· 系列讲座 ·

# 临床流行病学

## 第五讲　时间效应分析（一）

章扬熙

时间效应分析，又称生存分析，在临床流行病学中有其重要的地位。比如甲药治疗某病，治愈率为80%，乙药治疗某病，治愈率也是80%，能不能说两种药物的疗效是一样呢？不能，因为治疗效果是时间效应，如果甲药平均3天治愈80%的某病病人，而乙药平均7天才治愈80%的某病病人，则甲药比乙药好。治疗的效果随时间而异，事实上，每一种药治疗某病都有一条时间效应曲线，科学地对某一药物（或疗法）的疗效进行评价，则应是对这条时间效应曲线进行全面评价，即进行时间效应分析，而不仅是对某个时点的治愈率（或生存率）进行评价。

时间效应分析的方法，首先用于生存分析，所以也称为生存分析的方法。一个人从出生到死亡的时间叫作生存时间，即寿命。生存时间的长短，与许多因素相联系，研究诸因素与生存时间的联系及其程度，称为生存分析。在生存分析随访研究过程中，一部分研究对象可观察到死亡，可以得到准确的生存时间，所提供的信息是完全的。但往往有一部分人，或中途失访，或到观察结束时仍存活，对这些人无法知道准确的生存时间，只知道其生存时间比观察到的时间要长，这一部分数据叫作截尾数据（censored data），它提供不完全信息。生存分析方法可以同时分析有结局的生存时间数据和没有结局的截尾数据，能较充分地利用资料信息。

如果改变出生和死亡的含义，可使生存分析方法得到广泛的应用，即时间效应分析。比如，以接种某病疫苗代替出生，以某病发病代替死亡，用时间效应分析研究接种某病疫苗后的保护时间。又比如，以暴露于某危险因子代替出生，以某病发病代替死亡，用此分析技术研究暴露于某危险因子后多长时间发病的概率。再比如，以应用某疗法治疗某病人代替出生，以治愈代替死亡，可用此分析技术研究某病应用该疗法后的不同时间的治愈情况，全面了解该疗法的时间效应过程，如此等等。

常用的时间效应分析方法，分单因素与多因素两类。单因素时间效应分析有 Kaplan Meier 法和寿命表法，多因素时间效应分析有 Cox 模型法。

**一、Kaplan Meier 法**：这种方法又称乘积限估计法（product-limit estimates），它适用于小样本，观察时间单位越小，精度越高，是一种非参数估计法。计算方法用实例来说明。

例1　对15例新发多菌型（MB）麻风病例应用多药联合化疗治疗，随访观察其治愈的月数（Ti）和截尾数据（Ci 用其数据后带“+”号表示），结果如下：

（Ti）2，3，9，10，10，15，16，30

（Ci）$12^+$，$15^+$，$18^+$，$21^+$，$36^+$，$40^+$，$45^+$

试估计治愈率及其标准误，平均治愈病期及其标准差（表1）。

具体步骤如下：

1. 将非截尾的生存数据 $T_i$ 按时间从短到长排列，其秩次 $i=1, 2\cdots, k$，列于第①栏；其观察时间 $t_i$ 列于第②栏；期初病例数列于第③栏；$n_1=k$，其余 $n_i=n_{i-1}-h_{i-1}-w_{i-1}$，治愈人数 $h_i$ 列于第④栏；失访或末次随访仍未愈人数 $w_i$ 列于第⑤栏；治愈概率 $p_i=h_i/n_i$，列于第⑥栏，未治愈概率 $q_i=1-p_i$，列于第⑦栏；$t_i$ 时未治愈率 $t_iq_o=\prod_{i=1} q_i$，即第⑦栏的数值自上而下连乘，列于第⑧栏；$t_i$ 时治愈率 $t_ip_o=1-t_iq_o$，列于第⑨栏。治愈率的标准误 $S_{ti}p_o$ 用下式计算，结果列于第⑩栏。

$$S_{ti}p_o = t_iq_o\sqrt{\sum_{i=1}\frac{h_i}{n_i(n_i-h_i)}} \qquad (1)$$

比如，$i=2$ 时，$t_2=3$

作者单位：中国医学科学院皮肤病防治研究所　江苏南京　210042

表1 MDT治疗MB型麻风病例的治愈率分析

| 秩次 $i$ ① | 观察时间（月）$t_i$ ② | 期初病例数 $n_i$ ③ | 治愈人数 $h_i$ ④ | 失访或终访未愈人数 $w_i$ ⑤ | 治愈概率 $p_i$ ⑥ | 未治愈概率 $q_i$ ⑦ | 从开始到 $t_i$ 时未治愈率 $_{t_i}q_o$ ⑧ | 从开始到 $t_i$ 时治愈率 $_{t_i}p_o$ ⑨ | 治愈率标准误 $s_{t_i}p_o$ ⑩ |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 2 | 15 | 1 | 0 | 0.067 | 0.933 | 0.933 | 0.067 | 0.064 |
| 2 | 3 | 14 | 1 | 0 | 0.071 | 0.929 | 0.866 | 0.134 | 0.088 |
| 3 | 9 | 13 | 1 | 0 | 0.077 | 0.923 | 0.800 | 0.200 | 0.103 |
| 4 | 10 | 12 | 2 | 1 | 0.167 | 0.833 | 0.666 | 0.334 | 0.121 |
| 5 | 15 | 9 | 1 | 1 | 0.111 | 0.889 | 0.592 | 0.408 | 0.128 |
| 6 | 16 | 7 | 1 | 2 | 0.143 | 0.857 | 0.507 | 0.493 | 0.135 |
| 7 | 30 | 4 | 1 | 3 | 0.250 | 0.750 | 0.380 | 0.620 | 0.149 |

$$S_2p_0=0.134\sqrt{\frac{1}{15\ (15-1)}+\frac{1}{14\ (14-1)}}=$$

0.088 余类推。

应当说明，生存分析研究的起点、结局、生存时间在不同研究中的情况不同。本例的结局是好的结局，如果是以复发或畸残或死亡为结局时，则为坏的结局，计算复发率或畸残率或生存率的标准误仍用公式（1）计算。

计算平均治疗病期及其方差的步骤见表2。表2的第①、②、③、⑤、⑥栏来自表1；由于本例的最大值t是截尾数据，故把此值 $t_n$ 在第②栏末行列出，若最大值非截尾数据，则不列；第④栏是由第②栏 $x_{i+1}-x_i$ 得到，比如 $i=4$ 时 $t_5-t_4=15-10=5$；第⑦栏是第③、④栏的乘积，比如 $i=4$ 时 $_{10}q_o$. $(t_5-t_4)=0.666\times5=3.330$；第⑧栏是第⑦栏由下而上累加求得，比如 $i=4$ 时，$A_4=7.098+0.590+3.330=11.018$；第⑨栏由第⑤、⑥、⑧栏计算，比如 $i=4$ 时，$h_4A_4^2/[n_4(n_4-1)]=2\times11.018^2/(12\times11)=2.023$。余类推。

表2 平均治疗病期及其方差的计算

| 秩次 $i$ ① | 观察时间（月）$t_i$ ② | t时未治愈率 $_{t_i}q_o$ ③ | $t_{i+1}-t_i$ ④ | 治愈人数 $h_i$ ⑤ | 期初病例数 $n_i$ ⑥ | $_{t_i}q_o\ (t_{i+1}-t_i)$ ⑦ | $A_i$ ⑧ | $h_iA_i^2/[n_i\ (n_i-h_i)]$ ⑨ |
|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 1.000 | 2 | | | 2.000 | | |
| 1 | 2 | 0.933 | 1 | 1 | 15 | 0.933 | 17.949 | 1.534 |
| 2 | 3 | 0.866 | 6 | 1 | 14 | 5.196 | 17.140 | 1.591 |
| 3 | 9 | 0.800 | 1 | 1 | 13 | 0.800 | 11.818 | 0.895 |
| 4 | 10 | 0.666 | 5 | 2 | 12 | 3.330 | 11.018 | 2.023 |
| 5 | 15 | 0.592 | 1 | 1 | 9 | 0.590 | 7.688 | 0.821 |
| 6 | 16 | 0.507 | 14 | 1 | 7 | 7.098 | 7.098 | 1.200 |
| 7 | 30 | 0.380 | | 1 | 4 | | | |
| | 45+ | | | | | | | |
| 合计 | | | | | | 19.947 | | 8.064 |

平均治疗病期用下式计算：

$$\mu=\sum_{i=0}^{k-1}{}_{ti}q_o(t_{i+1}-t_i)+{}_{tk}q_o(t_n-t_k)\qquad(2)$$

本例 $k=7$，$\mu=2.000+0.933+\cdots+7.09+0.38$（45−30）=25.647（月）

平均治疗病期的方差用下式计算：

$$S^2=[k/(k-1)]\sum\{h_iA_i^2/[n_i(n_i-h_i)]\}\qquad(3)$$

本例 $S^2=[7/(7-1)](8.064)=9.408$（月）

标准差为　$S=\sqrt{9.403}=3.07$（月）

**二、寿命表法**：该法适用于大样本，它是把全部随访研究对象，以等间隔时间，如每隔一年计算研究对象处理后各年的生存概率，然后按照概率乘法法则，将各时期生存率相乘，即可得到自观察开始到各时点的生存率。

例 2　以例 1 的资料应用寿命表法求治疗后各年治愈率及其标准误。计算结果列于表 3。

**表 3　用寿命表法计算治愈率**

| 治疗后年数 $t_i\sim$ ① | 期初病例数 $n_i$ ② | 治愈人数 $h_i$ ③ | 失访或终访未愈人数 $w_i$ ④ | 校正人数 $n_i'$ ⑤ | 治愈概率 $p_i$ ⑥ | 未治愈概率 $q_i$ ⑦ | 从开始到 $t_i$ 时未治愈率 ${}_{ti}q_o$ ⑧ | 从开始到 $t_i$ 时治愈率 ${}_{ti}p_o$ ⑨ | 治愈率的标准误 $S_{ti}p_o$ ⑩ |
|---|---|---|---|---|---|---|---|---|---|
| 0～1 | 15 | 5 | 0 | 15 | 0.333 | 0.667 | 0.667 | 0.333 | 0.121 |
| 1～2 | 10 | 2 | 3 | 8.5 | 0.235 | 0.765 | 0.510 | 0.490 | 0.142 |
| 2～3 | 5 | 1 | 1 | 4.5 | 0.222 | 0.778 | 0.397 | 0.603 | 0.145 |
| 3～4 | 3 | 0 | 3 | 15 | 0 | 1 | 0.397 | 0.603 | 0.145 |

1. 期初病例数的第一行数值即治疗总人数，本例为 15，以后各行的 $n_i$ 用下式计算：

$$n_i = n_{i-1} - h_{i-1} - w_{i-1} \tag{4}$$

2. 治愈人数 $h_i$ 是从原始数据 $T_i$ 点数而得。

3. 撤出的失访和终访未愈人数 $w_i$ 是从原始数据 $C_i$ 点数而得。

4. 第⑤栏校正人数 $n_i'$ 用下式求得：

$$n_i' = n_i - \frac{w_i}{2} \tag{5}$$

5. 第⑥栏治愈概率 $p_i$，第⑦栏的未治愈概率 $q_i$，分别用下两式求得：

$$p_i = h_i/n_i' \tag{6}$$

$$q_i = 1 - p_i \tag{7}$$

6. 从开始到 $t_i$ 时未治愈率 ${}_{ti}q_o$ 与治愈率 ${}_{ti}p_o$，分别用下两式计算：

$$ {}_{ti}q_o = \prod_{i=1}^{k} q_i \tag{8}$$

$$ {}_{ti}p_o = 1 - {}_{ti}q_o \tag{9}$$

比如，组数 $k=4$，${}_{t4}q_o=0.667\times0.765\times0.778\times1=0.397$

${}_{t4}p_o=1-0.397=0.603$，余类推。

7. 治愈率的方差与标准误，用下两式计算

$$S^2_{ti}p_o = {}_{ti}q_o \sum_{i=1}^{k} \frac{h_i}{n_i'(n_i'-h_i)} \tag{10}$$

比如，本例 $k=4$ 时，四年生存率为

$S^2_{t4}p_o=0.397^2\times$（$\frac{5}{15\times10}+\frac{2}{8.5\times6.5}+\frac{1}{4.5\times3.5}+\frac{0}{15\times(15\times0)}=0.145$　余类推。

平均治疗病期 $\hat{\mu}$ 用下式计算

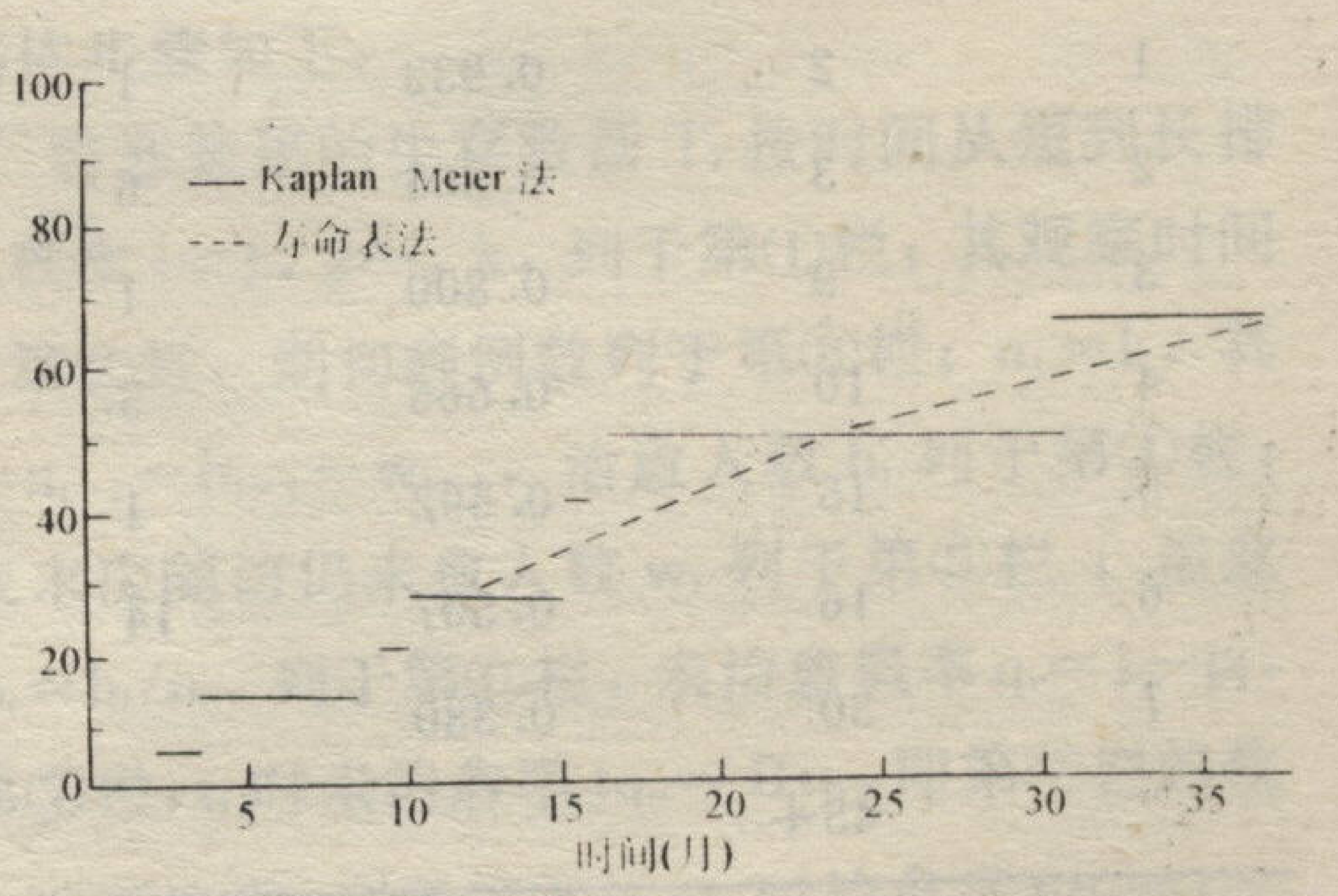
$$\hat{\mu} = \sum_{i=1}^{k} {}_{ti}q_o I_i \tag{11}$$

式中 $I_i$ 为第（1）栏治疗后年数的组距。本例 $k=4$，各组距皆为 1，故

$\hat{\mu}=0.667\times1+0.510\times1+0.397\times1+0.397\times1=1.97$（年）

**三、时间效应曲线**：它能全面反映效应率在时间上的动态变化，所以比计算某时点的效应率要好，通常以观察时间 t 作为横坐标，以相应的效应率（比如，生存率、治愈率、畸残率等）作为纵坐标而得到一条时间效应曲线。

例 3　以例 1、例 2 的结果，绘制时间效应曲线。




附图　例 1、例 2 数据的时间效应曲线

时间效应曲线是很有用的工具，可直视不同时期、不同对象、不同疗法的动态效应。**（待续）**

（收稿：1996-09-15）