Sains Malaysiana 53(10)(2024): 3535-3543
http://doi.org/10.17576/jsm-2024-5310-25

Predictive Factors for Diabetes Mellitus: Insights from Complete Blood Count Analysis

(Faktor Ramalan untuk Diabetes Melitus: Suatu Pandangan daripada Analisis Kiraan Darah Lengkap)

AGNES AYANG KENYANG[1,2], NURLIYANA JUHAN[3,*], YONG ZULINA ZUBAIRI[4], NORNAZIRAH AZIZAN[5] & HO CHONG MUN[1]

[1]Faculty of Science and Natural Resources, Universiti Malaysia Sabah, 88400 Kota Kinabalu, Sabah, Malaysia
[2]College of Computing, Informatics and Mathematics, Universiti Teknologi MARA, 94300 Kota Samarahan, Sarawak, Malaysia
[3]Preparatory Centre for Science and Technology, Universiti Malaysia Sabah, 88400 Kota Kinabalu, Sabah, Malaysia
4Institute for Advanced Studies, University of Malaya, 50603 Kuala Lumpur, Malaysia
[5]Department of Pathology and Microbiology, Faculty of Medicine & Health Sciences, Universiti Malaysia Sabah, 88400 Kota Kinabalu, Sabah, Malaysia

Received: 4 April 2024/Accepted: 19 August 2024

ABSTRACT

The 10th edition of the International Diabetes Federation reports that 537 million people worldwide had diabetes in 2021. In Southeast Asia, countries like Malaysia are facing a growing burden of diabetes. This highlights the urgent need for innovative and resourceful approaches to diabetes management. As the prevalence of diabetes continues to rise in these countries, tailored strategies are necessary. To identify and evaluate the potential prognostic indicators for diabetes mellitus, this study involved a dataset consisting of 500 entries, comprising demographic information and selected blood cells from the Complete Blood Count (CBC) test results obtained from the Clinical Laboratory Faculty of Medicine & Health Sciences, Universiti Malaysia Sabah. Using univariate and multivariate logistic regression analysis, the prognostic predictors for diabetes mellitus were identified. In the univariate analysis, all variables are statistically significance at 5% level of significance. However, at multivariate analysis, only age, mean corpuscular hemoglobin concentration (MCHC), white blood cells (WBC) and hematocrit (HCT) emerged as significant predictors of diabetes mellitus. Notably, the abnormal level in WBC exhibited the greatest association with diabetes mellitus, reflecting a 114.7% increased risk compared to a normal WBC level. The statistic value obtained from Hosmer-Lemeshow was 0.944 indicating a well-fitting model. Additionally, the receiver operator characteristic (ROC) curve has a value of 0.7, indicating a strong performance of the model. In conclusion, CBC parameters can be accurate markers and useful in assisting clinical decision-making when properly applied and interpreted.

Keywords: Complete Blood Count (CBC); diabetes mellitus; prognosis; regression analysis

ABSTRAK

Edisi ke-10 International Diabetes Federation melaporkan bahawa 537 juta orang di seluruh dunia menghidap diabetes pada tahun 2021. Di Asia Tenggara, negara seperti Malaysia menghadapi beban diabetes yang semakin meningkat. Kadar peningkatan ini telah mendorong kepada keperluan untuk merangka strategi yang inovatif dan berkebolehan bagi mengawal diabetes. Untuk mengenal pasti dan menilai penunjuk prognostik yang berpotensi untuk diabetes melitus, kajian ini melibatkan set data yang terdiri daripada 500 entri meliputi maklumat demografi dan sel darah terpilih daripada keputusan ujian Kiraan Darah Lengkap (CBC) yang diperoleh daripada Makmal Klinikal Fakulti Perubatan & Sains Kesihatan, Universiti Malaysia Sabah. Analisis regresi logistik di peringkat univariat dan multivariat telah dilaksanakan untuk mengenal pasti pemboleh ubah yang berpotensi untuk menjadi peramal prognostik untuk diabetes melitus. Dalam analisis univariat, semua pemboleh ubah adalah signifikan pada aras keertian 5%. Walau bagaimanapun, pada analisis di peringkat multivariat, hanya umur, purata kepekatan hemoglobin korpuskular (MCHC), sel darah putih (WBC) dan hematokrit (HCT) muncul sebagai peramal penting diabetes melitus. Keputusan kajian menunjukkan tahap abnormal dalam sel darah putih menunjukkan perkaitan yang paling besar dengan diabetes melitus, mencerminkan peningkatan risiko 114.7% berbanding tahap normal sel darah putih. Ukuran kesesuaian dan diskriminasi yang baik diperoleh daripada nilai Hosmer-Lemeshow dan graf Penerima Ciri Operasi (ROC) masing-masing ialah 0.944 dan 0.7. Pemboleh ubah di dalam CBC boleh menjadi penanda aras yang tepat dan berguna bagi membantu membuat keputusan klinikal sekiranya digunakan dan ditafsirkan dengan cara yang betul.

Kata kunci: Analisis regresi; diabetes melitus; Kiraan Darah Lengkap (CBC); prognosis

INTRODUCTION

In recent years, the prevalence of diabetes mellitus (DM) has increased in both developed and advanced countries, making it one of the most significant global public health issues (Ganasegaran et al. 2021; Sun et al. 2021). According to Sun et al. (2021), the anticipated prevalence of diabetes among adults aged 20 to 79 worldwide in 2021 was 10.5%, or 536.6 million people. This number is expected to rise to 12.2% (783.2 million) by 2045 (Sun et al. 2021). The global increase was due to lifestyle factors, including obesity and insulin resistance (Yaribeygi et al. 2021). Diabetes mellitus is a metabolic condition marked by unnecessarily increased blood glucose levels. Type 1 diabetes (T1DM), type 2 diabetes mellitus (T2DM), and gestational diabetes mellitus (GDM) are the three primary subtypes of DM. Diabetes that is not controlled raises the danger of blood, cellular, and metabolic changes that might cause cancer, vascular problems, and other fatalities (Tomic, Shaw & Magliano 2022).

Malaysia has the highest rate of diabetes in the Western Pacific area and one of the highest rates worldwide, with yearly costs of roughly 600 million US dollars (Ganasegaran et al. 2021). Unsurprisingly, as the prevalence of diabetes rises globally, so do healthcare costs to treat it; by 2045, these costs are projected to reach more than $1 trillion USD (Sun et al. 2021). Given this alarming trajectory, early detection of diabetes mellitus (DM) becomes paramount, urging all adults to undergo laboratory routine screening. The blood glucose level (BSL) is essential during the laboratory-based screening for diabetes. The fluctuation of blood glucose levels caused by diet, exercise, medications, and pathogenic processes can be identified as patterns according to blood glucose monitoring. Capillary blood glucose (CBG) testing is the term used to describe BSL monitoring carried out outside of clinic settings, such as at home. However, the reliability of results may vary in patients with hypoglycemia, anemia, altered hematocrit, hypotension, or critically ill patients (Mathew, Muhammad & Prasanna 2023). This means, the reliability of CBG testing is contingent upon accurate and consistent measurements, and healthcare practitioners must be attuned to the factors that may influence these readings.

This has led investigators to find alternative screening techniques. The strategic utilization of available data, transforming it into valuable insights, holds paramount importance in the realm of medicine, especially for enhancing risk prognostication and identifying patients for early intervention and more rigorous follow-up. Complete blood count (CBC) data is the most common and simple laboratory test, provides extensive information about a person's health status. It typically includes eight to ten parameters, such as Haemoglobin (Hb), indices of red blood cells, such as Red Blood Cells (RBCs), Mean Corpuscular Volume (MCV), Mean Corpuscular Haemoglobin (MCH), Mean Corpuscular Haemoglobin Concentration (MCHC), Red cell Distribution Width (RDW), and parameters of white blood cells, such as White Blood Cell Count (WBC) and differential count of neutrophils (Aamir & Safaa 2022; Ahmed et al. 2020; Aytan et al. 2020; Seo & Lee 2022). Several studies have been conducted globally to investigate the potential of CBC parameters as valuable tools for predicting chronic illnesses such as Gestational Diabetes Mellitus (GDM), Type 1 Diabetes Mellitus (T1DM) and Type 2 Diabetes Mellitus (T2DM), COVID-19 diagnosis, cardiovascular and metabolic diseases, and sepsis (Agnello et al. 2021; Ahmed et al. 2020; Seo & Lee 2022; Wynants et al. 2020). The severity of diabetic patients may be lessened by considering all factors in early diabetes mellitus prognosis utilizing the results of the CBC test, with the ultimate goal of enhancing health outcomes.

The logistic regression model is commonly used in medical research (Juhan et al. 2019, 2018; Zabor et al. 2022). The integration of statistical modelling is crucial in predicting the outcome of critical illness and making informed clinical decisions (Pawitra et al. 2022; Wang et al. 2022; Yadav & Akhter 2021). Numerous studies have consistently demonstrated well-fitted when applying multivariate logistic regression to predict patients' mortality risk or recovery (Akyüz et al. 2020; Juhan et al. 2020; Kim et al. 2021; Rahman et al. 2021; Thapa et al. 2022). On the other hand, recent research has paid extensive attention to CBC data. Studies have indicated that CBC parameters can be utilized to predict various conditions such as cancer, cardiovascular diseases, arteriosclerosis, type 2 diabetes (T2DM), and metabolic syndrome (Seo & Lee 2022). Zhao et al. (2023) showed that CBC-derived inflammatory biomarkers were positively correlated with the prevalence of psoriasis by using weighted logistic and Cox regression analyses. Similarly, Hanna et al. (2021) showed that CBC was useful in detecting leukocytes in children with acute gastroenteritis despite being prone to biological, technical and environmental factors. In a study related to cancer, Clara et al. (2022) discovered that 50% use of CBC during pre-treatment sessions for colorectal cancer patients can be considered a valuable prognostic indicator.

Efforts to address diabetes mellitus in Malaysia include prevalence, risk factors, medication, disease management and diabetes literacy (Awang et al. 2022, 2020; Iqbal et al. 2024; Lee et al. 2020; Rawi 2023; Saddki, Hashim & Mohamad 2022; Saidi 2023; Teng, Chan & Wong 2022). To the best of our knowledge, no prior research has been done to identify significant predictors of DM, particularly in Sabah, Malaysia, using the CBC parameters. Therefore, this study aims to develop a prognostic model of diabetes mellitus disease among the community living in Sabah, Malaysia, using a logistic regression to address these gaps in the literature. The dependent variable is the patient's health condition status, expressed in binary 'diabetic' and 'non-diabetic'. When a binary outcome is to be predicted from one or more independent predictive variables, logistic regression is commonly considered (Read et al. 2023; Wang et al. 2022; Zabor et al. 2022). The proposed model will utilize several patient health databases and describe the patterns and factors that may enhance the diagnosis of diabetes mellitus based on the results of CBC tests. These results are crucial for developing diabetes mellitus public health policies and diabetes mellitus screening and management strategies, especially considering the growing usage of CBC data, which are readily available and usable in all nations.

MATERIALS AND METHODS

SOURCE OF DATA

A total of 500 anonymised patient data presenting with CBC test results were obtained from the Clinical Laboratory Faculty of Medicine & Health Sciences, Universiti Malaysia Sabah, for the 5-year period from 2018 to 2022. The dependent variable in this study was the patient's health status, whether they had diabetes mellitus or not. Patient demographics such as gender (male, female), age (less than 60 years old, at least 60 years old), and ethnicity (Malay, non-Malay) and eight CBC parameters, which consist of Haemoglobin (Hb), Mean Corpuscular Volume (MCV), Mean Corpuscular Haemoglobin (MCH), Mean Corpuscular Haemoglobin Concentration (MCHC), Platelet count, Red Blood Cell (RBC), White Blood Cell (WBC) and Haematocrit (HCT) were included as independent factors in this investigation. Each parameter was categorised as normal if the test result fell within the normal range and abnormal if the test result fell beyond the normal range.

ETHICAL CLEARANCE

The study was approved by the Medical Research Ethics Committee, Universiti Malaysia Sabah.

LOGISTIC REGRESSION MODEL

A logistic regression model is frequently utilized in health sciences studies because it is particularly suitable for data involving illness conditions with binary outcome (diseased or healthy) and decision-making (yes or no) (Aydin, Şaylan & Ekiz İşcanlı 2022; Pawitra et al. 2022; Rahman et al. 2021; Yong et al. 2020).

Quantitatively, the relationship between the patient's health status and its dependency on eleven independent factors can be computed using Equation (1) as follows (Zabor et al 2022):

$$P(\text{event}) = \frac{e^{\beta_0+\beta_1 x_1+\beta_2 x_2+...+\beta_n x_n}}{1+e^{\beta_0+\beta_1 x_1+\beta_2 x_2+...+\beta_n x_n}} \quad (1)$$

where P(event) is the probability of the event occurring, in this case the probability of having diabetes mellitus. For each patient in the dataset, one can calculate the probability that the patient has diabetes mellitus from the natural exponent of the sum of the product of each of the covariates and their corresponding parameter estimates as shown in Equation (1).

The general equation of a logistic regression model is given by Equation (2):

$$Y = \text{logit}(P) = \ln\left(\frac{P}{1-P}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_n x_n \quad (2)$$

where Y is the probability of a patient is having diabetes mellitus; are the estimated coefficients from the sample data; is the number of independent variables; and are the independent factors. From Equation (2), the logit function was expressed as the 'odds ratio (OR)' (Sainani 2021). The probability of an event happening divided by the probability that one does not happen is how the odds are calculated. If the OR is greater than 1, the likelihood that an event will occur is greater than the likelihood that it will not, and vice versa when the value of the corresponding independent variable increases by one unit.

Steps to conduct logistic regression: 1). Selecting the dependent variable. In this study, the dependent variable was the health condition status of a patient which was expressed as 'diabetic' or 'non-diabetic'. 2). Selecting the independent variables. There were 11 categorical independent variables used as predictors such as gender, age, ethnicity, Hb, MCV, MCH, MCHC, Platelet, RBC, WBC and HCT. 3). Performing univariate data analysis. This analysis was going to find out the unadjusted link between outcome and independent variables. If the p-value is less than 0.05, it suggests that the variable was statistically significant, implying a meaningful association between the variable and the response and will be considered in the multivariate analysis. 4). Performing multivariate analysis. In the multivariate analysis, a preliminary logistic regression model was fitted. In this step, the interaction terms and multicollinearity between variables were evaluated. If the p-value for the interaction term is less than 0.05, it indicates that the interaction is significant and shall be removed from the preliminary model. The multicollinearity was detected using variation inflation factor (VIF). All the non-collinear variables (with VIF < 2) were included in the logistic regression model (Thapa et al. 2022). 5). Evaluation of the logistic regression model. The goodness of fit for the final model was assessed using the Hosmer and Lemeshow test. This test follows a chi-squared distribution and compared the observed and expected outcomes based on the model's. A larger p-value > 0.05 indicated a fit model (Sainani 2021). On the other hand, the model's performance was assessed by the area under a receiver operator characteristic (ROC) curve (AUC) of the logistic regression model. The higher value of AUC indicates how well the binary outcome can be predicted (Sainani 2021). 6) Calculating the odds ratio (OR). The value of odds ratio can be calculated by exponentiating the coefficient, which is given by.

RESULTS AND DISCUSSION

The demographic information and CBC parameters of 500 patients with and without diabetes were recorded in Table 1. Out of 500 patients, 174 had diabetes (case group), while 326 did not (control group). This study included 315 (63%) male patients and 406 (81.2%) patients under 60 years old. Out of 500 participants, 297 (59.4%) were Malay. Examining the findings of univariate data analysis for all demographic characteristics and CBC parameters statistically related to diabetes mellitus. Thus, all variables were included in the preliminary model.

According to Table 2, all the significant variables from the preliminary model were considered to obtain the final model. The results of the final logistic regression model showed that four variables namely age (OR = 1.833, 95% CI = 1.078, 3.118), MCHC (OR = 1.557, 95% CI = 1.055, 2.296), WBC (OR = 2.147, 95% CI = 1.442, 3.196) and HCT (OR = 0.206, 95% CI = 0.108, 0.393) were highly associated with having diabetes (, Table 2). Based on the values of VIF, all four variables were found to be non-collinear. This is desirable in regression analysis because it allows for a clearer interpretation of the individual effect of each predictor on the response variable. Among four variables, patients with an abnormal level of WBC had the greatest associations with diabetes mellitus because the risk of having diabetes mellitus increased by times compared to patients with a normal WBC level (, Table 2). The risk of having diabetes mellitus is times higher among patients who are 60 years old or older compared to being less than 60 years old (, Table 2). Patients with abnormal MCHC levels were   times more likely to have diabetes mellitus than patients with normal MCHC levels (, Table 2). Having abnormal HCT decreases the odds of having diabetes mellitus by a factor of compared to normal HCT (, Table 2).

According to Table 3, the values of Hosmer-Lemeshow and AUC were 0.944 and 0.7, respectively, indicating a well-fitted and good-discriminating model. The null model has a -2 log-likelihood value of 624.358, while the final model has a -2 log-likelihood value of 600.50. There was a significant decrease in the -2 log-likelihood value for the final model compared to the -2 log-likelihood value for the null model. In that case, this indicated the absolute fit of the final model to the data. Thus, the fitted binary multivariable logistic function is shown in Equation (3).

$$Y = 1.332 + 0.606(\text{Age}) + 0.442\ (\text{MCHC}) + 0.764\ (\text{WBC}) - 1.579\ (\text{HCT}) \tag{3}$$

This study underscores the importance of considering all factors, including the results of a CBC blood test, for an improved prognosis for early diabetes mellitus. Majority of the participants were males, 315 (63%) for both diabetic patients and control groups. There were 174 diabetes patients, and 121 (69.5%) were male. According to our findings, diabetes was more common in males than in women. The same findings were reported in additional nationwide studies conducted by Arkew et al. (2022), Kaul et al. (2022), and Narjis et al. (2021) which showed that males had greater incidences of pre-diabetes and diabetes during both the screening and follow-up periods. Males are less likely than females to follow diabetes screening recommendations. As a result, males have greater rates of prediabetes and diabetes, regardless of age, residence (urban or rural), or material condition (Kaul et al. 2022).

When age groups were compared, our findings showed that the prevalence of diabetes mellitus increased by 1.833 times among study participants 60 years of age or older compared to those under 60 years old. Similar results from a meta-analysis study by Akhtar et al. (2022) showed that the prevalence of diabetes in Malaysia dramatically rises with age, from 3.16% in the 20-29 age group to 33.45% in the 60-year-old and above age group. Another meta-analysis study by Mantovani et al. (2020) among severe COVID-19 patients showed that pre-existing diabetes was significantly higher among patients above the age of 60 than among those under 60 (23.30% vs. 8.79%).

In the current study, we observed that WBC was the most strongly associated with diabetes mellitus. Specifically, we found that for every one-unit increase in WBC levels, the average change in the log odds of having diabetes mellitus increased by 0.764 units. The risk of having diabetes mellitus increased by 2.147 times compared to patients with normal WBC levels. Elevated WBC levels are a well-establishment hallmark of inflammation (Edathodu et al. 2023; Narjis et al. 2021; Rafaqat & Rafaqat 2023). Previous epidemiological studies support this finding and provide evidence of a positive correlation between WBC count and the risk of developing diabetes. A study conducted by Mansoori et al. (2023) among Iranian adults aged 35 to 65 years recruited as part of the Mashhad stroke and heart atherosclerotic disorder (MASHAD) showed adults with a WBC had a four times higher prevalence of diabetes than those without it. Korean individuals living in the community who are not obese had T2DM levels considerably higher in the highest quartile of WBC count groups (Park et al. 2021).

MCHC and HCT are essential parameters for evaluating red blood cells' health. Our analysis showed that patients with abnormal HCT levels have significantly decreased compared to the control group. The beta value for HCT was specifically -1.579, indicating that people with abnormal HCT levels associated with a 79.4% (1 – 0.206 = 0.794) reduction in the odds of having diabetes mellitus. These results are consistent with other previous studies carried out by Mansoori et al. (2023) and Narjis et al. (2021) where patients with diabetes had HCT levels that were incredibly low.

The risk of patients having diabetes mellitus increased by 1.557 times when the level of MCHC was abnormal. This finding is aligned with other previous studies on diabetic group and the control group, where the MCHC percentages in diabetes patients were considerably higher than those in

TABLE 1. Demographic and CBC characteristics of 500 patients

No.	Variables		Patient health status Non-diabetic (n=326) (%)	Diabetic (n=174) (%)	p -value
1.	Gender	Male	194 (61.6%)	121 (38.4%)	.027
		Female	132 (71.4%)	53 (28.6%)	
2.	Ethnicity	Malay	180 (60.6%)	117 (39.4%)	.009
		Non-Malay	146 (71.9%)	57 (28.1%)	
3.	Age	60 years old	71 (75.5%)	23 (24.5%)	.020
		60 years old	255 (62.8%)	151 (37.2%)	
4.	Hb	Abnormal	116 (71.6%)	46 (28.4%)	.037
		Normal	210 (62.1%)	128 (37.9%)	
5.	MCV	Abnormal	232 (62.2%)	141 (37.8%)	.016
		Normal	94 (74.0%)	33 (26.0%)	
6.	MCH	Abnormal	115 (71.9%)	45 (28.1%)	.032
		Normal	211 (62.1%)	129 (37.9%)	
7.	MCHC	Abnormal	163 (60.8%)	105 (39.2%)	.027
		Normal	163 (70.3%)	69 (29.7%)	
8.	Platelet	Abnormal	249 (62.9%)	147 (37.1%)	.033
		Normal	77 (74.0%)	27 (26.0%)	
9.	WBC	Abnormal	161 (60.3%)	106 (39.7%)	.014
		Normal	165 (70.85)	68 (29.2%)	
10.	RBC	Abnormal	156 (60.5%)	102 (39.5%)	.022
		Normal	170 (70.2%)	72 (29.8%)	
11.	HCT	Abnormal	76 (85.4%)	13 (14.6%)	.000
		Normal	250 (60.8%)	161 (39.2%)	

TABLE 2. Variables in the final multivariate regression model

Variables	Characteristics	β	SE	Odds Ratio, $e\beta$ (95% Confidence Interval, CI)	p-value	Variance Inflation Factor (VIF)
Age	≥ 60 years old < 60 years old	0.606	0.271	1.833 (1.078, 3.118)	.025	1.004
MCHC	Abnormal Normal	0.442	0.198	1.557 (1.055, 2.296)	.026	1.002
WBC	Abnormal Normal	0.764	0.203	2.147 (1.442, 3.s196)	.000	1.055
HCT	Abnormal Normal	-1.579	0.329	0.206 (0.108, 0.393)	.000	1.052

TABLE 3. Model checking

Statistics	Value
Hosmer Lemeshow	0.944
AUC	0.7
-2 log-likelihood of the null model	624.358
-2 log-likelihood of the final model	600.050

non-diabetic patients (Aamir & Safaa 2022; Signing et al. 2020). Arkew et al. (2021) and Farooqui, Afsar and Afroze (2019) found that the values of MCHC and HCT levels among diabetic individuals were abnormal when compared to non-diabetic individuals.

Aamir and Safaa (2022), Arkew et al. (2022), Aytan et al. (2020), and Narjis et al. (2021) discovered that there were significant correlations existed between RBC parameters and T2DM patients. In comparison to the control group, T2DM patients had significantly lower mean values for RBC count, HCT, and Hb (Aamir & Safaa 2022; Arkew et al. 2022; Rafaqat & Rafaqat 2023). However, these results differed from our findings for the RBC and Hb levels. A different study by Wang et al. (2021) showed that several variables may affect the association between RBC and Hb levels and diabetes mellitus, including age, sex, race, lifestyle, food, comorbidities, medications, and genetic predisposition.

In evaluating both preliminary and final models, multicollinearity and interactions were examined. The analysis showed no significant interactions and multicollinearity between the independent variables were shown. Model validation and performance are crucial to assess how well the final model fits the data upon which it was generated. Application of modelling techniques without a later evaluation of the resultant model's performance might lead to poorly fitted findings and incorrectly forecast results on new subjects. To ensure the robustness of the results, the final model is assessed using the Hosmer-Lemeshow goodness of fit and the area under the ROC curve. For our results in Table 3, the Hosmer-Lemeshow goodness of fit test computed is 0.944, with a corresponding p-value computed from the chi-square distribution with 7 degrees of freedom is 2.340. The large p-value signifies no significant difference between the observed and predicted outcome values, affirming that the well-fitted of the final model.

The findings of this study offer valuable insights and significant implications for public health by making it possible to identify the prognostic indicator from the CBC parameters that differentiate between diabetes from non-diabetics. However, limitations of the study should be acknowledged including the study utilised only a single referral centre, and the absence of other parameters such as glucose, glycated hemoglobin (HbA1c), and red blood cell distribution width (RDW) readings in the patient's CBC profile. It is worthwhile to note that some of these factors were frequently used to predict diabetic complications and act as inflammatory prognoses and markers in diabetes patients (Aamir & Safaa 2022; Farooqui, Afsar & Afroze 2019). This could potentially impact on the conclusions' validity and dependability, as well as the recommendations' applicability. Therefore, future studies should replicate the study across diverse settings with larger and more varied populations to enhance the model's predictive ability.

CONCLUSION

In conclusion, our findings from the logistic regression analysis suggested that age, WBC, MCHC, and HCT parameters of the CBC are significant predictors of diabetes mellitus. These significant CBC parameters can serve as precise markers and are helpful in guiding clinical decision-making for diabetes mellitus management. The logistic regression modelling presented in this study can be further extended by including other predictors such as HbA1c and glucose levels. Also, future studies should replicate this research across diverse settings with larger and more varied populations to ensure broader applicability.

ACKNOWLEDGMENTS

This research was supported by the Malaysian Ministry of Higher Education through Fundamental Research Grant Scheme (FRGS) (Project Code: FRGS/1/2023/STG06/UMS/02/2).

REFERENCES

Aamir, M.A.A. & Safaa, H.A. 2022. Study the effect of diabetes mellitus (Type 1 and 2) on complete blood count in diabetic patients. *International Journal of Applied Sciences and Technology* 4(4): 112-124.

Agnello, L., Giglio, R.V., Bivona, G., Scazzone, C., Gambino, C.M., Iacona, A., Ciaccio, A. M., Lo Sasso, B. & Ciaccio, M. 2021. The value of a Complete Blood Count (CBC) for sepsis diagnosis and prognosis. *Diagnostics* 11(10): 1881.

Ahmed, M.M., Ghauri, S.K., Javaeed, A., Rafique, N., Hussain, W. & Khan, N. 2020. Trends of utilization of Complete Blood Count parameters for patient management among doctors in Azad Kashmir. *Pakistan Journal of Medical Sciences* 36(5): 999-1004.

Akhtar, S., Nasir, J.A., Ali, A., Asghar, M., Majeed, R. & Sarwar, A. 2022. Prevalence of type-2 diabetes and prediabetes in Malaysia: A systematic review and meta-analysis. *PLoS ONE* 17(1): e0263139.

Akyüz Ünsal, A.İ., Key, Ö., Güler, D., Kurt Omurlu, İ., Anık, A., Demirci, B. & Dündar, S. 2020. Can Complete Blood Count parameters predict retinopathy of prematurity? *Turkish Journal of Ophthalmology* 50(2): 87-93.

Arkew, M., Asmerom, H., Tesfa, T., Tsegaye, S., Gemechu, K., Bete, T. & Haile, K. 2022. Red blood cell parameters and their correlation with glycemic control among type 2 diabetic adult patients in Eastern Ethiopia: A comparative cross-sectional study. *Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy* 15: 3499-3507.

Arkew, M., Yemane, T., Mengistu, Y., Gemechu, K. & Tesfaye, G. 2021. Hematological parameters of type 2 diabetic adult patients at Debre Berhan Referral Hospital, Northeast Ethiopia: A comparative cross-sectional study. *PLoS ONE* 16(6): e0253286.

Awang, H., Muda, R., Rusli, N., Rahman, M. & Embong, K. 2022. Epidemiology of poor glycaemic control among patients with Type 2 diabetes mellitus in Terengganu State of Malaysia. *European Journal of Medical and Health Sciences* 4(5): 89-94.

Awang, H., Ja'afar, S., Ishak, N., Yusof Zainal, M., Aminuddin, A. & Dollah, Z. 2020. Poor glycemic control: Prevalence and risk factors among patients with type 2 diabetes mellitus in Northeast State of Peninsular Malaysia. I*nternational Journal of Human and Health Sciences* 4(3): 206.

Aydin, M., Şaylan, B. & Ekiz İşcanlı, İ.G. 2022. Factors associated with mortality in younger and older (≥75 years) hospitalized patients with community-acquired pneumonia. Annals of Saudi Medicine 42(1): 45-51.

Aytan, P., Bozkurt, B.S., Sakarya, Ö., Çiftçi, R.S. & Aytan, H. 2020. Can a simple Complete Blood Count predict gestational diabetes mellitus? *Journal of Contemporary Medicine* 10(3): 336-341.

Clara, N., Finbarrs-Bello, E., Ugwuene, F., Maduka, L., Odurukwe, O., Aneke, E., & Ugwuishi, E. 2022. The use of complete blood count as inflammatory biomarkers in patients with colorectal cancers in Enugu State University of Science and Technology Teaching Hospital, Parklane Enugu. *Asian Journal of Medicine and Health* 20(12): 88-117.

Edathodu, J., Alsugair, A., Al-Bugami, M., Alomar, I., Alrasheed, A., Fadel, R., Albalawi, W., Alshammary, A., Alsuhaim, A., Alghayti, S., Alkadi, A., Peedikayil, M., Aldakhil, H., Albedah, N. & Mohamed, G. 2023. Predictors of disease severity in patients hospitalized with coronavirus disease 2019. A*nnals of Saudi Medicine* 43(4): 254-261.

Farooqui, R., Afsar, N. & Afroze, I.A. 2019. Role and significance of hematological parameters in diabetes mellitus. *Annals of Pathology and Laboratory Medicine* 6(3): A158-162.

Ganasegeran, K., Hor, C.P., Jamil, M.F.A., Suppiah, P.D., Noor, J.M., Hamid, N.A., Chuan, D.R., Manaf, M.R.A., Ch'ng, A.S.H. & Looi, I. 2021. Mapping the scientific landscape of diabetes research in Malaysia (2000-2018): A systematic scientometrics study. *International Journal of Environmental Research and Public Health* 18(1): 318.

Hanna, W.W., Jan, N., Michał, D., Paula, S., Mariusz, S., Katarzyna, C., Patrycja, K.J. & Jaroslaw, W. 2021. Reliability of capillary complete blood count in children with acute gastroenteritis. *Frontiers in Pediatrics* 9.

Iqbal, M.Z., Alqahtani, S.S., Mubarak, N., Shahid, S., Mohammed, R., Mustafa, A., Khan, A.H. & Iqbal, M.S. 2024. The influence of pharmacist-led collaborative care on clinical outcomes in type 2 diabetes mellitus: A multicenter randomized control trial. *Frontiers in Public Health* 12: 1323102.

Juhan, N., Zubairi, Y.Z., Khalid, Z.M. & Zuhdi, A.S.M. 2020. A comparison between Bayesian and frequentist approach in the analysis of risk factors for female cardiovascular disease patients in Malaysia. *ASM Science Journal 13*. http://dx.doi.org/10.32802/asmscj.2020.sm26(1.1)

Juhan, N., Khalid, Z.M., Zubairi, Y.Z., Zuhdi, A.S.M. & Ahmad, W.A.W. 2019. Risk factors of cardiovascular disease among ST-elevation myocardial infarction male patients in Malaysia from 2006 to 2013. *Jurnal Teknologi* 81(3). https://doi.org/10.11113/jt.v81.12194

Juhan, N., Zubairi, Y.Z., Zuhdi, A., Khalid, Z.M. & Ahmad, W.A.W. 2018. Gender differences in mortality among ST-elevation myocardial infarction patients in Malaysia from 2006 to 2013. *Annals of Saudi Medicine* 38(1): 1-7.

Kaul, P., Manh Chu, L., Dover, D.C., Yeung, R.O., Eurich, D.T. & Butalia, S. 2022. Disparities in adherence to diabetes screening guidelines among males and females in a universal care setting: A population-based study of 1,380,697 adults. *Lancet Regional Health America* 14: 100320.

Kim, S., Hairi, N.N., Mustapha, F.I., Mohd Yusof, K., Mohd Ali, Z. & Foong, M.M. 2021. Poorer attainment of hemoglobin A1C, blood pressure and LDL-cholesterol goals among younger adults with type 2 diabetes. *Sains Malaysiana* 50(12): 3631-3645.

Lee, C., Chee, S., Arasu, K., Kwa, S., Ali, S. & Seremban, N. 2020. Diabetes literacy and knowledge among patients with type 2 diabetes mellitus attending a primary care clinic in Seremban, Malaysia. *Malaysian Journal of Nutrition* 25(3): 435-444.

Mansoori, A., Sahranavard, T., Hosseini, Z.S., Soflaei, S.S., Emrani, N., Nazar, E., Gharizadeh, M., Khorasanchi, Z., Effati, S., Ghamsary, M., Ferns, G., Esmaily, H. & Mobarhan, M.G. 2023. Prediction of type 2 diabetes mellitus using hematological factors based on machine learning approaches: A cohort study analysis. *Scientific Reports* 13(1): 663.

Mantovani, A., Byrne, C.D., Zheng, M.H. & Targher, G. 2020. Diabetes as a risk factor for greater COVID-19 severity and in-hospital death: A meta-analysis of observational studies. *Nutrition, Metabolism and Cardiovascular Diseases* 30(8): 1236-1248.

Mathew, T.K., Muhammad, Z. & Prasanna, T. 2023. *Blood Glucose Monitoring*. StatPearls Publishing LLC.

Narjis, M., Noreen, M., Safi, S.Z., Ilahi, N.E., Alomar, S.Y. & Alkhuriji, A.F. 2021. Cross talk between complete blood count and progression of type II diabetes mellitus. *Journal of King Saud University-Science* 33(6): 101492.

Park, J.M., Lee, H.S., Park, J.Y., Jung, D.H. & Lee, J.W. 2021. White blood cell count as a predictor of incident type 2 diabetes mellitus among non-obese adults: A longitudinal 10-year analysis of the Korean genome and epidemiology study. *Journal of Inflammation Research* 14: 1235-1242.

Pawitra, A.S., Diyanah, K.C., Latif, M.T. & Amiruddin, Z. 2022. Determinants of COVID-19 prevention equipment readiness in families during pandemic in Indonesia. *Sains Malaysiana* 51(6): 1945-1955.

Rafaqat, S. & Rafaqat, S. 2023. Role of hematological parameters in pathogenesis of diabetes mellitus: A review of the literature. *World Journal of Hematology* 10(3): 25-41.

Rahman, T., Khandakar, A., Hoque, M.E., Ibtehaz, N., Kashem, S.B., Masud, R., Shampa, L., Hasan, M.M., Islam, M.T., Al-Maadeed, S., Zughaier, S.M., Badran, S., Doi, S.A.R. & Chowdhury, M.E.H. 2021. Development and validation of an early scoring system for prediction of disease severity in COVID-19 using Complete Blood Count parameters. *IEEE Access: Practical Innovations, Open Solutions* 9: 120422-120441.

Rawi, S. 2023. Risk perception and its predictors towards type 2 diabetes mellitus among students in Universiti Kebangsaan Malaysia. *Medicine* 102(43): e35520.

Read, A.J., Zhou, W., Saini, S.D., Zhu, J. & Waljee, A.K. 2023. Prediction of gastrointestinal tract cancers using longitudinal electronic health record data. *Cancers* 15(5): 1399.

Saddki, N., Hashim, M. & Mohamad, N. 2022. Factors associated with oral health literacy among patients with type 2 diabetes mellitus attending Hospital Universiti Sains Malaysia. *Pesquisa Brasileira em Odontopediatria e Clínica Integrada* 22: e210143.

Saidi, S. 2023. Determining knowledge and practice of self-foot care and adherence of patient on self-foot assessment among patients with diabetes mellitus. *The Malaysian Journal of Nursing* 14(04): 62-72.

Sainani, K.L. 2021. Multinomial and ordinal logistic regression. *PMR* 13(9): 1050-1055.

Seo, I.H. & Lee, Y.J. 2022. Usefulness of Complete Blood Count (CBC) to assess cardiovascular and metabolic diseases in clinical settings: A comprehensive literature review. *Biomedicines* 10(11): 2697.

Signing, A.T., Marbou, W.J.T., Beng, V.P. & Kuete, V. 2020. Haematological Features and Urologic Pathologies of Diabetic Subjects at Bafoussam Regional Hospital: A Cross-Sectional Study. *International Journal of Chronic Diseases* 2020: 1-10.

Sun, H., Saeedi, P., Karuranga, S., Pinkepank, M., Ogurtsova, K., Duncan, B.B., Stein, C., Basit, A., Chan, J.C.N., Mbanya, J.C., Pavkov, M.E., Ramachandaran, A., Wild, S.H., James, S., Herman, W.H., Zhang, P., Bommer, C., Kuo, S., Boyko, E.J. & Magliano, D.J. 2022. IDF diabetes atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. *Diabetes Research and Clinical Practice* 183: 109119.

Thapa, J., Budhathoki, S.S., Niraula, S.R., Pandey, S., Thakur, N. & Pokharel, P.K. 2022. Prehypertension and its predictors among older adolescents: A cross-sectional study from eastern Nepal. *PLOS Global Public Health* 2(9): e0001117.

Teng, C., Chan, C. & Wong, P. 2022. Medication adherence of persons with type 2 diabetes in Malaysia: A scoping review and meta-analysis. *Journal of The ASEAN Federation of Endocrine Societies* 37(1): 75-82.

Tomic, D., Shaw, J.E. & Magliano, D.J. 2022. The burden and risks of emerging complications of diabetes mellitus. Nature Reviews. *Endocrinology* 18(9): 525-539.

Wynants, L., Van Calster, B., Collins, G.S., Riley, R.D., Heinze, G., Schuit, E., Bonten, M.M.J., Dahly, D.L., Damen, J.A.A., Debray, T.P.A., de Jong, V.M.T., De Vos, M., Dhiman, P., Haller, M.C., Harhay, M.O., Henckaerts, L., Heus, P., Kammer, M., Kreuzberger, N., Lohmann, A., Luijken, K., Ma, J., Martin, G.P., McLernon, D.J., Andaur Navarro, C.L., Reitsma, J.B., Sergeant, J.C., Shi, C., Skoetz, N., Smits, L.J.M., Snell, K.I.E., Sperrin, M., Spijker, R., Steyerberg, E.W., Takada, T., Tzoulaki, I., van Kuijk, S.M.J., van Bussel, B., van der Horst, I.C.C., van Royen, F.S., Verbakel, J.Y., Wallisch, C., Wilkinson, J., Wolff, R., Hooft, L., Moons, K.G.M. & van Smeden, M. 2020. Prediction models for diagnosis and prognosis of COVID-19: Systematic review and critical appraisal. *BMJ* 369: m1328.

Wang, Y., Li, X., Xu, J. & Zhou, Q. 2022. A complete blood count-based multivariate model for predicting the recovery of patients with moderate COVID-19: A retrospective study. *Scientific Reports* 12(1): 18262.

Wang, Y., Yang, P., Yan, Z., Liu, Z., Ma, Q., Zhang, Z., Wang, Y. & Su, Y. 2021. The relationship between erythrocytes and diabetes mellitus. *Journal of Diabetes Research* 2021: 6656062.

Yadav, S.K. & Akhter, Y. 2021. Statistical modeling for the prediction of infectious disease dissemination with special reference to COVID-19 spread. *Frontiers in Public Health* 9: 645405.

Yaribeygi, H., Maleki, M., Sathyapalan, T., Jamialahmadi, T. & Sahebkar, A. 2021. Obesity and insulin resistance: A review of molecular interactions. *Current Molecular Medicine* 21(3): 182-193.

Yong, K.C., Noor Azimah Muhammad, Lim, W.Y. & Amar-Singh, HSS. 2020. Tuberculosis and associated factors among type 2 diabetic patients in Perak: A case control study. *Sains Malaysiana* 49(5): 1081-1088.

Zabor, E.C., Reddy, C.A., Tendulkar, R.D. & Patil, S. 2022. Logistic regression in clinical studies. *International Journal of Radiation Oncology, Biology, Physics* 112(2): 271-277.

Zhao, Y., Yang, X. T., Bai, Y. P., & Li, L. F. 2023. Association of complete blood cell count-derived inflammatory biomarkers with psoriasis and mortality. *Clinical, Cosmetic and Investigational Dermatology* 16: 3267-3278.

*Corresponding author; email: liyana87@ums.edu.my