Proceedings of ICBioMed 2025 Symposium: AI for Healthcare: Advanced Medical Data Analytics and Smart Rehabilitation
DOI: 10.54254/2753-8818/2025.AU25462

Analysis of Factors Influencing Overweight in Adolescents Based on Logistic Regression Modeling

Lai Wei

Massey Institute, Nanjing University of Finance and Economics, Nanjing, China
weilai931@163.com

Abstract. Since the 21st century, adolescent overweight has emerged as a critical global public health challenge. Understanding its determinants is essential for effective intervention strategies. Using logistic regression models, this study analyzes nationally representative 2019 data from Colombia, Mexico, and Peru to identify key factors influencing adolescent overweight. This article examines 13 factors that influence overweight in adolescents and draws the following conclusions. Three factors showed significant positive associations with overweight: family history of obesity, self-monitoring of caloric intake, and passive transportation modes (e.g., car use). Conversely, higher meal frequency, physical activity levels, and reduced screen time were negatively associated with overweight. No significant associations were found for gender, high-calorie food consumption, vegetable intake, smoking, water consumption, or alcohol use. This paper proposes four evidence-based recommendations: (1) targeted awareness campaigns on overweight risks, (2) dietary guidelines tailored to adolescents, (3) promotion of physical activity, and (4) cross-national adaptation of successful obesity prevention programs.

Keywords: overweight, adolescents, logistic regression

1. Introduction

Globally, adolescent overweight has reached epidemic proportions, posing a critical public health challenge. According to the World Health Organization (WHO), over 390 million children and adolescents aged 5–19 were overweight or obese in 2022—a stark increase from 8% prevalence in 1990 to 20% in 2022. Notably, this trend is no longer confined to high-income nations; low- and middle-income countries (LMICs), particularly in Latin America, now face accelerating rates of adolescent overweight. Adolescent overweight is linked to immediate physical health risks and psychosocial consequences. These effects extend into adulthood, exacerbating chronic disease burdens and socioeconomic disparities. Given adolescence's pivotal role in lifelong health trajectories, addressing overweight during this stage is urgent. While prior studies have identified dietary and lifestyle factors in adult obesity, adolescent-specific drivers—especially in LMICs—remain understudied. This paper examines nationally representative 2019 data from Colombia, Mexico, and Peru to: Quantify associations between behavioral, dietary, and socioeconomic factors and adolescent overweight. Identify region-specific determinants that diverge from global trends. Propose targeted interventions for LMIC contexts.

© 2025 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-8818/2025.AU25462

Using logistic regression modeling, this paper analyzes 15 potential influencing factors, including family history, physical activity, and screen time. Our findings aim to inform scalable, evidence-based policies to curb adolescent overweight in rapidly transitioning economies.

2. Literature review

The first category of literature relevant to this article focuses on exploring the issue of overweight and obesity in adolescents. Overweight and obesity in adolescents are influenced by a variety of factors. Wei Jia et al. [1] found that air pollution has become an important risk factor for overweight and obesity, and its concentration is positively correlated with the risk of overweight and obesity. Renato Mikio Moriya et al. [2] further found that adverse childhood experiences may also be a key risk factor for triggering overweight and obesity in adolescents. In addition, Wei Jia et al. [3] showed that the level of exposure to secondhand smoke is the other important factor influencing overweight and obesity in adolescents. In particular, adolescents who are frequently exposed to secondhand smoke in their homes are at significantly increased risk for overweight and obesity. Other articles talk about the relationship between overweight and obesity and physical and mental illness in adolescents. Specific to physical problems, Youn Huh and Hye Soon Park [4] noted that overweight and obese adolescents exhibit a higher risk of cardiometabolic clustering. In terms of psychological problems, Anna Lewandowska et al. [5] found a significant association between adolescents' weight status (normal, overweight and obese) and their well-being, with normal weight adolescents having a higher level of well-being.

The second category of literature related to this paper focuses on the exploration of logistic regression modeling in medical applications. Logistic regression models are widely used in the field of medical research, especially in oncology. Xiaoyuan Wang et al. [6] found by logistic regression analysis that factors such as hypertension, white blood cell count, and platelet count have an important effect on tumor size in colorectal cancer. Another study by Yulin Shi et al. [7] analyzed the characteristics of tongue diagnosis data of benign and malignant lung nodules based on the logistic regression method and established an early warning model for lung cancer risk. Furthermore, in other areas of medicine, Kale Kawu Kale & Babagana Modu [8] explored the relationship between diabetes mellitus and its associated risk factors using logistic regression analysis, stating that blood glucose level and age are the most important predictors of diabetes mellitus. Yijun Chen et al. [9] investigated the relationship between dietary fiber intake (DFI) and sleep disorders through logistic regression modeling, finding that increasing DFI has a protective effect on reducing the risk of sleep disorders. Yongliang Feng et al. [10], on the other hand, based on logistic regression analysis, found that gender, dosage, and number of dialysis sessions are factors affecting hepatitis B vaccination in HD patients.

Based on the above literature, this paper may have the following two contributions. First, in terms of topic selection, few studies have focused on the use of logistic regression modeling to explore the influencing factors of adolescent overweight. This paper takes a sample data of individual adolescent overweight and their influencing factors in Mexico, Peru, and Colombia to explore in depth the correlation between each influencing factor and adolescent overweight. Secondly, in the context of contemporary society, the problem of overweight among adolescents has gradually become a hot topic of social concern as the rates of overweight and obesity continue to climb. Therefore, the conclusions and recommendations of this paper not only have important theoretical significance and practical value, but also provide useful cases and solutions for other developing countries.

DOI: 10.54254/2753-8818/2025.AU25462

3. Characterizing facts on overweight among adolescents in Mexico, Peru and Colombia

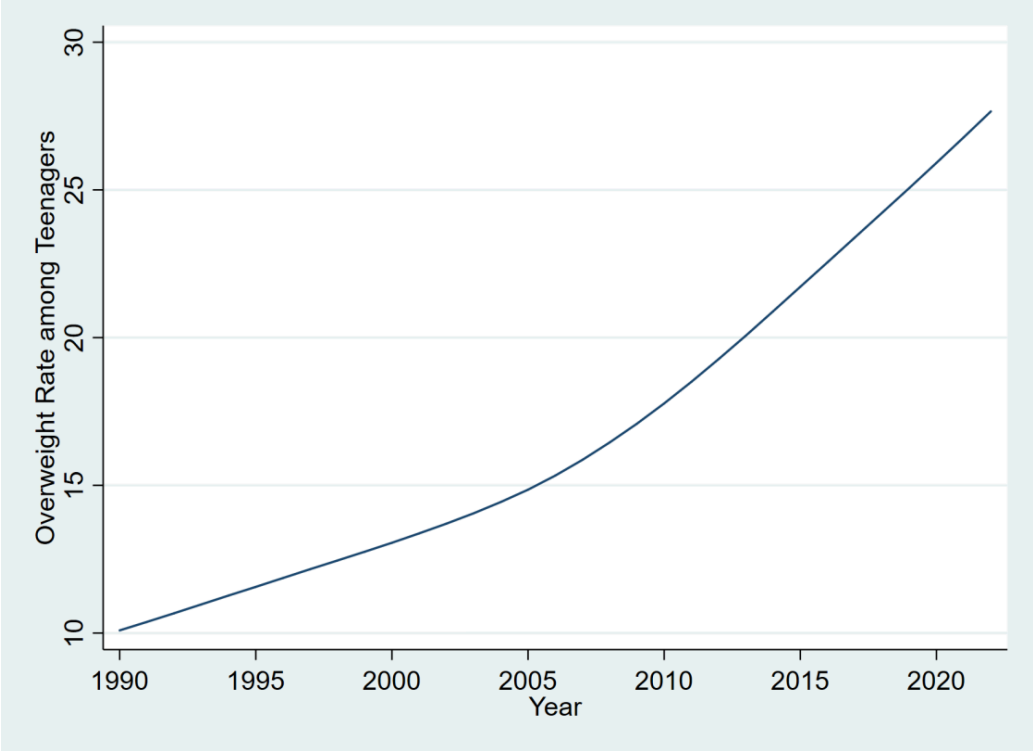

Figure 1. Trends in overweight rates among Colombian adolescents

Figure 1 illustrates the trend in overweight prevalence among adolescents aged 5-19 years in Colombia from 1990 to 2022. Analyzed in Figure 1, the trend shows that the rate of overweight shows a yearly increase. As of 2022, that overweight rate has risen nearly threefold, reflecting the continuing deterioration of overweight among Colombian adolescents. From the point of view of the rate of development, the rate of growth of overweight among Colombian adolescents accelerated significantly after 2005, to approximately twice the rate prior to 2005. It can be seen that the problem of overweight adolescents has become increasingly prominent in the twenty-first century and has become a public health challenge in Colombia that requires urgent attention and resolution.

DOI: 10.54254/2753-8818/2025.AU25462

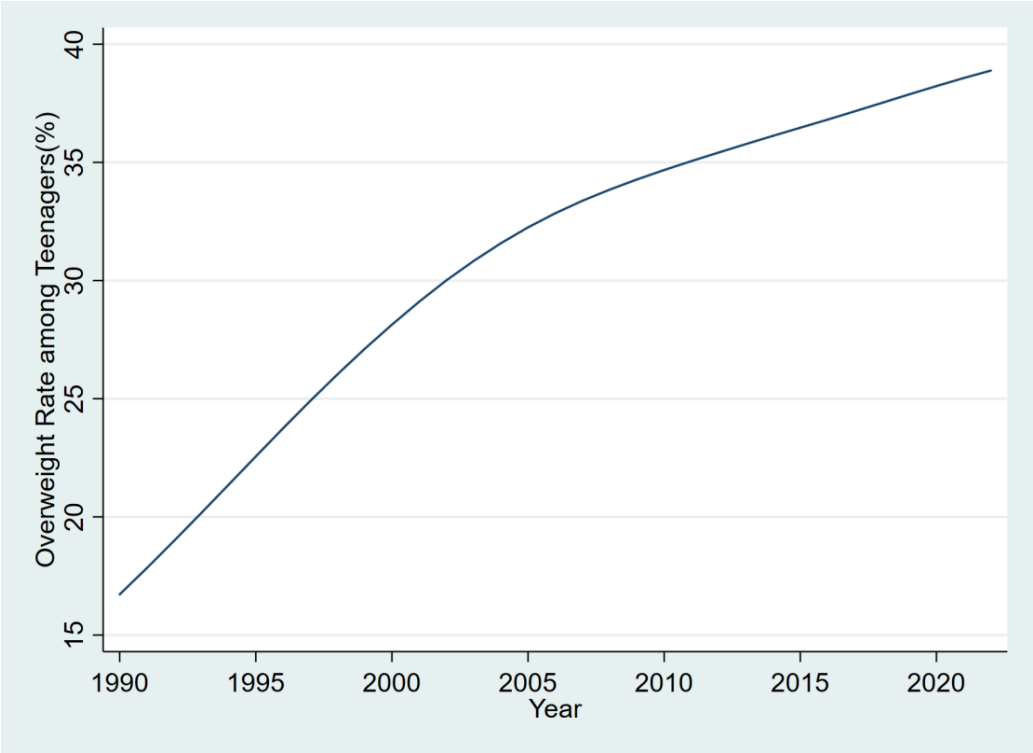

Figure 2. Trends in overweight rates among Mexican adolescents

Figure 2 depicts the trend in overweight prevalence among adolescents aged 5-19 years in Mexico between 1990 and 2022. As can be seen in Figure 2, in terms of trends, the rate of overweight among Mexican adolescents has been rising year after year, with an increase of almost 2.5 times between 1990 and 2020. In terms of the rate of development, Mexico's growth rate slowed down significantly after 2005, which may indicate that the Mexican government has taken effective measures to control and ameliorate the problem of overweight adolescents. However, the initial overweight rate in Mexico in 1990 was already over 15 per cent, reaching the threshold set by the World Health Organization for the need for systematic national intervention. This shows that the problem of overweight in Mexico has been considerable since 1990. Thus, despite the slowdown in recent years, the problem of overweight adolescents in Mexico can not be ignored.

DOI: 10.54254/2753-8818/2025.AU25462

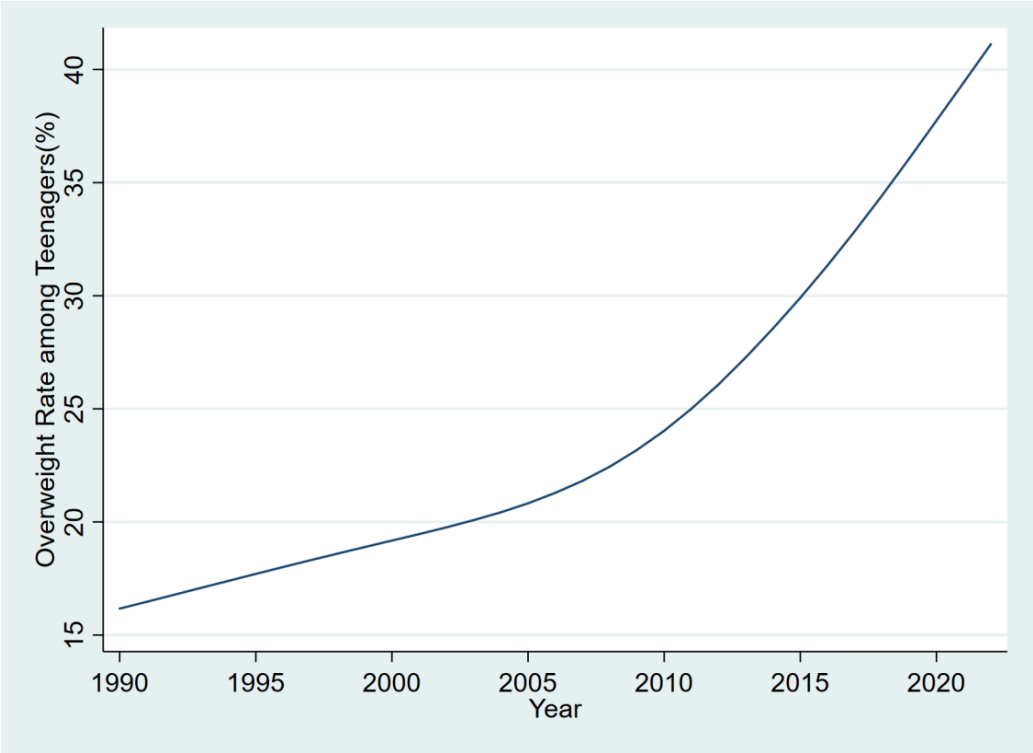

Figure 3. Trends in overweight rates among Peruvian adolescents

Figure 3 illustrates the escalating prevalence of overweight among Peruvian adolescents aged 5–19 from 1990 to 2022. The data reveal two critical trends: Persistently High Baseline: Peru's adolescent overweight rate exceeded 15% as early as 1990, surpassing the WHO's early warning threshold for public health risks. Recent Acceleration: The growth rate nearly doubled after 2009, with the 2009–2022 period showing a sharper increase compared to 1990–2009. This suggests a rapid deterioration in recent decades, likely driven by urbanization, dietary shifts, or socioeconomic changes (cite relevant studies if available). While Colombia, Mexico, and Peru all exhibited adolescent overweight rates exceeding 20% by 2022 (meeting WHO's "high prevalence" classification), Peru's trajectory is distinct: Earliest Onset: Peru was the only country where rates surpassed 20% before 1995, indicating a prolonged and unmitigated public health challenge. Nonlinear Progression: In contrast to Colombia and Mexico, Peru's overweight rates not only increased but accelerated over time, implying compounding systemic drivers (e.g., lack of policy intervention, cultural norms around diet). Peru's historical data underscore the urgency of: Primary Prevention: Targeting younger cohorts to disrupt the intergenerational cycle of obesity. Structural Interventions: Addressing food environments (e.g., junk food marketing regulations) and urban planning (e.g., walkability).

4. Research design

4.1. Variable selection

This paper uses 2019 adolescent overweight data from Colombia, Mexico, and Peru provided by the UCI Machine Learning Repository for analysis. The data contains 13 variables representing factors that influence overweight in adolescents, and the response variable is whether or not the adolescent is overweight.

DOI: 10.54254/2753-8818/2025.AU25462

4.2. Model design

The research model in this paper is a logistic regression model, specifically as shown in Table 1. The descriptive statistics of this paper are shown in Table 2.

Table 1. Interpretation of model variables

Variable	Coding	explanation
Overweight	0,1	whether overweight or not, 0 is not overweight
Gender	0,1	0 is female
Family_History	0,1	family history of obesity, 0 is no family history
FAVC	0,1	frequent consumption of high-calorie foods, 0 is infrequent
FCVC	1-3	frequency of eating vegetables in meals1 is never, 2 is sometimes, 3 is always
NCP	1-4	number of main meals per day
CAEC	1-4	consumption of food between meals1 no, 2 sometimes, 3 frequently, 4 always
SMOKE	0,1	whether smoking or not, 0 is no
CH2O	1-3	daily water intake1 is less than 1L, 2 is between 1 and 2L, 3 is more than 2L
SCC	0,1	self-monitoring of caloric intake, 1 is yes
FAF	0-3	physical activity frequency0 is not have, 1 is 1 or 2 days, 2 is 2 or 4 days, 3 is 4 or 5 days
TUE	0-2	time using electronic devices1 is 0-2 hours, 2 is 2-5 hours, 3 is more than 5 hours
CALC	1-3	alcohol consumption1 is not drink, 2 is sometimes, 3 is frequently
MTRANS	1-5	mode of transportation1 is automobile, 2 is bike, 3 is motorbike, 4 is public_transportation, 5 is walking

Table 2. Descriptive statistical analysis

Variable	Obs	Mean	Std.Dev.	Min	Max
Overweight	537	0.482	0.500	0	1
Gender	537	0.460	0.499	0	1
Family_history	537	0.655	0.476	0	1
FAVC	537	0.844	0.364	0	1
FCVC	537	2.389	0.622	1	3
NCP	537	2.790	0.859	1	4
CAEC	537	2.199	0.480	1	4
SMOKE	537	0.006	0.075	0	1
CH2O	537	1.929	0.668	1	3
SCC	537	0.112	0.315	0	1
FAF	537	1.242	0.877	0	3
TUE	537	0.899	0.656	0	2
CALC	537	1.676	0.535	1	3
MTRANS	537	3.739	0.934	1	5

DOI: 10.54254/2753-8818/2025.AU25462

5. Empirical studies

5.1. Univariate analysis

Table 3. Statistical tests for individual predictors

Explanatory Variable	Co-Efficient β	Standard Error	P-Value	Wald Test Wj	Odds Ratio OR
Gender	-0.182	0.251	0.468	-0.730	0.834
Family_history	3.460	0.337	<0.001	10.26	31.82
FAVC	-0.534	0.351	0.128	-1.520	0.586
FCVC	0.191	0.218	0.381	0.880	1.210
NCP	-0.371	0.173	0.032	-2.140	0.690
CAEC	-2.064	0.319	<0.001	-6.480	0.127
SMOKE	1.454	1.861	0.435	0.780	4.280
CH2O	0.268	0.197	0.174	1.360	1.307
SCC	1.096	0.415	0.008	2.640	2.992
FAF	-0.446	0.158	0.005	-2.830	0.640
TUE	-0.626	0.200	0.002	-3.120	0.535
CALC	0.408	0.242	0.092	1.690	1.504
MTRANS	0.954	0.176	<0.001	5.410	2.596

Table 3 demonstrates the results of the logistic regression statistical tests for individual factors. According to Table 3, out of the 13 influencing factors, Gender, FAVC, FCVC, SMOKE, CH_2O, and CALC are not significant at the 5% level of significance, which may indicate that they do not have a significant effect on the problem of overweight among adolescents. The other six influences are significant at 5% level of significance. The regression coefficients of family history of obesity, SCC and MTRANS were 3.460 ($p<0.001$), 1.096 ($p=0.008$), and 0.954 ($p<0.001$), respectively. This means they are significantly positive at 5% level of significance, and had a positive effect on adolescent overweight. And the regression coefficients of CAEC, NCP, FAF, and TUE are -2.064 ($p<0.001$), -0.371 ($p=0.032$), -0.446 ($p=0.005$), and -0.626 ($p=0.002$), separately, which were significantly negative at the 5% level of significance. So they are the negative influences on adolescent overweight.

5.2. Overall model evaluation and goodness-of-fit test

Table 4. Overall model evaluation and goodness-of-fit statistics

Test	Categories	X^2	DoF	P-value
Overall Model Evaluation	Likelihood Ratio Test	307.34	13	0.0000
	Wald Test	156.27	13	0.0000
Goodness of Fit Test	Hosmer and Lemeshow Test	5.19	3	0.1585

In order to test whether the logistic regression model is valid or not, this paper chooses to test the validity of the model by likelihood ratio test and Wald test. According to Table 4, the results of both

DOI: 10.54254/2753-8818/2025.AU25462

tests were 307.34 ($p<0.001$) and 156.27 ($p<0.001$), which are significant at 5% level of significance, indicating that the model for testing the 13 factors influencing overweight among adolescents is more effective than the null model. In addition, Table 4 demonstrates the Hosmer and Lemeshow goodness-of-fit test. This statistical test measures the correspondence between actual and predicted values, where the dependent variable is the rate of overweight adolescents. This test assumes that the null hypothesis is that there is no difference between predicted and actual values and alternative hypothesis is that there is a difference between predicted and actual values. From Table 4, the p-value of the Hosmer and Lemeshow test is 0.1585, which is greater than 0.05, so there is no enough reason to reject the null hypothesis. Therefore, the difference between the actual and expected values is not significant, indicating that the model fits the data well.

5.3. Model explanatory power

Table 5. Assessment of the explanatory power of the model

Test	Cox & Snell R^2	Nagelkerke R^2
Value	0.436	0.581

In order to test whether the logistic regression model in this paper can well explain or account for changes in the dependent variable adolescent overweight, the Nagel Kerk R2 and Cox & Snell R2 tests are used in this paper. Table 5 shows that Cox & Snell R2 value is 0.436 and Nagel Kerk R2 value is 0.581, meaning that the fitted model is able to explain or account for 58.1% of the variation in the dependent variable. This suggests that the model in this paper is good.

6. Summary

6.1. Main conclusions

Based on the above empirical results, this study finds that differences in whether adolescents had a family history of obesity, whether they monitored their caloric intake, and the mode of transportation they used for travel are positively associated with whether they were overweight; whereas, the frequency of food intake between meals, the number of main meals consumed per day, the frequency of physical activity, and the number of hours spent using electronic devices are negatively associated with whether or not an adolescent was overweight. Furthermore, analysis of the data in this paper does not reveal significant associations between the following seven factors and adolescent overweight: gender, high-calorie food intake, frequency of vegetable intake, whether or not one smokes, amount of water consumed, and frequency of alcohol intake.

6.2. Policy recommendations

The policy recommendations of the above conclusions are the following four. Firstly, government should strengthen its publicity work on the hazards of overweight among adolescents, promote public health education and enhance adolescents' own health awareness. Secondly, regulation and guidance on the dietary health of young people should be reinforced, the dietary environment should be improved and a healthy dietary culture needs to be promoted. Thirdly, government should increase the number of physical activity facilities in the community and in schools, and encourage adolescents to increase their physical activities and reduce their time spent on electronic devices and

DOI: 10.54254/2753-8818/2025.AU25462

sedentary behaviors. Fourthly, the government could learn from the experience of other countries in effectively alleviating the problem of overweight among adolescents and develop measures to prevent and control overweight among adolescents that are in line with the countries' own national conditions.

References

[1] Wei, J., Luo, J. Y., Chen, Y. H., Wang, F., Yang, F., Li, Y. M., & Luo, M. Y. (2025). Long-term exposure to outdoor air pollution correlated with overweight/obesity in children and adolescents: A cross-sectional real-world study. Journal of Environmental Management, 387, 125912.

[2] Moriya, R. M., Oliveira, C. E. C., Reiche, E. M. V., Passini, J. L. L., & Nunes, S. O. V. (2024). Association of adverse childhood experiences and overweight or obesity in adolescents: A systematic review and network analysis. Obesity Reviews, 25(11), e13809.

[3] Wei, J., Luo, J. Y., Chen, Y. H., Wang, F., Tan, Y. Q., Luo, M. Y., & Li, X. J. (2025). [The association between secondhand smoke exposure and overweight/obesity among children and adolescents]. Zhonghua Yu Fang Yi Xue Za Zhi [Chinese Journal of Preventive Medicine], 59(1), 69–75.

[4] Huh, Y., & Park, H. S. (2024). Associations of overweight and obesity with cardiometabolic risk factor clusters among Korean adolescents. Scientific Reports, 14(1), 3581.

[5] Lewandowska, A., Rudzki, G., Lewandowski, T., Bartosiewicz, A., Próchnicki, M., Stryjkowska Góra, A., ... & Pavlov, S. (2025). Overweight and obesity among adolescents: Health-conscious behaviours, acceptance, and the health behaviours of their parents. BMC Public Health, 25(1), 418.

[6] Wang, X., Li, J., Zhang, X., & Lu, M. (2024). Exploring the risk factors of tumor size in colorectal cancer based on multifactorial logistic regression model. Chinese Medical Science, 14(24), 139–143.

[7] Shi, Y., Yi, C., Liu, J., Liu, L., & Xu, J. (2024). An early warning model of lung cancer risk based on logistic regression of tongue features. Chinese Journal of Traditional Chinese Medicine Information, 31(10), 149–156.

[8] Kale, K. K., & Modu, B. (2025). Assessing diabetes risk factors using logistic regression: A Kaggle-based study. Asian Journal of Probability and Statistics, 27(4), 132–141.

[9] Chen, Y., Zhao, Z., Ding, W., Zhou, Z., & Xiao, M. (2024). Association between dietary fiber intake and sleep disorders: Based on the NHANES database. Brain and Behavior, 14(11), e70123.

[10] Feng, Y., Wang, J., Shao, Z., Chen, Z., Yao, T., Dong, S., ... & Liang, X. (2021). Predicting related factors of immunological response to hepatitis B vaccine in hemodialysis patients based on integration of decision tree classification and logistic regression. Human Vaccines & Immunotherapeutics, 17(9), 1-7.