MEDS Clinical Medicine (2025)
Clausius Scientific Press, Canada
DOI: 10.23977/medsc.2025.060218
ISSN 2616-1907 Vol. 6 Num. 2

Study on Spatio-Temporal Evolution of College Students' BMI and Its Correlation with Cardiopulmonary Health Indicators

Meiyan Han

School of Clinical Medicine, Guangdong Medical University, Dongguan, 523808, Guangdong, China
1499600731@qq.com

Keywords: College Students, Physical Health, BMI, Cardiopulmonary Function

Abstract: Based on the national college students 'physical health monitoring data from 2014 to 2023, this study systematically analyzed the spatio-temporal evolution characteristics of college students' body weight (BMI) and its association with cardiopulmonary function. The research results show that the overall BMI of Chinese college students (CCS) is in the normal range (BMI=22.22), but the individual differences are significant, and there is a two-way abnormal body weight phenomenon. In terms of time, BMI is increasing year by year, and the growth rate will accelerate after 2020, showing the profound impact of the epidemic on the weight of college students; in terms of space, BMI shows a geographical gradient of "high in the north and low in the south", reflecting regional diet, climate and The combined effect of lifestyle differences. In terms of cardiopulmonary function, BMI has a positive effect on vital capacity, but has a negative effect on endurance running performance, suggesting that it has a two-way effect on cardiopulmonary function with dimensional differences. Regression modeling results showed that height was the most stable predictor, while the effects of BMI and weight were heterogeneous among different groups. Gender and grade have a significant regulatory effect on the prediction model, and the correlation between physical fitness and exercise performance of male and lower-grade students is more significant. The study finally proposed that stratified and classified health intervention strategies should be formulated for different groups and regions to improve the overall physical level and cardiopulmonary health status of college students.

1. Introduction

1.1 Research Background and Significance

In recent years, with changes in lifestyle and eating habits, the body mass index (BMI) of college students has shown a significant trend of change. The body mass index (BMI) of college students and its relationship with cardiopulmonary metabolic health indicators have gradually attracted widespread Attention in the academic community. Research shows that in 2010, 76.84% of college students in Henan Province had a BMI in the normal range (18.5 - 24 kg/m^2), but some students still have low weight or overweight and obesity (Yang et al., 2013) [1]. BMI is not only an important indicator to measure an individual's nutritional status, but is also closely related to multiple health factors. Studies have pointed out that there is a significant negative correlation between BMI and physical health indicators such as lung capacity, 800-meter/1000-meter running, and standing long jump, indicating that excessive or low BMI may indicate a decline in physical health (Zhao, 2018)[2]. In addition, body shape indicators such as height and weight show a certain linear relationship with college students 'comprehensive physical health scores. By analyzing college students' height, weight, BMI and other indicators, a prediction model can be constructed to assist in health evaluation and risk screening (Xu, 2022)[3]. However, there have been relatively few systematic studies on the spatio-temporal evolution of college students 'BMI and its association with cardiopulmonary and metabolic health indicators. Exploring the spatio-temporal changes of college students 'BMI will help understand the differences in health status of students in different regions, genders and grades, and provide a basis for formulating targeted health intervention measures. At the same time, studying the association between BMI and cardiopulmonary function indicators can provide scientific basis for preventing cardiovascular diseases and improving college students 'physical health. Based on the national college student physical health test data from 2014 to 2023, this study systematically analyzed the development trend of college students 'BMI and its predictive effect on cardiopulmonary health from three aspects: temporal evolution, spatial distribution and population heterogeneity. By constructing a basic regression model, the mechanism of the impact of BMI on cardiopulmonary function in different gender and grade groups was identified, providing a practical reference for college student group health assessment and intervention strategies.

1.2 Literature Review

Body mass index (BMI) is an important indicator for assessing individual weight status and is widely used in public health and clinical medicine. As a result, the relationship between BMI and cardiopulmonary health indicators has become a research hotspot. This paper aims to review the spatio-temporal evolution of college students 'BMI and its association with cardiopulmonary health indicators, in order to provide reference for related research and intervention measures.

The temporal evolution of college students 'BMI reflects the impact of social and economic development, lifestyle changes and public health policies. Multiple studies have shown that BMI among college students has shown an upward trend in the past few decades. For example, Dong(2018) analyzed the long-term changes in the BMI of college students in China and found that from 1985 to 2014, the average BMI of college students increased significantly, and the rates of overweight and obesity also increased [4]. This trend is consistent with the global obesity epidemic. Research by Ogden et al. (2016) shows that BMI among college students has continued to increase over the past 30 years, especially among male college students. The increase in obesity rates is more significant [5].These research results suggest that the temporal evolution of college students 'BMI is closely related to socioeconomic development, changes in diet structure and reduced physical activity.

The spatial evolution of college students 'BMI reflects differences in geographical, cultural and economic factors. Zhang et al. (2018) conducted a study on the distribution of BMI among college students in different regions of China and found that the BMI of college students in the eastern coastal areas is generally higher than that in the central and western regions. This difference may be related to the higher level of economic development in the eastern region, the Westernization of the diet structure and reduced physical activity [6]. In cross-country comparisons, Popkin et al. (2012) found that there are significant differences in the distribution of BMI among college students in developed and developing countries. Obesity rates among college students in developed countries are generally higher, while malnutrition and overweight problems coexist among college students in developing countries [7]. These research results show that the spatial evolution of college students 'BMI is influenced by geographical, cultural and economic factors.

The relationship between BMI and cardiovascular health indicators has attracted much attention. Multiple studies have shown that too high or too low a BMI is associated with an increased risk of cardiovascular disease. For example, Wilson et al. (2002) found that BMI was positively correlated with blood pressure, lipid levels, and arteriosclerosis index [8]. The risk of cardiovascular disease among overweight and obese college students is significantly higher than that of normal-weight college students. In China, Chen et al. (2015) also showed that BMI is closely related to cardiovascular disease risk factors such as hypertension, hyperlipidemia and diabetes. The blood pressure and lipid levels of overweight and obese college students are significantly higher than those of normal weight college students [9]. These findings suggest that BMI is an important predictor of cardiovascular health.

The relationship between BMI and lung function has also received widespread attention. Studies have shown that too high or too low a BMI may affect lung function. For example, Chen et al. (2015) found that the vital capacity and maximum ventilation of obese college students were significantly lower than those of normal-weight college students. Obesity may cause limited movement of the chest wall and diaphragm, which may affect lung function [9]. On the other hand, the lung function of low-weight college students may also be affected. Dong et al. (2018) found that the lung capacity and maximum ventilation of low-weight college students were significantly lower than those of normal-weight college students. Malnutrition may lead to reduced respiratory muscle strength, which may affect lung function[4]. These research results show that BMI is an important factor in lung function. Cardiopulmonary endurance is an important indicator for evaluating cardiopulmonary health, and the relationship between BMI and cardiopulmonary endurance has also attracted much attention. Multiple studies have shown that too high or too low a BMI may affect cardiopulmonary endurance. For example, a study by Jackson et al. (1990) found that the maximum oxygen uptake of obese college students was significantly lower than that of normal-weight college students. Obesity may lead to increased cardiopulmonary stress, which may affect cardiopulmonary endurance [10]. In China, Zhang et al. (2018) also showed that BMI is negatively correlated with cardiopulmonary endurance. The maximum oxygen uptake and exercise endurance of overweight and obese college students are significantly lower than those of normal-weight college students[6]. These findings suggest that BMI is an important predictor of cardiopulmonary endurance.

The spatio-temporal evolution of college students 'BMI and its relationship with cardiopulmonary health indicators reflect the impact of social and economic development, lifestyle changes and public health policies. Time-evolution research shows that college students 'BMI has shown an upward trend in the past few decades, and the rates of overweight and obesity have increased significantly. Spatial evolution research shows that there are significant geographical, cultural and economic differences in the distribution of college students 'BMI. Studies on the relationship between BMI and cardiopulmonary health indicators have shown that too high or too low BMI may increase the risk of cardiovascular disease, affect lung function and cardiopulmonary endurance. These research results provide an important basis for formulating health intervention measures for college students.

Future research should further explore the spatio-temporal evolution mechanism of college students 'BMI and its relationship with cardiopulmonary health indicators, and provide scientific basis for formulating health intervention measures for college students. In terms of data dimension, this study uses a large sample of national college student physical examination data spanning ten years, which has the advantages of large time span, wide geographical range, and complete variable dimensions. In terms of method path, it adopts gender-grade-region multi-level cross-analysis design, and builds a basic prediction model of BMI on cardiopulmonary indicators such as lung capacity and endurance running to fill the gaps in this kind of cross-sectional research among domestic college students; In terms of practical orientation, the research focuses on "key indicator screening" and "high-risk group identification" in health tests, providing a simplified and efficient reference for colleges and universities to carry out precise health management.

1.3 Research Questions and Research Goals

Based on the rich sample resources of the National College Student Physical Health Test Database from 2014 to 2023, this study plans to focus on the following two core issues:

2014-2023 During the year 2000, did the BMI of China college students show significant temporal and spatial differences? Are there statistically significant differences in the trend of BMI change among different gender, grade and regional groups?

Among body shape indicators such as height, weight, and BMI, which indicators have stronger predictive power on college students 'cardiopulmonary function (lung capacity, endurance running performance, etc.)? Does this predictive ability vary depending on factors such as gender, grade, etc.?

This research plans to achieve the following two goals:

Based on the national college students 'physical health test data from 2014 to 2023, the spatio-temporal distribution characteristics of BMI are analyzed, and a geographical map is constructed to explore the change patterns of BMI in different regions, genders, and grades, and reveal the spatio-temporal evolution laws of college students' BMI.

Build a BMI-cardiopulmonary health prediction model to explore the predictive effects of body shape indicators such as height, weight, and BMI on cardiopulmonary function, identify the most predictive indicators, and analyze the heterogeneity of gender and grade.

2. Data and Methodology

2.1 Data Sources

The data used in this study comes from the National College Student Physical Health Test Database from 2014 to 2023. The database is uniformly organized and implemented by colleges and universities in various provinces and cities in accordance with the "National Student Physical Health Standards (Revised in 2014)", covering physical health testing data of college students from different regions, genders, and grades across the country. The data content includes core variables such as basic body shape indicators (such as height, weight, BMI), and cardiopulmonary function indicators (such as lung capacity, endurance running performance). It has the characteristics of large sample size, wide coverage, and strong continuity, and can be relatively comprehensive. It reflects the physical health status and changing trends of my country's college student population over the past decade. Based on strict data cleaning and desensitization, this study selected data samples with complete records and meeting analysis requirements for further statistical modeling and analysis.

2.2 Definition of Key Variables

The definitions of key variables in this study are shown in Tables 1 and 2 below.

In order to ensure the rationality of data analysis, indicators such as BMI, vital capacity, and endurance running will be unit converted and missing values will be processed during the variable processing process; categorical variables such as gender and region will be coded before modeling; all variables will pass descriptive statistics to test the rationality and consistency of their distribution.

Table 1 Interpretative variable (independent variable)

NO.	Var. name	Type	Unit	Explanation
1	BMI	Continuous	kg/m ²	Calculated as weight / height ²
2	height	Continuous	cm	Original variable
3	weight	Continuous	kg	Original variable
4	gender	Binary	—	0 = Female; 1 = Male
5	grade	Categorical	—	1 = Fr.; 2 = So.;3 = Jr.; 4 = Sr.
6	year	Continuous	Year	Range: 2014–2023
7	region	Categorical	—	Classified into Eastern / Central / Western / Northeastern

Table 2 Interpreted variable (dependent variable)

NO.	Var. Name	Type	Unit	Explanation
1	Vital capacity	Continuous	ml	Indicates lung function
2	Vital capacity score	Continuous	points	Score calculated from vital capacity
3	Vital capacity level	Categorical	—	National fitness grading (Excellent/Good/Pass/Fail, etc.)
4	Endurance run (1000m for males, 800m for females)	Continuous	seconds	Different items used based on gender
5	Endurance score	Continuous	points	Score calculated from endurance event
6	3000m/1500m run (if provided)	Continuous	seconds	Optional variable
7	Total physical fitness test score	Continuous	points	Overall physical fitness evaluation indicator
8	Standardized score (weighted total score)	Continuous	points	Used for standardized result comparison

2.3 Data Analysis Method

In order to systematically answer research questions and achieve research goals, this study uses methods such as descriptive statistics, time series analysis, spatial visualization, difference testing and multiple regression to build an analytical framework based on the nature of variables and analysis levels.

2.3.1 Descriptive Statistical Analysis

First, the core variables of college students from 2014 to 2023, including body shape indicators (such as BMI, height, weight) and cardiopulmonary function indicators (such as lung capacity, endurance running performance), were analyzed statistically on the central trend and dispersion. The indicators included mean, standard deviation, maximum and minimum. In order to further reveal the differences in basic characteristics between different groups, cross-sectional comparisons were conducted according to classification variables such as gender, grade, and region to explore the overall distribution pattern in terms of physical characteristics.

2.3.2 Spatio-temporal Evolution Analysis

In the time dimension, linear trend analysis and local weighted regression (LOESS) methods are used to fit and model the annual changes in BMI, and time series graphs are drawn to identify the trend and fluctuation rules of BMI changes over time. In terms of spatial dimensions, combined with place of birth information, a spatial distribution map of college students 'BMI in each region is constructed, and regional differences are visualized through the Python visualization module. At the same time, one-way analysis of variance (ANOVA) was used to analyze the statistical significance of BMI differences among different genders, grades and regional groups, and Tukey's post-hoc test was also used.

2.3.3 Predictive Modeling of Cardiopulmonary Function

In order to explore the predictive ability of body shape indicators to cardiovascular and pulmonary function of college students, Pearson and Spearman correlation analysis were first used to evaluate the correlation between dependent variables such as BMI, height, weight and lung capacity, and endurance running performance. On this basis, a multiple linear regression model was constructed, with lung capacity and endurance running time as dependent variables, and BMI, height, and weight as core independent variables, and variables such as gender, grade, and were controlled to identify their independent predictive effects. In order to improve the accuracy and simplicity of the model, step-wise regression or LASSO regression methods are further used to screen out the key morphological variables with the most explanatory power. In addition, in order to compare the applicability of the prediction model among different populations, hierarchical regression analysis was performed on gender and grade subsamples to explore the heterogeneity of prediction power.

3. Results

3.1 Temporal and Spatial Evolution Characteristics of College Students 'BMI

3.1.1 Overall Distribution of College Students 'BMI

According to summary statistics on the physical health data of college students nationwide from 2014 to 2023, a total of 123,281 valid subjects in the sample recorded BMI values. The BMI values of the overall sample ranged from 14.10 to 45.93, showing a large distribution range. The average BMI is 22.22 (standard deviation 3.42), indicating that overall the weight status of the college student population is within the "normal weight" range (World Health Organization standards of 18.5 - 24.9). It can be seen from the minimum and maximum values that the sample includes both severely underweight individuals and severely overweight individuals, indicating that there is a certain degree of difference in weight structure within the group. The standard deviation reaches 3.42, which further reflects the high degree of dispersion of the BMI distribution and significant individual differences among different students. See Table 3. Overall, the results show that the overall BMI control of college students in my country has been within a relatively reasonable range in the past ten years, but some students are at risk of weight imbalance, especially the existence of high BMI individuals, which should attract the attention of health education and weight intervention work in colleges and universities.

Table 3 Descriptive statistics

	N	Minimum	Average	Standard deviation
BMI	123281	14.1044	22.224644	3.4242619
Valid N (list status)	123281			

3.1.2 Temporal Trend Analysis of BMI

Figure 1 shows the annual change trend of the average BMI of college students from 2014 to 2023. It can be observed from the figure that the average BMI of college students shows an overall upward trend year by year, but this trend is not completely linear, but shows certain periodic fluctuations. During the period from 2014 to 2018, the overall average BMI increased slowly, rising from approximately 21.85 to approximately 22.27, and the growth rate was relatively stable. There was a slight decline in 2019, and the average BMI fell back to 2016 levels. However, since 2020, the BMI level has increased significantly. After 2020, the average BMI has jumped rapidly and reached its highest value at this stage in 2023, exceeding 22.70. This trend shows that college students 'BMI will show a significant upward turning point after 2020, which is highly consistent with the background of epidemic prevention and control since 2020, and may be closely related to lifestyle factors such as home study, reduction in physical exercise, and changes in diet structure. Judging from the trend trend, the rate of increase in BMI between 2020 and 2023 is significantly higher than that between 2014 and 2018, indicating that the rate of weight gain of college students has accelerated in recent years, posing a potential risk warning for physical health. The finding suggests that colleges and universities should strengthen monitoring and intervention in students 'physical condition in the later stages of the epidemic, especially in the reconstruction of physical exercise after normal teaching order is restored.

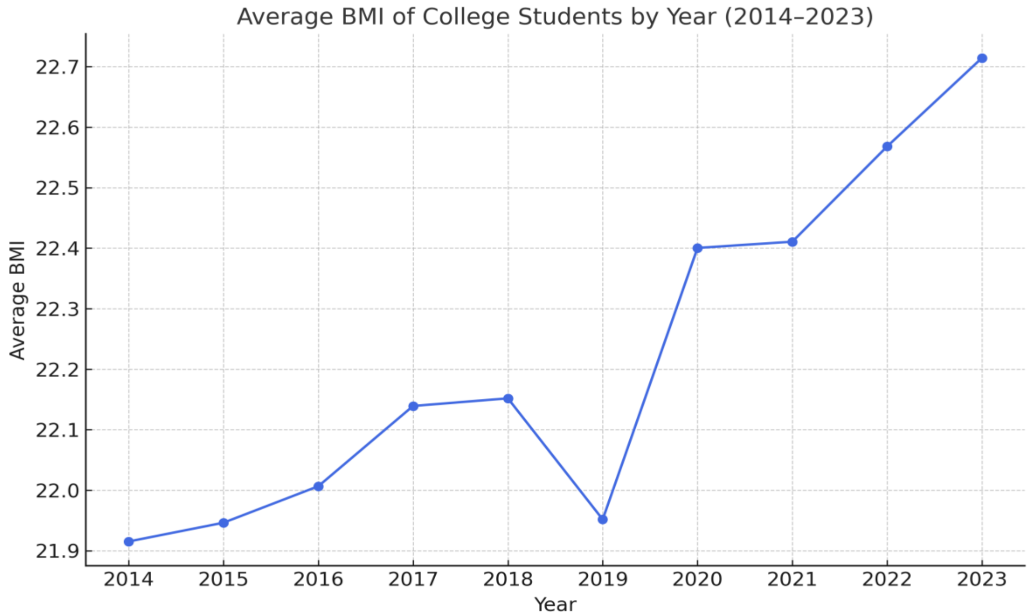

Figure 1 Trend chart of mean BMI among CCS

In order to test whether there is a significant annual difference in college students 'BMI between 2014 and 2023, this paper uses One-Way ANOVA to conduct a statistical test. The results are as shown in the table. The difference in mean BMI between different years was statistically significant ($F = 85.533$, $df = 9$, $p < .001$). This shows that over the past decade, the BMI level of college students has not been stable, but has shown systematic fluctuations over time. Specifically, the inter-group variance is much larger than the intra-group variance (inter-group mean square = 996.776; intra-group mean square = 11.654), further indicating that the factor "year" has a significant impact on BMI.

3.1.3 Spatial Distribution and Regional Differences in BMI

In order to explore the distribution characteristics of college students 'BMI in spatial dimensions in our country, this paper uses python to calculate and draw a nationwide thermal map of the spatial distribution of college students' BMI mean (see Figure 2) based on the birthplace information of each province. The darker the color in the figure, the higher the mean BMI, and the lighter the color, the lower the mean, showing a clearer geographical gradient difference.

Overall, the mean BMI in my country shows a significant spatial pattern of "high in the north and low in the south". The average BMI of college students in Northeast China (Heilongjiang, Jilin, Liaoning), North China (Inner Mongolia, Hebei, Shanxi), and some northwest provinces (Gansu, Ningxia, etc) is relatively high, generally higher than 22.6, showing a high weight level. Especially in Heilongjiang, Jilin and other provinces, the average BMI is close to or even exceeds 22.8, ranking among the top in the country.

In contrast, the average BMI in the southeastern coastal areas and southern regions is relatively low. For example, in Guangdong, Fujian, Hainan, Jiangxi, Zhejiang and other provinces, the average BMI is mostly below 22.0, and in some regions such as Hainan and Fujian, it is close to 21.5, showing the overall trend of lower weight among college students.

The BMI levels in central regions such as Henan, Hubei, Hunan and other provinces are relatively middle, with the mean concentration at around 22.2, which is neither a high value area nor significantly low. In addition, due to geographical and dietary differences, BMI performance in plateau areas such as Qinghai is relatively moderate and no extreme values occur.

This spatial distribution pattern may be affected by multiple factors, including regional dietary structure, living habits, climatic conditions, physical exercise methods and social and economic levels. For example, winters in the northern region are generally cold, staple foods are high in calories, and differences in physical activity patterns may lead to high BMI; while the southern region has a hot and humid climate, rich vegetable intake, and residents generally pay more attention to body shape control, which may lead to the overall low BMI level.

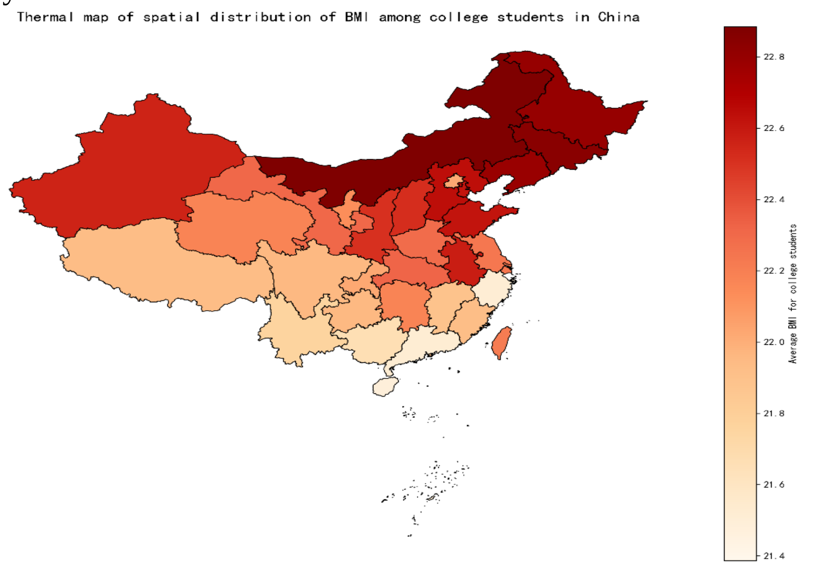

Figure 2 Heatmap of the spatial distribution of BMI among CCS.

In order to further clarify the differences in BMI of college students in different regions in our country, based on the national standard regional division plan, this paper divides 34 provincial administrative units into four major regions: the east (such as 13 provinces and cities such as Beijing, Shanghai, and Jiangsu), the middle (such as Henan, Hubei, and Hunan), the west (such as 12 provinces and regions such as Sichuan, Shaanxi, etc), and the northeast (Liaoning, Jilin, and Heilongjiang), and calculates the mean BMI of college students in each region (see Table 4). One-way analysis of variance (One-Way ANOVA) was then used to test significance for differences in mean between the four major regions. The analysis results showed that there were statistically significant differences in the mean BMI of college students among the four major regions ($F = 3.866$, $p = 0.01889$).

Further Tukey's post-hoc test results showed that the differences in mean BMI between Northeast and East, and between Northeast and Central were significant ($p < .05$), the differences between West and East were also significant, and the differences between Central and East were marginally significant or not significant.

Table 4 Mean BMI by region

Area	Mean BMI
Northeast	22.812128
Central	22.298242
Eastern	22.023628
Western	22.157077

These results are highly consistent with the spatial gradient trend shown by the heat map, further confirming that the BMI of college students in my country has systematic regional differences. The formation of differences may be related to the combined effect of multiple factors such as regional economic development level, dietary structure, allocation of physical education resources, and life rhythm. Especially when implementing policies to promote college students 'physical health, it is urgent to combine regional characteristics and adapt local conditions to improve the adaptability and accuracy of intervention strategies.

3.1.4 Population Heterogeneity Analysis

1) Comparison of mean BMI of male and female students (t-test)

Table 5 shows that there is a significant difference in the mean BMI between male and female students ($t = 86.79$, $p < 0.01$). The mean BMI of male students was significantly higher than that of female students, which may be related to differences in male and female students 'physical structure, physiological characteristics, and living habits (such as diet and exercise).

Table 5 Gender BMI comparison (t-Test)

Variable	Group 1 (male)	Group 2 (female)	T-statistic	P-value
BMI	84057	39224	86.79328191	0

2) Comparison of mean BMI among students in different grades (ANOVA)

Table 6 shows that there are significant differences in mean BMI between different grades ($F = 463.39$, $p < 0.01$). With the increase of grade, there is a certain trend in the mean BMI of students, see Figure 3.

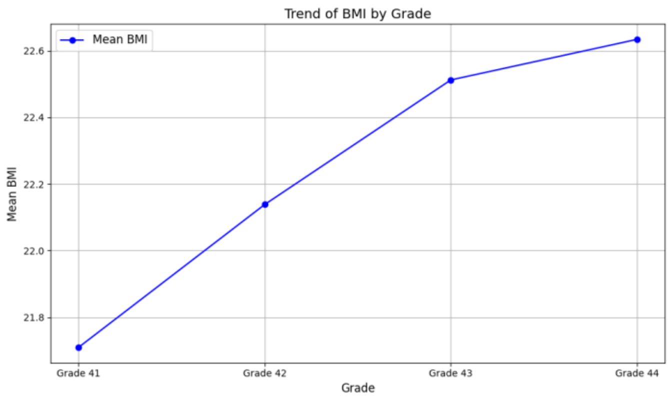

Figure 3 Trend of BMI by grade

Table 6 Grade BMI comparison (ANOVA)

Grade	Sample size	Mean BMI	F-Statistic	P-Value
41	33108	21.70996401	463.3872069	0.0000000
42	32858	22.13834805		
43	29283	22.51194316		
44	28032	22.63355443		

3.2 Differences in Cardiopulmonary Function among BMI Categories

In this study, samples were divided into four groups: Underweight, Normal, Overweight and Obese based on the BMI classification standard to explore the differences in the two cardiopulmonary function indicators of Vital Capacity and Endurance Score between different BMI groups.

To test whether the differences between BMI groups were statistically significant, one-way ANOVA was performed on lung capacity and endurance scores. ANOVA results showed that there were very significant differences in both indicators between different BMI groups (vital capacity: $F = 5496.10$, $p < 0.001$. Endurance score: $F = 5112.39$, $p < 0.001$), indicating that BMI classification has a significant impact on cardiopulmonary function.

Tukey HSD post-hoc multiple comparisons were further used to analyze specific inter-group differences. The Tukey test results for lung capacity are as follows: The vital capacity of the obese group was significantly higher than that of the normal weight group (Mean difference =707.72 ml, $p < 0.001$), the overweight group was also significantly higher than that in the normal weight group (Mean difference =510.32 ml, $p < 0.001$), vital capacity in the lean group was significantly lower than that in the normal group (Mean difference =-440.23ml, $p < 0.001$), the obesity group was significantly higher than the lean group (Mean difference = 1147.95ml, $p < 0.001$), the overweight group was significantly higher than the lean group (Mean difference = 950.55ml, $p < 0.001$), and the obesity group was significantly higher than the overweight group (Mean difference = 197.40ml, $p < 0.001$). It can be seen that lung capacity shows a trend of increasing with the increase of BMI. The median and maximum lung capacity in the obese group are significantly higher than other groups, while the lean group is generally lower. See Figure 4.

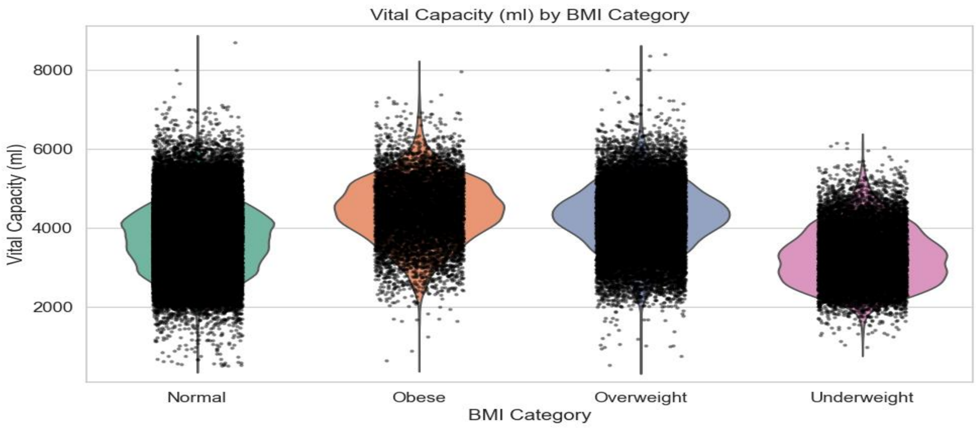

Figure 4 Vital capacity(ml) by BMI category

There were also significant differences in endurance scores and showed an opposite trend to lung capacity. It can be observed from the violin chart that the endurance scores in the normal weight group are the most concentrated and have the highest scores, while the endurance scores in the obese group and the lean group are significantly lower, the range of score fluctuations is more scattered, and the proportion of individuals with extremely low distribution values at the bottom is higher, suggesting that there are obvious deficiencies in the performance of these two groups in endurance running. See Figure 5.

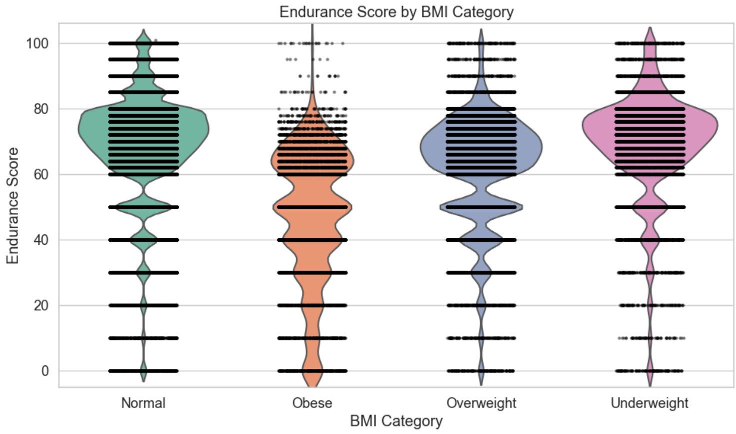

Figure 5 Endurance score by BMI category

To sum up, BMI has a two-way impact on the cardiopulmonary function of adolescents: obese and overweight individuals generally have higher lung capacity, but their endurance performance is poor; lean individuals are at a disadvantage in both lung capacity and endurance indicators. This result suggests that different dimensions of cardiopulmonary function are affected by BMI in different ways, and weight control should strike a balance between promoting lung function and endurance.

3.3 Predictive efficacy of body shape indicators on cardiopulmonary function

3.3.1 Correlation Analysis between Variables

Table 7 Morphological& cardiopulmonary indicators relationship

Morph. indicator	Cardiopul. indicator	Pearson correlation coefficient	Pearson p-value	Spearman correlation coefficient	Spearman p-value
BMI	Vital capacity (ml)	0.372744681	0	0.384587906	0
BMI	3000/1500m run (s)	0.36086874	0	0.352711175	0
Height (cm)	Vital capacity (ml)	0.711794184	0	0.717510913	0
Height (cm)	3000/1500m run (s)	0.604520584	0	0.544677856	0
Weight (kg)	Vital capacity (ml)	0.644488939	0	0.680903006	0
Weight (kg)	3000/1500m run (s)	0.577123164	0	0.569381206	0

In order to explore the predictive power of college students 'physical shape indicators on cardiopulmonary function, this paper selected three shape indicators: height, weight, and BMI, and conducted Pearson and Spearman grade correlation analysis with lung capacity and endurance running (3000 meters/1500 meters running seconds) respectively. The results are shown in Table 7.

The results showed that the overall trend between the two types of correlation coefficients was consistent, and both showed significant statistical significance ($p < 0.001$), indicating that the correlation between variables was stable. Among the Pearson correlations, height and vital capacity had the highest correlation coefficient ($r = 0.712$), followed by weight ($r = 0.644$) and BMI ($r = 0.373$); Spearman analysis also showed a similar ranking. In terms of the negative relationship with endurance running (the shorter the running time, the better the physical fitness), the correlation between height and endurance running is also strong ($r = 0.605$), followed by weight and BMI.

In summary, height showed a high positive correlation with both of the three morphological indicators, especially with vital capacity, indicating that its role in predicting cardiopulmonary function is more significant. In addition, although BMI is commonly used in medicine to classify physical fitness, its correlation with cardiopulmonary function is relatively weak, suggesting that it should be used with caution when used as a predictor.

3.3.2 Multiple Linear Regression Model Results

1) Modeling methods and regression equations

In order to explore the extent and direction of the influence of physical fitness indicators on college students 'vital capacity and endurance running performance, this study used Multiple Linear Regression model and used Ordinary Least Squares (OLS) to fit the following two models:

Model 1: Predicted vital capacity (unit: ml) is the dependent variable, and independent variables are BMI, height (unit: cm) and weight (unit: kg).

Model 2: Predicted endurance running time (unit: seconds, refers to 1500 meters (male) or 3000 meters (female)) is the dependent variable, and the independent variable is the same as above.

The basic form of the model is as follows:

$$Y=\beta_0 +\beta_1 \cdot BMI+\beta_2 \cdot Height +\beta_3 \cdot Weight +\varepsilon \tag{1}$$

Where, Y represents the dependent variable (vital capacity or running time), β_0 is a constant term, β_1, β_2, and β_3 are the regression coefficients of their respective variables, and ε is the error term.

Prior to modeling, the data has been preprocessed as follows: missing values and extreme outliers are eliminated, units are normalized, and there is no complete multicollinearity between variables. The model was built using the statsmodels module in Python, and the significance test criterion was set to $\alpha=0.05$.

2) Model results and interpretation

The effects of BMI, height and weight on vital capacity are shown in Table 8 below.

Table 8 Effects of BMI, height and weight on vital capacity

Variable	Regression Coefficient	P-value	Relationship Direction	R-squared	F-statistic
Const	-11338.91	<0.001	—	0.577	41800
BMI	103.46	<0.001	Positive		
Height (cm)	79.28	<0.001	Positive		
Weight (kg)	-11.25	<0.001	Negative		

$R^2=0.577$, indicating that the model can explain about 57.7% of the variability in lung capacity. The F statistic is 41800, $p < 0.001$. The model is overall significant, and all independent variables reach statistical significance at the 0.001 level.

Vital capacity is significantly positively correlated with BMI and height, indicating that people with higher body mass index or taller stature have relatively strong lung function; while the negative coefficient of the weight variable shows that after controlling for BMI and height, the increase in weight itself may be related to the decrease in vital capacity, which may reflect the existence of factors that are detrimental to lung function in the weight composition (such as high fat content). However, since the model does not include variables such as body fat ratio, this explanation still needs to be treated with caution.

BMI, The influence of height and weight on endurance running time is shown in Table 9.

Table 9 Effects of BMI, height and weight on endurance running time

Variable	Regression coefficient	P-value	Relationship direction	R-squared	F-statistic
Const	-4454.99	<0.001	—	0.445	24580
BMI	71.95	<0.001	Positive		
Height (cm)	28.08	<0.001	Positive		
Weight (kg)	-18.36	<0.001	Negative		

R^2=0.445, the model can explain about 44.5% of the variance in endurance running time, the F statistic is 24580, $p < 0.001$, which is significant overall, and all variables also reach the significance level.

BMI and height have a positive impact on endurance running time, indicating that the higher these two indicators, the longer the running time is to complete, that is, the relatively poor exercise performance; this trend may be related to weight burden, exercise coordination or other physiological factors. Body weight has a negative effect on running time. This result is inconsistent with common sense expectations and may reflect the existence of multiple collinearities or interactions among independent variables in the model, resulting in disturbance in the direction of the regression coefficient. This result suggests introducing body fat rate and muscle mass in subsequent analyses or standardizing variables to further test robustness.

Both models showed good fit, and the influence direction of each variable on vital capacity and running time was consistent. BMI and height were generally positive effects, while weight was negative in both models. It is necessary to further combine sex, body fat rate and other control variables to do hierarchical regression to improve the interpretation strength.

3.3.3 Analysis of Model Applicability under Different Groups

In order to explore the applicability of the model in different groups, this study conducted group modeling and analysis of sample data based on gender and grade. In the regression model, BMI, Height and Weight were used as independent variables to predict Vital Capacity and Running Time respectively, and differences in model adaptability were analyzed by comparing the fitting effects and the significance of variables in different groups. The following are the specific analysis results:

1) Gender-based modeling

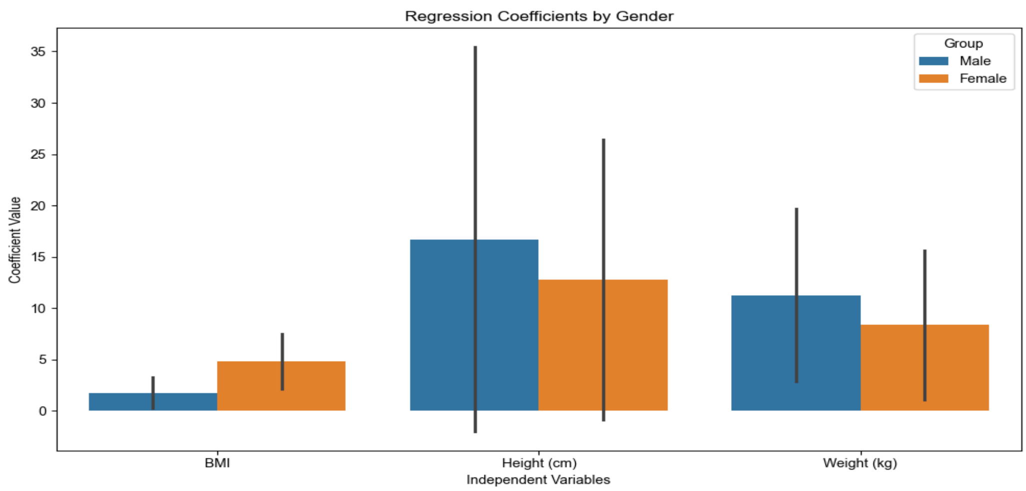

Figure 6 Regression coefficients by gender

By modeling male and female students separately, the fitting effect of the models and the significance of variables were compared. Figure 6 shows the comparison of the regression coefficients of BMI, height, and weight on predictive variables among male and female students.

It can be seen from the figure that the regression coefficient of BMI is positive among both male and female students, indicating that an increase in BMI will lead to longer running time, but the regression coefficient of men is slightly higher than that of women, indicating that BMI has a more significant impact on men's running performance. Effect of height on vital capacity: When predicting vital capacity, the regression coefficient of height for both male and female students is positive, and the influence coefficient of men is significantly higher than that of women, indicating that height contributes more to male vital capacity. Role of weight: When weight predicts running time and lung capacity, the regression coefficients are negative, indicating that weight gain will reduce running efficiency or lung capacity level, but the difference between male and female students is small. In addition, it can be seen from Figure 7 of residual error distribution that when predicting vital capacity in the male population, the model residual error is approximately normal, indicating that the model fitting effect is good.

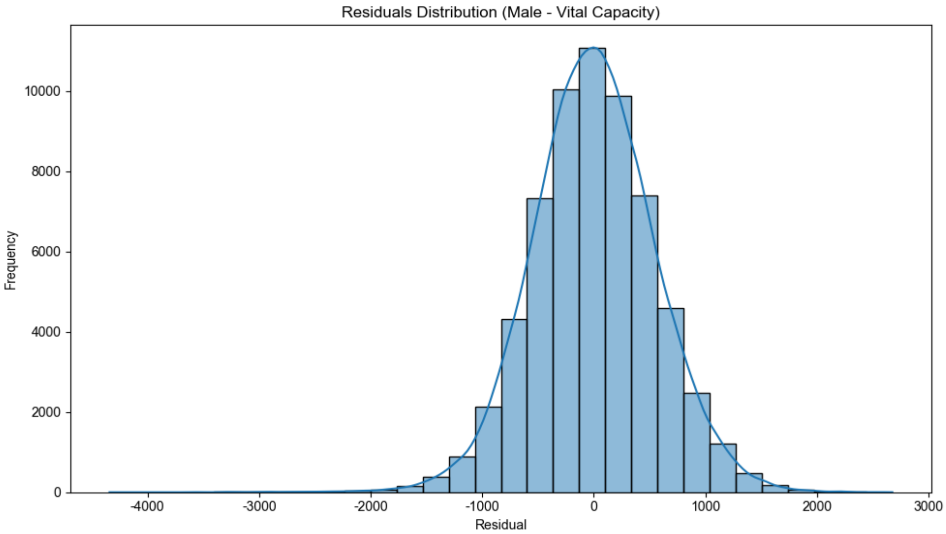

Figure 7 Residuals distribution (male-vital capacity)

2) Modeling by grade

There may be significant differences in physical characteristics and exercise performance among students in different grades, so further modeling will be carried out by grade to explore the differences in model adaptability between lower and upper grade groups. See figure 8. By comparing the regression analysis results of different grades, the main findings are as follows:

The regression coefficient of BMI varied significantly between different grades. The regression coefficient from Grade 41 to Grade 43 gradually increased to a maximum of about 120, indicating that among lower grade students, BMI has a more significant impact on predictive goals such as running time or lung capacity. The regression coefficient for Grade 44 dropped significantly, to only about 20, indicating that the importance of BMI has been significantly reduced for senior students. The regression coefficient of height is relatively stable among different grades, ranging from 60 to 100. The regression coefficient from Grade 41 to Grade 43 is slightly higher than that from Grade 44, but the overall difference is small, indicating that the impact of height on students 'sports performance is consistent. This suggests that height, as a physical characteristic, has a relatively stable contribution to sports performance throughout the college stage. The regression coefficient of weight showed negative values among different grades, but the change trend was obvious. The regression coefficient from Grade 41 to Grade 43 is negative, indicating that weight gain has a inhibitory effect on exercise performance, such as running time. The regression coefficient of Grade 44 turned from negative to positive, indicating that among senior students, the negative impact of weight on predictive goals has weakened, and may even begin to show a positive effect.

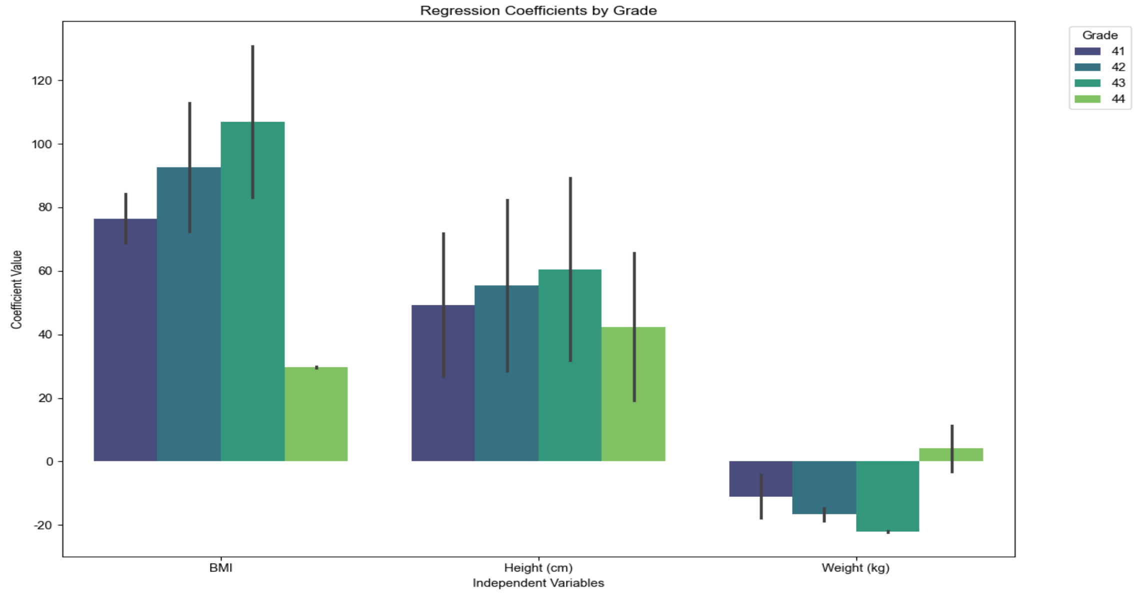

Figure 8 Regression coefficients by grade

To sum up, there are significant differences in model adaptability among different genders and grades. Among men, BMI and height have significant effects on exercise performance, while weight has a weak effect. Grade grouping analysis showed that the physical characteristics of lower-grade students had a more direct impact on their exercise performance, while the model fitting effect of upper-grade students decreased slightly, which may be related to the increased interference from other external factors. This result provides data support for the formulation of intervention measures for college students 'physical health, and recommends that differentiated measures be taken for different genders and grades when designing intervention strategies.

4. Discussed

Based on the physical health data of college students nationwide from 2014 to 2023, this study systematically analyzed the spatio-temporal evolution characteristics of BMI and its association with cardiopulmonary function indicators, revealing the dynamic changes, regional differences and impact on health of college students. The following is based on the research findings and discussed from the following four aspects:

4.1 Potential Factors Influencing the Temporal Trend of BMI

The study found that the average BMI of college students showed an overall upward trend in the past ten years, especially after 2020, there was a significant growth inflection point. This phenomenon may be closely related to epidemic prevention and control measures since 2020. Lifestyle adjustments such as home study, reduced physical exercise and changes in diet structure may be the direct causes of the rapid rise in BMI. In addition, the study also found that BMI growth rates accelerated after 2020, suggesting that the long-term impact of the epidemic on college students 'weight should not be ignored. After colleges and universities resume normalized teaching, they need to focus on the reconstruction of students 'physical exercise habits to curb the continuous rise in BMI. This finding is consistent with international research results. For example, Han et al. (2021) A comparative study of physical fitness before and after the epidemic conducted by many universities in China pointed out that college students generally lacked medium to high-intensity physical exercise during the epidemic, and their BMI increased significantly (average increase of 0.51, $p < 0.001$), especially among male students [11].

4.2 Regional differences in Spatial Distribution of BMI and Its Causes

Research reveals the spatial pattern of college students 'BMI being "high in the north and low in the south" in my country. The mean BMI in Northeast, North and Northwest regions is significantly higher than that in southeastern coastal areas and southern provinces. This difference may be caused by the combination of multiple factors: the northern region has cold winters and the diet structure is biased towards high-calorie staple foods, while the southern region has a hot and humid climate, more vegetable intake and residents pay more attention to body shape management. In addition, differences in regional economic development levels, allocation of physical education resources and living rhythm may also have an impact on the distribution of BMI. This finding is consistent with the research results of Wang and Hu (2017)[12]. This finding suggests that health intervention policies in colleges and universities need to adapt to local conditions and formulate differentiated measures based on regional characteristics.

4.3 Predictive efficacy of Body Shape Indicators on Cardiopulmonary Function

This study used multiple linear regression models to find that height, weight and BMI have significant effects on college students 'vital capacity and endurance running performance. Height plays a more significant role in predicting cardiopulmonary function, especially when it is highly correlated with lung capacity. Although BMI is commonly used in medicine to classify physical fitness, its correlation with cardiopulmonary function is relatively weak, suggesting that it should be used with caution when used as a predictor. In addition, model applicability analysis in different groups shows that BMI and height have a more significant impact on exercise performance in male groups, while physical characteristics of lower grade students have a more direct impact on exercise performance. These findings provide a scientific basis for universities to further optimize the physical health assessment system and formulate personalized exercise intervention plans. This finding is consistent with researcher Chen et al. The conclusions of (2022) are consistent. Chen et al. analyzed the body shape and lung function of China adolescents (aged 18 - 22) and found that BMI was positively correlated with vital capacity ($r = 0.35$, $p < 0.01$), but was significantly negatively correlated with the 12-minute endurance running performance, suggesting that lung ventilation capacity and aerobic endurance were affected by BMI in different directions [13].

4.4 Inspiration of Group Heterogeneity on Health Intervention

Studies have shown that changes in the regression coefficient of BMI for senior students suggest that as grade increases, other external factors (such as academic pressure, employment anxiety, etc.) may gradually increase their impact on weight and health. There are significant differences in the mean BMI of male and female students, and the changing trends of BMI among students in different grades are different. The higher mean BMI in men may be related to their physiological characteristics and living habits; this finding is consistent with Wang et al. (2020), which pointed out that among college students in China, the physical fitness indicators (BMI, weight) of the male group are more correlated with cardiopulmonary function, suggesting that future exercise intervention strategies should consider gender-specific design [14].These findings suggest that health interventions in colleges and universities need to adopt hierarchical strategies based on gender and grade differences, such as designing more strength training programs for boys and providing psychological counseling and health management support for senior students.

5. Conclusions and Limitations

5.1 Study Summary

Based on the national college student physical health test data from 2014 to 2023, this study systematically analyzed the spatio-temporal evolution characteristics of college students 'BMI and its association with cardiopulmonary function indicators. The main research conclusions are as follows: The overall BMI of college students is within a reasonable range, but the individual differences are significant. The average BMI of the overall sample is 22.22, which is within the World Health Organization's "normal weight" standard interval, but the standard deviation is large (3.42), indicating that college students have significant differences in weight structure, and some students are at risk of being thin or obese. BMI shows an increasing trend year by year in the time dimension, and its growth accelerates after 2020. Analysis shows that BMI has shown a significant upward trend after the outbreak of the epidemic in 2020, suggesting that the epidemic has a profound impact on college students 'lifestyles and is the focus of current physical intervention work. The spatial distribution of BMI shows a pattern of "high in the north and low in the south". The BMI of students in Northeast, North and Northwest China is significantly higher than that in southeastern coastal areas and southern provinces, indicating that regional factors (such as diet, climate, economic development level, etc.) have an important impact on the weight status of college students. BMI, height, and weight have significant predictive effects on cardiopulmonary function, but the direction of influence is different. BMI is positively correlated with vital capacity, but negatively correlated with endurance running performance, indicating that weight gain is beneficial to ventilation but not conducive to endurance performance. Height is a positive influencing variable in both models and is a stable indicator of cardiovascular health. Gender and grade have regulatory effects on the physical-health model. The correlation between BMI and exercise performance is stronger in the male group, and the physical fitness of lower-grade students has higher predictive power on cardiopulmonary function than in upper-grade students, suggesting that intervention strategies need to be differently designed due to group heterogeneity.

To sum up, this study reveals the dynamic evolution mechanism of college students 'BMI and its two-way impact on cardiopulmonary health, and provides data support and policy basis for physical health management and intervention in colleges and universities.

5.2 Study Limitations

Although this study has certain advantages in terms of data size and depth of analysis, it still has the following limitations: The data source is limited to existing physical health testing systems. Although the sample size is large, some variables (such as eating habits, mental health, family background, etc.) are missing, which limits a comprehensive explanation of the mechanism of BMI change. The study design was a cross-sectional + annual serial analysis. Although trends could be observed, it was not possible to confirm the dynamic action path between changes in BMI and cardiopulmonary function from the causal perspective. The cardiopulmonary function indicators are relatively single and mainly rely on two indicators: lung capacity and endurance running. In the future, more dimensions (such as maximum oxygen uptake, static heart rate, etc.) can be introduced to improve the comprehensiveness of the assessment. Not including lifestyle variables as control factors, such as physical exercise frequency, sleep quality, diet structure, etc., may have a confusing impact on the study results.

Based on the above limitations, future research can be expanded in the following directions: Researchers can build a longitudinal tracking database to conduct follow-up studies on BMI and health indicators during the university years, aiming to reveal causal relationships between weight changes and health outcomes. Researchers may integrate multimodal health data, such as data from physical activity sensors, nutritional intake records, and psychological assessments, to construct a higher-dimensional analytical framework linking health behaviors, physical fitness status, and health outcomes among university students. It is recommended to strengthen spatial modeling and intervention simulations by combining GIS-based spatial analysis with data on regional resource allocation, in order to simulate the effectiveness of health intervention strategies at universities in different regions and improve the precision of related policies. Researchers can introduce machine learning methods to optimize prediction models. By using existing physical health data, it is possible to develop individualized BMI prediction and cardiopulmonary function evaluation models, providing intelligent decision support for health interventions.

References

[1] Yang, M., Lou, X., Peng, Y., Wang, R., Li, L., & Guo, W. Correlation between BMI and physical fitness indicators among college students. Chinese School Health, 2013, 34(9), 1093–1095, 1098. (in Chinese)

[2] Zhao, L., Guo, W., Zhang, Y., et al. Correlation analysis between body mass index and the indicators of National Student Physical Health Standard. Sport Boutique (Academic Edition), 2018, 37(2), 19–20, 22. https://doi.org/10.3969/j.issn.1004-2644.2018.02.009 (in Chinese)

[3] Xu, W. A study on the correlation between BMI index and physical function and fitness indicators among college students—A case study of North University of China. Master's thesis, North University of China, Shanxi,2022. (in Chinese)

[4] Dong, Y., Jan, C., Ma, Y., Dong, B., Zou, Z., Yang, Z., ... & Ma, J. Prevalence and trends in overweight and obesity among Chinese children and adolescents, 1985–2014. International Journal of Environmental Research and Public Health, 2018, 15(8), 1738.

[5] Ogden, C. L., Carroll, M. D., Lawman, H. G., Fryar, C. D., Kruszon-Moran, D., Kit, B. K., & Flegal, K. M. Trends in obesity prevalence among children and adolescents in the United States, 1988–1994 through 2013–2014. JAMA, 2016, 315(21), 2292–2299.

[6] Zhang, Y., Zhang, H., Ma, X., & Wang, Y. Regional differences of physical fitness and overweight and obesity prevalence among college students before and after the COVID-19 pandemic since the double first-class initiative in China. BMC Public Health, 2018, 18(1), 105.

[7] Popkin, B. M., Adair, L. S., & Ng, S. W. Global nutrition transition and the pandemic of obesity in developing countries. Nutrition Reviews, 2012, 70(1), 3–21.

[8] Wilson, P. W., D'Agostino, R. B., Sullivan, L., Parise, H., & Kannel, W. B. Overweight and obesity as determinants of cardiovascular risk: the Framingham experience. Archives of Internal Medicine, 2002, 162(16), 1867–1872.

[9] Chen, Y., Hsieh, P. L., & Chen, Y. L. The relationship between obesity and forced vital capacity among university students. Nutrition Hospitalaria, 2015, 31(5), 2202–2204.

[10] Jackson, A. S., Blair, S. N., Mahar, M. T., Wier, L. T., Ross, R. M., & Stuteville, J. E. Prediction of functional aerobic capacity without exercise testing. Medicine and Science in Sports and Exercise, 1990, 22(6), 863–870.

[11] Han, G., Li, S., Wang, W., et al. Effects of physical inactivity on physical fitness and overweight in Chinese university students during the COVID-19 pandemic. Journal of Sport and Health Science, 2021,10(5), 601–608. https://doi.org/10.1016/j.jshs.2021.05.003

[12] Wang, L., & Hu, J. An analysis of the relationship between national physical fitness and social/natural factors based on geographically weighted regression (GWR) model. Journal of Jilin Institute of Physical Education,2017, 33(6), 15–21. (in Chinese)

[13] Chen, X., Huang, Y., & Sun, M. Body mass index and its associations with lung function and cardiorespiratory fitness among Chinese college students. BMC Public Health, 2022, 22, 187. https://doi.org/10.1186/s12889-022-12541-0

[14] Wang, D., Chen, T., & Zhang, Y. Gender differences in body composition and physical fitness among Chinese college students. International Journal of Environmental Research and Public Health, 2020, 17(7), 2375. https://doi.org/10.3390/ijerph17072375