Development Of Predictive Model for Diabetes Medication Adherence Using Algorithms.

By

Anum Mustafa

(Registration No: 00000399938)

Department of Biomedicine

Atta-ur-Rahman School of Applied biosciences (ASAB)

National University of Sciences & Technology (NUST)

Islamabad, Pakistan

(2024)

Development Of Predictive Model for Diabetes Medication Adherence Using Algorithms.

By

Anum Mustafa

(Registration No: 00000399938)

A thesis submitted to the National University of Sciences and Technology, Islamabad,

in partial fulfillment of the requirements for the degree of

Master of Science in
Healthcare Biotechnology

Supervisor: Prof. Dr. Peter John

Atta-ur-Rahman School of Applied Biosciences

National University of Sciences & Technology (NUST)

Islamabad, Pakistan

(2024)

FORM TH-4

National University of Sciences & Technology

MS THESIS WORK

We hereby recommend that the dissertation prepared under our supervision by **Anum Mustafa, 00000399938** Titled: ***"Development of Predictive Model for Diabetes Medication Adherence using Algorithms"*** be accepted in partial fulfillment of the requirements for the award of **MS Degree in Healthcare Biotechnology** with (A grade).

Examination Committee Members

1. Name: **Prof. Dr. Attya Bhatti** Signature: _____
2. Name: **Dr. Maria Shabbir** Signature: _____
3. Name: **Saeed Mughal** Signature: _____

Supervisor's name: **Prof. Dr. Peter John** Signature: _____ Date: 05/08/2024

Prof. Dr. Peter John
Deptt of Healthcare Biotechnology
Atta-ur-Rahman School of Applied Biosciences (ASAB), NUST Islamabad

Head of Department: _____ Date: 22/08/2024

COUNTERSINGED

Date: 22/08/2024

Dean/Principal

A/Principal & Dean
Atta-ur-Rahman School of Applied Biosciences (ASAB), NUST, Islamabad

THESIS ACCEPTANCE CERTIFICATE

Certified that content and forms of thesis entitled "***Development of Predictive Model for Diabetes Medication Adherence using Algorithms***" submitted by **Anum Mustafa** has been found excellent for the requirement of the degree.

Supervisor: _____

Prof. Dr. Peter John
Deptt of Healthcare Biotechnology
Atta-ur-Rahman School of Applied Biosciences (ASAB), NUST Islamabad

Professor Dr. Peter John

Atta-ur-Rahman School of Applied Biosciences (ASAB)

National University of Sciences and Technology (NUST)

Head of Department: _____ 22/08/2024

Atta-ur-Rahman School of Applied Biosciences (ASAB)

National University of Sciences and Technology (NUST)

Principal ASAB: _____

A/Principal & Dean
Atta-ur-Rahman School of Applied Biosciences (ASAB), NUST, Islamabad

Atta-ur-Rahman School of Applied Biosciences (ASAB)

National University of Sciences and Technology (NUST)

AUTHOR'S DECLARATION

I Anum Mustafa hereby state that my MS thesis titled "Development Of Predictive Models for Diabetes Medication Adherence Using Algorithms" is my own work and has not been submitted previously by me for taking any degree from National University of Sciences and Technology, Islamabad or anywhere else in the country/ world.

At any time if my statement is found to be incorrect even after I graduate, the university has the right to withdraw my MS degree.

Name of Student: Anum Mustafa

Date: _____

DEDICATION

To my mother who gave me the confidence to achieve everything I have today (though I carry a piece of your heart with me but still there is a void without you) to my father who gave me all the support I needed and always being there for me and to my siblings for always cheering me on and encouraging me.

ACKNOWLEDGEMENTS

"I am extremely grateful to Allah Almighty for completing my thesis and always guiding me to the best in most unexpected ways. I would like to express my sincere gratitude to Dr. Peter John for his invaluable guidance, unwavering support, and expert supervision throughout this project. His mentorship and encouragement have been instrumental in shaping this research. I would like to express my heartfelt gratitude to my father for always prioritizing my future. I also want to thank my family for their constant love, understanding, and encouragement throughout this journey. Their support has been a source of strength and motivation, especially my brothers (Faizan and Bhai Nadeem) who guided me throughout and my sister Rabbia for all the support and Tehrim, Api Maryam, Ayesha and Bhai Farrukh for all the encouragement. I would like to thank federal government polyclinic hospital for collaborating with us. A special thanks to the participants whose data contributed to this study. Their participation was essential to the success of our research. I would like to express my gratitude to my GEC members (Saeed Mughal, Dr Attya Bhatti, Dr Maria Shabbir) and my MS lab fellows (Hijab, Zainab, Afnan, Pakeeza, Maimona, Nisha) for their unwavering support.

Anum Mustafa.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS **7**

TABLE OF CONTENTS **8**

LIST OF TABLES **10**

LIST OF FIGURES **11**

LIST OF SYMBOLS, ABBREVIATIONS AND ACRONYMS **12**

ABSTRACT **14**

CHAPTER 1: INTRODUCTION **15**
1.1 Type 2 Diabetes Mellitus **15**
1.2 Pathophysiology of Type 2 diabetes mellitus **17**
1.3 Diagnosis test for Diabetes mellitus. **19**
1.4 Medication Adherence. **21**
1.5. Metformin **23**
1.5.1. Mechanism of action against diabetes 24
1.6. AI and Predictive models in diabetes medication adherence **26**

CHAPTER 2: LITERATURE REVIEW **28**
2.1 Hemoglobin A1c (HbA1c) as a Marker and Predictor in Type 2 Diabetes Mellitus **30**
2.2. Current Approaches to Medication Adherence in T2DM **31**
2.2.1. Empowering Patients through Education 32
2.2.2. Role of Gene Expression Analysis in Predictive Models 34
2.2.3. Existing Predictive Models for T2DM Medication Adherence 35
2.3. Dysregulation of GLUT4 Expression in Type 2 Diabetes Mellitus (T2DM) **37**
2.4. Metformin's Mechanisms: Enhancing Glucose Metabolism Through GLUT4 Translocation **37**

CHAPTER 3: METHODOLY **40**
3.1. Sampling and Data Collection **40**
3.1.1. Data analysis 40
3.2 Expression Analysis **40**
3.2.1 RNA Extraction 40
3.2.2 Complementary DNA (cDNA) synthesis 41
3.2.3 Semi quantitative PCR 41
3.2.4 Gel Electrophoresis 42
3.2.5 Primers 42

3.2.6 Real Time PCR 43
3.2.7. Statistical analysis and Results validation 43
3.3. Web based application Development 43
3.3.1. Import essential dependencies for data analysis/preprocessing and machine learning 43
3.3.2. Read the .csv file (read the data set) 43
3.3.3. Printing the head of the dataset 43
3.3.4. Dimensions of dataset (Rows, Columns) 44
3.3.5. Statistics of dataset of each column 44
3.3.6. Dropping rows having null values from your dataset 44
3.3.7. Encoding 44
3.3.8. Converting all dataset into float type 44
3.3.9. Separating the output feature and input features 44
3.3.10. Machine learning model 45
3.3.11. Checking the credibility 45
3.3.12. Prediction system 45
3.3.13. Created a streamlit app 45
3.4. Accuracy of application 45

CHAPTER 4: RESULTS 47
4.1. Demographic and clinical characteristics of study subject 47
4.2. Statistics of the selected categorical predictors 52
4.3. Prediction performance of the ML model 53
4.4. Development of application 54
4.5. Primer designing 54
4.5.1. Primers 54
4.5.2. Primer optimization 56
4.6. Expression analysis by Real time PCR 56
4.7. Incorporation in application 58
4.8. Success rate of application 59

CHAPTER 5: DISSCUSION 60

REFERENCES 62

LIST OF TABLES

Table 1.1 Diagnosis tests for Diabetes mellitus……..20

Table 1.2. Common medication prescribed for diabetes…….22

Table 3 Performance of the decision tree regressor.....54

LIST OF FIGURES

Figure 1-1 Insulin in T2DM, pancreatic β-cells are damaged by hyperglycemia, hyperlipidemia, cytokines, and amyloids. Although pancreatic β-cells produce insulin, the insulin level is insufficient to compensate for insulin resistance, resulting in a relative insulin deficiency, leading to hyperglycemia. 16
Figure 1.2. Signaling pathways involved in insulin secretion in β-cells in physiological conditions…….11
Figure 1.3. Insulin secretion in pancreatic β-cells…….12
Figure 2.1. personalized diabetes care……15
Figure 2.2. HbA1c to estimated average glucose (MG/DL)…….24
Figure 2.3. A systems epidemiology approach investigates how interactions between the exposome (comprising all non-genetic exposures) and measurable components of the human physiome contribute to our understanding…….28
Figure 2.4. Metformin treatment led to inhibited glucose uptake, which was demonstrated through the use of phloretin, an inhibitor of glucose transporters (data not shown).
Figure 4.2. Comparison of HbA1c and frequency of complications. 47
Figure 4.3. Relative frequency of complications in diabetes patients 48
Figure 4.4. Age distribution of patients with T2DM 48
Figure 5.4. Comparison of glycemic control in patients diagnosed with T2DM 49
Figure 4.6. Comparison between insulin dependency and HbA1c levels..... 49
Figure 4.7. Comparison of compliance to the medication and HbA1c levels in patients. 50
Figure 4.8. Relation of dosage adjustment and HbA1c levels..... 51
Figure 4.9. Comparison of HbA1c levels in male and female..... 51
Figure 4.10. Comparison between self-management and HbA1c levels. 52
Figure 4.11. comparison of physical activity and HbA1c levels. 52
Figure 12. Figure of gel containg ladder of 1kbp in 1st well and beta actin in 2nd to 4th well while Glut-4 in 12th to 14th wells. 56

LIST OF SYMBOLS, ABBREVIATIONS AND ACRONYMS

T2DM	Type 2 diabetes mellitus
IR	Insulin resistance
GLUT	Glucose transporters
ATP	Adenosine Triphosphate
ADP	Adenosine Diphosphate
cAMP	Cyclic Adenosine Monophosphate
RYR	Ryanodine Receptor
P2Y	Purinergic Receptor P2Y
P2X	Purinergic Receptor P2X
ER	Endoplasmic Reticulum
IP3	Inositol Triphosphate
FFA	Free Fatty Acid
UPR	Unfolded Protein Response
IAAP	International Association of Applied Psychology
IL	Interleukin

FDA	Food and Drug Administration
AI	Artificial Intelligence
HbA1c	Hemoglobin A1c
ML	Machine Learning
DL	Deep Learning

ABSTRACT

Type 2 diabetes (T2DM) is a persistent metabolic disorder characterized by its complex interaction with both environmental factors and genetic predisposition. It poses a significant global health challenge, steadily increasing in prevalence and presenting substantial difficulties for healthcare systems and individuals alike. Managing T2DM effectively requires a comprehensive approach involving lifestyle changes, medications, and sometimes insulin therapy to regulate blood sugar levels and mitigate associated complications. In Pakistan, the prevalence of Type 2 Diabetes Mellitus (T2DM) has risen to concerning levels, posing a formidable health issue nationwide.

Our study centers on developing a web-based application utilizing a decision tree regressor to forecast patients' HbA1c levels and medication adherence. Validation of the application includes analyzing gene expression of GLUT4. Additionally, association studies involving expression are conducted to potentially integrate this markers into future models.

The model achieved an accuracy of 80% with metrics showing a mean squared error of 0.143, mean absolute error of 0.15, and an R^2 value of 0.88. Future studies could explore incorporating GLUT4 expressions to enhance predictive accuracy further.

Keywords: Machine learning models, HbA1c prediction, medication adherence, Decision tree regressor, GLUT4 expression, Diabetes prediction by ML

CHAPTER 1: INTRODUCTION

The World Health Organization (WHO) reports that over 420 million people are currently living with diabetes, a number expected to rise to 578 million by 2030. Shockingly, half of adults with type 2 diabetes are undiagnosed. Despite insulin's discovery a century ago, roughly half of those who need insulin treatment for type 2 diabetes still do not get it. (Kaku, 2010). More than 90% of diabetes cases belong to type 2 diabetes mellitus (T2DM), which is marked by insufficient insulin secretion from pancreatic islet β-cells, tissue insulin resistance (IR), and an insufficient compensatory insulin response. (Scheen, 2003). Type 2 diabetes mellitus (T2DM) is a multifaceted condition influenced by a mix of genetic, epigenetic, and environmental factors. (Himanshu et al., 2020; Scheen, 2003).

In 2019, the International Diabetes Federation (IDF) reported that diabetes caused 4.2 million deaths and affected 463 million adults aged 20 to 79. This figure is expected to rise to around 700 million by 2045.(Kaku, 2010).

1.1 Type 2 Diabetes Mellitus

The development of Type 2 Diabetes Mellitus (T2DM) is mainly due to two factors: impaired insulin secretion by pancreatic β-cells and the inability of insulin-sensitive tissues to respond to insulin properly. Insulin resistance, a condition where cells fail to respond effectively to insulin, leads to the inhibition of lipoprotein lipase activity and changes in lipid and apolipoprotein metabolism. This results in increased lipolysis and the release of free fatty acids (FFA) into the bloodstream, contributing to the disease's progression. Obesity is a significant factor in the development of insulin resistance. Even mild obesity (with a BMI over 25) increases the risk of diabetes by 4 to 5 times. (Goldstein, 2002). Additionally, there is a significant genetic predisposition to T2DM, with the risk increasing by up to 40% if a close relative has the condition. This risk doubles if the mother has the disease. Dysfunctional β-cells also contribute to reduced

insulin secretion, which is insufficient to maintain normal blood glucose levels. (Galicia-Garcia et al., 2020).

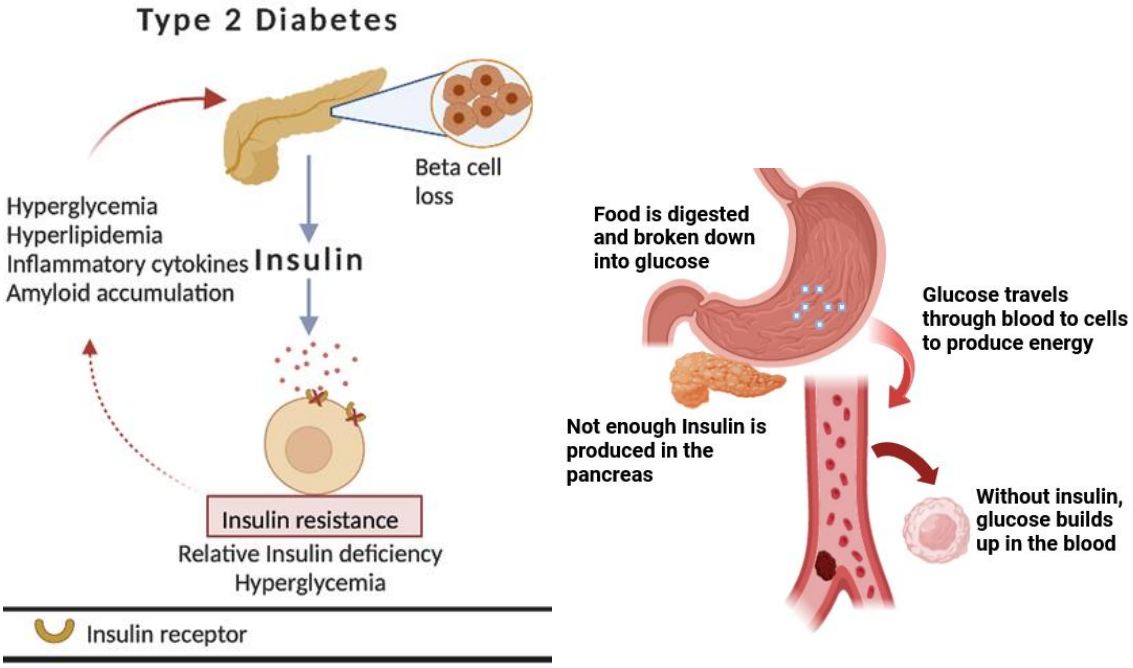


Figure 1-1 *Insulin in T2DM, pancreatic β-cells are damaged by hyperglycemia, hyperlipidemia, cytokines, and amyloids. Although pancreatic β-cells produce insulin, the insulin level is insufficient to compensate for insulin resistance, resulting in a relative insulin deficiency, leading to hyperglycemia.*

There are numerous risk factors associated with Type 2 Diabetes, including:

1. Stress-related factors:

- Overeating, particularly excessive consumption of simple sugars

- Smoking

- Increased alcohol intake

- Nervous and endocrine system disorders, such as elevated cortisol levels and abnormalities in sex hormone secretion

2. Reduced energy expenditure due to lack of exercise

3. Genetic predisposition

4. Aging

1.2 Pathophysiology of Type 2 diabetes mellitus

Cellular integrity and well-regulated mechanisms are essential for proper β-cell function. β-cells create pre-proinsulin, which is converted into proinsulin during synthesis. Mature insulin is stored in granules and released in response to high glucose levels. Glucose enters β-cells through glucose transporter 2 (GLUT2), initiating glucose catabolism, raising the intracellular ATP/ADP ratio, and allowing Ca2+ entry. This triggers the exocytosis, priming, and fusion of insulin-containing granules to the plasma membrane.

Several cellular signals influence insulin release from β-cells, including the ryanodine receptor (RYR), which amplifies Ca2+ signals and mediates Ca2+-driven Ca2+ release. Other significant players in the insulin secretion process are cAMP, extracellular ATP, and purinergic signaling via P2Y and P2X receptors. RYR enhances Ca2+ signals when activated by messenger molecules from food metabolism or ligand binding. Additionally, P2Y receptors can mediate intracellular Ca2+ mobilization in response to the synthesis of inositol-1,4,5-triphosphate. (Ostenson, 2001).

The traditional link between β-cell malfunction and β-cell death is likely due to a complex network of interactions between environmental factors and various molecular pathways in cell biology. In conditions like obesity, hyperglycemia, and hyperlipidemia, which favor insulin resistance and chronic inflammation, β-cells are particularly vulnerable to islet integrity loss due to their genetic predisposition to stress. Toxic stresses such as inflammation, ER stress, metabolic/oxidative stress, and amyloid stress can compromise β-cell function.

Obesity-related lipotoxicity, glucose toxicity, and glucolipotoxicity induce oxidative and metabolic stress that damage β-cells. High levels of saturated FFAs can trigger the unfolded protein response (UPR) pathway by inhibiting sarco/endoplasmic reticulum Ca2+ ATPase (SERCA), which is essential for ER Ca2+ mobilization, activating IP3 receptors, or directly disrupting ER homeostasis. Prolonged high glucose levels in β-cells increase proinsulin biosynthesis and islet amyloid polypeptides (IAAP), leading to the accumulation of misfolded insulin and IAAP, as well as increased reactive oxygen species (ROS) formation through oxidative protein folding.

These alterations disrupt the normal mobilization of ER Ca2+, promoting proapoptotic signals, degradation of proinsulin mRNA, and the release of IL-1, which attracts macrophages and exacerbates local islet inflammation.(Chatterjee et al., 2017; Galicia-Garcia et al., 2020).

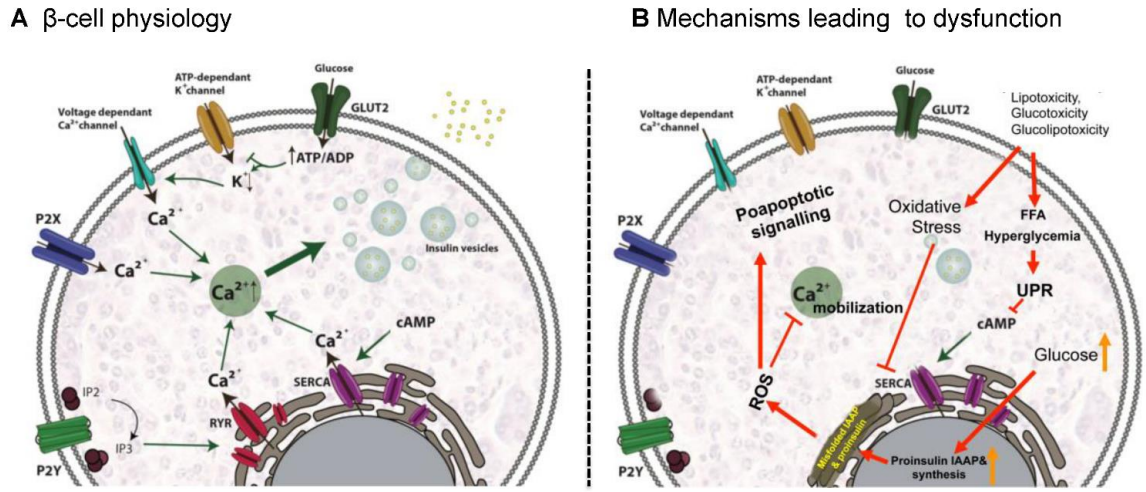


Figure 1-2: *Signaling pathways involved in insulin secretion in β-cells in physiological conditions.*

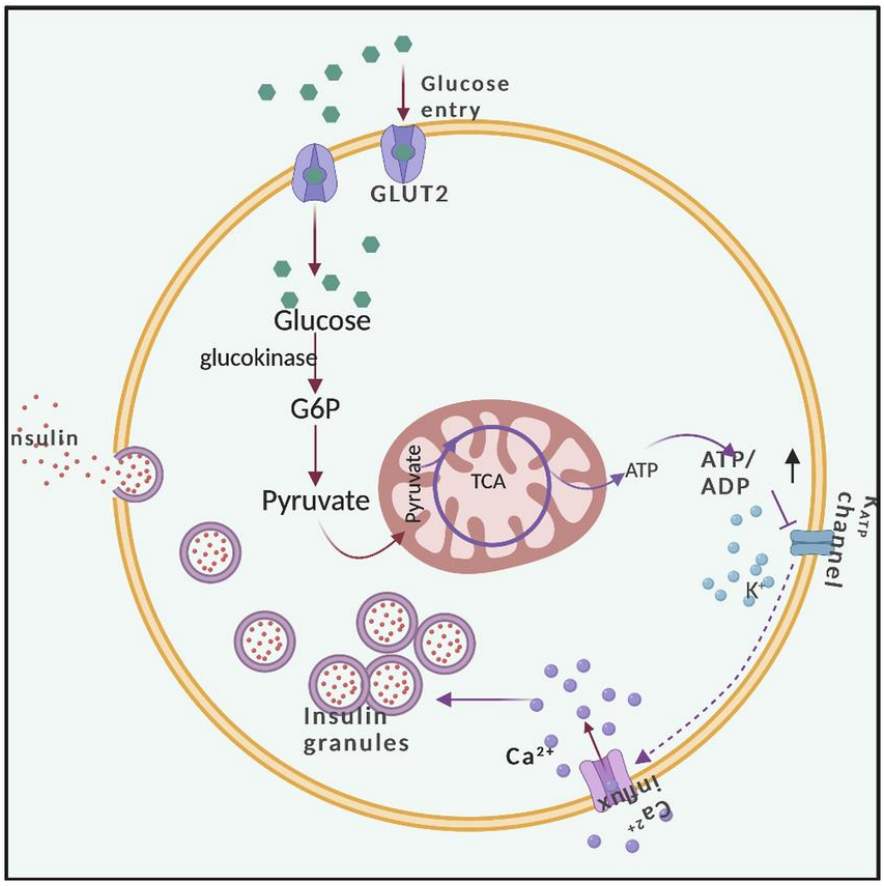


Figure 1.3. *Insulin secretion in pancreatic β-cells is a regulated process involving several signaling pathways under normal conditions. When glucose levels rise, glucose enters β-cells through GLUT2 transporters on the cell membrane. Inside the cell, glucose is converted to glucose-6-phosphate (G6P) by glucokinase. G6P undergoes glycolysis to produce pyruvate, which enters the mitochondria and generates ATP via the TCA cycle. This increase in ATP levels raises the ATP/ADP ratio, causing closure of ATP-dependent potassium channels and membrane depolarization. Depolarization prevents potassium efflux, leading to calcium influx into the cell. Elevated calcium levels trigger insulin granule exocytosis, releasing insulin into the bloodstream.*

1.3 Diagnosis test for Diabetes mellitus.

In order to monitor blood glucose levels and evaluate glycemic control, a number of tests are frequently employed in the diagnosis of type 2 diabetes mellitus (T2DM). To diagnose type 2 diabetes mellitus and evaluate glycemic control in people with diabetes

who are already diagnosed, these diagnostic tests are utilized in conjunction with clinical symptoms and risk factors. Depending on variables like patient characteristics, clinical presentation, and local recommendations, several tests may be selected. It's crucial to consult medical experts about appropriate testing and result interpretation (Florkowski, 2013).

	Hemoglobin A1C	Fasting Blood Sugar Test	Oral Glucose Tolerance Test (OGTT)	Random Blood Sugar Test
Normal	Less than 5.7% (39 mmol/mol	Less than 100 mg/dL (5.6 mmol/L)	A 2-hour blood glucose level of less than 140 mg/dL (7.8 mmol/L)	
Prediabetes	5.7-6.4% (39-47 mmol/mol)	100-125 mg/dL (5.6-6.9 mmol/L)	A 2-hour blood glucose level of 140-199 mg/dL (7.8-11.0 mmol/L)	
Diabetes	6.5% (48 mmol/mol) or higher on 2 separate occasions	126 mg/dL (7 mmol/L) or higher on 2 separate tests	A 2-hour blood glucose level of 200 mg/dL (11.1 mmol/L) or higher	200 mg/dL (11.1 mmol/L) or higher

Table 1.1 Diagnosis tests for Diabetes mellitus

1.4 Medication Adherence.

Diabetes management is a challenging, lifelong procedure that calls for significant patient effort. More than any other healthcare professional, the patient is the key to effective management. A few significant problems might arise from poor management. Due to this, diabetes patients' non-adherence to therapeutic regimens has remained a challenge for both patients and healthcare professionals.

The quick and cost-effective data production caused by the significant expansion of biotechnology and high throughput computing has pushed computational biology research into the big data sphere.

The capacity of the appropriate viewpoints to discover the data model formation contributes to the efficacy and consistency of these procedures. Diabetes mellitus (DM) is a life-threatening disease that can lead to a variety of other health issues. (Lindenmeyer et al., 2006).

In order to enhance glycemic control, prevent complications, and lower total health expenditures, people with T2DM must adhere to their medications well, according to studies. Medication adherence can be greatly increased by interventions like integrative health coaching. Medication adherence in T2DM patients may be influenced by a variety of variables, such as the severity of the condition, mental health, anxiety, sadness, irritability, smoking, cost, and so forth. These variables change depending on the study's design and geographic location. That may be resulted from the differences in the selection of variables, patients' income levels, education levels, cultural characteristics, living habits, and so on (Brunton & Polonsky, 2017).

Higher risk of type 2 diabetes is linked to higher body weight. Increased morbidity and mortality are linked to both type 2 diabetes and obesity. Although food modification, exercise, and lifestyle changes can help manage type 2 diabetes, there are additional

There are various medication options available to help maintain desired blood sugar levels in the treatment of Type 2 Diabetes Mellitus (T2DM), which primarily focus on:

1. Raising insulin levels:

- Oral medications that stimulate insulin secretion: These include insulin secretagogues like sulfonylureas.
- Direct insulin administration.

2. Increasing tissue sensitivity:

- Insulin sensitizers: These include biguanide metformin and thiazolidinediones (TZDs).

3. Reducing the rate of carbohydrate absorption from the gastrointestinal tract:

- α-glucosidase inhibitors: Such as acarbose.
- Agents that decrease gastric motility (Tanase et al., 2020).

	Mechanism of action	**Molecular target**	**% Decrease in HbA_{1c}**
Sulfonylurea (SU)/repaglinide	Increase in insulin secretion	SU receptor/ATP-K^+ channel	1.5–2.0
Metformin	Increase in muscle insulin sensitivity; decrease in hepatic glucose output	Unknown	1.5–2.0
Troglitazone (T)/rosiglitazone	Increase in muscle insulin sensitivity;	PPARγ	1.0–1.2

(R)/pioglitazone	decrease in hepatic glucose output		
Acarbose	Decrease in GI absorption	α-glucosidase	0.7–1.0

Table 1.2. Common medication prescribed for diabetes.

1.5. Metformin

The biguanide antihyperglycemic medication metformin is used to manage blood sugar levels in Type 2 Diabetes Mellitus (T2DM) in conjunction with diet and exercise. As the first-line treatment for T2DM, metformin is valued for its ability to lower blood glucose levels without causing hypoglycemia, classifying it as an antihyperglycemic medication. Often referred to as an "insulin sensitizer," metformin effectively reduces insulin resistance and plasma fasting insulin levels in a clinically relevant manner. Another notable benefit of metformin is its association with modest weight loss, making it an attractive option for obese T2DM patients. Metformin received FDA approval in the United States in 1995, after initially being approved in Canada in 1972.(Rena et al., 2013, 2017).

The mechanisms of action of metformin differ from those of other oral hypoglycemic medications. Metformin lowers blood sugar levels and improves insulin sensitivity by:

1. Reducing intestinal glucose absorption.

2. Decreasing hepatic glucose production (gluconeogenesis).

3. Enhancing peripheral glucose uptake and utilization.

Metformin is known to inhibit mitochondrial complex I activity, which is believed to be a key factor in its potent anti-diabetic effects. These actions collectively lower blood sugar levels, helping to manage Type 2 Diabetes Mellitus and improve glycemic control.(Grisouard et al., 2010; Hundal et al., 2000).

1.5.1. Mechanism of action against diabetes

After consumption, metformin is taken up into hepatocytes (liver cells) by the organic cation transporter-1 (OCT1). Due to its positive charge, metformin accumulates in cells and mitochondria, driven by the membrane potentials between the plasma membrane and the mitochondrial inner membrane. By inhibiting mitochondrial complex I, metformin reduces mitochondrial ATP production, raising cytoplasmic ADP: ATP and AMP:ATP ratios. These changes activate AMP-activated protein kinase (AMPK), an enzyme crucial for regulating glucose metabolism. Additionally, AMPK can be activated through a lysosomal mechanism involving other activators. (Klip & Leiter, 1990; LaMoia & Shulman, 2021).

Increased AMP:ATP ratios lead to the inhibition of fructose-1,6-bisphosphatase, which suppresses gluconeogenesis. This also results in the inhibition of adenylate cyclase, reducing cyclic adenosine monophosphate (cAMP) production, a molecule involved in cell signaling. Metformin's activation of AMPK inhibits fat production and promotes fat oxidation by phosphorylating two isoforms of the acetyl-CoA carboxylase enzyme, which lowers hepatic lipid storage and improves insulin sensitivity in the liver.

Metformin also enhances anaerobic glucose metabolism in enterocytes (intestinal cells), reducing net glucose absorption and increasing lactate transport to the liver. Recent studies suggest that the gut may be the primary site of action for metformin, rather than the liver, in patients with Type 2 Diabetes Mellitus. Additionally, metformin may stimulate glucose metabolism by increasing glucagon-like peptide-1 (GLP-1) levels and improving gastrointestinal glucose utilization. (Hundal et al., 2000; Jackson et al., 1987).

In addition to the pathways mentioned, metformin's mechanism of action may involve other mechanisms that have been extensively studied in recent years. While its primary effects are understood, ongoing research continues to explore additional pathways and interactions that contribute to its overall impact on glucose metabolism and insulin sensitivity. (Foretz et al., 2019).

1.6. Artificial intelligence in diabetes treatment

Artificial intelligence (AI) has the potential to significantly enhance diabetes management and treatment in several ways:

1. Risk Identification: AI algorithms can analyze patient data, including medical history, genetics, and lifestyle factors, to identify individuals at risk of developing diabetes.
2. Early Diagnosis: By examining blood glucose levels and patient symptoms, machine learning models can aid in the early detection of diabetes, allowing for prompt intervention and better disease management.

These applications can lead to more personalized and proactive care, potentially improving outcomes and reducing the burden of diabetes. (Ellahham, 2020; Guan et al., 2023). Machine learning (ML) and deep learning (DL) are prominent AI methodologies used in various applications, including diabetes management. In “supervised machine learning”, a system is trained using a dataset of labeled examples, which means each data point in the training set is already annotated with the correct outcome or classification. The models learn from these reference instances to make accurate predictions or classifications on new, unseen data based on the patterns they have learned. This approach relies on having a well-curated database with relevant and high-quality features to enable the system to effectively generalize and apply its learning to new cases. (Gautier et al., 2021; Nomura et al., 2021).

As a result of supervised machine learning, ML algorithms become adept at analyzing data and performing tasks based on the knowledge and examples they have been trained on. Starting from specific problems, these algorithms can develop an inductive hypothesis, which serves as a resolution model for broader issues.

Deep learning (DL) is based on artificial neural networks, which consist of nodes (or neurons) organized in layers. The architecture includes:

- Input Layer: Receives raw input data, with each input feature corresponding to a node.

- Hidden Layers: Process the input data by applying weighted sums of inputs from the previous layer, generating output signals for the next layer.
- Output Layer: Produces the final output of the network, which could be a classification label, a regression value, a probability distribution, or another type of output depending on the network's design.

The network's performance is optimized through the training process, which involves adjusting weights and biases, including a bias term added to the weighted sums.

DL techniques have been employed to develop noninvasive diabetes risk forecasting models by analyzing morphological features, such as tongue or retinal fundus images, or by assessing specific patterns of body fat distribution using imaging techniques like abdominal computed tomography (CT) or magnetic resonance imaging (MRI) (Contreras & Vehi, 2018; Ellahham, 2020).

1.6. AI and Predictive models in diabetes medication adherence

AI generally refers to computer systems that replicate mental processes unique to humans, such as reasoning, finding meaning, generalizing, or learning from experience, to achieve goals without explicit programming for each specific task. AI is commonly categorized into three types:

1. Strong AI: Systems that aim to fully mimic human reasoning and behavior.

2. Weak AI: Systems that produce outcomes comparable to human performance but may use different methods.

3. "In-Between" AI: Systems that are inspired by or guided by human reasoning, which is where much significant work is currently focused.

In medicine, AI is divided into:

1. Virtual AI: Includes informatics and deep learning systems used for data analysis and decision support.

2. Physical AI: Encompasses robot-assisted systems that perform physical tasks or procedures.AI comprises the use of a computerised system (hardware or software) to mimic intelligent behaviour with little human participation. AI has a variety of potential applications in the healthcare sector, including supporting the early detection, diagnosis, management, and treatment of medical conditions, enhancing patient engagement and boosting medication adherence, assisting the elderly, promoting health, providing counselling, managing administrative tasks, and even assisting healthcare professionals in their education and learning (Alanazi et al., 2018; Ceriello et al., 2014).

Since the last ten years, predictive models have been employed to forecast the course of the disease and inform future choices. Any predictive model's primary goal is to accurately forecast the course of a disease. The term "accurate prediction" refers to a situation when new disease outcomes can be determined using historical data. The results of disease prediction aid medical personnel and clinicians in making decisions in the future. Prior to the development of these prediction models, medical professionals and doctors used their professional judgement and experiences to estimate or anticipate the course of a disease. The earlier predictive techniques have evolved into effective predictive models to choose the patient's disease therapy as a result of the development of new technology (Adua et al., 2021; Ozery-Flato et al., 2013).

CHAPTER 2: LITERATURE REVIEW

Diabetes has been recognized since ancient times, with references dating back to as early as 1500 BC, noting excessive urination in historical manuscripts. Type 2 diabetes, labeled a global epidemic by the World Health Organization, is also known as adult-onset or non-insulin-dependent diabetes. It arises from the body's inefficient use of insulin, leading to chronic metabolic disorder. Factors such as sedentary lifestyle and obesity contribute significantly to its prevalence, with genetic predisposition also playing a role. While Type 2 diabetes was traditionally associated with adults, recent years have seen an alarming rise in cases among children (Himanshu et al., 2020).

The worldwide incidence of Type 2 diabetes has surged over the decades, with approximately 392 million people diagnosed in 2015 compared to a mere 30 million in 1985. Unlike Type 1 diabetes, symptoms of Type 2 diabetes may be subtler, often leading to delayed diagnosis and subsequent complications. These complications encompass a range of serious health issues, including cardiovascular disease, strokes, diabetic retinopathy, kidney failure, and impaired circulation in the extremities, sometimes necessitating amputations (Ma & Tong, 2024).

Recognizing the gravity of this health crisis, lifestyle modifications and medication are crucial for managing diabetes and mitigating its adverse effects. Given its pervasive impact, diabetes earns its reputation as a global epidemic (Jude et al., 2022).

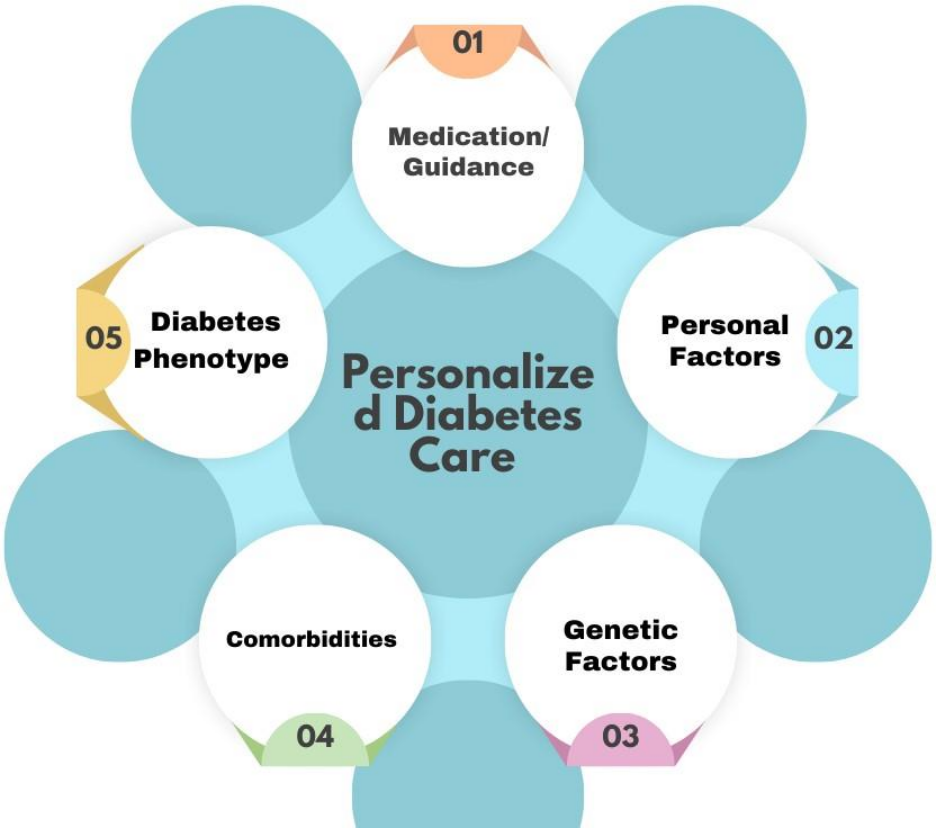


Figure 2.1. *personalized diabetes care. This figure summarizes the key considerations that are needed when contemplating the choice of diabetes pharmacotherapy for a patient with T2D.*

Type 2 diabetes mellitus (T2DM) arises from abnormalities in the metabolism of proteins, lipids, and carbohydrates, often due to either insulin resistance or insufficient insulin production. T2DM accounts for over 90% of diabetes cases, surpassing both type 1 diabetes mellitus (T1DM) and gestational diabetes. Recent advances have deepened our understanding of T2DM's onset and progression.

The primary cause of T2DM is a gradual decline in pancreatic β-cell function, typically occurring alongside existing insulin resistance in the liver, adipose tissue, and skeletal muscle. Before overt hyperglycemia develops, individuals may experience prediabetes, which significantly increases the risk of progressing to T2DM. Prediabetes is characterized by impaired glucose tolerance (IGT), elevated glycated hemoglobin A1c (HbA1c) levels, and impaired fasting glucose (IFG).

IFG is indicated by elevated fasting plasma glucose levels that do not meet diabetes criteria and reflects hepatic insulin resistance and diminished early-phase insulin secretion. In contrast, IGT involves muscle insulin resistance and reduced late-phase insulin secretion following meals. Individuals with prediabetes have HbA1c levels ranging from 5.7% to 6.4%, and the annual conversion rate from prediabetes to T2DM ranges from 3% to 11%.(Franks et al., 2013; Lin & Sun, 2010).

2.1 Hemoglobin A1c (HbA1c) as a Marker and Predictor in Type 2 Diabetes Mellitus

HbA1c, or glycosylated hemoglobin, is frequently used to indicate one's blood sugar level. When glycated hemoglobin (HbA1c) was initially identified in diabetic patients, it was referred to as an unusual hemoglobin (Sirsikar et al., 2016). The HbA1c value reflects the net mean blood glucose level over the preceding 1 or 2 months, and does not reflect immediate changes in the blood glucose profile after treatment. Hemoglobin A1c, or HbA1c, is a vital marker and predictor in the field of Type 2 Diabetes Mellitus (T2DM), providing important information about glycemic control and long-term disease treatment. Glycated hemoglobin, or HbA1c, is a dependable indicator of total glucose management since it shows the average blood glucose levels throughout the two to three months prior (Bennett et al., 2007; Vijayakumar et al., 2016).

The key to understanding the significance of HbA1c is that, in contrast to oral glucose tolerance testing or fasting plasma glucose testing, which only provide a moment in time view of glucose levels, it can measure glycemic control over a prolonged period. For this reason, HbA1c is the recommended method for determining glycemic status and directing therapy choices in T2DM. Hemoglobin A1c, or HbA1c, is a vital marker and predictor in the field of Type 2 Diabetes Mellitus (T2DM), providing important information about glycemic control and long-term disease treatment. Glycated hemoglobin, or HbA1c, is a dependable indicator of total glucose management since it shows the average blood glucose levels throughout the two to three months prior (Bennett et al., 2007; Florkowski, 2013).

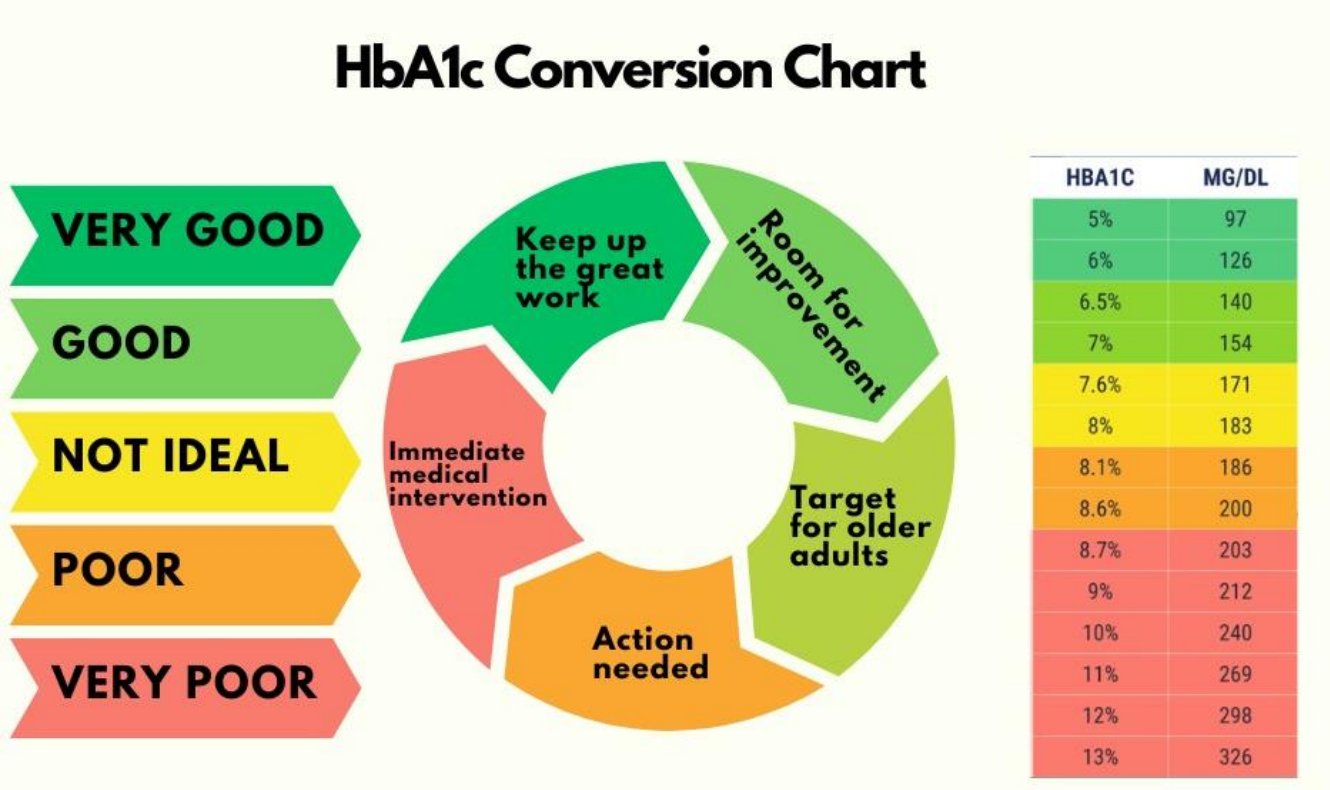


Figure. 2.2. *HbA1c to estimated average glucose (MG/DL). Categorizing the medication adherence and self-control on the basis of these values.*

2.2. Current Approaches to Medication Adherence in T2DM

Scholars document an extensive array of impediments to patient adherence, encompassing health system, provider, and patient issues. Psychological distress has long been known to have an impact on health results, and depression's effect on medication adherence may be one way that mood disorders influence health outcomes. Patients with depression have numerous risk factors that could lead to non-adherence, including low energy, lack of desire, social disengagement, hopelessness, and altered expectations about the advantages or disadvantages of therapy. Patients who are depressed may also find it harder to communicate with healthcare providers and feel less satisfied with their treatment. (Grenard et al., 2011)

Adhering to medication is crucial for effectively managing type 2 diabetes mellitus (T2DM), as it aids in controlling blood sugar levels and lowers the risk of complications.

Here are some contemporary strategies to enhance medication adherence in T2DM: (Brunton & Polonsky, 2017).

2.2.1. Empowering Patients through Education

Providing comprehensive education to patients about their condition, the importance of medication adherence, and the consequences of non-adherence can empower them to take control of their health. This includes understanding the purpose of each medication, potential side effects, and strategies for managing them (Markovič et al., 2023).

1. **Simplified Regimens:** Insulin injection therapy, requiring multiple daily injections, is commonly prescribed to patients with type 2 diabetes who exhibit severe hyperglycemia (MDI). When on the MDI (multiple daily injections) regimen for years, a significant portion of patients have overtreatment, which is defined as administering a treatment even when its prospective risks outweigh its benefits. There are no set guidelines, however the treatment might be shortened if the glucose toxicity goes away. Adherence can be increased by streamlining prescription regimens by lowering the quantity of tablets and frequency of doses. Combination medications, which combine several prescriptions into one tablet, can lessen the difficulty of taking pills and facilitate patients' adherence to their treatment schedule (Taybani et al., 2019).
2. **Personalized Treatment Plans:** To achieve target glycemic control, previous guidelines for treating individuals with type 2 diabetes mellitus (T2D) primarily depended on strict algorithms for the sequential addition of pharmacotherapies. Newer guidelines advocate for a more personalized approach to diabetes treatment to improve quality of life, medication adherence, patient satisfaction, and overall health outcomes. Clinicians should collaborate with patients to develop individualized treatment plans that address weight management, targeted glycemic control, comorbidity prevention and treatment, and avoiding complications such as hypoglycemia. The Italian Association of Medical Diabetologists has developed a novel tailored algorithm for type 2 diabetes management, available online. In Pakistan, creating specific care plans is not very common. (Williams et al., 2022).
3. **Use of Technology:** Medication adherence, according to the World Health Organization (WHO), is a complex phenomenon involving individuals, their healthcare professionals,

and the act of taking medication. Over the past ten years, information technology (IT) has become more and more important in the field of health care. It has been demonstrated that a number of cutting-edge IT tools, including electronic health records (EHRs), e-prescribing, and electronic drug monitoring (EDM) for clinicians to track patients' medication adherence, are useful in offering seamless and efficient solutions for enhancing diabetes patient outcomes and healthcare practices (Shrivastava et al., 2023). Apps are also made to help individuals take their medications more consistently. The apps employed e-dairy, learning instructions, reminders, communication with a healthcare practitioner, and medication adherence promotion strategies. Reminding patients to take their medications via a reminder or a motivational activity with patient education features was offered by four applications (Islam et al., 2022).

4. **Behavioral Interventions:** Cognitive-behavioral therapy, problem-solving skills training, and motivational interviewing are a few behavioral interventions that can assist address psychological barriers to drug adherence, including fear of side effects, pharmaceutical beliefs, and forgetfulness (Lambrinou et al., 2019; Lindenmeyer et al., 2006).
5. **Regular Follow-up and Monitoring:** Diabetes management requires routine blood glucose monitoring and follow-up care. To attain glycemic control and stop the course of the condition, doctors need to know the results of blood glucose monitoring to prescribe the right medication and offer personalized lifestyle recommendations (Li et al., 2022). DSME is an ongoing process that assists patients in acquiring the competencies, know-how, and skills necessary to effectively manage their diabetes on their own. It has been demonstrated to improve patient outcomes and is a crucial part of diabetic therapy. Physical activity, a healthy diet, medication adherence, monitoring, problem-solving, risk reduction, and healthy coping are the seven self-care behaviors that the American Association of Diabetes Educators (AADE) has identified as trustworthy outcome measures of diabetes self-management education (DSME). By employing a combination of these approaches, healthcare providers can help improve medication adherence and ultimately enhance the management of T2DM (Brunton & Polonsky, 2017; Lindenmeyer et al., 2006).

2.2.2. Role of Gene Expression Analysis in Predictive Models

Gene expression analysis plays a crucial role in diabetes prediction models by providing insights into the molecular mechanisms underlying the disease and identifying biomarkers associated with diabetes risk. Here are several keyways gene expression analysis contributes to diabetes prediction models (Franks et al., 2013).

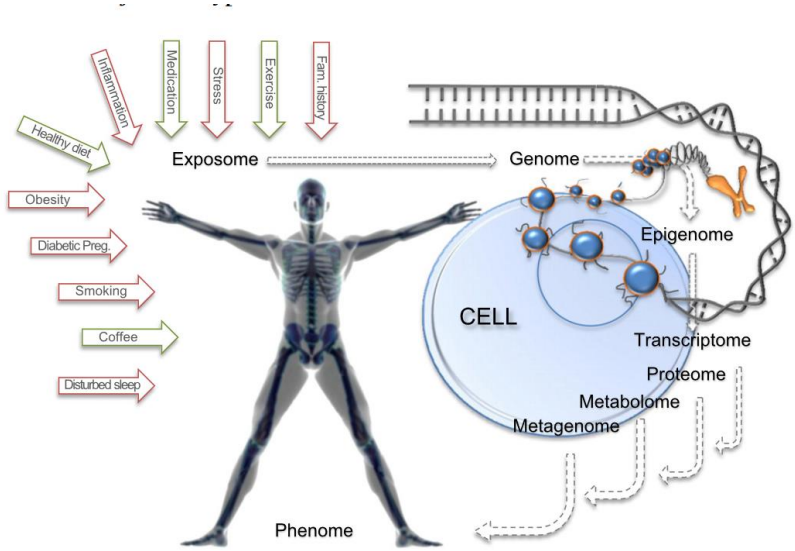


Figure 2.2. *A systems epidemiology approach investigates how interactions between the exposome (comprising all non-genetic exposures) and measurable components of the human physiome contribute to our understanding.*

Identification of Biomarkers: Using gene expression analysis, it is possible to pinpoint genes whose expression levels are linked to the onset of diabetes. Through gene expression profile comparisons between individuals with and without diabetes, researchers can identify differentially expressed genes that may be linked to an increased risk of developing the illness (Fan et al., 2021).

Subtyping Diabetes: Diabetes is a diverse illness with several subtypes that are identified by unique molecular markers. Based on their molecular profiles, individuals can be classified into several subtypes with the aid of gene expression analysis, allowing for the development of more specialized prediction models aimed at particular subgroups (Keller et al., 2008).

Predictive Modeling: Clinical and demographic data can be combined with gene expression data to create more precise predictive models for diabetes risk assessment. Gene expression patterns and other pertinent characteristics can be analyzed by machine learning algorithms to predict an individual's risk of getting diabetes within a specific time frame (Adua et al., 2021; Ozery-Flato et al., 2013).

Understanding Pathophysiology: Understanding the underlying molecular mechanisms involved in the etiology of diabetes is possible through the investigation of gene expression. Researchers can gain a better understanding of the mechanisms behind the course of disease and discover new treatment targets by clarifying how gene expression patterns alter in response to metabolic imbalance (Adua et al., 2021; Szabo et al., 2018).

Early Detection and Prevention: Even before clinical symptoms appear, changes in gene expression may take place in the early stages of diabetes development. Thus, gene expression analysis can help identify people who are at a high risk of getting diabetes early on, enabling prompt intervention and the development of preventative measures to slow the progression of the illness (Samsom et al., 2016).

Monitoring Treatment Response: Gene expression profiling can be used to track an individual's response to medication or lifestyle modifications employed in the treatment of diabetes. Clinicians can evaluate the effectiveness of treatment and customize therapeutic approaches for specific patients by monitoring changes in gene expression after the start of treatment (Thipsawat, 2021).

Overall, molecular insights into disease causes, individualized risk assessment, and therapeutic intervention guidance for better patient outcomes are some of the ways that gene expression analysis improves the accuracy and usefulness of diabetes prediction models.

2.2.3. Existing Predictive Models for T2DM Medication Adherence

Several predictive models exist for Type 2 Diabetes Mellitus (T2DM) medication adherence, aiming to enhance patient outcomes through improved adherence rates. Here are some common approaches:

1. **Machine Learning Models**: Machine learning techniques, such as decision trees, random forests, support vector machines, and neural networks, have been employed to

predict medication adherence. These models utilize patient demographic data, clinical variables (e.g., HbA1c levels, comorbidities), medication history, and psychosocial factors (e.g., depression, health literacy) to forecast adherence behavior (Fregoso-Aparicio et al., 2021; Xiong et al., 2019).

2. **Health Behavior Models**: These models integrate health behavior theories, such as the Health Belief Model or Social Cognitive Theory, to understand and predict medication adherence. They assess factors like perceived susceptibility to complications, perceived benefits of adherence, self-efficacy, and social support to predict patient adherence behavior (Karimy et al., 2016; Velu et al., 2023).
3. **Electronic Health Records (EHR) Based Models**: Leveraging data from electronic health records, these models identify patterns and predictors of medication adherence. They utilize patient demographics, clinical variables, medication history, and healthcare utilization patterns to predict adherence behavior (Nguyen et al., 2019; Zheng et al., 2017).
4. **Mobile Health (mHealth) Apps**: Mobile apps and wearable devices are increasingly used to monitor medication adherence in real-time(Istepanian et al., 2017). These apps collect data on medication intake, physical activity, glucose levels, and other relevant parameters to predict adherence behavior. Machine learning algorithms may be employed to analyze this data and provide personalized adherence predictions (Årsand et al., 2012; Muralidharan et al., 2017).
5. **Natural Language Processing (NLP) Models**: NLP techniques analyze unstructured data, such as patient-provider communication or social media posts, to identify predictors of medication adherence. By extracting relevant information from textual data, NLP models can predict patient adherence behavior based on linguistic cues and contextual factors (Misra-Hebert et al., 2020; Vidyadharan et al., 2021).
6. **Patient Engagement Models**: These models focus on patient engagement strategies to promote medication adherence. By analyzing patient engagement metrics, such as app usage patterns, communication frequency with healthcare providers, and participation in educational programs, these models predict adherence behavior and recommend tailored interventions to improve adherence rates (Fioravanti et al., 2015).

7. **Integrated Care Models**: Integrated care models combine clinical data with patient-reported outcomes and behavioral assessments to predict medication adherence. These models often involve interdisciplinary care teams, including physicians, nurses, pharmacists, and behavioral health specialists, to address the multifaceted determinants of adherence behavior (Silva et al., 2020).

Each of these predictive models offers unique advantages and may be tailored to specific patient populations or healthcare settings to optimize medication adherence and improve T2DM management outcomes.

2.3. Dysregulation of GLUT4 Expression in Type 2 Diabetes Mellitus (T2DM)

Insulin-responsive facilitative glucose transporter GLUT4 is mostly expressed in skeletal muscle, cardiac muscle, and adipose tissue. In muscle and fat cells, GLUT4 is sequestered in intracellular vesicles when insulin levels are low. Insulin causes GLUT4 to translocate from these vesicles to the plasma membrane, which causes a sharp rise in glucose uptake (Alam et al., 2016). GLUT4 transporters are introduced and made available for transferring glucose as the vesicles fuse with the plasma membrane, increasing the absorption of glucose. Whole-body insulin-mediated glucose homeostasis and GLUT4 expression levels are connected, and studies have linked GLUT4 expression dysregulation to diabetes and obesity. Mice carrying the GLUT4 heterozygous mutant allele (GLUT4±) exhibit elevated blood glucose and insulin levels, decreased muscle glucose absorption, hypertension, and diabetic liver and heart histopathologies. The role of GLUT4 in the development of diabetes was demonstrated by the ability of overexpression of GLUT4 to partially improve in vivo glucose tolerance in diabetic mice (Galicia-Garcia et al., 2020; Kaku, 2010).

2.4. Metformin's Mechanisms: Enhancing Glucose Metabolism Through GLUT4 Translocation

GLUT4 is a glucose transporter primarily found in adipose tissue and skeletal muscle cells. Its translocation to the cell membrane allows for the uptake of glucose from the bloodstream into these cells, thereby lowering blood glucose levels. Here's how metformin induces GLUT4 translocation and affects metabolism:

1. **AMP-activated Protein Kinase (AMPK) Activation**: Metformin activates AMPK, a key cellular energy sensor. AMPK activation leads to various metabolic effects, including increased glucose uptake and utilization in skeletal muscle cells. AMPK activation also stimulates GLUT4 translocation to the cell membrane, enhancing glucose uptake (Jackson et al., 1987).
2. **Insulin Sensitization**: Metformin improves insulin sensitivity in peripheral tissues, such as skeletal muscle and adipose tissue. By enhancing insulin sensitivity, metformin promotes GLUT4 translocation in response to insulin, allowing for efficient glucose uptake into cells.
3. **Reduction of Hepatic Glucose Production**: Metformin inhibits hepatic gluconeogenesis, the process by which the liver produces glucose. By reducing hepatic glucose production, metformin decreases blood glucose levels, which in turn may lead to decreased insulin secretion and improved insulin sensitivity, facilitating GLUT4 translocation (Hundal et al., 2000).
4. **Mitochondrial Effects**: Metformin alters mitochondrial function, leading to changes in cellular energy metabolism. These changes can indirectly influence GLUT4 translocation and glucose uptake by modulating cellular energy status and signaling pathways involved in glucose metabolism.
5. **Adiponectin Secretion**: Metformin has been shown to increase adiponectin secretion from adipose tissue. Adiponectin is an adipokine that plays a role in regulating glucose and lipid metabolism. Increased adiponectin levels may enhance insulin sensitivity and promote GLUT4 translocation in target tissues (Foretz et al., 2019; Jackson et al., 1987). Overall, metformin induces GLUT4 translocation and improves glucose metabolism through multiple mechanisms, including AMPK activation, insulin sensitization, reduction of hepatic glucose production, mitochondrial effects, and modulation of adipokine secretion. These effects contribute to the therapeutic efficacy of metformin in the management of T2DM.

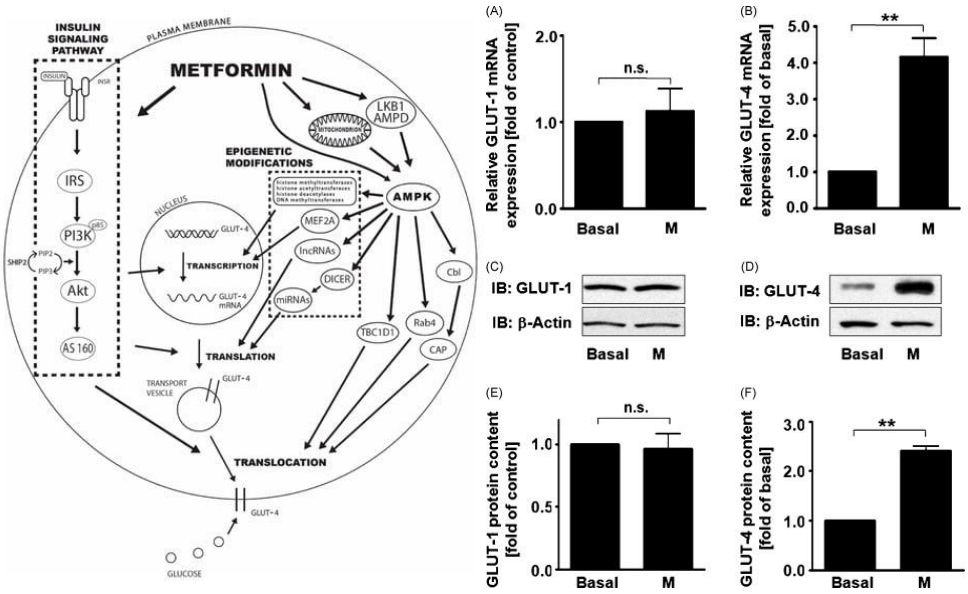


Figure 2.3. *Metformin treatment led to inhibited glucose uptake, which was demonstrated through the use of phloretin, an inhibitor of glucose transporters (data not shown). There was no observed change in GLUT-1 mRNA expression due to metformin treatment (Fig. 3A). However, GLUT-4 mRNA expression increased significantly by 4.2 ± 0.5-fold ($p < 0.01$ compared to basal levels) (Fig. 3B). Similarly, at the protein level, metformin did not alter the amount of GLUT-1 (Fig. 3C and E), whereas GLUT-4 protein content showed a notable increase of up to 2.6-fold ($p < 0.01$ compared to basal levels) (Fig. 3D and F).*

CHAPTER 3: METHODOLY

3.1. Sampling and Data Collection

Study protocols were scrutinized by Institutional Review Board ASAB and Ethical Committee of Federal Government polyclinic hospital, Islamabad. A total of 513 patients were interviewed to fill the questionnaires designed to access the medication adherence of the patients. Also 65 blood samples were collected from the patients who met the inclusion criteria of the study.

3.1.1. Data analysis

Patients from all districts of Pakistan and AJK were included in the study. Weight in kilograms and height in feet were recorded to calculate BMI. Patients with previous gestational diabetes, type 1 diabetes (distinguished from T2DM by age at onset and ketoacidosis), and other forms of diabetes were excluded from the final analysis. Each patient underwent a thorough interview to collect information on their medical history, demographics, and clinical details. This included age, sex, disease duration, family medical history, treatment type, weight, height, body mass index (BMI), blood pressure, overall health status, any existing micro- and macro-vascular complications, hypertension (defined as patients already on hypertension medication or with blood pressure >140/80 mm Hg), medication compliance, and self-management practices. Statistical analysis was performed using Microsoft Excel data analysis tools.

3.2 Expression Analysis

3.2.1 RNA Extraction

For RNA extraction, 200 µL of blood sample was added to a 1.5 mL Eppendorf tube, followed by the addition of 750 µL of Trizol reagent. The solution was vortexed for a few seconds to achieve a homogenized mixture. The tube was then centrifuged at

12,000 rpm for 10 minutes at 4°C. The middle pink layer containing Trizol was transferred to a new 1.5 mL microcentrifuge tube. This mixture was incubated for 5 minutes at room temperature, and 20 µL of 5N glacial acetic acid was added. After mixing vigorously for 15 seconds, the mixture was incubated for another 5 minutes at room temperature. Next, 200 µL of chloroform was added, mixed vigorously for 15 seconds, and incubated for 10 minutes at room temperature. The mixture was centrifuged at 12,000 rpm for 15 minutes at 4°C. The supernatant containing RNA was transferred to a new microcentrifuge tube, and 500 µL of isopropanol was added and vortexed. The mixture was centrifuged at 12,000 rpm for 10 minutes at 4°C. The supernatant was removed, and the pellet was washed by adding 1 mL of 75% ethanol followed by centrifugation at 7500 rcf for 5 minutes at room temperature. After centrifugation, the ethanol was removed, and the pellet was air-dried. Finally, the pellet was dissolved in 20 µL of Diethylpyrocarbonate (DEPC) treated water. The RNA was quantified using a Thermo Scientific NanoDrop 2000.(Heidary & Kakhki, 2014).

3.2.2 Complementary DNA (cDNA) synthesis

The extracted RNA was converted into cDNA using RevertAid Reverse Transcriptase (Thermo Scientific). The reaction mixture contained the following components:

- 1 µL of template RNA
- 1 µL Oligo dT primers
- 1 µL RiboLock RNase inhibitor
- 4 µL of 5X reaction buffer
- 2 µL of 10 mM dNTPs mixture
- 1 µL reverse transcriptase

The total volume was brought up to 20 µL with nuclease-free water. The mixture tube was given a short spin and then incubated at 37°C for 5 minutes. The reaction conditions in the thermocycler were set to 60 minutes at 37°C followed by 10 minutes at 70°C. After the reaction was completed, the cDNA was stored at -20°C. (Mannhalter et al., 2000).

3.2.3 Semi quantitative PCR

Semi-quantitative PCR was performed to confirm the synthesis of cDNA, using primers for β-actin. The reaction mixture included the following components:

- 0.5 µL of cDNA template
- 2.5 µL of 10X Taq buffer
- 2.0 µL of 25 mM MgCl2
- 0.5 µL of 2 mM dNTP mix
- 1 µL of each primer (forward and reverse, 1 µM)
- 0.4 µL of Taq DNA polymerase (Thermo Scientific)
- Volume raised to 20 µL with PCR-grade water

  The reaction profile in the thermocycler was set as follows:

  1. Initial denaturation at 95°C for 3 minutes
  2. 35 cycles of:
     - Denaturation at 95°C for 20 seconds
     - Annealing at 60°C for 35 seconds
     - Extension at 72°C for 35 seconds
  3. Final extension at 72°C for 10 minutes(Mannhalter et al., 2000).

3.2.4 Gel Electrophoresis

Semi quantitative PCR product of β actin specific region was confirmed by gel electrophoresis. For this, 2% agarose gel was prepared by dissolving 1g of agarose in 50 mL 1X TAE. Ethidium bromide 3 µL was added into it for staining of DNA. Three microliters of PCR product was mixed with 1 µL of loading dye and loaded into the well. The amplicon size was compared with 100 bp ladder, which was run along with sample, on Dolphin doc gel documentation system (Mannhalter et al., 2000).

3.2.5 Primers

Primers for ß-Actin has already been reported in literature. Primers for GLUT 4 has been designed using NCBI Primer Blast software (https://www.ncbi.nlm.nih.gov/tools/primerblast/index.cgi?INPUT_SEQUENCE=%20EU563945.2LINK_LOC=nuccore). The mRNA sequences were taken from the NCBI. After checking the self-complementarity and other properties. Using USCS in-silico PCR

the primers were confirmed for non-specific binding (https://genome.ucsc.edu/cgi-bin/hgPcr).

3.2.6 Real Time PCR

For gene expression analysis, real time PCR was performed on 7300 Real Time PCR System (Applied Biosystems) with fluorescence based SYBR Green assay. To normalize the data, β actin was used as reference gene. The reaction mixture included 40 ng/μL (final concentration) template, 0.25 μL (10 μM) of each forward and reverse primer, 6.25 μL of Maxima SYBR Green qPCR master mix (Thermo Scientific), and total volume was raised up to 12.5 μL by adding nuclease free water. Thermal profile was set as 95°C for 10 min, followed by 40 cycles of 95°C for 20 seconds, annealing at 55°C (for GLUT 4) and 60°C (for β actin) for 35 seconds and extension at 72°C for 35 seconds. Dissociation stage profile was added for melt curve analysis as 95°C for 15 seconds, 60°C for 30 seconds and 95°C for 15 seconds. Each sample reaction was done in duplicates.

3.2.7. Statistical analysis and Results validation

ΔCt was obtained using excel and after analysis of app results graph was plotted.

3.3. Web based application Development

3.3.1. Import essential dependencies for data analysis/preprocessing and machine learning

For python different libraries were imported for data analysis, preprocessing and machine learning using different commands.

3.3.2. Read the .csv file (read the data set)

To read a CSV file into a Pandas DataFrame, we used the `read_csv()` function.

3.3.3. Printing the head of the dataset

Using .head() method to get the idea of how our dataset look like.

3.3.4. Dimensions of dataset (Rows, Columns)

By using the `.shape` attribute of the DataFrame the dimensions of the dataset (i.e., the number of rows and columns) were retrieved.

3.3.5. Statistics of dataset of each column

By using the `.describe()` method generated descriptive statistics of the numerical columns in our DataFrame. This displayed various statistics such as count, mean, standard deviation, minimum, maximum, and quartiles for each numerical column in dataset. The dataset containing non-numeric columns, will be excluded from the output of `.describe()`.

3.3.6. Dropping rows having null values from your dataset

First and foremost part of the data-analysis is Dropping rows having null values from your dataset.

3.3.7. Encoding

Replaced the string text with a number, e.g Male will be replaced with 1 and female will be replaced with 0, to give it float/integer.

3.3.8. Converting all dataset into float type

By using .astype() method (because if we don't it'll rise an error that datatype doesn't match with output feature (which is in float)).
Using `.astype()` method is a common approach to convert data types in a Pandas DataFrame. If you have a DataFrame named `df` and you want to convert all columns to float type.

3.3.9. Separating the output feature and input features

To separate the output feature (often referred to as the target variable or dependent variable) from the input features (independent variables), you can use pandas indexing. Assuming your DataFrame has columns representing both input features and the output feature.

3.3.10. Machine learning model

Divided the whole dataset into 2 parts (Train and Test) (Train part should be more like 70-80-90 percent). We defined variables for four part (Train >> Input Feature , Train >> Output Feature ; Test >> Input Feature , Test >> Output Feature). Once defined these, Initiatialize ML model (Our ML model is DecisionTreeRegressor, it gives good score in our case (continous input, output feature). Using .fit() to fit (parsing the input output TRAIN only features).

3.3.11. Checking the credibility

We're using certain metrices and parameters i.e Mean Squared Error, Mean Absolute Error, R2 Score etc (Remember these metrices are for CONTINOUS TYPE OF DATA only, If our data would've been Discrete (integer) we would use `Accuracy` metrix).

3.3.12. Prediction system

Created a prediction system for Command Line Interface (non-GUI based). Created a .py file which transforms the whole notebook (ML Model) into a file with extension `.joblib`. because we want to freeze/saved the ML Model which can be used for further uses and transform it to an executable, so we won't have to run notebook everytime unless we want to change something in our model).

3.3.13. Created a streamlit app

Created a streamlit web app and added input fields and prediction button. Also created a session to save the progress of all pervious predictions.

3.4. Accuracy of application

To evaluate the accuracy of the application, a dataset comprising 515 patients was utilized. The predicted values generated by the application were compared with the actual HbA1c values of these patients. The standard deviation was calculated to assess the

dispersion of the prediction errors, and a graph was plotted to visualize the comparison between the predicted and actual HbA1c values.

CHAPTER 4: RESULTS

4.1. Demographic and clinical characteristics of study subject

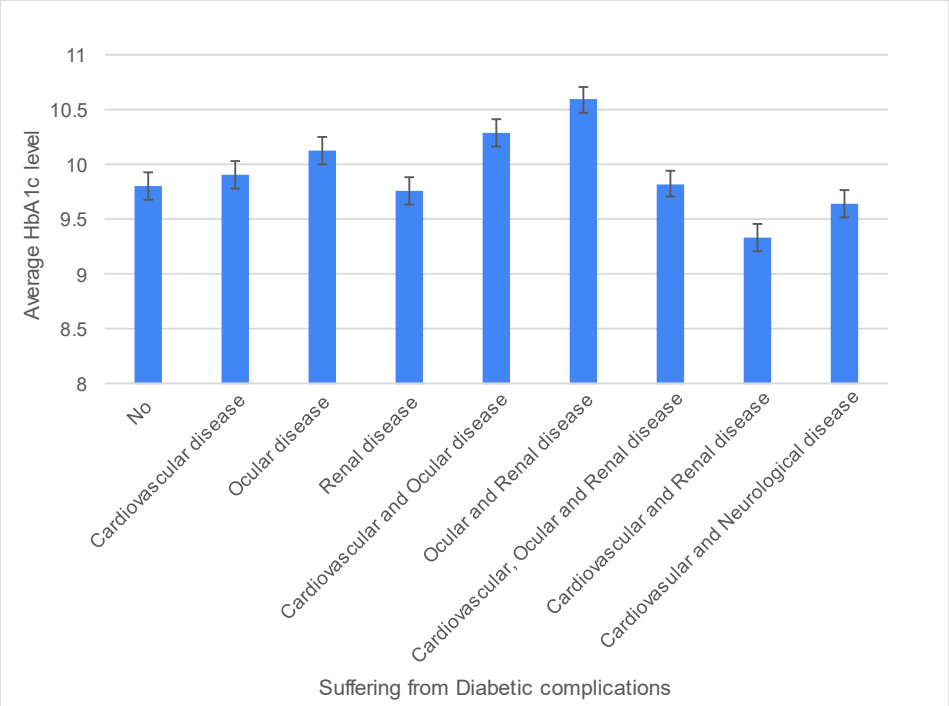


Figure 4.2. Comparison of HbA1c and frequency of complications.

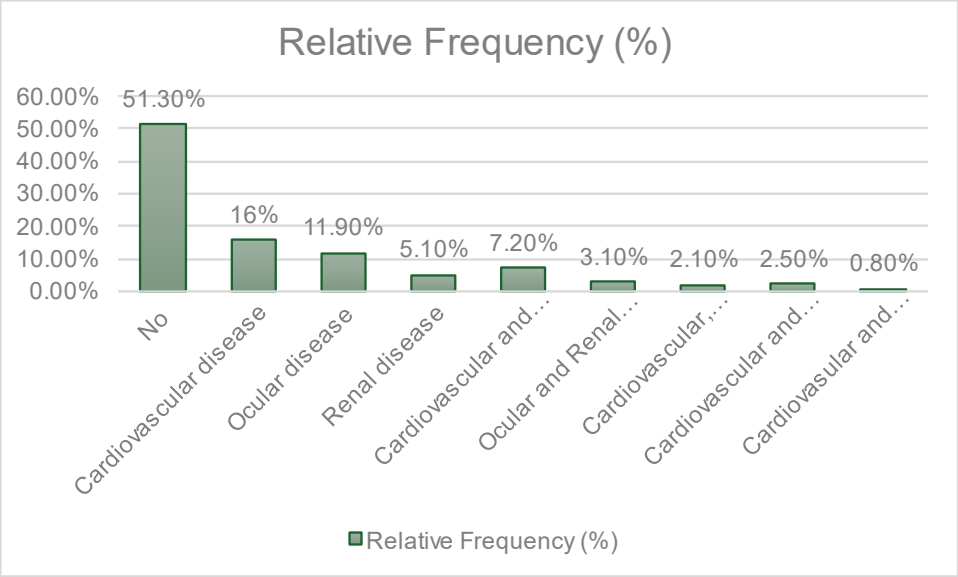


Figure 4.3. Relative frequency of complications in diabetes patients

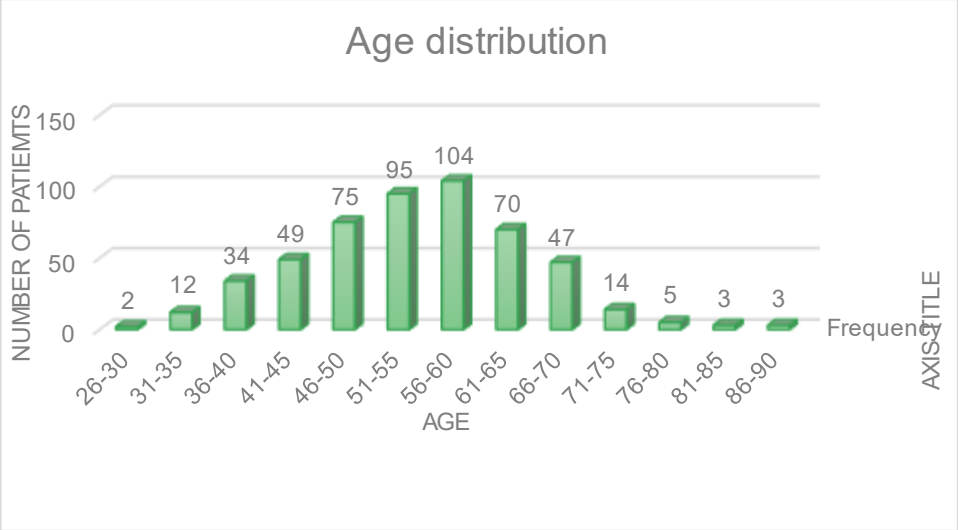


Figure 4.4. Age distribution of patients with T2DM

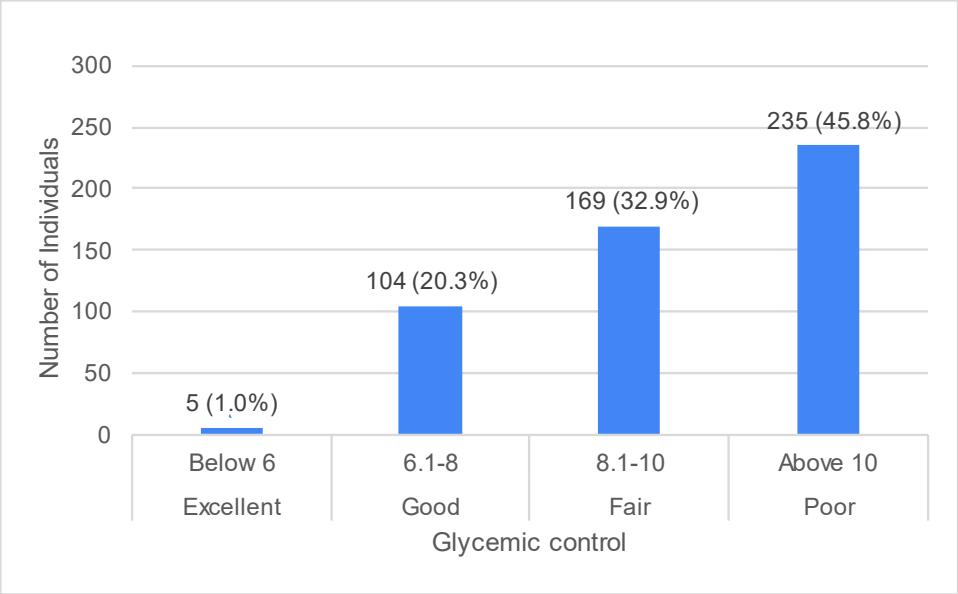


Figure 5.4. Comparison of glycemic control in patients diagnosed with T2DM

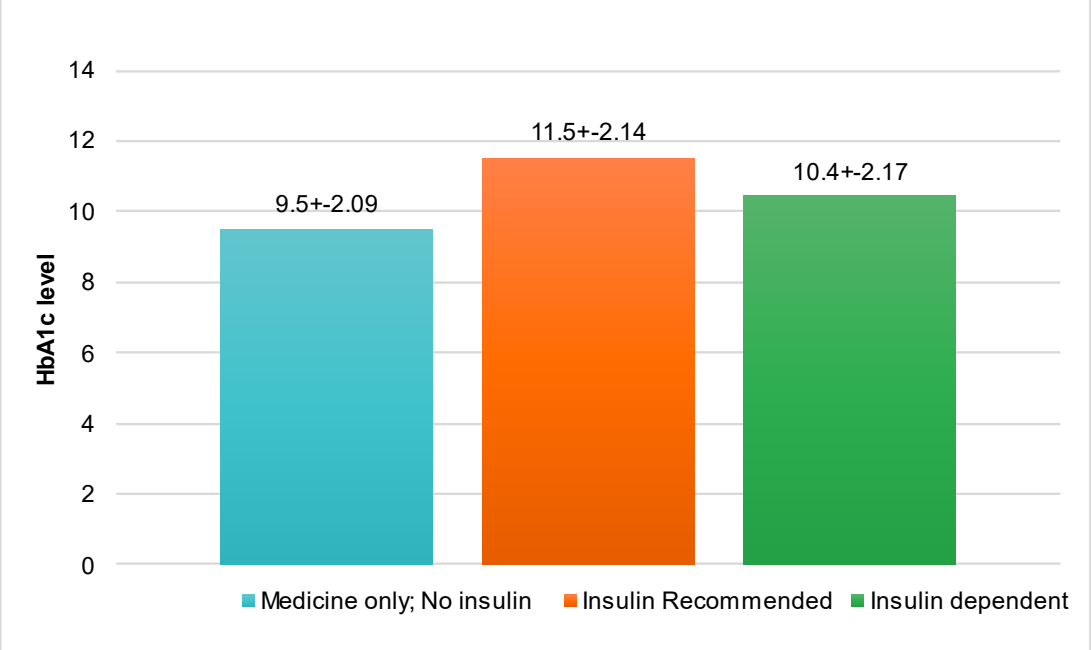


Figure 4.6. Comparison between insulin dependency and HbA1c levels

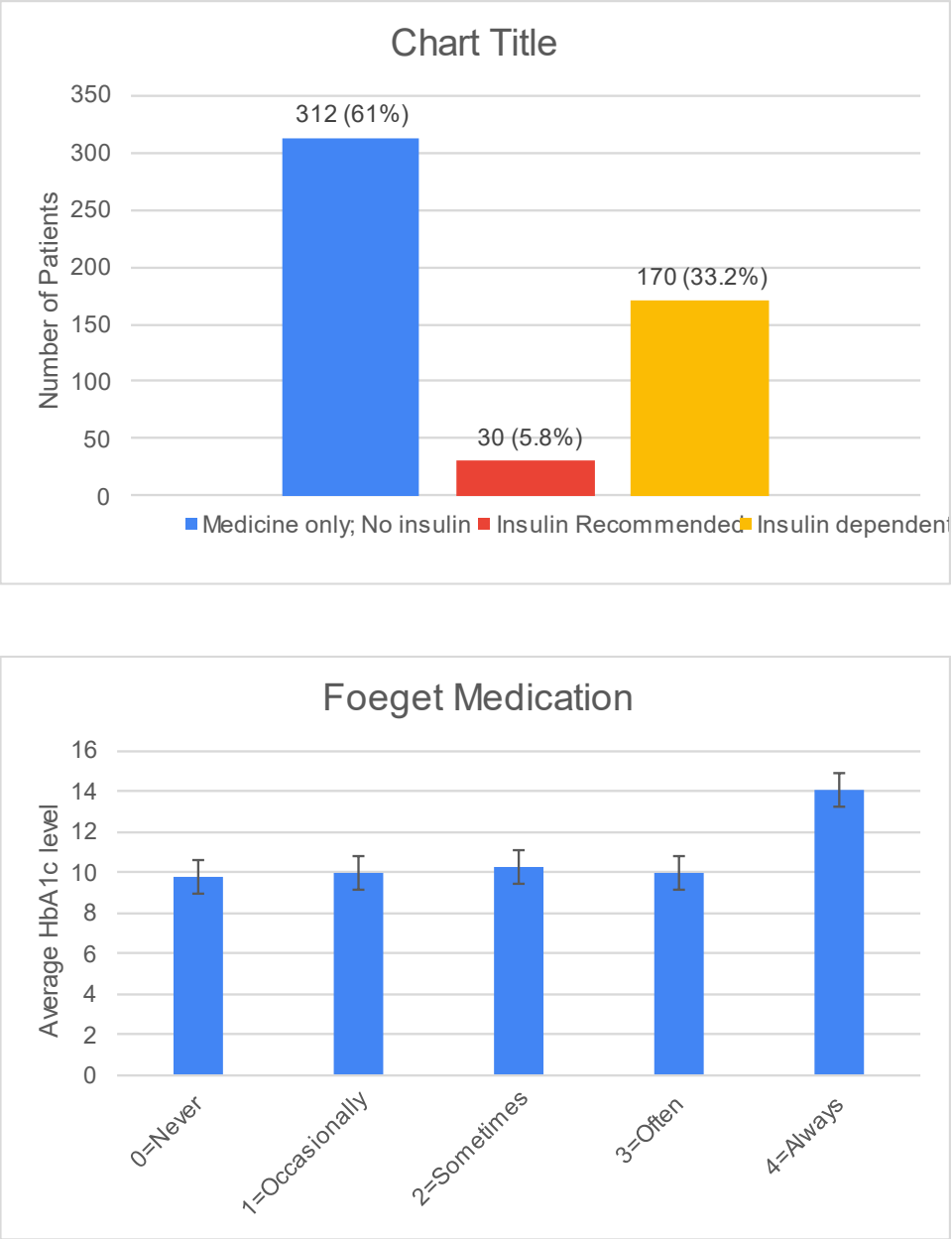


Figure 4.7. Comparison of compliance to the medication and HbA1c levels in patients

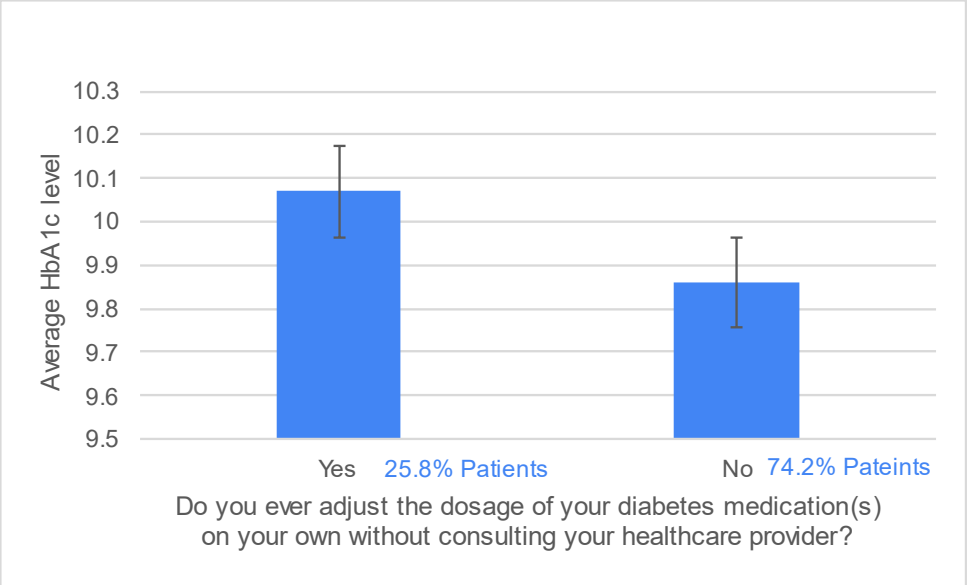


Figure 4.8. Relation of dosage adjustment and HbA1c levels

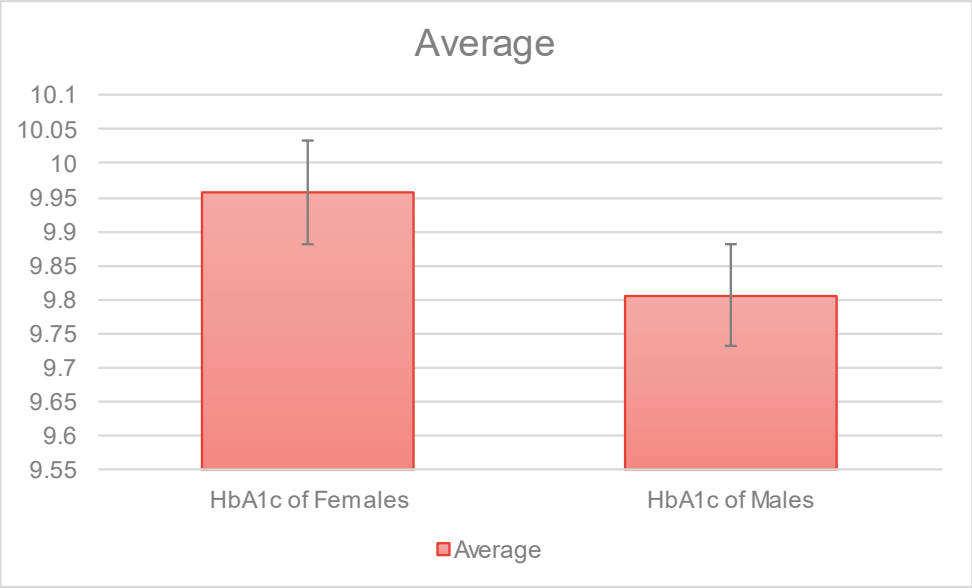


Figure 4.9. Comparison of HbA1c levels in male and female

Figure 4.10. Comparison between self-management and HbA1c levels.

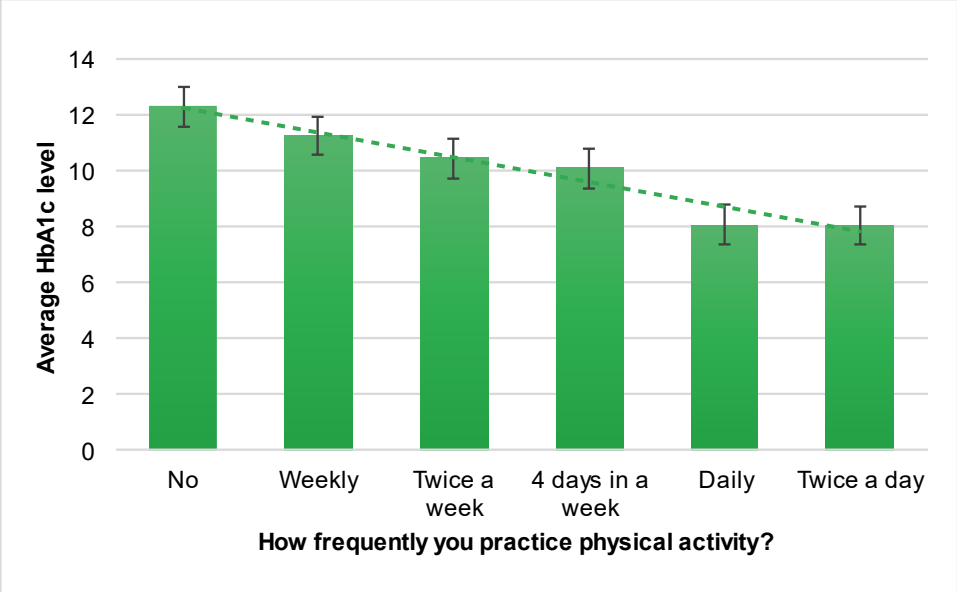


Figure 4.11. comparison of physical activity and HbA1c levels.

4.2. Statistics of the selected categorical predictors

The selected data had 9 parameters. Total of 513 patients participated from Pakistan in our study and total of 0.1 million records were selected from the Kaggle database. The 513 records from Pakistan were used to study different correlations and trends in the Pakistani population of different factors. The following 9 factors were selected based on those trends.

	gender	age	hypertension	heart_disease	smoking_history	bmi	HbA1c_level	blood_glucose_level	diabetes
0	Female	80.0	0.0	1.0	never	25.19	6.6	140.0	0
1	Female	54.0	0.0	0.0	No Info	27.32	6.6	80.0	0
2	Male	28.0	0.0	0.0	never	27.32	5.7	158.0	0
3	Female	36.0	0.0	0.0	current	23.45	5.0	155.0	0
4	Male	76.0	1.0	1.0	current	20.14	4.8	155.0	0

Dimensions of Dataset

Figure 4.11. *Dimensions of dataset (gender (Male, female, other), age, hypertension(0,1), heart disease (1,0), smoking history (never, current, former, not current), body mass index(BMI), diabetes status)*

	age	hypertension	heart_disease	bmi	HbA1c_level	blood_glucose_level	diabetes
count	100513.000000	100510.000000	100512.000000	100513.000000	100513.000000	100512.000000	100513.00000
mean	41.953753	0.077296	0.040682	27.338273	5.549894	138.844188	0.08967
std	22.491265	0.267062	0.197553	6.641393	1.123745	42.589711	0.28571
min	0.080000	0.000000	0.000000	10.010000	3.500000	80.000000	0.00000
25%	24.000000	0.000000	0.000000	23.650000	4.800000	100.000000	0.00000
50%	43.000000	0.000000	0.000000	27.320000	5.800000	140.000000	0.00000
75%	60.000000	0.000000	0.000000	29.630000	6.200000	159.000000	0.00000
max	90.000000	1.000000	1.000000	95.690000	18.300000	642.000000	1.00000

Counts of instances of classes in various attributes

Figure 4.12. *Descriptive statistics of the dataset, count of the instances, mean, STD, min, max 25%, 50%, 75%, max values of all the factors*

4.3. Prediction performance of the ML model

The prediction performance of the ML models were assessed on the major three parameters mean squared error mean absolute error and R^2 values. Decision tree regressor gave the best results for our provided data. Mean squared, mean absolute errors should be closer to zero as close as possible (ours are 0.143, 0.15 ish). R2 score should be closer 1.00 as close as possible (ours is 0.88).

```
# Accuracy on Training Data

input_features_train_prediction = decision_tree.predict(input_features_train)

meanSquaredError = mean_squared_error(output_feature_train, input_features_train_prediction)
meanAbsoluteError = mean_absolute_error(output_feature_train, input_features_train_prediction)
r2 = r2_score(output_feature_train, input_features_train_prediction)
```

```
print("Mean Squared Error", meanSquaredError)
```

```
Mean Squared Error 0.14304941779839453
```

```
print("Mean Absolute Error", meanSquaredError)
```

```
Mean Absolute Error 0.14304941779839453
```

```
print("R2 Score", r2)
```

```
R2 Score 0.8858752871660557
```

ML model	Mean squared error	Mean absolute error	R^2
Decision tree regressor	0.143	0.143	0.88

Table 3 Performance of the decision tree regressor.

4.4. Development of application

An application was developed using the decision tree regressor model, which serves as an effective tool for predicting HbA1c levels. Based on these predictions, results are categorized into four levels of medication adherence ranging from poor to excellent.

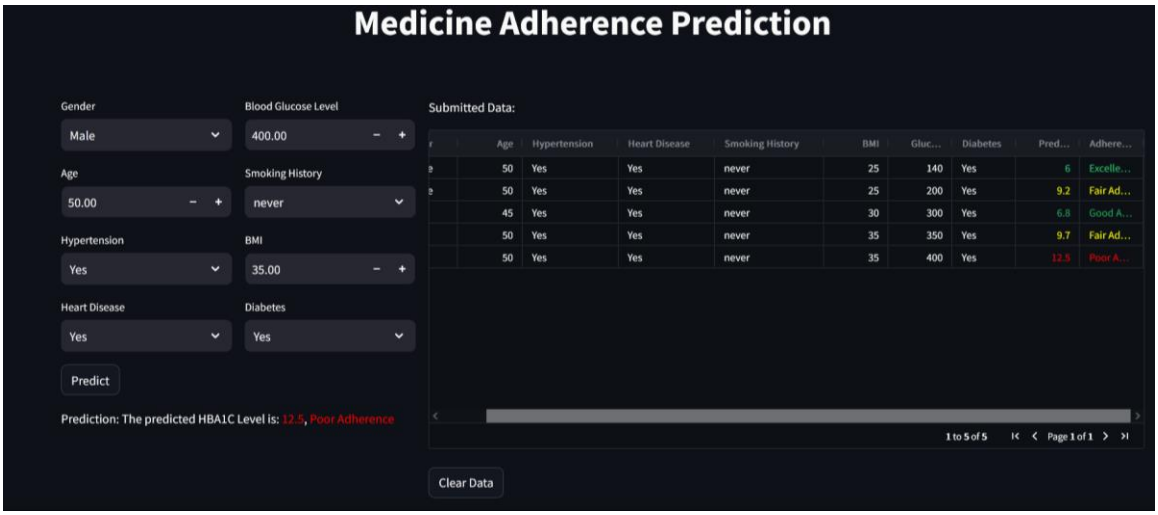


Figure 4.13. *Image of the website to predict HbA1c levels and make medication adherence prediction by integrating decision tree model using Gender, BGLs, Age, smoking history, hypertension, BMI, and heart disease status.*

4.5. Primer designing

The primers were designed using primer3 plus and by using primer BLAST and UCSC In-silico PCR sequence homology of both primers were confirmed.

4.5.1. Primers

Products on target templates

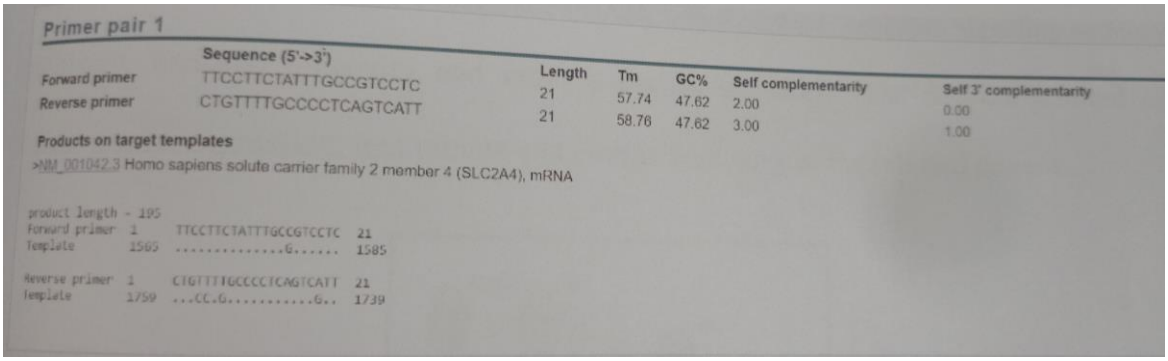
>XM_054342575.1 PREDICTED: Homo sapiens pleckstrin (PLEK), transcript variant X2, mRNA

```
product length = 113
Forward primer  1    AAGAAGGGGAGCGTGTTCA  19
Template        620  ..................  638

Reverse primer  1    AGCGGGATCATTCCTTTGG  19
Template        732  ..................  714
```

UCSC In-Silico PCR

>chr2:68390892+68391004 113bp AAGAAGGGGAGCGTGTTCA AGCGGGATCATTCCTTTGG
ttacAGGGGAGCGTGTTCAatacgtggaaacccatgtgggttgtattgtt
agaagatggaattgaattctataagaagaaaagtgacaacagccCCAAAG
GAATGATCCCGCT



Oligo name	Sequence (5’ to 3’)	Product size	Length	Tm (°C)	GC content
H SLC2A4 R HSLC2A4 F	CTGTTTTGCCCCTCAGTCATT TTCCTTCTATTTGCCGTCCTC	195	21	60.6	47.62 47.62

4.5.2. Primer optimization

The primers were optimized using two main approaches: varying the concentration of cDNA and adjusting the annealing temperature. For GLUT4, the optimal annealing temperature was determined to be 60.6°C. Additionally, a 1:100 concentration of cDNA was found to be necessary for RT-PCR.

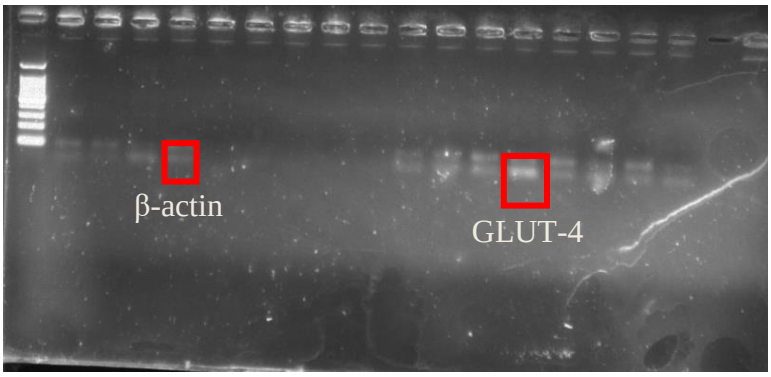


Figure 12. Figure of gel containg ladder of 1kbp in 1st well and beta actin in 2nd to 4th well while Glut-4 in 12th to 14th wells.

4.6. Expression analysis by Real time PCR

For expression analysis, five samples were run in triplicate using real-time PCR. The results obtained are as follows:

Patients 1 and 5: High ΔCt values (6.4), high HbA1c (10.7, 10.6) suggest poor GLUT4 expression and poor glycemic control showing poor medication adherence.

Patient 3: Moderate ΔCt (5.5), moderate HbA1c (9.5) indicates moderate GLUT4 expression and glycemic control showing Fair medication adherence.

Patient 4: Low ΔCt (-2.0), low HbA1c (6.2) suggests high GLUT4 expression and good glycemic control as well as good medication adherence.

Patient 2: Low ΔCt (0.8), relatively low HbA1c (7.7) indicates high GLUT4 expression and good glycemic control and medication adherence.

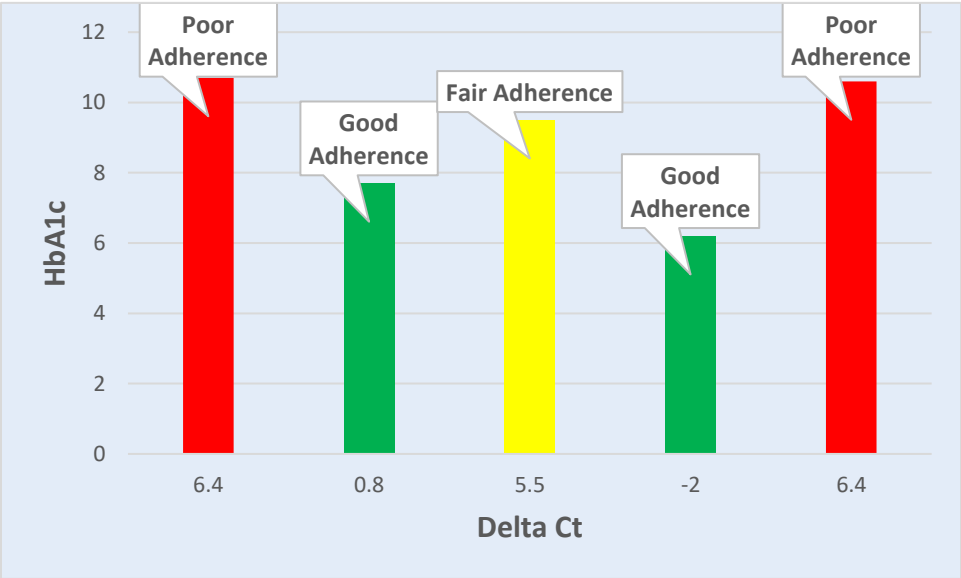
	beta actin (Ct)	Glut-4 (Ct)	Delta Ct	HbA1c
1	**23.1**	**29.5**	**6.4**	**10.7**
2	**22.7**	**23.5**	**0.8**	**7.7**
3	**22.5**	**28**	**5.5**	**9.5**
4	**23**	**21**	**-2**	**6.2**
5	**23.1**	**29.5**	**6.4**	**10.6**

These results indicate that GLUT4 expression may serve as an additional marker for assessing patient adherence to prescribed medication regimens. The observed relationship between GLUT4 expression levels and HbA1c values suggests that GLUT4 expression could provide insights into whether patients are following their treatment plans effectively. Lower GLUT4 expression is associated with poorer glycemic control, which may signal issues with medication adherence, while higher GLUT4 expression correlates with better glycemic control and potentially better adherence. Therefore, GLUT4

expression could be a useful biomarker for evaluating and improving patient adherence to diabetes management strategies.

4.7. Incorporation in application

HbA1c	delta CT	Medication Adherence	Predicted by App
10.7	6.4	Poor	10, Poor Adherence
7.7	0.8	Poor	7.7, Good Adherence
9.5	5.5	Fair	9.5, Fair Adherence
6.2	-2	Good	5.9, Good Adherence
10.6	6.4	Good	10.6, Poor Adherence

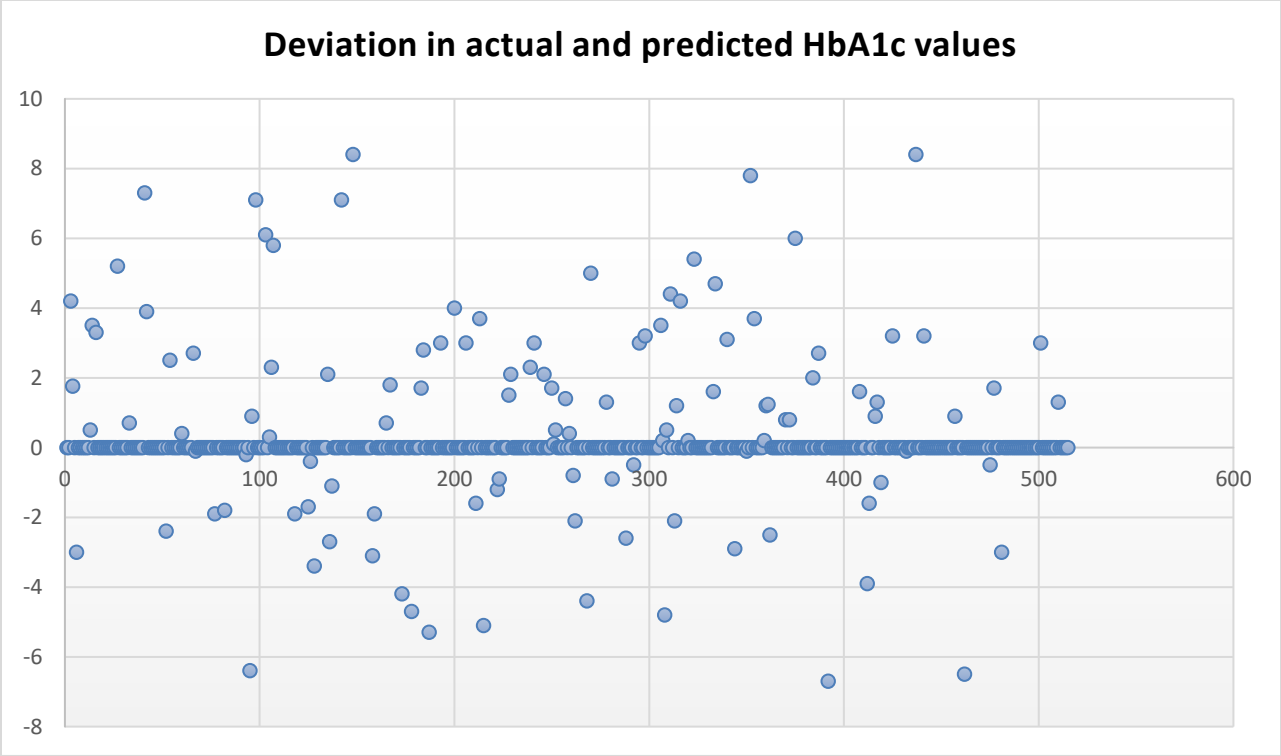


Expression profiling indicates that HbA1c levels are influenced by GLUT4 expression. Additionally, medication adherence significantly affects GLUT4 expression. Patients who adhere properly to their prescribed medication regimen typically exhibit higher GLUT4 expression. Conversely, those with poor adherence often show reduced GLUT4

expression. Thus, effective medication adherence is associated with better GLUT4 expression. Based on the observed relationship between GLUT4 expression and medication adherence, we can hypothesize that GLUT4 may serve as a key factor in the development of future applications designed to monitor and improve medication adherence.

4.8. Success rate of application

By comparing the actual and predicted values of HbA1c, the following graph illustrates the deviation of the predicted values from the actual values. The predicted values exhibited a mean deviation of 0.20, indicating that the app achieved a success rate of 80%.



CHAPTER 5: DISSCUSION

Numerous studies have investigated machine learning (ML) models for predicting diabetes, employing diverse algorithms and techniques to improve prediction accuracy. Comparative analyses have shown that logistic regression, (Joshi & Dhakal, 2021; Rajendra & Latifi, 2021) artificial neural networks (ANNs) (Bukhari et al., 2021; Butt et al., 2021), and decision tree models achieve accuracies ranging from 73% to 78% in predicting diabetes or prediabetes using common risk factors. Gradient boosting algorithms, such as Friedman's gradient boosting decision tree method, have demonstrated success in clinical settings by accurately predicting variables like BMI with an accuracy of 0.91 (Nusrat et al., 2020; Rufo et al., 2021; Seto et al., 2022).

Ensemble learning methods, including decision tree and decision forest regression, have also exhibited strong performance in predicting continuous variables such as blood glucose levels (Mahesh et al., 2022). Deep learning models, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), excel in capturing complex patterns within large datasets. Support vector machines (SVMs) and random forest models have proven effective in classifying individuals as diabetic or non-diabetic.

Furthermore, longitudinal data analysis leveraging RNNs has played a critical role in forecasting diabetes onset by identifying temporal patterns (Afsaneh et al., 2022; Gupta et al., 2022; Refat et al., 2021). These studies underscore ongoing efforts to enhance the accuracy and reliability of ML models for diabetes prediction through diverse methodologies and continuous advancements in research.

In our study, we conducted an analysis using a decision tree regressor, which emerged as the most effective model among those evaluated. Our research leveraged health records from a substantial cohort of 100,513 patients. Emphasizing usability and accessibility, we carefully selected the most pertinent parameters to ensure practical application of the model.

The decision tree model we developed represents a robust tool for predicting HbA1c levels and elucidating correlations with medication adherence among patients. This

predictive capability holds significant promise for enhancing patient care outcomes and optimizing healthcare resource allocation.

As a pivotal aspect of our initiative, we have transformed this model into a web application, as depicted in Figure. We envision this application as a valuable asset that will deliver substantial benefits to both patients and the broader healthcare system. By facilitating streamlined access to predictive insights, healthcare providers can make informed decisions promptly, thereby improving patient management and treatment outcomes.

Furthermore, our endeavor underscores a commitment to advancing precision medicine through innovative technological solutions. By harnessing the power of machine learning and data-driven approaches, we aim to empower healthcare professionals with tools that facilitate personalized and proactive patient care.

In conclusion, the deployment of our decision tree regressor model as a web application marks a significant milestone in bridging the gap between research findings and practical healthcare applications. We are optimistic that this initiative will contribute positively to the ongoing evolution of healthcare delivery, ultimately benefiting patient health and well-being on a broader scale.

REFERENCES

1. Adua, E., Kolog, E. A., Afrifa-Yamoah, E., Amankwah, B., Obirikorang, C., Anto, E. O., Acheampong, E., Wang, W., & Tetteh, A. Y. (2021). Predictive model and feature importance for early detection of type II diabetes mellitus. *Translational Medicine Communications*, *6*(1), 17. https://doi.org/10.1186/s41231-021-00096-z

2. Afsaneh, E., Sharifdini, A., Ghazzaghi, H., & Ghobadi, M. Z. (2022). Recent applications of machine learning and deep learning models in the prediction, diagnosis, and management of diabetes: A comprehensive review. *Diabetology & Metabolic Syndrome*, *14*(1), 196. https://doi.org/10.1186/s13098-022-00969-9

3. Alam, F., Asiful Islam, Md., Ibrahim Khalil, Md., & Hua Gan, S. (2016). Metabolic Control of Type 2 Diabetes by Targeting the GLUT4 Glucose Transporter: Intervention Approaches. *Current Pharmaceutical Design*, *22*(20), 3034–3049.

4. Alanazi, H. O., Abdullah, A. H., Qureshi, K. N., & Ismail, A. S. (2018). Accurate and dynamic predictive model for better prediction in medicine and healthcare. *Irish Journal of Medical Science*, *187*(2), 501–513. https://doi.org/10.1007/s11845-017-1655-3

5. Årsand, E., Frøisland, D. H., Skrøvseth, S. O., Chomutare, T., Tatara, N., Hartvigsen, G., & Tufano, J. T. (2012). Mobile Health Applications to Assist Patients with Diabetes: Lessons Learned and Design Implications. *Journal of*

Diabetes Science and Technology, *6*(5), 1197–1206. https://doi.org/10.1177/193229681200600525

6. Bennett, C. M., Guo, M., & Dharmage, S. C. (2007). HbA1c as a screening tool for detection of Type 2 diabetes: A systematic review. *Diabetic Medicine*, *24*(4), 333–343. https://doi.org/10.1111/j.1464-5491.2007.02106.x

7. Brunton, S. A., & Polonsky, W. H. (2017). Medication adherence in type 2 diabetes mellitus: Real-world strategies for addressing a common problem. *Journal of Family Practice*, *66*(4), S46–S46.

8. Bukhari, M. M., Alkhamees, B. F., Hussain, S., Gumaei, A., Assiri, A., & Ullah, S. S. (2021). An Improved Artificial Neural Network Model for Effective Diabetes Prediction. *Complexity*, *2021*(1), 5525271. https://doi.org/10.1155/2021/5525271

9. Butt, U. M., Letchmunan, S., Ali, M., Hassan, F. H., Baqir, A., & Sherazi, H. H. R. (2021). Machine Learning Based Diabetes Classification and Prediction for Healthcare Applications. *Journal of Healthcare Engineering*, *2021*(1), 9930985. https://doi.org/10.1155/2021/9930985

10. Ceriello, A., Gallo, M., Candido, R., De Micheli, A., Esposito, K., Gentile, S., & Medea, G. (2014). Personalized therapy algorithms for type 2 diabetes: A phenotype-based approach. *Pharmacogenomics and Personalized Medicine*, *7*, 129–136. https://doi.org/10.2147/PGPM.S50288

11. Chatterjee, S., Khunti, K., & Davies, M. J. (2017). Type 2 diabetes. *The Lancet*, *389*(10085), 2239–2251. https://doi.org/10.1016/S0140-6736(17)30058-2

12. Choi, S. A., Suh, H. J., Yun, J. W., & Choi, J. W. (2012). Differential gene expression in pancreatic tissues of streptozocin-induced diabetic rats and genetically-diabetic mice in response to hypoglycemic dipeptide cyclo (His-Pro) treatment. *Molecular Biology Reports*, *39*(9), 8821–8835. https://doi.org/10.1007/s11033-012-1746-1

13. Contreras, I., & Vehi, J. (2018). Artificial Intelligence for Diabetes Management and Decision Support: Literature Review. *Journal of Medical Internet Research*, *20*(5), e10775. https://doi.org/10.2196/10775

14. Ellahham, S. (2020). Artificial Intelligence: The Future for Diabetes Care. *The American Journal of Medicine*, *133*(8), 895–900. https://doi.org/10.1016/j.amjmed.2020.03.033

15. Fan, Y., Long, E., Cai, L., Cao, Q., Wu, X., & Tong, R. (2021). Machine Learning Approaches to Predict Risks of Diabetic Complications and Poor Glycemic Control in Nonadherent Type 2 Diabetes. *Frontiers in Pharmacology*, *12*. https://doi.org/10.3389/fphar.2021.665951

16. Fioravanti, A., Fico, G., Salvi, D., García-Betances, R. I., & Arredondo, M. T. (2015). Automatic messaging for improving patients engagement in diabetes management: An exploratory study. *Medical & Biological Engineering & Computing*, *53*(12), 1285–1294. https://doi.org/10.1007/s11517-014-1237-8

17. Florkowski, C. (2013). HbA1c as a Diagnostic Test for Diabetes Mellitus – Reviewing the Evidence. *The Clinical Biochemist Reviews*, *34*(2), 75–83.

18. Foretz, M., Guigas, B., & Viollet, B. (2019). Understanding the glucoregulatory mechanisms of metformin in type 2 diabetes mellitus. *Nature Reviews Endocrinology*, *15*(10), 569–589. https://doi.org/10.1038/s41574-019-0242-2

19. Franks, P. W., Pearson, E., & Florez, J. C. (2013). Gene-Environment and Gene-Treatment Interactions in Type 2 Diabetes: Progress, pitfalls, and prospects. *Diabetes Care*, *36*(5), 1413–1421. https://doi.org/10.2337/dc12-2211

20. Fregoso-Aparicio, L., Noguez, J., Montesinos, L., & García-García, J. A. (2021). Machine learning and deep learning predictive models for type 2 diabetes: A systematic review. *Diabetology & Metabolic Syndrome*, *13*(1), 148. https://doi.org/10.1186/s13098-021-00767-9

21. Fu, Y., Xu, L., Zhang, H., Ding, N., Zhang, J., Ma, S., Yang, A., Hao, Y., Gao, Y., & Jiang, Y. (2023). Identification and Validation of Immune-Related Genes Diagnostic for Progression of Atherosclerosis and Diabetes. *Journal of Inflammation Research*, *16*, 505–521. https://doi.org/10.2147/JIR.S393788

22. Galicia-Garcia, U., Benito-Vicente, A., Jebari, S., Larrea-Sebal, A., Siddiqi, H., Uribe, K. B., Ostolaza, H., & Martín, C. (2020). Pathophysiology of Type 2 Diabetes Mellitus. *International Journal of Molecular Sciences*, *21*(17), 6275. https://doi.org/10.3390/ijms21176275

23. Gautier, T., Ziegler, L. B., Gerber, M. S., Campos-Náñez, E., & Patek, S. D. (2021). Artificial intelligence and diabetes technology: A review. *Metabolism*, *124*, 154872. https://doi.org/10.1016/j.metabol.2021.154872

24. Goldstein, B. J. (2002). Insulin resistance as the core defect in type 2 diabetes mellitus. *The American Journal of Cardiology*, *90*(5, Supplement 1), 3–10. https://doi.org/10.1016/S0002-9149(02)02553-5

25. Grisouard, J., Timper, K., Radimerski, T. M., Frey, D. M., Peterli, R., Kola, B., Korbonits, M., Herrmann, P., Krähenbühl, S., Zulewski, H., Keller, U., Müller, B., & Christ-Crain, M. (2010). Mechanisms of metformin action on glucose transport and metabolism in human adipocytes. *Biochemical Pharmacology*, *80*(11), 1736–1745. https://doi.org/10.1016/j.bcp.2010.08.021

26. Guan, Z., Li, H., Liu, R., Cai, C., Liu, Y., Li, J., Wang, X., Huang, S., Wu, L., Liu, D., Yu, S., Wang, Z., Shu, J., Hou, X., Yang, X., Jia, W., & Sheng, B. (2023). Artificial intelligence in diabetes management: Advancements, opportunities, and challenges. *Cell Reports Medicine*, *4*(10), 101213. https://doi.org/10.1016/j.xcrm.2023.101213

27. Gupta, H., Varshney, H., Sharma, T. K., Pachauri, N., & Verma, O. P. (2022). Comparative performance analysis of quantum machine learning with deep learning for diabetes prediction. *Complex & Intelligent Systems*, *8*(4), 3073–3087. https://doi.org/10.1007/s40747-021-00398-7

28. Heidary, M., & Kakhki, M. P. (2014). *TRIzol-based RNA Extraction: A Reliable Method for Gene Expression Studies*. *25*(1).

29. Himanshu, D., Ali, W., & Wamique, M. (2020). Type 2 diabetes mellitus: Pathogenesis and genetic diagnosis. *Journal of Diabetes and Metabolic Disorders*, *19*(2), 1959–1966. https://doi.org/10.1007/s40200-020-00641-x

30. Hundal, R. S., Krssak, M., Dufour, S., Laurent, D., Lebon, V., Chandramouli, V., Inzucchi, S. E., Schumann, W. C., Petersen, K. F., Landau, B. R., & Shulman, G. I. (2000). Mechanism by which metformin reduces glucose production in type 2 diabetes. *Diabetes*, *49*(12), 2063–2069. https://doi.org/10.2337/diabetes.49.12.2063

31. Islam, S. M. S., Mishra, V., Siddiqui, M. U., Moses, J. C., Adibi, S., Nguyen, L., & Wickramasinghe, N. (2022). Smartphone Apps for Diabetes Medication Adherence: Systematic Review. *JMIR Diabetes*, *7*(2), e33264. https://doi.org/10.2196/33264

32. Istepanian, R. S. H., Casiglia, D., & Gregory, J. W. (2017). Mobile health (m-Health) for diabetes management. *British Journal of Healthcare Management*, *23*(3), 102–108. https://doi.org/10.12968/bjhc.2017.23.3.102

33. Jackson, R. A., Hawa, M. I., Jaspan, J. B., Sim, B. M., DiSilvio, L., Featherbe, D., & Kurtz, A. B. (1987). Mechanism of Metformin Action in Non-Insulin-Dependent Diabetes. *Diabetes*, *36*(5), 632–640. https://doi.org/10.2337/diab.36.5.632

34. Joshi, R. D., & Dhakal, C. K. (2021). Predicting Type 2 Diabetes Using Logistic Regression and Machine Learning Approaches. *International Journal of Environmental Research and Public Health*, *18*(14), Article 14. https://doi.org/10.3390/ijerph18147346

35. Jude, E. B., Malecki, M. T., Gomez Huelgas, R., Prazny, M., Snoek, F., Tankova, T., Giugliano, D., & Khunti, K. (2022). Expert Panel Guidance and Narrative Review of Treatment Simplification of Complex Insulin Regimens to Improve Outcomes in Type 2 Diabetes. *Diabetes Therapy*, *13*(4), 619–634. https://doi.org/10.1007/s13300-022-01222-2

36. Kaku, K. (2010). Pathophysiology of type 2 diabetes and its treatment policy. *Japan Medical Association Journal*, *53*, 41–46.

37. Keller, M. P., Choi, Y., Wang, P., Davis, D. B., Rabaglia, M. E., Oler, A. T., Stapleton, D. S., Argmann, C., Schueler, K. L., Edwards, S., Steinberg, H. A., Neto, E. C., Kleinhanz, R., Turner, S., Hellerstein, M. K., Schadt, E. E., Yandell, B. S., Kendziorski, C., & Attie, A. D. (2008). A gene expression network model of type 2 diabetes links cell cycle regulation in islets with diabetes susceptibility. *Genome Research*, *18*(5), 706–716. https://doi.org/10.1101/gr.074914.107

38. Klip, A., & Leiter, L. A. (1990). Cellular Mechanism of Action of Metformin. *Diabetes Care*, *13*(6), 696–704. https://doi.org/10.2337/diacare.13.6.696

39. Lambrinou, E., Hansen, T. B., & Beulens, J. W. (2019). Lifestyle factors, self-management and patient empowerment in diabetes care. *European Journal of*

Preventive Cardiology, *26*(2_suppl), 55–63. https://doi.org/10.1177/2047487319885455

40. LaMoia, T. E., & Shulman, G. I. (2021). Cellular and Molecular Mechanisms of Metformin Action. *Endocrine Reviews*, *42*(1), 77–96. https://doi.org/10.1210/endrev/bnaa023

41. Li, Y., Zhong, Q., Zhu, S., Cheng, H., Huang, W., Wang, H. H. X., & Li, Y.-T. (2022). Frequency of Follow-Up Attendance and Blood Glucose Monitoring in Type 2 Diabetic Patients at Moderate to High Cardiovascular Risk: A Cross-Sectional Study in Primary Care. *International Journal of Environmental Research and Public Health*, *19*(21), Article 21. https://doi.org/10.3390/ijerph192114175

42. Lin, Y., & Sun, Z. (2010). Current views on type 2 diabetes. *Journal of Endocrinology*, *204*(1), 1–11. https://doi.org/10.1677/JOE-09-0260

43. Lindenmeyer, A., Hearnshaw, H., Vermeire, E., Van Royen, P., Wens, J., & Biot, Y. (2006). Interventions to improve adherence to medication in people with type 2 diabetes mellitus: A review of the literature on the role of pharmacists. *Journal of Clinical Pharmacy and Therapeutics*, *31*(5), 409–419. https://doi.org/10.1111/j.1365-2710.2006.00759.x

44. Ma, R. C. W., & Tong, P. C. Y. (2024). Epidemiology of Type 2 Diabetes. In *Textbook of Diabetes* (pp. 55–74). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119697473.ch5

45. Mahesh, T. R., Kumar, D., Vinoth Kumar, V., Asghar, J., Mekcha Bazezew, B., Natarajan, R., & Vivek, V. (2022). Blended Ensemble Learning Prediction Model for Strengthening Diagnosis and Treatment of Chronic Diabetes Disease. *Computational Intelligence and Neuroscience, 2022*(1), 4451792. https://doi.org/10.1155/2022/4451792

46. Mannhalter, C., Koizar, D., & Mitterbauer, G. (2000). *Evaluation of RNA Isolation Methods and Reference Genes for RT-PCR Analyses of Rare Target RNA*. *38*(2), 171–177. https://doi.org/10.1515/CCLM.2000.026

47. Markovič, R., Grubelnik, V., Završnik, T., Blažun Vošner, H., Kokol, P., Perc, M., Marhl, M., Završnik, M., & Završnik, J. (2023). Profiling of patients with type 2 diabetes based on medication adherence data. *Frontiers in Public Health, 11*. https://doi.org/10.3389/fpubh.2023.1209809

48. Misra-Hebert, A. D., Milinovich, A., Zajichek, A., Ji, X., Hobbs, T. D., Weng, W., Petraro, P., Kong, S. X., Mocarski, M., Ganguly, R., Bauman, J. M., Pantalone, K. M., Zimmerman, R. S., & Kattan, M. W. (2020). Natural Language Processing Improves Detection of Nonsevere Hypoglycemia in Medical Records Versus Coding Alone in Patients With Type 2 Diabetes but Does Not Improve Prediction of Severe Hypoglycemia Events: An Analysis Using the Electronic Medical Record in a Large Health System. *Diabetes Care*, *43*(8), 1937–1940. https://doi.org/10.2337/dc19-1791

49. Muralidharan, S., Ranjani, H., Anjana, R. M., Allender, S., & Mohan, V. (2017). Mobile Health Technology in the Prevention and Management of Type 2

Diabetes. *Indian Journal of Endocrinology and Metabolism*, *21*(2), 334. https://doi.org/10.4103/ijem.IJEM_407_16

50. Nguyen, B. P., Pham, H. N., Tran, H., Nghiem, N., Nguyen, Q. H., Do, T. T. T., Tran, C. T., & Simpson, C. R. (2019). Predicting the onset of type 2 diabetes using wide and deep learning with electronic health records. *Computer Methods and Programs in Biomedicine*, *182*, 105055. https://doi.org/10.1016/j.cmpb.2019.105055

51. Nomura, A., Noguchi, M., Kometani, M., Furukawa, K., & Yoneda, T. (2021). Artificial Intelligence in Current Diabetes Management and Prediction. *Current Diabetes Reports*, *21*(12), 61. https://doi.org/10.1007/s11892-021-01423-2

52. Nusrat, F., Uzbaş, B., & Baykan, Ö. K. (2020). Prediction of Diabetes Mellitus by using Gradient Boosting Classification. *Avrupa Bilim ve Teknoloji Dergisi*, 268–272. https://doi.org/10.31590/ejosat.803504

53. Ostenson, C. G. (2001). The pathophysiology of type 2 diabetes mellitus: An overview. *Acta Physiologica Scandinavica*, *171*(3), 241–247. https://doi.org/10.1046/j.1365-201x.2001.00826.x

54. Ozery-Flato, M., Parush, N., El-Hay, T., Visockienė, Ž., Ryliškytė, L., Badarienė, J., Solovjova, S., Kovaitė, M., Navickas, R., & Laucevičius, A. (2013). Predictive models for type 2 diabetes onset in middle-aged subjects with the metabolic syndrome. *Diabetology & Metabolic Syndrome*, *5*(1), 36. https://doi.org/10.1186/1758-5996-5-36

55. Rajendra, P., & Latifi, S. (2021). Prediction of diabetes using logistic regression and ensemble techniques. *Computer Methods and Programs in Biomedicine Update*, *1*, 100032. https://doi.org/10.1016/j.cmpbup.2021.100032

56. Refat, M. A. R., Amin, Md. A., Kaushal, C., Yeasmin, M. N., & Islam, M. K. (2021). A Comparative Analysis of Early Stage Diabetes Prediction using Machine Learning and Deep Learning Approach. *2021 6th International Conference on Signal Processing, Computing and Control (ISPCC)*, 654–659. https://doi.org/10.1109/ISPCC53510.2021.9609364

57. Rena, G., Hardie, D. G., & Pearson, E. R. (2017). The mechanisms of action of metformin. *Diabetologia*, *60*(9), 1577–1585. https://doi.org/10.1007/s00125-017-4342-z

58. Rena, G., Pearson, E. R., & Sakamoto, K. (2013). Molecular mechanism of action of metformin: Old or new insights? *Diabetologia*, *56*(9), 1898–1906. https://doi.org/10.1007/s00125-013-2991-0

59. Rufo, D. D., Debelee, T. G., Ibenthal, A., & Negera, W. G. (2021). Diagnosis of Diabetes Mellitus Using Gradient Boosting Machine (LightGBM). *Diagnostics*, *11*(9), Article 9. https://doi.org/10.3390/diagnostics11091714

60. Samsom, M., Trivedi, T., Orekoya, O., & Vyas, S. (2016). Understanding the Importance of Gene and Environment in the Etiology and Prevention of Type 2 Diabetes Mellitus in High-Risk Populations. *Oral Health Case Reports*, *2*(1), 112.

61. Scheen, A. J. (2003). Pathophysiology of type 2 diabetes. *Acta Clinica Belgica*, *58*(6), 335–341. https://doi.org/10.1179/acb.2003.58.6.001

62. Seto, H., Oyama, A., Kitora, S., Toki, H., Yamamoto, R., Kotoku, J., Haga, A., Shinzawa, M., Yamakawa, M., Fukui, S., & Moriyama, T. (2022). Gradient boosting decision tree becomes more reliable than logistic regression in predicting probability for diabetes with big data. *Scientific Reports*, *12*(1), 15889. https://doi.org/10.1038/s41598-022-20149-z

63. Shrivastava, T. P., Goswami, S., Gupta, R., & Goyal, R. K. (2023). Mobile App Interventions to Improve Medication Adherence Among Type 2 Diabetes Mellitus Patients: A Systematic Review of Clinical Trials. *Journal of Diabetes Science and Technology*, *17*(2), 458–466. https://doi.org/10.1177/19322968211060060

64. Silva, K. D., Lee, W. K., Forbes, A., Demmer, R. T., Barton, C., & Enticott, J. (2020). Use and performance of machine learning models for type 2 diabetes prediction in community settings: A systematic review and meta-analysis. *International Journal of Medical Informatics*, *143*, 104268. https://doi.org/10.1016/j.ijmedinf.2020.104268

65. Sirsikar, M., Supriya, S., Mohanty, S., & Pinnelli, V. (2016). Role of glycated hemoglobin (HBA1c) as a dual marker to predict glycemic status and dyslipidemia in type II diabetes mellitus. *International Journal of Research in Medical Sciences*, 4524–4529. https://doi.org/10.18203/2320-6012.ijrms20163322

66. Szabo, M., Máté, B., Csép, K., & Benedek, T. (2018). Genetic Approaches to the Study of Gene Variants and Their Impact on the Pathophysiology of Type 2 Diabetes. *Biochemical Genetics*, *56*(1), 22–55. https://doi.org/10.1007/s10528-017-9827-4

67. Tanase, D. M., Gosav, E. M., Costea, C. F., Ciocoiu, M., Lacatusu, C. M., Maranduca, M. A., Ouatu, A., & Floria, M. (2020). The Intricate Relationship between Type 2 Diabetes Mellitus (T2DM), Insulin Resistance (IR), and Nonalcoholic Fatty Liver Disease (NAFLD). *Journal of Diabetes Research*, *2020*, 1–16. https://doi.org/10.1155/2020/3920196

68. Taybani, Z., Bótyik, B., Katkó, M., Gyimesi, A., & Várkonyi, T. (2019). Simplifying Complex Insulin Regimens While Preserving Good Glycemic Control in Type 2 Diabetes. *Diabetes Therapy*, *10*(5), 1869–1878. https://doi.org/10.1007/s13300-019-0673-8

69. Thipsawat, S. (2021). Early detection of diabetic nephropathy in patient with type 2 diabetes mellitus: A review of the literature. *Diabetes and Vascular Disease Research*, *18*(6), 14791641211058856. https://doi.org/10.1177/14791641211058856

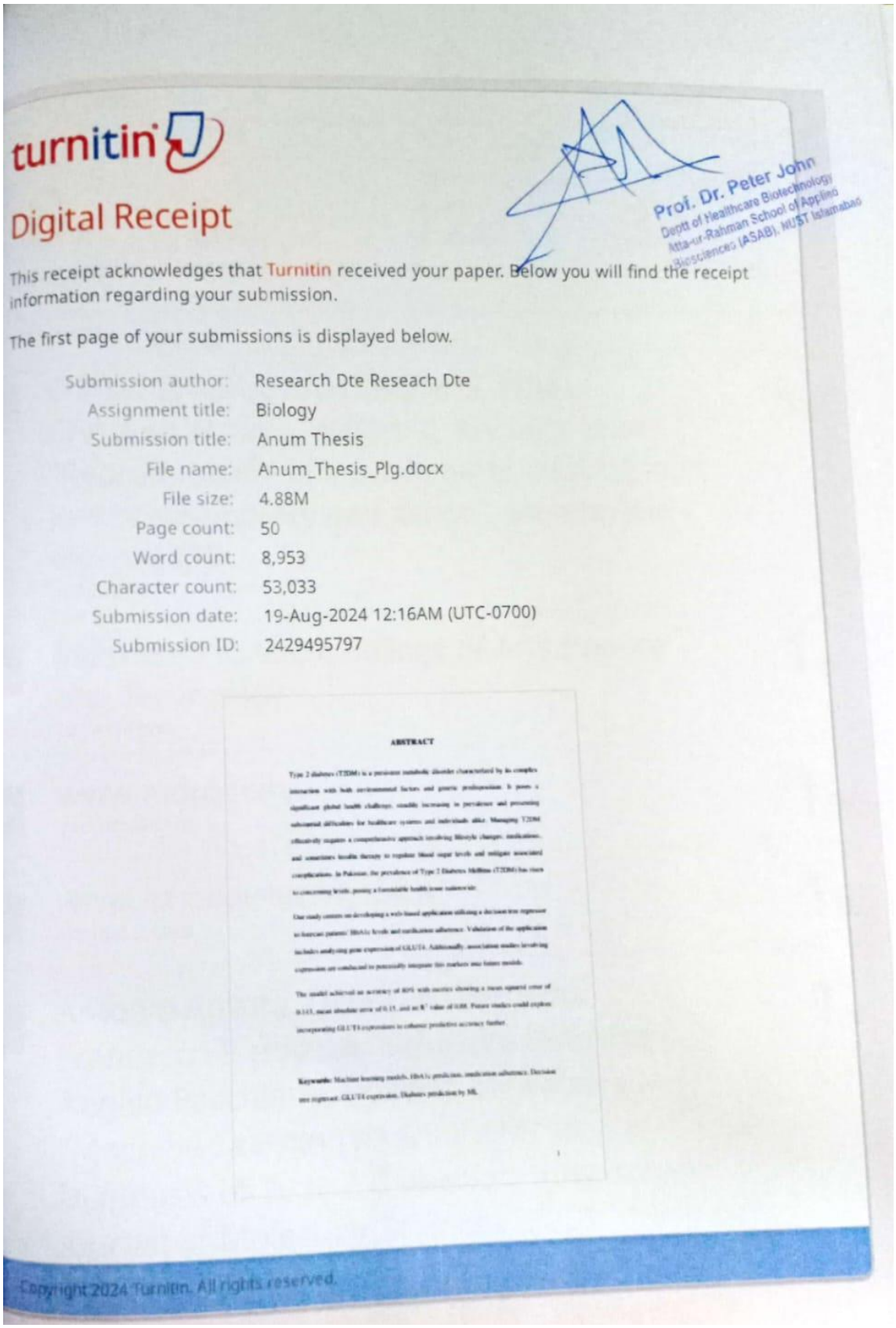
70. Vidyadharan, V., Hamdan, M., & Zalzala, A. M. S. (2021). An Evidence-Based Study of Diabetes Prevention and Management with NLP and Deep Learning. *2021 IEEE Symposium Series on Computational Intelligence (SSCI)*, 1–8. https://doi.org/10.1109/SSCI50451.2021.9660000

71. Vijayakumar, P., Nelson, R. G., Hanson, R. L., Knowler, W. C., & Sinha, M. (2016). HbA1c and the Prediction of Type 2 Diabetes in Children and Adults. *Diabetes Care*, *40*(1), 16–21. https://doi.org/10.2337/dc16-1358

72. Williams, D. M., Jones, H., & Stephens, J. W. (2022). Personalized Type 2 Diabetes Management: An Update on Recent Advances and Recommendations. *Diabetes, Metabolic Syndrome and Obesity*, *15*, 281–295. https://doi.org/10.2147/DMSO.S331654

73. Xiong, X., Zhang, R., Bi, Y., Zhou, W., Yu, Y., & Zhu, D. (2019). Machine Learning Models in Type 2 Diabetes Risk Prediction: Results from a Cross-sectional Retrospective Study in Chinese Adults. *Current Medical Science*, *39*(4), 582–588. https://doi.org/10.1007/s11596-019-2077-4

74. Zheng, T., Xie, W., Xu, L., He, X., Zhang, Y., You, M., Yang, G., & Chen, Y. (2017). A machine learning-based framework to identify type 2 diabetes through electronic health records. *International Journal of Medical Informatics*, *97*, 120–127. https://doi.org/10.1016/j.ijmedinf.2016.09.014



turnitin

Digital Receipt

Prof. Dr. Peter John
Deptt of Healthcare Biotechnology
Atta-ur-Rahman School of Applied Biosciences (ASAB), NUST Islamabad

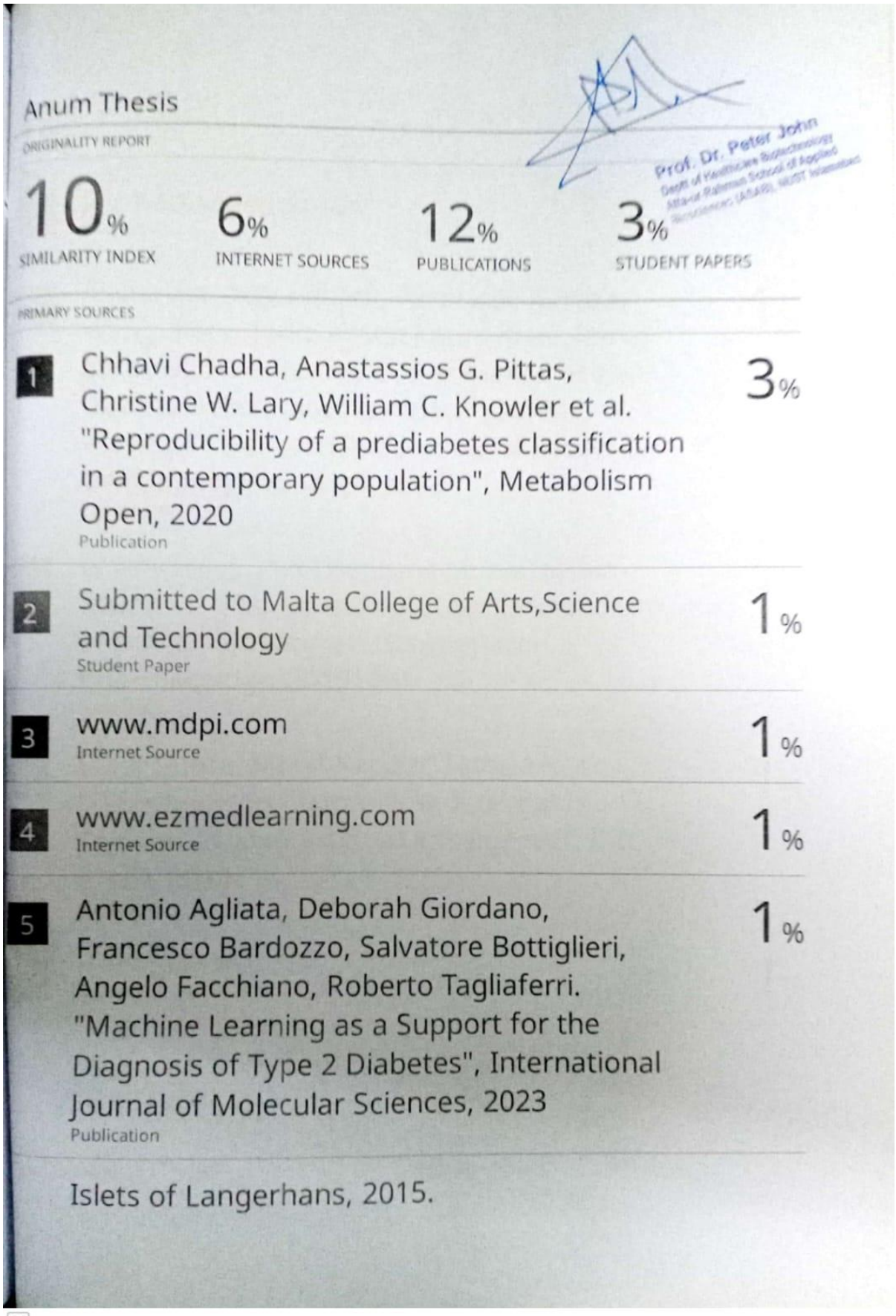
This receipt acknowledges that Turnitin received your paper. Below you will find the receipt information regarding your submission.

The first page of your submissions is displayed below.

Submission author:	Research Dte Reseach Dte
Assignment title:	Biology
Submission title:	Anum Thesis
File name:	Anum_Thesis_Plg.docx
File size:	4.88M
Page count:	50
Word count:	8,953
Character count:	53,033
Submission date:	19-Aug-2024 12:16AM (UTC-0700)
Submission ID:	2429495797

ABSTRACT

Copyright 2024 Turnitin. All rights reserved.



Anum Thesis

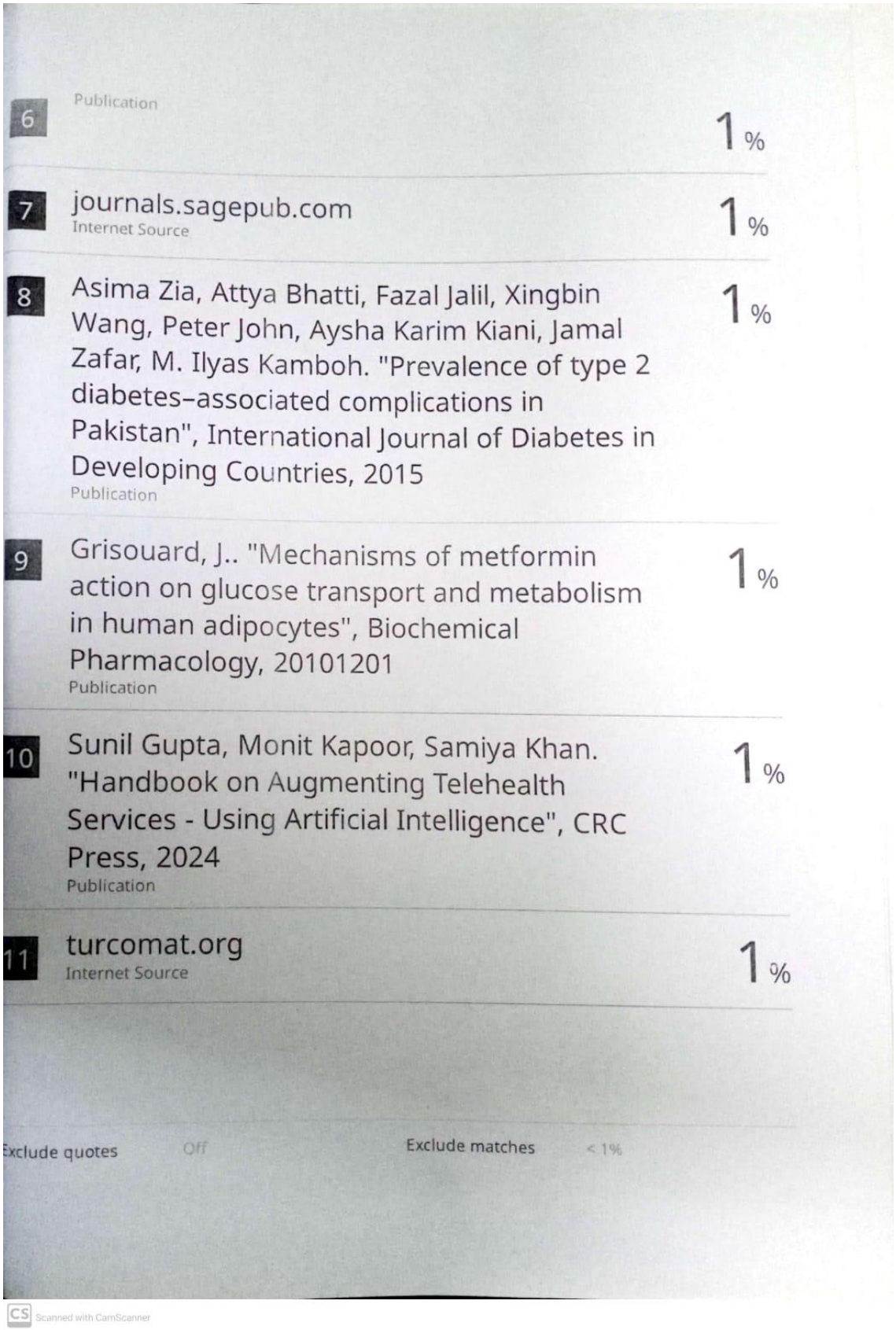
ORIGINALITY REPORT

10%	6%	12%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Chhavi Chadha, Anastassios G. Pittas, Christine W. Lary, William C. Knowler et al. "Reproducibility of a prediabetes classification in a contemporary population", Metabolism Open, 2020 Publication	3%
2	Submitted to Malta College of Arts,Science and Technology Student Paper	1%
3	www.mdpi.com Internet Source	1%
4	www.ezmedlearning.com Internet Source	1%
5	Antonio Agliata, Deborah Giordano, Francesco Bardozzo, Salvatore Bottiglieri, Angelo Facchiano, Roberto Tagliaferri. "Machine Learning as a Support for the Diagnosis of Type 2 Diabetes", International Journal of Molecular Sciences, 2023 Publication	1%

Islets of Langerhans, 2015.



6	Publication	1 %
7	journals.sagepub.com Internet Source	1 %
8	Asima Zia, Attya Bhatti, Fazal Jalil, Xingbin Wang, Peter John, Aysha Karim Kiani, Jamal Zafar, M. Ilyas Kamboh. "Prevalence of type 2 diabetes–associated complications in Pakistan", International Journal of Diabetes in Developing Countries, 2015 Publication	1 %
9	Grisouard, J.. "Mechanisms of metformin action on glucose transport and metabolism in human adipocytes", Biochemical Pharmacology, 20101201 Publication	1 %
10	Sunil Gupta, Monit Kapoor, Samiya Khan. "Handbook on Augmenting Telehealth Services - Using Artificial Intelligence", CRC Press, 2024 Publication	1 %
11	turcomat.org Internet Source	1 %

Exclude quotes Off Exclude matches < 1%