Research on the Current Situation of Chronic Disease Medical Treatment for Primary and Secondary School Teachers and the Optimization of Education Management System—— Based on Structural Equation Model and Kmeans Algorithm

Yunru Zhou[1,a], Yiyin Zheng [2,b], Xiao Zhang [1,c*]

[1]*School of Public Health, Southeast University, Nanjing, China*
2 School of Economics, Wuhan University of Technology
Yunru Zhou and Yiyin Zheng contributed equally to this work and should be considered co-first authors.
[a]*Yunruzhou9766@163.com,* [b]*295727@whut.edu.cn,* [*c]*zhangxiao@seu.edu.cn*

Abstract

In order to protect the physical and mental health of teachers and explore the factors hindering teachers' normal medical treatment, a series of educational reforms such as the "double reduction" policy have further increased the burden on primary and secondary school teachers, using spss for descriptive statistics, constructing a structural equation model based on deconstructed planning behavior (DTPB) theory, and targeting the shortcomings of the traditional kmeans clustering algorithm. The kmeans++ algorithm based on compound distance is proposed, and the initial value selection is optimized, which has a better clustering effect, and further extracts the characteristics of the teacher group who have not gone to the doctor, and the research conclusion shows that the perfect performance management policy and system support of the school is the biggest driving force for medical treatment, and the game of health benefits and cost expenditures in cognitive demand is the biggest resistance to the treatment of chronic diseases; based on this, it is necessary to adhere to the dual line of policy care and institutional guarantee, and carry out the reform of the teaching system. Form a teaching management and teaching team that is compatible with the environment.

Keywords: *Educational Reform; K-Means++ Algorithm; School Human Capital Management; Multigroup Structural Equation Model*

1. INTRODUCTION

In order to effectively improve the level of school education, comprehensively consolidate the accumulation of talents, serve the national strategic needs and the great cause of national rejuvenation, the "double reduction" policy came into being, behind the education reform is the increase in the burden of primary and secondary school teachers fighting in the front line of education, and the performance-oriented, quantitative, non-professional and administrative assessment leads to the physical and mental burden of teachers, thereby increasing the risk of chronic diseases and the difficulty of medical treatment. According to the "China Medium- and Long-term Plan for the Prevention and Treatment of Chronic Diseases (2017-2025)", the proportion of chronic disease deaths in China is currently as high as 86.6%. Among them, among the chronic disease groups, primary and secondary school teachers accounted for 69.2%, becoming the main group.

At present, the academic community is very concerned about the health status of teachers, investigated the current situation of teachers' health and the factors of teachers' illness, and found that teachers' professional cognition, teachers' pressure status, teacher workload, and teachers' professional security are the main influencing factors, but at present, there are still certain defects in the process of education management

A. El-Hashash et al. (Eds.): IEIT 2022, ACSR 100, pp. 1041–1046, 2023.
https://doi.org/10.2991/978-94-6463-058-9_163

and education reform in Primary and Secondary Schools in China, and teacher human resources are the key reserves in the sustainable development of schools, ensuring the health of teachers, improving the sick leave management system and teacher force construction. It is directly related to the teaching level and long-term development of the school.

Although the existing research on the empirical and theoretical research on the medical behavior of chronic diseases has formed a certain system, but in the research group mainly focuses on the vulnerable groups in society, for primary and secondary school teachers, one of the main incidence groups of chronic diseases, there are few discussions on the medical behavior of this group at home and abroad, and the current statistical analysis methods are difficult to deeply express the interaction and path between the influencing factors of medical behavior, and lack of deeper discussion. According to this, based on DTPB, the structural equation model is constructed, and the deeper psychological perception elements of individual behavior are deeply analyzed, in order to analyze the subjective decision-making basis of teachers' medical behavior, aiming at the shortcomings of the traditional kmeans clustering algorithm, the kmeans++ algorithm based on composite distance is proposed, and the initial value selection is optimized, which has a better clustering effect, explores the influence of different sociodemographic characteristics, and provides a realistic reference for formulating a sound school education management system and ensuring the physical and mental health of the teacher group.

2. RESEARCH METHODOLOGY

2.1. Method selection

In this study, the questionnaire was entered using Epi Data 3.0 software and statistical analysis was performed with SPSS 24 software, and the statistical analysis method was mainly descriptive analysis. Structural equation model construction and fitting were performed with Amos software, and parametric estimation of the measurement model was performed using great likelihood estimation (ML). Aiming at the shortcomings of the traditional kmeans clustering algorithm, the kmeans++ algorithm based on composite distance is proposed, which considers the spatial similarity and morphological similarity of the power curve, and optimizes the initial value selection, which has a better clustering effect.

The traditional kmeans algorithm uses a random method for the selection of the initial value, which makes the multiple clustering results very unstable [11]. The kmeans++ method is used to optimize the selection of the initial value. The core idea is that the ith cluster center should be chosen as farthest away from the selected cluster center [2].

2.2. Main process

2.2.1 Structural equation modeling

Structural equation modeling (SEM) is a widely used linear statistical modeling technique commonly used in sociology, psychology, economics, and other fields.

(1) Model building. Based on the planned behavior theory, various internal factors and external factors are placed in the three core concepts of behavior attitude, subjective norms and perceptual behavior control, respectively, to integrate and study behavioral willingness and actual behavior, and further deconstruct behavioral attitudes into perceptual ease of use, perceptual usefulness, and compatibility, subjective norms are deconstructed into peer influence and superior influence, and perceptual behavior control is deconstructed into self-efficacy, resource convenience and technical convenience. Compared with the traditional rational behavior theory and planned behavior theory, DTPB theory has a broader vision and has obvious advantages in explaining people's behavioral motivations and willingness. Widely used in the field of medical social science research. Therefore, the analysis of normative beliefs is based on the formation of social and close groups, and the analysis of beliefs is controlled by the definition of self-efficacy and convenience. Combined with the content and purpose of the actual research, the influence on behavioral attitudes was creatively incorporated into the previous medical experience as an external factor.

A measurement model is a model that uses the relationship between observed variables and potential variables to construct potential variables. The specific form is as follows:

$$x=\Lambda x\xi+\delta \quad (1)$$

$$y=\Lambda y\eta+\varepsilon \quad (2)$$

where: x is an exogenous identifier; Λx is the factor load matrix connecting x variables to ξ variables; ξ is an exogenous latent variable; Measurement error of δ x; y is an endogenous logo; Λy is the factor load matrix connecting the η variables of the y variable; η is an endogenous latent variable; The measurement error ε y.

Structural models reflect relationships between potential variables that cannot be directly measured, as follows:

$$\eta=B\eta+\Gamma\ \xi+\zeta \quad (3)$$

Where: η is an endogenous latent variable; B is the relationship between endogenous subliminal variables; Γ is the influence of exogenous submersible variables on endogenous subliminal variables; ξ is an exogenous latent variable; ζ is the interference term in the equation.

2.2.2. K-means ++

The principle of the elbow method and the calculation formula of the typical user screening model are as follows：

$$F(X)=1n\sum i=1kfmin[(Xi-Pi)2]F(X)=1n\sum i=1kfmin[(Xi-Pi)2] \quad (4)$$

$$F[dis(A2,B2)]min=fmin[ED(Xk,ck)+ FD_{\partial}^{N}(X_k,c_k)] \quad (5)$$

The algorithm steps are as follows：

(1) An object Xi is randomly selected from the input set as the initial cluster center c1.

(2) Calculate the shortest distance from the existing cluster center of each object in the sample, expressed in D(Xi), i.e

(3) Calculate the probability P (e.g., equation (6)) of each object in the sample to become the next cluster center, and select the largest probability as the next cluster center.

(4) Repeat steps (2) and (3) until the number of selected clusters reaches the required k.

3. RESULTS AND DISCUSSION

3.1. Reliability test

In this paper, SPSS24.0 software was used to test the reliability and validity of the questionnaire. As shown in Table 3, the Cronbach'αvalues of the potential variables are greater than 0.6 in the reliability test, indicating that the internal consistency and stability of the scale are good. In the validity test, the KMO values of the potential variables are above the threshold of 0.6, and the significance P-value of bartlett′s spherical test is 0.000<0.05, and the scale can be analyzed for factors. This paper tests the validity of the model by means of validation factor analysis. The results show that the combined reliability (CR) is greater than 0.7, indicating that the structural validity of each potential variable is good, and the model data passes the reliability and validity tests.

Table1: Reliability and validity results

Latent variable	Observation variable	Factor load	cronbach'a coefficient	Combined reliability (CR)	KMO
PU	PU1	0.811	0.793	0.892	0.774
	PU2	0.886			
PE	PE1	0.755	0.748	0.708	0.632
	PE2	0.725			
SI	SI1	0.756	0.727	0.715	0.695
	SI2	0.844			
	SI3	0.692			
CI	CI1	0.798	0.758	0.748	0.744
	CI2	0.864			
	CI3	0.813			
SE	SE1	0.816	0.883	0.829	0.869
	SE2	0.836			
	SE3	0.835			
RC	RC1	0.81	0.845	0.846	0.847
	RC2	0.732			
	RC3	0.721			

3.2. Structural equation path analysis

Table2: Table of factor loads for the model of medical behavior

Path analysis	Ustd	S.E.	C.R.	P
BI→RA	0.153	0.023	6.534	***
PE→AT	0.107	0.026	6.612	***
PU→AT	0.18	0.031	4.853	***
PS→PE	0.132	0.045	4.312	***
PS→PU	0.137	0.039	4.437	***
AT→BI	0.122	0.02	3.950	***
SN→BI	0.151	0.033	4.401	***
PBC→BI	-0.058	0.026	-1.732	0.083
PBC→RA	0.114	0.019	3.512	***

In the dimension of subjective specification, the standardized path coefficient of SI→SN is greater than that of CI→ SN. This conclusion suggests that in the decision-making process, teachers are more dependent on the population in the environment in which they work. When teachers find that their social groups are widely involved in attending classes with illness, their willingness to seek medical treatment is greatly reduced. The results of the interview showed that the procedures for teachers to take sick leave were cumbersome, and after the implementation of performance pay, each school had a provision for deducting the corresponding salary for sick leave, and there was no official mechanism for teachers to take leave to substitute classes, most of which were decided by teachers' private

communication. In addition, some teachers are worried that medical treatment will have an impact on students' learning progress and teaching quality, so they choose to take sick classes, and this phenomenon mostly occurs in the senior teacher group.

3.3. k means++ results

1. Determine the optimal number of clusters

In this paper, the K value of the number of teacher group clusters is determined by the elbow rule. The principle of calculation of the elbow function is the cost function, which is expressed as the sum of the degrees of distortion of various types, and the smaller the degree of distortion, the more compact the members inside such classes are with each other.

In this regard, this paper draws the elbow plot as shown in Figure 1, the distortion coefficient change is shown in Table 3, and its abscissa is the number of cluster categories K, and the ordinate coordinate is the degree of distortion. As can be seen from the figure, when the number of classes increases from 1 to 3, the total degree of distortion decreases rapidly, but when the number of categories exceeds 3, the total degree of distortion changes slowly. In this regard, this paper determines that k=3 is the optimal number of clusters.

Table3: Table of distortion coefficient changes

The number of clusters	Distortion coefficient
1	15.59
2	12.73
3	8.47
4	7.84
5	5.85
6	5.76
7	4.65
8	3.42

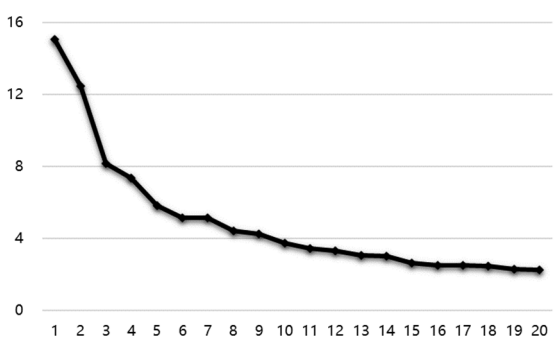

Figure 1: Cluster elbow plot

According to the clustering operation results, the teacher's feature analysis is carried out, and the characteristics of the three groups on each attribute are analyzed separately, so as to summarize the group characteristics of each category.

Based on the characteristic description, the three groups are named as "teachers with higher quality of life teachers", "general quality of life teachers" and "teachers with lower quality of life", based on the characteristic description. Among them, the characteristics of each teacher category are as follows:

Teachers with higher quality of life: This type of teacher group is about 45-60 years old, has a medium income level, is mainly engaged in teaching work, and is more willing to seek medical treatment than the other two groups. The possible reason is that this group of people has a higher standard of living, has sufficient time cost and economic affordability to seek medical treatment, and attaches more importance to physical health.

General quality of life teachers: This group of young teachers is the backbone of the school's teaching and scientific research work, because of the busy work and life, lack of time, and a strong sense of professionalism, even if they are sick, they are likely not to go to the doctor.

Teachers with low quality of life: Most of these teachers hold administrative positions, and in addition to teaching work, they also undertake administrative, publicity, teaching and research management and other work. This reflects another reason for not going to medical treatment: due to the rigidity of the management system such as the restriction of the school's teacher establishment, the school cannot hire substitute teachers, so the sick teacher's class has to be substituted by the teacher of the school, and the workload is multiplied, which often makes the substitute teachers tired and forms a vicious circle.

4. CONCLUSION

(1) Establish a scientific sick leave approval system and develop a teacher information management system

(2) Rationally and balanced allocation of teachers' human resources, and integrate multi-party forces outside the school

(3) Strengthen the construction of medical institutions and provide convenient health services

REFERENCES

[1] CX Miao, HH Li, JX Jiang, et al, “Analysis on the willingness of urban residents to seek medical treatment for chronic diseases and their influencing factors”, *Chinese Public Health,* vol. 37, no.1, pp.

123-127, 2021.

[2] Dubey A, Choubey A, "A systematic review on k-means clustering techniques", *International Journal of Scientific Research Engineering & Technology,* vol. 6, no. 6, pp. 624-627, 2017

[3] Jain A K, Murty M N, Flynn P J, "Data clustering: a review", *ACM Computing Surveys,* vol. 3, no.3, pp. 264-323, 1999

[4] JN Liu, XW Jiang, LY Yan, et al, "Pathway analysis of chronic disease health knowledge, lifestyle and workplace health service utilization in a certain area of Beijing", *Chinese Journal of Disease Control.* vol 24, no. 1, pp. 52-56, 2020.

[5] L Wang ,ZH Wang, "Investigation and analysis of health status of primary and secondary school teachers", *Modern Primary and Secondary Education,* vol. 30, no. 11, pp. 106-110, 2014.

[6] Moriya T, Roth H R, Nakamura S, et al, "Unsupervised pathology image segmentation using representation learning with spherical k-means", *Medical Imaging 2018 : Digital Pathology*, 2018.

[7] Nasir A, Mashor A S, Mohamed M, et al, "Enhanced kmeans clustering algorithm for malaria slide image segmentation", *Journal of Advanced Research in Fluid Mechanics and Thermal Sciences,* vol. 42, pp. 1-15, 2018.

[8] Pu Wen, Yang Min, "Analysis on the decision-making status and influencing factors of medical institutions for elderly patients with chronic diseases in Wuhan", *Modern Preventive Medicine,* vol. 48, no. 18, pp. 3384-3388, 2021.

[9] SS Jing, AT Yin, QY Meng, et al, "Study on the medical choice behavior of rural residents with chronic diseases", *China Health Economics,* vol. 9, no. 2, pp. 32-34, 2010.

[10] Tang J L, Wang D, Zhang Z G, et al, "Weed identification based on K-means feature learning combined with convolutional neural network", *Computers and Electronics in Agriculture,* vol. 135, pp. 63-70, 2017.

[11] Tang J L , Zhang Z G, Wang D, et al, "Research on weeds identification based on K-means feature learning", *Soft Computing,* vol. 22, no.22, pp. 7649-7658, 2018.

[12] XE Xiao, PH Zhong, X Lin, et al, "Investigation and research on the health status of primary and secondary school teachers in a district of Chongqing", *Chongqing Medicine,* vol. 46, no.26, pp. 3682-3684, 2017.