

单纯病例研究方法在流行病学研究中的应用

俞顺章 蔡琳 穆丽娜 张作风

流行病学在公共卫生决策与计划、预防疾病和促进健康方面正在发挥着愈来愈重要的作用。分子流行病学研究的兴起为揭开疾病的“黑匣子”之谜创造了条件,为阐明致病因素与发病机制提供了有利的工具,与此同时也促进了该领域方法学的研究,并取得了令人瞩目的进展。其中,单纯病例研究(case-only study)方法在探索环境-基因交互作用、评价环境暴露和生物效应以及寻找突发病因的研究中已崭露头角,并在不断发展和完善之中。单纯病例研究的基本特征是研究对象仅用病例而不用对照。按其应用范围可分为四类:病例分布研究(case-distribution study)、病例-病例研究(case-case study)、病例-系列研究(case-series study)和病例-交叉研究(case-crossover study),现分别介绍如下。

1. 病例分布研究:随着传染病和寄生虫病被控制,慢性非传染病的发病率和死亡率在不断上升,成为威胁人类健康的主要疾病。慢性非传染病有复杂的原因,是涉及遗传和环境的多因子疾病(multifactorial disease)。因此,研究致病基因和遗传易感基因的分布特征与疾病的关系,对阐明慢性疾病发生的规律、制定防制策略和措施有着重要意义。病例分布研究,尤其是基因型分布研究(case-genotype study)近年来已被很好地应用到分子流行病学的研究中。

依据孟德尔式群体(Mendelian population)的遗传平衡定律(law of genetic equilibrium),许多遗传学指标常按遗传特性或基因多态规律分布。病例分布研究把调查所得的病例的基因型分布与源人群(source population)的基因型分布进行比较,分析危险因素(遗传易感基因型或致病基因型)与疾病的联

系,测定暴露频率的危险比值。该方法由于仅用病例作为研究对象,避免了对照的选择,降低了研究的费用和因选择对照而引起的选择偏倚。

例1 已知某遗传基因的等位基因B与b和3个基因型以及源人群该基因型的分布(表1)。

表1 基因型频率(%)分布

对象	BB	Bb	bb
人群	0.25	0.50	0.25
病例	0.40	0.40	0.20

显而易见,纯合子基因型BB与疾病有关,BB基因型增加了疾病的危险。当连锁基因的实际频率偏离群体的等位基因的期望频率而呈现连锁不平衡(linkage disequilibrium)时往往与疾病有关。该研究也可分析不同基因型在发病中可能的易感性差异。此研究的关键在于获得人群的基因型分布,这些分布可从已发表的文献中查得,或在人群中抽样检测。该研究的优点是能很快筛检出疾病有关的基因型,但是不能下统计学结论。

日本已报道了25种基因多态的研究,最近又检测了241例非癌症患者50种基因多态,为基因分布特征的研究提供了基础资料^[1]。毫无疑问,世人注目的人类基因组计划的完成将为基因与疾病的研究建立极其丰富而宝贵的信息库,也增加了病例分布研究的可行性和有效性。

2. 病例-病例研究:该方法主要用于测定交互作用(interaction,或修饰效应,modification effect)。在分子流行病学研究中,基因环境交互作用的评估是研究的主要目标。交互作用是研究者极力期望发现的客观现象,可以为多因素综合作用所致疾病的病因探讨提供科学依据。但是,评估交互作用所需的样本量一般是评估相同大小主效应所需样本量的4倍。例如,评估交互作用的比值比(OR)是2.0,所需样本是1200例,相当于评估相同大小(OR=2.0)的主因素效应所需样本量($n=300$)的4倍(表2)。鉴于分子流行病学研究的费用高,亟待发展一种所需样本量较小的新的研究方法。由此,病例-病例研究

基金项目:全国优秀博士学位论文专项资金资助项目(200157)

作者单位:200032 复旦大学公共卫生学院预防医学研究所(俞顺章、穆丽娜);福建医科大学流行病学教研组(蔡琳);University of California, Los Angeles School of Public Health, California, USA(张作风)

应运而生。

我们需假设基因和环境之间相互独立。例如已知基因和环境的暴露率都是 50% ,检验把握度(1 - β)为 80% ,检验显著性水平为 0.05 ,交互作用的 OR 值为 2.0。采用常规的病例对照研究方法需要样本 1 200例(600 名对照 ,600 例病例) ,而应用病例-病例研究所需的样本数(病例数)仅为 300 例。由此可见 ,病例-病例研究所需样本量明显减少(表 2)。

表2 病例对照研究与病例-病例研究所需样本量比较

类型	主效应(OR = 2.0)	交互效应(OR = 2.0)
病例-对照	病例(n = 150)	病例(n = 600)
	对照(n = 150)	对照(n = 600)
病例-病例	-	病例(n = 300)

在评估交互作用时 ,常规的病例对照研究一般用 2 × 4 表来估计联合效应和两个主效应(环境、基因) ,其模式如表 3。

表3 病例对照研究模式表

环境	基因	病例组	对照组	OR 值
+	+	a	b	OR ₁₁ = ah / bg
+	-	c	d	OR ₁₀ = ch / dg
-	+	e	f	OR ₀₁ = eh / fg
-	-	g	h	OR ₀₀ = 1

在此 ,OR₀₀为共同对照组 ,OR₁₀为环境的主效应 ,OR₀₁为基因的主效应 ,OR₁₁为环境-基因的联合作用。应用表 3 的结果 ,可以估计交互作用 OR_{int}为 :

$$OR_{int} = OR_{11} / OR_{01} OR_{10} \\ = (ah / bg) [(ch / dg) \times (eh / fg)] \\ = (ag / ce) (bh / df)$$

因为(ag / ce)等于病例的 OR(OR_{case}) (bh / df) 等于对照的 OR(OR_{control}) ,所以交互作用的 OR 等于两者之比 ,即 OR_{int} = OR_{case} / OR_{control}。按照评估交互作用的前提假设 ,基因和环境相互独立 ,OR_{control}应等于 1 ,所以 OR_{int} = OR_{case}。由此可见 ,病例-病例研究可有效地代替病例对照研究用于交互作用的评估。

例 2 5 ,10-亚甲基四氢叶酸还原酶(MTHFR)

和 β 胱硫醚合酶基因多态(CBS)与神经管缺陷(NTDs)的研究^[2](表 4)。

表4 MTHFR 和 CBS 与 NTDs 危险的病例对照研究

MTHFR	CBS	病例组例数	对照组例数	OR 值(95% CI)
+	+	7	5	5.2(1.4 ~ 21.2)
+	-	19	34	2.1(1.1 ~ 39.0)
-	+	16	76	0.8(0.4 ~ 1.4)
-	-	85	315	1.0

假设 MTHFR 与 CBS 基因频率相互独立 ,采用常规的病例对照研究 OR_{int} = 3.1(95% CI : 0.8 ~ 13.1) ,采用病例-病例研究 OR_{int} = 2.0(95% CI : 0.6 ~ 6.0)。NTDs 属多基因遗传 ,有家族聚集性 ,基因的交互作用分析提示携带较多的易感基因的个体增加疾病发生的风险。

基因还可修饰机体对环境因素的易感性。

例 3 MTHFR677 与饮用水交互作用分析见表 5 和表 6。

以上的结果显示基因(MTHFR677)与环境暴露(饮用水)的交互作用。病例对照研究结果 ,调整后的交互项 OR_{int}为 2.74(0.99 ~ 7.59) ,病例-病例研究调整的 OR_{int}为 2.36(1.04 ~ 5.36)。病例-病例研究所得的 OR_{int}略降低 ,但 95% CI 有缩小的趋势 ,对交互作用评估的精确性较高。

应用病例-病例研究前提条件是需假设环境和基因或基因和基因间相互独立。相对于常规的病例对照研究 ,所需的样本量小。病例-病例研究是评估相乘交互作用的有效方法 ,但不能评价主效应。

3. 病例-系列研究 :该研究主要应用于测定病因的异质性(etiological heterogeneity) ,并能评估机体对环境暴露的生物效应。在肿瘤分子流行病学研究中 ,生物标志检测方法相对比较复杂 ,组织标本采集受到某些限制。比如在很多情况下 ,组织标本仅能从病例中获得 ,而不能从对照中获得。采用病例-系列研究可解决这一难题。

表5 MTHFR677 与饮用水交互作用的病例对照研究

MTHFR 基因型	饮用水	病例组例数(%)*	对照组例数(%)*	粗 OR 值(95% CI)	调整 OR 值(95% CI)
C/C	无	35(19.8)	8(26.3)	1.00	1.00
C/C	有	9(5.1)	2(7.7)	0.88(0.38 ~ 2.07)	0.84(0.35 ~ 2.02)
Any T	无	83(46.9)	18(53.9)	1.16(0.73 ~ 1.85)	1.22(0.75 ~ 1.99)
Any T	有	50(28.2)	4(12.1)	3.10(1.76 ~ 5.48)	2.79(1.54 ~ 5.08)
交互项	-	-	-	3.04(1.14 ~ 8.11)	2.74(0.99 ~ 7.59)

* 构成比(%)

表6 MTHFR677 与饮水水交互作用的病例-病例研究

饮水水	MTHFR 677Any T 例数 (%) [*]	MTHFR 677C/C 例数 (%) [*]	粗 OR 值 (95% CI)	调整 OR 值 (95% CI)
	有	5(37.6)		
无	8(62.4)	35(79.6)	2.3(1.04~5.28)	2.3(1.04~5.36)

* 构成比 (%)

例4 研究报道膀胱癌与吸烟有关($OR = 2 \sim 3$) 40% ~ 60% 的膀胱癌组织中检测到 p53 基因突变。为探讨以下问题: ① p53 基因突变膀胱癌与 p53 基因未突变膀胱癌的病因是否一致? ② p53 基因突变膀胱癌是否与吸烟有关? ③ p53 基因未突变膀胱癌是否与吸烟无关? ④ 吸烟与 p53 基因突变在膀胱癌发生中是否有关? 可采用病例对照研究设计。但因 p53 基因突变是肿瘤标志, 仅能从膀胱组织和正常组织检测, 很难或不可能从对照中取样, 所以 p53⁺ 和 p53⁻ 基因组用同一对照组比较, 表7 是吸烟、p53 基因突变与膀胱癌的研究实例。

表7 吸烟、p53 突变与膀胱癌^[3]

吸烟状态	病例组例数		对照组例数
	p53 突变 (p53 ⁺)	p53 未突变 (p53 ⁻)	
吸	34(a ⁺)	43(a ⁻)	81(b)
不吸	10(c ⁺)	21(c ⁻)	64(d)

分别计算估计相对危险度为:

$$OR_{(p53^+)} = (a^+ / d) / (c^+ / b) = (34 / 64) / (10 / 81) = 2.69$$

$$OR_{(p53^-)} = (a^- / d) / (c^- / b) = (43 / 64) / (21 / 81) = 1.62$$

传统的病例-对照研究: $OR_{(p53^+)} / OR_{(p53^-)} = 1.66$

病例-系列研究:

$$OR_{(p53^+)} / OR_{(p53^-)} = (a^+ + c^-) / (c^+ + a^-)$$

以上数据如应用 logistic 回归, 其结果亦相同(表8)。

表8 吸烟、p53 突变对膀胱癌影响的 logistic 回归分析

研究类型	参数	粗 OR 值 (95% CI)	调整 OR 值 [*] (95% CI)
病例-对照	OR_{p53^+}	2.6(1.23 ~ 5.87)	3.6(1.41 ~ 9.29)
	OR_{p53^-}	1.6(0.87 ~ 3.00)	2.1(0.98 ~ 4.54)
	$OR_{(p53^+)} / OR_{(p53^-)}$	1.6(0.69 ~ 4.01)	1.7(0.66 ~ 4.43)
病例-系列	OR	1.6(0.69 ~ 4.01)	1.7(0.63 ~ 4.66)

* 调整年龄; OR_{p53^+} 指 p53 突变, OR_{p53^-} 指 p53 未突变

结果说明, 吸烟对 p53⁺ 膀胱癌的作用比对 p53⁻ 膀胱癌的作用强, 吸烟可能与 p53 基因突变有关。应用病例-系列研究结果与病例对照研究结果一致, 调整混杂因素后两种研究方法结果相同。

应用病例-系列研究方法仅用病例组, 样本量相对缩小。该方法在测定生物标志物、研究机体对环

境暴露的生物反应(如获得性 p53 基因突变)时尤为有效。

4. 病例-交叉研究: 把病例在暴露危险时间段与病例在疾病发生前的另一时间段的暴露分布进行比较, 以判断暴露危险因素与某病(或事件)有无关联的研究方法称病例-交叉研究, 该方法类似于交叉设计的病例对照研究。

这种研究设计把最接近疾病(或事件)发生的一个短时间段作为暴露危险时间段(risk period), 把疾病发生前的一个或多个时间段作为与病例配比的“对照”阶段(control period), 即把在此阶段中病例的暴露作为与病例配比的“对照”的暴露, 分析比较病例在危险时间阶段和对照时间阶段的暴露频率差异, 获得估计相对危险度 OR 值(图1)。

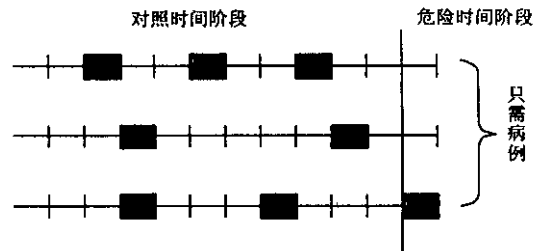


图1 病例-交叉研究模式图

该研究的前提条件是假设混杂因子在整个观察期间不发生系统改变; “暴露”必须有变化, 且断续发生。如眼睛的颜色和血型是稳定不变的, 就不可用此方法研究; 暴露效应必须在暴露后的短时间内产生(诱导期短), 即效应是由时间上接近的暴露触发产生, 并且可被测量; 无延期效应(carryover effect), 避免将过去的暴露作为此疾病发生的原因, 否则需考虑两次暴露间的洗脱期(wash-out period)。

病例-交叉研究主要用于突发事件和短暂期暴露效应的病因研究^[4, 5]。

例5 应用病例-交叉设计研究移动电话使用与车祸发生关系。以 699 例车祸病例为研究对象, 调查发现他们在车祸发生一周内共打过 20 513 次电话, 发生车祸前 10 min 内(risk period) 24.32% 打过电话, 车祸发生一周前的 10 min 内(control period) 5.29% 打过电话。配对分析结果 OR 值为 6.5(95% CI 4.5 ~ 10.0)(表9)。

吃安眠药与发生车祸亦可用这种方法来进行分析^[6]。性活动与心肌梗死、咖啡因摄入、酒精消耗、药物暴露、剧烈活动的关系也可应用病例-交叉方法来研究。在病例-交叉方法研究中, 每个病例及其配

比的对照都自动地在个体不会改变的所有特征上配比(如性别、种族、出生日期等)。因此,不论是否对这些特征做了测量,可能的混杂因素都得到控制。

表9 打移动电话与车祸的发生关系*

	risk period(车祸的发生期)		
	未打电话次数	曾打电话次数	
control period	未打电话次数	505	157
对照期	曾打电话次数	24	13

注:RR = 157/24 = 6.5(95% CI 4.5 ~ 10.0);* 摘自 NEJM, 1997, 336:453-458.

综上所述,单纯病例研究具有病例对照研究的优点,适用于罕见疾病的病因研究,相对省时省力,花费较少,易于实施。因为用来估计或检验效应的研究对象只有病例,避免了病例对照研究中对照的选择偏倚,也避免了病例与对照间可能存在的不均衡现象。单纯病例研究的病例选择与病例对照研究的研究对象选择原则是相同的。该方法的各类设计有其独特的长处(以上已分别做了讨论),应用时应

注意各类设计的前提假设、适用范围和局限性。随着该方法的不断发展和完善,将在对人类复杂疾病的病因研究中发挥应有的作用。

参 考 文 献

- 1 Hamajima N, Saito T, Matsuo K, et al. Genotype frequencies of 50 polymorphisms for 241 Japanese non-cancer patients. *Jap J Epidemiol*, 2002, 12:229-236.
 - 2 Yang Q, Khoury MJ, Sun F, et al. Case-only design to measure gene-gene interaction. *Epidemiology*, 1999, 10:167-170.
 - 3 Zhang ZF, Cordon-Cardo C, Littlefield LG. Methodological issues in the use of tumour markers in cancer epidemiology. In: Toniolo P, Boffetta P, Shuker DEG, et al. eds. *Application of biomarkers in cancer epidemiology*. IARC Scientific Publications, 1997, 142:201-214.
 - 4 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.
 - 5 Redelmeier DA, Tibshirani RJ. Association between cellular telephone calls and motor vehicle collisions. *New England J Med*, 1997, 336:453-458.
 - 6 Barbone F, McMahon AD, Davey PG, et al. Association of road-traffic accidents with benzodiazepine use. *Lancet*, 1998, 352:1331-1336.
- (收稿日期 2002-08-14)
(本文编辑 张林东)

· 疾病控制 ·

一起毒蜘蛛伤人事件的流行病学调查

赵连国 张天庆 王振家 赵春贵 曹志馨 依不拉英

2002年6月份以来,在新疆哈密东天山南部平原地区的部分乡村和团场,连续发生了多起毒蜘蛛伤人事件。据初步调查分析,有60多人受到毒蜘蛛的袭击。现将此事件的流行病学调查结果报道如下。

1. 流行病学调查 我们对在地区中心医院和农十三师红星医院住院和急诊科就诊的36例患者进行了流行病学个案调查,结果:①被咬伤的36例患者中,男性占61.11%,女性占38.89%。因为男性从事户外劳动和各种活动较多,所以被毒蜘蛛袭击的概率较高。②以农民患者最多,占58.33%;其次是民工占16.67%,学生占13.89%,工人占11.11%。③以7月份毒蜘蛛伤人最多,占63.89%;6月份占27.78%;8月份占8.33%。考虑与毒蜘蛛的生活习性和该季节农田活动较多有关。④地区分布,以哈密市为中心,东起哈密市大泉湾乡(距市区50多公里),西至农十三师火箭农场(距市区20多公里),南到哈密市南湖乡(距市区30多公里),共13个区、乡、团场均有被毒蜘蛛咬伤的患者,总面积约1000 km²,涉及总人口10万多人。⑤经专家现场调查鉴定,当地群众发现的黑色毒蜘蛛是间斑寇蛛(俗称黑寡妇),属剧毒蜘蛛。

7、8月份是间斑寇蛛的繁殖期,母蛛在一旁守护,每受侵扰,即奋起反抗。间斑寇蛛毒液中含有一种蛋白神经毒素,能作用于脊椎动物神经突触,引起神经递质大量释放。人被咬伤后有针扎感,伤口处可呈苍白、发红或起荨麻疹,局部疼痛或压痛。全身症状有乏力、软弱、头晕、恶心、呕吐、出汗、腹肌痉挛、抽搐、发热、白细胞计数增高、双足麻木、刺痛感以及特征性的足跟部烧灼感。严重病例血压先升高后降低,呈休克状态,并有呼吸困难、谵妄,小儿常出现惊厥症状。与本地区患者有许多相似之处。

2. 采取的预防措施及效果 经专家指导,提出了以下预防措施:①在毒蜘蛛密集的农田劳动时,要穿好裤袜,扎紧裤管,不要在田间、地头、打麦场、杂草、植物丛中休息。②一旦被咬伤,要立即在伤口近心脏一侧用绳或布带扎紧,然后用嘴反复吸吮伤口或用烟头(火)烧灼伤口,并立即送就近的医院治疗。③采用人工捕杀的方法,消灭家庭院落和田间地头的毒蜘蛛及其卵袋,降低毒虫密度。

以上措施经电视、广播、传单等形式大力宣传后,7月下旬患者明显减少,到8月初,延续两个多月的毒蜘蛛伤人事件基本平息。

作者单位 839000 新疆哈密地区卫生防疫站防疫科(赵连国、张天庆、王振家、曹志馨、依不拉英);兵团农十三师卫生防疫站(赵春贵)

(收稿日期 2002-11-10)

(本文编辑:尹廉)