Analysis of Influencing Factors of HPV Vaccination Willingness based on Principal Component Analysis (PCA) and Multiple Linear Regression (MLR)

-- Take Anhui Province as an Example

Wenxu Jiao[1, *], Xinyu Tu[2], Wenli Liu[3]

[1]School of Statistics and Applied Mathematics, Anhui University of Finance and Economics, Bengbu 233030, China

[2]School of International Business and Economics, Anhui University of Finance and Economics, Bengbu 233030, China

[3]School of Law, Anhui University of Finance and Economics, Bengbu 233030, China

*1714290897@qq.com

Abstract

Background: The current HPV vaccination rate among Chinese women of appropriate age is low. This study aimed to take Anhui Province as an example to explore the factors that affect people's willingness to vaccinate against HPV, and to explore suggestions for promoting HPV vaccine. Methods: Based on the results of the questionnaire survey as data support, principal component analysis and multiple linear regression analysis are used to study the degree of influence of each variable on the population's willingness to receive vaccination. Conclusion: Vaccine quality effects (B=0.632, p<0.05) and External influences (B=0.275, p<0.05) are significantly and positively correlated with vaccination willingness, while Users' sense of actual experience (B=-0.541, p<0.05) and Users' own condition (B=-0.188, p<0.05) are negatively correlated with willingness to inoculate.

Keywords

Human Papillomavirus; Vaccination; Influencing Factors; Principal Components; Multiple Linear Regression.

1. Introduction

Cervical cancer is the second most common female malignant tumor in China, and in recent years, it has shown an increasing incidence and a trend of younger incidence. According to the World Health Organization (WHO), there are an average of 570,000 new cases of cervical cancer and about 311,000 deaths due to cervical cancer worldwide each year. Continuous and close infection of high-risk HPV virus is a necessary cause of cervical cancer, and HPV vaccination is one of the effective measures to prevent cervical cancer [1]. In the Global Strategy for Accelerating the Elimination of Cervical Cancer, the World Health Organization stated the goal of achieving 90% of girls to be vaccinated against HPV by age 15 by 2030. However, in the actual vaccination work, the current HPV vaccination rate of school-age girls in China is less than 1% due to the lack of public awareness, difficulties in making vaccine appointments, unreasonable price setting, and concerns about the safety of the vaccine, etc. Among them, Anhui province is located in the southeast of China, the regional economic development is relatively backward, the health and medical conditions are relatively poor, the appropriate age group HPV vaccination rate is at a low level. Therefore, it is important to explore the factors influencing people's willingness to receive HPV vaccination in Anhui Province as an example, in order to

better guide the public to prevent HPV infection and advocate early vaccination for people of appropriate age.

2. Methodology

2.1. Data Sources

By consulting the data and combining the opinions of experts, we conducted a simple random sampling of 61.13 million permanent residents in Anhui Province from October 2021 to December 2021, and carried out the survey by means of self-filling online questionnaires, first conducting a pre-survey in a small area, and modifying and improving the questionnaire according to the pre-survey results. The purpose of the survey was explained to the study participants before the survey and informed consent was provided. After obtaining the consent of the respondents, the survey was completed through online publication and questionnaires were collected. A total of 957 questionnaires were distributed, and 818 were validly returned, with a valid return rate of 85.48%.

2.2. Indicators

The questionnaire was designed with reference to domestic and foreign literature [2,3,4], including basic information of the population, HPV vaccination status, vaccination willingness and its influencing factors.

2.3. Ethical Considerations

Informed consent was provided to the study subjects, all subjects participated voluntarily, and all responses were collected and analyzed without identifiers.

2.4. Statistical Analysis

SPSS 25.0 software is used for statistical analysis of the data, and values of $P<0.05$ are considered to be statistically significant. Principal component analysis and multiple linear regression analysis are used to explore the related factors affecting HPV vaccination willingness. Linear regression analysis requires that independent variables are independent of each other, but in practical problems, there is often a certain correlation between different indicators, due to the large number of indicators and the correlation between each other, greatly increasing the complexity of the analysis problems. PCA starts from the mutual relationship between multiple variables, and uses the idea of dimension reduction to turn multiple indicators into a few unrelated comprehensive indicators [5]. After eliminating the multicollinearity in the regression model, the extracted principal component variables are used as independent variables for regression analysis to explore the degree of influence of each principal component on HPV vaccination willingness.

3. Results and Discussion

Cervical cancer, as one of the most common cancers in women, has been attracting much attention. Chen Yamei (2020) pointed out in the study that vaccine price and people's knowledge of HPV are closely related to vaccination willingness [6]; in a study by Zhang Xi and Chen Hui et al. (2021), it was established that the safety, efficacy and side effects of the vaccine, the support of family and friends, and one's own economic status would all affect people's vaccination intention [7]; She Xi et al. (2020) believed that HPV vaccine vaccination location, vaccine source and whether medical personnel would significantly affect respondents' acceptance of HPV vaccine [8]; in Liu Qianqian's (2020) study, SEM was used to Determine the impact mechanism of government and community support and guarantees on vaccination willingness [9].

Based on a large number of literature at home and abroad, this paper further explores the influencing factors of HPV vaccination willingness. Twelve main influencing factors are identified: $X_1, X_2 \ldots, X_{12}$ respectively represent Vaccine safety, Vaccine adverse reactions, Actual preventive effect, Vaccine price, and Convenience of vaccination (many people reflect that vaccine supply is insufficient in their own areas and vaccination time is not suitable for themselves), Appointment efficiency (At present, China's vaccine market is in short supply, and nine-valent and four-valent vaccines are still imported from abroad. In many places, it is difficult to get a needle and make an appointment), Age characteristics (refers to people who are not in HPV vaccine applicable age range of 9-45 years old), Knowledge level, Government support, Family and friends' support, Economic status and Degree of standardization of the vaccine market (currently there is vaccination chaos in the vaccine market. Scalpers snatch vaccines on behalf of others, and institutions bid up the price of vaccines).

3.1. Applicability Test

KMO test results: Sampling suitability measure (KMO)>0.5, correlation exists between variables, and Bartlett's sphericity test P<0.05, indicating that the sampling data can be analyzed by principal component analysis.

Table 1. KMO and Bartlett's Test

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.769
Bartlett's Test of Sphericity	Approx. Chi-Square	351.218
	df	245
	Sig.	0.000

3.2. Principal Component Analysis

The data of the 12 influencing factors listed above are imported into SPSS 25.0 for principal component analysis. The output results are shown in Tables 2 and 3.

Table 2. Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative%	Total	% of Variance	Cumulative%
1	4.109	32.350	32.350	4.109	32.350	32.350
2	2.372	21.942	54.292	2.372	21.942	54.292
3	1.636	18.114	72.406	1.636	18.114	72.406
4	1.118	14.387	86.793	1.118	14.387	86.793
5	0.891	4.054	90.847			
6	0.823	2.418	93.265			
7	0.770	2.103	95.368			
8	0.504	1.602	96.970			
9	0.264	1.183	98.153			
10	0.188	1.022	99.175			
11	0.129	0.514	99.689			
12	0.058	0.311	100.00			

The output shows that the cumulative variance contribution rate of the 4 PCs (principal component) generated by this model is 86.793%, which can explain most of the information of the original data and has the most ideal effect.

Table 3. Component Matrix

Variable	Component			
	1	2	3	4
X_1	0.814	0.274	0.056	0.004
X_2	0.950	0.173	0.048	0.068
X_3	0.748	0.218	0.157	0.023
X_4	0.364	0.891	0.337	0.099
X_5	0.094	-0.849	-0.028	0.025
X_6	-0.326	0.798	0.026	-0.144
X_7	0.151	-0.318	-0.863	0.009
X_8	0.314	-0.280	0.716	-0.170
X_9	-0.064	-0.139	0.315	0.804
X_{10}	-0.118	0.299	0.078	-0.739
X_{11}	0.360	0.148	-0.413	0.708
X_{12}	-0.383	0.108	0.138	0.902

For the 4 PCs, the loadings represent the importance of the variable on each principal component. Among them, the three variables of Vaccine safety, Vaccine adverse reactions and Actual prevention effect have relatively high loadings on PC1, which is summarized as Vaccine quality effect Y_1. The loadings of Vaccine price, Convenience of vaccine and Appointment efficiency are relatively high on PC2, so it is called User's actual experience Y_2. The loadings of Age characteristics and Knowledge level are higher on PC3, which is called the User's own situation Y_3. While the Government support, Family and friends' support, Economic status and the Degree of standardization of the vaccine market have a high load on PC4, so they are collectively referred to as External influences Y_4.

The PCA shows that the influencing factors of people's HPV vaccination willingness can be divided into four levels: Vaccine quality effect, User's actual experience, User's own situation and External influences.

3.3. Multiple Linear Regression Analysis

Assuming that the loading coefficients of the eigenvectors are respectively $a_i(a_{1i}, a_{2i}, \cdots, a_{12i})$ $i = 1,2,3,4$, then the principal components

$$Y_i = a_{1i}X_1 + a_{2i}X_2 + \cdots + a_{12i}X_{12} \quad i = 1,2,3,4$$

In the linear regression model, we use the 4 PCs (Y_1, Y_2, Y_3, Y_4) as independent variables and the HPV vaccination willingness (Z) as dependent variable to establish the regression model as follows:

$$Z = \beta_0 + \beta_1 Y_1 + \beta_2 Y_2 + \beta_3 Y_3 + \beta_4 Y_4 + \varepsilon$$

Table 4. Model Summary.

R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
0.977^a	0.954	0.953	1.114	2.317

a. Predictors: (Constant), Y_1, Y_2, Y_3, Y_4
b. Dependent Variable: Z

SPSS 25.0 software is used to obtain the output results shown in Table 4.

Table 5. ANOVA.

	Sum of Squares	df	Mean Square	F	Sig.
Regression	5026.715	4	1256.679	1012.634	0.000^b
Residual	241.995	195	1.241		
Total	5268.710	199			

Table 6. Coefficients

	Unstandardized coefficients		Standardized coefficients	t	Sig.
	B	Std. Error	$Beta$		
(Constant)	3.21	2.849	-	12.187	0.006
Y_1	0.632	0.094	0.714	6.718	0.000
Y_2	-0.541	0.125	-0.302	-4.326	0.000
Y_3	-0.188	0.023	-0.236	-8.173	0.013
Y_4	0.275	0.089	0.495	3.132	0.039

The regression results show that ①the value of the correlation coefficient R is equal to 0.977, and the adjusted R^2value is equal to 0.953, which is close to 1, indicating that the model fits well and has a high degree of explanation; ②the overall significance test F value=1012.634, p-value<0.05, indicating that the overall effect of the model estimation is good; ③the p-values of the significance tests of the constant term and the four explanatory variables are all less than the significance level, indicating that the 4 PCs can explain the variation of the dependent variable well; ④DW test value=2.317, there is no serial autocorrelation in the model, and the comprehensive model is significantly valid and statistically significant.

From the regression results, we find that Vaccine quality effect (B=0.632, p<0.05) and External influences (B=0.275, p<0.05) have a significant positive effect on people's willingness to receive HPV vaccine, while Users' actual experience feeling (B=-0.541, p<0.05) and Users' own situation (B=-0.188,p<0.05) have negative impact, i.e., the safety of the vaccine, adverse reactions, vaccine price, people's knowledge of HPV vaccine, the difficulty of vaccine appointment and vaccination, the degree of vaccine market standardization, family government support, and personal economic status are important influencing factors on people's willingness to receive HPV vaccine.

The less people know about HPV and HPV vaccine, the more reluctant they are to get vaccinated. At present, the supply of vaccine in the market exceeds the demand, making appointments difficult in many areas, and the market chaos is frequent. However some people are concerned about the safety and efficacy of vaccines and whether there are adverse reactions. Therefore, improving the quality of vaccines, strengthening government, community and family support for vaccination, and appropriately adjusting the price of vaccines are conducive to increasing people's willingness to receive vaccines.

4. Conclusion and Recommendations

4.1. Conclusion

This paper aims to use PCA and MLR to model the relationship between HPV vaccination willingness and 12 influencing factors in Anhui Province. Based on the data collected from the questionnaire, 4 PCs (explaining 86.793% of the original information) are retained after extraction by PCA to avoid over-extraction and under-extraction. Regression analysis is then

used to explore the relationship between the 4 uncorrelated PCs and vaccination intention using component scores. The results show that Vaccine quality effect and External influences have a significant positive effect on people's willingness to receive HPV vaccine, while Users' actual experience and Users' own situation have a negative effect on willingness to receive vaccination.

4.2. Recommendations

In view of the current problems in the HPV vaccine market and the actual needs of the public, ①it is recommended to strengthen the popularization and education of the public on HPV knowledge, and conduct extensive publicity on the occurrence, prevention, treatment, and harm of cervical cancer, so as to comprehensively improve the public's awareness of HPV vaccine. Recognize and strive to develop vaccine types that are suitable for a wider range of ages, so that people are no longer limited by age for vaccination; ②suggest that domestic vaccine R&D enterprises actively introduce foreign vaccine R&D technologies, break through R&D challenges as soon as possible, complete the clinical trial stage as early as possible, expand the market scale and increase the supply of vaccines; ③recommend setting reasonable vaccination prices, consider including HPV vaccine in health insurance coverage or providing free or subsidized vaccination for the population; ④it is recommended to strengthen vaccine supervision, strictly control the quality of vaccines, improve the actual preventive effect of vaccines, ensure the safety and reliability of vaccination, reduce the incidence of adverse reactions, and ensure higher cervical cancer prevention efficiency; ⑤it is suggested that the state vigorously crack down on chaotic behaviors in the vaccine market, purify the HPV vaccine market, and at the same time strengthen education and propaganda encouragement for the general public in the community, schools and government.

References

[1] Lin Daner, Huang Yong, Li Weinan, Zhu Qi. Analysis of HPV vaccination willingness and influencing factors among school-age women in Baiyun District, Guangzhou [J]. China Public Health Management, 2021,37(06):774-777+782.

[2] Lai Jiangyi, Wu xiaqiu. Survey on human papillomavirus and its vaccine awareness and vaccination intention among female university students in Hangzhou[J]. China Vaccine and Immunization, 2019, 25 (03):303-307.

[3] Li Yajun, Shen Jie. Analysis of cervical disease screening results and HPV vaccination willingness among women in Jiaxing City [J]. Chin J Public Health Management, 201,37(06):778-782.

[4] Zhang Jie, Zhang Jiayi, Pang Minhui, Yang Jianwei, Chen Xiongfei, Dong Xiaomei. Guangzhou college students' intention and influence factors analysis of HPV vaccine immunization [J]. Chinese journal of disease control, 2018, 22 (9): 965-967 + 971. DOI: 10.16462/ j. carol carroll nki ZHJBKZ. 2018. 09.022.

[5] Zhang Qian, Shen Shulin. Based on principal component analysis of influencing factors of housing industrialization development [J]. Journal of engineering economy, 2021, 31 (5): 76-80. The DOI: 10.19298/ j. carol carroll nki. 1672-2442.202105076.

[6] Chen Yamei. A study on the awareness of HPV vaccine knowledge, willingness to vaccinate and factors influencing vaccination [D]. Nanchang University, 2020. DOI:10.27232/ d. cnki. gnchu. 2020. 000221.

[7] Zhang Xi, Chen Hui, Sun Liying, Yuan Changzheng, Zhou Jing, Feng Xiaoyu, Si Mingyu, Su Xiaoyou, Li Jing. A Study on influencing factors of female college students' HPV vaccination willingness in midwest region [J]. Modern Preventive Medicine, 2021, 48(23):4265-4271.

[8] She Xi, Ding Yibiao, Wang Yong, Yao Yuping. A survey on awareness of HPV and vaccines among community workers and medical staff in Shanghai [J]. International Journal of Reproductive Health/ Family Planning, 2020, 39(01): 40-44.

[9] Liu Qianqian. Study on the influencing factors of HPV vaccine acceptance based on SEM model [D]. Central China Normal University, 2020. DOI: 10.27159/d.cnki.ghzsu.2020.000679.

[10] Xie Wenbo. HPV Vaccine: Behind "One shot is hard to find" [N]. Health News, 2022-02-15(002).