

SAS 在辐射剂量-效应研究中的应用

刘长安

(北京大学第三医院肿瘤与放射病科, 北京 100083)

摘要: 简述辐射剂量-效应研究中常用的 13 种数学模型。应用 SAS 6.12 for windows 统计软件包编写的程序可以方便、快捷地实现辐射剂量-效应曲线的拟合并筛选出最佳的数学模型。

关键词: SAS; 数学模型; 辐射; 剂量-效应曲线

中图分类号: TP317; R852.7 **文献标识码:** A **文章编号:** 1002-221X(2001)01-0010-03

Application of SAS in the study on radiation dose-effect

LIU Chang-an

(Department of Oncology and Radiation Sickness, Peking University Third Hospital, Beijing 100083, China)

Abstract: Thirteen mathematical models usually used in study on radiation dose-effect are described in this paper. Radiation dose-effect curve can be fitted conveniently and quickly and an optimal mathematical model for it can be selected with computer programs programmed with Statistical Analysis System (SAS) software package, version 6.12 for windows.

Key words: Statistical Analysis System (SAS); Mathematical model; Radiation; Dose-effect curve

1 引言

外周血淋巴细胞染色体畸变和微核分析已广泛用于急性事故照射的剂量估算和筛检,是较好的近期照射的生物剂量计。辐射损伤及放射生物剂量学研究中某些效应指标与剂量的关系,需要有相应的数学模型作定量描述。剂量效应曲线的人工拟合非常复杂费时。涂开成等把放射损伤研究中描述剂量-效应关系的 12 个模式利用 BASIC 语言编制了计算机程序用于曲线拟合^[1]。我们进一步应用 SAS (Statistical Analysis System) 6.12 版编写了辐射剂量-效应曲线拟合程序,在 Windows 操作系统下运行,人机对话界面友好,简便快捷,有利于迅速得到各种数学模式的待定参数、 F 值、曲线拟合度 (R_{adj}^2) 等参数并筛选出最佳的数学模型。可供毒理学、放射生物剂量学及劳动卫生学研究人员参考。

2 数学模型及 SAS 程序

得到剂量、效应原始数据后,先建立数据集。利用以下数学模型及 SAS 程序逐一进行剂量-效应曲线的拟合。得到各种数学模式的待定参数、 F 值等,根据拟合度 (R_{adj}^2) 筛选出最佳剂量-效应曲线。程序主要用 REG 过程或 NLIN 进行回归分析。

数据步: Data ED;

```
input E D @@;  
cards;  
.....  
;  
run;
```

E 为某效应指标值, D 为吸收剂量 (Gy), 下同。假如数据存在 A 盘或 D 盘中,可写为: Data ED; infile A: \ * . *; input E D; run;。

辐射剂量-效应模型按不同的 LET, 不同剂量率可分为以下几种模型。其中前 4 种为 WHO (1973) 提出的数学模型^[1~3]。

2.1 线性无阈模型 (“一次击中”模型)

这种模型在高 LET 低剂量率照射时可出现,其模型的表达式为: $E = a + bD$ 。程序为: data ED1; set ED; proc reg; model E=D; run;。

2.2 幂函数模型

表达式为: $E = aD^b$ 。程序步为: data ED2; set ED; proc nlin; parms a=1 b=1; model E=a * D ** b; run;。

2.3 二次方程模型 (“二次击中”模型)

表达式为: $E = a + bD^2$ 。程序为: data ED3; set ED; D2=D * D; proc reg; model E=D2; run;。

2.4 二次多项式模型

低 LET 高剂量率照射时可出现,表达式为: $E = a$

收稿日期: 2000-06-11; 修回日期: 2000-08-21

作者简介: 刘长安 (1968-), 男, 陕西富平人, 助理研究员, 在职硕士生, 研究方向为造血干细胞损伤与修复。

+ $bD+cD^2$ 。程序为: data ED4; set ED; $D2=D*D$; proc reg; model $E=D D2$; run;。

2 5 线性平方模型

低 LET 低剂量率照射时可能出现, 表达式为: $E=bD+cD^2$ 。程序为: data ED5; set ED; $D2=D*D$; proc reg; model $E=D D2$ /nonint; run;。

2 6 指数模型

低 LET 高剂量率照射所致的剂量-效应关系, 可用该模型描述: $E=a*\exp(\lambda D)$ 。程序为: data ED6; set ED; proc nlin; parms $a=1 f=1$ (注: f 为 λ); model $E=a*\exp(f*D)$; run;。

2 7 修正的指数模型

低 LET 高剂量率照射所致的剂量-效应关系, 也可用该模型描述: $E=a+b*\exp(\lambda D)$ 。程序为: data ED7; set ED; proc nlin; parms $a=1 b=1 f=1$ (注: f 为 λ); model $E=a+b*\exp(f*D)$; run;。

2 8 S 型(实际阈值)模型

低 LET 高剂量率照射时出现的效应遵循该模型, 表达式为: $E=(bD+cD^2)\exp(\lambda D)$ 。程序为: data ED8; set ED; $D2=D*D$; proc nlin; parms $b=1 c=1 f=1$ (注: f 为 λ); model $E=(b*D+c*D2)*\exp(f*D)$; run;。

2 9 实际阈值模型

低 LET 高剂量率照射时亦可用下述模型描述: $E=(bD+cD^2)\exp[-(\lambda_1 D+\lambda_2 D^2)]$ 。程序为: data ED9; set ED; $D2=D*D$; proc nlin; parms $b=1 c=1 f_1=1 f_2=1$ (注: f_1, f_2 分别代表 λ_1, λ_2); model $E=(b*D+c*D2)*\exp(-f_1*D-f_2*D2)$; run;。

2 10 凸向下模型

高 LET 高剂量率照射进可出现下列模型形式的效应: $E=bD^2+c*\exp(-\lambda D)$ 。程序为: data ED10; set ED; $D2=D*D$; proc nlin; parms $b=1 c=1 f=1$ (注: f 代表 λ); model $E=b*D2+c*\exp(-f*D)$; run;。

2 11 无阈值性修正模型

该模型可用于描述低 LET 高剂量率照射时的剂量-效应曲线, 其式为: $E=a+bD+c*\exp(-\lambda D)$ 。程序为: data ED11; set ED; proc nlin; parms $a=1 b=1 f=1$ (注: f 为 λ); model $E=a+b*D+c*\exp(-f*D)$; run;。

2 12 无阈值线性平方修正模型

该模型可用于描述低 LET 高剂量率照射时的剂量-效应曲线, 即 $E=(a+bD+cD^2)\exp(-\lambda D)$ 。

程序为: data ED12; set ED; $D2=D*D$; proc nlin; parms $a=1 b=1 c=1 f=1$ (注: f 为 λ); model $E=(a+b*D+c*D2)*\exp(-f*D)$; run;。

2 13 logistic “S” 型模型

形式如下: $E=E_{\max}/[1+a*\exp(-\lambda D)]$ 。程序步为: data ED13; set ED; proc nlin; parms $E_{\max}=1 a=1 f=1$ (注: f 代表 λ); model $E=E_{\max}/(1+a*\exp(-f*D))$; run;。

NLIN 程序中, parms $a=1 b=1 c=1 f=1$; 设置参数 a, b, c, f 的初值。参数初值的设置有时并不是一件简单的事。NLIN 过程将用迭代的方法求得满意的最小二乘解。如果初值选择不好将找不到满意的最小二乘解, 迭代过程就不能满足收敛的要求。一般可根据观察数据的情况或用其他简单方法求得一个初步解来估计参数的初值^[4]。模型初值的确定可参考文献[5]。

3 应用及讨论

我们应用 WHO (1973) 所提出的 4 种数学模型及相应的 SAS 程序对文献[6]中的数据进行了剂量-效应曲线的拟合。文献用不同剂量快中子(吸收剂量率为 0.501Gy/min)、X 射线(吸收剂量率为 0.38Gy/min)和 γ 射线(吸收剂量率为 0.38Gy/min)照射离体人血, 观察双核淋巴细胞微核, 以建立不同辐射类型的剂量-效应曲线。上述类型辐射诱发的微核率结果见表 1。

拟合结果输出微核的剂量-效应回归方程式、回归系数检验、拟合度(R_{adj}^2)见表 2。运行全过程耗时约 4 分钟。NLIN 迭代过程(Gauss-Newton 法)均已达到收敛标准。

曲线拟合的好坏可由剩余平方和(Q)的大小来衡量, 如果 Q 对 L_{yy} (总平方和)的比例越小, 说明实际观察值与估计值越接近, 曲线拟合得越好, 这种拟合度可用相关指数(R^2)来表示。 R^2 越接近于 1, 表示拟合得越好。由于各数学模型自变量数目不同, 因此比较不同模型曲线的拟合度时, 校正的 R^2 (R_{adj}^2)比 R^2 准确。从结果上看, 12 条曲线均成立(回归系数显著性检验, P 值均小于 0.01)。根据相关指数(R^2), 在 0.1~5.0Gy 范围内, 拟合的最佳数学模型: 180kV X 射线为幂函数($E_x=150.618 0D^{1.322 9}$), ^{60}Co γ 射线为二次多项式($E_\gamma=17.972 8+33.252 5D+42.908 6D^2$), 2.14 Mev 快中子为二次多项式($E_n=55.157 2+563.614 5D-43.422 3D^2$)。

表 1 不同辐射类型照射离体人血 诱发淋巴细胞微核率 ($P \pm s_p, \%$)^[6]

剂量 (Gy)	180kV X 射线		⁶⁰ Co γ 射线		2 14Mev 快中子	
	分析细胞数	微核率 %	分析细胞数	微核率 %	分析细胞数	微核率 %
0 0	12 000	9 2±0 9	15 000	16.9±1 1	12 000	7.0±0 8
0 1	6 000	21.2±1 9	8 000	19.0±1 5	3 000	198.7±8 1
0 2	—	—	8 000	25.1±1 8	—	—
0 3	6 000	33.3±2 4	8 000	32.8±2 0	—	—
0 5	5 000	51.6±3 2	4 000	37.0±3 0	3 000	279.0±9 6
1 0	3 000	133.3±6 7	3 000	111.7±6 1	3 000	565.0±13. 7
2 0	3 000	340.0±10. 6	1 500	259.3±13. 0	—	—
2 5	—	—	—	—	1 530	1 235.9±28. 4
3 0	3 000	718.3±15. 5	1 500	494.0±18. 1	—	—
3 5	—	—	—	—	1 500	1 460.0±31. 2
4 0	3 000	911.7±17. 4	1 000	835.0±29. 0	—	—
5 0	3 000	1 265.0±20. 5	1 000	1 261.0±36. 0	900	1 790.9±44. 6

表 2 回归方程式、回归系数与拟合度检验

回归方程式	F 值	拟合度(R ² _{adj})
180kV X 射线		
(1)E _x =-49.252 5+248.535 7D	376	0.979 2
(2)E _x =150.618 0D ^{1.322 9}	940	0.994 1 *
(3)E _x =71.668 7+51.293 5D ²	210	0.967 8
(4)E _x =-7.251 1+150.767 3D+21.083 1D ²	447	0.991 1
⁶⁰ Co γ 射线		
(1)E _γ =-63.135 5+231.251 9D	151	0.943 7
(2)E _γ =76.020 1D ^{1.739 0}	2 637	0.997 8
(3)E _γ =34.253 1+49.634 8D ²	4 924	0.998 2
(4)E _γ =17.972 8+33.252 5D+42.908 6D ²	11 160	0.999 6 *
2.14Mev 快中子		
(1)E _n =145.853 6+358.375 0	161	0.963 8
(2)E _n =605.829 2D ^{0.690 0}	508	0.987 0
(3)E _n =360.395 7+67.330 9D ²	23	0.820 1
(4)E _n =55.157 2+563.614 5D-43.422 3D ²	376	0.992 1 *

注：①E_x、E_γ、E_n分别为X、γ、中子诱发的微核率(%)，D为剂量(Gy)。②回归系数显著性检验：各式P值均小于0.01。
③*：最佳拟合数学模型。

由此可见，我们应用 SAS 6.12 for windows 统计软件包编写的程序可以方便、快捷地实现辐射剂量-效应曲线的拟合并筛选出最佳的数学模型，从而避免了十分复杂费时的人工运算过程，大大提高了工作效率。该程序可供毒理学、放射生物剂量学及劳动卫生

学研究人员参考。
SAS 系统是一个用来管理、分析数据和编写报告的组合软件系统。SAS 研究所每年都在改进 SAS 系统，目前SAS 已由一个流行的功能强大的统计分析软件包发展成为用途广泛的第四代高级编程语言。系统有以下特点：处理方式灵活；信息存取简单；语言编程能力强大，程序书写自由简捷；可对数据连续处理；统计分析方法丰富；可以漂亮清晰的格式输出结果；有较强的宏命令；容易学习。SAS 软件已广泛用于政府办公、教育、科学和商业等从事数据管理与数据分析处理的诸多领域中^[4]。随着该软件近年在我国的迅速推广应用，必将大大提高医学科研人员的计算机应用水平及信息处理的工作效率和工作质量。

参考文献：
[1] 涂开成，曹珍山，叶常青，等. 辐射剂量-效应研究中的数学模型及应用[J]. 中国辐射卫生，1993，2(3)：107.
[2] WHO. Methods for the analysis of human chromosome aberrations. Geneva: WHO, 1973. 66.
[3] 王继先，金瑾珍，白玉书，等. 放射生物剂量学[M]. 北京：原子能出版社，1997. 216.
[4] 董大钧，金丕焕，陈康. SAS 统计分析[M]. 沈阳：辽宁科学技术出版社，1996. 137.
[5] 冯士雍. 回归分析法[M]. 北京：科学出版社，1985. 45.
[6] 白玉书，黄琦龙，关树荣，等. 快中子对X和γ射线诱发淋巴细胞微核效应的比较研究[J]. 中华放射医学与防护杂志，1998，18(4)：259.