病例交叉法和时间序列法在臭氧对心血管短期影响中的应用比较*

谷少华[1] 王 永[1] 陆蓓蓓[1] 段东辉[1] 王小丽[2] 王爱红[1△]

【提 要】 **目的** 应用基于条件泊松模型的病例交叉法和时间序列法评估臭氧对心血管疾病发病与死亡的短期影响。**方法** 收集宁波市2015—2021年的心血管疾病死亡、住院和环境因素监测数据，采用时间分层的病例交叉法结合条件泊松回归模型定量评估日最大8小时臭氧浓度与心血管疾病住院和死亡人数的短期关联，并与时间序列法的分析结果进行比较。**结果** 病例交叉法和时间序列法的分析结果非常接近，均显示臭氧暴露当天（lag0）心血管疾病住院风险上升，两种方法计算获得的超额住院率分别为0.56%（95%*CI*：0.01%~1.11%）和0.49%（95%*CI*：0.00%~0.98%）。暴露滞后1~4天时出现心血管疾病死亡风险上升，两种方法获得的累积5天（lag0~4）超额死亡率分别为0.51%（95%*CI*：0.11%~0.92%）和0.48%（95%*CI*：0.10%~0.87%）。**结论** 病例交叉法和时间序列法的分析结果非常接近，均提示短期臭氧暴露与心血管疾病发病和死亡风险上升相关。

【关键词】 病例交叉 时间序列 条件泊松 臭氧 心血管疾病
【中图分类号】 R195.1 【文献标志码】 A **DOI** 10.11783/j.issn.1002-3674.2026.05.015

病例交叉研究是一种自我匹配的病例对照设计方法，通过比较研究对象在某疾病或事件发生时与未发生时的暴露情况，评估短暂暴露对急性疾病或事件的瞬时影响[1]。根据“对照期”选取方式的差异，可分为单向、双向和时间分层三种类别，其中时间分层的病例交叉设计能够较好地控制时间混杂因素，获得参数的无偏估计，更被研究者所认同[2-3]。近年来病例交叉法作为时间序列法的平行替代，也常应用于环境流行病学研究，在群体水平上评估大气污染物、气温等环境因子对人群健康的急性影响[4]。

群体水平的病例交叉研究通常基于条件逻辑回归建立多因素模型，可利用R软件中survival程序包的coxph函数实现数据分析过程，但是此方法需要的数据集较为复杂，且不能解决数据过度离散和自相关问题[4-5]。此外，R软件中glm包的glm函数利用泊松回归模型也可以实现病例交叉分析，然而该模型中参数较多，计算效率低[5-6]。Armstrong等[5]在既往研究基础上提出基于条件泊松回归的病例交叉分析方法，通过设置哑变量方法进行时间分层，节省了数据整理时间，同时解决了过度离散、计算效率低等问题。但是目前该方法在国内应用较少，并缺少与其他研究方法的比较。

本文拟利用宁波市环境与健康历史监测数据，分别采用基于条件泊松模型的病例交叉法与时间序列法分析地面臭氧浓度与心血管疾病死亡和发病的短期关联，比较两种方法研究结果的差异。同时文中提供示例数据集和相应R语言代码，期望给予其他研究者更多启发性和实用性参考。

对象与方法

1. 数据收集与整理

从宁波市疾病预防控制中心收集2015—2021年宁波市人群死因监测个案数据，变量包括：死亡时间、根本死因和根本死因的ICD-10编码。从宁波市某三甲医院收集2015—2021年的住院个案数据，变量包括住院时间、住院诊断和住院诊断的ICD-10编码，该医院患者主要来源于主城区及周边区县的常住居民。为排除新冠疫情对患者就诊意愿的影响，2020年1月—3月的住院数据设置为缺失值。根据ICD-10编码筛选疾病种类，心血管疾病分类为I00-I99。

同期的气象数据收集自宁波市气象服务中心，来源于位于主城区的国家基准站点，变量包括日均气温和日均相对湿度。同期的空气质量监测数据收集自宁波市生态环境监测中心，涵盖8个位于主城区的国控点位，选用区域平均日最大8小时臭氧浓度作为人群暴露的代表指标。

将环境监测数据和健康结局数据整理为按日排序的时间序列数据，并按照日期合并。分析数据集示例见表1。

2. 数据分析

分别采用病例交叉法与时间序列法进行统计分析。由于死亡和住院均为小概率事件，且可能存在过度离散，因此均采用准泊松分布（quasi-Poisson）作为模型的链接函数。选择分布滞后非线性模型（distributed lag non-linear model）拟合臭氧和气温对健康影

* 基金项目：浙江省自然科学基金项目（LTGY24H260008）；宁波市医疗卫生品牌学科（PPXK2024-09）；宁波市公益类科技计划（2022S069）
1. 宁波市疾病预防控制中心（315000）
2. 宁波卫生职业技术学院
△通信作者：王爱红，E-mail：77828079@qq.com

表1　分析数据集（前10条记录）

日期	年份	月份	星期	死亡（例）	住院（例）	O_3（μg/m^3）	日均气温（℃）	日均相对湿度（%）
2015-1-1	2015	1	4	53	3	66.88	3.50	34.00
2015-1-2	2015	1	5	50	11	57.75	2.20	49.00
2015-1-3	2015	1	6	43	8	73.00	5.50	57.00
2015-1-4	2015	1	7	50	8	75.25	10.10	57.00
2015-1-5	2015	1	1	41	10	72.50	14.90	76.00
2015-1-6	2015	1	2	50	9	58.63	9.60	68.00
2015-1-7	2015	1	3	37	8	48.13	4.70	66.00
2015-1-8	2015	1	4	45	7	65.13	4.70	58.00
2015-1-9	2015	1	5	42	6	64.88	4.20	61.00
2015-1-10	2015	1	6	32	3	28.50	6.40	64.00

响的滞后性和非线性[7]。参考以往文献，臭氧的暴露-反应关系使用线性拟合，日均气温的的暴露-反应关系使用自然立方样条函数（ns）进行非线性拟合，自由度设为3。两者的最长滞后时间分别选择7天和14天，滞后-反应关系均采用对数等间距的样条函数拟合，设置2个节点[8-9]。

病例交叉法以月份进行时间分层，病例期和对照期为同一年、同一月的同一个星期几[2]。时间序列法采用ns拟合时间（time）变量（1，2，3，……，2557）以控制长期趋势和季节趋势，自由度设为7/年；同时将星期几（dow）按照哑变量模式纳入模型以控制星期几效应。日均相对湿度（RH）作为潜在的混杂因素，采用ns拟合后纳入模型，自由度为3[8]。

为评价结果的稳定性，通过修改臭氧的最长滞后时间、臭氧滞后-反应关系的样条函数节点数量、气温暴露-反应关系的自由度、时间序列法time变量的自由度等方法进行敏感性分析。两种方法计算获得的健康效应均采用超额危险度（excess risk，ER）表示，意为臭氧每升高10μg/m^{3}造成死亡或住院率增加的百分比。

应用R软件（R 4.2.1）进行数据分析，其中条件泊松回归的病例交叉法使用gnm程序包，分布滞后非线性模型使用dlnm程序包，样条函数使用splines程序包。检验水准均设为0.05。

3. R软件实现

主要程序代码如下：

```
data<-read.csv("cardiovascular.csv",colClasses=c(date="日期"))#读取分析数据集
library(gnm);library(dlnm);library(splines)#加载gnm等程序包
cb_o3<-crossbasis(data$O3,lag=7,argvar=list(fun="lin"),arglag=list(knots=logknots(7,2)))#设置臭氧的交叉基函数
cb_temp<-crossbasis(data$日均气温,lag=14,argvar=list(fun="ns",df=3),arglag=list(knots=logknots(14,2)))#设置气温的交叉基函数
#病例交叉法#
data$month<-as.factor(data$月份)#转换月为因子向量
data$year<-as.factor(data$年份)#转换年为因子向量
data$dow<-as.factor(data$星期)#转换星期几为因子向量
data$stratum<-with(data,as.factor(year:month:dow))#按照年、月、星期进行时间分层
data$ind<-tapply(data$死亡,data$stratum,sum)[data$stratum]#各时间层的健康结局数求和
model1<-gnm(死亡~cb_o3+cb_temp+ns(RH,3),data=data,family=quasipoisson,subset=ind>0,
eliminate=stratum)#建立臭氧与心血管疾病死亡的条件泊松回归模型
#时间序列法#
data$time<-1:length(死亡$日期)#新生成time变量
model2<-glm(死亡~cb_o3+cb_temp+ns(RH,3)+ns(time,7*7)+as.factor(dow),data=data,family=quasipoisson)#建立臭氧与心血管疾病死亡的时间序列回归模型
```

结　果

研究期间共收集心血管疾病死亡85367例，日均33.39±9.28例；住院26974例，日均10.94±5.07例。臭氧的日均浓度为94.84±41.16μg/m^3，范围4.88~254.88μg/m^3；日均气温和日均相对湿度的日均值分别为18.15±8.25℃和76.88±12.52%，详情见表2。死亡人数、臭氧和日均气温均有明显的季节性波动，死亡人数呈现为夏季少、冬季多，臭氧和日均气温则呈现为夏季高、冬季低；住院人数的季节性趋势不明显。

表 2 2015—2021 年宁波市心血管疾病健康结局和环境因素数据情况

变量	$\bar{x}$±s	最小值	P_{25}	中位数	P_{75}	最大值
心血管疾病(例)						
每日住院例数	10. 94±5. 07	0. 00	7. 00	10. 00	14. 00	39. 00
每日死亡例数	33. 39±9. 28	11. 00	27. 00	32. 00	39. 00	75. 00
环境因素						
臭氧(μg/m^3)	94. 84±41. 16	4. 88	66. 25	90. 25	120. 00	254. 88
日均气温(℃)	18. 15±8. 25	−4. 20	11. 10	18. 90	25. 10	33. 20
日均相对湿度(%)	76. 88±12. 52	32. 00	69. 00	77. 00	86. 00	100. 00

研究显示,病例交叉法和时间序列法的分析结果非常接近。臭氧暴露当天出现心血管疾病住院风险上升,暴露滞后 1~4 天时出现心血管疾病死亡风险上升,滞后−反应关系结果详见图 1。臭氧每升高 10μg/m^3,病例交叉法和时间序列法得到当天(lag0)的超额住院率分别为 0. 56%(95%*CI*: 0. 01%~1. 11%)和 0. 49%(95%*CI*: 0. 00%~0. 98%)。累积 5 天(lag0~4)的超额死亡率分别为 0. 51%(95%*CI*: 0. 11%~0. 92%)和 0. 48%(95%*CI*: 0. 10%~0. 87%)。

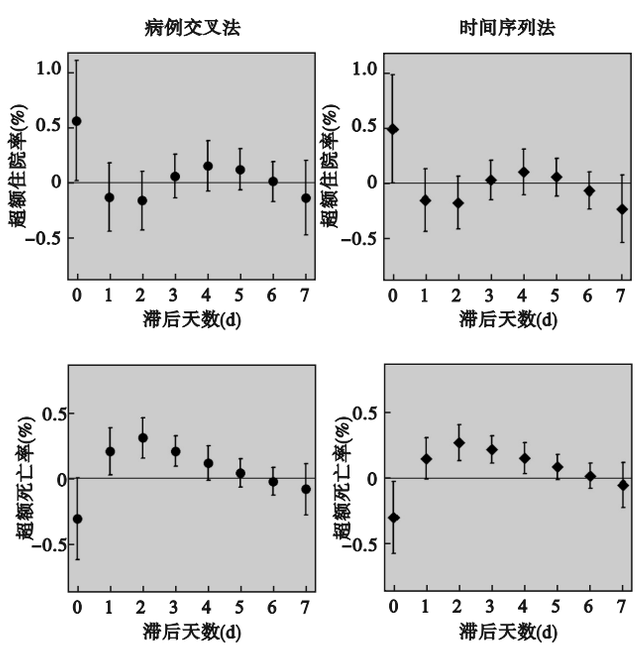


图 1 臭氧与心血管疾病住院和死亡的滞后−反应关系图

修改模型中的相关参数后两种分析方法的结果均较为稳定,表明短期内臭氧暴露与人群心血管疾病住院和死亡的风险上升相关。相较而言,时间序列法的结果容易受到 time 变量自由度选择的影响,结果稳定性比病例交叉法略差,见表 3。

表 3 臭氧对心血管病住院和死亡影响的超额危险度(ER)和 95% 置信区间(%)

分组	病例交叉法	时间序列法
修改臭氧最长滞后时间为 5 天		
住院	0. 70(0. 13,1. 28)	0. 60(0. 09,1. 11)
死亡	0. 60(0. 18,1. 03)	0. 55(0. 15,0. 95)
修改臭氧滞后−反应关系的样条函数节点为 3		
住院	0. 68(0. 09,1. 27)	0. 55(0. 03,1. 08)
死亡	0. 55(0. 14,0. 95)	0. 50(0. 12,0. 89)
修改日均气温的暴露−反应关系自由度为 4		
住院	0. 55(0. 01,1. 10)	0. 47(−0. 02,0. 96)
死亡	0. 47(0. 06,0. 87)	0. 44(0. 05,0. 83)
修改时间序列法时间变量(time)的自由度为 6/年		
住院	0. 56(0. 01,1. 11)	0. 39(−0. 09,0. 87)
死亡	0. 51(0. 11,0. 92)	0. 56(0. 19,0. 94)
修改时间序列法时间变量(time)的自由度为 8/年		
住院	0. 56(0. 01,1. 11)	0. 51(0. 02,1. 00)
死亡	0. 51(0. 11,0. 92)	0. 42(0. 03,0. 82)

注:住院的效应为滞后 0 天(lag0),死亡的效应为累积 5 天(lag0~4)

讨 论

本文同时采用两种分析方法,均发现宁波市臭氧浓度升高与心血管疾病的住院和死亡风险上升相关,该结果与以往研究一致。一项 meta 分析指出我国地面臭氧每升高 10μg/m^{3}可造成人群心血管疾病死亡率增加 0. 60%(95%*CI*: 0. 48%~0. 72%)[10];另一项 70 城市的研究表明臭氧与多种心血管疾病住院风险增加相关,超额住院率分布在 0. 40%~0. 75%[11]。近年来我国实施的清洁空气行动使得颗粒物等污染物浓度明显降低,已取得极大的健康收益,但是臭氧浓度仍维持在较高水平,且其导致的健康风险不断上升,需引起更多关注[12]。

病例交叉法优点是采用自身对照设计,可以有效控制遗传因素、吸烟、饮酒等个体差异。通过合理选择“对照期”也可以去除季节趋势、星期几效应等时间混杂偏倚。研究基于匹配数据,能够提高统计分析效率,相对节约样本量,且设计简单易行,成本较低[4,6]。利用条件泊松回归模型开展病例交叉研究,既降低了数据集整理难度,又解决了滞后性、非线性、过离散和自相关等问题,同时大幅减少模型中参数量,缩短运算时间,并保证分析结果的准确性[5]。该方法在环境流行病学领域拥有广阔的应用前景,包括扩展为“时−空”分层的病例交叉设计替代当前的“两阶段”时间序

列分析方法，以开展多城市数据研究。或者利用“分层”分析方法评估年龄、城乡等人群特征的修饰效应研究等[4,13]。

个别文献指出，尽管病例交叉法和时间序列法的效应值比较接近，但是前者的置信区间较宽，对结果估计的精准度可能低于后者[14-15]。然而，本研究显示基于条件泊松模型的病例交叉法和时间序列法的分析结果基本等价，效应值大小和置信区间均非常接近。Guo等[16]认为病例交叉法的结果稳定性主要受“对照期”选择影响，而时间序列法则主要受到时间变量自由度选择的影响。本文的敏感性分析结果也支持上述观点，两种方法的差异主要来自模型中参数的选择，尤其是对时间偏倚的控制方法。需要注意的是，病例交叉法和时间序列法在开展群体水平分析时均属于生态学研究，可能存在“生态学谬误”，并且两者均存在与时间变化相关特征引起的偏倚，例如气温、湿度等，需要在模型中进行控制。此外，本研究仅纳入一家三甲医院的住院病例来评估臭氧对心血管疾病发病的影响，可能导致选择偏倚，且限制了研究结果的外推性，仍需要更全面的分析结果进行验证。

综上所述，本研究提示基于条件泊松模型的病例交叉法与时间序列法的研究结果非常接近，均表明短期内臭氧浓度升高与心血管疾病发病和死亡风险上升相关，并且前者的敏感性分析结果更为稳定。基于条件泊松模型的病例交叉法简单易行、适用范围多样，未来可以在环境流行病学等公共卫生领域推广应用。

参　考　文　献

[1] Maclure M. The case-crossover design: a method for studying transient effects on the risk of acute events[J]. Am J Epidemiol, 1991, 133(2): 144-153.

[2] Janes H, Sheppard L, Lumley T. Case-crossover analyses of air pollution exposure data: referent selection strategies and their implications for bias[J]. Epidemiology, 2005, 16(6): 717-726.

[3] Carracedo-Martinez E, Taracido M, Tobias A, et al. Case-crossover analysis of air pollution health effects: a systematic review of methodology and application[J]. Environ Health Perspect, 2010, 118(8): 1173-1182.

[4] Tobias A, Kim Y, Madaniyazi L. Time-stratified case-crossover studies for aggregated data in environmental epidemiology: a tutorial [J]. Int J Epidemiol, 2024, 53(2): dyae020.

[5] Armstrong BG, Gasparrini A, Tobias A. Conditional Poisson models: a flexible alternative to conditional logistic case cross-over analysis[J]. BMC Med Res Methodol, 2014, 14: 122.

[6] 张彩霞，刘志东，张斐斐，等. 时间分层病例交叉研究的R软件实现[J]. 中国卫生统计, 2016, 33(3): 507-509.

[7] Gasparrini A, Armstrong B, Kenward MG. Distributed lag non-linear models[J]. Statistics in Medicine, 2010, 29(21): 2224-2234.

[8] Krishnan B, Gasparrini A, Hajat S, et al. Time series regression studies in environmental epidemiology[J]. International Journal of Epidemiology, 2013, 42(4): 1187-1195.

[9] 谷少华，贺天锋，陆蓓蓓，等. 基于分布滞后非线性模型的归因风险评估方法及应用[J]. 中国卫生统计, 2016, 33(6): 959-962.

[10] 彭星宇，王彦丁，张新民，等. 臭氧短期暴露与我国人群死亡之间关系的meta分析[J]. 环境与职业医学, 2022, 39(12): 1391-1397.

[11] Jiang Y, Huang J, Li G, et al. Ozone pollution and hospital admissions for cardiovascular events[J]. Eur Heart J, 2023, 44(18): 1622-1632.

[12] Xiao Q, Geng G, Xue T, et al. Tracking $PM_{2.5}$ and O_3 Pollution and the Related Health Burden in China 2013-2020[J]. Environ Sci Technol, 2022, 56(11): 6922-6932.

[13] Wu Y, Li S, Guo Y. Space-Time-Stratified Case-Crossover Design in Environmental Epidemiology Study[J]. Health Data Sci, 2021, 2021: 9870798.

[14] 郑浩，汪庆庆，熊丽林，等. 时间序列和病例交叉方法在环境流行病学中的应用比较[J]. 中华预防医学杂志, 2022, 56(10): 1446-1451.

[15] Slama A, Sliwczynski A, Woznica-Pyzikiewicz J, et al. The short-term effects of air pollution on respiratory disease hospitalizations in 5 cities in Poland: comparison of time-series and case-crossover analyses[J]. Environ Sci Pollut Res Int, 2020, 27(19): 24582-24590.

[16] Guo Y, Barnett AG, Zhang Y, et al. The short-term effect of air pollution on cardiovascular mortality in Tianjin, China: comparison of time series and case-crossover analyses[J]. Sci Total Environ, 2010, 409(2): 300-306.

（责任编辑：张　悦）