

Time Series Analysis of the Number of Deaths from Respiratory Diseases in Karamay, China from 2014 to 2019

Yun Huang¹ Weimin Zhu¹ Shuaiyin Zheng² Jian Liu² Peidi Chen^{2*}

1. Karamay City Center for Disease Control and Prevention, Karamay, Xinjiang, 834000, China

2. Xinjiang Second Medical College, Karamay, Xinjiang, 834000, China

Abstract

Objective: To analyze and predict the characteristics of deaths from respiratory diseases in Karamay from 2014 to 2019, so as to provide reference for the prevention and treatment of diseases. **Methods:** Use ARIMA mathematical model to analyze the trend of mortality. **Results:** ARIMA (0,0,0) (2,1,0) is the optimal prediction model; the number of monthly deaths during 2014-2019 was mainly higher in January, February, November and December, and the monthly deaths decreased in line with the trend of the predicted value. **Conclusion:** ARIMA (0,0,0) (2,1,0) has some guidance for short-term prediction analysis of respiratory diseases in Karamay city; the number of deaths tends to decrease year by year, and the prevention and treatment work needs to be strengthened in autumn and winter.

Keywords

respiratory diseases; number of deaths; ARIMA mode prediction

2014—2019年中国克拉玛依市呼吸系统疾病死亡数的时间序列分析

黄芸¹ 朱伟敏¹ 郑帅印² 刘健² 陈佩弟^{2*}

1. 克拉玛依市疾病预防控制中心, 中国·新疆 克拉玛依 834000

2. 新疆第二医学院, 中国·新疆 克拉玛依 834000

摘要

目的: 对2014—2019年克拉玛依市因呼吸系统疾病死亡人数特征进行分析预测, 为疾病的防治提供参考依据。**方法:** 运用ARIMA数学模型, 分析疾病死亡人数趋势变化特征。**结果:** ARIMA (0, 0, 0) (2, 1, 0) 为最优预测模型; 2014—2019年期间的疾病月死亡人数主要在每年的1、2、11、12月份较多, 月死亡值逐年呈下降趋势与预测值变化趋势相符。**结论:** ARIMA (0, 0, 0) (2, 1, 0) 对克拉玛依市呼吸系统疾病的短期预测分析具有一定的指导性; 死亡数逐年趋于减少趋势, 秋冬季节仍需要继续加强防治工作。

关键词

呼吸系统疾病; 死亡数; ARIMA模型预测

1 引言

呼吸系统疾病列为全球4大慢性非传染性疾病之一^[1], 对人群健康带来影响^[2], 成为慢性病防治工作的重要组成部

【基金项目】新疆维吾尔自治区高校科研计划项目(项目编号: XJEDU2022P147); 新疆第二医学院青年科学基金项目(项目编号: QK202211)。

【作者简介】黄芸(1989-), 女, 中国湖南宁远人, 本科, 讲师, 从事慢性病流行病学研究。

【通讯作者】陈佩弟(1996-), 男, 中国山东单县人, 硕士, 讲师, 从事慢性病预防与控制研究。

分。据世界卫生组织统计显示, 在2019年的全球人口中, 每年约超过4100万人死因归于慢性疾病, 至2030年预计将超过5500万人^[3], 将进一步加重居民疾病负担。在中国患呼吸系统疾病的人数在患病人群中占比极高, 由于慢性呼吸系统疾病引发的死亡人数占整体的11%^[4], 成为导致居民死亡的重要原因之一^[5]。所以, 应加强疾病动态监测机制。近年来, 呼吸系统疾病成为影响克拉玛依市居民健康因素之一^[6,7]。研究表明, 该市居民慢性病死亡占疾病总死亡的88.85%^[8]。因此, 开展克拉玛依市呼吸系统疾病的动态监测, 在推动“健康新疆、健康克拉玛依”的战略举措背景下, 具有重要的现实意义。ARIMA模型^[9]常被用于分析季节性变化序列并可对序列的未来发展趋势进行有效预测。该模型也常被应用于医学领域^[10], 对疾病的防治颇有成效。

本研究利用 ARIMA 模型,了解克拉玛依市呼吸系统疾病死亡流行特征及变化趋势,以期为制定疾病预防和干预提供依据。

2 资料与方法

2.1 资料来源

2014—2019 年呼吸系统疾病死亡资料由克拉玛依市疾病预防控制中心提供。

2.2 方法

2.2.1 原序列平稳性检验

运用增广迪基—富勒检验方法,评估初始序列的差异、走势以及季节变化。

2.2.2 原序列的平稳化处理

若 $P < 0.05$,推断经过差分处理后的时间序列表现出稳定性。若 $P > 0.05$ 时,需要对时间序列再次进行差分处理,直至序列平稳。

2.2.3 最优模型识别与参数估计

运用最佳准则函数定阶方法与博克思 - 詹金斯定阶方法相融,对平稳序列 X_t 进行辨识。根据准则函数值的大小进行评价,得出函数值最小的模型,即为原序列的最佳模型。

基于时间序列 X_t 经过平稳化处理后的数据,推出自相关序列图(自相关函数, ACF)和偏自相关序列图(偏自相关函数, PACF)。然后,挑选最合适的 ARIMA (p, d, q) (季节性部分: P, D, Q, s) 模型,在选取 p 和 q 的时候,需要运用最小二乘法进行参数估计。

2.2.4 模型与参数显著性检验

若原序列中提取几乎全部的信息,认定是最佳模型;假设残差次序呈现与白噪声不同特性,则信息尚未彻底提取,指示难以进行有效预测分析。

如果经过检验统计量具有显著性 ($P < 0.05$),表明该参数与独立变量以及因变量之间存在明显影响,应在拟合模型中予以保留。

2.2.5 模型优化与预测

经过参数显著性检验后,若拟合模型参数显著,有效拟合观察值序列波动,并且需要使用施瓦茨贝叶斯准则 BIC,最小信息量准则 AIC 和经过改良后的 Akaike 信息准则 (AICc) 来进行确定,在模型的各项取值中,数值最小则为最优模型。对经过优化的最佳模型进行验证,一旦经过平均绝对百分比误差 MAPE 的评估,便可进行短期预测。MAPE 计算公式为:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{|W_t - \hat{W}_t|}{W_t} \times 100\%$$

其中, W_t 为呼吸系统死亡数; \hat{W}_t 为 t 时刻模型预测的死亡数; n 为观测值数量。

2.3 统计学分析

运用 EpiData 软件,构建 2014—2019 年间克拉玛依市

呼吸系统疾病每月死亡数据库,使用 R 统计软件中“tseries”以及“forecast”软件包进行模型构建,运用最小二乘法进行模型参数估计,并以 t 检验评估参数显著性, $\alpha = 0.05$ 。

3 结果

3.1 时序图与季节指数

2014—2019 年克拉玛依市居民呼吸系统疾病死亡数整体呈下降趋势,在 1、2、11、12 月其季节指数均呈现上升趋势,展现季节性变化。在 2017 年,呼吸系统疾病的死亡人数峰值较高,见图 1。



图 1 克拉玛依市居民呼吸系统疾病死亡数时序图

3.2 白噪声检验

对序列白噪声检验, $P < 0.01$,差异具有统计学意义,说明为非白噪声序列,具有研究意义。

3.3 平稳性检验

采用 ADF 单位根检验法,对原序列进行检验,得出 $P=0.0145 < 0.05$,说明原序列为平稳序列,不需进行差分处理。

3.4 模型识别与参数估计

随着阶数增大,自相关序列图趋近于 0,表现为拖尾特征,偏自相关序列图表现为截尾特征。

运用枚举法,列出所有可能模型。选取 AICc 指标值最小为最佳模型,最终确定 ARIMA (0, 0, 0) (2, 1, 0) [12] 为最佳模型。针对各项参数进行显著性检验,推断该模型内的参数差异具有统计学意义,系数均显著且不为零。对于此残差序列进行白噪声验证,检测结果显示 $P=0.1788 > 0.05$,认为在统计学上不显著。残差序列的平均值在置信区间 0 的范围内波动,经过白噪声检验,得出模型残差呈现白噪声特性,表明原始序列信息已充分提取,确定该序列最优模型为 ARIMA (p, d, q) (P, D, Q) s, 为 ARIMA (0, 0, 0) (2, 1, 0) [12], 其中 AIC=308.67, AICc=307.91, BIC=314.95。

3.5 参数显著性检验

遵循 $\alpha=0.05$ 进行检验,得出 SAR (1)、SAR (2) 中 t 值分别为 7.20、3.47, P 值均小于 0.01,表明所有系数均具有显著性,该参数与其对应的自变量与因变量之间存在明显关联。

3.6 模型拟合与预测

根据 2014—2019 年克拉玛依市居民呼吸系统每月死亡数,采用最优模型拟合分析,并与 2019 年实际值对比,显

示实际值与预测值之间的绝对误差和相对误差非常接近，相对误差较小，MAPE=16.67%，说明该模型拟合和预测效果相当良好，可以采用该模型，见图2。利用最优模型，对2020年进行预测，发现每年1、2、11、12月死亡数较其他月份高，见表1。

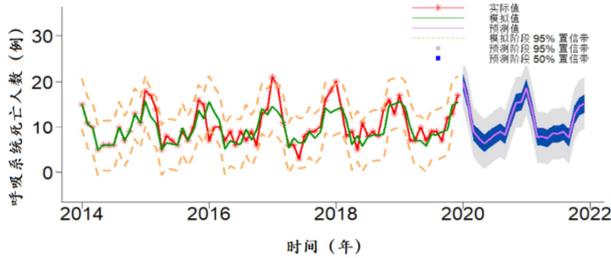


图2 克拉玛依市居民呼吸系统疾病死亡数模型拟合及预测结果

表1 克拉玛依市呼吸系统疾病死亡人数实际值与预测值及其95%置信区间

时间(年-月)	点估计(人)	95%置信区间
2020-1	19.77	14.17~25.36
2020-2	15.66	10.06~21.25
2020-3	8.91	3.32~14.51
2020-4	7.48	1.89~13.08
2020-5	6.38	0.78~11.98
2020-6	7.22	1.62~12.82
2020-7	8.20	2.61~13.80
2020-8	9.00	3.40~14.60
2020-9	8.17	2.57~13.77
2020-10	12.11	6.51~17.71
2020-11	15.39	9.80~21.00
2020-12	15.66	10.07~21.26

4 讨论

本研究确定ARIMA(0, 0, 0)(2, 1, 0)₁₂模型适用于该市呼吸系统疾病每月死亡人数的短期预测。2020年短期预测结果与2014—2019年情况相符，呈现季节性周期变化，秋冬季节是高发时段，增加居民疾病负担，可能因为克拉玛依市在秋冬季节天气较为寒冷干燥，导致人体呼吸系统对排除细菌和病毒能力减弱，使有害物质在呼吸道内停留并繁殖，降低机体疾病的抵抗力，且对于婴幼儿及老年群体较为敏感，例如：慢性气流受限性肺疾病和其他呼吸系统疾病^[11-13]。在秋冬季节，浮尘增多，取暖时间延长，若居民在户外暴露时间加长^[14]，导致身体更容易吸附各类有害微粒，进而导致暴露风险上升。因此，在户外空气质量不佳时，宜减少外出活动时间，外出时建议佩戴口罩，以降低有害物质吸入的可能性，从而避免或减少疾病的发生。

从逐年趋势来看，疾病死亡人数呈下降趋势。然而，在2017年，疾病死亡人数增幅明显超过往年，可能与2017年克拉玛依市空气质量状况^[15]有关，因大气污染物粒径较小，体表面积较大，活性强，易吸附有害物质^[16]，透过呼

吸进入呼吸道直至肺泡区域^[17]，增加人群患下呼吸道疾病、流感以及肺炎的风险。通过研究发现，虽然克拉玛依市居民呼吸系统疾病死亡数逐年呈下降趋势，与乌鲁木齐市相比，呼吸系统疾病死亡数较低^[18]，体现了当地政府实施“蓝天工程”方案^[19]的有效性，以及“绿水青山就是金山银山”理念的有效落实，但仍需加强大气污染防治，以降低疾病导致死亡风险，推进健康中国建设。

参考文献

- 丁璐璐.1990—2019年启东市呼吸系统疾病死亡趋势及减寿分析[J].预防医学,2021,33(11):1149-1152+1156.
- 毛叶挺,单利玲.2006—2017年南通市新发职业病疾病谱分析[J].职业与健康,2018,34(23):3198-3202.
- 左建辉,易军波.慢性病管理的实践与挑战[J].内科理论与实践,2019,14(6):49-51.
- 陈浩.基于统计年鉴的2002—2019年中国呼吸系统疾病死亡趋势分析[J].滨州医学院学报,2022,45(6):454-459.
- 郭杰.2011—2015年沈阳市呼吸系统疾病死亡的趋势分析[J].职业与健康,2018,34(10):1343-1347.
- 卢萌萌.典型石油型城市大气环境质量影响因素的多元统计模型研究——以大庆市为例[J].四川环境,2012,31(3):116-121.
- 郭凤娟.克拉玛依市PM_{2.5}质量浓度分布及其影响因素分析[J].气象与环境学报,2020,36(4):52-58.
- 王丽萍.PM_{2.5}对呼吸系统疾病的影响及其相关机制[J].临床肺科杂志,2019,24(1):148-151.
- 尚猛.郴州市循环系统疾病死亡人数预测模型建立[J].疾病监测与控制,2017,11(1):4-7.
- 钞虹.2007—2015年齐齐哈尔市恶性肿瘤死亡趋势分析及预测[J].中国慢性病预防与控制,2018,26(6):480.
- 何智敏.南通市空气污染对小学生健康危害的现状调查[J].职业与健康,2020,36(8):1095-1098+1103.
- 葛挺.宁波市城区大气PM_{2.5}污染与哮喘门诊量的关联性研究[J].预防医学,2019,31(6):568-572.
- 张晓伟,等.雾霾细颗粒物PM_{2.5}对COPD患者气道炎性反应的影响[J].职业与健康,2020,36(2):241-243.
- 王玉雯,等.天津社区居民PM_{2.5}中多环芳烃室外呼吸暴露健康风险评估[J].环境与健康杂志,2019,36(4):358-361.
- 闫琪.新疆4城市大气污染对呼吸系统及心血管系统影响的危险因素分析[D].乌鲁木齐:新疆医科大学,2019.
- 谢惠芳,陈佩弟,孙高峰.2015—2019年乌鲁木齐市大气污染物(PM_{2.5}和PM₁₀)对呼吸系统疾病死亡的影响[J].职业与健康,2021,37(17):2406-2411.
- 何晓庆,王小红,罗进斌.大气PM₁₀与呼吸系统疾病死亡的关系研究[J].预防医学,2019,31(10):987-991.
- 徐学哲.从“蓝天工程”看油城克拉玛依[J].中国环境管理干部学院学报,2017,27(5):95.
- 陈华.浅谈环境保护对新疆克拉玛依的重要性[J].节能与环保,2020(8):34-35.