

## 生物组织的形态计量研究中应特别注意的两个问题：参照空间的“陷阱”和样本含量问题

杨正伟  
(组胚教研室)

**关键词** 体视学 参照空间 样本含量

随着体视学的发展和推广，随着对形态计量的重要性的认识，应用体视学方法的形态计量研究逐渐增多。但是，其中相当部分的研究设计不完善，结果的分析不严谨，因而结论站不住脚，尽管不少的研究已发表在国际、国内的正式杂志上。之所以如此，常常是因为对体视学的理解不够，忽视了一些简单但重要的问题。本文探讨两个普遍存在的问题：参照空间的“陷阱”和样本含量问题，以提醒同行重视，提高计量研究的质量和值。

### 一、参照空间的“陷阱”

#### 1、问题的提出 试看以下三个有问题的例子：

**例 1：**某种实验条件下，主动脉壁横断面上中膜内平滑肌细胞核的面积分数和轮廓密度（即单位面积中膜内平滑肌细胞核的面积和数目），显著高于对照组，这提示中膜内平滑肌细胞（核）有增生。

**例 2：**某种实验条件下，肾脏肾小球内毛细血管的长度密度和体积分数（即单位体积肾小球内毛细血管的长度和体积）无显著变化，这提示肾小球内毛细血管的长度和体积无显著增减。

**例 3：**大鼠胰腺胰岛 B 细胞内的分泌颗粒的体积分数和数目密度（即单位体积 B 细胞内分泌颗粒的体积和数目），在白昼逐渐增高，黑夜中逐渐降低，这表明 B 细胞的分泌颗粒主要在白昼合成，而在黑夜逐渐排放。

以上三个例子可如下一般化：

某种组织结构（参照空间）内所测结构的密度值（如体积分数、表面积密度、长度密度、积分平均曲率密度、数目密度等）的升降，就意味着该种组织结构内所测结构的绝对值（即总体积、总表面积等）的升降。

**问题是：**上述结论合适吗？答案很明显：不能如上作结论！也许绝对值实际上无升降，也许绝对值的升降幅度实际上更大。由于不知道参照空间的体积，得不到绝对值，就不能作结论。这就是参照空间的“陷阱”。试看以下三个例子：

**例 4：**小鼠小肠上、下段的绒毛的表面积密度（绒毛表面积/肠壁体积）无显著性差异，但这并不说明小肠上、下段的绒毛同样丰富。事实上，上段小肠壁显著厚于下段，上段小肠绒毛的表面积与浆膜面的表面积之比（相当于二者的表面积密度之比）显著高于下段，这说明上段小肠的绒毛更丰富<sup>[1]</sup>。

**例 5：**生精小管内精原细胞核的数目密度，在 15 天龄的大鼠达高峰，70 天龄时下降了十倍，但这并不能说明生精小管内精原细胞核的多少。事实上，不同龄大鼠睾丸内生精小管的总体积不同；每个睾丸内精原细胞核的数目，在 15 至 30 天龄之间逐渐略有增加，以后在 30 至 70 天龄之间无显著变化<sup>[2]</sup>。

**例 6：**苯巴比妥对大鼠肝细胞质内的内质网的体积分数和表面积密度无显著影响，但这并不说明内质网的多少无显著变化。事实上，苯巴比妥使肝细胞质的总体积显著增加，因而细胞质内的内质网（的体积和表面积）亦显著增加<sup>[3]</sup>。

**2、问题的解决** 办法很简单：测量或估计参照空间的体积，将之与密度值相乘，获得绝对值。例 4、5、6 就这样做了。例 1、2、3 的问题可如下解决：

**例 1：**估计主动脉壁横断面上中膜的总面积，从而得中膜内细胞核轮廓的总面积和总数目，这较有比较价值。

**例 2：**估计肾小球的平均体积，从而得每个肾小球内毛细血管的平均总长和总体积。或者，估计肾小球轮廓的平均面积，从而估计每个肾小球轮廓内毛细血管断面的平均总数目和总面积。最好的办法是：估计整个肾脏内肾小球的总体积，由此得整个肾脏内肾小球毛细血管的总长和总体积。肾小球的总体积，可根据肾脏的体积和肾脏内肾小球的体积分数估计（二者相乘即得）。

**例 3：**估计 B 细胞的平均体积，从而得每个 B 细胞内分泌颗粒的平均总体积和总数目。或者，估计 B 细胞轮廓的平均面积，从而估计每个 B 细胞轮廓内分泌颗粒断面的平均总面积和总数目。最好的办法是：估计整个胰腺内胰岛 B 细胞的总体积，从而估计出整个胰腺内胰岛 B 细胞的分泌颗粒的总体积和总数目。胰腺内胰岛 B 细胞的总体积，可根据胰腺的体积、胰腺内胰岛的体积分数和胰岛内 B 细胞的体积分数估计（三者相乘即得）。

为获得参照空间的体积，时常需进行阶段抽样估计。电镜下的计量研究，一般需先进行光镜下的计量研究，见上述例 3 的解决办法。忽视光镜，一步“跳”到电镜，就很可能“跳”进参照空间的“陷阱”。

有时，参照空间的体积难以获得，如皮肤、神经、脊髓、脊髓灰质内板层等。有时，参照空间的体积无法知道，如病理检查中利用活检组织块的研究。有时，所测组织结构无明确的界限，如大脑皮质的分区，因此无法定义一个明确的参照空间。总而言之，若参照空间体积未知，人们就无法从密度值估计绝对值。密度值常无多大的实质性比较价值，估计密度值常是为了进一步估计绝对值。不过，人们可选择其它参照系或其它参数，使研究仍有比较价值。例如，上述例 1 的解决办法和例 4 就是这样的。又如，相同部位的神经的横断面可当作参考系，组织内的某些粒子的平均体积甚至平均轮廓面积亦可用作参照系（见上述例 2、3 的解决办法）。为反映上皮细胞微绒毛的丰富程度，可选择微绒毛的表面积与其“基底面”（即假想的各微绒毛基底部的连接面，其在切片上的迹线即微绒毛底端的人为连接线）的面积之比这个参数，比较例 4。或者，选择不必知道参照空间的体积的参数，如粒子的平均体积，薄膜的平均厚度，结构的表面积体积比，平均曲率等。或者，选

择不必知道参照空间的体积的体视学方法, 如分合法<sup>[4]</sup>。当然, 不知道参照空间的体积, 并不意味着就不能估计密度值。有时, 尤其是活检材料的病理检查时, 密度值仍有一定的甚至满意的价值。例如, 血管的长度密度就可定义为组织内的血管密度, 反映组织的血供情况。不过, 无论在什么情况下, 若能获得参照空间的体积和绝对估计值, 则研究更全面, 更有意义; 参照空间的大小本身也是一个基本的形态计量特征。

## 二、样本含量问题

假设一项研究中设立了实验组和对照组, 每组取 5 个器官(动物), 每个器官切 5 张切片, 每张切片摄 4 张显微照片。为统计结果并进行  $t$  检验, 一种不少见的做法是: 从每张照片得一个估计值, 从每组的 100 张照片得一个平均值和标准差(样本含量为 100), 由此进行实验组和对照组之间的统计检验。

**问题是:** 上述做法合适吗? 试看下面的一个例子:

**例 7:** 著者在定量研究点的空间分布中有这样一个内容: 利用 4 个成年男性小脑, 从每个小脑随机获取若干切片和视野, 借助描绘目镜描绘出蒲肯野细胞层的迹线及蒲肯野细胞体的轮廓, 测量各轮廓中心点至其两端的最近的一个轮廓中心点之间的迹线长度(即最小邻距)。结果: 从 4 个小脑共测量 247 个最小邻距, 其平均为  $105 \pm 70$  微米(平均  $\pm$  标准差, 以下同); 从 4 个小脑共测量 50 个视野, 各视野的平均最小邻距的平均为  $117 \pm 50$ ; 从各个小脑所测的最小邻距的平均分别为  $134 \pm 88$ 、 $111 \pm 69$ 、 $83 \pm 56$ 、 $94 \pm 59$ , 这 4 个均值的平均(即 4 个小脑之间的平均)为  $106 \pm 22$ 。

这是一个简单的统计学问题, 问题的答案很明确: 上述做法不合适。正确的做法是, 从各个器官的 20 张照片得相应的一个估计值, 从每个组的 5 个器官得一个平均值和标准差(样本含量为 5), 由此进行实验组和对照组之间的统计检验。例 7 中后一种统计结果的方法就是这样的。这样处理的理由如下:

把 100 当作样本含量, 相当于把每组的那 5 个器官(或其所有可能获取的无穷多张照片)当成总体, 因此样本均值只是那 5 个器官的真值的估计, 其统计检验只能说明实验组的那几个器官与对照组的那几个器官之间有无显著性差异。这种结果无实质性的意义, 因为人们通过实验研究想说明的是: 实验条件对那一类器官(动物)有无显著影响, 而不仅仅针对所研究的几个器官, 从而为那一研究领域提供参考资料。换句话讲, 人们不是为了研究那几个器官而研究那几个器官, 不是为研究样本而研究样本。

研究所要显示的, 是某种实验条件对各器官(整体)有无影响, 从而推论对那一类器官(动物)有无影响, 而不是看对切片(局部)有无影响。若的确要观察对器官的不同部位有无影响, 则需采用分层抽样, 把各个部位当作亚整体来研究。切片仅是器官的样本, 器官才是所研究的那一类器官(动物)的样本。如各个器官的照片数不同, 或照片(参照系)的大小不同, 或各照片内测量长度值的数目不同, 则所有照片的(平均)估计值的平均, 不等于各器官估计值之间的平均, 见例 7。事实上, 从每张照片得一个估计值也是不合适的, 从一个器官的所有照片, 才能得到该器官的一个估计值。

把所有照片当作一个样本, 把各照片当作样本的元素, 使样本含量变大, 人为降低误差(标准误), 从而会使统计检验得出有显著性差异的结论的机会偏大, 即得出错误统计结论的机会增加。

受社会经济现象的统计的影响, 人们常常希望利用大样本进行估计, 但生物组织器官的形态计量研究, 不像测量身高那样简便、迅速。通常情况下, 人们只抽选 5~10 个器官(动物)为一组进行形态计量研究。多数情况下, 这应当是足够的了, 但应注意动物品系、性别、年龄、体重甚至获取样本的时间的**选择**。理论上**讲**, 样本含量是否足够, 取决于总体的变异情况以及人们所希望获得的**精确度**。简而言之, 误差系数(标准误/均值)不大于 5% 的估计几乎可总被认为是足够精确、满意的; 有时, 误差系数稍大于 5% 也是可以接受的。同样条件下的同种组织结构的变异, 常常不是很大, 这应当归功于基因的调节。若事先不知道动物间的变异情况, 不妨先选用 5 个动物为一组, 理由之一是: 如所有 5 个动物中都发现**某种**参数增大或变小, 那么这种发现不符合事实的机会**可能**仅为  $(1/2)^5 < 5\%$ , 因此实验仍可得出该发现的结论<sup>[5]</sup>。倘若待测总体的某一参数的变异很大, 与其大大增加样本含量以得到满意的精确度, 也许倒不如另选一个变异较小的参数或者对个体重新分组。参数变异很大本身就说明这个参数不能很好地反映总体的平均情况。

### 参 考 文 献

1. 朱家媛 等. 小鼠小肠绒毛表面积的定量研究. 川北医学院学报 1990; 5(3): 1.
2. Yang Zhengwei et al. A quantitative study of spermatogenesis in the developing rat testis. Biol. Reprod. 1990; 43: 629.
3. Staubli W. et al. Correlated morphometric and biochemical studies on the liver cell. II. Effects of phenobarbital on rat hepatocytes. J. Cell Biol. 1969; 42: 92.
4. Gundersen H. J. G. Stereology of arbitrary particles, a review of some unbiased number and size estimators and the presentation of some new ones. J. Microsc. 1986; 143: 3.
5. Cruz-Orive L. M. et al. Recent stereological methods for cell biology: a brief survey. Am. J. Physiol. 1990; 258 (Lung Cell Mol. Physiol. 2): L148.

#### 2007 年 4 月本文作者注:

该 PDF 文稿是根据发表的论文原文重新录入、重新排版后制作的, 其中红色的字、字母和标点是根据原文含义修改或补充的。主要修改或补充的内容如下:

- 原文中的“陷阱”在该 PDF 稿中被更正为“**陷阱**”, “轮廓面”被改为“**轮廓**”, “核”被改为“**平滑肌细胞核**”, “特征物”被改为“**结构**”, “相对的浆膜面”被改为“**浆膜面的**”, “说明了”被改为“**说明**”, “曲细精管”被改为“**生精小管**”, “天龄至 30 天”被改为“**至 30 天龄**”, “总体积”被改为“**体积**”, “总体积”被更正为“**分泌颗粒的总体积**”, “组织块”被更正为“**组织**”, “的表面基底部”被改为“**基底部**”, “得取”被更正为“**取**”, “各”被改为“**每**”, “象”被更正为“**像**”, “协调”被改为“**选择**”, “严格地讲”被改为“**讲**”, “精确度(假设器官的抽样是单纯随机的)”被改为“**精确度**”, “某一种”被改为“**某种**”, “仅为”被改为“**可能仅为**”。
- “**所测**”、“**整个**”、“**根**”、“**个**”, “**取**”、“**切**”、“**摄**”、“**组**”为补充的字。