

· 基础理论与方法 ·

指数模型拟合法在乙型肝炎疫苗接种后
抗体持久性预测中的应用

罗凤基 董春明 沈永刚 黄长江

【摘要】 目的 探讨乙型肝炎疫苗接种后抗体水平的指数模型拟合法。方法 利用国产重组(酵母)乙型肝炎疫苗(YDV)接种后 5 年的观察数据拟合指数模型,用第 8 年的观察数据对模型的预测值进行验证。结果 根据国产 YDV 接种 5 年后的数据所拟合的指数模型为: $\hat{Y} = 165.67 \exp(-0.019X)$; 决定系数 R^2 为 0.98; 到第 8 年时,实际测定值为 35 mIU/ml,模型预测的结果为 27 mIU/ml,比实际值低 8 mIU/ml。接种 12 年时,其预测的平均抗体水平为 10.74 mIU/ml,仍维持在 10 mIU/ml 的保护性水平以上。结论 在疫苗接种后,对符合指数线性趋势的抗体数据可拟合指数模型,利用该模型在一定的时间范围内可预测抗体水平持久性。实例拟合的结果表明,由于在第 8 年时,模型预测值低于实际测定值,因此,模型预测的结果具有可靠性。

【关键词】 乙型肝炎疫苗;指数模型;抗体持久性预测

Application of exponent curve model to study the hepatitis B DNA recombinant yeast derived vaccine antibody levels LUO Feng-ji*, DONG Chun-ming, SHEN Yong-gang, HUANG Chang-jiang.
*Chaoyang Center for Disease Control and Prevention, Beijing 100021, China

【Abstract】 Objective To introduce exponent curve model methods in the study of the hepatitis B vaccine antibody level. **Methods** After the China made vaccine of hepatitis B DNA recombinant yeast derived vaccine (YDV) had been carried out for 5 years, data on the anti-HBsAg's titer were used to construct an exponent curve model. When the vaccination program had been carried out for 8 years, the predicating results of the model were further tested by observed number. **Results** The exponent curve model was $\hat{Y} = 165.67 \exp(-0.019X)$ and the R^2 was 0.98. After 8 years, the practical observed number became 35 mIU/ml, and the predicating result of the model was 27 mIU/ml, 8 mIU/ml lower than the observed number. When the vaccine had been carried out for 12 years, the predicating results of the model became 10.74 mIU/ml, still higher than 10 mIU/ml but was still in the effective range. **Conclusion** An exponent curve model could be constructed, as long as the data of the antibody's titer was in accordance with the tendency of exponent curve. The model could be used to predict the persistence lever of vaccine antibody under certain conditions. The results showed that after 8 years, the predicting results of the model were reliably lower than the observed number.

【Key words】 Hepatitis B vaccine; Exponent curve model; Persistence lever of vaccine antibody predicate

在疫苗的免疫持久性研究中,常需用数年随访的实测数据拟合出一个最适模型,以概括全部实测数据的信息,并预测在随访时间以外的抗体水平及其变化趋势。本文介绍的是如何根据实测数据用最小二乘法拟合出最适的疫苗免疫持久性的指数模型。

基本原理

1. 确定模型的曲线类型^[1,2]:用实测数据在普通坐标纸上绘制散点图以观察实测曲线是否符合指数曲线的类型。在疫苗免疫持久性研究中,常用的曲线是指数曲线。指数曲线方程的一般形式为 $\hat{Y} = k + a \times \exp(bX)$,自变量 X 按等差级数变化时, Y 按等比级数变化。

指数曲线的基本示意图形见图 1。在图 1 中,曲线的位置、方向及曲率的大小分别由其方程中的常数项 k 、参数 a 和 b 的正负和大小来确定。实际应

作者单位:100021 北京市朝阳区疾病预防控制中心(罗凤基);
卫生部北京生物制品研究所流行病学研究室(董春明);安徽省疾病
预防控制中心(沈永刚);安徽省巢湖市疾病预防控制中心(黄长江)

用时可根据实测数据所绘制的图形确定所要拟合的曲线方程。

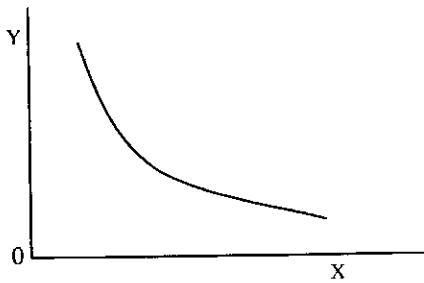


图1 指数曲线 $\hat{Y} = k + a \times \exp(-bX)$

2. 曲线直线化 将实测数据在半对数坐标纸(指数曲线)上再次绘制散点图。若散点已呈直线趋势,可直接对应变量 Y 取对数后求其方程;若散点不呈直线趋势时,可对 X 或 Y 加减不同的 k 值后反复尝试,直到散点呈直线趋势为止,所得 k 值即为所拟曲线方程的常数项 k 。 k 值也可以通过公式 $k = (Y_1 \times Y_2 - Y_3^2) / (Y_1 + Y_2 - 2Y_3)$ 来估计,式中 Y_1 、 Y_2 为散点图上所取的两个相距较远而距线性趋势较近的观察点 X_1 和 X_2 所对应的 Y 值。 Y_3 为 $X_3 = (X_1 + X_2) / 2$ 时所对应的 Y 值。

3. 求方程:

(1)对自变量 X 及应变量 Y 或 $(Y - k)$ 分别或单独取自然对数。

(2)求自变量 X 的离均差平方和 L_{xx} 、应变量 Y 的离均差平方和 L_{yy} 及两变量的离均差积和 L_{xy} 。 L_{xx} 和 L_{yy} 也可用计算器直接求出相应的方差后再乘以相应的样本例数减 1 来求得,即 $L_{xx} = S_x^2(n_x - 1)$ 。式中 S_x^2 、 n_x 分别为自变量 X 的方差和样本例数, L_{yy} 的求法亦然。

(3)在求得上述指标后,即可带入式 $b = L_{xy} / L_{xx}$ 算得参数 b ,再将 b 带入式 $Y = a + bX$ 即可求出参

数 a 的值。

(4)将 k 、 a 、 b 分别带入所求曲线方程的自然对数形式:

$$\text{Ln}(\hat{Y} - k) = \text{Ln} a + bX$$

(5)将上式还原为指数曲线方程 $\hat{Y} = k + a \times \exp(bX)$

(6)拟合优度分析:计算确定系数 R^2 , R^2 的取值范围在 0~1 之间; R^2 越接近 1 时,表明曲线的拟合效果越好^[3]。

$$R^2 = 1 - \frac{\sum(Y_i - \hat{Y}_i)^2}{\sum(Y_i - \bar{Y}_i)^2}$$

实例分析

现以国产重组(酵母)乙型肝炎疫苗(YDV)接种后 5 年(1994~1999 年)的观察数据拟合其指数模型,并用第 8 年的实测数据与模型的预测值进行比较。表 1 为 YDV 接种后 5 年乙肝表面抗体的平均滴度和指数曲线的拟合过程。

拟合过程:

1. 曲线直线化:按实测数据在普通坐标纸上绘制散点图后,曲线基本符合指数曲线的类型,将实测数据在半对数坐标纸(指数曲线)上再次绘制散点图,散点已呈直线趋势,因此可直接对应变量 Y 取对数后求其方程。

2. 对应变量 Y 单独取自然对数,见表 1 第 4 栏。

3. 求自变量 X 的离均差平方和 L_{xx} 、应变量 Y 的离均差平方和 L_{yy} 及两变量的离均差积和 L_{xy} ,见表 1 第 5、6、7 栏合计。

$$\bar{X} = \sum X_i / n_{Xi} = 181 / 6 = 30.17, \bar{Y} = \sum Y_i / n_{yi} = 604 / 6 = 100.67$$

$$\text{Ln} \bar{Y} = \sum \text{Ln} Y_i / n_{yi} = 27.22 / 6 = 4.54$$

$$L_{xx} = \sum X_i^2 - (\sum X_i)^2 / n_{Xi} = 7921 - (181)^2 / 6 = 2460.83$$

同理得 $L_{yy} = 0.92$

表1 接种国产 YDV 后 5 年的平均抗体滴度及曲线拟合的计算过程

观察时间(月)	观察人数	平均抗体滴度*	$\text{Ln} Y_i$	$X(\text{Ln} Y_i)$	X_i^2	$(\text{Ln} Y_i)^2$	模型测算值	残方差
X_i	N_i	Y_i					\hat{Y}_i	$(\hat{Y}_i - Y_i)^2$
(1)	(2)	(3)	(4)=Ln(3)	(5)=(1)×(4)	(6)=(1) ²	(7)=(4) ²	(8)	(9)
1	34	160	5.08	5.08	1	25.81	162.0	4.00
12	34	143	4.96	59.52	144	24.60	132.0	121.00
24	34	98	4.59	110.16	576	21.07	105.0	49.00
36	34	83	4.41	158.76	1296	19.45	83.6	0.36
48	34	67	4.21	202.08	2304	17.72	66.6	0.16
60	34	53	3.97	238.20	3600	15.76	53.0	0.00
合计 181	-	604	27.22	773.80	7921	124.41	-	174.52

* mIU/ml

$$L_{xy} = \sum X_i (\ln Y_i) - [\sum X_i \sum (\ln Y_i)] / n = 773.8 - (181 \times 27.22) / 6 = -47.34$$

4. 在求得上述指标后,即可带入式 $b = L_{xy} / L_{xx}$ 算得参数 $b = -47.34 / 2460.83 = -0.019$;再将 b 带入式 $\ln \hat{Y} = a + bX$ 即可求出参数 a 值为 5.11。

5. 将 a 、 b 分别带入所求指数曲线方程的自然对数形式:

$$\ln \hat{Y} = \ln a + bX \quad \ln a = \ln 5.11 = 165.67$$

6. 将上式还原为指数曲线方程 $\hat{Y} = a \times \exp(bX)$ 则得指数方程为:

$$\hat{Y} = 165.67 \exp(-0.019X)$$

由该方程求得表 1 中第 8 栏的模型预测值,与第 3 栏的实测值相比十分接近。

7. 拟合优度分析:计算确定系数 R^2 , R^2 越接近 1,则曲线拟合的效果越好, R^2 的计算公式如下:

$$R^2 = 1 - \sum (Y_i - \hat{Y}_i)^2 / L_{yy}$$

本例中: $\sum Y_i^2 = 69\ 840$; $\sum (Y_i - \bar{Y}_i)^2 = 174.52$, 见表 1 第 9 栏。

$$L_{yy} = \sum Y_i^2 - (\sum Y_i)^2 / n_i = 69\ 840 - (604)^2 / 6 = 9037.33$$

则决定系数 $R^2 = 1 - 174.52 / 9037.33 = 0.98$ 。通过残差分析计算,决定系数 R^2 为 0.98,非常接近于 1,表明方程拟合非常理想。

8. 预测结果的可靠性分析:为检验模型预测值的可靠性,2002 年也就是在上述研究完成后的第 3 年,我们又按照原试验的设计要求,对前 5 年的 34 名试验对象再次进行了血清学随访。在 34 名被研究者中,失访 8 人,其余 26 人的乙肝表面抗体血清学的检测结果表明,平均抗体滴度为 35 mIU/ml,而模型预测的结果为 27 mIU/ml,仅比实际值低 8 mIU/ml。表明模型预测的结果在疫苗接种 8 年

后仍较准确。

讨 论

1. 本文介绍的指数曲线的拟合方法是最小二乘法,其基本条件为实测数据经对数、倒数或平方根转化后使曲线方程直线化。直线化的效果越好则模型拟合的效果也越理想。

2. 对于难于直线化的曲线模型,其拟合的方法比较复杂,可根据曲线的类型选取最适的模型,采用牛顿迭代法在计算机上完成。

3. 在对加强免疫后的疫苗持久性进行模型拟合时,其拟合用的数据应从加强免疫后开始。

4. 所拟合的模型曲线是 X 轴的一条渐近线,当 X 的取值为无穷大时,所对应的 Y 值趋于 0,但永远不等于 0。因此,应用时可根据实际情况确定所需的阶段范围。

5. 所拟合的模型可用于观察疫苗持久性的发展趋势和预测随访时间以外抗体水平。但其结果仅可作为预测值或理论值,在实际应用时还需参考其他资料。

6. 任何预测方法都有局限性,一般说来,预测的时间越长,结果的可靠性越差,这是很正常的。因此,在判定预测结果的可靠性时,除了要考虑模型的误差外,还应考虑预测时间的长短。

参 考 文 献

- 1 倪宗瓿,主编.医学统计学.北京:人民卫生出版社,1995.112-116.
- 2 杨树勤,主编.中国医学百科全书医学统计学.上海:上海科技出版社,1985.163-165.
- 3 杨树勤,主编.卫生统计学.北京:人民卫生出版社,1990.115.
(收稿日期:2003-12-04)
(本文编辑:张林东)

· 消息 ·

中华微生物学和免疫学杂志英文版创刊

《Journal of Microbiology and Immunology》(中华微生物学和免疫学杂志英文版),于 2003 年 11 月创刊,为中国科学技术协会主管,中华医学会主办的中华医学系列杂志之一。中国标准刊号:CN 11-4986/R;国际刊号:ISSN 1672-4666。本刊主要报道国内外医学微生物学和免疫学的基础理论研究以及临床与实用技术研究论文、简报、述评、综述、新书介绍、会议信息等。欢迎国内外科研工作者踊跃投稿、订阅本刊。

投稿要求:除同中华微生物学和免疫学杂志中文版外,每一篇文章需附中文摘要及中文关键词。欢迎同时附中文稿。本刊为季刊,大 16 开,每期 64 页,每季中出版。国内外公开发行。

编辑部地址:北京市朝阳区三间房南里 4 号 中华微生物学和免疫学杂志英文版编辑部 电话:010-65756595;
Email: cjmic@163.com