病例—对照研究中的偏倚、混杂和因果关联概念

中国医学科学院肿瘤研究所 刘韵源

编者按：本文系刘韵源同志根据IARC于1980年邀请Breslow和Day编写出的《Statistical Methods in Cancer Research Volumel-the analysis of Case-control studies》一书的译稿为基础，并参阅近期文献编写出《病例—对照研究统计方法讲座》，共分十讲。为适应我国读者的要求和建议，本刊将分期连续刊载。原中国医学科学院肿瘤研究所将不再成册出版，特告拟订阅者注意。十讲题目如下：

（一）病例—对照研究中的偏倚、混杂和因果关联概念

（二）相对危险度和归因危险度

（三）两暴露水平资料及统计推断

（四）2×2表的合并信息与混杂控制

（五）多暴露水平分层资料的分析

（六）两暴露水平1：1配对数据的分析

（七）两暴露水平1:M配对数据的分析

（八）多暴露水平1：1配对数据的统计推断

（九）多层资料的无条件Logistic回归模型（上），（下）

（十）配对数据的条件Logistic回归分析

自从Mantel和Haenszel[1]首次发表论述病例—对照研究资料统计分析的经典性文献以来，已经过去了二十年。他们的方法深受流行病学和统计学工作者的欢迎，在癌症与其它疾病的病因与防治研究中广为应用，成为流行病学文献中最常引证的论文之一，影响深远且至今不衰。另一方面，随着时间的推移，一些新的统计概念和方法也日益为人们所接受和采用，并大有后来居上之势，使得病例—对照研究资料的分析多姿多彩。病例—对照研究方法是一项年轻技术，仍在不断发展和完善，但不言而喻的事实是，它已成为现代癌症流行病学的基础。本讲座以深入浅出的方式向读者较系统地介绍这方面的知识，共分十讲，包括经典的列联表分析技术和先进的Logistic模型描述方法。病例—对照研究的一个重要方面，即实验设计问题，以及与Logistic回归技术有着异曲同工之妙的Log线性模型和比例危险模型(proportional hazards-model)，限于篇幅，本讲座不予论述，有兴趣的读者可参阅文献〔2〕—〔6〕。应指出的是，虽然在现代文献中，许多作者致力于流行病学研究中队列方法与病例—对照技术间的比较，但它们实际上具备共同的概念基础，因而适用于其一的统计方法，只需稍加改变甚至原封不动，就可在另一场合采用。无疑，就调查人群大小，研究持续时间，以至更重要的，由于病例的选择和通过访问或其它回顾性调查方式，来确定暴露史所引起的偏倚问题等方面来说，病例—对照方法的确不同于队列研究。然而，用以描述发病率及其与各种环境或遗传危险因素间关联的统计模型，对于两种研究却是相同的，这一共同点大大拓广了分析方法的应用范围。

简单说来，病例—对照研究乃是选择一批某种疾病患者作为指示病例，并寻找未患此病的可比个体作为对照，调查他们对该疾病各种可能危险因素的暴露史，以便评价一个或数个危险因素是否为此种疾病之病因的假设。对一项病例—对照研究资料进行分析，目的就在于鉴别所查因素中哪些与疾病危险具有显著的关联。分析时需予考虑的基本问题有：疾病危险度与各因素间关联的强度，所见关联的可信范围，即偏倚、混杂和某些偶然事件对结果的影响，以及在多大程度上能将此种连带描述为因果关系。

（一）选择偏倚和观察偏倚

队列和病例—对照研究都属于比较研究[7]，这就涉及到一个根本性问题，即研究资料的可比性。队列研究的资料是依据疾病状况确定前的暴露得来的，对于在给定时期内出现的所有病例都应进行调查，因此病例与对照的有关暴露资料是可比的，能够借以导出不同亚人群中发病率的无偏估计。而在病例—对照研究中，有关暴露的资料是于疾病状况确立之后获得的，病例与对照代表着抽自于总人群的样本。倘若选择病例和对照的程序不同，或者不能按可比方法从这两组人中取得暴露资料，例如由于对同一询问表中的问题，病例和对照有不同质量的回答，发病率的估计都将是有偏的，前者造成选择偏倚，后者导致观察偏倚（又称为回忆偏倚）。可见，偏倚是设计方案本身带来的后果，也就该尽量在设计阶段加以消除，在数据分析时常常很难控制偏倚效应。病例—对照研究的缺点最终在于，研究者必须求助于调查对象的密切合作，或者仅能半定量地论证，从病例和对照那儿得来的资料，在来源和质量上具有等效性。于是，成功地进行一次病例—对照研究的主要问题，在很大程度上取决于是否能避免偏倚。选择偏倚尤其会降低病例与对照间的可比性，是病例—对照研究中最为严重的潜在疑难。

所谓选择偏倚，系指依据病例（或对照）显示出来与待查危险因素暴露史有关的某些特性，而将它们列入一项研究之中（或排除在此研究之外）所导致的偏差。对于病例来说，这些特性往往就是癌症患者的体征或症状，然而却未必总是由此癌症本身引起的。病例组应选择患有某种恶性肿瘤且病因学上尽可能同源的一组个体。每次集中探讨一种可能的病因较之同时考察多种病因更为容易。例如在子宫癌的流行病学研究中，仅能获得十分有限的知识；但若将子宫体腺癌与子宫颈鳞状上皮细胞癌加以区分，并分别进行研究，就能大有改观，收到立竿见影的良好效果。对病例特征作出若干必要的限制，除了能获得病因学上同源的病例组外，还可以带来其它好处，其中之一是，所研究的可能致病因素范围较窄，就更有希望把一些似是而非的"病因"排除在考虑之外，此类虚假"病因"与癌症并无本质关联，或者仅有由于混杂所引起的连带；另一特别重要的得益就是可以减小选择偏倚。通常使用的病例有三种类型：发病、患病和死亡病例。死亡病例只能通过亲属来调查，往往带来很大的记忆偏倚，因此只是偶然采用。发病病例，即新近确诊的病例，往往被优先选用，他们有患病（以往已确诊）病例所不具备的一些优点。例如，发病时间更接近于病因暴露时间，所以发病病例比之患病病例，能更好回忆自己的有关经历和暴露史。并且，与以往的记录案宗相比，新近的病例资料、职业暴露或其它记录更易于得到，所含信息也更为丰富。此外，不同于患病病例，一组发病病例尚未受到各种决定生存之因素的影响，即是说，任何时刻的患病病例，是在以往时期确诊的大批发病病例中的残存者。倘若影响两者的危险因素各异，采用患病病例的研究将导致错误结论。采用发病病例的第三个优点是，疾病本身的效应表现要弱于病因。如果某个病例是存活多年的患病者，在此期间他可能曾以数种方式变更了其环境条件和生活习惯。因此，除非小心翼翼地详细询问发病前的情况，否则将出现虚假的特征。发病病例的最后一个优点在于，他们与病因学研究的通常目标关系更为直接；即是说，通过发病病例，我们能估计癌症的发病率（而不是患病率）与暴露间的关联方式。

对照组类型、数目和容量的确定以及具体对照对象的选择问题，也许是设计一项病例—对照研究时最为重要和困难的任务，文献〔8〕中有详细论述，这里只指出下列几点。首先，影响对照组选择的主要因素是指示病例的各种特性和来源、对有关癌症危险因素的知识，以及这些因素与所研究的暴露间可能怎样混杂或交互作用的认识，总之，要确保研究资料的可比性。一般说来，如果已从基本人群中收集到了病例组，该人群的一随机或（年龄和性别）分层随机样本，将证明是一适宜的对照组。如果病例取自于某个受到限制的来源，则从同一来源中选择对照往往是合理的。其次，应该指出，某些研究者认为，除了所关心的疾病外，对照与病例必须在各方面都相似，这种错误概念的历史渊源是清楚的，它来自于实验研究所遵循的准则，要求对照组在各方面都必须象实验组那样加以处理。但在病例—对照研究中，此公理并不适用，现在已经认识到，除了非危险因素外，要求对照在与待查暴露有关的某些方面也和病例相似，即所谓"过份配对（Over-matching）"，会导致不良后果，这个问题将在叙述配对时讨论。

（二）混杂及其控制方法

在比较研究中，除了所考查的危险因素外，对于病例和对照组来说，还有众多的背景因素可能不同，其中有些背景因素会混淆或掩盖危险因素的真

© 1994-2014 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

实效应，导致后者的估计失真。同时满足下列条件的背景因素称为混杂变量（Confounding variable）：①在病例和对照组中，它们的取值明显不同；②这些背景因素本身影响比较结果；即是说，同时与疾病和危险因素暴露状况有关联的背景因素。如果在一项比较研究的设计或分析阶段忽视了混杂的控制，将影响结果的真实性，因为它们对结果的混杂效应可能被误认为是危险因素所致，严重时会导致似是而非的结论，给思维布下陷井，把人引入岐途。我们来看看下述两种简单情况。其一是，若有一混杂因素包含两个水平，疾病和暴露的分布如下：

混杂因素 { 水平 1　疾病有高的危险度，暴露重度流行；
水平 2　疾病有低的危险度，暴露轻度流行，作为例子，假定疾病为肺癌，暴露主要指兰领工人的某种职业，混杂因素是吸烟。至少在美国，兰领工人要比管理人员和专业工作者更常吸烟。在此情况下考查职业暴露与肺癌的关系时，忽视吸烟这一混杂因素，将会使暴露与疾病间的关联比考虑混杂效应时的更强，高的疾病危险度对应于重度流行的暴露，暴露轻度流行将与疾病有低的危险度一起出现。可能遇到的第二种情况是：

混杂因素 { 水平 1　疾病有高的危险度，暴露轻度流行；
水平 2　疾病有低的危险度，暴露重度流行。

我们可把为减缓绝经期症状而使用雌激素与乳癌关系的研究作为例子，此时绝经年龄为混杂因素。绝经期开始较早能降低乳癌危险度，却导致更多地使用补充雌激素。这儿，忽视混杂变量的影响，将会使暴露与疾病间的关联显得比实际情况偏低。

通常要注意区分两类混杂效应，一类引起暴露与疾病间的虚假关联，称为正混杂；另一类则是掩盖暴露与疾病间的真实关联，叫做负混杂。此外应该指出，仅从统计学角度出发是不能透彻讨论混杂现象的。当一个因素与疾病的关联看来至少能由第二个因素（它同时与疾病和第一个因素有关联）得到部分解释，但此两因素实质上测度了同一事件，或者第二个因素只是第一个因素的某种后果时，也可能出现混杂。但在此情况下，认为第二个因素混杂了第一个因素与疾病间关联的看法却并不适宜，这涉及到过份配对问题，下面再来讨论。

如上所述，混杂反映了研究人群中诸变量间复杂的因果关联，在队列和病例—对照研究中有类似表现。如果能对混杂因素进行观测和鉴别，它们所引起的效应就可显著地减小，甚至消除。控制混杂的基本方法是分层与配对，这既可在研究设计阶段考虑，也能在数据分析阶段进行。为了使得病例和对照资料能在一定程度上达到匹配，当我们发现某个变量看来会影响疾病与暴露间的关联时，就应该把它作为一混杂因素来处理。下面各讲中将系统介绍分层资料和配对数据的统计分析方法。

在癌症研究中，一般都实现年龄和性别的配对，因为这两个因素都可能混杂其它因素的效应。许多研究还要求对一些附加变量实现配对，并且常常在各个层次中，每一病例可以有多个相应的对照（1:M配对），以便改善由分层分析所得有关相对危险度估计的精度。某些配对因素，例如居住场所或血缘关系，代表着由若干因素组成的复合变量。我们很难甚至无法确定此类复合变量的各种具体成分及其性质，因而不能一一实现每个组成因素的配对。由于复合因素的配对等效于用其各组成因素来进行分层，所以这时的配对目的，正是为了消除组成该复合因素的种种只能含糊推测之变量的混杂效应。在这种情况下，配对可能成为消除危险度估计中偏倚的一种重要手段。但是须注意，配对要求并非越严格越好，倘若是通过在不同变量上都极好地实现匹配而完成配对，则对于所研究的暴露，病例和对照实际上也将无甚区别，结果是适得其反，耗费了时间与钱财，却无所收获，甚至会得出有害的错误结论。因此必须避免过份配对，即防止关于下述变量的配对，它们与所研究的暴露相关，但不是独立危险因素，因而不可能是混杂因素。

区分两类过份配对是有益的。一类出现于这样的情况，即调查者所配对的因素本身属于待查之因素致癌机理的组成部分，作为例子，考虑在研究外原性雌激素与子宫内膜癌的关系时，就病例和对照的子宫内膜增生与否进行配对之后果。增生是由外原性雌激素引起的一种状态，有可能演变为癌症。这样的配对将使对照与病例在暴露史方面极为相似，因此既便采用适当的分析方法，配对数据也将导致相对危险度估计向1偏倚。当取本身并非独立危险因素，也不是因果机理的组成部分，但和所查危险因素相关之变量进行配对时，则会出现第二类过份配对。在此情况下，采用适当的分析方法，能够给出推理上成立的相对危险度估计，但研究是低效的，即不精确，所得估计难以置信。

（三）确定因果关系的标准

© 1994-2014 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

首先要指出，在癌症流行病学研究中，将必要性与充分性概念引入因果关系的讨论是无益的。除了某些偶然见到的极端职业性暴露外，至今尚未识别出这样一些相互影响的因素，它们的存在与否必然地决定着癌症的是否发生。因此我们将在概率意义下采用"病因"这个词。当说及某因素为某种疾病的病因之一时，只是意味着，由于此因素的存在而导致该疾病的危险有所增加。从这一观点看来，一种疾病可以有众多的病因，其中的一些因素能够起协同作用。借助多阶段模型(multistage model)来进行思考，并把病因视为这样的因素是有益的，在癌瘤从一个阶段演变到其后阶段的发展过程中，它能直接增加此疾病的一种或数种率。一种因素可能需有另一因素存在时方能起作用（即所谓的多因素综合作用），在此情况下，应该断然地将两者的联合出现归为一种病因。由于癌症的病因通常是多因素的，就使研究工作中的因果推断变得更为复杂，也要求有关癌症研究的流行病学和生物统计学同样具有完善的先进方法，以便揭示癌症发生、发展的规律。从这个角度看，多因素统计概念具有重要意义，多元分析方法是得力的分析手段。

其次要注意，分析得出的显著关联只是一种统计联系，并不一定就说明因果关系，但它能为我们深入探讨病因提供重要线索。一般说来，确定因果关系有以下几条标准：

①剂量—反应关系

人们总是期待，一种真实联系的强度应随暴露水平和暴露持续时间的增大而上升。剂量—反应关系的确证是因果关联的重要指示，而缺乏这种关系则成为后者的反证。

②危险度对于疾病各亚类的特异性

如下面例子所表明的，确认联系仅限于疾病的某些亚类，可以成为因果关联的有力证据。在较早时期，吸烟引起肺癌的作用尚未完全确立，但调查中包括了有说服力的发现，那就是不吸烟的肺癌患者往往属于比较少见的腺癌。又如，苯暴露与白血病间的联系仅限于一些特殊细胞类型，即急性非淋巴细胞性白血病。联系的特异性也许是能将它视为因果关联的主要理由。

③危险度对于暴露各亚类的特异性

倘若能够证明，对于某些类型的暴露或若干类别的个体，疾病与暴露间的关联较强，也能提高我们对此关联因果性解释的置信程度。在暴露愈严重的人中危险度也愈高的剂量—反应关系就是一个明显的例子。各种交互作用也有助于洞察疾病的发病机制。例如暴露于电离辐射后乳癌危险度的例子：分别在20岁时和30岁以后受到辐射的妇女中，所观察到的乳癌危险度明显不同，前者有更高的危险度。以后的研究表明，在月经初期前两年内或初次妊娠期间，受到照射的妇女中，乳癌危险度实际上有很大增加。在妇女一生中的这两个时期，乳腺组织迅速增生。

④关联的强度

疾病与暴露间的关联愈强，人们便更有可能将此联系视为因果关联。在病例—对照研究中，病例和对照组的选择能够导致偏倚。但关联愈强，用偏倚来解释此种关联的根据就变得愈加脆弱。有关母亲在妊娠期使用烯雌酚，将导致女儿发生阴道腺癌可能性的研究可作为例子，研究中包括8名病例，每一病例有4个配对对照；8名病例中有7例的子宫曾暴露于已烯雌酚，与此相反，32例对照无一有过此暴露史。此联系的强度使人相信它的因果性质，虽然对于20余年前，某种处理的回忆是一种潜在的严重偏倚来源。

⑤危险度与暴露的时间关系

对于大多数上皮瘤，可以期待至少有15年的潜伏期。在持续暴露的典型情况下，于初次暴露后的10—15年内危险性并不大，但此后相对危险度便开始上升，经过30年或更长时间后达到平稳状态。对于辐射引起的白血病来说，危险度上升速度较快；而在器官移植的受者中，某些淋巴瘤的危险度，在一年内就有很大增加。虽然暴露后危险度随时间的演变情况对于队列研究评价因果关系的作用更大，但在病例—对照研究中，更多地注意到调查资料的这一性质，也将证明是有益的。

⑥缺乏其它可能的解释

在所要分析的数据中，必须证明感兴趣的暴露与疾病间的关系，并非某个本身与疾病和此暴露都有关联之因素的混杂效应所致。虚假联系也能因病例或对照的有偏选择，以及由病例或对照组取得资料的方式不当引起。与混杂相比，偏倚问题通常更难通过在数据内部作调整的方法来解决。然而，如果选择对照或获得资料的方式与疾病状况无关，则能用偏倚来解释观察结果的程度常常会有减小。

⑦一项研究结果的旁证

实践中很少见到单次研究就能给出因果关系的充分证据，通常还需用不同方法在各类人群中进行其它研究。暴露停止后危险度会相应下降的事实是

© 1994-2014 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

癌组织中IgG的免疫荧光观察

解放军总医院病理科 纪小龙 李维华 张笑明

在肿瘤免疫学中，逐步弄清肿瘤个体的免疫状态以及免疫与肿瘤间相互作用的各种表现，并揭示其内在实质，无疑是肿瘤免疫要解决的重要课题。自1969年Hellström[1]首先报导患瘤动物血清存在着影响细胞免疫作用的因素以来，对肿瘤和宿主间的复杂的相互作用进行了广泛研究。1971年Hellström[2]进一步用体外方法证实人体血清中具有影响抗肿瘤淋巴细胞细胞毒作用的封闭因子，随后有关肿瘤细胞表面具有抗体的报道逐渐增多[3~7]，并且证明这种结合于肿瘤细胞的免疫球蛋白的生物学活性与肿瘤抗体和／或肿瘤抗原抗体复合物相似，在体内具有促进肿瘤生长的作用[8]，并能在体外对淋巴细胞的细胞毒作用产生封闭[9]。Ioachim等[10]则从肺癌组织中及胸腔渗出液中提出IgG、IgM、IgA。我们则应用较简便的免疫荧光抗体技术直接对肿瘤组织进行免疫球蛋白的检测，以证明癌组织中免疫球蛋白的存在及表现形式，为肿瘤免疫提供可能有益的线索。

材料和方法

收集手术切除的消化道癌标本14例（食道鳞癌5例，胃贲门腺癌6例、胃窦腺癌3例）选取无坏死的癌组织及邻近的正常组织作冰冻切片，厚4微米、吹干、不固定，作为抗原片保存于冰箱备用（2天内应用）。将冰冻切片剩余组织再作常规石蜡切片，H、E染色与荧光所见作对照观察。

每例于手术前取患者静脉血2毫升，分离血清后放冰箱备用（一周内应用）。

每例冰冻切片均做直接免疫荧光及间接免疫荧光检查。

组织切片直接免疫荧光法：抗原片上滴加经稀释至染色效价为1:10的FITC标记的羊抗人IgG荧光抗体（上海生物制品研究所、批号80—1—1）。置温箱37℃染色45分钟，经pH8.0的PBS液换洗

更具说服力的证据，虽然危险度未曾减小并非无因果关系的指示，在实验室中开展对某些因素致癌力的研究，从生物学角度确证其真实性或似是而非的程度，能提供有力的旁证。

欲使一种关联的因果性质得到公认，往往需要满足比上述标准更为广泛的条件，包括一系列确证研究和排除生物学上似是而非的假象。尽管如此，即使单次研究难能提供因果关联的确实证据，分析时仍应力求取得此项研究所能提供的支持或否定因果性质的全部证据。

参考文献

〔1〕Mantel, N. & Haenszel, W, 1959), Statistical aspects of the analysis of data from retrospective studies of disease. J. natl Cancer Inst., 22, 719—748

〔2〕Schlesselman, J. J.(1982), Case-Control Studies: Design, Conduct, Analysis. New York: Oxford University Press, 1982

〔3〕Breslow, N. E., Smith, P. G., Day, N. E., et al. (1982), Short Course on Statistical Methods in Cancer Epodemiology, IARC, Lyon, 1982

〔4〕刘韵源（1982），队列和病例—对照研究对样本容量的要求（Ⅰ）和（Ⅱ）。（待发表）

〔5〕艾沃日特著，刘韵源，周家丽译（1980），列联表分析，科学出版社，1980

〔6〕Fienberg, S. E.(1977), The Analysis of Cross-Classified Categorical Data, The MIT Press, 1977.

〔7〕Anderson, S., Auguier, A. et al (1980), Statistical Methods for Conparative Studies, Techniguos for Bias Reduotion, John Wiley & Sons, 1980

〔8〕Breslow, N. E. & Day, N. E. (1980), Statistical Methods in Cancer Research, volumel-Jhe Analysis of Case-Control Studies. IARC, Lyon, 1980.

〔9〕刘韵源（1982），癌症研究中临床试验数据的分析（Ⅰ）、（Ⅱ）和（Ⅲ）。（待发表）

（1982年10月23号收稿）

© 1994-2014 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net