

• 讲座 •

如何正确进行生物医学科研设计

II. 如何合理选用单因素实验设计

李长平, 胡良平

编者按

生物医学科研的研究对象是生物体, 而生物体具有极大的变异性。科研的目的就是要从表面上看似杂乱无章的事物或现象中找出规律性的东西来, 以便正确地解释和揭示事物或现象的变化、发展乃至消亡的规律, 从而达到认识人类社会并促进其发展、认识自然、改造自然以至于征服自然之目的。人类要完成此重任, 仅靠热情和苦干是不够的, 必须加上巧干。所谓巧干, 就是要充分利用人类已经积累起来的宝贵经验, 加上人类特有的创造性思维和勇于实践的精神, 在确立了可行的奋斗目标之后, 对未知事物或现象进行调查或试验。要想使调查或试验得出的结论真实可信, 同时, 花费的人力、物力、财力和时间又比较少, 就需要有科学完善严谨的科研设计方案的指导。本刊特邀军事医学科学院生物医学统计学咨询中心主任胡良平教授, 以“如何正确进行生物医学科研设计”为题, 撰写系列统计学讲座。希望该系列讲座能对生物医学科研工作者制定科研设计方案有所帮助。

本讲座第一讲中提到统计研究设计渗透到生物医学科研工作的各个环节, 一个完善的统计研究设计对于科研工作来说, 就如同水对于鱼儿一样重要。在现实的科研工作中, 若熟练掌握了统计研究设计的相关知识和技能, 对于科研工作者来说能起到事半功倍的效果, 从而可以大大提高科研课题的质量。调查设计与实验(含临床试验)设计有个性也有共性, 从研究所涉及的因素及其组织形式上看, 两类研究对应的设计类型是统一的。由于各种科研设计的要点几乎都包含在实验设计之中, 所以在本文中主要介绍实际科研工作中使用频率非常高并易于被实际工作者接受单因素实验设计。

1 单因素实验设计的概念及适用范围

单因素实验设计是指实验中仅涉及 1 个实验因素, 未对其他任何重要非实验因素做有计划的安排, 仅仅希望通过随机化分组方法来平衡所有非实验因素对观测结果的影响。因此, 当欲考察的实验因素仅有 1 个、对重要的非实验因素不需要特意在实验设计中做出具体安排或可以凭借随机化来达到组间具有良好均衡性时, 可考虑采用单因素实验设计。

单因素实验设计主要包括单组设计、配对设计、成组设计和单因素多水平设计等 4 种设计类型^[1-2]。下面将通过具体实例, 介绍为实现何种研究目的, 可选择何种具体的单因素实验设计类型。

1.1 适合选用单组设计的场合

有定量观测指标的标准值而且不需要考察其他因素对观测结果的影响时, 适合选用单组设计。

例 1 某研究者欲分析吸烟与非吸烟的健康成年男子之间平均牙菌斑指数的差异有无统计学意义, 从符合条件的众多吸烟健康成年男子中随机抽查了 30 人, 获得了每个人的牙

菌斑指数(最低为 0.87, 最高为 2.69)。已知非吸烟健康成年男子的平均牙菌斑指数约为 1.23。请问: 该资料所对应的是何种实验设计类型?

实验设计类型的判定: 该研究的目的是分析吸烟健康成年男子平均牙菌斑指数的大小与其标准值或理论值(1.23)之间的差异有无统计学意义。在该研究中体现了如下特点: ①只选取了 30 名吸烟的健康成年男子, 未对他们按任何因素(如吸烟时间、年龄等)进行分组; ②仅涉及 1 个实验因素的 1 个特定水平(即“是否吸烟”中的吸烟); ③提供了进行假设检验需要的理论值或标准值(本例为非吸烟健康成年男子的平均牙菌斑指数 1.23)。故该实验设计属于“单组设计”, 该资料属于单组设计一元定量资料。进行数据分析时, 若数据满足参数检验的前提条件(即独立性和正态性), 可选用单组设计一元定量资料的 t 检验; 否则, 需要选用符号秩和检验。

单组设计就是对一组符合研究目的的受试对象(其总体应有明确的定义, 样本含量应通过专门公式计算出来)不按任何其他因素进行分组, 仅在实验因素的某特定水平(本例为在吸烟时间和吸烟量上符合一定标准的健康男性吸烟者)下观测其定量指标的数值, 定量指标有多少个, 就有多少组观测值。若希望对单组设计定量资料进行统计分析, 则必须提供定量观测指标的标准值或理论值。

需要指出的是, 上述实验设计中最关键的问题是研究者已有充足的理由(由充分的专业知识决定), 认为可以不考虑“吸烟年限”、“吸烟量”、“吸烟种类”、“刷牙情况”、“洗牙情况”等对“牙菌斑指数”的影响。否则, 用上述的单组设计和相应的统计分析方法来实现预先确定的研究目的是很不科学的。

作者单位: 100850 北京, 军事医学科学院生物医学统计学咨询中心

1.2 适合选用配对设计的场合

若实验中欲分析 1 个具有两水平的实验因素，若实际情况具备配对条件，可采用配对设计。

例 2 某研究者为了考察牙周炎基础治疗对患者龈沟液中弹性蛋白酶（EA）水平的影响，测得 8 例患者在进行牙周炎基础治疗前、后龈沟液中 EA 水平的数据如表 1 所示，请判断该资料所对应的实验设计类型。

表 1 8 例牙周炎患者基础治疗前、后龈沟液中 EA 水平的检测结果

患者编号	1	2	3	4	5	6	7	8
治疗前	1.236	0.646	0.085	0.887	1.62	0.588	0.96	0.049
治疗后	1.088	0.597	0.098	0.016	1.399	0.337	0.203	0.022

实验设计类型的判定：该研究的目的是考察牙周炎基础治疗是否具有一定的疗效。在该研究中体现了如下特点：①涉及 1 个实验因素（是否进行牙周炎治疗）的 2 个水平（治疗前、治疗后）；②每对数据来自同一个个体。观察 8 例患者接受治疗前、后 EA 的差异是否有统计学意义，属于自身配对设计。故该实验设计属于自身配对设计，该资料属于自身配对设计一元定量资料。进行数据分析时，若各对数据的差量满足参数检验的前提条件（即独立性和正态性），可选用配对设计一元定量资料的 *t* 检验；否则，应选用符号秩和检验。

1.3 适合选用成组设计的场合

若实验中仅涉及 1 个具有两水平的实验因素，实际情况不具备配对条件，并未对其他任何重要非实验因素进行有计划的安排，仅仅靠随机化的方法来平衡非实验因素的干扰和影响时，可采用成组设计。

例 3 某研究者欲考察急性缺氧条件下肺泡表面 2 种磷脂（A 和 B）的相对含量，选取大鼠作为受试对象，将 40 只大鼠随机分入正常对照组和急性缺氧组，每组 20 只，检测 2 种磷脂的相对含量。整理后的数据如表 2 所示。请判断该资料所对应的实验设计类型。

表 2 急性缺氧和正常对照组大鼠 2 种磷脂相对含量测定结果（ $\bar{x} \pm s$ ）

组别	鼠数	A	B
正常对照组	20	3.63 ± 3.25	56.38 ± 9.68
急性缺氧组	20	3.13 ± 2.92	45.61 ± 17.00

实验设计类型的判定：该研究的目的是考察急性缺氧对大鼠体内 2 种磷脂相对含量的影响。该研究有如下特点：①涉及 1 个实验因素（是否处于急性缺氧状态）；②将受试对象随机分为 2 组；③有 2 个定量观测指标（A 和 B）。故该实验设计属于完全随机分配受试对象的成组设计，该资料属于成组设计二元定量资料。若 A、B 两定量指标在专业上有联系，则属于二元方差分析问题；否则，可分别采用成组设计定量资料一元方差分析或秩和检验（两组定量资料

不满足方差分析前提条件，即独立性、正态性和方差齐性）。

1.4 适合选用单因素多水平设计的场合

若在实验研究中所考察的实验因素只有 1 个，并且该实验因素的水平数 $k \geq 3$ ，未对其他任何重要非实验因素进行有计划的安排，仅仅靠随机化的方法来平衡非实验因素的干扰和影响时，可采用单因素多水平设计，它是成组设计的扩大。

例 4 某研究者欲考察莪术水煎剂对未孕大鼠子宫肌电活动的影响，将 40 只大鼠随机均分成 4 组，对照组每只大鼠灌服生理盐水 10 ml/kg，3 个不同剂量莪术组每只大鼠分别灌服 25%、50%、100% 莪术水煎剂。观察每只大鼠子宫肌电爆发波的峰面积、持续时间和个数。资料如表 3 所示。请分析该资料所对应的实验设计类型。

表 3 莪术水煎剂对大鼠子宫肌电活动的影响（ $\bar{x} \pm s$ ）

莪术剂量（g/kg）	峰面积（mV·ms）	持续时间（s）	爆发波个数
0.0	142.65 ± 34.64	3.50 ± 1.23	0.54 ± 0.10
2.5	337.74 ± 250.91	8.82 ± 2.78	1.06 ± 0.34
5.0	449.50 ± 154.45	9.82 ± 2.89	1.42 ± 0.47
7.5	651.08 ± 98.20	13.07 ± 1.88	1.40 ± 0.25

实验设计类型判定：该研究的目的是考察不同剂量莪术液的效应大小。该实验有如下特点：①涉及 1 个实验因素（莪术水煎剂）的 4 个水平（莪术水煎剂的不同剂量，即 0、2.5、5.0、7.5 g/kg）；②将受试对象完全随机地均分到 4 个处理组中去；③有 3 个定量观测指标（子宫肌电爆发波的峰面积、持续时间、爆发波个数）。故该实验设计为单因素四水平设计。对资料进行统计分析，可回答 4 个剂量组各定量指标平均数之间的差异是否具有统计学意义。该实验资料为单因素四水平设计三元定量资料。若 3 个定量观测指标之间在专业上有密切联系，需要作为整体同时考虑的话，若资料满足参数检验的前提条件（即独立性、正态性、方差齐性），则需要采用单因素四水平设计定量资料三元方差分析；若各指标之间彼此没有十分密切的联系，且资料满足参数检验的前提条件，可分别选用单因素四水平设计定量资料一元方差分析；若资料不满足参数检验的前提条件，则可分别选用单因素四水平设计定量资料的 Kruskal-Wallis 秩和检验。

2 单因素实验设计的误用

单因素实验设计使用起来比较方便，人们习惯上首选单因素实验设计。但是事实上，任何一个科研课题，都很难想象能仅依赖于单因素实验设计及其对应的统计分析方法便可获得正确的结论。

当每次试验都涉及 2 个或 2 个以上实验因素时，将其他因素固定在各自的某个水平上，仅允许其中 1 个实验因素取一系列水平，做一批试验，得到一个被认为比较好的结果，从而确定该因素的一个“最佳水平”；再依次确定其他实验因素的“最佳水平”；最后将全部实验因素的“最佳水

平”组合在一起,就认为是进行此项实验研究的“最佳实验条件”。这种做法叫做“用多次单因素实验设计取代某种应采取的多因素实验设计”,当实验因素之间存在不可忽视的交互作用时,其结论的可信度低,甚至是完全错误的。

当研究者施加于受试对象的实验因素只有 1 个,但来自外界环境的干扰、受试者自身的条件或状况或心理状态甚至研究者的手法或态度或心理状态等可能对观测结果存在不可忽视的影响时(例如,不同日期做实验,无法保证实验条件完全相同;受试者患病的时间长短和病情轻重无法通过分层随机化使其在各组间完全均衡可比;研究者在不同时间做实验的手法无法做到前后一致),在分析实验数据时仍应采用相应的多因素实验设计类型或相应设计定量资料的协方差分析。否则,其结论的可信度也会下降。

有的实验从表面上看是“单因素实验”,而从本质上看却是“多因素实验”,不妨称这样的实验设计为“假单因素实验设计”或叫做“多因素非平衡组合实验”^[3-4]。请看下面的例子。

例 5 某研究者欲研究乙醇对不同性别小鼠体重的影响,选取体重 20~30 g 小鼠 24 只(雌雄各半),采用随机化的方法随机分为 3 组,分别以水、5% 乙醇、10% 乙醇为唯一饮料连续喂养 60 d,然后测量各组小鼠的体重增加量,最后用单因素方差分析得出随乙醇浓度的提高,雄性小鼠的体重明显增加($P < 0.05$),而雌性小鼠体重明显减轻($P < 0.05$)。根据研究者的实验分组、统计分析方法的选择和所陈述的统计和专业结论判断,请问:研究者在实验设计阶段犯了哪些错误?所选用的统计分析方法与实际存在的实验设计类型是否吻合?所得出的专业结论可信度高吗?

对差错的辨析和释疑:研究者在实验设计阶段犯了如下错误。

1. 实验设计中未正确贯彻实验设计的四原则,且误用统计分析方法。该研究的目的是探讨不同浓度乙醇对不同性别小鼠体重的影响,但根据所描述的分组方法,可知研究者并没有把 2 个欲考察的实验因素综合起来考虑。该实验首先涉及到了乙醇浓度(3 个水平:0、5%、10%)、性别(2 个水平:雌、雄)2 个因素,共有 6 种水平组合,研究者仅采用随机化的方法将 24 只小鼠随机分为 3 组是不妥当的,最好按性别进行分层随机化,即把雌、雄小鼠分别随机均分入 3 个不同乙醇浓度组,使每种组合下各有 4 只

小鼠。另外,在该实验中还涉及到了测定时间的问题。等实验结束后再把不同性别、饮用不同浓度乙醇的小鼠单独提出来进行分析,事实上割裂了整个实验设计,这样不能正确分析出不同浓度下、不同性别小鼠的体重增加量之间的差异有无统计学意义,更不能分析出不同乙醇浓度、性别和测定时间之间交互作用有无统计学意义。

2. 忽视了重要非实验因素对观测结果的影响。根据常识可知,动物的进食量对其体重增加量有不可忽视的影响,如果不同组小鼠的平均进食量明显不同,就不能说明体重增加量的不同是因为摄取不同浓度乙醇引起的。而该研究正是忽略了各组动物平均进食量这样一个重要的非实验因素对观测结果的影响。在实验设计时应提前考虑到这个因素,记录实验期间各组每只小鼠每餐的进食量,这样在实验结束时,才有可能通过采用协方差分析的方法消除各组小鼠平均进食量不同对观测结果评价所造成的不利影响,从而得出可靠的结论。换句话说,研究者在进行统计分析时,不仅少考虑了定性的实验因素,还少考虑了定量的实验因素,所采用的单因素设计定量资料方差分析与实际存在的实验设计是不吻合的。

由于研究者用假的单因素实验设计取代了多因素实验设计,其研究结果是不可信的。

参考文献

- [1] Hu LP. Modern statistics and the application of SAS. Beijing: Press of Military Medical Sciences, 2000:74-77. (in Chinese)
胡良平. 现代统计学与 SAS 应用. 北京:军事医学科学出版社, 2000:74-77.
- [2] Hu LP. Application of triple-type theory of statistics in experimental design. Beijing: People's Military Medical Press, 2006:44-63, 215-261. (in Chinese)
胡良平. 统计学三型理论在实验设计中的应用. 北京:人民军医出版社, 2006:44-63, 215-261.
- [3] Hu LP. The actual training of medical statistics. Beijing: Press of Military Medical Sciences, 2007:181-319. (in Chinese)
胡良平. 医学统计实战练习. 北京:军事医学科学出版社, 2007:181-319.
- [4] Hu LP. The design of scientific research and statistical analysis application in stomatology. Beijing: People's Military Medical Press, 2007:191-215. (in Chinese)
胡良平. 口腔医学科研设计与统计分析. 北京:人民军医出版社, 2007:191-215.

本期广告目次

北京源德生物医学工程有限公司 封三

国药励展展览有限责任公司 封四