

随访研究中的失访及其控制

河南省医学科学研究所 赵国龙

随访研究中常有失访，致使经典试验设计和经典统计学方法不能直接用于这类研究，其设计、测量和评价因此而复杂化。本文回顾失访的成因和后果，提出对失访进行控制的必要性和方法，并讨论了其伦理学方面。

随访与失访

随访记录见附表，终点是死于宫颈癌。结果1表明随访已完成。2是失访，原因不明，或为迁徙、退出等，留下不完整随访史。3~8均形成不完整随访史。偶尔，工作失误、资料丢失、方案破坏等在后果上也类似失访。如果这是最后一次记录，9~12也形成不完整史。

失访和截尾

由完整随访史得到一个精确生存时间 t_j ，而不完整史不能提供生存时间的直接信息，唯知下限，写作 t_j^* ，称右截尾。这是随访研究中主要的截尾形式，简称截尾。精确生存时间为非截尾值。结果2~8出现在试验期间任何时候，称失访截尾；9~12由试验结束形成，为终末截尾。两者在设计时分别估计，在测量中各自形成，在评价中同等对待。这样，每个样本分成非截尾值 t_j 子样本和截尾值 t_j^* 子样本，其条件分布分别为 F_D 和 F_C ，评价时需估计生存分布 F 。若用经典方法评价，只好将 t_j^* 当作 t_j 或将 t_j^* 去掉。前者显然低估了生存期，后者等于以 F_D 代替 F ，导致对生存率的严重低估。

截尾数的表达

用生存分析可免除上述偏性。但截尾越多，生存率或其他统计量方差越大，结果越不可靠。着眼于评

价，Simon R.提出过一个不可评价率，这指失访截尾。他认为该率 $\geq .20$ 说明病例选择不当，成为偏性的来源，建议该率不宜大于.15。着眼于设计，笔者曾提出截尾率： $R_i = (n_i - d_i) / n_i$ ，这里 $i=c, e$ ，分别表示对照和试验组， d_i 为非截尾值数， $n_i - d_i$ 为截尾值数。设病人达终点和失访时间均为指数分布，危险率分别为 θ_i 和 ϕ_i ，那么失访截尾率为 $l_i = \phi_i / (\theta_i + \phi_i)$ ，终末截尾率为 $r_i = 1 - (1 - l_i)^{-1} (1 - R_i)$ 。前者对应于不可评价率，后者推导自Rubinstein[3]公式(A2)。

对失访的控制

与Simon相呼应，建议以 $l \leq .1$ 为轻度失访， $.1 < l \leq .2$ 为中度失访， $l > .2$ 为重度失访，作为失访控制的尺度。一般不宜取重度。用软件CHL6-IV可辅助选取失访截尾率，并以此为设计参数调整样本容量和试验期限。控制措施是1.完善病人进入登记手续，特别是适当的选择条件和同意表条款；2.制定周密随访计划，汇总每次失访数；3.建立对试验的监测计划，即早了解失访是否可能超出预定失访截尾率；4.最好有中期分析计划，按此调整随访计划，必要时追加随访。经验告诉我们，工作越细，失访越少。至于那些意外损失，可通过监测计划及时发现和补救。终末截尾则由试验结束早晚加以调节。

结 论

以往设计方法文献不考虑失访控制。由他们的方法推算出的失访截尾率有的高达.67。这难以保证结果的可靠性。采用本文的控制方法可减小样本，节约经费，保证结果的可靠性。

附表

宫颈癌登记病例随访结果分类

1.死，宫颈癌存在	7.死，第二种原发癌和宫颈癌存在
2.失访	8.死，没有关于死亡原因的资料
3.死于一种间发病，宫颈癌不存在	9.活，没有宫颈癌复发的征象
4.死于一种间发病，无宫颈癌存在与否的资料	10.活，宫颈癌存在
5.死，间接归因于治疗	11.活，第二种原发癌存在而宫颈癌不存在
6.死，第二种原发癌存在，宫颈癌不存在	12.活，第二种原发癌和宫颈癌存在