

第五章

统计表与统计图



人民卫生出版社
PEOPLE'S MEDICAL PUBLISHING HOUSE

目录

第一节 统计表

第二节 统计图



重点难点

掌握

统计表和统计图的主要用途、统计表的基本制作原则、以及使用统计图表的注意事项。

熟悉

统计图的主要类型，如直方图、箱式图、线图、半对数线图、误差条图、散点图、条图和圆图等，以及各种图型的应用。



第一节

统计表



人民卫生出版社
PEOPLE'S MEDICAL PUBLISHING HOUSE



概念

- **统计表 (statistics table) : 是把统计资料和统计分析结果用表格的形式表达, 其目的是简洁、清晰和直观, 方便对比和阅读。**

优势

- **统计表用于数据的精确展示。**



一、统计表的编制原则和结构

(一) 统计表的编制原则

1. 重点突出，简单明了。一张表表达一个中心内容和一个主题。
2. 主谓分明，层次清楚。统计表虽然是表格的形式，但其内涵是若干完整的文字语句。
3. 数据表达规范、文字和线条尽量从简。



(二) 统计表的结构

- 从外形上看, 统计表可由**标题**、**标目** (包括横标目、纵标目)、**线条** (三线图)、**数字**和**备注**5部分构成。

Diagram illustrating the structure of a statistical table with callouts:

- 标题** (Title): 表1试验药与对照药治疗帕金森病临床疗效的有效率比较
- 横标目** (Horizontal Header): 组别
- 纵标目** (Vertical Headers): 例数, 有效例数, 无效例数, 有效率(%)

组别	例数	有效例数	无效例数	有效率(%)
对照组	103	17	86	16.51
试验组	105	74	31	70.48*
合计	208	91	117	43.75

备注: *代表具有显著差异



(二) 统计表的结构

1. 标题

它是统计表的总名称，放在表的上方中间位置，并编号，简明扼要地说明表的主要内容，包括时间、地点和研究内容，如xx年某医院xx病人数。

2. 标目

用来说明表内数据涵义。**横标目**位于表的左侧，说明各行数据的意义，**纵标目**位于表头右侧，说明每列数据的意义，**总标目**主要是对纵标目内容的概括，在需要时才设置。





3. 线条

目前一般采用三线表，表的顶线和底线把表的主要内容与标题分隔开，中间一条线把纵标目与数据分隔开。

4. 数字

用阿拉伯数字表示，位数对齐，小数位数一致。表内不留空格，无数字用“-”表示，缺失数字用“...”表示，最好以备注的形式说明。若数字是“0”，则填写“0”。

5. 备注

需要说明时可用“*”号标出，将说明文字写在表格的下面。



总标目
表2 某三个医院试验药与对照药治疗帕金森病临床疗效的有效率比较
总标目

医院	对照组 (安慰剂)			试验组		
	例数	有效例数	有效率(%)	例数	有效例数	有效率(%)
医院1	36	6	16.67	36	27	75.00
医院2	31	4	12.90	34	22	64.71
医院3	36	7	19.44	35	25	71.43
合计	103	17	16.50	105	74	70.48



二、统计表的种类

- **简单表**：简单表的标目只有一个层次，主语按一个标志排列，一般用作横标目，而纵标目为统计指标名称。

每一行可用一个完整的句子表达：

第一行：对照组试验的例数为103例，其中有效例数为17例，无效例数为86例，有效率为16.5%。

第二行：试验组的例数为105例，其中有效例数74例，无效例数为31例，有效率70.48%。

表3试验药与对照药治疗帕金森病临床疗效的有效率比较

组别	例数	有效例数	无效例数	有效率(%)
对照组	103	17	86	16.51
试验组	105	74	31	70.48*
合计	208	91	117	43.75



- **复合表：**复合表有两个或三个分组标志，即主语按多个标志排列。一般把其中主要的或分项较多的一个作为横标目，而其余的则安排在纵标目与总标目上。

表4 试验药物与安慰剂治疗帕金森病的多中心临床试验结果

医院	对照组 (安慰剂)			试验组		
	例数	有效例数	有效率(%)	例数	有效例数	有效率(%)
医院1	36	6	16.67	36	27	75.00
医院2	31	4	12.90	34	22	64.71
医院3	36	7	19.44	35	25	71.43
合计	103	17	16.50	105	74	70.48



不建议

表5 试验药物与安慰剂治疗帕金森病的多中心临床试验结果

医院	医院1			医院2			医院3		
	例数	有效例数	有效率 (%)	例数	有效例数	有效率 (%)	例数	有效例数	有效率 (%)
对照组 (安慰剂)	36	6	16.67	31	4	12.90	36	7	19.44
试验组	36	27	75.00	34	22	64.71	35	25	71.43



三、编制统计表的注意事项

- **统计表不一定是唯一的**，同一数据经过标目重排或分解组合，可以根据需要构造不同形式的统计表，表中标目位置的安排最不易把握，通常**将要比较的主体放在表格最左边**，右边则是相应的统计学指标和实验数据，左右两边位置不宜颠倒，但有时也要视具体情况决定。





表7 试验组和对照组血压的基线情况(mmHg)

分组	例数	舒张压		收缩压	
		$\bar{X} \pm S$	<i>M</i>	$\bar{X} \pm S$	<i>M</i>
试验组	61	98.64±4.73	98.0	150.56±13.27	149.0
对照组	62	98.31±5.29	96.0	148.23±12.43	146.0

表8 试验组和对照组血压的基线情况(mmHg)

指标	实验组 ($n_1=61$)		对照组 ($n_2=62$)	
	$\bar{X} \pm S$	<i>M</i>	$\bar{X} \pm S$	<i>M</i>
舒张压	98.64±4.73	98.0	98.31±5.29	96.0
收缩压	150.56±13.27	149.0	148.23±12.43	146.0



三、编制统计表的注意事项

- **避免内容混杂，表达不清，结构混乱。**若标目层次多于三个以上，统计表就会变得繁冗，因此尽量少用。如果需要表达多个内容，可以分成多个表进行描述。



例题

➤ **例5-1** 某地某年111例钩端螺旋体病患者发病季节、年龄和职业构成资料如下表所示，请按照编制统计表的基本要求，检查此表编制是否合适，如不合适，加以修改。

表9 流行病学有关的主要因素 (原表)

季 节							
8月上旬		8月中旬		8月下旬		9月上旬	
人数	%	人数	%	人数	%	人数	%
7	6.3	70	63.1	28	25.2	6	5.4
年 龄 (岁)							
15以下		15~		45以上			
人数	%	人数	%	人数	%		
12	10.8	95	85.6	4	3.6		
职 业							
农民		学生		待业青年		其他	
人数	%	人数	%	人数	%	人数	%
106	95.5	1	0.9	3	2.7	1	0.9



表10 某地某年钩端螺旋体病患者发病季节分布 (修改表)

季节	人数	构成比(%)
8月上旬	7	6.3
8月中旬	70	63.1
8月下旬	28	25.2
9月上旬	6	5.4
合计	111	100.0

表11 某地某年钩端螺旋体病患者发病年龄分布 (修改表)

年龄 (岁)	人数	构成比(%)
15以下	12	10.8
15~	95	85.6
45及以上	4	3.6
合计	111	100.0



表12 某地某年钩端螺旋体病患者发病职业分布 (修改表)

职业	人数	构成比(%)
农民	106	95.5
学生	1	0.9
待业青年	3	2.7
其他	1	0.9
合计	111	100.0



例题

- **例5-2** 某医院用麦芽根糖浆治疗急慢性肝炎患者161例，疗效资料如下表，指出其缺点并加以改进。

表13 麦芽根糖浆治疗急慢性肝炎疗效观察 (原表)

效果	有效						无效	
	小计		近期痊愈		好转			
	例	%	例	%	例	%	例	%
总例数	108	67.1	70	43.5	38	23.6	53	32.9



表14 某年某医院麦芽根糖浆治疗急慢性肝炎的疗效观察 (修改表)

效果 (疗效)	例数	构成比(%)
无效	53	32.9
好转	38	23.6
近期痊愈	70	43.5
合计	161	100.0

第二节

统计图



人民卫生出版社
PEOPLE'S MEDICAL PUBLISHING HOUSE



概念

- 统计图 (statistics chart) : 是把数据资料以图示的形式表达, 使数据对比更加形象、直观、一目了然。

优势

- 统计图利用**点**的位置、**曲线**的变化、**直条**的长短和**面积**的大小等几何图形来表达统计资料和统计指标, 将研究对象的特征生动形象地表达, 易于比较和理解研究结果。





一、统计图的制作

(一) 统计图的制作原则

1. 必须根据资料的性质、分析目的选用适当的统计图，由于统计图不能精确地显示数据大小，所以经常需要与统计表一起使用。
2. 一个图通常只表达一个中心内容和一个主题，即一个统计指标。
3. 绘制图形应注意准确、美观，图线粗细适当，定点准确，不同事物用不同线条（实线、虚线、点线）或颜色表示，给人以清晰的印象。



(二) 统计图的结构

包括标题、图域、标目、图例、刻度。

1. 标题

其作用是简明扼要地说明资料的内容、时间和地点，一般位于**图的下方中央位置并编号**，便于引用和说明。

2. 图域

即制图空间，除圆图外，一般用直角坐标系第一象限的位置表示图域，或者用长方形的框架表示。

3. 标目

分为纵标目和横标目，表示**纵轴和横轴数字刻度的意义**，一般有度量衡单位。





4. 图例

对图中不同颜色或图案代表的指标注释。图例通常放在横标目与标题之间，如果图域部分有较大空间，也可以放在图域中。

5. 刻度

即纵轴与横轴上的坐标。刻度数值按从小到大的顺序，纵轴由下向上，横轴由左向右。绘图时按照统计指标数值的大小，适当选择坐标原点和刻度的间隔。



(三) 统计图的种类

- 医学中常用的统计图有直条图、百分条图、圆图、普通线图、半对数线图、直方图、箱式图、误差条图和散点图等；
- 还有一些特殊的统计图，如表达多个试验中心研究结果的森林图、生存分析中的**生存曲线图**、聚类分析的树形图等。
- 一般应根据资料的性质和分析目的选择适当的图形。以下仅介绍医学中常用的统计图。

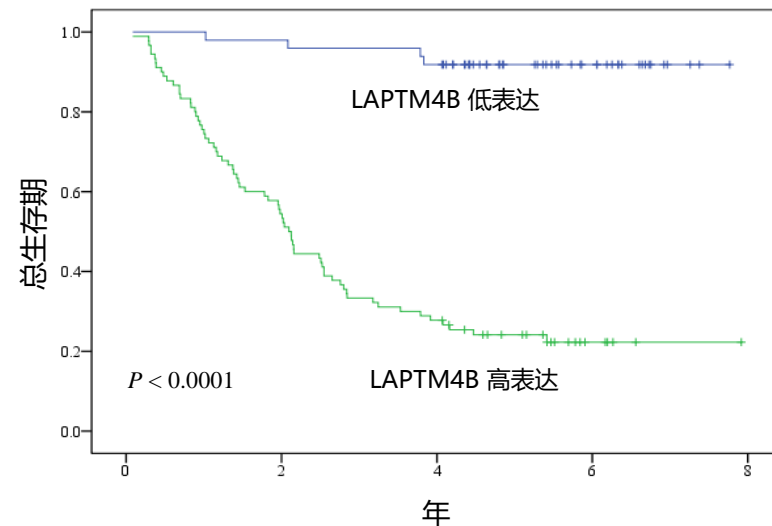


图1 LAPT M4B对卵巢癌化疗的影响



二、描述定量数据的统计图

(一) 直方图

- 直方图 (histogram) 用直条条段高度代表各组频数, 各条段高度总和代表频数的总和。它主要用于表示连续变量频数分布情况。



表1 广州市150名3岁女孩身高的频数分布

组段 (cm)	频数
80 ~	1
82 ~	3
84 ~	8
86 ~	10
88 ~	19
90 ~	23
92 ~	26
94 ~	24
96 ~	17
98 ~	10
100 ~	6
102 ~	2
104 ~ 106	1

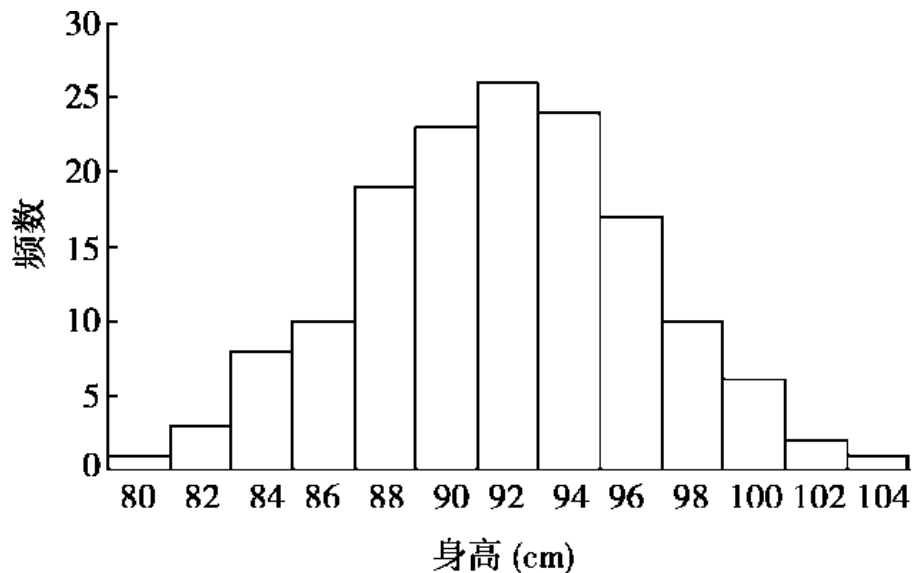


图2 广州市150名3岁女孩身高频数分布



(二) 线图

统计指标随时间变化的趋势

- ▶ **线图 (line chart)** 是用线段的升降来表示指标 (变量) 的连续变化情况, 适用于描述一个变量随另一个变量变化的趋势。通常纵坐标是统计指标, 横坐标是时间变量。普通线图的坐标都按算术尺度标记刻度, 若纵坐标按对数尺度 (通常采用常用对数), 横坐标按算术尺度, 制作的线图为**半对数线图 (semi-logarithmic linear chart)**。

--描述研究指标变化的速度

- ▶ 绘制线图时应注意: 横轴和纵轴的刻度都可以不从“0”开始; 用短线依次将相邻各点连接即得线图, 不应将折线描成光滑曲线。线图中若只有一条折线, 称为**单式线图**; 若有两条及以上的线条, 称为**复式线图**。在绘图时, 需要注意纵横轴比例, 比例不同, 给人的印象也不同。





表3 2002—2015年中国城乡地区急性心肌梗死死亡率 (1/10万)

年份	死亡率		年份	死亡率	
	农村	城市		农村	城市
2002	12.00	16.46	2009	44.95	43.14
2003	17.36	23.20	2010	43.19	38.09
2004	15.77	23.16	2011	48.53	47.36
2005	11.30	21.50	2012	39.32	42.50
2006	18.40	25.53	2013	66.62	51.46
2007	28.10	23.42	2014	68.60	55.32
2008	34.12	39.72	2015	70.09	56.38

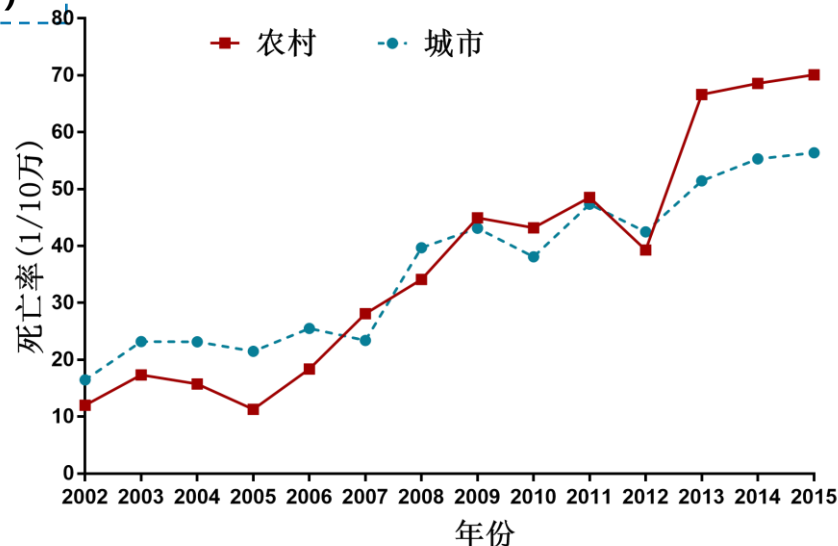


图4 2002—2015年中国城乡地区急性心肌梗死死亡率 (1/10万)



表4 某地2006—2016年细菌性痢疾、阿米巴性痢疾发病率 (1/10万)

年份	细菌性痢疾		阿米巴性痢疾	
	发病率	对数值	发病率	对数值
2006	3.95252	1.37435	2.8546	1.04893
2007	5.61003	1.72455	3.45233	1.23904
2008	10.7860	2.37825	4.2298	1.44216
2009	2.37845	0.86644	1.5135	0.41446
2010	3.434778	1.23395	1.6163	0.48017
2011	3.22061	1.16957	1.61030	0.47642
2012	2.40770	0.87867	1.00321	0.00320
2013	1.80036	0.58798	0.80016	-0.22294
2014	1.79461	0.58479	0.99700	-0.00299
2015	1.19426	0.1775	0.79617	-0.22793
2016	1.38531	0.32592	0.59370	-0.52136



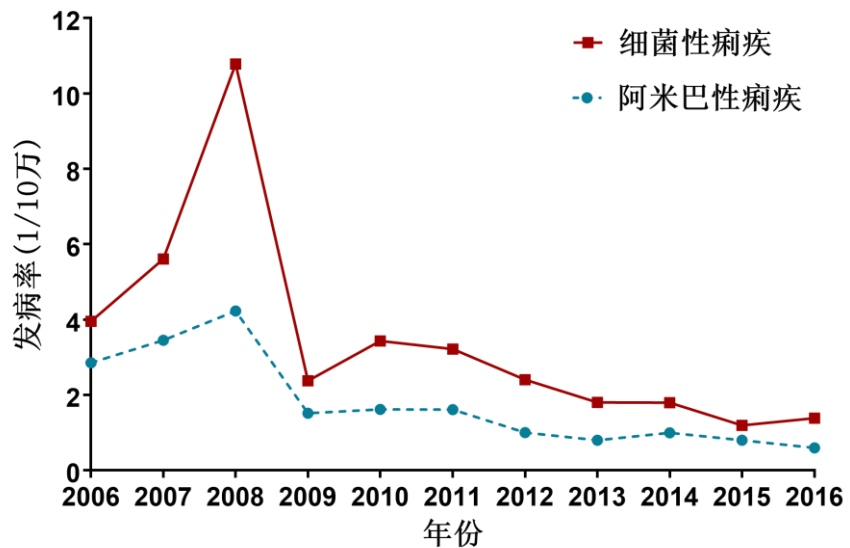


图5 某地2006—2016年细菌性痢疾和阿米巴性痢疾发病率线图

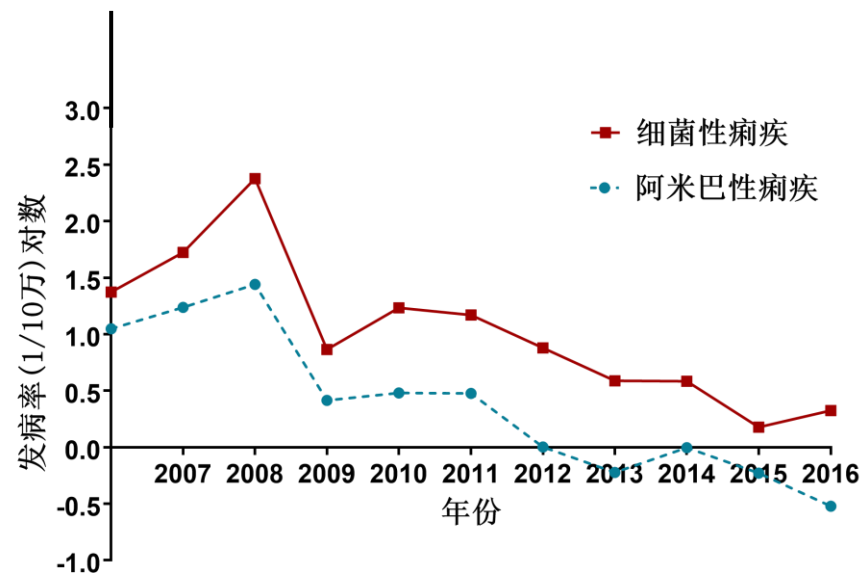


图6 某地2006—2016年细菌性痢疾和阿米巴性痢疾发病率半对数线图



(三) 箱式图

- **箱式图 (box plot)** 用于比较两组或多组资料的平均水平 (集中趋势) 和变异程度 (离散趋势), 各组数据均可呈现其**平均水平**、**四分位数间距**、**最小值**和**最大值**。
- 主要适用于描述偏态分布的资料。

表5 大白鼠正常肝和肝肿瘤中的递甲氨酶活性(CPM)

正常肝	肿瘤肝
19	227
30	339
43	130
70	592
64	405
91	104
35	211
68	346
15	133
6	814



正常	6	15	19	30	35	43	64	68	70	91
肿瘤	104	130	133	211	227	339	346	405	593	814

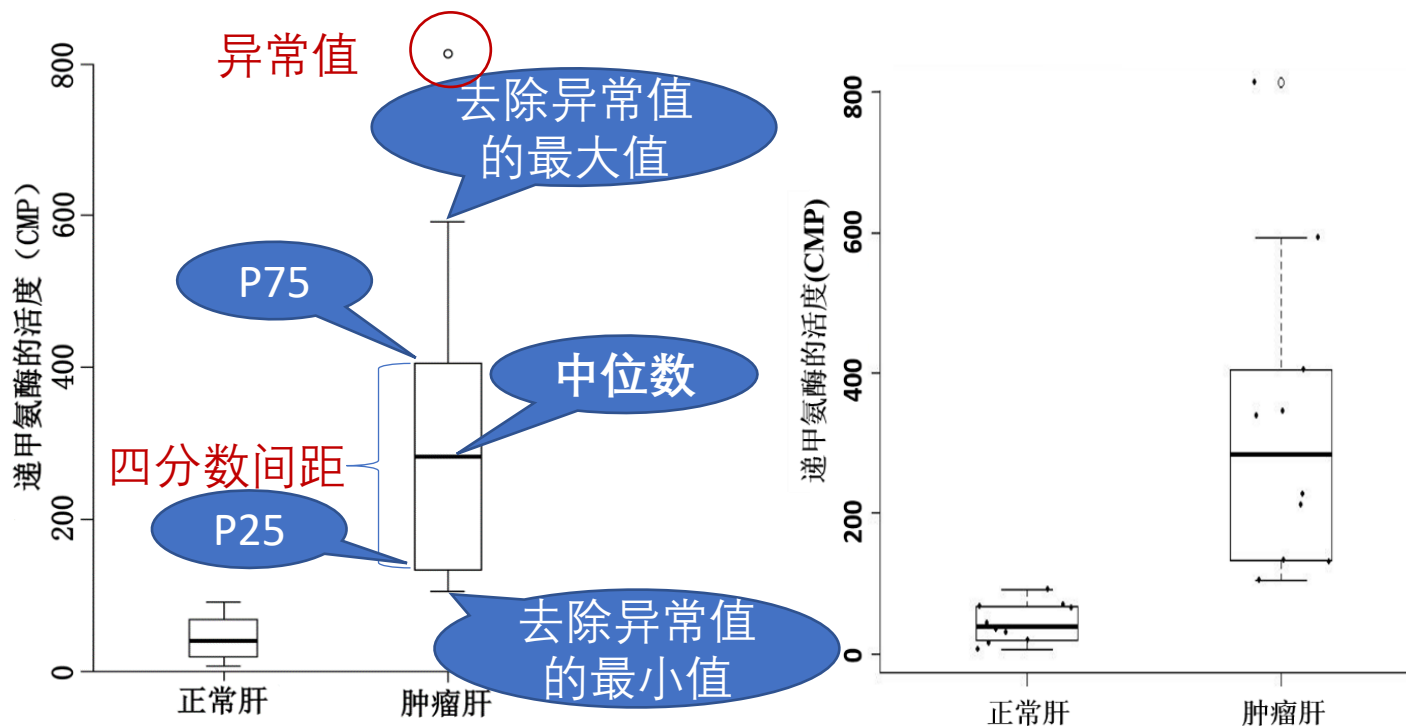


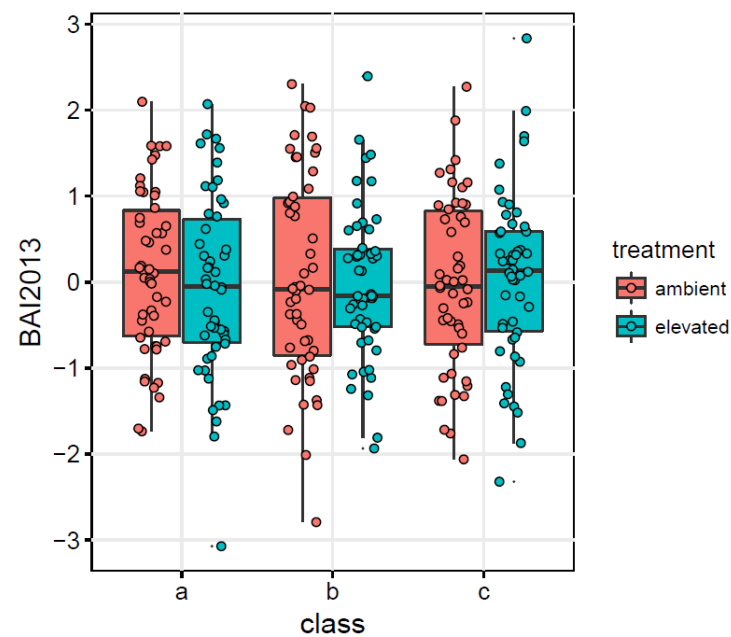
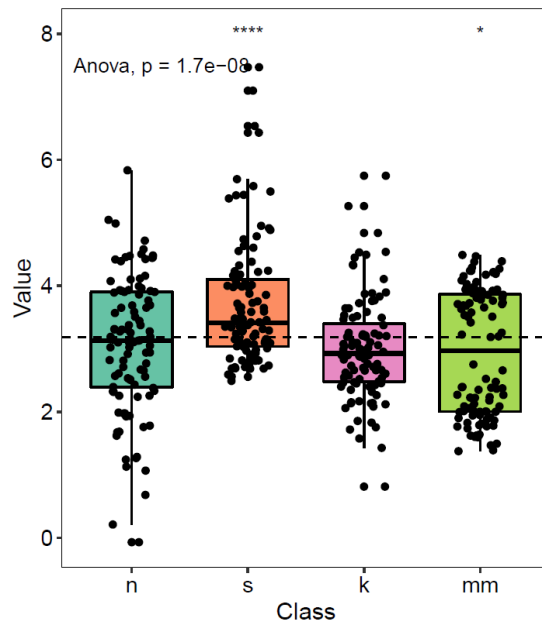
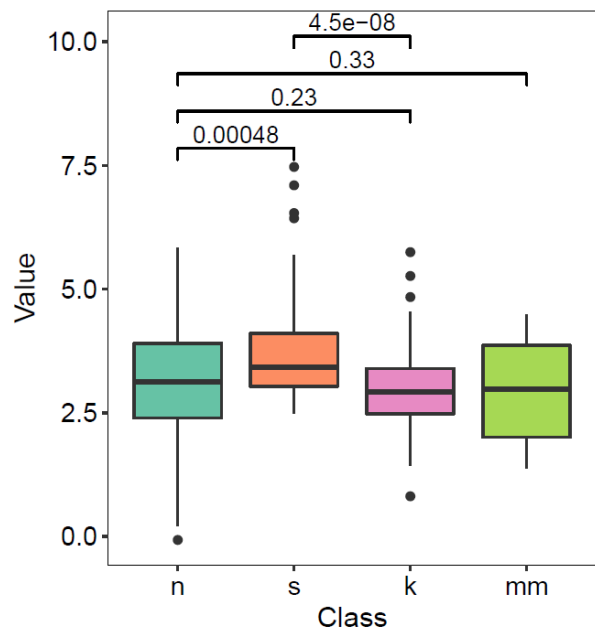
图7 大白鼠正常肝和肿瘤肝中递甲氨酶(CPM)的活度箱式图

- 图8 大白鼠正常肝和肿瘤肝中递甲氨酶(CPM)的活度散点图与箱式图



- 箱式图的中间横线表示中位数，箱体的长度表示四分位数间距，两端分别是 P_{75} 和 P_{25} 。最外面两端连线有两种表示方法：一种是表示**最大值和最小值**；另一种是**去除离群值后的最大值和最小值**，对离群值另作标记。
- **离群值 (异常值)**： 大于 $P_{75} + 1.5Q$ 或小于 $P_{25} - 1.5Q$
- **极端值**： 大于 $P_{75} + 3Q$ 或小于 $P_{25} - 3Q$
- 箱体越长表示数据离散程度越大；中间横线若在箱子中心位置，表示数据分布对称，中间横线偏离箱子正中心越远，表示数据分布越偏离中位数。
- 绘制箱式图时纵轴起点不一定从“0”开始。







(四) 误差条图

- 误差条图 (error bars) 用于比较多组资料的均值和标准差 (或标准误、可信区间)。
- 用线条的高度表示均值的大小, 可用“ \perp ”字部分表示均数的标准差或标准误, 也可用“工”表示可信区间, 上端“—”表示可信区间的上限, 下端“—”表示可信区间的下限, 中间“|”的长度表示可信区间的长度。
- 绘制误差条图时应注意: 如果以横轴为基线, 横轴表示各个分组或标志, 纵轴则表示各标志相应的均值和标准误 (或可信区间); 纵轴尺度必须从“0”开始, 表示均值的各直条的宽度要相等, 直条之间的间隔不必等距。



表6 试验组和对照组6周后血压的变化情况 (mmHg)

试验组		对照组		试验组		对照组	
舒张压	收缩压	舒张压	收缩压	舒张压	收缩压	舒张压	收缩压
12	17	10	10	8	20	7	10
12	8	-4	-7	-7	1	6	10
10	19	8	21	9	26	8	8
0	24	12	13	14	44	-4	6
16	9	10	-11	6	11	8	15
10	20	10	6	14	15	8	16
14	19	16	25	16	26	11	10
14	19	12	12	20	27	-4	3
11	14	10	7	9	10	11	25
2	8	2	12	2	9	9	5

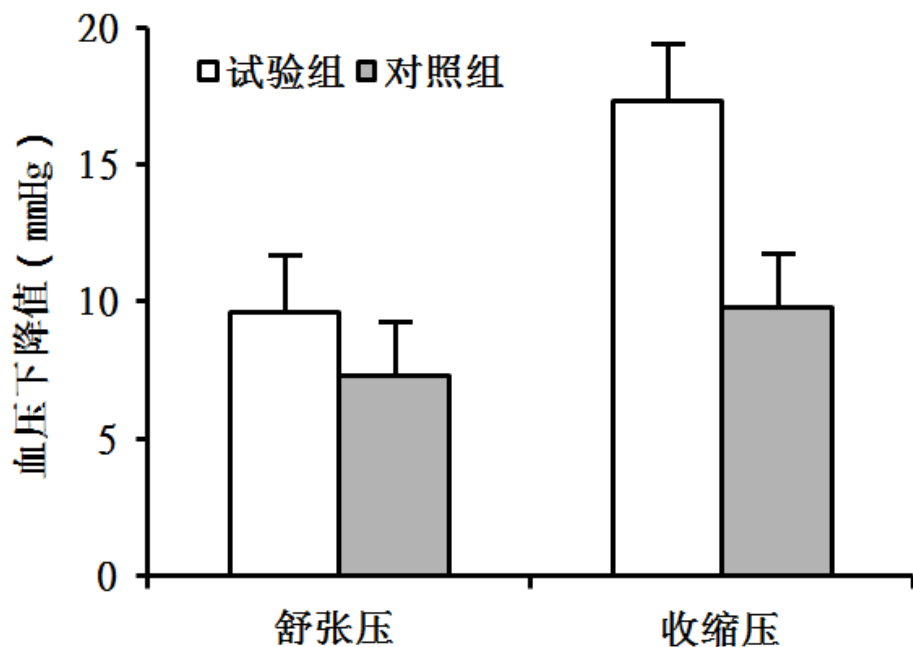


图9 试验组和对照组血压的平均变化情况(均数±标准误)

标准差

$$S = \sqrt{\frac{1}{n-1} \sum (x_i - \bar{x})^2}$$

标准误

$$S_{\bar{x}} = \frac{S}{\sqrt{n}}$$

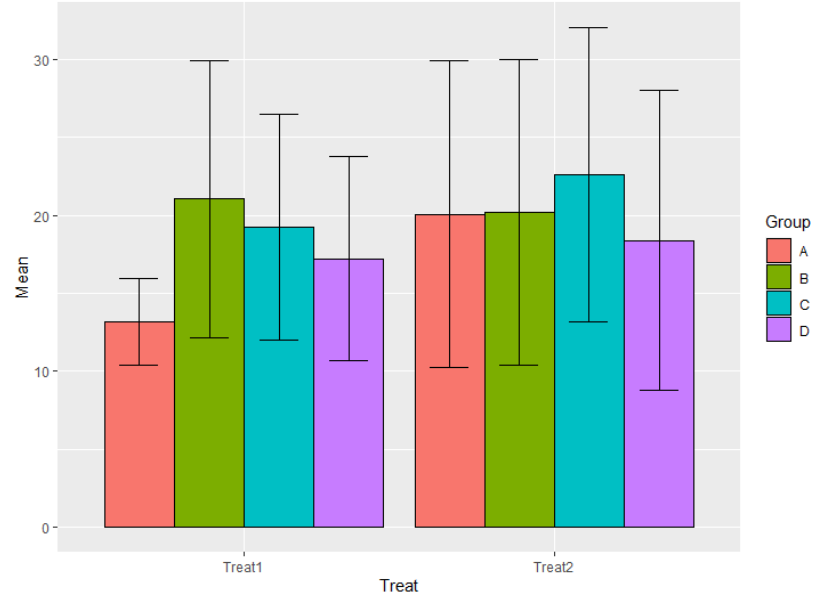
试验组的舒张压

$$\bar{X} = \frac{\sum X}{n} = 192 / 20 = 9.6$$

$$S = \sqrt{\frac{\sum (X - \bar{X})^2}{n-1}} = \sqrt{\frac{\sum X^2 - (\sum X)^2 / n}{n-1}}$$

$$= \sqrt{\frac{2624 - 1843.2}{19}} = \sqrt{41} = 6.4$$

$$SE = \frac{S}{\sqrt{n}} = 6.4 / 4.5 = 1.4$$





(五) 散点图

- 散点图 (scatter plot) : 用点的密集程度和变化趋势表示**两指标之间的直线或曲线关系**。
- 一般横轴代表**自变量**, 纵轴代表**因变量**, 纵轴和横轴的起点, 不一定从“0”开始。

表7 高血脂患者载脂蛋白B与低密度脂蛋白的关系

Apo-B(g/L)	LDL(mmol/L)	Apo-B(g/L)	LDL(mmol/L)
0.80	2.98	0.66	2.28
0.60	2.76	0.67	2.34
0.90	3.22	1.12	4.84
0.80	3.84	0.97	3.89
0.90	3.12	0.80	2.14
0.90	4.18	0.94	3.60
0.70	2.00	0.82	2.30
0.70	2.94	0.79	3.30
0.80	3.12	0.76	2.20
1.21	5.33	0.80	2.12

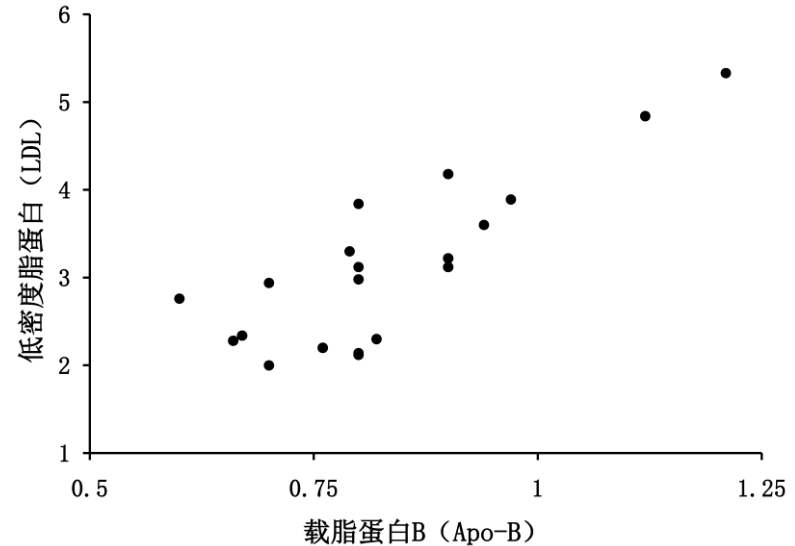
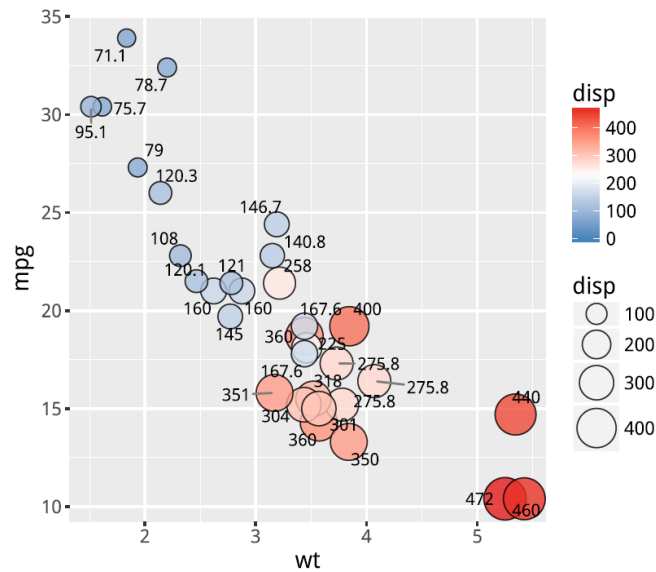
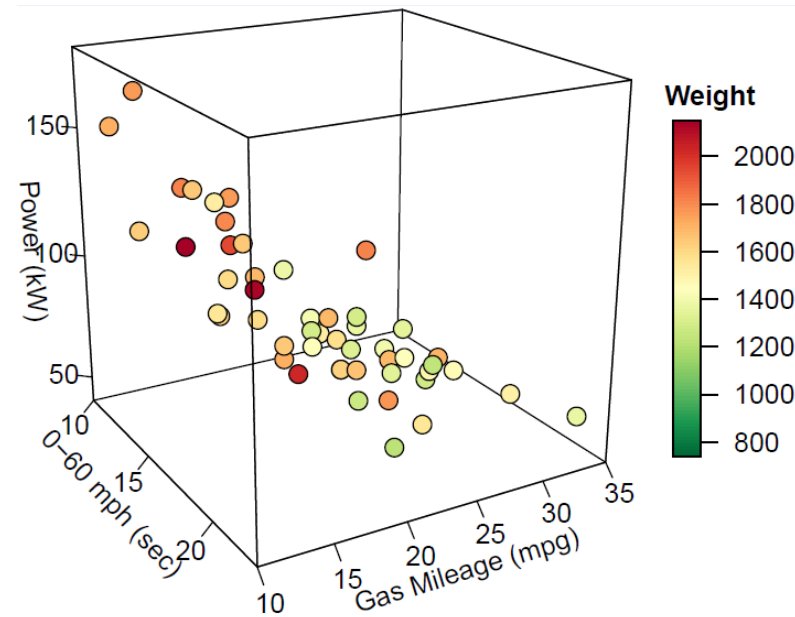
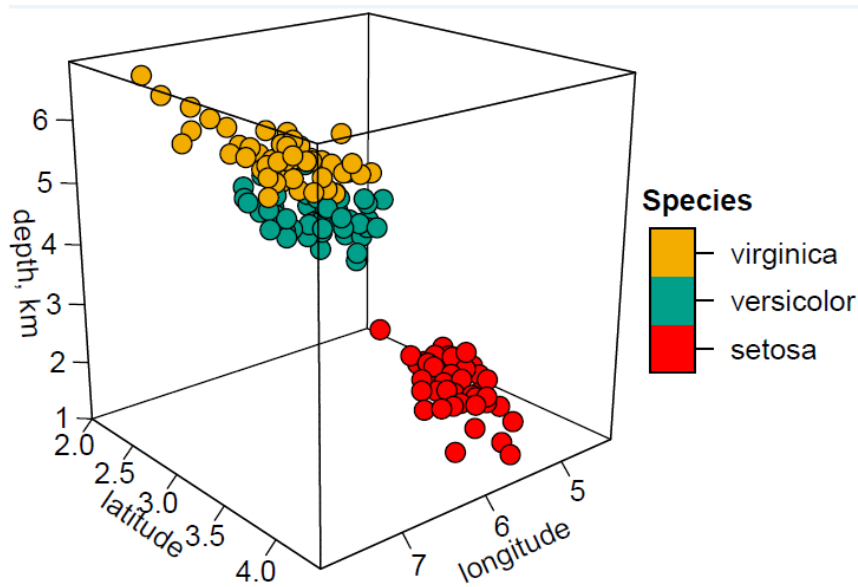


图10 载脂蛋白B与低密度脂蛋白关系的散点图



气泡图



三维散点图



(六) 热图

- **热图 (heat map)** 是用不同的颜色 (或者深浅) 表示观测值的大小, 生物信息学中也常用热图描述基因表达谱。
- 如图是乳腺癌5种分子分型的35个基因**聚类分析**的结果, 红色代表基因的表达量高, 蓝色代表基因的表达量低。该图揭示了乳腺癌5种分子分型不同的基因表达特点。

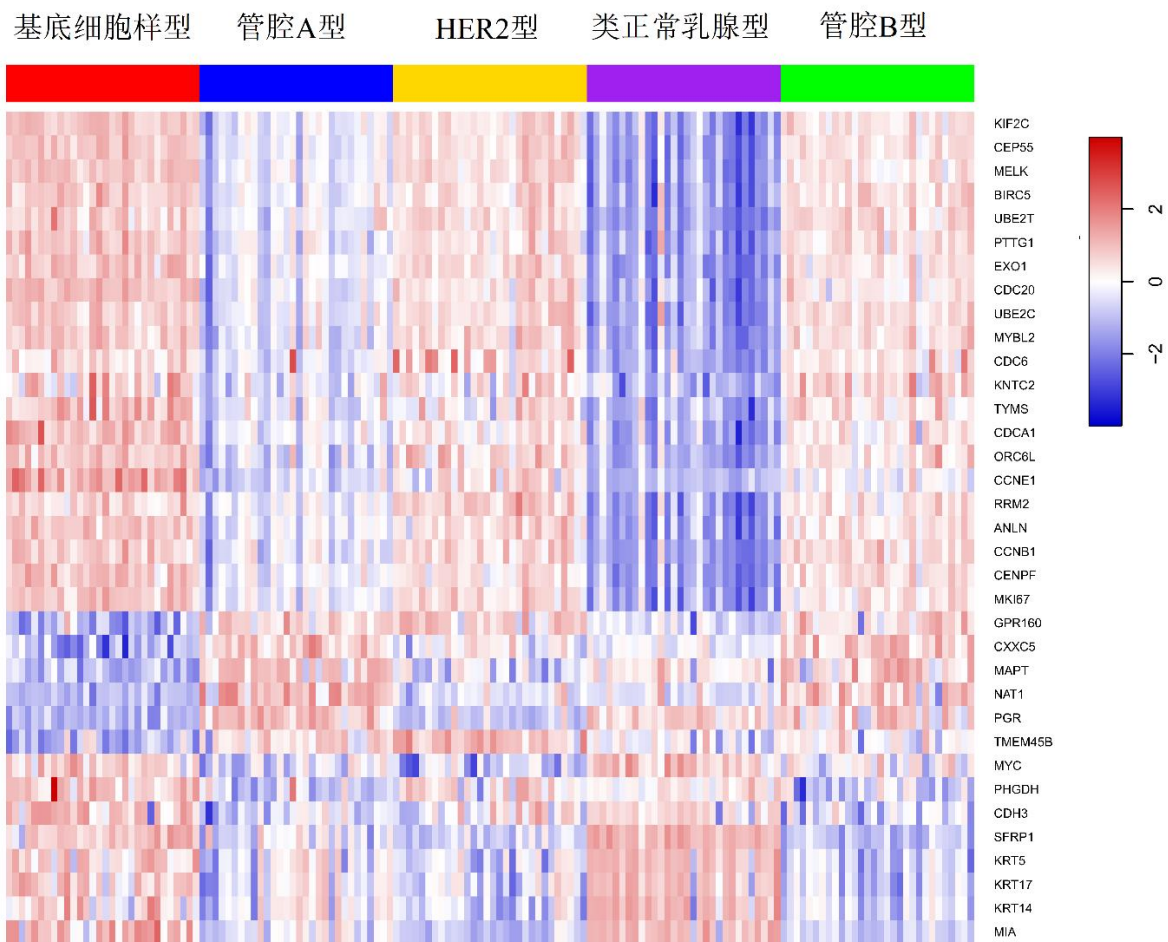
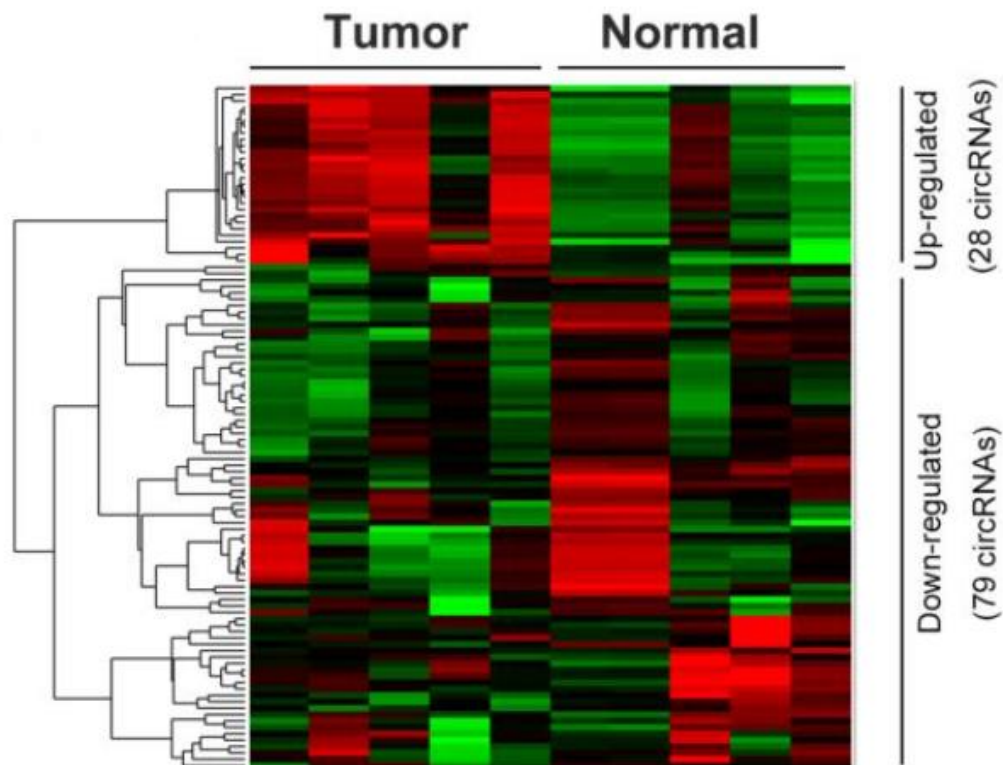
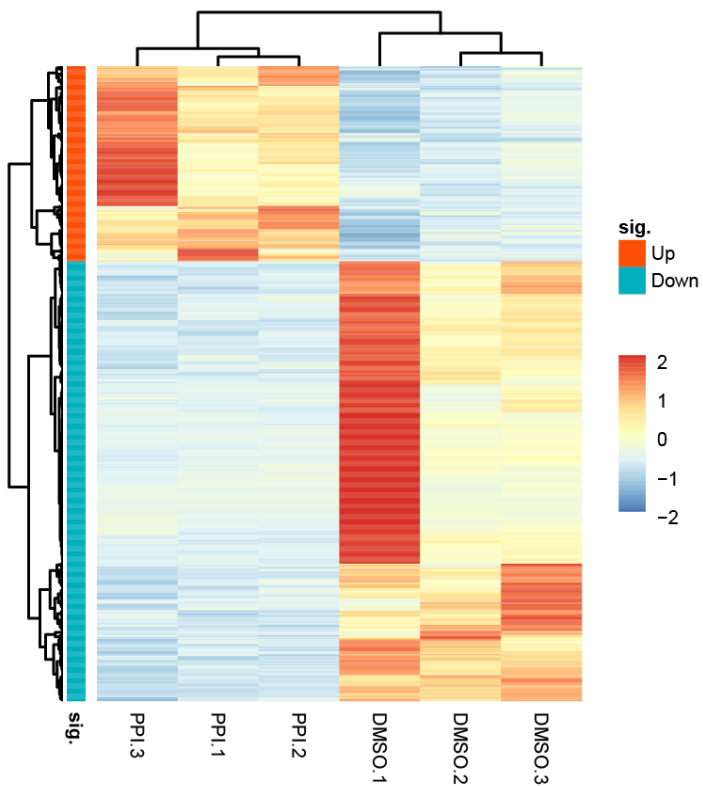


图11 不同乳腺癌分型的35个基因表达热图



热图



热图可以直观展示重点研究对象的差异变化情况。

对差异数据的表现体现在：

每个小方格表示每个基因，其颜色表示该基因表达量大小，表达量越大颜色越深（红色为上调，绿色为下调）。

每行表示每个基因在不同样本中的表达量情况，每列表示每个样品中所有基因的表达量情况。

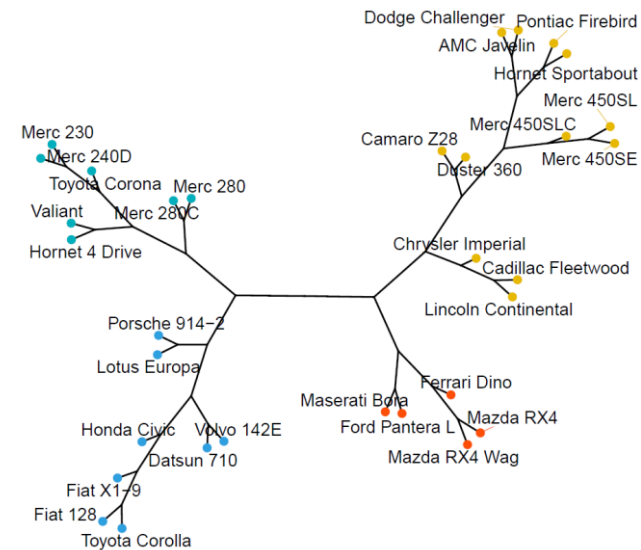
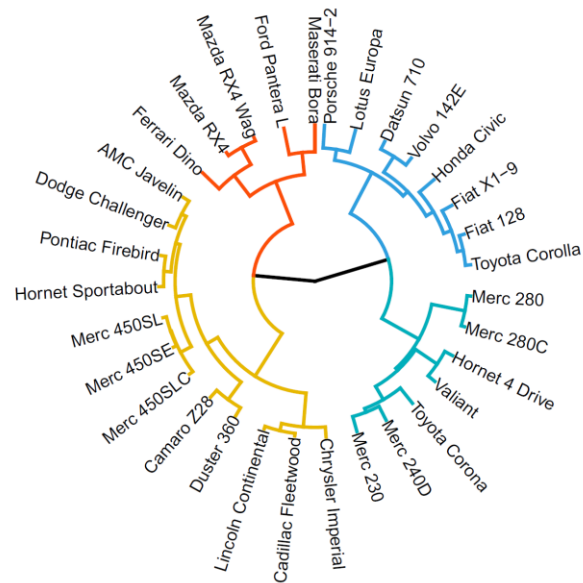
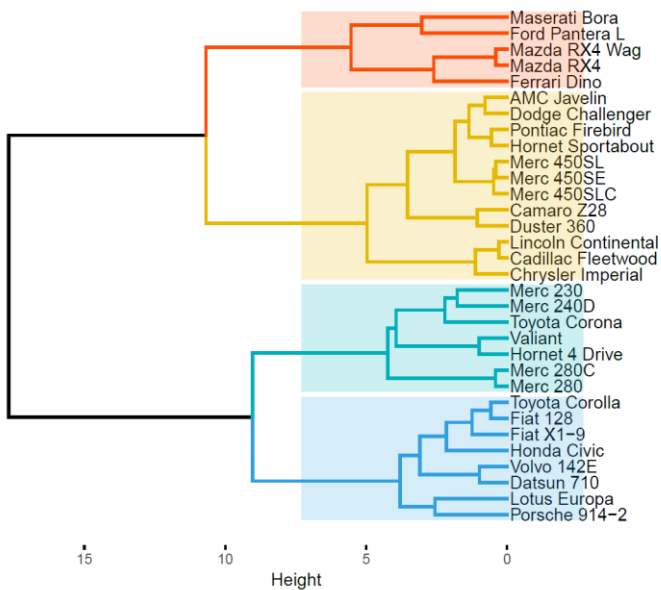
上方树形图表示对来自不同**样本**的聚类分析结果，左侧树状图表示不同**基因**的聚类分析结果。



对基因聚类：考虑不同基因，如基因1和基因2在每个样本中的表达模式一致，都高或都低，则基因1和基因2可聚为一类。

对样本聚类：考虑每个样本，如样本1和样本2，其中的基因表达情况较一致，如样本1中各基因的表达情况与样本2中的各基因表达情况一致，则样本1和样本2可聚为一类。

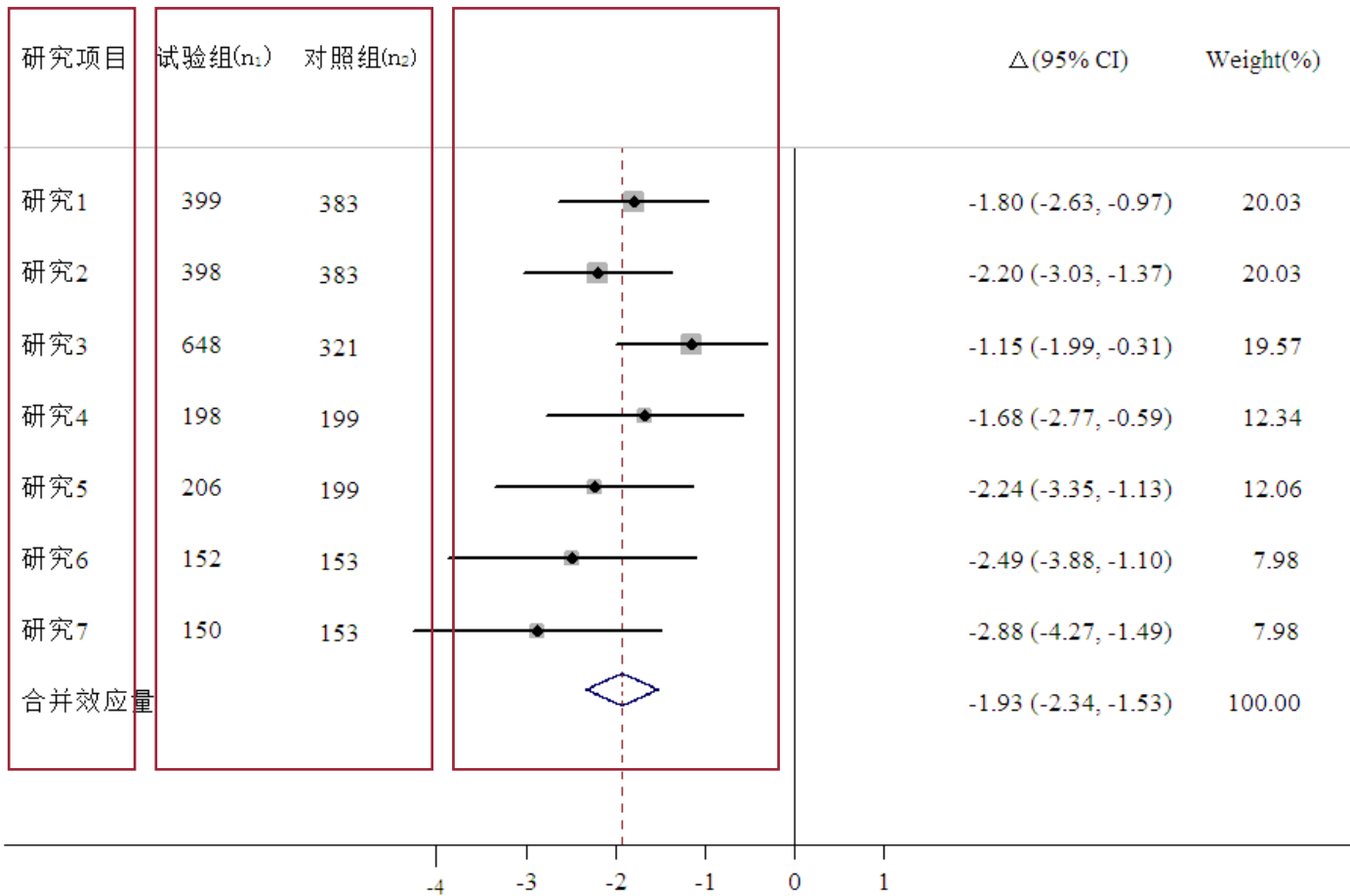
相关性分析 (第十一章)





(七) 森林图

- **森林图 (forest plot)** 是以统计指标和统计分析方法为基础, 用数值运算结果绘制出的图型。
- 它在平面直角坐标系中, 以一条垂直线为基础, 用平行于横轴的多条线段描述了每个被纳入研究的**效应量**和**可信区间**, 用一个**菱形** (或其他图性) 描述多个研究的**合并结果**。



简单来说，就是采用平面直角坐标系，以当中一条垂直的无效线(横坐标刻度为1或0)为中心，用平行于横轴的多条横线描述每个被纳入研究的效应量和可信区间，末尾采用菱形或其它图形描述多个研究合并的效应值和可信区间。

7项多奈哌齐 (Donepezil) 治疗阿尔兹海默 (AD) 研究文献的综合分析



权重

森林图的原理：根据各个研究结果的加权平均值(权重)，来估计某一项干预措施的效果，以提供可靠的证据。

这里，介绍下**权重**(weight)，它表示各个试验结果在总体结果中所占的百分比，一般病例数越多，权重越大，即：

大样本研究，由于效应估计值更为准确，**权重较大**；

小样本研究，由于估计值不够准确，**权重较小**。



森林图的图示介绍:

在平面直角坐标系中以一条垂直的无效线(横坐标刻度为1或0)为中心, 用平行于横轴的多条线段描述了每个被纳入研究的效应量和CI, 以及Meta分析合并效应量和CI;

用一个点(或其他图形)描述了每个研究的结果, 点的面积代表该研究在Meta分析中被赋予的权重, 同时有一根水平线向点的两端延伸代表CI(通常是95%CI); CI描述的是与研究结果相一致的疗效/相关性的可变范围, 表示单个研究的结果间是否有统计学差异;

较大的点意味着较大的权重, 即更窄的可信区间, 也决定了最终的计算结果。

简单来说, 就是采用平面直角坐标系, 以当中一条垂直的无效线(横坐标刻度为1或0)为中心, 用平行于横轴的多条横线描述每个被纳入研究的效应量和可信区间, 末尾采用菱形或其它图形描述多个研究合并的效应值和可信区间。



1. 二值变量的森林图

当研究对象为二值变量(如发生与不发生)时, 采用RR和OR作为统计学指标。

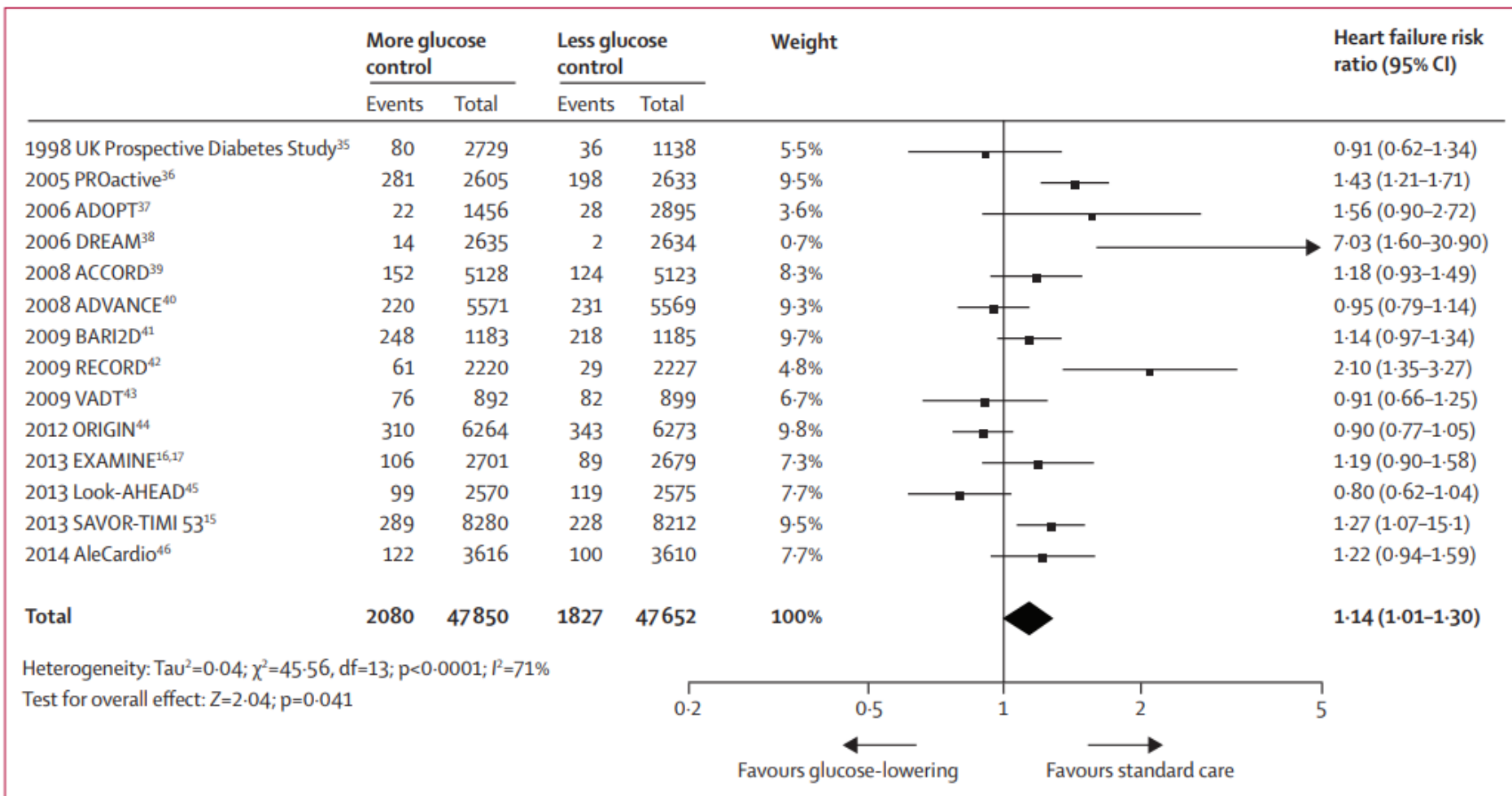
- ① 当某研究RR的95%CI包含了1, 即在森林图中其95%CI的横线与无效竖线相交, 可认为实验组发生率与对照组发生率相等, 试验因素无效。
- ② 以1为界限, 该研究RR的95%CI上下限若均 >1 , 即在森林图中横线落在无效线右侧, 可认为试验组的发生率大于对照组的发生率。若研究者所研究的事件是不利事件(如发病、患病、死亡等)时, 试验组的试验因素会增加该不利事件的发生, 试验因素为不利因素;若是有益事件(如有效、缓解、生存等)时, 试验因素会增加有益事件的发生, 试验因素为有益因素。
- ③ 反之亦然, 若研究的95%CI上下限若均 <1 , 即在森林图中横线落在无效线左侧, 可认为试验组的发生率小于对照组的发生率。若研究的是不利事件, 试验因素则为保护因素, 若研究的是有益事件, 试验因素为有害因素。



2. 连续性变量的森林图

当分析指标是连续变量, 也称数值变量时, 可选择加权均数差(WMD)或标准化均数差(SMD)为合并统计量。

- ① 当某研究的95%CI包含了0, 即在森林图中其95%CI横线与无效竖线(横坐标刻度为0)相交时, 可认为试验组某指标的均数与对照组相等, 试验因素无效。
- ② 当某研究的95%CI上下限均大于0, 即在森林图中, 其95%CI横线不与无效竖线相交, 且该横线落在无效线右侧时, 可认为试验组某指标的均数大于对照组, 若研究者所研究的指标是不利事件时, 试验因素为有害因素(危险因素);若研究者所研究的事件是有益事件时, 试验因素为有益因素。
- ③ 当某研究的95%CI上下限均小于0, 即在森林图中, 其95%CI横线不与无效竖线相交, 且该横线落在无效线左侧时, 可认为试验组某指标的均数小于对照组, 若研究者所研究的指标是不利事件时, 试验因素为有益因素(保护因素);若研究者所研究的事件是有益事件时, 试验因素为有害因素。





三、描述定性数据的统计图

(一) 直条图

- 直条图又称条图 (bar chart) , 用等宽直条的长短来表示相互独立的统计指标数值大小和它们间的对比关系。
- 直条图按直条是横放还是竖放分为卧式和立式两种, 按分组因素的多少分为单式和复式两种。
- 绘制直条图时应注意: 纵轴尺度必须从“0”开始且要等距; 直条的宽度要相等, 直条的间隔要等距, 间隔的距离通常与直条宽度相等或略小。



发病率

不同疾病

1. 单式条图：具有一个统计指标，一个分组因素。

表1我国大陆2010年和2013年四种虫媒传染病发病率(/千万人)

传染病	2010年	2013年
出血热	71.37	94.61
登革热	1.67	34.44
疟疾	55.36	28.77
乙脑	19.04	16.09

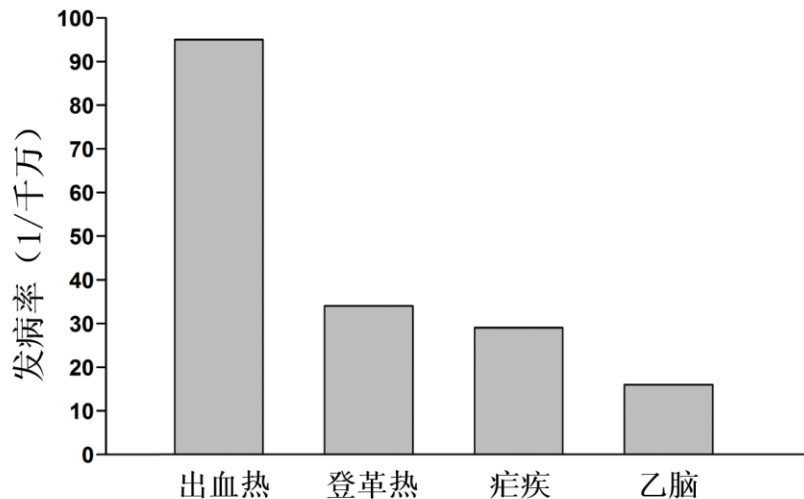


图1 2013年我国大陆四种虫媒传染病发病率(/千万人)



2. 复式条图: 具有一个统计指标, 两个分组因素。

发病率

不同疾病

年份

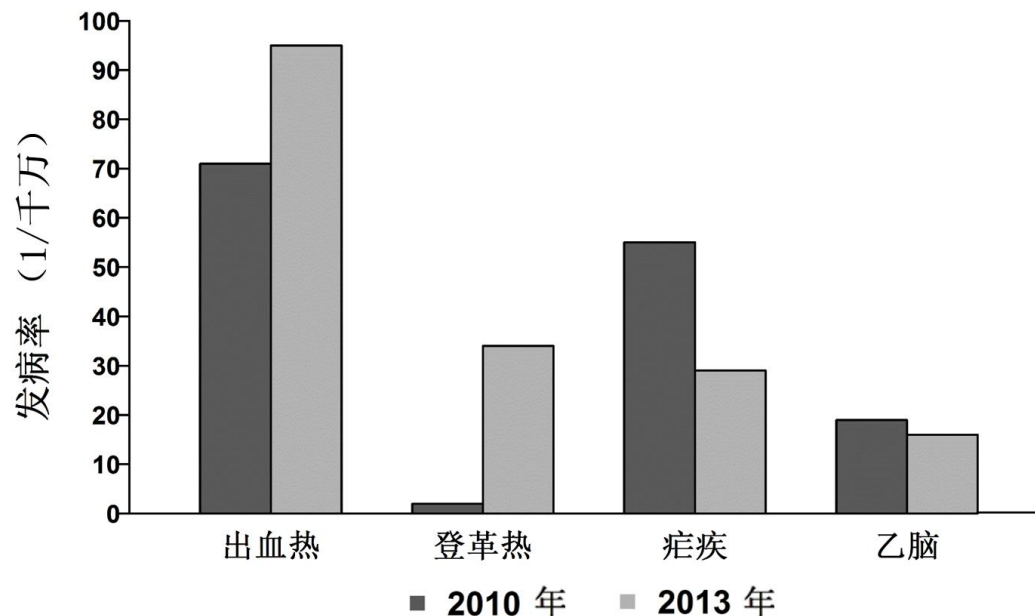


图2 2010年和2013年我国大陆四种虫媒传染病发病率(/千万人)



(二) 构成图

➤ 构成图常用于描述构成比资料。常用的构成图有：

➤ **圆图 (pie chart)**

圆图把圆的总面积作为100%，表示事物的全部，而圆内**各扇形面积**用来表示全体中**各部分所占的比例**。

➤ **百分条图 (percentage chart)**

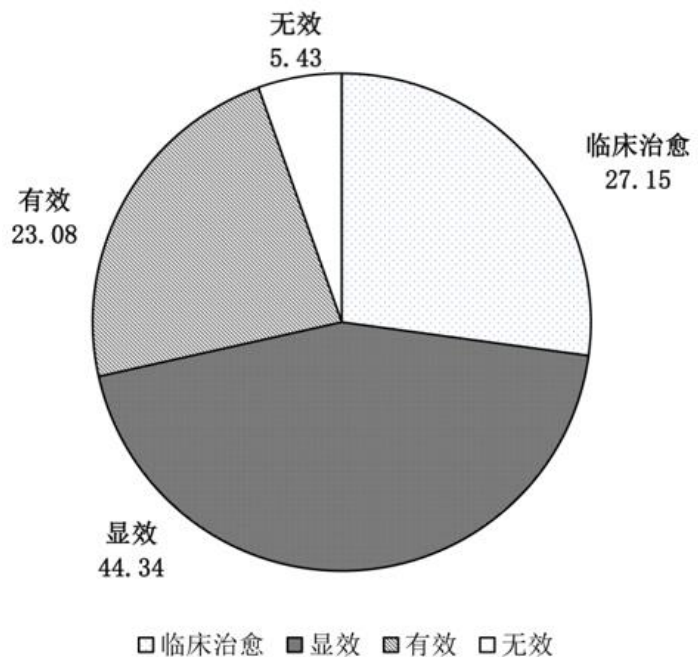
百分条图用矩形直条的长度表示100%，而用其中**分割的各段**表示各**构成部分的百分比**。



复方猪胆囊治疗老年性气管炎近期疗效比较

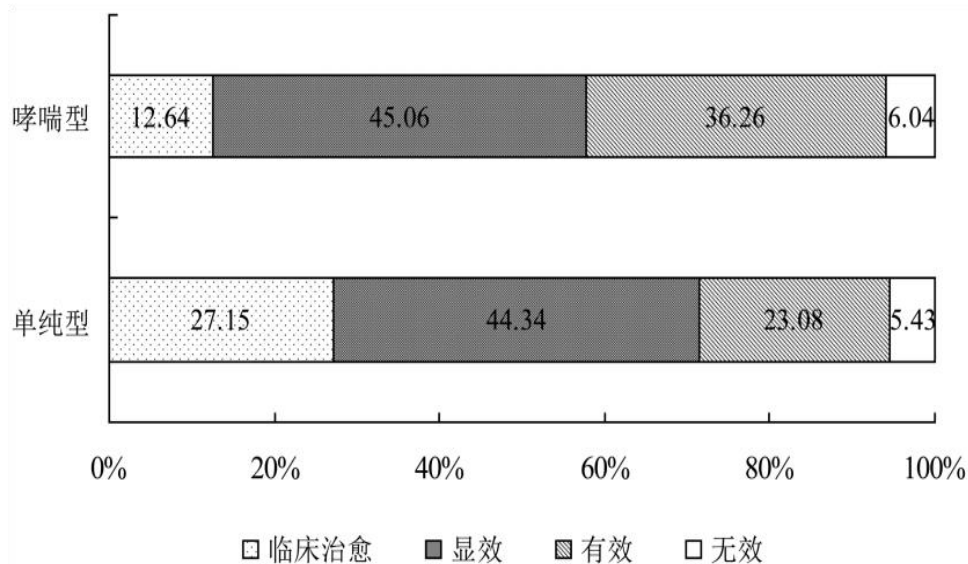
近期疗效	单纯型		哮喘型	
	例数	构成比(%)	例数	构成比(%)
临床治愈	60	27.15	23	12.64
显 效	98	44.34	82	45.06
有 效	51	23.08	66	36.26
无 效	12	5.43	11	6.04
合 计	221	100.00	182	100.00

圆图

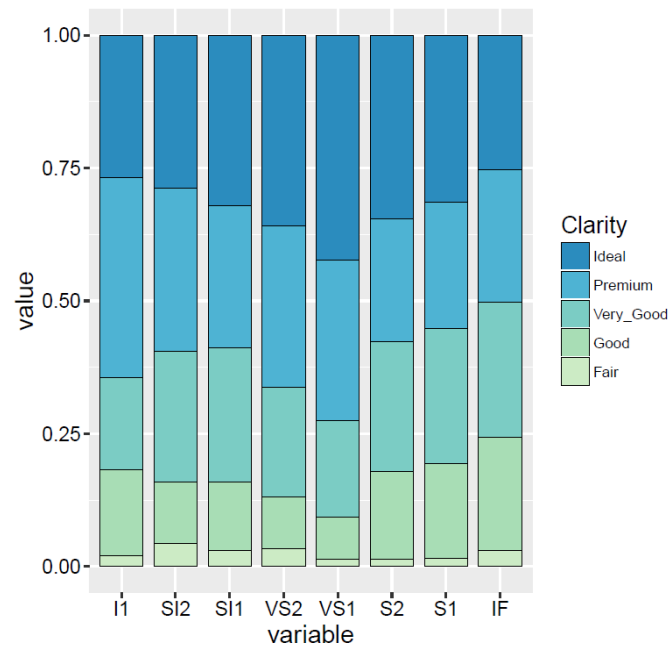
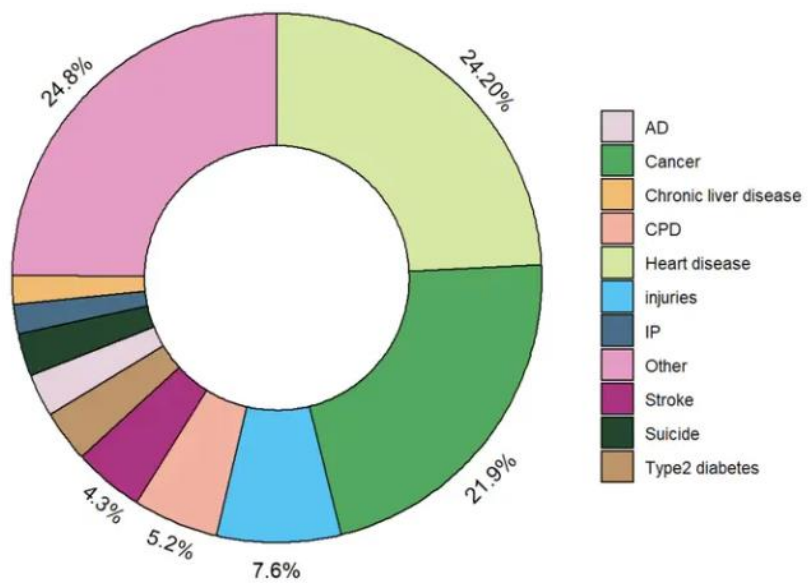


复方猪胆囊治疗单纯型老年气管炎近期疗效不同结果的构成比

百分条图



复方猪胆囊治疗老年性气管炎近期疗效比较





圆图绘制方法:

- 以圆形的 360° 角为100%，将各构成部分的构成比 (%) 乘以 3.6° ，即得各组成部分应占的圆心角度数；再以某刻度（比如相当于时钟12点的位置）为起点，顺时针按圆心角度数大小或者自然顺序排列各个扇形，不同扇形可以用不同颜色或花纹区别，也可简要注明文字和百分比。

百分条图绘制方法:

- 绘制一条等宽的水平直条，其长度为100%，根据各部分所占百分比，按其大小或资料的自然顺序把直条分出若干段，然后在直条的各分段上标出百分比。



本章小结

1.统计表和统计图是统计描述的重要工具，可清晰、直观地表达统计分析结果，方便分析对比和阅读。图表的运用对撰写论文的质量至关重要。

2.规范的统计表为三线表，在结构上包括5个部分，即标题、标目、线条、数字和备注；统计图通常由标题、图域、标目、图例和刻度5个部分组成。



本章小结

3.常用的统计图形有条图、百分条图、圆图、线图、半对数线图、直方图、箱式图、误差条图、散点图等，应根据资料的性质和分析目的，并考虑表达的效果选择适当的统计图形。

4.实际中也可以根据需要自己设计统计图，如图形和表组合在一起作为一个图给出。



课后习题二-1

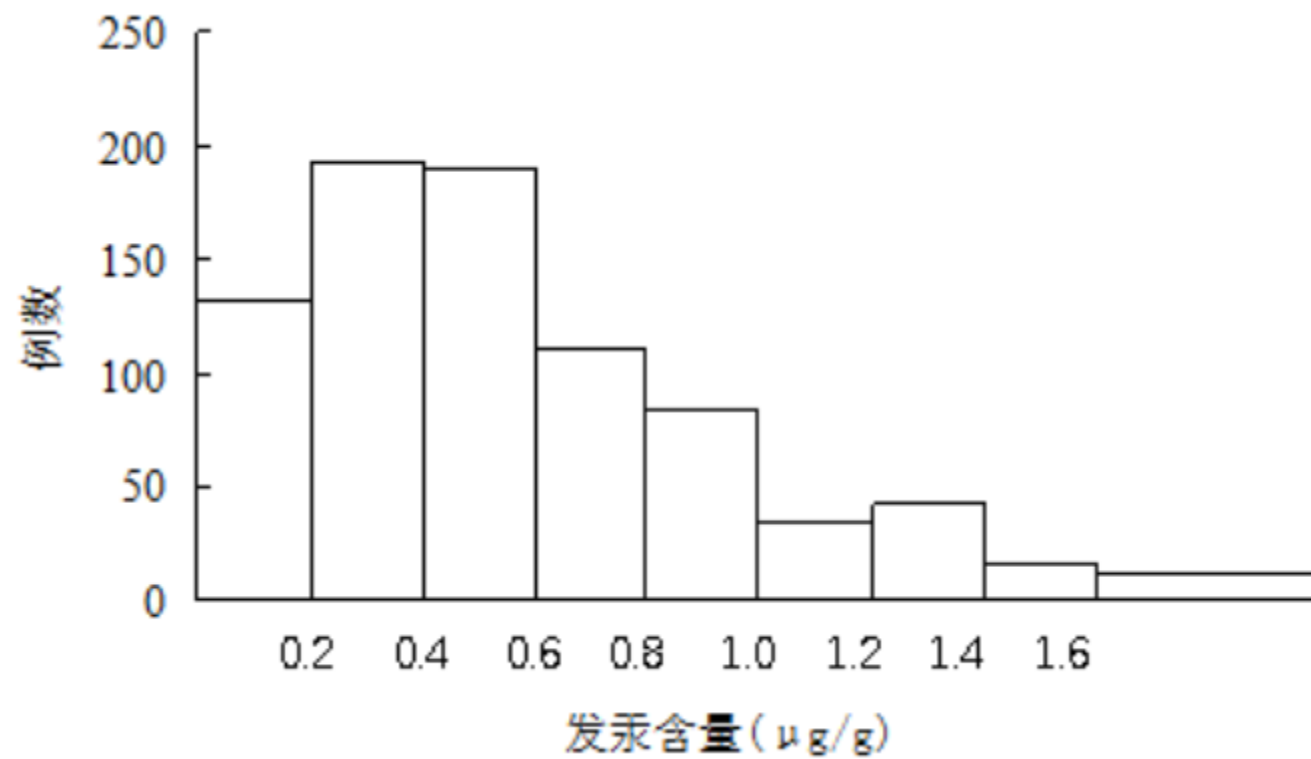
- 1、无标题
- 2、横标目顺序有问题，纵标目结构不一致
- 3、线条较多，小数位不统一

某地居民脾肿程度和疟疾血膜涂片检查结果的关系

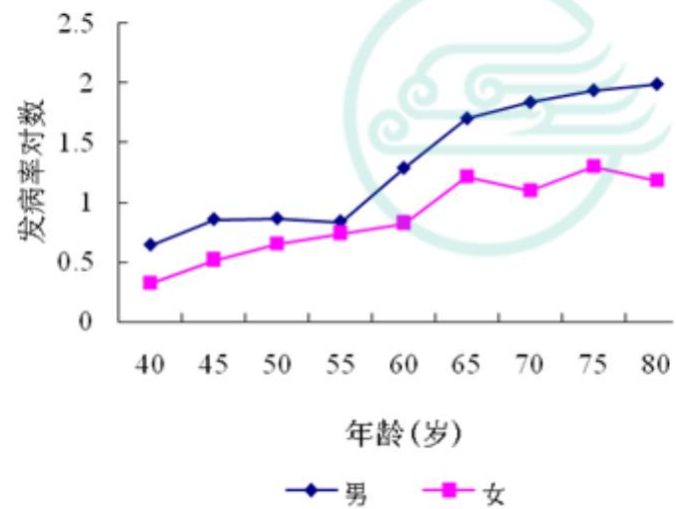
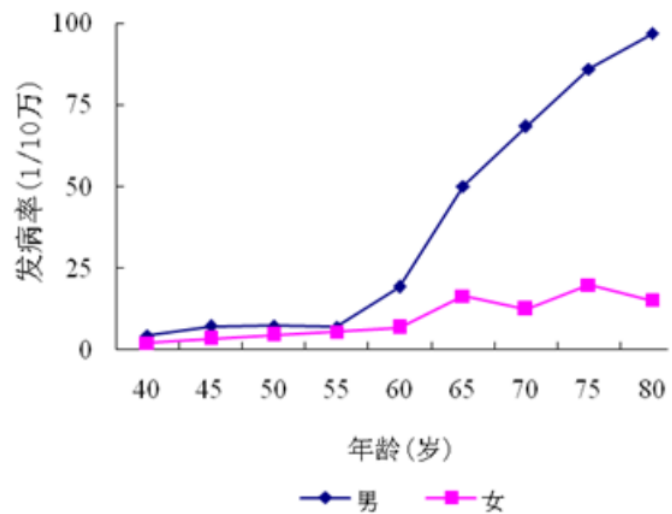
脾肿 程度	观察 例数	血膜阴性		恶性疟		间日疟		血膜阳性合计	
		例数	构成比(%)	例数	构成比(%)	例数	构成比(%)	例数	构成比(%)
脾Ⅰ	122	105	86.06	8	6.56	9	7.38	17	13.94
脾Ⅱ	70	51	72.86	14	20.00	5	7.14	19	27.14
脾Ⅲ	26	15	57.69	6	23.08	5	19.23	11	42.31
其他	4	3	75.00	0	0.00	1	25.00	1	25.00
合计	222	174	78.38	28	12.61	20	9.01	48	21.62



课后习题二-2



课后习题二-3





人民卫生出版社

PEOPLE'S MEDICAL PUBLISHING HOUSE

谢谢观看