



OPEN Geographical analysis of malignant tumor incidence and treatment in China

Yudong Bian^{1,4}, Xunyi Yu^{2,4}, Jiang Zhang^{3,4}, Yongsheng Zhao^{1✉}, Ming Zheng¹, Chunzhen Tang¹, Xiaolei Zhang¹ & Lili Yue¹

Malignant tumors have a severe impact on both patients and their families, and the treatment of malignant tumors is a fundamental component of public health security. Therefore, the issue of malignant tumors is a significant focus in the construction of “Healthy China.” There are notable spatial disparities in both the incidence and treatment of malignant tumors in China. Research from a geographical perspective and methodology can help improve local malignant tumor prevention efforts and optimize the national allocation of medical resources. This study collected case samples publicly disclosed by representative malignant tumor mutual assistance platforms in China. The analysis was conducted using prefecture-level regions as the basic unit, applying spatial correlation analysis, geographic detectors, multi-factor regression, geographic inequality decomposition, birth cohort analysis, and other methods. The key findings are as follows: (1) there is a strong spatial aggregation of six high-incidence malignant tumors: respiratory system, digestive system, breast system, reproductive system, urinary system, and endocrine system cancers; (2) natural and human geographical factors have significant explanatory power for the occurrence of malignant tumors, although the influencing factors vary widely among different types of tumors; (3) there are considerable disparities in the distribution of medical resources for malignant tumors across regions, and when patients choose to seek treatment in different locations, they take both geographical distance and the level of medical services into account; (4) the difference in the “treatment occurrence rate” of malignant tumors is substantial between the two sides of the “Hu Huanyong Line”; (5) the degree of inequality in the national “treatment occurrence rate” for malignant tumors is widening with the increasing number of patients born later, but is narrowing with the establishment of new hospitals. Finally, this paper proposes policy recommendations to improve the prevention, control, and treatment of malignant tumors based on local conditions.

Keywords Malignant tumors, Hospitals, Geographic inequality, Health geography, Healthy China, Insurance

Health is the result of the integrated effects of natural, cultural, and social factors, and it is also a geographical phenomenon in space¹. One manifestation of deteriorating health is the occurrence of malignant tumors, which impact residents’ health life expectancy and disrupt the normal lives of families. The treatment of malignant tumors is a fundamental public demand, and it is an area of high concern for both the public and the government. The issue of malignant tumor occurrence falls within the domains of medical and environmental geography^{2–4}, while the distribution of medical resources for the treatment of malignant tumors primarily belongs to human economic geography, both of which are key areas of focus in the construction of “Healthy China”¹.

Research in China’s medical and health geography field on endemic diseases is relatively mature^{5,6}. Existing literature has extensively studied malignant tumors in the respiratory system⁷, digestive system⁸, breast system⁹, reproductive system^{10,11}, urinary system, and endocrine system, along with related health indicators for residents^{12–21}. These studies, in terms of geographical spatial levels, primarily focus on provincial and higher research scales^{21,22}, with fewer studies analyzing prefecture-level cities and county units^{11–20}. Other literature has analyzed finer-grained geographical units within specific regions^{13–16}, such as more than 300 streets in Beijing¹⁰, towns in Jingyuan County, Ningxia¹², districts in Xi’an¹⁴, rural areas in Jinan¹⁵, and districts in Wuhan¹⁶.

¹Liaoyang Petrochemical General Hospital, Liaoyang 111000, China. ²Affiliated Hospital of Liaoning University of Traditional Chinese Medicine, Shenyang 110000, China. ³Key Laboratory of TCM Viscera Theory and Application of Ministry of Education, Liaoning University of Chinese Medicine, Shenyang 110000, China. ⁴Yudong Bian, Xunyi Yu and Jiang Zhang contributed equally to this work. ✉email: 13998044361@163.com

Regarding the spatial allocation of medical resources, existing literature mainly studies indicators such as the number of medical and health institutions per capita^{24–27}, the number of hospital beds per capita²⁴, and the number of doctors per capita^{25–27}. An increasing number of studies utilize real patient data from specific regions to investigate the accessibility of medical resources, such as in Beijing²⁸, Nanjing²⁹, and Shenzhen³⁰.

This study collects information from over 200,000 individuals and constructs a sample database based on China's prefecture-level regions. It maps the geographical distribution of six high-incidence malignant tumors and analyzes the impact of geographical factors on the occurrence of malignant tumors, which will enrich research in medical and health geography to some extent¹. By describing the geographical distribution of malignant tumor medical resources and patients' treatment destinations, this paper integrates the "occurrence" and "treatment" of malignant tumors^{12,13}. It investigates the geographical inequality in medical service supply and demand at the prefecture level and performs birth cohort analysis separately for the "occurrence side" (patients) and the "treatment side" (hospitals), thereby examining the evolution of geographical inequality. This research also extends studies on the spatiotemporal status of medical and health resources^{24–26}.

Geography can play an important role in at least 18 fields related to the construction of "Healthy China," and this paper addresses four of these: malignant tumor monitoring and prevention, health promotion in key regions, optimization of healthcare resource allocation, and the accessibility and equity of health services. This research provides decision-making references for improving the prevention, control, and treatment of major diseases in China and its regions. It also offers data support for the implementation of the report from the "20th National Congress of the Communist Party of China," which calls for the "promotion of the expansion of high-quality medical resources and their regional balanced distribution."

Research data and methods

Malignant tumors and hospitals

Identification of malignant tumors and hospitals

The research data is derived from the publicly disclosed information of one of China's most representative malignant tumor mutual assistance platforms (<https://www.zgdbjz.org.cn/>). The platform has been operating since October 16, 2018, with an average membership of over 80 million per session³¹. By the end of March 2022, the platform had disclosed a total of 202,092 cases. This study utilized an automated case query tool to capture screenshots of the disclosed materials, and text recognition software was used to extract textual information.

The types of malignant tumors covered by the mutual assistance platform are the same as those commonly found in malignant tumor disease insurance. The tumors are classified based on the body site where they initially occur. This study selects the six most common types of malignant tumors, which are: respiratory system malignant tumors, digestive system malignant tumors, breast system malignant tumors, reproductive system malignant tumors, urinary system malignant tumors, and endocrine system malignant tumors. These six types represent 17.45%, 13.97%, 12.13%, 6.50%, 6.38%, and 4.43% of all the cases in the sample, respectively. These six malignant tumor types are highly representative.

The location of malignant tumor occurrence is defined as the patient's usual place of residence. This study uses prefecture-level (and above) regions (excluding data from Hong Kong, Macao, and Taiwan) as the study area, excluding regions with small populations and few cases. In the analysis of proportion-based indicators, only 308 regions with more than 30 cases are included (Fig. 1). The names of hospitals in the selected data were standardized, resulting in a total of 6,432 hospitals that issued diagnostic certificates to the patients in the sample. Based on the standardized names of these hospitals, the corresponding prefecture-level (and above) regions where these hospitals are located were identified as the treatment locations, totaling 348 regions³².

Tumor classification scheme

In accordance with the International Classification of Diseases for Oncology (ICD-10) and the GLOBOCAN classification system used by the International Agency for Research on Cancer (IARC), we mapped the six major categories of malignant tumors into specific anatomical sites as follows (Table 1): Respiratory system cancers (ICD-10: C30–C39): including nasal cavity (C30), nasopharynx (C11), oropharynx (C10), larynx (C32), trachea (C33), bronchi and lungs (C34); Digestive system cancers (ICD-10: C15–C26): including esophagus (C15), stomach (C16), small intestine (C17), colon (C18), rectum (C20), liver (C22), gallbladder (C23), pancreas (C25); Breast cancers (ICD-10: C50): female and male breast malignancies; Reproductive system cancers (ICD-10: C51–C58 for females; C60–C63 for males): including cervical cancer (C53), ovarian cancer (C56), endometrial cancer (C54), prostate cancer (C61), testicular cancer (C62); Urinary system cancers (ICD-10: C64–C68): including kidney (C64), renal pelvis (C65), ureter (C66), bladder (C67); Endocrine system cancers (ICD-10: C73–C75): including thyroid gland (C73), adrenal gland (C74), other endocrine glands (C75). The classification of each case into these broad categories was based on the diagnostic information disclosed on the mutual aid platform. ICD codes were approximated based on tumor anatomical site names and expert mapping.

Inclusion and exclusion criteria

The dataset used in this study was obtained from publicly disclosed diagnostic cases on a national malignant tumor mutual aid platform. The inclusion criteria were as follows:

1. Patients with a clearly indicated diagnosis of malignant tumor, as determined by a physician and supported by a diagnostic certificate;
2. Tumors classified as malignant based on standard clinical definitions (ICD-10 codes), regardless of histological subtype;
3. Patients with available data on residence location and hospital of diagnosis;
4. Cases diagnosed between October 2018 and March 2022.

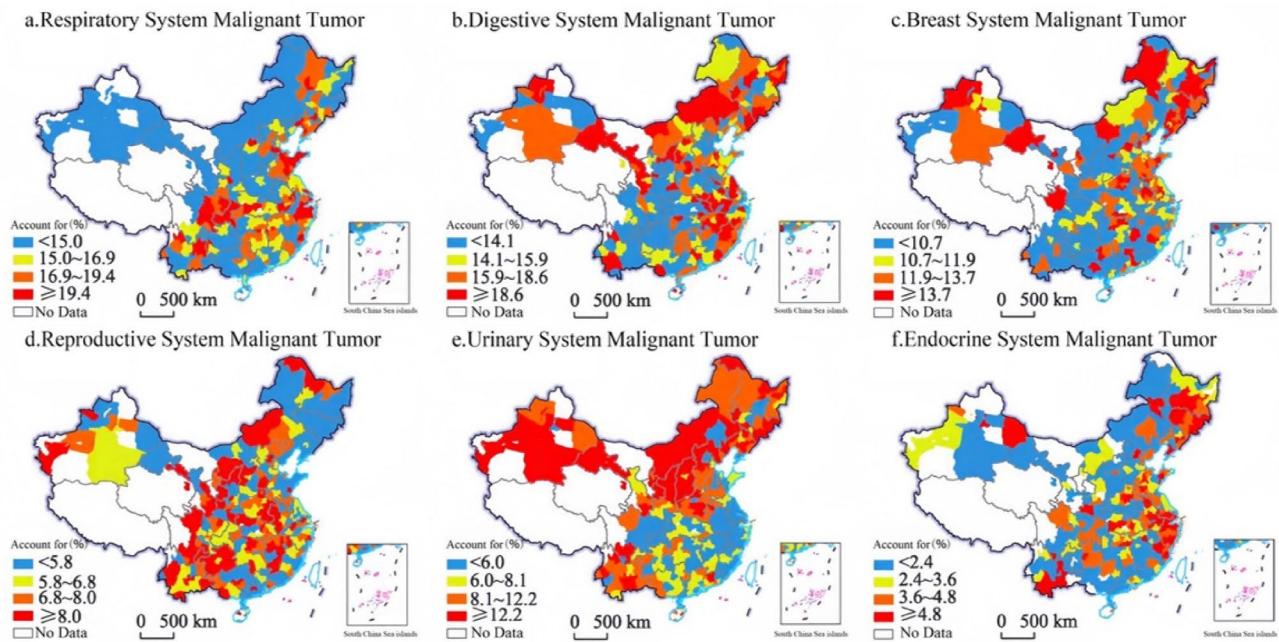


Fig. 1. Proportion in the occurrence of six types of malignant tumors in prefecture-level regions from October 2018 to March 2022. *Note:* The base map was created using the standard map with the approval number GS(2020)4630 issued by the Ministry of Natural Resources, and the boundary of the map was not modified. To better reflect the spatial variation in the incidence proportion of each tumor type, we used customized classification intervals (cutoffs) for each map in Fig. 1. This approach allows for the clear visualization of intra-type spatial clustering even for tumor types with lower national average proportions (e.g., urinary and endocrine systems). The color scales are therefore not uniform across maps and are optimized for each tumor type's data distribution. This is a common approach in thematic mapping when the goal is to visualize spatial patterns rather than directly compare magnitudes across different variables.

Tumor category	ICD-10 range	Representative sites	No. of cases	% of total cases
Respiratory system	C30–C39	Nasal cavity, lung, trachea, larynx	35,260	17.45%
Digestive system	C15–C26	Esophagus, stomach, colon, liver, pancreas	28,242	13.97%
Breast	C50	Female and male breast	24,516	12.13%
Reproductive system	C51–C58, C60–C63	Cervix, ovary, uterus, prostate, testis	13,139	6.50%
Urinary system	C64–C68	Kidney, bladder, ureter	12,888	6.38%
Endocrine system	C73–C75	Thyroid, adrenal glands	8951	4.43%
Total	–	–	202,092	100%

Table 1. Composition of six major tumor categories by ICD-10 code and sample size.

The exclusion criteria were:

1. Patients with missing or ambiguous tumor site data that could not be mapped to ICD-10 codes;
2. Duplicate entries from the same patient;
3. Benign, borderline, or in-situ tumors;
4. Pediatric cases under 18 years old (due to different pathophysiological characteristics and treatment patterns).

Importantly, due to the aggregated and anonymized nature of the data disclosed by the platform, histological subtypes (e.g., adenocarcinoma vs. squamous cell carcinoma in lung cancer) were not available for analysis. Therefore, the classification of cancers in this study is based on anatomical site only, as is common in population-based cancer surveillance systems (e.g., GLOBOCAN, National Cancer Registry).

Hospital classification and grouping

In our study, a total of 6,432 hospitals were identified based on diagnostic certificates issued to patients. We classified these institutions into three broad categories based on publicly available hospital registration and naming information:

1. Specialised cancer hospitals (n = 183): institutions exclusively dedicated to oncology treatment and research;
2. General hospitals with dedicated oncology departments (n = 2346): large tertiary or secondary hospitals that operate oncology departments within a multiprofile setting;
3. General or local hospitals without clear oncology departments (n = 3903): hospitals where cancer treatment is provided, but oncology is not clearly distinguished as a separate unit.

Due to the nature of the data obtained from the mutual aid platform, some classifications are inferred based on hospital names (e.g., “Tumor Hospital,” “Oncology Department,” or known tertiary-level classification), and may contain minor uncertainties. However, all institutions included have officially issued tumor-related diagnostic certificates and therefore meet the criteria of providing oncology care. As shown in Supplementary Table S1.

Malignant tumor occurrence intensity by region

To measure the intensity of a certain type of malignant tumor in a region, the proportion of this type of malignant tumor in all malignant tumors in the region is usually used. This paper defines the proportion of malignant tumor k in region i = the number of cases of malignant tumor k in region i / the number of cases of all malignant tumors in region i . “Gender” and “age” are two basic variables that affect the disease³³. Therefore, in order to obtain a more robust explanatory power of geographical factors, gender and age factors are removed before analysis. The formula for calculating the intensity of 6 types of malignant tumors after removing gender and age factors in each region is as follows:

$$y_Remove\ gender\ and\ age\ factors_i^k = y_i^k - y_Baseline_i^k \\ = y_i^k - \sum_{Gender, Age} (w_{i, Gender, Age} \times y_{Gender, Age}^k) \quad (1)$$

In order to control for demographic structure differences and better isolate the effect of geographical factors, we applied a method analogous to direct standardization, a commonly used approach in spatial epidemiology and health geography. Specifically, we reweighted the sex- and age-specific tumor incidence in each region according to the national population distribution across four age groups and both sexes. This allowed us to calculate the adjusted tumor proportion by region while removing demographic bias³⁴. A detailed step-by-step example of this adjustment process is provided in the Supplementary Table S2 for clarity and reproducibility.

In the formula: $y_Baseline_i^k$ is the proportion of malignant tumor k in area i when assuming that area i has a national overall sex-specific occurrence state, which is jointly determined by $w_{i, Gender, Age}$ and $y_{Gender, Age}^k$; $w_{i, Gender, Age}$ is the proportion of all cases in region i by gender and age group. Gender is divided into male and female, and age is divided into 4 groups according to the age of the year, which are <40 years old, 40–50 years old, 50–60 years old and ≥60 years old.

Geographical factors

This study analyzes the impact of 8 natural geographical factors and 6 human geographical factors on the occurrence of malignant tumors, based on existing theories and literature, as well as the availability of data at the prefecture level.

Natural geographical factors

Average Elevation and Terrain Relief. The altitude directly determines air pressure, influencing climate conditions, and has an effect on lung function and overall health. Data on altitude is obtained from the DEM (Digital Elevation Model) data of terrain surface morphology, released by the Ministry of Natural Resources’ Data Service Center, and processed using ArcGIS 10.8 (Esri) software (<https://www.gscloud.cn>). Terrain relief affects the comfort of the natural environment, population density, and so on, which in turn influences disease occurrence³⁵. The data for terrain relief is referenced from the work of Feng Zhiming et al. The average elevation and terrain relief are based on the measurements from January 2020.

Meteorological Factors: (1) January average temperature: Affects the respiratory system, lung function, and thyroid hormone levels, etc.; (2) July average temperature: Affects the respiratory system, lung function, and thyroid hormone levels, etc.; (3) Annual precipitation: Affects the respiratory and digestive systems, among others; (4) Relative humidity: Affects the respiratory system, etc.; (5) Annual sunshine hours: Affects the respiratory system¹⁷; (6) Average wind speed: Affects lung function, respiratory diseases, etc. Meteorological data is obtained from the surface weather station observation data released by the National Meteorological Information Center of China Meteorological Data Network. The data used is from the nearest weather station to the geometric center of each sample region. Considering the birth years of patients in the sample, the meteorological data used for this study covers the average values from 1970 to 2021 (<https://data.cma.cn>).

Human geographical factors

(1) **Income Level:** Income level reflects the overall economic capacity of a region and may influence the development of healthcare infrastructure, availability of medical services, and regional investment in early detection and public health initiatives. In this study, regional income level is measured by per capita Gross Domestic Product (GDP) (Collected from the China City Statistical Yearbook (various years), published by the National Bureau of Statistics). (2) **Urbanization Rate:** Urbanization affects working and resting conditions, population density, and the accessibility of early intervention measures, which in turn impacts health. Studies have found significant differences in health and disease status between urban and rural residents³⁶. The urbanization rate of a region is measured by the proportion of the urban population to the total population. (3) **Education Level:** Education

affects individuals' thinking and behaviors. Education level has a significant relationship with diseases in the digestive system, endocrine system, reproductive system, etc.³⁷. Given the relatively low level of health literacy among Chinese residents and the large regional disparities, education level is more likely to influence health behaviors and health status. The education level of a region is measured by the proportion of college students per 10,000 people (Sourced from the Ministry of Education's national higher education database and city-level yearbooks). (4) Public Health Resources: Diseases can be prevented and controlled, and an important aspect of public health is the control of chronic diseases and malignant tumors. The significant impact of public health resources on residents' health status has been supported by several studies³⁸. Public health resources in a region are measured by the number of hospital and health center beds per 10,000 people (Obtained from the China Health Statistical Yearbook). (5) Water Pollution: Water pollution affects water quality and soil, and in turn influences foodborne diseases, digestive system diseases, reproductive system diseases, and overall public health. The water pollution situation in a region is measured by the per capita industrial wastewater discharge (Data collected from the China Environmental Statistical Yearbook published by the Ministry of Ecology and Environment of China). (6) Air Pollution: Air pollution has been a topic of significant concern in recent years and has a notable impact on respiratory system diseases, digestive system diseases, reproductive system diseases, and population health³⁹. The air pollution situation in a region is measured by the per capita industrial sulfur dioxide (SO₂) emissions (Data collected from the China Statistical Yearbook on Environment published by the Ministry of Ecology and Environment of China).

Research methods

Spatial autocorrelation test

A rigorous method to determine the spatial correlation of variables is to perform a spatial autocorrelation test⁴⁰. The global Moran's I index is calculated for the occurrence proportion of each type of malignant tumor, and its formula is as follows:

$$I = \frac{\sum_{i=1}^N \sum_{i'=1}^N w_{ii'} (y_i - \bar{y}) (y_{i'} - \bar{y})}{(\sum_{i=1}^N \sum_{i'=1}^N w_{ii'}) \sum_{i=1}^N (y_{ii'} - \bar{y})^2 / N}, \bar{y} = \frac{1}{N} \left(\sum_{i=1}^N y_i \right), i \neq i' \quad (2)$$

In the formula: $w_{ii'}$ indicates whether regions i and i' are geographically adjacent (0 or 1). The $w_{ii'}$ values for all pairs of regions form the spatial weight matrix; y_i and $y_{i'}$ represent the proportions of malignant tumor cases of a particular type in regions i and i' among all malignant tumor cases in those regions, respectively; $\bar{y}$ is the average of y_i across all regions; N is the number of regions in the sample. The value of Moran's I index ranges from -1 to 1. A higher value indicates a stronger positive spatial correlation of the variable at the national level.

Geographic detector

This study uses the factor detection method from geographic detectors⁴¹ to analyze the explanatory power of individual geographical factors on the occurrence of malignant tumors. The detection power is denoted as q , and its calculation formula is as follows:

$$q = 1 - \frac{\sum_{h=1}^L \frac{N_h \sigma_h^2}{N \sigma^2}}{N \sigma^2} = 1 - \frac{SSW}{SST} \quad (3)$$

In the formula: h represents a specific stratum of the factor, with each factor in this study divided into 10 strata, where $h = 1, 2, \dots, 10$; N_h and σ_h^2 represent the sample size (number of regions) and variance within the h stratum, respectively; N and σ^2 represent the total sample size (total number of regions) and overall variance, respectively; SSW is the sum of the within-stratum variances; SST is the total variance. The value of q ranges from [0, 1]. The higher the value, the stronger the association between the factor and the occurrence of malignant tumors.

Multiple regression analysis

To overcome the multicollinearity issue and identify the most significant factors, this study uses the stepwise regression method⁴². Initially, all influencing factors are included as independent variables in the regression model. The factor with the highest p value is removed, and the process continues until all remaining factors have p values smaller than 0.1. The regression model is as follows:

$$y_i^k = \alpha^k + X1_std_i' \beta^k + X2_std_i' \gamma^k + \varepsilon_i^k \quad (4)$$

In the formula: y_i^k represents the occurrence intensity of malignant tumor type k in region i ; $X1_std_i'$ represents the natural geographical factors, $X2_std_i' \gamma^k$ represents the human geographical factors. Both values have been standardized to visually reflect the magnitude of the coefficients of different variables; β^k and γ^k are the coefficients to be estimated for the natural and human geographical factors, respectively; α^k is the constant term; ε_i^k is the random error term.

To further evaluate potential multicollinearity among the geographic variables, we calculated the Variance Inflation Factor (VIF) for all independent variables included in the regression models. As shown in Supplementary Table S3, all VIF values were below 5, indicating no serious multicollinearity problems.

Theil index and its regional decomposition

The Theil index is used to measure the geographical inequality of the "treatment-to-incidence ratio" of malignant tumors in China. The calculation formula is as follows⁴³:

$$Theil = \frac{1}{N} \sum_{i=1}^N \frac{Z_i}{\bar{Z}} \ln \frac{Z_i}{\bar{Z}} \tag{5}$$

In the formula: Z_i represents the "treatment-to-incidence ratio" of malignant tumors in region i ; $\bar{Z}$ is the average "treatment-to-incidence ratio" of malignant tumors across all regions; N is the number of regions. The larger the value of the *Theil* index, the higher the degree of inequality. Geographical division in China is quite distinct, and the geographical inequality at the national level includes both within-region and between-region disparities⁴⁴. The *Theil* index at the national level is decomposed into two parts, with the calculation formula as follows:

$$\begin{aligned} Theil &= Theil_{Within-region} + Theil_{Between-region} \\ &= \sum_{j=1}^J \left(\frac{N_j}{N} \times \frac{\bar{Z}_j}{\bar{Z}} \right) Theil_j + \sum_{j=1}^J \left(\frac{N_j}{N} \times \frac{\bar{Z}_j}{\bar{Z}} \right) \ln \frac{\bar{Z}_j}{\bar{Z}} \end{aligned} \tag{6}$$

In the formula: The "within-region" variable j represents different regions; J represents the number of regions; N_j represents the number of regions within region j ; N represents the total number of regions in the country; $\bar{Z}_j$ represents the average "treatment-to-incidence ratio" of malignant tumors in region j ; $\bar{Z}$ represents the average "treatment-to-incidence ratio" of malignant tumors across all regions in the country. For the "between-region" part, in addition to the variables used in the within-region part, there is also $Theil_j$, which represents the *Theil* index of the "treatment-to-incidence ratio" of malignant tumors in region j , reflecting the differences between regions within region j .

Geographic analysis of malignant tumor incidence
Geographic distribution of malignant tumor incidence

There is a noticeable spatial variation in the incidence of six types of malignant tumors in China (Fig. 1). The incidence of respiratory system malignant tumors is highest in East China at 18.58%, followed by Southwest China at 16.67%, and lowest in Northwest China at 9.32%. The incidence of digestive system malignant tumors is higher in Northeast, East, and North China, with rates of 16.50%, 16.22%, and 16.09%, respectively, and lowest in Southwest China at 11.59%. The spatial variation in the incidence of breast system malignant tumors is the smallest, with the highest incidence in Northeast China at 12.26%, and lower incidences in Northwest and Southwest China at 9.00% and 8.26%, respectively. The incidence of reproductive system malignant tumors is higher in the western regions, with rates of 9.00% in Northwest China and 8.26% in Southwest China, while Northeast China has the lowest incidence at 4.56%. The incidence of urinary system malignant tumors is higher in the northern regions, with Northwest, North, and Northeast China having incidences of 14.91%, 13.96%, and 11.35%, respectively. The incidence of endocrine system malignant tumors is highest in East China at 4.26%, while South China and Northwest China have lower incidences of 0.97% and 1.09%, respectively.

The Moran's *I* index for the spatial clustering of the incidence of the six types of malignant tumors is significant at a level below 0.001 (Table 2), indicating a strong spatial clustering for all six types of malignant tumors. In terms of the Moran's *I* index for different tumor systems, the urinary system malignant tumors have the highest index at 0.5917, while breast system malignant tumors have the lowest index at 0.1807. The spatial clustering of the six types of malignant tumors is ranked from strongest to weakest as follows: urinary system malignant tumors, respiratory system malignant tumors, endocrine system malignant tumors, digestive system malignant tumors, reproductive system malignant tumors, and breast system malignant tumors.

Factors affecting the incidence of malignant tumors

Geographic detector analysis

Table 3 presents the detection power values for 14 geographic factors on the incidence of six types of malignant tumors. Overall, there is a strong association between geographic factors and the occurrence of malignant tumors, with detection power values in most cells being statistically significant (at the 5% level). Comparing natural geographic factors with human geographic factors, it can be observed that the average association of eight natural geographic factors with the six types of malignant tumors is 11.10%, which is higher than the average association of six human geographic factors with the tumors (7.84%). Among the geographic factors, average elevation has the highest association with malignant tumor occurrence, followed by topographic variation and

Type of malignant tumor	Respiratory system	Digestive system	Breast system	Reproductive system	Urinary system	Endocrine system
Moran's I	0.4116	0.3653	0.1807	0.2165	0.5917	0.3790
Z-value	10.2431	8.9393	4.3279	5.6663	14.9529	9.3000
Significance level	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Spatial pattern	Strong clustering	Strong clustering	Strong clustering	Strong clustering	Strong clustering	Strong clustering

Table 2. Spatial autocorrelation test of six types of malignant tumors' occurrence proportion in prefecture-level regions.

Geographic factor	Respiratory system malignant tumor	Digestive system malignant tumor	Breast system malignant tumor	Reproductive system malignant tumor	Urinary system malignant tumor	Endocrine system malignant tumor	Average
Natural geographic factors							
Average elevation	25.41	8.82	11.51	9.65	20.82	16.63	15.47
Topographic variation	18.99	3.81	12.05	9.77	20.23	9.78	12.44
Average temperature in January	18.18	5.57	7.59	3.98	24.71	7.94	11.33
Average temperature in July	16.65	4.28	5.52	6.93	25.75	9.36	11.42
Annual precipitation	24.03	6.67	5.78	3.64	23.64	6.19	11.66
Average annual relative humidity	21.01	2.70	4.81	3.19	27.46	8.09	11.21
Annual sunshine hours	19.31	4.02	5.80	1.37	24.99	6.01	10.25
Average annual wind speed	4.26	4.46	7.27	4.91	3.52	5.70	5.02
Human geographic factors							
Income level	7.53	4.14	10.85	3.80	8.23	23.49	9.67
Urbanization rate	12.90	4.74	23.28	6.19	7.29	17.99	12.07
Education level	9.14	2.83	8.81	4.04	8.78	9.01	7.10
Public health resources	4.48	2.25	4.48	4.00	3.16	8.06	4.41
Water pollution	10.09	3.68	11.69	5.53	8.43	9.00	8.07
Air pollution	8.79	4.58	4.30	3.51	9.64	3.35	5.70
Average	14.34	4.47	8.84	5.04	15.48	10.04	

Table 3. Detection power of natural and human geographical factors on the occurrence proportion of six types of malignant tumors in prefecture-level regions (%). The q-values reflect the explanatory power of each factor. Although no formal *p* values are reported, a higher q-value generally indicates stronger association. As a rule of thumb, q-values greater than 0.1 are often interpreted as having moderate to strong explanatory power in spatial analysis.

urbanization rate⁴⁵. In contrast, average annual wind speed and public health resources show a significantly weaker association compared to other geographic factors. In terms of specific tumor types, the association between urinary system malignant tumors and respiratory system malignant tumors with geographic factors is relatively high, while the association between digestive system malignant tumors and reproductive system malignant tumors with geographic factors is lower. Regarding the combinations of geographic factors and tumor types, average elevation has a particularly high association with respiratory system malignant tumors, reaching 25.41%, which is significantly higher than the average association of average elevation (15.47%) and the average association of respiratory system malignant tumors (14.34%). This suggests that average elevation has an impact on the incidence of respiratory system malignant tumors. For example, the association between public health resources and urinary system malignant tumors is low, at 3.16%, which is below the average association of public health resources (4.41%) and the average association of urinary system malignant tumors (15.48%), indicating that public health resources have little impact on the occurrence of urinary system malignant tumors¹⁵.

Multiple regression analysis

Based on knowledge of geography and regional development, and by calculating the Pearson correlation coefficients between pairs of factors, it was found that many influencing factors are significantly correlated. The correlation coefficients between natural geographic factors are higher than those between human geographic factors. Among the factors with the highest absolute correlation coefficients with other factors, the top three are annual precipitation, average temperature in January, and average temperature in July. Stepwise regression was used to estimate the influencing factors for six types of malignant tumors, and the results are shown in Table 4.

Respiratory System Malignant Tumors. For respiratory system malignant tumors, the significant influencing factors identified in the regression model include topographic variation (negative), January average temperature (negative), annual precipitation (negative), and income level (positive). Topographic variation can promote physical activity, enhancing the survival ability of lung cells. Additionally, regions with larger topographic variation generally have better environmental quality, so the effect of topographic variation is negative^{5,6}. Lower winter temperatures cause residents to stay indoors more, reducing exposure to outdoor air pollution. Cold temperatures also reduce the spread of respiratory viruses, which helps lower the risk of lung cancer. Precipitation can wash and dilute pollutants in the air, further helping to reduce the risk of lung cancer. High-income areas in China typically have more industrial pollution and occupational exposure risks, which increases the risk of lung cancer.

Digestive System Malignant Tumors. Higher humidity in living environments accelerates the growth of bacteria and molds, and increases the spread of aerosol substances in the air, which is detrimental to digestive system health. Intense ultraviolet radiation affects the intestinal mucosa and immune system, increasing the risk of digestive system malignant tumors^{26–28}. The development of public health initiatives promotes healthier diets and improved control of infectious diseases, such as the active treatment of *Helicobacter pylori* infections. Water pollution leads to increased consumption of heavy metals, microorganisms, and parasites, which accumulate carcinogenic factors in the digestive system.

Variable	Respiratory system malignant tumor	Digestive system malignant tumor	Breast system malignant tumor	Reproductive system malignant tumor	Urinary system malignant tumor	Endocrine system malignant tumor
Constant	0.150*** (0.003)	0.158*** (0.003)	0.107*** (0.002)	0.074*** (0.004)	0.100*** (0.004)	0.030*** (0.001)
Average elevation				0.021 (0.013)	0.033*** (0.009)	−0.005*** (0.002)
Topographic variation	−0.014** (0.005)		−0.005 (0.002)			
Average temperature in January	−0.028*** (0.005)				−0.026** (0.009)	0.010* (0.003)
Average temperature in July						0.009* (0.003)
Annual precipitation	−0.028*** (0.005)				−0.026** (0.009)	0.010* (0.003)
Average annual relative humidity		0.019*** (0.004)			0.005 (0.004)	−0.003 (0.002)
Annual sunshine hours		0.015*** (0.004)			0.010 (0.006)	0.005* (0.002)
Average annual wind speed				0.005* (0.002)		
Income level	0.007** (0.003)	−0.004 (0.003)	0.005 (0.003)		−0.006** (0.002)	0.009*** (0.001)
Urbanization rate			0.015*** (0.004)	−0.009 (0.005)		
Education level	0.003 (0.002)		0.004* (0.002)	−0.003 (0.002)	−0.008** (0.003)	0.004*** (0.001)
Public health resources		−0.011*** (0.001)		−0.014* (0.006)		0.002 (0.001)
Water pollution		0.003* (0.002)				
Air pollution			0.006** (0.002)		−0.003 (0.003)	
Goodness of fit (R ²)	0.37	0.16	0.36	0.21	0.55	0.41
Sample size	294	294	294	294	294	294

Table 4. Multivariable regression analysis on the proportion of malignant tumors occurrence in prefecture-level regions. A stepwise regression analysis was used, where the least significant influencing factor was removed in each step until all influencing factors had a *p* value less than 0.1. ***, **, and * represent the significance of the influencing factors at the 0.1%, 1%, and 5% levels, respectively.

Breast System Malignant Tumors. Adequate exposure to ultraviolet B (UVB) radiation from sunlight is essential for vitamin D synthesis, which may play a protective role in breast cancer development through its influence on hormonal pathways and cell proliferation. However, in regions with colder climates—especially in northern China—sunlight exposure in winter tends to be limited, contributing to higher rates of vitamin D deficiency^{46,47}. It is therefore plausible that breast cancer risk may be influenced by both geographic variation in UV exposure and regional mitigation measures such as supplementation. In our study, we found a weak but significant relationship between annual sunshine hours and breast cancer incidence, suggesting a possible link worthy of further investigation. The detection rate of breast diseases is higher in urban areas, where residents with higher education levels tend to undergo more early screening for breast diseases, leading to earlier detection of breast cancer. Particulate matter and chemicals in air pollution can cause oxidative stress and inflammation. Polycyclic aromatic hydrocarbons, volatile organic compounds, heavy metals, and other substances can damage the breast tissue.

Reproductive System Malignant Tumors. The causes of reproductive system diseases include bacterial infections, viral infections, hormonal imbalances, endocrine dysfunction, genetic and congenital factors, immune and inflammatory responses, and lifestyle factors. Higher wind speeds can alter the diffusion range and concentration of airborne pollutants. Strong winds may carry environmental pathogens (such as bacteria and viruses), which, if they come into contact with the reproductive organs or enter the body through the respiratory tract, may indirectly cause infections (such as pelvic inflammatory disease or increased HPV transmission risk)^{14–16}. Strong winds are often accompanied by sharp temperature drops or dryness, which can weaken local immune barriers (such as the defenses of the reproductive tract mucosa), increasing the risk of infections or inflammation and thereby affecting the incidence of reproductive system malignant tumors. Regions with better public health resources tend to conduct more health education and outreach activities, encouraging residents to adopt healthy lifestyles and preventing the occurrence of reproductive system malignant tumors.

Urinary System Malignant Tumors. The hypoxic environment of high-altitude regions leads to increased blood viscosity and reduced renal blood flow, which may trigger a decline in glomerular filtration rate and hypoxic damage to renal tubules. This can potentially result in glomerulosclerosis, proteinuria, or acute kidney

injury. The intensity of ultraviolet radiation at high altitudes is 1.5 to 2 times greater than at lower elevations, and prolonged exposure may suppress both skin and systemic immune responses, leading to reduced immunity and possibly increasing the risk of urinary system malignant tumors. Precipitation can purify the air, helping to reduce inflammation and oxidative stress in the urinary system, thereby lowering the risk. Higher-income groups tend to consume low-fat, low-sugar, and low-salt diets and are less likely to have smoking or drinking habits, resulting in a lower risk of urinary system malignant tumors. On average, individuals with higher education levels are more aware of the risk factors for cardiovascular diseases and the corresponding preventive measures, enabling them to seek treatment interventions for urinary system malignant tumors when early symptoms appear.

Endocrine System Malignant Tumors. Regions at high altitudes generally have better environmental quality, and the low air pressure can protect the immune system, thereby reducing the risk of thyroid cancer. The impact of high temperatures on thyroid cancer has not yet been strongly supported by research evidence^{29–32}. Precipitation may contain radioactive substances such as radioactive iodine, which can promote the growth and reproduction of fungi, potentially increasing the risk of thyroid cancer. Areas with high levels of sunlight have stronger ultraviolet radiation, which can cause some damage to the thyroid. The detection rate of thyroid cancer is higher in areas with better access to health checks, and higher-income and higher-education residents are more likely to undergo routine screenings and detect early thyroid cancer.

Literature Review on Geographic Factors and Malignant Tumors. Previous studies have examined the impact of one or more geographic factors on the occurrence of specific types of malignant tumors. However, there is often considerable variation in the findings regarding the “degree” of impact of a specific geographic factor on a certain type of malignant tumor, with some studies showing opposite directions of influence. This inconsistency in conclusions is largely due to the complex and long-term nature of malignant tumor development, influenced by a range of factors, as well as the limitations of the study samples. Therefore, the macro-level results of this study, combined with existing research findings, provide valuable insights for public health decision-making in China and its regions.

Comparison of Goodness of Fit (R^2) in Table 4. By comparing the goodness of fit (R^2) of the six regression equations in Table 4, the degree to which geographic factors explain the incidence of the six types of malignant tumors is ranked as follows, from highest to lowest: urinary system malignant tumors, respiratory system malignant tumors, endocrine system malignant tumors, breast system malignant tumors, reproductive system malignant tumors, and digestive system malignant tumors. This ranking is only slightly different from the average detection power values obtained from the geographic detector in Table 3.

To account for gender and age differences, the variable “y_adjusted for gender and age factors k_i ” in Eq. (1) was used as the dependent variable in Eq. (4) for regression analysis (Table 5). The results of Table 4 are similar to those in Table 3, with the main differences being that topographic variation was excluded in the regression for respiratory system malignant tumors, and in the regression for breast system malignant tumors, average temperature in January and education level were excluded, while average wind speed was included²². The goodness of fit (R^2) rankings in Table 4 for the six types of malignant tumors are consistent with those in Table 3, with the only difference being that the goodness of fit for respiratory system malignant tumors is higher than that for breast system malignant tumors. In summary, after considering gender and age factors, the majority of the findings and conclusions in Table 3 are quite robust.

Geographic characteristics of malignant tumor treatment

Geographic distribution of hospitals and treatment numbers

Hospitals are the primary providers of medical services for malignant tumors, and their distribution is influenced by natural, economic, social, and political factors, showing significant regional variation. Figure 2a shows that hospitals providing malignant tumor treatment are concentrated in North and East China, followed by Central, South, and Northeast China, and then Southwest China, with Northwest China having the fewest hospitals. The geographical differences between the two sides of the “Hu Huanyong Line” are quite evident. The number of malignant tumor patients shows a distribution similar to that of the hospitals (Fig. 2b). The classification intervals in Fig. 2b were determined using the natural breaks (Jenks) method / quantile method, which groups similar values together while maximizing differences between classes. The Theil index was used to calculate the inequality in the number of hospitals and the number of patients treated in different regions, yielding values of 0.3582 and 1.0780, respectively. This indicates that regions with more concentrated malignant tumor medical resources tend to have higher hospital technical levels and stronger patient attraction, as well as larger average treatment scales^{19–22}.

Different types of malignant tumors have varying degrees of urgency in terms of treatment, difficulty, and patient mobility. Therefore, this study calculated the Pearson correlation between the treatment of six types of malignant tumors in different regions (Table 6). (1) Overall, the geographic distribution of treatment for the six types of malignant tumors is significantly positively correlated, with all pairwise correlation coefficients being significant at the 0.01 level. (2) The correlation coefficients between the number of hospitals and the number of treated patients for different malignant tumors show little variation. Among them, respiratory system malignant tumors, digestive system malignant tumors, and breast system malignant tumors have higher correlation coefficients with other types of malignant tumors than urinary system malignant tumors. (3) The correlation coefficients between different malignant tumors and the number of hospitals are noticeably smaller than those with the number of treated patients. This is because the correlation with the number of hospitals reflects the supply-side of medical services, while the correlation with the number of treated patients reflects the dual correlation of both the supply and demand sides of medical services.

Variable	Respiratory system malignant tumor	Digestive system malignant tumor	Breast system malignant tumor	Reproductive system malignant tumor	Urinary system malignant tumor	Endocrine system malignant tumor
Constant	0.150*** (0.003)	0.158*** (0.003)	0.107*** (0.002)	0.074*** (0.004)	0.100*** (0.004)	0.030*** (0.001)
Average elevation					0.020* (0.009)	−0.007*** (0.002)
Topographic variation						
Average temperature in January	−0.016*** (0.003)	−0.009* (0.004)				−0.005** (0.002)
Average temperature in July						
Annual precipitation	−0.026*** (0.004)				−0.024** (0.008)	0.012*** (0.003)
Average annual relative humidity		0.019*** (0.004)			0.005 (0.004)	−0.003 (0.002)
Annual sunshine hours		0.014*** (0.003)			0.011* (0.005)	0.006** (0.002)
Average annual wind speed			0.005*** (0.001)	0.005* (0.002)		
Income level	0.013*** (0.003)		0.004 (0.002)		−0.001 (0.004)	0.006*** (0.001)
Urbanization rate			0.016** (0.003)	−0.012 (0.007)	−0.008 (0.005)	
Education level	0.004 (0.002)			−0.003 (0.002)	−0.006** (0.002)	0.003*** (0.001)
Public health Resources		−0.006*** (0.001)	−0.008*** (0.001)	−0.008*** (0.002)	0.011* (0.005)	
Water pollution		0.003* (0.002)				
Air pollution			0.005** (0.002)		−0.004 (0.003)	
Goodness of fit (R^2)	0.36	0.14	0.36	0.19	0.55	0.41
Sample size	294	294	294	294	294	294

Table 5. Multivariable regression analysis on the proportion of malignant tumors occurrence in prefecture-level regions. Gender and age factors were removed, and stepwise regression analysis was performed. In each step, the least significant influencing factor was removed until all influencing factors had a p value less than 0.1. ***, **, and * represent the significance of the influencing factors at the 0.1%, 1%, and 5% levels, respectively.

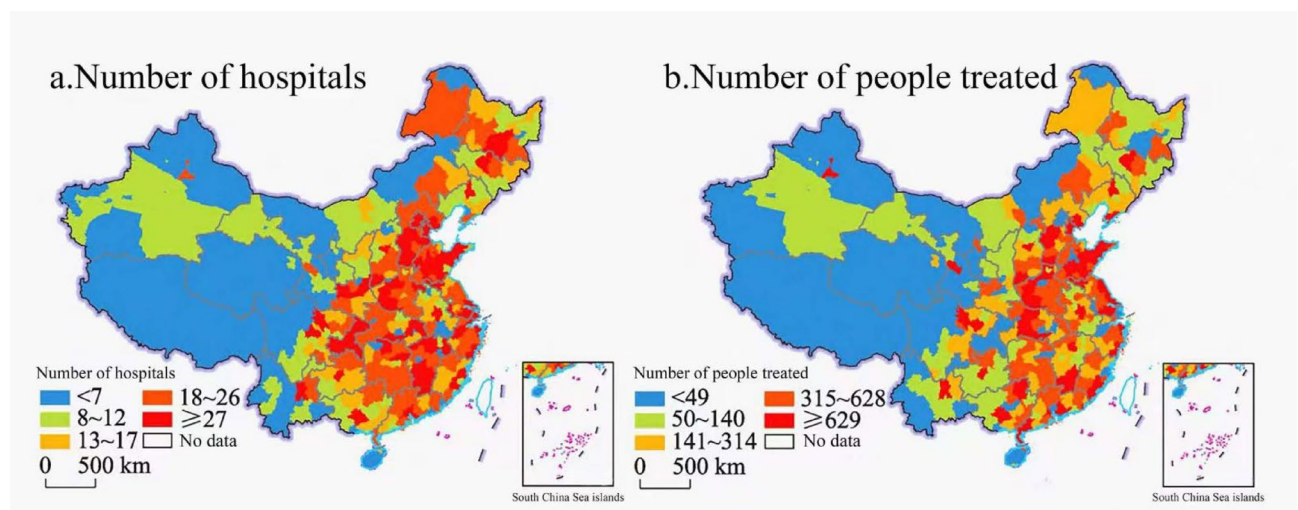


Fig. 2. Number of hospitals and patients of six types of malignant tumors in prefecture-level regions from October 2018 to March 2022. *Note:* The map was created based on the standard map provided by the Ministry of Natural Resources, with the approval number GS(2020)4630. The boundary of the base map has not been modified.

Malignant tumor	Respiratory system	Digestive system	Breast system	Reproductive system	Urinary system	Endocrine system
Respiratory system	1, 1					
Respiratory system	0.63, 0.96	1, 1				
Respiratory system	0.65, 0.96	0.62, 0.96	1, 1			
Respiratory system	0.55, 0.94	0.57, 0.92	0.59, 0.94	1, 1		
Respiratory system	0.44, 0.77	0.38, 0.84	0.35, 0.90	0.38, 0.82	1, 1	
Respiratory system	0.45, 0.95	0.47, 0.92	0.42, 0.95	0.40, 0.93	0.27, 0.79	1, 1

Table 6. Correlation of hospital numbers and patient numbers among prefecture-level regions. The numbers before the comma represent the correlation coefficient for the number of hospitals, while the numbers after the comma represent the correlation coefficient for the number of treated patients.

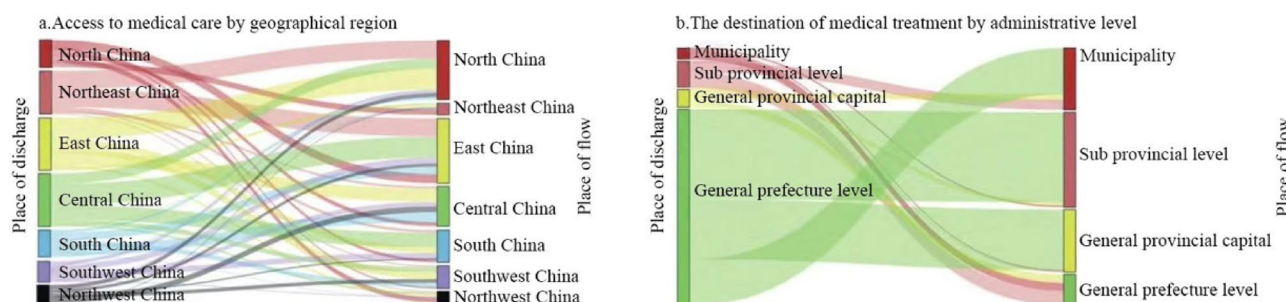


Fig. 3. Geographical distribution of out-of-region medical treatment destinations for malignant tumors patients. *Note:* North China includes Beijing, Tianjin, Shanxi, Hebei, and central Inner Mongolia; Northeast China includes Liaoning, Jilin, Heilongjiang, and the eastern part of Inner Mongolia (Hulunbuir, Xing'an, Tongliao, Chifeng); East China includes Shanghai, Jiangsu, Zhejiang, Anhui, Shandong, and Fujian; Central China includes Henan, Hubei, Hunan, and Jiangxi; South China includes Guangdong, Guangxi, and Hainan; Southwest China includes Chongqing, Sichuan, Guizhou, Yunnan, and Tibet; Northwest China includes Shaanxi, Gansu, Qinghai, Ningxia, Xinjiang, and Alashan in Inner Mongolia.

Out-of-town medical treatment destinations

If the distribution of malignant tumor incidence does not match well with the spatial distribution of quality medical resources, more patients may need to seek medical treatment in other regions, which increases both the economic and psychological burden on patients and may affect the quality of medical care. To optimize the spatial allocation of medical resources, it is necessary to analyze the medical destinations chosen by malignant tumor patients in different regions.

Figure 3a, which divides patients by seven geographic regions, shows the following: (1) North and East China are net inflow areas for out-of-town medical treatment, both of which are regions rich in high-quality medical resources. (2) The out-of-town treatment destinations for patients in North China are more dispersed, while patients from Northeast China mostly go to North China. Patients from East China tend to go to both North and Central China, while patients from Central China more often go to Central and South China. Patients from South China typically go to East and Central China for treatment, and Southwest China patients often seek treatment in South and East China. Finally, patients from Northwest China mostly go to Central and North China. The study shows that when patients choose their medical destinations, they primarily consider the geographic distance and the level of medical services.

Figure 3b divides regions into four administrative levels. By comparing the left and right sides of the chart, it can be observed that municipalities directly under the central government (4), sub-provincial cities (15), and regular provincial capitals (23) are net inflow areas for out-of-town medical treatment, while regular prefecture-level cities are net outflow areas. The allocation of high-quality medical resources is significantly influenced by the administrative level of the city. It is worth noting that patients from sub-provincial cities tend not to go to regular provincial capitals for treatment, and patients from provincial capitals also seldom go to sub-provincial cities. This may be because the geographic distance between sub-provincial cities and provincial capitals is generally far, and their overall development level and medical resource availability are similar. For these regions, patients' out-of-town medical destinations are more often municipalities directly under the central government (in search of better medical services) or regular prefecture-level cities (possibly for family companionship).

Geographic imbalance in the treatment-to-incidence ratio of malignant tumors

To more accurately reflect the extent to which a region's medical services for malignant tumors can meet its own needs, this study defines a region's treatment-to-incidence ratio for malignant tumors as the number of treated cases divided by the number of diagnosed cases in that region. A ratio greater than 1 indicates a surplus of

medical resources for treating malignant tumors—suggesting that the region is also serving patients from other areas. A ratio less than 1 indicates a shortage, meaning that the region cannot fully meet its own demand for treatment. Here, although “medical resources” refers to quantitative indicators such as the number of doctors, hospital beds, and equipment, it is assumed that these resources also meet the qualitative standards required for treating malignant tumors.

This study uses the *Theil* index to examine geographic imbalances in the treatment-to-incidence ratio across three common regional classification schemes: (1) The seven traditional geographic regions (North China, Northeast China, East China, Central China, South China, Southwest China, and Northwest China); (2) North vs. South, based on the Qinling Mountains–Huai River line (excluding the Tibet Autonomous Region); (3) Northwest vs. Southeast, based on the Hu Huanyong Line. (4) To analyze the evolution of geographic imbalances in the treatment-to-incidence ratio, a cohort analysis was conducted from both the incidence side and the treatment side: (4) On the incidence side, the patients’ birth cohorts were studied. (4) On the treatment side, the establishment period of hospitals was analyzed.

Analysis from the incidence side (patients)

Patients are the demand side of malignant tumor medical services. Differences in birth period, place of residence, and disease characteristics influence healthcare-seeking behaviors. Based on the principle of approximately equal case distribution—while considering that younger people have lower incidence and the very elderly are relatively fewer—the patients were divided into five birth cohorts: 1949–1959 (19.8%); 1960–1969 (29.5%); 1970–1979 (21.7%); 1980–1989 (19.6%); 1990–2021 (10.01%).

When calculating the *Theil* index of the treatment-to-incidence ratio by region and birth cohort, it was found that the index remained relatively stable across the first three cohorts. However, for those born between 1990 and 2021, the *Theil* index increased significantly—reaching 0.12 (Fig. 4). This suggests that patients born after 1990 are more likely to seek treatment in more developed regions, resulting in a greater geographic imbalance in treatment.

The decomposition of *Theil* index by the seven geographic regions (Fig. 4a) reveals that: Southwest, Northwest, and South China have consistently had larger within-region disparities, while East China and Central China have had relatively smaller disparities. These findings indicate that when allocating medical resources, underdeveloped areas in Southwest, Northwest, and South China should be prioritized. Over time, within-region disparities have increased across most regions. Notably, Northeast and Northwest China show the most rapid growth in internal disparities, implying that younger patients from these regions are more likely to seek treatment elsewhere. As today’s younger generations gradually become the primary group of malignant

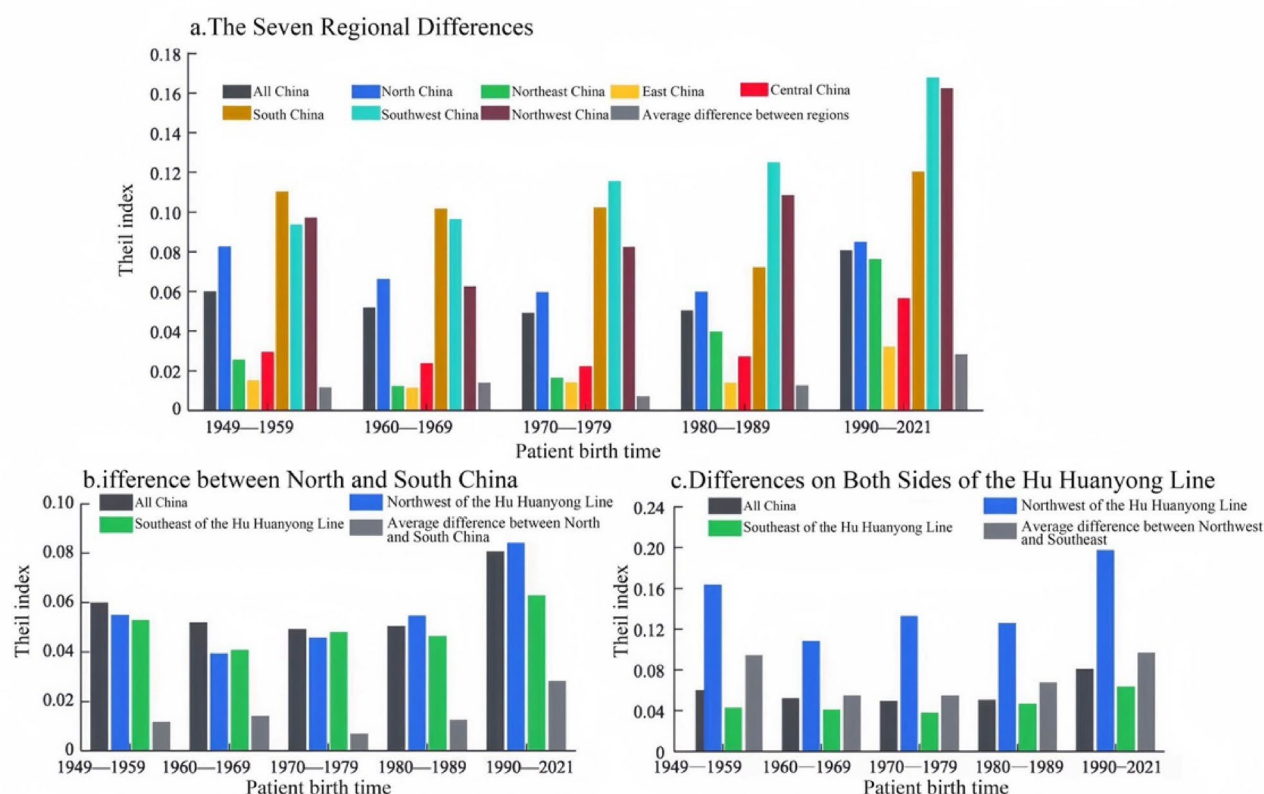


Fig. 4. Inequality in treatment-to-occurrence ratio of malignant tumors among prefecture-level regions grouped by patient birth periods.

tumor patients, the mismatch between the supply and demand for high-quality cancer care in northern regions is likely to become more pronounced—a trend that deserves close attention.

The analysis of the northern and southern regions (Fig. 4b) reveals that, overall, the internal disparity within the northern region is similar to that of the southern region, and both are larger than the average disparity between the north and south. Therefore, the gap in the supply and demand of malignant tumor medical services between the north and south is relatively small. The internal disparity in the northern region began to increase rapidly starting from the 1960–1969 birth cohort, while the increase in the southern region's internal disparity was relatively slower. The average disparity between the north and south started to rise more quickly from the 1970–1979 birth cohort. This indicates that the shortage of high-quality malignant tumor medical services in the northern region will become more pronounced.

The analysis of the "Huhuan Yong Line" (Fig. 4c) shows that, overall, the internal disparity within the northwest half is greater than that of the southeast half, indicating that certain areas in the northwest are more lacking in medical resources. Comparing this with Fig. 4b, it is evident that the differences between the two halves of the "Huhuan Yong Line" are significantly larger than the north–south disparity, with the former being approximately three times larger. In terms of different birth cohorts, the increase in internal disparity in the northwest half is more pronounced, while the rise in internal disparity in the southeast half and the average disparity between the northwest and southeast is relatively smaller. This suggests that, as the living and healthcare patterns of the current younger generation gradually become the norm, there is a more urgent need to increase the supply of malignant tumor medical services in the northwest region.

Malignant tumor treatment (hospital) analysis

Hospitals are the main providers of malignant tumor treatment services, and the hospitals where patients in the sample sought treatment were established between 1844 and 2021. Based on the principle of equally dividing the cases, four time points were chosen for the hospital establishment years, resulting in five groups: Group 1 consists of hospitals established before 1912; Group 2 consists of hospitals established before 1940; Group 3 consists of hospitals established before 1950; Group 4 consists of hospitals established before 1963; and Group 5 consists of hospitals established before 2021. The percentage of cases treated by hospitals in these five groups in relation to the total number of sample cases is 20.6%, 39.4%, 60.5%, 80.5%, and 100%, respectively. If hospitals were established in regions with high incidence rates of malignant tumors, the national inequality in malignant tumor treatment occurrence would decrease; conversely, inequality would increase.

For each group of hospitals, the residential location (i.e., the disease occurrence location) of the treated patients was identified, and the occurrence ratios of malignant tumor treatment in each region were calculated. The Theil index for the national malignant tumor treatment occurrence ratio was then computed. The data were further decomposed by geographical region (Fig. 5). The rapid decline in the Theil index indicates that the establishment of new hospitals has significantly alleviated the geographical imbalance in the supply and demand for malignant tumor medical services. Among the seven geographical regions, the largest internal disparity is in the northwest, followed by the northeast, while the smallest disparities are in Central China and East China. With the establishment of new hospitals, the relative disparities within the seven regions have been narrowing, with the internal disparity in the southwest region decreasing from relatively high to middle. The average disparity between the seven regions remains relatively stable; in the early twentieth century, it was much smaller than the internal disparities within the seven geographical regions, but by the 2020s, it surpassed the internal disparities in Northeast, East, and Central China. Since the early twentieth century, the geographical pattern of China's economic and social development has undergone several changes, and after 1949, the establishment of new hospitals considered the supply and demand disparities for malignant tumor treatment across regions.

The analysis of the north and south regions (Fig. 5b) shows that the internal disparity in the northern region has always been larger than in the southern region. After 1949, the difference between the internal disparities of the north and south has remained at a relatively low level. The average disparity between the north and south has always been smaller than the internal disparities within both the northern and southern regions. Therefore, the internal disparities within the north and south are the dominant factors contributing to the national imbalance.

The analysis of the differences between the two halves of the "Huhuan Yong Line" (Fig. 5c) reveals that the internal disparity in the northwest half has always been greater than that in the southeast half, and the gap between the two is narrowing rapidly. The average disparity between the northwest and southeast halves is decreasing slowly, but it has already exceeded the internal disparity in the southeast half and is very close to the internal disparity in the northwest. Therefore, the disparity in medical resources between the two halves of the "Huhuan Yong Line" is worthy of attention, and when establishing new hospitals, more consideration should be given to the northwest region. The average disparity between the two halves of the "Huhuan Yong Line" is greater than the average disparity between the north and south (Fig. 5b), further indicating that the differences between the two halves of the "Huhuan Yong Line" are larger than the geographical differences between the north and south.

Conclusion Findings

The occurrence and treatment of malignant tumors are major concerns for both the public and the government and are among the key priorities in building a "Healthy China." Both the incidence and treatment of malignant tumors exhibit significant spatial heterogeneity, underscoring the need for enhanced geographic research. This study collected and analyzed data on over 200,000 individuals from a representative malignant tumor mutual aid platform in China, leading to the following findings:

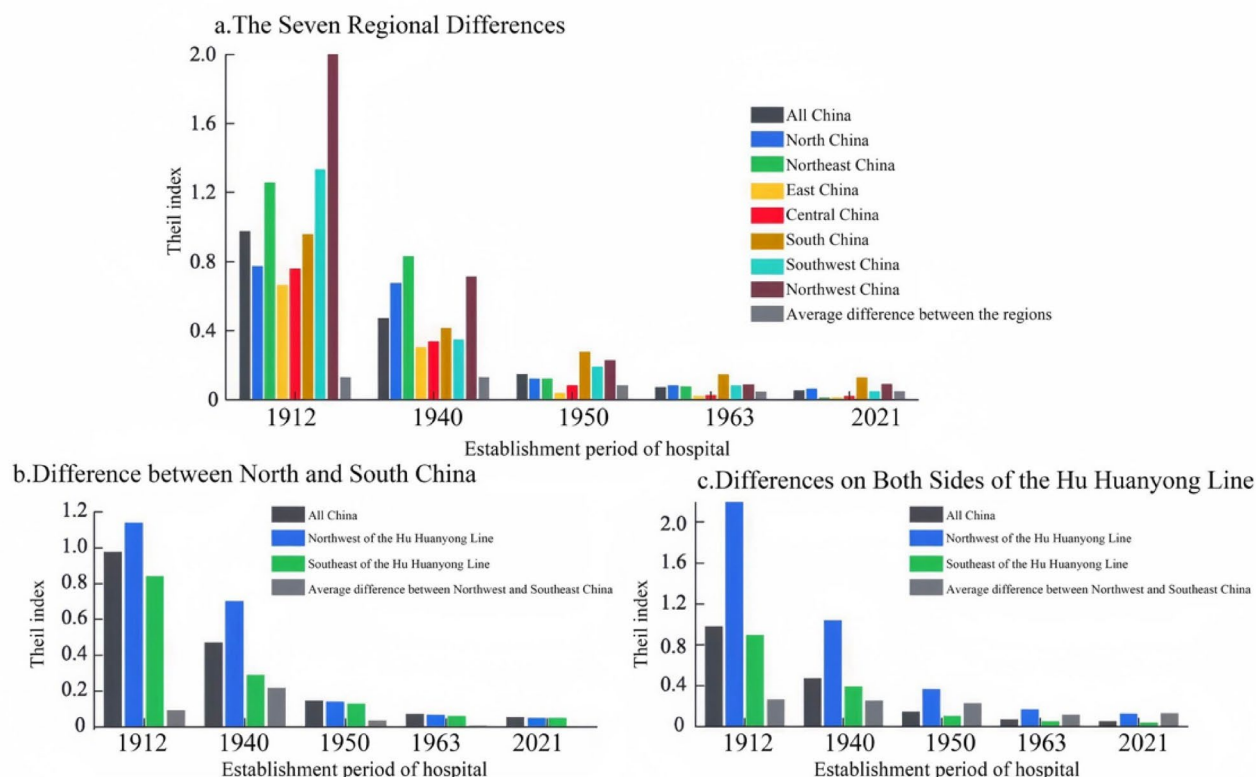


Fig. 5. Inequality in treatment-to-occurrence ratio of malignant tumors among prefecture-level regions grouped by hospital establishment periods.

1. For lung cancer, high incidence areas are not only correlated with environmental exposures such as air pollution, but also with disparities in access to early detection methods—most notably low-dose computed tomography (LDCT) for high-risk individuals. The uneven availability of such diagnostic tools across regions likely contributes to the geographic inequality observed.
2. In the case of breast and cervical cancers, regional differences may be partially explained by inconsistent implementation of screening programs. While some provinces have initiated organized screening for these cancers—such as the “two-cancer” screening program and HPV vaccination for adolescent girls—coverage remains uneven, particularly in rural and underdeveloped areas. These disparities in preventive infrastructure are reflected in the spatial clustering patterns we observe.
3. Colorectal cancer, which is rising in incidence nationally, currently lacks a standardized nationwide screening protocol. This gap further amplifies interregional inequality, especially in areas with lower public health investment.
4. For digestive system cancers such as gastric and liver cancers, environmental exposures (e.g., dietary patterns, *Helicobacter pylori* prevalence, hepatitis B infection) and socioeconomic conditions both play a role. However, the absence of systematic early detection programs remains a key barrier to improving outcomes in high-incidence regions.

International comparison with GLOBOCAN data

To place our findings in a global context, we compared the cancer incidence patterns observed in this study with international data from the IARC GLOBOCAN 2020 database (<https://gco.iarc.fr/today>). Similar to global trends, lung cancer (C34) remains the most common cancer in China, especially among males, reflecting the high prevalence of smoking and air pollution—consistent with international studies. Likewise, digestive system cancers (e.g., stomach, colorectal, and liver cancers) continue to account for a large proportion of cases, as observed in other East Asian countries such as Japan and South Korea.

However, notable differences exist. For example, while breast cancer (C50) is the most common cancer among women globally, in China, its spatial variability is more pronounced, and the overall incidence rate is lower than in high-income Western countries, possibly due to differing reproductive behaviors and screening coverage. Additionally, the incidence of thyroid cancer (C73) in China shows strong regional clustering, which may reflect overdiagnosis in well-resourced areas—a pattern also noted in countries like South Korea.

These similarities and differences indicate that while China's cancer burden reflects many global patterns, it also has distinct regional and systemic characteristics, underscoring the importance of localized prevention and resource allocation strategies.

Suggestions

Malignant tumor prevention and early management should be tailored to local conditions. This study proposes the following three policy recommendations:

1. **Localized Prevention Strategies:** Local governments should focus on malignant tumors that are highly correlated with regional geographic characteristics in their prevention efforts. This is particularly important for tumor types with large regional disparities, such as respiratory, urinary, and endocrine system malignancies.
2. **Targeted Screening Services:** Regions should establish targeted medical screening services for locally prevalent malignant tumors to enable timely intervention and prevent disease progression.
3. **Differentiated Health Insurance Design:** In terms of health insurance policies, except for the region-specific products available in the Guangdong-Hong Kong-Macao Greater Bay Area, there is little regional differentiation in the design of terms and premiums. Within the ethical framework of commercial insurance, insurance providers can leverage geographic factors to design differentiated premiums for various regions.

Recommendations

1. **Regional Prioritization in Resource Allocation.** Priority should be given to the northwest regions of China, especially those beyond the “Hu Huanyong Line,” followed by the northeastern regions. Within the broader western region, cities should be differentiated internally to reduce disparities. Increasing the number of hospitals or expanding the scale of existing ones are both effective strategies for addressing geographic gaps in cancer care access.
2. **Tumor-Specific Medical Resource and Screening Investment.** When expanding high-quality medical resources, targeted efforts should prioritize both clinical treatment capacity and screening coverage for cancer types with high incidence and proven screening effectiveness. These include: Specialized hospitals and low-dose CT screening access for lung cancer (pulmonology); Facilities offering mammographic screening and breast surgical care for breast cancer; Enhanced HPV vaccination and cervical screening infrastructure for cervical cancer, and Expanded capacity for colorectal cancer screening and care (gastroenterology). These investments are especially urgent in provinces with high incidence and low early detection rates, or with large young or mobile populations.
3. **Explore Regional Health Insurance Responsiveness to Spatial Disparities.** Although this study does not include direct economic variables, our findings reveal significant interregional disparities in access to cancer treatment. These disparities suggest a need to further explore whether health insurance mechanisms can be regionally adjusted to better support patients in underserved areas. For example, differentiated reimbursement rates, expanded cross-regional coverage, or targeted subsidy programs might help reduce the treatment burden for patients who must travel long distances for care. This topic warrants dedicated health policy research and is proposed here as a conceptual direction for future reforms rather than a data-driven conclusion of this study.

Limitations

1. **Histological information** such as tumor subtype was not available in the dataset, which limits more granular analyses (e.g., smoking-related differences in small-cell vs. non-small-cell lung cancer). Future studies could integrate pathology-level data for more precise modeling.
2. **Lack of Birthplace Data.** This study considers only the place of residence and treatment. Since many major illnesses are influenced by genetic factors, having access to information on patients’ birthplaces would allow for a decomposition method to differentiate between inborn and acquired factors, thereby enriching conclusions about the geographic influences on health.
3. **Sample Representativeness.** The data used in this study comes from a national malignant tumor mutual aid platform (similar to a disease insurance program). Participants in such platforms may differ from the general population in terms of economic status, risk preferences, and physical condition.
4. **Limitations in Available Data.** Due to data limitations, this study does not address the geographical disparities in medical costs, treatment outcomes, or mortality rates, all of which also exhibit significant spatial variation.
5. **This study does not include data on health insurance coverage, costs, or reimbursement rates, and therefore cannot evaluate the impact of insurance design on cancer incidence or access to care.**

Data availability

All the raw data are available upon request to the corresponding author.

Received: 1 April 2025; Accepted: 25 August 2025

Published online: 01 September 2025

References

1. Xiong, J. et al. Asymmetry between ecosystem health and ecological quality from an Earth observation perspective. *Sci. Rep.* **15**, 10143 (2025).
2. Zhang, W. et al. Negative binomial regression analysis of factors influencing the number of distinct mosquito species in Zhejiang Province, China, 2023. *Sci. Rep.* **15**, 10433 (2025).
3. Wang, S. et al. Mapping the landscape and roadmap of geospatial artificial intelligence (GeoAI) in quantitative human geography: An extensive systematic review. *Int. J. Appl. Earth Obs. Geoinf.* **128**, 103734 (2024).

4. Chaffin, B. C., Floyd, T. M. & Anzollitto, P. Environmental governance networks and geography: A research agenda at the confluence of critical concepts for navigating rapid environmental change. *Ann. Am. Assoc. Geogr.* **114**, 1718–1730 (2024).
5. Lun, Z. R. et al. Visceral leishmaniasis in China: an endemic disease under control. *Clin. Microbiol. Rev.* **28**, 987–1004 (2015).
6. Huang, J. et al. Identification of health risks of hand, foot and mouth disease in China using the geographical detector technique. *Int. J. Environ. Res. Public Health* **11**, 3407–3423 (2014).
7. Vázquez-González, J. et al. Effects of smoking marijuana on the respiratory system: a systematic review. *Subst. Abuse* **44**, 249–260 (2023).
8. Thomson, J. R. & Friendship, R. M. Digestive system. In *Diseases Swine* 234–263 (2019).
9. McKinney, S. M. et al. International evaluation of an AI system for breast cancer screening. *Nature* **577**, 89–94 (2020).
10. Kalantaridou, S. N. et al. Stress and the female reproductive system. *J. Reprod. Immunol.* **62**, 61–68 (2004).
11. Hess, R. A. et al. A role for oestrogens in the male reproductive system. *Nature* **390**, 509–512 (1997).
12. Drolet, R. Urinary system. In *Diseases Swine* 408–424 (2019).
13. Huang, H. et al. Micro/nano plastics in the urinary system: Pathways, mechanisms, and health risks. *Environ. Int.* **109**, 109109 (2024).
14. Chen, H. J. et al. Role of hydrogen sulfide in the endocrine system. *Front. Endocrinol.* **12**, 704620 (2021).
15. Xiao, Y. et al. Progress and application of intelligent nanomedicine in urinary system tumors. *J. Pharm. Anal.* **100**, 964 (2024).
16. Li, R. et al. The relationship between the network of non-coding RNAs-molecular targets and N6-methyladenosine modification in tumors of urinary system. *Cell Death Dis.* **15**, 275 (2024).
17. Orchard, I. & Lange, A. B. The neuroendocrine and endocrine systems in insect—Historical perspective and overview. *Mol. Cell. Endocrinol.* **580**, 112108 (2024).
18. Guarnotta, V. et al. Impact of chemical endocrine disruptors and hormone modulators on the endocrine system. *Int. J. Mol. Sci.* **23**, 5710 (2022).
19. Qi, X. et al. The impact of the gut microbiota on the reproductive and metabolic endocrine system. *Gut Microbes* **13**, 1894070 (2021).
20. Clarke, S. A., Abbata, A. & Dhillon, W. S. Impact of COVID-19 on the endocrine system: A mini-review. *Endocrinology* **163**, bqab203 (2022).
21. Pal, R. & Banerjee, M. COVID-19 and the endocrine system: exploring the unexplored. *J. Endocrinol. Investig.* **43**, 1027–1031 (2020).
22. Clifton, K. J. et al. Representing pedestrian activity in travel demand models: Framework and application. *J. Transp. Geogr.* **52**, 111–122 (2016).
23. Chen, C. L. & Hall, P. The impacts of high-speed trains on British economic geography: a study of the UK's InterCity 125/225 and its effects. *J. Transp. Geogr.* **19**, 689–704 (2011).
24. Pedigo, A. S. & Odoi, A. Investigation of disparities in geographic accessibility to emergency stroke and myocardial infarction care in East Tennessee using geographic information systems and network analysis. *Ann. Epidemiol.* **20**, 924–930 (2010).
25. Freed, G. L. et al. Relation of per capita income and gross domestic product to the supply and distribution of pediatricians in the United States. *J. Pediatr.* **144**, 723–728 (2004).
26. Scheffler, R. M. et al. Forecasting the global shortage of physicians: an economic-and needs-based approach. *Bull. World Health Organ.* **86**, 516–523B (2008).
27. Gong, S. et al. Evaluating healthcare resource inequality in Beijing, China based on an improved spatial accessibility measurement. *Trans. GIS* **25**, 1504–1521 (2021).
28. Lu, C., Zhang, Z. & Lan, X. Impact of China's referral reform on the equity and spatial accessibility of healthcare resources: A case study of Beijing. *Soc. Sci. Med.* **235**, 112386 (2019).
29. Qian, T. et al. Evaluating spatial accessibility to general hospitals with navigation and social media location data: A case study in Nanjing. *Int. J. Environ. Res. Public Health* **17**, 2752 (2020).
30. Chen, L. et al. Spatial accessibility evaluation and location optimization of primary healthcare in China: A case study of Shenzhen. *GeoHealth* **7**, e2022GH000753 (2023).
31. Zhu, L. et al. Assessing spatial accessibility to medical resources at the community level in Shenzhen, China. *Int. J. Environ. Res. Public Health* **16**, 242 (2019).
32. Ran, X. et al. Geographic, sex and socioeconomic disparities in esophageal cancer incidence in China: A population-based study. *Int. J. Cancer* **154**, 477–487 (2024).
33. Zheng, R. et al. Cancer incidence and mortality in China, 2016. *J. Natl. Cancer Cent.* **2**, 1–9 (2022).
34. Desai, R. J. & Franklin, J. M. Alternative approaches for confounding adjustment in observational studies using weighting based on the propensity score: A primer for practitioners. *bmj* 367 (2019).
35. Zhang, S. et al. Cancer incidence and mortality in China, 2015. *J. Natl. Cancer Cent.* **1**, 2–11 (2021).
36. Mainous, A. G. III. & Kohrs, F. P. A comparison of health status between rural and urban adults. *J. Commun. Health* **20**(5), 423–431 (1995).
37. Qi, X. et al. The impact of the gut microbiota on the reproductive and metabolic endocrine system. *Gut Microbes* **13**(1), 1894070 (2021).
38. Philbin, M. M. et al. Public libraries: A community-level resource to advance population health. *J. Commun. Health* **44**(1), 192–199 (2019).
39. Manisalidis, I. et al. Environmental and health impacts of air pollution: A review. *Front. Public Health* **8**, 14 (2020).
40. Dormann, C. F. et al. Methods to account for spatial autocorrelation in the analysis of species distributional data: A review. *Ecography* **30**, 609–628 (2007).
41. Ni, X. et al. Socioeconomic inequalities in cancer incidence and access to health services among children and adolescents in China: a cross-sectional study. *Lancet* **400**, 1020–1032 (2022).
42. Zeng, H. et al. Changing cancer survival in China during 2003–15: A pooled analysis of 17 population-based cancer registries. *Lancet Glob. Health* **6**, e555–e567 (2018).
43. Qi, J. et al. National and subnational trends in cancer burden in China, 2005–20: An analysis of national mortality surveillance data. *Lancet Public Health* **8**, e943–e955 (2023).
44. Li, M. et al. Mapping overdiagnosis of thyroid cancer in China. *Lancet Diabetes Endocrinol.* **9**, 330–332 (2021).
45. Wei, W. et al. Cancer registration in China and its role in cancer prevention and control. *Lancet Oncol.* **21**, e342–e349 (2020).
46. Wacker, M. & Holick, M. F. Sunlight and Vitamin D: A global perspective for health. *Dermato-Endocrinology* **5**(1), 51–108 (2013).
47. Hu, Y. et al. Vitamin D status for Chinese children and adolescents in CNNHS 2016–2017. *Nutrients* **14**(22), 4928 (2022).

Author contributions

Y.B. and X.Y.: Project administration, Writing—review & editing, Writing—original draft, Investigation. J.Z.: Validation, Conceptualization. Y.Z.: Visualization. M.Z., C.T. and P.W.: Funding acquisition, Project administration. X.Z., L.Y. Supervision, Writing—review & editing, Resources. All authors reviewed the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-17452-w>.

Correspondence and requests for materials should be addressed to Y.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025