

# 如何合理选择统计分析方法 处理实验资料（III）

胡良平

编者按

生物医学期刊是宣传和反映生物医学科研与临床研究成果的重要媒体，是培养年轻科研工作者的摇篮，也是一个国家科研实力的重要象征。期刊中学术论文的质量是期刊存在的重要保证，而学术论文质量高低的重要标志之一是科研设计和统计分析质量的高低。本刊在 2007 年中，拟邀请军事医学科学院生物医学统计学咨询中心主任、博士生导师胡良平教授，以“如何合理选择统计分析方法处理实验资料”为题，撰写 6 篇文章，每期发表 1 篇，较系统地介绍在生物医学论文写作中，如何正确地应用医学统计学知识，从而提高学术论文的质量。需要指出的是，论文中统计学应用正确，并不能说明科研课题做得一定正确。广大作者和读者更应高度重视科研工作之前的科研设计的质量。事实上，由一个错误的科研设计产生出来的实验结果，即使其论文写得再漂亮，统计分析方法用得再正确，对于一个国家科技事业的发展和人才培养都是有害无益的。

合理选择统计分析方法处理定量资料最大的障碍就是准确判定其所对应的实验设计类型，有些实际问题看上去很复杂，因此，要想做到正确辨析其设计类型，需要通过解决尽可能多的实际问题来积累经验，为此，本文再举几个生物医学科研中的实例，以帮助读者提高识别实验设计类型的能力。

例 1 为了研究肌内注射灭活 SARS 病毒早期对 BALB/c 小鼠黏膜部位淋巴细胞构成比的影响，研究者将 20 只 BALB/c 小鼠完全随机地均分为肌内注射组和盐水对照组，在每只动物的 4 个部位“即 NALT、NP、肺脏和 PP”检测 3 项定量指标（即 CD4+％、CD8+％和 CD3+％）的数值，见表 1。请根据表 1 中的定量资料辨析实验设计类型。

表 1 肌内注射灭活 SARS 病毒早期对 BALB/c 小鼠  
黏膜部位淋巴细胞构成比的影响

部位	组别	CD4+%	CD8+%	CD3+%
NALT	肌内注射	35.05±7.67	8.34±3.25	66.41±12.13
	盐水对照	34.26±6.06	5.53±2.77	64.97±7.02
NP	肌内注射	8.81±2.24	2.10±0.74	17.99±4.93
	盐水对照	4.01±0.57	0.85±0.16	9.09±2.60
肺脏	肌内注射	30.07±4.13	11.05±2.00	49.38±9.58
	盐水对照	34.40±5.96	15.01±2.44	42.38±10.42
PP	肌内注射	13.49±4.91	3.46±1.39	41.98±10.27
	盐水对照	14.26±2.04	2.72±1.25	41.99±8.39

分析与解答：表 1 包含 2 个实验因素。其一是“部位”，具有 4 个水平：NALT、NP、肺脏、PP；其二是组别，具有 2 个水平：肌内注射和盐水对照。这里，“肌内注射”的全称为“肌内注射灭活 SARS 病毒”；而“盐水对照”的全称为“肌内注射生理盐水”。

如果研究者是从每只小鼠的 4 个不同部位上分别观测 3 个指标的取值，则该实验所对应的实验设计类型是“具有一个重复测量的两因素设计”，即在 4 个不同部位上进行重复测量；如果仅观测每只小鼠的 1 个部位，该实验所对应的实验设计类型是“两因素析因设计”。表 1 不符合统计表编制要求，不便于辨析其实验设计类型，修改结果见表 2。

表 2 肌内注射灭活 SARS 病毒早期对 BALB/c 小鼠  
黏膜部位淋巴细胞构成比影响的测定结果

肌内注射物 种类	注射 部位	CD4+%	CD8+%	CD3+%
灭活 SARS 病毒	NALT	35.05±7.67	8.34±3.25	66.41±12.13
	NP	8.81±2.24	2.10±0.74	17.99±4.93
	肺脏	30.07±4.13	11.05±2.00	49.38±9.58
	PP	13.49±4.91	3.46±1.39	41.98±10.27
生理盐水	NALT	34.26±6.06	5.53±2.77	64.97±7.02
	NP	4.01±0.57	0.85±0.16	9.09±2.60
	肺脏	34.40±5.96	15.01±2.44	42.38±10.42
	PP	14.26±2.04	2.72±1.25	41.99±8.39

表 1 与表 2 似乎一样，仅前两列的位置互换了一下。其实，这两张表有本质区别。表 1 似乎是将“部位”视为“实验分组”因素，即全部动物先按部位分成 4 组，在每组中再按肌内注射物种类分成 2 个小组，然后，对小组中的每只动物检测 3 项指标的数值；而实际实验是按表 2 的顺序来做的，即先将全部动物完全随机地分成 2 组，分别肌内注射不同种类的物质（即灭活 SARS 病毒、生理盐水），然后，对注射不同物质的每只动物在其 4 个部位上观测 3 项指标的数值。用表 2 表达该实验，其实验设计类型还表

表 3 肌内注射灭活 SARS 病毒早期对 BALB/c 小鼠黏膜部位淋巴细胞构成比影响的测定结果

肌肉注射物种类	CD4+%				
	注射部位:	NALT	NP	肺脏	PP
灭活 SARS 病毒		35.05 ± 7.67	8.81 ± 2.24	30.07 ± 4.13	13.49 ± 4.91
生理盐水		34.26 ± 6.06	4.01 ± 0.57	34.40 ± 5.96	14.26 ± 2.04

表 4 辐照前后红细胞溶血率的比较

	0 d 辐照组溶血率 (%) (n = 11)						21 d 辐照组溶血率 (%) (n = 9)		
	0 d	7 d	14 d	21 d	28 d	35 d	21 d	28 d	35 d
对照	0.03 ± 0.01	0.05 ± 0.02	0.07 ± 0.02	0.12 ± 0.04	0.17 ± 0.05	0.27 ± 0.10	0.12 ± 0.04	0.17 ± 0.05	0.25 ± 0.05
10 Gy	0.03 ± 0.01	0.06 ± 0.01	0.08 ± 0.01	0.14 ± 0.04	0.21 ± 0.06	0.32 ± 0.10	0.12 ± 0.04	0.20 ± 0.05	0.29 ± 0.07
25 Gy	0.04 ± 0.01	0.07 ± 0.01	0.11 ± 0.02	0.18 ± 0.05	0.29 ± 0.08	0.38 ± 0.10	0.14 ± 0.05	0.22 ± 0.06	0.35 ± 0.06
35 Gy	0.04 ± 0.01	0.10 ± 0.03	0.16 ± 0.03	0.24 ± 0.04	0.34 ± 0.08	0.42 ± 0.10	0.16 ± 0.06	0.26 ± 0.06	0.39 ± 0.04
45 Gy	0.03 ± 0.01	0.12 ± 0.05	0.24 ± 0.08	0.35 ± 0.14	0.43 ± 0.16	0.54 ± 0.19	0.18 ± 0.08	0.31 ± 0.05	0.44 ± 0.04

达得不是特别清晰,若将其中一个定量指标按表 3 的形式表达出来,则一目了然。

表 3 就是表达“具有一个重复测量的两因素设计”的标准模式,写在表左边的因素叫做“实验分组因素”,写在表头横线下方的因素叫做“重复测量因素”。

例 2 为了研究动物在不同剂量辐照下体内红细胞溶血率的变化规律,研究者将动物随机分入辐照“0 d 组”和“21 d 组”2 个大组,在每个大组中,再将动物随机分为 5 个剂量组,然后,对每个剂量组中的每只动物在多个时间点上重复观测“溶血率(%)”,设计和实验结果见表 4。请根据表 4 中的定量资料辨析实验设计类型。

分析与解答:第 1 种考虑问题的方法,是将表 4 中“0 d 辐照组溶血率(%)”和“21 d 辐照组溶血率(%)”看作是 2 个定量观测指标。当观察“0 d 辐照组溶血率(%)”指标时,实验涉及到 2 个因素。其一是“辐照剂量”,具有 5 个水平;其二是“观测时间”,具有 6 个水平。在时间因素上重复测量了红细胞溶血率,故实验设计类型是“具有一个重复测量的两因素设计”。当观察“21 d 辐照组溶血率(%)”指标时,实验涉及到 2 个因素。其一是“辐照剂量”,具有 5 个水平;其二是“观测时间”,具有 3 个水平。在时间因素上重复测量了红细胞溶血率,故实验设计类型是“具有一个重复测量的两因素设计”。其标准的列表格式参见例 1 的表 3。

第 2 种考虑问题的方法,是将 0 d 辐照和 21 d 辐照看作 1 个大的分组因素的 2 个水平,而“观测时间”因素只取“21 d、28 d 和 35 d”3 个水平,此时,所对应的实验设计类型为“具有一个重复测量的三因素设计”。其标准的列表格式见表 5。

例 3 为了研究安君宁对吗啡依赖大鼠部分脑结构 PENK mRNA 表达的影响,研究者将全部动物随机分为 5 组,即正常组、P12EC 组、P12E 组、P30EC 组和 P30E 组,对每组中的每只动物在其 3 个部位检测部分脑结构 PENK mRNA 表达。请根据表 6 中的定量资料辨析实验设计类型。

分析与解答:第 1 种考虑,认为纵向的 5 个组之间在

表 5 两种辐照时间长度以及不同辐照剂量条件下红细胞溶血率的测定结果

辐照 时间 (d)	辐照 剂量 (Gy)	溶血率 (%)			
		*	21	28	35
0	0	0.12 ± 0.04	0.17 ± 0.05	0.27 ± 0.10	
	10	0.14 ± 0.04	0.21 ± 0.06	0.32 ± 0.10	
	25	0.18 ± 0.05	0.29 ± 0.08	0.38 ± 0.10	
	35	0.24 ± 0.04	0.34 ± 0.08	0.42 ± 0.10	
	45	0.35 ± 0.14	0.43 ± 0.16	0.54 ± 0.19	
21	0	0.12 ± 0.04	0.17 ± 0.05	0.25 ± 0.05	
	10	0.12 ± 0.04	0.20 ± 0.05	0.29 ± 0.07	
	25	0.14 ± 0.05	0.22 ± 0.06	0.35 ± 0.06	
	35	0.16 ± 0.06	0.26 ± 0.06	0.39 ± 0.04	
	45	0.18 ± 0.08	0.31 ± 0.05	0.44 ± 0.04	

注: \* 代表“观测时间(d)”

专业上具有可比性。则可按如下方法分析:表 6 涉及到 2 个实验因素。其一为“部位”因素,具有 3 个水平:伏隔核、前额皮质和 VTA;其二为分组因素,具有 5 个水平:正常组、P12EC 组、P12E 组、P30EC 组和 P30E 组。2 种实验因素的水平共有 15 种组合。所以该实验对应的设计类型是“具有一个重复测量的两因素设计”,其中,“部位”是重复测量因素。

第 2 种考虑,认为纵向的 5 个组之间在专业上不具有可比性。则可按如下方法分析:需要对 5 组进行拆分,可形成以下 3 种组合。

组合 1: 正常对照组、P12EC 组、P30EC 组(分组本身为单因素 3 水平设计结构);

组合 2: 正常对照组、P12E 组、P30E 组(分组本身为单因素 3 水平设计结构);

组合 3: P12EC 组、P12E 组、P30EC 组、P30E 组[分组本身是两因素析因设计结构,两因素分别是“剂量(P12 或 P30)”和“类型(EC 或 E)"]。

结合“部位”因素,与组合 1 和组合 2 对应的实验设计类型为“具有一个重复测量的两因素设计”,其标准列表

表6 安君宁对吗啡依赖大鼠部分脑结构PENK mRNA表达的影响(×10<sup>-2</sup>,  $\bar{x} \pm s$ )

部位	正常组	P12EC组	P12E组	P30EC组	P30E组
伏隔核	20.18 ± 7.58	9.47 ± 0.87	14.74 ± 4.81	12.36 ± 6.95	18.46 ± 8.27
前额皮质	32.14 ± 5.47	10.87 ± 3.96	20.47 ± 7.03	19.78 ± 7.08	28.58 ± 6.29
VTA	29.90 ± 3.79	5.94 ± 0.91	16.09 ± 4.94	10.31 ± 0.78	25.73 ± 8.24

格式参见本文例1中的表3;与组合3对应的实验设计类型为“具有一个重复测量的三因素设计”,其标准列表格式参见本文例2中的表5。

例4 为了研究不同免疫时间TGI-310与G3PDH表达量的比值主要受哪些因素的影响,研究者设计了如下的实验,其设计和实验结果见表7。请根据表7中的定量资料辨析实验设计类型。

表7 不同免疫时间TGI-310与G3PDH表达量的比值

	胃	肠	PP
Normal	2.52 ± 2.11	2.03 ± 0.61	2.09 ± 0.96
PBS	2.77 ± 1.15	2.73 ± 5.53	2.36 ± 0.86
Ty21a 一免后第2天	4.64 ± 6.00	3.88 ± 2.63	2.63 ± 1.10
Ty21a 一免后第5天	3.79 ± 2.38	3.17 ± 3.98	2.35 ± 2.21
Ty21a 二免后第2天	2.04 ± 0.43	1.51 ± 1.33	1.65 ± 1.52
Ty21a 二免后第5天	1.60 ± 1.64	1.58 ± 0.75	1.48 ± 0.30
Ty21a 三免后第2天	1.93 ± 1.05	1.57 ± 1.52	1.06 ± 0.30
Ty21a 三免后第5天	1.58 ± 0.75	1.34 ± 0.30	1.18 ± 0.53

分析与解答:表7看起来象是一个单因素8水平设计,其实涉及到多个实验因素。应该用拆分法处理。拆分的

结果如下。  
组合1: Normal、PBS;  
组合2: Ty21a 一免后第2天、Ty21a 一免后第5天、Ty21a 二免后第2天、Ty21a 二免后第5天、Ty21a 三免后第2天、Ty21a 三免后第5天;  
组合1包含2个实验因素。其一是组别,具有2个水平: Normal 和 PBS。其二是部位,具有3个水平: 胃、肠和 PP。所以组合1所对应的实验设计类型是“具有一个重复测量的两因素设计”,其中,重复测量因素为“部位”。  
组合2包含3个实验因素。第1个因素是“免疫等级”,具有3个水平: 一免、二免、三免;第2个实验因素是“免疫后天数”,具有2个水平: 第2天和第5天;第3个实验因素是“部位”,具有3个水平: 胃、肠和 PP。组合2所对应的实验设计类型是“具有三个重复测量的三因素设计”,因为每个受试对象都要在“免疫等级”、“观测时间”和“部位”上被重复观测,即从每个个体身上要重复观测18个数据,其表头结构为: 一免后[第2天(胃、肠、PP)、第5天(胃、肠、PP)]; 二免后[第2天(胃、肠、PP)、第5天(胃、肠、PP)]; 三免后[第2天(胃、肠、PP)、第5天(胃、肠、PP)]。

• 读者 • 作者 • 编者 •

关于一稿两投和一稿两用问题处理的声明

《中国医药生物技术》杂志编辑部

为维护《中国医药生物技术》的声誉和广大读者的利益,特就一稿两投和一稿两用问题的处理声明如下。

1. 本声明中所涉及的文稿均指原始研究的报告或尽管2篇文稿在文字的表达和讨论的叙述上可能存在某些不同之处,但其主要数据和图表是相同的。所指文稿不包括重要会议的纪要、疾病的诊断标准和防治指南、有关组织达成的共识性文件、新闻报道类文稿以及在一种刊物发表过摘要或初步报道而将全文投向另一种期刊的文稿。
2. 如果一篇文稿已以全文方式在某刊物发表,除非文种不同,否则不可再将该文投寄给他刊。
3. 请作者所在单位在来稿推荐信中注明该文稿有无一稿两投问题。
4. 凡来稿在接到本刊编辑部回执后满2个月未接到退稿,则表明稿件仍在处理中,作者欲投他刊,应事先与本刊编辑部联系。
5. 编辑部认为文稿有一稿两投嫌疑时,将认真收集有关资料并仔细核实。确认属于一稿两投时,将立即进行退稿处理。期刊编辑部与作者双方意见发生分歧时,将由上级主管部门或有关权威机构进行最后仲裁。
6. 一稿两用一经证实,本刊将择期在杂志中发布撤销该论文的声明;对该文作者作为第一作者所撰写的一切文稿,2年内拒绝在本刊发表;就此事件向作者所在单位和该领域内的其他科技期刊进行通报。