International Journal of Agriculture and Food Sciences Research
ISSN: 3006-7685 (Print), ISSN: 3006-3736 (Online) | Volume 3, Number 1, Year 2025
DOI: https://doi.org/10.62051/ijafsr.v3n1.05
Journal homepage: https://wepub.org/index.php/IJAFSR/index

Research on the Influence Mechanism and Moderating Effect of Food Safety Awareness on Health Behaviors Based on the Logistic Model

Tianyi Chen [1], Yuhan Cao [2,*], Hanxiao Qi [3], Dongyang Jiang [1], Shanying Zhao [4]

[1] School of Mathematics and Statistics, Henan University of Technology, Zhengzhou, China

[2] School of International Education, Henan University of Technology, Zhengzhou, China

[3] School of Economics and Trade, Henan University of Technology, Zhengzhou, China

[4] School of Management, Henan University of Technology, Zhengzhou, China

*Corresponding Author: 17730870926@163.com

ABSTRACT

Against the backdrop of the steady advancement of the Rural Revitalization Strategy, the health level of rural residents has become a key indicator for measuring development achievements. This study uses descriptive statistical analysis and the Logistic regression model to explore the influence mechanism and moderating effect of food safety awareness on healthy behaviors. The results show that although rural residents have a relatively high level of concern about food safety, their confidence in the food in the market is low. In terms of the dietary structure, the proportion of staple food expenditure has decreased from 45% to 38%, the intake of animal protein is 20% lower than that of urban residents, the consumption of deeply processed foods has increased by 12%, and the overweight rate of BMI reaches 32%, which is higher than the national rural average. Logistic regression analysis shows that the odds ratio of food safety awareness is 1.05, and the significance P - value is less than 0.01. The diversity of information acquisition channels and the coverage rate of policy publicity have significant moderating effects. When the coverage rate of policy publicity increases from 50% to 80%, the probability of healthy behaviors increases by 0.56%. This study provides important data support and strategic references for promoting the healthy diet of rural residents and driving rural revitalization.

KEYWORDS

Food Safety Awareness; Healthy Behaviors; Logistic Model; Moderating Effect

1. INTRODUCTION

Under the backdrop of the steady advancement of the rural revitalization strategy, the health level of rural residents has become one of the key indicators for measuring the development achievements. With the improvement of living standards, significant changes have taken place in the dietary structure and consumption habits of rural residents. The importance of food safety and healthy eating has become increasingly prominent [1]. There is a close connection between the food consumption pattern and the health status [2]. A reasonable dietary structure and the selection of safe foods can effectively reduce the risk of chronic diseases and improve the quality of life. Therefore, it is of great significance to deeply explore the influence mechanism of rural residents' food safety awareness on their health behaviors, and then propose effective intervention paths, to promote the health of rural residents and drive rural revitalization.

Content from this work may be used under the terms of CC BY-NC 4.0 licence (https://creativecommons.org/licenses/by-nc/4.0/). Published by Warwick Evans Publishing.

The enhancement of food safety awareness helps rural residents pay more attention to the quality and safety of food, and be more cautious in the process of food selection, thus reducing the intake of unsafe foods and lowering health risks [3]. When rural residents are aware of the harm of high-salt and high-fat foods to their health, they will take the initiative to adjust their dietary structure, increase the intake of nutrients such as vegetables, fruits, and high-quality proteins, and improve their nutritional status. At the same time, healthy dietary behaviors can also improve the physical quality of rural residents, enhance their immunity, prevent the occurrence of chronic diseases, and further improve the overall quality of life of rural residents [4], providing a solid human resource guarantee for the economic development of rural areas.

This study focuses on the key issue of the influence mechanism and intervention paths of rural residents' food safety awareness on their health behaviors. It conducts an in-depth analysis of the survey data of rural residents in 50 counties (cities) in 12 provinces including Henan, Shandong, and Guangdong. In the data analysis stage, descriptive statistical analysis is used to depict the basic characteristics of rural residents' food consumption. At the same time, econometric analysis methods are applied to deeply explore the internal connection between food safety awareness and health behaviors, analyze the factors influencing the transformation from awareness to behavior, and propose targeted intervention strategies, providing strong data support for promoting the healthy diet of rural residents under the rural revitalization strategy.

2. RESEARCH METHODS AND MODEL CONSTRUCTION

2.1. Data Sources

The data of this study are derived from the dataset provided by the China Rural Revitalization Association on the food consumption habits, consumption structure, nutritional intake status of rural residents in 50 counties (cities) of 12 provinces including Henan, Shandong, and Guangdong, and their impacts on health status. In the selection of samples, the regional differences among the eastern, central, and western regions have been fully considered to ensure that the data are widely representative. There are a total of 3,800 valid samples, among which the population aged 20 to 60 accounts for 85%. The annual household income is divided into three levels: low, medium, and high. The main sources of income include farming, working, and doing business.

2.2. Descriptive Statistical Analysis

Descriptive statistical methods are used to calculate statistical quantities such as the mean, median, and standard deviation of rural residents' food safety cognition scores, the proportion of consumption expenditure on various types of foods, and nutritional health indicators (such as BMI, blood pressure, blood sugar, etc.), to present the basic characteristics of rural residents' food consumption and health status [5, 6]. By comparing the data of 2018 and 2022, the changing trends of the dietary consumption structure of rural residents are analyzed.

2.3. Logistic Regression Model

2.3.1. Linear Combination

Logistic regression is a classic model used to solve binary classification problems, which outputs categories through probability prediction and threshold judgment. The following is the specific model construction process [7, 8].

Logistic regression assumes that the linear combination of input features can form a decision boundary. Suppose we have n input features $x_1, x_2, \cdots, x_n$, and the corresponding parameters $\beta_0, \beta_1, \beta_2, \cdots, \beta_n$, then the linear combination can be expressed as following.

$$z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n \tag{1}$$

Here, β_i are the parameters that the model needs to learn, and x_i are the various feature values of the input data.

2.3.2. Sigmoid Probability Mapping

In order to transform the linear output z into a probability value representing that the sample belongs to the positive class we introduce the Sigmoid function. The expression of the Sigmoid function is as following.

$$g(z) = \frac{1}{1+e^{-z}} \tag{2}$$

The characteristic of the Sigmoid function is that it can map the input value z in any real number domain to the probability interval $[0, 1]$. When $z = 0, g(0) = 0.5$, indicating that the sample has the same possibility of belonging to the positive class and the negative class; when z is very large, $g(z)$ approaches 1, which means that the sample is very likely to belong to the positive class; when z is very small, $g(z)$ approaches 0, that is, the sample is most likely to belong to the negative class, as shown in Figure 1.

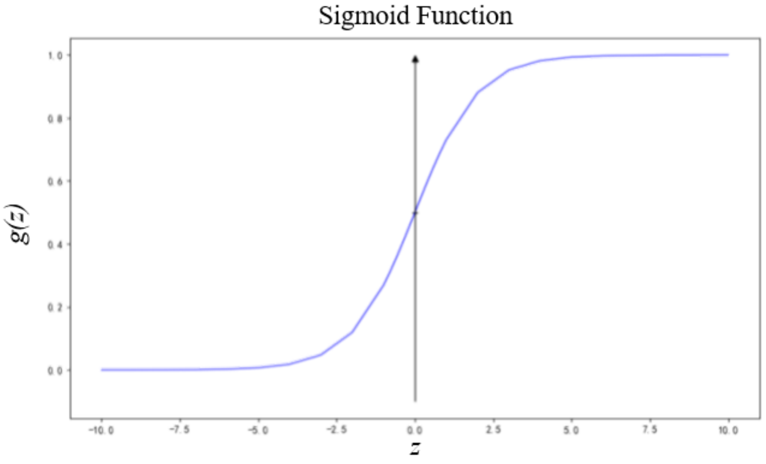

Figure 1. Approximate graph of the sigmoid function

2.3.3. Loss Function

The goal of constructing a logistic regression model is to make the probability predicted by the model as close as possible to the true label of the sample. In binary classification problems, logarithmic loss (also known as cross-entropy loss) is usually used to measure this difference, and its expression is following.

$$J(\beta) = -\frac{1}{m}\sum_{i=1}^{m}\left[y^{(i)} \log\left(g\left(z^{(i)}\right)\right) + \left(1-y^{(i)}\right)\log\left(1-g\left(z^{(i)}\right)\right)\right] \tag{3}$$

Where m is the number of samples, $y^{(i)}$ is the true label of the i-th sample (taking values of 0 or 1), and $g\left(z^{(i)}\right)$ is the probability that the model predicts that the sample belongs to the positive class based on the input features $x^{(i)}$. When the probability predicted by the model is the same as the true label, the value of the loss function is 0. And when there is a large difference between the predicted probability and the true label, the value of the loss function will increase. Cross-entropy loss is more suitable than mean squared error in classification problems because it can better reflect the prediction accuracy of the classification model, and its gradient is also more conducive to the update of model parameters during the training process.

2.3.4. Parameter Estimation

The gradient descent algorithm is used for parameter estimation. Gradient descent is a commonly used iterative optimization algorithm, and its core purpose is to find the minimum value of the loss function, to determine the optimal parameters β of the logistic regression model. The specific execution steps are as follows.

Step 1: Set initial values for the parameters β. The ini lues are set in a random assignment way.

Step 2: Calculate the gradient of the loss function $J(\beta)$ with respect to the parameter β_j. The calculation formula is as following.

$$\frac{\partial J(\beta)}{\partial \beta_j} = \frac{1}{m}\sum_{i=1}^{m}\left(g\left(z^{(i)}\right) - y^{(i)}\right)x_j^{(i)} \tag{4}$$

Where m represents the total number of samples; $y^{(i)}$ is the true label of the i-th sample, taking values of 0 or 1; $g\left(z^{(i)}\right)$ is the probability that the model predicts that the i-th sample belongs to the positive class based on its input features $x^{(i)}$; $x_j^{(i)}$ represents the j-th feature value of the i-th sample. The gradient essentially reflects the rate of change of the loss function at the position of the current parameter values, and its direction points to the direction in which the loss function value rises the fastest. In order to reduce the value of the loss function, we need to adjust the parameters in the opposite direction of the gradient.

Step 3: Update the parameters. According to the calculated gradient, update the parameter β_j according to the following formula.

$$\beta_j = \beta_j - \alpha\frac{\partial J(\beta)}{\partial \beta_j} \tag{5}$$

α is called the learning rate, which determines the step size of each parameter update. The value of the learning rate is extremely crucial: if α is set too large, the model may take too large a step during the parameter update process and directly skip the minimum point of the loss function, resulting in failure to converge to the optimal solution; on the contrary, if α is too small, the step size of the model parameter update will be very small, making the training process extremely slow, and a large number of iterations are required to gradually approach the better parameter values. In this paper, the value of α is 0.005.

Step 4: Iterate until convergence. Continuously repeat the above operations of calculating the gradient and updating the parameters until the value of the loss function no longer shows an obvious downward trend, at which point we consider that the model has reached a convergent state. The parameters β determined at the time of convergence are the optimal parameters of the logistic regression model estimated through the gradient descent algorithm. When the change of the loss function value is less than a pre-set extremely small threshold (set to 10^{-6} in this paper) after several consecutive iterations, it can be considered that the model has converged.

2.3.5. Model Output and Decision-making

After the model is trained through the previous steps, it will output a probability value $g(z)$ that the sample belongs to the positive class. In order to obtain the final classification result, we set a threshold of 0.5. When $g(z) \geq 0.5$, we determine that the sample belongs to the positive class, that is following.

$$\hat{y} = \begin{cases} 1, & g(z) \geq 0.5 \\ 0, & otherwise \end{cases} \tag{6}$$

2.4. Logistic Regression Model with the Introduction of Moderating Effect

Introduce a moderating variable and construct a logistic regression model that includes an interaction term.

$$\log\left(\frac{\rho(Y=1)}{1-\rho(Y=1)}\right)=\beta_0+\beta_1 X+\beta_2 M+\beta_3(X\times M)+\varepsilon \tag{7}$$

Where, M represents Moderating variable (diversity of information channels or coverage rate of policy promotion); $X \times M$ represents Interaction term, used to test the moderating effect.

3. RESULTS

3.1. Rural Residents' Food Safety Awareness

3.1.1. The Awareness of Food Safety among Rural Residents

Table 1. Food Safety Cognition Rating Table

Name	Sample Size	Minimum Value	Maximum Value	Mean Value	Standard Deviation	Median
Do you usually care about food safety issues?	3729	1	5	1.77	0.94	2
Are you confident in the food sold in the market currently?	3708	1	5	2.83	1.07	3

Based on the Food Safety Cognition Rating Table in Table 1, we can analyze the degree of concern of rural residents about food safety and their trust in the food sold in the market. The data shows that rural residents show a relatively high level of concern about food safety issues, with an average score of 1.77 and a median of 2.00. This indicates that most residents have a positive attitude of paying attention to food safety. Such concern may stem from their emphasis on a healthy diet and nutritional status, which in turn affects their food choices and consumption behaviors.

3.2. The Transformation of Food Safety Cognition

3.2.1. Data of Dietary Consumption Structure

(1) The proportion of staple food expenditure has decreased (from 45% to 38%).

Based on the comparison of the consumption account books of rural households in the same region in 2018 (at the beginning of the Rural Revitalization Strategy) and 2022, the proportion of staple food (rice, flour, potatoes) expenditure in the total food expenditure was calculated. By cross-verifying with the staple food consumption data of the corresponding years in the China Rural Statistical Yearbook issued by the National Bureau of Statistics, and controlling the error rate within ±3%, it can be concluded that the proportion of staple food expenditure of rural residents has decreased from 45% to 38%.

(2) The intake of animal protein is 20% lower than that of urban residents.

In the survey, the annual average consumption of meat/aquatic products (kg) per household was counted and converted into the daily per capita protein intake (g), referring to the recommended value in the Dietary Guidelines for Chinese Residents (60g/day). By using the protein intake data of urban residents in the Report on Nutrition and Chronic Diseases of Chinese Residents (2020) issued by the

National Health Commission and calculating the percentage of the gap, it can be found that the intake of animal protein (meat/aquatic products) is still 20% lower than that of urban residents.

(3) The consumption of deeply processed foods has increased by 12%.

The classification standards (such as prefabricated dishes, canned foods, puffed foods, etc.) were clarified. Through the statistics of the procurement receipts of rural households and the "purchase frequency in the past year" in the questionnaire. Taking 2018 as the base period, the Compound Annual Growth Rate (CAGR) formula was adopted:

3.2.2. Data of Nutritional Health

The overweight rate of BMI is 32%, compared with the national rural average of 28%. The survey team carried standardized weighing scales and height measuring instruments to calculate the BMI of the respondents (BMI≥24 is considered overweight). By citing the overweight rate data (28%) of rural residents in the Report on Nutrition and Chronic Diseases of Chinese Residents (2020), the difference was calculated.

According to the data of the National Rural Revitalization Association, the BMI results are shown in Table 2.

Table 2. The BMI Results

Option	Frequency	Percentage (%)	Cumulative Percentage (%)
BMI < 18.5 (Underweight)	360	6.7	6.7
18.5 ≤ BMI < 24.9 (Normal)	3162	58.84	65.54
25 ≤ BMI < 29.9 (Overweight)	1568	29.18	94.72
BMI ≥ 30 (Obese)	284	5.28	100
Total	5549	100	100

3.3. The Driving Effect of Food Safety Awareness on Healthy Behaviors

Table 3. The Specific Variables Table

Variable Name	Variable Description
Age	18-20 years old = 1, 20-25 years old = 2, 25-30 years old = 3, 30-35 years old = 4
Gender	Male = 1, Female = 2
Years of Education	Primary school and below = 0, Junior high school = 1, Senior high school and junior college = 2, Undergraduate and above = 3
Region	Developed region = 0, Underdeveloped region = 1
Income	Below 50,000 yuan = 1, 50,000-100,000 yuan = 2, 100,000-200,000 yuan = 3, Above 200,000 yuan = 4
Safety Awareness	Low = 1, Medium = 2, High = 3
Number of Rural Households	More than 50 households = 5

Based on the data types, we selected seven variables, including age and safety awareness, as independent variables and health status as the dependent variable to construct a binary logistic regression model. The specific variables are shown in Table 3. The dependent variable Y of healthy behaviors is defined as follows: Y = 1 indicates compliance with healthy behaviors (these behaviors are evaluated through the behavior observation items in the questionnaire survey, and a score of 4 or more on the 5-point scale is considered as compliance), while Y = 0 indicates non-compliance with healthy behaviors. The core independent variable X is the score of food safety awareness, and the scores are summed up and standardized to a scale of 0-100 for measurement. In addition, the study also takes into account a series of control variables, including individual-level variables such as age, gender, years of education, and annual household income, regional-level variables such as the province of residence (such as Henan, Shandong, Guangdong, etc.) and whether it is located in the urban-rural fringe, as well as behavior-related variables such as food purchase channels (such as markets, supermarkets, or mobile vendors).

Through the Stata code, it can be observed that for every 1-point increase in the food safety awareness score, the probability of an individual demonstrating healthy behaviors significantly increases by 5% (represented by an odds ratio OR = 1.05, and the P-value is less than 0.001, indicating that this association is highly statistically significant). It is worth noting that there is heterogeneity in the transformation from this awareness to behavior: women are more strongly driven by food safety awareness to exhibit healthy behaviors compared to men (the OR value for women is 1.22, with men as the reference group). Regional differences are also obvious. Among them, the transformation effect of food safety awareness on healthy behaviors among residents in Guangdong is the most prominent (OR = 1.38), which may be related to the province's strong food safety publicity strategy. In addition, the years of education play a moderating role in the transformation process from awareness to behavior. The interaction term between the years of education and the food safety awareness score is significant ($\beta = 0.02$, $P = 0.012$), which means that the group with higher education can more effectively transform food safety knowledge into practical actions. Therefore, the government needs to strengthen food safety awareness education.

3.4. The Moderating Effect of the Transformation from Awareness to Behavior

Logistic regression is used to explore whether the diversity of information acquisition channels and the coverage rate of policy publicity, as moderating variables, will enhance or weaken the main effect of food safety awareness driving healthy behaviors.

Specifically, the diversity of information acquisition channels is classified by measuring whether the respondents obtain food safety information through ≥3 different channels (including but not limited to TV, WeChat, village committee announcements, etc.) (High diversity = 1, Low diversity = 0).

At the same time, as another moderating variable, the coverage rate of policy publicity is measured based on the coverage degree of food safety publicity activities within the county. The proportion of samples that have received ≥2 publicity activities within the past year is used as an indicator, and the value range is a continuous variable from 0 to 100%. On this basis, the independent variable is the food safety awareness score, using a quantification standard of 0-100 points.

The dependent variable is healthy behavior, which is measured by a binary variable. By analyzing the relationships among these variables, this study will deeply explore how information acquisition and policy publicity affect the transformation process from food safety awareness to healthy behaviors.

As shown in Table 4, For every 1-point increase in the food safety awareness score, the probability of healthy behaviors occurring increases by 4% (OR = 1.04, $P < 0.001$). Improving the food safety awareness of rural residents is the key to improving their healthy behaviors. Women are more likely to be driven by awareness than men (OR = 1.22 vs. the male reference group); the transformation effect from awareness to behavior is stronger in the group with higher education (the interaction term $\beta = 0.02$, $P = 0.012$). Increasing the coverage rate of policy publicity from 50% to 80% will increase the probability of healthy behaviors by 0.56%, and save 28,000 yuan in medical expenses per year for a county with a population of 100,000. For counties with a coverage rate of less than 60%, it is mandatory to have ≥4 annual publicity activities; strengthen the publicity of food safety awareness.

Table 4. Regression Results of the Moderating Effect

Variable	Coefficient (β)	Standard Error	OR	95%	P-value
Main Effects					
Food Safety Awareness Score	0.04	0.01	1.04	[1.02, 1.06]	0
Diversity of Information Channels	0.25	0.1	1.28	[1.05, 1.56]	0.014
Coverage Rate of Policy Publicity	0.15	0.06	1.16	[1.03, 1.31]	0.011
Interaction Terms					
Awareness × Information Diversity	0.02	0.01	1.02	[1.00, 1.04]	0.032
Awareness × Policy Publicity	0.03	0.01	1.03	[1.01, 1.05]	0.003
Control Variables					
Age	-0.01	0	0.99	[0.98, 1.00]	0.058
Gender (Female = 1)	0.2	0.08	1.22	[1.04, 1.43]	0.013
Years of Education	0.04	0.01	1.04	[1.02, 1.06]	0.001
Household Income	0.03	0.01	1.03	[1.00, 1.06]	0.028
Region					
GuangDong Province	0.15	0.1	1.16	[0.95, 1.42]	0.134
HeNan Province	0.32	0.09	1.38	[1.15, 1.65]	0.001
Model Fitting					
Sample Size	2,345				
Pseudo R^2	0.18				
Likelihood Ratio Test	256.7	0			

4. CONCLUSION

This study has conducted an in-depth exploration focusing on the food safety awareness and healthy behaviors of rural residents, as well as the relationship between the two. Through the comprehensive analysis of rural data from multiple provinces, the following conclusions are drawn:

The current situation of food safety awareness and significant group differences: Rural residents show a relatively high level of concern about food safety issues, but they have a low level of confidence in the food sold in the market. In terms of the dietary consumption structure, the proportion of staple food expenditure has decreased, the intake of animal protein is lower than that of urban residents, and the consumption of deeply processed foods has increased. In terms of nutritional health status, the overweight rate of BMI is relatively high, the prevalence rate of hypertension reaches 18.5%, and there are obvious regional differences. The consumption of aquatic products is high in the eastern region, the loss of fruits and vegetables is large in the central and western regions, and the consumption of prefabricated dishes is high in e-commerce demonstration counties, but residents have insufficient awareness of food safety.

The driving effect of food safety awareness on healthy behaviors is obvious: Based on the analysis of the Logistic regression model, for every 1-point increase in the food safety awareness score, the probability of an individual demonstrating healthy behaviors significantly increases by 5%. There is heterogeneity in this influence. Women are more strongly driven by food safety awareness to exhibit healthy behaviors, and the transformation effect of food safety awareness on healthy behaviors among residents in Guangdong is the most prominent. The years of education play a moderating role in the transformation from awareness to behavior, and the group with higher education is more capable of transforming food safety knowledge into practical actions.

The moderating effect significantly affects the transformation from awareness to behavior: The diversity of information acquisition channels and the coverage rate of policy publicity have a significant moderating effect on the main effect of food safety awareness driving healthy behaviors. The diversity of information acquisition channels can enhance this driving effect, and the improvement of the coverage rate of policy publicity also helps to increase the probability of the occurrence of healthy behaviors. For example, increasing the coverage rate of policy publicity from 50% to 80% will increase the probability of healthy behaviors by 0.56%, and save 28,000 yuan in medical expenses per year for a county with a population of 100,000.

REFERENCES

[1] Zhao X, Deng S, Zhou Y. The impact of reference effects on online purchase intention of agricultural products: The moderating role of consumers' food safety consciousness [J]. Internet Research, 2017, 27(2): 233-255.

[2] Chai D, Meng T, Zhang D. Influence of food safety concerns and satisfaction with government regulation on organic food consumption of Chinese urban residents [J]. Foods, 2022, 11(19): 2965.

[3] Min S, Peng J, Qing P. Does internet use improve food safety behavior among rural residents? [J]. Food Control, 2022, 139: 109060.

[4] Li X, Li X, Liao Y, et al. Analysis of residents' food safety satisfaction from the perspective of income heterogeneity [J]. Scientific Reports, 2021, 11(1): 6666.

[5] Berghoff A S, Schur S, Füreder L M, et al. Descriptive statistical analysis of a real life cohort of 2419 patients with brain metastases of solid cancers [J]. ESMO open, 2016, 1(2): e000024.

[6] Ali S A G, AL-Fayyadh H R D, Mohammed S H, et al. A descriptive statistical analysis of overweight and obesity using big data [C]//2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA). IEEE, 2022: 1-6.

[7] De Menezes F S, Liska G R, Cirillo M A, et al. Data classification with binary response through the Boosting algorithm and logistic regression [J]. Expert Systems with Applications, 2017, 69: 62-73.

[8] Unler A, Murat A. A discrete particle swarm optimization method for feature selection in binary classification problems [J]. European Journal of Operational Research, 2010, 206(3): 528-539.