

现代流行病学

第四讲 流行病学研究中的混杂作用

曾光

一、混杂作用的概念：

1. 混杂的定义：在对某病的病因学研究中，当对所关心的某种暴露因素（E）与这种疾病（D）之间的联系定量估计时，由于其它因子对疾病的混合作用存在，致使E和D之间的联系的真实性被歪曲，其联系程度被放大或缩小，这种混入的作用被称作混杂作用（Confounding effect），如此起到了混杂作用的因子为混杂因子（Confounder，简称F）。

混杂因子成立的条件为：①混杂因子是所研究疾病的危险因子；②混杂因子与暴露因子间存在统计学联系；③混杂因子不是暴露因子—疾病因果关系链其中任何一个环节上的因子，混杂与暴露相互之间也不存在因果关系。

2. 暴露（E）、疾病（D）、混杂（F）三者之间的关系：

① E、D之间：在E、D的因果关系研究中，人们主要关心的是探讨清楚E对D的发生是否有作用以及作用的程度如何。从逻辑的角度以下三种结果都可能是研究的结论：即E是D的病因；E、D之间仅存在着一般的统计学联系；E、D之间不存在起码的统计学联系，或称二者之间无关系。无论E、D间的真实关系属于那种，表现出的联系，都有可能受到外部因子F的影响。

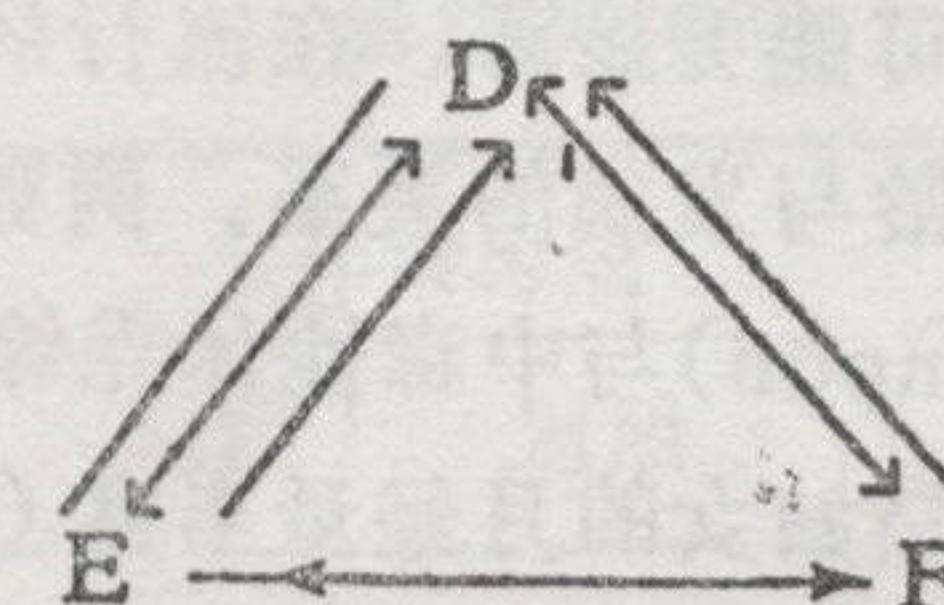
② F、D之间：F与D至少存在一般的统计学联系，也就是说，混杂因子本身应该是所研究的危险因子。这种联系的本质有两种可能：F是D的病因；F、D之间仅存在一般的统计学联系。

当F、D之间存在统计学联系时，还有其它两种逻辑可能：D是F的病因；和F是E、D的因果关系链中的一个变量。一旦发现这两种情况之一，都不能认为F是混杂因子，因为这时F已不符合混杂因子成立的条件。发生上述情况的原因在于，研究者事先对D产生的原因缺乏足够的认识。

③ F、E之间：首先，F、E间需要存在某种联系，

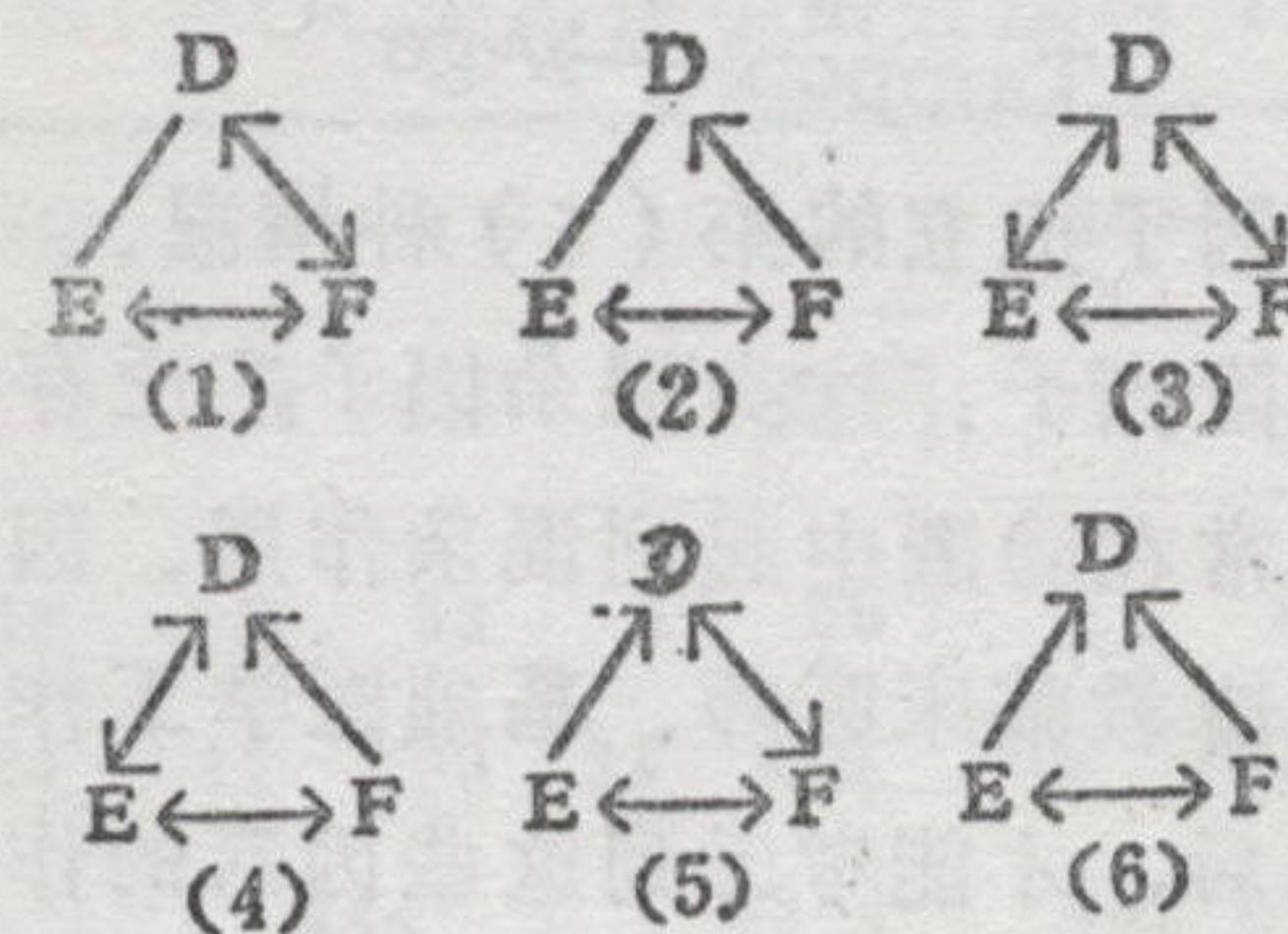
否则F起不到混杂因子的作用；其次，F不能是E的病因，E也不能是F的病因，无论是直接病因还是间接病因都不允许，否则，由于F、D都同一条因果关系链相联系，因而否定了F作为混杂因子成立的条件。这样只剩下最后一种可能——F、D之间只能存在也必须存在一般的统计学联系。

3. 混杂因子成立的各种形式：D、E、F三者之间可能成立的关系，可用下图来表示：



图中，“——”表示不存在联系；“ \longleftrightarrow ”表示存在非因果关系的统计学联系；“ \rightarrow ”表示存在因果关系，箭头所示，从因到果。

如将上图展开，可得到混杂因子成立的全部6种形式：



其中(1)、(3)、(5)三种形式意味着F起到混杂作用的原因不明，需要做出合理的解释。

4. 混杂因子的特征推论：从混杂因子成立的条件出发，可得出以下三条特征推论。

本文作者单位：100021 北京市，中国预防医学科学院流行病学微生物学研究所流行病研究室

①无论F、D之间的关系为因果关系抑或为一般统计学关系，只要F在不同暴露强度层的分布不一致，都将同时影响暴露组(E^+)和非暴露组(E^-)中疾病发生的真实性或真实程度。

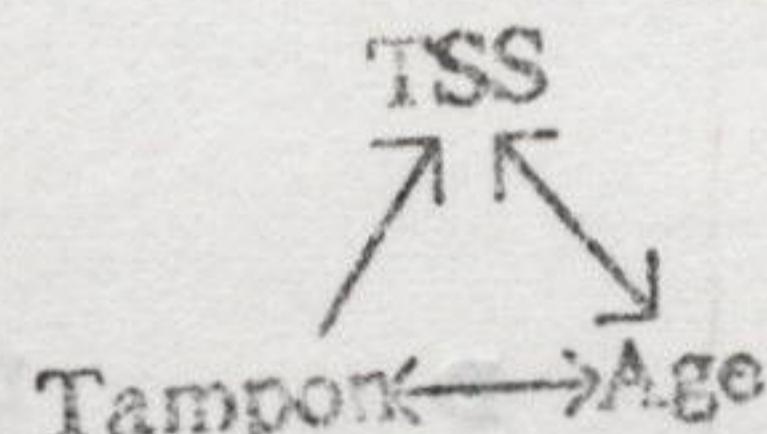
②将F作为变量适当分层后，再对E分层，可以将E、D之间的真实关系表现出来。

③混杂因子有特异性，在某项疾病病因学研究中起混杂作用的因子，在另一项研究中可以不起任何作用。

5. 常见的混杂因子：常见的混杂因子有人口统计学因子和与发病有关的E以外的其他危险因子。

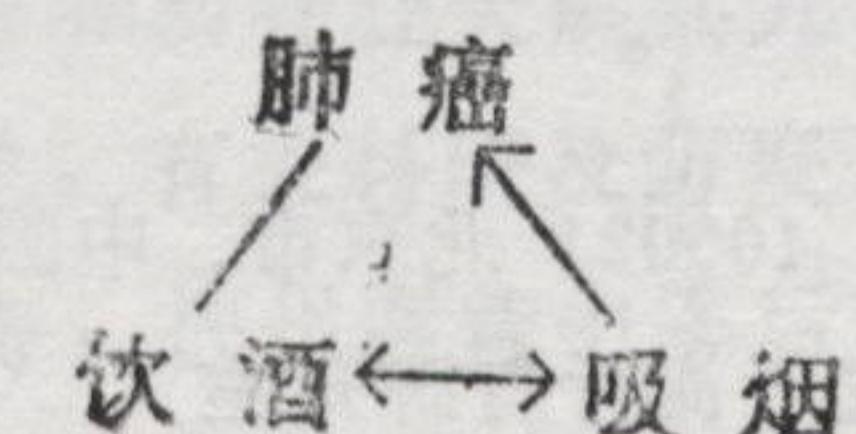
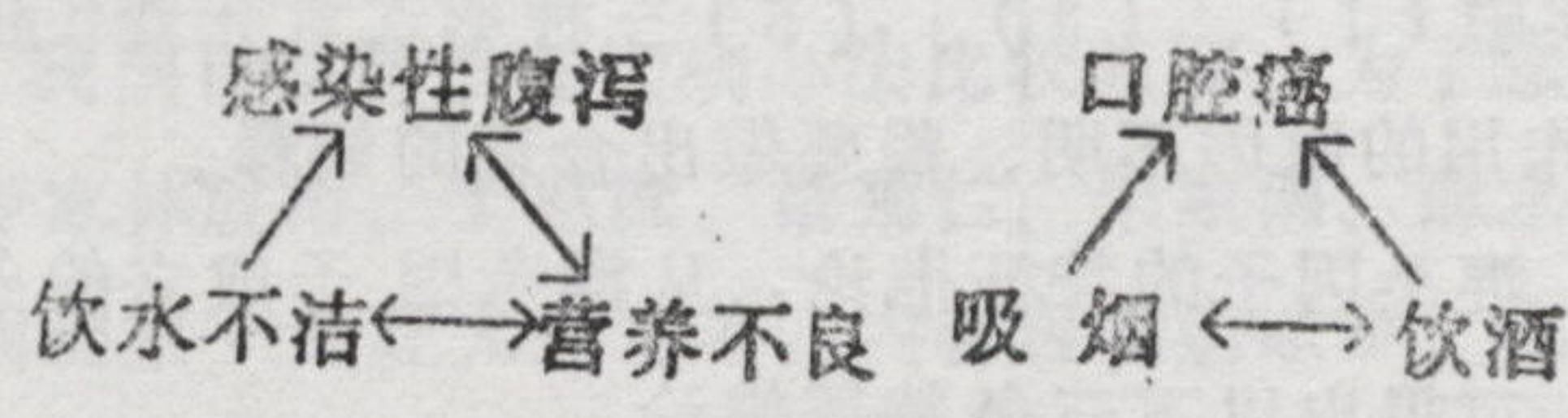
①人口统计学因子：年龄、性别、种族、职业、经济收入、文化水平等人口统计学指标，是经常遇到的混杂因子。

以年龄为例，不同年龄对疾病的易患程度不同。随着年龄的增长，人体会发生一系列生理、生化、解剖和免疫学的变化，人的社会地位会发生变化，在心理和行为上的变化可能更显著。在流行病学调查中，往往要首先考虑年龄与发病的关系。例如，在研究妇女使用月经棉(Tampon)与中毒性休克综合征(TSS)的关系时，因为年轻妇女的月经量丰富(生理变化)，她们比中年妇女更喜欢使用高吸收度的月经棉(行为选择)，而高吸收度的月经棉比一般的月经棉更容易导致发生中毒性休克综合征，因此年龄(Age)与暴露于月经棉和发生中毒性休克综合征都存在统计学联系，具备了混杂因子成立的条件。三者间的关系为：



符合混杂因子成立的第(5)种情况。

②其它危险因子：这类混杂因子广泛存在，表现形式多样，常常在隐匿中起到混杂作用。因此流行病学家往往需要请教临床医师、基础医学工作者和其他公共卫生专家后，才能考虑到这些因素的作用。下面举三个例子。



上述三种情况：营养不良、饮酒和吸烟各自都可以起到混杂因子的作用，分别属于混杂因子成立的第(5)、第(6)和第(2)种形式。

6. 正混杂与负混杂：导致人们过高地估计暴露与疾病联系的混杂因子称为正混杂(Positive confounder)。特点为粗OR值或RR值出现比实际的OR值或RR值远离1的偏差。例如，暴露于饮酒与肺癌的发生无关，但是饮酒者中吸烟者的比例很高，而且吸烟是肺癌的真正病因，由此造成了饮酒与肺癌的发生高度相关的假象，吸烟就是本研究中的正混杂因子。导致过低估计暴露与疾病间关系的因子为负混杂因子(Negative confounder)，特点为粗OR值或RR值出现比实际的OR值或RR值向1靠近的偏差。例如，在研究气功锻炼与冠心病的关系时，性别就是潜在的混杂因素。因为不锻炼气功的人群，男性比女性冠心病发生率高，在男、女合计的分析时，将会低估练习气功对冠心病的保护作用。

二、混杂的识别与度量：

1. 混杂的识别：

①专业知识和经验的初判：在一项流行病学因果关系调查开始之前，首先要依靠专业知识，从研究的疾病和暴露出发，列出所有能想到的混杂因子，按性质或相互联系归类。例如经济状况和文化水平之间的相关性很强，可归为同类。然后凭借经验对每一类中混杂因子潜在作用的大小作出初步判断。

②分层分析：为分析阶段首选应用的方法，分析的步骤为首先估计暴露与疾病间的粗的(未分层)联系；然后，将某一特定的混杂因素分层后，分别估计各层中暴露与疾病的联系，并与粗联系进行比较，以确定混杂作用的大小。

③多元分析方法：在分层分析的基础上，将选出的各类中重要混杂因子与暴露因子一起作为自变量，与疾病这一因变量做多元分析，分别估计各种混杂因子对疾病发生的作用大小。

④依靠专业知识最终判定：在上述初判和分析的基础上，通过专业知识做出分析和解释，最终确定有意义的混杂因子，并描述其产生混杂作用的原因和作用的大小。

2. 混杂的度量：有人认为，一个变量的混杂作用，只能截然分成“存在”和“不存在”两种情况，即具有“全或无”的性质(all or nothing property)，并完全依靠统计学显著性检验来判断某个因素

与暴露和疾病的联系，从而判断是否是混杂因素。这种看法和作法是不正确的和片面的。因为几乎所有能想到的潜在混杂因子，只要在不同暴露强度层中的分布不完全一致，或多或少都能度量到混杂作用，只是大小不同而已。而样本的大小对显著性检验的结果有决定性作用，再小的作用在样本增大到一定程度时会显示出显著性差异，因此，用显著性检验来度量混杂作用是不可取的。正确的方法，应该通过计算“混杂作用度指数”(Index of the degree of confounding)来度量。

$$RF_i = \frac{OR_i \text{ (或 } RR_i)}{OR_c \text{ (或 } RR_c)}$$

式中：RF_i为按混杂因素分层第i层的混杂作用度指数；OR_i(RR_i)为按混杂因素分层第i层的比值比(相对危险性)；OR_c(RR_c)为不分层时，粗的比值比(相对危险性)。因此RF_i实质是OR_c相对于OR_i的比例。

1979年Shapiro等人调查了234例心肌梗塞(MI)女病人，在入院前一个月使用口服避孕药(OC)的情况，并同时调查了1742名对照，结果如表1。

表1 近期使用口服避孕药与心肌梗塞的关系

	MI组	对照组	比值比
OC使用			
是	29	135	OR=1.68
否	205	1607	
合计	234	1742	

如果按年龄分层，分别计算各年龄组的比值比，则发现除35~39岁年龄的OR值略低于粗OR值1.63外，其余4组的OR值是粗OR值的2~5倍。最初算得的粗OR值明显估计过低，给人以口服避孕药对发生心肌梗塞影响不大的假象。

粗OR发生虚假性偏低的原因可从表3分析，第一，MI组与对照组的年龄构成不同，对照组的5个年龄组都占20%左右，而MI组占70%的为40~49岁的高龄妇女。第二，无论MI组还是对照组都是低年龄组妇女使用OC的比例高，随着年龄增大这一比例明显递减。在这里年龄起到了负混杂的作用。

正确的方法，可应用Mantel-Haenszel法对年龄这一混杂因子做分层分析，重新计算比值比(OR)。为了便于计算将表2数据改写为表4的形式。

表2 按年龄分组近期口服避孕药与心肌梗塞的关系

年龄组(岁)	使用OC	MI组	对照组	OR	RF
25~29	是	4	62	7.2	4.3
	否	2	224		
30~34	是	9	33	8.9	5.3
	否	12	390		
35~39	是	4	26	1.5	0.9
	否	33	330		
40~44	是	6	9	3.7	2.2
	否	65	362		
45~49	是	6	5	3.9	2.3
	否	93	301		
合计		234	1742		

表3 病例与对照组和年龄构成及各年龄组的暴露率(%)

年龄组(岁)	年龄构成		OC暴露率	
	MI	对照	MI	对照
25~29	2.6	16.4	67	22
30~34	9.0	24.3	43	8
35~39	15.8	20.4	11	7
40~44	30.3	21.3	8	2
45~49	42.3	17.6	6	2
合计	100.0	100.0		

表4 MI与OC关系按年龄分层M-H法计算表

OC 使用	25~29岁		30~34岁		35~39岁		40~44岁		45~49岁	
	MI	对照								
是	4	62	9	33	4	26	6	9	6	5
否	2	224	12	390	33	330	65	362	93	301
Ni	292		444		393		442		405	
Wi	0.425		0.892		2.183		1.324		1.148	
Vi	0.771		0.227		0.332		0.296		0.381	
OR	7.2		8.9		1.5		3.7		3.9	

Ni、Wi、Vi的计算公式如下：

$$Ni = ai + bi + ci + di; Wi = bici / ni$$

$$Vi = (ai + ci) / aici + (bi + di) / bidi$$

例如，对25~29岁组：

$$N = 4 + 62 + 2 + 224 = 292; W = 62 \times 2 / 292 = 0.452$$

$$V = (4+224)/4 \times 224 + (62+2)/62 \times 2 \\ = 0.771$$

OR_{MH}计算公式为：

$$OR_{MH} = [\sum (a_{idi}/n_i) / (\sum (b_{ici}/n_i))] \\ = 23.71/5.97 \\ = 3.97$$

计算 OR_{MH} 自然对数 (lnOR_{MH}) 方差的估计值：

$$Var(lnOR_{MH}) = \sum (W_i V_i)^2 / (\sum W_i)^2 \\ = 2.876 / (5.972)^2 \\ = 0.0806$$

lnOR_{MH} 95% 可信限为：

$$ln 3.97 \pm 1.96 \sqrt{0.0806}$$

为 0.822~1.935

取自然对数的反对数，OR_{MH} 为 95% 可信限为 2.28~6.93

相应的 RF_{MH} 值为：3.97/1.68=2.36 其上下限为：1.36~1.13

即如将年龄这一混杂因子加以控制，近期口服避孕药发生心肌梗塞的作用，是不考虑年龄因素的 2.36 倍，其 95% 可信限范围为 1.36~4.13 倍。

三、混杂作用的控制：原则上讲，所有能考虑到的混杂因子都能得到控制。控制混杂因子的方法，包括在研究设计阶段采取预防措施，以及在分析阶段通过各种技术处理予以控制的方法。

1. 在设计阶段控制：应该优先考虑在研究设计阶段预防混杂因子发生的控制措施。然而，能否作到这一点，不但取决于研究者的专业知识和经验，也取决于研究条件，病例和对照选取的难易。

①限制：限制指对研究对象选择的条件加以限制。例如，在研究口服避孕药与心肌梗塞的关系时，考虑到年龄是可能的混杂，即只对 35~44 岁年龄组的妇女做调查。在研究气功锻炼对高血压发生保护作用时，只先取 40~50 岁男性做调查，以排除性别和年龄的混杂作用。在这里性别为离数的二值变量，从中选取一种性别，年龄为连续变量，从中截取一个狭窄的范围，年龄限制的越窄，残余混杂作用越小，但可得到的样本也越小，需在两者之间平衡。

采用限制的方法控制混杂因子，一方面只能针对个别最重要的混杂因子，先决条件为限制后仍然能保证适当的样本数。另一方面，限制的结果使得研究者不可能对混杂作用的大小和交互影响加以分析和度

量，只能确保对暴露与疾病之间关系的研究。

限制的缺点还在于，在限制混杂因素的同时，对暴露和疾病发生的范围也做了限制，难于观察到暴露对疾病影响的全貌，在做结论时应该慎重。

②随机分配：随机分配只适用于实验流行病学的研究，即将受试者随机分配于暴露强度不同的各组，观察疾病产生的不同结果。随机分配的目的在于，使各组人群具备混杂因素的情况均匀分布，使得暴露强度的差别成为各组间的唯一差别。随机分配又可分为简单随机分配和分层随机分配 (Stratified randomization) 两种方式。

③匹配 (Matching)：匹配是最经常用于控制混杂因子的方法。匹配可分为群体频数匹配和个体匹配两种，两种方式在病例对照调查和定群调查中都经常应用。前者指混杂因子发生的频度在不同组应大致平衡，后者指按一个至数个混杂因子分层，为病例或暴露选择同层的对照，为一个病例配的对照数一般为 1~2 个。匹配使用得当可以有效地控制混杂因子的作用，提高研究结果的真实性，并可在减少总样本数的情况下得到结论，即提高了研究的效率。

近年来，人们发现匹配也存在一些缺点。首先人们难以对匹配掉的混杂因子及交互影响做深入分析；其次，在病例对照调查中，用匹配的方法控制混杂是困难的，经常低估暴露对疾病的作用。有时还会引出过度匹配 (Overmatching) 的问题，掩盖了暴露的真实作用。还有，过分苛刻的匹配，会使得部分病例找不到对照，致使信息浪费，使研究的效率反而降低。

2. 在分析阶段控制混杂：

①分层分析：分层分析是按混杂因素分层后，分别就暴露对疾病的联系做分析，如应用 Mantel-Haenszel 法在分析阶段控制混杂因子。

但是由于某些资料常有各层样本数大小之间的悬殊，病例与对照或者暴露与非暴露在各层中的比例失调等问题，为如何分层和分层后分析带来困难。在这种情况下，常使人们不得不对混杂做限制后再做分析，或者直接应用多元分析的方法。

②标准化方法：标准化的方法是对分层分析方法的补充手段。当不同暴露强度组间混杂因子分布不均匀时，可以选择一种标准构成，来调整原来分布的不均匀性。率的标准化分直接标准化和间接标准化两种方式，在卫生统计著作中有详细介绍。对相对危险性 (RR)

也可以标化，得到标化的相对危险性(SRR)。标化的实质在于考虑某因素分层后的权重，由此推论， OR_{MH} 也可以认为是标化后的计算结果。

③多因素分析法：在多因素分析时，暴露因子和每个混杂因子，都被放在同样地位做分析。常用多元分析法，有Logistic回归模型，Cox模型和对数线性模型(Loglinear models)等。多因素分析法的结

果有时并不可靠，因为从理论上讲，多元分析方法要求各自变量符合等同协方差矩阵的多元正态分布，且各变量间相互独立，而一般资料难以满足要求。此外，引入的自变量不同，相互之间的交互作用也不同，其结果可以改变同一自变量对疾病作用的大小，甚至作用的方向也可能发生变化。

衢州市钩端螺旋体菌苗预防效果的研究

陈作良 傅桂明 何粟海 吕跃民 余樟友 俞年生

浙江省钩端螺旋体(钩体)病流行已有数十年历史，是广大农村常见病之一。应用钩体菌苗预防钩体病已收到明显预防效果。但是，由于钩体菌苗接种剂量较大，较易出现局部反应而影响接种率。为此，1991年6月在衢州市所辖六个县(市、区)全部推行一针一毫升一次性免疫，其中选择1990年发病较多的20个乡镇作为重点调查对象，进一步探讨钩体苗免疫前及后一个月抗体水平、菌苗接种与否与发病率关系。现将结果报告如下。

一、材料与方法：

1. 菌苗：钩体双价苗，菌苗型为“浙江甲型”(黄、流各占50%)、“浙乙型”(黄、波各50%)、“浙丁型”(黄、七各50%)，由卫生部上海生物制品研究所生产，抗原浓度为600条/400×。

2. 接种对象：衢州市所辖六个县(市、区)中20个乡镇，总人口262278人，15~50岁健康应接种人群138602人，除1ml接种组人群外，均为对照人群。

3. 免疫方法：皮下注射1ml一次性免疫，禁忌症例外。

4. 抗体测定：随机抽样，接种对象免疫前、后一月(同批)采静脉血2ml分离血清进行显凝(MAT)试验，分别测定“浙丁型”的黄疸出血型和七日热型的抗体滴度，免前<1:20为阴性，免后一月抗体滴度≥1:20和免后一月抗体滴度是免前的4倍或以上者为阳转。

5. 钩体病人诊断：在流行季节临床诊断或疑似钩

体病人，均采静脉血2ml，取双份血清进行显凝试验，抗体滴度4倍或以上增长者为阳性。单份血清抗体滴度在1:400及以上者为阳性。

二、结果与分析：

1. 菌苗接种率：实际接种117990人，接种率为85.13%，比1990年(2ml一次性免疫)提高51.5%。

2. 血清抗体：用MAT测定免疫人群抗体水平，黄疸出血群免前GMT为29.21，免后为276.64，经统计处理($P<0.01$)，免后GMT为免前的9.5倍；七日热型免前GMT为14.35，免后为86.29，经统计处理($P<0.01$)，免后GMT为免前6.0倍；两种抗体阳转率分别为76.47%(黄疸)和75.63%(七日热)，经统计处理 $t=0.152$ ， $P>0.05$ 。

3. 流行季节20个观察乡(镇)经实验室确诊共发生钩体病人84例。其中有菌苗接种者发病11人，发病率为9.32/10万；无接种者发病73例，发病率为50.59/10万，两者经统计处理 $\chi^2=34.53$ ， $P<0.01$ 。以上结果说明浙江省应用钩体菌苗(对号入座)接种1ml一次性免疫是比较理想的。

(参加调查工作还有常山县卫生防疫站，方桂珍、王苟俊、朱水源；开化县防疫站邱家全、徐志林；衢县卫生防疫站王正立、杨根友、徐相卿；江山市卫生防疫站诸葛强、赵渭荣)

本文作者单位：310009 杭州市，浙江省卫生防疫站(陈作良、傅桂明、何粟海)；衢州市卫生防疫站(吕跃民、余樟友、俞年生)