

遗传毒理学短期试验统计评价的要点

周宗灿 王纪宪¹

北京医科大学卫生毒理学教研室 ¹生物数学和生物统计学教研室 北京 100083

遗传毒理学短期试验(STT)的目的是检测化学品及环境样品的遗传毒性,评定对哺乳动物和人体细胞及性细胞的遗传危险性,以及,预测对哺乳动物和人的致癌性。统计学观点及方法在 STT 的设计和结果评价中起关键的作用,STT的发展和广泛应用也促进了生物统计学的发展⁽¹⁻³⁾。本文介绍遗传毒理学STT统计学评价的要点及进展。

1. STT设计的统计学要求^(1,2,4)

STT的设计应遵循随机、重复及对照三个原则,要求各观察值具有代表性,并且是相互独立的。STT的设计具体涉及到剂量点数目及间隔,每个剂量点的实验单位数,每个实验单位接种及计数的细胞数,对照组的设置等。实验单位的确定是很重要的,以培养物或动物作为实验单位要比以细胞作为实验单位更为可取。

样本的代表性要求具有同质性,即各处理组和对照组的非实验因素的条件均衡一致,为此,各实验单位(培养物或动物)的分组及整个实验的全部操作都应遵循随机化原则。在利用形态学指标的STT,必须采用盲法观察结果,以消除实验执行者观察结果的偏性。样本应有足够的大小和适当的重复次数,以估计处理之间、实验室内和实验室间的变异性。一般可根据显著性水准、检测把握度、容许误差、总体标准差等来估计样本的大小。

STT 常用的对照有4种:① 空白对照组,不施加任何处理因素,可确定试验体系

的生物学性状及自发突变率。② 阴性(溶剂)对照,不给处理因素但给以必须的实验因素(溶剂),以排除此实验因素(溶剂)的影响,阴性对照是作为与处理组比较的基础。③ 阳性对照,用已知的致突变物检测试验体系的有效性。阳性对照组应与受试物用相同的溶剂、染毒途径及采样时间。④ 历史对照,由本实验室过去多次实验的对照组数据组成,上述三种对照都可构成相应的历史对照。历史对照的最好用途是通过同质性检验检查试验体系的稳定性,即进行实验室质量控制和保证。

严格执行 STT 设计的上述要求,才可能得到可靠的、重复性和再现性良好的结果,也是进行正确的统计学评价的基础。

2. STT结果的数据结构和统计学分布

STT的数据通常是由剂量水平和相应观察值组成的二维关系型数据。绝大多数STT的观察值为计数资料(如突变体数、畸变数、SCE数)或是相对数(如存活细胞的突变率),因此 STT 结果的统计学主要是对离散性资料的统计学推断。

从理论上来说,突变是罕见事件,服从泊松分布,但实际上,STT数据分布模型很复杂,也可服从二项分布、负二项分布等。样本分布的研究可用拟合优度检验,如 χ^2 检验、拟似然比检验、精确检验、Kolmogorov-Smirnov 检验及 Fisher 推荐的离差检验。STT资料通常比常规的统计学分布模型具有更大的变异性,样本数据的均数与泊松分布

模型拟合较好, 而方差显著大于模型的方差, 即出现超离差 (overdispersion)。超离差可假定泊松分布的均数服从某一特定的分布来表达, 即构成复合分布 (compound distribution)⁽⁶⁾。如负二项分布可记为泊松 * 伽马复合, 其他还有多种复合分布, 如泊松 * 反伽马复合、泊松 * 正态复合等。忽视超离差可使假设检验的第一类错误概率增加、参数估计的可信区间过窄。因此, 超离

差是STT资料统计学分析的重要问题。

3. STT资料的统计学分析

对 STT 资料的统计学分析主要是检验各处理组与阴性对照组之间的差别, 研究剂量-反应关系和定量遗传毒性强度。常规 STT 已报告的统计学检验方法见附表⁽¹⁻⁴⁾, 其中以对 Ames 试验结果的统计学研究资料最丰富。

表 常规 STT 已报告的统计学检验方法

统计学方法	SAL	GMT	CGT	MNT	SCE	YER	SLRL	UDS	DLT
处理组与阴性对照组比较									
1. 变换后 t 检验/多重比较	✓	✓	✓	✓		✓		✓	
2. 变换后方差分析	✓	✓		✓	✓				✓
3. χ^2 /K-B表/精确检验			✓	✓	✓	✓	✓		✓
4. 秩检验	✓		✓	✓					✓
5. Kolmogorov-Smirnov 法	✓		✓						
剂量-反应关系定性评价									
1. 线性趋势检验	✓	✓	✓	✓	✓	✓	✓		✓
2. 保序回归	✓								
3. Jonckheer 法	✓	✓		✓					✓
4. Page 检验	✓								
剂量-反应关系定量评价									
1. 线性回归	✓		✓			✓			
2. 适应法线性回归	✓								
3. 广义线性模型	✓		✓	✓	✓				
4. 非线性模型	✓	✓							

注: 常规 STT 的缩写为: SAL, Ames 试验; GMT, 哺乳动物细胞突变试验; CGT, 细胞遗传学试验; MNT, 微核试验; SCE, 姐妹染色单体交换试验; YER, 啤酒酵母有丝分裂重组试验; SLRL, 果蝇伴性隐性致死试验; UDS, 程序外DNA合成试验; DLT, 显性致死试验

对 STT 处理组与阴性对照组观察值均数的比较, 如果数据服从泊松分布或二项分布, 可以根据条件选用相应的 u 检验、Kastenbaum-Bowman 表 (K-B 表)、 χ^2 检验或 Fisher 精确检验。更常用的是进行数据转换, 以满足正态性和方差齐性, 利用 t 检验、多重比较 (Dunnett 检验和 Bonferroni 校正) 或方差分析。如果任何变换都不能改善数据的分布, 可能存在个别可疑值, 应予以识别和剔除⁽⁵⁾。另一方面, 可使用不依

赖总体分布模型的非参数统计分析, 如 Wilcoxon 秩和检验、Mann-Whitney 秩检验、H 检验及两样本 Kolmogorov-Smirnov 检验等。

对 STT 结果的统计学分析更为重要的是剂量-反应关系研究。一些弱遗传毒物的各剂量组与阴性对照组比较差别均无显著性, 但可具有阳性剂量-反应关系。剂量-反应关系研究分为定性和定量研究两类, 定量研究还可用来评价受试物的遗传毒性强度。

定性研究方法主要是趋势检验，定量研究方法包括线性和非线性模型拟合。趋势检验和广义线性模型的应用是 STT 统计学方法最主要的进展。

4. 趋势检验(trend test)

趋势检验是定性评价 STT 资料剂量-反应关系的重要工具，研究随剂量增加反应是否存在单调性上升或下降的趋势。参数检验主要用线性趋势检验(Cochran-Armitage 检验)，此检验适用于泊松分布和二项分布，也可推广至负二项分布。非参数检验有 Jonckhere 检验、Page 检验及置换检验等。

5. 广义线性模型 (generalized linear model)

广义线性模型(GLM)⁽⁶⁾是一般线性模型的推广。此模型由两部分组成，即随机部分和结构部分。随机部分规定数据的统计分布，包括正态分布、泊松分布、二项分布和负二项分布，并可推广到超离差模型。结构部分表现了剂量-反应关系。设剂量 d_i 时，观察值的均值为 μ_i ，则 μ_i 和 d_i 之间的关系在模型中可表示为 $\mu_i = \eta(\beta_0 + \beta_1 d_i)$ 。 η 是一个平滑的函数，称为连接函数。例如，对于泊松分布， $\mu_i = \exp(\beta_0 + \beta_1 d_i)$ 。在定量评价剂量-反应关系时，首先要对模型进行拟合，即估计参数 β_0 和 β_1 。当 β_1 显著地不为零时，可以认为存在剂量-反应关系。利用此模型可预测剂量 d 时 μ 的估计值及其可信区间。模型拟合的方法可用经验权重法和最大似然解法，目前最常用的是迭代加权最小二乘法。这些方法可以利用统计程序包如 SAS, Genstat 等来实现。

在 GLM 中用拟似然函数(quasi-likelihood function)代替模型的随机部分，就可以简化方法，根据观察值的均值和方差即可进行模型拟合。利用拟似然函数的另一个优点是适用于多种超离差分布，这就建立

了一个适应性很强的统一的模型。超离差模型的拟合仍可采用迭代加权最小二乘法，但在加权中要计入超离差的作用。模型中除系数 β_0 和 β_1 外，还含有超离差系数，此系数可以用矩方程进行估计。

6. 对统计学分析的评价

目前，对各个 STT 常规试验分别推荐了几种以至十余种统计学分析方法(见表及文献^(1,2))，一种 STT 可以有若干种正确的统计学分析方法，但可能不存在唯一正确的方法。其原因主要是表面上不同的统计学分析方法常以相同的统计学概念和模型为基础。另一方面，利用不同的统计学方法来评价 STT 资料缺乏比较研究，在这方面 Mont Carlo 法是很有用的工具。⁽⁷⁾

关于 STT 阳性结果判定，一般的规则是如果有 2 个或 2 个以上的剂量水平有超过阴性对照的统计学显著的增加，此增加与剂量相关，并超过根据历史对照预先规定的界限，即可判为阳性。处理组与阴性对照组的比较和剂量-反应关系研究应同时进行，综合评价。Galloway 等⁽⁸⁾对体外细胞遗传学试验结果的趋势检验和与阴性对照组比较鉴别有显著性的剂量组数结合起来，将结果分为 5 种：阴性、阳性有剂量相关趋势、阳性无剂量-反应关系、弱阳性、可能阳性。此方法有待发展和验证。

每个实验都必须予以重复，如果 2 次实验的结果得到矛盾的结果，则应进行第 3 次实验。这就有 6 种可能的结果：++，+-，-+，--，+-+ 和 -+-，前 3 种结果可判为阳性。也有学者提出对重复实验的结果应进行一致性检验，如是，则可合并这些资料，并进行合并的显著性检验⁽²⁾。

对同一受试物在实验室内或实验室间同一种 STT 的多次实验资料进行综合评价和定量统计合并具有重要的意义，Meta 分析⁽⁹⁾

的理论和方法可能是很有价值的。对于多种STT结果的综合评价, ICP EMC 第1委员会专家组⁽¹⁰⁾以化学品的最低有效剂量和最高无效剂量进行变换, 并给以一系列权重因子, 得到遗传毒性强度的综合指数, 并发展了有关的统计学方法, 用此综合比较化学品的遗传毒性, 并预测基于遗传学改变的毒性(癌和可遗传的突变)。

对实验结果的解释应综合考虑生物学意义和统计学意义。统计检验的假设是关于总体特征的假设, 检验方法是以统计量的抽样分布为根据的, 得到的结论是概率性的, 不是绝对的肯定或否定, 不等同于有或无生物学意义。对实验结果作出科学的判断和解释, 应该根据统计学分析的结果、生物知识和经验。一般来说, 具有统计学意义是具有生物学意义的必要条件之一。正确地利用统计学假设检验的结果有助于确定实验结果的生物学关联。也有一些实验结果具有生物学意义但无统计学意义, 这可能是由于该事件的发生是极端罕见的, 但更常见的是因为实验设计不良所致。

7. 小结

本文简要介绍了STT的实验设计和统计学要求。STT数据主要是离散性资料(计数资料或相对数), 统计学分布模型很复杂, 常出现超离差现象。对常规的STT已推荐了多种统计学分析方法, 用此比较处理组和阴性对照组的差别及剂量-反应关系研究。近年来, 趋势检验和广义线性模型在STT统计学分析中取得较大的成功, 微型计

算机和统计学软件的普及为这些方法的应用提供了条件。对STT实验结果的评价和解释应综合考虑生物学意义和统计学意义。

参考文献

1. Eider L. Statistical methods for short-term tests in genetic toxicology; The first fifteen years. *Mutat Res*, 1992; 277(1): 11.
2. Kirkland DC. Statistical evaluation of mutagenicity test data. Cambridge; Cambridge Univ Press, 1989.
3. Hothorn L. Statistical Methods in Toxicology. Berlin; Springer, 1991.
4. Ehrenberg L. Aspects of statistical test procedure, in Kilbey BJ, et al, eds. Handbook of mutagenicity test procedures. 2nd, Amsterdam; Elsevier, 1984: 775-822.
5. Barnett V, Lewis T. Outlines in statistical data 2nd, New York: Wiley, 1984.
6. McCullagh P, Nelder JA. Generalized linear models, 2nd, London: Chapman and Hall, 1989.
7. Rubinstein RY. Simulation and the Monte Carlo method. New York: Wiley, 1981.
8. Galloway SM, Bloom AD, Resnick M, et al. Development of a standard protocol for in vitro cytogenetic testing with CHO cells; comparison of results for 22 chemicals in two laboratories. *Environ Mutagenesis*, 1985; 7(1): 1.
9. Hedges LV, Oikin I. Statistical methods for Meta analysis. New York; Academic Press, 1984.
10. Brusick DJ, Ashby J, de Serres FJ, et al. A method for combining and comparing short-term genotoxicity test data. *Mutat Res*, 1992; 266(1): 1.