive control for complex physiological systems: Automated insulin delivery in diabetes," *Annual Reviews in Control*, 2020.

[61] P. Colmegna, F. D. Bianchi, and R. Sánchez-Peña, "Automatic glucose control during meals and exercise in type 1 diabetes: Proof-of-concept in silico tests using a switched lpv approach," *IEEE Control Systems Letters*, vol. 5, no. 5, pp. 1489–1494, 2020.

[62] G. M. Steil, "Algorithms for a closed-loop artificial pancreas: the case for proportional-integral-derivative control," *Journal of diabetes science and technology*, vol. 7, no. 6, pp. 1621–1631, 2013.

[63] N. S. Tyler and P. G. Jacobs, "Artificial intelligence in decision support systems for type 1 diabetes," *Sensors*, vol. 20, no. 11, p. 3214, 2020.

[64] G. Vishwakarma, A. Sonpal, and J. Hachmann, "Metrics for benchmarking and uncertainty quantification: Quality, applicability, and best practices for machine learning in chemistry," *Trends in Chemistry*, 2021.

[65] P. Dangeti, *Statistics for machine learning*. Packt Publishing Ltd, 2017.

[66] J. Zhang and X. Wang, "Selecting the best regression equation via the p-value of f-test," *Metrika*, vol. 46, no. 1, pp. 33–40, 1997.

[67] D. A. York *et al.*, "Metabolic regulation of food intake." *Metabolic regulation of food intake.*, vol. 34, no. Supplement 25, pp. 33–39, 1990.

[68] D. Unwin, G. Livesey, and D. Haslam, "It is the glycaemic response to, not the carbohydrate content of food that matters in diabetes and obesity: The glycaemic index revisited," *Journal of Insulin Resistance*, vol. 1, no. 1, pp. 1–9, 2016.

[69] S. Wen, C. Wang, M. Gong, and L. Zhou, "An overview of energy and metabolic regulation," *Science China Life Sciences*, vol. 62, no. 6, pp. 771–790, 2019.

[70] S. S. Farabi, D. W. Carley, and L. Quinn, "Glucose variations and activity are strongly coupled in sleep and wake in young adults with type 1 diabetes," *Biological research for nursing*, vol. 19, no. 3, pp. 249–257, 2017.

[71] K. C. Summa and F. W. Turek, "Chronobiology and obesity: interactions between circadian rhythms and energy regulation," *Advances in nutrition*, vol. 5, no. 3, pp. 312S–319S, 2014.

[72] B. Trümper, K. Reschke, and J. Molling, "Circadian variation of insulin requirement in insulin dependent diabetes mellitus the relationship between circadian change in insulin demand and diurnal patterns of growth hormone, cortisol and glucagon during euglycemia," *Hormone and metabolic research*, vol. 27, no. 03, pp. 141–147, 1995.

[73] M. D. Breton and B. P. Kovatchev, "One year real-world use of the control-iq advanced hybrid closed-loop technology," *Diabetes Technology & Therapeutics*, 2021.