

遗传流行病学研究中的几种设计

郝艳晖 姜庆五

许多常见慢性疾病(如癌症、心血管疾病等)的病因都是极其复杂的,遗传因素和环境因素都可能和这些疾病有关。识别这些复杂的病因是流行病学和遗传学研究者共同的目标,两门学科的相互渗透形成了遗传流行病学。相对来说,遗传流行病学是一个比较新的研究领域,它将遗传学和流行病学的设计和方法结合起来,以探索遗传因素和环境因素对疾病的单独作用以及它们对疾病的联合作用。其主要任务是评价环境因素与疾病状态的关联;评价疾病的家庭聚集性;通过分离分析判断疾病的遗传模式;从而最终通过和遗传标记连锁定位出致病基因。

为便于描述,首先给出一个病因的框架。遗传因素 G (泛指无法测量的遗传因素和可检测的候选基因或遗传标记等因素)和环境因素 E (泛指可测量的外环境因素和可传递的家庭教养环境)以及它们之间的交互作用 GE 共同影响疾病状态 D (或称疾病性状、表型)。 D 可以是连续的,如血压、血脂水平、出生体重等;也可以是离散的,如发病($D=1$)或不发病($D=0$)、某病的不同亚型;还可以是包含截尾的生存数据,如发病年龄等。以二分类性状为例, G 和 E 对 D 的关系可由 logistic 回归模型描述:

$$pr(D=1/E, G) = \frac{\exp(\alpha + \beta E + \gamma G + \delta EG)}{1 + \exp(\alpha + \beta E + \gamma G + \delta EG)} \quad (1)$$

式中 $\alpha, \beta, \gamma, \delta$ 是未知参数 α 表示遗传因素 G 和环境因素 E 为 0 时个体的基准危险。 β, γ, δ 分别是环境因素 E 、遗传因素 G 和两者的交互作用 EG 对疾病状态的对数 OR。它们的大小是我们病因研究中的首要关心的问题。在流行病学中,模型(1)常用于分析环境因素对疾病的影响;在遗传学中,该模型表示致病基因对疾病表达的外显率。

为探察疾病病因,需要进行系列研究。本文主要介绍几种遗传流行病学研究中常用的设计。

1. 以人群基础的病例对照研究设计(population-based case-control study design) 病例对照研究设计是一种以人群为基础的设计,所以它的研究结果可以推广到目标人群。对罕见疾病,由于病例对照设计具有省时、省力、省费用等优点,在慢性病病因研究中倍受流行病学研究者的青睐^[1,2]。但传统的流行病学研究着重环境因素 E 而忽视遗传因素 G 的作用,这样 D 和单独的 E 关联的概率模型就变成在 G 上的边际 logistic 回归,表示为:

$$pr(D=1/E) = \sum_G pr(D=1/E, G) f(G/E) = \sum_G \frac{f(G/E) \exp(\alpha + \beta E + \gamma G + \delta EG)}{1 + \exp(\alpha + \beta E + \gamma G + \delta EG)}$$

除非没有遗传作用和遗传/环境交互作用,否则上式不服从我们通常假设 D 和 E 关联的 logistic 回归,所以忽略遗传因素的作用,必然导致可测量的环境因素的有偏估计。如果遗传因素 G 对 D 有重要影响,并且和 E 不存在交互作用,忽视 G 可能使 β 的估计朝向无效假设。如果交互作用存在,忽视 G 使 β 估计朝向或背离无效假设^[3]。但是在病例对照研究中,对不能测量到的遗传因素是无法估计其对疾病影响的,更无法评价基因和环境的交互作用。为了克服这个缺陷,一个替代的办法是把家族史作为一个危险因素放入模型,一方面可以识别疾病是否存在家庭聚集性,另一方面可以用来调整遗传因素的影响。但应该格外注意的是,家族史这个变量和病例对照研究中其他可测的危险因素有所不同,首先家族史并不是病例和对照自身的一个特征,它依赖于一些外部因素,包括家庭大小、指示病例亲属的生物学关系、这些亲属的年龄分布和疾病的患病率等因素^[4]。结果导致这个变量很容易错分,而这种错分必然影响到参数估计^[5]。例如即使某病没有遗传病因,当疾病患病率 $\pi=0.05$,某个体的一级亲属数 $n=5$ 时,该个体有阳性家族史的概率 $1-(1-\pi)^n$ 为 0.23;如果患病率不变, $n=10$ 时,个体阳性家族史的概率为 0.40。所以如果家庭成员数目大小在病例和对照中的分布不一致时,就会导致 OR 值估计的错误。其次阳性的家族史本身也受遗传和环境因素在家庭内聚集的影响,所以单纯调整家族史是不够的。和后面所述的研究设计相比,病例对照设计在遗传流行病学研究中还有一个重要的局限性就是它的资料一般并不能用于遗传流行病学随后的病因学研究(如通过分离分析确定疾病的遗传模式和遗传标记连锁定位致病基因)。当然随着近年来分子生物学技术的飞速发展,病例对照设计已推广到研究已定位的候选基因对疾病的影响以及基因/环境交互作用和基因/基因交互作用^[6,7]。特别是应用基因组扫描技术,可同时研究大量的遗传标记位点和疾病的关联,从而绘制疾病的关联图谱^[8]。然而,病例对照设计在关联分析中最大的问题是如果研究人群是具有不同遗传特征的人群的混合,不同遗传背景的人群与疾病的相对危险和基因突变频率可能存在差异,导致候选基因或遗传标记和疾病状态的关联将会受到人群结构的混杂而出现假阳性^[9]。

2. 遗传学中传统的家系研究设计(family study design): 遗传学研究设计和传统流行病学设计最大的不同是抽样单位,后者抽样的单位往往是独立的个人,而前者常常以家庭

为单位抽样。患病的个体称为先证者,进一步确认他的家庭成员,并调查他们的疾病状态,收集的资料为家系资料。由于家庭成员之间往往具有相似的遗传物质和共同的教养环境,所以家系成员的疾病状态 D 之间可能具有某种家族聚集性模式。这种家庭聚集性是否由于遗传因素引起,可通过对家系资料进行分离分析来判断。分离分析基于似然函数:

$$L(\underline{D}) = \prod_K \left[\sum_G f(\underline{D}/\underline{G}) f(\underline{G}) \right] \quad (2)$$

\underline{D} 和 \underline{G} 分别表示所有家庭的疾病状态和假定致病基因的所有可能取值, K 表示家庭总数。 $f(\underline{G})$ 是假设服从 Hardy-Weinberg 平衡和 Mendelian 遗传的致病基因的等位基因频率, $f(\underline{D}/\underline{G})$ 是外显率(即给定基因型后表型的概率)和剩余家庭聚集参数(反映消除致病基因影响后的疾病的家庭聚集性)。如果分离分析显示存在主基因遗传,可利用遗传标记连锁分析定位致病基因,连锁分析基于似然函数:

$$L(\underline{D}, \underline{M}) = \prod_K \left[\sum_G f(\underline{D}/\underline{G}) f(\underline{G}, \underline{M}) \right] \quad (3)$$

$f(\underline{G}, \underline{M})$ 是在假定遗传标记和给定致病基因后疾病的表型无关的情况下,由等位基因频率和致病基因与遗传标记的重组率决定。利用家系资料可以分析疾病的家族聚集性,致病基因的遗传模式和定位致病基因。一般用于连锁分析的资料应尽可能选择携带致病基因概率大的家庭,如选择有多个个体发病的家庭,所以连锁分析多为有偏抽样,它的参数估计对抽样是稳健的^[10]。但是基于模型基础的连锁分析要求预先指定传递概率、外显率和等位基因频率等参数,而这些参数需由分离分析估计。所以,任何影响分离分析参数估计的因素都会导致遗传模型参数的错误指定,进一步影响到连锁分析的效率。传统的遗传学研究往往忽视环境因素的存在,但如果确实存在影响疾病的环境因素,或存在基因/环境的交互作用,那么模型(2)和(3)中的外显率函数应该由模型(1)来考虑。所以如果忽视环境因素,在丧失评价环境因素及其基因/环境交互作用的同时,还会影响到分离分析和连锁分析的效能及参数估计的准确性。如果环境因素没有家庭聚集性,即使不造成分离分析和连锁分析参数的有偏估计,也会降低分析的效能,如果环境因素存在家庭聚集性,则不仅降低分析效能,还会带来参数的有偏估计^[3]。除此之外,另一个不容忽视的问题就是此种设计抽样的非随机性^[11],由于这种以家庭为基础的设计收集到的资料只是通过某些病例确认的,因而导致确认偏倚的存在。尽管这种家庭可能包含更多的致病基因的信息,但在研究环境因素和疾病关联以及疾病表型的家庭聚集性时,如不考虑家庭确认问题,其结果能否推广到目标人群值得考虑^[11,12]。

3. 以家系为基础的病例对照研究设计(family-based case-control study design)如上所述,在研究遗传标记与疾病关联的时候,传统的病例对照设计易受人群结构混杂的影响,造成人群分层偏倚。为避免这个问题,许多研究人员提出用和病例有相同遗传背景的亲属作对照,即以家庭为基础的病例

对照研究设计,其中主要包括病例-同胞设计(case-sibling design)和病例-父母设计(case-parent design)等。在病例同胞设计中,用一个或多个未发病的病例同胞做为对照。一般来说,当研究发病年龄不确定的复杂疾病时,对照应选择到病例发病年龄(称指示年龄)时仍未发病的同胞^[13,14]。对在指示年龄没有发病但后来发病的同胞也不排除在外,否则会造成遗传效应估计偏离无效假设^[15]。对所有病例和对照收集在指示年龄时的环境危险因素和生物学标本。如果只有近来的新发病例,那么按照年龄匹配的要求就限制了只能选择比病例年龄大的同胞作为对照。这可能产生有长期趋势或出生顺序效应的环境暴露效应的混杂。这种现象一般不会影响遗传因素的效应评价,但在评价环境的作用和基因/环境交互作用时可以考虑为一种潜在的偏倚来源。有些病例可能没有满足对照条件的同胞,所以只能从分析中排除掉。原则上通过构造似然函数可以对这个问题加以校正,如似然函数中包括还未到指示年龄的同胞在到达指示年龄时仍未受累的概率。但是如果包含这些年龄较小的对照,很难控制相应的协变量没有时间依赖性,这在分析时也将带来很多不便。此外,病例同胞设计的有效性要依赖假设:在父母基因型的条件下,如何选取病例和对照与他们的基因型无关。与病例对照设计相比,在检验候选基因和环境因素的作用时,由于同胞对的过度匹配可能使病例同胞设计的效能偏低^[16]。

在病例-父母研究设计中,收集病例和父母的生物学标本以及有关病例的环境因素的信息。通过比较父母基因型中传递给病例的基因型频率和没有传递的基因型频率或比较父母基因型中传递的等位基因和未传递的等位基因的分布,来判断候选基因对疾病的影响^[17,18]。所以事实上这种设计并没有选择实际存在的对照,只是将父母基因型中没有传递给病例的基因型作为“对照”(或相当于伪同胞对照)。故这种分析方法的有效性依赖假设:人群中父母的等位基因以相同的独立的概率传递给子代。在这种设计中,无需调查父母的疾病状态和关于环境暴露的信息。因为分析时的对照相当于“伪同胞”,他们除了基因型与病例不同,其他环境因素都相同,所以该设计不能分析环境因素和疾病的关联。但是通过比较有暴露病例和无暴露病例的候选基因的传递频率有无差别,可以检验是否存在基因/环境交互作用^[19]。而且在检验基因/环境交互作用方面,由于病例父母设计的弱假设条件和可以通过分层的方法控制家庭内依赖性,这种设计是值得推荐的^[20],尤其在疾病易感基因很罕见时,用病例父母设计比病例对照设计检验基因/环境交互作用会有更高的效率^[19]。似然函数表明在检验基因的主效应时只利用了基因型为杂合子的父母的信息。和病例对照设计相比,病例的父母有更高概率为杂合子,所以在检验显性基因作用时,病例父母设计和病例对照设计效率几乎相等,但检验隐性基因作用时,效能会更高。但某些学者也认为,不能完全利用资料的信息,这也是病例父母设计的一大缺

陷^[8]。而使用这种设计更大的局限是受到疾病发病年龄的限制,如果是晚发性疾病,通常很难收集到父母的信息。

4. 以人群为基础的病例对照家系设计(population-based case-control family design):在遗传流行病学的发展中,很自然的将流行病学中以人群为基础的病例对照设计和遗传学中的家系设计结合起来,形成了病例对照家系设计。在过去十年,这种设计已经得到许多遗传流行病学研究者的重视^[21-23]。类似于病例对照研究,在人群中选择患有研究疾病的病例和没有患此疾病的对照,按习惯称病例(对照)先证者,收集先证者及其家庭成员的有关疾病状态和环境因素的信息。与传统病例对照设计相比,无论从抽样还是分析的角度来看,这种设计都是以家庭作为单位,比较病例家庭和对照家庭成员的疾病状态,这个区别在实际应用中和统计推断的有效性方面都有重要的含义。对于病例对照家系资料,我们首先关心病例亲属是否比对照亲属有更高的疾病风险,也就是家庭风险。病例家系和对照家系的聚集程度有无差别等等。而这些都是可以同时考虑环境因素对疾病的影响,从而得到更准确的估计。考虑到可能与疾病有关的环境因素,利用该设计资料可以建立模型直接估计再发风险(即在一个家庭中给定某个人发病的情况下,其他成员发病的概率),这可用于遗传咨询中;另外,假设对照是一般人群的代表,可以得到再发风险率比(家庭中有一人发病时其他成员的再发风险和一般人群中的疾病风险的比值)的近似估计。这个值在多个受累同胞对的连锁分析中样本量的计算上起着关键作用。在研究环境因素和疾病的关系方面,利用这种以人群为基础的病例对照研究设计可以同时考虑疾病的家庭聚集后更精确的测量环境因素对疾病的影响,包括可以得到更准确的 OR 值。同时在调整了环境因素后如果疾病仍存在家庭聚集,提示需研究假定的致病基因或其他环境危险因素。该家系资料可进一步用于分离分析,以识别该疾病是否存在主基因效应,在控制了主基因效应和已知环境因素后,是否还存在剩余家庭聚集。此外,如果具有家系成员某候选基因的信息,还可评价候选基因和环境因素的交互作用,候选基因和假定致病基因的交互作用。如果疾病存在各种亚型,也可研究环境因素对各种亚型的影响和各亚型是否具有同质性。这些因素的研究往往给基因的寻找和定位提供非常有用的信息。总之,分离分析可以提供一般人群中致病基因的外显率,等位基因频率等参数的无偏估计。调整环境因素和假定致病基因后,可得到更准确的剩余家庭聚集,为下一步的病因研究提供最有效的信息。

尽管病例对照家系设计研究疾病的家庭聚集性分析和假定致病基因的分离分析中有很多的优良性质,但是它最大的局限性就是通过人群为基础选择的病例对照家庭中可能真正携带致病基因的家庭并不多,所以如果对所有参加者收集生物学标本进行连锁分析不仅花费大且可行性低,而且可能根本就是无效的。但是以病例对照的家系调查为依据,可以选择包含三个或更多病例的家庭,收集这些有较高概率

携带致病基因的家庭成员的生物学标本,进行连锁分析以搜寻假定的致病基因。

5. 多阶段抽样设计(multi-stage sampling design):遗传流行病学的最终目的是寻找致病基因的位置,所以收集参加成员的生物学标本是必不可少的。当资料较难收集且花费很高时,通过两个或更多阶段的抽样可能会使整个研究更有效。Whittemore^[24,25]提出多阶段抽样设计的家系研究,它扩展了流行病学中两阶段设计到家庭中。以前列腺癌的遗传研究为例,一次大规模的病例对照研究中(Stage 1)调查了前列腺癌发病风险与饮食和其他生活方式的关系。病例是最近诊断患有前列腺癌患者的随机样本(称为病例先证者),对照是没有患前列腺癌的人群中的随机样本(称为对照先证者)。为避免大量的繁琐的调查,对先证者只是简单的询问是否有父亲或兄弟也患有前列腺癌(称为家族史阳性或阴性)。依据简单的家族史资料,可以将样本分为家族史阳性的病例、家族史阴性的病例、家族史阳性的对照、家族史阴性的对照共四层。从每一层中按一定的比例随机抽样得到第一个样本中的亚样本。对亚样本详细调查每个先证者的家族史情况,如家庭大小及结构、家庭中所有成员发生前列腺癌的年龄或截尾年龄、发生前列腺癌的地点和诊断日期,每个成员的危环境因素方面的信息等等情况。利用亚样本的家庭资料可以进行家庭聚集性的测量和分离分析(Stage 2)。最后对包含三个或更多个临床诊断为前列腺癌患者的家庭,收集每个家庭成员的血样或其他生物学标本,用于遗传标记的检测,进行连锁分析定位易感基因(Stage 3)。在多阶段抽样设计中,首先要确定每阶段抽样的比例,例如在上述第二阶段中从每层抽取的亚样本比例。为达到最优化设计,抽样的比例要满足:分离分析时参数估计的方差达到最小,连锁分析时前列腺癌的假定易感基因有分离的家庭数目达到最大。例如 Whittemore^[25]的模拟研究曾表明,在上述第二阶段中最好的策略是抽取所有的家族史阳性的病例,大部分家族史阳性的对照,较少的家族史阴性的病例,和非常少的家族史阴性的对照。在上述第三阶段中最好的策略是家族史阳性的病例和对照,少量家族史阴性的病例,而家族史阴性的对照对连锁分析并不提供任何信息。但为了参数估计的一致性,需要对所有层都进行抽样,折中的策略是从家族史阴性的对照中也抽取极少数的亚样本。与以人群为基础的单纯随机抽样的病例对照家系设计相比,如按照相同含量的病例和对照确定的家系资料进行分离分析时,多阶段抽样设计具有更高的效率。

6. 一套以人群为基础的家系研究设计(a class of population-based family study designs):在多阶段设计的思想促使下和出于分析方法的考虑,Zhao 等^[3]提出一套以人群为基础的家系研究设计,包括三个阶段。每个阶段可以有若干种设计方案选择(表 1)。

第一阶段称为关联和聚集阶段,抽样主要出于流行病学考虑,主要目的是评价环境危险因素和疾病的关系。如收集

表1 一套以人群为基础的家系研究设计^[3]

阶段 A Ass/Agg	阶段 S Ass/Seg/Agg	阶段 L Ass/Linkage/Seg/Agg
病例对照研究	以家系为基础的	连续系列抽样
不匹配的	病例对照研究	有偏倚抽样
匹配的	先证者固定的亲属	按照家族史
频数匹配的	连续系列抽样	按照家族史和危险因素
病例系列研究	有偏倚抽样	
队列研究	按照家族史	
监测研究	按照家族史和危险因素	

注:Ass/Agg(Association and Aggregation):关联和聚集;Ass/Seg/Agg(Association, Segregation and Aggregation):关联、分离和聚集;Ass/Linkage/Seg/Agg(Association, Linkage, Segregation and Aggregation):关联、连锁、分离和聚集

到候选基因或遗传标记信息,也可评价它们和疾病的关联以及基因/环境因素的交互作用。基于家族史资料,可考虑疾病的家庭聚集性。在研究发病率低,而可能的暴露频率不太低的疾病时,病例对照设计可能是最有效的。队列研究也可以用于评价暴露和疾病的关联。此外,病例系列研究设计近年来有越来越多的应用^[26]。仅收集病例的环境因素的信息,所以病例系列研究不能评价环境因素和疾病的关联,但可以评价环境因素和疾病亚型的关联。在对照的生物学标本很难收集的情况下,只根据病例的候选基因信息,也可在环境因素和候选基因独立的假设下探索基因/环境的交互作用^[26]。

第二阶段称为分离阶段,主要目的是推断调整了环境因素后人群中假定致病基因的特征和调整了环境因素和假定致病基因后的剩余家庭聚集。该阶段主要是在第一阶段的基础上收集家系资料(包括亲属的疾病状态和有关环境因素的信息),可以收集第一阶段所有参加者的亲属信息,如病例的父母或同胞信息,或者包括病例和对照的所有亲属。也可以采用有偏倚抽样,如只收集家族史阳性的先证者亲属的信息,或收集具有(没有)某危险因素的先证者亲属信息,如研究非吸烟者中肺癌的遗传模式。还有一种方式是对每个家庭中固定数目的亲属连续抽样。如先收集病例的一级亲属的信息,再收集一级亲属中的病例的一级亲属的信息,这样连续抽样,直至没有新的病例发现或没有可继续加入的亲属。这种抽样方式在确认携带假定致病基因的家庭时是非常有效的,但在一般人群中实施这种设计的可行性问题还应考虑。

第三阶段称为连锁阶段,其主要目的是完成遗传流行病学的最终任务,定位假定致病基因。所以遗传设计主要考虑抽取最可能携带致病基因的家庭,以便提高连锁分析的效率。如前所述,连续家庭抽样是不错的选择。此外,根据家族史或高危环境因素的有偏倚抽样可能有效的搜寻到与危险因素有交互作用的假定致病基因。对最终确定的家庭成员收集生物学标本。

所有上述三个阶段收集到的信息都可用于关联、连锁、分离和聚集分析。特别是第一、二阶段的信息可以分析环境因素和疾病的关联,假定致病基因的特征和剩余的家庭聚

集。虽然连锁分析对抽样的随机性并无要求,而估计也是稳健的,但它的效率直接依赖前两阶段估计的参数的准确性。所以需要以人群为基础的研究设计。此组设计与上述的多阶段设计有一些共同的特征,但它们最主要的区别在于有不同的倾向性:多阶段设计实际上主要考虑抽样操作的逻辑性;以人群为基础的家庭研究设计主要集中在关联、聚集、分离和连锁分析的效率上。它们并不互相排斥,而是可以互相补充。

7. 结语 探索复杂疾病的病因是一项长期而艰巨的任务,选择科学合理的设计方案可以大大提高病因学分析的效率。每种设计方案有各自的适用侧重,例如在识别环境因素和慢性疾病的关系时,传统的病例对照设计功不可没;而在疾病遗传因素的研究中,家系设计也有不可替代的作用;在研究疾病的家庭聚集性和识别疾病的遗传模式时,病例对照家系设计可能是不错的选择;而在对假定致病基因定位时,则需要有偏的选择携带致病基因概率高的家庭;当致病基因定位后,需要研究候选基因对疾病的影响以及基因/环境交互作用时,可选择家系基础的病例对照设计或人群基础的病例对照设计。但是每种设计方案又存在各自的局限性,虽然有些局限可以通过统计分析方法的改进来克服^[27-31]。随着人们对病因认识的不断深入,需要综合考虑疾病各种影响因素的作用,这也需要对设计方案和分析方法的修正和结合。例如由于分子生物学技术的迅速发展,特别是染色体扫描的可能,为提高分析效能,可以先利用病例对照设计研究大量的遗传标记和疾病的关联,对阳性的标记位点再用以家系为基础的病例对照设计进行关联分析和连锁分析^[28]。另外,多阶段抽样设计^[25]和一套以人群为基础的家系研究设计^[3]的提出也是各种设计方案取长补短的应用。由于设计方案的复杂性,必然也带来传统的统计分析方法的不适应,对使用各种设计形式以及各种设计形式的结合收集到的资料的统计分析方法也在发展之中。总之,疾病的病因是复杂的,认识疾病病因的过程也必然是艰辛和漫长的,但是在以往的研究中,我们认识了吸烟是肺癌的危险因素,也证实了BRCA1、BRCA2是乳腺癌的易感基因,我们有理由相信依靠遗传学、流行病学、统计学、分子生物学、计算机科学等等各个学科工作者共同的努力,可以最终揭开复杂疾病的病因之谜。

参 考 文 献

- Schlesselman JJ. Case control study. New York :Oxford University Press, 1982.
- 沈福民,主编. 流行病学原理与方法. 上海:复旦大学出版社, 2001.
- Zhao LP, Hsu L, Davidov O, et al. Population-based family study designs: an interdisciplinary research framework for genetic epidemiology. Genet Epidemiol, 1997, 14: 365-388.
- Khoury MJ, Beaty TH, Cohen BH. Fundamentals of genetic epidemiology. London :Oxford University Press, 1993.
- Szatmari P, Jones MB. Effects of misclassification on estimates of relative risk in family history studies. Genet Epidemiol, 1999, 16: 368-381.

6 Ottman R. Gene-environment interaction: definitions and study designs. *Prev Med*, 1996, 25:764-770.

7 Hwang SJ, Beaty TH, Liang KY, et al. Minimum sample size estimation to detect gene-environment interaction in case control designs. *Am J Epidemiol*, 1994, 140:1029-1037.

8 Zhu XF, Zhang SL, Zhao HY, et al. Association mapping, using a mixture model for complex traits. *Genet Epidemiol*, 2002, 23:181-196.

9 Lander ES, Schork NJ. Genetic dissection of complex traits. *Science*, 1994, 265:2037-2048.

10 Elston RC. Introduction and overview. *Statistical Methods in Medical Research* 2000, 9:527-541.

11 Chakraborty R, Hanis CL. Non-random sampling in human genetics: estimation of familial correlations, model testing, and interpretation. *Statistics in Medicine*, 1987, 6:629-646.

12 Olson JM, Cordell HJ. Ascertainment bias in the estimation of sibling genetic risk parameters. *Genet Epidemiol*, 2000, 18:217-235.

13 Weinberg C, Umbach D. Choosing a retrospective design to assess joint genetic and environmental contributions to risk. *Am J Epidemiol* 2000, 152:197-203.

14 Witte JS, Gauderman WJ, Thomas DC. Asymptotic bias and efficiency in case control studies of candidate genes and gene-environment interactions: basic family designs. *Am J Epidemiol*, 1999, 148:693-705.

15 Lubin J, Gail M. Biased selection of controls for case-control analysis of cohort studies. *Biometrics*, 1984, 40:63-75.

16 Elston RC, Olson JM, Palmer L. Biostatistical genetics and genetic epidemiology. John Wiley & Sons LTD 2002. 267-275.

17 Spielman RS, McGinnis RE, Ewens WJ. Transmission test for linkage disequilibrium: the insulin gene region and insulin dependent diabetes mellitus (IDDM). *American Journal of Human Genetic*, 1993, 52:506-516.

18 Zhu XF, Elston RC. Transmission/disequilibrium tests for quantitative traits. *Genet Epidemiol* 2001, 20:57-74.

19 Schaid DJ. Case-parents design for gene-environment interaction. *Genet Epidemiol* 1999, 16:261-273.

20 Thomas DC. Case-parents design for gene-environment interaction by Schaid. *Genet Epidemiol* 2000, 19:461-463.

21 Claus EB, Risch NJ, Thompson WD. Age at onset as an indicator of familial risk of breast cancer. *Am J Epidemiol*, 1991, 131:961-972.

22 Mettlin C, Croghan I, Natarajam N, et al. The association of age and familial risk in a case-control study of breast cancer. *Am J Epidemiol*, 1991, 131:973-983.

23 Liang KY, Beaty TH. Statistical designs for familial aggregation. *Statistical Methods in Medical Research* 2000, 9:543-562.

24 Whittemore AS. Multi-stage sampling designs and estimating equations. *Journal of the Royal Statistical Society, Series B*, 1997, 59:589-602.

25 Whittemore AS, Halpern J. Multi-stage sampling in genetic epidemiology. *Statistics in Medicine*, 1997, 16:153-167.

26 Khoury MJ, Flanders WD. Nontraditional epidemiologic approaches in the analysis of gene-environment interaction: case-control studies with no controls. *Am J Epidemiol*, 1996, 144:207-213.

27 Martin RB, Alda M, Maclean CJ. Parental genotype reconstruction: applications of haplotype relative risk to incomplete parental data. *Genet Epidemiol*, 1998, 15:471-490.

28 Martin ER, Kaplan NL. A monte carlo procedure for two-stage tests with correlated data. *Genet Epidemiol* 2000, 18:48-62.

(收稿日期:2002-12-23)
(本文编辑:尹廉)

· 疾病控制 ·

贵州省龙里县新闻媒体曝光前后餐饮具消毒效果监测

王培文 磨剑 杨桦 窦启平

搞好餐饮具消毒是预防食源性疾病传播的重要措施。餐饮具消毒效果监测是食品卫生监督管理工作中的一项重要内容,为此我们对 2002 年抽检的 1411 份餐饮具进行分析,并比较新闻媒体曝光前后餐饮具消毒效果。

随机抽取龙里县城区饮食行业大型酒店宾馆 44 家,中型餐馆 103 家,学校、机关食堂 21 家,饮摊和小型餐馆 370 家的各类餐饮具。

采用大肠菌群定性检测。大肠菌群快速检验纸片(25 cm² × 2)为南京三爱实业有限公司产品,卫生部食品卫生监督检验所监制。现场随机抽取准备使用的餐饮具,每家餐馆采样 5~10 份,每份样品内贴纸片 2 张(1 份)。用无菌生理盐水湿润大肠菌群检测用纸后,立即贴于食具内侧表面 30 s 后取下置于无菌塑料袋内。筷子以 5 只为一份样品,用毛细吸管吸取无菌生理盐水湿润纸片后,立即将筷子一端(约 5 cm)抹拭纸片,每份样品抹拭 2 张,放入无菌塑料袋内送检。将已采样的纸片置 37℃ 温箱培养 16~18 h,若纸片保持紫蓝

色不变为大肠菌群阴性,纸片变黄并在黄色背景上呈现红色斑点或片状红晕为阳性。按照 GB14934-1994 进行评价,阴性为合格,阳性为不合格。

新闻媒体曝光前(3~6 月份)随机抽取 154 家饮食业餐饮具 763 份,合格 244 份,合格率 32%,合格率较低。通过新闻媒体宣传曝光之后(7~10 月份)随机抽取 132 家饮食业餐饮具 648 份,合格 479 份,合格率为 74%,差异有非常显著性意义($\chi^2 = 253.5, P < 0.01$)。并对城区各类型餐饮具合格率进行比较,大型酒店合格率为 86%(219/264),好于其他形式餐饮业,其中以饮食摊餐具消毒合格率最低为 30%(111/370),经统计学处理差异有非常显著性意义($\chi^2 = 434.3, P < 0.01$)。

本次调查显示,通过新闻媒体曝光之后餐饮具合格率明显提高,在市场经济条件下饮食经营业主最怕因卫生不合格曝光而影响经济效益。各类餐饮业以大型酒店餐具合格率较高,饮食摊最低,主要是经营者法制观念淡薄,不重视餐饮具消毒,对餐饮具消毒保洁设施投入不足,各项消毒制度不健全,措施不落实,无专人负责,部分从业人员卫生意识缺乏。

(收稿日期:2003-07-21)
(本文编辑:张林东)

作者单位:551200 贵州省龙里县卫生监督所(王培文、窦启平);
龙里县疾病预防控制中心(磨剑、杨桦)