International Journal of Academic and Applied Research (IJAAR)
ISSN: 2643-9603
Vol. 10 Issue 7 July - 2026, Pages: 87-93

Total Cholesterol Modeling Based on Blood Glucose: A Local Linear Estimator Approach in Diabetes Mellitus Patients

Marisa Rifada*, Roro Indrianingtyas, Audi Amaris, Akhbar Ramadhan Putra, Bitasya Nanda Dwita

Department of Mathematics, Faculty of Science and Technology, Airlangga University, Surabaya, Indonesia
Corresponding author: marisa.rifada@fst.unair.ac.id

Abstract: *Diabetes mellitus is a major global health threat often accompanied by dyslipidemia, which elevates cardiovascular risks. While rising cases demand precise predictive tools for cholesterol-related complications, existing studies largely rely on simple linear models that fail to capture the nonlinear, heterogeneous nature of clinical data. This study modeled total cholesterol levels based on blood glucose in diabetes patients using a flexible nonparametric local linear estimator. Secondary data from 42 outpatients at RSUD Kabupaten Karangasem. The proposed approach was compared against parametric linear, quadratic, and cubic regression models. Bandwidth selection employed Cross-Validation, Generalized Cross-Validation, and Rule of Thumb across seven kernel functions. Results revealed a significant positive correlation between blood glucose and cholesterol levels ($r = 0.710$, $p < 0.05$). Furthermore, the local linear estimator utilizing a Gaussian kernel and Cross-Validation bandwidth ($h = 1.49$) demonstrated superior predictive performance. It achieved an MSE of 104.56, a MAPE of 2.27%, and an R-squared of 90.51%, vastly outperforming parametric models, which yielded R-squared values of 40.70–48.77%. In conclusion, the nonparametric local linear approach offers a highly accurate, flexible method for predicting cholesterol levels from blood glucose data. It presents a robust decision-support tool for clinicians managing diabetes-related cardiovascular complications.*

Keywords: blood glucose; diabetes mellitus; local linear estimator; nonparametric regression; total cholesterol

1. Introduction

Diabetes mellitus (DM) is a global health threat with a significant increase in mortality rates in the last two decades [1]. One of the riskiest comorbidity complications for people with DM is lipid metabolism disorders or dyslipidemia, which is generally characterized by high total cholesterol levels [2]. Insulin resistance conditions in DM patients trigger an increase in bad cholesterol (LDL) and a decrease in good cholesterol (HDL), which in turn accelerates the formation of atherosclerotic plaques and multiplies the risk of cardiovascular diseases such as heart attack and stroke [3], [4]. Locally, the urgency of handling this condition is very visible at the Karangasem Regency Hospital, where the number of DM outpatient visits jumped dramatically from 3,344 patients in 2023 to 3,480 patients in 2024, and has exceeded 879 patients in the first quarter of 2025 alone.

The escalation of such alarming medical cases demands precise analytical interventions, where statistical modeling plays a crucial role as a policy-making tool for medical personnel to map and predict fluctuations in the risk of complications in patients [5]. Observations of clinical medical data show that relationships between variables are often very heterogeneous, complex, and rarely form constant linear curves, so a nonparametric regression approach using local linear estimators is present as a highly reliable adaptive analytical solution to minimize bias in the boundary effects area [6], [7]. The realization of an accurate predictive model is essential to formulate rapid and efficient clinical interventions [8], while contributing directly to the achievement of the Sustainable Development Goals (SDGs) Target 3.4 to reduce premature mortality due to Non-Communicable Diseases (NCDs) [9].

Empirically, various previous studies have proven the existence of a clinical relationship between blood sugar levels and cholesterol. Research conducted by [10], [11], and [12] consistently confirms a significant correlation between increased blood glucose and deterioration of lipid profiles in DM patients. Furthermore, previous research conducted specifically at Karangasem Regency Hospital has also laid an empirical foundation that there is a positive correlation between these two variables in the outpatient population in hospitals [13].

Although correlations between variables have been widely proven, previous series of studies have been largely limited to simple linear association testing and have not yet made a step towards the formulation of comprehensive predictive functional models. The use of conventional parametric regression that strictly requires the assumption of certain curve shapes is prone to resulting in biased and less precise predictive estimates when faced with clinical data in the field [14]. Based on these methodological gaps, this study aims to formulate "Modeling of Total Cholesterol Levels Based on Blood Sugar Levels in Patients with Diabetes Mellitus at Karangasem Regency Hospital with Local Linear Estimators" in order to obtain flexible estimates that are able to follow the actual data pattern.

2. Methods

2.1 Data Source

This study utilized secondary data obtained from RSUD Karangasem. The dataset consisted of 42 patient records. The use of secondary data allows the analysis to be conducted based on existing clinical records without direct data collection. Such data are commonly used in statistical modeling studies [15], including studies examining the relationship between blood glucose and cholesterol levels in diabetes mellitus patients [11], [12].

2.2 Research Variables

The variables used in this study consist of one independent variable and one dependent variable. The independent variable is denoted by X, while the dependent variable is denoted by Y. The definitions of these variables are as follows:

Table 1: Research Variables

Variable	Definitions	Units
X	Blood glucose	mg/dL
Y	Total cholesterol	mg/dL

The independent variable (X) functions as a predictor, whereas the dependent variable (Y) represents the response variable influenced by changes in X [16].

2.3 Statistical Analysis

2.3.1 Parametric Regression

Parametric regression was employed as an initial approach to identify the general pattern of the relationship between the independent and dependent variables. The models considered in this study include linear, quadratic, and cubic regression models [17].

The linear regression model is expressed as:

Table 2: Parametric Regression

Regression	Models
Linear	$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$
Quadratic	$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \varepsilon_i$
Cubic	$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \beta_3 X_i^3 + \varepsilon_i$

where Y_i represents the response variable, X_i is the predictor variable, β_0, β_1, β_2, and β_3 are regression coefficients, and ε_i is the error term.

2.3.2 Local Linear Estimator

To capture nonlinear patterns, a local linear estimator was applied. This method is a nonparametric regression technique that does not assume a global functional form [18]. This flexibility allows the model to adapt to local trends in the data that may be too complex for a global parametric model [19].

The estimator is obtained by minimizing the weighted least squares function:

$$m(x) = \sum_{j=0}^{p} K\left(\frac{X_i - x}{h}\right)\left[Y_i - \left(\beta_0 + \beta_1 (X_i - x)\right)\right]^2 \quad (1)$$

where K is the kernel function and h is the bandwidth parameter. This approach allows the model to adapt locally to the structure of the data [19].

2.3.3 Kernel Function

Kernel functions are used to assign weights to observations based on their distance from the estimation point [20]. In kernel regression, the choice of bandwidth is generally considered more critical than the choice of kernel function itself [21]. The general form is:

$$K_h(x) = \frac{1}{h} K\left(\frac{x}{h}\right), \qquad \text{for } -\infty < x < \infty, h > 0 \quad (2)$$

Some commonly used kernel functions are defined as follows:

Table 3: Kernel Functions

Kernel	Functions
Uniform	$K(x) = \frac{1}{2}$ for $\lvert x\rvert \le 1$
Triangle	$K(x) = (1 - \lvert x\rvert)$ for $\lvert x\rvert \le 1$
Epanechnikov	$K(x) = \frac{3}{4}(1 - x^2)$ for $\lvert x\rvert \le 1$
Kuartik	$K(x) = \frac{15}{16}(1 - x^2)^2$ for $\lvert x\rvert \le 1$
Triweight	$K(x) = \frac{35}{32}(1 - x^2)^3$ for $\lvert x\rvert \le 1$
Cosinus	$K(x) = \frac{\pi}{4}\cos\left(\frac{\pi x}{2}\right)$ for $\lvert x\rvert \le 1$
Gaussian	$K(x) = \frac{1}{\sqrt{2\pi}}\left(\frac{1}{2}(-x^2)\right)$ for $-\infty < x < \infty$

2.3.4 Bandwidth Selection

Bandwidth is a smoothing parameter that controls the trade-off between bias and variance in nonparametric regression [18], [22].

Cross Validation (CV):

$$CV(h) = \left(\frac{1}{n}\right)\sum_{i=1}^{n}\left(y_i - \hat{g}_h^{-i}(x_i)\right)^2 \quad (3)$$

Generalized Cross Validation (GCV):

$$GCV(h) = \frac{\sum_{i=1}^{n}\left(y_i - \hat{g}(x_i)\right)^2}{(n^{-1} trace[\boldsymbol{I} - \boldsymbol{H}(h)])^2} \quad (4)$$

Rule of Thumb:

$$h = c \cdot \sigma \cdot n^{-\frac{1}{5}} \tag{5}$$

Bandwidth selection plays a crucial role in determining the smoothness of the estimated curve.

2.3.5 Model Evaluations

Model performance is evaluated using several statistical measures [23] (Hastie et al., 2009):

Mean Square Error (MSE):

$$MSE = \frac{\sum_{i=1}^{N}(y_i - \hat{y})^2}{n} \tag{6}$$

Coefficient of Determination (R^2):

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - \widehat{m}(x_i)\right)^2}{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2} \tag{7}$$

Mean Absolute Percentage Error (MAPE):

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{A_i - F_i}{A_i}\right| \times 100\% \tag{8}$$

The best model is selected based on the smallest MSE and MAPE values and the largest R^2 value [24] (Nabillah & Ranggadara, 2020).

3. Results and Discussion

3.1 Results

3.1.1 Descriptive Statistics

Descriptive statistics for the blood sugar and cholesterol variables are presented in **Table 4** below.

Table 4: Descriptive Statistical Analysis

Var	Count	Mean	StDev	Variance	SUM	Min	Max
X	42	183,29	58,28	3396,99	7698,18	84	273
Y	42	242,62	33,61	1129,46	10316,04	155	295

Next is a scatterplot, which is used to examine the pattern of the relationship, the direction of the relationship (positive/negative), the nature of the relationship (linear or nonlinear), and the presence of outliers before building a regression model.

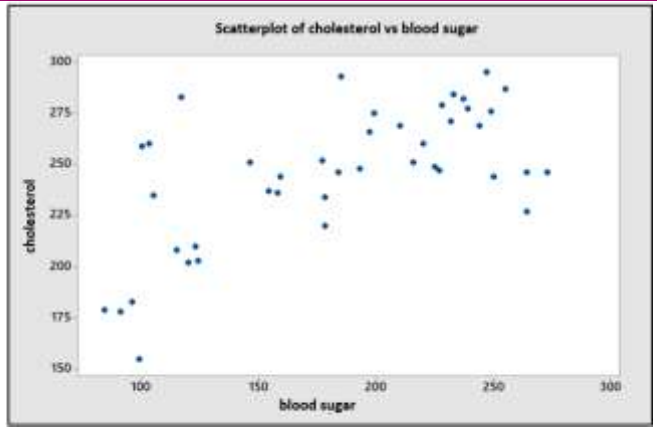

Fig. 1. *Scatterplot of total cholesterol levels by blood sugar level.*

Based on the figure, it can be seen that cholesterol levels in patients with diabetes mellitus form a scattered plot with an upward trend. Next, a correlation test will be conducted to determine the linear relationship between the predictor variable (blood glucose level) and the response variable (cholesterol level), with the following hypothesis:

H_0 : There is no linear relationship between blood sugar levels and cholesterol levels in patients with diabetes mellitus.

H_1 : There is a linear relationship between blood sugar levels and cholesterol levels in patients with diabetes mellitus.

The critical region rejects H_0 if p-value < 0,05.

The following is the output from the correlation test between blood sugar levels (X) and cholesterol levels (Y) in patients with diabetes mellitus.

Table 5: Output of the Correlation Test Results for Variables X and Y

Pearson Correlation	0,710
P-value	0,000

Based on the table, the p-value (0,000) is less than 0,05. Therefore, the decision from this correlation test is to reject H_0, and it can be concluded that there is a statistically significant linear relationship between blood glucose levels and cholesterol levels in patients with diabetes mellitus. The positive correlation coefficient indicates that as blood sugar levels increase, cholesterol levels also tend to increase.

3.1.2 Parametric Linear Regression Method

The linear regression estimate for total cholesterol levels based on blood sugar levels is shown in the following figure.

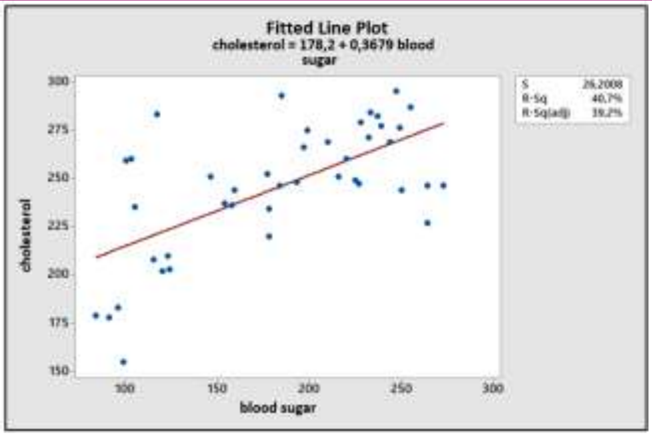

Fig. 2. *Linear regression plot of total cholesterol levels based on blood sugar levels.*

Based on the figure, we obtain

$$y = 178{,}2 + 0{,}3679x \quad (9)$$

The equation yields an estimated R-squared value of 40,7% the accuracy of this estimate falls into the moderate category, with an MSE of 653,78 and a MAPE of 8,71%, which is considered highly accurate.

3.1.3 Parametric Quadratic Regression Method

The quadratic regression estimate for total cholesterol levels based on blood sugar levels is shown in the following figure.

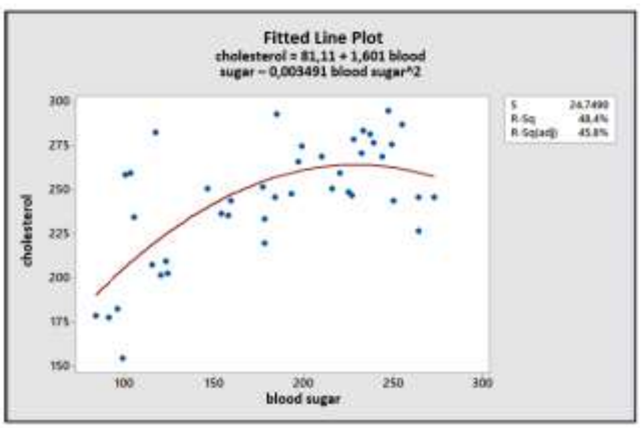

Fig. 3. *Quadratic regression plot of total cholesterol levels based on blood sugar levels.*

Based on the figure, we obtain

$$y = 81{,}11 + 1{,}601x - 0{,}003491x^2 \quad (10)$$

The equation yields an estimated R-squared value of 48,41% the accuracy of this estimate falls into the moderate category, with an MSE of 568,76 and a MAPE of 8,15%, which is considered highly accurate.

3.1.4 Parametric Cubic Regression Method

The cubic regression estimate for total cholesterol levels based on blood sugar levels is shown in the following figure.

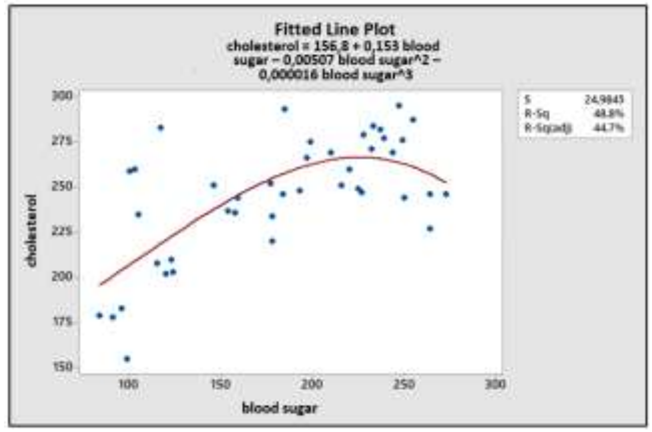

Fig. 4. *Cubic regression plot of total cholesterol levels based on blood sugar levels.*

Based on the figure, we obtain

$$y = 156{,}8 + 0{,}153x + 0{,}00507x^2 - 0{,}000016x^3 \quad (11)$$

The equation yields an estimated R-squared value of 48,77% the accuracy of this estimate falls into the moderate category, with an MSE of 564,77 and a MAPE of 8%, which is considered highly accurate.

3.1.5 Non-Parametric Local Linear Regression Method

The selection of the optimal bandwidth was performed using three methods, that is Cross-Validation, Generalized Cross-Validation, and Rule of Thumb. These methods were evaluated across seven types of kernels, namely Gaussian, Epanechnikov, Triangle, Cosine, Triweight, Quartic, and Uniform. The complete comparison of the bandwidth selection methods and kernel types is presented in **Table 6** below.

Table 6: Comparison of Methods and Kernels on Overall Data

Metode	Kernel	MSE	MAPE	R²	Bandwidth
CV	Gaussian	104,5665	2,2743 %	0,9051 %	1,4999
GCV	Gaussian	514,0905	7,3545 %	0,5337 %	20
	Epanechnikov	345,4494	5,3383 %	0,6866 %	12
	Triangle	310,0902	5,0348 %	0,7187 %	12
	Cosinus	341,6526	5,3046 %	0,6901 %	12

International Journal of Academic and Applied Research (IJAAR)
ISSN: 2643-9603
Vol. 10 Issue 7 July - 2026, Pages: 87-93

	Triweight	342,4260	5,3240 %	0,6893 %	12
	Quartic	333,2169	5,2550 %	0,6977 %	12,5
	Uniform	385,1376	5,7393 %	0,6506 %	9
Rule of Tumb	Gaussian	316,1112	5,1404 %	0,7132 %	4,3187
	Epanechnikov	344,2209	5,4166 %	0,6878 %	9,5608
	Triangle	301,0148	4,9921 %	0,7269 %	10,5031
	Cosinus	341,3248	5,3883 %	0,6904 %	9,8248
	Triweight	322,5531	5,1975 %	0,7074 %	12,8617
	Quartic	326,7838	5,2291 %	0,7036 %	11,3264

Based on the results, the CV method with the Gaussian kernel produced the smallest MSE and MAPE values and the largest R^2 value. This combination of CV and the Gaussian Kernel yielded an MSE of 104,56, a MAPE of 2,27%, and an R^2 of 90,51%, with an optimal bandwidth (h) of 1,49.

The optimal bandwidth value obtained from the Cross-Validation (CV) method was used in R software to estimate the parameter $\hat{\beta}$ in the local linear estimator. The estimation results were then used to generate a prediction curve for total cholesterol levels based on blood glucose levels. A plot of the estimation results using the Gaussian kernel is presented in the following figure.

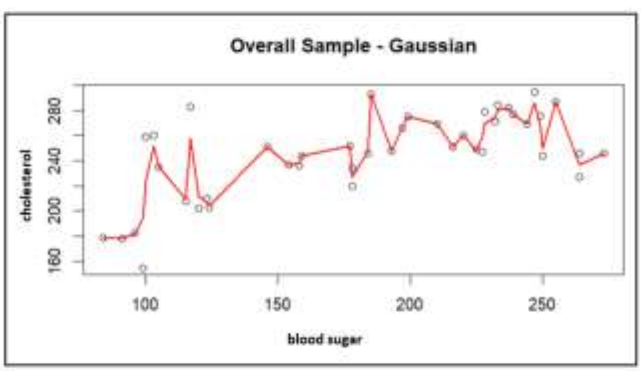

Fig. 5. *Comparison of actual total cholesterol levels and predicted cholesterol levels based on blood glucose levels in the overall data set.*

Based on Fig. 5, the red line connecting the data points represents the optimal model. The flexibility of this model is considered optimal because it captures the main trends in the data without becoming overly complex. These characteristics give the model high predictive accuracy, both when applied to existing data and when faced with new data. Therefore, the resulting model has proven to be balanced, efficient, and highly generalizable.

3.2 Discussion

Based on the results of the analysis to select the best regression model, which was based on the MSE, MAPE, and R-squared values obtained earlier. The findings are presented in the following table.

Table 7: Comparison of Regression Models

Value	Parametric Regression			Non Parametric Regression
	Linear	Quadratic	Cubic	
MSE	653,79	568,76	564,77	104,56
R^2	40,70%	48,41%	48,77%	90,51%
$MAPE$	8,71%	8,15%	8,00%	2,27%

Based on the MSE, MAPE, and R-squared values in the table above for cholesterol level data of patients with diabetes mellitus based on blood glucose levels, it was found that nonparametric regression using a Gaussian kernel is the best regression model for analyzing the data, with an MSE of 104,56, a MAPE of 2,27%, and an R-squared of 90,51%.

4. Conclusion

This study confirms that blood glucose levels have a significant positive relationship with total cholesterol levels in diabetes mellitus patients at RSUD Kabupaten Karangasem, as indicated by a Pearson correlation of 0.710 ($p < 0.05$). Comparison among parametric regression models (linear, quadratic, and cubic) revealed that all three produced relatively similar and limited predictive accuracy, with R-squared values ranging from 40.70% to 48.77%, indicating that the relationship between the two variables does not follow a simple linear or polynomial pattern. The nonparametric local linear estimator, using the Gaussian kernel with bandwidth selected through the Cross-Validation method (h = 1.49), substantially outperformed all parametric models, yielding an MSE of 104.56, a MAPE of 2.27%, and an R-squared of 90.51%. These results confirm that the local linear estimator is more capable of capturing the heterogeneous and nonlinear pattern of clinical data without being constrained by a fixed functional form, thereby producing a more flexible, accurate, and generalizable predictive model. Practically, this model can serve as a decision-support tool for medical personnel at RSUD Kabupaten Karangasem in anticipating dyslipidemia risk in diabetes mellitus patients based on their blood glucose levels, allowing for earlier and more targeted clinical intervention. Future research is recommended to apply this local linear estimator approach to larger and more diverse datasets, as well as to explore its extension to multiple predictor variables, in order to strengthen the model's generalizability and support broader efforts toward achieving SDGs Target 3.4 on reducing premature mortality from non-communicable diseases.

5. Acknowledgments

The authors would like to express their sincere gratitude to RSUD Kabupaten Karangasem for granting access to the secondary patient data. Special acknowledgment is also given to Ida Ayu Nyoman Widya Puspita, whose 2025 scientific paper served as the source of the dataset utilized in this study. Furthermore, the authors extend their appreciation to Airlangga University for the institutional support provided throughout the completion of this research.

6. References

[1] PERKENI. (2021). *Pedoman Pengelolaan dan Pencegahan Diabetes Melitus Tipe 2 di Indonesia 2021*. Perkumpulan Endokrinologi Indonesia.

[2] Alam, S., & Aijaz, M. (2024). Complications of cardiovascular disease: The Impact of diabetes, dyslipidemia, and metabolic disorders. World J Pharm Res, 13, 321-56.

[3] Bahri, S. L., Darmaningrat, C. I. A. A., Putra, I. W. M., Adzuba, K. K., Paramita, L. K. D. L., Alfarisi, M. D., & Irmayanto, T. (2023). Complications of Diabetes Mellitus: A Review Article. *Green Medical Journal*, *5*(3), 128–136. https://doi.org/10.26717/BJSTR.2022.43.006869

[4] Bereda, G. (2022). Pathophysiology and Management of Dyslipidaemia. *Biomedical Journal of Scientific & Technical Research*, *43*(2), 34369–34375. https://doi.org/10.26717/bjstr.2022.43.006869

[5] Babarinde, A. O., Ayo-Farai, O., Maduka, C. P., Okongwu, C. C., & Sodamade, O. (2023). Data analytics in public health, A USA perspective: A review. World Journal of Advanced Research and Reviews, 20(3), 211-224.

[6] Alfiani, M. L., Nur, I. M., & Utami, T. W. (2014). MODEL REGRESI NONPARAMETRIK BERDASARKAN ESTIMATOR POLINOMIAL LOKAL KERNEL PADA KASUS PERTUMBUHAN BALITA. *Statistika*, *2*(1), 34–39.

[7] Nurhuda, G., Wasono, & Nohe, D. A. (2022). Nonparametric Regression Modeling Based on Spline Truncated Estimator on Simulation Data. *Jurnal Matematika, Statistika Dan Komputasi*, *19*(1), 172–182. https://doi.org/10.20956/j.v19i1.21534

[8] Eloranta, S., & Boman, M. (2022). Predictive models for clinical decision making: Deep dives in practical machine learning. *Journal of internal medicine*, *292*(2), 278-295.

[9] Freihat, O., Sipos, D., Aamir, M., & Kovacs, A. (2025). Global burden and future projections of non-communicable diseases (2000–2050): Progress toward SDG 3.4 and disparities across regions and risk factors. PLoS One, 20(12), e0336036.

[10] Lestari, Zulkarnain, & Sijid. ST Aisyah. (2021). Diabetes Melitus: Review Etiologi, Patofisiologi, Gejala, Penyebab, Cara Pemeriksaan, Cara Pengobatan dan Cara Pencegahan. *Prosiding Biologi Achieving the Sustainable Development Goals with Biodiversity in Confronting Climate Change*, *7*(1), 237–241. http://journal.uin-alauddin.ac.id/index.php/psb

[11] Rakhmawati, A. (2024). Korelasi kadar glukosa darah dengan kolesterol total pada pasien diabetes melitus tipe 2 di Puskesmas Purwokerto Selatan. Jurnal Kesehatan dan Science, XX(2), 858–4616.

[12] Julianti, M. (2020). Hubungan kadar gula darah dengan kadar kolesterol pada pasien diabetes militus rawat jalan di Rumah Sakit Bhayangkara Kota Palembang. Politeknik Kesehatan Palembang.

[13] Puspita, I. A. N. W. (2025). *Hubungan Kadar Gula Darah Dengan Kadar Kolesterol Total Pada Pasien Diabetes Mellitus Rawat Jalan Di Rsud Kabupaten Karangasem*. Poltekkes Kemenkes Denpasar.

[14] Arredondo Montero, J. (2026). Clinical prediction models: from foundational concepts to practical application. Diagnosis, (0).

[15] Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). Introduction to linear regression analysis (5th ed.). John Wiley & Sons.

[16] Draper, N. R., & Smith, H. (1998). *Applied regression analysis* (3rd ed.). John Wiley & Sons.

[17] Kroc, E., & Olvera Astivia, O. L. (2025). The case for the curve: Parametric regression with second- and third-order polynomial functions of predictors should be routine. Psychological Methods, 30(6), 1277–1293. https://doi.org/10.1037/met0000629

[18] Fan, J., & Gijbels, I. (1996). Local polynomial modelling and its applications (Monographs on Statistics and Applied Probability, Vol. 66). Chapman & Hall.

[19] Eguasa, O., Edionwe, E., & Mbegbu, J. I. (2023). Local Linear Regression and the problem of dimensionality: A remedial strategy via a new locally adaptive bandwidths selector. Journal of Applied Statistics, 50(6), 1283–1309. https://doi.org/10.1080/02664763.2022.2026895

[20] Wand, M. P., & Jones, M. C. (1995). Kernel smoothing. Chapman & Hall.

[21] Sukarsa, I. K. G., & Srinadi, I. G. A. M. (2012). Estimator kernel dalam model regresi nonparametrik. Jurnal Matematika, 2(1), 19–30.

[22] Turlach, B. A. (n.d.). Bandwidth selection in kernel density estimation: A review.

[23] Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: Data mining, inference, and prediction (2nd ed.). Springer.

[24] Nabillah, I., & Ranggadara, I. (2020). Mean Absolute Percentage Error untuk evaluasi hasil prediksi komoditas laut. JOINS (Journal of Information System), 5(2), 250–255.

[23] Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: Data mining, inference, and prediction (2nd ed.). Springer.

[24] Nabillah, I., & Ranggadara, I. (2020). Mean Absolute Percentage Error untuk evaluasi hasil prediksi komoditas laut. JOINS (Journal of Information System), 5(2), 250–255. https://doi.org/10.33633/joins.v5i2.3900

[25] De Brabanter, K., Tabassum, M., & Wang, Y. (2025). Bandwidth selection for multivariate local linear regression with correlated errors. TEST, 34(4), 1104–1139. https://doi.org/10.1007/s11749-025-00988-4

[26] Dong, H., Otsu, T., & Taylor, L. (2023). Bandwidth selection for nonparametric regression with errors-in-variables. Econometric Reviews, 42(4), 393-419.

[27] Li, J., Li, P., Li, P., Tang, L., Zhang, X., & Wu, Q. (2022). Self-position awareness based on cascade direct localization over multiple source data. IEEE transactions on intelligent transportation systems, 25(1), 796-804.

[28] WHO. (2024). *Diabetes*. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/diabetes

[29] Yavuz, E., & Şahin, M. (2022). Investigation of parametric, non-parametric and semiparametric methods in regression analysis. Sakarya University Journal of Science, 26(6), 1111-1116.