Revista ARACÊ
ISSN: 2358-2472

LINEAR REGRESSION AND FUZZY LOGIC IN THE ANALYSIS OF FACTORS INFLUENCING GLYCATED HEMOGLOBIN IN CHILDREN AND ADOLESCENTS WITH TYPE 1 DIABETES MELLITUS

https://doi.org/10.56238/arev6n4-393

Submitted on: 11/23/2024 **Publication Date:** 12/23/2024

Fábio Fontolan Sampaio[1], Henrique Villa Chagas[2], Jesselina Francisco dos Santos Haber[3], Luís Roberto Almeida Gabriel Filho[4] and Eduardo Federighi Baisi Chagas[5].

ABSTRACT

This article presents a study on the analysis of factors that influence glycated hemoglobin (HbA1c) levels in children and adolescents with type 1 diabetes mellitus (DM1). After an initial characterization of the dataset, linear regression and fuzzy logic were used to model and interpret these factors, with an additional focus on nutritional status and insulin delivery method. The study covers a sample of 80 patients between 4 and 19 years of age and is based on data collected at a medical outpatient clinic in a city in the interior of São Paulo. The results indicate a significant difference in HbA1c levels between insulin delivery methods (SICI and MDI), but not between different nutritional statuses. The linear regression model points to the time of diagnosis and total cholesterol as the most significant factors for the model. Thus, the fuzzy logic complemented the analysis of the study, allowing a three-dimensional visual representation of the variables and showing that the two variables significantly affect the levels of HbA1c.

Keywords: Glycated hemoglobin. DM1. Linear regression. Fuzzy Logic.

---

[1] Bachelor of Statistics (UNESP)Faculty of Science and Engineering of Tupã, São Paulo State University (UNESP)
E-mail: fabio.sampaio@unesp.br
ORCID: https://orcid.org/0009-0004-3919-4209
Lattes: http://lattes.cnpq.br/9430261139765585
[2] Undergraduate student in Medicine (UNIMAR)Interdisciplinary Center in Diabetes (CENID), University of Marília (UNIMAR)
E-mail: henriquevillachagas@gmail.com
ORCID: https://orcid.org/0000-0002-6690-127X
Lattes: http://lattes.cnpq.br/5781672795516191
[3] Doctorate student in Law (UNIMAR)Interdisciplinary Center in Diabetes (CENID), University of Marília (UNIMAR)
E-mail: haber.jesselina@gmail.com
ORCID: https://orcid.org/0000-0002-8060-5928
Lattes: http://lattes.cnpq.br/4504732146184670
[4] Dr. in Agronomy (UNESP)Faculty of Science and Engineering of Tupã, São Paulo State University (UNESP)
E-mail: gabriel.filho@unesp.br
ORCID: https://orcid.org/0000-0002-7269-2806
Lattes: http://lattes.cnpq.br/5255054247692798
[5] Dr. in Human Development and Technologies (UNESP)Interdisciplinary Center in Diabetes (CENID), University of Marília (UNIMAR)
Faculty of Medicine of Marília (FAMEMA)
Post-Doctoral Program at the Faculty of Sciences and Engineering of Tupã (UNESP).
E-mail: efbchagas@unimar.br
ORCID: https://orcid.org/0000-0001-6901-9082
Lattes: http://lattes.cnpq.br/0168500869625770

Revista ARACÊ
ISSN: 2358-2472

INTRODUCTION

Glycated hemoglobin (HbA1c) is a very important biological factor for monitoring the control of type 1 diabetes mellitus (DM1) (SOCIEDADE BRSILEIRA DE DIABETES, 2019). This protein, which is formed by the link between glucose and hemoglobin, is represented by the average of blood glucose levels in the last two to three months. Keeping HbA1c within the recommended limits is essential to prevent long-term complications, such as retinopathy, nephropathy, and neuropathy (MINISTRY OF HEALTH, 2020).

If we direct our attention to children and adolescents with T1D, maintaining HbA1c at adequate levels is essential to reduce the risk of acute and chronic complications. A study by Lind et al. (2019) highlighted that inadequate glycemic control considerably increases the risk of early development of microvascular complications, suggesting that controlling HbA1c from youth has long-lasting effects on preventing these conditions.

Considering that, complications of DM1 can manifest early, with cumulative impacts, affecting future health and quality of life in adulthood. These children and adolescents with DM1 already face significant emotional and social challenges due to the routine control of the disease. High HbA1c values may indicate difficulties in managing the disease, leading to higher hospitalization rates and increasing stress, anxiety, and worry for both patients and families. Studies show that children and adolescents with good glycemic control report greater emotional well-being and lower anxiety, since the risk of hypoglycemic and hyperglycemic episodes is reduced. As highlighted by Smith et al. (2020), the quality of life of young people with T1D is strongly related to glycemic control, with better HbA1c levels associated with a lower negative impact on daily activities and social and academic development.

Analyzing the factors that affect glycated hemoglobin is not simple, especially in children and adolescents. We have seen that many factors, anthropometric and ambulatory measurements, for example, can influence glycemic control and HbA1c values. According to Chiang et al. (2018), these variables have a complex relationship that, for their analysis, requires the use of statistical and mathematical tools. The applications of these techniques aim to simplify and facilitate the control of the indexes, through data modeling, serving as support for the patients themselves, their families and the health professionals who accompany them.

Linear regression and fuzzy logic are examples of valuable statistical tools for the analysis of glycated hemoglobin in children and adolescents with DM1. By combining these

Revista ARACÊ
ISSN: 2358-2472

techniques, it is possible to obtain a more complete and accurate view of the relationship between HbA1c and other factors, which, according to Castro and Novaes (2019), their results can help in making better and more accurate decisions about glycated hemoglobin control, aiming to contribute to improving quality of life and reducing the risk of complications. Linear regression would provide a basis for modeling the relationships between factors and fuzzy logic is very useful for situations in which we deal with uncertainties, inaccuracies and partial truths, as Massad et al. (2004) say, as occurs in glycemic control.

Thus, in view of this complex relationship between glycated hemoglobin and the various factors that influence it, the study aimed to analyze the use of linear regression and Fuzzy logic to identify factors that influence HbA1c values in children and adolescents with DM1.

METHODOLOGY

This is an analytical cross-sectional observational study, whose data were obtained through access to the database of patients at the Specialty Medical Outpatient Clinic of a University in the interior of the State of São Paulo, in the years 2019 and 2020, and stored in the database of an Interdisciplinary Center for Diabetes (CENID).

Considering the eligibility criteria, 80 patients of both sexes (male=47; female=33), aged between 04 and 19 years and diagnosed with DM1 for at least 12 months, were included in the study.

Other data were obtained to characterize the sample, in addition to gender (male and female), age (in years) and time of diagnosis (in years). In this study, data from anthropometric measures such as weight (in kilograms) and height (in meters) were used to calculate the Body Mass Index Z-Score (BMI-z) and categorized them according to Nutritional Status (NE) as underweight, eutrophic, overweight and obese (according to the recommendations of the World Health Organization (ONIS et al., 2007).

The Insulin Delivery Method (MAI) was grouped into Continuous Insulin Infusion System (SICI), known as insulin infusion pump, and patients with multiple doses of insulin (MDI).

Data were collected from laboratory tests of fasting glucose, HbA1c, total cholesterol (TC), LDL-C, HDL-c and triglycerides (TG).

Revista ARACÊ
ISSN: 2358-2472

Glycemic control was assessed by glycated hemoglobin (HbA1c%) by the high-performance liquid chromatography (HPLC) method. Fasting glucose was analyzed by the colorimetric enzymatic method and values <100 mg/dL were considered normal. The HbA1c% values were categorized, due to the profile of the population served, as less than 7%, from 7% to 8% and greater than 8% (El Sayed et al., 2023). Total cholesterol, HDL-c and TG were analyzed using the colorimetric method and LDL-c was calculated using the Friedewald equation (SIBAL et al., 2010).

Due to the definition of the initial parameters in the use of fuzzy logic, the data of the laboratory tests were categorized according to reference values suggested by the literature. Thus, for diagnostic interpretation, lipid parameters were categorized into normal, borderline, and elevated (FERRANTI & NEWBURGER, 2023). However, according to the Practical Update Manual of the Brazilian Society of Pediatrics, for the Brazilian population between 02 and 19 years old, it is recommended to use borderline values with a cutoff point to identify altered values and the presence of dyslipidemia is defined by the presence of at least one altered lipid parameter considering: TC ≥ 170 mg/dL; LDL-c ≥ 110 mg/dL; HDL-c ≤ 45 mg/dL; TG ≥ 75 mg/dL (0 to 9 years) and ≥ 90 mg/dL (10 to 19 years). However, to allow comparison with other studies, lipid parameters were also classified as elevated when: TC ≥ 200 mg/dL; LDL-c ≥ 130 mg/dL; HDL-c < 40 mg/dL; TG ≥ 100 mg/dL (0 to 9 years) and ≥ 130 mg/dL (10 to 19 years).

The descriptive statistics of the data were presented according to their types, quantitative or qualitative, through means and standard deviations or by absolute and relative frequencies.

Knowing the factors that influence glycated hemoglobin (physiological assumption of glycemic metabolism), the study will begin through the application of the statistical technique of linear regression, where our objective will be to model this relationship between factors (independent variables) and HbA1c (dependent variable). Thus, first, the model will include all the factors mentioned and, progressively, using the backward elimination method, all those factors that will contribute the least to the final model, that is, less statistically significant, will be removed. Finally, through the coefficient of determination, R2, the performance of the model (value closest to 1) will be verified, in which the best possible model (with fewer independent variables) that will explain a significant proportion of the HbA1c variation will be sought.

Revista ARACÊ
ISSN: 2358-2472

Once the best linear regression model for glycated hemoglobin was found, fuzzy logic techniques were applied to complement the analysis of the results. The purpose of this complementation was to minimize possible estimation errors due to uncertainties and inaccuracies in the data set. Thus, in the first stage, fuzzification was carried out, which is the creation of fuzzy sets, represented by linguistic concepts (classifications), through the transformation of independent variables (numerical values) into pertinence functions (degrees between 0 and 1) (see Figures 1, 2 and 3). After fuzzification, the fuzzy classification rules were defined, which are the relationships between the input variables (variables of the linear regression model) and the output variable (HbA1c), in which the degrees of relevance are presented (see Table 5). To conclude the analysis, the defuzzification step was performed, which converts the results to more precise numerical values (in the prediction ratio of these values) (see Table 6).

The level of significance was set at 5% and the data were analyzed using JAMOVI version 2.3.38, MATLAB version 24.2.0 and MINITAB version 21.2.

RESULTS

Table 1 presents the descriptive statistics of the characteristics of the sample in relation to the quantitative variables. Regarding HbA1c, which represents the outcome of interest, the results of the descriptive statistics indicate that the sample is composed of patients at different levels of glycemic control.

Table 1: Descriptive statistics of the study variables for 80 patients aged 4 to 19 years with a diagnosis of DM1

Variables	Average	Standard deviation	Minimum	Maximum
Age (years)	12,563	3,579	4,00	19,00
Weight (kg)	49,119	17,178	18,70	112,40
Height (m)	1,534	0,175	1,07	1,85
IMC-z (indics)	0,160	1,213	-4,50	2,90
Time of diagnosis (years)	4,275	2,977	1,00	14,00
CT (mg/dL)	165,263	34,011	87,00	246,00
TG (mg/dL)	82,349	52,816	21,72	343,00
LDL (mg/dL)	89,137	27,498	24,00	171,00
HDL (mg/dL)	55,285	10,553	23,00	76,00
HbA1c (%)	8,574	2,280	4,91	15,30

Source: Prepared by the author

Considering HbA1c as the outcome of interest, association analyses were performed to categorize HbA1c in relation to gender distribution, method of insulin administration, and nutritional status (Table 2).

Revista ARACÊ
ISSN: 2358-2472

It is important to note that, after talking to experts, it was decided to expand the number of HbA1c categories proposed by ElSayed et al. (2023), from three to four groups. This decision was made to identify more refined situations between the groups and their associations with the factors studied, because, considering the HbA1c data and their relationship with mean blood glucose, HbA1c values were categorized according to the target glycemic zones as <7%, 7% to 8%, >8% to 10% and > 10%, with the denominations of "Ideal", "Acceptable", "High" and "Very High", respectively.

In the evaluation, no association was observed between gender and HbA1c category. A significant association was found between HbA1c categories with the method of insulin administration and nutritional status. Among patients with HbA1c <7% and 7 to 8%, a higher proportion of patients using SICI was observed. Among patients with HbA1c >8%, a higher proportion of overweight was observed.

Table 2: Table of Contingencies for the categories of Glycated Hemoglobin and the variables Sex, Method of Insulin Administration and Nutritional Status.

Variables		**HbA1c**					**P-value**
		< 7%	**7% a 8%**	**> 8% a 10%**	**> 10%**	**Total**	
Sex	Male	10 21,3%	12 25,5%	15 53,2%	10 21,3%	47 100%	0,254
	Female	10 30,3%	3 9,1%	10 30,3%	10 30,3%	33 100%	
Insulin Delivery Method	SICI	9 40,9%	6 27,3%	5 22,7%	2 9,1%	22 100%	0,045*
	MDI	11 19%	9 15,5%	20 34,5%	18 31%	58 100%	
Nutritional status	Underweig ht	1 11,1%	2 22,2%	2 22,2%	4 44,4	9 100%	0,020*
	Eutrophic	13 25,5%	13 25,5%	12 23,5%	13 25,5%	51 100%	
	Overweight	4 22,2%	0 0%	11 61,1%	3 16,7%	18 100%	
	Obese	2 100%	0 0%	0 0%	0 0%	2 100%	

Note 1: Absolute values and percentage per line are displayed
Note 2: * indicates significant association by the Chi-square test for p-value < 0.05
Source: Prepared by the author

To identify the independent variables with the best performance to predict HbA1c values, a multiple linear regression model was built using the backward elimination method. The final model chosen (general) was presented in Table 3, which indicates a significant effect of time of diagnosis and total cholesterol on HbA1c. To increase the sensitivity of the model in the search for explanatory variables, the regression model was explored without the regression constant.

Table 3: Presentation of the linear regression models and coefficients of determination, for the data set of 80 patients aged 04 to 19 years, with a diagnosis of DM1, general and separated by the types of categorical predictors

Situation		Linear Regression Model	R2
General		HbA1c = 0.0822 * Time of diagnosis + 0.04888 * Total cholesterol	93,83%
State Nutritional	Underweight	HbA1c = 0.0541 * Time of diagnosis + 0.04453 * Total cholesterol	94,25%
	Eutrophic	HbA1c = 0.0541 * Time of diagnosis + 0.04453 * Total cholesterol + 1.169	
	Overweight	HbA1c = 0.0541 * Time of diagnosis + 0.04453 * Total cholesterol + 0.759	
	Obese	HbA1c = 0.0541 * Diagnosis Time + 0.04453 * Total Cholesterol – 1.52	
Method Administration of Insulin	SICI	HbA1c = 0.0998 * Time of diagnosis + 0.04016 * Total cholesterol	94,92%
	MDI	HbA1c = 0.0998 * Time of diagnosis + 0.04016 * Total cholesterol + 1.969	

Source: Prepared by the author

Also in Table 3, the linear regression models using the predictors were presented. Starting with the predictor Nutritional Status, it was found that for each of the identified types, a distinct linear regression model was obtained, changing only the values of the regression constant. In the case of the predictor Insulin Administration Method, different results were also observed for each of the methods, distinguishing them by the regression constant.

Once the influencing factors on HbA1c were defined, the application of Fuzzy Logic was initiated by the fuzzification stage, where the pertinence functions of the input variables (Time of Diagnosis and Total Cholesterol), output (HbA1c) and the fuzzy classification rules related to the interactions between the input and output variables were defined. Finally, with the defuzzification stage, the linguistic results (classifications) were converted into numerical values and the three-dimensional graphic representation was constructed.

For the so-called Fuzzy-Glycated Model, the fuzzification step with the input variable Time to Diagnosis (in years), with domain [0; 15], represented by the ranges from 0 to 5 years old and from above 5 to 15 years old, with the linguistic terms "Less than 5 years"

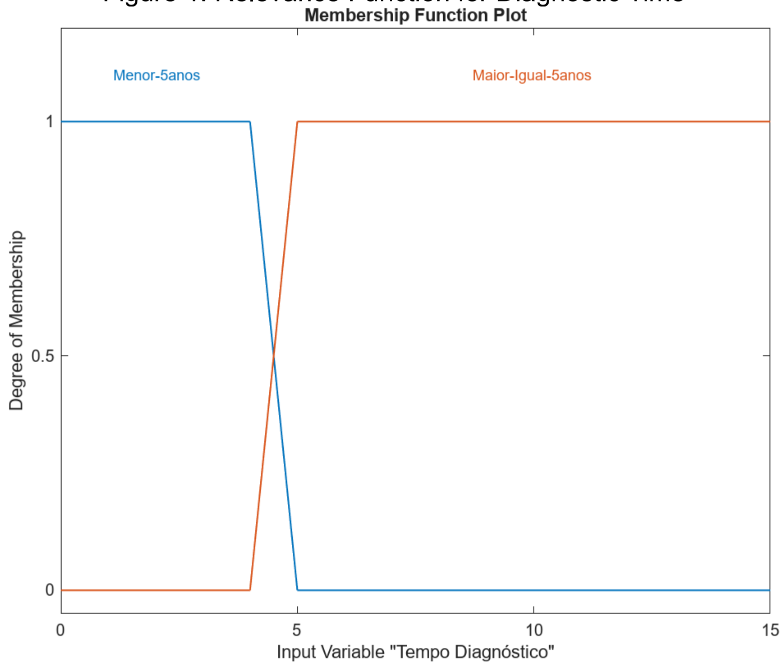
Revista ARACÊ
ISSN: 2358-2472

and "Greater Equal to 5 years", respectively, the trapezoidal pertinence functions were created, as shown in Figure 1.

For the input variable Total Cholesterol (TC, in mg/dl), with domain [0; 400], represented by the ranges of less than 170 mg/dl, from 170 to 190 mg/dl and greater than 190 mg/dl, with the linguistic terms "Ideal", "Acceptable" and "Undesirable", respectively, according to the Brazilian Society of Pediatrics, the trapezoidal pertinence functions were created, as shown in Figure 2.

Finally, for the output variable Glycated Hemoglobin (in %), with domain [0; 20], represented by the ranges of less than 7%, from 7% to 8%, greater than 8% to 10% and greater than 10%, with the linguistic terms "Ideal", "Acceptable", "High" and "Very High", respectively, according to the relationship of HbA1c with the average blood glucose and the target zones of glycemia, trapezoidal pertinence functions were created, as shown in Figure 3.

Figure 1: Relevance Function for Diagnostic Time



Revista ARACÊ
ISSN: 2358-2472

Figure 2: Pertinence Function for Total Cholesterol

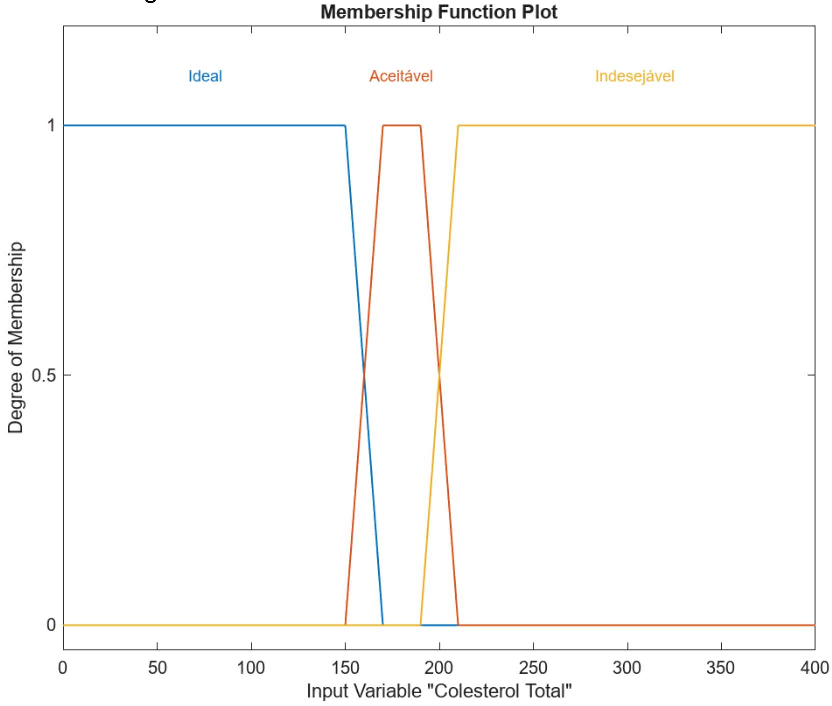
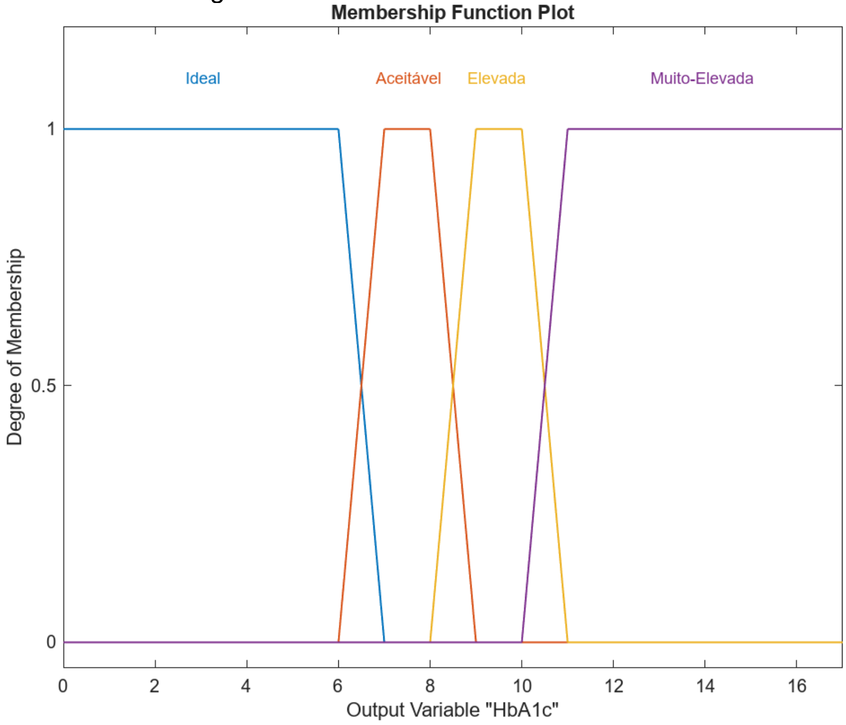


Figure 3: Relevance Function for HbA1c



Combining the three fuzzy sets, relating the two input variables with the output variable, of the generated possibilities, a total of 06 rules is observed.

After dialogues with the specialist on the subject, it was decided to create two fuzzy classification rules, as shown in Table 4, called Rule 1 Situation and Rule 2 Situation.

Revista ARACÊ
ISSN: 2358-2472

The difference between the situations is related to the Time to Diagnosis groups. For the Rule 1 Situation, the values for the classification of the output variable are less rigid for the group with time of diagnosis less than 05 years. And this rule is reversed to the Rule 2 Situation, moving the group with a diagnosis time of over 05 years old to a less rigid situation.

Table 4: Situations for the fuzzy classification rules for the relationship between the input variables Time of diagnosis and Total Cholesterol and the output variable HbA1c, according to definitions for the Fuzzy-Glycated Model

Situation	Diagnosis time	Colesterol Total	HbA1c
Rule 1	< 5 years	Ideal	Ideal
	< 5 years	Undesirable	Acceptable
	< 5 years	High	High
	≥ 5 years	Ideal	Acceptable
	≥ 5 years	Undesirable	High
	≥ 5 years	High	Very High
Rule 2	< 5 years	Ideal	Acceptable
	< 5 years	Undesirable	High
	< 5 years	High	Very High
	≥ 5 years	Ideal	Ideal
	≥ 5 years	Undesirable	Acceptable
	≥ 5 years	High	High

Source: Prepared by the author

Figures 4 and 5 present the three-dimensional graphic representations of the defuzzification stage, for the relationship between the variables Time of Diagnosis and Total Cholesterol, to explain the variable Glycated Hemoglobin. The type of defuzzification used was centroid.

Revista ARACÊ
ISSN: 2358-2472

Figure 4: Three-dimensional graphical representation of the variables Time of diagnosis, TC and HbA1c, under the Rule 1 Situation

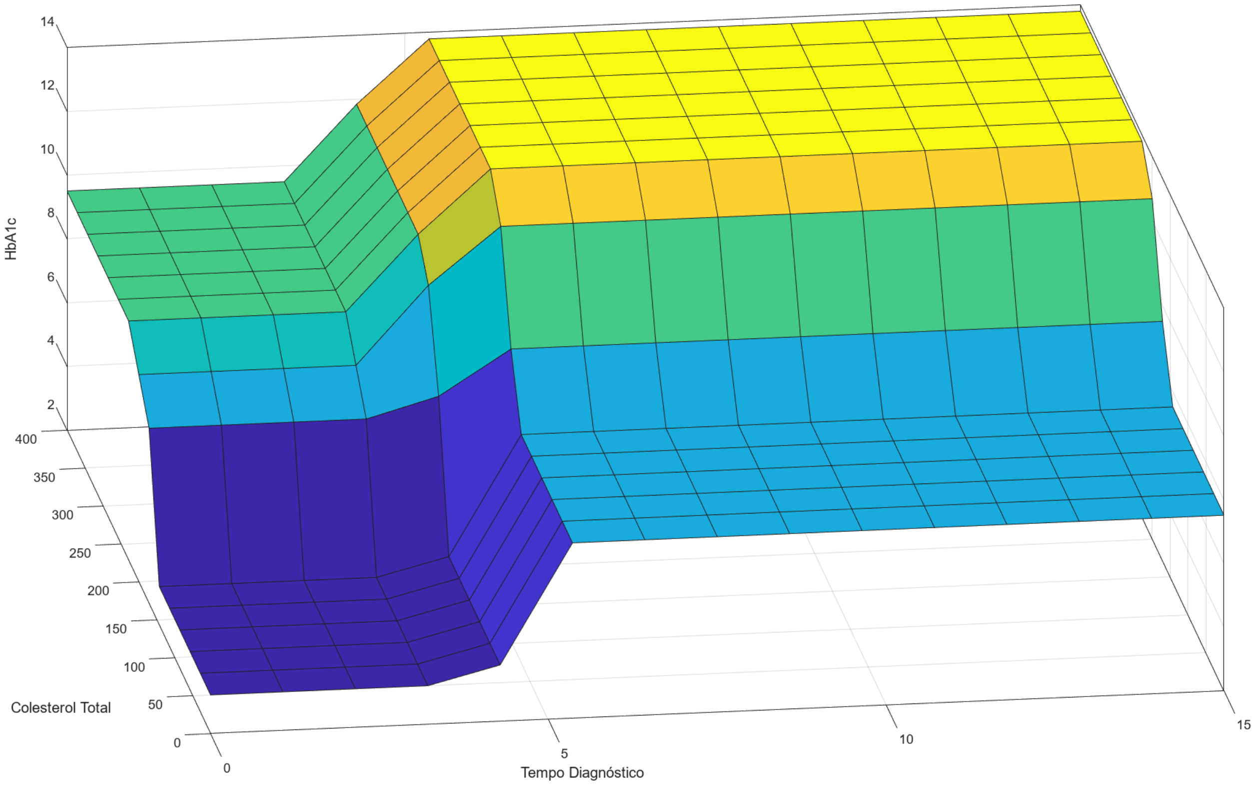
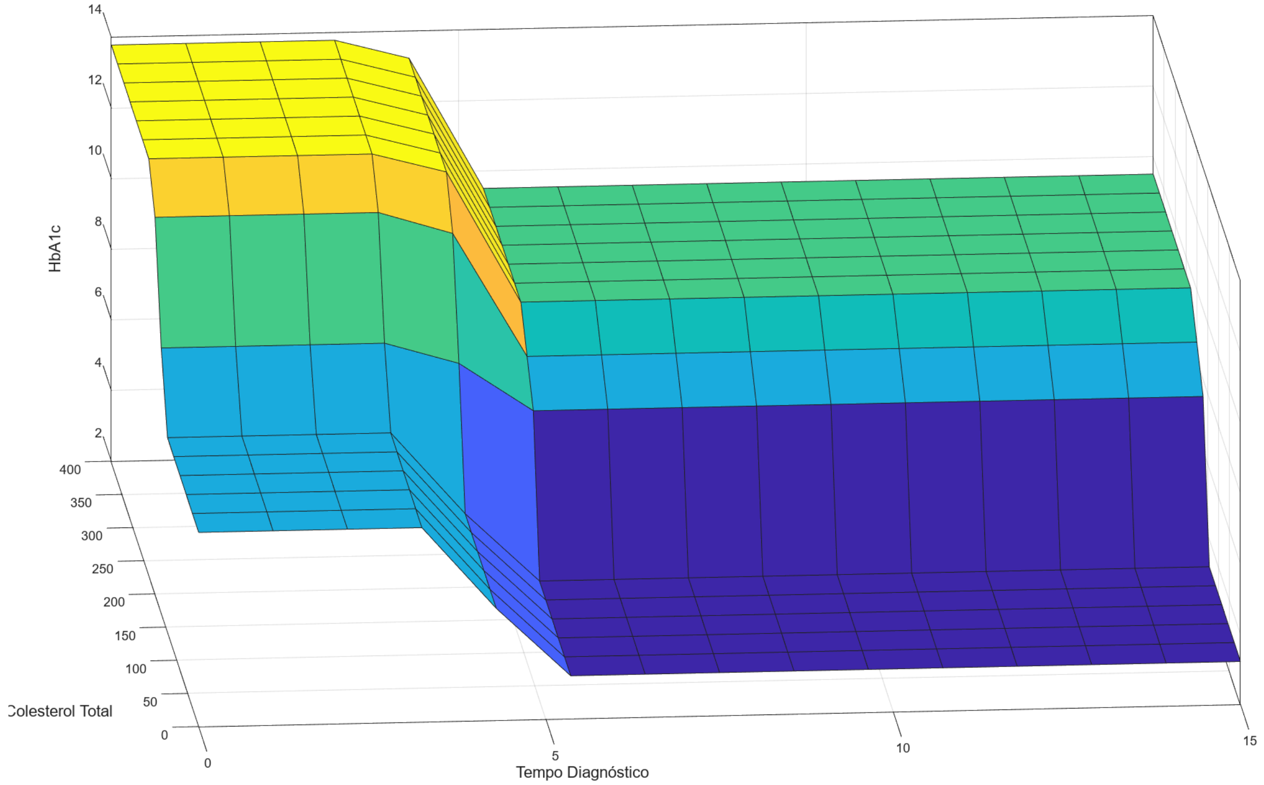


Figure 5: Three-dimensional graphical representation of the variables Time of diagnosis, TC and HbA1c, under the Rule 2 Situation



Revista ARACÊ
ISSN: 2358-2472

Defuzzification in fuzzy models converts the results of fuzzy rules to more accurate numerical HbA1c values, making it easier to apply and even compare with other studies. For the purpose of exemplification, the data collected and the data observed after applying the statistical techniques of 10 patients (randomly chosen) of the study were presented (Table 7), through their values in the range of 0-20%, where they were classified in each interval, in which they represent the classification of the level of Glycated Hemoglobin.

The results presented in Table 7 indicate that both the regression model and the Fuzzy models presented limitations for predicting HbA1c values and levels considering total cholesterol and time of diagnosis. In the regression model, it was verified that the HbA1c level was correctly predicted in only 40% of the cases. In the Fuzzy model of rule 1, it was verified that the system was able to correctly predict the level of HbA1c in only 20% of the cases. In the Fuzzy model of rule 2, the system can correctly predict the level of HbA1c in only 30% of cases.

Table 7: Comparison between the classifications of glycated hemoglobin levels for the cases in the collected data and the data modeled by linear regression and fuzzy logic

Data Collected				Modeled Data					
				Linear Regression		Fuzzy – Regra 1		Fuzzy – Regra 2	
TD	CT	HbA1c	HbA1c Level	HbA1c	HbA1c Level	HbA1c	Level HbA1c	HbA1c	Level HbA1c
1	201,00	5,88	1	9,91	3	8,59	3	12,8	4
3	224,00	15,20	4	**11,20**	**4**	9,50	3	**13,8**	**4**
3	154,00	10,20	4	7,77	2	3,58	1	7,94	2
4	172,92	8,55	3	**8,78**	**3**	7,50	2	**9,50**	**3**
5	179,00	6,90	1	9,16	3	9,50	3	7,50	2
6	203,00	7,20	2	10,42	4	13,1	4	8,77	3
6	170,00	7,30	2	8,80	3	9,50	3	**7,50**	**2**
7	212,00	14,20	4	**10,94**	**4**	**13,8**	**4**	9,50	3
9	139,00	7,60	2	**7,53**	**2**	**7,50**	**2**	3,21	1
11	184,00	6,00	1	9,90	3	9,50	3	7,50	2

Note: For interpretation at the HbA1c level, category 1 (<7%), category 2 (7 to 8%), category 3 (>8% to 10%) and category 4 (>10%) were considered.

Source: Prepared by the author

DISCUSSION

In the sample, there was no significant difference in the distribution of proportion between the sexes. Only records of consultations with complete data were included in the study. In the data collection period, there was an increase in the number of queries with complete data. The lower number of consultations in 2019 to 2020 compared to 2022 and 2023 is due to the COVID-19 pandemic period and the gradual return of consultation

Revista ARACÊ
ISSN: 2358-2472

routines. The largest proportion of the sample used MDI as a method of insulin administration (Table 1).

Glycated hemoglobin (HbA1c) is considered an important indicator in monitoring glycemic control in people with diabetes. It reflects the average of blood glucose levels in the last 2-3 months and is a determining parameter for assessing the risk of diabetes-related complications (CHIANG et al., 2018). Lack of glycemic control, characterized by elevated HbA1c values, is related to increased oxidative stress and the formation of advanced glycation products (AGEs). These mechanisms, as described by Yamagishi et al. (2015), are interconnected and play central roles in the development of microvascular and macrovascular complications.

Oxidative stress is defined as an imbalance between the production of reactive oxygen species (ROS) and the antioxidant capacity of the body. Hyperglycemia increases the generation of ROS by activating the mitochondrial respiratory chain in glucose-dependent cells such as endothelial and renal cells. Excess intracellular glucose activates metabolic pathways such as polyol (sorbitol), hexosamine, and protein kinase C (PKC), contributing to increased oxidative stress. As described by Wang and Zhang (2024), this condition results in damage to lipids (lipid peroxidation), proteins (oxidation), and DNA (mutations and dysfunctions). Endothelial dysfunction due to a reduction in the bioavailability of nitric oxide (NO) promotes vasoconstriction and inflammation, aggravating tissue lesions.

Advanced Glycation End Products (AGEs) are compounds formed from non-enzymatic reactions between reducing sugars and proteins, lipids or nucleic acids, promoted by chronic hyperglycemia. The formation of AGEs is accelerated under conditions of hyperglycemia and oxidative stress, with an impact on structural proteins such as collagen, reducing elasticity and promoting tissue stiffness. The binding of AGEs to RAGE (Receptor for Advanced Glycation End-products) receptors activates inflammatory intracellular cascades, such as NF-κB, promoting additional oxidative stress. This condition stimulates the production of pro-inflammatory cytokines (TNF-α, IL-6), which leads to chronic inflammation aggravating microvascular and macrovascular dysfunction (references). These mechanisms reinforce the need for early interventions to prevent the damage caused by persistent hyperglycemia by protecting target tissues such as the kidneys, eyes, nerves, and cardiovascular system. Therefore, T1D therapy is essentially based on insulin administration and blood glucose monitoring to promote the maintenance

Revista ARACÊ
ISSN: 2358-2472

of blood glucose at acceptable values. In addition, a nutritional approach, which involves carbohydrate counting and changes in habits in relation to physical activity, contributes significantly to the control of the disease. Thus, the management of T1D requires a multidisciplinary approach and an active involvement of the patient and his family (CHIANG et al., 2018).

Although adherence to insulin therapy is the main factor related to good glycemic control, other factors such as time of diagnosis, body composition, and lipid profile are pointed out as factors that can influence glycemic control and increase the risk of complications. Studies, such as the one by Rossaneis et al. (2019), highlight the importance of investigating the impact of these variables for a better understanding of the mechanisms involved in controlling the disease.

However, the mathematical models used to analyze the effect of risk factors on health indicators and disease control can provide different results and perspectives. Considering the complexity of the interaction of physiological mechanisms on the behavior of a given outcome, multiple linear regression and logistic regression models are widely used for studies that seek to investigate this cause-effect relationship (COX, 2021).

Linear regression can be used to identify and quantify the relationship between HbA1c and influencing factors, such as age, time of diagnosis, insulin doses, frequency of glycemic monitoring, adherence to treatment, physical activity, and psychosocial factors. This analysis model is relatively simple and allows the direct interpretation of coefficients when they present linear relationships. However, they have limitations to identify nonlinear relationships between variables.

Fuzzy logic is another mathematical model that can be used for the analysis of influencing factors on an outcome of interest, although its use is still limited in studies of patients with DM1. Fuzzy Logic handles uncertainty and imprecise data well, assigning degrees of pertinence (values between 0 and 1) instead of rigid categorizations. As shown by Hasan et al. (2024), this approach is useful for modeling complex or nonlinear relationships between variables and for incorporating qualitative factors such as quality of life or perceived adherence to treatment. Thus, it has the advantage of flexibility to deal with uncertainties, adaptive interpretation, and integration of qualitative and quantitative variables.

Although Fuzzy Logic seems advantageous over regression models, the integrated use of the two approaches can provide interpretations that are closer to reality. As

Revista ARACÊ
ISSN: 2358-2472

demonstrated by Sahoo and Chakraverty (2024), linear regression can provide an initial insight into associations between variables, while Fuzzy Logic can capture additional nuances and complement analyses. However, it is worth noting that the complexity of biological interactions in relation to the number of influencing factors must be considered, as none of the models is able to capture all the factors that determine the response to an outcome of interest. It is worth noting that the present study does not aim to discuss physiological aspects of glycemic control, but to analyze the application of the use of regression and Fuzzy Logic for the investigation of factors influencing HbA1c.

The results indicate that the linear regression presented a high coefficient of determination (R^2 = 93.83%), suggesting a good fit to the data. The high value of R2 is due to the regression model without the regression constant, which is not indicated, because the regression constant provides an important adjustment of the error and increases the external validity (reference). However, to increase the chance of identifying variables with potential for Fuzzy Logic, we opted for the use of regression without the constant. Thus, we do not recommend that the R2 values presented be used for the quantification of the cause-effect relationship or for the analysis of the quality of the regression model.

The limitation of the regression model without the constant can be seen in the results of table 7, in which the accuracy in predicting HbA1c levels was limited, correctly identifying only 40% of the cases, which suggests a low external validity of the model. This is seen in the studies by Schisterman et al. (2006), who discuss the limitations of regression models and propose methods to deal with these values below the established limits. This reflects the complexity of the interactions between the factors studied (time of diagnosis and total cholesterol) and other potential influencers that may not have been captured by the linear model. Previous studies indicate that variables such as adherence to treatment, physical activity, and daily glycemic control can play significant roles in HbA1c control and were not included in this model. For example, Abdollahi et al. (2022) demonstrated that self-control, including treatment adherence, physical activity, and daily glycemic monitoring, has a direct impact on HbA1c levels. In addition, Gomes and Negrato (2016) showed that adherence to insulin therapeutic regimens in patients is essential to maintain adequate glycemic control.

Fuzzy models were developed to translate the complex relationships between input and output variables. Despite its flexibility to deal with uncertainties, the results demonstrated limitations in terms of accuracy, with Rule 1 correctly predicting only 20% of cases and Rule 2 correctly predicting only 30% of cases. As discussed by Zimmermann

Revista ARACÊ
ISSN: 2358-2472

(2021), the limitations may be related to the choice of pertinence functions and fuzzy rules, which play a central role in modeling. Sivanandam et al. (2007) also point out that the simplicity of the criteria defined for linguistic variables can negatively influence the effectiveness of the model, suggesting greater attention to these aspects in order to generate better results.

In this sense, studies applied in various clinical contexts are increasingly careful in certain choices for the prediction of glycated hemoglobin. For example, in the study by Kalpana and Kumar (2019), the idea was to incorporate clinically relevant variables for glycated prediction, combining multiple fuzzy rules and pertinence functions, to achieve meaningful results. Bressan et al. (2020), in their studies of the use of Fuzzy Logic, the premises were related to the combination of decision tree tools. Therefore, some limitations were observed, even based on the help of specialists. The predictions based on the simplicity of the criteria defined for the linguistic variables and pertinence functions affected the study, making the integration of variables, functions and rules more robust.

Finally, Araujo et al. (2019), considered a wide range of variables for the prediction, including family history to eating habits, to have a more comprehensive fuzzy rule creation and fuzzification process, based on symptoms and risk factors. Comparing this study with those mentioned, it is possible to perceive the use of some ideas of each of them. From Kalpana and Kumar (2019), the idea of incorporating only relevant variables. De Bressan et al. (2020), with the help of experts, created the fuzzy rules. And Araujo et al. (2019), symptoms and, especially, risk factors were taken into account.

Analyzing the variables, the models identified that time of diagnosis and total cholesterol were the main predictors of HbA1c. Increased time to diagnosis is associated with a higher risk of inadequate glycemic control, possibly due to changes in treatment adherence and immune response over time (reference). Similarly, elevated total cholesterol reflects potential insulin resistance and higher cardiovascular risk, negatively influencing HbA1c levels. (MELO et al., 2023; MENEGUCCI et al., 2023)

The findings reinforce the importance of a multidimensional approach in the management of DM1. The Fuzzy systems can be improved to include factors such as treatment adherence and psychosocial support. Although the results indicated limitations of the model, the results suggest the need for attention to lipid profile control strategies and continuing education, especially in patients with longer time of diagnosis. In addition, information related to the use of technologies such as CGM (continuous glucose monitoring) and hybrid systems can provide more detailed data to feed the models.

As previously mentioned, it is necessary to consider the complexity of the interaction between the factors, since both linear regression and fuzzy models did not capture behavioral and environmental factors that influence HbA1c, suggesting the need to expand the variables studied. Thus, the improvement of models may depend on the integration of advanced techniques such as fuzzy neural networks and machine learning, which can improve forecasting by dealing with more complex datasets and interdependent variables.

CONCLUSION

The results demonstrate that, although linear regression showed a good statistical fit, fuzzy logic offers greater potential to incorporate the complexity inherent in glycemic control in children and adolescents with DM1. The combination of the two methodologies, combined with additional variables, can contribute to more effective and personalized clinical interventions. Thus, the individualization of the treatment and the clinical experience of the therapist must be taken into account for the best clinical decision-making. However, the identification of influencing factors, even if discretely, by mathematical models can contribute to a more assertive clinical practice, which collaborates with the control of the disease and its complications.

ACKNOWLEDGMENTS

I would like to thank the University of Marília (UNIMAR), the Postdoctoral Program of the Faculty of Science and Engineering of Tupã of the São Paulo State University "Júlio de Mesquita Filho" – UNESP, and the National Council for Scientific and Technological Development (CNPq) for granting the Productivity Scholarship to the author Luís Roberto Almeida Gabriel Filho (Proc. 317061/2023-2).

Revista ARACÊ
ISSN: 2358-2472

REFERENCES

1. Abdollahi, B., Ghaffari, M., Doroste, A. P., & Rakhshanderou, S. (2022). The role of diabetes-related self-management in type II diabetes mellitus and impact on HbA1c. Journal of Diabetes Metabolism and its Complications. Available at: http://dx.doi.org/10.31487/j.JDMC.2022.01.01. Accessed on: December 18, 2024.

2. Araujo, M. D. A., Moreira, L. Y. M. R., de Araujo, A. V. C., & de Brito, R. X. (2019). A lógica fuzzy como recurso auxiliador na predição da diabetes do tipo 2. In Escola Regional de Computação do Ceará, Maranhão e Piauí (ERCEMAPI), 7., São Luís/MA. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 17-24.

3. Bauman, C. D., Bauman, J. M., Mourão, D. M., Pinho, L. de, Brito, M. F. S. F., Carneiro, A. L. G., Silveira, M. F., & Silva, C. S. de O. e. (2020). Dyslipidemia prevalence in adolescents in public schools. Revista Brasileira de Enfermagem, 73(3), e20180523. Available at: https://doi.org/10.1590/0034-7167-2018-0523. Accessed on: December 18, 2024.

4. Bjornstad, P., Donaghue, K. C., & Maahs, D. M. (2018). Macrovascular disease and risk factors in youth with type 1 diabetes: Time to be more attentive to treatment? The Lancet Diabetes & Endocrinology, 6(10), 809–820. Available at: https://doi.org/10.1016/S2213-8587(18)30035-4. Accessed on: December 18, 2024.

5. Bressan, G. M., Azevedo, B. C. F. de, & Souza, R. M. de. (2020). A fuzzy approach for diabetes mellitus type 2 classification. Brazilian Archives of Biology and Technology, 63, e20180742. Available at: https://doi.org/10.1590/1678-4324-2020180742. Accessed on: December 18, 2024.

6. Castro, A. P. P., & Novaes, J. F. de. (2019). Increased ApoB/ApoA1 ratio is associated with excess weight, body adiposity, and altered lipid profile in children. Jornal de Pediatria (Versão em Português), 95(2), 238–246. Available at: https://doi.org/10.1016/j.jpedp.2018.03.007. Accessed on: December 18, 2024.

7. Chiang, J. L., Maahs, D. M., Garvey, K. C., Hood, K. K., Laffel, L. M., Weinzimer, S. A., Wolfsdorf, J. I., & Schatz, D. (2018). Type 1 diabetes in children and adolescents: A position statement by the American Diabetes Association. Diabetes Care, 41(9), 2026–2044. Available at: https://doi.org/10.2337/dci18-0023. Accessed on: December 18, 2024.

8. Cox, L. A. Jr. (2021). Toward practical causal epidemiology. Global Epidemiology, 3, 100065. Available at: https://doi.org/10.1016/j.gloepi.2021.100065. Accessed on: December 18, 2024.

9. Departamento Científico de Endocrinologia SBP. (2020a, May). Dislipidemia na criança e no adolescente - Orientações para o pediatra. Guia Prático de Atualização SBP, 8. Available at: https://www.sbp.com.br/fileadmin/user_upload/22336c-GPA_-_Dislipidemia_Crianca_e_Adoles.pdf. Accessed on: December 18, 2024.

Revista
ARACÊ
ISSN: 2358-2472

10. Elhenawy, Y. I., Abdelmageed, R. I., Zaaafar, D. K., & Abdelaziz, A. W. (2022). Adherence to insulin therapy among children with type 1 diabetes: Reliability and validity of the Arabic version of the 4-item Morisky medication adherence scale. Patient Preference and Adherence, 16, 1415–1421. Available at: https://doi.org/10.2147/PPA.S341061. Accessed on: December 18, 2024.

11. Fagundes Melo, R., et al. (2024). Investigating the incidence of dyslipidemia among Brazilian children and adolescents diagnosed with type 1 diabetes mellitus: A cross-sectional study.

12. Ferranti, S. D. de, & Newburger, J. W. (2023a). Dyslipidemia in children and adolescents: Definition, screening, and diagnosis. UpToDate. Available at: https://www.uptodate.com/contents/dyslipidemia-in-children-and-adolescents-definition-screening-and-diagnosis. Accessed on: December 18, 2024.

13. Ferranti, S. D. de, & Newburger, J. W. (2023b). Dyslipidemia in children and adolescents: Definition, screening, and diagnosis. UpToDate. Available at: https://www.uptodate.com/contents/dyslipidemia-in-children-and-adolescents-definition-screening-and-diagnosis. Accessed on: December 18, 2024.

14. Gomes, M. B., & Negrato, C. A. (2016). Adherence to insulin therapeutic regimens in patients with type 1 diabetes: A nationwide survey in Brazil. Diabetes Research and Clinical Practice, 120, 47–55. Available at: https://doi.org/10.1016/j.diabres.2016.07.011. Accessed on: December 18, 2024.

15. Hasan, N., Islam, M. S., & Hayat, M. T. (2024). A fuzzy logic approach for predicting gestational diabetes mellitus using risk factors. In International Conference on Electrical Engineering and Information & Communication Technology, 6th, Dhaka. Proceedings [...]. Dhaka: IEEE, 1479-1484.

16. Holm, H., Kennback, C., Laucyte-Cibulskiene, A., Nilsson, P. M., & Jujic, A. (2024). The impact of prediabetes and diabetes on endothelial function in a large population-based cohort. Blood Pressure, 33(1). Available at: https://doi.org/10.1080/08037051.2023.2298309. Accessed on: December 18, 2024.

17. Izar, M. C. de O., Fonseca, F. A. H., Faludi, A. Á., Araújo, D. B. de, Valente, F., & Bertoluci, M. (2022). Manejo do risco cardiovascular: Dislipidemia. In Diretriz Oficial da Sociedade Brasileira de Diabetes (1st ed., pp. 1–28). Available at: https://doi.org/10.29327/557753.2022-19. Accessed on: December 18, 2024.

18. Kalpana, M., & Kumar, A. V. S. (2011). Fuzzy expert system for diabetes using fuzzy verdict mechanism. International Journal of Advanced Networking and Applications, 3(2), 1128.

19. Lind, M., et al. (2014). Glycemic control and excess mortality in type 1 diabetes. New England Journal of Medicine, 371(21), 1972-1982.

20. Massad, E., Ortega, N. R. S., & Silveira, P. S. P. (2004). Métodos quantitativos em medicina. São Paulo: Manole.

Revista ARACÊ
ISSN: 2358-2472

21. Melo, K. F. S., et al. (2023). Tratamento do diabetes mellitus tipo 1 no SUS. In Diretriz da Sociedade Brasileira de Diabetes. São Paulo: Conectando Pessoas, 1–29.

22. Menegucci, T., et al. (2023). The influence of body fat and lean mass on HbA1c and lipid profile in children and adolescents with type 1 diabetes mellitus. Diseases, 11(4), 125.

23. Monteiro, A., et al. (2022). Lipid profile of children and adolescents with type 1 diabetes mellitus. Residência Pediátrica, 12(2), 1-6.

24. Onis, M., et al. (2007). Development of a WHO growth reference for school-aged children and adolescents. Bulletin of the World Health Organization, 85(9), 660-667.

25. Rodacki, M., et al. (2022). Classificação do diabetes. In Diretriz Oficial da Sociedade Brasileira de Diabetes. São Paulo: Conectando Pessoas, 1–28.

26. Rossaneis, M. A., et al. (2019). Fatores associados ao controle glicêmico de pessoas com diabetes mellitus. Ciência & Saúde Coletiva, 24(3), 997-1005.

27. Sahoo, M., & Chakraverty, S. (2024). Possibilistic linear and quadratic regression analysis for fuzzy random data and application. Journal of Statistics and Management Systems, 27(6), 1095-1115.

28. Schisterman, E. F., et al. (2006). The limitations due to exposure detection limits for regression models. American Journal of Epidemiology, 163(4), 374-383.

29. Shen, C. Y., et al. (2024). Advanced glycation end-products acting as immunomodulators for chronic inflammation, inflammaging and carcinogenesis in patients with diabetes and immune-related diseases. Biomedicines, 12(8), 1699.

30. Sibal, L., et al. (2010). Friedewald equation underestimates low-density lipoprotein cholesterol at low concentrations in young people with and without Type 1 diabetes. Diabetic Medicine, 27(1), 37-45.

31. Silva, T. P. R., et al. (2023). Total cholesterol and low-density lipoprotein alterations in children and adolescents from Brazil: A prevalence meta-analysis. Archives of Endocrinology and Metabolism, 67(1), 19-44.

32. Sivanandam, N., Sai, S., & Deepa, S. N. (2007). Introduction to fuzzy logic using MATLAB. Berlin: Springer.

33. Smith, J. K., Johnson, L., & Brown, R. (2020). Impact of glycemic control on quality of life in pediatric patients with type 1 diabetes. Journal of Pediatric Endocrinology and Metabolism, 33(5), 713-721.

34. Sociedade Brasileira de Diabetes. (2019). Diretrizes da Sociedade Brasileira de Diabetes 2019-2020. São Paulo: Clannad.

35. Wang, N., & Zhang, C. (2024). Oxidative stress: A culprit in the progression of diabetic kidney disease. Antioxidants, 13(4), 455.

36. Yamagishi, S., et al. (2015). Advanced glycation end products: A molecular target for vascular complications in diabetes. Molecular Medicine, 21, S32-S40.

37. Zimmermann, H. J. (2001). Fuzzy set theory – and its applications. Boston: Springer.