

中医临床疗效评价的统计方法^{*}

□易丹辉^{**} 李芃芃 童小军

(中国人民大学统计学院 中国人民大学应用统计科学研究中心 北京 100872)

摘要:临床疗效评价的目的是寻找针对某种疾病、某类人群的有效治疗方案、方法、手段、药方。在临床治疗的实际过程中,通过试验设计,可以选定一些人群,采用随机分组控制一些影响因素,由于预期结果并不仅仅是研究因素,还有许多混杂因素,这些修饰因素作为解释变量或称自变量、协变量参与临床疗效研究,运用统计模型对疗效做出分析评价。Logistic 模型、Cox 比率危险率模型、纵向数据模型、多层线性模型都是可以用于临床疗效评价的统计方法。

关键词:临床疗效评价 Logistic 模型 Cox 比率危险率模型 纵向数据模型 多层线性模型

一、中医临床疗效评价的理解

临床疗效评价的目的是寻找针对某种疾病、某类人群的有效治疗方案、方法、手段、药方。其关键一是针对性,即通过评价寻找到适应人群;二是有效性,即通过评价确认实际治疗的效果。在临床治疗的实际中,通过试验设计,可以选定一些人群,采用随机分组可以控制一些影响因素,但预期结果实现的影响因素并不仅仅是研究因素,也就是所采用的治疗方案、方法、手段和药方等,往往还有混杂因素。由于这些因素的存在,无法采用在同一时间点进行统计检验的方法做出疗效评价,而需要将这些因素作为修饰因素,即作为解释变量或称自变量、协变量,运用一些统计模型对

疗效做出评价。

中医治疗的特点是辨证论治、方证结合、病证结合,其临床治疗效果的反映常常是患者的变化,有理化指标的变化,更有生存质量、症状的变化,因而反映治疗有效性的指标类型不单单是可测量的连续变量值,还有离散的变量值,又可分为分类的、有序的等。有效性指标类型不同,采用的方法不同。

二、Logistic 模型^[1]

Logistic 模型属于广义线性模型,主要包含两类:一是针对二维响应变量的 Logistic 模型,另一是针对多维响应变量的 Logistic 模型,又可以分为多分类无序和多分类有序两种形式。

1. 模型形式

(1) 二维响应变量的 Logistic 模型。

组稿日期:2007-07-28

^{*} 国家“十一五”科技计划项目(2006BAI08B05-09):“中医药诊疗与评价技术研究”之课题“中医药诊疗信息处理技术研究”子课题“中医临床疗效评价技术探索”子课题负责人:易丹辉。世界中医药联合会临床疗效评价专业委员会成立大会暨首届国际学术交流会议论文。

^{**} 联系人:易丹辉:中国人民大学统计学院教授、博士生导师,研究方向:预测与决策,风险管理与保险,Tel:010-82500172(O),Fax:010-62515246,Email:xueyi905@yahoo.com.cn。

[World Science and Technology/Modernization of Traditional Chinese Medicine and Materia Medica] 81

有效性指标为二分类的,如病情是否好转、症状是否改善等,可以采用该模型。实际数据建立的 Logistic 模型形式一般为:

$$\log\left(\frac{\pi_i}{1-\pi_i}\right) = \beta_0 + \beta_1 * x_{i1} + \cdots + \beta_p * x_{ip} \quad (1)$$

假定 y 服从取值为 0,1 的二项分布,即 $y \sim B(1, \pi)$ 。对不同观测中的 y_i ,有: $P(y_i = 1) = \pi_i, P(y_i = 0) = 1 - \pi_i$ 。所以 y_i 的期望和方差形式为:

$$E(y_i) = \mu_i = \pi_i, \text{var}(y_i) = \sigma^2(\mu_i) = \sigma^2(\pi_i) = \pi_i * (1 - \pi_i)。$$

η_i 和 μ_i 之间的联系函数及响应函数为:

$$\eta_i = \log(\mu_i / (1 - \mu_i)) \text{ 与 } \mu_i = \exp(\eta_i) / (1 + \exp(\eta_i))$$

联系函数或者响应函数一般根据研究的需要以及计算的简便来选取。定义 logit 函数为:

$$\text{logit}(x) = \log(x / (1 - x)),$$

(2) 多维无序响应变量的 Logistic 模型。

有效性指标为多分类的,如中医证候(虚、实、寒、热)的变化,可以考虑运用该模型。多维无序是指响应变量 Y 的多个响应水平间地位平等,无法按大小、高低、重要程度等排序。在二维响应变量的情形下,选用一个联系函数来刻画因变量的期望与自变量间的关系。在 k 维响应变量的情形下,一般需选用 $k-1$ 个联系函数来描述因变量期望与自变量间的关系。最常用的做法是选择一个参照水平,将其他水平与参照水平相比较来构造联系函数。这就是基准类 logits 模型(baseline - category logits)。以第 k 个水平为参照水平,其联系函数形式为:

$$\log\left(\frac{\pi_i}{\pi_k}\right) = \beta_{i0} + \beta_{i1} * x_i + \cdots + \beta_{ip} * x_p \quad (2)$$

$$i = 1, 2, \cdots, k-1。$$

因此,该模型含有 $k-1$ 个 logit 方程,每个方程含有不同的参数,这表明对不同的响应水平,自变量的影响效果是不同的。由(2)式,可以得到响应概率的表达式:

$$\pi_i = \frac{\exp(\beta_{i0} + \beta_{i1} * x_1 + \cdots + \beta_{ip} * x_p)}{\sum_h \exp(\beta_{h0} + \beta_{h1} * x_1 + \cdots + \beta_{hp} * x_p)} \quad (3)$$

$$i = 1, 2, \cdots, k-1$$

(3) 多维有序响应变量的 Logistic 模型。

有效性指标为多分类有等级的,如症状程度、量表得分等的变化,可以运用该模型。多维有序是指响应变量 Y 的多个水平间可以按照一定的标准排序。对具有有序响应水平的变量建模,联系函数应当反映这种有序性。不同的联系函数产生出不同的 Logistic 模型。最常用的 Logistic 模型为累积 logit 模型。

累积 logit 是用累积概率构造 logit 函数。对于响应变量 Y , 累积概率为:

$$P(Y \leq j) = \pi_1 + \pi_2 + \cdots + \pi_j \quad j = 1, 2, \cdots, k。$$

$P(Y \leq 1) \leq P(Y \leq 2) \leq \cdots \leq P(Y \leq k) = 1$, 所以累积概率能够反映响应变量的有序性。由于 $P(Y \leq k)$ 是肯定为 1 的, 所以利用累积概率来构造 logit 函数, 只用到前面 $k-1$ 个累积概率。累积 logit 模型形式为:

$$\text{logit}(P(Y \leq j)) = \log\left(\frac{P(Y \leq j)}{1 - P(Y \leq j)}\right) = \beta_{j0} + \beta_{j1} * x_1 + \cdots + \beta_{jp} * x_p \quad j = 1, 2, \cdots, k \quad (4)$$

2. 模型的参数估计及检验

(1) 参数估计。

Logistic 模型的参数通常采用最大似然估计得到。由 y 的特定的指数族分布形式(二项分布)得到似然函数,通过最大化似然函数来得到参数 $\beta_m (m = 1, 2, \cdots, p)$ 的估计。这里并不具体叙述估计的过程。有兴趣者,可以参考有关广义线性模型或 Logistic 模型的文献。

(2) 检验。

模型的检验主要有两类:一是模型显著性检验;二是模型的拟合优度检验。

① 模型显著性检验。

模型显著性检验可以从两个角度进行。一是检验参数整体是否显著不为 0, 即响应变量和自变量之间的线性关系在整体上是否成立;二是检验单个参数是否显著不为 0, 即该参数对应的自变量与响应变量之间的线性关系是否显著。一个好的模型应该是整体显

著的,其包含的每个参数也应是显著的。

模型整体显著性检验通过对所有系数是否同时为零进行检验得到结论;每个参数的显著性检验通过对每一个参数是否为零进行检验得到结论。

②模型拟合优度检验。

模型的拟合优度检验用来判断模型的估计值与原始数据的拟合情况,是判定模型适用性的重要标准。有时,模型通过了显著性检验,但无法通过拟合优度检验,说明其对原始数据的拟合程度差,模型仍然是不可取的。

模型的拟合优度检验通常通过构造服从 χ^2 分布的 pearson 统计量和 Deviance 统计量进行。

三、Cox 比率危险率模型^[2]

Cox 比率危险率模型(Cox proportional hazards regression model)可以用于有效性指标为生存时间、改善(好转)时间、延缓时间、缓解(复发)时间等的疗效评价。它是一种检验各种因素(协变量)以及他们的交叉项是否对因变量(有效性指标)有显著影响的方法。

四、纵向数据模型^{[3][4]}

纵向数据是指对任意个体的同一特性在一段时间或某几个时间点作重复性观测得到的数据。重复观测的有效性指标可以是连续性测量值,如各种客观检测的理化指标值、利用量表测量得到的指标值等;可以是离散性测量值,如将糖尿病患者测量的空腹血糖值按大小分为重症和非重症等。纵向数据研究的主要目的是描述响应变量随时间的变化规律,以及这些变化是如何相关于感兴趣的协变量。在数据分析上,纵向数据有两点区别于其它数据的主要特征:首先,由于在不同时间点上对同一个体进行了多次测量,因此研究者可以直接的研究该个体随时间的进展情况。但要注意同一个体的不同时点测量值通常是高度相关的,这些相关性应该纳入分析之中。此外,对于同一个体而言,在时间上相近的观测之间会比相距较远的观测表现出更紧密的联系,因此纵向数据的时间先后顺序对于分析推断有着重要的意义。纵向数据分析可以研究两种治疗方法下疗效随时间变化的模式是否有显著差

别,从而动态地判断出治疗效果。

纵向数据分析常用的模型有:边际模型、转移模型、线性混合模型。

边际模型在纵向数据,特别是多水平离散纵向数据的分析方法中占有重要的地位。所谓边际模型,指的是直接考察响应变量边际分布与协变量关系的模型。“边际”两字暗示了子集的边际期望值仅依赖于协变量,而非之前时点的结果或其他随机因素。边际模型是在广义线性模型的基础之上构建起来的,它直接沿用了广义线性模型中通过连接函数将因变量与回归方程相连的形式;边际模型也是广义线性方程的推广,体现为在模型中加入了同一个体观测值之间的相关性假设,从而解决了广义线性模型无法将纵向数据重复测量值关系有效表达的问题。边际模型主要研究在给定协变量 X 后,响应变量 Y 随时间的变化情况,其侧重于对总体平均变化水平的研究。

转移模型即条件模型,是一个条件回归模型。每个个体第 i 期的响应变量不仅取决于其自身的协变量,而且与之前的第 i-1 期、第 i-2 期……的响应变量有关,主要探讨序列间的自相关关系。转移模型中讨论的纵向数据重复观测之间的时间间隔是等距的。

线性混合效应模型认为,不同个体之间的协变量对响应变量的影响是有差异的,将其分为固定效应和随机效应两部分,固定效应是所有个体所共有的,而每个个体都有自己特定的随机效应。同一个体的数据之间的强相关关系的本质就是其自身的随机效应。随机效应模型可以更好的体现纵向数据同一个体数据间的相关关系和不同个体之间的差异,能够对每个个体随时间的变化情况进行更准确描述和预测,而且便于不同个体之间的比较。

五、多层线性模型^[5]

有效性指标具有分层或嵌套结构,如不同类型患者,如不同医院患者,根据影响临床研究方案有效性的多种因素将模型分为不同层次,在模型中考虑多因素对方案有效性的干扰。多层线性模型可以包含固定效应和随机效应两部分。固定效应衡量患者因其自身特

征不同所引起的治疗前后有效性指标的差异,主要由两部分组成,一是个体特征差异,例如患者的病程、体重、性别等;二是治疗方法的差异,例如患者可能会采用中西医结合疗法或纯中医法进行治疗;随机效应则衡量个体特征因其所属的组别(层)的不同而对响应变量产生的影响。

模型的形式通常有零模型、完整模型、随机效应模型。

(1) 零模型 (The Null Model)。

零模型没有任何解释变量。主要目的在于把总体变异分解成由个体差异造成的部分和层差异造成的部分。

$$\text{第一层: } Y_{ij} = \beta_{0i} + \varepsilon_{ij} \quad \varepsilon \sim N(0, \sigma^2)$$

$$\text{第二层: } \beta_{0i} = \gamma_{00} + b_{0i} \quad b \sim N(0, \tau_{00}^2) \quad \varepsilon, b \text{ 相互独立}$$

$$\text{综合: } Y_{ij} = \gamma_{00} + b_{0i} + \varepsilon_{ij}$$

$$\text{方差: } \text{var}(Y_{ij}) = \text{var}(b_{0i} + \varepsilon_{ij}) = \text{var}(b_{0i}) + \text{var}(\varepsilon_{ij}) = \tau_{00}^2 + \sigma^2 \quad (5)$$

其中, Y_{ij} 的下标 j 代表第一层的单位;下标 i 代表第一层的个体所隶属的第二层的单位; γ_{00} 是 β_{0i} 的均值,是其固定部分; b_{0i} 是 β_{0i} 的随机部分,代表第二层单位之间的差异。

(2) 完整模型 (The Full Model)。

完整模型是在每一层都包含协变量的模型。这样可以解释 Y 的总体变异是怎样受到第一层(式(6)中的第一个模型)和第二层(式(6)中的第二个模型)因素的影响的,综合模型(式(6)中的第三个模型)包含了交互效应。

$$\text{第一层 } Y_{ij} = \beta_{0i} + \beta_{1i} X_{ij} + \varepsilon_{ij} \quad \varepsilon \sim N(0, \sigma^2)$$

$$\text{第二层 } \beta_{0i} = \gamma_{00} + \gamma_{01} W_{ij} + b_{0i}$$

$$\begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \sim N \begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} \tau_{00}^2 & \tau_{01}^2 \\ \tau_{01}^2 & \tau_{11}^2 \end{pmatrix} \quad \varepsilon, b \text{ 相互独立}$$

$$\beta_{1i} = \gamma_{10} + \gamma_{11} W_{ij} + b_{1i}$$

$$\text{综合 } Y_{ij} = \gamma_{00} + \gamma_{01} W_{ij} + \gamma_{10} X_{ij} + \gamma_{11} W_{ij} X_{ij} + b_{0i} + X_{ij} b_{1i} + \varepsilon_{ij}$$

$$\begin{aligned} \text{方差 } \text{var}(Y_{ij}) &= \text{var}(b_{0i} + X_{ij} b_{1i} + \varepsilon_{ij}) \\ &= \text{var}(b_{0i}) + 2X_{ij} \text{cov}(b_{0i}, b_{1i}) + X_{ij}^2 \text{var}(b_{1i}) \\ &+ \text{var}(\varepsilon_{ij}) \\ &= \tau_{00}^2 + 2X_{ij} \tau_{01} + X_{ij}^2 \tau_{11}^2 + \sigma^2 \end{aligned} \quad (6)$$

(3) 随机效应模型 (Random Effect Model)。

随机效应模型在第二层模型中没有协变量。其主要着眼点为第二层的变异。在找不到合适的第二层协变量时,可以采用这种模型。

$$\text{第一层 } Y_{ij} = \beta_{0i} + \beta_{1i} X_{ij} + \varepsilon_{ij} \quad \varepsilon \sim N(0, \sigma^2)$$

$$\text{第二层 } \beta_{0i} = \gamma_{00} + b_{0i} \quad \begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \sim N \begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} \tau_{00}^2 & \tau_{01}^2 \\ \tau_{01}^2 & \tau_{11}^2 \end{pmatrix}$$

ε, b 相互独立

$$\beta_{1i} = \gamma_{10} + b_{1i}$$

$$\text{综合 } Y_{ij} = \gamma_{00} + \gamma_{10} X_{ij} + b_{0i} + b_{1i} X_{ij} + \varepsilon_{ij}$$

$$\begin{aligned} \text{方差 } \text{var}(Y_{ij}) &= \text{var}(b_{0i} + X_{ij} b_{1i} + \varepsilon_{ij}) \\ &= \text{var}(b_{0i}) + 2X_{ij} \text{cov}(b_{0i}, b_{1i}) + X_{ij}^2 \text{var}(b_{1i}) \\ &+ \text{var}(\varepsilon_{ij}) \\ &= \tau_{00}^2 + 2X_{ij} \tau_{01} + X_{ij}^2 \tau_{11}^2 + \sigma^2 \end{aligned} \quad (7)$$

实际应用中,可以根据研究的需要以及数据的情况选择合适的模型。

临床疗效评价通常都强调随机性,主要是为避免造成试验组与对照组的人群偏移,保证结果分析的一般性即科学性。但即使随机,也很难避免一些非研究因素的影响,运用统计模型,将这些因素作为协变量同时纳入分析,可以更明确对有效性指标影响的主要因素,使分析评价的结果更科学和可靠。

参考文献

- 1 易丹辉:北京市居民医疗消费行为及意愿研究,中国人民大学出版社,2004年11月,85~94.
- 2 陈家鼎:生存分析与可靠性,北京大学出版社,2005年11月
- 3 Geert Verbeke, Geert Molenberghs: Linear Mixed Models for Longitudinal Data, Springer, New York, 2000.
- 4 Geert Molenberghs, Geert Verbeke: Models for Discrete Longitudinal Data, Springer, 2005.
- 5 张雷, 雷雳, 郭伯良: 多层线性模型应用, 教育科学出版社, 2005年6月第2版, 18~23.

Statistical Methods for Evaluating Clinical Therapeutic Effects

Yi Danhui, Li Pengpeng, Tong Xiaojun

(The School of Statistics, RUC, Beijing 100872)

The purpose of evaluating clinical therapeutic effects is to confirm the effectiveness of a therapeutic plan, method, and recipe in treating diseases, in the context of selected patients. In the course of clinical treatment, we may choose some patients groups, and control some influence factors using random grouping. However, the realization of anticipated results has to be determined by a range of mixed factors. As a result, it is difficult to make an evaluation merely using a statistical method with an identical time. Fortunately, we can treat them as decorative factors, i. e., explanation variables, independent variables, or covariates, before making the evaluation using statistical models. As index types that reflect the efficacy of clinical treatment are different, statistical models have to be different too. Logistic model, Cox proportional hazards regression model, longitudinal models, and multilevel linear model are useful statistical approaches for evaluating clinical therapeutic effects.

Keywords: clinical therapeutic effects evaluation; logistic model; Cox proportional hazards regression model; longitudinal models; multilevel linear model

(责任编辑:王 瑀, 责任译审:邹春申)

(Continued from Page 73)

- 31 Jakupovic J, Jeske T, Morgenstern T, *et al.* Diterpenes from *Euphorbia segetalis*. *Phytochemistry*, 1998, 47(8): 1583 ~ 1600.
- 32 Gerd Vogg, Elke Mattes, Jan Rothenburger, *et al.* Tumor promoting diterpenes from *Euphorbia leuconeura* L. *Phytochemistry*, 1999, 51(12): 289 ~ 295.
- 33 Ilona Mucsi, Joseph Molnar, Judit Hohmann, *et al.* Cytotoxicities and

Anti - Herpes Simplex Virus Activities of Diterpenes Isolated from *Euphorbia species*. *Planta Med*, 2001, 67(7): 672 ~ 674.

- 34 M. Gundidza, B. Sorg, E. Hecker. A skin irritant principle from *Euphorbia matabelensis* Pax. *J Ethnopharmacology*, 1993, 39(11): 209 ~ 212.
- 35 Amir Reza Jassbi. Chemistry and biological activity of secondary metabolites in *Euphorbia* from Iram. *Phytochemistry*, 2006, 67(12): 1977 ~ 1984.

Progresses in Studying Diterpene Esters Structures and Associated Toxicity

Su Shulan

(Jiangsu Key Laboratory for TCM Formulae Research, Nanjing University of Chinese Medicine, Nanjing 210046, Jiangsu University School of Pharmacy, Zhenjiang 212013)

Duan Jinao, Ding Anwei

(Jiangsu Key Laboratory for TCM Formulae Research, Nanjing University of Chinese Medicine, Nanjing 210046)

Authors analyzed the relationships between the structures of Diterpene esters and associated toxicities based on past study results. The distribution of Diterpene esters elements in *Euphorbia* plants was reviewed. Authors also discussed theoretical aspects of toxicity constituents, in an attempt to provide evidences for clinical safety.

Keywords: Diterpene Esters; toxicity; progress

(责任编辑:张述庆, 责任译审:邹春申)

[World Science and Technology/Modernization of Traditional Chinese Medicine and Materia Medica] 85