

应用时间序列分析研究秋冬型恙虫病 时间分布特征及趋势

丁磊 丁淑军 张萌 王显军 李忠 赵仲堂

【摘要】 目的 应用时间序列分析探讨秋冬型恙虫病的时间分布特征及趋势。方法 基于山东省疾病报告信息系统 2006—2011 年恙虫病监测数据,对以年为单位的监测数据采用频谱分析、移动平均分析,预测 2012—2014 年该病发病数;以月为单位的数据进行乘法季节效应分析,进行差分自回归移动平均模型 (ARIMA) 建模拟合,预测 2011 年 11、12 月秋冬型恙虫病发病情况,并与实际发病数进行比较。结果 频谱分析结果显示秋冬型恙虫病流行周期为 3 年;移动平均分析表明其流行强度将持续增强,2012—2014 年各年预测发病数分别为 310、337、366 例,年增长率约 9%。乘法季节效应分析显示该病符合秋冬型流行特点,10、11 月季节效应指数分别为 8.454、2.230,其他月份指数均 < 1.000; ARIMA(0, 1, 1)(0, 1, 0)₁₂ 模型为 $(1-B)(1-B^{12})X_t = (1-0.811B)u_t$, 残差序列为白噪声, Box-Ljung 统计量为 3.116, $P=0.999$, 模型拟合效果良好。应用该模型预测 2011 年 11、12 月发病数,预测值与实际发病数相符,实际观测值均在预测值的 95% CI 内。结论 秋冬型恙虫病流行可能具有周期性,总体流行强度将持续增强,2012—2014 年该病年增长率约为 9%;每年以 10 月流行最严重,应用 ARIMA(0, 1, 1)(0, 1, 0)₁₂ 模型可较好地预测发病情况。

【关键词】 恙虫病; 时间序列分析; 流行趋势

Study on the characteristics of temporal distribution and the epidemic trend of autumn-winter type scrub typhus under time series analysis DING Lei¹, DING Shu-jun², ZHANG Meng¹, WANG Xian-jun², LI Zhong², ZHAO Zhong-tang¹. 1 Institute of Epidemiology and Health Statistics, School of Public Health, Shandong University, Jinan 250012, China; 2 Institute of Virus Disease, Shandong Provincial Center for Disease Control and Prevention, Shandong Provincial Key Laboratory of Communicable Disease Control and Prevention

Corresponding author: ZHAO Zhong-tang, Email: ztzhao@sdu.edu.cn

This work was supported by a grant from the National Natural Science Foundation of China (No. 30972515).

【Abstract】 Objective To explore the characteristics of temporal distribution and epidemic trend of autumn-winter type scrub typhus using the time series analysis. Methods Based on the data of scrub typhus collected from Shandong Diseases Reporting Information System from 2006 to 2011, both spectral analysis and moving average analysis were used to analyze the annual data of scrub typhus while scrub typhus incidence in 2012–2014 was forecasted. Seasonal decomposition analysis was applied to analyze the monthly data from January of 2006 to October of 2011, followed by Autoregressive Integrated Moving Average Model (ARIMA) which was constructed to forecast case number in November and December of 2011 and compared to the actual incidence. Results The results of spectral analysis showed that the prevalence of autumn-winter type scrub typhus had a feature of '3-year-periodicity'. A long-term up-trend was confirmed by method of moving average analysis, with annually case numbers of 310, 337 and another number of 366 forecasted for 2012 to 2014, respectively, with the annual increase rate as 9% per-year. Data from analysis of monthly data of scrub typhus showed that through multiple seasonal decomposition analysis, the results indicated that the prevalence of this disease possessed a typical autumn-winter type. The seasonality indexes for scrub typhus in October and November were 8.454 and 2.230, respectively, while others were less than 1.000. The ARIMA (0, 1, 1)(0, 1, 0)₁₂ model of $(1-B)(1-B^{12})X_t = (1-0.811B)u_t$ that was used to forecast the prevalence of autumn-winter type scrub typhus and was constructed with the

DOI: 10.3760/cma.j.issn.0254-6450.2012.07.012

基金项目: 国家自然科学基金(30972515)

作者单位: 250012 济南, 山东大学公共卫生学院流行病学与卫生统计学研究所(丁磊、张萌、赵仲堂); 山东省疾病预防控制中心病毒病所 山东省传染病预防控制重点实验室(丁淑军、王显军、李忠)

丁磊、丁淑军同为第一作者

通信作者: 赵仲堂, Email: ztzhao@sdu.edu.cn

residual error of 16 lags as white noise. The *Box-Ljung* test statistic for the model was 3.116, giving a *P* value of 0.999. The model fitted the data well. Good accordance was achieved between the observed values and the forecasted values of scrub typhus in November and December of 2011 which was produced by the ARIMA model, and all observed values were within the forecasted 95% CI. **Conclusion** The prevalence of autumn-winter type scrub typhus showed a 3-year-periodicity, with a long-term up-trend, and the case numbers of 2012 to 2014 were forecasted, rising on the end with an increasing rate of 9% per year, which occurred seasonally with October as the peak time in every year. The ARIMA (0, 1, 1) (0, 1, 0)₁₂ model seemed to be quite appropriate in predicting the autumn-winter type scrub typhus.

【Key words】 Scrub typhus; Time series analysis; Epidemic trend

恙虫病是由恙虫病东方体(*Ot*)感染引起的一种急性自然疫源性疾病,我国为重点疫区之一^[1-4]。1985年以前,该病主要流行于北纬31°以南的广大地区,以夏季流行为主。之后在北方地区不断蔓延,流行形式呈秋冬型。目前秋冬型恙虫病已波及我国北方约3/4的地区^[5],然而许多地区尚缺乏完整的流行病学资料。为了解秋冬型恙虫病的时间分布特征与流行趋势,本研究以山东省2006—2011年监测数据为基础,应用时间序列分析进行探讨。

资料与方法

1. 资料来源:山东省疾病报告信息系统2006—2011年恙虫病监测数据,各年发病数分别为216、242、352、230、251、450例,共1741例。诊断标准:①流行病学史;②临床表现:高热、特征性焦痂或溃疡、淋巴结肿大、皮疹、肝大或脾大;③间接免疫荧光试验显示*Ot*抗体滴度升高 ≥ 4 倍;④变形杆菌株*OXk*株外斐试验显示凝集度 $> 1:160$ 。满足以上3项者即诊断为恙虫病^[1,6]。

2. 分析方法:

(1)建模过程:以恙虫病监测数据为基础建立数据库,应用SPSS 16.0软件分别对以年、月为单位的发病数据进行时间序列分析;建模过程:①序列特征及平稳性:应用时间序列图、自相关系数函数图、偏自相关系数函数图分析序列特征(趋势性、季节性)及平稳性。若序列平稳,进行平稳序列的时间序列分析。②序列平稳化:对于非平稳序列,结合序列特征,采用合适的预处理方法,包括数据转换、普通差分、季节差分等,实现序列平稳化后进行分析。模型识别、定阶与参数估计参照文献[7]。阶数识别与定阶采用自相关系数、偏自相关系数及*Box-Ljung*法,参数估计采用最小二乘或非线性估计法。若序列不具有任何相关性,呈白噪声,则结合原序列特征,选择合适的确定性分析方法,包括趋势分析、平滑分析及季节效应分析。③模型拟合优度检验:创建ARIMA模型后,若残差序列呈白噪声,则认为所

建模型具有统计学意义;根据序列特征创建所有可能的ARIMA模型并比较分析,选择Akaike信息准则、Schwarz贝叶斯准则或贝叶斯信息准则(BIC)统计值最小者为最优模型。

(2)时间序列分析:①以年为单位的发病资料进行频谱分析^[8]和移动平均分析^[7],了解恙虫病发生是否具有周期性及趋势性,预测2012—2014年发病情况。②以月为单位的发病资料进行乘法季节效应分析^[7],了解恙虫病发生是否具有季节特征,并创建ARIMA(p, d, q) (P, D, Q)_s模型以预测恙虫病短期发病情况,利用最优模型对2011年11、12月恙虫病发病情况进行拟合预测。

结 果

1. 以年为单位的发病资料分析:

(1)序列特征及平稳性:恙虫病发病自相关系数与偏自相关系数均在随机区间内,*Box-Ljung*统计量为0.265, $P=0.607$,提示该序列呈平稳性,不存在短期相关性,无法创建ARIMA模型。

(2)频谱分析:当频率为0.35时,恙虫病发生出现周期峰值,提示该病的发生具有1/0.35年(约3年)的周期性。

(3)移动平均分析:由于恙虫病流行周期约为3年,因此进行3期移动平均分析,结果显示该病流行存在上升趋势,2012—2014年预测发病数分别为310、337、366例,年增长率约为9%。

2. 以月为单位的发病资料分析:

(1)序列特征及平稳性:以月为单位的发病资料分析显示,恙虫病发病呈明显季节性(图1);在滞长为12时,发病数自相关系数明显超过随机区间置信限,表明该序列存在季节相关性,需进行差分处理使其平稳化(图2)。

(2)模型建立:根据恙虫病发病数自相关系数图,采用一次季节差分($\nabla_{12} x_t$)与一次普通差分加一次季节差分($\nabla \nabla_{12} x_t$)平稳化处理,分别进行分析。

$\nabla_{12} x_t$ 差分处理:发病数自相关系数落入随机区

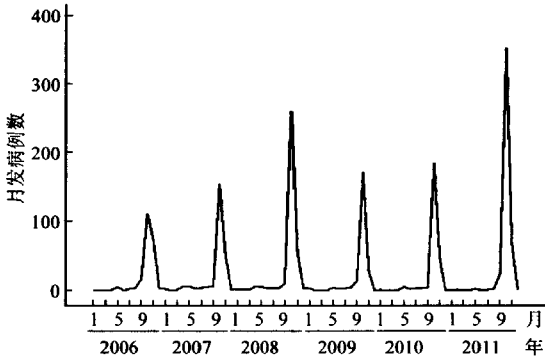


图1 2006—2011年山东省恙虫病的时间序列

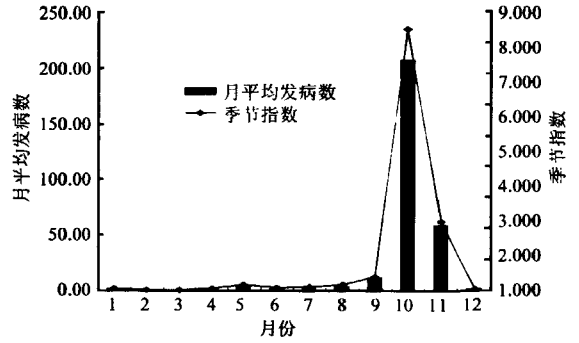


图3 2006—2011年山东省恙虫病季节效应分析

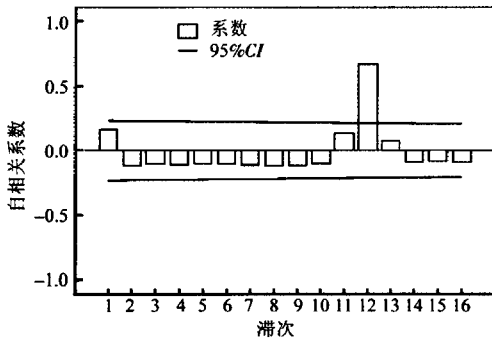


图2 2006—2011年山东省恙虫病的自相关系数

间内, *Box-Ljung* 统计量为 1.340, *P* 值为 0.247, 序列呈现平稳性。平稳化后序列不具有任何相关性, 结合原序列呈现的季节性, 采用季节效应分析。经过乘法季节效应分析, 结果显示, 10、11 月发病季节指数值分别为 8.454、2.230, 其他月份季节指数均 < 1.000。做残

差自相关系数图进行拟合优度检验, *Box-Ljung* 统计量为 4.116, *P* < 0.05, 残差序列呈非白噪声, 提示该模型不能对恙虫病进行准确拟合预测(图3)。

$\nabla \nabla_{12} x_t$ 差分处理:

①恙虫病发病数自相关系数及偏自相关系数落入随机区间(图4), 序列呈现平稳性。②模型识别、定阶及参数估计: 分析平稳化后序列的自相关系数、偏自相关系数图(图4)。结果显示, 自相关系数在滞次 $q=1$ 处较大且截尾, *Box-Ljung* 统计量为 4.064, *P* =

0.044, 偏自相关系数在 $p=1$ 处较大并拖尾。根据以上特点, 建立 $ARIMA(0, 1, 1)(0, 1, 0)_{12}$ 、 $ARIMA(0, 1, 0)(0, 1, 1)_{12}$ 、 $ARIMA(0, 1, 1)(0, 1, 1)_{12}$ 3 种模型进行拟合、比较, 进而选择最优模型。结果显示, 利用 3 种模型对恙虫病发生进行拟合预测后, 残差序列均呈白噪声, 所建模型均具有统计学意义; 其正态化 BIC 值分别为 6.992、7.185、7.015, 选用 BIC 值最小的 $ARIMA(0, 1, 1)(0, 1, 0)_{12}$ 模型为最优模型, 模型中参数 MA 的估计值为 0.811, 标准误为 0.135, $t=0.600$, $P < 0.01$, 模型表示为: $(1-B)(1-B^{12})X_t = (1-0.811B)u_t$, 式中 X_t 表示恙虫病月发病数。

(3)模型拟合优度检验: 残差序列的自相关系数及偏自相关系数均在 95%CI 以内, 残差呈白噪声; *Box-Ljung* 统计量为 3.116, $P=0.999$, 表明残差为随机性误差, 所建立的模型为最终模型(图5)。

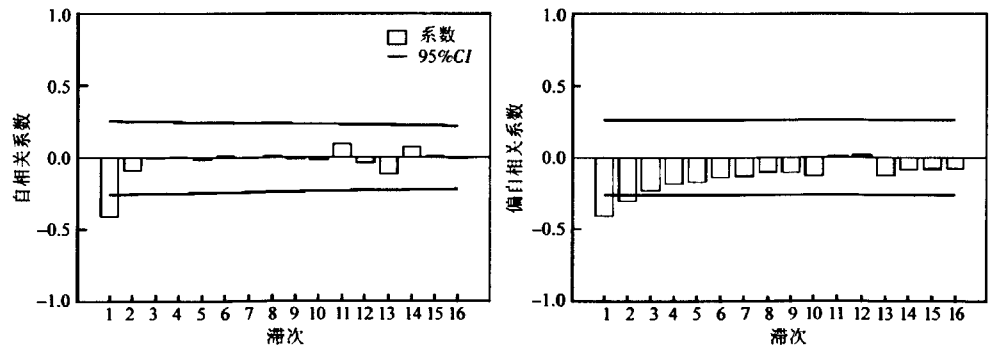


图4 2006—2011年山东省恙虫病数据经 $\nabla \nabla_{12} x_t$ 差分处理后序列的自相关、偏自相关系数

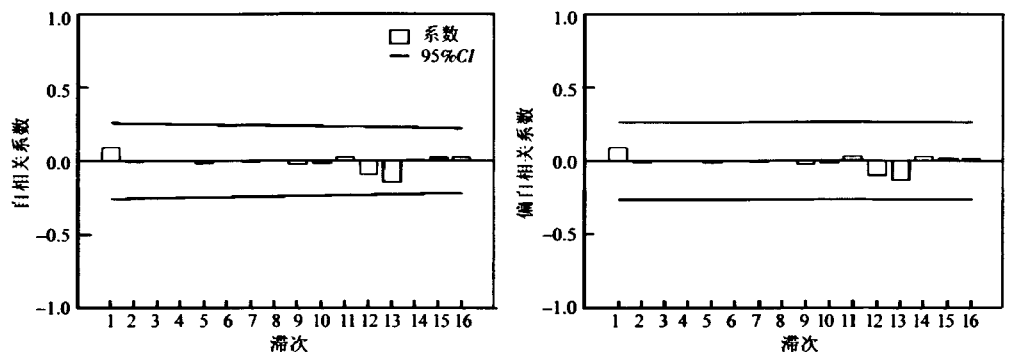


图5 模型拟合优度检验的残差自相关、偏自相关系数

3. 恙虫病发病预测:应用所建模型对2006年1月至2011年10月恙虫病发病时间序列进行回代预测,该时间段恙虫病发病数的观测值与拟合值基本相符,均位于拟合值置信限范围内。对2011年11、12月恙虫病发病数进行预测,预测值(95%CI)分别为85(23~146)与38(0~101),实际发病数分别为68、0例,均在预测值的95%CI内(图6)。

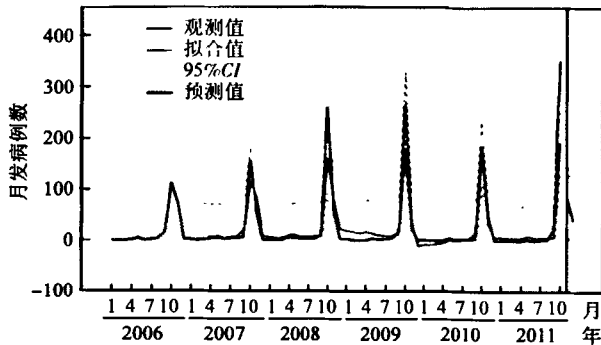


图6 ARIMA(0,1,1)(0,1,0)₁₂模型对秋冬型恙虫病发病的短期预测

讨 论

疾病监测数据往往具有时间相关性,采用对数数据要求具有独立性和随机性的传统分析方法进行分析,易丢失很多有价值的时间信息。时间序列分析能较好地解决这一问题,一方面能分析、识别数据间存在的时间属性信息,另一方面也能利用该属性信息,通过建立相应的模型拟合预测未来一定时间疾病的发病情况。在时间序列分析中,频谱、季节效应等确定性时序分析,能够提取序列具有的明显规律性信息,包括趋势性、季节性等;ARIMA等随机时序分析可利用序列存在的短期时间相关性,对疾病的发生进行拟合预测。

山东省为秋冬型恙虫病的典型疫区,自1986年始发现该病流行以来,流行范围不断扩大,全省已有80%的地区发生流行^[9,10],目前我国北方大多地区尚缺乏完整恙虫病发病资料,以山东省监测资料为基础进行分析具有一定的借鉴意义。

对以年为单位的资料进行分析,发现秋冬型恙虫病流行具有周期性,约每3年发生一次较高强度流行,这可能与人群抗体水平及易感人群数量变化有关,也可能与该病流行相关的自然、地理要素、储存宿主及媒介密度周期性变化相关。然而,本研究监测仅使用6年的数据探讨该病流行的周期性,时间跨度短,所得结论仅供参考,需以不断积累的资料数据进一步验证。采用移动平均分析消除周期性影响后,显示秋冬型恙虫病存在长期上升趋势,2012—2014年年度预测发病数分别为310、337、366例,年

增长率约为9%,提示该病流行强度将不断增强,应予关注。对以月为单位的资料进行分析,结果显示10月该病季节指数最高,与其秋冬型流行特点相符,可能与此间Ot储存宿主及媒介处于活跃期等有关^[11,12]。通过多种模型拟合,最终选取正态化BIC最小的复合季节模型[ARIMA(0,1,1)(0,1,0)₁₂]对恙虫病进行拟合预测,该模型能对恙虫病进行较准确的短期预测,可作为开展秋冬型恙虫病防治工作的参考。此外恙虫病作为一种自然疫源性疾,其流行受自然、地理及社会等因素的影响;研究发现该病在丘陵地区存在聚集现象^[13]。因此要对该病发生与流行进行更为准确的预测,还需在模型中纳入主要影响因素,进行进一步研究。

参 考 文 献

- [1] Peng WW. Infectious Diseases. 6th ed. Beijing: People's Medical Publishing House, 2004: 118-121. (in Chinese)
彭文伟. 传染病学. 6版. 北京:人民卫生出版社, 2004: 118-121.
- [2] Liu YX, Feng D, Suo JJ, et al. Clinical characteristics of the autumn-winter type scrub typhus cases in south of Shandong province, Northern China. BMC Infect Dis, 2009, 9: 82.
- [3] Zhang SY, Song HB, Liu Y. Scrub typhus in previously unrecognized areas of endemicity in China. J Clin Microbiol, 2010, 48(4): 1241-1244.
- [4] Koh GC, Maude RJ, Paris DH, et al. Review: diagnosis of scrub typhus. Am J Trop Med Hyg, 2010, 82(3): 368-370.
- [5] Zhao ZT. An epidemiologic study on the autumn-winter type scrub typhus. 2011 Cross-Strait Conference on Tropical Medicine, 2011: 50-51. (in Chinese)
赵仲堂. 秋冬型恙虫病流行病学研究. 2011两岸热带医学学术交流研讨会汇编, 2011: 50-51.
- [6] Yang LP, Zhao ZT, Li Z, et al. Comparative analysis of nucleotide sequences of *Orientia tsutsugamushi* in different epidemic areas of scrub typhus in Shandong, China. Am J Trop Med Hyg, 2008, 78(6): 968-972.
- [7] Box GE, Jenkins GM, Reinsel GC. Time Series Analysis: Forecasting and Control(时间序列分析: 预测与控制). 顾岚, 译. 北京: 中国统计出版社, 1999.
- [8] Zhang WT. Advanced Course of SPSS Statistical Analysis. Beijing: Higher Education Press, 2004: 266-267. (in Chinese)
张文彤. SPSS统计分析高级教程. 北京: 高等教育出版社, 2004: 266-267.
- [9] Zhang M, Wang XJ, Zhao ZT. Current epidemic status and issues on prevention and control of scrub typhus. Chin J Epidemiol, 2011, 32(4): 419-423. (in Chinese)
张萌, 王显军, 赵仲堂. 中国恙虫病流行态势及预防控制. 中华流行病学杂志, 2011, 32(4): 419-423.
- [10] Zhang SY, Jiang DY, Yu DZ. Endemic characteristics of autumn scrub typhus and the epidemic situation in China. Foreign Med Sci: Section of Medgeography, 2009, 30(3): 133-135. (in Chinese)
张守印, 蒋德勇, 俞东征. 我国秋冬型恙虫病地方特点及流行状况分析. 国外医学: 医学地理分册, 2009, 30(3): 133-135.
- [11] Liu YX, Zhao ZT, Yang ZQ, et al. A study on the animal reservoirs of scrub typhus in Shandong area. Chin J Epidemiol, 2004, 25(2): 183. (in Chinese)
刘运喜, 赵仲堂, 杨占清, 等. 山东地区秋冬型恙虫病宿主动物流行病学调查. 中华流行病学杂志, 2004, 25(2): 183.
- [12] Zhang Q, Liu YX, Zhao ZT, et al. Study on the molecular epidemiology regarding the natural infection of *Orientia tsutsugamushi* in 4 species of dominant chiggers collected in various seasons from the foci of Shandong province. Chin J Epidemiol, 2006, 27(7): 600-603. (in Chinese)
张倩, 刘运喜, 赵仲堂, 等. 山东地区不同季节优势恙螨自然感染恙虫病东方体的分子流行病学调查. 中华流行病学杂志, 2006, 27(7): 600-603.
- [13] Ding L, Li Z, Wang XJ, et al. Analysis of epidemic features of scrub typhus between year 2006 and 2010 in Shandong province, China. Chin J Prev Med, 2012, 46(4): 338-342. (in Chinese)
丁磊, 李志, 王显军, 等. 2006—2010年山东省恙虫病流行特征分析. 中华预防医学杂志, 2012, 46(4): 338-342.

(收稿日期: 2012-01-18)

(本文编辑: 万玉立)