

# 国际人用药品注册技术要求协调会

## ICH协调指导原则

### 《临床试验的统计学原则》指导原则的增补 临床试验中的估计目标与敏感性分析 E9(R1)

最终版本

2019 年 11 月 20 日采纳

该指导原则由相应的 ICH 专家工作组制定，按照 ICH 进程，已通过药品监管机构讨论。在 ICH 进程第四阶段，推荐 ICH 地区监管机构采纳最终草案。

## E9(R1)

## 文件历史

编码	历史	日期
E9(R1)	在第4阶段，ICH代表大会监管机构会员采纳（文件日期：2019年11月17日）	2019 年 11 月 20 日
E9(R1)	在第2阶段，ICH代表大会批准并发布公开征求意见。	2017 年 8 月 30 日

**法律声明：**本文件受版权保护，在始终承认ICH版权的前提下，基于公共许可可以使用、复制、在其他作品中引用、改编、修改、翻译或传播。如对本文件进行改编、修改或翻译，必须采取合理措施来清晰地标明、界定或以其他方式标记对本文件做了改变。应避免对本文件的改编、修改或翻译由ICH认可或发起的任何误导。

本文件根据现有内容提供，不作任何保证。ICH或本文作者在任何情况下均不对使用本文件产生的索赔、损失或其他责任负责。

上述许可不适用于由第三方提供的内容。因此，对第三方拥有版权的文件需获得版权所有人的复制许可。

## ICH 协调指导原则

### 《临床试验的统计学原则》指导原则的增补

### 临床试验中的估计目标与敏感性分析

#### ICH E9(R1)

#### ICH 共识指导原则

#### 目录

<b>A.1. 目的和范围</b>	<b>1</b>
<b>A.2. 将计划、设计、实施、分析和解释协调一致的框架</b>	<b>4</b>
<b>A.3. 估计目标</b>	<b>5</b>
A.3.1. 临床问题中反映的伴发事件	6
A.3.2. 在定义临床问题时解决伴发事件的策略	8
A.3.3. 估计目标的属性	12
A.3.4. 构建估计目标的考量	13
<b>A.4. 对试验设计和实施的影响</b>	<b>18</b>
<b>A.5. 对试验分析的影响</b>	<b>21</b>
A.5.1. 主要估计	21
A.5.2. 敏感性分析	24
A.5.2.1. 敏感性分析的作用	24
A.5.2.2. 敏感性分析的选择	25
A.5.3. 补充分析	25
<b>A.6. 估计目标和敏感性分析的记录</b>	<b>26</b>
<b>词汇表</b>	<b>28</b>

## A.1. 目的和范围

为了给制药公司、监管机构、患者、医生和其他利益相关方的决策提供正确的信息，应明确描述特定医疗条件下治疗（药物）的获益和风险。如果不能对此进行明确描述，报告的“治疗效应”可能会被误解。本增补提出了一个结构化的框架，以加强参与制定临床试验目的、设计、实施、分析和解释的多学科间的交流，并加强申办方和监管机构之间关于临床试验中治疗效应的沟通。

构建相应临床问题的“估计目标”（见词汇表；A.3.）有助于精确描述治疗效应，这就需要深思熟虑地定义“伴发事件”（见词汇表；A.3.1.），如终止分配的治疗，使用额外或其他治疗，或终末事件（如死亡）等。估计目标的描述应该反映出与这些伴发事件相关的临床问题，并且本增补介绍了反映不同临床问题的策略。在描述临床问题时，策略的选择可能会影响到如何反映试验的更加常规的属性，例如治疗、人群或相关的变量（终点）。

临床试验数据的统计分析应当与估计目标对应。本增补阐明了“敏感性分析”（见词汇表）在探索主要统计分析结论稳健性中的作用。

本增补中，对原始 ICH E9 的引用采用 x.y 格式，对本增补的引用采用 A.x.y.格式。

本增补就以下若干方面澄清和扩展了 ICH E9。第一，ICH E9 介绍了随机对照试验中对应于疗法策略的意向治疗（ITT）原则，据此对受试者进行随访、评估和分析，而不考虑其是否依从计划的治疗过程，这表明保持随机化为统计学检验提供了一个坚实的基础。ITT 原则具有以下三个含义。首先，试验分析应包括与研究问题相关的所有受试者。其次，受试者应按随机化时的分配纳入分析。最后，根据 ITT 原则（见 ICH E9 词汇表）的定义，无论是否依从预定的治疗过程，都应对受试者进行随访和评估，并在分析中使用这些评估。毫无疑问，随机化是对照临床试验的基石，分析时应最大限度地利用随机化的这一优势。然而，根据 ITT 原则估计治疗效应能否总是代表与监管和临床决策最相关的治疗效应，这个问题仍然悬而未决。本增补中概述的框架为描述不同的治疗效应提供了基础，并提出了试验设计和分析需考虑的要点，以便估计治疗效应，为决策提供可靠依据。

第二，本增补重新审视了通常归为数据处理和“缺失数据”（见词汇表）的一些问题，并提出了两个重要的区别。首先，增补对终止随机分配的治疗和退出研究加以区分。前者代表一个伴发事件，需通过在试验目的中对估计目标的精确说明加以解决；后者导致缺失数据，需在统计分析中加以解决。例如，考虑在肿瘤学试验中转组治疗的受试者，以及由于试验完成而无法观测到结局事件的受试者。前者代表伴发事件，关于该事件的临床问题应明确。后者属于管理性删失，需要在统计分析中作为缺失数据问题加以解决。估计目标的清晰性为计划需要收集哪些数据提供

了依据，以及哪些数据如果未被收集到即为缺失数据问题，需要在统计分析中加以解决。然后，可以选择解决缺失数据问题的方法，以与估计目标一致。其次，增补强调了不同伴发事件的不同影响。诸如终止治疗、转组治疗或使用额外药物等事件可能导致变量的后续观测值即使可以收集到数据也与估计目标不相关或难以解释。而对于死亡的受试者，死亡后的观测值是不存在的。

第三，在框架中考虑了与分析集概念相关的问题。第 5.2.节强烈建议优效性试验的分析基于全分析集，即尽可能包括所有随机化受试者的分析集。然而，试验往往包括对同一受试者的重复观测。某些受试者按计划收集的观测值可能被认为是无关的或难以解释的，剔除这些观测值，与从全分析集中完全剔除受试者可能具有类似的后果，即没有完全保留最初的随机化。这样做的一个后果是，随机化赋予关于治疗效应的检验假设的理论优势获益以及平衡基线混杂因素的实际获益可能被削弱。另外，有意义的结局变量取值可能不存在，例如当受试者已死亡。第 5.2.节没有直接阐明这些问题。这些问题要在考虑伴发事件的前提下，通过仔细定义关注的治疗效应来进行明确，既要确定要包括在治疗效应估计中的受试者人群，又要确定每个受试者包括在分析中的观测值。本增补也重新审视了使用符合方案集来分析的意义和作用，尤其是，是否需要用比分析符合方案集更能减少偏倚、更有可解读性的方式，来研究方案违背和偏离的影响。

最后，在敏感性分析部分进一步讨论了稳健性的概念（见 1.2.）。特别区分了所选分析方法的假设的敏感性，以及分析方法选择上的敏感性。通过精确说明已达成共识的估计目标，以及与估计目标一致的分析方法且其预先设定的细节描述达到能使第三方精确地重现分析结果的程度，这样，监管机构对于一个特定分析可聚焦于假设偏离和数据局限的敏感性。

无论是基于有效性或安全性的治疗效应估计，还是对治疗效应相关假设的检验，本增补中概述的原则均适用。虽然主要关注的是随机临床试验，但这些原则也同样适用于单臂试验和观察性研究。该框架适用于任何数据类型，包括纵向数据、首次事件发作时间数据和复发事件数据。对于确证性临床试验和用于产生确证性结论的跨试验整合数据，监管部门对所述原则的应用将更为关注。

## A.2. 将计划、设计、实施、分析和解释协调一致的框架

试验计划应按顺序进行（图 1）。应通过定义合适的估计目标，将明确的试验目的转化为关键的临床所关注的问题。估计目标根据特定的试验目的定义估计的目标（即“要估计什么”，见 A.3.），然后可以选择合适的估计方法（即分析方法，称为主“估计方法”，见词汇表）（见 A.5.1.）。主估计方法将以特定假设为基础，为了探索根据主估计方法所作推断对偏离其基本假设的稳健性，应针对同一估计目标采用一种或多种形式进行敏感性分析（见 A.5.2.）。

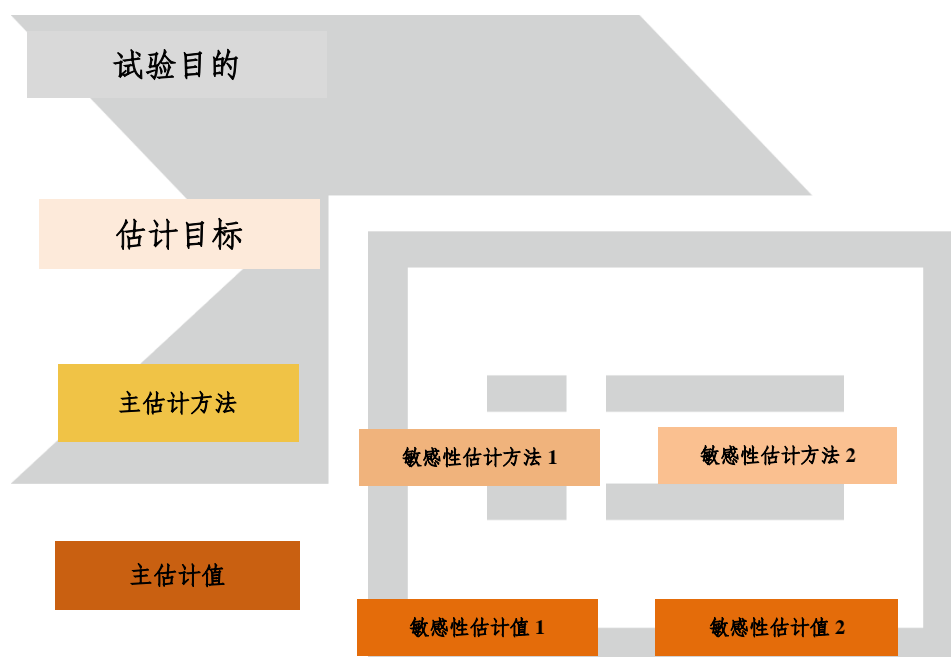


图 1：协调估计的目标、估计的方法和敏感性分析，使其与给定试验目的对应

该框架有助于制定适当的试验计划，以明确区分估计的目标（试验目的，估计目标）、估计的方法（估计方法）、数值结果（“估计值”，见词汇表）和敏感性分析。这将有助于申办方的试验计划制定和监管机构的审评工作，并在双方讨论临床试验设计的适宜性和临床试验结果的解释时增强交流。

指定适当的估计目标（见 A.3.）通常是试验设计、实施（见 A.4.）和分析（见 A.5.）方面的主要决定因素。

### A.3. 估计目标

药物开发和批准的核心问题是明确治疗效应是否存在，并估计其大小：如何比较相同受试者接受不同治疗的结局（即，如果受试者未接受治疗或接受不同治疗）。估计目标是对治疗效应的精确描述，反映了既定临床试验目的提出的临床问题。它在群体层面上总结了同一批患者在不同治疗条件下比较的结果。估计的



目标将在临床试验之前定义。一旦定义了估计的目标，即可设计试验以可靠地估计治疗效应。

估计目标的描述涉及特定属性的精确说明，这些属性不仅应基于临床考虑而制定，还应基于所关注的临床问题中如何反映伴发事件。第 A.3.1.节介绍了伴发事件。第 A.3.2.节介绍了各种策略，来描述与伴发事件有关的问题。第 A.3.3.节描述了估计目标的属性，第 A.3.4.节则提出了估计目标构建的考虑要点。理解不同策略之间的差异，并精确阐明哪些策略用于构建估计目标，这一点至关重要。

#### A.3.1. 临床问题中反映的伴发事件

伴发事件是指治疗开始后发生的事件，可影响与临床问题相关的观测结果的解读或存在。在描述临床问题时，有必要阐明伴发事件，以便准确定义需要估计的治疗效应。

在描述治疗效应时需要考虑伴发事件，因为变量的观测结果可能受伴发事件的影响，而伴发事件的发生可能取决于治疗。例如，两名患者可能最初暴露于相同的治疗并提供相同的结局观测值，但如果其中一名患者接受了其他药物治疗，则两名患者之间，观测值所反映的治疗的信息会有所不同。此外，患者接受的治疗会影响到他们是否需要服用其他用药，以及是否可以继续接受治疗。与缺失数据不同，伴发事件不应被认为是临床试验中需要避免的缺陷。在临床试验中发生的终止既定治疗、使用其他药

物和其他此类事件在临床实践中也可能发生，因此在定义临床问题时需要明确考虑这些事件发生的可能。

可影响观测结果解释的伴发事件包括终止所分配的治疗，和使用额外或其他治疗。使用额外或其他疗法可以有多种形式，包括改变基础治疗或合并治疗、转组治疗。影响观测结果存在的伴发事件包括终末事件，例如死亡和腿截肢（当评估糖尿病性足溃疡的症状时），而且这些事件不是变量本身的一部分。当某些临床事件的发生或不发生定义了一个主层时（见 A.3.2.），这些事件也可以是伴发事件。例如，肿瘤领域中在评估缓解持续时间疗效时定义客观缓解的肿瘤缩小；对于初始未感染的接种疫苗受试者在评估感染严重程度疗效时的感染发生。

伴发事件可能仅由事件本身确定，如终止治疗，或可能有更详细的定义。详细的定义例如，可明确说明事件发生的原因，如因毒性作用终止治疗，或因缺乏疗效而终止治疗；事件可能需要达到一定量级或程度，如使用超过规定时间或剂量的其他药物；或明确说明事件发生的时机，可能与其对变量评估的接近程度有关。一些事件会无限期地影响结局观测值的解释，例如终止治疗，而另一些事件只会暂时影响，例如短期使用其他治疗。事实上，额外或其他治疗可以是多样的；可以是替代或补充受试者获益不足时的治疗，或作为对既定治疗不耐受的另一种选择，或作为控制疾病暂时急性发作的短期急性治疗。在临床试验中，额外或其他治疗通常是指诸如基础治疗、补救药物和禁用药物，要区

分它们的不同作用以对其分别考虑。如果要使用不同的策略，则需要额外的详细信息，确定不同的伴发事件。例如，如果伴发事件不仅取决于未继续治疗，还取决于与未继续治疗相关的原因、程度或时机，则应在临床试验中准确定义和记录该附加信息。理论上，描述伴发事件可能体现治疗和随访非常具体的细节，例如长期治疗的单次漏服或日间服药的错误时间。如果预期这些具体标准不会影响对变量的解释，则不需要将它们作为伴发事件处理。

如上所述，在构建估计目标时需要考虑伴发事件。因为估计目标要在试验设计之前进行定义，所以无论是退出研究还是其他缺失数据的原因（例如生存结局的试验中的管理性删失）本身都不是伴发事件。退出试验的受试者在退出前可能已经发生了伴发事件。

### A.3.2. 在定义临床问题时解决伴发事件的策略

下面列出了多种策略，每种策略又体现了对于特定伴发事件的不同临床问题。无论是否使用如下命名规则，构建估计目标时策略的选择都必须清晰明确。无需使用相同的策略来处理所有的伴发事件。事实上，通常会使用不同的策略来明确体现不同伴发事件的临床问题。第 A.3.4.节给出了一些在构建估计目标时策略选择上的考虑。

### 疗法策略

疗法策略下伴发事件的发生与定义治疗效应无关，即无论是否发生伴发事件，均会使用相关变量的值。例如，将使用其他药物治疗作为伴发事件时，规定无论患者是否服用其他药物，都使用相关变量的值。

对于患者是否继续治疗以及患者的其他治疗（基础或合并治疗）是否有变化等伴发事件，在疗法策略中被视为治疗的一部分。基于这种情况的比较就体现了 ICH E9 所阐述的 ITT 原则，比较结果亦是疗法策略下的治疗效应。

一般情况下，对于终末事件类型的伴发事件，不能采用疗法策略，原因是该类伴发事件后变量的值不再存在。例如，在死亡之后变量是无法观测的，因此不能基于此策略构建估计目标。

### **假想策略**

假想策略设想一种没有发生伴发事件的情景：此时，体现临床问题的变量值是在所假设的情景下采用的变量值。

存在各种各样的假设情景，但其中有些情景更具临床或监管意义。例如，在与可实施试验条件不同的条件下的治疗效应可能具有临床或监管重要性。具体而言，当出于伦理原因必须提供额外药物治疗时，可能要考虑未提供额外药物情形下的治疗效应。一个非常不同的假设情景可能是假定伴发事件不会发生，或者会发生不同的伴发事件。例如，对于因发生不良事件而终止治疗的受试者，可考虑同一受试者没有发生不良事件或即使发生不良事

件仍然继续治疗的情景。这种假设情景的临床和监管意义有限，并且通常需要清楚地理解伴发事件或其后果在临床实践与临床试验中为什么不同以及如何不同。

如果提出了一个假想策略，应该明确具体的假设情景是什么。举例来说，诸如“如果患者未服用额外药物”之类的措辞可能会导致混淆，因为不清楚患者是因为没有额外药物可用而未服用，还是该患者不需要服用额外药物而未服用。

### **复合变量策略**

复合变量策略与关注的变量有关（见 A.3.3.）。伴发事件本身可提供关于患者结局的信息，因此将其纳入变量的定义之中。例如，由于毒性而终止治疗的患者可能被认为治疗失败。如果变量已被定义为成功或失败，因毒性终止治疗将被认为是另一种形式的失败。复合变量策略不仅限于二分类变量，也可以是连续型变量。例如，在观测生理功能的试验中，死亡的受试者可以用某一数值代表生理功能缺失。当变量原始观测值可能不存在或没有意义，但是伴发事件本身能够体现患者结局（如患者死亡）时，可将复合变量策略视为遵循意向治疗原则的策略。

终末事件，如死亡，可能是需要采用复合策略的最突出例子。如果某种治疗可以挽救生命，可能会关注其对存活患者的各种指标的作用，但是，如果汇总指标仅关注存活患者的一些数值指标的平均值是不够的，要同时关注数值指标和是否生存。例

如，肿瘤试验中的无进展生存期衡量了肿瘤生长和生存组合在一起的治疗效应。

### **在治策略**

在治策略关注在伴发事件发生之前的治疗效应。该策略的具体术语将取决于相关伴发事件；例如，当将死亡视为伴发事件时，可以称为“在世策略”。

如果一个变量被重复测量，则伴发事件发生前的所有观测值都可能被认为与临床问题相关，而不是所有受试者在相同固定时间点的值。这也适用于二分类结局在伴发事件之前发生的情况。例如，处于终末期的受试者可能会因为死亡而终止对症治疗，但可以根据死亡前症状的缓解情况评估治疗效果。还有一种情形，受试者可能终止治疗，此时评估其暴露于治疗期间药物不良反应的风险是值得关注的。

因此，在治策略与复合变量策略类似，会影响变量的定义。在这种情况下，在治策略通过将相应的观测时间限制在伴发事件之前来影响。如果各治疗组间的伴发事件的发生率不同，则尤其需要谨慎（见 A.3.3.）。

### **主层策略**

主层策略与人群有关（见 A.3.3.）。可认为目标人群是会发生伴发事件的“主层”（见词汇表）。或者，目标人群是不会发生伴发事件的主层。临床问题仅在该主层中与治疗效应相关。例

如，在接种疫苗后仍然感染的患者主层中，可能需要了解针对感染严重程度的治疗效应。或者，毒性可能会使一些患者无法继续接受试验药物，但需要了解能够耐受试验药物的患者的治疗效应。

区分“主层”和子集很重要。“主层”（见词汇表）是基于潜在的伴发事件（例如，若分配到试验组将终止治疗的受试者），而“子集”是基于实际发生的伴发事件（终止既定治疗的受试者）。在试验组发生伴发事件的受试者子集通常与对照组发生相同伴发事件的受试者子集不同。比较这些子集的结局而定义的治疗效应，会混杂不同治疗间的真实效应和可能由于受试者不同特征导致的结局差异。

#### A.3.3. 估计目标的属性

下述属性用于构建估计目标，定义相关的治疗效应。

**治疗（处理）：**相关的**治疗**条件，以及适用时进行比较的其他治疗条件（在本文件其余部分中称为“治疗”）。这些可能是单独的干预措施，也可能是同时进行的干预措施的组合（例如作为加载治疗），或者是一个复杂干预序列组成的整体方案。（请见 A.3.2.下的疗法策略和假想策略）。

**人群：**临床问题所针对的患者**人群**。可以是整个试验人群，也可以是按某种基线特征定义的亚组，或由特定伴发事件的发生



（或不发生，视具体情况而定）定义的主层（参见 A.3.2.下的主层策略）。

**变量（或终点）：**为解决临床问题从**每个患者**获得的**变量（或终点）**。变量定义可能包括患者是否发生伴发事件（参见 A.3.2.下的复合变量策略和在治策略）。

**其他伴发事件：**在申办方与监管机构关于相关临床问题的交流中，治疗、人群和变量的精确说明有助于解决一些伴发事件。针对任何**其他伴发事件**的临床相关问题，通常采用疗法策略、假想策略或在治策略来反映。

**群体层面汇总：**最后，应规定变量的**群体层面的汇总统计量**，为不同治疗之间的比较提供基础。

在定义治疗效应时，重要的是能够明确效应是由治疗引起的，而不是由潜在的混杂因素如观察期或患者特征的差异等引起的。

#### A.3.4. 构建估计目标的考量

临床问题及与之相关联的估计目标，应当在计划临床试验的初始阶段予以明确。大多数临床试验目的的精确说明，需要体现终止治疗、使用额外治疗或其他治疗的影响。在某些情况下，还应说明死亡一类的终末事件。有些试验目的只能参照临床事件来描述，例如获得应答的受试者的应答持续时间。



构建一个估计目标，应该考虑在特定医疗环境下特定治疗的临床相关性。需考虑的因素包括：所研究的疾病、临床情况（例如可供选择的其它治疗）、治疗方式（例如一次性给药、短期治疗或长期给药）和治疗目的（例如预防、疾病改善、症状控制）。同样重要的是，能否估计出可靠的治疗效应供决策之用。例如，在临床结局发生前，无论是否使用其他治疗情况下的治疗效应和假设没有额外药物可用情况下的治疗效应，是不同的，它们可能都是值得关注的临床问题。但是，在这两种情况下，为估计这些治疗效应，相应的临床试验设计通常会考虑到在医学上需要使用额外药物的可能性。对于前一问题，使用额外治疗后的观测值是有意义的。对于后一问题，额外治疗后的观测值则无直接相关性，因为这些数值也反映了额外治疗的影响。在估计目标最终确定之前，应确认能够得到可靠的估计，包括后一个问题中，用什么方法替代未被分析使用的观测值。

在构建估计目标时，有必要清楚地了解相关临床问题所涉及的治疗（见 A.3.3.）。对治疗的明确说明可能已经反映了多个相关的伴发事件。具体而言，治疗可能已经反映了临床关注问题所涉及的以下变化：基础治疗、合并用药、使用额外或后线治疗、转组治疗和预处理方案。例如，可以将治疗指定为 *干预 A 加基础治疗 B，并按需给药*。这种情况下，无需将基础治疗 B 剂量的变化视为伴发事件。但是，需要将额外治疗视为伴发事件。如果治疗还涉及额外药物的使用，例如在使用疗法策略时，可将治疗定义为 *干预 A 加基础治疗 B，按需给药，并按需使用额外药物*。或

者，如果治疗定义为干预 A，那么基础治疗的变化和额外治疗的使用都将视为伴发事件。

还应讨论是否通过明确人群和变量属性来说明相关临床问题的伴发事件。然后可以考虑有关任何其他伴发事件的策略。通常需要反复讨论来确定对于决策具有临床相关性，并且能得出可靠估计值的估计目标。一些估计目标，特别是那些观测值与临床问题相关的估计目标，通常可以通过很少的假设做出稳健的估计。而有些估计目标的分析方法可能需要更具体的假设，这些假设可能更难以论证，并且可能对假设的合理变化更敏感（见 A.5.1.）。对于某一特定的估计目标，如果在试验设计的合理性或者估计值的可靠性方面存在明显不足的话，就需要考虑另一种估计目标、试验设计和分析方法。

省略或过度简化讨论和构建估计目标的过程，会产生导致试验目的、试验设计、数据收集和分析方法之间不一致的风险。当无法得出可靠的估计值时，可能会妨碍某些策略的选择，重要的是要从试验目的和对临床相关问题的理解出发，而不是为了选择数据收集和分析方法，来确定估计目标。

还应考虑试验现状。如果根据临床试验方案对受试者的管理（例如，因不耐受而进行剂量调整，因应答不足而做的补救治疗，临床试验评估所带来的负担）被证实与临床实践中所预期的不同，这可能要在估计目标的构建中有所体现。

估计目标的构建，应该明确和清晰地定义估计的目标。以终止治疗伴发事件为例，最重要的是区分基于假设接受试验药物就能继续治疗的患者主层的治疗效应与实际持续治疗期间的效应。此外，如果所有患者都能继续治疗，则这两种情况都不能用来反映效应。

如上所述，当使用假想策略时，某些情形更可能为监管决策所接受。因此，所描述的假想情形应该能合理地用来量化可解释的治疗效应，为监管机构做决策和临床实践中药物的使用提供相关信息。假如未获得补救药物，变量值会是多少也许是一个重要问题。相反，如果假想情形中因药物不良反应而终止治疗的受试者实际上继续接受治疗，那么相关变量值在这一假想情形下会是多少这个问题可能不具有临床或监管意义。假如所有受试者都能够继续接受治疗，但没有对他们会继续这一假想情形进行充分讨论的话，则基于该效应的临床问题的定义是不充分的。药物不耐受本身可能就构成了无法达到有利结局的证据。

使用基于疗法策略的估计目标来描述获益效应以支持监管决策，可能更被普遍接受，特别是在某些情况下，尽管基于其他策略的估计目标可能被认为更具临床意义，但是无法找到其公认的能支持可靠估计值或稳健推断的主估计方法和敏感性估计方法。基于疗法策略的估计目标仍然可能得到具有临床相关性的可靠估计值。在这种情况下，建议还包括那些被认为具有更大临床相关性的估计目标，并给出所得到的估计值，以及关于该特定方法在

试验设计或统计分析方面的局限性的讨论。在构建基于疗法策略的估计目标时，可以对使用该策略的每个伴发事件定义额外的估计目标和分析来补充推断；例如，对比各治疗组中治疗对症状评分的治疗效应和使用额外药物的受试者比例。类似的，使用在治策略的估计目标通常应该有伴发事件发生时间分布的附加信息，基于主层策略的估计目标通常有关于该主层中患者比例的信息（如果有）。

基于非劣效性或等效性目的构建估计目标以支持监管决策的考虑，可能与基于优效性目的的估计目标不同。正如在 ICH E9 中所解释的，非劣效性或等效性研究与优效性研究相比，监管机构在决策中面临的问题是不同的。在第 3.3.2.节中指出，这类试验本质上不保守，重要的是尽量减少方案违背和偏离、不依从和退出研究的数量。在第 5.2.1.节中指出，全分析集（FAS）的结果一般不是保守的，它在这类试验中的作用应该被认真考虑。使用疗法策略来说明由一个或多个伴发事件构建的估计目标，对于非劣效性和等效性试验而言，存在的问题与 ITT 原则下使用 FAS 分析存在的相关问题类似。终止随机分配的治疗或因各种原因使用另一种药物，两个治疗组的应答情况可以表现得更相近，而这与最初的随机分配的治疗组间的相似度无关。可以构建估计目标来直接说明那些可能导致治疗组之间差异被弱化的伴发事件（例如终止治疗和使用额外药物）。在选择策略时，区分用于检测出含有相同或相似药物活性成分的治疗之间是否有差异的试验（例如，将生物类似物与参比治疗进行比较）与采用非劣效性或等效性假

设来建立和量化有效性证据的试验是很重要的。为适用于监管决策，可以构建一个针对治疗效应的估计目标，优先考虑能更灵敏地检测出治疗间效应的差异。

#### A.4. 对试验设计和实施的影响

试验设计需要与反映试验目的的估计目标相一致。一种适用于某个估计目标的试验设计，不一定适用于其他具有潜在重要性的估计目标。治疗效应的量化依赖于估计目标，而估计目标的明确定义应当为试验设计的选择提供相关信息。这包括定义目标人群的入选和排除标准，治疗（包括方案中允许的药物和禁用药物），以及患者管理和数据收集等方面。例如，如果关注的是不管是否发生特定伴发事件的治疗效应，则该试验应收集所有受试者的变量。或者，如果用于支持监管决策而制定的估计目标不需要收集伴发事件后的变量值，则应权衡为其他估计目标收集此类数据的益处与数据收集的复杂性和潜在缺陷。

应尽可能收集所有与估计相关的数据，包括伴发事件的特征、发生情况和时间。然而数据不是总能被收集到，不能违背受试者意愿将其留在试验中，而且在某些试验中，受试者数据的缺失按照设计是不可避免的，如生存结局试验中的管理性删失。相反，如发生终止治疗、转组治疗或使用额外药物等伴发事件，尽管这些测量结果可能并不相关，并不意味着在事件之后无法测量变量。对于死亡等终末事件，变量不能在伴发事件后进行观测，但这些数据通常都不应被视为缺失。

不收集评估估计目标所需的数据，将导致后续统计推断中的缺失数据问题。统计分析的合理性可能取决于不可验证的假设和缺失数据的比例，这些可能会削弱结果的稳健性（见 A.5.）。制定一个前瞻性计划收集详细的数据缺失的原因，将有助于区分伴发事件的发生与缺失数据。这样可改进分析，并可使敏感性分析的选择更合理。例如，将“失访”记录为“因缺乏疗效而终止治疗”可能更准确。在将其定义为伴发事件的情况下，可以选择相应策略，而不是将其作为缺失数据问题来处理。为减少缺失数据，可采取措施将受试者保留在试验中。然而，减少或避免临床实践中通常会发生的伴发事件的措施存在降低试验外部有效性的风险。例如，如果在临床实践中不会通过选择试验人群，或使用滴定方案，或使用合并用药来减轻毒性的影响，那么在试验中采取这些措施也就不合适了。

随机化和盲法仍然是对照临床试验的基石。避免偏倚的设计方法请见第 2.3.节。某些估计目标可能需要或受益于试验设计的运用，如导入期或富集设计、随机撤药设计或滴定设计。在对受试者进行随机化分组之前，利用导入期识别对药物耐受的受试者主层，这个办法是可取的。监管机构和申办方之间的交流需要考虑拟定的导入期是否适合于确定目标人群，以及后续试验设计（例如，洗脱期、随机化）的选择是否支持目标治疗效应的估计和相关推断。这些考虑可能会限制这些试验设计的使用，以及特定策略的使用。

对治疗效应的精确描述应当为样本量计算提供信息。在参照历史研究时应特别谨慎，这些历史研究可能隐含或明确地报告了基于不同估计目标的治疗效应或变异的估计值。当所有受试者均能为分析提供信息，且在目标效应量和预期方差中已考虑了相应的策略来反映伴发事件的影响时，则在计算样本量时通常不需要按预期退出试验的受试者比例额外增加样本量。

第 7.2.节说明了关于跨临床试验数据汇总的问题。强调对所关注的变量应有一致的定义，并且这可延伸用于估计目标的构建。因此，为了从多个临床试验获得综合性证据，在计划阶段应构建一个合适的估计目标，使其包含在项目里的所有试验方案中，并反映在相关试验的设计选择中。类似的考虑适用于 meta 分析的设计，使用已完成的试验中估计的效应量来确定非劣效界值，或者使用外部对照组来诠释单臂试验。在没有考虑和说明每个数据呈现或统计分析所对应的估计目标的情况下，不同来源数据之间的简单比较，或者来自多个试验的数据整合可能会产生误导。

总的来说，一个试验有可能将多个目的转化为多个估计目标，每个估计目标都与统计检验和估计相关联。此时，应当考虑其中的多重性问题。

## A.5. 对试验分析的影响

### A.5.1. 主要估计

对于治疗组相较于对照组的治疗效应，估计目标通过比较具有相似受试者的试验组和对照组的结果来进行估计。对于给定的估计目标，应采用与其相一致的分析方法（或估计方法），使所得估计值可以支持对结果的可靠解读。该分析方法还应能计算置信区间并进行统计学显著性的检验。可解释的估计值是否存在，一个重要的考虑因素是分析中需要作出的假设的程度。关键假设应与估计目标及其主要估计方法和敏感性估计方法一起明确说明。假设应是合理的，要避免不恰当的假设。对于潜在的偏离假设的情形，结果的稳健性应通过与估计目标相对应的敏感性分析进行评估（见 A.5.2.）。如果一个估计依赖于多种假设或强假设，则需要更广泛的敏感性分析。如果偏离假设所造成的影响不能通过敏感性分析进行全面研究，那么这种特定的估计目标和分析方法的组合可能无法用于决策。

所有的分析方法都依赖于假设，而且即使对应于相同的估计目标，不同的分析方法也可能依赖于不同的假设。尽管如此，对于使用每种不同策略的估计目标，相对应的所有方法中都存在一些固有的假设。例如，预测在假想情形下本可观察到的结局的方法，或者在主层策略下确定合适目标人群的方法。下面列举了一些关于针对不同伴发事件采用不同策略的例子。在申办方和监管机构就估计目标、主要分析和敏感性分析达成一致之前，这里所强调的问题将是双方讨论的关键。



与针对某一伴发事件的疗法策略相对应的分析，根据试验的设计和实施，可能需要或强或弱的假设。如果在相应的伴发事件（如治疗终止）之后仍对大多数受试者进行随访，那么缺失数据问题可能相对较小。相反，如果伴发事件后停止观察（对于该策略显然是不提倡的），那么假设终止治疗受试者的（未观察到的）结局与继续治疗的受试者的（观察到的）结局相似通常是不合理的。此时在处理缺失数据时需要对其他方法进行论证，并进行敏感性分析。

与假想策略相对应的分析方法所涉及的结局与实际观测的结局不同；例如，尽管实际上给予了补救药物，但需要估计未给予补救药物时的结局。在给予补救药物之前的观测值与不需要补救药物的受试者的观测值可能提供有效信息，但需要更强的假设。

复合变量策略可以通过将伴发事件的发生视为结局的组成部分来避免对伴发事件后的数据做出统计学假设。这种情况下，潜在的担忧往往不在于估计中相应的假设，而在于对治疗效应估计结果的解释。为使估计目标得以合理解释，如果将伴发事件的发生认定为失败，并给出评分，这个评分应该能够有效地反映出患者获益的缺乏程度（例如，对死亡与因不良事件导致处理终止的反映也许不同）。

如果伴发事件发生之前的结局已被收集，那么依据在治策略所构建的估计目标就可以被估计。同样的，此处的关键假设将影响到结果的解释。以终止治疗为例，治疗过程中结局可能会改

善，但同时治疗也可能因为引发、延迟、终止治疗等原因，使得治疗期缩短或延长。此类影响应在解释和评估临床获益时予以考虑。

与主层策略相对应的分析通常需要较强假设。例如，一些主层方法是基于受试者的基线特征而推断出来的，但这种推断的正确性可能难以评估。而且，这种困难不能通过简化的方法来避免。例如，假设伴发事件与处理无关，并简单地比较试验组和对照组中未出现伴发事件的受试者，这是非常难以论证的。

即使以恰当的方式定义了解决伴发事件的估计目标，并努力收集估计所需的数据（见 A.4.），一些数据仍然可能缺失。例如，生存结局试验中的管理性删失。未能收集到相关数据不应与选择不收集或选择收集但不使用（因伴发事件变得无关的数据）相混淆。例如，对于基于疗法策略的估计目标，在终止试验药物后的数据仍应被收集，如果未收集，则视为数据缺失；然而，对另一种策略而言，相同的数据点可能不相关，因此，对于相应的估计目标，此类未收集数据则不会被视为缺失。如果数据收集不完整，则有必要在统计分析中对缺失数据的处理做出一些假设。缺失数据的处理应基于临床上的合理假设，并在可能的情况下以估计目标描述中采用的策略为指导。采取的方法可能基于个体受试者和与其相似受试者所观测到的协变量和基线后数据。识别相似受试者的标准可能包括是否发生伴发事件。例如，对于终止治

疗但未收集更多数据的受试者，可使用终止治疗但继续收集数据的其他受试者的数据来建模。

## A.5.2. 敏感性分析

### A.5.2.1. 敏感性分析的作用

基于特定估计目标的统计推断，应该对数据的局限以及主估计方法统计模型中假设的偏离具有稳健性。这种稳健性应通过敏感性分析来评价。对于所有用于监管决策和说明书制定的估计目标的主估计方法，都应有相应的敏感性分析计划。此问题需要在申办方和监管机构之间讨论并达成一致。

支持主估计方法的统计假设应明确记录。对于同一估计目标，应该预先规定一项或多项分析来评估这些假设，目的是验证根据主估计方法得出的估计值是否对假设偏离具有稳健性。其衡量标准可以是对假设不同程度的偏离是否会改变结果的统计学或临床意义（如临界点分析）。

敏感性分析旨在探索偏离假设时分析结果的稳健性，与此不同的，为了更全面地研究和理解试验数据而进行的其他分析可称为“补充分析”（见词汇表，A.5.3.）。如果申办方和监管机构就所关注的主要估计目标达成一致，并预先明确规定了主估计方法，且敏感性分析也验证了估计值的结果解释是可靠的，则补充分析在结果评估中通常不被优先考量。

### A.5.2.2. 敏感性分析的选择

当计划和实施敏感性分析时，同时改变主要分析的多个方面可能难以确定由哪些假设导致了目前所观测到的潜在差异。因此，通常采用结构化的方法，指定不同分析背后的假设的变化，而不是简单地基于一组不同的假设比较不同分析的结果。应根据具体情况考虑是否需要同时改变多个假设的分析。在评估不同分析的解释和相关性时，区分可验证的和不可验证的假设可能是有帮助的。

在本文所设立框架中，进一步明确了对缺失数据进行敏感性分析的必要性和重要性。缺失数据应依据特定估计目标进行定义和考虑（见 A.4.）。对应于特定估计目标的缺失数据，以及与特定估计目标不直接相关的数据，两者之间存在区别，由此在分析中产生了不同类别的假设，需要通过敏感性分析来检查。

### A.5.3. 补充分析

试验结果的解释应侧重于对应每个估计目标的主估计方法，并通过敏感性分析验证相应估计值的稳健性。除了主要分析和敏感性分析之外，还可以对估计目标进行补充分析，以提供对治疗效应更全面的了解。补充分析在解释试验结果方面的作用通常较小。每项试验均需考虑补充分析的必要性和作用。

第 5.2.3.节指出，同时基于全分析集（FAS）和符合方案集（PPS）的分析计划通常是适当的，从而它们之间的差异会成为

讨论和结果解读的关键。如果基于 **FAS** 分析和 **PPS** 分析的结果一致，则可增强试验结果的可信度。第 5.2.2.节还指出，基于 **PPS** 的结果可能会产生严重偏倚。就本增补中提出的框架而言，可能无法构建与 **PPS** 分析相对应的估计目标。如上所述，**PPS** 分析不能实现在任何主层（例如，在能够耐受并继续接受试验药物的受试者）中估计效应的目的，因为 **PPS** 所比较的受试者在不同治疗组之间可能不具有可比性。

即使没有发生伴发事件，方案违背和偏离（例如，在时间窗外进行访视）也可能使受试者从 **PPS** 中被排除。同样，受试者可能发生伴发事件（例如死亡）但却没有偏离方案。尽管违背和偏离方案与伴发事件之间存在差异，在估计目标的描述中仍应考虑可能影响观测结果解释或存在的事件。通过构建估计目标和相应的分析方法，可能更好地反映与 **PPS** 分析相关的目标。此时，**PPS** 分析也许不能提供额外的信息。

## A.6. 估计目标和敏感性分析的记录

试验方案应当定义并明确说明与主要试验目的相对应的主要估计目标。方案和分析计划中应预先规定与主要估计目标一致并对应主要分析的主估计方法，以及当假设偏离时用来探索结果稳健性的合适的敏感性分析。对于可能支持监管决策的次要试验目的（例如，与次要变量相关），相应的估计目标也应当明确定义和说明，并且每个估计目标都有相应的主估计方法和合适的敏感

性分析。还可以考虑属于探索性质的额外的探索性试验目的，此时也会产生额外的估计目标。

主要估计目标的选择通常是试验设计、实施和分析的主要决定因素。按照常规，这些信息应在试验方案中详细记录。如果次要估计目标同样需要重点关注，这些考虑可扩展到相应的估计目标，并同样在方案中记录。除此之外，对于试验设计、实施和分析的常规考虑仍然保持不变。

尽管明确阐明估计的内容对申办方有益，但监管部门并不要求对每一个探索性目的的估计目标都进行记录。

在临床试验报告中应系统报告主要分析、敏感性分析和补充分析的结果，同时详细说明每项分析是否为预先规定的、在试验仍处于盲态时引入进行的，还是事后进行的。应汇总报告各处理组中各类伴发事件的数量和出现时间。

试验期间改变估计目标可能是有问题的，这样做会降低试验的可信度。对于在设计阶段未预见但在试验实施过程中发现的伴发事件，不仅要讨论分析方法的选择，还要讨论它们对估计目标的影响，即对所估计的治疗效应描述的影响，和对试验结果解释的影响。估计目标的改变通常应通过修订试验方案来体现。

## 词汇表

### 估计目标:

对治疗效应的精确描述，反映了针对临床试验目的提出的临床问题。它在群体水平上汇总比较相同患者在不同治疗条件下的结局。

### 估计值:

由估计方法计算得出的数值。

### 估计方法:

采用临床试验数据计算估计目标的估计值的分析方法。

### 伴发事件:

治疗开始后发生的事件，可影响与临床问题相关的观测结果的解释或存在。在描述相关临床问题时，需解决伴发事件，以便准确定义需要估计的治疗效应。

### 缺失数据:

对于既定估计目标的分析有意义、但未收集到的数据。它应该与不存在的数据，或由于伴发事件而被认为没有意义的数据区分开来。

### 主分层:

根据所有治疗中伴发事件的潜在发生情况，对受试者进行的分类。以两种治疗为例，针对特定的伴发事件，有四个主层：任一治疗期间均不会发生事件的受试者，在 A 治疗期间会发生事件但在 B 治疗期间不会发生事件的受试者，在 B 治疗期间会发生事件但在 A 治疗期间不会发生事件的受试者，以及在两种治疗期间均会发生事件的受试者。在本文件中，主层是指主分层定义的任何分层（或分层组合）。

### **敏感性分析：**

针对模型假设的偏离和数据局限，探索主估计方法统计推断的稳健性的一系列分析。

### **补充分析：**

对于主要分析和敏感性分析之外的分析的一般描述，目的是更多地了解治疗效应。