

# 2018~2021年中国食源性疾病地理分布及影响因素分析

杨钰隆, 曾庆\*

重庆医科大学公共卫生学院, 重庆

收稿日期: 2024年1月20日; 录用日期: 2024年4月1日; 发布日期: 2024年4月8日

## 摘要

目的: 分析2018~2021年中国内地食源性疾病的时空分布特征, 为预防和控制提供相应的理论依据。方法: 基于2018~2021年《中国卫生健康统计年鉴》, 采用空间自相关分析、时空扫描等方法, 研究中国内地居民食源性疾病的空间分布格局及其变化。结果: 国内内地报告的食源性疾病在2.71~3.24/10万。全局空间自相关分析结果显示仅2020年食源性疾病发病率存在空间自相关。时空扫描结果显示在2018~2021期间有三个时空聚集区。公共卫生间建设数量、气温对食源性疾病呈正相关; 居民个人食品支出、健康支出对食源性疾病呈现负相关。地理加权回归(Geographic Weighted Regression, GWR)模型能够更好地解释模型效果。结论: 对食源性疾病采用空间聚集性分析能够较好地探测出食源性疾病的聚集点。对开展有关防控措施提供依据。

## 关键词

食源性疾病, 空间自相关, 时空扫描, 回归分析, 地理加权回归

# Analysis on the Geographical Distribution and Influencing Factors of Foodborne Diseases in China, 2018~2021

Yulong Yang, Qing Zeng\*

School of Public Health, Chongqing Medical University, Chongqing

Received: Jan. 20<sup>th</sup>, 2024; accepted: Apr. 1<sup>st</sup>, 2024; published: Apr. 8<sup>th</sup>, 2024

\*通讯作者。

文章引用: 杨钰隆, 曾庆. 2018~2021年中国食源性疾病地理分布及影响因素分析[J]. 统计学与应用, 2024, 13(2): 297-306. DOI: 10.12677/sa.2024.132030

## Abstract

**Objective:** To analyze the spatial-temporal distribution characteristics of foodborne diseases in mainland of China from 2018 to 2021, and provide a theoretical basis for the prevention and control of foodborne diseases. **Methods:** Based on the *China Health and Wellness Statistical Yearbook* from 2018 to 2021, spatial autocorrelation analysis and space-time scanning were used to study the spatial distribution pattern of foodborne illnesses and their changes in mainland China residents. **Results:** The reported foodborne diseases in China ranged from 2.71 per 100,000 to 3.24 per 100,000. The results of global spatial autocorrelation analysis showed that only the incidence of foodborne diseases in 2020 had spatial autocorrelation. The space-time scanning results showed that there were three spatial-temporal clusters from 2018 to 2021. The temperature has an impact on the increase in foodborne diseases. The number of public toilets and temperature were positively correlated with foodborne diseases. Personal food expenditure and health expenditure were negatively correlated with foodborne diseases. The geographic weighted regression (GWR) model could better explain the model effect. **Conclusion:** Spatial clustering analysis could better detect the clustering points of foodborne diseases, which can provide evidence for prevention and control measures.

## Keywords

Foodborne Diseases, Spatial Autocorrelation, Space-Time Scanning, Regression Analysis, Geographic Weighted Regression

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

食源性疾病是指因摄入由细菌、病毒、寄生虫或化学物质污染的食物或水引起的疾病[1]。食源性疾病全球一大公共卫生问题,也是影响我国食品安全的主要因素,据世界卫生组织报告每年食源性疾病高达6亿例[2],食源性疾病也是发达国家一大重要卫生问题,据估计仅在美国每年就有7600万例食源性疾病病例导致325,000人住院和5000人死亡[3]。而我国同样受到食源性疾病的困扰,国内食源性疾病的扩散性和突发性趋势更加明显[4]。

近年来,随着食源性疾病监测系统的完善,以及食源性疾病快速诊断能力的提高,食源性疾病致病因子复杂繁多,制定有效的公共卫生措施对食源性疾病的防控更加重要[5][6][7]。通过分析食源性疾病的空间特征,对防治控制食源性疾病的发展,制定相关措施有重要意义。

## 2. 资料与方法

### 数据来源和研究区域

2018~2021年期间食源性疾病报告情况数据来源于《中国卫生统计年鉴》。气象和卫生环境资料来源于《中国统计年鉴》。研究区域包含中国内地31个省份和直辖市。需要注意的是西藏数据存在缺失,故在研究中剔除。

### 研究方法

#### 空间自相关分析:

空间自相关分析是研究空间单元的属性值是否与相邻空间单元的属性值相关, 揭示研究区域属性值的空间分布类型的统计方法。它被分为全局它被分为全局空间自相关和局部空间自相关, 前者用于探索整个区域的空间依赖性, 后者用于进一步确定具有空间自相关的局部格局和地点。全局指数(范围: -1 到 1)用于全局空间自相关。当和时存在正空间自相关, 当和时存在负空间自相关[8] [9]。在空间自相关聚集图的局域指标中, 局域县区之间的空间关联模式可分为 4 类: 高-高聚集、低-低聚集、低-高离群值和高-低离群值。值得注意的是, 高-高聚集区是戊型肝炎传播和发生的重要“热点”[10]。空间权重矩阵采用一阶 QUEEN 准则。

#### 空间和时空聚类分析:

基于离散泊松概率模型, 进行纯空间和时空扫描分析, 以发现在区/县一级具有高比率的聚类。空间和时空扫描统计量分别由对应于地图上地理区域的圆形窗口和具有圆形地理基底且高度对应于时间的圆柱形窗口定义。将扫描窗口在空间和(或)时间上进行移动, 得到无数个大小和形状不同的重叠窗口, 共同覆盖整个研究区域。每个圆或圆柱都是一个可能的候选集群。最大地理圆大小设置为总人口的 25%, 和最大时间扫描周期设置为整个研究周期的 50%。默认情况下, 不会报告地理位置重叠的集群。对于每个扫描窗口, 零假设是窗口内外的风险是相同的, 备择假设是窗口内的风险升高。采用对数似然比(LLR)作为检验统计量, 值采用 999 个重复的蒙特卡罗假设检验, 得到最大的窗口作为最有可能的聚类, 以及在其他窗口中具有统计学意义的为次要聚类。并按 LLR 值进行排序。当取最大值且时, 可以认为该区域存在聚集特征[11] [12]。

#### 逐步回归

在回归分析中, 普通最小二乘法(OLS)是估计应变量与自变量之间线性回归的传统方法。简单 OLS 是对  $y_i$  和  $x_i$  两个变量之间的线性关系的估计, 公式如下:

$$y = a + \beta x_i + u_i \quad i = 1, 2, \dots, n \quad (1)$$

其中  $y_i$  表示因变量  $y$  的第  $i$  个观测值, 即食源性疾病发病率。  $x_i$  表示自变量  $x$  的第  $i$  个观测值。我们通过收集相关省份的气象和卫生环境资料, 剔除共线性较高的变量(VIF > 10)。研究影响食源性疾病病例数升高的因素(表 1)。采用逐步回归的逐步选择策略, 依次添加最有贡献的预测变量, 从而选择出最优的线性回归模型。

**Table 1.** Summary statistics of meteorological factors and health environment in China, 2018~2021

**表 1.** 2018~2021 年全国气象因子和卫生环境基本汇总统计说明

变量	变量解释	均值	方差	最小值	最大值
Case	病例数(log(n + 1))	6.641	1.044	3.892	8.958
Public lavator	公共卫生间数量(log(n + 1))	8.372	0.7373	6.568	9.635
Garbage disposal	垃圾无害化处理量(万吨) (log(n + 1))	6.413	0.7345	4.657	8.115
Humidity	年平均湿度(%) (log(n + 1))	4.183	0.1755	3.807	4.431
Sunlight	年日照时常(小时) (log(n + 1))	7.593	0.2861	6.755	8.138
Precipitation	年降水量(毫米) (log(n + 1))	6.707	0.6421	4.980	7.808
Temperature	年平均气温(摄氏度) (log(n + 1))	2.628	0.4033	1.629	3.250
Expenditure for food	食品支出(元) (log(n + 1))	8.914	0.2092	8.456	9.463

续表

Health expenditure	医疗保健支出(元) (log(n + 1))	7.727	0.2241	7.107	8.436
Flood control area	除涝面积(公顷) (log(n + 1))	5.204	2.3247	0	8.418

### 地理加权回归

地理加权回归的模型的一般形式如下:

$$y_i = \beta_0(u_i, v_i) + \beta_1(u_i, v_i)x_{i1} + \beta_2(u_i, v_i)x_{i2} + \dots + \beta_p(u_i, v_i)x_{ip} + \varepsilon_i \quad (2)$$

$$i = 1, 2, \dots, n; \quad j = 1, 2, \dots, p$$

其中,  $y_i$  与  $x_{i1}, x_{i2}, \dots, x_{ip}$  是因变量  $y$  与自变量  $x_{i1}, x_{i2}, \dots, x_{ip}$  在  $(u_i, v_i)$  处的观测值,  $(u_i, v_i)$  为第  $i$  个采样点的坐标,  $\beta_k(u_i, v_i)$  是第  $i$  个采样点上的第  $k$  个回归参数, 点  $i$  被称为回归点.  $\varepsilon_i \sim N(0, \sigma^2)$ , 当  $i \neq j$  时,  $Cov(\varepsilon_i, \varepsilon_j) = 0$  [13]。可以简便记为:

$$y_i = \beta_{i0} + \sum_{k=1}^p \beta_{ik} x_{ik} + \varepsilon_i \quad (3)$$

$$i = 1, 2, \dots, n$$

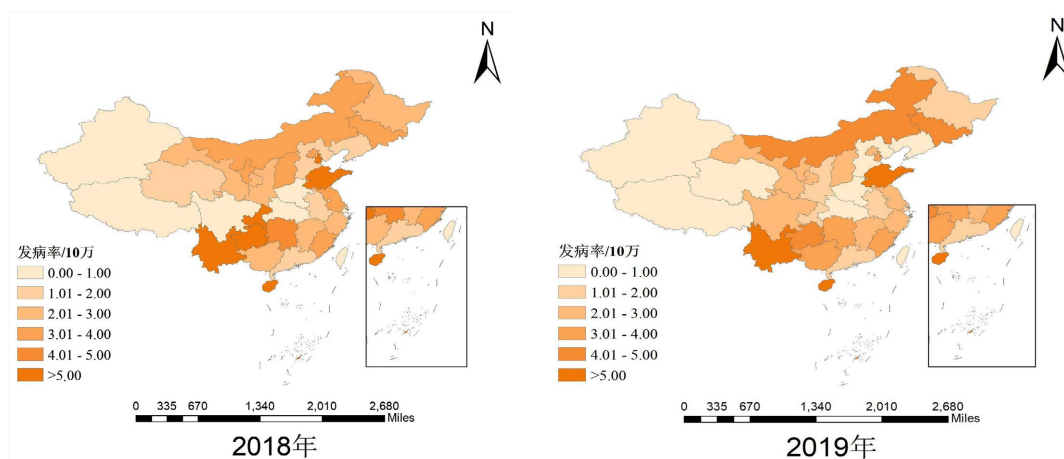
地理加权回归会得到出  $n$  个回归方程, 对应每个回归点  $i$ , 都会有一个回归方程, 若  $\beta_{1k} = \beta_{2k} = \dots = \beta_{nk}$ , 则地理加权回归模型退化为普通线性回归模型。

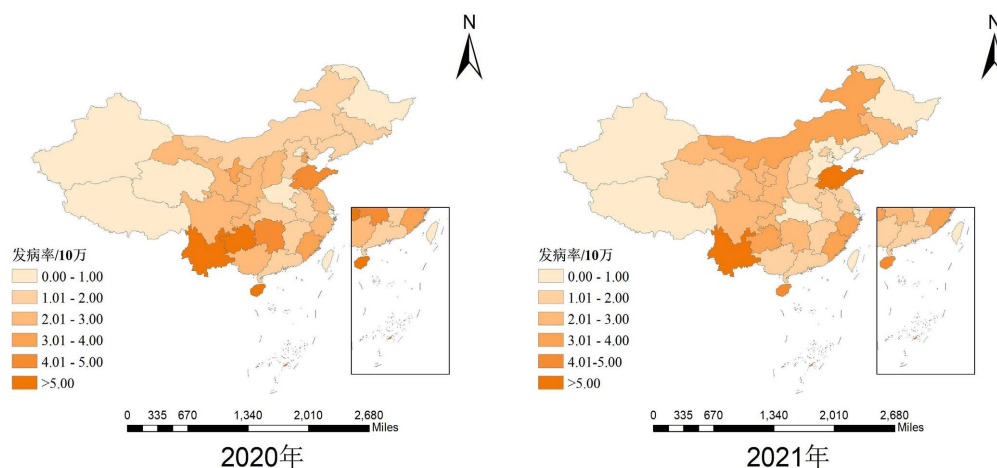
统计软件: 采用 Geoda1.12 进行全局空间自相关和局部空间自相关, SaTScan 10.1 软件(Martin Kulldor, National Cancer Institute, Bethesda, MD, USA)进行时空聚类分析, R 软件进行逐步回归, 分析结果通过 arcgis10.2.2 (ESRI 公司, Redlands, CA, USA)进行地理加权回归(Geographic Weighted Regression, GWR)以及结果展示。

## 3. 结果

### 流行病学特征

2018~2021 年中国内地共累计报告食源性疾病事件 25,493 起, 发病 150,335 人。年均报告发病率 2.75/10 万人, 发病率从 3.24/10 万人下降到 2.28/10 万人。发病率的空间分布间图如下图, 发病率较高的省份主要集中在西南地区如贵州省、云南省、东部地区主要是山东省、北部内蒙古自治区食源性疾病发病率较高(图 1)。





注: 该图基于自然资源部标准底图服务网站下载的审图号为 GS (2022) 1061 号的标准地图制作, 底图边界无修改。

Figure 1. Spatial distribution of foodborne disease incidence from 2018 to 2021

图 1. 2018 年至 2021 年食源性疾病发病率空间分布

### 全局空间自相关

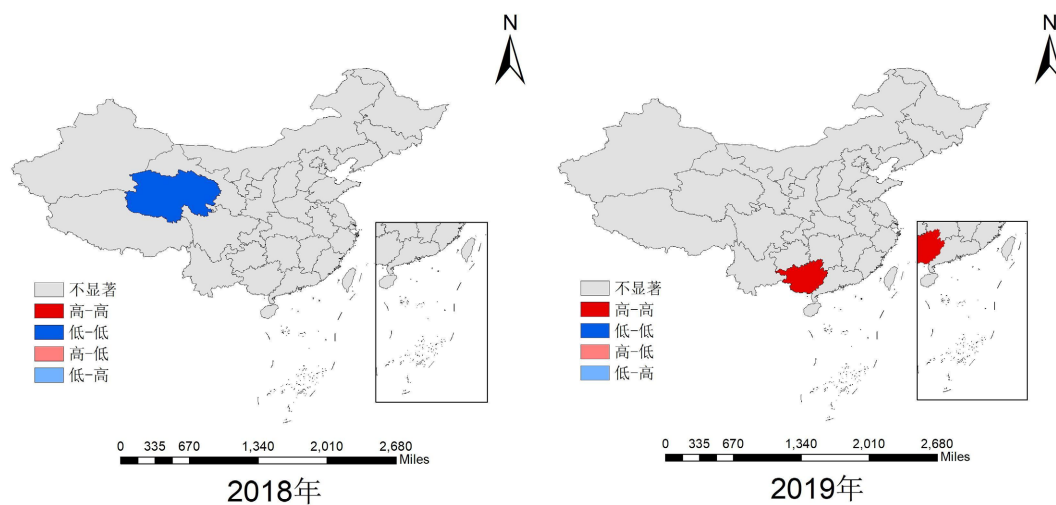
分析结果显示 2018~2021 年期间, 仅 2020 年食源性疾病发病率在空间上具有全局自相关性(表 2)。因此, 2020 年存在一个正的空间自相关, 其他年份上的食源性疾病的空间分布总体上是随机的。

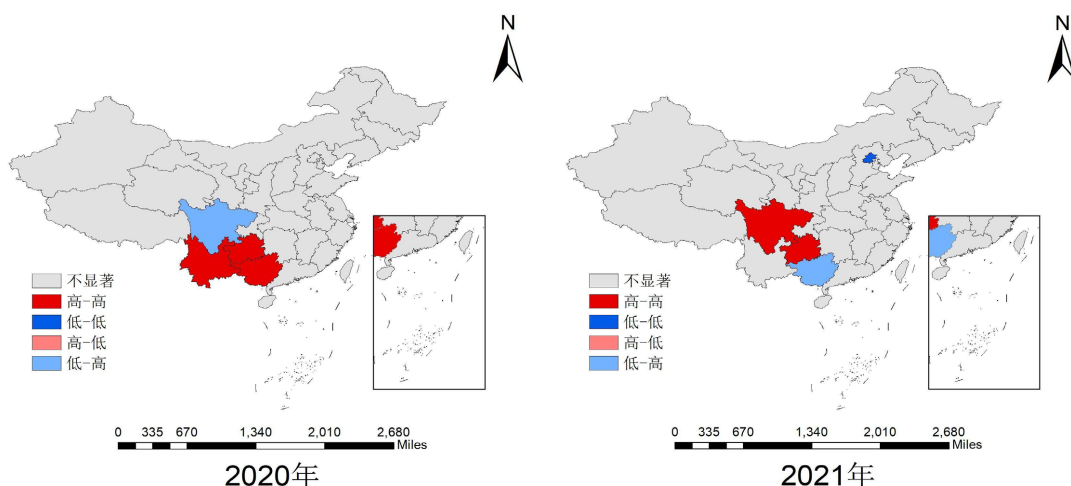
Table 2. Global spatial autocorrelation of incidence in China, 2018~2021

表 2. 2018~2021 年发病率的全局空间自相关

年份	2018	2019	2020	2021
Moran'I	-0.023	0.029	0.182	0.035
P-value	0.463	0.251	0.025	0.23

通过对我国内地省、直辖市的食源性疾病发病率的局部空间自相关分析, 确定了具有统计学意义的局部空间聚类(图 2)。四个空间集群的位置和区域每年不完全相同, 高高地区主要集中在西南地区, 主要包括贵州省、云南省和广西壮族自治区。



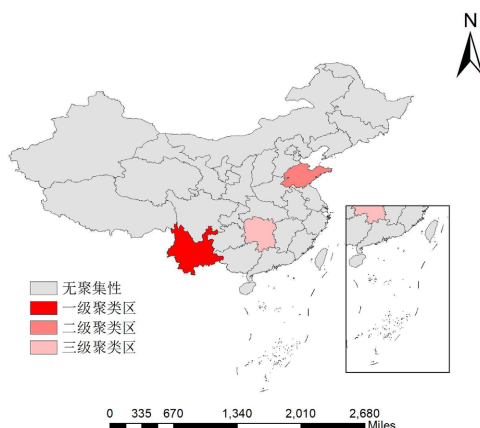


注：该图基于自然资源部标准底图服务网站下载的审图号为 GS (2022) 1061 号的标准地图制作，底图边界无修改。

**Figure 2.** Local spatial autocorrelation cluster maps of foodborne diseases in China, 2018~2021  
**图 2.** 2018~2021 年食源性疾病发病率的局部空间自相关图

### 空间聚类分析

分别对 2018~2021 年各省的食源性疾病报告情况进行纯空间扫描。从 2018 年~2021 年，使用时空扫描检测到三个聚类区(图 3)。其中食源性疾病时空聚集性最强的为 2019~2020 年云南省(RR = 5.30, LLR = 10339.05,  $P < 0.001$ )，其次为 2018 年~2019 年湖南省(RR = 2.88, LLR = 5572.68,  $P < 0.001$ )和 2020 年山东省(RR = 1.74, LLR = 391.52,  $P < 0.001$ ) (表 3)。



注：该图基于自然资源部标准底图服务网站下载的审图号为 GS (2022) 1061 号的标准地图制作，底图边界无修改。

**Figure 3.** Spatiotemporal clustering of foodborne diseases in China, 2018~2021  
**图 3.** 2018~2021 年食源性疾病的时空聚类图

**Table 3.** Results for spatial clusters of foodborne diseases in China, 2018~2021  
**表 3.** 2018~2021 年食源性疾病时空聚类结果

	集群时间	县(n)	预期观察值	半径(km)	相对风险	对数似然比	P 值
一级聚类	2019/01~2020/12	1	4.94	0	5.30	10339.05	<0.001
二级聚类(I)	2018/01~2019/12	1	2.70	0	2.88	5572.68	<0.001
二级聚类(II)	2020/01~2020/12	1	1.73	0	1.74	391.51	<0.001

### 逐步回归

回归分析结果显示, 公共卫生间数量、温度对食源性疾病报告人数的增加存在相关性; 而食品支出、医疗保健指出对食源性疾病报告人数的增加存在负相关性(表 4)。

**Table 4.** Stepwise regression results

**表 4.** 逐步回归结果

	系数	标准误	t 值	Pr (> t )
截距	15.9890	3.4139	4.684	<0.001
Public lavator	0.4706	0.1356	3.471	<0.001
Temperature	1.2336	0.2520	4.894	<0.001
Expenditure for food	-1.0870	0.4348	-2.500	<0.05
Health expenditure	-0.9377	0.3916	-2.395	<0.05
Flood control area	0.0607	0.0416	1.459	0.1474

### 地理加权回归(Geographic weighted regression)

我们采用地理加权回归(Geographic Weighted Regression, GWR)以及时空地理加权回归(Geographically and Temporally Weighted Regression, GTWR)模型来考察戊型肝炎与其因子之间是否存在时空关系。我们认为食源性疾病的发病率在研究区域内存在聚类性和变化性。因此 GWR 模型中的自适应核是合适的(the adaptive kernel in GWR models is appropriate)。通过采用 AICc (Akaike information criterion)方法来找到使 AICc 值最小的带宽。模型的 AICc 值越低, 观测数据提供的拟合度就越好。如表 5 所示, OLS, GWR, GTWR 的  $R^2$  值分别为 0.52、0.79、0.77。对应的 AICc 值为 273.18、261.07、279.19。综合评定结果认为地理加权回归(GWR)相对于逐步回归和时空地理加权回归(GTWR)有着更好的效果(表 5)。

**Table 5.** Regression analysis information

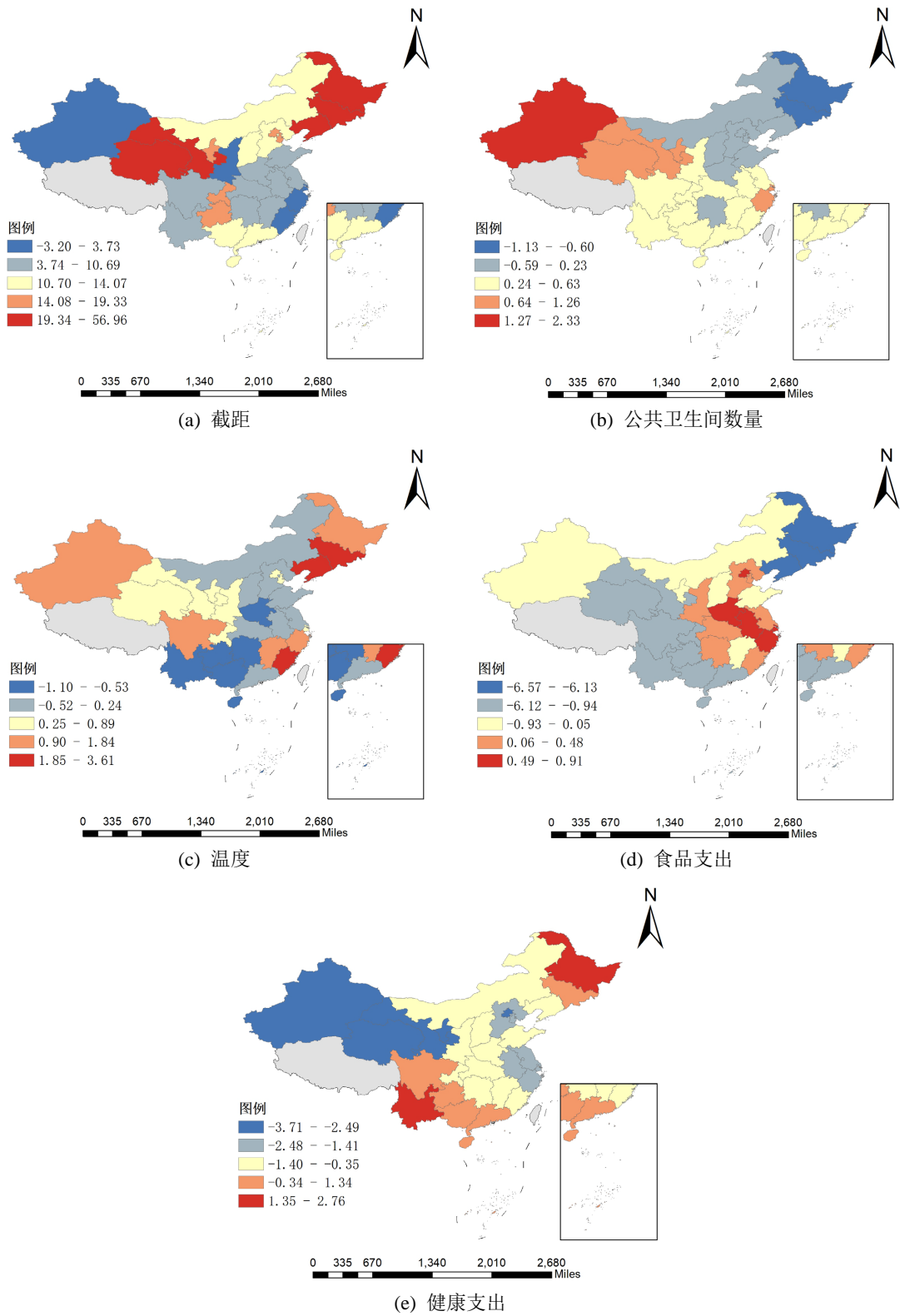
**表 5.** 回归分析信息

模型	AICc	$R^2$	Adjust $R^2$	$\sigma^2$
OLS	273.18	0.52	0.50	0.7371
GWR	261.07	0.79	0.78	0.4732
GTWR	279.19	0.77	0.76	0.5029

对于发病率的空间变化, 我们通过 arcgis 的 GWR 模型进行分析, 生成的结果采用自然间断点分级法(Jenks)分为 5 个类。图 4 包含了基于逐步回归中五个具有意义的印象因素的系数空间分布图。可以看到各省和直辖市均有差异。

1) 公共卫生间建设数量对食源性疾病影响的空间变异特征。从回归系数的空间分布(图 4(b))来看, 回归系数在中国北部地区由东向西部逐渐增大, 而中国中部和南部地区系数值均为正且系数绝对值较小。这可能是由于受北部气候不同的影响, 在东北部地区建立公共卫生间将有利于提高粪便管理以减少食源性疾病的污染和传播, 而西部地区则因减少或规范不合格公共卫生间。

2) 温度对食源性疾病影响的空间变异特征。从回归系数的空间分布(图 4(c))结果显示, 西南地区温度与食源性疾病呈现负相关性, 而沿海地区如广东、福建、辽宁和吉林受温度影响较大, 这可能原因是夏季海产品等易发生变质, 副溶血性弧菌等细菌最适温度在  $30^{\circ}\text{C}\sim 37^{\circ}\text{C}$ , 在夏季更易繁殖从而引起食源性疾病的发生[14] [15]。



注：该图基于自然资源部标准底图服务网站下载的审图号为 GS (2022) 1061 号的标准地图制作，底图边界无修改。

**Figure 4.** Geographically weighted model regression spatial distribution coefficient  
**图 4.** 地理加权模型回归空间分布系数



3) 食品支出对食源性疾病影响的空间变异特征。从回归系数的空间分布(图 4(d))结果显示, 在国内沿海较发达地区, 增加食品支出对减少食源性疾病的发生效果不大。而在内陆地区, 提高居民食品支出对降低食源性疾病病例数量具有较高的意义。因提高食品质量, 提高内陆地区居民对食品安全的意识、减少通过不合理、不规范的家庭自制食品所引起的食物中毒。

4) 在健康卫生方面, 除黑龙江, 云南两地。全国绝大部分地区提高医疗卫生支出将有利于减少食源性疾病的发生, 尤其以西北部地区显著。这表明西部地区医疗发展水平相对落后, 因提高食源性的疾病治疗水平。

## 4. 讨论

2018~2021 年中国内地共累计报告食源性疾病事件 25,493 起, 发病 150,335 人。年均报告发病率 2.75/10 万人。国内食源性疾病在 2018~2021 年期间发病率逐渐下降, 这可能与加强食品监管的提升有着重要的关系。

食源性疾病的时空分布探索是深入了解我国食源性疾病流行状况和趋势的重要手段。通过采用自相关分析, 来识别疾病在地理空间中存在的模式和趋势, 从而发现存在的聚集或分散趋势[8]。时空扫描统计可以对疾病数据进行危险性分析, 近年来, SaTScan 可帮助确定地理区域和时间范围内的疾病聚类[16]。通过扫描可以识别聚类区域, 了解疾病的地理分布的变化趋势。通过识别异常高发地区, 为疾病监测和防控提供重要信息, 从而可以由针对性地分配医疗资源, 以更有效地管理疾病风险。空间流行病学在多个领域都取得了显著的成就, 为食源性疾病的研究提供了新的视角和方法[17] [18] [19]。我们的研究结果显示全国各地区之间的食源性疾病发病率在 2018, 2019 和 2020 年没有空间相关性。而 2020 年存在空间相关性。

食源性疾病的时空分布受多种因素的影响, 包括地理环境、气候条件、人口密度、饮食习惯等[20]。我们的研究显示公共卫生间数量、温度对食源性疾病的发生具有正相关性, 可能是公共卫生间的卫生管理尚不完善, 蚊蝇滋生等情况易导致食源性疾病的传播, 而温度的升高容易使食物发生变质、以及细菌滋生。食品支出以及卫生健康支出对食源性疾病的发生具有负相关性, 可能的原因是健全食品卫生质量, 加强市场监督管理以及居民食品健康意识的提高从而降低了食源性疾病的发生。而医疗支出的提高与居民健康意识同样息息相关, 及时有效地治疗进一步避免了食源性疾病的传播。此外, 我国幅员辽阔, 包含热带季风气候、亚热带季风气候、温带季风气候、温带大陆性气候、高原山地气候。同时各地区饮食习惯差异较大。对食源性疾病的发病类型进一步研究将有利于减少食源性疾病的发生。

对于食源性疾病的流行现状, 我们建议: (一) 建立健全的食源性疾病监测系统。(二) 加强健康教育, 开展多种形式的食品安全卫生宣传教育工作。(三) 加强食品监管, 提高食品生鲜市场的监管力度。尤其是对食源性高发季节以及人口密集场所的食品安全监管。(四) 根据不同地区的饮食习惯, 生活习俗以及气候等因素, 针对性的制定公共卫生策略。

## 参考文献

- [1] 杨艳丽, 梁小勇, 韩宗辉. 2017~2021 年西安市食源性疾病监测结果的流行特征分析[J]. 食品安全导刊, 2023(33): 46-48.
- [2] 世界卫生组织. 世卫组织加强行动, 以改善食品安全, 保护人们免受疾病侵害[Z]. 2021.
- [3] Newell, D.G., Koopmans, M., Verhoef, L., et al. (2010) Food-Borne Diseases—The Challenges of 20 Years Ago Still Persist While New Ones Continue to Emerge. *International Journal of Food Microbiology*, **139**, S3-S15. <https://doi.org/10.1016/j.ijfoodmicro.2010.01.021>
- [4] 曾梦晴, 张壮, 陈有华. 中国食源性疾病省际时空差异与成因研究[J]. 美食研究, 2023, 40(4): 28-36.
- [5] 陆姣, 王晓莉, 吴林海. 国内外食源性疾病防控的研究进展[J]. 中华疾病控制杂志, 2017, 21(2): 196-199.

- [6] 张智芳, 廖冬冬. 1985~2022 年国内食源性疾病研究热点和前沿趋势——基于 CiteSpace 可视化分析[J]. 公共卫生与预防医学, 2023, 34(4): 21-25.
- [7] 陈婷, 周庆琼, 戚平, 等. 2011~2020 年中国食源性疾病流行病学分析[J]. 中国食品卫生杂志, 2023, 35(10): 1545-1550.
- [8] Anselin, L., Syabri, I. and Kho, Y. (2006) *GeoDa: An Introduction to Spatial Data Analysis*. *Geographical Analysis*, **38**, 5-22. <https://doi.org/10.1111/j.0016-7363.2005.00671.x>
- [9] Anselin, L. and Rey, S.J. (2014) *Modern Spatial Econometrics in Practice: A Guide to GeoDa, GeoDaSpace and PySAL*. GeoDa Press, Chicago.
- [10] Anselin, L. (1995) Local Indicators of Spatial Association—LISA. *Geographical Analysis*, **27**, 93-115. <https://doi.org/10.1111/j.1538-4632.1995.tb00338.x>
- [11] Kulldorff, M. (2022) SaTScan™ User Guide. <https://www.satscan.org>
- [12] Wangdi, K., Penjor, K., Tobgyal, *et al.* (2021) Space-Time Clustering Characteristics of Malaria in Bhutan at the End Stages of Elimination. *International Journal of Environmental Research and Public Health*, **18**, Article 5553. <https://doi.org/10.3390/ijerph18115553>
- [13] 王增铮, 张福浩, 赵阳阳, 等. 区域地理加权回归分析方法[J]. 测绘通报, 2023(12): 81-87.
- [14] 刁文丽, 王凯琳, 宋蕴奇, 等. 2014~2019 年辽宁省食源性疾病流行病学分析[J]. 中国食品卫生杂志, 2021, 33(4): 451-455.
- [15] 张丽萍, 董莉萍, 王玉娜. 2014~2015 年吉林市食源性疾病流行特征和影响因素分析[J]. 铁路节能环保与安全卫生, 2017, 7(1): 43-46.
- [16] Kulldorff, M., Heffernan, R., Hartman, J., *et al.* (2005) A Space-Time Permutation Scan Statistic for Disease Outbreak Detection. *PLOS Medicine*, **2**, e59. <https://doi.org/10.1371/journal.pmed.0020059>
- [17] 岳玉娟, 刘小波, 任东升, 等. 中国大陆 2005~2020 年登革热流行病学特征分析[J]. 中国媒介生物学及控制杂志, 2023, 34(6): 761-766, 818.
- [18] 陈艳伟, 李旭, 李夫, 等. 2013~2022 年北京市人间布鲁氏菌病流行特征和空间分布分析[J]. 疾病监测, 2023, 38(10): 1168-1173.
- [19] 陈丽敏. 基于监测医院报告病例的江西省食源性疾病时空特征及发病趋势预测研究[D]: [硕士学位论文]. 南昌: 南昌大学, 2023.
- [20] 张雪佩, 滕艾颖, 王珊珊, 等. 气象因素对食源性疾病发病情况影响的研究[J]. 公共卫生与预防医学, 2024, 35(1): 45-48.