

## · 基础理论与方法 ·

## 病例-时间-对照设计

王涛 詹思延 胡永华

药物流行病学研究中,如果已知或预知药物的疗效,药物实际上是被选择性地给予特定严重程度病人,因而药物的使用与疾病的严重程度存在很强的关联,同时疾病的严重程度还与研究的结局,如药物不良反应密切相关,研究中如果不考虑这些因素,结果就会出现指示性混杂(confounding by the indication)<sup>[1]</sup>。指示性混杂对非实验性研究会产生致命的影响。为此,1995年 Suissa<sup>[2]</sup>提出了病例-时间-对照设计。该设计的应用条件是药物暴露随时间发生变化,暴露在两个或多个时间点进行了测量。

1. 病例-时间-对照设计的基本思想:采用传统病例对照研究时,疾病严重程度造成的混杂往往不能完全控制。这是因为一般情况下,疾病的严重程度没有精确的测量方法,无法肯定疾病严重程度在病例和对照两组间分布一致。

根据对照选择的原则,最适合的对照并不是具有相同疾病和预后因素的病人,而是病例自身。根据该观点, Maclure<sup>[3]</sup>提出病例交叉设计来研究短暂或急性效应。每个病例以自身在另一时间点上的暴露数据为对照,疾病严重程度造成的倚倚自然得到了控制。但是,病例交叉设计仅适用于短暂效应的研究<sup>[4]</sup>,因为信息完全来源于病例,如果将该设计扩展至研究慢性暴露,OR值可能会受到影响。这是因为随时间的推移,药物的使用可能会“自然增加”。药物使用的“自然增加”不仅与研究的事件相关,而且与医疗措施的改变,对药物好处的认识加深,对使用该药物信心的增加,适应症的扩大,病人对药物依赖的增加,市场的推广等均有关。这样,药物使用的自然变化趋势会混合到由病例交叉分析所得的OR值中。另设一组对照,对照组中每个研究对象也观测两次,则可以消除该影响。这种在病例交叉设计中结合传统病例对照研究设计即为病例-时间-对照设计(case-time-control design)。

2. 病例-时间-对照模型:下面从最简单的病例对照研究情况加以考虑,即感兴趣事件的发生与单一药物暴露(暴露为二分变量)的关系。

每个研究对象,无论是病例还是对照,药物暴露与否分别记为  $E = 1$  和  $E = 0$  (1 使用, 0 未使用)。暴露在两段进行测量,记作  $j = 0, 1$ , 这两段时间可以是彼此相连,也可以彼此间隔。一个研究对象在研究时的暴露相对应的时间段

称为当前期(current time),记作  $j = 1$ ,当前期之前的一段时间称为参照期(reference period),记做  $j = 0$ 。当前期长度的规定与研究的假设有关。这里规定:

$$L_{ijkl} = \logit [P(E_{ijkl} = 1)]$$

式中,  $E_{ijkl}$  代表  $i$  组, 时间  $j$ , 结局  $k$ , 研究对象  $l$  的暴露情况。 $P$  为概率,  $\logit$  为自然对数。 $i$  等于 0、1 表示对照和病例,  $j$  等于 0、1 分别代表当前时间和参照时间,  $k = 0、1$  代表事件未发生或发生,  $l = 1 \dots n$  代表  $i$  组内  $l$  研究对象。 $k = 1$  表示在当前时间内的病例组研究对象。

病例-时间-对照的模型为:  $L_{ijkl} = \mu + s_{il} + \pi_j + \theta_k$  (1) 式中  $\mu$  代表全部暴露造成的效应,  $s_{il}$  是  $i$  组中  $l$  研究对象的效应,  $\pi_j$  为时间  $j$  造成的效应,  $\theta_k$  是发生  $k$  事件的效应, 它是我们感兴趣的参数。病人的严重程度本质上是  $s_{il}$  的一部分。

对于病例组有公式(2)和(3):

$$\text{当病例组在当前期时有: } L_{111l} = \mu + s_{1l} + \pi_1 + \theta_1 \quad (2)$$

$$\text{当病例组在参照期时有: } L_{100l} = \mu + s_{1l} + \pi_0 + \theta_0 \quad (3)$$

对于对照组有公式(4)和(5):

$$\text{当对照组在当前期时有: } L_{010l} = \mu + s_{0l} + \pi_1 + \theta_0 \quad (4)$$

$$\text{当对照组在参照期时有: } L_{000l} = \mu + s_{0l} + \pi_0 + \theta_0 \quad (5)$$

时间的净效应  $\delta_\pi = \pi_1 - \pi_0$ , 感兴趣的暴露净效应  $\delta_\theta = \theta_1 - \theta_0$ 。在病例组中仅能得到  $\delta_\pi + \delta_\theta$ , 而在对照组可以估计  $\delta_\pi$ , 这样暴露对事件发生的效应  $\delta_\theta$  可以因此得到估计。

3. 病例-时间-对照模型参数的估计:要估计因为暴露对事件发生产生的效应和时间所造成的效应,可以利用传统病例对照设计中配对数据的计算方法。处于不同期间每个研究对象自身构成了一个对子。假设在一个对子中暴露是条件独立的,应用条件 logistic 回归模型就可以估计时间和结局效应,而且避免了由于  $s_{il}$  存在而需要估计的参数。模型 1 中暴露并非独立,因为要与标准的计算方法相符合,这里将暴露视作独立变量。

在条件 logistic 回归模型中分别定义变量  $E$ 、 $G$ 。 $E$  是指暴露(1 暴露, 0 不暴露),  $G$  代表研究的组别(1 病例, 0 对照)。如果将结局事件的概率记做  $R$ , 用于估计 OR 值的回归模型则是:

$$\logit(R) = \beta_0 + \beta_1 E + \beta_2 E \times G$$

与暴露相关的风险为  $\beta_2$  的反自然对数, 即  $OR = \exp(\beta_2)$ , 由暴露和研究对象组别的交互项所获得, 这里  $\beta_2$  恰恰等于式 4-5 与 2-3 之差。与时间相关的效应为对照组 ( $G = 0$ ) 中计算得到的  $OR_{co} = \exp(\beta_1)$ , 相当于  $\logit 4 - \logit 5$ ; 另一方面从病例组 ( $G = 1$ ) 中计算得到的  $OR_{ca} =$

$exp(\beta_1 + \beta_2)$  相当于  $logit2 - logit3$ , 除时间效应外还包括了与药物有关的效应。

4. 范例 Suissa<sup>[2]</sup> 为说明病例-时间-对照设计, 用它重新分析一项有关哮喘流行病学研究的数据<sup>[5]</sup>。该研究是一项关于吸入  $\beta$  阻滞剂治疗哮喘与致死性或接近致死性哮喘关系的巢式病例对照研究, 其目的是评价在治疗哮喘时长期吸入  $\beta$  阻滞剂所带来的风险。

研究者利用加拿大某省的健康保险数据库, 从中找到了 12 301 例年龄在 5 ~ 54 岁的哮喘患者, 并在 1980 ~ 1987 年期间进行了随访, 其中发现了 129 例发生致死或接近致死的哮喘患者。在巢式病例对照研究设计中, 研究者选择了 655 名对照, 利用传统的病例对照研究非条件 logistic 回归模型进行分析。研究者将所感兴趣的暴露记作  $E$ , 以一年 12 瓶  $\beta$  阻滞剂为分类点, 将暴露划分为二分变量 ( $\leq 12$  为低,  $> 12$  为高)。为了进一步评价剂量反应关系, 暴露又划分为三分变量 (低为  $\leq 12$ ,  $> 12$  为中,  $\geq 24$  为大量) 及保持连续变量进行分析 (每月使用的量)。虽然在最初的设计中, 对照与病例是匹配的, 但是配对的效率实际上微乎其微, 因此, 研究者在这里采用了非配对的分析 (表 1)。表 1 中 A 行为以二分变量划分的暴露, 粗 OR 值 4.4 (95% CI 2.9 ~ 6.7)。对当前期协变量调整后, OR 值降至 3.4 (95% CI 2.2 ~ 5.3), 进一步将参照期的协变量加以调整, OR 降至 3.1 (95% CI 1.8 ~ 5.4)。从传统的方法来看, 可以认为 3.1 即为最佳估计值。

表1 在三种模型和三种形式的暴露下采用传统非条件 logistic 回归分析的结果

当前期 $\beta$ 阻滞剂使用情况 (瓶/年)	初值		调整协变量			
	OR 值	95% CI	当前期*	95% CI	当前期及参照期**	95% CI
A 二分变量						
> 12/ $\leq 12$	4.4	2.9~6.7	3.4	2.2~5.3	3.1	1.8~5.4
B 三分变量						
13~24/ $\leq 12$	2.9	1.8~4.7	2.3	1.4~3.9	2.2	1.3~3.8
> 24/ $\leq 12$	7.3	4.5~11.8	5.4	3.3~9.0	4.9	2.3~10.2
C 连续变量#	1.9	1.6~2.2	1.7	1.5~2.0	2.1	1.6~2.9

\* 对当前期因哮喘住院的次数和口服可的松进行了调整; \*\* 对当前期因哮喘住院的次数和口服可的松, 以及在参照期内使用  $\beta$  阻滞剂和口服可的松进行了调整; # OR 值以每月每瓶计算

然而研究者认为这并不足以消除由于疾病严重程度带来的指示性混杂, 因此, 接下来采用了病例交叉分析, 所得的 OR 值为  $OR_{ca} = 3.2$  (95% CI 1.5 ~ 6.8)。此时, 病例交叉分析用到了慢性效应, 虽然由于用病例自身作为对照, 将严重程度造成的偏倚加以了控制, 但是引入了由于药物使用在时间上的自然变动趋势所造成的影响。

最后, 研究者用病例-时间-对照方法进行了分析。分析结果见表 2。表 2 中 A 行是暴露为二分变量的情况, 药物本身的效应为 1.2 (95% CI 0.5 ~ 3.0), 自然的时间效应 OR 值为 2.6 (95% CI 1.6 ~ 4.1)。由于存在这一时间效应, 因此, 不难理解为什么病例交叉设计所得到的超额相对危险为 3.2

(95% CI 1.5 ~ 6.8)。显然, 病例-时间-对照方法的结果更为合理。表 2 中 B 和 C 行为暴露划分为三分变量和连续变量的分析结果。

表2 在三种模型和三种形式的暴露下病例交叉和病例-时间-对照的分析结果

$\beta$ 阻滞剂使用 (瓶/年)	病例交叉		病例-时间-对照	
	OR 值	95% CI	OR 值	95% CI
A 二分变量				
> 12/ $\leq 12$	3.2	1.5~6.8	2.6	1.6~4.1
$\beta$ 阻滞剂 $\times G$			1.2	0.5~3.0
B 三分变量				
13~24/ $\leq 12$	2.7	1.2~5.8	2.7	1.7~4.2
> 24/ $\leq 12$	6.4	1.9~21.7	2.1	1.0~4.3
$\beta$ 阻滞剂 $\times G$				
13~24/ $\leq 12$			1.1	0.4~2.5
> 24/ $\leq 12$			3.1	0.7~12.9
C 连续变量 (瓶/月)	2.8	1.6~4.5	1.6	1.2~2.2
$\beta$ 阻滞剂 $\times G$			1.7	0.9~3.0

注:  $G$  表示研究对象组:  $G=1$  为病例组,  $G=0$  为对照组

5. 优点及局限性: 病例-时间-对照设计为解决疾病严重程度等因素造成的指示性混杂提供了一种办法, 它将研究对象作为自身的对照, 并利用对照组对药物使用自然的变化加以调整, 因此, 即使在疾病的严重程度并不能够测量时, 也可以得到药物的净效应。

与传统的方法比较, 病例-时间-对照分析结果发生了二项变化。首先, 比传统方法估计的 OR 值低。第二, 病例-时间-对照设计的结果看起来精确度下降了。这种精确度的损失, 部分因为研究对象内部存在着相关性, 使有效样本量实际上减少所致。

与病例交叉设计比较, 病例-时间-对照设计将研究范围扩展到了用于慢性暴露的研究。事实上, 如果没有药物使用的自然变化趋势, 除外时间窗口更长外, 该设计可以简化为病例交叉设计。

病例-时间-对照设计暴露期的长度取决于药物的作用机制和药物暴露的流行病学知识。暴露期长短的精确规定对分析是至关重要的。缺少精确的日期估计时, 各种合理的时间都可以进行尝试, 然后进行综合考虑。参照期是疾病过程前的一段时间, 也需要对其间的暴露加以确定。参照期必须与当前期一样长, 它可以为当前期开始前的一个时间段, 也可以为更远的时间, 这时需要对研究对象有更为严格的入选标准。所以, 采用当前期和参照期彼此靠近的方法较为简单。另一方面, 也尽可能防止了疾病严重程度改变造成的偏倚, 如果时间间隔过长, 疾病严重程度很可能发生改变。简单而言, 选择参照期必须理智地对各种因素加以权衡。

与病例交叉设计一样, 该方法也存在一定的局限性, 这些局限性与病例交叉设计类似。首先, 严重程度不变时, 以研究对象自身为对照, 研究对象本身造成的偏倚可以消除。然而, 严重程度随时间发生改变, 在病例组和对照组中变化

速度不同时,对药物净效应的估计则可能存在某种程度的偏倚。如果 OR 值接近 1,那么这种影响较小。其二,该模型是建立在研究对象内暴露在当前期和参照期条件独立,不存在参照期影响到当前期的效应。如果不独立,可能对结果有影响,但是如果 OR 接近 1,该影响则微不足道。其三,与病例交叉设计不同,该设计仍需要选择对照组,存在选择性偏倚。因为对照在该设计中仍起着至关重要的作用,因此,选择对照时需要倍加小心。

Greenland<sup>[6]</sup>甚至提出,病例-时间-对照设计与传统的研究一样存在混杂偏倚,其中包括了指示性混杂。通过病例自身配对只是消除了那些由于固定特征造成的混杂,对于变化的特征,病例-时间-对照设计使用不同时间的自身对照可能产生一些新的混杂,其结果与传统的研究结果相比可能更准确,也可能更糟糕。Suissa<sup>[7]</sup>亦进一步指出了病例-时间-对照设计用于解决指示性混杂时,有其严格定义的模型,需要满足其假设和一些条件。

无论如何,病例-时间-对照设计提供了一种解决指示性混杂的办法,这种偏倚在药物流行病学研究中广泛存在,在其他的流行病学问题中也大量存在。因此,怀疑存在指示偏倚时,不妨用病例-时间-对照加以分析,可能有所帮助。但是

在使用该方法时需要严格注意其使用条件,清楚它的优势和可能存在的问题,在下结论时加以综合考虑。

参 考 文 献

- 1 Psaty BM, Koepsell TD, Lin DJ, et al. Assessment and control for confounding by indication in observational studies. *Am Geriatr Soc*, 1999, 47:749-754.
- 2 Suissa S. The case-time-control design. *Epidemiology*, 1995, 6:248-253.
- 3 Maclure M. The case-crossover design: a method for studying transient effects on the risk of acute events. *Am J Epidemiol*, 1991, 133:144-153.
- 4 张政,詹思延. 病例交叉设计. *中华流行病学杂志*, 2001, 22: 304-306.
- 5 Spitzer WO, Suissa S, Ernst P, et al. The use of beta-agonists and the risk of death and near death from asthma. *N Engl J Med*, 1992, 326:501-506.
- 6 Greenland S. Confounding and exposure trends in case-crossover and case-time-control designs. *Epidemiology*, 1996, 7:231-239.
- 7 Suissa S. The case-time-control design: further assumptions and conditions. *Epidemiology*, 1998, 9:441-445.

(收稿日期 2001-08-24)  
(本文编辑:杨莲芬)

· 短篇报道 ·

都匀市 1999 ~ 2001 年食品卫生抽检结果分析

武建华

为进一步了解都匀市食品卫生状况,有效防范肠道传染病的流行,我们于 1999~2001 年对全市各类食品进行抽样检测,按照国家标准判定结果。

在 1999~2001 年抽检的食品中,年平均合格率以酒类(98.43%)、酱腌菜(95.08%)为最高,非发酵豆制品(40.00%)、熟肉(68.75%)为最低。结果见表 1。

都匀市 1999~2001 年食品卫生监测合格率均保持在 78.00% 以上。餐具监测合格率呈逐年上升趋势,反映出我市在食品卫生管理及贯彻《食品卫生法》方面取得了一定成

绩,但非发酵豆制品及熟肉制品合格率还很低,尤其是豆腐,2001 年合格率仅为 45.00%。不合格食品的主要问题是微生物污染严重,这种状况构成了肠道传染病发生及流行的潜在危险,必须引起重视。

为使我市食品卫生抽检合格率稳步提高,建议针对食品卫生薄弱环节加强管理,对抽检合格率较低食品要加强监督监测力度,特别是在夏秋季节,肠道传染病易于流行,增加抽检频次,加大执法力度,不允许卫生不合格食品上市销售,是防止肠道传染病流行的重要环节。

表1 都匀市 1999~2001 年各类食品卫生监测合格率(%)统计

年份	植物油	熟肉及制品	牛奶及制品	罐头	冷饮	酒类	酱油	非发酵豆制品	酱腌菜	糕点	其他	餐具
1999	80.00 (16/20)	75.00 (15/20)	100.00 (15/15)	94.73 (54/57)	66.66 (64/96)	100.00 (20/20)	70.00 (7/10)	65.00 (13/20)	100.00 (20/20)	83.67 (41/49)	100.00 (3/3)	80.41 (234/291)
2000	90.00 (18/20)	66.66 (6/9)	66.66 (6/9)	71.11 (32/45)	76.85 (83/108)	100.00 (24/24)	100.00 (10/10)	10.00 (2/20)	100.00 (20/20)	95.65 (44/46)	100.00 (1/1)	80.59 (191/237)
2001	90.00 (18/20)	65.00 (13/20)	42.85 (6/14)	70.00 (14/20)	77.35 (41/53)	95.00 (19/20)	90.00 (9/10)	45.00 (9/20)	85.71 (18/21)	83.87 (52/62)	93.18 (41/44)	80.75 (235/291)
合计	86.66 (52/60)	68.75 (44/64)	71.05 (27/38)	81.96 (100/122)	73.15 (188/257)	98.43 (63/64)	86.66 (26/30)	40.00 (24/60)	95.08 (58/61)	87.26 (137/157)	93.75 (45/48)	80.58 (660/819)

注:表中分母为检测数,分子为合格数

作者单位 558000 贵州省都匀市食品卫生监督检验所

(收稿日期 2001-11-28)  
(本文编辑:杨莲芬)