

高级药理学讲义

(三) 生物统计学

(师资及研究生用)

上海第一医学院
基础医学部药理学教研室

一九八二年九月

萬葉詩序

(二) 1114丁巳

《萬葉詩序》

上古第一集
中古世間集

生物统计

目 录

第一章 总 论	1 - 1
第一节 生物统计的目的	1 - 1
第二节 概率与概率分布	1 - 2
第三节 样本与总体	1 - 6
第四节 随机抽样与分组	1 - 7
第五节 误差分类	1 - 8
第六节 实验指标的选择	1 - 10
第七节 实验指标的分类	1 - 11
第八节 精密度与正确度	1 - 13
习 题	1 - 15
第二章 平均数与差异估量	2 - 1
第一节 算术平均数和中位数	2 - 1
第二节 几何平均数和调和平均数	2 - 2
第三节 标准差	2 - 3
第四节 标准误	2 - 6
第五节 差异系数	2 - 7
第六节 可信限	2 - 8
第七节 质反应率的标准误及可信限	2 - 10

习 题	2-12
第三章 两组间比较	3-1
第一节 假设检验与显著性水平	3-1
第二节 双侧与单侧检验	3-4
第三节 t 检验	3-6
第四节 属于质反应的两阳性百分率比较	3-11
习 题	3-13
第四章 多组间比较	4-1
第一节 单因子或单向分组的方差分析法	4-1
第二节 多组比较的 LSD 方法	4-6
第三节 k 个用药组与一个对照组比较	4-7
习 题	4-11
第五章 χ^2 检 验	5-1
第一节 四格列联表(2×2 表)	5-1
第二节 四格表的配对检验	5-4
第三节 四格表的精确概率计算法	5-6
第四节 R×C 列联表	5-8
习 题	5-10
第六章 方差分析	6-1
第一节 随机化区组设计的方差分析	6-1

第二节	拉丁方设计的方差分析	6 - 4
第三节	平衡不完全区组设计的方差分析	6 - 7
第四节	析因设计的方差分析	6 - 12
第五节	重复数不等的两因子方差分析	6 - 16
	习 题	6 - 19
第七章	两指标间的联系——直线回归与相关	7 - 1
第一节	一个直线回归的模型	7 - 2
第二节	一个应变量有重复的直线回归的模型	7 - 5
第三节	直线回归的方差分析	7 - 8
第四节	相关系数	7 - 13
第五节	相关系数的可信限及两相关系数的差别检验	7 - 17
	习 题	7 - 19
第八章	非参数统计方法	8 - 1
第一节	符号检验法	8 - 1
第二节	符号等级检验	8 - 3
第三节	两组比较的等级检验	8 - 4
第四节	多组比较的等级检验	8 - 5
第五节	等级单位分析法	8 - 7
	习 题	8 - 10
第九章	半数效量法	9 - 1
第一节	求 $L.D_{50}$ 的克贝尔法	9 - 3

第二节	求 LD_{50} 的加权近似法	9 - 5
第三节	求 LD_{50} 的序贯法	9 - 9
第四节	用半数致死量测药物的相对毒力	9 - 13
第五节	药物的安全度指标——治疗指数	9 - 16
	习 题	9 - 19
第十章 生物检定		10 - 1
第一节	直接检定	10 - 2
第二节	属于量反应的平行线检定	10 - 5
第三节	属于质反应的平行线检定	10 - 14
	习 题	10 - 21

生物统计讲义

第一章 总 论

第一节 生物统计的目的

医学科学的研究对象是生物，而变异 (variation) 是生物的固有特征，在相同条件下生活的人，身体的发育程度及生理指标的测量值可不相同，在相同条件下的动物实验，其结果可有很大的差别。

这种变异的存在，常使人们不能对所看到的现象作出肯定的回答，但实践证明尽管存在变异，大量的同类现象却蕴藏着必然规律性。

生物统计学是数理统计在生物科学方面的应用，它是研究大量生物变异中的必然规律性的一门科学。譬如在药物研究中，尽管影响药物作用的因素繁多，但在一定条件下，药物作用总是遵循着一定规律。适当剂量的肾上腺素不论对人或动物都可引起一定的血压升高，这是因为各个体间虽存有很大差别，但它们的心血管系统对肾上腺素的反应基本上是相同的。一个升压药能使血压升高的反应是属于本质和基本的现象，也称必然现象，而使本质和基本现象出现复杂化并改变其程度（偶然也会改变其性质）是属于非本质和次要的现象，也称偶然现象。医学研究的目的是要从试验中获得本质和基本现象的规律，即必然规律性，生物统计学能帮助试验者了介和控制影响必然规律性的偶然现象，也就是可借助生物统计学在实验的安排与设计上，预先设法减轻偶然因素的纠缠，使实验经济有效，减少浪费，并以最科学的方法对实验结果作出应有的评价。

利用生物统计学可对取得的实验数据进行综合分析，它是结果判断的有力工具。但试验者必须善于运用这个工具，才能使研究工作收到事半功倍的效果，这里应注意以下几点。

试验者运用生物统计应铭记统计方法与演运过程来自数理统计的概念与原理，学习生物统计的同时，要求基本熟悉这些概念和原理，否则易犯机械套用公式的毛病，当套用错误时，必将导致判断的

错误。

2. 应用生物统计的同时，不可忘记专业知识的重要性，生物统计只是对可靠的实验数据用数理统计的方法加以分析，对试验结果的判断起到一个辅助工具的作用，但最终还必须结合专业知识进行全面的综合分析才能得出一个科学的结论。

3. 生物统计的应用从来不是在实验结束后仅仅对取得的实验资料进行处理，而是在实验开始前，就应在生物统计的理论指导下经行周密设计，这样可使试验消耗最少的人、物力及时间以获得最多的资料。反之，一堆未经设计的实验数据，实验本身不合理、资料不正确，则任凭什么“高深”的生物统计也决不能从中找出正确的答案来。

4. 生物统计有时需用冗长的繁复计算。为此，试验者常会偏面地追求一些简单的或粗糙的方法代替，甚至为了减少计算工作，不惜放弃一些有用资料。须知应用生物统计方法对最终的实验数据所消耗的演算劳动和用于整个实验劳动相比，毕竟只占很小部分。试验者应把生物统计的计算看作是整个实验过程的一个重要环节，要充分利用所有有用资料，不厌其烦地将资料分析与演算工作完成得正确无误。

第二章 概率与概率分布

要说明什么是概率 (probability)，先对事件 (event) 的概念作出解释。事件可分为必然事件，不可能事件与随机事件。事件发生的可能性用一个 0 到 1 的数来度量。这个可能性的度量称该事件发生的概率。通常用字母 P 表示，这里用一个简单的例子说明。

人要衰老——必然事件；

使老人变为小孩——不可能事件；

某抗癌药能治愈癌症——随机事件。

必然事件就是通常人们讲的 100% 发生的事件。把这类事件发生的概率定义为 $P = 1$ ；不可能事件也就是通常人们讲的 0% 不发生的事件，或者讲发生的可能性是 0%，把这类事件发生的概率定义为 $P = 0$ ；随机事件是可能发生也可能不发生的事件，这类事件发生

概率自然应定义为 0 与 1 之间，它的大小应根据事件发生的可能性大小来决定。

在医学研究中，观察到的生物现象一般皆为随机事件，对随机事件发生的概率这里不拟象数理统计或概率论一样对它下严格的定义，而直观地描述如下：

一个随机事件在 n 次试验中出现 r 次，不出现 $n - r$ 次则当 n 充分大时，这个随机事件出现的概率可估计为

$$p = \frac{r}{n} \quad (1-1)$$

$$\text{不出现的概率估计为 } q = 1 - p = \frac{n-r}{n} \quad (1-2)$$

由公式 (1-1) 与 (1-2) 可知随机事件的概率估计与通常所用的百分率是一致的。譬如某事件发生率为 8.0%，也就是这个事件发生的概率估计为 $p=0.8$ ，不发生的概率估计为 0.2，由此可知一个事件发生与不发生的概率和恒等于 1，且公式 (1-1) 与 (1-2) 中的 p 与 q 都不会大于 1，也不会是负数。特别指出的是一个随机事件出现的概率应是一个固定的常数。但由公式 (1-1) 与 (1-2) 所估计的概率或百分率却不会是常数，试验次数 n 愈小，估计概率 p 与 q 的变异愈大，可以肯定随着 n 的增大，估计概率或百分率接近真正概率（常数）的可能性也愈大。例如以某药一定剂量注射于 1000 只小鼠，结果有 800 只死亡，我们可以说死亡率

$$p = \frac{r}{n} = \frac{800}{1000} = 80\% = 0.8$$
。这个 0.8 就是该药在应用这个剂量时使小鼠死亡的概率估计。由于 $n = 1000$ 很大，可以认为 $p = 0.8$ 接近真正的概率的可能性很大。简言之，将 $p = 0.8$ 作为小鼠的死亡概率不致有很大出入。反之，当 n 较小，譬如 $n=10$, $r=8$ ，同样可得出 $p = 0.8$ ，但由于 n 较小，这个 0.8 与真正概率有较大差别的可能性就很大，因此我们就很少有把握认为将它作为真正概率的估计是可靠的。这里对充分大的 n ($= 1000$)，作出概率的估计

p ($= 0.8$) 有以下 2 种实际意义：

第一，以同样剂量给予一只小鼠，该鼠死亡的可能性为 80%；

第二，试验者对用药后的一群小鼠，预估有 80% 死亡。

与概率相联系的为概率分布 (distribution of probability) 或简称分布。对分布的涵义可作以下说明：

某药以剂量 x 给予一生物系统，产生某种反应的概率(y)应是剂量(x)的函数，可用 $y = f(x)$ 表示，它也称为发生某反应的概率 y 在剂量 x 上的分布函数。譬如用洋地黄静脉滴注于猫，观察其产生心跳停止的总剂量。可以设想一定有一个剂量可使猫产生心跳停止的概率最大。假定这个剂量为 $0.65 \sim 0.66 \text{ mg/kg}$ ，也就是大多数的猫会在这个剂量范围内使心跳停止，小于 0.65 mg/kg 和大于 0.66 mg/kg 剂量的洋地黄使猫心跳停止的概率就较小，剂量距 $0.65 \sim 0.66 \text{ mg/kg}$ 越远，使猫心跳停止的概率也越小。图 1-1 a 代表猫发生心跳停止反应的概率分布曲线。如果要用实验方法得出这样的分布曲线，可将剂量 x 分成连续区间，用猫的死亡百分率来估计概率，于是可得出如图 1-1 a 中用虚线表示的直方图，它与分布曲线是近似的。可以设想，如果将剂量区间缩小，试验的动物数增多，直方图就会越来越接近分布曲线。

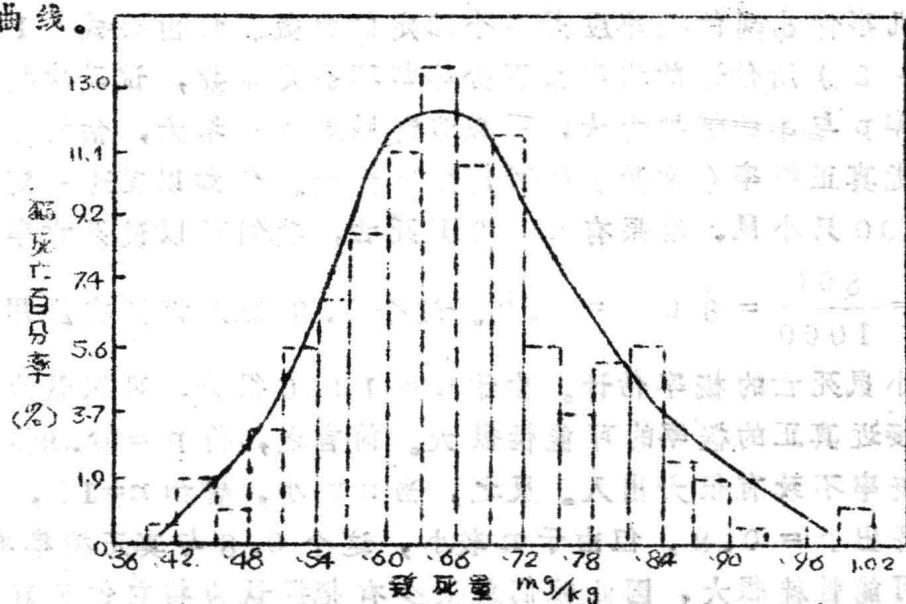


图 1-1a 静脉注射洋地黄即使猫产生心跳停止的概率分布曲线
(猫总数 216 只)

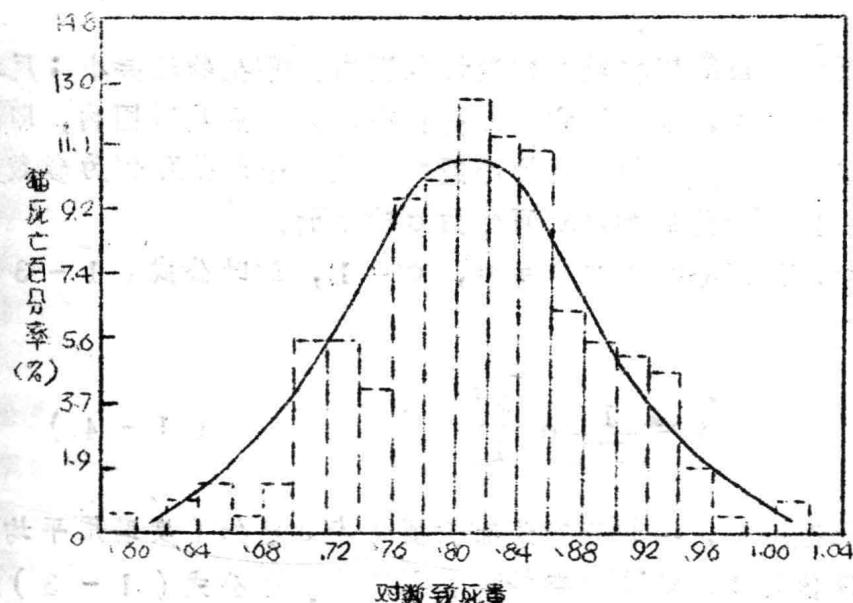


图 1-1b 将剂量转换对数后的概率分布曲线(猫总数 216 只)
(仿 Bliss, CI: J Amer Pharmaceut Asso., 33, 225, 1944.)

图 1-1a 的概率分布曲线左右不对称, 右侧拖着一个对药物具有耐受性的尾巴。如果将剂量转换为对数, 则呈左右对称的倒立钟状分布曲线, 这就是在统计学占极重要地位的正态分布 (normal distribution)。在医学研究中, 很多测量指标是正态或近似正态分布。例如正常成年人的身高, 血压, 血中钙、钠、钾等的测定等。读数或称量误差也服从正态分布。生物差异更是普遍地接近正态分布或对数正态分布 (lognormal distribution), 即图中将剂量转换对数后才是正态的分布。正态分布的曲线可用以下公式:

$$y = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (1-3)$$

μ 与 σ 为正态分布的两个参数, μ 代表分布曲线的高峰在横轴上的坐标, 因此它标志分布中心的位置; σ 为分布曲线的“腰”在横轴上距离 μ 的大小, 大的 σ 标志分布曲线峰矮而底宽; 小的 σ 标志分布曲线峰高而底窄。如果以生物差异来解释分布曲线, 图 (1-1b) 是以对数的最小死亡量作横坐标, 出现死亡的百分率作纵坐标, 分布曲线峰

高底窄表示死亡反应集中在较小的剂量范围内，即生物差异小；反之，分布曲线峰矮而底宽，表示死亡反应分散在较大的剂量范围内，即生物差异大。服从正态分布的对数最小死亡剂量的平均数可作为参数 μ 的估计，而标志离散度的标准差可作为 σ 的估计。

一个特殊的正态分布是当 $\mu = 0$, $\sigma = 1$ ，这时公式(1-3)可化为：

$$y = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (1-4)$$

以上称标准正态分布，一切正态分布只要 μ 与 σ 已知（实验用平均数与标准差分别估计），就可由变换 $z = \frac{x - \mu}{\sigma}$ ，使公式(1-3)简化为公式(1-4)。公式(1-4)的函数值可列表查出。

在生物统计中，除去正态分布外，尚有很多其他类型的分布，如二项分布，普阿松分布等，这里不拟一一介绍。

第三节 样本与总体

科学试验的目的常以少量的试验结果推及全体，譬如以一定剂量的某降压药试验20只大鼠，平均血压下降20 mm Hg，试验者就认为该药用这个剂量可普遍使大鼠的血压下降，且下降程度在20 mm Hg左右。这里用20只大鼠的结果推及全体大鼠是一个非常重要的推理过程，它也称为统计推断(statistic reference)。试验是在少数大鼠身上做的，结论却要用到全体大鼠身上去，前者为样本(sample)，后者为总体(population)。既然要由样本推及总体，则所用的样本就必须代表总体。这在逻辑上是显而易见的，但在实行时却并不简单。首先，如果样本太少，就很难代表总体。譬如用少量大鼠作试验，使结论适用于全体大鼠必须包括不同性别、不同年龄，不同体重，不同饲养条件等，如果不限制种系，则须包括不同种系，这样就势必要加大样本，增加实验负担，若能使样本既代表总体，又要使样本不致过大，除应用一些必要的生物统计方法以外，对

总体实行某些限制是必要的。如限制大鼠的种系，体重，年龄甚至性别。但试验结果只能适用于限制后的总体，而决不能应用于非样本所代表的另一个总体上去。从雌性大鼠身上所获得的结论一般不宜推及雄性大鼠；从幼年大鼠所获得的结论一般不适用于老年大鼠。不仅如此，样本的分布形态、差异程度也必须与总体一致，否则所作出的统计推断可能会是错误的。

试验结果只适用于样本所代表的总体这一事实并不意味动物试验会失去实用价值，因为尽管人与动物的种属不同，然而人与动物的组织结构，解剖特征，心血管系统，神经系统以及生化、代谢等方面皆有一定的类似之处，特别是哺乳类动物与人较为接近，因此它们对外界的刺激如药物引起的反应与人同质的较多，不同质的较少。一般讲，动物试验的结果虽不能推及到人，然而是过渡到人体试验前的一个有价值的参考。

第四节 随机抽样与分组

试验者所关心的往往是一个试验组与另一试验组或几个试验组间的比较，这就要求分配到各试验组的受试对象应来自同一总体，否则试验组间的差别就可以由不同总体间的差别引起，这样的试验必定归于无效。例如试验长期应用雌性素类药物对小鼠的致癌作用，小鼠以给不同雌性素及不同剂量分组，这类试验要求试验时间较长，每组动物数较多，可以设想，如果各组动物的来源不同，如一组为本实验室繁殖，一组为外单位供应，一组来自外地，一组由采购员购买，如此等等，这样分组试验显然无法得到有价值的资料，试验最终必然徒劳。但怎样才能使各试验组来自同一总体？这里将涉及两个问题，第一是总体的对象，第二是对总体如何抽样。某医院要试验两种抗冠心病药物的效果对比，试验对象是以主诉有心绞痛且有心电图异常的冠心病病人，于是总体就是有这种疾病后来院求诊的病人；某实验室用热板法试验延胡索对小鼠的镇痛作用，动物选自本动物饲养所并规定体重为18—20克的雌性小鼠，于是总体就是本动物饲养所内所有符

合这类规格的小鼠。总体确定后，就是如何由总体抽样以保证各小组样本能代表同一总体，这就是生物统计中有关抽样的一个极重要原则，即随机抽样 (random sampling)。所谓随机抽样就是总体内每个个体被抽到的机会相等。于是总体内少的抽到的机会也少，多的抽到的机会也多。这样所抽得的样本最能代表总体，其间尚有的误差，仅由抽样造成称抽样误差 (sampling error)，这在统计分析中是可以被估计的。要做到随机抽样首先要无成见地由总体抽样及分组，但事实上仅做到这一点仍可能会破坏随机抽样，譬如刚才提到临水上要比较两种抗冠心病药物的效果，规定甲医师收到的病员皆应用A药治疗，乙医师收到的病员皆用B药治疗，这种分组法似乎没有主观偏见但由于病人对医生的挑选，医生给予病人的嘱咐，暗示以及态度的不同，可使原来由相同总体来的两组病员却不能代表同一总体。再又如大笼内随机抽取小鼠，如果试验者认为无成见地闭目“乱抽”就可达到随机抽样的目的，则往往会事与愿违，因为闭目“乱抽”一般先抽到的是不活泼或体重较大的小鼠，而后抽到的是较活泼或体重较轻的小鼠，这种虽排除了主观倾向却仍存在定向误差的抽样当然不是随机抽样。而样本也不会有代表性。在生物统计中，这样的样本称为偏性样本，造成偏性样本的因素称为偏因 (bias)。

生物统计学家曾介绍一些随机抽样技术，使不受主观和系统因素的影响以保证样本能代表总体。如常用的随机数字表等。

第五节 误差分类

误差是指实验观察指标与真值的偏离，这个偏离越大，则实验的观察结果与客观事实相差越远，偏离越小，则与客观事实相差越小，因此一切实验必须尽可能减少误差，使我们的实验结果能反映实际，否则容易得出错误结论。误差分随机误差 (random error) 与系统误差 (systematic error) 两类，随机误差往往不能为人力所控制，例如测量时必然会产生测量误差，动物对药物反应所表现的个体差异以及一些无法被人们所觉察的很多因素对实验观察值所造成

的微小波动都属于随机误差。但随机误差可以通过主观努力使其缩小，例如测量仪器的改进，实验动物的精心挑选，一切实验条件的控制都可使随机误差缩小到最低限度，而且这个最低限度的随机误差还可用生物统计方法从实验结果中“提取”出来。系统误差也称偏因，它是由于实验本身存在问题而造成一种定向的错误。例如一架未校正好的天平所称得药物的重量；一个对某药有主观偏见的医生来评价这个药物所得出的不正确结果；未用随机分配的原则进行分配的小组间的差别等。系统误差往往可使整个实验归于无效，且这种误差也无法用生物统计方法进行“提取”。因此任何实验必须避免系统误差。

现将实验中最易发生的误差列举如下：

(1) 测量误差 (error of measurement) 血球计数、目测比色、血压、呼吸、脉搏以及血液或尿液生化指标的测量往往可以由测量者受外界条件的影响(光线、视觉、听觉、情绪、音响等)而产生误差，测量误差服从标准正态分布，因此重复测量的平均值接近真值。

(2) 抽样误差 (sampling error) 每次抽样所获得的结果不会一样，例如从大笼内取 5 只大鼠的平均体重为 250 克，重复以上试验，平均体重不会都是 250 克，它们之间的差异称抽样误差，抽样误差一般无法避免，但可用两种手段缩小抽样误差：(a) 加大样本；(b) 随机抽样。

(3) 顺序误差 (error of sequence) 若试验总是按固定的顺序进行，则可产生与顺序有关的误差。例如有 A、B、C、D 4 种药物皆可使离体的大鼠子宫收缩，如果 4 个药物在同一子宫进行比较，若给药顺序固定为 A、B、C、D，则很可能子宫对药物敏感性的改变而使比较存在偏因。正确的方法应该是每条子宫接受药物的顺序是随机的，更好的按排应该是用拉丁方排列，这样可抵消顺序误差。又例如在慢性动物试验中，顺序的先后会涉及气候、季节、温度、动物的机体状况，对药物的敏感性以及动物本身对实验环境的习惯等条件而引起误差。因此，除控制各种条件外，用平行对照以消除这类误差。

(4) 非均匀误差 (error of unhomogeneity) 样本必须代表总体，否则就是非均匀样本，例如要测量正常大鼠的血压，如果都用幼年大鼠，测得血压就偏低；如果都为老年大鼠，测得的血压就偏高。这类由于非均匀样本所引起的误差称非均匀误差，补救的办法应在抽样时严格地应用随机抽样的原则。

(5) 条件误差 (conditional error) 由各次实验条件不同引起的差异，例如上、下午室温不同可影响痛阈的测定，激素的测量可由于季节不同引起差别；不同仪器的测量值可有很大不同，两个测量者测量同一指标可以有区别。因此必须控制实验条件一致，以保证不产生条件误差。

(6) 分配误差 (error of allocation) 由于不按随机分组，各组可能会存在分配误差。但可用一些人为方法使分配均匀，例如体重的分配，性别的分配等。

第六节 实验指标的选择

医学研究的实验指标必须满足以下几点。

1. 合理性：所谓合理性是指实验指标必须针对研究的问题。一切非针对研究问题的指标是不合理的，譬如研究降压药以肾型高血压犬的血压为指标；研究抗结核病药以豚鼠的肺结核病变作指标；测定注射药的致热质用家兔。体温作指标等，这些指标的合理性是显然的有些非直接指标由于易于测定，且在一定程度上能反映研究问题的看来也是合理的。例如象伤寒这样的传染病，体温的变化一般可标志病情的变化，因此用体温作指标就比较合理。但是以体温作为研究肺结核病的指标就不合理了，因为体温正常并不能说明结核病已痊愈，体温不正常也不能说明肺结核病一定比正常体温的来得严重。对于研究的问题既不能从质又不能从量的关系来说明的指标当然就不合理了。

2. 特异性：指标的特异性与合理性有关，一个特异性高的指标一定是合理的，但合理指标是否有高的特异性要视指标的性质而定。例如小鼠的耐缺氧试验可作为研究抗冠心病药物的指标，由于观察方

便，且可间接衡量小鼠用药后对心肌缺血的耐受力，因此是合理的，但特异性不高，因为对冠心病无效但有中枢抑制作用的药物同样可提高小鼠的耐缺氧能力。以体温作为伤寒病情的指标，虽合理但特异性也不高。若用血液及大便伤寒杆菌的培养作指标，特异性就大大提高了。

3. 正确性：指标的测量力求正确，一个不能正确测量的指标不仅误差大，且易引出错误的结果，指标的正确与否与指标的测定者，仪器的精度及指标的性质有关。客观的与计量的指标一般可正确测定，动物的血压可切开颈动脉用记录仪描记，血液中钠、钾、钙等的含量可应用火焰光度计测定，体内微量化学物质的分布可利用放射性同位素在盖革计数器测定。由仪器测定的指标一般要受到仪器的稳定性左右，仪器的灵敏度越高，受干扰的因素也越多，误差也大；仪器的灵敏度不高，测定的正确性就差。对微小的差别就不易辨清。

由主观作出判断的指标不易正确，测定痛的指标不可避免会受到被测者或测定者的心理、情绪等因素的影响。目测比色法不如用光电比色法，因前者会受到主观偏见的影响，而后者是客观的测量。

4. 重现性：一个合理、客观和正确的指标一定有很好的重现性。它可以经得起重复测定的考验。当然，重现性还会受到生物差异大小的影响，但这里指的是指标本身的性质而言。非客观测定的指标重现性较差，有时会达到不能使用的地步。例如用评分法来评定结扎犬冠状动脉后心肌梗塞区的大小，用分级制衡量动物行为等，除非经周密设计，控制主观因素，否则很难有较好的稳定性。

一个未经校正好的仪器或一个有偏爱的测定者所测定的指标，重现性虽好但却无用。

第七节 实验指标的分类

指标的表示可分为量反应 (quantitative response)，质反应 (qualitative response) 与时反应 (time response)，量反应是以连续量表示的指标，如血压以 mmHg，血糖浓度以 mg%，