

## · 基础理论与方法 ·

## 用广义估计方程估计数量性状的家庭相关

郜艳晖 姜庆五

**【摘要】** 目的 研究数量性状家庭相关的测量方法,并对身高家系资料进行分析。方法 应用广义估计方程(GEE2)估计数量性状的均数结构和关联结构的边际回归模型。所有估计都可在软件MAREG中实现。并用身高家系实例说明方法的应用。结果 GEE2可同时考虑协变量对性状的影响和性状的内部相关,得到回归系数和相关系数的稳健估计。对身高家系资料分析,调整性别、18岁前主要居住地和出生年代后,亲子相关( $r = 0.459$ )和同胞相关( $r = 0.671$ )高于配偶相关( $r = 0.184$ ),有统计学意义。同一类型亲属对中,同性别相关(如父子 $r = 0.603$ ,母女 $r = 0.456$ ,兄弟 $r = 0.947$ ,姐妹 $r = 0.681$ )大于异性别相关(如父女 $r = 0.431$ ,母子 $r = 0.364$ ,兄妹或姐弟 $r = 0.530$ )。结论 GEE2可灵活的估计各种家庭相关系数和协变量对性状均数的影响,且参数估计稳健,因此可作为评价数量性状家庭聚集性的标准方法之一。

**【关键词】** 家庭相关;广义估计方程2;身高;数量性状

**Estimating family correlation of quantitative traits using generalized estimating equation** GAO Yan-hui, JIANG Qing-wu. Department of Epidemiology, College of Public Health, Fudan University, Shanghai 200032, Chian

**【Abstract】 Objective** To study the method for measuring familial correlations of quantitative trait and analyzing family data set of body height. **Methods** Generalized estimating equation (GEE2) was employed to estimating both regression coefficients and the familial correlation. Analyses was carried out on software MAREG. A example from height pedigrees illustrated the method. **Results** GEE2 provided robust estimations of regression coefficients and familial correlations simultaneously. In body height the correlations between parents and offspring ( $r = 0.459$ ) and between siblings ( $r = 0.671$ ) were significantly higher than those between two parents ( $r = 0.184$ ) after adjusting gender, residence and birth age. Of the same types of relative pairs, the correlation between pairs with individuals of the same gender (eg. father-son  $r = 0.603$ , mother-daughter  $r = 0.456$ , male sibling  $r = 0.947$ , female sibling  $r = 0.681$ ) was higher than those individuals of different gender (eg father-daughter  $r = 0.431$ , mother-son  $r = 0.364$ , sibling with different gender  $r = 0.530$ ). **Conclusion** GEE2 should be considered a standard method for the investigation of familial aggregation due to its flexibility and robustness.

**【Key words】** Family correlation; Generalized estimating equation 2; Body height; Quantitative trait

对于疾病相关性状的资料,我们常常关心:①可疑危险因素对性状指标的影响大小;②性状指标在家庭成员中的相似性,这种家庭相似性可以用家庭各成员关系间的相关来衡量。在调整危险因素对性状指标的影响后测量到的家庭相关可以为该性状是否受遗传因素的影响提供线索。一般来说,家庭相关越大,性状具有遗传因素影响的证据越强。进一步可以研究家庭相似性的模式,例如,如果一级亲属间相关高于二级亲属间相关,提示遗传因素有重要的作用;如果同胞之间的相关高于父母相关,提示可

能存在未知基因的显性效应;如果母亲和子女的相关高于父亲和子女的相关,提示遗传机制可能不是简单的常染色体遗传等。类似于家系资料等非独立数据的统计方法研究近年来受到统计学者的重视<sup>[1,2]</sup>。其中,广义估计方程(generalized estimating equation, GEE)的方法有越来越多的应用<sup>[3,4]</sup>。如果研究的兴趣主要集中在回归系数上,可以利用 Liang Zeger<sup>[5]</sup>提出的广义估计方程(GEE1)。但在某些研究领域如遗传流行病学中,除了对回归系数感兴趣外,家庭相关往往更值得关注。Prentice<sup>[6]</sup>扩展了GEE1,建立第二套估计方程(GEE2),同时估计均数结构和关联结构的边际模型。我们主要将GEE2应用于数量性状的家系资料,并对327个核心家系的身高资料进行了家庭相关分析。

基金项目 国家自然科学基金资助项目(39930160)

作者单位 200032 上海,复旦大学公共卫生学院流行病学教研室

### 原理与方法

1. 均数和关联结构的边际模型: 设  $Y_i = (Y_{i1}, Y_{i2}, \dots, Y_{in_i})'$  ( $i = 1, 2, \dots, N$ ) 表示第  $i$  个家庭的响应变量向量,  $X_i = (X_{i1}, X_{i2}, \dots, X_{in_i})'$  表示相应的  $n_i \times p$  维的协变量矩阵。本文主要考虑数量性状, 因此用恒等连接函数连接第  $i$  个家庭中第  $j$  个个体的响应变量  $Y_{ij}$  的边际均数和协变量  $X_{ij}$

$$\mu_{ij} = E(Y_{ij} | X_i) = E(Y_{ij} | X_{ij}) = X_{ij}\beta \quad (1)$$

式中  $\beta$  为  $p$  维回归系数向量, 反映协变量对响应变量的影响大小。在第  $i$  个家庭中, 第  $j$  和第  $k$  个个体的样本相关系数用 Pearson's 乘积矩相关系数表示为  $Z_{ijk} = (Y_{ij} - \mu_{ij})(Y_{ik} - \mu_{ik}) / \sigma_{ij}\sigma_{ik}$ ,  $\sigma_{ij} = [\text{var}(Y_{ij} | X_i)]^{1/2}$ ,  $Z_{ijk}$  的均数通常用 Fisher's  $z$  变换的倒数<sup>[7]</sup>

$$\rho_{ijk} = \text{corr}(Y_{ij}, Y_{ik} | X_i) = \frac{\exp(e'_{ijk}\alpha) - 1}{\exp(e'_{ijk}\alpha) + 1} \quad (2)$$

或表示为

$$\logit\left(\frac{1 + \rho_{ijk}}{2}\right) = e'_{ijk}\alpha$$

式中  $e_{ijk}$  为第  $i$  个家庭中第  $j$  和第  $k$  个个体共有的  $q \times 1$  维亲属关系变量和协变量向量,  $\alpha$  为相应的  $q$  维关联参数向量, 反映亲属相关的大小和协变量对亲属相关的影响。式(2)保证了相关系数介于  $(-1, 1)$  之间, 更符合通常意义的解释。

求解下列估计方程可得到参数  $(\beta, \alpha)$  的估计

$$u\left(\begin{matrix} \hat{\beta} \\ \hat{\alpha} \end{matrix}\right) = \sum_{i=1}^N \begin{pmatrix} D_i & 0 \\ 0 & C_i \end{pmatrix} \begin{pmatrix} V_i & 0 \\ 0 & W_i \end{pmatrix}^{-1} \begin{pmatrix} Y_i - \mu_i \\ Z_i - \rho_i \end{pmatrix} = 0 \quad (3)$$

式中  $\mu_i = (\mu_{i1}, \mu_{i2}, \dots, \mu_{in_i})'$  为  $n_i \times 1$  维向量,  $Z_i = (Z_{i12}, Z_{i13}, \dots, Z_{i23}, \dots, Z_{i(n_i-1)n_i})'$ ,  $\rho_i = (\rho_{i12}, \rho_{i13}, \dots, \rho_{i23}, \dots, \rho_{i(n_i-1)n_i})'$ , 均为  $n_i(n_i - 1)/2 \times 1$  维向量。

$D_i = \partial\mu_i / \partial\beta$ ,  $C_i = \partial\rho_i / \partial\alpha$  分别是给定  $X_i$  时  $Y_i$  和  $Z_i$  的均数对  $\beta$  和  $\alpha$  的一阶导数。 $V_i$  和  $W_i$  分别是给定  $X_i$  时  $Y_i$  和  $Z_i$  的“工作”协方差矩阵。 $V_i = A_i^{-1/2} R_i(\alpha) A_i^{-1/2}$  其中  $A_i = \text{diag}\{\text{var}(Y_{ij} | X_i)\}$  完全由边际分布决定, 即只和  $\beta$  有关。 $R_i(\alpha) = \text{corr}(Y_i)$  是  $Y_i$  的相关矩阵。通常在应用中只将  $W_i$  作为“工作”矩阵, 为方便设为  $n_i(n_i - 1)/2$  维单位矩阵。对式(3)的求解可通过交互修正 Fisher scoring 算法。可以证明, 只要正确指定式(1)和(2), 在常规条件下  $(\hat{\beta}, \hat{\alpha})$  是  $(\beta, \alpha)$  的强一致估计, 并且  $\hat{\beta}$  和  $\hat{\alpha}$  是近似联合正态分布, 其均数向量为 0, 参数的稳健方差矩阵

(或称 Huber 或 Sandwich 方差矩阵) 参见文献 [5, 6]。

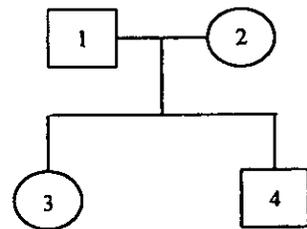
上述所有参数的估计都可在软件 MAREG 中实现, 该软件专门用于分析边际回归模型, 目前它可以在互联网 <http://www.stat.unimuenchen.de> 上免费得到。

2. 关联结构设计矩阵的构建: 在遗传流行病学中, 除了关心回归系数的大小外, 家庭各成员关系间的相关更受重视。式(2)中的  $e_{ijk}$  可由用户灵活定义, 它可以是表示家庭成员关系的变量, 也可以是亲属对共有的协变量。根据家系资料结构和研究兴趣, 建立适当的关联结构设计矩阵。

当每个家庭大小相等, 结构相同时, 可以指定  $n(n-1)/2 \times q$  维设计矩阵 ( $n$  为每个家庭成员数) 为各个家庭共享。当家庭大小不等或结构不同时, 每个家庭的设计矩阵不同, 要根据每个家庭具体情况构建。本文用 SAS macro 构建设计矩阵。

(1) 只有同胞的资料: 假设同胞间的相关是相同的, 即相关矩阵为可交换的 (exchanged), 此时式(2)相当于只有截距项, 设计矩阵为全是 1 的向量。假设不同性别的同胞可能具有不同的相关, 可令  $e_{ijk} = 1(0)$  表示不同(相同)性别的同胞对, 或者直接指定  $e_{ijk1} = 1$  为兄弟关系,  $e_{ijk2} = 1$  姐妹关系,  $e_{ijk3} = 1$  兄妹(或姐弟)关系, 否则为 0。对应的关联参数  $(\alpha_1, \alpha_2, \alpha_3)$  反映不同同胞关系的相关大小。

(2) 核心家系资料: 核心家系指仅包括父母和子女的家庭, 这种数据是遗传流行病学中最常见的资料类型。以图 1 的家系结构为例说明不同研究兴趣时的设计矩阵(表 1)。



方框表示男性, 圆框表示女性  
方框或圆框内数字为家庭成员的编号

图1 父母及两个子女的核心家系

4 个成员的家庭共构成 6 个亲属对, 根据研究兴趣构建关联结构的设计矩阵。例如研究兴趣 1 考察家庭内配偶相关, 亲代-子代相关和同胞相关。假设家庭所有成员共享相同的教养环境, 如果同胞相

表1 不同研究兴趣时的关联结构设计矩阵

亲属对	研究兴趣 1			研究兴趣 2				研究兴趣 3							
	$\rho_{pp}$	$\rho_{po}$	$\rho_{oo}$	$\rho_{pp}$	$\rho_{jo}$	$\rho_{mo}$	$\rho_{oo}$	$\rho_{pp}$	$\rho_{js}$	$\rho_{fd}$	$\rho_{ms}$	$\rho_{md}$	$\rho_{ss}$	$\rho_{sd}$	$\rho_{dd}$
1~2	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0
1~3	0	1	0	0	1	0	0	0	0	1	0	0	0	0	0
1~4	0	1	0	0	1	0	0	0	1	0	0	0	0	0	0
2~3	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0
2~4	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0
3~4	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0

注： $\rho_{pp}$ ,  $\rho_{po}$ ,  $\rho_{oo}$  分别表示配偶相关、亲代-子代相关和同胞相关； $\rho_{jo}$ ,  $\rho_{mo}$ ,  $\rho_{oo}$  分别表示配偶相关、父亲-子代相关、母亲-子代相关和同胞相关； $\rho_{pp}$ ,  $\rho_{js}$ ,  $\rho_{fd}$ ,  $\rho_{ms}$ ,  $\rho_{md}$ ,  $\rho_{ss}$ ,  $\rho_{sd}$ ,  $\rho_{dd}$  分别表示配偶相关、父子相关、父女相关、母子相关、母女相关、兄弟相关、兄妹(或姐弟)相关和姐妹相关

关远远大于配偶相关,则增强了性状由遗传决定的证据。研究兴趣 2 将亲代-子代相关进一步细分为父亲-子代相关和母亲-子代相关,通过比较  $\rho_{jo}$  和  $\rho_{mo}$ ,可检验性状是否为母系遗传机制或父系遗传机制。研究兴趣 3 全面考虑家庭各成员对的性别差异,为性状是否有性染色体遗传机制提供线索。

除利用设计矩阵识别家庭成员关系外,设计矩阵中还可以包括其他协变量来检验某些特定的假设。例如,如果想检验家庭中有血缘关系的亲属对(如父子、母子、同胞等)间相关和配偶间相关是否相等  $H_0: \rho_{pp} = \rho_{po}, \rho_{pp} = \rho_{oo}$ ,可将设计矩阵中的第一列替换为常数项,如全为 1 的向量。此外,还可以检验亲属对共有的协变量对相关的影响,例如可添加家庭别的协变量[如研究的家庭来自不同人群(如种族),可添加种族协变量],研究不同性质的家庭对家庭相关的影响。

(3) 扩展家系资料 利用扩展家系资料可以构建更多类型亲属对的设计矩阵,如祖父母/外祖父母-孙子女/外孙子女、叔伯姑/姨舅-侄子女/外甥女、堂兄弟姐妹/表兄弟姐妹等关系类型。进一步也可以检验一级亲属相关和二级亲属相关是否相等,为遗传因素的作用提供证据。

实例分析

1. 身高家系资料的家庭相关分析:327 个核心

家系,共计 1 166 人,对不满 18 周岁的子女删去不计,剩余 1 150 人。每个家庭成员数为 3~8 人,相应的家庭数分别是 202(61.7%)、95(29.05%)、22(6.73%)、3(0.92%)、4(1.22%)、1(0.31%)。调查因素包括身高、性别、出生年月,18 岁前主要居住地(城市或农村)等。父母亲共 654 人,平均年龄为 50.99 岁 ± 5.46 岁。子女共计 496 人,平均年龄为 24.29 岁 ± 5.18 岁。子女中男性为 223 人(44.96%),女性为 273 人(55.04%)。

应用软件 MAREG 对 327 个核心家系身高资料同时拟合均数和关联结构的边际模型。关联结构的设计矩阵用 SAS macro 编程产生,保存为一个文本文件。分析结果见表 2~4。

由表 2 均数结构的参数估计可知,男女的平均身高相差 0.118 m,城市居民比农村居民平均高 0.015 m,身高的性别差异,城乡差异均有显著性意义。1955 年以前出生者、1955 年到 1975 年间出生者和 1975 年以后出生者相比,平均低 0.014 m 和 0.006 m,虽后者差别无统计学意义,但也可看出随年代增长,人类平均身高的增加趋势。表 2 拟合了配偶、亲子和同胞的关联结构,三类亲属相关均有显著性意义。配偶相关符合常理对身高的选型婚配。同胞相关系数、亲子相关系数均大于配偶相关系数,提示身高的遗传作用。为进一步分析可能是父系遗传还是母系遗传,采用关联结构 2 重新拟合边际模

表2 327 个家系、1 150 个个体身高的均数结构和关联结构 1 的广义估计方程结果

变量	编码	参数估计	s	P 值	相关系数(r)
均数结构					
截距	1	1.602	0.004	<0.001	
性别	1 = 男性 $\rho$ = 女性	0.118	0.002	<0.001	
居住地	1 = 城市 $\rho$ = 农村	0.015	0.003	<0.001	
出生年代 1*	1 = 1955 年以前 $\rho$ = 其他	-0.014	0.002	<0.001	
出生年代 2*	1 = 1955~1975 年 $\rho$ = 其他	-0.006	0.003	0.072	
关联结构 1					
配偶	1 = 是 $\rho$ = 否	0.372	0.114	0.001	0.184
亲代-子代	1 = 是 $\rho$ = 否	0.993	0.080	<0.001	0.459
同胞	1 = 是 $\rho$ = 否	1.624	0.238	<0.001	0.671

\* 出生年代以 1975 年及以后为参照

型。结果见表 3。

表3 327 个家系、1 150 个个体身高的均数结构和  
关联结构 2 的广义估计方程结果

变量	亲属 对数目	参数 估计	s	P 值	相关系数 (r)
均数结构					
截距		1.602	0.004	<0.001	
性别		0.118	0.002	<0.001	
居住地		0.015	0.003	<0.001	
出生年代 1		-0.014	0.002	<0.001	
出生年代 2		-0.006	0.003	0.062	
关联结构 2					
配偶	327	0.372	0.114	0.001	0.184
父亲-子代	496	1.120	0.121	<0.001	0.508
母亲-子代	496	0.873	0.108	<0.001	0.411
同胞	234	1.628	0.238	<0.000	0.672

注:变量编码同表 2

表 3 中的均数结构的参数估计和表 2 基本相同,说明不同的相关结构对均数参数的估计是稳健的,这也是 GEE 的特点和优点。从关联结构的参数估计中可以看到,配偶相关系数和同胞相关系数几乎同表 2,而父亲-子代相关系数略大于母亲-子代相关系数。提示身高的遗传中,父亲对子代的遗传作用可能更大。

采用关联结构 3 进一步分析性别差异对身高相关的影响。结果见表 4。

表4 327 个家系、1 150 个个体身高的均数结构和  
关联结构 3 的广义估计方程结果

变量	亲属 对数目	参数 估计	s	P 值	相关系数 (r)
均数结构					
截距		1.600	0.004	<0.001	
性别		0.117	0.003	<0.001	
居住地		0.019	0.003	<0.001	
出生年代 1		-0.014	0.002	<0.001	
出生年代 2		-0.004	0.003	0.161	
关联结构 3					
配偶	327	0.376	0.115	0.001	0.186
父子	223	1.396	0.210	<0.001	0.603
父女	273	0.922	0.149	<0.001	0.431
母子	223	0.763	0.154	<0.001	0.364
母女	273	0.983	0.153	<0.001	0.456
兄弟	55	3.605	2.617	0.168	0.947
兄妹或姐弟	67	1.181	0.263	<0.001	0.530
姐妹	112	1.662	0.456	<0.001	0.681

注:变量编码同表 2

可见,均数结构的参数估计仍然是稳健的。关联结构中的配偶相关也近似表 2 和表 3。表 3 中父亲-子代的相关分解为表 4 中的父子和父女相关,且父子的相关系数略大于父女的相关系数(检验)。同样表 3 中母亲-子代的相关分解为表 4 中的母子和

母女相关,且母女略大于母子相关系数(检验)。同胞相关分解为兄弟、兄妹或姐弟和姐妹相关,其中兄弟相关系数最大,兄妹或姐弟相关系数最小(检验)。总之,所有有血缘关系的同类型亲属对相关中,同性别的亲属对都比异性别的亲属对相关系数要大。表明身高的家庭相关中,性别可能是主要的影响因素。

在表 4 中,我们注意到虽然兄弟间具有最高的相关系数 0.947,但关联参数检验的 P 值是 0.168,并没有统计学意义。这个反常的现象可能和兄弟对的数目较少( $n = 55$ )有关。此外,结合表 2~4,我们还发现同一类型的亲属对相关系数近似为各亚型亲属对相关系数的加权平均,权重为亚型亲属对的数目。例如父亲-子代相关系数为 0.508,近似等于  $(0.603 \times 223 + 0.431 \times 273) / (223 + 273)$ 。

## 讨 论

家系资料中研究性状之间不独立,所以不能使用传统的线性模型来估计协变量对性状的影响。我们采用统计学者和流行病学者都比较熟悉的 Pearson's 乘积矩相关系数来测量家庭相关。这个指标用于数量性状具有简单直观,易解释等优点。如果是仅包含双胞胎的设计研究,可直接用相关系数估计遗传度,即由于一个或多个常染色体座位导致的遗传方差占总变异的比例。如果仅包含同胞的设计研究,同胞对相关实际上是同胞对间的方差占同胞对内方差与同胞对间方差之和的比例,它可以被看成所有可能的同胞对的平均类内相关<sup>[8]</sup>。如果是核心家系资料,可同时估计类内相关(intraclass correlation)如同胞相关,配偶相关)和类间相关(interclass correlation)如亲子相关等。这种家庭相关也可以用混合效应模型<sup>[2]</sup>采用迭代广义最小二乘法来估计,但此模型前提要对多个不独立的响应变量的分布做出假设。而使用广义估计方程 2,无须指定响应变量的联合分布,只要正确指定均数和相关系数的连接函数,就可同时对均数结构和关联结构进行参数估计,且估计结果稳健<sup>[5,6]</sup>。进一步通过关联结构设计矩阵的修改,GEE2 可以灵活的估计各种亲属关系类型间的相关和协变量对相关的影响。因此该法可考虑作为数量性状的家庭聚集性研究的一种标准方法。

虽然性状的家庭相关可由共享的环境或遗传引起,但强的家庭相关还是可以为遗传作用提供证据。特别是特殊的家庭相关模式可为遗传机制提供线

索。利用 GEE2 的方法分析身高家系资料,在调整了性别、18 岁前主要居住地和出生年代对身高均数的影响后,同胞相关和亲子相关都大于配偶相关(通过改变关联结构的设计矩阵进行统计学检验  $P < 0.001$  结果未列出)。同一性质亲属关系类型中,同性别的亲属对间相关大于异性别的亲属对间相关。在亲子间和同胞间,又分别以父子,兄弟间相关为大,是否提示在身高的遗传机制中,性染色体的作用?当然,本研究中三口之家占了很大比例(61.7%),子女中男女比例的不均衡(223:273),致使同胞对数目(特别是兄弟对数目)相对较少。另外,调查中未考虑身高的其他影响因素等原因是否会影响相关参数的估计,还有待于进一步研究。

## 参 考 文 献

- 1 徐勇勇,陈长生,曹秀堂,等.医学与卫生统计资料的系统结构数据.中国卫生统计,1995,12:12-15.
- 2 陈峰,任仕泉,陆守曾.非独立计量资料的内部相关性研究.现代预防医学,1998,25:269-271.
- 3 熊林平,曹秀堂,徐勇勇,等.纵向观测计数数据的对数线性模型.中国卫生统计,1999,16:68-71.
- 4 李康,魏韦,王滨友,等.医学诊断试验评价的 ROC 分析.中国医院统计,2001,8:3-6.
- 5 Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. Biometrika, 1986, 73:13-22.
- 6 Prentice RL. Correlated binary regression with covariates specific to each binary observation. Biometrics, 1988, 44:1033-1048.
- 7 Lipsitz SR, Laird NM, Harrington DP. Generalized estimating equations for correlated binary data: using the odds ratio as a measure of association. Biometrika, 1991, 78:153-160.
- 8 Liang KY, Beaty, TH. Statistical designs for familial aggregation. Statistical methods in medical research, 2000, 9:543-562.

(收稿日期 2003-01-03)

(本文编辑:张林东)

## · 疾病控制 ·

## 新疆石油勘探区人群莱姆病血清流行病学调查

田 楨 万康林 侯学霞 郝琴

为查明新疆石油勘探地区莱姆病感染情况,对该地区 5 577 名人群进行了莱姆病血清流行病学调查。

1. 材料与方 法 河南石油勘探局新疆石油勘探区地处天山南麓焉耆盆地腹心偏南,属巴音郭楞蒙古自治州的焉耆和博湖两县,盆地地貌,一般海拔 1 000 m 以上,植被以人工植被为主。采集辖区内 5 577 名人员血清标本。

间接免疫荧光抗体试验(IFA)抗原片由中国疾病预防控制中心传染病预防控制所(传染病所)提供,抗原菌株用  $FP_1$ ;血清用 3% 卵黄 PBS(0.01 mol/L pH 7.2)液,按倍比稀释,每稀释度各取 15  $\mu$ l 加于抗原片上,37℃ 保温盒内孵育 30 min,取出后用 PBS 泡 10 min,蒸馏水洗涤 3 次,室温风干;每孔加入 15  $\mu$ l 羊抗人荧光抗体 IgG 或羊抗人 IgM,37℃ 保温盒内孵育 30 min,取出后用 PBS 泡 10 min,蒸馏水洗涤 3 次,室温风干;用 90% 甘油封片,荧光显微镜检查,每视野(40×10)有 50% 螺旋体染色且清晰可见者为阳性。IgG 阳性稀释度达 1:128 或以上者,IgM 阳性稀释度达 1:64 或以上者,可判断为莱姆病血清抗体阳性。酶联免疫吸附试验(ELISA)抗原板由传染病所提供,抗原菌株用  $FP_1$ ;血清用 PBS-Tween 1:500 稀释,取 50  $\mu$ l 加于抗原反应孔内,同时加入辣根过氧化物酶标记的 SPA 或羊抗人 IgM 50  $\mu$ l,37℃ 孵育 40 min,取出洗涤 5 次,室温风干;加入 2N  $H_2SO_4$  终止反应;用 ELISA 检测仪测定,波长 450 nm,空白对照调零,  $A = 0.13$  判为阳性。对检出的阳性血清进一步检测梅毒螺旋体抗体和钩端螺旋体抗体。

梅毒抗体检测试剂盒由吉比爱生物技术(北京)有限公司生产,操作方法按试剂盒说明进行。钩端螺旋体标准菌 13 群 15 型菌株抗原由传染病所提供,用 MAT 检测钩端螺旋体。

2. 结果 5 577 名人群中,莱姆病抗体阳性血清 750 份,阳性率为 13.45%(750/5 577)。根据工作性质将 19 个单位分为 3 个观察组,Ⅰ组为首先进入疫区的野外勘探作业单位,Ⅱ组为野外开发建设施工单位,Ⅲ组为后勤服务单位。结果显示,Ⅰ、Ⅱ、Ⅲ组作业单位感染率分别为 13.72%(255/1 859)、13.71%(378/2 757)、12.17%(117/961)。不同作业人员莱姆病感染率不同,但差异无显著性( $\chi^2 = 1.62, P > 0.05$ )。感染者中年龄最小为 6 岁,最大年龄为 61 岁。将所有人员分为 5 个年龄组(2~、20~、30~、40~ 和 50~ 岁),阳性率分别为 15.13%(18/119)、12.77%(175/1 370)、13.85%(290/2 094)、15.63%(198/1 267)和 9.49%(69/727),各年龄组间的感染率差异无显著性( $\chi^2 = 3.91, P > 0.05$ )。用梅毒抗体检测试剂盒和钩端螺旋体病 MAT 对莱姆病抗体阳性血清进行检测,均未检出阳性结果。

3. 结论 采用 IFA 和 ELISA 法对新疆石油勘探区 5 577 名职工进行了莱姆病感染调查,感染 750 例,平均感染率为 13.45%,高于 1992 年全国流行病学调查的人群感染率,与我国莱姆病主要流行区东北林区相似;本次调查为国内首次在新疆地区的石油勘探区开展莱姆病流行病学调查。结果提示,新疆地区为莱姆病高流行区。

(收稿日期 2002-10-29)

(本文编辑:张林东)

作者单位 473132 南阳,河南石油勘探局卫生防疫站(田楨);中国疾病预防控制中心传染病预防控制所(万康林、侯学霞、郝琴)