Cihan University-Erbil Scientific Journal (CUESJ)

RESEARCH ARTICLE

Apply Binary Logistic Regression Model to Recognize the Risk Factors of Diabetes through Measuring Glycated Hemoglobin Levels

Nozad H. Mahmood[1], Rebaz O. Yahya[2], Saya J. Aziz[1]

[1]Department of Business Administration, Cihan University Sulaimaniya, Kurdistan Region, Iraq, [2]Department of Business Administration, Cihan University-Erbil, Kurdistan Region, Iraq

ABSTRACT

This study aimed to identify the diabetes risk factors in the Kurdistan Region of Iraq and explain why diabetes is rapidly spreading there, which examined some sociodemographic characteristics factors that might affect type 2 diabetes such as age, gender, alcohol consumption, and smoking, diabetes family history, and body weight. The data were collected from the hospital of diabetes named the center of diabetes in Sulaymaniyah city of Iraq, in which 218 diabetic cases were used for that purpose. According to the findings, some factors influence type 2 diabetes, such as gender, smoking, and body weight. For gender, females are more likely to have diabetes than males. Furthermore, someone that smokes is more likely to have diabetes than those who do not smoke. Furthermore, with increasing each kilogram of body weight, the diabetes degree increases as well. On the other hand, regarding the results, some factors such as age, consumption of alcoholic, and diabetes family history do not affect type 2 diabetes. Depending on the findings, it is recommended that people engage in regular physical activity and consume nutritious foods to minimize weight gain, which is one of the primary causes of diabetes as well as they should quit smoking.

Keywords: Binary logistic regression, diabetes, glycated hemoglobin, HbA1c, risk factors of diabetes

INTRODUCTION

Diabetes disease is one of the most widespread infections worldwide, affecting millions of people. Besides that, disease rates are increasing quickly, and it is a genuine global health emergency, imposing crucial trouble on many countries.[1] According to the WHO reports, it predicts that the global diabetes epidemic will nearly double by 2030. One American is killed by diabetes every 3 min, making it the leading cause of blindness, kidney disease, and even death in the United States today. Regarding the International Diabetes Federation reports, there will be 700 million diabetics worldwide by 2045.[2] There has been an epidemic of diabetes mellitus in both developed and developing countries.[3]

There are two primary varieties of diabetes. With having type 1, our body will stop making any insulin, a hormone that allows the body to utilize glucose from food sources.

Type 2 diabetes is a severe disease where the body cannot properly utilize food as an energy source. Our bodies use glucose, or sugar, to make energy from most of the food we consume. The pancreas produces a hormone that helps glucose enter cells which is insulin. There is a lack of insulin production and an inability to utilize insulin properly in type 2 diabetes. Thus, insulin sensitivity resistance and pancreatic cell failure are two characteristics of type 2 diabetes, leading to many health issues and death.[4] In general, type 2 diabetes accounts for about 95% of all diabetic cases. There has been a rise in this kind of diabetes incidence among younger people due to obesity and a family history of diabetes.[5]

Increased morbidity and mortality might be expected in persons with diabetes mellitus who do not control their blood sugar.[6] According to recent estimates, 50% of type 2 diabetes mellitus (T2DM) requires insulin in the first 6 years following diagnosis. Insulin therapy is often necessary for many people with T2DM to achieve their desired blood sugar levels.[7] As a result of insulin's ability to keep blood sugar levels in a healthy

Corresponding Author:
Nozad H. Mahmood, Department of Business Administration, Cihan University Sulaimaniya, Kurdistan Region, Iraq.
E-mail: Nozad.mahmood@sulicihan.edu.krd

Received: November 29, 2021
Accepted: December 25, 2021
Published: January 30, 2022

DOI: 10.24086/cuesj.v6n1y2022.pp7-11

range, it prevents a wide range of health issues that might arise from hyperglycemia.

The problem statement of conducting this study was that many people in Iraq's Kurdistan Region were diagnosed with type 2 diabetes at various ages, and the main research questions that led this study are as follows:

1. Why is diabetes rapidly increasing in the Kurdistan Region of Iraq?
2. Is there any relationship between type 2 diabetes and the factors of gender, age, consumption of alcoholic, smoking, diabetes family history, and body weight?

The Objectives of the Study

The essential goals of this research are to determine the factors that affect diabetes disease in the Kurdistan Region of Iraq and explain the reasons for its rapidly increasing prevalence in that region.

The Significance of the Study

The significance of this research is for identifying all the factors that lead to diabetes in Kurdistan so that people can be aware of the causes of the disease and better control diabetes in Kurdistan.

The Hypothesis of the Study

1- There is no statistically significant difference between type 2 diabetes and gender (male or female).
2- There is no statistically significant difference between type 2 diabetes and the aging of people.
3- There is no statistically significant difference in type 2 diabetes based on the consumption of alcohol.
4- There is no statistically significant difference between type 2 diabetes and smoking.
5- There is no statistically significant difference in type 2 diabetes and diabetes family history.
6- There is no statistically significant difference in type 2 diabetes and body weight.

LITERATURE REVIEW

Glycemic control was more likely to be compromised in patients with younger age, female gender, longer disease duration, high glycated hemoglobin (HbA1c) levels, obesity, and lack hypertension.[8] Another study showed HbA1c levels affected by sociodemographic factors in such a way that the younger and less educated participants recorded higher HbA1c levels than their older and more informed counterparts.[9]

The previous study revealed a significant association between smoking habits and diabetes mellitus.[10,11] The available evidence shows that the smoking habit increases insulin resistance, and it has further been associated with the risk of chronic pancreatitis and pancreatic cancer.[12]

Body mass index and waist circumference were the other variables that showed significant associations in the multivariable analysis. Overweight people were 4.1 times more likely to have diabetes than ordinary people. This finding is consistent with research from India.[13] Similarly, participants with a high waist circumference were more likely to be diabetic than their counterparts, with a prevalence of 11.2% diabetes.

It has been proposed that increased abdominal fat stores interfere with insulin action by releasing free fatty acids. Furthermore, fat cells secrete signaling factors such as interleukin and tumor necrosis factor, both of which influence the development of diabetes.[14] Furthermore, most countries' life expectancy has increased, with the global life expectancy at birth increasing from 65.6 years of age in 1990 to 73 years in 2017.[15] When these global aging trends are combined with the global diabetes epidemic and recent advances in diabetes patient survival, it is no surprise that older age is now recognized as an increasingly crucial diabetes risk factor.[15-18] Aging raises the risk of diabetes by impairing insulin secretion and increasing insulin resistance through obesity and sarcopenia.

Moreover, older adults have higher rates of insufficient physical activity, increasing sedentary time, and decreasing physical activity.[19] At present, older adults (65 years and older) have the highest T2DM prevalence of any age group. Diabetes cases in this age group globally are expected to increase from 122.8 million in 2017 to 253.4 million by 2045.[16]

T2DM family history has been recognized as an essential non-modifiable diabetic risk factor, providing a conveniently assessed marker of the fundamental diabetic genetic component.[7] Diabetic family history is a predictor of type 2 diabetes in both males and females.[20,21] Notably, some research found a high correlation between diabetic family background and T2DM risk. In addition, studies from a large Swedish cohort of 60-year-old men and women demonstrate that a family history of diabetes in combination with being overweight appears to be highly risky in men, raising the chance of acquiring T2DM synergistically.[21]

Methodology

The data were collected from 218 patients that recorded their information in the hospital of diabetes in Sulaymaniyah city of Iraq, which included sociodemographic variables such as age, gender, weight, and clinical characteristics such as diabetes family histories. Similarly, some information obtains on the patients' smoking status and alcohol consumption. The logistic regression model was used to identify risk variables that affect type 2 diabetes. Besides, testing the odds ratios, likelihood ratio, and 95% confidence intervals estimate for each factor using JMP Pro, a statistical tool, version 16, was used to analyze the data.

HbA1c

Diabetes mellitus is a major epidemic affecting both developing and developed countries. The American Diabetes Association suggests using HbA1c instead of fasting blood glucose to diagnose diabetes. HbA1c is a long-term glycemic management measure that can reflect 2–3 months of glycemic history. For healthy people, HbA1c levels should be <6.5%, and studies recognized that diabetic problems are controllable if the HbA1c level is below 6.5%, and above 6.5% will be out of control.[22]

The relationship between plasma glucose and HbA1c can be shown by the following formula to make sure whether the diabetes case is under control or not. Then assign value 1 to diabetic cases that are controllable (HbA1c < 6.5%) and value 0 to diabetic cases that are uncontrollable (HbA1c ≥ 6.5%). In other words, the dependent variable (HbA1c) has 1 and 0 values, that is, why a binary logistic regression model is used.

$$\text{HbA1c} = \frac{46.7 + Plasma\,Glu\cos}{28.7} \quad (1)$$

Thus, the variables of the model represent as below:

RESULTS AND DATA ANALYSIS

The report of Table 2 provides the results of each iteration and criteria used to determine whether or not the model has converged. Throughout this case, the model is a convergent gradient that was iterated 5 times.

Table 3 illustrates to choose the best model. According to the results shown in the table, since $P < 0.001$, the model is significant of the Chi-squared test. In other words, the full model is better than the reduced one (model with only intercept parameter) and indicates that at least one of the parameters affects the response variable. On the other hand, A minimum Akaike's information criterion (AIC) is used to select between two models in that way, the model for which AIC is smallest represents the "best" with this formula.

$$AIC = -2\log likelihood + 2k + 2k\left(\frac{k+1}{n-k+1}\right) \quad (2)$$

Where k is equal to the number of parameters in the model and n is a sample size

Thus,

AIC=238.7672 for full model

AIC=256.5974 for null model

Therefore, the full model is the best one.

Furthermore, Table 3 shows the R-square, which is the ratio of the negative log-likelihood values for the difference and reduced models. Its ranges from 0 for no improvement to 1 for a perfect fit. A nominal model rarely has a high R-square, and it has an R-square of 1 only when all the probabilities of the events that occur are 1. In our case, R-square is 11.92%.

It is now time to determine which factors affect type 2 diabetes and which do not. According to the findings in Table 4, age, alcohol, and family are statistically insignificant. In other words, these three variables have no significant impact on diabetes. Then, we can try rerunning the regression model without any of those variables to see how that affects prediction accuracy. The regression equation is as follows:

Log (odds of a control case)=−4.3217463 +1.163 (gender) + 1.053 (smoking) + 0.066 (weight)

As with any regression, the positive coefficients indicate a positive relationship with the dependent variable. 1.163 is the increment to log odds of a worse outcome for females; the odds ratio $e^{1.163}$=3.20 indicates that females are 3.20 times more likely to achieve a high level of diabetes (out of control) than males. For the smoking case, the odds ratio $e^{1.053}$=2.87 indicates that a person with smoking is 2.886 times more likely to achieve a high level of diabetes (out of control) than one with no smoking. Furthermore, with increasing 1 kg of weight body, the diabetes degree increases by $e^{0.066}$=1.068.

Table 5 shows 95% confidence intervals for parameters (beta) and odd ratios, and we can note that confidence intervals of age, alcohol consumption, and diabetes family history have 0 between the lower and upper levels. It indicates that the coefficient of these variables might become 0, which is why age, alcohol, and family are not significant, as we explained before.

When it comes to odds ratio, it is clear that

log (odds ratio)=beta, and odds ratio=EXP(beta)

Thus, regarding results in Table 5, the 95% confidence interval of the odds ratio of gender, smoking, and weight (1.582827, 6.622663), (1.370512, 6.347617), and (1.030501,

Table 1: The variables of the model

Variables	Types of variables	Codes
HbA1c	Outcome (Response)	HbA1c≥6.5%=0, HbA1c<6.5%=1
Gender	Independent	Male=0, female=1
Age	Independent	Continuous
Alcohol consumption	Independent	Alcoholic=1, no alcoholic=0
Smoking status	Independent	Smoker=1, no smoker=0
Diabetes family histories	Independent	Yes=1, No=0
Body weight	Independent	Continuous

Table 2: Iterations of convergent gradient the model

Iteration	Objective	Relative gradient	Norm gradient
0	151.10608536	8.2627136759	4632.4497299
1	113.96878959	1.8226836288	638.86115249
2	112.15745738	0.2820389174	84.319608151
3	112.11696969	0.0077978033	2.2874541839
4	112.11693927	6.1496927e-6	0.0017917235
5	112.11693927	3.859414e-12	1.1109202e-9

Table 3: Whole the model

Model	Log-likelihood	DF	Chi-square	Prob.> Chi-Sq.
Difference	15.17251	6	30.34501	<.0001*
Full	112.11694			
Reduced	127.28945			
R-square (U)	0.1192			
AIC	238.767			
Observations	218			

Table 4: Parameter estimates

Term	Estimate (beta)	Std. error	Chi-square	Prob.>Chi-sq.	Exp (beta)
Intercept	−4.3217463	1.5420851	7.85	0.0051*	−11.7478
Gender	1.16270774	0.3636579	10.22	0.0014*	3.1601
Age	−0.0076878	0.0131161	0.34	0.5578	−.0021
Alcohol	0.37083796	0.7383133	0.25	0.6155	1.0080
Smoking	1.05325849	0.38882	7.34	0.0068*	2.8631
Family	−0.1939148	0.3332303	0.34	0.5606	−0.5271
Weight	0.06587092	0.0188393	12.23	0.0005*	0.1791

Table 5: 95% confidence intervals for parameters and odds ratios

Parameters	95% confidence intervals				
	For parameters (beta's)		For odds ratios		
	Lower 95%	Upper 95%	Odds ratio	Lower 95%	Upper 95%
Gender	0.45921259	1.89049762	3.198583	1.582827	6.622663
Age	−0.0336268	0.01805573	0.992342	0.966932	1.01822
Alcohol	−0.9918519	1.98504948	1.448948	0.370889	7.279408
Smoking	0.31518408	1.84807951	2.866978	1.370512	6.347617
Family	−0.8550631	0.45598815	0.823728	0.425256	1.577732
Weight	0.030045	0.1041753	1.068089	1.030501	1.109795

1.109795), respectively, indicates that these parameters are significant because their CI does not contain 1, but the others (age, alcohol, and family) are not significant because their CI contains 1 and log (1)=0

Table 6 is about Wald test results, which is the same as the likelihood ratio test of Table 4. Here, we can note that the *P*-values of the variables (gender, smoking, and weight) are <0.05, and it tells us that these variables are significant, while the other variables are not significant because their *P*-values are more than 0.05.

Variance is a measure to show how much the data are spreading around the mean point, and it is equal to the square of the standard deviation. Table 7 shows variance and covariance matrix for estimates of parameters, and the outcomes in the table explain that the off-diagonal elements contain the covariances of each pair of variables, and the diagonal ones of the covariance matrix represent the variances of each variable.

According to the results, the covariance between each pair of (gender and alcohol), (gender and smoking), and (gender, body weight) is approximately 0.0539, 0.0312, and 0.0011, respectively, which indicates that the relationships are positive between these pairs of variables. In contrast, the covariance between gender and age is −0.0008 and between gender and family is −0.0071. The negative values indicate that there is some negative relationship between these variables. Likewise, the relationships between pair variables of (age, family) and (age, body weight) are positive. On the other hand, the covariance estimates of the parameters (age, alcohol), (age, smoking), (alcohol, smoking), (alcohol, family), (alcohol, body weight), (smoking, family), (smoking and body weight), and (family, body weight) show some negative relationships between these pair of variables.

Table 6: Effect of Wald test

Source	No. of parameters	DF	Wald Chi-square	Prob.>Chi-sq.
Gender	1	1	10.2224534	0.0014*
Age	1	1	0.34355643	0.5578
Alcohol	1	1	0.25228239	0.6155
Smoking	1	1	7.33791774	0.0068*
Family	1	1	0.33863607	0.5606
Weight	1	1	12.2253152	0.0005*

Table 7: Covariance of estimates of parameters

Parameters	Gender	Age	Alcohol	Smoking	Family	Weight
Gender	0.1322	−0.0008	0.0539	0.0312	−0.0071	0.0011
Age	−0.0008	0.0002	−0.0010	−0.0004	0.0003	0.0000
Alcohol	0.0540	−0.0010	0.5451	−0.0619	−0.0219	−0.0001
Smoking	0.0312	−0.0004	−0.0619	0.1512	−0.0075	−0.0001
Family	−0.0071	0.0003	−0.0219	−0.0075	0.1110	−0.0004
Weight	0.0011	0.0001	−0.0009	−0.0001	−0.0004	0.0002

CONCLUSIONS

The objective in this study is to illustrate the factors that affecting on the type 2 diabetes in Sulaymaniyah city of Iraq, according of the results, the factors of gender, smoking status, and weight body have high effecting on diabetes, in which for gender, the females have more likely to have high level of diabetes than the males. This may be due to the fact that in our society, women are more likely to do housework, but men are more likely to work outside the home, so men are more active compared to women. Furthermore, for smoking case, the person with smoking is almost 3 times more likely to have diabetes than those with no smoking. Furthermore, with each kilogram of body weight gained, the risk of developing diabetes increases. On the other hand, the variables of age, alcohol consumption, and diabetes family history do not have any effect on diabetes.

RECOMMENDATION

1- It is necessary to perform additional type 2 diabetes research in all other cities in the Kurdistan Region of Iraq to identify the fundamental risk factors that significantly influence the development of type 2 diabetes in that area.
2- It is suggested that more encouragement be given to people to engage in some regular physical activities and be warned about consuming unhealthy foods that might lead to weight gain and eventually diabetes.
3- Men in the Kurdish community have always been expected to fulfill greater responsibilities outside the home, while women have traditionally been supposed to stay home more. As a result, to avoid developing diabetes, women must engage in more physical activity and exercise.

REFERENCES

1. S. J. Ghabban, B. Althobaiti, I. M. Farouk, M. Al Hablany, A. Ghabban, R. Alghbban, S. Harbi and A. E. Sr. Albalawi. Diabetic complications and factors affecting glycemic control among patients with Type II diabetes mellitus attending the chronic illness clinics at Tabuk, Saudi Arabia. *Cureus,* vol. 12, no. 11, p. e11683, 2020.
2. International Diabetes Federation. *IDF Diabetes Atlas.* 9th ed. Brussels, Belgium: International Diabetes Federation, 2019. Available from: https://diabetesatlas.org/en [Last accessed on 2021 August 20].
3. P. Zimmet, K. G. Alberti, D. J. Magliano and P. H. Bennett PH. Diabetes mellitus statistics on prevalence and mortality: Facts and fallacies. *Nature Reviews Endocrinology,* vol. 12, no. 10, pp. 616-622, 2016.
4. American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diabetes Care,* vol. 37, no. Suppl 1, pp. S81-S90, 2014.
5. D. M. Yigazu and T. A. Desse. Glycemic control and associated factors among Type 2 diabetic patients at Shanan Gibe Hospital, Southwest Ethiopia. *BMC Research Notes,* vol. 10, no. 1, p. 597, 2017.
6. H. King, R. E. Aubert and W. H. Herman. Global burden of diabetes, 1995-2025: Prevalence, numerical estimates, and projections. *Diabetes Care,* vol. 21, pp. 1414-1431, 1998.
7. Y. Zheng, S. H. Ley and F. B. Hu. Global aetiology and epidemiology of Type 2 diabetes mellitus and its complications. *Nature Reviews Endocrinology,* vol. 14, no. 2, pp. 88-98, 2018.
8. A. Ali Mansour, N. T. Y. Alibrahim, H. A. Alidrisi, A. H. Alhamza, A. M. Almomin, I. A. Zaboon, M. B. Kadhim, R. N. Hussein, H. A. Nwayyir, A. G. Mohammed, D. K. J. Al-Waeli, I. H. Hussein and FDEMC Study Group. Prevalence and correlation of glycemic control achievement in patients with Type 2 diabetes in Iraq: A retrospective analysis of a tertiary care database over a 9-year period. *Diabetes and Metabolic Syndrome: Clinical Research and Reviews,* 14, no. 3, pp. 265-272, 2020.
9. S. Safarini, H. Hilali, S. Elsheikh and A. Omer. Analytical study of 594 Type 2 diabetic patients attending a private diabetes and endocrine center at Dallah Hospital in Riyadh, KSA. *Journal of Diabetes and Metabolism,* vol. 3, p. 234, 2012.
10. D. Kapoor, A. K. Bhardwaj, D. Kumar and S. K. Raina. Prevalence of diabetes mellitus and its risk factors among permanently settled tribal individuals in tribal and urban areas in Northern state of sub-Himalayan region of India. *International Journal of Chronic Diseases,* vol. 2014, p. 380597, 2014.
11. S. A. Chang. Smoking and Type 2 diabetes mellitus. *Diabetes and Metabolism Journal,* vol. 36, no. 6, pp. 399-403, 2012.
12. S. B. Aynalem and A. J. Zeleke. Prevalence of diabetes mellitus and its risk factors among individuals aged 15 years and above in Mizan-Aman town, Southwest Ethiopia, 2016: A cross sectional study. *International Journal of Endocrinology,* vol. 2018, p. 9317987, 2018.
13. J. Ahmad, M. A. Masoodi, M. Ashraf, R. Rashid, R. Ahmad, A. Ahmad and S. Dawood. Prevalence of diabetes mellitus and its associated risk factors in age group of 20 years and above in Kashmir, India. *Al Ameen Journal of Medical Sciences,* vol. 4, no. 1, pp. 38-44, 2011.
14. C. E. Ekpenyong, U. Akpan, J. O. Ibu and D. E. Nyebuk. Gender and age specific prevalence and associated risk factors of Type 2 diabetes mellitus in Uyo metropolis, South Eastern Nigeria. *Diabetologia Croatica,* vol. 41, no. 1, pp. 17-28, 2012.
15. GBD 2017 DALYs and HALE Collaborators. Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990-2017: A systematic analysis for the global burden of disease study 2017. *Lancet,* vol. 392, no. 10159, pp. 1859-1922, 2018.
16. NCD Risk Factor Collaboration. Worldwide trends in diabetes since 1980: A pooled analysis of 751 population-based studies with 4 4 million participants. *Lancet,* vol. 387, pp. 1513-1530, 2016.
17. N. H. Cho, J. E. Shaw, S. Karuranga, Y. Huang, J. D. da Rocha Fernandes, A. W. Ohlrogge and B. Malanda. IDF diabetes atlas: Global estimates of diabetes prevalence for 2017 and projections for 2045. *Diabetes Research and Clinical Practice,* vol. 138, pp. 271-281, 2018.
18. International Diabetes Federation. *IDF Diabetes Atlas.* 8th ed. Brussels, Belgium: International Diabetes Federation, 2017. Available from http://www.diabetesatlas.org [Last accessed on 2021 August 20].
19. S. R. Colberg, R. J. Sigal, J. E. Yardley, M. C. Riddell, D. W. Dunstan, P. C. Dempsey, E. S. Horton, K. Castorino and D. F. Tate. Physical activity/exercise and diabetes: A position statement of the American Diabetes Association. *Diabetes Care,* vol. 39, no. 11, pp. 2065-2079, 2016.
20. F. Ning, Z. Pang, T. Laatikainen, W. Gao, S. Wang, L. Zhang, J. Tuomilehto, Q. Qiao and Qingdao 2006 Diabetes Survey and FINRISK 2002 Study. Joint effect of family history of diabetes with obesity on prevalence of Type 2 diabetes mellitus among Chinese and Finnish men and women. *The Canadian Journal of Diabetes,* vol. 37, no. 2, pp. 65-71, 2013.
21. C. Wikner, B. Gigante, M. L. Hellénius, U. de Faire and K. Leander. The risk of Type 2 diabetes in men is synergistically affected by parental history of diabetes and overweight. *PLoS One,* vol. 8, no. 4, p. e61763, 2013.
22. S. I. Sherwani, H. A. Khan, A. Ekhzaimy, A. Masood and M. K. Sakharkar. Significance of HbA1c test in diagnosis and prognosis of diabetic patients. *Biomarker Insights,* vol. 11, pp. 95-104, 2016.