

doi:10.3969/j.issn.1672-5166.2018.06.02

某市空气污染对慢性阻塞性肺部疾病住院人数的发病影响研究

丁中印^① 高 辉^② 范计朋^{③△} 张遵真^④ 何明杰^③

文章编号: 1672-5166 (2018)06-0619-06 中图分类号: R-34 ; R563 文献标志码: A

摘要 **目的** 探讨自贡市室外大气污染与慢性阻塞性肺疾病(COPD)发生的关系,为获得早期预警提供依据。**方法** 收集2014年12月18日—2016年7月1日自贡市各城区医院电子病历病案首页数据,结合空气污染物(包括PM_{2.5}、PM₁₀、SO₂、CO、NO₂、O₃)浓度,使用基于聚合机制和滞后机制的皮尔森相关系数分析方法、单因子泊松回归和多因子泊松回归分析方法,分析自贡市空气污染对慢性阻塞性肺部疾病住院人数的发病影响。**结果** 单因子模型分析结果为:各污染物与COPD住院人数有显著相关关系,空气中的SO₂、NO₂每增加10μg/m³,COPD住院率增加0.3783%、0.2576%;多因子模型分析结果为:空气中的SO₂、NO₂每增加10μg/m³,住院人数分别增加0.3605%、0.2083%。**结论** 自贡市空气污染物浓度对COPD住院人数有影响且存在累积效应和滞后效应,尤以SO₂、NO₂最为显著。

关键词 空气污染 慢性阻塞性肺部疾病 累积效应 滞后效应

Study on the Impact of Air Pollution on the Incidence of Hospitalization of Chronic Obstructive Pulmonary Disease in Zigong City

DING Zhongyin, GAO Hui, FAN Jipeng, ZHANG Zunzhen, HE Mingjie

Zi Gong first People's Hospital, Zigong 643000, Sichuan, China

Abstract **Objective** To explore the relationship between outdoor air pollution and the occurrence of chronic obstructive pulmonary disease (COPD) in Zigong City, and provide evidence for early warning. **Methods** we collected the data of the electronic medical records of hospitals in each district of Zigong city from December 18, 2014 to July 1, 2016, combined with concentrations of air pollutants (PM_{2.5}、PM₁₀、SO₂、CO、NO₂、O₃). The Pearson correlation coefficient analysis method based on aggregation mechanism and hysteresis mechanism, single factor Poisson regression analysis methods, and multifactor Poisson regression analysis methods were used to analyze the impact of air pollution on the number of

基金项目: 国家自然科学基金(项目编号: 71661167005); 四川省科技计划项目: 基于云平台的医学影像质量评价体系研究与应用示范(项目编号: 2017SZ0016)

① 自贡市第一人民医院, 自贡市, 643000

② 电子科技大学, 成都市, 610041

③ 成都金盘电子科大多媒体技术有限公司, 成都市, 610041

④ 四川大学华西公共卫生学院, 成都市, 610041

作者简介: 丁中印(1962—), 男, 院长; 研究方向: 医院管理 E-mail: 842640499@qq.com

通信作者: 范计朋(1987—), 男, 硕士; 研究方向: 医学影像与人工智能 E-mail: fanjipeng@gdpacs.com

△ 通信作者

hospitalized patients with chronic obstructive pulmonary disease in Zigong city. **Results** The analysis results of the single factor model showed that there was a significant correlation between each pollutant and the number of COPD hospitalizations. A $10 \mu\text{g}/\text{m}^3$ increase in the concentration of sulfur dioxide and nitrogen dioxide in the air corresponded to a 0.3783% and 0.2576% increase in the hospitalization rate of COPD respectively. The results of multi-factor analysis showed that every $10 \mu\text{g}/\text{m}^3$ increase in the concentration of sulfur dioxide and nitrogen dioxide in the air corresponded to an increase of 0.3605% and 0.2083% in the number of COPD hospitalizations, respectively. **Conclusions** The concentration of air pollutants in Zigong city has an impact on the number of COPD hospitalizations and has cumulative effects and hysteresis effects, sulfur dioxide and nitrogen dioxide are the most significant.

Key words air pollution; COPD; cumulative effect; hysteresis effect

慢性阻塞性肺疾病 (chronic obstructive pulmonary diseases, COPD) 是一种以气道炎症为核心, 由多种因素构成的进展性疾病^[1-2]。据世界卫生组织推测, 到2030年, COPD 将从1990年的第6位跃升至全球第3位致死病因, COPD 已成为全球范围内重要的公共卫生问题, 以及国内外学者研究的重点^[3-4]。庞大的患者基数和疾病本身的高致死率, 给人类实现 COPD 有效规范性防治带来了巨大挑战。

近年的流行病学研究认为, 大约30%的呼吸道疾病病因可能与空气污染有关, 空气污染对 COPD 的健康风险存在地区和人群差异^[5-6]。目前, 国内外关于空气污染物与 COPD 住院人数关系的研究报道也不尽相同。Yusef 等研究表明^[7], O_3 浓度每增加 $10 \mu\text{g}/\text{m}^3$, 呼吸道疾病死亡率和 COPD 风险分别增加 1.25% 和 0.86%, 但其潜在的致病机制尚不明确。在芬兰、美国、加拿大的研究也得到类似的结果^[8], 即 SO_2 、 NO_2 、 O_3 与 COPD 的急性发作存在着较强的相关性。但朱瑞霞等人的一项研究显示, 无论在单污染物模型还是双污染物模型, COPD 急性加重与空气污染物并未发现任何相关^[9]。

目前, 大量的科学研究, 主要集中在深层次挖掘与分析空气污染物对 COPD 住院人数的滞后效应^[10-12]。北京市的一项研究结果表明^[13], 针对空气污染物的滞后效应, 应用时间序列分析, 探讨主要空气污染物与居民的日就诊入院人数之间的关系, 使得呼吸系统疾病急诊就诊人数的相对危险度值达到最大, 以推断空气污

物和 COPD 住院人数的关联程度。此外, 也有部分研究采用时间序列分析方法, 挖掘历史平均空气污染物浓度对当天 COPD 住院人数的综合影响^[14]。但是, 由于数据缺乏地点的一致性, 数据收集方法也不尽一致, 往往会对研究结果产生一定的偏倚, 难以推断空气污染物与 COPD 住院人数之间的潜在影响关系。更重要的是, 很少有研究分析污染物在时间维度上的累积效应, 对 COPD 住院人数的影响。基于此, 本研究对污染物浓度在时间维度上应用聚合处理, 分析污染物与 COPD 住院人数的相关性, 探讨不同污染物对 COPD 住院人数影响的累积效应和滞后效应。

1 资料与方法

1.1 资料来源

1.1.1 患者数据

来源于自贡市各城区医院的病案记录, 研究获取 2014 年 12 月 18 日—2016 年 7 月 1 日期间每日 COPD 患者数据, 共 82485 条。患者数据的具体信息包括: 疾病编码、住院和出院时间、性别、年龄等。由住院和出院时间可统计得到该期间 COPD 住院人数 561 人, 出院人数 548 条。

1.1.2 空气污染物数据

应用网络爬虫技术从中国空气质量在线监测分析平台 (<https://www.aqistudy.cn/>), 爬取自贡市空气质量数据。具体内容, 包括自贡市 2014 年 12 月 18

日—2016年7月1日期间每日空气污染物的日平均浓度。其中,空气污染物主要包括PM_{2.5}、PM₁₀、二氧化硫(SO₂)、一氧化碳(CO)、二氧化氮(NO₂)、臭氧(O₃)。

1.2 资料数据预处理

1.2.1 患者住院数据预处理

由于以上患者住院数据资料是包括自贡市医院各种疾病的住院患者数据,而本研究主要是针对COPD患者,故首先通过疾病编号筛选出自贡市2014年12月18日—2016年7月1日期间每日的COPD住院患者,然后通过患者住院日期统计出每日COPD住院人数,以此作为研究的主要统计数据。

1.2.2 空气污染数据预处理

国内外众多研究表明,空气污染物浓度上升当日不一定引起人体立即发病或死亡,需累积达到一定程度,经过一定的时间后作用于人体,这就是致病因子的累积效应和滞后效应。为了将各空气污染物的累积效应与滞后效应纳入考量,本研究提出了一种对空气污染物聚合和滞后的处理方法,应用不同的时间粒度,遍历空气污染物在不同累积和滞后情况下,与COPD住院人数之间的相关系数,以此揭示空气污染物对COPD住院人数的影响。污染物X在当前时刻t之前d天向前聚合m天的浓度之和X(t,d,m)的计算公式如下:

$$X(t,d,m) = X_{t-d} + X_{t-d-1} + \dots + X_{t-d-m+1}$$

其中, X_i表示空气污染物X在时刻i时的浓度, t表示当前时刻, d表示空气污染物X的滞后天数, m表示空气污染物X的前向聚合天数。见图1。

皮尔森相关系数(Pearson correlation coefficient, 简称PCC)是一种线性相关系数,是用来反映两个变量线性相关程度的统计量,绝对值越大表示相关性越强。本文研究采用皮尔森相关系数检验,准确度量不同空气

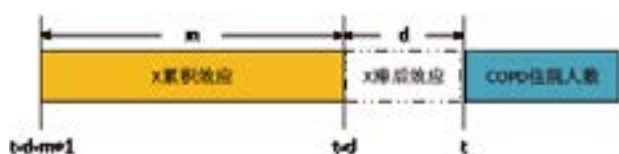


图1 聚合、滞后处理示意

污染物浓度在聚合和滞后效应上,与COPD住院人数之间的关系密切程度。皮尔森相关系数的定义如下:

$$r(d,m) = \frac{\sum_{t=1}^n (X(t,d,m) - \overline{X(t,d,m)})(Y_t - \overline{Y_t})}{\sqrt{\sum_{t=1}^n (X(t,d,m) - \overline{X(t,d,m)})^2 \cdot \sum_{t=1}^n (Y_t - \overline{Y_t})^2}}$$

其中, r(d,m)表示在聚合m天和滞后d天基础上污染物与COPD的皮尔森相关系数, X(t,d,m)表示针对聚合m天和滞后d天,得到的污染物X在时刻t时的浓度, $\overline{X(t,d,m)}$ 表示,时间t在[1,n]期间每天对应的X(t,d,m)的均值; Y_t表示t时刻COPD住院总人数; $\overline{Y_t}$ 表示,时间t在[1,n]期间每天对应的Y_t的均值, n为时间跨度。

在考虑聚合效应和滞后效应的基础上,将空气污染物数据X(t,d,m)与COPD住院人数看作是随时间变化的一组序列(X(t,d,m), Y_t)。本研究通过使用贪心算法遍历计算污染物在不同(d,m)参数下与住院人数的相关系数,将二者相关系数最大时的参数(d,m)作为该空气污染物的最佳聚合粒度和滞后时间。计算公式如下:

$$(d,m) = \arg \max_{d,m} |r(d,m)|$$

1.3 泊松回归

在统计分析中,泊松回归是以一定时间内事件发生次数为研究对象,研究风险暴露期和其它协变量对事件发生率的影响。研究采用基于时间序列的泊松回归模型筛选出与COPD住院率有统计学关联的污染物因素。泊松回归模型的一般形式为:

$$\ln Y_t = \beta_0 + \sum_i \beta_i X_i(t,d,m)$$

其中, Y_t表示t时刻COPD住院人数, i ∈ {PM_{2.5}, PM₁₀, SO₂, CO, NO₂, O₃}, 通过皮尔森相关系数r(d,m)确定最佳聚合粒度m和滞后时间d, X_i(t,d,m)表示在时间为t-d-m+1到t-d期间污染物浓度的累积。偏回归系数β_i表示在控制其它变量的条件下, X_i(t,d,m)变化一个单位,将给目标变量带来对数均值上的变化,可用exp(β_i)来反映X_i(t,d,m)变化一个单位时,期望计数的倍数变化。

在数据预处理的基础上,结合空气污染物对COPD住院人数的累积效应、滞后效应,构建单污染物和多污

染物模型^[15],应用时间序列的泊松回归分析,以判断自贡市 COPD 住院人数和空气污染物的累积、滞后浓度是否有显著相关性,以及其与各空气污染物相关程度与相关方向。在每段时间区间内,COPD 住院人数是一种计数过程,相对于人群来说是一种小概率事件,因此,在每段时间区间内对 COPD 住院人数数据拟合泊松分布是合适的^[16]。将空气污染物变量与 COPD 住院人数进行泊松回归,根据 P 值判断各空气污染物对 COPD 住院人数是否有显著意义,据此得到各空气污染物与 COPD 住院人数的泊松回归结果。

2 结果

2.1 描述性分析

2014 年 12 月 18 日—2016 年 7 月 1 日期间,自贡市 6 种主要空气污染物 PM_{2.5}、PM₁₀、SO₂、CO、NO₂、O₃ 的日平均浓度分别为:75 μg/m³、110 μg/m³、16 μg/m³、1 μg/m³、30 μg/m³、67 μg/m³。根据我国的《环境空气质量标准》(GB3095—2012)规定:PM_{2.5}、PM₁₀、SO₂、CO、NO₂、O₃ 日平均浓度二级标准分别为 75 μg/m³、150 μg/m³、150 μg/m³、4 μg/m³、80 μg/m³、160 μg/m³。其中 PM_{2.5}、PM₁₀ 超过标准的天数分别占:39.26%、20.26%,可知自贡市空气污染较严重,主要表现为可吸入颗粒物污染。见表 1。

2.2 相关性分析

2014 年 12 月 18 日—2016 年 7 月 1 日期间,自贡市每日 COPD 住院人数与空气污染物因素相关分析,见表 2。SO₂、NO₂ 均与 COPD 住院人数呈正相关,相关系数分别为 0.107、0.169,差异有统计学意义 (P < 0.05); PM_{2.5}、PM₁₀、CO 与 COPD 住院人数呈正相关,但差异无统计学意义 (P > 0.05); O₃ 与 COPD 住院人数呈负相关,相关系数为 -0.259,差异有统计学意义 (P < 0.05)。

考虑到空气污染物对人体健康影响的滞后效应,本研究探讨在不同的滞后时间 (0~15 天) 条件下,

主要空气污染物与每日 COPD 住院人数的相关性,见图 2。可以看出:PM_{2.5}、PM₁₀、NO₂ 随滞后时间的延长,其与 COPD 住院人数的皮尔森相关系数 r 值均表现出不同程度的增长趋势。其中,SO₂、CO 与每日 COPD 住院人数的相关性峰值分别出现在滞后 10 天和滞后 13 天时;O₃ 与每日 COPD 住院人数出现较强的负相关性,且其皮尔森系数随滞后时间的延长而逐渐减小,故 O₃ 浓度的变化对当天 COPD 住院人数的影响最大。

滞后效应无法全面评价空气污染物对 COPD 住院人数的综合影响。因此,在引入滞后效应的基

表 1 2015—2016 年自贡市空气污染物浓度分布

污染物	均值 ± 标准差	最小值	四分位			最大值
			25th	50th	75th	
PM _{2.5}	74.81 ± 45.48	15.0	42.0	62.0	95.0	316.0
PM ₁₀	109.90 ± 58.03	28.0	64.75	94.5	136.0	388.0
SO ₂	16.41 ± 7.59	5.0	11.0	15.0	20.0	73.0
CO	0.95 ± 0.37	0.30	0.70	0.90	1.20	2.10
NO ₂	30.40 ± 9.47	13.0	24.0	29.0	34.0	76.0
O ₃	67.26 ± 29.16	17.0	46.0	63.0	83.0	149.0

表 2 2015—2016 年自贡市空气污染物与 COPD 住院人数 Pearson 相关分析

指标	空气污染物					
	PM _{2.5}	PM ₁₀	SO ₂	CO	NO ₂	O ₃
PM _{2.5}	1.000					
PM ₁₀	0.984	1.000				
SO ₂	0.497	0.506	1.000			
CO	0.769	0.748	0.514	1.000		
NO ₂	0.566	0.569	0.389	0.521	1.000	
O ₃	-0.155	-0.120	-0.072	-0.270	-0.104	1.000
COPD	0.062	0.058	0.107	0.070	0.169	-0.259

BIG DATA OF HEALTH CARE 健康医疗大数据专栏

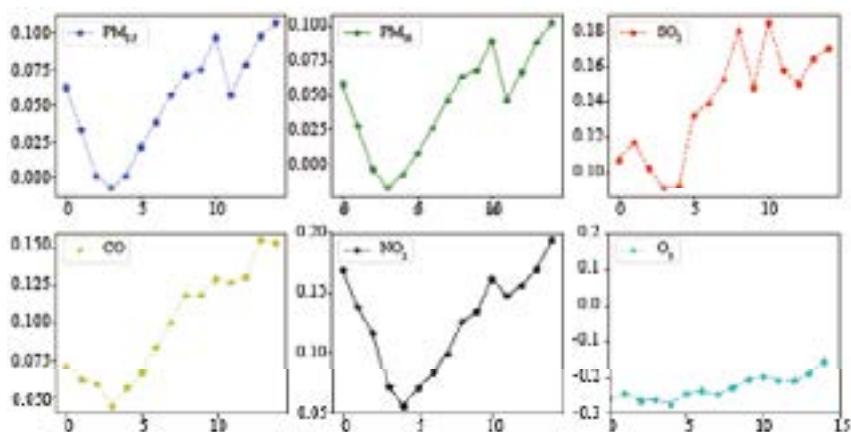


图2 污染物对COPD发病的滞后效应

基础上,考虑多天污染物浓度水平对COPD住院人数的累积效应,故应用聚合方法处理每日空气污染物浓度数据。综合考量各空气污染物的累积效应、滞后效应后,单一空气污染物与COPD住院人数的皮尔森相关性结果见图3。结果揭示了各空气污染物对每日COPD住院人数的累积效应、滞后效应。以PM_{2.5}为例,分析得到PM_{2.5}在滞后0天,前向聚合14天后的数据与每日COPD住院人数达到最高相关性,r值为0.26,其它污染物与每日COPD住院人数达到最高相关性时的效应机制见表3。

通过上述分析可看出,空气中各污染物浓度与COPD住院人数存在相关性。此外,在引入空气污染物的累积效应与滞后效应后,各污染物与COPD住院人数的相关性得到了显著提升,其中SO₂与COPD住院人数相关性最高,达到0.47。以皮尔森相关性水平为依据,计算得到各污染物累积效应与滞后效应。为深入研究自贡市空气污染物浓度与COPD住院人数的具体相关

程度,将在相关性分析的工作基础上应用泊松回归分析,分析各污染物对COPD住院人数的影响程度。

2.3 泊松回归分析

在得到各空气污染物的累积效

应和滞后效应的基础上,将空气中的PM_{2.5}、PM₁₀、SO₂、CO、NO₂、O₃等污染物浓度,分别与相应的COPD住院人数进行单因素泊松回归分析。结果表明:空气中PM_{2.5}、PM₁₀、SO₂、CO、NO₂、O₃与COPD的相关关系有显著性,SO₂、NO₂每增加10μg/m³,COPD的住院率将分别增加0.3738%、0.2576%。见表4。

在多污染物模型中,以PM_{2.5}、PM₁₀、SO₂、CO、NO₂、O₃为自变量,COPD住院人数为因变量,结合污染物的最佳滞后效应和累积效应,对COPD住院人数进行多因素泊松回归分析。结果表明:当引入其他污染物后,大气PM_{2.5}、PM₁₀、

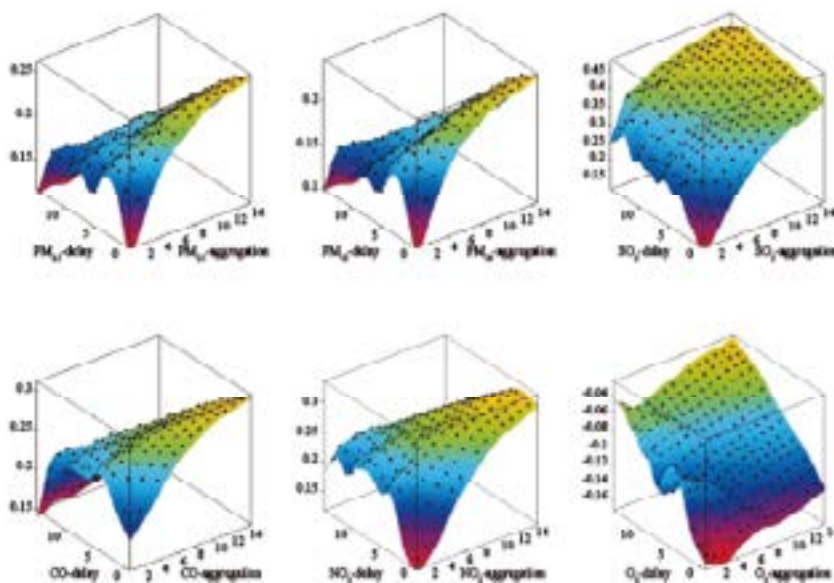


图3 聚合污染物浓度与COPD住院人数相关性分析

表3 各污染物对COPD住院人数的滞后效应、累积效应(天)

	PM _{2.5}	PM ₁₀	SO ₂	CO	NO ₂	O ₃
累积效应	14	14	14	14	14	4
滞后效应	0	1	12	0	2	0
r	0.26	0.24	0.47	0.31	0.33	-0.17

表4 各污染物每增加 10 μg/m³ 对 COPD 住院率增加的百分比 (单因素分析)

泊松回归	PM _{2.5}	PM ₁₀	SO ₂	CO	NO ₂	O ₃
COPD	0.2018*	0.1854*	0.3738*	0.2077*	0.2576*	0.0604*

CO 对 COPD 住院人数的影响消失 ($P>0.05$), 仅 SO₂、NO₂、O₃ 三种污染物与 COPD 住院人数有显著性关系, 所以在多污染物模型中移除 PM_{2.5}、PM₁₀、CO 三种关系不显著的变量, 仅将自变量设为 SO₂、NO₂、O₃。空气中 SO₂、NO₂、O₃

每增加 10 μg/m³, COPD 的住院率分别增加 0.3605%、0.2083%、0.1364%, 见表 5。多污染物模型泊松回归公式表达如下:

$$\ln Y_1 = 1.346 + 1.282X_{SO_2}(t, 12, 14) + 0.734X_{NO_2}(t, 2, 14) + 0.311X_{O_3}(t, 0, 4)$$

表5 各污染物每增加 10 μg/m³ 对 COPD 住院率增加的百分比 (多因素分析)

泊松回归	PM _{2.5}	PM ₁₀	SO ₂	CO	NO ₂	O ₃
多因素模型	-	-	0.3605*	-	0.2083*	0.1364*

3 讨论

本研究通过自贡市主要大气污染物与 COPD 住院人数等数据, 在时间刻度上, 将污染物数据进行聚合, 客观分析空气污染物浓度与 COPD 住院人数的关系, 为进一步预测 COPD 住院人数提供有效依据。尽管本文给出了各空气污染物对 COPD 住院人数的滞后效应、累积效应, 但其具体原因有待进一步分析、完善。■

参考文献

[1] Zhijing Lin, Yutong Gu, Cong Liu, et al. Effects of ambient temperature on lung function in patients with chronic obstructive pulmonary disease: A time-series panel study[J]. Science of the total environment, 2018, 619: 360-365.
 [2] Sujee Lee, Sijie Wang, Philip A. Bain, et al. Modeling and Analysis of Postdischarge Intervention Process to Reduce COPD Readmissions. IEEE Transactions on Automation Science

and Engineering.
 [3] 钱轶峰, 韩明, 靳文正, 等. 上海市大气污染与慢性阻塞性肺病死亡的关联性[J]. 环境与职业医学, 2015, 32(12): 1093-1097.
 [4] Sun Xian Wen, Chen Pei Li, Ren Lei. The cumulative effect of air pollutants on the acute exacerbation of COPD in Shanghai, China[J]. Science of the total environment, 2018, 622: 875-881.
 [5] 崔娟, 殷鹏, 王黎君, 等. 1990年与2013年中国大气臭氧污染导致慢性阻塞性肺疾病的疾病负担分析[J]. 中华预防医学杂志, 2016, 50(5): 391-396.
 [6] 李春艳, 周丽婷, 王树越, 等. 大气污染对慢性阻塞性肺疾病住院影响的Meta分析[J]. 中国老年学杂志, 2016, 36(5): 1141-1145.
 [7] Khaniabadi Yusef Omid, Hopke Philip K, Goudarzi Gholamreza. Cardiopulmonary mortality and COPD attributed to ambient ozone[J]. Environment research, 2017, 152: 336-341.

[8] 周骥. 浅析气象因子及大气成分对 COPD 发病的影响 [A]. 中国环境科学学会. 2014 中国环境科学学会学术年会 (第三章) [C]. 中国环境科学学会, 2014, 32: 9.
 [9] 朱瑞霞, 陈亚红, 邓芙蓉, 等. 北京市大气污染物对慢性阻塞性肺疾病预后的影响 [J]. 国际呼吸杂志, 2012, 32(9): 665-671.
 [10] 赵凤敏, 吴一峰, 吴峰. 大气污染物浓度与上呼吸道疾病门诊就诊的关联性研究 [J]. 浙江预防医学, 2016, 28(2): 165-167, 170.
 [11] 吴筱音, 李国星, 王旭英, 等. 北京市大气污染与呼吸系统疾病死亡的相关性—基于卫星遥感数据的时空分析 [J]. 北京大学学报 (医学版), 2017, 49(3): 409-417.
 [12] 李阳. 杭州市主要空气污染物浓度与呼吸系统疾病的关系研究 [D]. 浙江农林大学, 2017.
 [13] 张莹, 宁贵财, 康延臻, 等. 北京市大气污染物与呼吸系统疾病入院人数的时间序列 [J]. 兰州大学学报 (自然科学版), 2015, 59(1): 87-92.
 [14] 张莹, 邵毅, 王式功, 等. 北京市空气污染物对呼吸系统疾病门诊人数的影响 [J]. 中国环境科学, 2014, 34(9): 2401-2407.
 [15] 张晓云. 气象要素和大气污染物对心脑血管疾病影响分析 [A]. 中国气象学会. 第 33 届中国气象学会年会 S16 气候环境变化与人体健康 [C]. 中国气象学会, 2016: 8.
 [16] 王德征, 江国虹, 顾清, 等. 采用时间序列泊松回归分析天津市大气污染物对心脑血管疾病死亡的急性影响 [J]. 中国循环杂志, 2014, 29(6): 453-457.

[收稿日期: 2018-09-25
 修回日期: 2018-11-27]