BMJ Open

Geospatial analysis of hospitalization rate for rheumatoid arthritis patients in Hunan：A cross-sectional Chinese study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2023-075088
Article Type:	Original research
Date Submitted by the Author:	29-Apr-2023
Complete List of Authors:	Ge, Yan; The Second Xiangya Hospital of Central South University Wang, Shiwen; Central South University, Department of Epidemiology and Medical Statistics Shi, Qianshan; Information Statistics Center of Health Commission of Hunan Province Shi, Jingcheng; Central South University, Department of Epidemiology and Medical Statistics Tian, Jing; Central South University, Department of Rheumatology and Immunology
Keywords:	Epidemiology < TROPICAL MEDICINE, RHEUMATOLOGY, Hospitalization, Protocols & guidelines < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Health Services for the Aged

Geospatial analysis of hospitalization rate for rheumatoid arthritis patients in Hunan：A cross-sectional Chinese study

Yan Ge[1#], Shiwen Wang[2#], Qianshan Shi[3], Jingcheng Shi[2*], Jing Tian[1*]

[1]Department of Rheumatology and Immunology, The Second Xiangya Hospital, Central South University, Changsha, Hunan, China

[2]Department of Epidemiology and Medical Statistics, Xiangya School of Public Health, Central South University, Changsha, Hunan, China

[3]Information Statistics Center of Health Commission of Hunan Province, Changsha, Hunan, China

[#]These authors contributed equally to this work and share first authorship

*** Correspondence:**
Jingcheng Shi, Department of Epidemiology and Medical Statistics, Xiangya School of Public Health, Central South University, Changsha, Hunan, China. Email:shijch@csu.edu.cn; Jing Tian, Department of Rheumatology and Immunology, The Second Xiangya Hospital, Central South University, Changsha, Hunan, China. E-mail: tianjing001@csu.edu.cn

Keywords: spatial analysis, rheumatoid arthritis, hospitalization rate, average temperature, per capita GDP, medical resource

Strengths and limitations of this study:

1. We applied spatial analysis to explore the geographic variation of HR for RA patients at the district and county level, which may fill the gap in knowledge of HR for RA patients at local levels in Hunan, providing important information for health care planning. ation.

2. Our study indicated that geospatial analysis may be a promising in-deep approach in the study of RA, providing more information of genetic, socioeconomic and environmental inform.

3. It was a cross-section study from Hunan province used provincial administrative databases from secondary or tertiary hospitals as the source.

4. There are numerous determinants of the HR for RA patients, a number of risk factors such as air pollutants have not been investigated in the present study due to the lack of data, which may be considered as a limitation of the present study.

Abstract

Background: Scant data was known about spatial variability of hospitalization rate (HR) of Rheumatoid arthritis (RA) patients worldwide especially in China.

Methods: A cross-sectional study was conducted among RA patients admitted in Hunan Province. Global Moran's I and Local Indicators of Spatial Association (LISA) were performed to explore the geospatial pattern of HR for RA patients. Generalized estimating equation (GEE) analysis and geographically weighted regression (GWR) were used to identify the potential influencing factors of HR for RA patients.

Results: There was a total of 11,599 admissions and the average HR was 1.57 per 10,000 population in Hunan. LISA detected different clusters pattern of HR for RA patients. Age, ethnicity, average temperature, average temperature range, average rainfall, regions, gross domestic product per capital (PGDP), doctors and hospitals per 10,000 people were risk factors of HR. However, only average temperature, PGDP and hospitals per 10,000 people showed different regression coefficients on HR provincially. The increase of hospitals increased the probability of HR from the east to the west in Hunan with a positive coefficient, while temperature drops increased the risk of HR from the south to the north negatively, similarly, the growth of PGDP decreased the probability of HR from the southwest to the northeast.

Conclusion: A non-random spatial distribution of HR for RA patients was demonstrated in Hunan, and average temperature, PGDP and hospitals per 10,000 people showed different regression coefficients on HR provincially. Our study indicates spatial and geostatistics may be useful approaches for further study in RA patients.

1 Introduction

Rheumatoid arthritis (RA) is an autoimmune disease, characterized by abnormal synovial hyperplasia and progressive destruction of cartilage and bone, affecting approximately 0.5–1% of the adult population worldwide[1, 2]. Although the etiology and pathogenies of RA remains unclear, it is generally believed that multifactorial factors are involved combining genetic predispositions [3, 4] and a variety of environmental factors, such as smoking, infection, and microbiota, air pollution and climate factors[5-9].

Differences in the incidence of RA are known to exist, Alamanos et al.[10] showed that RA varies geographically in areas of the world, with southern European countries having lower median incidence rates than northern European countries and North America. Geographic variation in incident RA has also been observed at the regional level in USA[11, 12]. The findings suggest women living in higher latitudes may be at greater risk for RA. Further, RA risk may be greater for locations that occur earlier in residential histories. Similar geographic clusters were also observed in other autoimmune diseases such as systemic lupus erythematosus, dermatomyositis, polymyositis, and vasculitis[13-15]. Recently, Liu found a notable rural–urban variation in RA prevalence in Alberta, Cal Canada. However, little is known about what is behind these geographic differences. Spatial and geostatistics are emerging tools that combine statistical, geographical and epidemiological methodologies in describing the spatial characteristics of both infectious and non-infectious diseases, which may be a promising in-deep approach in the study of RA, providing more information of genetic, socioeconomic and environmental information while simultaneously taking into account their spatial diversity. For most of the previous work on the pathogenesis and prevalence of RA explores potential risk factors using classical statistical techniques.

Hospitalization rates (HR) is an important index which can reflect the local medical and health service capacity and to some extent may indicate the prevalence and onset of disease[16]. Until now, the literature lacks in the application of geospatial analyses in HR of RA patients and its risk factors, especially in China. Therefore, we conducted a geospatial study to explore the spatial distribution of HR for RA patients in Hunan province from a global to local scale, and to investigate the potential risk factors of geographic, socioeconomic, environmental and health-care while simultaneously considering their spatial diversity.

2 Material and methods

2.1 Study location

Hunan province is located in the southern-central China (24°38'N-30°08'N, 108°47'E-114°15'E), with a total area of 211,800 km^2(http://www.gov.cn/guoqing/2005-09/13/content_5043917.htm). The terrain is a horseshoe-shaped landform, surrounded by mountains on three sides and facing the north, and it belongs to a subtropical monsoon climate. Hunan Province has jurisdiction over 14 prefecture-level administrative regions, including 13 prefecture-level cities and 1 autonomous prefecture. There are 36 municipal districts, 19 county-level cities, 60 counties and 7 autonomous counties, totalling 122 county/district-level divisions. In 2019, Hunan province had a total of population of 73.1953 million people and a GDP of 3975.21 billion yuan (http://tjj.hunan.gov.cn/). Dongting Lake region is an area around Dongting Lake, which located in the south of the middle reaches of the Yangze River and the north of Hunan Province. In terms of administrative regions, it involves 21 counties and districts belonging to City Yueyang, Yiyang and Changde in Hunan Province, with a land area of about 31,700 square kilometers, accounting for 15% of the province's total area. Dongting Lake region is a famous "land of fish and rice", rich in natural resources, with a large population and relatively high economic level.

2.2 Data collection

The study collected patients aged 16 or older, who were admitted to secondary or tertiary hospitals with main diagnosis as RA from January 1,2019 to December 31,2019 through the *statistical information direct reporting system of Hunan Health Commission*. RA-related records were identified as those with the International Classification of Diseases (ICD), Tenth Revision (ICD-10) codes of M05.X or M06.X. The admission times of patients were collected as well as their individual demographic data, including age, sex, permanent address, native place, residential post-code, and vital status. In this study, we used the permanent address to define the patient’s location rather than his or her native place. The names, ID numbers and telephone numbers of patients were desensitized in the data collection phase. Data confidentiality principles were followed throughout the study. Patients or the public were not involved in the design, or conduct, or reporting, or dissemination plans of our research.

Meteorological data (temperature, air pressure, rainfall) were collected from *China Meteorological Data Network* (https://data.cma.cn/). Map information came from the Standard Map service system(http://bzdt.ch.mnr.gov.cn/). Spatiotemporal Kriging interpolation analysis was used to complement the daily meteorological data for each district/county of Hunan Province in 2019, then the average annual value was calculated as the total daily value divided by 365 days. The data of population, economy (Gross Domestic Product, GDP; GDP per capital, PGDP) and health resources (numbers of doctors, numbers of beds, numbers of hospital) of each district/county were collected from *Hunan Statistical Yearbook 2019* (http://tjj.hunan.gov.cn/). The HR for RA patients in Hunan and in each district/county were calculated per 10,000 population (the total RA admissions divided by the number of permanent residents), as well as GDP, PGDP, numbers of doctors, numbers of beds and numbers of hospitals. ArcGIS 10.8 were used to make a spatial distribution map of HR and other variables in Hunan Province at district/county level. The same color with different shades indicates the variables, the darker the color, the higher the variables in this area.

Data checking and data screening were also carried out: ①logical verification: removing cases with obvious logical errors and duplication such as individuals with 0 hospitalization cost. Patients with multiple admissions in the same day were counted as a single admission; ② missing data: cases with missing of vital information (such as present residence, household registration address or discharge diagnosis information) were deleted;③outliers: graphical representation, Q-test were used to discriminate univariate outliers, and multivariate outliers were found out by Mahalanobis distance. The discovered outlier data were verified to delete or convert.

2.3 Statistical methods

In this study, quantitative data were presented as the median and the 25th–75th percentile interquartile range (IQR). Qualitative data like demographic data, meteorological conditions, socioeconomic factors, medical resources were described as frequency (percentage)percentiles. Chi-square method was used for intergroup comparison of categorical variables, and the statistical significance level was corrected by Bonferroni correction. Global spatial autocorrelation analysis and Local spatial autocorrelation analysis were used to determine whether the HR of RA was clustered in the province and analysis the index of local spatial autocorrelation. Generalized estimating equation(GEE) in two level was used (district/county level and person level) to identify the influencing factors of HR for RA patients, odds ratios (ORs) and 95% confidence intervals (Cl) were calculated, and variables with a *p*-value of <0.05 were considered as candidates for the Geographically Weighted Regression (GWR) model, which was used to reveal the spatially varying relationships between the RA admission rates and the influencing factors, and a set of location specific parameter estimates. Data were analyzed using the SPSS statistical software package (version 26.0, IBM). District and county served as the basic unit for spatial analysis (n=122).

2.3.1 Global spatial autocorrelation analysis

The global spatial autocorrelation analysis started from the macro-level of the whole province, compared the mean value of the attribute in the general area and the attribute value on each spatial unit to obtain the average degree of association between the HR of RA in various districts /counties across the province. That was, to determine whether the HR of RA patients was clustered within the province. Global Moran's I is a commonly used global correlation index ranging from -1 to 1[17]. $I>0$ indicates a positive spatial correlation at the given significance level. The larger the value, the more pronounced the spatial correlation, high observations accumulate in space with high observations, and low observations accumulate spatially. $I<0$ indicates a negative spatial correlation. The smaller the value, the more significant the spatial difference, the tendency of high observations to cluster with low observations. If $I=0$, the observations are randomly distributed in space[18].

2.3.2 Local spatial autocorrelation analysis

Local indicator of spatial association (LISA) is the analysis index of local spatial autocorrelation analysis[19]. In this study, it was used to measure the degree of spatial dependence between the RA admission rates of a district /country and the RA admission rates of its neighboring districts /counties and identify its spatial cluster pattern in a local space. There are four spatial correlation patterns: High-High cluster, Low-Low cluster, High-Low cluster, and Low-High cluster. High-High cluster and Low-Low cluster indicate that the observations have a strong positive spatial correlation. The two patterns respectively indicate that the high-rate areas are adjacent to the high-rate areas, and the low-rate areas are adjacent to the low-rate areas. The High-Low cluster and Low-High cluster signify that the observations have a strong negative spatial correlation. The two patterns mean that high-rates areas are adjacent to low-rates areas. The results of local spatial autocorrelation and global spatial autocorrelation analysis were displayed on the map with ArcMap (version 10.8, ESRI Inc., Redlands, CA, USA).

2.3.3 Two-level GEE

The GEE is used to estimate the parameters of the generalized linear model, which was first proposed by Liang and zeger in 1986[20]. It is a regression model for analyzing correlation data by using quasi likelihood estimation method to estimate parameters in the generalized linear model and repeated measurement data. In addition to the normal distribution, GEE uses the connection function to fit the stress variables of binomial distribution, poisson distribution, gamma distribution and other distributions into corresponding statistical models, which solves the problem of non-independence of repeated measurement data and can obtain robust parameters.

In this study, we used GEE in two level (district/county level and person level) to study the influencing factors of HR for RA patients. The dependent variable was HR of the RA patients, and according to literatures review, independent variables at district/county level included:① environmental factors: the average temperature, the average temperature range, the average air pressure, the average rainfall, whether Dongting Lake area or not and seasons;②socioeconomic indicators: PGDP, GDP, population; ③health-care resources: number of doctors, number of beds and hospitals per 10,000 people. Independent variables at person level included: gender, age and ethnic group. The correlation among the variables were tested. A pair of variables with a correlation coefficient greater than 0.7 were considered to be strongly correlated, and a better variable will be selected to be included in the model according to the clinical significance and the goodness of fit of the model. Finally, gender, ethnicity, age, average temperature, average temperature range, average rainfall, whether Dongting Lake area or not, PGDP, number of hospitals, number of doctors were chosen to the final regression model as our preliminary analyses found that there was a high correlation between average temperature with average air pressure (r_s=-0.8674), number of beds and number of doctors (r_s=0.940), as well as GDP and PGDP (r_s=0.787).

2.3.4 GWR

GWR considers spatial dependency by performing several separated regressions, identifying significant geographic clusters within the studied area [32]. Additionally, GWR produces an estimate for the association between HR of RA patients and its predictor variables, from analysis of the local spatial variability for each district and county. It is assumed that the regression coefficient is a function of regional geographical location and changes with geographical location. In this study, meaningful independent variables from GEE results were selected into the GWR model. Akaike information criterion (AIC) and R^2 was adopted to build the best fit model for analysis, while the smaller AIC and larger R^2 are, the better model is. Based on Fotheringham et al.[21], the GWR model with one independent variable can be expressed as:

$$\boldsymbol{y_i = \beta_0(u_i, v_i) + \beta_1(u_i, v_i)x_i + \varepsilon_i}$$

Where u_i and v_i are the coordinates of the i th location, and β_0 (u_i, v_i) is the local intercept for i th location, β_1 (u_i, v_i)is the estimated local regression coefficient for the i th location and ε_i is the random error at the i th location. The differential spatial impact of each variable on HR of RA were displayed on the map with ArcMap (version 10.8, ESRI Inc., Redlands, CA, USA).

2.4 Patient and Public Involvement

Patients or the public WERE NOT involved in the design, or conduct, or reporting, or dissemination plans of our research.

Results

3.1 Characteristics of hospitalizations of RA patients in Hunan

Basic information and admissions of RA patients are shown in Table 1. In 2019, the total number of hospitalizations of RA patients in Hunan Province was 11,599, including 8223(70.9%) females and 3376 (29.1%) males (2.441:1). The average age of admitted RA patients was 63(54,70) years. 59.3% (6877) patients aged 60 years and above, 33.9% (4694) aged 19-59 years, followed by 0.2% (28) aged 16-18 years. The total HR was 1.57 per 10,000 people, and significant differences were found between gender, age groups, ethnic groups, regions (whether Dongting Lake area or not) and seasons ($p<0.001$). The HR for RA patients was much higher in female, aged people (⩾60 years) as well as in Han race and in region around Dongting Lake($p<0.01$). For the admission time, there was a relative lower HR for RA patients in spring compared to summer, autumn and winter($p<0.01$).

3.2 Distribution of HR for RA patients and demographic data, meteorological conditions, socioeconomic factors, medical resources among 122 counties and districts in Hunan

Supplement Table 1 list the basic information of demographic data, meteorological conditions, socioeconomic factors, medical resources and HR for RA patients at district/county level(n=122). Higher HR of RA patients could be seen in the central part, the northeast region and the Dongting Lake basin in Hunan Province (Figure 1). In contrast, the fewest HR was seen in regions mainly located in the mountain areas. There was only one RA patient admitted in Wulingyuan District, Zhangjiajie, with the minimum HR of 0.01 per 10,000 people, while in Xiangxiang, a city near the provincial capital, Changsha, the HR was 26.57 per 10,000 people. A significant variation in the distribution of HR for RA patients was noted in different districts and counties as well as the distribution of other climatic factors, socioeconomic indicator and medical resources (Figure 1).

3.3 Existence of Global and local spatial autocorrelation of HR for RA patients among 122 districts and counties in Hunan, China

In global spatial autocorrelation analysis, there was a globally spatiality of HR for RA patients in Hunan Province (Moran's I=0.2585, Z=4.0318, p<0.000). The LISA analysis detected two High-High (HH) pattern clusters (Ningxiang county, Shuangfeng county) in the central regions of Hunan, in which high RA admission rate close to neighbors also showing high RA admission rate. It also observed eleven Low-Low (LL) pattern clusters in the north-western and southern mountain areas in Hunan indicating districts and countries with low RA admission rate with neighbors with low RA admission rate (Figure 2, Supplement Table 2). It is worth noting that six LL clusters (four districts and two counties) belong to Hengyang city. We also found one High-Low pattern cluster and five Low-High pattern clusters in Hunan. The global and local spatial autocorrelation coefficients were considered significant when $p < 0.05$.

3.4 Factors including demographic indictors, meteorological conditions, socioeconomic conditions and medical resources were associated with HR of RA patients in Hunan by Two-level GEE analysis

In the two-level GEE analysis, it was found age, ethnicity, average temperature, average temperature range, average rainfall, whether Dongting Lake area or not, PGDP, the number of doctors and hospitals per 10,000 people were risk factors of HR for RA patients, except for gender (*p*=0.213, Table 2). Among them, average temperature range, number of hospitals per 10,000 have a slightly positive effect on the HR for RA patients (OR,1.459;95%CI,1.246-1.641; *p*<0.001; OR,1.167;95%CI,1.157-1.181; *p*<0.001;). Oppositely, the risk of HR was negatively correlated with temperature, rainfall, PDGP, number of doctors per 10,000 people (OR,0.876;95%CI,0.785-0.983;*p*<0.001; OR,0.997;95%CI,0.996-0.999;*p*<0.001;OR,0.953;95%CI,0.936-0.983;*p*<0.001; OR,0.965;95%CI,0.948-0.971; *p*<0.001). In subgroup analysis, it demonstrated the risk of HR in Han group, the older (≥60 years), the region of Dongting Lake was much higher than those in minority ethic, the younger (<60 years), and non-Dongting Lake region (*p*<0.001).

3.5 The impact of average temperature, PGDP and number of hospitals on HR for RA patients among 122 districts and counties in Hunan by GWR analysis

To further identify risk factors, which had the greatest geospatial impact on the spatial distribution of HR for RA patients, we performed a multivariate spatial regression analysis using GWR based on the above GEE analysis.

GWR model showed that average temperature, PGDP, hospitals per 10,000 people were independent influencing factors for HR of RA patients with spatial trend (R^2,0.521; adjustR^2,0.493; AIC,124.543). The distinctive spatial impact of each significant variable on HR of RA patients are indicated in the Figure 3. The effect of average temperature and PGDP on HR for RA patients in Hunan province was negative. The more temperature drops, the lower of the PGDP, the higher HR for RA patients was. On the whole, the influence of temperature on HR decreased from the south of Hunan to the north regions (regression coefficient, −0.0548–−0.0546), the HR of RA patients in the south was more likely to be affected by average temperature. Similarly, the influence of PGDP decreased from the southwest to the northeast regions in Hunan, which means the HR of patients was more likely to be affected by PGDP in the southwest (regression coefficient, −0.0637–−0.0634). In contrast, the impact of number of hospitals per 10,000 people on HR for RA patients was positive, which increased from the east to the west regions, the more of number of hospitals, the higher HR of people was, especially in the west Hunan (regression coefficient, 0.0257–0.0258). In all, the GWR analysis suggests that HR of RA patients was indeed affected by local temperature, PGDP and number of hospitals, and varies with regions.

3 Discussion

RA is a chronic autoimmune disease and a leading cause of working disability, and this situation will become more serious with the increase of aging. Although the etiology is not entirely clear, RA is a multifactorial disease involving genetics, immunology and environment. The ultimate therapeutic goal for RA is to achieve disease remission or low disease activity long-term. Thus, the management of RA needs the collaboration not only of patients and health provider, but also the corresponding medical policy formulation, medical resource allocation and support from the government. It is known to all there are big gaps in economic conditions, medical resource allocation and medical service efficiency among different regions in China and even different cities in the same province, also the meteorological conditions differ from regions to regions. In this study, the spatial distribution of HR for RA patients by district/county was first evaluated and its association with demographic, meteorological conditions, socioeconomic indictors and medical resources taking spatial dependence into account or not by GWR analysis and GEE analysis were investigated in Hunan, a southern-central province in China with a median national PGDP level. To the best of our knowledge, this is the first spatial analysis study on the HR for RA patients in China.

In 2019, the annual HR for RA was 1.57 per 10,000 people in the whole province, but we found a notable variation from 0.01 case/10,000 population to 26.57 cases/10,000 population between geographic areas. The HR of RA patients was significantly higher in the central and northeast areas and Dongting Lake region in Hunan. We identified 2 High-High clusters located in the north-eastern and central region and 11 Low-Low clusters located in mountain areas. Furthermore, we investigated the possible risk factors associated with disparities of HR for RA patients in Hunan. It's interesting to notice that age, ethnicities, temperature, temperature range, rain fall, whether Dongting Lake area or not, PGDP, number of doctors and hospitals per 10,000 people were risk variables (p<0.001) in the GEE analysis. However, only temperature, PGDP and number of hospitals per 10,000 people were related with the risk of HR for RA patients in GWR when considering the spatial dependence. In total, patients in regions with lower temperature, lower PGDP or more hospitals per capita had a higher risk of HR. Notably, the influence of temperature, PGDP and per number of hospitals on HR of RA patients in Hunan were different from each other and different in regions. These spatial variations and risk factors should be considered when designing evidence-based interventions and planning programs to improve RA care in Hunan. Our findings highlight the need for regional approaches to the planning and delivery of RA care as well as the management of RA patients. Further studies should investigate the underline pathogenesis for the difference between the higher HR and lower HR of RA patients in Hunan especially the relationship with temperature, PDGP and numbers of hospitals per 10,000 people.

A total of 11,599 RA patients admitted in Hunan in 2019, with a female predominate and the mean age was 63(54,70) years-old, which was similar with the results in other reports in China[9]. In our study, the HR of RA was 1.57 per 10,000 population, which was much lower than the prevalence of RA (approximately 0.5–1% of the adult population). In the past years, a decrease in RA hospitalization in western countries have been reported. The RA hospitalization in WA, USA was decreased from 202.6 per 100,000 population in 1995 to 87.3 per 100,000 population in 2002[16]. In New Zealand, the HR of RA was decreased from 11.3 to 8 per 100,000 inhabitants between 1995 and 2002[22]. In Canada, the annual HR of RA was decreased from 15.4 to 12.4 per 100,000 population between 2000 and 2002[23]. Due to the lack of data about the HR of RA patients in Hunan, even in China, it is hard to know the HR was decreasing or increasing, but there may be a in a dynamic change in recent years. On one hand, the HR may increase because the overall improvement of numbers of rheumatologists, departments in hospitals, access to information and coverage of health insurance system in low-income arears for patients. On the other hand, as more and more patients receive early detection and better management, a decline in HR of RA patients would happen.

Spatial distributions of infectious diseases have been extensively studied [24-26], for the spread of the pathogenic microbial agent. However, it seems that non-communicable diseases in humans also show clustering in space. Such is the case of heart diseases, diabetes and autoimmune inflammatory rheumatic diseases [15, 27-29]. Matej Sapina et al. found out that incidences of childhood IgA vasculitis and IgA vasculitis-associated nephritis in Croatia are not randomly distributed but clustered in space [30]. A huge variation of HR for RA patients between 122 districts and counties in Hunan were found, ranging from 0.01-26.57 per 10,000 population. In addition, existence of spatial autocorrelation was observed in HR of RA in our study. H-H clusters of HR for RA patients were found in Ningxiang county and Shuangfeng county in the north-eastern and central part, showing higher HR positive autocorrelations around those areas. Ningxiang county located in the capital city, Changsha. 5 L-L clusters were found dispersed partially in the northwestern, southeastern and southern mountain areas, while other 6 L-L clusters, belonging to the Hengyang city, located in the south-central part of Hunan. It's surprising to see Shuangfeng county, a H-H cluster, is near those L-L clusters in Hengyang city. One of the possible reasons may be their differences in geographical situation. The terrain of Shuangfeng is closer to the northern and central plains, but separated from the latter by the Mountain Hengshan. However, as a consequence, questions arise and should be discussed: why do these hotspots and coldspots happen and what happen behind?

Previous studies showed incident RA varies geographically in region level in USA using GAMs and GLS as well as in worldwide [11, 12]. The findings suggest women living in higher latitudes may be at greater risk for RA. Further, RA risk may be greater for locations that occur earlier in residential histories. Recently, a significant rural-urban variation in RA prevalence in Alberta, Canada was reported in 2021[31]. Authors believed factors such as obesity, socioeconomic factors, genetic and environmental factors such as aboriginal and immigrant status, local meteorological conditions may partially explain these variations.

In consistent with HR, the demographic characteristics, meteorological conditions, socioeconomic indictors and health care resources also demonstrated geospatial discrepancies at the districts and counties level in our study. Besides, we found that age, ethnicities, temperature, temperature range, rain fall, whether Dongting Lake or not, PGDP, number of doctors and hospitals per 10,000 people were risk factors for HR in patients with RA in the GEE analysis.

However, if we consider the spatial variation, only temperature, PGDP and number of hospitals per 10,000 people affected the risk of HR for RA patients with spatial dependence using GWR model. There have been many researches about the affect of temperature on RA incidence, joint pain and medical seeking behaviors such as clinic visits or admissions for a longtime, however the results are conflicting, which may be mainly explained by different statistical analysis and adjustment for climate covariates. Recently, Huang et al [2]. reported exposing to cold temperature increased the risk of RA admissions in Anhui using spatial-time series analysis with a distributed lag non-linear model. Notably, we found the influence of temperature on HR was negative whether considering spatial dependence or not. And a significant trend was demonstrated that the effect of temperature on HR was increased from the north to the south. When temperature drops, the risk of HR for RA people in the southern part increased slightly compared to the northern part. It indicated that geospatial analysis may provide a new approach in the future study to solve the conflicting influence of climate factors in RA and other autoimmune and inflammation diseases. Low temperatures were also associated with increased risks of all types of inflammation-related diseases in Yanchi county, Ningxia, China (RR: 4.344, 95% CI: 2.887, 6.535) [32]. Abundant evidences indicated that extreme high or low temperatures could increase risks of mortality and morbidity for cardiovascular and respiratory diseases [33, 34]. Some researchers believed that the underline reasons may be that inflammatory factors such as interlukin-6, c-reactive protein play an important role in triggering the adverse effects by the low or extreme temperatures [35, 36].

It's well known that lower socioeconomic status increased the risk of RA prevalence, function declines and overall mortality rate over times [37, 38]. In this study, we also demonstrated that lower PGDP had a higher risk of HR for RA patients, and the effect of GDP showed spatial variation in different regions in Hunan. the HR of RA patients seemed more likely to be affected by the PGDP in the southwest, while much less in the region of northeast. It is suggested that more attention should be paid to the economic situation of RA patients in the southwest region and more should be done to improve their poverty. In addition, the inequalities in healthcare, low accessibility to healthcare centers and limited health coverage offered to the patients may also be responsible for HR of RA. In this study, number of hospitals per 10,000 people had a positive effect on the HR for RA patients. In addition, the impact of number of hospitals per 10,000 people on HR for RA patients was positive, which increased from the east to the west regions in Hunan. The more of number of hospitals, the higher HR of people was, especially in the west Hunan. This was consistent with the fact that lack of healthcare centers in west region of Hunan. Thus, to increase the assess to the healthcare centers and the disrupt the barrier to healthcare delivery should do great favour to RA patients in the west Hunan in future. Large-scale studies based on epidemiological and physiological data are necessary to further determine whether the differences in HR of RA and the related mechanisms behind the influence of temperature, economic level and medical resources in RA patients. In all, the GWR analysis suggests the inequalities in healthcare, major disparities in individual income and local temperature may be responsible for HR of RA patients in Hunan. But there were also other factors that affect the HR of RA.

This study has strengths and limitations. First, we applied spatial analysis to explore the geographic variation of HR for RA patients at the district and county level, which may fill the gap in knowledge of HR for RA patients at local levels in Hunan, providing important information for health care planning. Second, our study indicated that geospatial analysis may be a promising in-deep approach in the study of RA, providing more information of genetic, socioeconomic and environmental information. However, several limitations that must be taken into consideration when interpreting the results. First, it was a cross-section study from Hunan province used provincial administrative databases from secondary or tertiary hospitals as the source. All of these could lead to biases. Previous studies suggest that the RA case definition using health administrative databases has excellent sensitivity when multiple physician claims are included [15]. Due to lack of rheumatologists in primary hospital in Hunan, we believe that these estimates may be largely represent the true number of hospitalizations for RA patients in Hunan Province. However, these data have inherent limitations and should be interpreted with caution. Second, we used the annual average data of climate factors for analysis, ignore their short-term influence on the patients, which needed to be further studied in subsequent studies. Third, there are numerous determinants of the HR for RA patients; nevertheless, a number of risk factors such as air pollutants have not been investigated in the present study due to the lack of data, which may be considered as a limitation of the present study. Finally, it is not possible to accurately acquire individual exposures, which may lead to certain measurement bias.

4 Conclusion

A non-random spatial distribution of HR for RA patients was demonstrated in Hunan, in addition, average temperature, PGDP and hospitals per 10,000 people showed different regression coefficients across the regions on HR of RA patients. Our study indicates spatial and geostatistics may be useful approaches for further study in RA to provide more information of genetic, socioeconomic and environmental risk factors while taking into account their spatial diversity.

5 Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

6 Author Contributions

Yan Ge, Shiwen Wang designed, analyzed the results and wrote the manuscript. Qianshan Shi supervised the manuscript and data curation. Jingcheng Shi, Jing Tian designed, supervised and edited the manuscript.

7 Funding

This work was supported by National Natural Science Foundation of China (No.81701622), National Science& Technology Foundational Resource Investigation Program of China (No.2018FY100900), Natural Science Foundation of Hunan (No.2022JJ40674, No.2020JJ4772 and No.2021JJ30893), Natural Science Foundation of Changsha (No. kq2202409).

8 Acknowledgments

We thank the Information Statistics Center of Health Commission of Hunan Province for their support with data access and all agencies who provided the meteorological data.

9 Data sharing statements

The data that support the findings of this study are available from the author, Qianshan Shi(hnsqs@qq.com), upon reasonable request.

10 Reference

1. Scott DL, Wolfe F, Huizinga TW. Rheumatoid arthritis. Lancet. 2010;376(9746):1094-108.

2. Huang LJ, Zha JJ, Cao NW, Zhou HY, Chu XJ, Wang H, Li XB, Li BZ. Temperature might increase the hospital admission risk for rheumatoid arthritis patients in Anqing, China: a time-series study. Int J Biometeorol. 2022;66(1):201-11.

3. Karami J, Aslani S, Jamshidi A, Garshasbi M, Mahmoudi M. Genetic implications in the pathogenesis of rheumatoid arthritis; an updated review. Gene. 2019;702:8-16.

4. Constantin A, Navaux F, Lauwers-Cances V, Abbal M, van Meerwijk JP, Mazieres B, Cambon-Thomsen A, Cantagrel A. Interferon gamma gene polymorphism and susceptibility to, and severity of, rheumatoid arthritis. Lancet. 2001;358(9298):2051-2.

5. Liu X, Tedeschi SK, Barbhaiya M, Leatherwood CL, Speyer CB, Lu B, Costenbader KH, Karlson EW, Sparks JA. Impact and Timing of Smoking Cessation on Reducing Risk of Rheumatoid Arthritis Among Women in the Nurses' Health Studies. Arthritis Care Res (Hoboken). 2019;71(7):914-24.

6. Arleevskaya MI, Albina S, Larionova RV, Gabdoulkhakova AG, Lemerle J, Renaudineau Y. Prevalence and Incidence of Upper Respiratory Tract Infection Events Are Elevated Prior to the Development of Rheumatoid Arthritis in First-Degree Relatives. Front Immunol. 2018;9:2771.

7. Zhang X, Zhang D, Jia H, Feng Q, Wang D, Liang D, Wu X, Li J, Tang L, Li Y, Lan Z, Chen B, Li Y, Zhong H, Xie H, Jie Z, Chen W, Tang S, Xu X, Wang X, Cai X, Liu S, Xia Y, Li J, Qiao X, Al-Aama JY, Chen H, Wang L, Wu QJ, Zhang F, Zheng W, Li Y, Zhang M, Luo G, Xue W, Xiao L, Li J, Chen W, Xu X, Yin Y, Yang H, Wang J, Kristiansen K, Liu L, Li T, Huang Q, Li Y, Wang J.

The oral and gut microbiomes are perturbed in rheumatoid arthritis and partly normalized after treatment. Nat Med. 2015;21(8):895-905.

8. Wu Q, Xu Z, Dan YL, Cheng J, Zhao CN, Mao YM, Xiang K, Hu YQ, He YS, Pan HF. Association between traffic-related air pollution and hospital readmissions for rheumatoid arthritis in Hefei, China: A time-series study. Environ Pollut. 2021;268(Pt A):115628.

9. Wang J, Yu L, Deng J, Gao X, Chen Y, Shao M, Zhang T, Ni M, Pan F. Short-term effect of meteorological factors on the risk of rheumatoid arthritis hospital admissions: A distributed lag non-linear analysis in Hefei, China. Environ Res. 2022;207:112168.

10. Alamanos Y, Voulgari PV, Drosos AA. Incidence and prevalence of rheumatoid arthritis, based on the 1987 American College of Rheumatology criteria: a systematic review. Semin Arthritis Rheum. 2006;36(3):182-8.

11. Vieira VM, Hart JE, Webster TF, Weinberg J, Puett R, Laden F, Costenbader KH, Karlson EW. Association between residences in U.S. northern latitudes and rheumatoid arthritis: A spatial analysis of the Nurses' Health Study. Environ Health Perspect. 2010;118(7):957-61.

12. Costenbader KH, Chang SC, Laden F, Puett R, Karlson EW. Geographic variation in rheumatoid arthritis incidence among women in the United States. Arch Intern Med. 2008;168(15):1664-70.

13. Tian J, Zhang D, Yao X, Huang Y, Lu Q. Global epidemiology of systemic lupus erythematosus: a comprehensive systematic analysis and modelling study. Ann Rheum Dis. 2022.

14. Muro Y, Sugiura K, Hoshino K, Akiyama M, Tamakoshi K. Epidemiologic study of clinically amyopathic dermatomyositis and anti-melanoma differentiation-associated gene 5 antibodies in central Japan. Arthritis Res Ther. 2011;13(6):R214.

15. Liu X, Barber CEH, Katz S, Homik J, Bertazzon S, Patel AB, Robert J, Smith C, Mosher D, Marshall DA. Geographic Variation in the Prevalence of Rheumatoid Arthritis in Alberta, Canada. ACR Open Rheumatol. 2021;3(5):324-32.

16. Almutairi K, Nossent J, Preen DB, Keen H, Inderjeeth C. The temporal association between hospital admissions, biological therapy usage and direct health care costs in rheumatoid arthritis patients. Rheumatol Int. 2022;42(11):2027-37.

17. Pfeiffer DU, Robinson TP, Stevenson M, Stevens KB, Rogers DJ, Clements AC. Spatial analysis in epidemiology: OUP Oxford; 2008.

18. Shaweno D, Karmakar M, Alene KA, Ragonnet R, Clements AC, Trauer JM, Denholm JT, McBryde ES. Methods used in the spatial analysis of tuberculosis epidemiology: a systematic review. BMC Med. 2018;16(1):193.

19. Rezaeian M, Dunn G, St Leger S, Appleby L. Geographical epidemiology, spatial analysis and geographical information systems: a multidisciplinary glossary. J Epidemiol Community Health. 2007;61(2):98-102.

20. Zeger SL, Liang KY, Albert PS. Models for longitudinal data: a generalized estimating equation approach. Biometrics. 1988;44(4):1049-60.

21. Raza O, Mansournia MA, Rahimi Foroushani A, Holakouie-Naieni K. Geographically Weighted Regression Analysis: A Statistical Method to Account for Spatial Heterogeneity. Arch Iran Med. 2019;22(3):155-60.

22. Collings S, Highton J. Changing patterns of hospital admissions for patients with rheumatic diseases. N Z Med J. 2002;115(1150):131-2.

23. Rai SK, Avina-Zubieta JA, McCormick N, De Vera MA, Lacaille D, Sayre EC, Choi HK. Trends in Gout and Rheumatoid Arthritis Hospitalizations in Canada From 2000 to 2011. Arthritis Care Res (Hoboken). 2017;69(5):758-62.

24. Chowell G, Rothenberg R. Spatial infectious disease epidemiology: on the cusp. BMC Med. 2018;16(1):192.

25. Caprarelli G, Fletcher S. A brief review of spatial analysis concepts and tools used for mapping, containment and risk modelling of infectious diseases and other illnesses. Parasitology. 2014;141(5):581-601.

26. Zhang Y, Liu M, Wu SS, Jiang H, Zhang J, Wang S, Ma W, Li Q, Ma Y, Liu Y, Feng W, Amsalu E, Li X, Wang W, Li W, Guo X. Spatial distribution of tuberculosis and its association with meteorological factors in mainland China. BMC Infect Dis. 2019;19(1):379.

27. Casper M, Kramer MR, Quick H, Schieb LJ, Vaughan AS, Greer S. Changes in the Geographic Patterns of Heart Disease Mortality in the United States: 1973 to 2010. Circulation. 2016;133(12):1171-80.

28. Silman A, Harrison B, Barrett E, Symmons D. The existence of geographical clusters of cases of inflammatory polyarthritis in a primary care based register. Ann Rheum Dis. 2000;59(2):152-4.

29. de Carvalho Dutra A, Silva LL, Pedroso RB, Tchuisseu YP, da Silva MT, Bergamini M, Scheidt J, Iora PH, do Lago Franco R, Staton CA, Vissoci JRN, Nihei OK, de Andrade L. The Impact of Socioeconomic Factors, Coverage and Access to Health on Heart Ischemic Disease Mortality in a Brazilian Southern State: A Geospatial Analysis. Glob Heart. 2021;16(1):5.

30. Sapina M, Frkovic M, Sestan M, Srsen S, Ovuka A, Batnozic Varga M, Kramaric K, Brdaric D, Milas K, Gagro A, Jelusic M. Geospatial clustering of childhood IgA vasculitis and IgA vasculitis-associated nephritis. Ann Rheum Dis. 2021;80(5):610-6.

31. Liu X, Barber CEH, Katz S, Homik J, Bertazzon S, Patel AB, Robert J, Smith C, Mosher D, Marshall DA. Geographic Variation in the Prevalence of Rheumatoid Arthritis in Alberta, Canada. ACR Open Rheumatology. 2021;3(5):324-32.

32. Wang Q, Zhao Q, Wang G, Wang B, Zhang Y, Zhang J, Li N, Zhao Y, Qiao H, Li W, Liu X, Liu L, Wang F, Zhang Y, Guo Y. The association between ambient temperature and clinical visits for inflammation-related diseases in rural areas in China. Environ Pollut. 2020;261:114128.

33. Green H, Bailey J, Schwarz L, Vanos J, Ebi K, Benmarhnia T. Impact of heat on mortality and morbidity in low and middle income countries: A review of the epidemiological evidence and considerations for future research. Environ Res. 2019;171:80-91.

34. Zhao Q, Li S, Coelho M, Saldiva PHN, Hu K, Arblaster JM, Nicholls N, Huxley RR, Abramson MJ, Guo Y. Geographic, Demographic, and Temporal Variations in the Association between Heat Exposure and Hospitalization in Brazil: A Nationwide Study between 2000 and 2015. Environ Health Perspect. 2019;127(1):17001.

35. Hampel R, Breitner S, Ruckerl R, Frampton MW, Koenig W, Phipps RP, Wichmann HE, Peters A, Schneider A. Air temperature and inflammatory and coagulation responses in men with coronary or pulmonary disease during the winter season. Occup Environ Med. 2010;67(6):408-16.

36. Schneider A, Panagiotakos D, Picciotto S, Katsouyanni K, Lowel H, Jacquemin B, Lanki T, Stafoggia M, Bellander T, Koenig W, Peters A, Group AS. Air temperature and inflammatory responses in myocardial infarction survivors. Epidemiology. 2008;19(3):391-400.

37. Izadi Z, Li J, Evans M, Hammam N, Katz P, Ogdie A, Suter LG, Yazdany J, Schmajuk G. Socioeconomic Disparities in Functional Status in a National Sample of Patients With Rheumatoid Arthritis. JAMA Netw Open. 2021;4(8):e2119400.

38. Chen CH, Huang KY, Wang JY, Huang HB, Chou P, Lee CC. Combined effect of individual and neighbourhood socioeconomic status on mortality of rheumatoid arthritis patients under universal health care coverage system. Fam Pract. 2015;32(1):41-8.

Table 1. Characteristics of hospitalizations of RA patients in Hunan in 2019

	Hospitalizations (n, %)	**Hospitalization rate（per 10,000 people）**	***P***
Total	11599(100)	1.57	
Gender			<0.001
Male	3376(29.1)	0.91	
Female	8223(70.9)	2.32	
Age groups in years			<0.001
16-18	28(0.2)	0.024[ab]	
19-59	4694(40.5)	0.86[b]	
≥60	6877(59.3)	10.71	
Ethnic group			<0.001
Han	11211(96.7)	1.88	
ethnic minorities	388(3.3)	0.58	
Region			<0.001
Dongting Lake	2491(21.48)	2.06	
Non-Dongtin Lake	9108(78.52)	1.61	
Seasons			<0.001
Spring (March-May)	2167(18.7)	0.29	
Summer (June-August)	3186(27.5)	0.43[c]	
Autumn (September-November)	3132(27.0)	0.42[c]	
Winter (December-February)	3114(26.8)	0.42[c]	

a: Compared to the 19-59 years group，$P<0.01$；b: Compared to 60 years and above group，$P<0.01$；c: Compared to the Spring (March-May)，$P<0.01$.

574 **Table 2. The two-level GEE analysis for risk factor of HR for RA patients in Hunan in 2019**

Variables	***OR***	***P***	**95% CI**	
			Lower	**Upper**
(Intercept)	417.991	<0.001	0.000	0.000
Gender				
Female	1.036	0.213	0.848	1.151
male	1			
Age				
16-18	0.089	<0.001	0.024	0.139
19-59	0.837	<0.001	0.742	0.913
⩾60	1			
Ethnic group				
Han	4.482	<0.001	3.341	5.868
Ethnic minorities	1			
Temperature	0.876	<0.001	0.785	0.983
Temperature range (℃)	1.459	<0.001	1.246	1.641
Rain fall(mm)	0.997	<0.001	0.996	0.999
Region				
Non-Dongting	1			
Dongting Lake	1.340	<0.001	1.145	1.511
PGDP	0.953	<0.001	0.936	0.983
Doctors (per 10,000 people)	0.965	<0.001	0.948	0.971
Hospitals (per 10,000 people)	1.167	<0.001	1.157	1.181

575 PGDP: GDP per capita; OR: odds ratio; CI: confidence interval. Estimated odds ratios (ORs) were
576 compared with the reference (OR=1) in each group.

577

578

579

580

581

582

583

584

585

586

587

588

589

Figure 1. Distribution of HR and average demographic data, meteorological conditions, socioeconomic factors and medical resources in 122 districts and counties in Hunan in 2019.

A non-random distribution of HR for RA patients was demonstrated in 122 districts and counties in Hunan(A), and distribution of demographic data (population per10,000 people, sex portion (F/M); B-C), meteorological conditions (temperature, air pressure, temperature range, rain fall; D-G), socioeconomic factors (GDP, PGDP; H-I) and medical resources (doctors, beds, hospitals per 10,000 people; J-L) were different in Hunan.

Figure 2.　Spatial distribution clusters of HR for RA patients in 122 districts and counties in Hunan in 2019 by Local Indicators of Spatial Association (LISA) analysis.

LISA analysis indicates clusters according to high and low HR distribution patterns.

Figure 3. The effect of average temperate, PGDP and hospitals per 10,000 on HR of RA patients among 122 districts and counties in Hunan by GWR analysis.

Average temperate, PGDP and hospitals per 10,000 show distinctive spatial impact on HR of RA patients.

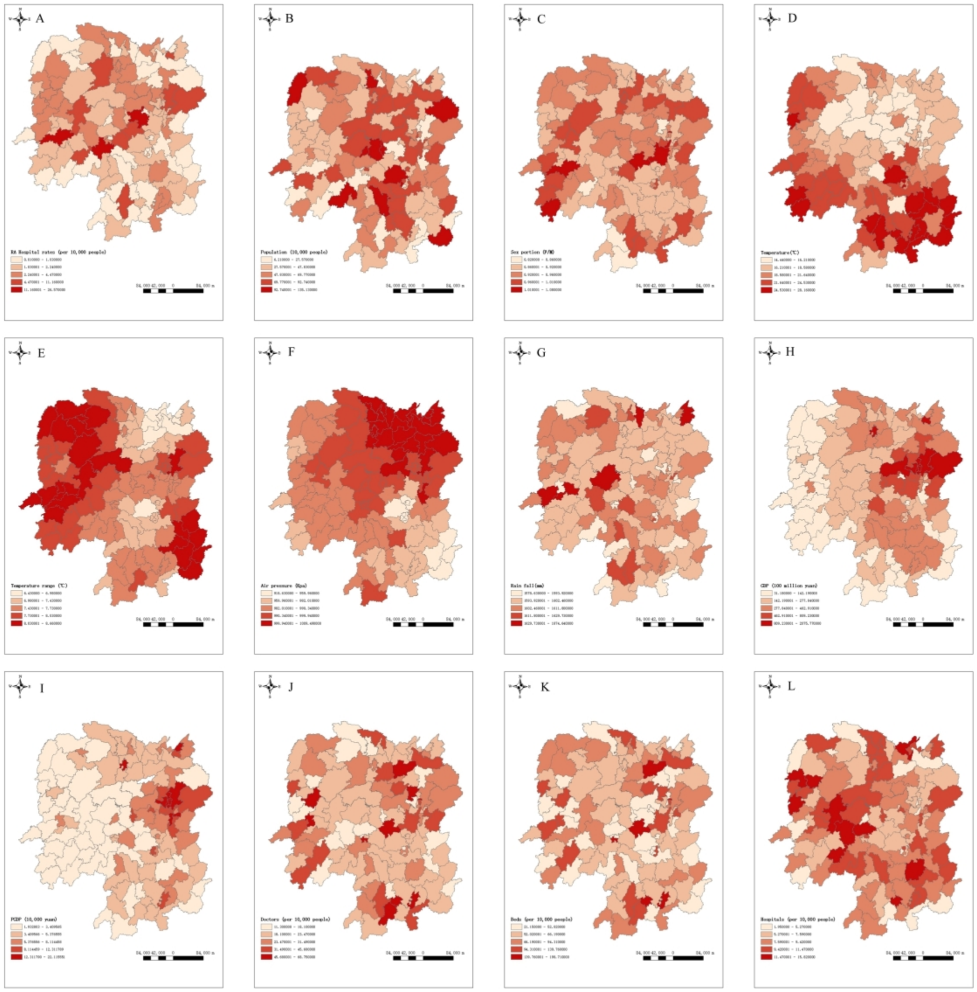

Figure 1. Distribution of the HR and average demographic data, meteorological conditions, socioeconomic factors and medical resources among 122 districts and counties in Hunan in 2019.

254x254mm (299 x 299 DPI)

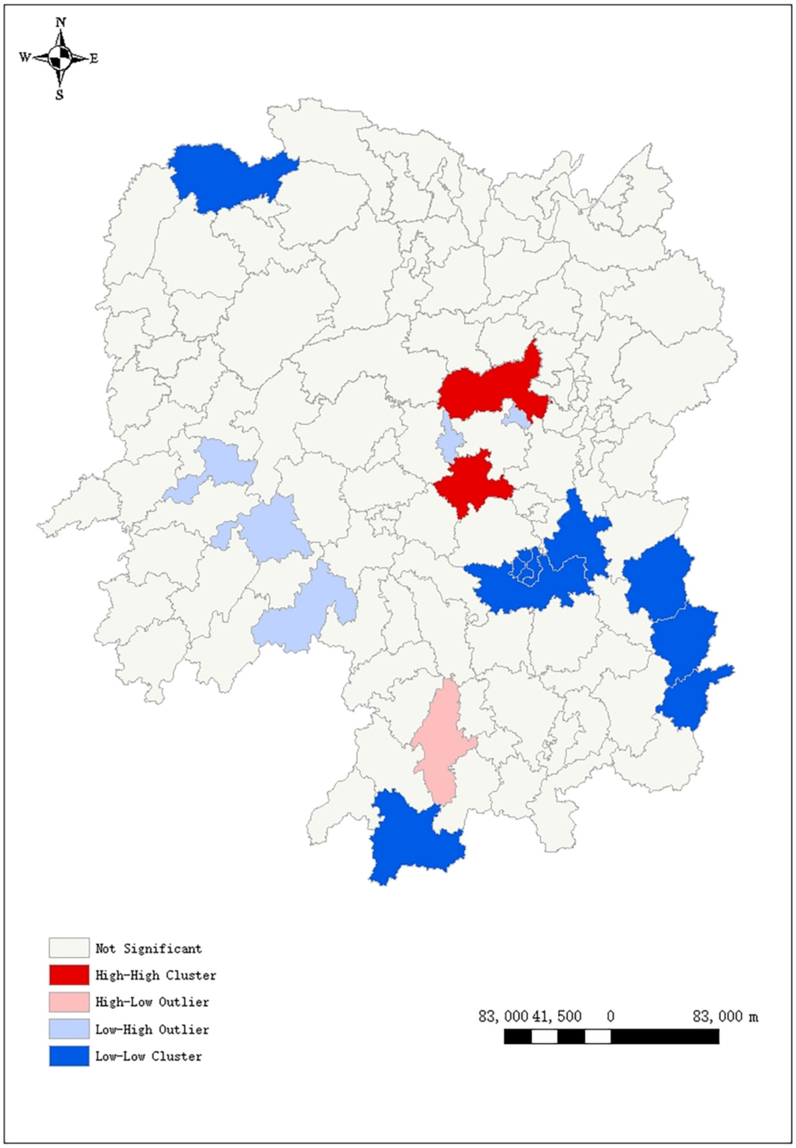

Figure 2. Spatial distribution clusters of the HR of RA patients among 122 districts and counties in Hunan in 2019 by local indicators of spatial association (LISA) analysis.

210x297mm (300 x 300 DPI)

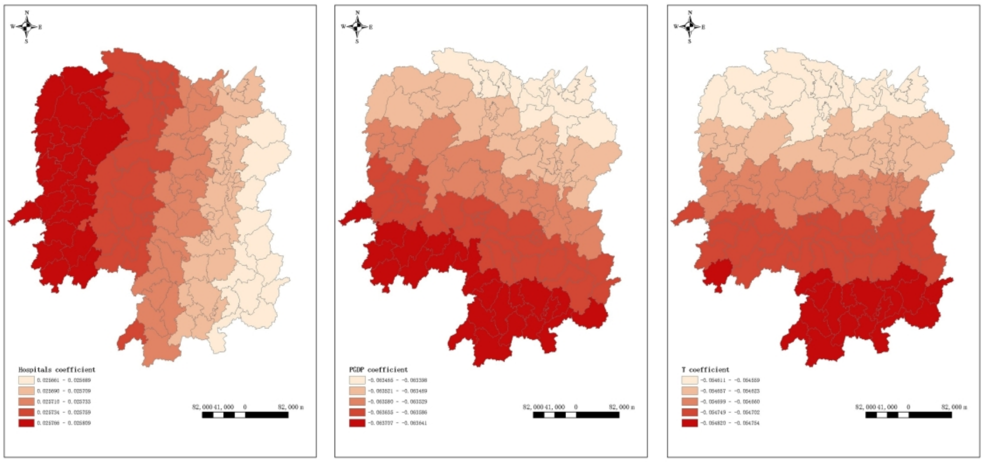

Figure 3. The effect of average temperate, PGDP and hospitals per 10,000 on HR of RA patients among 122 districts and counties in Hunan by GWR analysis.

252x119mm (300 x 300 DPI)

Supplemental Table 1. Characteristics of average demographic data, meteorological conditions, socioeconomic factors, medical resources and HR of RA patients among 122 districts and counties in Hunan in 2019

Variables	Percentile				
	X_{min}	P_{25}	P_{50}	P_{75}	X_{max}
Population (10,000 people)	6.21	34.60	57.06	81.07	135.13
Sex portion (F/M)	0.82	0.92	0.94	0.96	1.08
Age(years)	40	45	52	60	63
Temperature (℃)	14.44	16.49	19.17	23.20	28.16
Air pressure (Kpa)	918.63	982.28	990.19	997.53	1009.49
Temperature range (℃)	6.43	7.40	7.67	7.94	8.66
Rain fall (mm)	1578.63	1596.27	1598.92	1606.81	1874.64
GDP (100 million yuan)	31.18	117.59	234.72	364.87	2075.77
PGDP (10,000 yuan)	1.93	2.90	4.42	6.72	22.11
Doctors (per 10,000 people)	11.30	19.31	22.47	29.36	68.75
Beds (per 10,000 people)	21.15	50.75	58.69	78.39	198.71
Hospitals (per 10,000 people)	1.95	6.70	8.53	10.09	15.82
HR of RA patients (per 10,000 people)	.01	.88	1.65	2.75	26.57

F: female; M: male; GDP: Gross Domestic Product; PGDP: GDP per capita; HR: Hospitalization rates

Supplemental Table 2. Different cluster patterns of HR for RA patients in Hunan in 2019

District/county	Local Moran's *I*	*Z* Score	*P*	Patterns
Ningxiang county[a]	4.46×10^{-6}	2.16	0.04	HH
Shuangfeng county[b]	1.51×10^{-3}	2.85	0.02	HH
Chaling county[c]	1.86×10^{-5}	0.88	0.04	LL
Yanling county[c]	1.99×10^{-5}	0.87	0.04	LL
Zhuhui district[d]	8.80×10^{-5}	0.99	0.03	LL
Yanfeng district[d]	5.59×10^{-5}	1.17	0.00	LL
Shigu district[d]	1.04×10^{-4}	1.10	0.02	LL
Zhengxiang district[d]	1.07×10^{-4}	1.00	0.02	LL
Hengnan county[d]	5.51×10^{-5}	0.98	0.03	LL
Hengdong county[d]	3.21×10^{-5}	1.19	0.02	LL
Sangzhi county[e]	1.52×10^{-5}	0.85	0.00	LL
Guidong county[f]	1.04×10^{-5}	0.87	0.05	LL
Jianghua county (Yao Autonomous)[g]	2.98×10^{-6}	0.69	0.03	LL
Ningyuan county[g]	-1.53×10^{-4}	-1.15	0.04	HL
Shaoshan city[h]	-1.36×10^{-4}	-3.01	0.01	LH
Dongkou county[i]	-3.22×10^{-5}	-3.01	0.03	LH
Louxing district[b]	-6.75×10^{-5}	-3.55	0.02	LH
Xinning county[i]	-2.50×10^{-5}	-1.51	0.05	LH
Zhongfang county[j]	-3.36×10^{-5}	-2.17	0.05	LH

HR: Hospitalization rates; HH: high-high pattern; LL: low-low pattern; HL: high-low pattern; LH: low-high pattern; a: belongs to Changsha city; b: belongs to Loudi city; c: belongs to Zhuzhou city; d: belongs to Hengyang city; e: belongs to Zhangjiajie city; f: belongs to Chenzhou city; g: belongs to Yongzhou city; h: belongs to Xiangtan city; i: belongs to Shaoyang city; j: belongs to Huaihua city.

STROBE Statement—Checklist of items that should be included in reports of ***cross-sectional studies***

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	P1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	P1-2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	P2
Objectives	3	State specific objectives, including any prespecified hypotheses	P2
Methods			
Study design	4	Present key elements of study design early in the paper	P2
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	P2
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants	P3
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	P3
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	*P3*
Bias	9	Describe any efforts to address potential sources of bias	P3
Study size	10	Explain how the study size was arrived at	P2
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	P3
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	P3-5
		(*b*) Describe any methods used to examine subgroups and interactions	P3-5
		(*c*) Explain how missing data were addressed	P3-5
		(*d*) If applicable, describe analytical methods taking account of sampling strategy	P3-5
		(*e*) Describe any sensitivity analyses	P3-5
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	P5
		(b) Give reasons for non-participation at each stage	N/A
		(c) Consider use of a flow diagram	N/A
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	P5
		(b) Indicate number of participants with missing data for each variable of interest	P5
Outcome data	15*	Report numbers of outcome events or summary measures	P5-7
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	P5-7

		(*b*) Report category boundaries when continuous variables were categorized	P5-7
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	N/A
Discussion			
Key results	18	Summarise key results with reference to study objectives	P7-9
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	P9-10
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	P9-10
Generalisability	21	Discuss the generalisability (external validity) of the study results	P10
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	P10

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

Geospatial analysis of the hospitalization rate of rheumatoid arthritis patients in Hunan: A cross-sectional Chinese study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2023-075088.R1
Article Type:	Original research
Date Submitted by the Author:	22-Sep-2023
Complete List of Authors:	Ge, Yan; The Second Xiangya Hospital of Central South University, Department of Rheumatology and Immunology Wang, Shiwen; Central South University Xiangya School of Public Health, Department of Epidemiology and Medical Statistics Shi, Qianshan; Information Statistics Center of Health Commission of Hunan Province Shi, Jingcheng; Central South University Xiangya School of Public Health, Department of Epidemiology and Medical Statistics Tian, Jing; The Second Xiangya Hospital of Central South University, Department of Rheumatology and Immunology
Primary Subject Heading:	Rheumatology
Secondary Subject Heading:	Rheumatology, Public health
Keywords:	Epidemiology < TROPICAL MEDICINE, RHEUMATOLOGY, Hospitalization, Protocols & guidelines < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Health Services for the Aged

SCHOLARONE™
Manuscripts

Geospatial analysis of the hospitalization rate of rheumatoid arthritis patients in Hunan: A cross-sectional Chinese study

Yan Ge[1#], Shiwen Wang[2#], Qianshan Shi[3], Jingcheng Shi[2*], Jing Tian[1*]

[1]Department of Rheumatology and Immunology, The Second Xiangya Hospital, Central South University, Changsha, Hunan, China

[2]Department of Epidemiology and Medical Statistics, Xiangya School of Public Health, Central South University, Changsha, Hunan, China

[3]Information Statistics Center of Health Commission of Hunan Province, Changsha, Hunan, China

[#]These authors contributed equally to this work and share first authorship

*** Correspondence:**
Jingcheng Shi, Department of Epidemiology and Medical Statistics, Xiangya School of Public Health, Central South University, Changsha, Hunan, China. Email:shijch@csu.edu.cn; Jing Tian, Department of Rheumatology and Immunology, The Second Xiangya Hospital, Central South University, Changsha, Hunan, China. E-mail: tianjing001@csu.edu.cn

Keywords: spatial analysis, rheumatoid arthritis, hospitalization rate, average temperature, per capita GDP, medical resource

Abstract

Objective: Little is known about spatial variability of hospitalization rate (HR) of rheumatoid arthritis (RA) patients worldwide especially in China.

Methods: A cross-sectional study was conducted among RA patients admitted to hospitals in Hunan Province. Global Moran's I and local indicators of spatial association were used to explore the geospatial pattern of the HR of RA patients. Generalized estimating equation analysis and geographically weighted regression were used to identify the potential influencing factors of the HR of RA patients.

Results: There was a total of 11,599 admissions, and the average HR was 1.57 per 10,000 population in Hunan. We detected different cluster patterns of the HR among RA patients by local indicators of spatial association. Age, ethnicity, average temperature, average temperature range, average rainfall, regions, gross domestic product per capital, and doctors and hospitals per 10,000 people were risk factors for the HR. However, only average temperature, gross domestic product per capital and hospitals per 10,000 people showed different regression coefficients on the HR in different counties. The increase in hospitals increased the probability of HR from east to west in Hunan with a positive coefficient, while temperature decreases increased the risk of HR from south to north negatively. Similarly, the growth of gross domestic product per capital decreased the probability of HR from southwest to northeast.

Conclusion: A nonrandom spatial distribution of the HR of RA patients was demonstrated in Hunan, and average temperature, gross domestic product per capital and hospitals per 10,000 people showed different regression coefficients on the HR in different counties. Our study indicated that spatial and geostatistics may be useful approaches for further study among RA patients.

Strengths and limitations of this study:

1. This is the first spatial analysis of geographic variation in the hospitalization rate (HR) of RA patients at the district and county levels in Hunan, China.

2. The data came from secondary and tertiary hospitals from provincial administrative databases, which increased accuracy.

3. Geospatial analysis, a promising in-depth approach in the study of RA, was used.

4. Many risk factors, such as air pollutants, were not fully investigated due to a lack of data.

1 Introduction

Rheumatoid arthritis (RA) is an autoimmune disease characterized by abnormal synovial hyperplasia and progressive destruction of cartilage and bone, affecting approximately 0.5–1% of the adult population worldwide [1, 2]. Although the aetiology and pathogenesis of RA remain unclear, it is generally believed that multifactorial factors are involved, combining genetic predispositions [3, 4] and a variety of environmental factors, such as smoking, infection, microbiota, air pollution and climate factors [5-9].

There are differences in the incidence of RA. Alamanos et al. [10] showed that RA varies geographically, with southern European countries having lower median incidence rates than northern European countries and North America. Geographic variation in incident RA has also been observed at the regional level in the USA [11, 12]. The findings suggest women living at higher latitudes may be at greater risk for RA. Furthermore, the risk of RA may be greater for locations that occur earlier in residential histories. Similar geographic clusters were observed for other autoimmune diseases, such as systemic lupus erythematosus, dermatomyositis, polymyositis, and vasculitis [13-15]. Recently, Liu found a notable rural–urban variation in RA prevalence in Alberta, Canada. However, little is known about what is behind these geographic differences. Geostatistics are emerging tools that combine statistical, geographical and epidemiological methodologies in describing the spatial characteristics of both infectious and noninfectious diseases, which may be a promising in-depth approach in the study of RA, providing more information on genetic, socioeconomic and environmental information while simultaneously taking into account their spatial diversity. Most of the previous work on the pathogenesis and prevalence of RA explores potential risk factors using classical statistical techniques.

The hospitalization rates (HR) is an important index which can reflect the local medical and health service capacity and, to some extent, may indicate the prevalence and onset of disease [16]. In addition, hospitalization of patients consumes the majority of medical resources. To date, t the application of geospatial analyses in the HR of RA patients and its risk factors, especially in China, is lacking in the literature. Therefore, we conducted a geospatial study to explore the spatial distribution of the HR of RA patients in Hunan province from a global to local scale and to investigate the potential geographic, socioeconomic, environmental and health care risk factors while simultaneously considering their spatial diversity.

2 Material and methods

2.1 Study location

Hunan Province is located in the southern-central China (24°38'N-30°08'N, 108°47'E-114°15'E), with a total area of 211,800 km^2(http://www.gov.cn/guoqing/2005-09/13/content_5043917.htm). The terrain is a horseshoe-shaped landform surrounded by mountains on three sides and facing the north, and it has a subtropical monsoon climate. Hunan Province has jurisdiction over 14 prefecture-level administrative regions, including 13 prefecture-level cities and 1 autonomous prefecture. There are 36 municipal districts, 19 county-level cities, 60 counties and 7 autonomous counties, totalling 122 county/district-level divisions. In 2019, Hunan Province had a total population of 73.1953 million people and a GDP of 3975.21 billion yuan (http://tjj.hunan.gov.cn/). The Dongting Lake region is an area around Dongting Lake, which is located south of the middle reaches of the Yangze River and north of Hunan Province. In terms of administrative regions, the Dongting Lake region includes 21 counties and districts belonging to Yueyang, Yiyang and Changde in Hunan Province, with a land area of approximately31,700 square kilometres, accounting for 15% of the province's total area. The Dongting Lake region is a famous "land of fish and rice", rich in natural resources, with a large population and relatively high economic level. In this study, Hunan Province was chosen as the representative region of southern-central China, which has four distinct seasons with a varied climate throughout the year and contains many geographical elements, such as mountains and lakes. These geographical advantages are helpful for studying the effects of the HR of RA patients on population health and extrapolating the results.

2.2 Data collection

This study enrolled patients aged 16 or older who were admitted to secondary or tertiary hospitals with main diagnosis of RA from January 1,2019, to December 31,2019, using the *statistical information Direct Reporting System of the Hunan Health Commission*. RA-related records were identified as those with the International Classification of Diseases (ICD), Tenth Revision (ICD-10) codes of M05.X or M06.X. The admission times of patients were collected as well as their individual demographic data, including age, sex, permanent address, place of origin, residential postcode, and vital status. In this study, we used the permanent address to define the patient's location rather than his or her place of origin. The names, ID numbers and telephone numbers of patients were deidentified in the data collection phase. Data confidentiality principles were followed throughout this study. Patients and the public were not involved in the design, conduct, reporting, or dissemination plans of our research.

The independent variables were selected based on the literature review and a priori expert knowledge about the variables that were known to be associated with the time as well as on their availability in the dataset [17-20]. Meteorological data (temperature, air pressure, rainfall) were collected from the *China Meteorological Data Network* (https://data.cma.cn/). Map information came from the standard map service system(http://bzdt.ch.mnr.gov.cn/). Spatiotemporal kriging interpolation analysis was used to complement the daily meteorological data for each district/county of Hunan Province in 2019, then the average annual value was calculated as the total daily value divided by 365 days. The population, economy (gross domestic product, GDP; GDP per capital, PGDP) and health resource (number of doctors, number of beds, number of hospitals) data for each district/county were collected from *Hunan Statistical Yearbook 2019* (http://tjj.hunan.gov.cn/). The HR of RA patients in Hunan

119 and in each district/county were calculated per 10,000 population (the total RA admissions divided by the number of permanent residents), as well as GDP, PGDP, numbers of doctors, numbers of beds and numbers of hospitals. ArcGIS 10.8 was used to make a spatial distribution map of the HR and other variables in Hunan Province at the district/county level. The same color with different shades indicates the variables, the darker the color, the greater the values of the variables in this area.

124 Data checking and data screening were also carried out: ①logical verification was performed by removing cases with obvious logical errors and duplication, such as individuals with 0 hospitalization costs. Patients with multiple admissions on the same day were counted as a single admission; ② missing data, such as cases with missing vital information (such as present residence, household registration address or discharge diagnosis information) were deleted;③ for outliers, graphical representation and Q-test were used to discriminate univariate outliers, and multivariate outliers were determined by Mahalanobis distance. The discovered outlier data were verified for deletion or conversion.

132 **2.3 Statistical methods**

133 In this study, quantitative data were presented as the medians and the 25th–75th percentile interquartile ranges (IQRs). Data such as demographic data, meteorological conditions, socioeconomic factors, and medical resources were described as frequency (percentage)percentiles. The Chi-square method was used for intergroup comparisons of categorical variables, and the statistical significance level was corrected by Bonferroni correction. Global spatial autocorrelation analysis and local spatial autocorrelation analysis were used to determine whether the HR of RA patients was clustered in the province and to analyse the index of local spatial autocorrelation. The included variables were divided into county and individual levels. To analyse the influencing factors more scientifically, a generalized estimating equation (GEE) in two levels was used (district/county level and individual level) to identify the influencing factors of the HR of RA patients. Odds ratios (ORs) and 95% confidence intervals (CIs) were calculated, and variables with a p-value of <0.05 were considered candidates for the geographically weighted regression (GWR) model, which was used to reveal the spatially varying relationships between the RA admission rates and the influencing factors and a set of location-specific parameter estimates. We also conducted OLS and collinearity tests on variables that can be included in GWR to ensure that variables were not omitted. Data were analysed using SPSS statistical software (version 26.0, IBM) and ArcMap (version 10.8, ESRI Inc., Redlands, CA, USA). Districts and counties served as the basic units for spatial analysis (n=122).

150 **2.3.1 Global spatial autocorrelation analysis**

151 The global spatial autocorrelation analysis started from the macrolevel of the whole province and compared the mean value of the attribute in the general area and the attribute value for each spatial unit to obtain the average degree of association between the HRs of RA in various districts /counties across the province. That is, we conducted an analysis on whether the HR of RA patients was clustered within the province. Global Moran's I is a commonly used global correlation index ranging from -1 to 1[21]. I>0 indicates a positive spatial correlation at the given significance level. The larger the value is, the more pronounced the spatial correlation; a high HR of RA accumulates in space with a high HR of RA, and a low HR of RA accumulates spatially. I<0 indicates a negative spatial correlation. The smaller the value is, the more significant the spatial difference, and the greater the tendency of a high HR of RA to cluster with low HR of RA. If I =0, the observations are randomly distributed in space [22]. This method describes whether there is a clustering phenomenon but cannot confirm where the clustering area is.

2.3.2 Local spatial autocorrelation analysis

The local indicator of spatial association (LISA) is the analysis index of local spatial autocorrelation analysis [23]. In this study, LISA was used to measure the degree of spatial dependence between the RA admission rates of a district /county and the RA admission rates of its neighbouring districts /counties and identify its spatial cluster pattern in a local space. There are four spatial correlation patterns: high-high cluster, low-low cluster, high-low cluster, and low-high cluster. The high-high cluster and low-low clusters indicated that the HR of RA had a strong positive spatial correlation. The two patterns indicated that the high-HR RA areas were adjacent to the high-HR RA areas, and the low-HR RA areas were adjacent to the low-HR RA areas. The high-low cluster and low-high cluster signified that the HR of RA had a strong negative spatial correlation. The two patterns mean that high-HR RA areas are adjacent to low-HR RA areas.

2.3.3 Two-level GEE

The GEE is used to estimate the parameters of the generalized linear model, which was first proposed by Liang and zeger in 1986[24]. It is a regression model for analysing correlation data by using the quasi-likelihood estimation method to estimate parameters in the generalized linear model and repeated measurement data. In addition to the normal distribution, the GEE uses the connection function to fit the stress variables of binomial distribution, Poisson distribution, gamma distribution and other distributions into corresponding statistical models, which solves the problem of nonindependence of repeated measurement data and can obtain robust parameters.

In this study, we used The GEE at two level (district/county level and individual level) to study the influencing factors of HR of RA patients. The dependent variable was the HR of the RA patients, and according to the literatures review, independent variables at the district/county level included ① environmental factors, such as the average temperature, average temperature range, average air pressure, average rainfall, whether the location was the Dongting Lake region and seasons;② socioeconomic indicators, including PGDP, GDP, and population, and ③health-care resources, including the number of doctors, number of beds and hospitals per 10,000 people. Independent variables at the individual level included sex, age and ethnic group. The correlation among the variables were tested. A pair of variables with a correlation coefficient greater than 0.7 were considered to be strongly correlated, and we also took VIF＞5 as the basis for collinearity [25]. A better variable was selected to be included in the model according to the clinical significance and the goodness of fit of the model. Finally, sex, ethnicity, age, average temperature, average temperature range, average rainfall, Dongting Lake area, PGDP, number of hospitals, and number of doctors were chosen for the final regression model, as our preliminary analyses found that there was a high correlation between average temperature and average air pressure (r_s=-0.8674), number of beds and number of doctors (r_s=0.940), and GDP and PGDP (r_s=0.787).

2.3.4 GWR

GWR considers spatial dependency by performing several separated regressions, identifying significant geographic clusters within the studied area [23]. Additionally, GWR produced an estimate for the association between the HR of RA patients and its predictor variables from analysis of the local spatial variability for each district and county. It is assumed that the regression coefficient is a function of regional geographical location and changes in geographical location. In this study, meaningful independent variables from the GEE results were selected for the GWR model. The Akaike information criterion (AIC) and R^2 were adopted to build the best fit model for analysis; the smaller the AIC and larger the R^2 were, the better model was. Based on Fotheringham et al. [26], the GWR model with one independent variable can be expressed as follows:

$$y_i = \beta_0(u_i, v_i) + \beta_1(u_i, v_i)x_i + \varepsilon_i$$

Where u_i and v_i are the coordinates of the i th location, and β_0 (u_i, v_i) is the local intercept for i th location, β_1 (u_i, v_i)is the estimated local regression coefficient for the i th location and ε_i is the random error at the i th location.

2.4 Patient and Public Involvement

Patients or the public were not involved in the design, conduct, reporting, or dissemination plans of our research.

Results

3.1 Characteristics of hospitalizations of RA patients in Hunan

Basic and admission information of RA patients is shown in Table 1. In 2019, the total number of hospitalizations of RA patients in Hunan Province was 11,599, including 8223(70.9%) females and 3376 (29.1%) males (2.441:1). The median age of the admitted RA patients was 63(P_{25}=54, P_{75}=70) years. Patients aged 60 and above accounted for 59.3% (6877), those aged 19-59 accounted for 33.9% (4694), and those aged 16-18 accounted for 0.2% (28). The total HR was 1.57 per 10,000 people, and significant differences were found in terms of sex, age groups, ethnic groups, regions (whether the patients were located in the Dongting Lake region) and seasons (p<0.001). The HR of RA patients was much higher among female, elderly people (⩾60 years), those of Han ethnicity and in the region around Dongting Lake(p<0.01). Regarding admission time, there was a lower HR of RA patients in spring than in summer, autumn and winter(p<0.01).

3.2 Distribution of the HRs of RA patients and demographic data, meteorological conditions, socioeconomic factors, and medical resources among 122 counties and districts in Hunan

Supplement Table 1 list the basic information of demographic data, meteorological conditions, socioeconomic factors, medical resources and HRs for RA patients at the district/county level(n=122). Higher HRs of RA patients were observed in the central area, the northeast region and the Dongting Lake basin in Hunan Province (Figure 1). In contrast, the lowest HRs were observed in regions mainly located in the mountainous areas. There was only one RA patient admitted in Wuling yuan District, Zhangjiajie, with a low HR of 0.01 per 10,000 people, while in Xinxiang, a city near the provincial capital of Changsha, the HR was 26.57 per 10,000 people. A significant variation in the distribution of HRs among RA patients was noted in different districts and counties as well as the distribution of other climatic factors, socioeconomic indicators and medical resources (Figure 1).

3.3 Global and local spatial autocorrelation of the HRs of RA patients among 122 districts and counties in Hunan, China

Global spatial autocorrelation analysis showed a globally spatiality of the HR of RA patients in Hunan Province (Moran's I=0.2585, Z=4.0318, p<0.000). LISA analysis detected two high-high (HH) pattern clusters (Ning Xiang county and Shuang Feng county) in the central regions of Hunan, in which high RA admission rates close to neighbors were also associated with high RA admission rates. Eleven low-

low (LL) pattern clusters were also observed in the northwestern and southern mountainous areas in Hunan, indicating districts and counties with low RA admission rates with neighbors with low RA admission rates (Figure 2, Supplement Table 2). Notably, six LL clusters (four districts and two counties) were found in Hengyang city. We also found one high-low pattern cluster and five low-high pattern clusters in Hunan. The global and local spatial autocorrelation coefficients were considered significant at $p < 0.05$.

3.4 Factors including demographic indictors, meteorological conditions, socioeconomic conditions and medical resources were associated with the HR of RA patients in Hunan in two-level GEE analysis

In the two-level GEE analysis, age, ethnicity, average temperature, average temperature range, average rainfall, regions (whether the patients were located in the Dongting Lake region area or not), PGDP, and the number of doctors and hospitals per 10,000 people were risk factors for hospitalization among RA patients, sex was not a risk factor (p=0.213, Table 2). Among the factors, the average temperature range and number of hospitals per 10,000 had a slightly positive effect on the HR of RA patients (OR,1.459;95%CI,1.246-1.641; p<0.001; OR,1.167;95%CI,1.157-1.181; p<0.001;).In contrast, the risk of hospitalization was negatively correlated with temperature, rainfall, PDGP, number of doctors per 10,000 people (OR,0.876;95%CI,0.785-0.983;p<0.001;OR,0.997;95%CI,0.996-0.999;p<0.001;OR,0.953;95%CI,0.936-0.983;p<0.001; OR,0.965;95%CI,0.948-0.971; p<0.001). The subgroup analysis demonstrated the risk of hospitalization in the Han group, the elderly group (≥60 years), and the Dongting Lake region group was much higher than that in the minority group, the younger group (<60 years), and the non-Dongting Lake region group (p<0.001).

3.5 The impact of average temperature, PGDP and number of hospitals on the HR of RA patients among 122 districts and counties in Hunan by GWR analysis

To further identify risk factors that had the greatest geospatial impact on the spatial distribution of HRs among RA patients, we performed a multivariate spatial regression analysis using GWR based on the above GEE analysis.

The GWR model showed that average temperature, PGDP, hospitals per 10,000 people were independent influencing factors for the HR of RA patients with spatial trend (R^2,0.521; adjustR^2,0.493; AIC,124.543). The distinctive spatial impact of each significant variable on the HR of RA patients is indicated in the Figure 3. The effect of average temperature and PGDP on the HR of RA patients in Hunan Province was negative. The great reduction in temperature, the lower of the PGDP and the higher the HR of RA patients. Overall, the influence of temperature on the HR decreased from the south of Hunan to the northern regions (regression coefficient, −0.0548–−0.0546)and the HR of RA patients in the south was more likely to be affected by average temperature. Similarly, the influence of PGDP decreased from the southwest to the northeast regions in Hunan, which means that the HR of patients was more likely to be affected by PGDP in the southwest (regression coefficient, −0.0637–−0.0634). In contrast, the impact of the number of hospitals per 10,000 people on the HR of RA patients was positive and increased from the eastern to the western regions; the greater the number of hospitals was, the higher the HR of people was, especially in western Hunan (regression coefficient, 0.0257–0.0258). The GWR analysis suggested that the HR of RA patients was indeed affected by local temperature, PGDP and the number of hospitals and varied by region.

3 Discussion

RA is a chronic autoimmune disease and a leading cause of working disability, and this situation will become more serious with the increase of aging. Although the aetiology is not entirely clear, RA is a multifactorial disease involving genetics, immunology and environment. The ultimate therapeutic goal for RA is to achieve long-term disease remission or low disease activity. Thus, the management of RA requires collaboration not only between patients and health provider but also the formulation of corresponding medical policies, medical resource allocation and support from the government. There are large gaps in economic conditions, medical resource allocation and medical service efficiency among different regions in China and even different cities in the same province, and meteorological conditions differ by region. In this study, the spatial distribution of the HR of RA patients by district/county was first evaluated, and its association with demographics, meteorological conditions, socioeconomic indictors and medical resources, taking spatial dependence into account, by GWR analysis and GEE analysis was investigated in Hunan, a southern-central province in China with a median national PGDP level. To the best of our knowledge, this is the first spatial analysis study on the HR for RA patients in China.

In 2019, the annual HR for RA was 1.57 per 10,000 people in the whole province, but we found a notable variation from 0.01 case/10,000 population to 26.57 cases/10,000 population between geographic areas. The HR of RA patients was significantly higher in the central and northeast areas and the Dongting Lake region in Hunan. We identified 2 high-high clusters located in the northeastern and central region and 11 low-low clusters located in mountain areas. Furthermore, we investigated the possible risk factors associated with disparities of in the HR of RA patients in Hunan. Notably,age, ethnicity, temperature, temperature range, rain fall, Dongting Lake area, PGDP, and number of doctors and hospitals per 10,000 people were risk factors (p<0.001) in the GEE analysis. However, only temperature, PGDP and number of hospitals per 10,000 people were related to the risk of hospitalization among RA patients in GWR when considering spatial dependence. In total, patients in regions with lower temperature, lower PGDP or more hospitals per capita had a higher risk of hospitalization. Notably, the influence of temperature, PGDP and number of hospitals on the HR of RA patients in Hunan varied by region. These spatial variations and risk factors should be considered when designing evidence-based interventions and planning programs to improve RA care in Hunan. Our findings highlight the need for regional approaches for the planning and delivery of RA care as well as the management of RA patients. Further studies should investigate the underline pathogenesis ofthe difference in HRs among RA patients in Hunan, especially the relationship with temperature, PDGP and numbers of hospitals per 10,000 people.

A total of 11,599 RA patients were admitted to hospitals in Hunan in 2019, with a female predominate, and the mean age was 63(54,70) years-old, which was similar to the results in other reports in China[9]. In our study, the HR of RA was 1.57 per 10,000 population, which was much lower than the prevalence of RA (approximately 0.5–1% of the adult population). In recent years, a decrease in the number of RA hospitalization in Western countries has been reported. The number of RA hospitalization in WA, USA, decreased from 202.6 per 100,000 population in 1995 to 87.3 per 100,000 population in 2002[16]. In New Zealand, the HR of RA decreased from 11.3 to 8 per 100,000 inhabitants between 1995 and 2002[27]. In Canada, the annual HR of RA was decreased from 15.4 to 12.4 per 100,000 population between 2000 and 2002[28]. Due to the lack of data on the HR of RA patients in Hunan, even in China, it is difficult to know whether the HR will decrease or increase, but there has been dynamic change in recent years. On one hand, the HR may increase because of the overall improvement in the numbers of rheumatologists and departments in hospitals, access to information and health insurance system coverage in low-income arears for patients. On the other hand, as an increasing number of patients with RA are detected early and receive better management, a decline in the HR of RA patients might occur.

The spatial distributions of infectious diseases have been extensively studied [29-31], for the spread of the pathogenic microbial agent. However, noncommunicable diseases in humans also show clustering in space. Such is the case for heart diseases, diabetes and autoimmune inflammatory rheumatic diseases [15, 32-34]. Matej Sapina et al. found that incidences of childhood IgA vasculitis and IgA vasculitis-associated nephritis in Croatia were not randomly distributed but clustered in space [35]. An enormous variation in the HRs of RA patients among122 districts and counties in Hunan was found, ranging from 0.01-26.57 per 10,000 population. In addition, the existence of spatial autocorrelation was observed in the HR of RA in our study. H-H clusters of HR for RA patients were found in Ning Xiang county and Shuang Feng county in the northeastern and central areas, showing higher positive HR autocorrelations in those areas. Ning Xiang county located in the capital city of Changsha. Five L-L clusters were found dispersed partially in the northwestern, southeastern and southern mountainous areas, while other 6 L-L clusters, belonging to Hengyang city, were located in the south-central part of Hunan. Surprisingly, Shuang Feng county, an H-H cluster, was near the L-L clusters in Hengyang city. One of the possible reasons may be their geographical differences. The terrain of Shuang Feng is closer to the northern and central plains but separated from the latter by Mountain Hengshan. However, subsequent questions arise and should be discussed: why do these hotspots and cold spots occur, and how can they be explained?

Previous studies using GAMs and GLS have shown that incident RA varies geographically at the regional level in the USA as well as worldwide [11, 12]. The findings suggested that women living at higher latitudes may be at greater risk for RA. Furthermore, the risk of RA may be greater for older locations with a longer residential history. Recently, a significant rural-urban variation in RA prevalence in Alberta, Canada, was reported in 2021[36]. The authors reported factors such as obesity, socioeconomic factors, genetic and environmental factors such as aboriginal and immigrant status, and local meteorological conditions may partially explain these variations.

Consistent with the HR, the demographic characteristics, meteorological conditions, socioeconomic indictors and health care resources also demonstrated geospatial discrepancies at the district and county levels in our study. In addition, we found that age, ethnicity, temperature, temperature range, rain fall, Dongting Lake region, PGDP, and number of doctors and hospitals per 10,000 people were risk factors for hospitalization among patients with RA in the GEE analysis.

However, if we consider the spatial variation, only temperature, PGDP and number of hospitals per 10,000 people affected the risk of hospitalization among RA patients with spatial dependence using the GWR model. There have been many studies on the effect of temperature on RA incidence, joint pain and medical seeking behaviours, such as clinic visits or admissions; however, the results are conflicting, which may be mainly explained by different statistical analysis methods and adjustment for climate covariates. Recently, Huang et al [2]. reported the exposure to cold temperatures increased the risk of RA admissions in Anhui using spatiotemporal series analysis with a distributed lag nonlinear model. Notably, we found that the influence of temperature on the HR was negative regardless of whether spatial dependence was considered. A significant trend was demonstrated in which the effect of temperature on the HR increased from north to south. When the temperature dropped, the risk of hospitalization among RA patients in the southern area increased slightly compared to that in the northern area. This result indicated that geospatial analysis may provide a new approach in future studies to address the conflicting influence of climate factors in RA and other autoimmune and inflammation diseases. Low temperatures were also associated with increased risks of all types of inflammation-related diseases in Yan Chi County, Ningxia, China (RR: 4.344, 95% CI: 2.887, 6.535) [37]. Abundant evidence indicates that extreme high or low temperatures could increase risks of mortality and morbidity for cardiovascular and respiratory diseases [38, 39]. Some researchers believed that the underlying reasons may be that inflammatory factors such as interlukin-6 and Cc-reactive protein play an important role in triggering the adverse effects of low or extreme temperatures [40, 41].

It's well known that lower socioeconomic status increases the risk of RA prevalence, functional decline and overall mortality rate over time [42, 43]. In this study, we also demonstrated that lower PGDP was associated with a higher risk of hospitalization among RA patients, and the effect of GDP showed spatial variation in different regions in Hunan. The HR of RA patients seemed more likely to be affected by the PGDP in the southwest, while it was much lower in the northeast region. Greater attention should be given to the economic situation of RA patients in the southwest region, and more should be done to improve their poverty. In addition, the inequalities in healthcare, low accessibility to health care centres and limited health coverage offered to the patients may also be responsible for the HR of RA. In this study, the number of hospitals per 10,000 people had a positive effect on the HR of RA patients. In addition, the impact of the number of hospitals per 10,000 people on the HR of RA patients was positive, increasing from the eastern to the western regions in Hunan. The greater the number of hospitals was, the higher the HR, especially in western Hunan. This was consistent with the lack of health care centres in the western region of Hunan. Thus, increasing access to health care centres and the disrupting barriers to health care delivery would greatly benefit RA patients in western Hunan in the future. Large-scale studies based on epidemiological and physiological data are necessary to further determine whether there are differences in the HR of RA and the related mechanisms behind the influence of temperature, economic level and medical resources on the HR of RA patients. The GWR analysis suggested that inequalities in health care and major disparities in individual income and local temperature may be responsible for the HR of RA patients in Hunan. However, there are also other factors that affect the HR of RA.

This study has strengths and limitations. First, we applied spatial analysis to explore the geographic variation in the HR of RA patients at the district and county levels, which may fill the gap in the knowledge of the HR of RA patients at local levels in Hunan, providing important information for health care planning. Second, our study indicated that geospatial analysis may be a promising in-depth approach in the study of RA, providing more genetic, socioeconomic and environmental information. However, several limitations must be taken into consideration when interpreting the results. First, this was a cross-section study from Hunan Province that used provincial administrative databases from secondary or tertiary hospitals as the source. All of these factors may have led to biases. Previous studies suggest that the RA case definition using health administrative databases has excellent sensitivity when multiple physician claims are included [15]. Due to the lack of rheumatologists in primary hospitals in Hunan, we believe that these estimates may largely represent the true number of hospitalizations for RA patients in Hunan Province. However, these data have inherent limitations and should be interpreted with caution. Second, we used the annual average data of climate factors for analysis, ignoring their short-term influence on the patients, which needs to be further studied in subsequent studies. Third, there are numerous determinants of the HR of RA patients; nevertheless, a number of risk factors, such as air pollutants, were not investigated in the present study due to the lack of data, which may be considered as a limitation of the present study. Fourth, GWR can analyse the HR of RA from multiple spatial factors but ignores the different spatiotemporal scales of various factors. Thus, for a more reasonable analysis, we first took a two-level GEE to identify significant variables. Finally, it is not possible to accurately ascertain individual exposures, which may have led to certain measurement biases.

4 Conclusion

A nonrandom spatial distribution of the HR of RA patients was demonstrated in Hunan. In addition, average temperature, PGDP and hospitals per 10,000 people showed different regression coefficients across the regions on the HR of RA patients. Our study indicated that spatial and geostatistics may be useful approaches for the further study in RA to provide more information on genetic, socioeconomic and environmental risk factors while taking into account their spatial diversity.

5 Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

6 Author Contributions

Yan Ge, Shiwen Wang designed, analyzed the results and wrote the manuscript. Qianshan Shi supervised the manuscript and data curation. Jingcheng Shi, Jing Tian designed, supervised and edited the manuscript.

7 Funding

This work was supported by National Natural Science Foundation of China (No.81701622), National Science& Technology Foundational Resource Investigation Program of China (No.2018FY100900), Natural Science Foundation of Hunan (No.2022JJ40674, No.2020JJ4772 and No.2021JJ30893), Natural Science Foundation of Changsha (No. kq2202409).

8 Acknowledgments

We thank the Information Statistics Center of Health Commission of Hunan Province for their support with data access and all agencies who provided the meteorological data.

9 Ethics Approval statement

This study was approved by the Ethics Committee of the Second Xiangya Hospital of Central South University (xyeyy-kq-190119).

10 Data sharing statements

The data that support the findings of this study are available from the author, Qianshan Shi(hnsqs@qq.com), upon reasonable request.

11 Reference

1. Scott DL, Wolfe F, Huizinga TW. Rheumatoid arthritis. Lancet. 2010;376(9746):1094-108.

2. Huang LJ, Zha JJ, Cao NW, Zhou HY, Chu XJ, Wang H, Li XB, Li BZ. Temperature might increase the hospital admission risk for rheumatoid arthritis patients in Anqing, China: a time-series study. Int J Biometeorol. 2022;66(1):201-11.

3. Karami J, Aslani S, Jamshidi A, Garshasbi M, Mahmoudi M. Genetic implications in the pathogenesis of rheumatoid arthritis; an updated review. Gene. 2019;702:8-16.

4. Constantin A, Navaux F, Lauwers-Cances V, Abbal M, van Meerwijk JP, Mazieres B, Cambon-Thomsen A, Cantagrel A. Interferon gamma gene polymorphism and susceptibility to, and severity of, rheumatoid arthritis. Lancet. 2001;358(9298):2051-2.

5. Liu X, Tedeschi SK, Barbhaiya M, Leatherwood CL, Speyer CB, Lu B, Costenbader KH, Karlson EW, Sparks JA. Impact and Timing of Smoking Cessation on Reducing Risk of Rheumatoid Arthritis Among Women in the Nurses' Health Studies. Arthritis Care Res (Hoboken). 2019;71(7):914-24.

6. Arleevskaya MI, Albina S, Larionova RV, Gabdoulkhakova AG, Lemerle J, Renaudineau Y. Prevalence and Incidence of Upper Respiratory Tract Infection Events Are Elevated Prior to the Development of Rheumatoid Arthritis in First-Degree Relatives. Front Immunol. 2018;9:2771.

7. Zhang X, Zhang D, Jia H, Feng Q, Wang D, Liang D, Wu X, Li J, Tang L, Li Y, Lan Z, Chen B, Li Y, Zhong H, Xie H, Jie Z, Chen W, Tang S, Xu X, Wang X, Cai X, Liu S, Xia Y, Li J, Qiao X, Al-Aama JY, Chen H, Wang L, Wu QJ, Zhang F, Zheng W, Li Y, Zhang M, Luo G, Xue W, Xiao L, Li J, Chen W, Xu X, Yin Y, Yang H, Wang J, Kristiansen K, Liu L, Li T, Huang Q, Li Y, Wang J. The oral and gut microbiomes are perturbed in rheumatoid arthritis and partly normalized after treatment. Nat Med. 2015;21(8):895-905.

8. Wu Q, Xu Z, Dan YL, Cheng J, Zhao CN, Mao YM, Xiang K, Hu YQ, He YS, Pan HF. Association between traffic-related air pollution and hospital readmissions for rheumatoid arthritis in Hefei, China: A time-series study. Environ Pollut. 2021;268(Pt A):115628.

9. Wang J, Yu L, Deng J, Gao X, Chen Y, Shao M, Zhang T, Ni M, Pan F. Short-term effect of meteorological factors on the risk of rheumatoid arthritis hospital admissions: A distributed lag non-linear analysis in Hefei, China. Environ Res. 2022;207:112168.

10. Alamanos Y, Voulgari PV, Drosos AA. Incidence and prevalence of rheumatoid arthritis, based on the 1987 American College of Rheumatology criteria: a systematic review. Semin Arthritis Rheum. 2006;36(3):182-8.

11. Vieira VM, Hart JE, Webster TF, Weinberg J, Puett R, Laden F, Costenbader KH, Karlson EW. Association between residences in U.S. northern latitudes and rheumatoid arthritis: A spatial analysis of the Nurses' Health Study. Environ Health Perspect. 2010;118(7):957-61.

12. Costenbader KH, Chang SC, Laden F, Puett R, Karlson EW. Geographic variation in rheumatoid arthritis incidence among women in the United States. Arch Intern Med. 2008;168(15):1664-70.

13. Tian J, Zhang D, Yao X, Huang Y, Lu Q. Global epidemiology of systemic lupus erythematosus: a comprehensive systematic analysis and modelling study. Ann Rheum Dis. 2022.

14. Muro Y, Sugiura K, Hoshino K, Akiyama M, Tamakoshi K. Epidemiologic study of clinically amyopathic dermatomyositis and anti-melanoma differentiation-associated gene 5 antibodies in central Japan. Arthritis Res Ther. 2011;13(6):R214.

15. Liu X, Barber CEH, Katz S, Homik J, Bertazzon S, Patel AB, Robert J, Smith C, Mosher D, Marshall DA. Geographic Variation in the Prevalence of Rheumatoid Arthritis in Alberta, Canada. ACR Open Rheumatol. 2021;3(5):324-32.

16. Almutairi K, Nossent J, Preen DB, Keen H, Inderjeeth C. The temporal association between hospital admissions, biological therapy usage and direct health care costs in rheumatoid arthritis patients. Rheumatol Int. 2022;42(11):2027-37.

17. Venetsanopoulou AI, Alamanos Y, Voulgari PV, Drosos AA. Epidemiology of rheumatoid arthritis: genetic and environmental influences. Expert Rev Clin Immunol. 2022;18(9):923-931.

18. McInnes IB, Schett G. The pathogenesis of rheumatoid arthritis. N Engl J Med. 2011;365(23):2205-2219.

19. Smolen JS, Aletaha D, McInnes IB. Rheumatoid arthritis [published correction appears in Lancet. 2016 Oct 22;388(10055):1984]. Lancet. 2016;388(10055):2023-2038.

20. Wang J, Yu L, Deng J, et al. Short-term effect of meteorological factors on the risk of rheumatoid arthritis hospital admissions: A distributed lag non-linear analysis in Hefei, China. Environ Res. 2022;207:112168.

21. Pfeiffer DU, Robinson TP, Stevenson M, Stevens KB, Rogers DJ, Clements AC. Spatial analysis in epidemiology: OUP Oxford; 2008.

22. Shaweno D, Karmakar M, Alene KA, Ragonnet R, Clements AC, Trauer JM, Denholm JT, McBryde ES. Methods used in the spatial analysis of tuberculosis epidemiology: a systematic review. BMC Med. 2018;16(1):193.

23. Rezaeian M, Dunn G, St Leger S, Appleby L. Geographical epidemiology, spatial analysis and geographical information systems: a multidisciplinary glossary. J Epidemiol Community Health. 2007;61(2):98-102.

24. Zeger SL, Liang KY, Albert PS. Models for longitudinal data: a generalized estimating equation approach. Biometrics. 1988;44(4):1049-60.

25. A.G. Asuero, A. Sayago & A. G. González. The Correlation Coefficient: An Overview, Critical Reviews in Analytical Chemistry, 2006,36:1, 41-59.

26. Raza O, Mansournia MA, Rahimi Foroushani A, Holakouie-Naieni K. Geographically Weighted Regression Analysis: A Statistical Method to Account for Spatial Heterogeneity. Arch Iran Med. 2019;22(3):155-60.

27. Collings S, Highton J. Changing patterns of hospital admissions for patients with rheumatic diseases. N Z Med J. 2002;115(1150):131-2.

28. Rai SK, Avina-Zubieta JA, McCormick N, De Vera MA, Lacaille D, Sayre EC, Choi HK. Trends in Gout and Rheumatoid Arthritis Hospitalizations in Canada From 2000 to 2011. Arthritis Care Res (Hoboken). 2017;69(5):758-62.

29. Chowell G, Rothenberg R. Spatial infectious disease epidemiology: on the cusp. BMC Med. 2018;16(1):192.

30. Caprarelli G, Fletcher S. A brief review of spatial analysis concepts and tools used for mapping, containment and risk modelling of infectious diseases and other illnesses. Parasitology. 2014;141(5):581-601.

31. Zhang Y, Liu M, Wu SS, Jiang H, Zhang J, Wang S, Ma W, Li Q, Ma Y, Liu Y, Feng W, Amsalu E, Li X, Wang W, Li W, Guo X. Spatial distribution of tuberculosis and its association with meteorological factors in mainland China. BMC Infect Dis. 2019;19(1):379.

32. Casper M, Kramer MR, Quick H, Schieb LJ, Vaughan AS, Greer S. Changes in the Geographic Patterns of Heart Disease Mortality in the United States: 1973 to 2010. Circulation. 2016;133(12):1171-80.

33. Silman A, Harrison B, Barrett E, Symmons D. The existence of geographical clusters of cases of inflammatory polyarthritis in a primary care based register. Ann Rheum Dis. 2000;59(2):152-4.

34. de Carvalho Dutra A, Silva LL, Pedroso RB, Tchuisseu YP, da Silva MT, Bergamini M, Scheidt J, Iora PH, do Lago Franco R, Staton CA, Vissoci JRN, Nihei OK, de Andrade L. The Impact of Socioeconomic Factors, Coverage and Access to Health on Heart Ischemic Disease Mortality in a Brazilian Southern State: A Geospatial Analysis. Glob Heart. 2021;16(1):5.

35. Sapina M, Frkovic M, Sestan M, Srsen S, Ovuka A, Batnozic Varga M, Kramaric K, Brdaric D, Milas K, Gagro A, Jelusic M. Geospatial clustering of childhood IgA vasculitis and IgA vasculitis-associated nephritis. Ann Rheum Dis. 2021;80(5):610-6.

36. Liu X, Barber CEH, Katz S, Homik J, Bertazzon S, Patel AB, Robert J, Smith C, Mosher D, Marshall DA. Geographic Variation in the Prevalence of Rheumatoid Arthritis in Alberta, Canada. ACR Open Rheumatology. 2021;3(5):324-32.

37. Wang Q, Zhao Q, Wang G, Wang B, Zhang Y, Zhang J, Li N, Zhao Y, Qiao H, Li W, Liu X, Liu L, Wang F, Zhang Y, Guo Y. The association between ambient temperature and clinical visits for inflammation-related diseases in rural areas in China. Environ Pollut. 2020;261:114128.

38. Green H, Bailey J, Schwarz L, Vanos J, Ebi K, Benmarhnia T. Impact of heat on mortality and morbidity in low and middle income countries: A review of the epidemiological evidence and considerations for future research. Environ Res. 2019;171:80-91.

39. Zhao Q, Li S, Coelho M, Saldiva PHN, Hu K, Arblaster JM, Nicholls N, Huxley RR, Abramson MJ, Guo Y. Geographic, Demographic, and Temporal Variations in the Association between Heat Exposure and Hospitalization in Brazil: A Nationwide Study between 2000 and 2015. Environ Health Perspect. 2019;127(1):17001.

40. Hampel R, Breitner S, Ruckerl R, Frampton MW, Koenig W, Phipps RP, Wichmann HE, Peters A, Schneider A. Air temperature and inflammatory and coagulation responses in men with coronary or pulmonary disease during the winter season. Occup Environ Med. 2010;67(6):408-16.

41. Schneider A, Panagiotakos D, Picciotto S, Katsouyanni K, Lowel H, Jacquemin B, Lanki T, Stafoggia M, Bellander T, Koenig W, Peters A, Group AS. Air temperature and inflammatory responses in myocardial infarction survivors. Epidemiology. 2008;19(3):391-400.

42. Izadi Z, Li J, Evans M, Hammam N, Katz P, Ogdie A, Suter LG, Yazdany J, Schmajuk G. Socioeconomic Disparities in Functional Status in a National Sample of Patients With Rheumatoid Arthritis. JAMA Netw Open. 2021;4(8):e2119400.

43. Chen CH, Huang KY, Wang JY, Huang HB, Chou P, Lee CC. Combined effect of individual and neighbourhood socioeconomic status on mortality of rheumatoid arthritis patients under universal health care coverage system. Fam Pract. 2015;32(1):41-8.

Table 1. Characteristics of hospitalizations of RA patients in Hunan in 2019

	Hospitalizations (n, %)	Hospitalization rate（per 10,000 people）	*P*
Total	11599(100)	1.57	
Gender			<0.001
Male	3376(29.1）	0.91	
Female	8223(70.9)	2.32	
Age groups in years			<0.001
16-18	28(0.2)	0.024[ab]	
19-59	4694(40.5)	0.86[b]	
≥60	6877(59.3)	10.71	
Ethnic group			<0.001
Han	11211(96.7)	1.88	
ethnic minorities	388(3.3)	0.58	
Region			<0.001
Dongting Lake	2491(21.48)	2.06	
Non-Dongtin Lake	9108(78.52)	1.61	
Seasons			<0.001
Spring (March-May)	2167(18.7)	0.29	
Summer (June-August)	3186(27.5)	0.43[c]	
Autumn (September-November)	3132(27.0)	0.42[c]	
Winter (December-February)	3114(26.8)	0.42[c]	

a: Compared to the 19-59 years group，$P<0.01$；b: Compared to 60 years and above group，$P<0.01$；c: Compared to the Spring (March-May)，$P<0.01$.

610

611

612 **Table 2. The two-level GEE analysis for risk factor of the HR of RA patients in Hunan in**
613 **2019**

Variables		***OR***	***P***	**95% CI**	
				Lower	**Upper**
(Intercept)	216.868	——	<0.001	150.845	311.788
Gender					
Female		1.036	0.213	0.848	1.151
male		1			
Age					
16-18		0.089	<0.001	0.024	0.139
19-59		0.837	<0.001	0.742	0.913
⩾60		1			
Ethnic group					
Han		4.482	<0.001	3.341	5.868
Ethnic minorities		1			
Temperature		0.876	<0.001	0.785	0.983
Temperature range (℃)		1.459	<0.001	1.246	1.641
Rain fall(mm)		0.997	<0.001	0.996	0.999
Region					
Non-Dongting		1			
Dongting Lake		1.340	<0.001	1.145	1.511
PGDP		0.953	<0.001	0.936	0.983
Doctors (per 10,000 people)		0.965	<0.001	0.948	0.971
Hospitals (per 10,000 people)		1.167	<0.001	1.157	1.181

614 PGDP: GDP per capita; OR: odds ratio; CI: confidence interval. Estimated odds ratios (ORs) were
615 compared with the reference (OR=1) in each group.

616

617

618

619

620

621

622

623

624

Figure 1. Distribution of the HR and average demographic data, meteorological conditions, socioeconomic factors and medical resources among 122 districts and counties in Hunan in 2019.

A non-random distribution of HR for RA patients was demonstrated in 122 districts and counties in Hunan(A), and distribution of demographic data (population per10,000 people, sex portion (F/M); B-C), meteorological conditions (temperature, air pressure, temperature range, rain fall; D-G), socioeconomic factors (GDP, PGDP; H-I) and medical resources (doctors, beds, hospitals per 10,000 people; J-L) were different in Hunan.

Figure 2．Spatial distribution clusters of the HR of RA patients among 122 districts and counties in Hunan in 2019 by local indicators of spatial association (LISA) analysis.

LISA analysis indicates clusters according to high and low HR distribution patterns.

Figure 3. The effect of average temperate, PGDP and hospitals per 10,000 on the HR of RA patients among 122 districts and counties in Hunan by GWR analysis.

Average temperate, PGDP and hospitals per 10,000 show distinctive spatial impact on HR of RA patients.

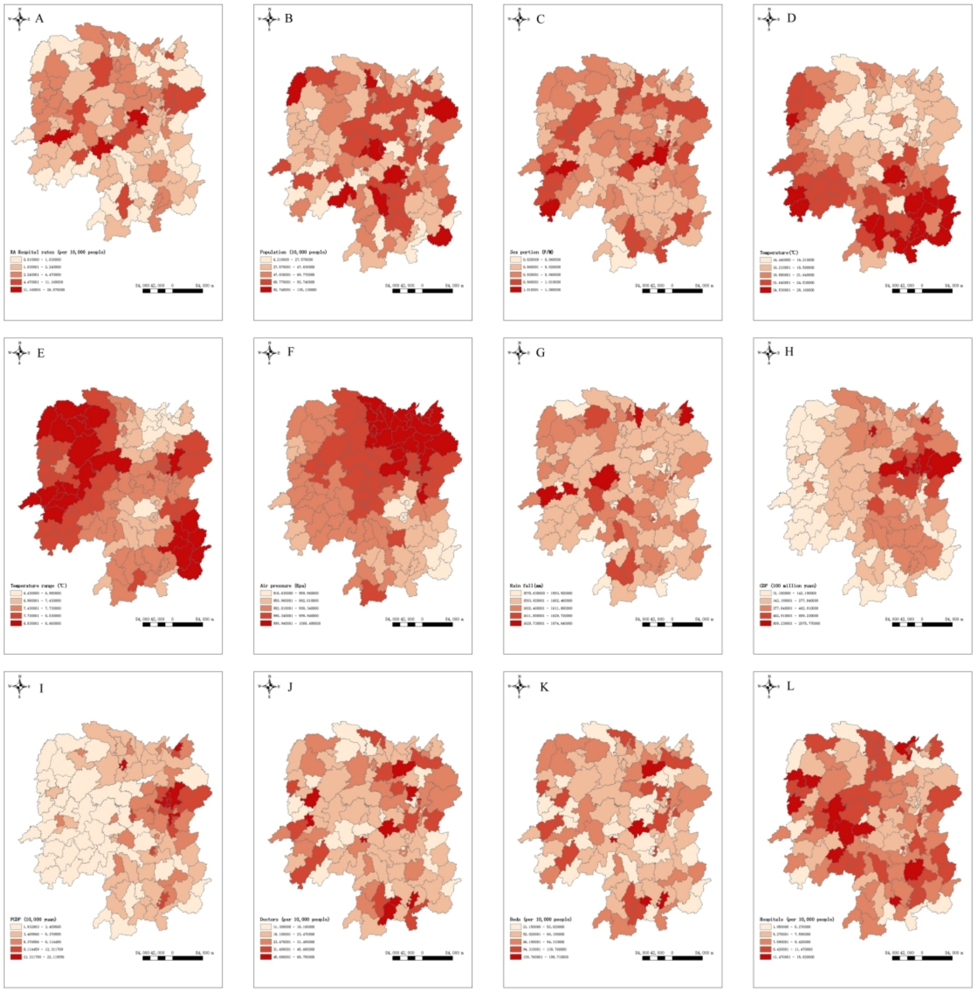

Figure 1. Distribution of the HR and average demographic data, meteorological conditions, socioeconomic factors and medical resources among 122 districts and counties in Hunan in 2019.

254x254mm (299 x 299 DPI)

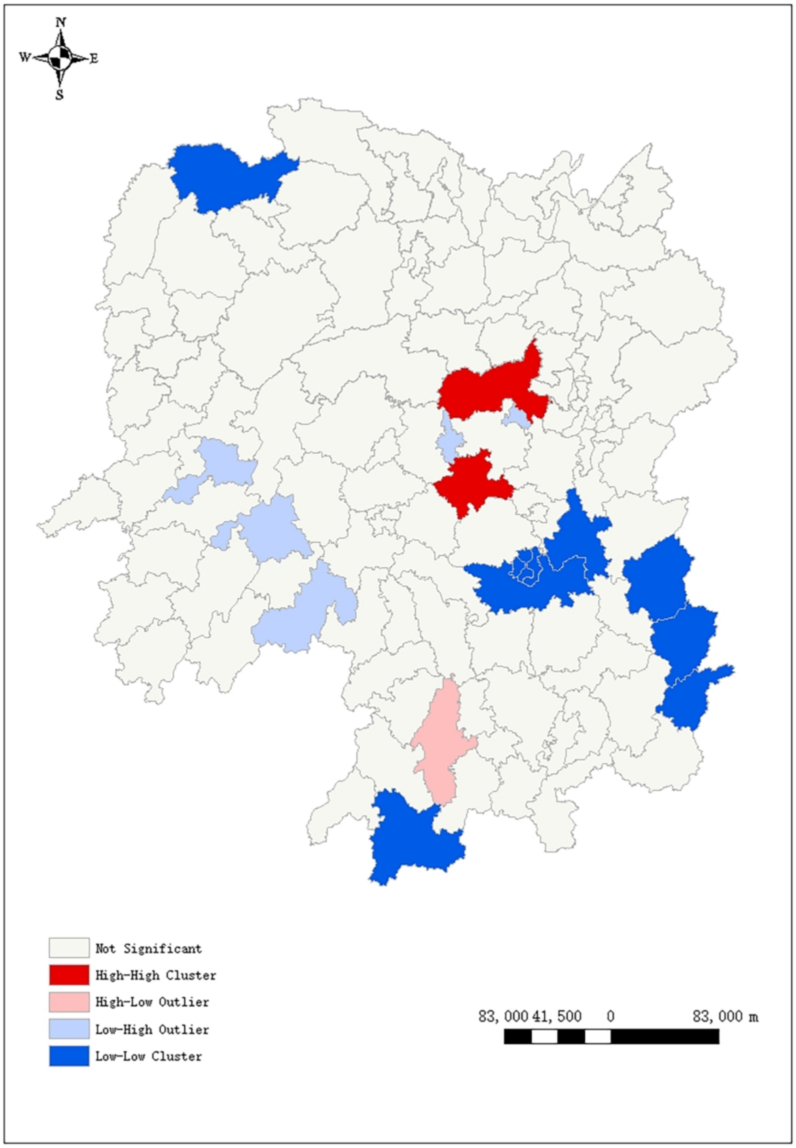

Figure 2. Spatial distribution clusters of the HR of RA patients among 122 districts and counties in Hunan in 2019 by local indicators of spatial association (LISA) analysis.

210x297mm (300 x 300 DPI)

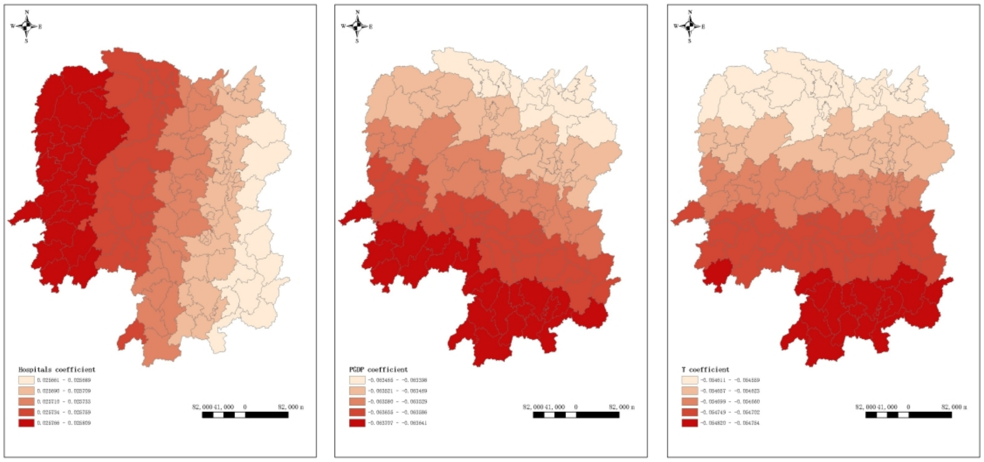

Figure 3. The effect of average temperate, PGDP and hospitals per 10,000 on the HR of RA patients among 122 districts and counties in Hunan by GWR analysis.

252x119mm (330 x 330 DPI)

Table S1. Characteristics of average demographic data, meteorological conditions, socioeconomic factors, medical resources and the HR of RA patients among 122 districts and counties in Hunan in 2019

Variables	Percentile				
	X_{min}	P_{25}	P_{50}	P_{75}	X_{max}
Population (10,000 people)	6.21	34.60	57.06	81.07	135.13
Sex portion (F/M)	0.82	0.92	0.94	0.96	1.08
Age(years)	40	45	52	60	63
Temperature (℃)	14.44	16.49	19.17	23.20	28.16
Air pressure (Kpa)	918.63	982.28	990.19	997.53	1009.49
Temperature range (℃)	6.43	7.40	7.67	7.94	8.66
Rain fall (mm)	1578.63	1596.27	1598.92	1606.81	1874.64
GDP (100 million yuan)	31.18	117.59	234.72	364.87	2075.77
PGDP (10,000 yuan)	1.93	2.90	4.42	6.72	22.11
Doctors (per 10,000 people)	11.30	19.31	22.47	29.36	68.75
Beds (per 10,000 people)	21.15	50.75	58.69	78.39	198.71
Hospitals (per 10,000 people)	1.95	6.70	8.53	10.09	15.82
HR of RA patients (per 10,000 people)	.01	.88	1.65	2.75	26.57

F: female; M: male; GDP: Gross Domestic Product; PGDP: GDP per capita; HR: Hospitalization rates

Table S2. Different cluster patterns of the HR of RA patients in Hunan in 2019

District/county	Local Moran's *I*	*Z* Score	*P*	Patterns
Ningxiang county[a]	4.46×10^{-6}	2.16	0.04	HH
Shuangfeng county[b]	1.51×10^{-3}	2.85	0.02	HH
Chaling county[c]	1.86×10^{-5}	0.88	0.04	LL
Yanling county[c]	1.99×10^{-5}	0.87	0.04	LL
Zhuhui district[d]	8.80×10^{-5}	0.99	0.03	LL
Yanfeng district[d]	5.59×10^{-5}	1.17	0.00	LL
Shigu district[d]	1.04×10^{-4}	1.10	0.02	LL
Zhengxiang district[d]	1.07×10^{-4}	1.00	0.02	LL
Hengnan county[d]	5.51×10^{-5}	0.98	0.03	LL
Hengdong county[d]	3.21×10^{-5}	1.19	0.02	LL
Sangzhi county[e]	1.52×10^{-5}	0.85	0.00	LL
Guidong county[f]	1.04×10^{-5}	0.87	0.05	LL
Jianghua county (Yao Autonomous)[g]	2.98×10^{-6}	0.69	0.03	LL
Ningyuan county[g]	-1.53×10^{-4}	-1.15	0.04	HL
Shaoshan city[h]	-1.36×10^{-4}	-3.01	0.01	LH
Dongkou county[i]	-3.22×10^{-5}	-3.01	0.03	LH
Louxing district[b]	-6.75×10^{-5}	-3.55	0.02	LH
Xinning county[i]	-2.50×10^{-5}	-1.51	0.05	LH
Zhongfang county[j]	-3.36×10^{-5}	-2.17	0.05	LH

HR: Hospitalization rates; HH: high-high pattern; LL: low-low pattern; HL: high-low pattern; LH: low-high pattern; a: belongs to Changsha city; b: belongs to Loudi city; c: belongs to Zhuzhou city; d: belongs to Hengyang city; e: belongs to Zhangjiajie city; f: belongs to Chenzhou city; g: belongs to Yongzhou city; h: belongs to Xiangtan city; i: belongs to Shaoyang city; j: belongs to Huaihua city.

STROBE Statement—Checklist of items that should be included in reports of ***cross-sectional studies***

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	P1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	P1-2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	P2
Objectives	3	State specific objectives, including any prespecified hypotheses	P2
Methods			
Study design	4	Present key elements of study design early in the paper	P2
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	P2
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants	P3
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	P3
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	*P3*
Bias	9	Describe any efforts to address potential sources of bias	P3
Study size	10	Explain how the study size was arrived at	P2
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	P3
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	P3-5
		(*b*) Describe any methods used to examine subgroups and interactions	P3-5
		(*c*) Explain how missing data were addressed	P3-5
		(*d*) If applicable, describe analytical methods taking account of sampling strategy	P3-5
		(*e*) Describe any sensitivity analyses	P3-5
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	P5
		(b) Give reasons for non-participation at each stage	N/A
		(c) Consider use of a flow diagram	N/A
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	P5
		(b) Indicate number of participants with missing data for each variable of interest	P5
Outcome data	15*	Report numbers of outcome events or summary measures	P5-7
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	P5-7

		(*b*) Report category boundaries when continuous variables were categorized	P5-7
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	N/A
Discussion			
Key results	18	Summarise key results with reference to study objectives	P7-9
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	P9-10
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	P9-10
Generalisability	21	Discuss the generalisability (external validity) of the study results	P10
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	P10

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.