

肿瘤流行病学研究资料的统计分析

项永兵

第四讲 生存资料单变量分析中危险率比估计

Logrank 检验^[1,2]是生存分析中最常用的非参数统计方法。在资料分析时,除计算假设检验统计量及相应 P 值外,还可以对研究因素的效应做出估计,即估计一些参数或指标,例如回归系数 β 或相对危险度 RR。单因素分析中常用的参数是相对危险度(RR)。在生存分析中该术语经常被称为危险率比(hazard ratio, HR)。国人在生存资料单因素分析的结果中仅报道假设检验统计量及 P 值,很少给出 RR 或 HR 的估计值,一般在 Cox 回归模型分析中列出它们的结果。所以如何在单变量分析情形下,基于 logrank 检验估计 RR 或 HR,是本讲的主要内容。

一、统计方法:以最常见二元协变量为例,即两组比较的情形。假定两组个体在时刻 t 时的危险率分别为 λ₁(t) 和 λ₂(t),根据 Cox 的比例危险假设^[1,2],存在下式

$$\lambda_1(t) = HR \times \lambda_2(t) \quad \text{或} \quad HR = \frac{\lambda_1(t)}{\lambda_2(t)} \quad (1)$$

式中 HR 即危险率比。它不依赖于时间,是个常数。回到上一讲(第三讲)中关于样本资料的一些假定,例如 2 × 2 表,这里不在重复,请参考。首先看无效假设,即 H₀: HR = 1。而假设检验统计量则是常用的 logrank 检验统计量,分别有 Mantel 和 Peto 氏统计量两种。即为上一讲中的公式(5)和(4)。为了叙述上的方便,这里重复这两个公式。Mantel 氏统计量^[1-7]和 Peto 氏统计量^[1-5, 8-11]分别为

$$\chi^2_M = \left[\sum_{k=1}^K d_{1k} - \frac{\sum_{k=1}^K E(d_{1k})}{\sum_{k=1}^K E(d_{1k})} \right]^2 \div \sum_{k=1}^K \text{Var}(d_{1k}) \quad (2)$$

$$\chi^2_P = \frac{\left[\sum_{k=1}^K d_{1k} - \frac{\sum_{k=1}^K E(d_{1k})}{\sum_{k=1}^K E(d_{1k})} \right]^2 \left[\sum_{k=1}^K d_{2k} - \frac{\sum_{k=1}^K E(d_{2k})}{\sum_{k=1}^K E(d_{2k})} \right]^2}{\sum_{k=1}^K E(d_{1k}) \sum_{k=1}^K E(d_{2k})} \quad (3)$$

其中 Peto 氏(简化式)统计量比 Mantel 氏统计量在假设检验上较保守^[1, 2, 10-12]。在(2)式的基础上,估计危险率比的公式^[8]为

$$HR_1 = \exp \left[\left(\sum_{k=1}^K d_{1k} - \frac{\sum_{k=1}^K E(d_{1k})}{\sum_{k=1}^K E(d_{1k})} \right) \div \sum_{k=1}^K \text{Var}(d_{1k}) \right] \quad (4)$$

式中分子为实际死亡数与期望死亡数的差,分母为方差。期望死亡数与方差的计算方法见上一讲。对数危险率比 log(HR) 的方差估计可以用下式

$$\text{Var}[\log(HR_1)] = 1 \div \sum_{k=1}^K \text{Var}(d_{1k}) \quad (5)$$

而与(3)式相对应的危险率比估计公式^[9, 14, 15]为

$$HR_2 = \frac{\sum_{k=1}^K d_{1k} \div \sum_{k=1}^K E(d_{1k})}{\sum_{k=1}^K d_{2k} \div \sum_{k=1}^K E(d_{2k})} \quad (6)$$

上式分母中的期望死亡数除可用上一讲中的方法计算外,还可用下式

$$\sum_{k=1}^K E(d_{2k}) = \sum_{k=1}^K d_{1k} + \sum_{k=1}^K d_{2k} - \sum_{k=1}^K E(d_{1k})$$

对数 HR₂ 的方差估计公式^[2]为

$$\text{Var}[\log(HR_2)] = \frac{1}{\sum_{k=1}^K E(d_{1k})} + \frac{1}{\sum_{k=1}^K E(d_{2k})} \quad (7)$$

或采用下式^[1]

$$\text{Var}[\log(HR_2)] = 1/W \quad (8)$$

$$\text{式中 } W = \sum_{k=1}^K \frac{l_{1k} \times l_{2k} \times D_k \times HR_2}{(l_{1k} \times HR_2 + l_{2k})^2}$$

估计 HR 方差的目的,是为了能估计 HR 的 95% 置信区间(CI),或做假设检验。

二、实例说明:以 Freireich 白血病临床试验数据^[1, 2, 16]为例,资料见表 1。研究因素是治疗方法,即对照组(安慰剂组)和 6-MP 治疗组,也即两组比较的情形。目的是比较两组病人的生存期(预后)有没有差别,或者为了说明采用 6-MP 药物治疗的病

人,其预后是否优于对照组。本文的重点在于估计研究因素的效应大小,即危险率比 HR。表 2 是分析该数据的假设检验统计量、危险率比及其对数的方差估计值等。为了与 Cox 回归模型^{1,2)}的分析进行对比,表中同时给出 Cox 回归模型的分析结果。以 logrank 检验为例,对照组相对于治疗组的危险率比为 5.15,且有高度统计学意义。表中的结果同时也说明了简化式统计量是较保守的。与 Cox 方法相比,基于 logrank 检验的参数估计值 HR 似有高估的倾向,而在简化式基础上推导的参数 HR 则低估。

表 1 Freieich 白血病临床试验病人随访资料 (对照组和 6- MP 治疗组)

对照组:	1, 1, 2, 2, 3, 4, 4, 5, 5, 8, 8, 8, 8, 11, 11, 12, 12, 15, 17, 22, 23
试验组:	6, 6, 6, 6+, 7, 9+, 10+, 10+, 11+, 13, 16, 17+, 19+, 20+, 22, 23, 25+, 32+, 32+, 34+, 35+

注:生存时间单位为周;“+”表示截尾生存时间(censoring times)。

表 2 假设检验统计量及危险率比估计值

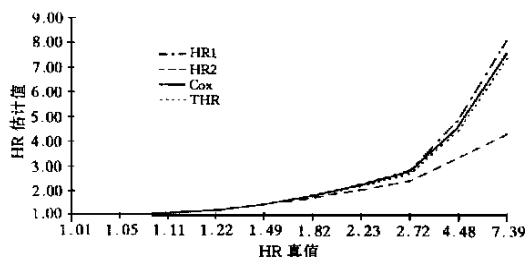
方法	统计量	P 值	HR	$\hat{var}[\log(HR)]$
Mantel 氏 logrank	16.79	0.000042	5.15	0.1598
Peto 氏(简化式)	15.23	0.000095	4.18	0.1450
Cox	15.93*	0.000070	4.52	-
	15.21#	0.000100		

* : Cox 比分析统计量; #: 似然比统计量。

表 3 计算机模拟随机样本的实验方法及条件

研究项目	实验条件	其它条件
样本大小(n)	50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000	(0.4, 50, 25%)
随访时间(α)	10, 20, 30, 40, 50, 60, 70, 80, 90, 100	(1000, 0.4, 25%)
截尾比例(%)	0, 10, 20, 30, 40, 50, 60, 70, 80, 90	(1000, 0.4, 50)
因素效应(β)	0.01, 0.05, 0.1, 0.2, 0.4, 0.6, 0.8, 1.0, 1.5, 2.0	(1000, 50, 25%)

注:生存时间假定为指数分布(均数为 α);样本含量为 n;因素的效应以回归系数 β 的大小表示。



附图 不同方法估计的 HR 偏性大小比较

比的偏性大小及方向。当 HR 取 10 个不同的值时,

为了对上述两种相对危险度的估计方法做出评价及说明它们的偏性大小,笔者利用计算机进行了模拟研究。生存时间以指数分布为例。具体做法见表 3。除考查样本大小的作用以外,其它情况均是产生一个含量为 1 000 的随机样本。其中因素效应的大小以参数 β 的大小来表示。需要说明的是括号内其它条件中的 0.4 为 β; 50 为指数分布(生存时间)的均数; 25% 为截尾数据所占比例大小。结果如下:

1. 样本大小: 随着样本含量增加,危险率比的估计越来越稳定,偏性也越小。而不同的估计值之间却相差极小,说明不同的方法不受样本含量的影响。

2. 截尾数据比例: 随着截尾数据所占的比例增加,危险率比的估计似出现越来越不稳定的趋势。但不同的估计方法似乎受其影响较小。

3. 因素效应: 随着研究因素的效应不断加强,两种估计值的偏性均越来越大,且 HR₁ 高估了真值,而 HR₂ 则低估了真值,似乎 HR₂ 的偏性大于 HR₁。从结果来看,当因素的作用不太强时(如 HR = 1~ 3),HR₁ 和 HR₂ 的偏性均很小。用 Cox 偏似然函数法估计的危险率比始终很稳定。

此外,对随访时间的长短的影响也进行了探讨,但对结果没有影响,且各种估计值均与真值非常接近。除样本大小以外,其它方面对样本方差估计的影响都很小。结果中仅观察到随样本含量的增加,对数 HR 的方差估计越来越小。

附图直观地显示了在用不同方法估计的危险率

即图中 X 轴。Y 轴则代表 HR 的估计值。图中共有四条线,分别为 HR₁ 估计值、HR₂ 估计值、Cox 法估计值和真值 THR。结果正如前述。

三、小结: 正如病例对照或队列研究资料的分析,人们并不仅仅限于计算假设检验统计量(例如 Mantel- Haenszel 卡方统计量和 P 值大小),同时要估计因素的作用大小,即估计计数比 OR 或相对危险度 RR。但在生存分析研究领域,这方面往往在实际应用中被忽略。国内更是如此,在生存资料单变量

分析时,主要强调了假设检验的问题(在结果中列出卡方统计量和 P 值),而没有给出相应的参数估计。

本讲介绍了生存资料分析中基于 logrank 检验的两种危险率比估计方法,同时与 Cox 的偏似然函数法估计进行了比较。文中给出了各种方法在分析一个数据中的实例应用。计算机模拟研究的结果表明,样本含量、截尾数据比例的大小对样本危险率比的估计有一定影响;而随访时间长短的影响很小;主要受研究因素效应大小的影响。在研究因素效应不

大时,他们均是很可靠的估计方法。而在因素效应较强时(如 $HR > 3$),均表现出了不同程度的偏性。其中,基于简化式的估计法的偏性更大,所以当 $HR > 3$ 时,危险率比的估计易采用更加可靠的方法,如 Cox 方法。

国外一些生物统计学家,例如 Peto^[8]、Pike^[9]、Mantel-Haenszel^[6,7,14,15]、Cox^[2,14]、Breslow^[12]等均对危险率比 HR 的估计方法进行过研究,均可参考。其中前三者的方法无需迭代,计算简便。

用彩超和心电图检测在校或幼儿园儿童心脏异常的发生率

杜忠东 李 竞 吴 进 邱 枫 李路萍 王增慧 郑 彤 刘 艳

笔者用心电图和彩色多普勒超声(彩超)对北京市儿童进行筛查,旨在了解我国城市儿童心脏的健康状况。

一、对象和方法:1996 年 9 月~1997 年 5 月,选北京市朝阳区中等条件的幼儿园两所,中小学各一所,共有 3~14 岁儿童 894 名。除去先天性心脏病术后儿童 2 例,进行性肌营养不良 1 例,先天性甲状腺功能低下 1 例,890 名无心脏病病史者组成本研究对象。男 460 名,女 430 名,平均年龄 7 岁。

全部小儿经询问病史、体检、心电图及完整彩超检查。对心律失常或可疑心律失常者进行心电图监测或记录 24 小时动态心电图。

二、结果:890 名小儿中,23 例有心脏异常,发生率 2.6%。其中男 9 例,女 14 例,性别间差异无显著性($\chi^2 = 2.401, P > 0.05$)。以心电图异常最常见,共 20 名,发生率 2.2%。其中室性早搏 4 例,1 例为频发,早搏 > 8 次/分;频发房性早搏 1 例;房室传导阻滞 2 例,1 例为 IV 度阻滞,另 1 例为 III 度阻滞;右束支传导阻滞 5 例,1 例为完全性阻滞;预激综合症 2 例;窦性心动过缓 7 例。5 例有早搏的小儿中,4 女,1 男,但性别间 Fisher 检验差异无显著性($P = 0.1404$)。

890 名小儿中,2 例发现先天性心脏病,1 例为 III 孔型房间隔缺损(5 岁女童),另 1 例为动脉导管未闭(6 岁女童)。2 例患儿均无临床症状,1 例身高低于该年龄组正常标准,另 1 例生长发育正常。房间隔

缺损,胸骨左缘第二肋间闻 2/6 收缩期杂音,第二心音轻度亢进。动脉导管未闭儿在胸骨左缘第二肋间有 2/6 级连续性杂音,无震颤。

全部受检儿中,仅 1 名 10 岁女童有二尖瓣脱垂,患儿心尖部有 2/6 级收缩期杂音。未发现风湿性心脏病、川崎病冠状动脉瘤或其它心脏异常。

三、讨论:先天性心脏病发病率为 8‰ 活产儿,出生后,多数畸形被诊治,部分轻度畸形如小室间隔缺损可自愈,还有少数患儿因表现不典型或无症状而被漏诊,及时发现这些患儿对先天性心脏病治疗水平的提高有重要意义。本研究首次对 890 名无心脏病史和症状的“正常”儿童进行彩超筛查,发现了 2 例心脏畸形,发生率 2‰。虽不能代表先心病的发病率,但提示在校或幼儿园儿童中,无症状的心脏畸形并非罕见。2 例患儿均有较典型的临床体征,仔细的体检可能会发现或怀疑心脏畸形,说明对在校或幼儿园儿童进行定期体检的必要性。

心电图异常发生于 2.2% 的在校或幼儿园儿童,以右束支传导阻滞、早搏较常见。其中 2 名小儿为频发的房性或室性早搏,2 名为房室传导阻滞,1 名完全型右束支传导阻滞,均为心肌炎诊断的指标。但此 5 名小儿临床上无症状,生长发育正常,彩超示心腔大小、心功能、心内血流正常。其心电图异常是否为病理性尚待进一步门诊。

本组小儿未发现有风湿性心脏病或风湿热等情况,与浙江的流调报告及 Lue 等的研究资料不符,可能与风湿热的地理分布不同或本组样本较小有关。

(收稿:1998-03-30 修回:1998-05-22)