

E10

人用药品注册技术要求  
国际协调会  
ICH 三方协调指导原则

临床试验中对照组的选择

**ICH 指导委员会**  
**2000 年 7 月 20 日**  
**ICH 进程第四阶段推荐采纳**

该指导原则草案由相应的 ICH 专家小组制定，按照 ICH 进程，在 ICH 进程第二阶段，草案被一致通过，并推荐给欧盟、日本和美国三方的管理机构以供内外讨论。

# 临床试验中对照组的选择

## 目 录

<b>1.引言</b> .....	(317)
1.1 指导原则的总体安排和目的.....	(317)
1.2 对照组的目的.....	(318)
1.3 对照的类型.....	(320)
1.4 临床试验的目的.....	(323)
1.5 评价非劣效性或等效性研究的药效灵 敏度和分析灵敏度.....	(326)
<b>2. 对照类型的详细考虑</b> .....	(333)
2.1 安慰剂对照.....	(333)
2.2 无治疗并行对照(见 1.3.2).....	(341)
2.3 量效并行对照(见 1.3.3).....	(342)
2.4 阳性药物对照.....	(345)
2.5 外部对照(历史对照).....	(348)
<b>3. 选择对照组</b> .....	(352)
<b>附录</b> .....	(353)

# 临床试验中对照组的选择和相关问题

## 1.引言

在设计临床试验时,选择对照组常常是一项关键性的决定。这一选择对研究的许多方面都会产生影响,包括:对试验的推理,在研究进行和分析中能最大限度地减少偏倚的程度,被招募的受试者类型和募集的速度,研究终点的种类,研究结果的公众可信度,研究结果对于管理当局的可接受性,以及研究的许多其他特点,包括其实施和解释。

### 1.1 指导原则的总体安排和目的

本指导原则的目的是描述在临床研究中为证明疗效选择对照组所考虑到的一般原则,并讨论试验设计和实施中的一些相关问题。本指导原则并非任何地区的管理要求,它只是描述各种设计方案的研究能说明什么。这些原则与所有对照试验有关,然而选择对照组对在药物研发中为证明疗效所进行的试验尤其的重要。在选择对照组时,应当考虑现行的标准治疗、有足够的证据支持所选方案,同时要考虑伦理方面的因素。

指导原则首先描述了选择对照组的目的是为证明有效性通常采用的对照类型。之后探讨了通过证明非劣效性/等效性的有效性活性对照试验,及其相关的重要设计和阐述问题(参阅1.5)。在某些情况下,非劣效性的结果不能被解释为有效的证

据。特别是，非劣效性的结果如果被解释为有效的证据，该临床研究必须能够区分有效、不太有效和无效。

指导原则将比较详细地描述每一种对照组(参阅 2.0~2.5.7)，并就下列问题加以讨论：

- 减少偏倚的能力；
- 与使用有关的伦理学问题和实际问题；
- 在特定情况下的用途和推理质量；
- 研究设计的修改或与其他对照结合起来解决道德的、实际的或推理问题；
- 总的优点和缺点。

下列几份 ICH 指导原则都与对照组的选择特别有关：

- E3：临床研究报告的结构与内容
- E4：支持药物注册的量效信息
- E6：临床试验管理规范
- E8：临床试验的一般考虑
- E9：临床试验统计学指导原则

尽管下面描述和讨论地各种对照组至少在某些情况下可作为注册用地基础研究，并被认可，但并不同时适用于所有情况。将在第 3 部分，图 1 和表 1 中列出选择对照类型的一般方法。

尽管指导原则主要针对的是企图评价有效性的临床试验，所探讨的许多方面也适用于评估特定的安全性假设，和两种治疗的安全性和有效性的比较。

## 1.2 对照组的目的是

对照组有一个主要目的，就是可以将受试药物给病人带来的结果(症状、体征或其他病状的改变)与其他因素，如疾病的自然进展、观察者或者病人的期望、其他治疗措施等造成的结果区分开来。对照组的结果可以告诉我们，假如没有接受受试药物，病人会发生什么情况(或者接受另外一种已知有效药物，病人会发生什么情况)。

假如在特定的病人人群中，疾病的进程是相同的，或者是根据病人的特征，就可以可靠地预测任何特定受试者或受试组的结果，那么，治疗的结果就可以简单地与未接受治疗的已知结果相比较。比如，可以假定疼痛会持续了一定时间；血压不会改变；抑郁将持续一定的时间；肿瘤将会进展；急性梗死后的死亡率和以前一样，等等。事实上，仅在少数情况下，疾病的病程在特定人群中是可以预测的，并且可以使用以往研究过的类似的病人组作为“历史对照”(参阅 1.3.5)。然而，在多数情况下，需要有一个并行的对照组，因为不可能预测出足够精确的研究结果。

并行的对照组是从与试验组相同的人群中选出的，并且作为研究受试药物的同一个试验的一部分按规定的方法接受治疗。试验组和对照组的所有基线值，以及除了试验治疗外的其他可能影响结果的有关变量都应当相似。达不到这样的相似性就可能在研究中引入偏倚。这里所说的偏倚(以及在 E9 所说的

偏倚),是指在临床试验的设计、实施、分析和对结果的解释等任何方面出现的系统倾向,从而使对治疗结果的估计偏离它的真实值。随机化和盲法是常用的两种技术,用以预防这种偏差的发生,以及保证试验组和对照组在开始研究时彼此相似,并在研究过程中接受相似的处理(参阅 ICH E9)。试验设计是否包括有这些特性是决定其质量和说服力的关键因素。

### **1.2.1 随机化**

将单一样本人群随机分入试验组或对照组能最好地确保两组的受试人群相似。随机化可以避免那些可能影响结果的变量的组间系统差异。不能消除系统差异是没有并行随机对照研究的一个主要问题(参阅 1.3.5, 外部对照试验)。随机化也为统计推理打下坚实的基础。

### **1.2.2 盲法**

除了接受受试药物和对照药物之外,试验过程中,各组不仅基线值应当相似,而且对其处理和观察也应是相似的。临床试验通常是“双盲”,是指受试者与研究者(包括数据分析者、申办者、其他临床试验人员)双方都不了解每个受试者接受的治疗,从而最大限度地减少由于受试者或研究人员了解治疗分配后引起的在管理、治疗或对病人的评价以及解释结果时出现的偏倚。

例如：

- 假如受试者知道他们服用的受试药物，受试者可能会报告有利的结果，因为他们期望得益于治疗，或者更有可能留在研究中。
- 观察者可能不太会辨别和报告非治疗组的反应，而对接受活性药物治疗的病人的有利结果或不良事件更加敏感。
- 了解治疗的分配情况可能会影响对研究或随访数据的收集。
- 了解治疗的分配情况可能会影响对病人是否应继续接受治疗或接受伴随用药或其他辅助治疗的决定。
- 了解治疗的分配情况可能会对决定是否应将某一受试者的结果列入分析产生影响。
- 了解治疗的分配情况可能对选择统计分析方法产生影响。

双盲的目的是为了确保主观评价及决定不会因了解治疗分配而受到影响。

### 1.3 对照的类型

根据以下两个关键的特性可以对临床试验中的对照组进行分类：(1) 所接受治疗的类型以及(2)决定谁将进入对照组的方法。治疗类型可以是下列四种中的任何一种：(1)安慰剂、(2)无治疗、(3)不同剂量或剂量方案的受试药物、(4)以及不同的活

性药物。决定谁将进入对照组的主要方法是随机化或选择一个有别于治疗人群的对照人群(外部的或历史对照)。本文将对照组划分为五类。前四类为并行对照(对照组与试验组从同一人群中挑选并同时进行治疗),通常是将受试者随机分配到各组,以接受上述不同类型的对照治疗而分类。不论接受何种比较治疗,外部(历史的)对照都归为第五类,出于对这些试验中保证试验组与对照组可比性能力及减少重大偏倚的严肃考虑,使得这种设计只能在特殊条件下使用。

研究中选用一类以上对照组的情况越来越普遍。每一类对照适合于某些情况,但没有一种可以用于或适于所有的情况。这五种类型分别是:

### **1.3.1 安慰剂并行对照**

在安慰剂对照研究中,受试者被随机分配到受试药物组或外表完全相同的无活性药物组。药物可被滴定到有效水平或耐受水平,也可以给予一个或多个固定的剂量。这类试验几乎总是双盲试验,受试者与研究者都不知道治疗分配情况。本类型的名称提示其目的是以"安慰剂"为参照标准(受试者的改善是因为他或她认为是在服药),但这不是它唯一或主要的优势。更准确地说,安慰剂并行对照设计,通过采用盲法和随机化,以及设立一个无治疗组,控制了除试验药物药理作用之外的所有对真实和表观病程的潜在影响。这些影响包括自发性的改变(疾病的自然史)、受试者或研究者的期望、使用其他治疗以及诊断



或评估中的主观因素。在研究药效时，安慰剂对照试验试图发现治疗间的差异，但在评价安全性指标时，也可能是试图证明没有(指定大小的)差异。在这种情况下，试验是否显示差异就成为关键性的问题（见 1.5）。

采用安慰剂对照组并不是说对照组不接受治疗。在许多安慰剂对照试验中，新的治疗方法或安慰剂是加在通常的标准治疗之上的（所谓的 ‘add-on’研究，见 2.1.5.2.1 部分）。

### **1.3.2 无治疗并行对照**

在无治疗的对照研究中，受试者被随机分配到试验治疗组或非（如，无）试验治疗组。这种设计和安慰剂对照试验的主要区别是，治疗分配对受试者和研究者都是公开的。因为双盲设计的优点，这种设计很可能仅仅是在下列情况下才需要和适用：即难以或不可能实行双盲(例如，药物治疗与手术治疗；容易识别毒性的治疗)；以及有理由确信研究终点是客观的，而且研究结果不大可能受 1.2.2 所列因素的影响。值得注意的是，即使整个试验不是双盲的，但通常仍可能运用盲态的终点评价。这是一种有价值的方法，在不能进行双盲的研究中应该经常考虑它，但它不解决已知治疗分配的情况下所带来的其他问题(参阅 1.2.2)。

### **1.3.3 剂量-响应并行对照**

在随机化的、固定剂量的量效研究中，受试者被随机分配到几个固定剂量组中的一组。受试者可能一开始就接受其固定剂量，或者逐渐升高到这一剂量，而要进行的组间比较是最终剂量。量效研究常常是双盲的。研究可能包括一个安慰剂(零剂量)和/或活性对照。在浓度控制试验中，治疗组的剂量被增量到数种固定的浓度窗；这种类型的试验在概念上类似于固定剂量的量效试验。在剂量方法对照试验中，受试者被随机分配到两个或多个研究药物剂量方法组中（如，每天一次同每天两次的比较，三天同柒天的比较）。

### **1.3.4 活性(阳性)并行对照**

在活性对照(或阳性对照)研究中，受试者被随机分配到受试药物组或阳性对照药组。这种试验通常是双盲的，不过有时不可能做到双盲；例如，许多肿瘤学研究，由于存在不同的给药方案、不同的给药途径(参阅 1.3.2)以及不同的毒性，双盲被认为是不可能的。阳性对照试验在证明试验药物疗效方面有两个不同的目标：(1)显示受试药物的疗效与某种已知的有效药物一样好(等于，不劣于)；或者(2)显示受试药物的效能优于阳性对照药。这种阳性对照也可用于以比较两种药物的疗效/安全性为主要目的的研究(参阅 1.4)。当用这种设计来证明等效/不劣于

或对药物进行比较时，就会遇到一个棘手的问题：这种试验是否能够鉴别活性药物和无活性药物(参阅 1.5)。

### **1.3.5 外部对照(包括历史对照)**

外部对照研究是将接受受试药物的一组对象与本研究以外的一组病人进行比较，而不是与分配到不同治疗组的相同人群病人组成的内部对照组进行比较。外部对照可以是早些时候接受过治疗的一组病人（历史对照）或是在同一时间但是在另一试验中的一组病人。外部对照可能是限定的(一组特殊的病人)或非限定的，即根据对结果的一般医学知识确定的比较组。利用后一种比较组是十分不可靠的(这种试验有时称为非对照试验)，因为一般印象中这种试验常常不准确。基线对照研究是指将受试者接受治疗后的状况与治疗前的状况相比较(例如血压，肿瘤大小)，没有内部对照和是不设对照的或是外部对照的一种（见 2.5 部分）。

### **1.3.6 多个对照组**

正如以下的进一步介绍(参阅 1.5.1)，常有可能在一项研究中采用一种以上的对照，而且很有好处。例如，采用阳性对照药和安慰剂。同样，试验可采用多种剂量的受试药以及多种剂量的阳性对照药，可以有或没有安慰剂。这种设计也可用于两

种相对强度尚未确定的活性药物的比较，或者是以确定相对强度为目的的试验。

## 1.4 临床试验的目的

应该区分临床试验的两个目的：(1)评价一种治疗的疗效和/或安全性；(2)评价两种治疗的相对疗效、安全性、效益/风险关系或用途。

### 1.4.1 疗效的证据

采用任何类型对照的一项研究，可能通过显示受试药优于对照药(安慰剂、低剂量、阳性药)证明受试药的疗效。此外，在某些情况下，阳性对照研究可能通过显示新药与已知的有效药物疗效相似而证明其有效。通过显示新药的疗效相似于某个标准药来证明该新药疗效，然而，只有在活性对照药物在试验中有效的前提下假设成立，因为如果两种药物在试验中都无效也可以有相似的表现（见 1.5 部分）。

通过证明一个新药同一个标准的制剂疗效相似来证明药物疗效的临床研究称为“等效性试验”。在多数情况下，实际上这些试验是非劣效性试验，试图显示新药的疗效与对照比较，不低于一个指定的量，疗效不差。

## 1.4.2 可比的疗效和安全性

在某些情况下，研究的焦点是与另一种药物进行比较，而不是受试药物本身的疗效。根据治疗的领域，可以将这些研究看成是为评价相对效益/风险提供所需的信息。所用的阳性比较药必须能被研究资料所在区域接受。根据情况，可能不需要证明其优于对照药物，和于之等效或非劣效；例如，一种疗效较差的药物可能在安全性方面有优点，因而被认为是有益的。

即使这种研究的主要焦点是对治疗进行比较，而不是证明疗效，也需考虑到实施和解释非劣效性试验的注意事项(参阅 1.5)。必要时，需要确定比较性试验检测治疗间差异的能力，因为一项不能将事实上不同的治疗加以区别的试验，是不能提供有用的比较性信息的。

## 1.4.3 比较的公平性

此外，对于可获取相对效益和风险资料的比较性试验，必须公平，也就是每种药物都必须有机会良好应用。实际上，为提供疗效证据而进行的阳性对照等效/非劣效性试验，也几乎总是需要与对照作公平的比较，因为任何对于对照药在研究中是否具有其通常作用的怀疑，都会削弱对试验分析灵敏度的信心(参阅 1.5)。需要指出的是，假如试验的目的是通过证明比对照优越来显示受试药的疗效(就是说，即使对照药没有很好使用，

受试药在试验中还是会显示其疗效，确不表示其比对照药优越)，那么，公平与否并不是一个问题。

在研究设计中，可能会偏向性地有利于其中一个治疗组的因素包括：剂量或病人人群的选择以及终点的选择和时间的确定。

#### **1.4.3.1 剂量**

在比较受试药物与阳性对照药物以便对相对效益/风险进行评价时，为对照药选择一个适当的剂量和用法是重要的。在检查两种药物的比较结果时，对以下的问题予以考虑是很重要的：是否使用了疗效显然较差而且剂量又太低的对照药，或者是否使用了耐受性显然较差而且剂量又太高的对照药。在某些情况下，为了能令人信服地显示出好的疗效或安全性，需要对对照药和/或受试药进行多种剂量的研究，除非受试药所选择的剂量优于对照药的任何剂量(或者是唯一推荐的剂量)，并至少有同样好的耐受性。

#### **1.4.3.2 病人人群**

阳性对照试验受试者的选择会影响结果：在评价试验说明什么问题时要仔细地考虑所研究的人群。例如，假如受试者是来自对标准药无反应的人群，那么就会产生有利于新药的偏倚。

这种研究结果不能推广到过去未经治疗的病人整个人群。不过，这种结果仍然是证明新药疗效的一个好证据。而且，在对其他治疗无反应的病人中进行的正式新药研究，将治疗失败者随机分配到新药组或无效治疗组(只要这样不会使病人发生危险)，能提供临床上有价值的观察，是体现新药在这些无反应者中的价值的极好证明(参阅附件)。

同样，有时可将对某一特定药物有或多或少的有利反应或不良反应的病人分为亚群。例如，黑人对 $\beta$ 阻断剂和血管紧张素转化酶抑制剂的降压作用反应不好，所以在这些病人中，用这些药物与新的降压药进行比较，就会倾向于显示新药的优越性。这样就不宜得出新药一定优越的结论。不过，在认识到其局限性及可能得出什么合适的结论的前提下，在亚人群中随机化进行的研究还是可以提供丰富的信息。

参阅附录中对“精选”研究设计的一般讨论，以及从总人群中选择一个亚组以增加研究的敏感性或回答一个专门而范围狭窄的问题。

### **1.4.3.3 终点选择和时间确定**

当使用两种药物治疗同一种疾病或同一种病情的时候，特别当它们是属于不同种类或不同方式的两种治疗时，它们可能对这种疾病的结果产生不同的影响。因此，假如在一项临床试验中对它们进行比较，不同的终点选择和时间确定就可能对不

同的治疗有利。例如，对于急性心肌梗死的病人，溶栓药可以降低死亡率，但会增加卒中的危险。假如将一种新的更加有效的溶栓药与一种比较老的溶栓药作比较，如果终点是死亡率，则新的溶栓药会显得更有效；但是，如果终点是死亡率和丧失能力的止痛药对牙痛的疗效时，将终点观察时间设定得比较早，起效比较快的药就会比持续长效减轻牙痛的药有利。

## 1.5 分析灵敏度

如上所述(1.4.1)，以阳性对照的非劣效性或等效性设计来证实疗效，有一个特别的问题，那就是不能说明两种药物之间的差别。通过显示新疗法不次于或等于已有的有效疗法来证实新药的疗效，或者更准确地说，通过显示它们两者之间的差异不大于限定的范围(界限)来证实新药的疗效，是建立在一个重要的假设之上的，那就是：假如这两种药物之间真地有差异，即假如新药的疗效比老药小得多或者没有疗效，那就不能得出结论认为两种药物之间疗效没有差异。这一假设又是建立在另一假设之上的，即阳性对照药在研究中有一定的疗效。假如这些假设不正确，就可能得出认为药物有效的错误结论，因为事实上非劣效性试验不会对此加以分析。某项试验中如果两种治疗之间存在差异，则测定这种差异的能力称为“检测灵敏度(assay-sensitivity)”。对于非劣效性试验，检测灵敏度要求阳性对照药在试验中有一定的疗效，只有这样，试验才有能力不会



将事实上是低劣的新药推断为不次于对照药。如上所述，由于试验中并不测定阳性对照的疗效的真实大小，因此只能推断检测灵敏度。本文中的“检测灵敏度”是某项试验属性，不同于“药效灵敏度”，后者的定义是：在某治疗领域内，设计并操作得当、使用某一阳性对照药(或具有类似疗效的其他药物)的试验，能在试验条件下可靠地反应受试药物最小药效的能力。药效灵敏度是根据历史经验确定的；通常是通过肯定在有足够功效时这类试验确能有规则地区别活性药物和安慰剂来确定。用这种方法确定药效灵敏度，意味着在设计 and 实施均良好的同类非劣效性试验中，有能力不会将无效的药物当成为不次于对照的药物。相反，检测灵敏度仅适用于某一试验，且要求阳性对照药真正有效，从而使试验真正具有能力不会将劣效药物宣布为非劣效药物。这种能力取决于具体试验的设计和实施的细节，以及具有药效灵敏度。

### **1.5.1 非劣效(等效)性试验的分析灵敏度**

现在使用的非劣效或等效试验的灵敏度分析可从一下两方面进行推断：

- 1) 药效灵敏度的历史性证据，如，过去类似的临床试验不断的区分出有效治疗同疗效差或无效治疗，和
- 2) 试验良好的进行，如，临床试验的实施不会削弱它区分有效治疗同疗效差或无效治疗的能力。

在开始非劣效性试验前，能够并应当评估药物疗效敏感性

的历史性证据。特别是，应当确定，在进行研究的特定治疗领域里，使用特定的活性对照药、或其他有相似作用的治疗的良好设计和进行的试验，可信地显示作用。最理想的情况是，准备用作活性对照的活性治疗已经被有力地证实优于安慰剂。如果情况如此，就有了类似设计的临床试验药效敏感性的历史性证据（见 1.5.1.1）。

临床研究的良好实施只有在活性对照非劣效性试验完成后可以进行评估。非劣效性试验不仅在设计上应当类似于决定药效敏感性历史证据的临床试验（如，入选标准、允许的伴随治疗）；而且，实际入选的病人群、伴随治疗等都应当进行评价，以保证试验的实施在事实上类似于以往的试验。试验应当高质量进行（如，好的依从性，失访的病人少）。同药效敏感性的历史性证据一起，良好的试验实施（1.5.1.2）给新的活性对照试验提供分析灵敏度的保证。

因此，非劣效性试验的设计和实施包含了一下四个关键步骤：

- 1) *确定存在药效敏感性的历史性证据*。如果这一点不确定，通过非劣效性证明药物疗效是不可能的，不可进行。
- 2) *设计试验*。试验的重要细节设计应当类似于决定药效敏感性历史证据存在的临床试验，如，研究人群，伴随用药，研究终点，导入期。

- 3) *选择界值*。应当确定一个可接受的非劣效性试验的界值，将历史性的数据和相关的临床的和统计学的因素考虑进去。
- 4) *进行试验*。试验应当严格按照历史性的试验进行，并且应当是高质量的。

#### 1.5.1.1 药效敏感性的历史性证据和选择非劣效的界值

如前所述，大多数阳性对照药的“等效”试验，其实是旨在确定药疗效的非劣效性试验。非劣效性试验结果的分析，在 ICH 文件 E9 和 E3 中加以讨论。简而言之，这种试验是将新疗法同已经确定的疗法作比较。在试验开始以前，要选定一个等效或非劣效的界值，有时称为“ $\delta$ ”。这一界值是划分受试药与对照药的差别是否达到统计学上打算排除程度。如果受试药与对照药之间的差的可信限排除试验药物劣效的程度等于或大于界值，就可以宣布受试药物不次于对照药物，因而是有效的；假如可信限范围包括了与界值一样大的差异，就不能宣布受试药物不次于对照药，而且不能认为受试药物有效。

给非劣效性试验选择的界值不能大于在计划进行试验的条件下与安慰剂相比阳性对照药有把握取得的最小疗效。假如阳性对照药与新药之间疗效的差异等于或超过界值，而且有利于对照药，则新药可能完全无效。所期望的活性药物的可靠的最小作用的确定，只有在存在药效敏感性的历史性证据时才可能，的确，确定界值需要证据。这条界值往往是根据以往在类似新试验条件下经过恰当设计的安慰剂对照试验确定的，但也可以

通过剂量反应或活性对照的有效研究获得支持。不管这些早期研究对照组的采用，在决定界值中更有价值的是测量活性治疗与对照比较的优效测量，不是非对照的测量，如同基线期比较的变化。必须注意，本文件并没有叙述如何准确地计算界值，而且也几乎没有人发表过如何计算的经验。

界值的确定要根据统计学推理和临床判断，要反映出赖以选定的证据的不确定性，而且应当适当保守。假如能恰当地做到这一点，当适当选定的界值不包括在新药与阳性对照药差异的可信限范围之内时，就可以保证新药的疗效大于零。实际上，选定的界值往往比表示阳性对照药最小期望疗效要小，因为这样可以使某些在临床上可以接受的疗效(或阳性对照药的部分疗效)仍然被保留。对于主要目的是要说明受试药与阳性对照药治疗上等效的试验也是一样(参阅 1.4.2)，这种试验通常是寻求证明受试药物与对照药物十分相似，而不是简单地证明新药有任何疗效。

界值的选定只有根据以往的经验这一事实，给了非劣效性试验一个历史性对照(外部对照)研究的成分。这种研究设计只有在对一种预期的药物疗效所作的历史性估计能够得到对照药的早先研究结果文献的充分支持时才是适用和可靠的。这些研究应当可以得出结论，在设计与其相似(病人人群、研究规模、研究终点、剂量、伴随治疗，等)的试验中，阳性对照药总是能够与安慰剂相区别，同时应当可以测出对照药的预期最小疗效的大小。假如类似设计的安慰剂对照试验不只是偶尔显示阳性

对照药与安慰剂无差异，而且不能用研究中的某些特点来加以解释，那么唯一可能的解释就是受试药物的优效。

根据历史经验得出的药效灵敏度是否适用于任何情况，从某种程度上看是一个判断的问题。不过，在许多情况下，在有良好对照的研究中，被认为有效的药物并不是都能显得优于安慰剂，因此，不能可靠地确定在某一试验中药物可能具有的最小疗效。这种情况往往包括安慰剂组出现实质性的改善和改变，和/或治疗作用小或易变，例如抑郁、焦虑、痴呆、心绞痛、症状性充血性心力衰竭、季节性过敏以及症状性胃食管反流疾病。

对所有这些情况，标准治疗无疑是有效的，因为对这些已经证明有效的药物中的每一种都进行过许多有良好对照的研究。然而，根据已有经验难以描述药物肯定会显示最小疗效的研究条件(即有药效灵敏度的条件)，以及能用来确定适当界值的条件。在某些情况下，赖以预测药效灵敏度的经验可能并不可靠，例如随着时间的推移，诊断标准和治疗标准发生了重大改变。如果使用非劣效设计不能支持为选定劣效界值而进行的研究的药效灵敏度，那么，非劣效的结果并不能被认为是证明了受试药的疗效或临床可比性/等效性。

值得注意的是，药效敏感性历史证据的确定只适用于一种特别设计的试验。对于一个计划的同药效敏感性类似的非劣效性试验，该试验在关键设计特性上同历史性试验相类似是非常

重要的。这些设计特性包括，如，入选标准（病情的严重程度、伴随的疾病状况、诊断方法），对照药物的剂量和剂型，采用的伴随治疗，终点的测量和评价时间，为排除特定病人群所采用的清洗期。当试验设计特征上不可避免差别时（如，由于技术或治疗的进步），应当仔细考虑采用不同的药效敏感性历史证据和选择界值。

#### **1.5.1.2 试验良好的实施**

即使历史经验表明，在某一特定治疗领域内进行的研究可能具有药效灵敏度，但是，在进行研究的特殊环境下，这种可能性会被削弱。为确保试验的分析灵敏性，试验的实施应当是高质量的和实际如组的病人、实际给予的治疗（除检验治疗之外的），和实际的评价应当类似于所依据的药效历史敏感性试验。

在试验进行中，有许多因素能减低在有效治疗、疗效差和无效治疗中被观察到的差别，从而降低试验的测定敏感性，例如：

- 1.对治疗的依从性差。
- 2.研究人群对药物作用的反应差。
- 3.同时应用的其他药物或其他治疗干扰了受试药物或减低了受试药物可能反应的范围或程度。
- 4.一部分受试人群有自发性改善疾病的倾向，因而不可能再有受试药物引起的改善。
- 5.诊断标准不明确(病人没有患有要研究的疾病)。

6.由于知道所有病人都服用了可能有效的药物而引起的对终点评价的偏倚，例如，在读血压反应时倾向于较实际血压低，这样就减少了受试药与对照药之间的差异。

临床研究人员和研究申办者希望进行高质量的研究，临床试验管理规范指导原则(ICH E6)的发布将会有助于研究质量的提高。不论怎样，必须明确，在要证明不同治疗的组间差异的试验中，很有必要采用良好的研究设计和尽量减少研究的误差，因为试验的不完善会使无法显示固存的治疗组之间差异的可能性增加。在安慰剂对照试验中，应当尽可能提高病人的依从性和病人人群对药物作用反应的可能性，以保证能将有效的治疗区别于安慰剂。然而，在许多临床研究中，虽然作了很多努力以确保研究的优异性和测定的灵敏度，但临床研究仍常常不能可靠地区别有效药物和安慰剂。

相反，在试图说明两种治疗之间差异不超过特定范围的试验(非劣效性试验)中，对确保差异的检测方面就不会太关注，因为不显示差异大于界值才是研究所期望的结果。某些研究误差会缩小观察到的治疗组之间的差异（如，依从性差、高安慰剂反应、一定的伴随治疗、结果额错误分级），对保护分析敏感度来说特别要关注。然而，当相信新药实际上优于对照药，就会有进行高质量研究的强刺激，非劣效性界值很可能被除外。同样应当注意到，某些研究误差却会加大变异，从而通过放宽可信限而减少显示非劣效的可能性，导致不能排除受试药与对

照药的差异大于阈值。因此，在这些研究中应采取一个有效的措施以减少变异，如可能由于测定技术差而引起的变异。

如所注意到的，确定一项非劣效性试验良好进行，不单要关注那些减低组间差异的因素，也要关注那些可能使试验区别于确定非劣效阈值基准的试验的因素。具体地说，应当确定是否存在入选人群、伴随治疗、治疗依从性、脱落的程度和原因的差异，影响到分析敏感度。即便当试验的设计和进行似乎类似于确定非劣效阈值基准的试验特，活性对照组的结果显示非同寻常（如，抗生素试验的治愈率异常的高或低）可提示重要的差异存在。

### **1.5.2 试图证明优效性的临床试验的敏感度**

分析敏感度的问题，尽管对非劣效性试验尤其的重要，事实上在任何不能检测出治疗组间区别的试验中都存在，包括安慰剂对照试验和剂量反应试验。如果一个治疗不能显示优于安慰剂，如，就意味着要么治疗无效，要么试验的设计和无法将有效治疗同安慰剂区分开。

在活性对照和安慰剂对照的临床试验中，一个有用的评估敏感度的方法是三组治疗试验（three-arm trial），包括安慰剂和一个已知的活性对照治疗，一种有许多优点的试验设计。这样的试验检测疗效大小（检测药物同安慰剂）和许可在一个确定了敏感度的系统下比较检测药物和活性对照治疗，敏感度是通过活性对照同安慰剂比较确定的（见 2.1.5.1.1）。



## 2. 对照类型的详细考虑

### 2.1 安慰剂对照

#### 2.1.1 描述(参阅 1.3.1)

在安慰剂对照研究中，受试者基本上是被随机分配到受试药或安慰剂组。安慰剂是一种“假”药，其物理特性例如颜色、重量、味道和气味都要尽可能地与受试药物相同。某些试验可能研究一个以上剂量的受试药，或者同时包括一个阳性对照药和一个安慰剂对照。在这些情况下，研究者应用一个以上的安慰剂（“双模拟”）可能要比做到使全部药物看上去都相同更容易。使用安慰剂有利于，也总是伴随于双盲试验。活性药物与安慰剂组之间所测得结果的差异就是在这种研究条件下这种药物作用的大小。在本节描述中，提出了可被成功采用的多种设计：平行或交叉设计(参阅 ICH E9)；活性药物组中的单一固定剂量或增量；以及数个固定剂量。值得特别注意的几种设计将介绍如下。需要指出的是，并非每一项含有安慰剂的研究都是安慰剂对照研究。例如，一项阳性对照研究可以为每一种药物采用一种安慰剂(双模拟)以利于设盲；这仍然是一项阳性对照研究，而不是安慰剂对照研究。安慰剂对照研究是比较安慰剂治疗与活性药治疗的研究。

应当注意并非所有的安慰剂都完全没有活性。如，局部皮肤使用药物研究中的某些中介体可能有一些有益的作用。这并不损害测量检测制剂特定效果的能力。当选择的中介体可能存

在不良的作用时，就会出现特别的问题。在这种情形下，‘无治疗’组被许可用来测量检测制剂和中介体的共同作用。

### **2.1.2 减少偏倚的能力**

采用随机化和盲法的安慰剂对照试验，一般能最大限度地减少受试者和研究者的偏倚，但是这种试验会受到由于通过识别出一种药物的药理作用而破盲的影响(也许在交叉设计中关系较大)；在这种情况下，盲态结果评价有利于减少偏倚。

### **2.1.3 伦理学问题**

一种新药用于尚无已知有效药物可以治疗的疾病进行试验同时，对新药和安慰剂进行比较研究通常不存在伦理学问题。是，如果在计划进行的试验中，被研究的疾病已有有效的治疗时，采用安慰剂对照可能会出现伦理学、可接受性以及可行性的问题。当现行治疗已知可以防止受试人群发生如死亡或不可逆转的病态等危重伤害时，一般不宜用安慰剂对照。但是偶尔也有例外，例如，标准治疗有严重的毒性以致许多病人会拒绝治疗。

在其他情况下，当停用或延迟有效治疗不会造成大的健康风险时，即使可能会导致病人感到不适，但只要他们的参加是非强迫性的，而且他们对可能有的治疗以及延迟治疗的后果完

全知情,要求病人参加安慰剂对照试验可以认为是合乎道德的。但是,这种试验可能会出现重要的实际问题。例如,病人或医生可能不接受对疼痛或其他症状的延缓治疗,他们可能不愿参加这样的研究。在有有效治疗的情况下,对新药进行安慰剂对照试验是否能被受试者和研究者所接受是一个由研究者、病人和 IRB/IEC 判断的问题,而且,可接受性在不同的 ICH 区域会有所不同。可接受性可能取决于研究的特别设计以及所选择的患者人群,这将在下面讨论(参阅 2.1.5)。

在某些情况下,一项具体的安慰剂对照试验是否符合伦理规范,可能取决于临床上是否认为已经证实,和试验的特定环境。例如,在轻度原发性高血压且不伴有终末器官疾病的病人中,进行新降压药的短期安慰剂对照研究,一般可以考虑接受,如果是长期的或要有较严重的病人参加的研究,则原则上不会被接受。

必须注意,采用安慰剂对照或无治疗对照并不意味着病人完全没有接受任何治疗。例如,在肿瘤研究中,如果是没有被批准的活性药物,在安慰剂/无治疗组以及受试药物组的病人都接受必要的姑息治疗,例如镇痛剂,和更好的支持护理。许多安慰剂对照试验被作为附加(add-on)试验进行,所有的病人接受一个标准的特定治疗或由研究医生/或机构选择的治疗(见 2.1.5.2.1)。

#### **2.1.4 安慰剂对照研究的用途和在特定情况下推论的质量/有**

## 效性

当用于体现一种药物的有效性时，安慰剂对照试验没有假设和不需要外部的信息(额外研究)。大多数研究设计上的问题和疏忽所引起的失误会导致难以证明治疗差异(从而不能确定疗效)，所以，对于一项优秀的研究，试验本身的因素很重要。即使研究的主要目的是比较两种活性药物，或者是评价量效关系，加入安慰剂就提供了一种内部标准，有助于根据其他对照药所进行的推断。

安慰剂对照试验也为区别药物不良反应和来自潜在疾病或并发疾病的不良反应提供了最大能力。应注意当安慰剂对照试验用于反应相似性时，如在证明没有不良反应时，与任何等效性或非劣效性试验一样，有检测灵敏度问题；在解释结果时，必须明确如果研究药物引起不良反应，应当已被观察到。

### **2.1.5 设计方案的修改以及结合其他对照药解决伦理的、实际的或推理问题**

通过对研究设计进行修改而仍保留其推断的优势，常常可以解决安慰剂对照试验中伦理或实际的局限。此外，通过加入、其他治疗组，例如多剂量的受试药物或已知的活性对照药物，可以从安慰剂对照试验中得到更多的信息。

#### **2.1.5.1 其他对照组**

#### 2.1.5.1.1 三组研究；安慰剂和阳性对照

正如第 1.5.1 节中提到的，当受试药物与安慰剂无法区别时，包括活性对照以及安慰剂对照组的三组研究，就可以容易地评价是试验药物无效还是仅仅由于研究缺乏鉴别活性药物的能力。这类试验中安慰剂-标准药物的比较可以提供有关检测灵敏度的内部证据。如果认为重要的话，可以设定活性药物组的样本大于安慰剂组，以改善活性药物比较的精确度。这一方法也会因为减少了随机分配到安慰剂组的机会，使研究对病人更具吸引力，

#### 2.1.5.1.2 其他剂量

除了安慰剂之外，几个固定剂量的受试药物组之间的随机化分配有助于评估量效关系，并特别有利于在比较试验中确保治疗间的公平比较。(参阅 IOF 曰：支持药物注册的量效信息)。

#### 2.1.5.1.3 析因/联合研究

析因/联合（反应面）设计可用于探索受试药单独使用时和与几种剂量的另一个药物合用时的几种剂量。单独一项这类研究可以确定多种排列组合的联合治疗的特征。这类研究常用于对新的抗高血压药物进行评价，但也可用于同时应用一种以上治疗的其他场合，例如，这类研究证明了阿司匹林与链激酶在预防心脏病发作后发生死亡的独立相加作用。

### 2.1.5.2 研究设计的改变

2.1.5.2.1 在标准治疗基础上的对照研究、安慰剂对照、替代研究

在标准治疗基础上的对照研究是指在接受标准治疗的人群中进行的新药的安慰剂对照研究。当已知标准治疗可以降低死亡率或不可逆转的病态时，不会对已知可以从中获益的人群停止治疗；当采用标准治疗作为阳性对照的非劣效性试验无法进行或难以对结果进行解释(参阅 1.5)时，这类研究是有效的。通常在研究抗癌药物、抗癫痫药物以及抗心力衰竭药物时采用这类方法。这一设计仅在标准药不完全有效时(然而，几乎总是如此)适用，其优点是可以证明临床结果有所改善(强于“只是”非劣效)。当然，这一研究中确立的疗效仅仅是联合用药时产生的疗效，而且单独用药时的剂量可能与联合用药时的有效剂量有所不同。一般来说，这种方法仅在新药与标准药物的药理学机制不同时适用，当然也有例外，例如，由于可以延缓耐药的发生，艾滋病的联合治疗可能显示了药理学相关药物的有利作用。

这一设计稍作修改后就成为替代研究，在时能获得单独用药的信息，特别适用于对慢性疾病治疗的研究；在这一设计中，将已接受有效剂量常规治疗的病人随机分组分别加用新药或安慰剂后，再通过逐渐减量撤除常规治疗，然后采用事先认定的有效标准观察受试药物和安慰剂维持受试者基线状态的能力。这一方法已用于在类固醇依赖性病人中采用替代物减少类固醇使用的研究，这样可以避免在停药洗脱期间初始的类固醇撤除以及症状复发。这一方法也已经用于研究抗癫痫药物的单独治疗。

#### 2.1.5.2.2 “及早脱离”；急救药物

可以为从无效治疗中“及早脱离”设计一项研究。及早脱离是指即时撤出那些临床状态恶化或改善未达到预定的水平(血压在预定时间内未得到控制，癫痫发作率大于规定的水平，血压升高到一定的水平，心绞痛发生的频率超过预设的水平，肝炎病人的肝脏酶未能在预定的时间恢复正常)的受试者、出现受试药物预计可以防止的异常受试者(不稳定心绞痛的首次复发，癫痫大发作，阵发性室上性心律不齐)，以及需要外加治疗的受试者。在这种情况下，研究终点变为需要改变治疗。必须详细设定这些终点是否出现的判断标准，而且测定的时间选择应当保证病人在疾病不能得到良好的控制时不至于没有活性药物进行治疗。这种试验设计的主要问题是它只能获得短期的有效性信息。而同样具有“及早脱离”特点的随机撤药试验(参阅2.1.5.2.4)，却可以获得长期的有效性信息。应该指出的是，当临床状况恶化时，同样可以使用常规的急救药物。

#### 2.1.5.2.3 有限的安慰时段

在长期活性对照试验中，短时间加入一组安慰剂对照能够确定检测灵敏度(至少对短期疗效而言)。随后，试验在没有安慰剂对照的情况下继续进行。

#### 2.1.5.2.4 随机撤药

在随机撤药研究中，接受一定时间受试药物治疗的对象被随机分配继续使用受试药物治疗或使用安慰剂(即停用活性药物)治疗。这种试验的受试者可以来自下列研究：有组织地开放

单组研究、一组现有的临床队列（但通常都要经过一个“清洗进入(wash-in)”期以确定开始治疗时的基线）、对照试验中的活性组，以及活性对照试验中的一组或全部两组。继续接受药物治疗组和安慰剂组之间出现的任何差异都可以证明活性药物的疗效。随机化前对药物的观察期无限制，从而当长时期安慰剂治疗无法接受时，这种方法可以用来观察长期的持续疗效。停药后的观察期可以是固定的时间，或采用及早脱离，或到发生事件的时间(例如抑郁症复发)。当采用及早脱离设计时，应特别注意对人的监控和对研究终点评价的程序，以确保迅速地发现没有接受指定治疗的病人。

随机撤药方法适合于以下几种情况。首先，它可适用于对付复发性疾病发作的药物(例如抗抑郁药)，在这种情况下，撤药研究实际上是一个防止复发的研究。其次，它可用于能够抑制症状或体征(慢性疼痛、高血压、心绞痛)的，而难以进行长期安慰剂对照试验的药物；在这种情况下该研究可以确立长期的有效性。再次，这一设计可以用于决定治疗的持续时间(例如，梗死形成后使用 $\beta$ 阻断剂进行治疗)。

随机撤药设计的优点是：当其采用诸如症状重新出现等“及早脱离”为终点时，病人必须度过的安慰剂使用阶段比较短。

这一类型的设计可以解决有关剂量的问题。当所有病人接受了一个初始的固定剂量后，在“停药期”可以将他们随机地分配到几个不同的剂量组(也可以是安慰剂组)。当由于根据药效动力学资料，或由于母体药物或活性代谢产物的半衰期长而



造成的活性药物大量蓄积，有理由相信初始剂量和维持剂量可能不同时，这一方法特别有用。值得注意的是，在初始的安慰剂对照滴定研究后，随机撤药设计可以用于评价剂量-反应。滴定研究对确定疗效是一种有效的设计，但不能提供良好的剂量-反应关系信息。在随机撤药期，将有反应者随机地分配到几个固定的剂量或安慰剂组，就可以精确地研究剂量-效应关系，且确保滴定设计的效率。

在采用随机撤药设计时，充分认识停药现象的可能性是重要的，明智的方法是相对缓慢地逐渐减少用药。病人可能对某一种药物产生耐药，这样药物不会产生效果，但药物的撤除可能会引起疾病加剧，从而对药物的持续疗效作出错误的结论。由于随机化停药研究中集中了有反应者(参阅附件)，因此在这些研究中观察到的疗效可能会比在一般人群中观察到的疗效要大，认识到这一点也是重要的。当研究中明确地仅包括对受试药物有反应的病人或仅包括完成了前一阶段研究的受试者时(它通常代表着一种良好的反应)，就会出现这一现象。

#### 2.1.5.2.5 其他设计考虑

在任何一种安慰剂对照研究中，不平衡的随机(例如受试药物与安慰剂的比例为 2: 1)能够加强安全性数据库，并能使研究对病人和/或研究者更具吸引力。

### 2.1.6 安慰剂对照试验的优点

#### **2.1.6.1 能可靠地证明药物的疗效**

就像其他一些显示差异的试验，在解释安慰剂对照研究时既不需要外在的药效灵敏度的假设，也不需要评价检测方法敏感性。当不能确定非劣效性研究是否具有检测灵敏度时，这些可能是唯一可靠的研究设计。

#### **2.1.6.2 检测“绝对”有效性和安全性**

安慰剂对照试验可以检测治疗的绝对作用，并可以区分不良事件是由药物引起还是由于某些潜在的疾病或“背景噪音”所造成。对于一个 3 组(试验组、安慰剂组、已知阳性对照组)试验，即使其主要目的是比较试验药物和阳性药物，有关药物绝对用大小的信息也是有用的。

#### **2.1.6.3 高效率**

安慰剂对照试验具有较高的效率，与其他任何并行对照研究相比，它只需要较小的样本数就可以检测出疗效。阳性对照试验可以体现某种新疗法的优势，但总的来说，这种研究所反应的组间差异比安慰剂对照试验中安慰剂与阳性对照组之间的差异要小，因而需要较大的样本数。非劣效性阳性对照试验也需要较大的样本数，因为它们对于对照药物的作用大小必须有一个保守的假设，以确保受试药物的非劣效性在事实上能证明药物的疗效。量效研究的设计者需要推测量效曲线的形状与位置，也可能不经济地将受试者分配到几个元反应的和位于反应平台的剂量组。

#### **2.1.6.4 受试者和研究者期望值的影响降到最低**

采用盲法安慰剂对照可能减少来源于受试者或研究者所期望的改善程度，因为他们都知道，有些受试者会接受非活性药物。这样可以使研究增加检测药物真实作用的能力。

### **2.1.7 安慰剂对照试验的缺点**

#### **2.1.7.1 伦理方面的问题(见 2.1.3 和 2.1.4)**

如果已知有某种有效的治疗方法能预防某一特定人群的损害，那么，从伦理上说，这一部分人群通常就不能接受安慰剂对照试验，是否的确如此还存在着争议。对伦理方面的关注也可以引导研究尽可能少地以患者作为受试者，或者，当研究对长期结果更有兴趣时，促使其考察短期的终点。当安慰剂对照试验不合乎道德，并且阳性对照试验不可信时，新药的研究就有可能非常困难。例如，要在曾发作过心肌梗死的病人中进行一项 $\beta$ 阻滞剂的安慰剂对照试验，就不认为是道德的，也难以得出非劣效性试验具有药效灵敏度的结论。在 2.1.5 中介绍的设计可能适用于其中的某些情况。

#### **2.1.7.2 病人和医师的顾虑**

即使双方都认为停止或延迟治疗不会导致伤害，医师和(或)病人可能仍不愿意接受某一病人将被分配到安慰剂组的可能性。自认为疾病没有得到改善的对象可能会退出试验，因为他们将无效归因于接受了安慰剂治疗，这使研究的分析复杂化。但应注意，有时可以将由于无效而退出试验作为研究的终点。

虽然，这样也能获得一些药物疗效方面的信息，但是，与受试者在临床状态下接受指定的治疗相比，这种信息的精确度要低一些。

### **2.1.7.3 通用性**

有争议认为，任何对照试验，特别是安慰剂对照试验，代表了一种人工的环境，由此得出的结果不同于真实的有效性。如果出于伦理或实际的考虑，在安慰剂对照试验中的研究人群无代表性，研究结果的通用性就会有问题。例如，具有严重疾病的病人可能因为试验方案，研究者或病人的选择等因素而被排除于安慰剂对照试验之外。在有些情况下，可能只有有限的病例数或中心愿意参加研究。至于这些考虑是否在实际上(相对于理论上)限制了研究结果的通用性还没有被确定。

### **2.1.7.4 无疗效比较资料**

缺乏阳性对照组的安慰剂对照试验几乎不能提供疗效比较方面的有利信息，这些信息在许多情况下都是倍受关注且非常重要的。交叉研究的比较无法可靠地获得这些信息，因为研究的条件可能是很不一致的。

## **2.2 无治疗并行对照(见 1.3.2)**

随机的无治疗对照的一般特性及其优缺点均与安慰剂对照试验相似。而与安慰剂对照试验不同是，它不可能是完全盲法，这就会影响试验的各个方面，包括受试者的保留，病人管理，以及所有的观察项目(见 1.2.2)。这种设计适合于这种情况：对

于一项安慰剂对照试验，只是因治疗本身差异很大，如放疗与外科手术，或者因为治疗的副作用非常不同而不可能采用盲法。当使用这种设计时，最好由对治疗分配处于盲态的观察者对受试者的合格性、终点测定或者管理中的某些改变作出及时的决定。与资料分析有关的决定，例如在分析数据集中所包含的病人，也应该由不知道治疗分配的人作出(进一步讨论见 ICH E9)。

## **2.3 量效并行对照(见 1.3.3)**

### **2.3.1 描述**

量效研究是指受试者被随机分配入若干剂量组中的一组，可以设或不设安慰剂对照组。量效研究可以建立剂量和疗效/不良反应之间的关系，并证明药物的疗效。前者将在 F4 中予以讨论；后者即是本指导原则的主题。药物疗效的证明可以通过配对比较研究中各剂量组之间以及各剂量组与安慰剂之间的显著性差异体现出来，或者即使两组间无显著差异，但随着剂量的增加，疗效也有显著增加的倾向。但在后一种情况下，可能需要进一步研究低剂量的有效性。正如 E9 中所述，用于分析主要药效的特定方法应该预先说明。

研究组间有剂型变化的试验，有同量效反应试验类似的问题。因为使用剂型对照试验来决定疗效的试验很少见，以下的讨论集中在量效反应试验。

在量效研究中设立安慰剂组(零剂量组), 会有以下一些优点。首先, 可以避免一些研究由于各剂量产生相似的作用而无法解释, 无法评价是否所有剂量都同样有效或同样无效。其次, 设立安慰剂组可以估计药物作用的绝对大小, 但当各剂量组都相对较小时, 这种估计可能不十分精确。其三, 由于药物与安慰剂组间的差异通常大于各剂量之间的差异, 设立安慰剂组可以减少样本数。不同剂量组的样本数不需要完全相同, 例如, 较大的样本数可以得到比较精确的低剂量疗效的信息, 也可以增加研究本身在证明预期最佳剂量的明确作用时的把握度。量效关系研究可以包括一个或多个剂量的阳性对照组。随机撤药设计也可以将对象分配到多个剂量水平。

### **2.3.2 减少偏倚的能力**

如果量效研究为盲法, 则和其他盲法设计一样, 能够减少受试者和研究者的偏倚。如果说药物的药理学作用可能打破某些病人或研究者的盲态, 那么, 量效关系研究可能比安慰剂对照试验更容易保持盲态。盲法试验需要多重模拟或一些外观一样的不同剂量的药物制剂。

### **2.3.3 伦理学问题**

量效研究的伦理问题或实际问题方面的考虑与影响安慰剂对照试验的问题相似。如果有一种治疗方法可以有效地防止死亡或防止不可逆性的发病，那么，从伦理上说，审慎地让患者随机分到疗效差的治疗并不比随机分到安慰剂组更易被接受。如果某种治疗措施是被用于不太严重的情况，或者药物的毒性与其治疗利益有实质性相关，那么，在量效关系研究中，病人和研究者可能可以接受较低的、效果略差剂量的或安慰剂治疗。

#### **2.3.4 量效研究的用途和特殊情况下推论的质量/有效性**

通常，当安慰剂对照研究适用并同样可靠时，盲法的量效研究可用于确定疗效和安全性(见 2.1.4)。

#### **2.3.5 修改设计和结合其他对照解决伦理的、实际的或推论问题**

通常，为减轻诸如伦理的、实际的或推论方面的问题，对安慰剂对照设计进行的修改也可以用于量效关系研究(见 2.1.5)

#### **2.3.6 量效试验的优点**

##### **2.3.6.1 高效率**

虽然，效率最高的是以完全有效的剂量组与安慰剂比较来证明药物的疗效，但这种设计可能产生难以接受的毒性，并且

不能够获得任何量效关系的信息。如果时关系是单一的，量效试验在证明药物疗效方面相当有效，并且也能获得量效关系的信息。如果药物的最佳有效剂量未知，那么较为谨慎的做法是选择一个剂量范围，而不是选择一个可能被证明是次佳或有毒性的单一剂量。

### **2.3.6.2 可能的伦理方面优点**

在有些情况下，特别是当药物的疗效和毒性与剂量相关时，量效研究可能代表一项在伦理上和实际上都可行的示差试验而此时甚至安慰剂对照试验也是不可行的，因为病人和研究者有理由接受较小的有效性以换取更大的安全性。

### **2.3.7 量效研究的缺点**

需要认识到的一个潜在的且有增加趋势的问题是：当药物的剂量与疗效结果之间有显著的相关性，而各剂量组间却无明显的配对差异时，这种相关性可以确定药物的疗效，但不能确定(除最大剂量之外)哪个剂量是实际上有效的剂量(当然，单剂量研究中对于低于所研究的剂量也有同样的问题)，因而无法获得这些剂量的任何信息。

另一个必须意识到问题是，在量效研究中，各剂量组之间经常没有差异；如果没有安慰剂组为疗效提供清楚证明，这种研究就是花费极大而没有任何结果的试验。



如果药物的治疗剂量范围完全未知，那么设计可能是无效的，因为许多病人可有会使用低于或超出治疗剂量范围的药物剂量。

在证明药物作用方面，量效关系研究设计可能没有安慰剂对照增量设计有效；但多数情况下，它们能获得较好的量效信息。（见 F4）

## **2.4 阳性药物对照(见 1.3.4)**

### **2.4.1 描述**

活性药物(阳性药物)对照试验是将受试药物与已知的活性药物进行比较。这类试验通常是随机和双盲的。设计试验最关键的问题是，该试验是用于证明两药之间的差异，还是证明非劣效性/等效性。申办者如果想利用一个相对于标准药物的非劣效性试验来说明药物的效果，就应如 1.5 中所讨论的那样，提出药效灵敏度以及试验检测灵敏度问题。在药物非劣效性/等效性试验中，阳性对照药物在所使用的剂量和研究条件下应该有确定的疗效(见 ICH E9 部分；临床试验中的统计原则)。通常，这就意味着该阳性对照药在申请研究的地区已被接受，且剂量和适应证与受试研究中的相同。另一方面，即使阳性对照药物的剂量太低或者阳性对照药的益处不明确(但不能有害)，试验药物的优效性研究也完全可以解释其疗效。但这种结果，即试

验中受试药物所表现的相对于对照药物的优效性，只有当阳性药物在最佳的给药剂量和给药程序下，用于合适的病人时才能判断为真正的(见 1.4.2)。如果由于不能确保研究的灵敏度而无法体现试验药物的优效性，那么，作为一个非劣效研究，如果没有合理地使用对照药物也会导致研究无意义。

### **2.4.2 减少偏倚的能力**

随机且盲法的阳性对照试验通常能最大限度地减小受试者和研究者的偏倚，但仍有必要提醒一点。在非劣效性试验中，研究者和受试者都清楚所有受试者会接受活性药物，但不知道具体用哪个药。这可能会导致对部分对象的评价产生偏倚的解释，也就是倾向于将效果处于临界状态的对象归入有效一类，如在抗抑郁药的研究中。这种偏倚可能减少变异和/或治疗间的差异，并因此增加得到等效的错误结论的可能性。

### **2.4.3 伦理学问题**

因为阳性药物对照试验中所有对象都接受活性药物治疗，所以一般认为这类试验所面临的伦理学问题及实际问题要比安慰剂对照试验少。但应该意识到，其中接受新药物治疗的受试者，并不是接受标准治疗(就好像安慰剂组一样)，也有可能接受的是一个无效的或有害的药物。如果已知阳性对照治疗能改

善存活或降低不可逆性损害时，这是一个重要的问题。因此，试验药物应有非常可靠的理论基础。如果没有充足的理由来预期新药至少与标准对照药物一样好，在条件许可时，采用在标准治疗基础上的对照试验(Add-on study)可能更合适(见 2.1.5.2.1)。

#### **2.4.4 阳性对照试验的应用及特殊情况下推论的质量/有效性**

当一个新药显示其优于阳性对照药物，且阳性对照药物确实无害时，该阳性对照研究就可以推断出其疗效相对于任何其他示差试验。当阳性对照试验是用于说明试验药物的非劣效性/等效性时，那么就应该特别考虑在 1.5 讨论过的药效灵敏度及检测灵敏度问题。如果已经确立了检测灵敏度，不管是以前建立的(参考以前对照药物的经验)还是同时设立安慰剂对照与阳性药物对照的试验，这种阳性对照试验就能评价药物的比较疗效。

#### **2.4.5 修改设计和结合其他对照解决伦理的、实际的问题或推论问题**

正如前面所讨论的(见 2.1.5)，阳性对照试验可以包括一个安慰剂组，多个试验药物剂量组，和/或其他阳性对照药物的剂量组。在止痛药试验中，同时有几个试验药物剂量组和几个阳性对照药物剂量组的比较性量效研究是非常典型的。在阳性对照药物研究中，剂量可以是固定的，也可以呈滴定形式，既可

以使用交叉设计，也可以使用平行设计。非劣效性试验的检测灵敏度有时可以从一个随机的安慰剂对照试验在结束时的撤药相获得(见 2.1.5.2.4)。在所选择的人群中(对其他治疗无响应者)的活性对照优效性研究是非常有用的，通常也易于解释(见附录)，但其结果可能不通用。

## **2.4.6 阳性对照试验的优点**

### **2.4.6.1 伦理的/实际的优点**

不论试验是否是为了证明非劣效性/等效性或优效性，阳性对照设计都能减少由于未使用被证明有重要健康利益的药物而引起的伦理方面的顾虑，也可以打消病人和内科医生因未能使用已证明有效的治疗方法而担忧。阳性对照试验有助于对象的招募和 IRB/IEC 对试验的批准，而且有可能进行较大样本的研究；并能减少由于缺乏疗效而退出试验的对象数量。

### **2.4.6.2 信息含量**

研究显示试验治疗的疗效优于阳性治疗时，阳性对照研究更易于解释疗效证据。所需要的较大样本量有时在阳性对照试验中更易达到和被接受，而且能提供更多的安全性信息。如果设计恰当，阳性对照研究也能提供关于相对疗效的信息。

## **2.4.7 阳性对照试验的缺点**

### **2.4.7.1 信息内容**

1.5 中讨论了非劣效性/等效性试验中检测灵敏度的问题以及试验对得出疗效结论的能力。甚至当方法灵敏度是可靠的，而且研究也适于检测疗效时，阳性对照研究仍不能直接评价绝对作用的大小，难以定量描述安全性结果。

### **2.4.7.2 样本量大**

在非劣效性试验中，通常较为保守地选择拒绝差异的界限值，首先是因为在试验中阳性对照的预计最小作用一般被作为阳性对照作用的估计值；其次，通常会有意排除多于对照药物作用的合理部分的丢失(loss)(见 1.5.2)，这都会减小差异界限。由于需要谨慎地估计对照药物的作用大小，样本量可能非常大。在显示差异的阳性对照试验中，两种药物之间的差异总是比药物与安慰剂间的预期差异小，而且常常是小得多，这也导致了大样本量。

## **2.5 外部对照(历史对照, 见 1.3.5 部分)**

### **2.5.1 描述**

外部对照试验是指对照组的病人并非属于受试组所在的同一随机试验，即不存在并行随机对照组。因此，对照组与接受治疗者并不完全来自同一人群。通常，对照组是早些时候(历史对照)在其他研究机构观察的且记录良好的病人，或在同一机构

观察但不属于同一研究的病人。外部对照研究可以是优效性研究或等效性研究。有时来源于一个较大规模研究的某些病人也可根据其与治疗组相似的某些特征被选择作为对照组；甚至可能为接受治疗的病人匹配特别的对照。

所谓“基线对照研究”是指多种外部对照试验；这些研究有时会被认为是使用“病人自身作对照”，但这在逻辑上是不正确的。事实上，对照组是对病人没有治疗的条件下可能发生的状况的一种估计。基线对照研究和那些使用更复杂的开-关-开(交叉)设计都属于这一类型，但不包括并行随机对照试验。值得注意的是，在这些研究中，总是将观察到的相对于基线的变化或研究周期期间的变化与所估计的无治疗干预时可能发生的状况相比较(至少是暗示)。通常是根据“总体认识”进行这种估计，且不参照特定的对照人群。尽管对于某些情况，例如当治疗作用很显著、接受治疗后立即发生、且不可能是自发作用时(例如：全身麻醉、心脏复率、可测量的肿瘤缩小)，这是十分合理的；但在多数情况下并不是很明显，这就需要寻找一个特定的历史经验。此类试验的设计者或分析者需要意识到这种对照的风险性，并且应该有充分理由支持对照的使用。

### **2.5.2 减少偏差的能力**

不能控制偏差是外部对照试验的主要和公认的缺陷，这一缺陷足以使该研究设计在许多场合不适用。要确立治疗与对照

组之间的可比性从而充分发挥对照组的主要作用(见 1.2)总是比较困难,在许多情况下是不可能的。除了研究药物之外,对照组与治疗组间有许多可能影响研究结果的因素可能不相似,这些因素包括人口学特征、疾病诊断标准和疾病的发展阶段或持续时间、伴随治疗以及观测条件(例如:评价结果的方法、研究者期望)。当使用外部对照时,不能用盲法和随机化来减少偏差。已经证明无治疗的历史对照组的结果往往比随机化研究中显然相似的对照组的结果更差,主要是因为选择上的偏差。在随机化研究中,对照组必须符合某些标准才能进入研究,这些标准比外部对照组的入选标准更为严格,而且能辨别比外部对照组患更轻度疾病的人群。确认外部对照组通常是回顾性的,这会导致选择上的潜在偏差。无法控制偏差的一个后果是,外部对照试验结果的说服力取决于获得比同期对照试验相更为极端(extreme)的统计显著性水平以及各治疗组间更大的估计差异。

无法控制偏差使外部对照设计限于在治疗作用十分显著以及疾病的一般过程是可以准确预测的条件下使用。另外,外部对照的使用应被限于终点是客观的,并且基线和治疗变量对终点的影响可准确描述的试验。

值得注意的是,缺少随机化和盲法,以及受试组和对照组之间缺乏可比性保障而导致的问题,可能使该设计本身存在实质性偏差并且不能定量。然而,一些设计和实施外部对照试验的具体方法也可能使试验更具说服力并减少偏差。应当选择有

详细信息的对照组，必要时包括各病例的人口统计学数据、基线状况、伴随治疗以及研究进程。对照病例应尽可能地与研究中预期接受试验药物治疗的人群相似，并且应在除研究治疗外的相同背景下以相同的方式接受治疗。研究观测方法在时间选择和方法学上应与在对照病例中使用的相似。为了减少选择误差，应在比较分析前选择对照组；这并非总是可行的，因为对照组的结果可能已经发表。任何选择标准的匹配或者由于人群差异所做的调整都应在选择对照和进行研究之前说明。当不存在一个明显的“最适宜”外部对照时，建议采用多个外部对照，只要在分析计划中适当地说明在作出推论时如何使用每一个对照(例如：研究组必须充分优于作用最强的对照组才能作出有效结论)。在某些情况下，让一组独立的评价人员根据通用标准以盲态的方式重新评价对照组和受试组中的终点是十分有用的。

### **2.5.3 伦理学问题**

当一种药物预期用于治疗尚无令人满意的治疗方法的严重疾病时，尤其是根据理论基础、动物数据或早期人体用药经验，该新药被认为非常有前途时，就会理解不愿意进行一个没有病人不能采用新疗法的同期对照组的比较研究。同时，运用没有任何实际机会获得可信疗效的治疗来实施研究是不负责的或不人道的。应该预料到许多有前途的治疗其疗效没有预期的那



么好，或者在对照试验中显示无效，在这些情况下研究者可能会面临极难以作出判断的局面。开始实施一个外部对照试验，希望获得令人信服的显著作用，在没有成为事实时立即转向随机化试验，这在一些特殊情况下是十分诱人的(tempting)。

在处理没有满意治疗方法的严重疾病，但疾病的进程也不能很好地预测时，通常较可取的，也是较完美的方法是最早期的研究也应当是随机。当研究是在有“治疗是有效的”这一印象之前实施时，这通常是可能的。研究应在独立的数据监察委员会的监控下进行，这样可尽早地检测到治疗的显著利益。尽管在外部对照研究中使用单一治疗组，在这种情况下安慰剂对照试验通常是较为有效的设计(需要较少的对象)。因为一般必须保守地估计对照组结果，这需要较大的样本规模。由于在治疗组和对照组之间可能存在着确定的和非确定的或不可测的且常常是有利于治疗组的差异，所以必须引起高度注意(例如：使用更严格的显著性水平)。同期对照试验能快速检出极端效应，另外，可以检出外部对照试验不能准确反应的较轻微的但仍有价值的作用。

#### **2.5.4 外部对照试验的用途以及在特别状态下推论的质量/有效性**

只有试验进行前就非常肯定地相信受试治疗优于所有现有的治疗，以至于其他设计不可接受，而且被治疗的疾病或状态

有良好的记录和高度可预测的进程时，通常才会考虑使用外部对照试验。即使在这些条件下，也很有可能选择性使用随机、并行对照设计(见 2.1.5 和附录)。

如果研究终点是客观的，治疗组结果与外部对照结果有显著差异，而且治疗-对照之间的比较有高水平的统计显著性，能很好地描述影响疾病结果的协变量的特征，对照组和研究组在所有已知相关的基线水平、治疗(研究药物除外)和观测变量都十分相似，这时外部对照设计是极可能有说服力的。然而，即使在这种条件下，也有外部对照试验导致错误结论的实例记录。

当考虑使用外部对照试验时，适当注意设计和操作有助于减少偏差（见 2.5.2）。

### **2.5.5 修改设计以及合并使用其他对照能解决伦理的、实际的或推论问题**

如上文所述，随机安慰剂对照的撤药通常与及早脱离规则相关(见 2.1.5.2.4)，而通过运用这一撤药相，可以使外部对照设计结合随机化和盲法的一些原理。因此，早期治疗的结果(鉴别对治疗有响应的对象并继续接受治疗)可通过一个严格的基本无假设、无偏差的研究来证实。

### **2.5.6 外部对照试验的优点**

外部对照试验的主要优点是所有病人都能接受有前途药物的治疗，这样可使研究对病患和医生有更强的吸引力。

因为所有病人都暴露于试验药物，试验设计具有潜在的科学性(需要较小的样本量)，这对一些稀有疾病尤其重要。但是，尽管在外部对照试验中采用单治疗组，应当保守地估计外部对照组的結果，可能会导致需要比安慰剂对照研究大的样本量。由于治疗组和对照组间的差异可能测定，也可能无法测定，需要特别小心（如，采用一个更加严格的显著性区间），结果通常有利于治疗组。

### **2.5.7 外部对照试验的缺点**

外部对照试验不能采用盲法设计，因此存在由病人、观测人员和分析人员引起的偏差，这是主要缺点。这些问题可能在某种程度上能够缓和，但即使是 2.5.2 所推荐的方法也不能彻底解决这些问题，因为治疗分配不是随机化的，开始治疗时对照组与治疗组的可比性以及试验期间受试者的可比性得不到保证或者不能很好地进行评价；外部对照试验有过高估计治疗有效性的趋势，这一点已经得到很好的证明。

## **3. 选择对照组**

图 1 和表 1 描述了在不同类型对照组中进行选择的决定程序。尽管图表集中说明了如何选择用于疗效分析的对照，一些

设计也允许试验药和对照药的配对。某些区域内是否有这种疗法或医学技术的限制，会影响对照的选择。

特定条件、特定目的下主要对照类型(安慰剂、阳性对照和剂量-效应对照)的可能用途见表 1;应该结合描述实现对照可能用途的特定条件细节的上下文来使用该表，假定所有条件下的设计都是适宜的。外部对照是特殊情况因此没有列入表中。该表中记号 P 表示需要证明研究具有分析灵敏度。

通常，证明优于同期对照治疗是最有说服力的疗效证据。如果由于伦理学或实际原因优效性试验不可行或不适宜，而且能够有规则地观察到阳性对照确定的治疗作用(例如在大多数条件下抗生素的作用)，可以使用非劣效性/等效性研究，而且是有说服力的。

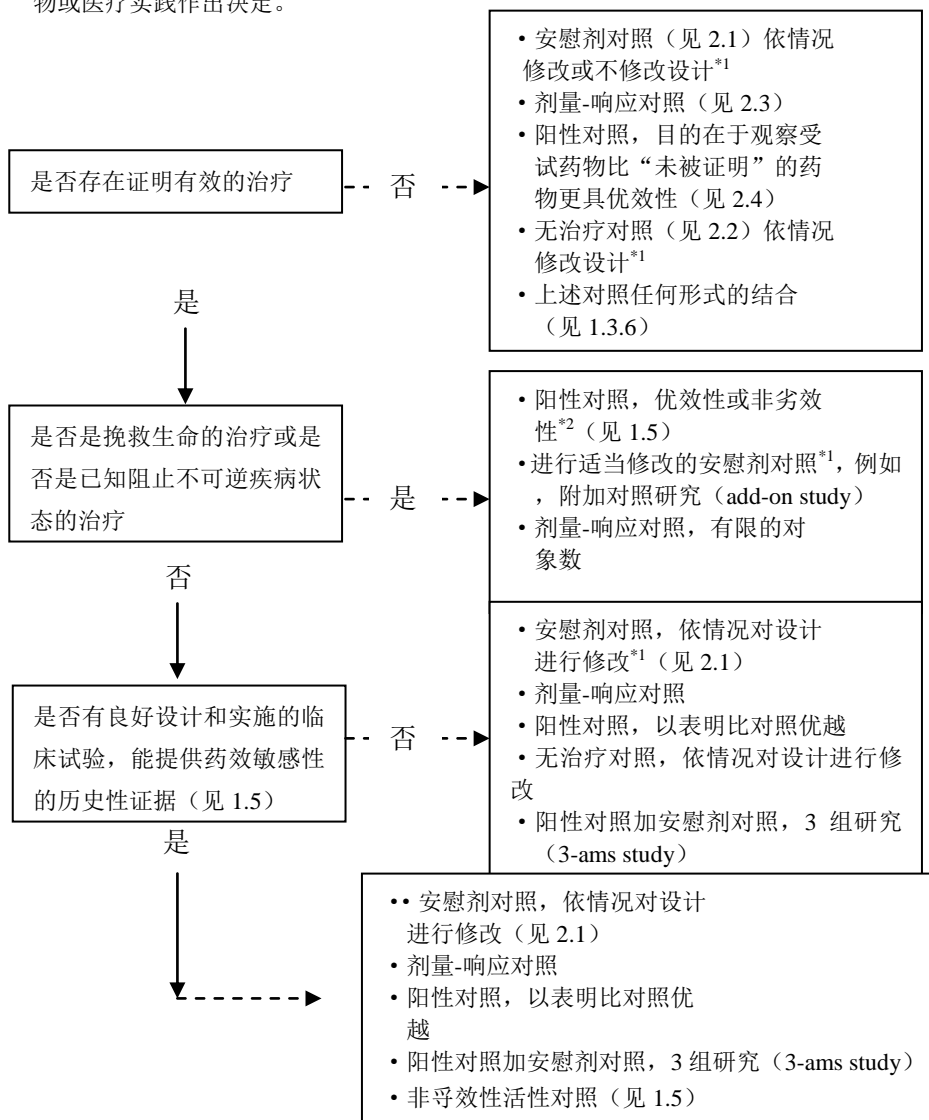
表 1 不同情况下各种对照类型的适用性

试验目的	对 照 类 型							
	安慰剂	阳性对照 非等效性	阳性对照 优效性	剂量效应 (D/R)	安慰剂+ 阳性对照	安慰剂 +D/R	阳性对 照+D/R	安慰剂+阳 性对照 +D/R
衡量绝对作用大小	Y	N	N	N	Y	Y	N	Y
观察是否有作用	Y	P	Y	Y	Y	Y	Y	Y
观察量效关系	N	N	N	Y	N	Y	Y	Y
治疗间的比较 (观察优效性、观察非劣效性)	N	P	Y	N	Y	N	P	Y
检测方法灵敏度*	Y	N	Y	Y	Y	Y	Y	Y

Y=是, N=否, P=可能, 根据是否有药物作用灵敏度的历史性证据

### 图 1 为观察疗效选择并行对照

该图表明选择对照组的基本逻辑;可以根据特定区域中可获得的药物或医疗实践作出决定。



\*1 附加对照、替换对照、早期脱离、简略安慰剂周期对照、以及随机撤出

对照（见 2.1.5.2）

\*<sup>2</sup> 如果已知能敏感地检出药物作用