

国际人用药品注册技术协调会

ICH 协调指导原则

临床试验的适应性设计

E20

草案

于 2025 年 06 月 25 日批准

在 ICH 进程的第 2 阶段，ICH 大会按照国家或地区程序，将相应 ICH 专家工作组商定的共识草案文本或指导原则转交给 ICH 地区的监管机构，供内部和外部征求意见。

ICH E20 的解释性说明：该指导原则草案对临床试验的适应性设计方法进行了说明。指导原则草案认为，适应性设计在不降低科学标准和监管标准的前提下，极有可能加速药物研发进程和优化资源配置。其中一些方法可能会影响工业界与监管机构在确证性试验计划与评估中的交流互动的性质和时机。最终指导原则将指出关键的适应性设计原则和方法，其中对适应性设计特征及其使用原理的讨论在计划阶段尤为重要。为了指导原则的最终确定，现特别就适应性设计的原则和方法及其对工业界-监管机构之间交流互动的影响征求反馈意见。在 ICH 第5阶段商定最终指导原则之前，不应将指导原则草案理解为 ICH 各方对其内容的完全监管认可，也不应该用其取代仍然有效的现行区域指导原则。现就该指导原则草案公开征求意见。

E20

文件历史

编码	历史	日期
E20	由 ICH 大会成员在第 2 阶段签署，并发布以供公开征求意见。	25/06/2025

法律声明： 本文受版权保护，除了 ICH 标志外，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、修正或翻译，必须采取合理措施来清晰地标识、区分或以其他方式标记对文件进行的修改。必须避免任何暗示 ICH 授权或支持对原版文件的改编、修改或翻译的行为。本文件“按原样”提供，不提供任何形式的担保。任何情况下，ICH 或原版文件作者不会对任何由使用本文件造成的索赔、伤害或其他责任负责。

上述许可不适用于由第三方提供的内容。因此，对版权归属第三方的文件，必须从该版权所有者处获得复制许可。

ICH协调指导原则

临床试验的适应性设计

E20

ICH 共识指导原则

目录

1.	前言和范围	1
2.	适应性设计的优势和挑战	1
3.	关键原则	3
3.1	开发计划的充分性.....	3
3.2	试验计划的充分性.....	3
3.3	限制得出错误结论的可能性.....	5
3.4	估计的可靠性.....	5
3.5	试验完整性的保持.....	6
4.	适应性设计类型	8
4.1	提前终止试验.....	8
4.2	样本量调整.....	10
4.3	人群选择.....	11
4.4	治疗选择.....	12
4.5	试验参与者分配调整.....	13
5.	专题讨论与注意事项	14
5.1	关于数据监查的特别考虑.....	14
5.2	模拟研究的计划、实施和报告.....	15
5.3	采用贝叶斯方法的适应性设计.....	18
5.4	以至事件发生时间为终点的适应性设计.....	20
5.5	探索性试验的适应性设计.....	21
5.6	操作性考虑.....	21
6.	文档管理	22
6.1	适应性设计确证性试验开展前的文档记录.....	22
6.2	适应性设计确证性试验完成后在上市申请中应包含的文档.....	23

1. 前言和范围

本指导原则为采用适应性设计的确证性临床试验提供指导，旨在整体开发计划的背景下评估给定医学疾病的治疗方法。就本指导原则而言，适应性设计是指允许根据试验参与者积累数据的期中分析结果，对试验的一个或多个方面进行前瞻性计划修改的临床试验设计。这里“前瞻性计划”是指在试验开始前，临床试验方案中预先设定了潜在的试验调整内容。本指导原则的范围不包括对设计进行计划外修改的试验，如独立的数据监查委员会（IDMC）根据非预期的期中分析结果提出的方案修订。该范围也不包括完全依据试验外部来源的信息做出的设计变更。对入组率、数据质量或试验参与者退出程度等操作方面的常规监测也超出该范围。

本指导原则目的在于对适应性设计试验的计划、实施、分析和结果解读提供指导，旨在确认治疗的有效性并支持获益-风险评估。重点聚焦于保障试验产生可靠且可解读信息的关键原则，以及采用适应性设计时需特殊考虑的要点。本指导原则不涉及具体统计方法的讨论。尽管本指导原则主要侧重于确证性临床试验，但所概述的原则与临床开发的所有阶段相关。

2. 适应性设计的优势和挑战

在确证性试验的计划阶段，即使经过药物开发早期阶段的仔细计划和实施，在设计方面仍可能存在不确定性，如适当的样本量。然而，对于非适应性设计，这些方面必须在试验开始前确定，并且在试验执行期间不能改变。适应性设计具有灵活性，可通过利用试验期间试验参与者的累积数据对设计进行预设的修改来避免不准确的假设。

适应性设计的灵活性可带来多方面的优势：首先，具有伦理方面的优势。例如，在成组序贯设计中，若已有充分证据表明治疗有效并且具有正向的获益-风险特征时，可提前终止试验，从而减少暴露于劣效对照治疗的试验参与者人数。其次，适应性设计可以提高试验的效率。例如，通过给定预期样本量提高试验的把握度。第三，适应性设计有助于提高对疗效和决策的理解。例如，在两阶段适应性确证性临床试验中，通过期中分析在两种剂量中选择更具优势的剂量，既可降低对最佳获益-风险剂量选择的不确定性，同时也能确证所选剂量的有效性。

然而，适应性设计也带来诸多挑战，可能会增加与第 3 章节中讨论的关键原则相关的复杂性和不确定性。例如，采用适应性设计可能会增加对期中分析结果保密的困难度，若处理不当，可能对试验完整性构成风险，最终导致试验结果不可靠以及试验结束时出现试验结果解读困难。此外，与未采用适应性设计的试验相比，适应性设计临床试验的合理计划与评估可能更复杂，并且需要更多的时间。具体而言，在适应性设计临床试验中，采用非适应性设计的常规分析方法，往往会导致 I 类错误概率升高及疗效估计的偏倚。例如，根据期中分析的疗效估计对样本量进行调整，最终分析时采用常规分析方法，I 类错误概率可能会增加一倍以上。再例如，由于试验会在观察到极端数据时优先停止，因此提前因疗效显著而停止试验可能会导致疗效估计出现偏倚。因此，通常需要采用考虑适应性设计的特殊分析方法进行假设检验和疗效估计。此外，一些采用适应性设计的试验提供的安全性信息可能较少，从而可能导致获益-风险评估过程中出现更多的不确定性。适应性设计也并非在所有临床试验环境中都能获益。例如，如果试验参与者入组速度快于适应性调整所依据的终点评估时间，或者无法及时获取数据以支持期中分析时作出可靠的调整决策，则适应性设计可能不宜采用。

临床试验中是否采用特定的适应性设计，需综合考虑包括上述因素在内的多项条件。后期临床试验的确证性质和在试验进行中对试验进行调整的建议之间可能存在矛盾。因此，在计划适应性设计时，必须详细论证调整试验的必要性，并评估所涉及调整类型、数量和复杂性的潜在影响。论证内容应包括临床和统计学方面的考虑。应当权衡设计的优势，以及拟调整内容可能对试验结果可靠性和可解读性带来的不确定性影响。例如，增加详细计划的因有效或无效而提前终止试验的期中分析，采用预设的合理终止规则，并确保提供足够信息以支持安全性和获益-风险评估，同时使用 IDMC 来保持试验完整性，这样可以最大限度地减少不确定性。另一方面，涉及多个试验要素调整的复杂设计可能会极大增加与保持试验完整性相关的不确定性。这种不确定性可能包括对信息流和数据访问规范充分性的存疑，或调整本身对试验实施及疗效可解释性造成的潜在影响。针对特定患者人群拟定治疗剂量的有效性和获益-风险特征，此类情况可能会给评估结果和监管决策带来挑战。拟采用的适应性设计必须提供清晰且具有说服力的论证。该论证应阐述适应性设计是否满足临床环境的固有需求，并评估其与替代设计（包括非适应性设计）相比的优势和局限性，包括比较不同候选设计之间重要的试验操作特征（比如把握度、预期样本量、调整决策的可靠性）。

3. 关键原则

就本指导原则而言，原则是指试验设计的特征，对于确保结果的可靠性和可解读性至关重要。本章节对适应性设计需要特别考虑的原则进行介绍。重点对确证性临床试验的适应性设计提供建议。无论采用何种调整类型或统计方法（比如频率论或贝叶斯方法），都应遵循所有这些原则。

3.1 开发计划的充分性

临床试验需要经过适当的设计、实施和分析，以在整体开发计划的背景下解决相关临床研究问题，这一点非常重要。对已完成的探索性试验进行仔细分析和评估，有助于为后续确证性试验的目标设定和设计选择提供依据，并最终生成监管决策所需的数
据。完整的开发计划应关注以下几个方面：描述剂量反应关系在有益及不利效应方面的特征；确定适当的治疗患者人群；选择具有临床意义和敏感性的终点；在预期患者人群中可靠地确认疗效并支持安全性和获益-风险评估。

在确证性试验阶段，适应性调整的次数和复杂性通常应受到限制。如果为了替代一系列的多项试验而增加调整次数或提高复杂性，可能会损害试验对重要临床问题的回答能力，也会影响通过仔细评估既往结果以最有效地设计进行计划开发的机会。在计划进行包含多次适应性调整的确证性试验之前，申办者应针对拟调整内容所涉及的问题，讨论是否还需开展额外的必要的探索性试验。

例如，考虑采用一种确证性两阶段适应性临床试验设计，通过期中分析在两种剂量中进行选择，并确认所选剂量的疗效。在进行剂量探索试验之后，仍不确定两个候选剂量哪一个最合适的情况下，这种设计可能有助于确定在预期患者人群中具有更好的获益-风险特征的剂量。但是，如果在开发计划的早期阶段没有进行适当的剂量探索试验，则可能无法很好地支持选择两种剂量进行确证性试验，从而增加了开发计划无法确定出合理剂量的风险。一般而言，适应性设计不应替代适当的剂量探索试验。通常，在进行确证性试验以确定疗效和评估获益-风险之前，需预计申办者已经完成了必要的、以评估更大范围和更多数量的剂量的试验。

3.2 试验计划的充分性

充分的计划对所有临床试验都很重要，以确保设计是预先规定的，试验实施和分析是

合理的，结果是可靠的和可解释的。如果确证性临床试验计划采用适应性设计，则通常应限制调整的次数和复杂性，并提供在药物开发的这个阶段对试验各方面进行调整的理由。在开始一项适应性设计的试验之前，除试验计划的常规内容外，还需额外明确以下要素并提供充分依据：包括期中分析的次数和时间、适应性调整类型、生成期中分析结果的统计方法、适应性决策的预期规则、与每个估计目标对应的主要分析方法、以及保持试验完整性的方法。对于计划在期中分析阶段选择估计目标的适应性设计，例如治疗或人群选择设计（第 4.3 和 4.4 章节），所有候选估计目标均须是完全预设的并与临床相关。

某些类型的适应性设计可能比其他类型需要更多的计划。例如，与采用盲态样本量调整的设计相比，非盲态样本量调整需要采用额外的方法来保持试验的完整性。如果模拟对于理解适应性设计的操作特征至关重要，则应仔细计划、实施和报告模拟研究（第 5.2 章节）。所有与适应性试验设计计划相关的细节都应妥善记录（第 6 章节）。

充分的计划有助于评估各类适应性调整所采用的统计方法是否恰当。例如，I 类错误概率的控制要求预先制定早期有效性终止的标准或跨阶段合并证据的规则。再例如，若在方案中规定盲态样本量调整内容以及调整规则，可进一步确保适应性选择的样本量不受非盲数据影响。充分的计划也有助于评估试验操作特征，并能够与 IDMC（比如参与适应性调整）进行有根据的讨论。申办者应与 IDMC 详细讨论适应性调整类型和预期的调整规则，以确认其理解和支持。这能够确保 IDMC 做好在试验期间审查期中分析结果并做出适应性调整建议的准备，同时保护试验参与者个体的安全。

对于适应性调整所依据的预期规则必须描述清晰。然而，在期中分析中需要多大程度遵循该预期规则，可能因适应性调整类型和所使用的统计推断方法的不同而有所差异。通常建议使用既能提供有效的推断结果，又允许在期中分析时根据整体获益-风险评估灵活偏离预期调整规则的分析方法。例如，考虑进行一项确证性两阶段适应性临床试验，在期中分析时选择两种剂量中的一种，以确认所选剂量的疗效并支持其获益-风险评估。在试验计划阶段，如果预计不会出现重大安全性问题，则可以制定一项基于有效性的期中剂量选择的规则。然而，可能会出现期中数据显示两种剂量疗效相似，但较高剂量存在非预期安全性问题的情况。若采用的统计方法允许在期中分析中灵活纳入此类获益-风险考虑，则预设的计划应承认偏离规则的可能性，并列出可能导

致此类偏离的因素。如果计划的统计方法要求严格遵守期中决策规则以确保推断的有效性（比如控制 I 类错误概率），则应在试验方案中说明遵守该规则的重要性。

3.3 限制得出错误结论的可能性

限制对拟定治疗的有效性、安全性和获益-风险特征得出错误结论的可能性非常重要。监管决策的一个基本要素是控制出现假阳性疗效结论（即真实无效的治疗被认为有效的结论）的可能性。常见的方法是采用频率学派方法来限制试验中假阳性疗效结论的可能性，此类方法将主要估计目标假设检验中的 I 类错误概率控制在预先设定的阈值内（ICH E9）。

对于大多数适应性设计，有必要使用特定的方法来控制 I 类错误概率。例如，若试验设计包含可能因有效性而提前终止的期中分析，则需预先设定合适的终止规则。当适应性设计包含采用多重检验方法来控制多个主要和/或次要终点的 I 类错误概率时，这些方法还应额外考虑因适应性调整可能导致 I 类错误概率增加的可能性。

尽管临床试验设计和分析的主要方法基于频率学派统计方法，但如果其他方法的使用理由明确且所得结论足够可靠时，也可被采用（ICH E9）。第 5.3 章节描述了在使用贝叶斯学派方法的适应性设计中，限制假阳性疗效结论可能性的重要考虑。

理解拟采取的适应性设计如何影响其他类型错误结论的可能性也很重要。这包括要求试验需要提供关于安全性、重要的次要疗效终点和相关患者亚组的充分信息，以确保能够进行可靠的获益-风险评估。例如，当计划一项可能因有效性结论而提前终止的试验时，需证明期中分析时的样本量和随访时间能够充分支持可靠的获益-风险评估。还包括评估适应性设计对期中分析结论的影响，以及适应性设计可能无法实现试验目的的风险。例如，申办者应基于疗效和获益-风险考虑，评估适应性剂量选择设计在期中分析时从两个剂量中选择更优剂量的能力。最后，适应性调整可以影响假阴性疗效结论的发生概率（即缺乏证明真实有效治疗的证据），因此评估试验是否具有足够的检验效能非常重要。

3.4 估计的可靠性

在确证性临床试验中，需要控制假阳性疗效结论的可能性（第 3.3 章节）。此外，对主要有效性终点的疗效与其他关键有效性和安全性结局进行可靠的估计，对于促进获益-

风险评估和为监管决策提供信息至关重要。因此，适应性设计试验的主要分析应提供可靠且与估计目标一致的疗效估计。申办者应评估疗效估计的偏倚和变异性，包括均方误差等指标。在偏倚和方差的权衡中，通常期望主要疗效估计中偏倚有限或没有偏倚。主要分析还应支持对不确定性的准确度量的计算，如目标覆盖概率的置信区间。

如果适应性设计试验在主要分析中采用的估计方法未考虑设计本身的适应性特征，则可能会出现不可靠的疗效估计和对不确定性不正确的估计（比如置信区间覆盖率不正确）。例如，在期中分析时，从多个治疗组中选择估计效应量最大的治疗，一般而言，会导致该治疗效应的过高估计。即使选择是基于预期可预测疗效的终点而非主要终点本身，这一偏倚依然存在。同理，在存在适应性调整的情况下，次要终点的疗效估计也可能存在偏倚。因此，采用适应性设计应评估疗效估计的偏倚和变异性，并为其可靠性提供支持。在部分情况下，可以解析计算偏倚和变异性。在其他情况下，必须依赖于模拟进行评估。对于某些设计，应采用已得出的可靠性更高的特定估计方法。以成组序贯设计为例，现有方法可通过调整估计量来减少或消除与早期终止可能相关的偏倚，并提升均方误差等指标的性能。

除了确保主要分析中疗效的可靠估计外，还需保证期中分析的估计结果能够支持促进可靠的适应性调整决策。例如，在适应性剂量选择设计中，若过早进行期中分析，可能导致估计结果变异过大，从而选择较差剂量。因此，申办者应评估设计的整体操作特征（比如选择更优剂量的概率），以便慎重确定期中分析的时间节点和调整规则。

3.5 试验完整性的保持

保持试验完整性对于确保试验可靠、符合伦理且及时的达成既定目标非常重要。第 3.3 章节和第 3.4 章节已对试验适应性调整对试验结果的统计有效性可能带来的影响进行了讨论。保持试验完整性还依赖于试验的恰当执行以及关于设想的调整对试验实施的潜在影响进行仔细评估，这是本章节的重点。

申办者、研究者或试验参与者对个体治疗分配、累积数据或某些试验变更的知晓会以难以预测且无法调整的方式影响期望和行为，从而影响试验完整性。对此类信息的知晓可能会对试验实施造成不易察觉的变化，例如入组试验参与者的速度和特征、研究治疗药物或其他药物给药的具体细节，或终点评估等的变化，这些都可能影响对试验结果的解读。例如，研究者和试验参与者根据正在进行的试验期间累积的数据了解到

治疗的估计疗效较低甚至呈负向趋势，可能会将此误解为治疗无效的可靠证据，导致试验参与者的入组率下降、依从性减弱和脱落率上升，最终导致结果不可靠，并在试验结束时难以解读。推荐的方法是使试验参与者、研究者和申办者不仅对个体治疗分配保持盲态，同时也对累积的汇总数据的分组标签（使用实际治疗或标签如 A 和 B）保持盲态，此举有助于最大程度避免此类信息对试验实施造成有意或无意的干扰。

许多类型的适应性设计的一个基本要素是需要获取一定程度的非盲态期中分析结果。能够接触累积的非盲态数据的人员应保持独立性，即不应存在利益冲突、不参与试验相关活动，并且应为独立于申办者的外部人员。为实现这一要求，当适应性设计需要获取非盲态期中数据时，应设立 IDMC 来对相关数据进行审核。在确证性试验中，通常已经预先设立 IDMC 以确保试验参与者的安全并保护试验的科学完整性。此时，IDMC 可承担额外的职责，即审查期中数据，以对是否实施预先计划的适应性调整做出建议。若尚未计划设立 IDMC，也可根据适应性设计的实施组建一个新的 IDMC，其设立目标和成员专业背景应与设计要求相匹配。此外，应制定标准操作程序和保密协议，以限制 IDMC 以外人员访问非盲态期中分析结果。关于 IDMC 和其他数据监查考虑的更多讨论内容，见第 5.1 章节。

即使仅知晓试验中发生了适应性调整这一信息，也可能导致研究者或试验参与者出现非预期的行为改变，甚至可能泄露非盲态期中分析结果的有关信息。例如，如果实施非盲态样本量调整，其中修改后的样本量是期中疗效估计量的函数，那么了解调整规则并知道修改后样本量的人可以推测出期中疗效的估计。因此，在确保符合伦理标准（比如提供充分的知情同意书）和操作可行性（比如保障充足的药品供应）的前提下，应采取相应措施，最大限度减少可被推测的信息内容；有关操作性考虑的进一步讨论内容，见第 5.6 章节。为降低试验过程中信息被推测的风险，可以采用特定的调整规则，将期中分析估计值划分为较大的区间范围，在同一区间内，无论估计值具体如何，都会触发相同的调整措施（比如样本量调整规则中，事先预设了一小部分可供选择的潜在样本量水平）。调整规则的细节可以保留在特定文件而非方案中，例如 IDMC 章程的机密附录，该文件仅限独立于临床试验管理和执行团队的申办者指定人员查阅。此外，可对申办者人员、研究者及试验参与者实施屏蔽，使其不获知特定的适应性变化。例如，在样本量调整后，可告知试验中心目标招募人数尚未达到，或仅告知试验中心或地区特定的入组目标，而非告知总体样本量目标。

申办者应在计划阶段与监管机构讨论适应性调整对试验实施的潜在影响，包括入组试验参与者的类型，以及试验结束时对结果的解读。讨论内容还应包括试验阶段的规模是否足以用来评估适应性调整的影响。申办者应建立保持试验完整性的方法。应记录流程以提高对这些方法的依从性并向利益相关者（比如监管机构和研究者）保持透明。为此需开展适当培训并制定周密计划，尽可能降低试验完整性受损风险。由于即使是最严格的流程也可能无法完全确保试验完整性，因此试验结束时的结果解读应考虑不同阶段试验结果之间的异质性、适应性设计的性质（比如调整次数、类型以及试验阶段的规模）、现行的流程及试验期间各类数据和信息的访问人员和权限，以及期中分析前后任何试验执行方面的显著变化（比如入组试验参与者类型变化）。若发现了非预期的异质性，申办者应进行讨论，因其可能影响对试验结果的解读。

上述保持试验完整性的原则，在开放性试验中尤为关键，因为在该类试验中，试验参与者和/或研究者已知晓每例试验参与者的个体治疗分配情况。需要强调的是，尽管这些试验中个体试验参与者的治疗分配信息已知，确保试验参与者、研究者及申办者无法获取按治疗组别累积的汇总数据依然是可行的，也是强烈建议的。

4. 适应性设计类型

本章节将讨论常见的适应性设计类型，并重点关注与第 3 章节所述原则相关的具体考虑，同时阐述第 2 章节中概述的适应性设计的一些优势和挑战。本章节的讨论聚焦于使用频率学方法进行统计分析的适应性设计，有关使用贝叶斯方法的适应性设计的特殊考虑，请参见第 5.3 章节。

4.1 提前终止试验

在临床试验进行期间，累积的数据可能提示继续开展试验已不再合适。为解决这一问题，申办者可以考虑采用一种包含前瞻性计划的序贯分析的试验设计，对累积的非盲态数据进行序贯分析，并按照预设的终止规则，在出现充分的有效性证据时终止试验（因有效性而终止），或在试验不太可能证明有效性时终止试验（因无效性而终止）。这种允许在试验过程中，基于预设时间点累积的各组试验参与者数据，通过序贯分析提前因疗效而终止试验的临床试验设计，称为成组序贯设计。

221 在设计允许因疗效而提前终止的试验时，应为序贯分析设置适当的终止边界，以控制 I
222 类错误概率。期中分析的时间点及具体的终止规则应基于多个因素进行确定，例如，
223 终止试验所需的早期结果的说服力、提前终止的概率，以及可能累积的预期和最大样
224 本量或事件数。可考虑采用允许偏离预设期中分析时间点的方法。这种方法有助于配
225 合 IDMC 在特定日历时间召开会议，使得期中分析时的实际样本量可以与预设的目标
226 略有差异。此外，应预设校正期中分析影响的主要疗效估计及其置信区间的计算方
227 法，以控制偏倚并提高在均方误差等指标上的性能（第 3.4 章节）。

228 因有效性而提前终止的试验，为安全性、重要次要疗效终点及相关患者亚组评估提供
229 的信息较少（比如样本量较小和/或随访期较短），而这些信息对整体获益-风险评估十
230 分重要。因此，期中分析时点的选择应考虑样本量足够大且随访时间足够长，以确保
231 有充分的信息用于决策。一般情况下，由于充分的安全性评估需要最低样本量和/或最
232 短随访时间，这会限制期中分析的最早实施时间或是否应该设置期中分析。这通常是一
233 个相关联的标准，例如，在预防性疫苗试验中，以及用于非危及生命疾病的长期治
234 疗药物需满足人群暴露程度的监管要求（ICH E1）时。此外，当存在充分的伦理依据
235 时（比如生存期为主要终点），更应考虑设置具有提前终止可能的期中分析，且有效性
236 终止规则通常要求在疗效的估计大小和证据强度两方面均需具有高度说服力的结果。

237 若符合期中分析终止规则并决定因有效性提前终止试验，在最终数据库锁定之前，试
238 验可能会继续收集除纳入期中分析之外的试验参与者数据。这可能是由于数据收集与
239 期中分析之间存在时间差，期间需完成数据裁决和清理工作所致。申办者应确保这些
240 额外的信息被恰当的记录，并报告期中分析结果和基于所有可用数据的分析结果，两
241 种结果对监管决策均具有重要意义。例如，若两次分析之间的疗效估计变化可能会影
242 响获益-风险评估，则需调查潜在原因，并可能对试验结果的解释存在挑战。

243 当试验设计包含无效性终止可能性时，虽然预期的无效规则应该预先设定和论证，但
244 通常建议采用非约束性无效性规则。这意味着无效性终止标准仅作为参考，根据期中
245 分析结果进行偏离时不会增加 I 类错误概率。这种灵活性至关重要，因为是否因无效性
246 而终止或继续试验的决策过程通常并非纯粹的算法化过程，可能需要考虑除主要有效
247 性终点以外的信息，如安全性或其他疗效数据。相比之下，虽然有建议采用约束性无
248 效性规则，并根据预设的无效性标准调整有效性决策标准，但此类方法的弊端在于要

求申办者必须遵守预设的无效性终止标准，否则将无法控制住 I 类错误概率并影响对试验结果的解读。

4.2 样本量调整

即使经过精心策划和实施的早期开发计划，影响临床试验样本量计算的参数假设中仍可能存在较大的不确定性。其中一个不确定性的来源是关于干扰参数的假设，这些参数并非主要关注参数，但可能会影响试验的样本量。干扰参数的示例包括连续结局的标准差和二分类结局中对照组的发生率，这些参数在某些疾病领域的不同试验中可能存在巨大差异。当干扰参数存在较大不确定性，且有合理理由的情况下，申办者可能会考虑根据干扰参数的期中分析估计结果来调整初始样本量，以确保试验具有足够的检验效能。参数假设中的另一个不确定性来源是对预期治疗效应大小的假设。当有合理依据认为仍存在不确定性（比如经过适当的探索性试验后；见第 3.1 章节）的情况下，申办者可能会考虑根据期中分析的疗效估计进行样本量调整。其目的是确保在一系列合理且具有临床意义的疗效范围内，试验均能保持足够的检验效能。

在设计包含样本量调整的试验时，应预先设定和论证潜在样本量的最小值与最大值、预期的样本量调整决策规则以及统计分析方法。需注意的是，最小样本量须仍能够为获益-风险评估（比如安全性、次要终点和亚组分析的评估）提供充分依据，这与早期终止试验时的考虑相一致（第 4.1 章节）。

基于干扰参数估计值进行样本量调整时应使用盲态数据，该方法不包含试验分组信息，可最大限度降低破坏试验完整性的风险。应预先设定预期的样本量调整规则，以确保调整的样本量不受非盲态数据的影响。预先设定的规则也有助于评估试验的操作特征（比如把握度和预期样本量）。申办者应提出一种能控制 I 类错误概率的检验方法并论证该方法的合理性。在某些情况下（比如采用连续终点且有合理样本量的双臂优越性试验），若依据充分，可采用适用于非适应性设计的常规分析方法进行主要分析。在其他情况下（比如采用连续终点的双臂非劣效性试验），使用这些常规方法可能会导致 I 类错误概率膨胀，此时需采用不同的分析方法。

对于基于非盲态数据的期中分析疗效估计进行样本量调整的试验，应建立 IDMC 和充分的控制措施来保持试验的完整性，包括最大限度降低从样本量调整中反推疗效信息的措施（第 3.5 章节）。由于此类设计通常会使样本量有所增加（相对于初始计划样本

量), 故而即使观察到的疗效弱于初始计划值仍有可能达到统计学显著性。因此, 在计划此类设计时, 需判断具有临床意义的疗效大小、论证额外试验参与者暴露的合理性, 并确保在该适应性设计下的潜在样本量从临床角度来看仍是合理的。

通常建议采用无需严格遵循预设调整规则的样本量调整方法, 例如采用基于预先设定的权重加权合并各试验阶段信息的假设检验方法。但尽管如此, 仍需预先设定预期调整规则, 以便评估试验操作特征 (比如预期样本量和检验效能), 并确保 IDMC 理解并认同该调整规则。

对于大多数基于期中分析疗效估计进行样本量调整的试验设计而言, 用于非适应性设计的常规检验方法并不适用, 需要使用特定的统计方法来控制 I 类错误概率。此外, 疗效大小的常规点估计可能存在偏倚, 常规的置信区间也可能无法达到预期的覆盖率。因此, 建议在试验设计阶段进行估计值的可靠性评估。该评估将有助于判断提出的适应性设计的可接受性, 或指导试验结果的解读。在某些情况下, 可采用调整估计值的方法来减少或消除适应性调整带来的偏倚, 通常此类方法在该情况下更为可取。

4.3 人群选择

在某些情况下, 对于哪些患者应该接受新疗法, 可能仍然存在不确定性。例如, 某种疗法可能会使整体人群中的某个目标亚群获益, 而非目标 (补充) 亚群的获益可能尚不明确。此类目标亚群可通过人口统计学特征、或假定与治疗作用机制相关的基因/病理生理学标志物等来界定。如果治疗对目标亚群确实有效, 但对补充亚群无效或疗效甚微, 则在整体人群中进行试验可能因为统计效能不足而无法确定疗效, 并且可能使试验参与者不必要地暴露于对他们无效的治疗。反之, 如果治疗对整体人群均有效, 则仅针对目标亚群进行的试验将无法获得补充亚群的疗效数据, 并将导致治疗适应症仅限于可获益的整体人群中的一个亚群。

此类不确定性通常可以通过探索性试验进行研究。然而, 在某些情况下, 也可考虑在整体人群中进行确证性试验, 并制定包含目标亚群有效性评价的分析计划 (比如采用多重检验方法来控制整体人群和亚群分析中的 I 类错误概率)。或者, 更高效的可能做法是考虑采用带有适应性人群调整选项的确证性试验设计, 即可根据非盲态的期中分析结果调整患者人群的试验设计。试验可能会在期中分析前从整体人群中招募试验参与者, 在期中分析时将决定是否继续在整体人群中招募, 或者将后续的招募限制在目

标亚群中。如果继续在整体人群中招募，则需要决定在试验结束时的分析中仅评估整体人群的疗效，还是同时评估整体人群和目标亚群的疗效。如果后续招募仅限于目标亚群，则试验结束时的分析将重点关注该亚群的疗效。在这种情况下，应适当合并期中分析前后累积的数据，以推断选定人群的疗效。

对此类设计的充分计划应包括：预先设定在期中分析中可能被选为后续招募目标的候选人群，在期中分析时需做出的关于统计推断目标人群的决策及其在试验结束时的分析方式，以及预期的适应性调整规则。还应制定相应计划，用于管理那些基于期中分析而停止进一步招募和评估的人群中已参与试验的人群。在根据期中疗效估计选择人群进行后续招募和分析的设计中，通常需要采用特定的统计方法来控制I类错误概率。通常建议采用允许灵活偏离预期适应性调整规则的方法，因为在期中分析时综合考虑所有的可用信息有助于确保人群选择的适当性。申办者还应确保期中估计有助于进行可靠的人群选择，包括在合适的时间点进行期中分析。此外，鉴于此类设计倾向于选择期中结果更有利的人群，试验结束时常规的疗效估计可能存在偏倚。应评估不同人群中疗效估计的可靠性，并考虑采用可减少或消除偏倚的调整后的估计。

值得注意的是，带有适应性人群调整的试验必须有合理的科学依据。例如，在整体人群中进行的试验包括期中分析，且可能将后续的招募和分析集中在特定亚群上，那么这种设计应基于既往试验结果和/或生物学证据，证明获益-风险特征在目标亚群中可能表现更优。对于此类试验，试验设计还需确保能够在期中分析后继续对整体人群进行招募，为可靠的决策提供支持。包括确保试验能够提供关于评估补充亚群获益-风险特征的足够信息；也包括制定具体标准（包含补充亚群中估计疗效的评估标准），以支持论证整体人群是否获益的结论。如果用于定义亚群的基线特征非二分类变量，则应在计划阶段对用于界定亚群的阈值进行合理论证。

4.4 治疗选择

部分试验旨在评估多种治疗方案，这些方案可能涉及不同药物或同一药物的不同剂量水平。例如，在药物开发的探索性阶段结束时，某药物的两种潜在有效剂量的获益-风险特征可能仍存在不确定性，此时可在确证性试验中将这两种剂量与对照组进行比较，以确认其疗效并在试验结束时选择最合适的剂量。在这种情况下，可考虑能基于期中分析累积的非盲态数据来进行剂量选择的试验设计：试验参与者最初将被随机分

配到两个剂量组或对照组；在期中分析时，选择一种或两种剂量进入第二阶段继续随机；最终分析旨在根据两个阶段的数据确认所选剂量的疗效并评估其获益-风险特征。

采用适应性治疗选择设计的临床试验应进行充分计划，包括明确待评估的治疗、在期中分析时需作出的决策、以及治疗选择的预期规则，同时包括对随机化方案和总样本量的影响。此外，还应制定方案以妥善管理那些正在接受治疗的参与者，而这些治疗在期中分析后决定停止后续评估。在基于期中疗效估计选择一种（或多种）治疗的设计中，需要采用特定的统计方法来控制 I 类错误概率。通常建议使用允许灵活偏离预期调整规则的方法，这种灵活性使得在期中分析时能够综合考虑所有可用信息，从而帮助支持更明智、更合理的治疗选择决策。申办者还应确保期中分析结果能够支持进行可靠的治疗选择，包括计划合适的时间点进行期中分析。鉴于此类设计倾向于选择期中分析结果更优的治疗，试验结束时常规的疗效估计可能存在偏倚。因此，应对估计的可靠性进行评估，并考虑采用可减少或消除偏倚的调整后的估计。

4.5 试验参与者分配调整

在随机试验中，通常根据固定的随机分配概率将试验参与者分配至各治疗组。或者也可以考虑采用其他的适应性分配方案，即根据早期入组试验参与者的数据动态调整分配。这些方法包括协变量适应性随机分配，即分配情况取决于累积的基线协变量数据，以及应答适应性随机分配，即分配情况取决于累积的结局数据。本章节重点讨论应答适应性随机分配（Response-Adaptive Randomization, RAR），即根据先前非盲状态下结局数据调整随机分配概率。其核心思想在于，将新试验参与者以更高概率分配至比当前其他组结局更佳的治疗组。

RAR 对试验参与者的优势体现在例如减少暴露于疗效较差治疗的试验参与者数量，或在二分类疗效终点试验中减少预期的试验参与者治疗失败事件数。然而，RAR 在确保有效统计推断方面也带来诸多挑战。其中最值得关注的是，顺时间趋势下 RAR 设计容易产生偏倚并导致 I 类错误概率膨胀。例如，如果早期入组试验参与者比晚期入组试验参与者更有可能被分配至对照组，且预后较差（比如由于背景治疗或试验参与者特征随时间变化），则 RAR 设计更有可能出现假阳性疗效结论。此外，采用基于疗效的算法修改随机化方案，可能会导致样本量不足，无法支持对疗效较弱但获益-风险比更优的治疗进行决策。

任何采用 RAR 的建议都应解决这些潜在问题。需预先明确具体的 RAR 方法并论证其合理性。应详细说明能控制 I 类错误概率并提供可靠估计的分析方法。还需解决由于时间趋势而造成混杂的可能性，这种混杂的程度可能取决于诸如试验预期持续时间、背景治疗或预后因素随时间变化的可能性（比如在快速演变的传染病场景中更为常见）。控制 I 类错误概率的一种方法是仅允许在单次或少数几次期中分析时调整随机分配比例，同时利用基于预设权重的适应性假设检验来整合各试验阶段的信息。针对时间趋势问题也可采用特定方法（比如重随机化检验）来处理，但使用此类检验的 RAR 设计可能不如固定随机化方案的设计强大。

为降低对试验完整性的风险，应预先制定一种无需申办者干预即可随时间推移实现随机化方案变更的方法。鉴于 RAR 方法和适应性选择的随机化比例可能泄露关于期中疗效估计的信息，应采取措施最大限度减少从调整中可推断出的内容（第 3.5 章节）。最后，该方法还可能面临其他挑战，例如确保持续及时获取高质量的期中数据，并将算法集成至随机化系统中。

此外，还存在某些非随机化的、确定性的调整，例如在双臂试验中，若试验参与者出现治疗应答，则下一例试验参与者继续接受相同治疗，若无应答则接受替代治疗。由于存在较高的偏倚风险以及可能预测后续治疗分配的问题，此类确定性调整方法不推荐使用（ICH E9）。

5. 专题讨论与注意事项

本章节将扩展探讨适应性设计中的其他若干专题，包括数据监查、模拟、贝叶斯方法的应用、至事件发生时间的终点指标、探索性试验及操作实施等。

5.1 关于数据监查的特别考虑

本章节重点探讨在基于累积非盲态数据进行期中分析的适应性设计确证性试验中，数据监查需特别考虑的事项。对于采用适应性设计的试验，IDMC 除履行常规职责（即在维护试验完整性的同时保障试验参与者个体安全）外，还应具备提出调整建议所需的全部专业能力。IDMC 应至少包含一名在期中监查及相关适应性设计与统计分析方法方面具备专业知识和经验的统计人员。IDMC 通常有权限获取非盲态有效性和安全性数据。应在指定的章程中明确操作规范，包括但不限于：拟提交报告的内容与提交

频率、会议日程与后勤保障、保密程序、监查计划的统计学要求，以及形成建议的具体流程。申办者需提前与 IDMC 就以下事项达成共识：试验目标与设计方案、对 IDMC 的履职要求（包括超出常规职责的部分）、适应性调整的类型及影响，以及预设的适应性调整规则。

应设立一个独立的统计团队，负责对累积的非盲态数据进行分析，并为 IDMC 生成期中分析报告。该团队不得包含监查委员会成员，且不应参与试验的其他相关工作。当有权限访问非盲态数据的该统计团队独立于申办者时，最能有效的维护试验的完整性。组成该团队的统计人员和程序员应具备实施适应性设计所需分析以及为 IDMC 提供支持的专业能力，并有权限访问履行其职责所需的全部试验数据。强烈建议非盲态的期中数据和分析结果仅由独立统计团队和 IDMC 掌握，同时需建立完善的保密流程（比如标准操作程序和保密协议）。

IDMC 在审查非盲态期中分析结果后，应将调整建议提交给与试验团队隔离的申办者指定人员。在 IDMC 建议提前终止试验的特定情况下，可向申办者披露相关充分信息（比如关键有效性和安全性结果）以支持其决定是否终止试验。但通常情况下，试验方案应设计为：申办者无需访问任何非盲态期中分析结果，即可实施 IDMC 关于试验调整的建议，例如在成组序贯设计中 IDMC 建议继续试验，或在样本量调整设计中确定具体样本量时。这要求申办者与 IDMC 在试验设计阶段进行充分讨论和计划，就监查流程和预期调整规则达成共识。

最大限度降低试验完整性风险的方式是完全限制申办者访问非盲态期中分析结果。但在特定情况下，申办者可能会提出有限访问非盲态数据。这一点应在试验计划阶段明确说明，且任何申办者的访问申请都需要提供令人信服的依据。若允许申办者接触数据，则应制定周密的措施来维护试验完整性，例如严格限制有访问权限的人数，确保有访问权限的人员完全独立于试验实施人员，并建立完善的保密流程。所有关于“何人接触何种数据”的信息都应被详细记录，以便监管机构在试验结束后评估适应性调整前后的结果时，能够确信试验完整性未受到损害。

5.2 模拟研究的计划、实施和报告

模拟研究通常在适应性设计的试验计划中发挥重要作用。模拟指重复执行大量的假设性临床试验，以了解在一系列特定配置的假设（场景）下试验设计的操作特征。模拟

可用于探究拟定适应性设计在不同场景下的操作特征，这些场景包括：基于不同疗效和干扰参数假设、存在不同脱落率或入组率、或者当分析方法的分析特性依赖于大样本量而采用特定样本量时的情况。例如，在假设无治疗获益的情况下模拟数据时，可通过计算导致得出“治疗有效”的错误阳性结论的模拟临床试验比例，来估计假阳性结论的概率。模拟有助于比较适应性和非适应性设计、不同适应性设计方案以及不同药物开发计划的差异（即比较一系列试验的差异）。模拟还可以为申办者关于试验后勤（比如研究中心选择和药物供应）的内部决策提供依据。本节重点介绍模拟的计划、实施和报告应遵循的原则，这些原则对于理解适应性设计试验的操作特征至关重要。

明确定义并聚焦模拟研究的关键目标非常重要。这些目标应具备明确性、相关性，并与模拟研究结果所做出的决策直接相关。为达成目标，需审慎选择一系列临床试验设计和分析方案，其中应包括基准设计和分析方法，即具有易于理解的操作特征的设计，比如非适应性设计或成组序贯设计。该设计范围可能还包括：期中分析次数与时点的选择、分阶段样本量设置、适应性调整类型、终止与调整规则以及假设检验和参数估计的统计方法等。设计方案的选择可能是一个随着操作特征的探索而调整的迭代过程，并且应足够广泛，以便对选定的适应性设计进行全面评估。评估模拟研究中所有设计方案的优势和局限性，对于理解设计选定时的权衡至关重要。

同时，还需定义和评估与模拟研究旨在解决的问题相匹配的关键操作特征。这些操作特征通常应包括：I类错误概率、预期样本量、预期试验持续时间、把握度、置信区间覆盖率以及疗效估计的偏倚和均方误差。根据试验设计和应用场景，可能还需要关注其他操作特征，如期中分析时因无效性或有效性而提前终止试验的概率。采用贝叶斯方法的适应性设计的操作特征考虑见第5.3章节。在某些情况下，还需关注单次试验以外的操作特征，比如选择合适剂量并随后确证其疗效的概率。虽然汇总模拟试验（重复）结果的平均值很重要，但评估结果分布的变异性、最小值、最大值或其他方面（比如具有提前终止或样本量调整可能性的试验中的样本量分布）也可能很重要。

模拟研究中包含的场景应涵盖合理的假设范围，以确保对拟定适应性设计的性能进行稳健评估。其中包括对疗效和干扰参数的假设（比如连续型结局的标准差），以及申办者可能有更大控制权的操作假设（比如入组率或脱落率）。假设的充分性应基于临床和统计学考虑进行论证，并需提供支持性文件证明。这些信息可来源于多种渠道，包括

既往试验数据、出版物、外推结果和专家意见等。申办者应使用所有可获取的相关信息，并尝试量化不确定性、识别潜在偏倚。使用假设网格（比如特定范围内的离散假设集）应基于现有临床知识进行论证，即证明网格中评估的范围涵盖所有可能的合理情况。此外，还需证明（比如基于单调性论证）网格足够精细（即范围内包含足够数量的不同假设），以对相关操作特征进行可靠估计。基于可靠证据且符合临床认知的信息来源，将使模拟研究结果更具可解释性和说服力。

模拟场景必须全面涵盖干扰参数配置的合理范围。例如，在采用模拟来考察 I 类错误概率时，即使在最简单的试验设计中，也不可能在符合无治疗获益的每种干扰参数设定下进行模拟。因此，对于依赖模拟来确定 I 类错误概率的设计，会存在额外的不确定性。考虑到这种额外的不确定性，预计需要提供更多的论证依据来支持此类设计。

应说明模拟研究的实施细节并论证其合理性，包括明确说明数据生成过程。在多数情况下，简单的统计模型可能适用，例如基于既往试验获得均值和方差的正态分布模型。在其他情况下，可以考虑基于早期试验结果（比如纵向患者特征）进行更复杂的模型拟合。同时，还需确定所需模拟重复次数，以达到关键操作特征的估计精度。对于某些操作特征或场景，可能需要达到更高的精度。例如，为达到 I 类错误概率的估计精度，每个情景可能需要进行 10 万次或更多次重复模拟；而对于把握度等操作特征，较少重复次数即可满足需求。所有算法均需进行记录，并采用可复现的方式生成随机数，如使用已记录的随机种子。

最后，必须详细记录模拟研究的设计、结果和结论。在进行试验之前，应在向监管提交的文件中纳入一份全面和结构化的模拟报告（第 6.1 章节）。临床和统计假设与模拟结果之间应有明确的联系。报告应符合本章节概述的考虑因素，并包含以下内容：

1. 模拟研究旨在解决的关键问题。
2. 模拟研究中评估的临床试验设计和分析方案。
3. 模拟研究中评估的操作特性的选择。
4. 现有知识以及任何支持数据或参考资料，为模拟场景的设置提供信息。

- 469 5. 用于模拟场景的参数组合配置，以及基于现有知识的临床论证，证明该组合能
470 充分覆盖各参数合理的取值范围。
- 471 6. 实施细节，包括数据生成过程和每个场景的重复次数，及其选择的依据。
- 472 7. 用于模拟的软件包，如果使用定制软件需提供模拟代码。提供代码时，应包含
473 充分的注释及详细的代码执行说明（比如调用示例和起始种子）。
- 474 8. 提供总体结果、解读和结论的摘要。摘要中应详细讨论拟定适应性设计及其在
475 各种情况下的估计操作特征。尽可能以交互式图表的形式汇总结果以提升可阅
476 读性。
- 477 9. 存在不同调整方案及结论的单次模拟临床试验的相关示例说明。例如，在具有
478 样本量调整的设计中，可能包括在期中分析时采用不同的样本量调整方案、并
479 获得阳性或阴性主要分析结果的情况，以便更好地理解潜在期中决策及其对试
480 验结果的影响。
- 481 10. 对限制模拟结果解读的任何方面的说明（比如假设或外推的不确定性）。
- 482 11. 关于模拟结果是否以及在何种程度上解决关键问题的临床讨论。
- 483 仔细记录模拟研究非常重要，因为模拟和相关结论的有效性将是试验结束时对结果进
484 行监管审查的一部分。

485 5.3 采用贝叶斯方法的适应性设计¹

486 ICH E9 指出，当采用贝叶斯方法的理由明确且所得结论足够稳健时，可以考虑在临床
487 试验中使用贝叶斯方法。贝叶斯方法是指结合先验概率分布和当前试验数据来获取感
488 兴趣的参数（比如疗效或估计目标）的后验概率分布的一系列统计方法。贝叶斯方法
489 可能适用于多种类型的适应性设计。无论采用何种统计方法，都应遵循第 3 章节中概

¹ 本章节关于适应性设计中贝叶斯方法的应用尚未完全达成共识。广泛运用贝叶斯方法于监管决策并非在所有情况下都是合理的。如 ICH E9 和本指导原则草案所述，只有当采用贝叶斯方法的理由明确且所得结论足够稳健时，方可在临床试验中考虑使用贝叶斯方法。现就以下方面公开征求意见：贝叶斯方法在哪些情况下符合适应性设计的核心原则，以及在哪些情况下可以考虑使用贝叶斯方法。

述的原则。贝叶斯方法在临床试验适应性设计中的应用形式多样，各自需考虑的要点也有所不同。

贝叶斯方法可用于指导试验中的适应性调整，同时应确保主要分析的决策标准能实现控制 I 类错误概率。例如，某项试验可以在期中分析时包含一个有预设的非约束性无效中止试验规则，这一规则可基于治疗无效的后验概率或试验结束时拒绝原假设的预测概率，而试验的主要疗效分析仍采用基于预设显著性水平下的频率学假设检验。对于此类设计，其操作特征应与不使用贝叶斯方法的适应性设计保持一致的预期表现。申办者应论证其所设定的先验分布、决策标准以及适应性设计要素（比如期中分析的次数和时间点以及适应性调整规则）可以实现预期的操作特征（比如效能、预期样本量、适应性调整决策的可靠性），同时控制 I 类错误概率。

一个特殊应用场景是在采用贝叶斯方法的临床试验中引入适应性设计要素，通过信息性先验分布借用外部信息，并基于相关估计目标的后验分布（比如有效性后验概率的阈值）设定主要分析的决策标准。借用外部数据辅助推断时，必须提供充分的科学依据，既要考虑不借用外部数据的替代方法的可行性（比如在不使用外部数据的条件下设计和开展有完全把握度的试验），又要确保外部数据的相关性和质量。先验分布设定不当会导致无法控制假阳性结论的概率。确保先验分布准确反映相关可用信息并解决先验分布和当前试验数据之间冲突的潜在问题，会引入使用不借用外部信息的频率论推断时所不存在的额外不确定性。

对于此类设计，申办者应在方案中讨论并记录用于生成先验分布的外部信息来源、外部信息与试验设计的相关性（比如人群和伴随治疗是否足够相似、终点是否相同）、所有潜在相关信息来源的列表，以及为何使用选定的信息来源而丢弃其他潜在相关来源。临床专业领域专家的意见对于评估外部信息的相关性至关重要。在考虑外部信息来源时，通常优先考虑随机对照试验数据和近期数据。通常要求提供患者层面的数据，因为这些数据有助于在计划阶段对外部信息的相关性进行全面评估，并有助于制定策略，在评估阶段解决先验和当前试验数据之间的潜在冲突。

申办者应预先说明并论证拟定先验分布的细节，包括外部数据的借用程度，以及试验成功的判定标准。该先验分布与决策标准须确保设计符合第 3.3 章节中的原则，包括控制出现假阳性结论的几率。对于先验分布的合理性说明，应包括讨论先验分布与试验

数据之间的平衡，并提出应对方案，以降低当前观察到的数据可能与先验分布产生冲突的风险。应有足够数量的当前试验数据来支持获益-风险评估。需进行模拟分析，以评估在先验分布与观察数据存在冲突的不同情景下，得出错误结论（包括假阳性结论）的可能性。在试验设计阶段，还应就以下内容展开讨论：最大可借用的数据量，以及实际观察到的冲突程度与借用强度之间的关系，包括在何种情况下应质疑外部数据的适用性，进而选择不借用数据。此外，还应计划进行敏感性分析，以评估试验结论在使用不同的合理先验分布设定下的稳健性。评估当前试验数据在不借用外部信息情形下所得到的结论，也同样重要。

5.4 以至事件发生时间为终点的适应性设计

对于主要终点为至某一事件发生时间的试验，还需注意额外的考虑。在此类以至事件发生时间为终点的试验中，试验的统计把握度取决于事件数，而非试验参与者人数。因此，此类试验在计划阶段计算样本量时，通常以固定的事件数为目标。此外，试验参与者的随访时间通常不明确指定，这意味着试验没有固定的观察期，所有试验参与者均随访直至发生一定数量的事件。因此，对于以至事件发生时间为终点的适应性设计试验，期中分析通常按目标事件数而非目标样本量来计划。此外，在以至事件发生时间为终点的试验中，基于期中疗效估计进行的样本量调整可能需要修改最初计划的事件数。例如，只需等待更长的时间让事件发生（允许更长的随访时间），而保持最初计划的试验参与者人数，即可实现比最初计划更多的事件数目标，或者增加试验参与者人数，亦或者二种方法兼用。在考虑增加试验参与者人数相较于增加事件数时，申办者应确保有足够的可用于获益-风险评估（比如理解药物的长期疗效以及评估患者人群的相关亚组，包括具有较低的基础事件风险的亚组人群）。

当每个试验参与者仅参与试验的一个阶段时，适应性设计最为直接。若期中分析前后收集的数据完全独立，则合并所有信息的统计分析可相对简单地进行。然而在以至事件发生时间为终点的试验中，部分试验参与者可能在某一阶段入组且未发生事件，但在后续阶段可能贡献事件数据。在期中分析中，利用已入组但尚未发生事件的试验参与者的信息（比如关于次要终点的信息）来指导潜在试验调整将违背独立性假设，并可能导致 I 类错误概率升高（即使采用适应性检验统计量亦如此）。因此，制定具有控制 I 类错误概率统计方法的计划很重要。一种方案是完全预先设定一个仅基于主要终点的调整规则，并且不允许有任何偏离。另一种方案是采用特殊方法，即根据期中分析

前后入组试验参与者划分试验阶段，并预先设定各阶段的固定随访时间或固定事件数。或者申办者可考虑采用设定更大目标事件数并包含期中分析时提前终止试验的设计，而非调整事件数。类似的概念问题及相应的考虑同样适用于纵向结局的适应性设计。这是因为，在期中分析时，若使用尚未完成所有随访的试验参与者的替代终点或中间结局信息会增加 I 类错误概率，除非采用适当的分析方法。

5.5 探索性试验的适应性设计

本指导原则重点介绍适应性设计在确证性临床试验中的应用。如果一项试验旨在确认有效性并支持获益-风险评估，那么遵循第 3 章节中的原则就至关重要。适应性设计也可用于药物开发早期的探索性试验，旨在获得治疗使用的广泛信息（比如剂量、用药方案、人群、终点的选择）。这一阶段的试验可能包含大量调整，以生成支持后续开发阶段重要决策的信息。本指导原则中所述原则也适用于早期探索性试验阶段，以确保试验结果的可靠性和可解释性，并支持基于此类试验的后续决策。

由于探索性试验的结果通常需要通过一项或多项单独试验进行独立验证，因此可能需要进行额外考虑。例如，方案可能只需阐明试验调整的一般原则即可，而无需详细规定具体的调整规则。比如在剂量递增试验中，需结合其他新出现的信息（比如不符合剂量限制性毒性标准的毒性反应信息）来综合考虑基于模型的剂量推荐。此外，采用适应性设计的探索性试验能够为其预期支持的决策提供可靠的信息，这一点非常关键。例如，为确证性试验中待研究的目标剂量的决策提供有力依据至关重要，因为得出的不恰当结论可能会对后续的开发计划产生严重的影响。在采用适应性设计的探索性试验中，保持试验完整性同样重要，但申办者在期中决策过程中的作用可能需要额外的考虑。例如，对适应性剂量探索试验进行监查，旨在为后续确证性试验确定合适剂量，可能需要做出多方面的调整决策，而这些决策需要来自申办者内部各学科领域的深度参与。申办者随后应当考虑试验拟解决的问题及其在开发计划中的定位，以及申办者参与监查过程与限制接触非盲结果以保持试验完整性之间的权衡。任何监查计划都应确保能够保护试验参与者的安全。

5.6 操作性考虑

适应性设计的应用可能会给临床试验的操作执行带来挑战，这些挑战应在试验计划阶段予以解决。例如，应采取措施最大限度减少可从期中分析中推断的信息，以保持试

575 验完整性（第 3.5 章节）。再例如，知情同意书应说明试验中进行适应性变更的可能
576 性。试验参与者应了解此类变更的原因（比如在期中分析时从多种剂量中选择具有最
577 佳获益-风险特征的剂量），明白这些变更反映了对研究治疗的认知提升，并且其权益
578 与安全仍受保障。此外，适应性设计试验所需的基础设施，如数据管理系统，可能不
579 同于非适应性设计试验。适应性设计临床试验通常使用交互式语音或网络随机化系统
580 来管理随机化，将试验参与者分配至治疗组。此类系统应与临床试验操作流程及药品
581 供应链机制充分整合。系统应内置预先设定的算法，以确保能够在最少申办者投入的
582 情况下处理可预见的情况（比如治疗组或随机化比例的变更）。此外，对样本量、治疗
583 组或试验参与者分配的调整可能会带来药物供应方面的挑战。其中一项挑战是药品生
584 产周期，因为快速调整可能会给药品供应链带来压力，如果无法及时获得充足的药品
585 供应，则会导致试验参与者的治疗延误。当一项适应性设计临床试验跨越多个国家甚
586 至区域时，这些挑战可能会更加严峻，因为需要将药品及时分发到这些地区。模拟可
587 能有助于支持试验计划和执行阶段与供应有关的决策。最后，应在计划阶段建立流
588 程，以确保相关期中数据能够得到适当验证和及时清理，为调整决策提供高质量的期
589 中数据支持。这可能包括要求正式的期中数据库锁定，以确保完成数据验证和清理活
590 动。

591 6. 文档管理

592 6.1 适应性设计确证性试验开展前的文档记录

593 文档记录是确证性试验充分计划的关键环节，有助于对拟定的适应性设计进行严格评
594 估。除临床试验方案或其他文件中通常包含的信息外，在适当的情况下，文档记录还
595 应包括以下内容：

- 596 1. 拟定的适应性设计的理论依据：该理论依据应同时包含临床与统计学考虑，论
597 证在确证性试验中采用适应设计的合理性，以及拟定的试验设计在临床开发计
598 划中的充分性。讨论与替代设计（包括非适应性设计）相比的优势和局限性，
599 将有助于监管机构评估拟定的适应性要素带来的额外不确定性的可接受程度。
- 600 2. 拟进行的适应性调整的描述：应包括试验中可能调整的各个方面、期中分析的
601 次数和时间点，以及预期的调整决策规则（比如在样本量调整设计中，需说明
602 根据期中疗效估计确定目标样本量的公式，包括潜在样本量的最小值和最大

值)。如果涉及在期中分析时选择估计目标(比如通过治疗或人群选择),则应对所有候选估计目标进行精确定义。

3. 统计分析方法的描述:应包括生成期中分析结果和指导调整决策的方法、主要和次要分析的统计方法(比如用于假设检验、估计疗效及相应不确定性)、以及重要的敏感性和补充分析。

4. 适应性设计实施的描述:应包括期中分析的执行人;审阅期中分析结果和提出调整建议和/或决策的负责人;相关委员会的成员、角色、职责和操作内容。

5. 保持期中分析结果的保密性和试验完整性的措施描述,以及操作执行的其他细节:应包括信息传输和访问流程;有权访问非盲态期中分析结果的人员;非盲态期中分析结果的访问控制;做出调整决策后发布信息的类型,信息发布人以及信息接受人;关于信息访问和发布的记录的保存位置。

6. 适应性设计重要操作特征的描述:如果模拟对于理解操作特征至关重要,则应包括一份描述模拟研究目的、设计、实施和结果的报告(第 5.2 章节)。

根据不同国家和地区的监管要求及惯例,这些信息应在试验开始前记录并将其纳入提交给监管的文件中。方案应涵盖试验设计所有的核心要素,包括:试验目的和相应的估计目标,以及试验设计、实施和统计分析的主要特征(包括所有适应性设计要素及其理论依据)。部分信息(比如 IDMC 操作和数据访问过程的详细信息)可记录在单独的文件中,如 IDMC 章程。在某些情况下,预设调整规则的详细信息应记录在具有访问限制的特定文件(而非方案)中,以保持试验完整性(第 3.5 章节)。

6.2 适应性设计确证性试验完成后在上市申请中应包含的文档

基于适应性设计确证性临床试验进行上市申请,应包括充足的文档,以便对试验结果进行全面审查。除了常规内容外,上市申请还应包括:

1. 第 6.1 章节中所述的所有前瞻性计划。

2. 适应性设计的实施信息,包括:期中分析的实际次数和时间、试验实施方面(比如基线特征、入组率、依从性、试验参与者保持率)在期中分析前后的显

- 628 著变化评估、用于作出调整决策的期中分析结果、试验不同阶段间结果的任何
629 显著异质性、实际作出的调整决策、预设的调整规则的遵循情况以及申办者揭
630 盲日期。如果实际执行与预设计划存在任何偏离（比如期中分析的次数或时间
631 方面，或预设调整规则的遵循方面），则应对以下内容进行分析讨论：偏离原
632 因、采取的任何以最大限度减少对试验完整性影响的措施、对试验结果解读的
633 其他潜在影响。
- 634 3. 关于数据访问计划流程的遵守和保持试验完整性的所有信息，比如：任何审计
635 结果，任何已知流程偏离的报告及其对潜在影响的讨论。
- 636 4. IDMC 的审议记录（比如所有闭门和开放的 IDMC 会议记录），包括任何与调整
637 决策相关的讨论记录。
- 638 5. 对适应性设计进行恰当分析的结果报告（比如经适当调整后的估计量、置信区
639 间和 P 值）。