SIN-CHN SCIENTIFIC PRESS PTE.LTD.

Molecular & Cellular Biomechanics 2025, 22(2), 1025.
https://doi.org/10.62617/mcb1025

Article

Comprehensive evaluation of health community construction based on biomechanical perspective

Xiang Hu[1,*], Yonghong Lan[2]

[1] School of Design and Arts, Hunan Institute of Engineering, Xiangtan 411104, China
[2] School of Automation and Electronic Information, Xiangtan University, Xiangtan 41105, China
*** Corresponding author:** Xiang Hu, 871782403@qq.com

CITATION

Hu X, Lan Y. Comprehensive evaluation of health community construction based on biomechanical perspective. Molecular & Cellular Biomechanics. 2025; 22(2): 1025. https://doi.org/10.62617/mcb1025

ARTICLE INFO

Received: 6 December 2024
Accepted: 27 December 2024
Available online: 22 January 2025

Abstract: The comprehensive evaluation of healthy community construction is an important means for managers and decision-makers to understand the development of community health within their jurisdiction, and also a scientific basis for formulating practical and feasible health policies and health development plans. Scientific evaluation of the overall level and differences in the construction of healthy communities has important theoretical significance and reference value. Based on the principal component clustering analysis method as well as the theoretical foundation, the multiple evaluation indicators in the evaluation index system of health community construction are standardized and dimensionally reduced. At the same time, the principal component factor load matrix clustering analysis is used, combined with the practical significance and evaluation direction of the indicator categories, to propose a design idea for the classification index system. The comprehensive principal component evaluation and clustering analysis are used to quantitatively examine the level of health community construction in specific regions and make horizontal comparisons, provide theoretical support and decision-making reference for the construction of healthy communities. The correctness and effectiveness of the proposed method system were verified by numerical examples.

Keywords: healthy community; comprehensive evaluation; principal component; clustering analysis

1. Introduction

Healthy community [1] conceptually means that all organizations and individuals in the community jointly promote the overall health of the community. The health of community refers to multidimensional health, including the physical and mental health of individuals and the health of the social environment, as well as the positive interaction between individuals and the social environment in peacetime and in disaster, covering the whole process of planning, construction, and governance. In view of the importance, persistence and complexity of healthy community planning, development, and construction, it is particularly important to evaluate the level of healthy community construction at each stage and provide useful reference for long-term planning of healthy community.

Canada put forward the construction program of healthy community in 1984 and encouraged its residents to participate in it on a voluntary basis, which promoted the development of early healthy community. In 1986, the World Health Organization (WHO) issued the Ottawa Charter, which was deemed an important symbol by some developed countries such as European countries and the United States. These countries took the lead in the rise of the healthy community movement. Later, healthy community was promoted to the world with the strong support of WHO. The United

States Department of Health and Human Services formalized the concept of healthy community in 1989 and formed the construction of healthy communities, cities and states across the country. After a large number of theoretical and practical studies, it has been proved that urban built environment is closely related to public health [2], including but not limited to the physical environment which affects the incidence of diseases [3] and human comfort [4], and directly affects residents' physical activity level [5] and mental health [6]. Research on the built environment index system of healthy communities mostly focuses on the method and model construction of the evaluation system of healthy communities [7]. The first evaluation system for community planning in the United States, effectively promotes the ecological and sustainable development of communities. Based on three theories of ecological approach, interactive worldview and naturalistic inquiry paradigm to human development, constructed a neighborhood assessment method suitable for residential and community, which helps planners to build healthier communities [8]. A logical model of housing satisfaction was proposed and pointed out that age, race, and various design elements of residential units would affect residents' satisfaction with the community and ultimately affect the livability of the community [9]. A 3D theoretical model [10] was used to analyze the environmental factors (usually diversity, design and density) that affect traffic, which has become the foundational theoretical model for the current research on urban areas and built environments. On this basis, bus transfer distance and traffic accessibility are added. The 3D model has then been developed into the 5D theory, constructing the basic health-oriented research of built environment. Related built environment research is mostly carried out on this basis [11].

Studies on the promotion of health by community environment mostly focus on the correlation between the community-built environment and physical activity [12]. The built environment factors were systematically analyzed from land layout [13], facility distribution [14], and functional mixing [15] and so on. At the same time, the influence of built environment on the physical activity of the elderly [16]. The impact of residential environment on residents’ epidemic diseases, respiratory diseases, stroke, cardiovascular and cerebrovascular diseases, depression and other related diseases was studied by Luque et al. [17]. Combined with subjects’ physical examination data collected in Changsha by the joint investigation of Hunan Health Management Quality Control Center and Health Care Research Center of Central South University, the results showed that the elderly were the key groups in the prevention and treatment of chronic diseases. As provincial capital of central China, Changsha had 1,436,200 people aged 60 and above by the end of 2019, accounting for 17.44% of its total population, exceeding the international standard of 10%. The onset age of chronic metabolic diseases such as overweight and obesity, dyslipidemia, fatty liver, and others becomes evidently younger, in which males are the majority. Such trends are closely related to the lifestyle of young people. Among the physical symptoms, people in Hunan most suffer from neck and lower-back pain—among chronic pains. The spinal health cannot be ignored as it is mainly related to occupational habits and insufficient physical activity.

The above-mentioned studies mainly focus on the construction of walkable space, the design of green open space and natural environment, the allocation of public service facilities and the construction of caring environment for the disadvantaged group in healthy communities. The quantitative evaluation and clustering analysis of various performance indicators of healthy communities have not been involved. Taking the establishment of healthy communities in Changsha as the research object, this paper will conduct standardization and dimensionality reduction of the multiple evaluation indicators in the evaluation index system of healthy community construction. At the same time, the principal component factor load matrix clustering analysis is used, combined with the practical significance and evaluation direction of the indicator categories. The comprehensive principal component evaluation and clustering analysis are used to quantitatively examine the level of health community construction in specific regions, to further provide theoretical support and decision-making reference for the construction of healthy communities.

2. Research data and methodology

2.1. Research period

This study started on 10 May 2023 and ended 10 September 2024, lasting for 1 year and 4 months.

2.2. Research scope

In this study, 5 districts, 11 sub-district and 15 communities with higher population density in Changsha are selected as research objects, as shown in **Table 1**, including Huangni Street Community, Fengquan Gujing Community and Zoumalou Community in Dingwangtai Subdistrict, Furong District; Qilimiao Community in Dongtang Subdistrict, Houjiatang Community in Zuojiatang Subdistrict, Ziyuan Road Community in Houjiatang Subdistrict, Fengyuan Community in Dongjing Subdistrict, Yuhua District; Yunluyuan Community and Meiyuan Community in Meixi Lake Subdistrict, Jinmao Community in Yuelu Subdistrict, and Zhongnan Community in Gaotangling Subdistrict, Yuelu District; Furong Road Community of Qingshuitang Subdistrict, Kaifu District; Taiping Street Community and Banxiang Street Community in Pozi Street Subdistrict, Tianxin District; and Jingwanzi Community in Qingyuan Subdistrict.

Table 1. Names of 15 communities.

No.	Names of Community	Sub-districts	Districts
1	Furong Road Community	Qingshuitang Subdistrict	Kaifu District
2	Huangni Street Community	Dingwangtai Subdistrict	Furong District
3	Fengquan Gujing Community	Dingwangtai Subdistrict	Furong District
4	Qilimiao Community	Dongtang Subdistrict	Yuhua District
5	Houjiatang Community	Zuojiatang Subdistrict	Yuhua District
6	Ziyuan Road Community	Houjiatang Subdistrict	Yuhua District
7	Jinmao Community	Yuelu Subdistrict	Yuelu District
8	Yunluyuan Community	Meixi Lake Subdistrict	Yuelu District
9	Meiyuan Community	Meixi Lake Subdistrict	Yuelu District

Table 1. (*Continued*).

No.	Names of Community	Sub-districts	Districts
10	Taiping Street Community	Pozi Street Subdistrict	Tianxin District
11	Fengyuan Community	Dongjing Subdistrict	Yuhua District
12	Zoumalou Community	Dingwangtai Subdistrict	Furong District
13	Jingwanzi Community	Qingyuan Subdistrict	Tianxin District
14	Zhongnan Community	Gaotangling Subdistrict	Yuelu District
15	Banxiang Street Community	Pozi Street Subdistrict	Tianxin District

2.3. Research methods

2.3.1. Field research and questionnaire distribution

By field research and questionnaire distribution in 15 communities of Changsha, the status quo of public space environment including spatial scale, greening, non-motorized transportation system, medical facilities and service facilities was investigated. Questionnaires covering these features (including residents' satisfaction and optimization willingness) were distributed, and the main spatial material elements of the community were extracted to form sample data of built environment indicators of healthy communities with high confidence.

2.3.2. Acquisition of big data images and photos

Through multiple surveys and comprehensive investigation of the built environment of communities in the main urban area of Changsha, combined with the high-resolution image data of Baidu Big Data, a total of 3425 data were obtained by selecting the panorama and the POI data of community micro-elements of 15 high-density communities in Changsha. After de-duplication, screening and space clipping, 1037 data were valid and retained, among which were 116 building data, 115 street data, 371 service facilities data, 108 medical facilities data, 138 community parks data, and 189 public green space data. The research data was obtained by crawling the AOI interest surface of Baidu Maps on 20 July 2023, and combining it with EasyPoi software to obtain relevant data.

2.3.3. Main test methodology

1) Building density

Building density refers to the coverage rate of buildings, specifically represents the ratio of the total area of the base of all buildings within the project's land scope and the area of the planned construction land (%). This can reflect the open space ratio and building density within a certain range of land.

2) Plot ratio

Plot ratio is an important indicator to measure the use intensity of construction land, referring to the ratio of the total construction area of all types of buildings above the ground to the base area within a certain area. The plot ratio in this study is derived from the pre-approval data of the "Regulations of Regionalized Management of Residential Land Density, Changsha".

3) Land use mixture

Land use mixture refers to the proportion of land functions that are closely related to people's life in a certain urban block. Generally, within a certain area, the more functions the block loads, the heavier the functional pressure and the higher the use mixture in this area. "Information entropy" is often used to measure the diversity and mixing of land use in a grid.

4) Coverage of Park Service Radius

Coverage of Park Service Radius (%) = Residential Land Area covered by the Park Service Radius in the Community (hm2)/Residential Land Area (hm2) × 100%

5) Coverage of Community Health Facilities (Buildings)

Coverage of Community Health Facilities (Buildings) = Base Area of Community Health Facilities/Community Land Area. For health facilities in this study, each community should have pharmacies and at least one in general hospitals, Centers for Disease Control, specialized hospitals, community healthcare centers and clinics. Only in this way, the community is deemed as "having" health facilities.

6) Greening rate

Greening rate refers to the ratio of the total greening area in the overall commercial land under planned construction to the overall land area under planned construction.

7) Coverage of Bus Stops within a 500-m Radius

The Coverage of Bus Stops within a 500-m Radius refers to the proportion of an area covered by a certain radius of all public transportation (buses, trams) in a certain area of a city, to the total area suitable for setting public transport stations/stops. That is, the ratio of "Coverage of Bus Stops within a 500-m Radius" to the area of "Study Area".

8) Road density

Road density refers to, within a certain region, the ratio of the total mileage of the road network to the area of region. The ArcGis method is used for calculation.

9) Pedestrian Route Directness (PRD)

Calculation is done via data scraping on Baidu ma, PRD = actual length/linear distance.

3. Calculation results and analysis

3.1. Descriptive analysis of spatial indicators of built environment in healthy communities

The results of data gap analysis of 15 communities are shown in **Table 2**. Some communities have problems with old communities, which have high density, low facility coverage, obvious shortage of medical and elderly care facilities, and low transportation accessibility. All of these have adverse effects on the health of the elderly population. Some communities have developed well and have the characteristics of high-density communities, with convenient living, but limited public greening and traffic congestion. Some communities have a high degree of mixed land use, good spatial green space and landscape construction, and relatively rich service facility layout, which has a positive impact on the physical and mental health of residents. The criteria for selecting effective data points mainly include the completeness, accuracy, and practicality of the data. Firstly, the data was deduplicated to ensure there were no duplicate data points. Then, screening was conducted based on the representativeness and practicality of the data; Finally, quality control was carried out on the selected data to ensure its accuracy and reliability. Through this series of screening and validation processes, valid data points were ultimately retained.

Table 2. Descriptive statistics of built environment indicators in communities.

Indicator	Range of Variation	Mean Value	Coefficient Variation	Median
Building Density	13.8–36.2	33.1	532.7	31.2
Plot Ratio	1.3–2.5	2.2	50.8	2.3
Land Use Mixture	0.27–0.83	0.61	16.1	0.59
Coverage of Dining and Shopping Facilities	53–95	76.6	962.4	78
Coverage of Community Health Facilities (Buildings)/%	51.63–78.79	66.7	860.5	68.8
Coverage of Community Elderly Care Facilities (Buildings)/%	10.39–15.88	13.3	173.5	13.2
Accessibility of Medical Facilities	0.12–1.0	0.35	5.2	0.32
Coverage of Greenway Service Radius in Community	20–45	32	781.9	35
Greening Rate/%	18–50	32	875.1	31
Coverage of Park Service Radius in Community/%	11–38	27	595.8	27
Gini Coefficient of Green Park Area per capita (based on communities)	0.90–0.78	0.68	7.8	0.83
Coverage of Bus Stops within a 500-m Radius/%	72.6–100	82	796.6	80.2
Road Density	4.62–6.27	6.8	41.2	6.7
Pedestrian Route Directness	1.2–2.3	1.9	22.4	1.8
Accessibility of Sports Facilities	0.45–1.0	0.56	3.2	0.76
Noise	20–70	50	43.2	50
Air quality	30–290	210	53.5	216
Infectious disease prevention	III	III	III	III

The data collection work has been carefully planned and implemented to ensure the timeliness and accuracy of the data. At the same time, the accuracy of the data has been verified through various methods, further improving the credibility of the data. Strict standards have also been established when screening valid data points to ensure the integrity and practicality of the data. This provides strong data support for subsequent analysis and application.

The variation coefficients of Accessibility of Medical Facilities, Gini Coefficient of Green Park Area per capita (based on communities), and Accessibility of Sports Facilities were 5.2%, 7.8% and 3.2% respectively, all less than 10%, showing a small dispersion degree. The coefficients of Accessibility of Medical Facilities and Accessibility of Sports Facilities were small, with relatively close value distributions. This might be related to the convergence of coverage of health and sports facilities in the 15 selected communities. In this study, the variation coefficient of the Gini Coefficient of Green Park Area per capita (based on communities) was low. This might be caused by calculating not only urban parks, community parks, and waterfront green space but also green space in residential areas and along the road, leading to a less significant distribution variance. The three indicators did not show significant differences in steric factors. The variation coefficients of the other 12 indicators were large, which indicated that the indicators of health quality of built environments in different communities had large differences, among which the variation coefficient of the Coverage of Dining and Shopping Facilities reached 960.2% with a large dispersion degree. And the measured values of indicators of each built environment factor varied greatly. Such results were related to the selection of urban fringe communities (such as Jingwanzi Community) and urban core communities (such as Taiping Street Community). Comparing the mean and median, it was found that the median of the community quality indicators was close to its mean, indicating that there were fewer data outliers. The above results showed that the measured values of all indicators of the built environment in the selected healthy communities were within the acceptable range, with less outliers. And the quality characteristics of the healthy communities demonstrated large differences, which had a certain universality and representativeness. In this study, all healthy communities were from the city area of Changsha, and the indicators and evaluation criteria of healthy communities were the same. Therefore, the community health quality derived from the differences among built environment indicators.

This data reflects the significant differences and fluctuations in community building density. Note that some of these data have undergone significant changes. For example, the coefficient of variation for Building Density is an astonishing 532.7%. The reasons and factors may be related to various factors such as differences in planning and design, differences in terrain and landforms, changes in population and demand, the influence of building types and heights, constraints of policies and regulations, and economic considerations of developers. These factors are intertwined and influence each other, jointly determining the differences and fluctuations in community building density.

3.2. Clustering analysis based on principal components

In practice, it is often necessary to collect as much information as possible on the research objects, to have a comprehensive understanding of the issue. However, due to the large number of variables, it will undoubtedly increase the analysis difficulty and complexity. For p-dimensional random vectors, the principal component analysis uses m ($m < p$) new variables to replace the original p variables, to minimize the loss of the original information and the number of variables. Clustering analysis uses quantitative statistical analysis to gradually aggregate samples according to their similarity of quality characteristics, prioritize the aggregation of the samples with the highest similarity, and finally realize the aggregation of multiple varieties based on the comprehensive nature of the categories. In this study, systematic clustering analysis was carried out by the method—Sum of Squares of Deviations—on 15 communities according to the built environment of communities.

The first step is to build the raw data matrix. Let the number of evaluation indicators of one evaluation indicator system be n. Systematic evaluation and data sampling are carried out on m evaluation objects. Let the original data evaluation matrix be $X = (X_{ij}))_{m \times n}$. In order to eliminate the effect of base unit and standardize the measurement range of evaluation indicators, all column vectors in the original data matrix X are normalized with normal distribution, and the standardized matrix Z can be obtained:

$$Z = \left(Z_{ij}\right)_{m\times n} = \begin{bmatrix} Z_{11} & Z_{12} & \cdots & Z_{1n} \\ Z_{21} & Z_{22} & \cdots & Z_{2n} \\ \vdots & \vdots & & \vdots \\ Z_{m1} & Z_{m2} & \cdots & Z_{mn} \end{bmatrix} \tag{1}$$

On the basis of the standardized matrix *Z*, the correlation coefficient matrix Σ which represents the relationship among elements in the matrix can be obtained:

$$\sum = \begin{bmatrix} \sigma_{z_1 z_2} & \cdots & \sigma_{z_1 z_n} \\ \vdots & \vdots & \vdots \\ \sigma_{z_n z_1} & \cdots & \sigma_{z_n z_n} \end{bmatrix} \tag{2}$$

of which,

$$\sigma_{Z_i Z_j} = \frac{\text{cov}(\,z_i, z_j)}{\sqrt{\text{Var}(z_i) Var(z_j)}} \tag{3}$$

cov(z_i,z_j) represents the covariance of vectors z_i and z_j; Var (z_j) represents the sample variance of the vector z_i. According to the relevant theory of Matrix Theory, it can be known that the symmetric positive matrix Σ must be orthogonally similar to the diagonal matrix Λ, that is

$$U^T \sum U = \Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \tag{4}$$

Among them, $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$, *U* is an orthogonal matrix made of eigenvectors corresponding to the eigenroots:

$$U = (u_1, u_2, \ldots u_n) \tag{5}$$

The linear combination of the original variables formed by the first-column elements of the matrix *U* has the greatest variance. It indicates the n-dimensional orthogonal vector $a1 = (a11, a21, \ldots, an1)^T$ if and only if $a1 = u_1$.

$$\text{Var}(F_l) = \text{Var}(a_1^T Z) \leq \lambda_l \tag{6}$$

Note that

$$\text{cov}\left(F_i, F_j\right) = u_j^T \sum u_j = 0 \tag{7}$$

$$\text{Var}(F_i) = Var\left(u_i^T Z\right) = u_i^T \sum u_i = \lambda_1 \tag{8}$$

So, the principal component of *Z*—F_i—is a linear combination having the eigenvestor of ***Σ*** as its coefficients, with no correlation. The variance is the eigen-root of Σ.

$$\text{Var}(F_1) \geq \text{Var}(F_2) \geq \text{Var}(F_n) > 0 \tag{9}$$

Thus, the variance contribution rate and cumulative variance contribution rate of the PC*i* can be defined as:

$$\omega_i = \lambda_i / \sum_{j=1}^{n} \lambda_j \quad (10)$$

$$\mathrm{p} = \sum_{i=1}^{m} \lambda_j \lambda_i / \sum_{j=1}^{n} \lambda_j \quad (11)$$

According to the general guideline for the retention of data information, the first *s* principal components are retained when ρ is greater than a certain threshold. According to

$$f = \omega_1\lambda_1 + \omega_2\lambda_2 + \cdots + \omega_s\lambda_s \quad (12)$$

the comprehensive PC evaluation function is generated, the final evaluation score is calculated, and the evaluation ranking result can be given.

On the other hand, *U* is defined as the principal component factor load matrix. Different values in *U* respectively correspond to the correlation coefficients between the first *s* principal components and the original evaluation indicator, and the distribution range is [−1, 1]. The positive and negative signs represent the positive correlation and the negative correlation respectively. The larger the absolute value, the higher the correlation is. A clustering method based on principal component factor load matrix is adopted, taking $|\mathrm{cov}(F_i,z_j)|$ greater than a certain threshold as the clustering criterion. Clustering recognition is carried out on *n* variables of evaluation indicators and the structure of evaluation system is generated.

3.3. Evaluation process of built environment indicators in healthy communities

Combined with the theory and evaluation model of principal component clustering analysis, the evaluation scheme of intelligent process of healthy community was designed to base on the PCA evaluation system. And the finalized evaluation process of built environment indicators of healthy communities is as **Figure 1** (Please see in the top of the next page).

The detailed evaluation process is as follows: First, a targeted evaluation indicator system is selected according to the target research area of healthy communities, and the built environment construction of different communities in urban Changsha is evaluated according to the established evaluation system. Then, the original evaluation data is preprocessed, and the normalized operation of multivariate normal distribution is carried out based on standardization. Next, the first-round judgment starts. If the principal component analysis fails, the evaluation indicator system should be re-selected and repeated the previous process. If it is feasible, the principal component analysis model is adopted to figure out solutions successively by establishing the correlation coefficient matrix, solving the eigenvalues and eigenroots, generating the principal component expression, and so on. Then the second-round judgement starts. Correlation among different PCs will be checked according to PCA coefficient matrix. If there exists correlation, a recalculation is required. If it shows no correlation, the number of PCA evaluation indicators is determined, and the subsequent comprehensive evaluation is carried out in two parts. One, the principal component analysis of the evaluation indicator system is carried out according to the principal component factor load matrix, to provide references for clustering analysis. Two, clustering analysis is carried out on the evaluation indicator system according to the principal component factor load matrix, and the original evaluation indicator system is reclassified according to the correlation. In this way, practical significance will be attributed to each indicator category, reflecting the multiple dimensions of the evaluation system structure of health community.

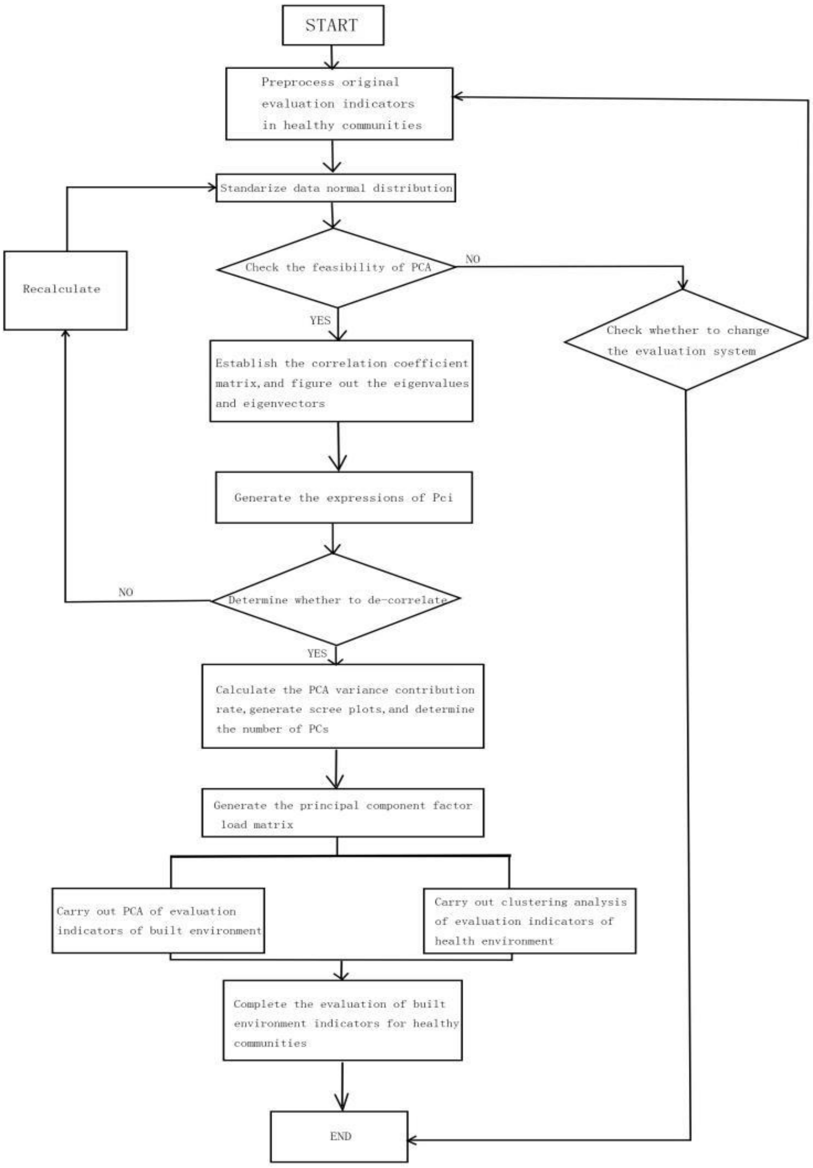

Figure 1. Evaluation process of built environment indicators in healthy communities.

3.4. Result analysis

According to the above results, three indicators namely Accessibility of Medical Facilities, Gini Coefficient Coefficient of Green Park Area per capita (based on communities) and Accessibility of Sports Facilities were excluded, and the PCA results of 12 quality indicators of 15 communities are shown in **Table 3**.

Table 3. Total PCA explanatory variable.

No. of Principal Component	Eigenvalue	Rate of Contribution	Cumulative Rate of Contribution
1	5.6232	0.4686	0.4686
2	2.8850	0.2404	0.7090
3	1.4078	0.1173	0.8263
4	0.5249	0.0437	0.8701
5	0.4549	0.0379	0.9080
6	0.4028	0.0336	0.9416
7	0.2765	0.0230	0.9646
8	0.1917	0.0160	0.9806
9	0.1277	0.0106	0.9912
10	0.0477	0.0040	0.9952
11	0.0467	0.0039	0.9991
12	0.0110	0.0009	1.0000

The scree plot (**Figure 2**) can be used to help determine the optimal number of principal components. In the Scree Plot, the x-axis represents the number of principal components, the y-axis represents the eigenvalue. And the linked steep part of the PC eigenvalue is the number of principal components to be taken. The optimal number of PC was determined by considering the characteristic value > 0.6, and the scree plot and variance contribution rate. As can be seen from **Figure 2**, the eigenvalues of the first 4 PCs are larger and the lines are steeper, indicating the first 4 PCs own the highest Rate of Contribution of the explanatory variables. It is more appropriate to extract 4 principal components, and the cumulative variance contribution rate is 87.01%. Most of the information of built environment indicators of healthy community was integrated.

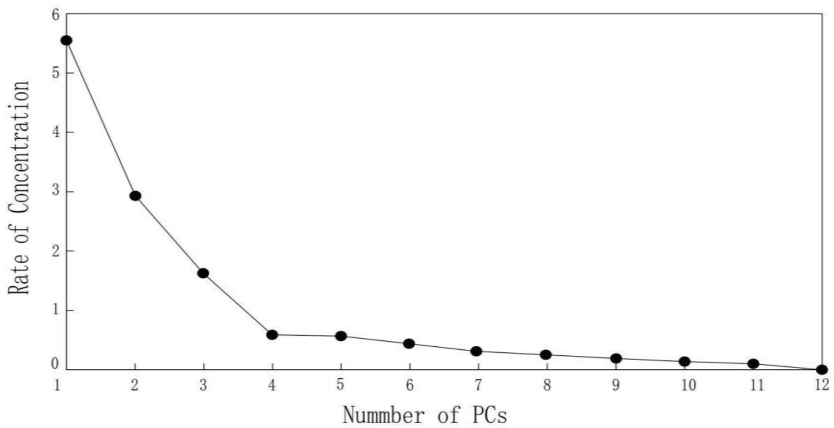

Figure 2. PCA scree plot.

After rotating the PC load matrix, the load coefficient was closer to 1 or 0, thus the resulting principal components can better explain and name the variables. In the Column PC1 in **Table 4**, 0.3783 and 0.3709 are relatively large, indicating that the principal component reflects the characteristics of the 11th and 3rd indicators. This could be applied in other columns. As can be seen from the results of **Table 4**, the PC1 had a strong positive correlation with z_{11}, z_3 and z_1 (Road Density, Land Use Mixture and Building Density). These variables comprehensively reflected the land use of built environment in communities, the level of development, activity, and diversity of the city or community. So, PC1 can be named as Community Development Activity Factor. The PC2 mainly integrated the information of z_7, z_8 and z_9 (Coverage of Greenway Service Radius in Community, Greening Rate, and Coverage of Park Service Radius in Community). PC2 showed a strong negative correlation with these indicators, mainly reflecting natural landscape information. PC2 can be named landscape factor. The PC3 showed a weak positive correlation with z_4, z_5, z_6, z_9 and z_{10}, and a strong positive correlation with z_{12} (Pedestrian Route Directness), indicating that service facilities had a certain inhibitory effect on the Pedestrian Route Directness. PC3 can be named as facility factor. The PC4 showed a strong positive correlation with z_3, z_7 and z_{11}, representing most information of community roads, and PC4 can be named as road factor.

Table 4. PCA Component load matrix after rotation.

	PC1	PC2	PC3	PC4
Building Density z1	0.3568	−0.0451	−0.2015	−0.2157
Plot Ratio z2	0.3123	−0.2443	−0.2938	−0.4357
Land Use Mixture z3	0.3709	0.1273	−0.0545	0.3551
Coverage of Dining and Shopping Facilities z4	0.3399	0.1767	0.0315	0.0865
Coverage of Community Health Facilities (Buildings)/% z5	0.3046	−0.3165	0.1600	0.2921
Coverage of Community Elderly Care Facilities (Buildings)/% z6	0.3208	−0.2404	0.0890	−0.4573
Coverage of Greenway Service Radius in Community/% z7	−0.1140	−0.5296	0.0072	0.3125
Greening Rate z8	−0.1958	−0.4879	−0.1217	0.1281
Coverage of Park Service Radius in Community/% z9	0.0386	−0.4276	0.4502	−0.0825
Coverage of Bus Stops within a 500-m Radius/% z10	0.3573	−0.039	0.1515	0.0177
Road Density z11	0.3783	0.0319	−0.0953	0.4537
Pedestrian Route Directness z12	0.0488	0.1739	0.7653	−0.0912

In this study, PC1 and PC2 contained 46.86%–24.04% of the original information. The PCA Score Plot was used to analyze the relationship between healthy communities and community indicators. The relationship between each community and PC1 & PC2 can be directly seen from **Figure 3**: 11, 13, and 15 (Fengyuan Community, Jingwanzi Community, and Banxiang Street Community) successively located in the negative interval of PC1 and PC2, with small values in residential land density, plot ratio, and greening rate. The facility accessibility and coverage rates were both low, with a serious shortage of health and elderly care facilities, reflecting the information of traditional old communities. 1, 2, and 7 (Furong Road Community, Taiping Street Community, and Huangni Street Community) successively located in the fourth section of PC1 and PC2, and the values in land use mixture, road density, and building density were relatively higher. The highly mixed land use mode reflects the concentration of facilities and the wide range of services, which helps promote the activity scope of residents and is beneficial to health. **Figure 3** also showed that communities 1, 2, 7, 11, 13 and 15 were in the negative interval of PC2. The values in community greening rate, coverage of park service radius in community, and coverage of greenway service radius in community were greater. These indicators were all negative, posing a certain inhibitory effect on health, which is required to improve the relevant indicators of green landscape to ensure the improvement of the comprehensive quality of healthy communities.

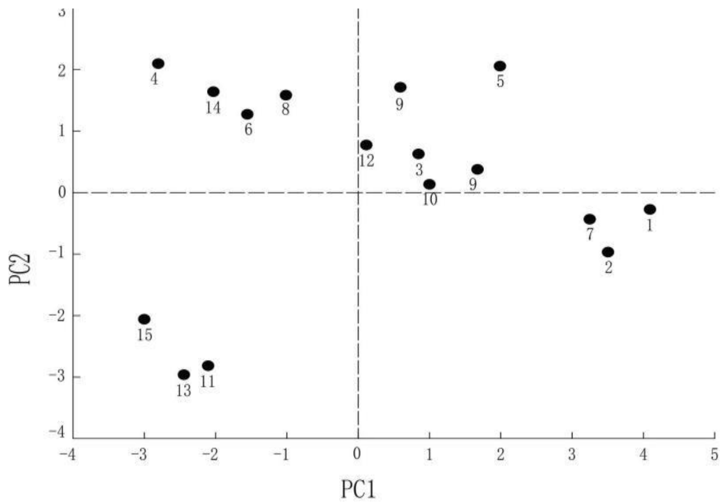

Figure 3. PCA score plot of PC1 & PC2.

The PC3 contained 11.73% of the original information. **Figure 4** shows the relationship between each community and PC2 & PC3: 3, 5, 7, 12 (Fengquan Gujing Community, Houjiatang Community, Jinmao Community, and Zoumalou Community) were in the positive interval of PC3. The values in the Coverage of Park Service Radius in Community and Coverage of Community Health Facilities (Buildings) were higher. The value of Pedestrian Route Directness was large, indicating that the paths to parks and medical facilities in these communities were blocked and the walking environment was poor. 1, 2, 11, and 13 (Furong Road Community, Huangni Street Community, Fengyuan Community, and Jingwanzi Community) were successively in the third section of PC2 and PC3. It is known from the above that the values in Pedestrian Route Directness, Coverage of Park Service Radius in Community, Building Density, and Plot Ratio of PC3 were higher than the average. The Greening Rate was negative, and the value of Coverage of Greenway Service Radius in Community was small, which indicates that in the urbanization process in Changsha, buildings and other facilities occupy the walking space and activity space of pedestrians and weaken the frequency and intensity of residents' physical activities. The Non-Motorized Transportation system in the community has an indirect effect on diseases commonly seen in the cities, such as cardiovascular and cerebrovascular diseases and obesity. It is necessary to improve the non-motorized transportation system of these 4 communities, improve the greenway and green area, and create a healthy living environment in the community. The distribution of 4 and 12 (Qilimiao Community and Zoumalou Community) was closest to and has little correlation with PC3. The Coverage of Park Service Radius in Community was higher, implying a better walking environment.

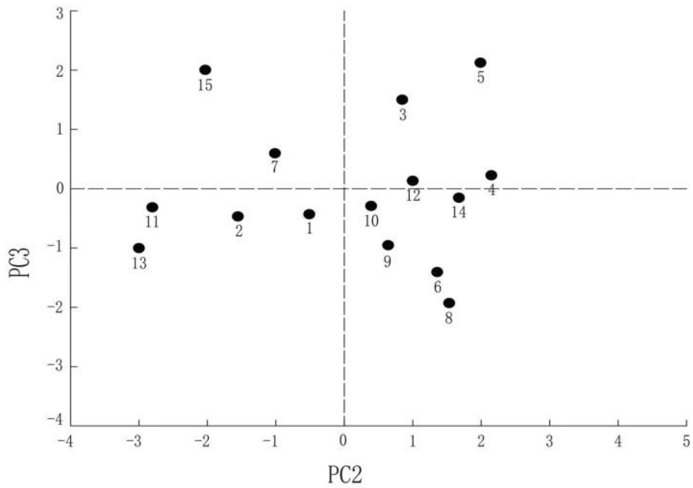

Figure 4. PCA score plots of PC2 and PC3.

Referring to the PCA score plots of PC4 and PC1, the values of greening rate and greenway coverage rate of 1, 2, 3 and 5 (Furong Road Community, Huangni Street Community, Fengquan Gujing Community and Houjiatang Community) are low. The residential land density and road density are high, and the road is crowded. Therefore, micro-transformation and improvement of greening and roads should be properly carried out. The remaining PCA score plots in this study will not be further listed.

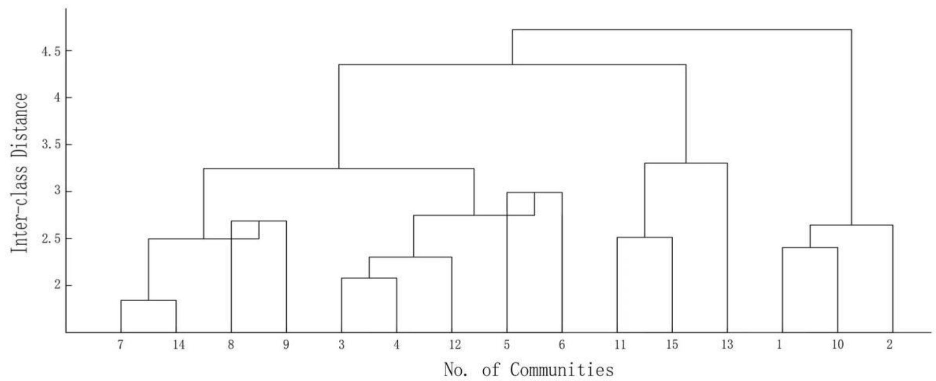

Figure 5. Clustering analysis dendrogram.

According to the Clustering Analysis Dendrogram (**Figure 5**), when the inter-class distance = 35, the 15 community samples were divided into 4 categories:

The first category included 4 communities, namely 7, 14, 8 and 9 (Jinmao Community, Zhongnan Community, Yunluyuan Community and Meiyuan community). This category mainly clustered communities with high values in Land Use Mixture, Coverage of Bus Stops within a 500-m Radius, Coverage of Greenway Service Radius in Community and Greening Rate. These communities are also of medium road density, good green space and landscape construction, relatively rich allocation of service facilities, and small neighborhood scale. Residents have a high degree of satisfaction and cohesion towards the community environment, which poses a positive impact on residents' physical and mental health. Such communities belong to commercial developmental communities. The result is mostly consistent with the results of negative interval of the PCA score plots of PC1 and PC3.

The second category included 5 communities: 3, 4, 12, 5 and 6 (Fengquan Gujing Community, Qilimiao Community, Zoumalou Community, Houjiatang Community and Ziyuan Road Community). In the urban expansion, these communities had developed insufficient facilities except catering and shopping, which affected travelling by walking to a certain extent. Such communities can increase the frequency and time of people's walking activities. But it is difficult for them to have a large space to walk. The health and elderly care facilities of these communities were lower than the average, and the residents' sense of belonging was not strong. The result is mostly consistent with the results of positive interval of the PCA score plots of PC2 and PC3. And such communities belong to mixed modern communities.

The third category included 11, 15 and 13 (Fengyuan Community, Banxiang Street Community and Jingwanzi Community). These communities are of high density and low coverage of various facilities. The health and elderly care facilities are obviously in short; the accessibility of transportation is low; and various built environment indicators have adverse effects on the health of the elderly group. In recent years, with the efforts of governmental departments, infrastructures like roads and traffic among the communities have shown evident improvement. However, the inconvenient transportation and the poor-quality greening (such as poor aesthetics, the limited cultivation of healthy plants, and so on) affect the comprehensive health quality of the communities. The result is basically consistent with the results of the third section of the PCA score plots of PC1 and PC2. And such communities belong to traditional old communities left over historically.

The fourth category included 3 communities—1, 10 and 2 (Furong Road Community, Taiping Street Community and Huangni Street Community), which are high-density communities and belong to core business developmental communities. The community plot ratio is large, and the rates of building density, land use mixture, and facility coverage are high. Life is convenient for residents, making these communities popular with young people. However, due to limited public greening and traffic congestion, the satisfaction of the communities is negatively affected. Green spaces and greening could be specifically explored while building healthy communities, to improve the greening. The planning and construction of secondary main roads and road network should be appropriately promoted, enabling residents to have more means of travelling. The result is basically consistent with the results of the third section of the PCA score plots of PC2 and PC3.

4. Discussion

A healthy built environment in communities allows residents to maintain good mental, physical and social health, thereby preventing and reducing the onset of mental and physical diseases. At present, there are so far no clear and systematic evaluation methods for the built environment in communities, either at home or abroad. And there are some differences in the selection of built environment indicators. In the selection of indicators in this study, both the quantity and specificity need to be further optimized. For example, community environmental factors that can prevent infectious diseases and healthy environmental factors of community sound were not involved. In the future, systematic research will be conducted on these important quality characteristics that affect community health.

Study, both the quantity and specificity need to be further optimized. For example, community environmental factors that can prevent infectious diseases and healthy environmental factors of community sound were not involved. In the future, systematic research will be conducted on these important quality characteristics that affect community health.

Principal component analysis and clustering analysis are increasingly used in quality analysis of multi-sample and multi-indicator studies by researchers. Principal component analysis (PCA) is a dimensionality reduction method that uses a small number of comprehensive indicators to replace most of the information of the original multiple indicators. It eliminates the unimportant part and retains the important information. Clustering analysis integrates research objects with closer relationships into one category, focuses on distinguishing intra-class and inter-class elements, and clarifies category boundaries. However, it does not delete information or distinguish the importance of elements. The inter-class importance is of no difference.

Based on this, this study carried out principal component analysis on 12 community indicators to extract more important information, and used the PCA score plot to analyze the relationship between each indicator of principal component, each indicator of healthy environment, and residents' health, which can provide intuitive theoretical basis for planning departments on the environmental planning of healthy communities. For example, traditional old communities, with high density and low facility coverage, are located in the third section of PC1 and PC2 score plots. Excessive density increases air pollution and discourages residents from physical activities, which rises the risk of chronic diseases such as obesity and stroke.

Clustering analysis can cluster the factors affecting the built environment of healthy communities, which helps identify the health hazard indicators of the built environment among communities of the same type. To prevent and control the health risks of local residents and improve the quality of their living environment, the traditional old communities should reduce the population density and avoid large-scale housing development. By optimizing the land use structure of the old urban area, reducing the proportion of residential land, increasing the land for public services, green spaces and squares, and road traffic, the old urban area could be renovated with a development mode of low building density and high intensity, so as to reduce the significant impact of the built environment on residents' chronic diseases.

5. Conclusions and suggestions

Based on the above analysis, we can draw the following conclusions:

1) 12 quality indicators collected from 15 communities in Changsha were measured, and the results were analyzed. The results show that in the high-density communities, the traditional old communities and mixed modern communities have a poorer quality of living environments, low coverage of health and elderly care facilities, relatively weak service facilities, and relatively poor functional attraction to walking activities, which need to be updated and improved. There

are deficiencies in the public greening and sidewalk systems of core business developmental communities and commercial developmental communities. Considering that greening and walking environments are important factors to improve the health of residents, it is urgent to improve the area and quality of natural landscape greening in these communities, and appropriately increase the planning and construction of secondary main roads and road network to create a healthy community environment.

While carrying out research and judgment on the quality of healthy communities, it is necessary to focus on the built environment indicators of healthy communities, for a more accurate and comprehensive evaluation of community quality. The results of descriptive analysis showed that the variation coefficients of the Accessibility of Medical Facilities, the Gini Coefficient of Green Park Area per capita (based on communities), and the Accessibility of Sports Facilities were 5.2%, 7.8% and 3.2%respectively, showing no intervarietal difference. The values of the three indicators ranged from 0.12 to 1.0, 0.90 to 0.78, and 0.45 to 1.0. The differences of these three indicators in other communities require further measurement and analysis in more communities.

2) Principal component analysis was carried out on the remaining 12 indicators, and 87.01% information of 4 principal components reflecting the original variables was extracted. Taking the first principal component PC1, the second principal component PC2 and the third principal component PC3 as examples, the score plot of principal components can intuitively obtain the relationship between various communities and principal components.

Communities with good facilities: 1, 2 and 10 (Furong Road Community, Taiping Street Community, Huangni Street Community) have large density values and good walkability, and the health and elderly care facilities meet the needs of residents; 11, 13, 15 (Jinmao Community, Jingwanzi Community, Banxiang Street community) have outdated facilities, lack of health and elderly care facilities, and low coverage of greenway, which does not meet the requirements of residents for healthy communities. It urgently needs micro-transformation to improve the quality of environmental indicators.

Community environment can support travel by walking and increase the frequency and duration of walking, carrying a positive effect. Among communities with high building density and high road density, the walkability is better in communities with relatively rich facilities. However, communities with relatively weak facilities have poorer functional attraction for walking activities.

3) Based on the community quality characteristics, 15 communities were divided into 4 categories by employing clustering analysis method, to preliminarily determine whether they meet the standard of built environment indicators of healthy communities. Clustering analysis and the score plot of principal components showed relatively consistent results of community categorization. There was a significant relationship between community quality and built environment indicators. Principal component analysis and clustering analysis can both be used to analyze the to-be-optimized indicators of the quality of healthy communities and the indicators determining better quality of built environments.

The built environment integrates green space, facilities, transportation and other indicators. According to these indicators, the built environments of different communities can be analyzed, and a forward-looking special renovation and optimization plan of the space environment can be formulated, together to build a healthier community space environment.

Finally, the following suggestions are given:

1) Optimize the allocation of community resources

Based on the results of principal component clustering analysis, areas with resource shortages or surpluses in the community can be identified. For regions with resource shortages, policy makers should prioritize investing resources, such as adding public facilities, improving transportation conditions, and enhancing greenery levels. For areas with excess resources, resource integration or optimization can be considered to avoid resource waste.

2) Promote balanced development of the community

Cluster analysis can help identify areas of uneven development within a community. Policy makers should take measures to promote balanced development in these regions, such as improving the economic development level and enhancing the quality of life of residents through policy guidance, capital investment, and other means.

3) Enhance the health level of the community

PC analysis can reveal key factors that affect community health. Policy makers should develop specific improvement measures based on these factors, such as strengthening health education, enhancing medical service levels, improving environmental sanitation, etc., in order to improve the overall health level of the community.

4) Strengthening community governance and resident participation

Cluster analysis can also reveal the problems and weak links in community governance. Policy makers should strengthen the construction of community governance systems and enhance governance capabilities. At the same time, encourage residents to participate in community affairs, enhance their sense of belonging and responsibility, and jointly promote community development.

5) Develop differentiated development strategies

According to the results of cluster analysis, different communities may have different characteristics and needs. Policy makers should develop differentiated development strategies and implement precise measures based on the characteristics and needs of different communities to achieve sustainable development.

6) Establish a monitoring and evaluation mechanism

To ensure the effective implementation of the above recommendations, policy makers should establish monitoring and evaluation mechanisms, regularly assess the development status of communities, and make adjustments and optimizations based on the evaluation results.

In summary, by applying the results of PC clustering analysis to the construction of healthy communities is of great significance for promoting urban planning and community development. Policymakers and planners should make full use of these analysis results and develop specific policy measures based on actual situations to achieve sustainable development of communities and the happiness and well-being of residents.

Author contributions: Methodology, XH; software and data curation, YL. All authors have read and agreed to the published version of the manuscript.

Funding: Key Research Projects funded by the Education Department of Hunan Province, 2022: Research on the Path to Improve the Spatial Quality of Healthy Communities in Hunan under the Regular Epidemic Prevention and Control, Project number: 22A0513.

Ethical approval: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of Comprehensive Evaluation of Health Community Construction of School of Design and Arts, Hunan Institute of Engineering (2023A15, 20230508).

Conflict of interest: The authors declare no conflict of interest.

References

1. Liu J F, Jia Y W. Users' intention to continue using online mental health communities: empowerment theory perspective. International journal of environmental research and public health, 2021, 18:9427.
2. Brownson R C, Paul C E. Revisiting the future of public health: the good, the bad, and the ugly. American Journal of Public Health, 2024, 114: 479-485.
3. Yang B, Min E, Bloom M S, et al. Community greenness, blood pressure, and hypertension in urban dwellers: The 33 Communities Chinese Health Study. Environment International, 2019, 126: 727-734.
4. Fayshal M A. Unveiling the impact of rapid urbanization on human comfort: a remote sensing-based study in Rajshahi Division, Bangladesh. Environment, Development and Sustainability, 2024, 33: 1-35.
5. Reou E. Thermal comfort in outdoor spaces and urban canyon microclimate. Renewable Energy, 2023, 55: 182-188.
6. Thomeer M B, Moody M D, Yahirun J. Racial and ethnic disparities in mental health and mental health care during the COVID-19 pandemic. Journal of racial and ethnic health disparities, 2023, 10 (2): 961-976.
7. Solano T, Bernal A, Mora D, et al. How bio-inspired solutions have influenced the built environment design in hot and humid climates. Frontiers in Built Environment, 2023, 9: 1267757.
8. Kim H, Rajbhandari A, Krile R, et al. Body mass index trajectories among the healthy communities study children: racial/ethnic and socioeconomic disparities in Childhood Obesity. Journal of Racial and Ethnic Health Disparities, 2024, 11(1): 203-215.
9. Liu Y, Huang X, Yang Q, et al. Effects of landscape on thermal livability at the community scale based on fine-grained geographic information: A case study of Shenzhen. Science of the Total Environment, 2023, 905: 167091.
10. Los C, Tsai H H. Design of 3D virtual reality in the metaverse for environmental conservation education based on cognitive theory. Sensors, 2022, 22 (21): 8329.
11. Lu H, Dong X X, Li D L, et al. Multimorbidity patterns and health-related quality of life among community-dwelling older adults: evidence from a rural town in Suzhou, China. Quality of Life Research, 2024, 33(5): 1335-1346.
12. Zhang W, Xia Z. A summary of the research on health-oriented community built environment construction and its enlightenment. Urban architecture, 2021, 29: 7-11.
13. Xue H, Cheng X, Jia P, et al. Road network intersection density and childhood obesity risk in the US: a national longitudinal study. Public Health, 2020, 178:31-37.
14. Bruno G, Cavola M, Diglio A, et al. Improving spatial accessibility to regional health systems through facility capacity management. Socio-Economic Planning Sciences, 2020, 71: 100881.
15. Bian H, Ren Z, Zou Y, Yuan S. Quantitative study on the proportion of functional composition of functional mixed Residential area--A case study of Tianjin. Southern architecture, 2021(01):28-34

16. Xie B, Zheng Y, Li Z, et al. Influence of high-density urban built environments on stroke in China: A case study of Wuhan. China City Planning Review, 2022, 3:23-29.
17. Luque G, Muxika J, Mendia O, et al. Green and blue space exposure and non-communicable disease related hospitalizations: A systematic review. Environmental Research, 2023, 4: 118059.