Analysis of the Differentiation of Medical Resource Allocation and Its Influencing Factors

——Based on the Number of Health Care Institutions in Each Province of the Country

Quanlv Guo[1,a], Xiuli Zhang[2,b,*]

[1]College of Mathematics and Statistics Chongqing Technology and Business University, Nanan, 400000, China
[2]College of Big Data Statistics Guizhou University of Finance and Economics, Guiyang, 550000, China

[a]1457800121@qq.com, [b*]2630132560@qq.com

Abstract. To allocate medical resources more reasonably and effectively, and to promote the sustainable development of health initiatives, this paper analyzes data on the number of medical and health institutions in China from 1990 to 2023. It explores the spatial and temporal differences in the allocation of medical resources through the analysis of Global and Local Moran's I, as well as non-parametric KDE. Additionally, Random Forest Regression model is employed to identify the key factors influencing the number of medical and health institutions. The findings indicate that the national level of healthcare resource allocation has increased significantly and stabilized after 2010, with the greatest degree of differentiation observed among provinces in the central and western regions. Furthermore, Per capita disposable income and population mortality rate are identified as key factors affecting the number of medical and health institutions.

Keywords: Global Moran's I; Local Moran's I; Non-parametric KDE; Random Forest Regression

1 Introduction

Rational allocation of health resources can more effectively promote the sustainable development of health initiatives and address the issue of regional disparities in the distribution of health resources in China.

Dong Enhong et al. (2022) used global and local spatial autocorrelation Moran's I index to analyze the aggregation index and spatial differentiation degree of the number of health institutions, beds, doctors, medical technicians, and nurses in 31 provinces of China. They found that the overall aggregation of health resources in China shows a trend of spatial regional differentiation[1]. Gao Dian et al. (2023) employed the DEA model to evaluate the efficiency of rural healthcare resource allocation in China, using the aggregation degree of health resources and GIS technology to spatially map the

S. M. Chung et al. (eds.), *Proceedings of 2024 4th International Conference on Public Management and Big Data Analysis (PMBDA 2024)*, Advances in Economics, Business and Management Research 318,
https://doi.org/10.2991/978-94-6463-656-7_6

allocation of rural healthcare resources, discovering that rural healthcare resources are concentrated in the eastern regions[2]. Zhu Yan et al. (2023) applied the entropy method for a comprehensive measurement of the level of health resource allocation, evaluated the spatial balance of health resource allocation based on the natural breaks classification method, calculated the degree and sources of regional differences in health resource allocation using the Theil index, and explored the dynamic evolution trend of health resource allocation through kernel density estimation, finding that the allocation of health resources in China is gradually becoming more balanced, although significant differences still exist between regions[3].

Research on the comprehensive model of medical resource allocation is continuously evolving, employing various approaches. It utilizes Global and Local Moran's I, non-parametric KDE, and random forest regression models to thoroughly examine the spatial and temporal variations in the distribution of medical resources and the influencing factors.

2 Data Preprocessing and Research Methods

2.1 Data Preprocessing

2.1.1 Data Source.

The number of medical and health institutions serves as a crucial indicator of the allocation of medical and health resources within a region. Data on the number of such institutions across each province and region of China has been collected for the period spanning from 1990 to 2023.

2.1.2 Data Cleansing.

Due to varying degrees of missing data on the number of medical and health institutions in each province of China from 1990 to 2023, the linear interpolation method is employed to estimate the missing values for the panel data. Using the total population of each province from 1990 to 2022, the index of medical and health institution concentration for each province during that period is ultimately calculated. This allows for an analysis of regional disparities and trends in the spatial distribution of healthcare resources based on the concentration index.

2.1.3 Index of Concentration.

To better explore the regional differentiation and spatial distribution trends of medical resource allocation in China, we utilized the index of medical and healthcare organization agglomeration for each region.

$$I_i = H_i / P_i \tag{1}$$

Where I represents Health care organization concentration index, H represents the number of medical and health institutions in the country, P represents total population for the year.

2.2 Research Methods

2.2.1 Global Moran's I.

$$Moran's\ I = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} W_{ij}(x_i - \overline{x})(x_j - \overline{x})}{s^2 \sum_{i=1}^{n}\sum_{j=1}^{n} W_{ij}} \tag{2}$$

Where, n represents the number of provinces, x represents the number of health institutions in the spatial province, W represents the value of the spatial weight matrix between the two provincial domains, s^2 represents variance.The value of Moran's I is between -1 and 1; I>0,I=0,I<0 represents positive spatial correlation, non-existent spatial correlation, and negative spatial correlation among provinces, respectively[4].

2.2.2 Local Moran's I.

Local Moran's I is a metric that quantifies local spatial correlation, allowing for an analysis of the spatial characteristics of the area under study. It evaluates the local spatial relationships between the area in question and its neighboring regions[5].

$$I_i = \frac{(x_i - \overline{x})}{s^2} \sum_{j=1} W_{ij}(x_j - \overline{x}) \tag{3}$$

For Local Moran's I, the values range between -1 and 1. According to the quadrant, they are categorized as High-High, Low-High, Low-Low, and High-Low.

2.2.3 Non-Parametric KDE.

KDE studies the data distribution characteristics from the data itself and overcomes the problem of subjectivity in setting the functional form in parameter estimation.

$$\begin{gathered} f(rf) = \frac{1}{nh}\sum_{j=1}^{n} K\left(\frac{rf_j - \overline{rf}}{h}\right) \\ K(rf) = \frac{1}{\sqrt{2\pi}} exp\left(-\frac{rf_j^2}{2}\right) \end{gathered} \tag{4}$$

Where, rf represents independently and identically distributed observations, i.e., the index of clustering of the number of healthcare facilities. K is kernel density function, In this paper, we take the Gaussian kernel density function. h represents the optimal bandwidth determined using the “thumb rule”[6].

3 Results & Discussion

3.1 Time-Space Difference

In this paper, Stata 17.0 is used as a spatial econometrics tool to analyze the spatial and temporal differences of health resources distribution in China between 1990 and 2023, including Global and Local Moran's I analysis and Non-parametric kernel density estimation.

Table 1. Index of concentration

Year	Moran's I	Sd(I)	Z	P
1990	-0.084	0.092	-0.548	0.292
1991	0.086	0.114	1.048	0.147
1992	0.126	0.112	1.421	0.078
1993	0.230	0.113	2.336	0.010**
1994	0.225	0.112	2.301	0.011*
1995	0.231	0.111	2.386	0.009**
1996	-0.029	0.072	0.060	0.476
1997	0.073	0.883	1.280	0.100
1998	0.234	0.118	2.262	0.012*
1999	0.280	0.118	2.651	0.004**
2000	0.144	0.118	1.499	0.047*
2001	-0.031	0.118	0.018	0.493
2002	0.264	0.117	2.546	0.005**
2003	-0.004	0.183	0.283	0.388
2004	0.356	0.116	3.347	0.000**
2005	0.044	0.118	0.662	0.254
2006	0.158	0.117	1.643	0.050*
2007	0.118	0.118	1.284	0.099
2008	-0.021	0.116	0.102	0.459
2009	0.282	0.118	1.988	0.023*
2010	0.235	0.119	2.260	0.012*
2011	0.137	0.118	1.442	0.075
2012	0.313	0.119	2.923	0.002*
2013	0.313	0.119	2.910	0.002*
2014	0.280	0.119	2.631	0.004*
2015	0.215	0.119	2.087	0.918
2016	0.225	0.119	2.182	0.015*
2017	0.232	0.118	2.245	0.012*
2018	0.052	0.188	0.785	0.216
2019	0.228	0.189	2.407	0.008**
2020	0.211	0.109	2.238	0.013*
2021	0.187	0.109	2.022	0.022*
2022	0.153	0.108	1.721	0.043*

Table 2. Areas where Local Moran's I passes the test of significance

Year	High-High	Low-High	Low-Low	High-Low
1990			Henan*, Anhui*	Guizhou***
2000	Tibet***, Qinghai**, Xinjiang*		Anhui**, Shandong*, Henan*, Hubei*	Zhejiang*, Hunan**
2010	Shanxi***, Inner Mongolia **, Hebei*, Liaoning*, Shaanxi**, Gansu**	Qinghai*, Ningxia***	Fujian**, Jiangxi*	
2020	Shanxi*, Inner Mongolia*, Sichuan*, Tibet*, Qinghai***, Gansu*	Yunnan*	Shanghai**, Jiangsu*, Zhejiang*	

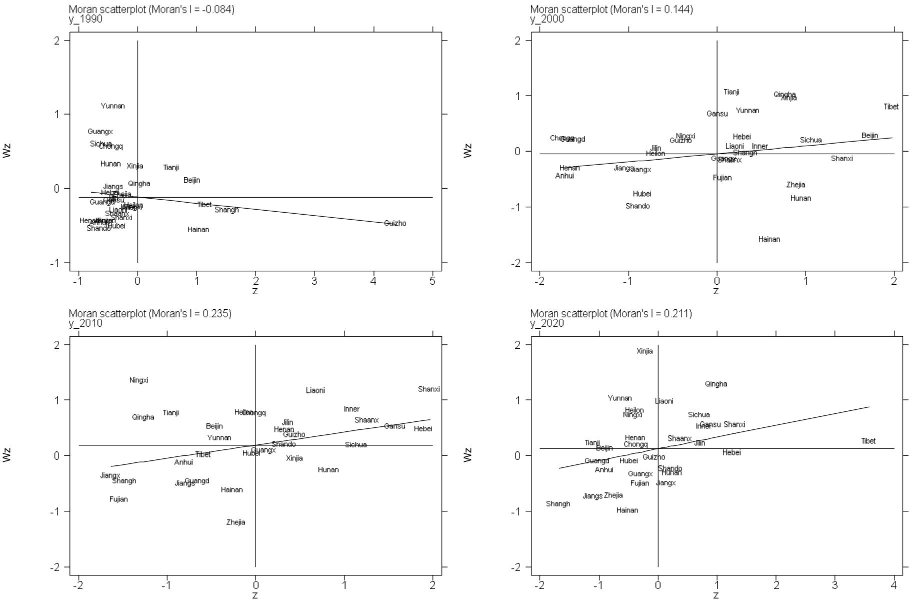

Fig. 1. Local Moran's I scatterplot

It can be seen from Table 1, from the Global Moran's I value and Z value, the Moran's I value of the clustering index of the number of healthcare institutions in China's provinces under the years of between1990 and 2022 showed positive values in the majority of the years, and showed significant spatial positive correlation under the significance level of P<0.05. Table 2 followed by Figure 1 plots the local Moran scatter plots and the regions where the local Moran index passed the significance test in 1990, 2000, 2010, and 2020, and found that the High-High type is mainly concentrated in the Northwest region, and it is increasing from 0 provinces in 1990 to 6 provinces in 2020, and it can be found from the local Moran's I scatterplots in 2020 that Shanxi, Inner Mongolia, Sichuan, Tibet, Qinghai, Gansu, Shanghai, Jiangsu and Zhejiang provinces show a significant positive spatial correlation, that is, it shows that there is a positive coordination relationship between these provinces and their provinces in the field, and the level of their medical resource allocation can drive the development of neighboring provinces.

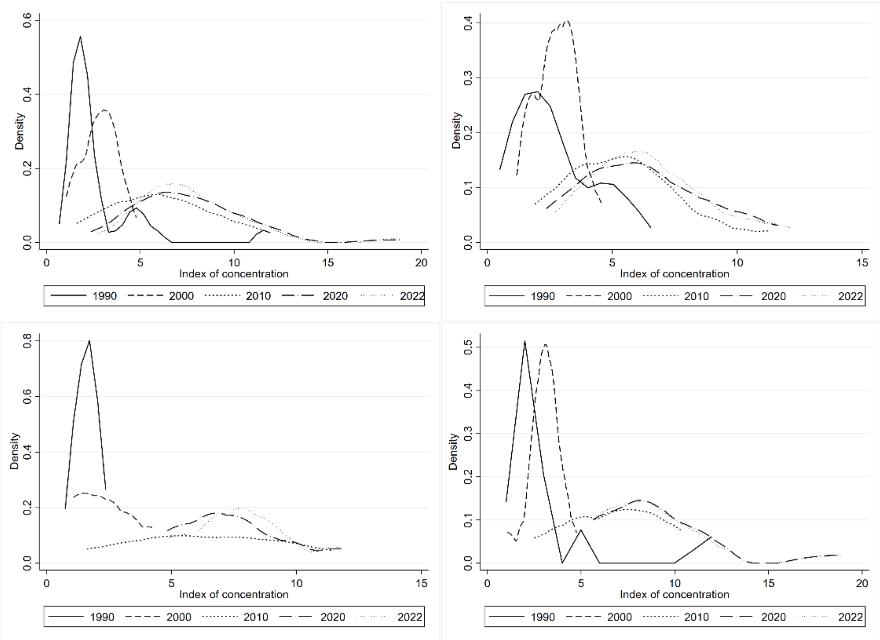

Fig. 2. Estimates of nuclear density

Figure 2 shows that the center position of the distribution curve of nuclear density in the whole country and the three major regions has been shifted to the right over time, indicating that the number of medical and health institutions and the level of medical allocation resources in each region have been significantly improved, and the changes in the center position of the curve and the degree of dispersion in the years of 2010, 2020, and 2022 have been small, which indicates that the changes in China's healthcare resource allocation have tended to be stabilized after 2010. The distribution curves of the whole country, the eastern region and the western region all show obvious “low side peak” and right side trailing characteristics, and the height of the wave peak gradually decreases, while the distribution curve of the central region has a gentle and inconspicuous wave peak with the longest wave peak width, which indicates that the level of healthcare allocation resources of the provinces and municipalities in this region is not uniform and has the greatest degree of differentiation.

3.2 Factors Analysis

This research utilizes Python to examine the factors that influence health care institutions. It focuses on four key areas: economic development levels, allocation of health care resources, population demographics, and the relevant policy direction of each province. An indicator system was developed, consisting of 4 main indicators and 12 sub-indicators, to analyze these influencing factors[7].

3.2.1 Indicator System.

The dimension of Economic development level uses Per Capita Disposable Income, GDP, and Fiscal revenue as indicators; Healthcare resource allocation selects the number of health personnel, licensed doctors, registered nurses, and the number of hospital beds as measurement indicators; Population structure is measured by Population density, Elderly dependency rate, and Population mortality rate; Policy Orientation includes the proportion of total health expenditure to GDP and general public budget expenditure as two indicators. The specific indicators are detailed in the Table 3.

Table 3. system of indicators

Primary Indicators	Secondary Indicators	Indicator Unit
Economic development level	Per Capita Disposable Income	Yuan
	GDP	Billion Yuan
	Fiscal revenue	Billion Yuan
Healthcare resource allocation	Number of health personnel	People
	Number of licensed doctors	People
	Number of registered nurses	People
	Number of beds	Zhang
Population structure	Population density	Per square kilometer
	Elderly dependency rate	%
	Population mortality rate	%
Policy Orientation	Total Health Expenditures as percentage of GDP	%
	General public budget expenditure	Billion Yuan

3.2.2 Model Parameters.

The optimal combination of parameters for the random forest regression model is as follows in Table 4.

Table 4. Model parameters

Parameter	Values
Data splitting	0.8
Cross-validation	5
Node splitting evaluation criteria	mae
Minimum sample size for internal node splitting	5
Minimum sample size for leaf nodes	3
Maximum depth of the tree	10
Maximum number of leaf nodes	70
Number of decision trees	90
Sampling with replacement	true

3.2.3 Model Evaluation.

Choose Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Coefficient of Determination (R^2) as model evaluation metrics. The metric results for the cross-validation set, training set, and test set are as follows in Table 5.

Table 5. Evaluation of indicator values

Data	MSE	RMSE	MAE	MAPE	R^2
Training set	0.011	0.106	0.071	1.165	0.881
Cross Validation Set	0.022	0.141	0.101	2.774	0.713
Test Set	0.064	0.254	0.179	2.226	0.843

3.2.4 Importance of Factor Variables.

The status of features in the random forest model is not equal, so it is necessary to assess the importance of the features in the model to measure the contribution of the features. In this paper, we use variable importance measures (VIM) to assess the importance of each feature, and features with high VIM have a greater impact on the differences in the number of health care institutions between provinces.

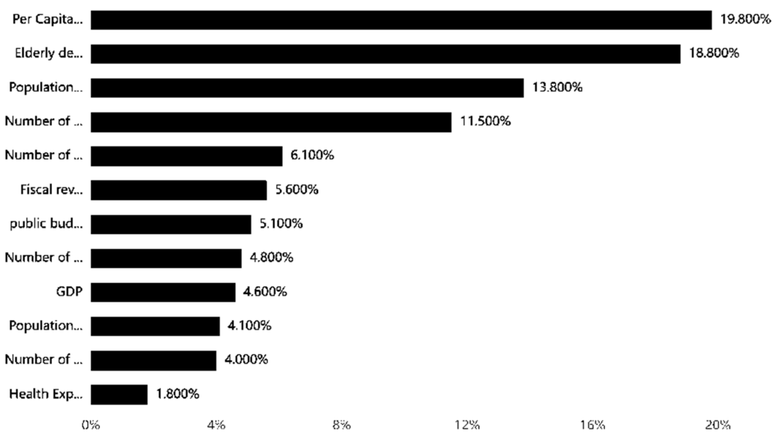

Fig. 3. Importance of variables

Figure 3 illustrates that per capita disposable income level is a significant variable in economic development, accounting for 19.8% of the influence on the number of medical and health institutions in China. This measure reflects the income level of the population and indicates the living standards and well-being of residents. Its strong correlation with the number of healthcare facilities suggests that residents' economic conditions are closely linked to their demand for and use of medical services. Additionally, the population mortality rate and the elderly dependency ratio contribute 13.8% and 18.8% respectively, underscoring the challenges posed by an aging population on healthcare demand. Furthermore, the number of practicing physicians and health personnel also play a crucial role, highlighting the importance of human resources in the healthcare system. The quantity and quality of practicing physicians, who directly provide medical services, are essential for ensuring the quality and accessibility of healthcare.

4 Conclusions

4.1 Conclusions

4.1.1 The Spatial Pattern of Health Resource Allocation.

The research results of the global and local Moran indices indicate that there is a spatial agglomeration effect in the allocation of healthcare resources in China. From 1990 to 2020, the level of healthcare resource allocation in our country has shown a continuous growth trend, but spatial imbalance still exists, and there are still differences in the distribution of medical resources among provinces. Analyzing the reasons: first, under the guidance of national policies, from the 2009 "Opinions on Deepening the Reform of the Medical and Health System," the 2015 "National Medical and Health Service System Planning Outline," to the "Key Work Tasks for Deepening the Reform of the Medical and Health System in 2022," the government has continuously favored the western regions with policies and funding, which has significantly improved the investment in healthcare resource allocation in economically underdeveloped areas, thereby narrowing the gap caused by local economic development levels to some extent; second, the "14th Five-Year Plan" for National Health released by the State Council proposes that the quality of medical and health services will continue to improve, the supporting capabilities related to medical and health services and the development level of the health industry will continue to enhance, the national health policy system will be further improved, and the average life expectancy will continue to increase by about one year based on the 2020 level.

4.1.2 Differences in the Allocation of Health Resources by Region.

The KDE results indicate that from 2011 to 2020, the kernel density curve of the regional allocation of health resources has slightly shifted to the right, while the central and western regions show a slight contraction trend. The degree of polarization has tended to moderate, but the degree of differentiation remains the most pronounced. The reasons for this are: first, for a long time, the national support for the central and western regions in areas such as the economy, health, and education has continued to increase, leading to significant improvements in the scale, quality, and efficiency of health resources in these regions, and the gap in the allocation of health resources within the region has also shown a certain degree of narrowing trend; second, the financial investment in the central and western regions has been relatively insufficient for a long time, resulting in weak financial support capabilities, which has led to inadequate medical service supply in these areas, and the uneven allocation of health resources among provinces has restricted the development of health service capacity in the region; third, the imbalance in development between regional and urban-rural dual structures has caused the allocation of health resources to tilt towards areas with higher economic levels, severely impacting some central and western regions.

4.1.3 Factors Affecting the Allocation of Health Resources.

The research results of the random forest regression model identified key factors affecting the number of healthcare institutions, including per capita disposable income, population mortality rate, elderly dependency ratio, and the number of practicing physicians and healthcare personnel. The "Healthy China 2030" Planning Outline and the "14th Five-Year Plan" for Health Talent Development propose measures to strengthen the construction of the health talent workforce to meet new health service demands. Economic development and human resources play a crucial role in the healthcare service system, while changes in population structure significantly impact the demand for healthcare services.

4.2 Recommendations

Improve the dynamic supply mechanism for the allocation of health resources. First, promote the implementation of talent introduction policies based on the health development plans of various regions to meet their health service needs, and establish a sound connection mechanism between medical institutions and the health human resources market; promote the high-quality development of health resources in various regions. In the eastern regions, due to the competitive relationship of resource competition, it is necessary to establish a regional integration resource-sharing mechanism and improve management systems centered on medical quality. For the central regions, it is essential to promote cross-regional cooperation in the healthcare industry and higher medical education, effectively plan and allocate health resources between regions, and formulate reasonable and efficient health management policies to provide a healthy socioeconomic environment for the stable development of health resource allocation, as well as to provide important support for the implementation of the Healthy China strategy. As the reform of the medical and health system continues to deepen, for the western regions, promoting the clustered development of health resources should leverage the "center" cities within urban agglomerations to form a high-quality cluster of diverse health resources, including national medical centers and regional medical centers, deeply integrating with and forming profound interactions with regional socio-economic development.

References

1. Dong Enhong, Yan Yue, Xie Yali. (2022) Research on the Regional Differences and Spatial Distribution Trends of Health Resource Allocation in China [J]. Chinese Health Policy Research, 15(06): 73-79.

2. Gao Dian, Shi Lushao Bo, Lin Jinhui. (2024) Research on the Efficiency and Equity of Rural Medical and Health Resource Allocation in China Based on DEA-GIS Method [J]. Chinese General Practice, 27(07): 849-856.
3. Zhu Yan, Ma Yulong, Wang Jiayi. (2023) Regional Differences and Dynamic E-volution of Health Resource Allocation in China from 2011 to 2020 [J]. Chinese Health Resources, 26(04): 417-423. DOI:10.13688/j.cnki.chr.2023.230095.
4. Moran, P. A. P. (1950) Notes on continuous stochastic phenomena. Biometrika, 37(1-2), 17-23.
5. Anselin L. (1995) Local Indicators of Spatial Association—LISA[J].Geographic-al Analysis, 27(2):93-115.DOI:10.1111/j.1538-4632.1995.tb00338.x.
6. Parzen, E. (1962) On Estimation of a Probability Density Function and Mode. The Annals of Mathematical Statistics, 33(3), 1065–1076.
7. Breiman L. (2001) Random forests, machine learning 45[J].Journal of Clinical Microbiology, 2:199-228.