Study on a New Model of Physical Exercise Status of College Students in a University in Xinjiang to Resist Subhealth

Chuyu Wang [1,*], Gege Zhouhe [1], Ke Liang [1], Yuhan Hu [2], Shijie Song [2]

[1] School of Geography and Remote Sensing, Xinjiang University, Xinjiang, China

[2] School of Software, Xinjiang University, Xinjiang, China

515964388@stu.xju.edu.cn

Abstract. Individual health has become the core of realizing a healthy China, and this paper takes students of a university in Xinjiang as an example to explore the effect of physical exercise of college students on warding off subhealth. The study adopts a composite approach, constructs an assessment model based on the three major categories of subhealth and 39 tertiary indexes, and applies FAHP, path analysis and other methods to explore the relationship between physical exercise and subhealth. The results showed that college students' physical fitness generally declined during the epidemic, but their awareness of health increased; most of them were in moderate subhealth; social resources and adaptation had the greatest impact on subhealth, and positive emotions helped improve subhealth; there was a significant correlation between physical exercise and subhealth, especially in the male and female students' physical fitness tests. The study suggests that publicity and education on sub-health should be strengthened, and scientific advice should be provided to prevent sub-health from various aspects, so as to contribute to the realization of "Healthy China".

Keywords: Physical measurements, sub-health, physical activity.

1. Introduction

The World Health Organization defines health as a state of complete physical, mental and social well-being. With the rapid economic development, environmental pollution, the opening of the fast-food era, fierce competition in the market and the complexity of interpersonal relationships, the factors affecting human health have changed dramatically, and the medical model has undergone a fundamental transformation: from a single biomedical model to a biosocial-psychological medical model. The proportion of "sub-healthy" people in the total population is as high as 70% to 80%, and the research on "sub-healthy" state can be regarded as an indispensable part of life science research in this century. With the increase of public health awareness and the improvement of living conditions, more and more people show their attention to active health management. To this end, the State has been promoting the importance of health for all through the formulation of relevant policies and strategies.[1]

College students are prone to subhealth problems such as emotional instability and abnormal social interactions due to irregular life routines and increased stress and lack of time for physical activity. Physical exercise not only plays an important role in improving individual physical fitness, but also promotes emotional and social health, which is crucial for the future development of the country. Therefore, it is of great significance to study the current situation of physical exercise among college students and its relationship with subhealth status to promote the healthy development of society.

2. Selection, collection and analysis of data sources

(i) Sources and selection of data

[1] 2021 The State Council recently issued the "Comprehensive Fitness Program (2021-2025)" and the "Healthy China 2030" Plan Outline, and the General Office of the State General Administration of Sports issued the "Key Points of Mass Sports Work in 2023" and the "Implementation Plan of the New Era National Fitness Venue and Facility Enhancement Project (2023-2025)".

According to the Q3 2021 "Mobile Apps Ranking" published by Aurora, Sina Weibo has a good representation with an annual average MAU of 384.1 million. The text data used in this web survey are Weibo comments. Using Python language to write a crawler program, call the Super Eagle platform interface to identify the CAPTCHA request web page, according to mid, uid, max_id to get the comment web page data, and get the text data. In this paper, the text data preprocessing statistics get 1806 total comments, 1296 effective comments, the text data efficiency is 71.76%.

(ii) MiniTagCloud Big Data Analytics

After data processing, the MiniTagCloud big data platform was utilized for preliminary word frequency, sentiment and word cloud analysis. The words "cancel", "physical test", "can't run" in the college students' physical test comments and "body", "problems", "health" in the subhealth comments are high-frequency words. "problems" and "health" are high-frequency words. This shows that many college students are in a state of sub-health, which may be related to their lack of physical exercise and poor performance in physical tests. From the words "fail", "sad" and "scared" in the comments on physical exercise and "mental" in the comments on sub-health, it can be seen that the state of sub-health may be related to the lack of physical exercise and poor physical test scores of college students. It can be seen that being in a state of sub-health may cause some psychological problems and make college students afraid of physical exercise.

Physical exercise, college students' physical testing and subhealth comments were sentiment analyzed. The negative situation of physical exercise accounts for 56.37%, the negative situation of college students' physical test accounts for 60.14%, and the negative situation of sub-health accounts for 51.96%, and the proportion of negative emotional comments of the three is similar, accounting for more than 50%. Indirectly, it can be concluded that there is a connection between physical exercise, college students' physical examination and sub-health, and the negative emotional attitudes of netizens towards the three are mostly negative.

The word cloud statistics of "sub-health comment" reflect that most netizens in sub-health state have physical health problems, from the words "life", "staying up late" and "work", it can be seen that many netizens believe that sub-health originates from life and work, and staying up late is the main cause of sub-health. From the words "life", "staying up late" and "work", it can be seen that many netizens think that sub-health originates from life and work, and staying up late is the main cause of sub-health. "The word cloud statistics of "College Students' Physical Test Comments" reflect the emotional attitude of college students towards physical test and their physical condition. Starting to exercise. From the words "staying up late", "after-effects" and "disease", it can be seen that today's college students have poor physical fitness. The word cloud statistics of "Physical Exercise Review" reflects the emotional attitude of netizens towards physical exercise, and some netizens think that students only do physical exercise because of the mandatory exams such as the midterm exams, and such exams should be canceled.

(iii) Data collection and testing

The number of college students enrolled nationwide in 2021 is approximately 35 million. Using the formula $n = \frac{NZ_{\alpha/2}^2 S^2}{Nd^2 + Z_{\alpha/2}^2 S^2}$ to calculate the sample capacity, using the data from the pre-survey, calculating the variance, the maximum variance has a bias, correction to get S^2=1.96, the relative error does not exceed 7.5% (i.e., the absolute error limit d=0.15), the confidence level is 95%, the calculation of the initial sample size is:

$$n_0 = \frac{NZ_{\alpha/2}^2 S^2}{Nd^2 + Z_{\alpha/2}^2 S^2} = 435 \quad (1)$$

Based on the valid response rate of the questionnaire r_0=73%, the sample size was readjusted:

$$n_0 = \frac{n_0}{r_0} = 596 \quad (2)$$

Considering the actual situation, it was initially determined that a minimum of 596 questionnaires could be released to meet the survey needs. This study was conducted through probability sampling, following the principle of simple random sampling, randomly selecting some students in some classes of a university in Xinjiang from the national college student population to form a sample for the survey.

The survey was conducted by means of online questionnaires. A total of 600 questionnaires were distributed and 553 questionnaires were returned, with an effective recovery rate of 92%. Sampling follows the principle of randomness, and the data collected meets the needs of the sample size. Before the survey, we refer to the mode of questionnaire SHMSV1.0, add trap questions, and at the same time, we define the questionnaires that are too fast to be filled in, and the answers that are inconsistent between the 18th question and the 28th question, "Are you satisfied with your current living situation?" At the same time, the questionnaires with inconsistent answers between question 18 and question 28 "Are you satisfied with your current life situation?" were defined as invalid questionnaires to achieve the effect of quality control, and finally 440 valid questionnaires were obtained, with an effective recovery rate of 73%. When finishing and analyzing the data, we dealt with the outliers and tested and analyzed the reliability of the questionnaire.

3. Descriptive Statistical Analysis of Physical Fitness and Sub-health in College Students

(i) Analysis of the level of morphological development and physical functioning of the body

The overall weight of students in a university in Xinjiang in 2019-2022 is normal, but the proportion of normal weight of male students is much lower than the proportion of normal weight of female students, and the weight of male students shows an increasing trend year by year. The obesity proportion of female students stays around 3.50%, the overweight proportion and normal proportion do not change too much, and the low weight shows a small upward trend. Reflecting the physical function through spirometry index, it can be seen that the spirometry of male and female college students both show a decreasing trend, and male college students show a larger decline.

(ii) Analysis of the current situation of physical exercise in a university in Xinjiang, 2019-2022

As far as each item is concerned, the level of standing long jump of male and female college students in a university in Xinjiang from 2019 to 2022 shows a trend of increasing and then decreasing, and the failure rate of male students is higher than that of female students. The passing rate has increased year by year, but the excellent and good rates of college students have been decreasing significantly. The level of standing long jump has been rising year by year, which indicates that the explosive power of leg step of college students in a college in Xinjiang is gradually increasing. The number of male pull-ups shows a downward trend, reflecting the decline in arm grip strength, upper limb strength, back strength and waist and abdominal strength of male college students in a university in Xinjiang. The number of sit-ups for female college students showed an upward trend, reflecting the increase in waist and abdominal strength of female college students.

In terms of the basic situation of college students' physical activity and physical fitness, the following key trends are worth studying. First, during the period 2019-2022, the physical fitness and health test shows a continuous decline in the proportion of male students achieving excellent and good levels, while the proportion of female students during the same period is relatively stable without significant fluctuations. Meanwhile, the number of passes increased year by year during this period, regardless of gender, and the proportion of boys and girls who failed the physical health test reached its highest during the new crown epidemic in 2020-2021. This phenomenon suggests that the physical fitness of university students was generally negatively affected in the early stages of the epidemic, with possible causes including lack of exercise, imbalanced diet leading to decreased immunity and deterioration of physical functions. By 2022, the situation had improved, possibly as a result of people beginning to value physical health and being physically active.

(iii) Analysis of the current situation of sub-health among university students

In this survey, the proportion of college students in sub-healthy state and disease state is 82.95%, in this part of the respondents, the respondents in sub-healthy state is as high as 4/5, which reflects that the overall health state of college students is poor, and the vast majority of them are in sub-healthy state, and the proportion of moderate sub-health is as high as 37.27%, which indicates that nearly half of the college students are in moderate sub-healthy and in poor physical condition.

The overall sub-health score of college students is 67.60, and the score of social sub-health is close to it, which reflects that college students are relatively better in physical fitness and organ functioning, and relatively worse in emotional self-regulation and psychological condition, which is also related to the rising rate of contemporary college students suffering from mental illness.

The overall subhealth scores of male college students are higher than those of female college students, but male college students are in better physical condition and worse psychological condition. Female college students have better ability in communicating with others and teamwork, meanwhile, female college students have more serious psychological problems and should pay attention to their mental health.

(iv) Analysis of the current situation of sub-health among university students

The four main dimensions that have an impact on physiological sub-health are physical symptoms, organ function, body movement function, and energy. The score of Physiological Sub-health Scale is 69.69, and the lowest score is organ function, which may be related to the irregular lifestyle of college students. The highest score is the physical motor function with 80.85, probably because college students are in their young adulthood and have better motor ability.

Positive emotions, psychological symptoms and cognitive functioning are the three main dimensions that have an impact on mental subhealth. Among the dimensions, the lowest score is cognitive function, which may be related to the lack of social connection of today's college students; the highest score is positive emotion, with 65.21, which can reflect that today's college students face life with a positive attitude.

Social adaptation, social resources and social support are the three main dimensions that affect social subhealth. The score of the Social Sub-health Scale is 67.80, of which the dimension with the highest score is social support, indicating that today's college students receive great support from the society; the lowest score is social resources, reflecting that today's college students have less access to social resources, which should be strengthened in favor of the college student group.

4. Exploration of the factors influencing the state of subhealth

From the perspective of college students' own health, at this stage, college students' cognition and understanding of sub-health is still far from adequate, and it is difficult to achieve a deeper understanding of their own sub-health and to deal with the prevention and treatment measures of sub-health. Based on the comprehensive knowledge of sub-health, this paper utilizes the FAHP evaluation method and path analysis method to comprehensively explore the influence of various indicators on college students' sub-health, in order to find a new path to resist sub-health.

(i) FAHP analysis of subfertility and influence factor scores

This paper derives 10 secondary indicators and 39 tertiary indicators to synthesize subfertility.

Validated factor analysis was first performed and the results are shown in the table with good extraction of the measures within their factors.

In order to study the scores of the factors, FAHP evaluation and analysis method was conducted as a way to explore the variability of the scores of the factors. Thirty-nine tertiary indicators were established as shown in Fig.

This was done by assigning the scores as a 5-point scale, in the order of very poor (20 points), poorer (40 points), fair (60 points), better (80 points), and very good (100 points). Frequency data were derived using physical characteristics as an example:

Derive the following matrix R_1 and calculate the scores for each of the three levels of indicators F_1.

$$R_1 = \begin{pmatrix} 0.008 & 0.030 & 0.2320.3220.405 \\ 0.019 & 0.096 & 0.2700.3760.236 \\ 0.086 & 0.179 & 0.2810.2500.202 \end{pmatrix} \quad (3)$$

$$F_1 = R_1 \cdot \begin{pmatrix} 20 \\ 40 \\ 60 \\ 80 \\ 100 \end{pmatrix} = \begin{pmatrix} 81.54 \\ 74.10 \\ 65.94 \end{pmatrix} \quad (4)$$

Similarly, indicator scores for physical motor function, energy, positive mood, social adaptation, social support, social resources, organ function, psychological symptoms, and cognitive function were derived in this manner.

(ii) Hierarchical analysis to estimate secondary indicator scores

(1) Constructing a judgment matrix

In the case of physical characteristics, for example, there are three constructed weight indicators, namely M11, M12, and M13,The factor loadings are used as indicator importance scores, and the importance comparison between indicators is determined by comparing the two with each other. It is denoted as T1.

$$T1 = \begin{pmatrix} 1.000 & 0.922 & 1.071 \\ 1.084 & 1.000 & 1.161 \\ 0.934 & 0.861 & 1.000 \end{pmatrix} \begin{matrix} M11 \\ M12 \\ M13 \end{matrix} \quad (5)$$

$$a_{ij} = \frac{Importance\ score\ for\ indicator\ i}{Importance\ score\ for\ indicator\ j} = \frac{w_i}{w_j} \quad (6)$$

The consistency test was performed on the judgment matrix and all were consistent. The arithmetic mean of the other rows is derived in this way so that the weights are obtained by normalization. As an example, the physical characteristics are shown in the table.

(2) Comparative analysis of the results of college students' subhealth and influencing factors scores

According to the number of weights calculated to obtain the physical characteristics factor score of 74.19, the same method can be obtained to derive several other secondary indicators scores, and organize the degree of conformity of each influence factor.

The physical motor function reached 84.91, much higher than other secondary indicators, which shows that college students have better physical motor function. The scores of energy and positive emotions are also higher, which is also in line with young people's high energy level. Cognitive function has the lowest level of compliance, only 68.53. Memory, thinking and processing problems are still lacking, thus indicating that cognitive sub-health is a great asset for defense and management of sub-health.

(iii) Path analysis of the relationships within the factors influencing subhealth

(1) Structural equation modeling and testing

The prevention and control of subhealth is the main purpose of studying subhealth, so it is especially important to explore the influence size of each indicator on subhealth. This paper establishes a structural equation model to establish a path between each secondary indicator and the overall score of subhealth one by one, and analyzes the difference of its influence by comparing the model coefficients.

In this paper, a least squares one-way linear regression model is developed to determine whether there is a direct linear relationship between the paths and to compare the regression coefficients of the path nodes. The SPSS analyses all passed the significance test, and the paths are valid, the

To further explore the link between physical characteristics, positive emotions, social adjustment and social resources, the following table was analyzed using SPSS:

From the above table, it is clear that the covariance relationships of the factors appeared to be significant and strongly correlated, with the highest correlation between social adjustment and positive emotions, followed by the correlation between social adjustment and social resources.

(2) Analysis of results

Analysis of the questionnaire data shows that the general health condition of most college students in a university in Xinjiang is poor, mainly in a subhealthy state. In particular, the proportion of people with moderate subhealth is relatively high, indicating that nearly half of the college students are in poor physical condition. From a gender perspective, male college students have a better physical state than females, but a poorer mental state, while female college students perform better in interpersonal communication and teamwork, but also face mental health problems.

The analysis of the factors influencing sub-health revealed several key factors. First, social resources and social adjustment had the greatest impact on subhealth, suggesting that strained social relationships may lead to anxiety and restlessness. Second, the standardized estimated coefficient of positive emotions on overall subhealth was 0.969, suggesting that the management of positive emotions had a significant effect on reducing subhealth. In addition, although physical characteristics did not have a significant positive effect on overall subhealth production, physical characteristics such as appetite, sleep quality, and satisfaction with one's appearance were also associated with subhealth status, emphasizing the importance of self-regulation. Finally, there is a complex relationship between physical, mental and social health that interacts with each other, which requires comprehensive interventions.

5. Analysis of factors affecting sub-health

(i) Structural equation modeling

Based on the validated factor analysis (CFA) of the components of physical subhealth, mental subhealth and social subhealth, this part further explores the mechanism of physical activity's influence on subhealth. This paper uses Amos28.0 to construct a structural equation model containing 1 independent variable, 3 latent variables, 10 factors and 48 observed variables.

(1) Modeling

According to Principles and Applications of Structural Equation Modeling (SEM), Structural Equation Modeling (SEM) indirectly measures unobservable latent variables by means of observable variables (i.e., indicators) and consists of measurement equations and structural equations, modeled as follows:

$$X = \Lambda_x \xi + \sigma \tag{7}$$

$$Y = \Lambda_y \eta + \varepsilon \tag{8}$$

$$\eta = \mathrm{B}\eta + \gamma\xi + \zeta \tag{9}$$

Equations (7) (8) are the observation equations, the X, and Y denote exogenous and endogenous latent variables, respectively. σ and ε denote the exogenous and endogenous latent variables, respectively, and X and Y the error terms of the measurements, the Λ_x, and Λ_y denote the correlation coefficient matrix; Eq. (9) is the structural equation, with η are endogenous latent variables; ξ are exogenous latent variables; B, and γ is the path coefficient matrix, and ζ is the error vector, the equation reflects the relationship between exogenous latent variables and endogenous latent variables.

Boys and girls with different physical activity were modeled as shown in the figure below, where 3 latent variables were at the first level, i.e., physical subhealth, psychological subhealth, and social subhealth, and 10 latent variables were at the second level, i.e., physical traits, organ functioning, bodily motor functioning, energy, positive emotions, psychological symptoms, cognitive functioning, social adaptation, social resources, and social support for the 48 observed variables in the Data.

(2) SEM fitness test for sub-health influencing factors

According to the results of the model fit test, CMIN/DF (chi-square degrees of freedom ratio) = 2.026, in the range of 1-3, and RMSEA (root mean square of error) = 0.063, in the good range of <0.08. The results of the tests of IFI, TLI, and CFI reached a good level of 0.8 or more, so the SEM for analyzing the influences on boys' subfertility has good fitness. The same is true for female students.

(3) Subhealth influences SEM pathway relationship construction test results

Path analysis of the factors affecting sub-health, get SEM path relationship test selected part of the results, the total score of the physical test and the standardized coefficient of social sub-health is the largest, the physical test scores can reflect the physical fitness situation, the physical fitness situation has a close relationship with social adaptation, social resources, social support. This responds that the relationship between human and society is interdependent. The total score of physical test and physiological sub-health standardized coefficient is large, the level of physical test scores and physiological sub-health conditions are closely related. College students should consciously strengthen physical exercise, practice more exercise programs such as 1000-meter running, and develop physical exercise into a habit.

(ii) Analysis of the correlation between physical activity and sub-health

In order to explore the mechanism of physical activity's impact on subfertility even further, this subsection uses gray correlation analysis (GRA) to explore the effect of physical activity representational variables on subfertility.

Firstly, the reference sequence reflecting the behavioral characteristics of the system and the comparison sequence affecting the system behavior are determined. In order to study the effect of physical activity on sub-health, this paper designates sub-health as the reference sequence, denoted as X_0. And each physical exercise factor index is multiplied by the corresponding factor loading weights respectively to get 8 data columns as the comparison sequence, denoted as X_1,,,,,,,,,,,,,, and X_2,,,,,,,,,,,,,, and X_3,,,,,,,,,,,,,, and X_4,,,,,,,,,,,,,,,,,,, and X_5,,,,,,,,,,,,,, and X_6,,,,,,,,,,,,,,,,,,,,,, and X_7 and X_8, obtaining the matrix as follows.

$$(X_0, X_1, X_2, \cdots, X_8) = \begin{pmatrix} x_0(1) & x_1(1) & \cdots & x_8(1) \\ x_0(2) & x_1(2) & & x_8(2) \\ \vdots & \vdots & \ddots & \vdots \\ x_0(430) & x_1(430) & \cdots & x_8(430) \end{pmatrix} \tag{10}$$

The next step is to perform the dimensionless processing of the reference and comparison series. Because the different magnitudes of the physical activity data for subhealth and each factor are not easy to compare, or it is difficult to get the correct conclusion when comparing, the data are usually processed with dimensionless when performing the gray scale correlation analysis. We used the mean value method to process the 8 columns of data in a dimensionless way. whereX_i is the first column in equation (10) i column of data, and $\overline{X_\iota}$ is also the mean of the first column of data in equation (10) i the mean value of the data in the columns is obtained through this step X_0', X_1', X_2', X_3', X_4', X_5', X_6', X_7' and X_8'.

$$\overline{X_\iota} = \frac{1}{n}\sum_{k=1}^{n} X_i(k) \tag{11}$$

$$X_i' = \frac{X_i(k)}{\overline{X_\iota}}, k = 0,1,2,\cdots,n \tag{12}$$

The size of the difference between the curves can be used as a measure of the degree of association. For the comparison series after the second step of dimensionless quantization (X_1', X_2', X_3', X_4', X_5', X_6', X_7' and X_8') are differenced from the mean sequence (X_0') to make a difference and take the absolute value to indicate the magnitude of the difference.

$$\Delta_i(k) = |X_0'(k) - X_i'(k)|, k = 1,2,\cdots,n \tag{13}$$

Correlation coefficient $\gamma_{oi}(k)$ is calculated by the following formula.

$$\gamma_{oi}(k) = \frac{\Delta(min) + \rho\Delta(max)}{\Delta_i(k) + \rho\Delta(max)} \tag{14}$$

Where ρ is the resolution coefficient, generally between 0 and 1, usually taken as 0.5;

$\Delta(min)$ is the minimum difference between the two levels, i.e.$\Delta(min) = \min_i \min_k \Delta_i(k)$

$\Delta(max)$ is the maximum difference between the two levels, i.e.$\Delta(max) = \max_i \max_k \Delta_i(k)$

Up to this point, the degree of association between each comparison series and the reference series is represented by the n association coefficients, and it is necessary to do centralization of the association information. The degree of association is the average of the gray association coefficients, i.e:

$$\gamma_{oi} = \frac{1}{n}\sum_{k=1}^{n}\gamma_{oi}(k) \tag{15}$$

The calculated correlations are sorted in descending order, as shown in the table below, which shows that 1000 meters (correlation of 0.936) has the strongest effect on male students' subhealth, which may be due to the fact that 1000 meters of running can have a beneficial effect on the body, regulating the heart's cardiac function, helping to promote blood circulation, helping to expel wastes in the blood, enhancing the function of the lungs, and improving the cervical spine, etc., which can all help to Enhance immunity, improve physical fitness, mental health and social adaptation of male college students and so on.

According to the ordering of the sub-correlations in Table 28, it can be seen that the total score (with a correlation of 0.889) has the strongest effect on the subhealth of female students, which may be attributed to the fact that physical exercise regulates the mood and invigorates the spirit at the same time by enhancing the physical fitness and the will to better deal with the social relationships of female students.

(iii) Fitted curves of sub-health and physical activity

As it can be seen in the previous section, the correlation between 1000 meters and boys' subfertility is higher, while the correlation between total score and girls' subfertility is higher, so this paper tries to build a nonlinear implicit function fitting curve using Origin2021.

(1) Modeling

For boys, the dependent variable is the total subhealth and the time of 1000 meters running is the independent variable to find the quantitative relationship between subhealth and physical activity; for girls, the dependent variable is the total subhealth and the total physical activity is the independent variable to find the quantitative relationship between subhealth and physical activity. Figure 26 and Figure 27 were obtained respectively after several fits and iterations, and their specific models are as follows:

$$\begin{cases} F(x,y) = (\frac{x-xc}{a})^2 + (\frac{y-yc}{b})^2 - 1 \\ F(x,y) = 0 \end{cases} \tag{16}$$

Where $F(x,y)$ denotes the independent variable$x$ and the dependent variable y. The functional relationship between xc, yc, a, and b are all constants.

(2) Fitting effect test

According to Table 30, it can be seen that the degree of freedom is high, SSE (sum of squares of residuals) = 0.00, which is in the range of <0.01, and both R2 and adjusted R2 are 1. Therefore, the combined results of this can indicate that the model has an excellent fit.

Therefore, it is possible to roughly predict boys' subhealth from the 1000-meter running score, and girls' subhealth from the total score of the physical test.

(iv) Regression analysis of sub-health and physical activity

(1) CART decision tree classification algorithm modeling

1) Let the training dataset of the node be D. Calculate the Gini index of the existing features for this dataset

a. For each feature A, for each of its possible values$a$, based on the sample points for A=a the test of "yes" or "no" to the D Split into D_1 and D_2 two parts;

$$Gini(D) = \sum_{i=1}^{k} p_i(1-p_i) \tag{17}$$

b. Calculated from equation (18) A=a Gini index at the time of

$$Gini(D,A) = \sum_{v=1}^{2} \frac{|D_v|}{|D|} Gini(D_v) \quad (18)$$

2) At all possible features A and all their possible cut-points a among all possible features and all their possible cut-points, select the feature with the smallest Gini index and its corresponding cut-point as the optimal feature and optimal cut-point. Based on the optimal feature and the optimal cut-point, two child nodes are generated from the parent node, and the training dataset is assigned to the two child nodes according to the features.

3) Recursively call (1) and (2) on both child nodes until the stopping condition is satisfied.

4) Generate a CART decision tree.

By using the results of the various items in the physical test, we can directly get a rough score of subhealth.

(2) Analysis of results

The data of 30% of subfertility and physical activity of each male and female students were selected as the test set, and Tables 31 and 32 were obtained. According to the results of the model evaluation in Tables 31 and 32, it can be seen that comparing the predicted values to the case of using only the mean, the results of the R2 are always greater than 0.9, which indicates that the two models are more accurate.

According to Figures 18 and 19 it can be seen again that the decision tree for the effect of physical activity on subfertility in boys and the decision tree for the effect of physical activity on subfertility in girls have a high degree of accuracy.

The physical test scores can reflect the health status of college students, especially the strongest correlation between the 1000-meter running score and subhealth, indicating that physical exercise plays an important role in improving the health level of college students. Male students' total physical test scores are closely related to social adaptability, while female students' total physical test scores are more associated with physical subhealth and social subhealth. These findings emphasize the key role of physical exercise in improving college students' subhealth, especially the importance of long-distance running in enhancing physical fitness and social adaptability.

6. Research-based recommendations for university students to counteract sub-health conditions

(i) Increased publicity to raise health awareness

For university students as a group, schools should promote physical and mental health education and counselling through the addition of course content on sub-health, the organization of regular knowledge lectures, and the production of publicity materials. Various forms of publicity, such as lectures, elective courses, and on-campus publicity media, can be used to enhance the popularization of mental health knowledge. In addition, in order to safeguard students' health, it is recommended that personal health records be established for each student and regular medical check-ups be implemented for early detection, prevention and quick treatment of diseases. When feeling unwell, students should take the initiative to seek counseling and treatment from doctors or specialists, and pay attention to self-relief of stress, participate in more outdoor activities and maintain a happy mood.

(ii) Specific measures to combat sub-health

Physical exercise plays a significant role in preventing sub-healthy state. College students should realize the importance of exercise and integrate aerobic exercise such as running into their daily life to develop a long-lasting habit. In addition, participating in a variety of sports activities not only strengthens physical fitness and will, but also regulates mood and invigorates the spirit. In terms of physiological health, it is crucial to adapt to the biological clock, adjust rest and sleep schedules, and change bad lifestyles and habits.

Mental health should not be ignored as well. Bad moods are directly related to the occurrence of sub-health. Therefore, it is very important to stay relaxed and happy and maintain emotional balance. Renewing belief systems and keeping emotions free contribute to the overall health of the body, mind

and spirit. In terms of social health, keeping in touch with friends and sharing joys and sorrows is not only conducive to emotional catharsis, but also an effective way of preventing the transformation of subhealth into disease. College students need to have good social adaptability and interpersonal skills to cope with the ever-changing modern society.

Scientific health management is a necessary trend in the development of modern society. From now on, raising the awareness of "investing in health" and putting it into action will not only add value to the long-term health of the individual, but also that of the family, and will fundamentally solve the problem of sub-health.

References

[1] Jiang Feng. Current status and prospect of subhealth research [J]. Henan Traditional Chinese Medicine, 2004 (02): 9-10.

[2] Zhao Rui-Qin. Progress in the research and treatment of sub-health [J]. China Clinical Physician, 2006 (11): 7-9.

[3] ZHAO Rui-Qin, HAN Xiu-Xia, LIANG Jing. Current status and countermeasures of domestic and international research on sub-health problems [J]. Foreign Medical Intelligence, 2005 (09): 8-11.

[4] Chen Li. Research on the Causes and Intervention Measures of "Sub-health" of Students in Some Colleges and Universities in Wuhan [D]. Wuhan Institute of Physical Education, 2009.

[5] Fan Xinxin. Research on the sub-health condition of civil servants and its influencing factors in Shenyang city [D]. Southern Medical University, 2011.

[6] LI Zehui, LUO Man, ZHANG Yanfen. Research progress of subhealth status [J]. Journal of Community Medicine, 2009, 7 (13): 36-38.

[7] Feng Lingyan. Study on the correlation between pregnancy and childbirth and the state of sub-health and reproductive diseases in women of childbearing age [D]. Chengdu University of Traditional Chinese Medicine, 2009.

[8] Luo Man. A study on the current status of nurses' subhealth status and their influencing factors in a tertiary hospital in Urumqi [D]. Xinjiang Medical University, 2010.

[9] Peng Si-Min. A study on the current status and correlation between physical activity, sedentary behavior, academic stress and neck subhealth among high school students [D]. Sichuan Normal University, 2022.

[10] Li Menran. Analysis of college students' subhealth status and influencing factors based on Logistic and Decision Tree Modeling [D]. Beijing Jiaotong University, 2015.

[11] XU Jun, FENG Liyi, LUO Ren, QU Jincai, ZHANG Jinhua, ZHAO Xiaoshan, WEI Qian. A study on the reliability and validity of the subhealth assessment scale [J]. Journal of Southern Medical University, 2011, 31 (01): 33-38.

[12] Jiang Yan. A study on the current situation of college students' physical fitness and influencing factors in Jiangsu Province from 2003 to 2006 [D]. Suzhou University, 2006.

[13] Xu Zhong. Investigation and analysis of physical fitness of medical school students--Taking Xuzhou Medical College as an example [J]. Journal of Harbin Institute of Physical Education, 2012, 30 (03): 118-121.

[14] Hou Liangyong. Comparative analysis of the body morphology of Min-Chinese students in higher vocational colleges in Urumqi [J]. Journal of Xinjiang Vocational University, 2008 (03): 45-47.

[15] Bo Yang, Gu Xinhong, Ni Wei. The effects of BMI abnormalities on college students' physical fitness indicators [J]. China Sports Science and Technology, 2005 (02): 140-141+145.

[16] YANG Xin, SHENG Yancha. Research on the correlation between body mass index and college students' physical health indicators [J]. Hubei Sports Science and Technology, 2006 (01): 44-46.

[17] Cao Lichun. The effect of body mass index on college students' physical health indicators [J]. Journal of Daqing Normal College, 2009, 29 (06): 133-135.

[18] XU Fu-Fu, LIU Xin-Yi, GAO Xi-Guang. Exploring the application in radar electronic reconnaissance based on machine learning [J]. Information Record Material, 2021, 22 (04): 19-22.

[19] YANG Xixi, LI Xiangfang, WU Qiong. Gray correlation analysis of the effect of development technology on recovery rate of gas reservoir stage [J]. Rocky Reservoirs, 2011, 23 (01): 107-110+122.

[20] Niu Ruoling, Luo Gaoyuan. Gray correlation analysis of agricultural industry structure in Zhejiang Province [J]. Hunan Agricultural Science, 2013 (13): 116-119.

[21] Liu, S. F.. Gray system theory and its application (9th edition) [M]. Beijing: Science Press, 2021.

[22] Dafoe W. Extremely Severe ME/CFS-A Personal Account. healthcare. 2021; 9 (5): 504.

[23] Knight SJ, Politis J, Garnham C, Scheinberg A, Tollit MA. school functioning in adolescents with Chronic Fatigue Syndrome. front Pediatr. (2018) 6: 302.