

药物临床试验中应用贝叶斯外部信息借用 方法的指导原则（试行）

国家药品监督管理局药品审评中心

2026年1月

目录

一、概述.....	1
二、理论基础.....	2
(一) 贝叶斯统计数学原理和应用背景	2
(二) 后验概率和后验预测概率	4
(三) 贝叶斯统计计算原理.....	5
(四) 贝叶斯信息借用方法.....	5
三、适用性评估.....	7
(一) 科学合理性评估.....	7
(二) 借用信息来源的数据质量与可靠性评估.....	8
(三) 实施可行性和对试验完整性影响的评估.....	8
(四) 分析结果的可解读性和稳健性评估	9
四、试验设计与统计分析	9
(一) 借用信息来源及评价	9
(二) 先验分布构建.....	11
(三) 试验假设和决策标准.....	13
(四) 样本量	15
(五) 统计分析的考虑.....	15
五、模拟技术要求	17
(一) 试验设计模拟.....	18
(二) 统计分析模拟.....	20

(三) 文档要求	21
六、结果的解读.....	22
七、沟通交流.....	23
附录 1、词汇表.....	25
附录 2、模拟研究报告示例文档	27

一、概述

贝叶斯方法是基于贝叶斯定理，将先验信息与观测数据整合，通过定量计算得到后验分布再进行推断的统计学方法。在临床试验中，贝叶斯方法可以高效地利用已有信息，对于提高药物研发效率和缩短研发周期具有重要作用和意义。

目前贝叶斯方法在临床试验中已有较为广泛的应用，尤其在探索性临床试验中的药物剂量选择、疗效预测、安全性评价等方面。本指导原则的适用范围是以贝叶斯方法作为主要分析方法的关键性临床试验，通常作为主要证据支持药物上市。探索性临床试验或者确证性临床试验中的敏感性分析、补充分析、亚组分析、次要疗效终点的贝叶斯统计分析可参考本指导原则。

贝叶斯方法在临床试验中所涉及的方法多样且复杂。为保证本指导原则更具针对性和实用性，结合目前药物临床试验发展现状，本指导原则主要关注贝叶斯外部信息借用方法在临床试验中的应用。关于贝叶斯外部信息借用方法的具体定义见本指导原则第二章第四节。

本指导原则在介绍贝叶斯统计相关基础理论和定义基础上，重点阐述了在临床试验中使用贝叶斯方法时应开展的适用性评估内容，以及具体试验设计和统计分析建议、模拟技术要求，并且介绍了贝叶斯框架下统计分析结果的临床解读和综合评价。最后说明了不同情形下沟通交流的要求。

本指导原则仅代表药品监管部门当前的观点和认识，不具有强制性的法律约束性。随着科学的研究的进展，本指导原则中的相关内容将不断完善与更新。应用本指导原则时，还请同时参考药物临床试验质量管理规范（GCP）、国际人用药品注册技术协调会（ICH）和其他国内外已发布的相关指导原则。

二、理论基础

本章简述了贝叶斯统计相关的基础理论和重要概念，其中重点介绍了贝叶斯信息借用的定义和方法。

（一）贝叶斯统计数学原理和应用背景

贝叶斯方法是一种基于贝叶斯定理的统计方法，它通过整合先验信息与观测数据得到后验分布以进行统计推断和决策。记观测数据为 x ，参数 θ 的后验分布 $\pi(\theta|x)$ 由以下贝叶斯公式给出：

$$\pi(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{m(x)} \propto f(x|\theta)\pi(\theta)$$

其中 $f(x|\theta)$ 表示给定参数 θ 的条件下，数据 x 的概率（或密度），也被称作似然。 $\pi(\theta)$ 是先验，反映在获得数据之前对于参数 θ 的信念。 $m(x) = \int f(x|\theta)\pi(\theta) d\theta$ 为边际似然，用于确保后验分布具有概率的归一性。后验分布 $\pi(\theta|x)$ 是在获得观测数据之后，于先验分布基础之上对参数信念的更新。 θ 可能包括冗余参数，即模型中存在但不直接关注的参数。此时

常通过积分处理冗余参数，从而得到关注参数的边际后验分布以支持进一步的统计推断。

贝叶斯统计和频率学统计是两种不同的统计推断框架，它们在统计思想和统计推断（包括统计决策和参数估计）方面存在差异。

在统计思想方面，贝叶斯统计视参数为随机变量，通过概率分布来描述参数的不确定性。频率学统计视参数为固定的未知常数，基于抽样分布理论量化参数估计的不确定性。

在统计决策方面，贝叶斯方法通过计算关注事件的后验概率评估假设的可信度。例如，当 θ 为试验组的有效率时，贝叶斯方法能够基于后验分布 $\pi(\theta|x)$ 给出 $\theta > 30\%$ 或 $\theta > 50\%$ 的后验概率；而频率学统计通常基于假设检验给出的 p 值来判断统计学显著性。在参数估计方面，贝叶斯方法计算出后验分布，之后基于后验分布给出点估计（例如后验均值估计）和可信区间；频率学统计通常使用最大似然估计等方法给出参数的点估计，并结合抽样分布理论构造置信区间。

在错判概率方面，本指导原则主要关注的“假阳性率”，定义为试验药物无效但决策为试验药物有效的概率；“假阴性率”定义为试验药物有效但决策为试验药物无效的概率。同理，定义“统计效能”为试验药物有效并决策为试验药物有效的概率。

总体而言，在药物研发和临床试验中，贝叶斯统计为处

理不确定性和整合已有信息提供了一种新工具、新方法，在适用场景中能够更加高效地达到试验目的，充分利用已有信息优化试验设计，提高决策可靠性和研发效率。此外，贝叶斯统计框架对关注内容给出概率形式结果，便于决策者进行风险获益分析，辅助综合评价治疗价值。

同时需要关注到贝叶斯统计使用过程中的挑战：其一，不恰当的先验分布（尤其是信息先验分布）的选择，可能会导致误导性的结果。其二，贝叶斯统计通常涉及复杂的数值计算，尤其是在高维参数空间中计算后验分布可能需要耗费较多的时间和计算资源。其三，出于对先验及模型稳健性评估的需要，贝叶斯临床试验设计往往需要大量的模拟研究支持，这会增加其应用的复杂度。

（二）后验概率和后验预测概率

在使用贝叶斯方法的临床试验中，基于后验分布的后验概率和基于后验预测分布的后验预测概率常作为决策的重要依据。后验分布反映了在现有证据下对参数的认识，量化了参数的不确定性，是贝叶斯统计推断的核心。后验概率是在后验分布基础上，对某一特定情况（例如药物有效率大于30%）的概率描述。后验预测分布是在后验分布基础上得到的未来样本的概率分布，在考虑参数不确定性的同时还反映了抽样的不确定性。后验预测概率是指在后验预测分布的基础上，预测未来某一观测值或临床结局发生的概率。

在一个具体试验中，后验概率常用于统计决策，而后验预测概率通常用于试验过程中预测试验后续达到试验成功标准的概率。

（三）贝叶斯统计计算原理

在一些简单情形中，关注参数的后验分布存在解析形式。例如，对于服从二项分布的随机变量，若以贝塔分布作为事件发生率的先验分布，则其后验分布仍为贝塔分布。这种后验分布与先验分布能够属于同一分布族的简单贝叶斯模型被称作共轭模型。虽然共轭模型具有简单的后验分布推导，但此类模型并不能适用于所有的统计分析需求。

更加复杂的贝叶斯模型，通常不能推导出后验分布的解析形式。此时可采用数值计算方法进行抽样近似，例如马尔可夫链蒙特卡洛方法。在得到抽样样本后，可从中得到关注参数的后验分布和估计量，包括点估计和区间估计等。

（四）贝叶斯信息借用方法

在贝叶斯统计中，参数的先验分布可以反映来自已有数据和/或专家意见的外部信息，从而实现在当前研究中借用信息的目的。

在信息借用方法中，可交换性假设是支撑统计模型合理性的核心理论基础。它是指观测序列的联合概率分布在所有可能的排列下保持不变。根据可交换性假设，不同亚组、外部数据或相关数据源的参数可被视为同一潜在分布的随机

实现，从而允许模型借用信息。可交换性假设的判断需结合科学背景，例如外部数据与当前研究人群特征的匹配程度，不同人群中的治疗效应一致性等。

信息借用按照信息来源可以分为外部信息借用和内部信息借用，按照借用方式可以分为静态信息借用和动态信息借用。

外部信息借用是指作为主要证据的临床试验数据外的信息借用，例如罕见病研究将少量受试者分配至对照组，并借用历史外部对照信息进行统计推断。

本指导原则主要介绍在外部信息借用中常使用的两项先验分布形式，包括幂先验和混合先验。

幂先验是一种简单直观的信息借用方法。具体地，记外部数据为 x_0 ，当前数据为 x ，幂先验服从 $\pi(\theta|x_0, \omega) \propto L(x_0|\theta)^\omega \pi(\theta)$ ，相应的参数后验分布可表示为 $\pi(\theta|x, x_0, \omega) \propto L(x|\theta)L(x_0|\theta)^\omega \pi(\theta)$ ，其中 $\pi(\theta)$ 为初始先验，表示在收集到外部数据前的先验信息。这样，通过预设参数 ω 实现控制外部信息的借用量。当 $\omega = 0$ 时，幂先验方法将不借用外部信息；当 $\omega = 1$ 时，幂先验方法等同于完全借用外部信息。

混合先验表示为 $\pi_{Mix}(\theta) = \omega\pi_1(\theta) + (1 - \omega)\pi_0(\theta)$ ，其中 $\pi_1(\theta)$ 为有信息先验， $\pi_0(\theta)$ 为弱信息先验或无信息先验。有信息先验是指拟开展试验之外相关数据的分布形式，例如通过 Meta 分析、Meta 分析预测方法构造的参数分布。混合

先验通过纳入弱信息先验或无信息先验和权重 ω 控制外部信息的借用量。当 $\omega = 0$ 时，混合先验将不借用信息先验；当 $\omega = 1$ 时，等同于完全借用信息先验。

内部信息借用是指作为主要证据的临床试验数据中各亚组/亚群间的数据信息借用。例如在篮式试验中对某适应症的疗效估计借用其他适应症的信息、在多区域临床试验中对区域疗效的估计借用其他区域的信息等。此类内部信息借用的统计分析结果常用于补充分析和亚组分析。

三、适用性评估

在临床试验中采用贝叶斯外部信息借用的设计阶段，应从科学合理性、可靠性、可行性、可解读性方面进行充分评估，并与监管部门积极沟通，以确保贝叶斯设计和分析方法在具体试验中的适用性。主要评估内容包含但不限于以下各项：

(一) 科学合理性评估

需要基于研究背景和临床开发计划，通过统计、临床医学、临床药理等多学科的综合评价，说明采用贝叶斯外部信息借用方法适用的理由，例如：无法纳入足够的受试者、存有高质量的具有相似性的临床试验数据等。评估应包含定量分析及模拟研究，以阐述说明本研究使用贝叶斯外部信息借用的科学合理性。此外，评估内容应重点阐述以下两方面内容：

1. 先验信息

通过评估外部信息与拟开展试验的可比性和一致性判断可交换性假设是否合理，以及先验信息和拟开展试验是否能够满足统计模型和统计分析方法的基本假设。

2. 假阳性率和假阴性率

通过评估试验中信息借用量、决策标准、样本量下的假阳性率和假阴性率表现，确保错判概率在一定的合理范围内。还需通过对假阳性率的评估确保信息借用量不会主导试验结果。

（二）借用信息来源的数据质量与可靠性评估

常见借用信息来源包括既往的药物临床试验数据、真实世界证据、同类药物的研究成果等。应详细展示和说明选择借用信息来源的过程和结果，并充分评估借用信息数据来源是否真实、可靠、全面，以及系统评估信息来源的证据强度和与拟开展试验的可比性。

相关评价内容和方法还可参考其他相关指导原则，例如真实世界数据相关指导原则中关于数据的适用性评价和预处理内容，以及有关外部数据相关指导原则等。

（三）实施可行性和对试验完整性影响的评估

需在试验前全面评估使用贝叶斯外部信息借用方法对试验设计、操作、模拟可行性的影响和要求。如果试验涉及适应性设计和/或期中分析，需参照相关指导原则以维持试验

的完整性，例如适应性设计相关指导原则等。

（四）分析结果的可解读性和稳健性评估

应事先在统计分析计划或者独立的贝叶斯分析计划中明确贝叶斯统计分析细节，并需确保统计分析结果及呈现方式具有临床可解读性。同时应预设合理且充分的敏感性分析以评估分析结果的稳健性。

四、试验设计与统计分析

贝叶斯外部信息借用的临床试验基于关注临床问题的完整全面信息形成先验分布，之后通过先验分布和拟开展临床试验的观测数据计算得到后验分布，并基于后验分布进行统计推断，以此对试验药物的风险获益进行评估。

此外，应充分评估存在借用信息和拟开展临床试验数据之间发生不一致的潜在风险，在试验设计时应考虑其对相关设计元素的影响并预设相应的处理方案。

总体试验设计和统计分析原则参考 ICH E9 及 E9 (R1) 等相关指导原则，额外需重点关注贝叶斯相关试验设计和统计分析内容，包括先验分布的构建、决策标准的确定、样本量的设计、后验分布的计算、贝叶斯相关统计推断等。

（一）借用信息来源及评价

借用信息须有明确的数据来源。应在构建借用信息的先验分布前，提供所选择数据来源的理由和依据，并充分评价信息来源的数据质量。评估的完整性和全面性，是确保其能

准确代表对于关注问题已有认识的基础。申办者应充分考察所有相关数据源，尤其是与当前研究问题具有高度相关性的内容。为避免在借用信息时产生潜在的选择性偏倚，不应仅选择对试验药物有利而排除对试验药物不利的数据来源，以确保现有认识不被过高估计，从而避免认知偏差导致的结论偏倚。

还需要重点评价借用数据来源的数据强度和与拟开展试验数据的可比性。信息来源的主要评价内容包括但不限于：

1. 临床背景

对于信息来源的临床背景主要从受试者特征和疾病治疗两方面进行评估。

受试者特征的评估包括入选标准、排除标准、诊断标准/方法、人口统计学特征、疾病特征、既往治疗特征等。

疾病治疗的评估，不仅包括研究药物，还应包括基础治疗和合并治疗。所关注内容包括药物制剂、剂量、给药途径、时间、用药频次和持续时间，以及剂量调整、中断、停药和依从性的具体规则。如必要且适用，还需从研究药物的PK/PD 特征等方面进行评价。

2. 研究设计

已完成的临床研究通常作为借用信息的主要数据来源，可以从研究设计和研究结果两方面，对借用信息来源的证据强度和与拟开展临床试验间的可比性进行评估。研究设计特

征包括研究类型（例如注册研究或非注册研究）、试验阶段（例如确证性或探索性临床试验）、设计类型（例如对照研究、单臂研究，干预或非干预研究）、盲法特征、比较类型、样本量、统计分析计划等内容。

观察指标结果通常为被借用的信息，因此需要重点关注观察指标的定义和评价标准，包括临床意义、评估标准、评估方法、评估结果等。同时还需关注具体统计分析方法、伴发事件处理策略、缺失值处理方式、协变量调整等。

3. 时效性

借用信息的数据来源具有一定时效性，其各方面可能会随着时间而改变。例如，受试者人群特征、生活方式，以及疾病反应或进展的判断标准、疾病的临床实践、治疗结局、观察指标的定义与评价、后续治疗的影响等方面发生变化，从而导致历史数据与当前数据存在差异。申办者应充分考虑信息来源的时效影响，慎重选择已有较长历史的信息来源。

通过以上内容对借用信息来源进行系统全面的证据强度评估以及可比性评价，为借用信息量的确定提供依据。通常情况下，以注册为目的的随机、对照、双盲的确证性临床试验被认为是证据强度较高的信息来源。

（二）先验分布构建

1. 信息借用方式

拟开展临床试验可以通过多种方式借用外部信息，例如

仅借用对照组（安慰剂或者阳性药物）信息，仅借用试验组信息，同时借用试验组和对照组信息，借用试验组和对照组的疗效差异信息等。选择合适的信息借用方式是开展贝叶斯外部信息借用临床试验的前提，这与拟开展临床试验需解决的临床问题、外部信息的可比性等方面有关。

例如，在儿童等特殊人群的药物研发中借用成人临床试验的数据，在针对罕见病的试验中借用历史对照数据，此类借用情形通常出于伦理考虑或者将有限临床资源集中于试验药物；在桥接试验中借用已完成的其他人群研究数据，在临床有效性比对试验中借用已完成的阳性药物临床试验数据，此类借用情形可以提高试验效率和参数估计精度。

2. 先验分布形式

先验分布的构建首先应基于数据类型（连续数据、离散数据、时间事件数据等）采用不同的分布类型，例如正态分布，贝塔分布，伽马分布等。

根据信息来源构建明确的先验分布，需要注意的是构建的先验分布应能准确表达关于参数的已有认识。当先验信息来自多个信息来源，应通过合适的统计学方法整合为一个信息先验分布。

先验分布形式的选择需要综合考虑各种因素，包括临床背景、信息来源的特征和质量、借用信息与拟开展试验数据的可比性、拟开展试验的试验设计和相关假设等。

3. 权重的意义与选择

由于借用信息和拟开展临床试验在试验背景、受试者人群特征、疗效评价等方面可能存在差异，外部信息并不能完全等同于和替代拟开展试验数据。或者外部信息的借用量过大将导致先验信息主导拟开展临床试验的试验结果。因此在借用外部信息时，需通过调整先验分布中的权重来控制外部信息借用量。

在不同先验分布形式中，权重表现形式略有不同：幂先验法的权重表现形式为幂次参数；混合先验中信息先验和弱信息（无信息）先验的权重比例反映了当前试验参数和拟采用的信息先验的一致性程度。需注意，由于权重在不同先验形式中意义不同，因此不同先验形式中使用相同权重值对试验的影响是不同的。

先验分布中的权重值不仅应体现外部信息在科学和临床层面的先验信念，也应体现先验信息对最终结果的影响程度。因此权重的选择，是结合外部信息的证据强度和可比性评估结果，与拟开展试验中参数操作特征的综合考虑，需要在研究开始前与审评机构就先验分布权重值进行沟通并达成一致。

（三）试验假设和决策标准

与常规试验设计相同，贝叶斯统计临床试验设计也需要在开展前针对关键问题提出明确试验假设，并在研究方案或

者统计分析计划中明确决策标准以判断研究是否达到预设试验目的。合理的决策标准有利于提高决策的准确率和可靠性，降低决策风险和确保研究的科学性。制定决策标准时应考虑拟开展试验的借用情形和决策目的。

在采用贝叶斯外部信息借用的临床试验中，需要决策的情形包括但不限于对主要观察指标的优效性/非劣效性/等效性的确证性结论、桥接一致性结论、基于期中分析数据的决策。例如，在优效性临床试验中，假设更高的疗效参数 θ 反映更好的疗效，使用观察指标的效应量大于预设效应值的后验概率，以其是否大于预设概率阈值作为达到试验目的的决策标准，即 $\Pr(\theta > \theta_0 | x) > T$ ，其中 θ_0 表示观察指标效应量的预设效应值， T 表示预设的概率阈值， x 表示观测到的试验数据。如果该后验概率超过预设的概率阈值，则认为当前试验结果支持达到预设试验目的的结论。除可用于支持试验的最终结论以外，后验概率也可能用于期中决策，后验预测概率可以提供期中分析的支持性证据。

概率阈值的选择应基于临床意义和当前决策标准下操作特征的综合评估。对阈值的设定特别要考虑其对错判概率的控制和影响，尤其关注其对假阳性率的控制。

申办者需充分评估决策标准的合理性、科学性、可靠性，并在研究开始前与审评机构就决策标准（包括预设的效应值和概率阈值）进行沟通并达成一致。

(四) 样本量

合理的借用数据，可以为拟开展临床试验的样本量设计提供多种选择和灵活性，常见情形有：1) 保持与频率学方法设计相同的样本量，借用数据可以提高拟开展试验的统计效能和估计可靠性；2) 保持与频率学方法设计相同的统计效能，通过借用信息减少拟开展临床试验的样本量，从而降低研究成本和/或缩短试验周期；3) 保持与频率学方法设计相同的样本量，但调整随机分配比例，将更多受试者分配到试验组，从而获得更多试验组的直接数据。例如，通过借用对照组信息减少对照组分配的样本量。

在采用贝叶斯外部信息借用方法的临床试验中，样本量的设计并非完全依赖于数学公式计算结果，而是需要结合实际研发背景、临床实践、安全性、有效性、先验信息、决策标准、试验设计等因素进行的综合评价。尤其是通过试验中重要参数的操作特征表现选择出合适的样本量。

需要认识到拟开展试验作为关注临床问题的直接数据来源，是确证药物有效的主要证据，应确保样本量设计充分满足基于贝叶斯设计的疗效分析及安全性评价的需求，并符合监管要求。

对于借用信息所代表的样本量，应进行充分评估，例如使用有效样本量。

(五) 统计分析的考虑

在使用贝叶斯外部信息借用的临床试验中，通过整合先验分布和观测数据的似然分布得到的后验分布，是贝叶斯统计推断的主要对象。因此需重点关注后验分布计算的正确性，包括解析公式的正确性、抽样模拟的模型正确性和结果收敛性。

贝叶斯统计临床试验通过计算后验概率并与预设概率阈值比较以确定试验成功与否，而不再依赖于 p 值判断试验成功与否，然后提供参数点估计和可信区间描述后验分布的特征，反映主要关注的观测指标估计值的不确定性范围。

在贝叶斯统计分析中，评估试验方案中预设内容对试验结论稳健性的影响非常重要，因此需要事先设置合理的敏感性分析，尤其关注先验信息及先验借用量的影响。例如通过临界点敏感性分析评估先验信息借用量的影响，计算不同先验权重值下的更新后验分布及得出的相应统计决策结论，找出将结果从成功变为失败的特定权重临界值。敏感性分析还可用于评估先验分布中超参数的预设值、不同的借用信息来源下统计推断的稳健性等。

需要注意，在研究结束后发现借用数据与当前试验数据存在冲突，即先验数据冲突，或借用信息来源不满足全面性，则需要从已知和未知的关键因素方面分析其原因，并根据试验预设分析计划评估其影响或校正偏倚。

在适应性设计中应用贝叶斯方法时，需重点关注多次决

策引起的假阳性率的膨胀。需事先在相关文档中明确期中分析的时间点和决策标准，并设置相应的假阳性率控制方案并体现在决策标准的概率阈值中。同样在处理多个比较时，也需考虑多重决策引起的假阳性率膨胀风险。

在进行期中分析决策时，需关注预设的借用信息量是否适用于期中分析。由于期中分析时所获得的试验数据量要少于最终分析，因此期中分析直接使用最终分析时的借用信息量，可能会造成借用信息主导期中分析结果的情况。因此应慎重考虑基于期中分析的有效性终止设计。

总体而言，在应用贝叶斯外部信息借用的临床试验中，先验分布构建的重要部分是确定先验信息的借用量，以体现信息来源的证据强度和拟开展临床试验数据的可比性，并防止借用量过大导致先验信息主导试验结果；决策标准是判断拟开展试验达到预设试验目的的直接标准，与试验错判概率具有直接关系，应重点关注概率阈值的设定；样本量是试验实施的主要关注内容，也通常是试验使用贝叶斯外部信息借用的主要原因，需要通过多因素综合评价从而选择合适的样本量。这些设计参数并非独立确定，也非先后确定，而是结合试验关注的临床问题和操作参数的表现选择合适的设计组合，例如通过错判概率的表现选择合适的先验分布权重值、决策预设概率阈值和样本量的组合。

五、模拟技术要求

临床试验模拟技术是复杂试验设计中重要的工具，在贝叶斯试验设计中尤为重要。模拟技术在贝叶斯临床试验中的应用主要分为两方面，分别是设计模拟和分析模拟。在试验设计层面，通过模拟不同假设情形下的试验表现优化试验设计；在统计分析计算层面，通过抽样模拟进行后验分布计算、统计推断等。

复杂模拟可能需要高性能计算资源，过于复杂的情形会导致模拟可行性较差。因此，申办者应合理规划硬件需求和时间成本以优化模拟计划。贝叶斯模型比较复杂，应保证统计编程的正确性并确保算法得以正确实施。

申办者在实施模拟技术之前，需在统计分析计划（或独立模拟计划）中详细说明终点数据类型、分布特征、先验分布、模型选择、抽样算法、抽样数量、计算方法、决策标准等细节从而确保模拟方法的适用性和结果的可靠性。

（一）试验设计模拟

在试验设计阶段，可以通过不同假设情形下的大量模拟研究，开展操作特征的优化、错判概率的控制、不同设计框架和信息借用模型的对比分析等，为试验设计中参数设置提供充分依据。

当使用模拟技术评估贝叶斯临床试验设计时，首先需明确说明模拟研究与拟开展试验的试验目的关系，以及需解决的问题。为更好理解和评估所提出的贝叶斯试验设计，申办

者可考虑在模拟研究中加入与其他试验设计方案的比较，例如贝叶斯方法与频率学方法的比较。

试验模拟应基于合理全面的参数假设范围，同时覆盖合理的冗余参数的变化范围（例如入组速度，脱落率）。除常规试验设计参数外，在贝叶斯外部信息借用试验设计中还需重点评估贝叶斯方法相关参数，例如：1) 不同的信息借用模型（例如单臂借用或双臂借用，疗效借用或疗效差异借用，幂先验分布或混合先验分布等）；2) 借用信息量；3) 拟开展试验样本量和借用信息所代表样本量；4) 试验假设下的决策标准。

模拟研究应针对试验设计的操作特征进行评价，包括但不限于假阳性率、假阴性率、试验成功率等。对于操作特征的评价，既要关注点估计，也需要关注估计的变异性和平滑度。

对于通过模拟研究验证试验假阳性率的可控性，应事先确定药物无效假设的研究范围以及冗余参数的变异范围。虽然在大多数情况下不可能穷尽模拟所有可能的假阳性情况，但应尽可能确保研究范围的合理性与全面性，并选择充分的无效场景，这对于证明试验的假阳性可控和合理性是至关重要的。如果模拟结果显示假阳性率过高，可以考虑变更信息借用量、决策标准等办法，或重新审视先验分布或相应的统计模型。

建议申办者在模拟研究中选择具有代表性的参数设定

场景进行详细说明，这将有助于更加精准地理解试验设计。模拟结果应尽可能使用简显易懂、可视化的形式进行展示。

需要注意，当临床试验设计发生变更或最终观测数据不在模拟研究预设范围内时，申办者应考虑进行补充模拟研究。

对于在各种试验情形下的操作特征优化研究，应考虑不同操作特征对精度的要求选择充分的模拟次数。例如，对于假阳性率的模拟研究精度要求会高于对试验检验效能的模拟研究，因此一般会要求更多的模拟次数（例如至少 100000 次）。

（二）统计分析模拟

在试验完成后，对试验结果进行统计分析，包括计算后验分布、参数估计、统计决策等。在某些情况下，比如先验分布为共轭分布时，可直接根据公式推导得到后验分布。在大部分情况下，尤其是复杂分布或者高维数据时，无法直接通过公式推导得到后验分布，此时一般需要通过模拟的方式得到后验分布。

分析模拟应重点关注抽样的正确性，模型的收敛性以及估计的准确性等。当使用 MCMC 方法时，需要注意产生的数值链应在预热期后收敛。在进行推断时，需要剔除来自预热期的样本，并使用收敛诊断工具检查收敛性。

贝叶斯方法的统计决策主要基于后验分布，参数估计包括点估计和区间估计。首先采用模拟技术得到后验分布的样

本，然后可用后验分布的模拟样本得到参数的点估计值和可信区间。应事先考虑模型适用性、模型假设、模型拟合度、迭代次数等内容，以保证后验分布估计的准确性。

(三) 文档要求

当基于试验模拟证明使用贝叶斯方法的合理性时，申办者应向监管方提交独立模拟研究报告，该报告也可作为研究方案或统计分析计划的附件提交。模拟研究报告应该详细展示试验设计、借用信息来源、模拟方法、模拟结果、模拟结论。模拟研究报告应包含以下要点，可根据不同试验情况做增删。模拟研究报告示例文档见附录 2。

- 模拟目的（明确模拟研究需解决的核心问题）
- 试验设计备选方法（例如，频率学方法或者无信息先验贝叶斯方法）
 - 先验分布和借用信息来源
 - 有效性假设与冗余参数（明确研究范围）
 - 研究设计（模拟数据生成模型，先验分布模型，统计分析模型，试验成功决策标准）
 - 模拟结果总结（例如，假阳性率、统计效能等；先验分布对结果的影响；与其他备选方法的比较等）
 - 模拟研究代表性场景分析与解读（选择模拟研究中的代表性场景进行说明，解释试验设计逻辑并介绍结果）
 - 总结与结论（结合模拟结果进行试验设计的适用性评

估，以及针对潜在问题的解决方案）

- 模拟代码与说明文件（用于验证模拟过程的可重复性）
分析模拟细节，包括应用的先验分布、数据模型、分析模型、计算细节、决策标准，也应该写入统计分析计划中。若在试验过程中修订统计分析计划，应确保贝叶斯计算相关内容进行相应的修改并确保最终分析严格按照计划进行。

在提交临床试验结果时，应包括随机种子、软件（版本）、运行环境说明、程序（包括程序说明文件和详细程序批注），以及其他必要信息，确保监管机构可以复现和检验模拟结果。建议采用表格和/或可视化图形的方式展示结果（如收敛性，主要结论，敏感性分析）。

六、结果的解读

对于使用贝叶斯信息外部借用方法的试验，正确呈现和解读统计分析结果并结合临床实践进行获益风险评估是监管决策的关键环节。与采用其他统计方法设计和决策的临床试验类似，在贝叶斯框架下基于方案中提前设定的决策标准进行统计决策，是评价试验成功的重要标准之一。

贝叶斯方法的统计决策主要基于后验概率是否大于预设概率阈值。临床试验中的概率阈值并非固定值，而是结合临床实践和统计表现（例如错判概率结果），共同确定概率阈值、样本量、信息借用量的组合。

贝叶斯方法的统计分析结果需结合临床意义进行综合

评价，重点说明治疗效果的点估计大小及其对患者的实际影响，结合可信区间讨论结果的不确定性及其对临床决策的影响。同时还需关注贝叶斯方法相关敏感性分析评估结论的稳健性，以及避免先验信息主导试验决策。若适用，还需通过预设分析计划评估和说明当前观测数据和先验数据的不一致情形。

为了更加全面的评价试验结果，在贝叶斯外部信息借用方法结果的基础上，可参考弱信息（无信息）先验或频率学方法的结果进行比对解读。

七、沟通交流

申办者在药物临床试验申请前、药物临床试验过程中以及药品上市许可申请前等关键阶段，可以就重大问题与审评机构进行沟通交流。关于沟通交流的具体要求和内容参考药品研发与技术审评沟通交流相关管理办法。

目前应用贝叶斯外部信息借用方法的临床试验涉及三种沟通交流情形：

第一种情形是在通过沟通交流对临床试验方案及相关统计分析计划、模拟计划达成一致意见，例如 pre-IND 沟通交流、II 期临床试验结束/III 期临床试验启动前沟通交流。在该阶段中，申办者需要与审评机构就临床试验中涉及的贝叶斯方法相关细节达成一致意见，包括前文中提到的适用性评估、试验设计、统计分析、模拟计算等内容。

第二种情形是药物临床试验期间方案变更的沟通交流。根据《药物临床试验期间方案变更技术指导原则》，对于不会显著增加受试者安全风险，但可能显著影响试验科学性以及数据可靠性的实质性变更，应与审评机构进行沟通交流。临床试验中涉及到的贝叶斯相关设计因素的变更，例如先验信息来源的变更、信息借用量的变更、统计分析方法的变更等，均属于实质性变更。申办者应说明变更原因，并提交充分模拟结果。

第三种情形是新药上市许可申请前沟通交流，即 pre-NDA 沟通交流。该情形下会议目的主要是，探讨当前临床试验数据和贝叶斯统计分析结果是否满足药品上市许可的技术要求。

附录 1、词汇表

中文	英文
贝叶斯定理	Bayes' theorem
先验分布	Prior distribution
似然	Likelihood
后验分布	Posterior distribution
后验概率	Posterior probability
后验预测分布	Posterior predictive distribution
后验预测概率	Posterior predictive probability
信念	Belief
归一性	Normalization
冗余参数	Nuisance parameter
边际后验分布	Marginal posterior distribution
可交换性	Exchangeability
共轭先验	Conjugate prior
共轭模型	Conjugate model
最大似然估计	Maximum likelihood estimation
置信区间	Confidence interval
可信区间	Credible interval
统计效能	Power
错判概率	Probability of erroneous conclusions

假阳性率	False positive rate
假阴性率	False negative rate
马尔可夫链蒙特卡洛	Markov chain Monte Carlo (MCMC)
操作特征	Operating characteristics
幂先验	Power prior
混合先验	Mixture prior
正态分布	Normal distribution
贝塔分布	Beta distribution
伽马分布	Gamma distribution
Meta 分析预测先验	Meta-analytic-predictive (MAP) prior
决策标准	Decision criterion
有效样本量	Effective sample size (ESS)
超参数	Hyperparameters
预热期	Burn-in period
先验数据冲突	Prior-data conflict

附录 2、模拟研究报告示例文档

一、试验目的和估计目标

简要介绍拟开展临床试验的研究目的与估计目标。

二、试验设计

简要介绍拟开展临床试验背景和设计、试验目的等内容。

可以图示方式呈现试验设计、流程和不同治疗组治疗情况。

三、分析计划

简要介绍拟开展临床试验的统计分析计划，重点阐述贝叶斯相关信息，包括介绍借用信息来源及其评价，先验分布的构建，后验分布的计算和基于后验分布的统计推断。

四、模拟目的

阐述模拟报告需解决的核心问题和目的。

五、模拟方法

(一) 研究设计选项

拟开展临床试验计划使用的试验设计方案及其备选方案，以及试验设计中重点参数设置计划。

(二) 模型设置

1、临床数据生成模型

基于不同预设临床情形的参数设置，生成模拟临床试验数据的模型。

2、统计分析模型

例如包括主要疗效分析模型、缺失数据处理等。

（三）分析和决策标准

模拟研究中的统计分析方法和决策标准。

（四）操作特征

用于评估试验设计表现的具体指标。

六、模拟结果

（一）操作特征结果

模拟操作特征结果的关键总结内容，例如假阳性率、功效和其他操作特征、先验分布对结果的影响、与其他备选方法的比较。

（二）代表性场景

选择模拟研究中的代表性场景进行说明，解释试验设计的逻辑和结果呈现与解读，描述是否满足决策规则。

七、总结和建议

（一）模拟过程总结

概括总结整个模拟过程，结合模拟结果进行试验设计的适用性评估，基于决策规则所获试验建议，以及针对潜在问题的解决方案。

（二）方案设计建议

对于试验方案设计的建议，重点阐述贝叶斯相关特征的设计结论。

八、参考文献

九、附录

建议包括：缩略语、模拟代码与说明文件（用于验证模
拟过程的可重复性）、模拟结果图表等。