

第十五章 研究中的统计设计

第一节 | 研究类型

一句话总览

医学研究只有一个根本分野：
你能不能、该不该“动手干预”？

一、医学研究的两大基本类型（总框架）

医学研究按研究者是否主动施加干预，分为两大类：

类型	核心特征	研究者角色
观察性研究	不施加干预	被动观察
实验性研究	主动施加干预	主动控制

◆ 这是所有研究设计判断的第一步。

二、观察性研究（Observational Study）

① 本质特征（一定要抓住）

研究因素在自然状态下已经存在，
研究者不能、也不应随机分配。

研究者做的事只有三件：

1. 记录暴露
2. 记录结局
3. 分析二者关系

2 为什么必须用观察性研究？（伦理与现实）

很多问题在伦理或现实上不允许实验：

- 不能随机让人：
 - 抽烟
 - 久坐
 - 不运动
- 不能人为制造：
 - 疾病
 - 有害暴露

👉 这类问题只能通过观察性研究回答。

3 观察性研究的三大类型（按时间顺序）

(1) 横断面研究 (Cross-sectional study)

关键词：

某一时点 / 现况 / 同时测量

定义：

- 在某一特定时间
- 对一定范围人群
- 同时收集暴露和结局信息

能回答什么？

- 患病率
- 暴露率
- 相关性线索

不能回答什么？

- 因果方向
- 发病风险

逻辑特征：

“有没有关系？”

但不问“谁先谁后”。

(2) 病例-对照研究 (Case-control study)

关键词：

由果推因 / 回顾性

基本思路：

1. 先按结局分组
 - 病例组
 - 对照组
2. 再回顾既往暴露情况

时间逻辑：

结局 → 既往暴露

优点：

- 适合：
 - 罕见病
 - 潜伏期长的疾病
- 省时、省钱

局限：

- 回忆偏倚

- 无法直接算发病率
-

(3) 队列研究 (Cohort study)

关键词:

由因寻果 / 前瞻性 / 纵向

基本思路:

1. 先按暴露分组
2. 随访
3. 观察结局发生

时间逻辑:

暴露 → 结局

优点:

- 因果顺序清晰
- 可直接计算:
 - 发病率
 - 相对危险度 (RR)

缺点:

- 时间长
 - 成本高
 - 易失访
-

🕒 三种观察性研究的“时间轴对照”

类型	分组依据	时间方向
横断面	无明确分组	同一时点

病例-对照 结局 回顾过去
队列 暴露 走向未来

三、实验性研究 (Experimental Study)

1 本质特征 (和观察性研究的根本区别)

研究者主动施加干预，
并尽可能控制非研究因素。

核心手段通常包括：

- 干预
 - 对照
 - 随机化
 - 盲法 (在合适情况下)
-

2 实验性研究的三大类型 (按研究对象)

(1) 动物实验 (Animal experiment)

研究对象：

- 动物
- 组织、细胞等生物材料

用途：

- 机制探索
- 毒性研究
- 初步疗效验证

特点：

可控性强
但外推到人需谨慎。

(2) 临床试验 (Clinical trial)

研究对象：

- 人（患者或健康志愿者）

研究目的：

- 验证药物或治疗方法：
 - 疗效
 - 安全性
 - 药代动力学

◆ 这是：

循证医学中证据等级最高的设计之一。

(3) 社区干预试验 (Community intervention study)

研究对象：

- 人群 / 社区整体

特点：

- 干预持续时间长
- 常用于公共卫生领域

目的：

- 评估：
 - 预防措施

- 健康促进策略
对人群健康结局的影响
-

四、观察性 vs 实验性：一眼区分法

问题	是	否
研究者能否随机分组?	实验性	观察性
研究者是否主动干预?	实验性	观察性
研究是否受伦理强限制?	多为观察性	相对少

五、这一节真正要你掌握的“判断逻辑”

① 能不能干预?

→ 不能: 观察性研究

② 若是观察性, 时间顺序如何?

→ 同时: 横断面
→ 由果推因: 病例-对照
→ 由因寻果: 队列

③ 若能干预, 研究对象是谁?

→ 动物 / 细胞
→ 个体 (临床)
→ 人群 (社区)

六、考试级一句话总结 (可直接写)

医学研究按研究者是否主动施加干预分为观察性研究和实验性研究。观察性研究中研究因素在自然状态下已存在, 研究者不进行干预, 主要包括横断面研究、病例-对照研究和队列研究; 实验性研究则由研究者主动施加干预并控制非研究因素, 主要包括动

物实验、临床试验和社区干预试验，不同研究类型对应不同的设计方法和因果推断能力。

第二节 | 观察性研究设计

一句话总览

观察性研究设计的核心不是“选哪种研究”，而是：

如何在不能随机化的前提下，把偏倚压到最低，让推断尽可能可信。

一、观察性研究三大类型 (When & What)

1 横断面研究 (Cross-sectional)

核心问题 (Purpose)

回答“**现在是什么水平**”：

疾病/健康状况、暴露/特征的现况分布。

能做什么 (能产出)

- 描述：患病率、暴露率、特征分布
- 探索关联：发现危险因素线索（相关 ≠ 因果）

设计关键点 (怎么做得好)

1. 先明确研究问题 → 决定调查方式

- 普查 (census)
- 典型调查 (typical survey)
- 抽样调查 (最常用)

2. 调查表 (问卷/测量表) 是成败关键

- 项目围绕目的 (提问/测量)
- 逻辑顺序清晰
- 填写尽量简单 (√、○、数字, 少文字)

◆ 最容易踩雷

横断面很容易被误读为因果。
它最多提供“可疑线索”，不是“病因结论”。

2 病例-对照研究 (Case-control)

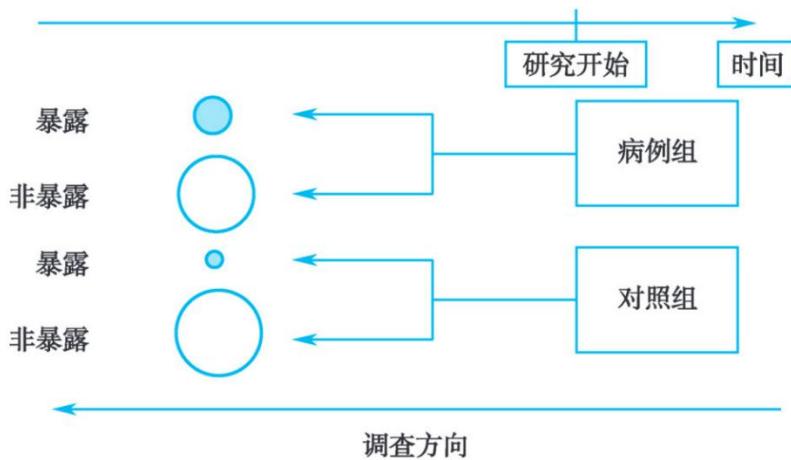


图 15-1 病例-对照研究示意图

(李康,贺佳主编, 2024, p. 159) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

核心问题 (Purpose)

探索“可能的危险因素”：
由果推因 (回顾性 retrospective)。

基本流程 (Structure)

1. 先按结局选人：病例 / 对照
2. 再回顾既往暴露史
3. 比较两组暴露差异 → 推断关联

优点 (为什么常用)

- 资料相对容易获取

- 周期短、成本低、出结果快
- 可同时研究一种病与多种暴露（多病因线索）

主要缺点（考试高频）

- 回忆偏倚（recall bias）
- 选择偏倚（selection bias）（对照选不好直接崩）

方法论定位

病例-对照是“找线索的强工具”，
但要证因果，常需进一步队列/干预试验验证。

3 队列研究（Cohort）

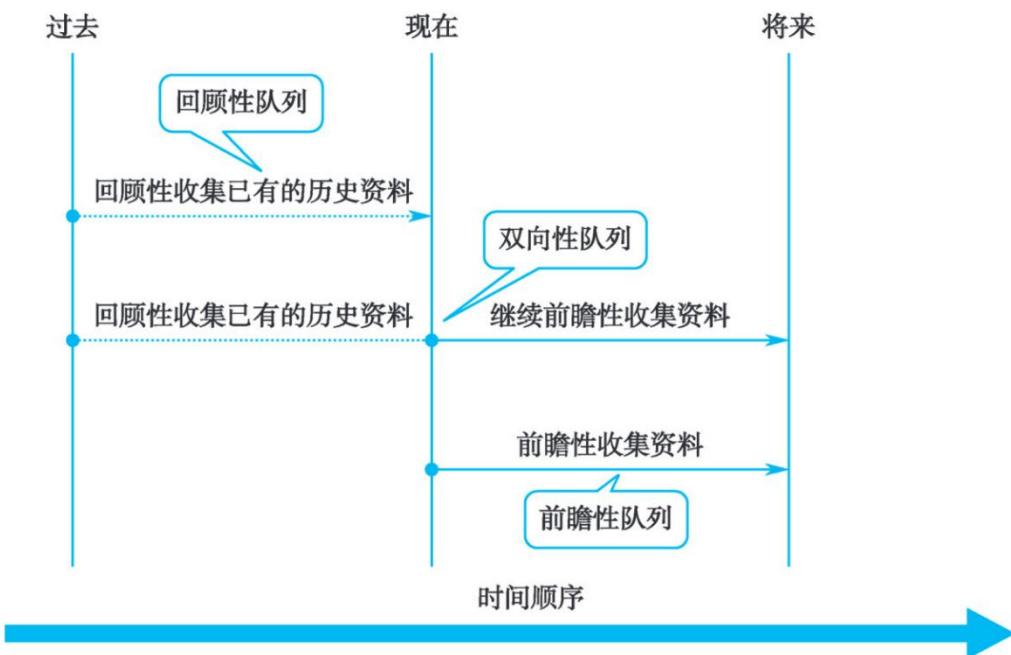


图 15-2 三种类型队列研究示意图

(李康,贺佳主编, 2024, p. 160) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

核心问题 (Purpose)

检验病因假设：
由因寻果 (纵向、常为前瞻性 prospective)。

基本流程 (Structure)

1. 研究开始时按暴露分组：暴露 / 非暴露 (自然形成, 非随机)
2. 随访
3. 比较发病/死亡/复发等发生率

能力与优势 (Why it's strong)

- 因果顺序更清楚 (暴露在前、结局在后)
- 可直接得到发病率/死亡率
- 假设检验能力较强，有时可较明确证实病因联系
- 还能获得“预期外的结局信息”

局限 (Why it's hard)

- 不适合罕见结局 (需要巨大样本)
- 随访时间长、成本高
- 失访增加 → 偏倚风险上升

队列研究的三种形态 (按时间起点)

① 前瞻性队列

- 开始时暴露已存在，结局尚未发生 → 未来随访

② 回顾性队列

- 暴露与结局都已发生 → 用历史记录“向后追踪”
- 优点：省钱省时
- 缺点：历史资料质量决定可信度

③ 双向性 (混合) 队列

- 暴露用历史档案确定

- 结局向未来随访确定
-

二、混合型设计：把队列和病例-对照“拼起来”

目的通常是：

兼顾效率（像病例-对照）与可比性（像队列）

尤其是生物标志物/标本成本很高时。

1 巢式病例对照 (Nested case-control)

核心结构

- 先建立一个队列并随访
- 队列里出现的新发病例 → 病例组
- 对照从同一队列中选（且在病例发病时仍未发病），常匹配
- 之后按病例-对照方法分析

最适用

- 随访开始后出现新的病因假说
 - 或研究某些“需要保存标本、成本很高”的前体指标
-

2 病例队列 (Case-cohort)

核心结构

- 先建全队列
- 随机抽取一个子队列作为对照组
- 全队列中所有病例作为病例组

最适用

- 巨大队列、结局很低发

- 需要计算发病率、标化死亡比、做外部比较
 - 想研究多个结局时更高效（子队列可复用）
-

三、抽样方法 (Sampling) : 代表性靠它兜底

1 调查方式三类

- 普查：全覆盖，成本高
 - 典型调查：方便但偏倚风险大
 - **抽样调查：最常用**（小投入换整体估计）
-

2 概率抽样 vs 非概率抽样 (底线区分)

- **概率抽样 (probability sampling)**
每个单位有已知非零抽中概率 → 才能做推断
- 非概率抽样：方便但推断性弱（偏倚风险大）

抽样框架 (sampling frame)

- 有限总体可列清单 → 适合概率抽样
 - 无限总体无法列框架 → 不适合概率抽样
-

3 四种常用概率抽样 (必考对比)

(1) 单纯随机抽样 SRS

- 每个个体等概率
- 计算简单
- 但：总体大且分散时编号困难、现场实施困难

(2) 系统抽样 systematic

- 等距抽样：间隔 $i = N/n$, 随机起点 k

- 优点：操作简单、样本分布均匀，误差常小于 SRS
- 缺点：总体存在周期性/单调趋势时会产生偏性

(3) 整群抽样 cluster

- 抽样单位是“群”（班级/社区/县区）
- 优点：省钱省力、组织方便
- 缺点：抽样误差大（群内相似导致有效信息少）
- 改善策略：**多抽群、群间差异尽量小、群大小相近**

(4) 分层抽样 stratified

- 先按重要特征分层（性别、年龄等）
- 各层内再抽样（可用系统/整群等）
- 优点：层内更同质 → 误差最小、精度最高
- 分层原则：**层间差别大，层内差别小**

👉 抽样误差一般大小规律（教材给的顺序）

整群 ≥ 单纯随机 ≥ 系统 ≥ 分层

同精度下：误差越大 → 需要样本量越多。

四、样本量估计 (Sample size) : 别盲目追求大, 关键是“够用且可信”

样本量估计本质是：

在给定精度与置信度下，求最小 n。

1 横断面研究样本量：先定三个参数

- 允许误差 δ
- 变异程度：定量用 σ , 定性用发生率 P
- 置信度 $1 - \alpha$ (常取 95%)

(1) 估计总体均数

$$n = \left(\frac{z_{\alpha/2} \sigma}{\delta} \right)^2$$

(2) 估计总体率

$$n = \frac{z_{\alpha/2}^2 P(1-P)}{\delta^2}$$

(P 越接近 0.5, 所需样本量越大)

【例 15-1】调查某地区学龄儿童贫血患病率情况,根据以往资料,该地学龄儿童贫血患病率约为 15%,若要求本次调查所得样本率与未知总体率相差不超过 3% 的可能性不大于 0.05,按照单纯随机抽样,估计需要调查多少名儿童?

本研究 $z_{0.05/2} = 1.96, P = 0.15, \delta = 0.03$, 计算:

$$n = \left(\frac{z_{\alpha/2}^2 P (1-P)}{\delta^2} \right) = \frac{1.96^2 \times 0.15 \times (1 - 0.15)}{0.03^2} = 545$$

即研究至少需要调查 545 名儿童。

(李康,贺佳主编, 2024, p. 163) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

2 病例-对照研究样本量 (成组设计)

$$n_1 = \frac{(1 + 1/C)\bar{P}(1 - \bar{P})(z_{\alpha/2} + z_{\beta})^2}{(P_1 - P_0)^2}, n_0 = Cn_1$$

其中:

- C : 对照/病例比例
- P_0, P_1 : 对照/病例暴露率
- $\bar{P} = (P_0 + P_1)/2$

并且:

$$P_1 = \frac{OR \times P_0}{1 + P_0(OR - 1)}$$

👉 直觉:

差异越小、效能要求越高 → 样本量越大。

【例 15-2】某研究者做病例-对照研究,对照组的某因素暴露率(P_0)估计为 0.4,希望 $OR \geq 2$,要求双侧 $\alpha = 0.05, \beta = 0.10$,对照组与病例组例数相同,样本例数估计为多少?

本研究 $z_{0.05/2} = 1.96, z_{0.10} = 1.282$,计算:

$$n_1 = \frac{(1+1/C)\bar{P}(1-\bar{P})(z_{\alpha/2} + z_{\beta})^2}{(P_1 - P_0)^2} = \frac{(1+1/1) \times 0.486 \times (1-0.486) \times (1.96 + 1.282)^2}{(0.571 - 0.4)^2} = 180$$

即病例组和对照组各需至少观察 180 例。

(李康,贺佳主编, 2024, p. 163) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

【例 15-3】设某研究者拟采用 1 : 1 的配对病例-对照研究,对照组某因素暴露率(P_0)估计为 0.4, OR 为 2,要求双侧 $\alpha = 0.05, \beta = 0.10$,样本例数估计为多少?

本研究 $z_{0.05/2} = 1.96, z_{0.10} = 1.282$,计算:

$$m = \frac{\left[z_{\alpha/2}/2 + z_{\beta}\sqrt{P(1-P)} \right]^2}{(P-0.5)^2} = \frac{\left[1.96/2 + 1.282\sqrt{0.67(1-0.67)} \right]^2}{(0.67-0.5)^2} = 87$$

需要的总对子数为:

$$M = \frac{m}{P_1(1-P_0) + P_0(1-P_1)} = \frac{87}{0.571 \times (1-0.4) + 0.4 \times (1-0.571)} = 170$$

即该研究至少需要 170 对观察对象。

(李康,贺佳主编, 2024, p. 164) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

3 队列研究样本量 (两组相等)

$$n_1 = n_0 = \frac{(z_{\alpha/2}\sqrt{2PQ} + z_{\beta}\sqrt{P_0Q_0 + P_1Q_1})^2}{(P_1 - P_0)^2}$$

其中:

- P_0, P_1 : 对照/暴露组发病率
- $Q_0 = 1 - P_0, Q_1 = 1 - P_1$
- $P = (P_0 + P_1)/2, Q = 1 - P$

【例 15-4】在某项队列研究中,将人群是否具有某种白细胞抗原分为两组,随访 5 年,观察每组中患某种疾病的人数。据有关资料估计相对危险度为 0.6,不具备该白细胞抗原者患病率约为 0.4,设双侧 $\alpha=0.05$, $\beta=0.10$,试计算其样本量。

本研究 $z_{0.05/2}=1.96$, $z_{0.10}=1.282$,计算:

$$n_1 = n_0 = \frac{(z_{\alpha/2}\sqrt{2PQ} + z_{\beta}\sqrt{P_0Q_0 + P_1Q_1})^2}{(P_1 - P_0)^2}$$
$$= \frac{(1.96\sqrt{2 \times 0.32 \times 0.68} + 1.282\sqrt{0.4 \times 0.6 + 0.24 \times 0.76})^2}{(0.24 - 0.4)^2} = 177$$

即本研究暴露组与对照组至少需要各随访 177 人。

(李康,贺佳主编, 2024, p. 164) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

五、真实世界研究 (RWS) : 理念与边界

1 什么是 RWS?

数据来自真实医疗环境 (real-world data), 目的在于:

评价干预措施在真实条件下的效果 (effectiveness) 、风险与价值
(包括经济学评价)

2 与随机对照试验 (RCT) 的差别 (考试高频)

- RCT: 更关注“理想条件下的效力 efficacy”
- RWS: 更关注“真实条件下的效果 effectiveness”
- RWS 纳入更宽泛、排除更少, 更贴近临床现实
- 常用大样本 (如 $n \geq 10,000$) 做亚组分析

3 RWS 的共同挑战: 混杂

观察性研究的共同特点就是混杂因素存在, 因此常用方法包括:

- 分层分析
- 多元回归

- 倾向评分 (propensity score)
- 工具变量 (instrumental variable)

👉 一句话定位：

RWS 不等于 “随便拿数据跑回归” ，
它更依赖严谨的混杂控制策略。

六、考试级一句话总结（可直接写）

观察性研究包括横断面研究、病例-对照研究和队列研究，分别用于描述现况、探索危险因素线索及检验病因假设。观察性研究常采用抽样调查以保证样本代表性，常用的概率抽样方法包括单纯随机抽样、系统抽样、整群抽样与分层抽样，其中分层抽样精度通常最高。不同研究目的需采用相应的样本量估计方法。真实世界研究利用真实医疗环境数据评价干预措施的实际效果，但必须重视混杂因素并采用分层、多元回归、倾向评分等方法进行控制。

第三节 | 实验设计（结构化总结）

一句话总览

实验设计的目标只有一个：

在可控条件下，把“处理因素 → 真实效应”从噪声里分离出来。

一、实验设计的三要素（缺一不可）

三要素 = 对象 + 处理 + 效应

没有其中任何一个，都不能叫实验性研究。

① 研究对象 (Study subject / participant)

核心要求：同质 + 可执行 + 合伦理

1. **明确纳入/排除标准** (保证同质性)
 - 动物实验：种属、品系、性别、体重、窝别...
 - 临床试验：诊断、病情、年龄、合并症...
2. **选择对处理因素“敏感”的对象**
 - 例如抗心律失常药：选“发作频繁者”，否则观察期内看不出差异
3. **选择依从性好的受试者** (否则偏倚)
4. **伦理底线**
 - 人：知情同意 (informed consent)，医疗优先
 - 动物：善待动物 (动物伦理)

❖ 一句话记忆：

对象选不对，统计方法救不了。

2 处理因素 (Treatment / study factor)

两个概念必须分清

- **处理因素**：你想研究的干预 (药物/剂量/手术/教育...)
- **非处理因素**：也会影响结局但不是你要研究的 (混杂因素 confounders)

处理因素设计的两条底线

1. **标准化 (standardization)**
 - 剂量、频次、批号、配制、采样时间点...都要写清楚并保持稳定
 2. **处理因素别太多，混杂要被控制**
 - 设计控制优先 (随机、分层、对照)
 - 统计校正是补救但**能力有限**
-

3 实验效应 (Experimental effect)

你观测什么，就决定你能证明什么。

指标选择三条准则

1. 尽量客观 (objective)
 - 主观指标 (量表、疼痛评分) 易受心理暗示影响 → 需要量化 + 盲法等控制
 2. 灵敏度 & 特异度高
 - 灵敏度高：更不容易漏掉真实效应 (少假阴性)
 - 特异度高：更不容易把噪声当效应 (少假阳性)
 3. 准确度 vs 精密度
 - 准确度 (accuracy)：接近真实值 (系统误差)
 - 精密度 (precision)：重复测量稳定 (随机误差)
-

二、实验设计的三原则 (控制偏倚的三把刀)

对照 + 随机化 + 重复

是实验性研究能“接近因果”的关键。

1 对照原则 (Control)

没有对照，就无法区分“自然变化/安慰剂效应/非处理因素”与真实效应。

常见对照形式 (重点掌握“适用场景”)：

1. 空白对照 (blank control)：不给任何处理
 - 动物实验常用；临床一般不宜 (心理差异+伦理)
2. 安慰剂对照 (placebo control)：伪药，外观一致
 - 控制心理偏倚很强，但可能有伦理问题 (等于不给有效治疗)
3. 标准对照 (standard control)：现有公认有效疗法/标准值
 - 新药 vs 现有一线药，最常见、最现实
4. 实验对照 (experimental control)：假注射/假手术/假针灸
 - 控制“操作本身带来的刺激与损伤”
5. 自身对照 (own control)：同一对象前后/同期不同处理
 - 典型用途：同一样本用两种方法检测，消除个体差异
6. 相互对照 (mutual control)：多组互为对照
 - 多药/多剂量比较

2 随机化原则 (Randomization)

随机化的本质：让已知/未知混杂在组间“平均化”，并且让推断有理论基础。

常见随机化方法（你需要记住“解决什么问题”）：

1. 简单随机化 (simple / complete)
 - 最简单，但可能出现组间人数/特征短期不平衡
 2. 区组随机化 (blocked / permuted block)
 - 解决：保证组间人数平衡
 3. 分层随机化 (stratified)
 - 解决：重要混杂因素分布均衡（但层别别太多，否则层内例数不够）
 4. 动态随机化 (dynamic: biased coin / urn / minimization)
 - 解决：在试验进行中实时调整分配概率，进一步平衡关键因素
-

3 重复原则 (Replication)

重复 = 足够的样本量与观察次数

没有重复就没有稳定规律，只剩偶然波动。

样本量估计的四要素（考试直接背这个框架就够）

1. α (I类错误)：越小 $\rightarrow n$ 越大
2. $1-\beta$ (power)：越大（通常 ≥ 0.80 ） $\rightarrow n$ 越大
3. 效应差 δ ：越小 $\rightarrow n$ 越大
4. 变异 σ / 率 π ：波动越大 $\rightarrow n$ 越大

（教材给的两均数比较公式你已有；关键是理解上面四点。）

三、临床试验的特殊性（人 → 心理与伦理）

1 盲法 (Blind method)

- **双盲**: 研究者+受试者+相关人员都不知道分组
- **单盲**: 通常受试者不知道
- **开放**: 不设盲, 但也要尽量减少人为干预

盲法的作用: 控制心理与观察者偏倚 (performance / assessment bias) 。

2 常见设计类型 (Design types)

1. **平行组 (parallel)** : 最常用; 简单但通常要更多样本
 - 双臂 (two-arm)、三臂 (three-arm: 安慰剂+阳性对照)
2. **交叉设计 (cross-over)** : 自身对照升级版
 - 必须有洗脱期 (washout), 否则残留效应污染结果
 - 适合: 缓解症状但不根治、可逆结局 (止痛、失眠、风湿等), 以及生物等效性
3. **析因设计 (factorial)**: 同时评价多个药物/剂量组合 + 看交互作用
4. **适应性设计 (adaptive)** : 中期根据累积数据调整方案
 - 关键风险: I类错误膨胀、结果可解释性下降
 - 常需独立数据监查委员会 (IDMC) 保持盲态与规范决策

3 比较目的三分法 (Superiority / Non-inferiority / Equivalence)

1. **优效性 (superiority)**: 试验药是否更好?
2. **非劣效 (non-inferiority)** : 不比对照差超过 Δ
 - 核心是 Δ 的临床意义: 太大 → 放水; 太小 → 做不出来
 - 常用单侧置信区间判断: 下限 $> -\Delta$ 才成立
3. **等效性 (equivalence)** : 差异在 $(-\Delta, \Delta)$ 内
 - 双侧区间完全落入界值范围才算等效

四、诊断试验 (Diagnostic accuracy study) 体系

一句话总览

诊断试验研究不是“看相关”, 而是:

新方法相对金标准的分类准确性。

1 设计要点 (最容易出偏倚的地方)

1. 金标准 (gold standard) 必须合理
 - 病理、手术所见、特定影像、培养、长期随访、公认综合诊断...
 2. 研究对象必须有代表性
 - “无病” 不一定是健康人，常应选“需要鉴别诊断”的类似病
 - “有病” 要覆盖典型+不典型、轻中重、不同分期
 3. 实施必须双向盲法
 - 做新方法的人不能知道金标准结果
 - 判金标准的人也不能受新方法影响
 4. 对象来源要写清
 - 专科门诊 vs 社区门诊 → 患病率不同 → PV 会完全不同
 5. 样本量
 - 单方法：按灵敏度估病例组 n，按特异度估对照组 n (单样本率思路)
 - 两方法比较：还要考虑配对一致性概率等
-

2 四格表与核心指标 (必须会解释)

设：

- a: 有病且阳性
- b: 有病但阴性
- c: 无病但阳性
- d: 无病且阴性

(1) 灵敏度 Se (sensitivity)

$$\widehat{Se} = \frac{a}{a + b}$$

适用语境：漏诊代价高、用于排除 (rule out)

(2) 特异度 Sp (specificity)

$$\widehat{Sp} = \frac{d}{c + d}$$

适用语境：误诊代价高、用于确诊 (rule in)

Se 与 Sp 往往此消彼长：提高一个常要牺牲另一个。

3 预测值 (PV) 为什么“会变” (重点理解)

阳性预测值 PV+

$$PV_+ = \frac{\pi Se}{\pi Se + (1 - \pi)(1 - Sp)}$$

阴性预测值 PV-

$$PV_- = \frac{(1 - \pi)Sp}{(1 - \pi)Sp + \pi(1 - Se)}$$

关键结论：PV 依赖患病率 π

同一个检验，放在普通人群（低 π ）PV+ 会显著下降；
放在高风险门诊（高 π ）PV+ 才会更高。
所以诊断试验论文必须交代样本来源。

4 ROC 曲线与 AUC (连续/等级指标的总评价)

- ROC：横轴 $1 - Sp$ (假阳性率)，纵轴 Se
- AUC (A)：0 到 1

常见参考：

- 0.5–0.7：价值较低
 - 0.7–0.8：有一定价值
 - 0.8–0.9：较高价值
 - 0.9：很高价值
-

五、这节的“脑内流程图”（实战版）

- ① 先把实验三要素写全：对象—处理—效应
- ② 用三原则控偏：对照—随机—重复（样本量）
- ③ 临床试验再加两层：盲法 + 比较类型（优效/非劣/等效）
- ④ 若是诊断试验：金标准 + 盲法实施 + Se/Sp/PV/ROC-AUC

[BIBLIOGRAPHY] Please click Zotero - Refresh in Word/LibreOffice to update all fields