

病例-病例-时间-对照设计

王晶 卓琳 詹思延

【导读】 自身病例对照设计(如病例交叉设计、自身对照病例系列研究)虽然能控制不随时间变化的混杂因素(可测量或不可测量的)的影响,但是不能控制暴露时间变化趋势所带来的混杂。双向病例交叉设计可以控制暴露时间变化趋势。但是在药物流行病学的研究中,疾病状态通常影响后续的药物使用,因此这种双向病例交叉设计往往不适用。Suissa提出的病例-时间-对照设计将病例交叉设计和病例对照设计相结合,很好地控制了暴露的时间变化趋势所带来的偏倚。但是当病例组和对照组不能良好地匹配时,可能会引入新的选择偏倚。本文介绍一种新的研究设计方法“病例-病例-时间-对照设计”,该研究设计是“病例-时间-对照设计”的衍生,即在选择对照组时并不采用外部对照,而是将未来病例作为当前病例的对照,从而避免了暴露的时间趋势所带来的混杂。文后还通过实例分析,对此方法的优点和局限性进行阐释。

【关键词】 病例-病例-时间-对照设计;自身病例对照设计;暴露的时间变化趋势

The case-case-time-control study design Wang Jing, Zhuo Lin, Zhan Siyan. Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing 100191, China
Corresponding author: Zhan Siyan, Email: siyan-zhan@bjmu.edu.cn

【Introduction】 Although the ‘self-matched case-only studies’ (such as the case-cross-over or self-controlled case-series method) can control the time-invariant confounders (measured or unmeasured) through design of the study, however, they can not control those confounders that vary with time. A bidirectional case-crossover design can be used to adjust the exposure-time trends. In the areas of pharmaco-epidemiology, illness often influence the future use of medications, making a bidirectional study design problematic. Suissa’s case-time-control design combines the case-crossover and the case-control design which could adjust for exposure-trend bias, but the control group may reintroduce selection bias, if the matching does not go well. We propose a “case-case-time-control” design which is an extension of the case-time-control design. However, rather than using a sample of external controls, we choose those future cases as controls for current cases to counter the bias that arising from temporal trends caused by exposure to the target of interest. In the end of this article we will discuss the strength and limitations of this design based on an applied example.

【Key words】 Case-case-time-control design; Self-matched case-only studies; Exposure-time trends

在病例对照研究中,面临的最大挑战就是如何选择合适的对照。理想的对照应该是来自于病例组源人群的一个代表性样本。但是,在实际的研究当中,这样理想的对照人群很难找到。数据库为药物流行病学研究提供了丰富的资料来源,然而大部分数据库中都缺乏对照组(如一项研究维生素与脑卒中发生风险的数据库中仅包括了脑卒中患者的信息)。因此,在研究一些短暂的暴露与某些急性发病的疾病之间的关系时,学者们往往采用病例组作为其自身的对照进行研究,如Maclure^[1]提出的病例交叉设计(case-cross-over design),Farrington^[2]提出的自身对照

病例系列研究(self-controlled case-series method)等。

虽然自身病例对照设计具有很多优点,但是同样也存在着一些局限性。Navidi^[3]首先提出病例交叉设计不能控制暴露时间趋势所带来的混杂,他指出Maclure的单向回顾性对照样本只研究了结局事件发生前的暴露情况,而他提出的对称性双向病例交叉设计可研究事件发生前、后的暴露情况,这样可以控制暴露时间趋势所带来的混杂。但是当结局事件的发生对患者的暴露情况造成影响时,这种方法也不适用(例如脑卒中发生以后,患者后续药物的使用会发生改变)。Suissa等^[4]提出的“病例-时间-对照”设计也可以解决暴露的时间趋势所带来的混杂。但是,如果对照组人群选择不合理,则会引入新的偏倚(例如选择来自其他数据源的对照组)。因为

此时对照组的暴露时间趋势并不能反映病例组的暴露时间趋势。究竟采用什么样的研究设计既能够控制暴露的时间趋势所带来的混杂,同时又能避免选择外部对照而引入新的偏倚呢?

基本原理

2011 年 Wang 等^[5]提出了一种新的流行病学设计方法“病例-病例-时间-对照设计”,这种研究设计既不需要研究结局事件发生之后的暴露情况,也不需要选取外部的非病人人群作为对照。它是在病例-时间-对照设计基础上的一种延伸设计。该研究设计主要适用于数据库资料,其应用的前提条件是:短暂暴露、急性结局事件,同时也要求这些急性结局事件在时间上不连续。Wang 等^[5]采用病例-病例-时间-对照设计重新分析了有关脑卒中患者的数据,具体说明了如何采用这种自身对照的研究设计,控制暴露时间变化趋势带来的混杂。

病例-病例-时间-对照设计与病例-时间-对照设计的不同之处在于对照组的选择,该设计的对照组人群是将要发病但当前未发病的人群,称之为“未来病例”。通过将已经患病的人群(病例组)和即将患病的人群(对照组)进行对比,从而控制暴露的时间变化趋势所带来的偏倚。该研究设计的假说是“未来病例”与“外部对照”相比,其提供的暴露情况或暴露的时间变化趋势能够更好地估计病例组人群的暴露率或暴露时间变化趋势。此外,这种研究设计还能减小选择外部对照时所带来的选择偏倚。

在患者数据库中,“病例”是指在选定的时点发生了所关心的结局事件的研究对象;“未来病例”是在选定的时点未发生结局事件但即将发生结局事件的研究对象。“当前期”(current time)是结局事件发生前的一段时间;“对照期”(reference time)是“当前期”之前的一段时间,在此时间段内,没有结局事件发生。病例和未来病例的“对照期”都是从研究对象暴露历史中选取的一段时间。此外,病例和未来病例的当前期和对照期是相匹配的。通过计算病例和未来病例的暴露比值比,可以估计出控制了暴露时间变化趋势之后暴露与所关注结局之间的关系,见图 1。

在统计学分析模型中,可采用如下的变量进行标记。定义变量 Y 为所关注的结局事件(1=事件发生,0=事件未发生);变量 T 代表时间段(1=当前期,0=对照期);变量 E 代表暴露水平(1=暴露,0=未暴露);变量 C 代表患病情况(1=病例,0=未来病例)。

病例-病例-时间-对照设计的统计学分析方法

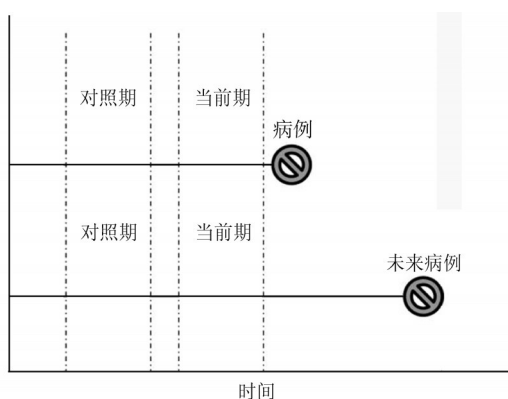


图 1 病例-病例-时间-对照设计

和病例交叉设计的统计学分析方法相近。只是增加了一组“未来病例”对照组,进而对暴露时间变化趋势进行调整。在病例-病例-时间-对照设计中,每个研究对象所抽取的当前期和对照期也是采用 1:1 的比例进行匹配的。病例-病例-时间-对照设计的模型为:

$$\text{logit}(\pi_{qi}) = \alpha_q + \beta_1 T_{qi} + \beta_2 T_{qi} \times C_{qi}$$

该模型应用的前提假设为:在病例组和未来病例组中,暴露随时间变化的趋势是相同的,在这种情况下,才能利用未来病例组的对照期对病例组的时间效应进行估计。

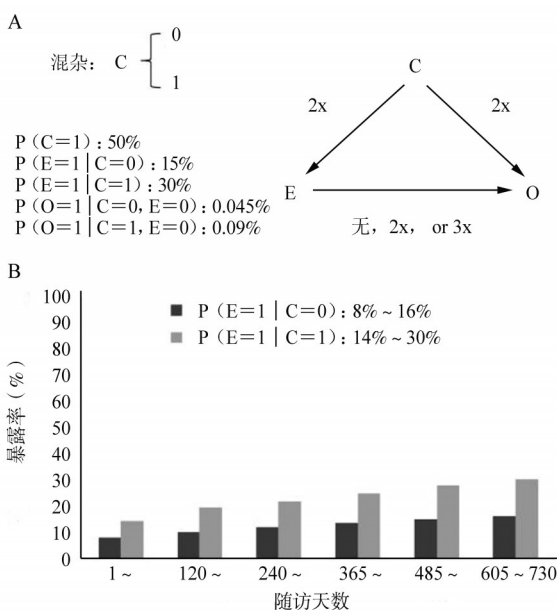
Wang 等^[5]通过一个模拟数据,将病例交叉设计、病例-病例-时间-对照设计和病例-时间-对照设计在 3 种不同场景下所得到的结果进行对比。模拟数据包括了 100 个队列(100 000 个短暂暴露的研究对象),随访时间为 2 年。此外,研究还设定了 3 个场景,场景 1 为无不可测量的混杂;场景 2 为存在不随时间变化的混杂;场景 3 为存在不随时间变化的混杂因素和逐渐增加的暴露率。研究假定在场景 2 和 3 中,二分类的且不随时间变化的混杂因素在源人群中出现的概率为 50%。如图 2A 所示,当存在混杂因素时(C=1),研究对象的暴露率和结局发生率均为无混杂时(C=0)的 2 倍。此外,研究还假定在场景 3 中,暴露率呈单调递增且在 2 年的随访期结束时,暴露率为随访初期的 2 倍(此假定在药物流行病学的实际研究中是合理的),如图 2B。研究假定暴露和结局之间存在 3 种真实联系,即 OR=1, 2, 3。该模拟队列是由 SAS 9.2 软件实现的。统计学分析在 Stata 10.0 软件下进行,见表 1。

结果发现,当假定不存在不可测量混杂因素且暴露测量良好时,病例交叉设计和病例-病例-时间-对照设计均可以得到无偏倚估计(D_β=0.0)。当存在不随时间变化的混杂因素时,由于病例交叉设

表 1 3 种研究设计下的分析结果

分析 方法	暴露(E)和结局(O)之间的真实联系								
	$\beta^a=0.0, OR=1.0$			$\beta=0.7, OR=2.0$			$\beta=1.1, OR=3.0$		
	Mean β	Mean s	D_β^b	Mean β	Mean s	D_β	Mean β	Mean s	D_β
病例交叉									
场景 1: 无混杂因素	0.0	0.03	0.0	0.7	0.02	0.0	1.1	0.02	0.0
场景 2: 不随时间变化的混杂因素	0.0	0.03	0.0	0.7	0.03	0.0	1.1	0.03	0.0
场景 3: 不随时间变化的混杂因素和逐渐增加的暴露率	0.0	0.04	0.1	0.8	0.03	0.1	1.2	0.03	0.1
病例-病例-时间-对照									
场景 1: 无混杂因素	0.0	0.05	0.0	0.7	0.04	0.0	1.1	0.04	0.0
场景 2: 不随时间变化的混杂因素	0.0	0.06	0.0	0.7	0.05	0.0	1.1	0.05	0.0
场景 3: 不随时间变化的混杂因素和逐渐增加的暴露率	0.0	0.06	0.0	0.7	0.05	0.0	1.1	0.05	0.0
病例-时间-对照									
场景 3: 不随时间变化的混杂因素和逐渐增加的暴露率	0.1	0.07	0.1	0.8	0.07	0.1	1.2	0.06	0.1

注: ^a β : 回归系数, $OR = \exp|\beta|$; ^b D_β : 估计 OR 值与真实 OR 值之间的差异; 0.1 表示有 10% 的差异; 0.2 表示有 20% 的差异



注: A: 混杂因素; B: 混杂因素和暴露率

图 2 不随时间变化的混杂因素和暴露率的变化

计和病例-病例-时间-对照设计采用了个体内部对照方法,在回归模型中调整了此混杂因素的影响,因此得到的也是无偏倚结果($D_\beta=0.0$)。但是,当暴露率随着时间变化逐渐增加时,无论暴露和结局之间是否存在关联,采用病例交叉设计方法所得到的结果均存在偏倚($D_\beta=0.1$);当采用病例-时间-对照设计时,如果对照组来自于源人群且其暴露率不随时间变化时,对照组的暴露时间变化趋势不能反映病例组的暴露时间变化趋势,此时会得到有偏倚的估计值($D_\beta=0.1$)。而采用病例-病例-时间-对照设计则能够很好地控制由于暴露时间变化趋势所带来的偏倚($D_\beta=0.0$)。

实例分析

以前文中提到的一项研究维生素使用和脑卒中

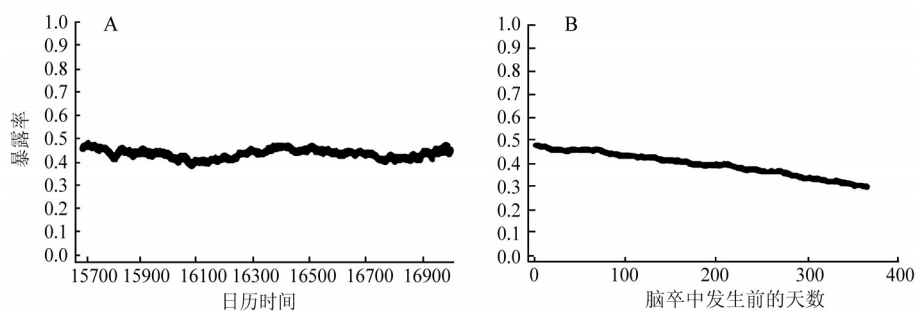
发生风险的研究为例。该研究采用的是美国退伍军人(Veteran's Health Administration, VA)数据库,该数据库包括了研究对象的药物处方、实验室检查、入院和出院等信息。研究者采用的数据是从 2003—2006 年入住某一家 VA 医院且诊断为脑卒中的所有患者的信息。对每例病例,均记录了维生素的开药日期以及用药天数。研究的暴露时间定义为维生素开药日期到停止发药日期之间的时间段。将患者因脑卒中入院前 30 d 至入院前 1 d 记为当前期,将入院前的 120 d 至 90 d 记为对照期。如果患者在当前期或对照期这 30 d 内,使用了维生素且持续时间 ≥ 3 d,则认为该患者在当前期或对照期暴露于维生素。当确定好病例组后,将在病例组当前期内未患病的患者作为对照组,即未来病例组。未来病例和病例采用了 1:1 的比例进行匹配,匹配的因素包括了年龄(± 1 岁)、性别等。

图 3 显示,通过观察脑卒中发生过程的日历时间,并未发现维生素使用率存在恒定的时间变化趋势(图 3A)。而当观察脑卒中发生前一年的这段时间时,却能发现暴露率随时间呈稳定的增长趋势(图 3B)。

采用病例交叉设计方法时,分析结果显示,当暴露于维生素时,脑卒中的基线患病风险可增加 50% ($OR=1.5, 95\%CI: 1.3 \sim 1.7$)。但是,采用病例-病例-时间-对照设计方法时,分析结果显示,维生素的使用并未增加脑卒中的患病风险 ($OR=1.1, 95\%CI: 0.9 \sim 1.3$)。

讨 论

在本实例中,通过病例交叉设计的方法得到的估计值提示当短时间暴露于维生素时,脑卒中的危险增加。这种情况可以解释为当患者的身体状况下降或出现脑卒中的预警症状时,他们往往倾向于去



注:A:日历时间;B:脑卒中发生前的365 d

图3 暴露的时间变化趋势

寻求治疗。这样就导致了在脑卒中发生前,患者暴露于多种药物治疗的可能性增加(检出症候偏倚)。通过计算与病例匹配的未来病例时,可以估计在结局事件发生以前药物治疗的变化趋势。这也就是病例-病例-时间-对照设计优于病例对照设计的原因。在本实例当中,通过使用未来病例调整了暴露时间变化趋势的影响之后,结果显示,尚不能证明短时间暴露于维生素能增加脑卒中的风险。

病例-病例-时间-对照设计通过研究对象的自身对照以及利用未来病例作为外部对照组,很好地控制了暴露时间变化趋势对暴露和结局之间关系的影响,从而得到了药物的净效应。

但是,在实施病例-病例-时间-对照设计的过程中,需要考虑以下几点。首先,病例-病例-时间-对照设计只适用于研究短时间暴露所引起的急性事件。即对照期的暴露在当前期没有残余作用。确定对照期和当前期的时间间隔主要是基于对暴露作用的现有生物学认识(例如:药效学)。此外,可以根据不同时间间隔进行敏感性分析,从而验证结果的稳定性。

其次,在实施病例-病例-时间-对照设计的过程中,需要考虑研究随访时间的长短。在随访期内,特别是在随访末期才出现健康结局的病例,由于即将发病的人数将会越来越少,因此没有办法找到合适的未来病例。这个特点将会严重影响该研究设计的统计学效力。在实际研究中,那些没有找到相匹配的对照病例将不能纳入分析。也就是说,每个病例所匹配的未来病例数目越多,其统计学效力越高。

此外,还要考虑病例和相匹配的未来病例出现结局事件的时间间隔。选取的当前期与未来病例结局事件的发生时点应该有足够长的时间间隔,这样才可以认为未来病例的暴露情况和结局事件发生是相互独立的。如果所研究的暴露和结局之间确实是有联系,且当前期和结局事件的发生时间点很相近时,则会得到有偏倚的结果。另一方面,当前期和结

局时间的发生时间点应该尽可能相近,在这种情况下,采用未来病例估计的暴露时间变化趋势才能反映病例组暴露的时间变化趋势。这方面的考虑十分重要,特别是当暴露时间变化趋势是非线性或变化很快时。此外,还可以根据不同的结局事件时间间隔进行敏感性分析,验证结果的稳定性。

在研究设计当中,病例和未来病例除了在时间上需要匹配以外,在年龄、性别或是居住地等其他方面也需要匹配。虽然多因素的匹配可能会提高估计值的真实性,但是如果匹配的因素过多,导致可以相匹配的病例减少,反而降低了研究的效力。

总之,病例-病例-时间-对照设计是在先前的病例交叉设计^[6]和病例-时间-对照设计^[7]的基础上发展起来的,它可以在不使用外部对照等情况下,调整暴露时间变化趋势。此外,该设计还能有效地控制检出症候偏倚。

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] Farrington CP. Control without separate controls: evaluation of vaccine safety using case-only methods [J]. *Vaccine*, 2004, 22(15): 2064-2070.
- [3] Navidi W. Bidirectional case-crossover designs for exposures with time trends [J]. *Biometrics*, 1998, 54(2): 596-605.
- [4] Suissa S. The case-time-control design [J]. *Epidemiology*, 1995, 6(3): 248-253.
- [5] Wang S, Linkletter C, Maclure M, et al. Future-cases as present controls to adjust for exposure-trend bias in case-only studies [J]. *Epidemiology*, 2011, 22(4): 568.
- [6] Zhang Z, Zhan SY. The case-crossover design [J]. *Chin J Epidemiol*, 2001, 22(4): 304-306. (in Chinese)
张政,詹思延. 病例交叉设计 [J]. *中华流行病学杂志*, 2001, 22(4): 304-306.
- [7] Wang T, Zhan SY, Hu YH. The case-time-control design [J]. *Chin J Epidemiol*, 2002, 23(2): 115-117. (in Chinese)
王涛,詹思延,胡永华. 病例-时间-对照设计 [J]. *中华流行病学杂志*, 2002, 23(2): 115-117.

(收稿日期:2014-06-17)

(本文编辑:王岚)