Diabetes Prediction Using Medical Variables: Analysis & Data Visualization

Anddrew Richmond Thezo
Business & Communication Department
Swiss German University
Tangerang, Indonesia
anddrew.thezo@student.sgu.ac.id

***Abstract*—Diabetes is a chronic illness that develops when the body either cannot use the insulin that the pancreas produces properly or does not produce enough of it. One hormone that controls blood sugar is insulin. Approximately 48% of all deaths caused by diabetes occurred before the age of 70 in 2019. Diabetes was the direct cause of 1.5 million deaths in 2019 based on the report from WHO (World Health Organization). This study shows the classification of whether someone has diabetes or not using the 8 datasets (medical variables) of age, gender, body mass index (BMI), hypertension (blood pressure), heart disease, smoking history, HbA1c level, and blood glucose level as the risk factors to predict diabetes in patients based on their medical history and demographic information. Furthermore, the result of this study will be presented with analysis and data visualization.**

Keywords—chronic illness, data visualization, diabetes prediction, medical variables

I. INTRODUCTION

Diabetes is a serious illness that has been recognized and described in ancient scripts, but it does not seem to have been a common occurrence for healers or physicians. The number of people with this condition has increased over the past few decades, which has an increasing impact on human health and development (Roglic, 2016). There are two types of diabetes. Type 1 diabetes, also referred to as juvenile or childhood-onset insulin-dependent diabetes, is characterized by insufficient insulin production and necessitates daily insulin administration. Nine million individuals worldwide had type 1 diabetes in 2017, with the bulk residing in affluent nations). The reason behind it is unknown, as are the ways to stop it. Type 2 diabetes, prevents the body from using insulin as it should, which, if left untreated, can result in elevated blood sugar levels. Type 2 diabetes, in particular, has the potential to seriously harm blood vessels and nerves in the body over time. Oftentimes, type 2 diabetes is avoidable. Being overweight, not exercising enough, and heredity are factors that can lead to the development of type 2 diabetes [2].

Numerous prospective and cross-sectional studies have verified the link between type 2 diabetes and obesity as risk factors. The majority of individuals with type 2 diabetes are overweight or obese; in southeast Scotland in 2005, over 85% of those with the disease had a body mass index (BMI) of more than 25, which is calculated by dividing weight in kilograms by height in meters squared. According to recent research, a high waist circumference may be a more reliable predictor of an increased risk of type 2 diabetes than a body mass index (BMI) [3] Besides, another study [4] shows that the prevalence of diabetes increased with age and body mass index and increased inversely with energy expenditure in both genders of males and females. Understanding the previous study results from the factors that influence diabetes, it is crucial to know the triggering factors; Therefore, this study will use a dataset of diabetes prediction with the medical variables ranging from age, gender, body mass index (BMI), hypertension (blood pressure), heart disease, smoking history, HbA1c level, and blood glucose level as the variables to predict diabetes in patients based on their medical history and show the result coloration of the variables between the outcome which is having diabetes or absence of diabetes presence that will be presented with data visualization.

II. LITERATURE REVIEW

Reducing the prevalence and subsequent complications of diabetes mellitus may be possible by focusing the control and prevention strategy on such modifiable risk factors linked to diabetes and prediabetes [5]. Here are the medical variables as the risk factors explanation used in the data to forecast a patient's risk of diabetes based on their medical history.

A. Age

Advanced age is a major risk factor for diabetes and prediabetes [6], [7], [8]. Compared to the young and middle-aged, the elderly have a higher prevalence of diabetes and prediabetes and are more likely to experience problems with their cardiovascular, retinal, and renal systems [6], [9], [10]. Based on previous studies, type 1 diabetes can occur at any age but is rare in the first year of life. In most populations, the incidence steadily increases with age up to puberty and is higher among those aged 35 years [11]. Significant personal health costs are associated with diabetes; for instance, a 60-year-old diagnosed with the disease shortens quality-of-life years by 11.8 years for women and 12.1 years for men [12]. Furthermore, a different study conduct mentioned that diabetes progressively increases with age. In 2005, the prevalence of diabetes in the United States was estimated to be 0.22% for persons <20 years of age and 9.6% for those >20 years of age. In individuals >60 years old, the prevalence of diabetes was 20.9% [13]. According to the Taiwan Nationwide Health Insurance database from 2000 to 2009, the diabetes prevalence rates were 21.97% for women aged 60–79 years and 23.97% for women aged 80 years and older. For men in the same age groups, the prevalence rates were 19.97% and 20.27%. The high prevalence of diabetes mellitus in the elderly challenges the medical systems [14].

B. Gender

Type 2 diabetes mellitus is becoming more common in both sexes, although men are typically diagnosed with the disease at a younger age and with less body fat than women. An estimated 17.7 million men and women worldwide suffer from diabetes mellitus. When type 2 diabetes is diagnosed, women seem to have more risk factors, particularly obesity.

 EISSN: 2961-872X ● PISSN: 2961-8924 ● DOI: 10.56741/esl.v3i01.472

Furthermore, women's diabetes risk may be more significantly influenced by psychosocial stress [15]. Men are more likely to have diabetes before puberty, whereas women are more likely to have diabetes after menopause and in old age. This gender disparity in diabetes prevalence is inversely correlated with reproductive life stages [16]. According to the Centers for Disease Control and Prevention (CDC), men are more likely to receive a diagnosis of diabetes than women. However, some research suggests that women with diabetes may be more likely to develop complications than men [17].

C. Body Mass Index (BMI)

The definition of overweight and obesity is an abnormal or excessive fat accumulation that poses a health risk. Over 25 is classified as overweight, and over 30 as obese based on body mass index (BMI) [18]. According recent study on BMI and Diabetes shows that type 2 diabetes is a condition that can be brought on by an excessive amount of body fat accumulation, and the risk of type 2 diabetes rises linearly with body mass index. As a result, the prevalence of type 2 diabetes has increased along with the global increase in obesity prevalence [19]. Another study revealed that those with a BMI greater than 35 kg/m2 had an increased risk of developing type II diabetes compared to those with a BMI lower than 23 kg/m2 [20] Compared to women of normal weight, women with a body mass index (BMI) of 30 kg/m2 have a 28-fold increased risk of developing diabetes. At 35 kg/m2, the risk of diabetes is 93 times higher [21].

D. Hypertension (Blood Pressure)

One of the main risk factors for the development of diabetes is hypertension. Individuals with normal blood pressure are not as likely to develop diabetes as those with hypertension, which increases the risk by 2-3 times [22]. Type 2 diabetes and hypertension are frequently co-occurring conditions. Patients with diabetes have twice as many cases of hypertension as people without the disease. Furthermore, compared to people with normotension, patients with hypertension frequently show signs of insulin resistance and are more likely to develop diabetes [23].

E. Heart Disease

A previous study analyzed the connection between Heart Failure and Diabetes mellitus in a group of older individuals and showed that having a history of Heart Failure is linked to a twofold rise in the likelihood of developing new-onset Diabetes mellitus within a span of three to four years, in comparison to a population without any previous Heart Failure [24]. Besides, this is also supported by a prior study showing that Diabetes and heart failure are closely connected: individuals with diabetes have a higher chance of developing heart failure, and those with heart failure are more likely to develop diabetes [25]. Diabetes mellitus is very common among patients with heart failure, particularly those with heart failure and preserved ejection fraction (HFpEF). Patients with both conditions have a greater risk of mortality compared to patients who do not have diabetes or heart failure [25].

F. Smoking History

Smoking has been shown as an independent risk factor for type 2 diabetes mellitus and has a combined effect with low insulin secretion and high insulin resistance in developing the condition. Based on the data used in the study, 80% of men were smokers as compared with only about 5% of women. After adjustment for many confounders, cigarette smoking was an independent risk factor for type 2 diabetes mellitus. [26]Another prior study conducted in Japan reported that Individuals who were currently smoking 16-25 cigarettes per day had a 3.27-fold increased risk of developing non-insulin-dependent diabetes mellitus during the follow-up period compared to individuals who had never smoked [27].

G. HbA1c Level

The HbA1c level is commonly used to regularly monitor the long-term glycemic status in patients with both type I and type II diabetes [28]. A glycated hemoglobin (HbA1c) range of 5.7–6.4% has been included as a category of increased risk for diabetes [29]. On top of that, HbA1c, which represents the average plasma glucose levels in the 2 to 3 months before the test, has been seen as a biomarker for the presence and severity of high blood sugar, indicating diabetes or pre-diabetes [30]. It is also seen as a marker for the risk factors of complications related to diabetes during treatment and management [31], [32]. Besides, previous studies show that HbA1c testing can assist in predicting the probability of patients developing diabetes in the future. Individuals with normal HbA1c levels have a low occurrence of diabetes and may not need to be reevaluated in a period of 3 years [33].

H. Blood Glucose Level

Chronic diabetes mellitus (DM) is brought on by a disruption in the metabolism of glucose, which results in abnormal blood glucose levels (BGLs) that can deviate from the normal range either higher (hyperglycemia) or lower (hypoglycemia): 3.9 mmol/L to 7.8 mmol/L, or 70 mg/dl to 140 mg/dl [34]. The risk of diabetes in overweight persons was significant when the blood sugar was above 140 mg/100 ml [35]. One of the main indicators of a later decline toward diabetes was the baseline blood glucose level. Furthermore, regardless of blood glucose levels, individuals exhibiting a reduced insulin response to glucose also demonstrated a higher incidence of disease progression [36]. Another study mentioned that tight blood glucose control is necessary to avoid complications from diabetes. Ideally, the level should be maintained between 4 and 8 mmol/l. The most common method of treating type I diabetes involves having the patient take multiple daily blood glucose readings, from which the ideal amount of insulin can be calculated [37].

III. Discussion And Analysis

In this report of the study, the diabetes prediction dataset was gathered from Kaggle presented by [38] with a total of 92,542 patients measured by age, gender, body mass index (BMI), hypertension (blood pressure), heart disease, smoking history, HbA1c level, and blood glucose level. The dataset gathered will be analyzed and will be presented in data visualization of bar chart, boxplot, and heatmap to see the variables' coloration using the Python 3 programming language default.

A. Total Patients by Gender

The visualization of the bar chart below shows both information of total gender on the dataset male – red bar (54,074) and female – purple bar (38,468). Besides the chart also shows the classification based on the dataset on how many patients by gender per history of Diabetes, it is indicated with the presence of diabetes (1)-green bar and without indication

 EISSN: 2961-872X ● PISSN: 2961-8924 ● DOI: 10.56741/esl.v3i01.472

of diabetes marked with (0)-blue bar. It shows that 2,766 female patients indicated the presence of diabetes while the others did not have it. On the other hand, 2978 male patients have diabetes.

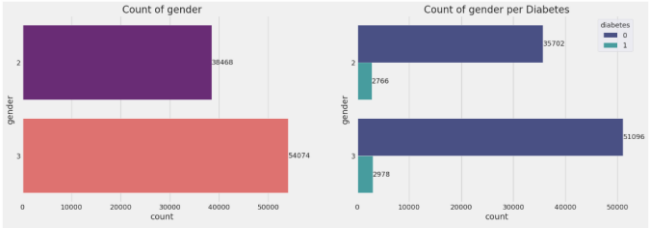

Fig. 1. Total Patients by Gender

B. Total Patients by Hypertension

From this chart below of patients by the hypertension rate, from all of the patients, 85,876 patients do not have hypertension (0 - purple bar) while 6,666 (1 – red bar) other patient has hypertension indicating the risk of getting diabetes. Furthermore, out of the 85,876 patients who do not have hypertension there are 4,344 of them who have diabetes, and 1,400 patients who have diabetes among the 6,666 patients that have hypertension.

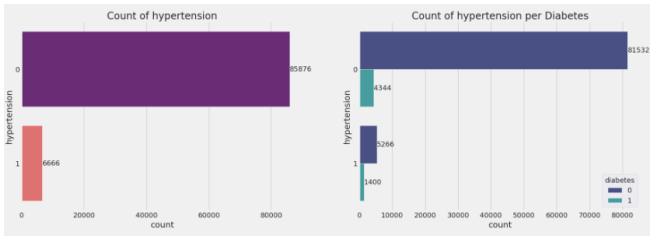

Fig. 2. Total Patients by Hypertension

C. Total Patients by Heart Disease

This figure below counts the total of the patient that has heart disease. 89,040 patients do not have Heart Disease while the other 3,503 patients have it. Interestingly, from 89,040 patients that do not have Heart Disease 4,875 among them have diabetes. On the other hand, 869 patients out of 3,502 who have Heart Disease also have Diabetes showing the colleration of 24.8% of patients with these diseases in the patient's data set.

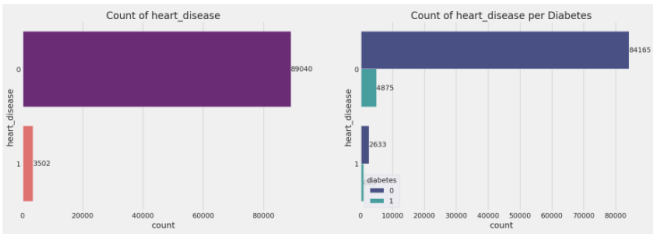

Fig. 3. Total Patients by Heart Disease

D. Total Patients by Smoking History

This bar chart counts the Smoking History based on the total number of patients. In this data, there are 6 cases of smoking information (2 – never smoke) with total patients of 32,975, (3 – no information) with total patients of 32,218, (4 – currently smoke) with total patients of 8,770, (5 – former) with total patients of 8,701, (6 – ever smoke) with total patients of 6,066, and (7 – not currently smoking) with total patients of 3,812. Compared with the data of smoking history per diabetes chart, it shows that patients who never smoke and do not have diabetes are at the amount of 30,704. While on the patients who were former smokers, 1,103 had diabetes. This indicates that smoking does influence the risk of getting diabetes.

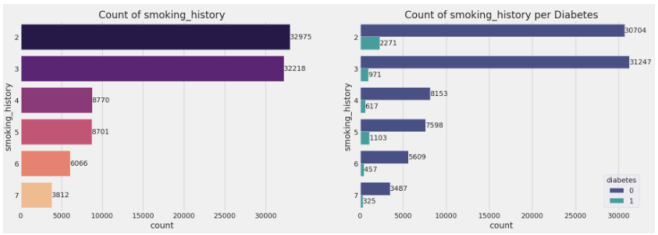

Fig. 4. Total Patients by Smoking History

E. Boxplot of Patients' age

This is the distribution of the age with the diabetes presence from the patients' dataset that ranges from 0-80. In this visualization, 0 indicates no presence of diabetes and 1 is an indication of diabetes. The range of patients that do not have diabetes starts at the age of 20 - 55 as the first quartile and third quartile and the median of it is at the age of 40. On the other hand, for the patients that have the presence of diabetes, the first quartile starts at the age of 52 and the third quartile of 72. The median of patients age that has diabetes is 62 years old.

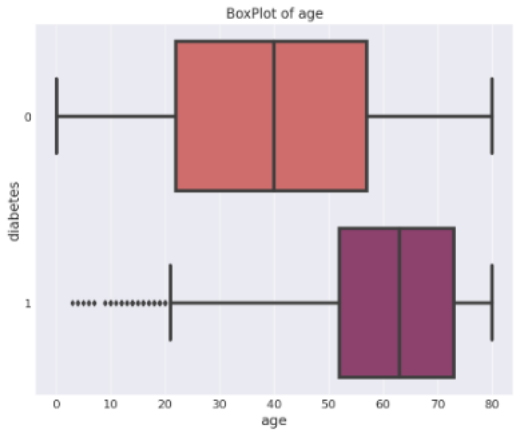

Fig. 5. Boxplot by Age

F. Boxplot of Patients'Body Mass Index (BMI)

In this figure boxplot below, is the distribution of the Body Mass Index. For patients with a body mass index below 18.5, they are classified as being in the underweight range. Those with a BMI between 18.5 and 24.9 are considered to have a healthy weight. Individuals with a BMI between 25 and 29.9 are placed in the overweight category, while those with a BMI of 30 or higher are categorized as being in the obese range. It shows that the median of patients who do not have diabetes is in the range of 23 – 30 for the BMI, and the median is 27. In contrast, the patients who have diabetes the median is 30 and the end of the quartile is reached 35 which indicates an obese range.

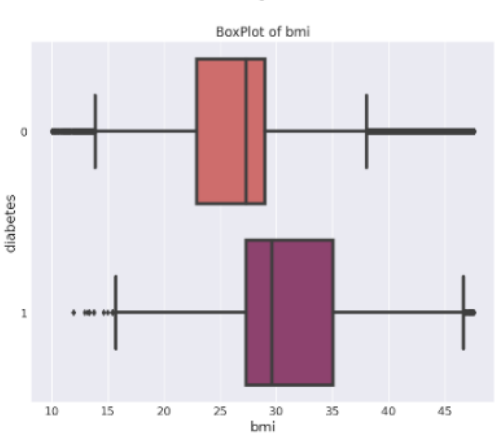

Fig. 6. Boxplot by BMI

G. Boxplot of Patients'HbAc1 Level

The figure below shows the boxplot of HbA1c level which is one of the indicators of diabetes prediction from the patients' dataset. To be classified as normal or within the non-diabetic range in an HbA1c test, the result should be below 5.7%. Individuals with HbA1c values ranging from 5.7% to 6.4% are categorized as prediabetic, whereas a diagnosis of diabetes is established when the HbA1c level is 6.5% or higher. In the data below, the first quartile of the patients who do not have diabetes starts at 4,7%, the median is 5,6%, and the third quartile is 6,2%. On diabetic patients, it shows that the first quartile starts at a rate of 6,4%, a median of 6,5%, and the third quartile for this is 7%. Therefore, patients that have diabetes have an HbA1c level above 6,6% based on the dataset.

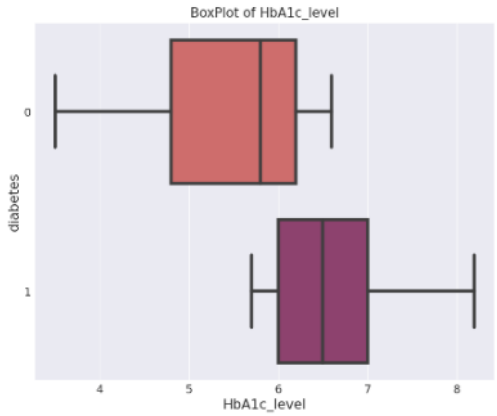

Fig. 7. Boxplot by HbAc1 Level

H. Boxplot of Patient's Blood Glucose Level

As the last variable of the indicator, this figure boxplot below shows the distribution number of the blood glucose level from the dataset of the patients. A normal fasting blood sugar level is below 100 mg/dL (5.6 mmol/L). Fasting blood sugar levels ranging from 100 to 125 mg/dL (5.6 to 6.9 mmol/L) are indicative of prediabetes. If the fasting blood sugar level is 126 mg/dL (7 mmol/L) or higher in two consecutive tests, it indicates diabetes. For the patients that do not have diabetes the blood glucose level starts in the first quartile of 100 mg/dL, the median of 130 mg/dL, and the third quartile is 155 mg/dL. On the other hand, most of the patient that has diabetes has the distribution of blood glucose level up to 225 mg/dL. The first quartile is 130 mg/dL, the median is 155 mg/dL, and the third quartile ends at 220 mg/dL.

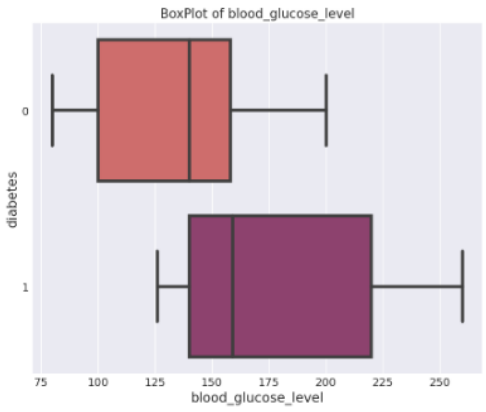

Fig. 8. Boxplot by Blood Glucose Level

I. Heatmap of All Variables

The data visualization below shows the heat map on selected medical variables such as age, hypertension, heart disease, BMI, HbA1c level, blood glucose level, and the outcome which is diabetes based on the color's saturation to indicate the strength of the relationship or the magnitude of the values.

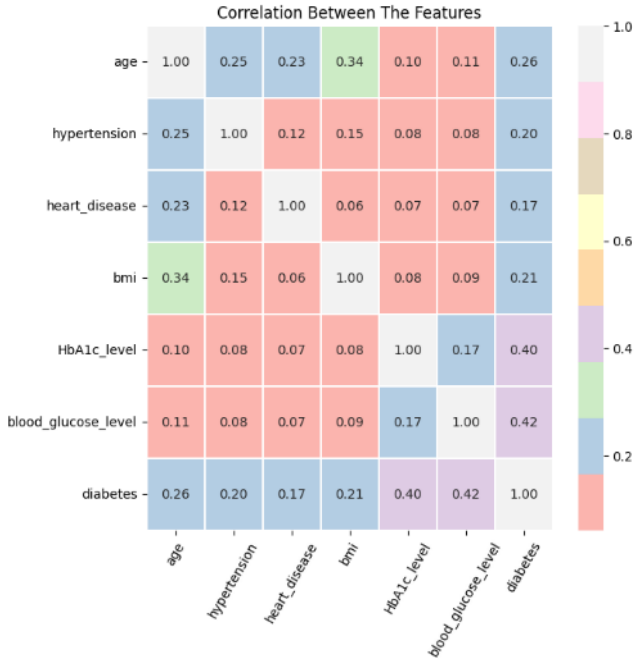

Fig. 9. Heatmap of Variables

Based on the heat map correlation, the rank of the most correlated variables to diabetes is blood glucose level (0.42) followed by HbA1c level (0.40), age (0.26), BMI (0.21), hypertension (0.20), and the least correlated is heart disease (0.17).

IV. CONCLUSION

The medical variables used in this study report start from age, gender, body mass index (BMI), hypertension (blood pressure), heart disease, smoking history, HbA1c level, and blood glucose level to determine the likelihood of getting diabetes. The prevalence of diabetes increased with age and body mass index and increased inversely with energy expenditure in both genders males and females. Data visualization is provided to learn the risk factors on the listed patient's data from having diabetes or not and the factors evaluated as well as to learn the patient's medical history such

 EISSN: 2961-872X ● PISSN: 2961-8924 ● DOI: 10.56741/esl.v3i01.472

as heart disease or smoking intake that is considered as the risk factors of diabetes. Based on this report, the most correlated risk factors for diabetes are blood glucose level and HbAC1 level; therefore, certain dietary, and physical exercises and regular medical check-ups are highly recommended as the risk of diabetes in overweight persons is significant when the blood sugar is above 140 mg/100 ml and HbA1c testing can assist in predicting the probability of patients developing diabetes in the future.

REFERENCES

[1] G. Roglic, "WHO Global report on diabetes: A summary," *Int J Noncommun Dis*, vol. 1, no. 1, p. 3, 2016, doi: 10.4103/2468-8827.184853.

[2] World Health Organization, "Diabetes." Accessed: Dec. 16, 2023. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/diabetes

[3] S. H. Wild and C. D. Byrne, "Risk factors for diabetes and coronary heart disease," *BMJ*, vol. 333, no. 7576, pp. 1009–1011, Nov. 2006, doi: 10.1136/bmj.39024.568738.43.

[4] B. C. K. Choi and F. Shi, "Risk factors for diabetes mellitus by age and sex: results of the National Population Health Survey," *Diabetologia*, vol. 44, no. 10, pp. 1221–1231, Oct. 2001, doi: 10.1007/s001250100648.

[5] T. Endris, A. Worede, and D. Asmelash, "Prevalence of Diabetes Mellitus, Prediabetes and Its Associated Factors in Dessie Town, Northeast Ethiopia: A Community-Based Study," *Diabetes Metab Syndr Obes*, vol. Volume 12, pp. 2799–2809, Dec. 2019, doi: 10.2147/DMSO.S225854.

[6] Z. Yan, M. Cai, X. Han, Q. Chen, and H. Lu, "The Interaction Between Age and Risk Factors for Diabetes and Prediabetes: A Community-Based Cross-Sectional Study," *Diabetes, Metabolic Syndrome and Obesity*, vol. Volume 16, pp. 85–93, Jan. 2023, doi: 10.2147/DMSO.S390857.

[7] K. Junker *et al.*, "The prevalence and correlates of pre-diabetes in middle- to older-aged Irish adults using three diagnostic methods," *PLoS One*, vol. 16, no. 6, p. e0253537, Jun. 2021, doi: 10.1371/journal.pone.0253537.

[8] M. Xia, K. Liu, J. Feng, Z. Zheng, and X. Xie, "Prevalence and Risk Factors of Type 2 Diabetes and Prediabetes Among 53,288 Middle-Aged and Elderly Adults in China: A Cross-Sectional Study," *Diabetes Metab Syndr Obes*, vol. Volume 14, pp. 1975–1985, May 2021, doi: 10.2147/DMSO.S305919.

[9] R. Amir, S. Suhl, and C. M. Alexander, "Renal Evaluation and Protection," *Clin Geriatr Med*, vol. 36, no. 3, pp. 431–445, Aug. 2020, doi: 10.1016/j.cger.2020.04.004.

[10] H. K. Fong *et al.*, "Sex disparities in cardiovascular disease outcomes among geriatric patients with prediabetes," *Prim Care Diabetes*, vol. 15, no. 1, pp. 95–100, Feb. 2021, doi: 10.1016/j.pcd.2020.06.005.

[11] N. G. Forouhi and N. J. Wareham, "Epidemiology of diabetes," *Medicine*, vol. 47, no. 1, pp. 22–27, Jan. 2019, doi: 10.1016/j.mpmed.2018.10.004.

[12] A. Sinclair, T. Dunning, and L. Rodriguez-Mañas, "Diabetes in older people: new insights and remaining challenges," *Lancet Diabetes Endocrinol*, vol. 3, no. 4, pp. 275–285, Apr. 2015, doi: 10.1016/S2213-8587(14)70176-7.

[13] S. B. Al-Mukhtar, N. N. Fadhil, and B. E. Hanna, "General and gender characteristics of type 2 diabetes mellitus among the younger and older age groups.," *Oman Med J*, vol. 27, no. 5, pp. 375–82, Sep. 2012, doi: 10.5001/omj.2012.94.

[14] Y.-D. Jiang, C.-H. Chang, T.-Y. Tai, J.-F. Chen, and L.-M. Chuang, "Incidence and prevalence rates of diabetes mellitus in Taiwan: Analysis of the 2000–2009 Nationwide Health Insurance database," *Journal of the Formosan Medical Association*, vol. 111, no. 11, pp. 599–604, Nov. 2012, doi: 10.1016/j.jfma.2012.09.014.

[15] A. Kautzky-Willer, M. Leutner, and J. Harreiter, "Sex differences in type 2 diabetes," *Diabetologia*, vol. 66, no. 6, pp. 986–1002, Jun. 2023, doi: 10.1007/s00125-023-05891-x.

[16] T. Ciarambino, P. Crispino, G. Leto, E. Mastrolorenzo, O. Para, and M. Giordano, "Influence of Gender in Diabetes Mellitus and Its Complication," *Int J Mol Sci*, vol. 23, no. 16, p. 8850, Aug. 2022, doi: 10.3390/ijms23168850.

[17] Centers for Disease Control and Prevention, "National Diabetes Statistics Report." Accessed: Dec. 17, 2023. [Online]. Available: https://www.cdc.gov/diabetes/data/statistics-report/index.html

[18] World Health Organization, "Obesity." Accessed: Dec. 17, 2023. [Online]. Available: https://www.who.int/health-topics/obesity#tab=tab_1

[19] S. Klein, A. Gastaldelli, H. Yki-Järvinen, and P. E. Scherer, "Why does obesity cause diabetes?," *Cell Metab*, vol. 34, no. 1, pp. 11–20, Jan. 2022, doi: 10.1016/j.cmet.2021.12.012.

[20] J. M. Chan, E. B. Rimm, G. A. Colditz, M. J. Stampfer, and W. C. Willett, "Obesity, Fat Distribution, and Weight Gain as Risk Factors for Clinical Diabetes in Men," *Diabetes Care*, vol. 17, no. 9, pp. 961–969, Sep. 1994, doi: 10.2337/diacare.17.9.961.

[21] G. A. Colditz, "Weight Gain as a Risk Factor for Clinical Diabetes Mellitus in Women," *Ann Intern Med*, vol. 122, no. 7, p. 481, Apr. 1995, doi: 10.7326/0003-4819-122-7-199504010-00001.

[22] E. Grossman and F. H. Messerli, "Hypertension and Diabetes," in *Cardiovascular Diabetology: Clinical, Metabolic and Inflammatory Facets*, Basel: KARGER, 2008, pp. 82–106. doi: 10.1159/000115189.

[23] J. R. Petrie, T. J. Guzik, and R. M. Touyz, "Diabetes, Hypertension, and Cardiovascular Disease: Clinical Insights and Vascular Mechanisms," *Canadian Journal of Cardiology*, vol. 34, no. 5, pp. 575–584, May 2018, doi: 10.1016/j.cjca.2017.12.005.

[24] M. Guglin, K. Lynch, and J. Krischer, "Heart Failure as a Risk Factor for Diabetes Mellitus," *Cardiology*, vol. 129, no. 2, pp. 84–92, 2014, doi: 10.1159/000363282.

[25] G. M. Rosano, C. Vitale, and P. Seferovic, "Heart Failure in Patients with Diabetes Mellitus," *Card Fail Rev*, vol. 03, no. 01, p. 52, 2017, doi: 10.15420/cfr.2016:20:2.

[26] N. H. Cho, J. C. N. Chan, H. C. Jang, S. Lim, H. L. Kim, and S. H. Choi, "Cigarette smoking is an independent risk factor for type 2 diabetes: a four-year community-based prospective study," *Clin Endocrinol (Oxf)*, vol. 71, no. 5, pp. 679–685, Nov. 2009, doi: 10.1111/j.1365-2265.2009.03586.x.

[27] D. Campagna *et al.*, "Smoking and diabetes: dangerous liaisons and confusing relationships," *Diabetol Metab Syndr*, vol. 11, no. 1, p. 85, Oct. 2019, doi: 10.1186/s13098-019-0482-2.

[28] C. Weykamp, "HbA1c: A Review of Analytical and Clinical Aspects," *Ann Lab Med*, vol. 33, no. 6, pp. 393–400, Nov. 2013, doi: 10.3343/alm.2013.33.6.393.

[29] J. C. Bae *et al.*, "Optimal range of HbA1c for the prediction of future diabetes: A 4-year longitudinal study," *Diabetes Res Clin Pract*, vol. 93, no. 2, pp. 255–259, Aug. 2011, doi: 10.1016/j.diabres.2011.05.028.

[30] World Health Organization, *Use of Glycated Haemoglobin (HbA1c) in the Diagnosis of Diabetes Mellitus: Abbreviated Report of a WHO Consultation*. 2011.

[31] X. Wang *et al.*, "Diabetes knowledge predicts HbA1c levels of people with type 2 diabetes mellitus in rural China: a ten-month follow-up study," *Sci Rep*, vol. 13, no. 1, p. 18248, Oct. 2023, doi: 10.1038/s41598-023-45312-y.

[32] S. Lee, T. Liu, J. Zhou, Q. Zhang, W. T. Wong, and G. Tse, "Predictions of diabetes complications and mortality using hba1c variability: a 10-year observational cohort study," *Acta Diabetol*, vol. 58, no. 2, pp. 171–180, Feb. 2021, doi: 10.1007/s00592-020-01605-6.

[33] D. Edelman, M. K. Olsen, T. K. Dudley, A. C. Harris, and E. Z. Oddone, "Utility of hemoglobin A1c in predicting diabetes risk," *J Gen Intern Med*, vol. 19, no. 12, pp. 1175–1180, Dec. 2004, doi: 10.1111/j.1525-1497.2004.40178.x.

[34] M. Z. Wadghiri, A. Idri, T. El Idrissi, and H. Hakkoum, "Ensemble blood glucose prediction in diabetes mellitus: A review," *Comput Biol Med*, vol. 147, p. 105674, Aug. 2022, doi: 10.1016/j.compbiomed.2022.105674.

[35] J. B. O'Sullivan, "Blood Sugar Levels, Glycosuria, and Body Weight Related to Development of Diabetes Mellitus," *JAMA*, vol. 194, no. 6, p. 587, Nov. 1965, doi: 10.1001/jama.1965.03090190009002.

[36] T. Kadowaki *et al.*, "Risk factors for worsening to diabetes in subjects with impaired glucose tolerance," *Diabetologia*, vol. 26, no. 1, Jan. 1984, doi: 10.1007/BF00252262.

[37] P. Kok, "Predicting blood glucose levels of diabetics using artificial neural networks," 2005. [Online]. Available: http://www.vexel.nl/igc

[38] M. Mustafa, "Diabetes prediction dataset." Accessed: Dec. 18, 2023. [Online]. Available: https://www.kaggle.com/search?q=Diabetes+prediction+dataset