



**国家药品监督管理局**

National Medical Products Administration

## 国家药监局关于发布体外诊断试剂临床试验技术指导原则的通告 (2021 年第 72 号)

---

发布时间：2021-09-27

为指导体外诊断试剂临床试验工作，根据《体外诊断试剂注册与备案管理办法》（国家市场监督管理总局令第 48 号），国家药监局组织制定了《体外诊断试剂临床试验技术指导原则》（见附件），现予发布。该技术指导原则自发布之日起实施，原国家食品药品监督管理总局发布的《体外诊断试剂临床试验技术指导原则》（原国家食品药品监督管理总局通告 2014 年第 16 号）废止。

特此通告。

附件：体外诊断试剂临床试验技术指导原则

国家药监局

2021 年 9 月 16 日

## 附件

# 体外诊断试剂临床试验技术指导原则

## 一、适用范围

体外诊断试剂临床试验是指在相应的临床环境中，对体外诊断试剂的临床性能进行的系统性研究。临床试验的目的在于证明体外诊断试剂能够满足预期用途要求，并确定产品的适用人群及适应证。临床试验结果为体外诊断试剂安全有效性的确认和风险受益分析提供有效的科学证据。

体外诊断试剂通过体外检测人体样本提供检测结果，用于单独或与其他信息共同辅助判断受试者的目标状态（健康状态、疾病状态、疾病进程或其他可用于指导临床处置的疾病/健康状态等）。体外诊断试剂的“临床性能”即指体外诊断试剂由预期使用者在预期使用环境中使用，针对目标人群获得与受试者目标状态相关的检测结果的能力。

本指导原则适用于按照医疗器械管理的体外诊断试剂在中国境内进行的、用于中国境内注册申请的临床试验。

本指导原则旨在明确临床试验的基本原则和临床试验中需要考虑的关键因素，并对临床试验质量管理提出基本要求，用于指导申办者的临床试验工作，也为技术审评部门对临床试验资料的审评提供参考。

由于体外诊断试剂产品具有发展快、专业跨度大、临床预期用途各异的特点，不同产品的临床试验方法及内容不尽相同。申办者应根据产品具体情况，制定合理的临床试验方案，本指导原则内容也将根据体外诊断试剂发展的需要，适时修订。

## 二、基本原则

### （一）伦理原则

临床试验应当遵循《世界医学大会赫尔辛基宣言》的伦理准则和国家涉及人的生物医学研究伦理的相关要求，应当经伦理委员会审查并同意。

研究者需考虑临床试验用样本，如血液、尿液、痰液、脑脊液、粪便、阴道分泌物、鼻咽拭子、组织切片、骨髓、羊水等的获得和试验结果对受试者的风险，提请伦理委员会审查，确保临床试验不会将受试者置于不合理的风险之中，并按要求获得受试者（或其监护人）的知情同意。

### （二）科学原则

临床试验的开展应建立在临床前研究的基础上，具有充分的科学依据和明确的试验目的。应根据产品预期用途、相关疾病的流行病学背景和统计学要求等，对临床试验进行科学的设计，同时最大限度控制试验误差、提高试验质量，对试验结果进行科学合理的分析。在保证试验结果科学、准确、可信的同时，尽可能做到高效、快速、经济。

### （三）依法原则

本指导原则是在《医疗器械监督管理条例》（国务院令第

739号)、《体外诊断试剂注册与备案管理办法》(国家市场监督管理总局令第48号)和《医疗器械临床试验质量管理规范》的法规框架下制定的。体外诊断试剂临床试验的开展应符合相关法规、规章的要求。

### 1. 临床试验机构和人员

体外诊断试剂临床试验应当在具备相应条件且按照规定备案的医疗器械临床试验机构开展。

体外诊断试剂临床试验应按照同一临床试验方案在多家临床试验机构开展。申办者负责选择、确定体外诊断试剂临床试验的协调研究者,协调研究者供职的医疗机构为组长单位。协调研究者承担临床试验中各中心的协调工作。

组长单位伦理委员会负责审查试验方案的伦理性和科学性,参加试验的其他临床试验机构伦理委员会在接受牵头单位伦理委员会审查意见的前提下,审查该项试验在本临床试验机构的可行性,包括研究者的资格与经验、设备与条件等,一般情况下不再对试验方案设计提出修改意见,但是有权不批准在其临床试验机构进行试验。

体外诊断试剂临床试验机构应当具备临床试验所需的专业技术水平、组织管理能力,并能够开展伦理审查工作,具有与所开展临床试验相适应的人员、设施和条件等。具体包括但不限于:常规开展相关检测项目和/或疾病诊疗项目,具有相关诊断结果解读和疾病处置的能力,具有防范和处理临床试验中突发事件和

严重不良事件的应急机制和处置能力；具有能够满足临床试验需要的受试人群；具有必备的实验室检测条件，满足相关的检测实验室资质认定要求（如有）等。临床试验机构应能够确保临床试验严格按照方案实施，并能够配合产品注册申报过程，包括进行必要的补充试验、配合申办者组织的监查和稽查，以及药品监督管理部门、卫生健康管理部门开展的检查等。

临床试验主要研究者应具有设计并实施相关临床试验的能力、具有试验体外诊断试剂临床试验所要求的专业知识和经验，应熟悉相关的临床试验法规要求。参与临床试验的人员经培训后应熟悉相关检测技术的原理、适用范围、操作方法等，并能够对检测结果进行正确判读。临床试验统计学负责人应为具有相关专业背景、专业能力的人员。

## 2. 临床试验方案和报告

### 2.1 临床试验方案

开展体外诊断试剂临床试验，申办者应根据试验目的，综合考虑试验体外诊断试剂的预期用途、产品特征和预期风险等，组织制定科学、合理的临床试验方案。临床试验方案经伦理委员会批准后应在临床试验全过程中严格遵循。

各临床试验机构应执行同一临床试验方案，方案中对试验设计类型、对比方法选择、受试者选择、评价指标、统计分析方法、样本量估算和质量控制要求等做出明确的规定，并根据各机构情况合理确定样本量分配计划。

有关临床试验方案内容的具体要求参见《医疗器械临床试验质量管理规范》及其附件体外诊断试剂临床试验方案范本。

## 2.2 临床试验小结与报告

临床试验结束后应分别总结各临床试验机构的临床试验数据，出具临床试验小结，并附临床试验数据表、临床试验中所采用的其他试验方法或其他体外诊断试剂等产品的基本信息等资料。临床试验小结正文应包括如下内容：

2.2.1 临床试验概述（试验体外诊断试剂信息、试验流程图、对比方法信息、其他相关检测方法/试剂信息、受试者入组情况、样本采集和保存、检测仪器、纳入统计的样本量、样本剔除情况说明、试验数据的总结和描述性分析等）；

2.2.2 执行方案的版本号及版本日期，临床试验人员信息，试验开始时间和结束时间；

2.2.3 临床试验质量管理情况，包括：临床试验前培训、临床试验过程质控、偏倚的控制措施、记录管理、数据管理、样本、试剂和仪器管理情况等；

2.2.4 临床试验伦理情况的说明；

2.2.5 不良事件和器械缺陷的发生以及处理情况；

2.2.6 方案偏离、方案修改情况说明；

2.2.7 其他。

临床试验数据表内容至少包括：唯一可追溯的样本编号、人口学信息（性别、年龄）、受试者临床诊断背景信息、样本类型、

检测结果等。需要时附临床试验原始图谱等。

临床试验数据表应由临床试验操作者、复核者签字，临床试验机构签章（封面章和骑缝章）。

申办者、协调研究者对所有临床试验数据进行汇总和统计分析后，完成临床试验报告。有关临床试验报告内容的具体要求参见《医疗器械临床试验质量管理规范》及其附件体外诊断试剂临床试验报告范本。

2.3 临床试验方案、临床试验小结应当由主要研究者签名、注明日期，经医疗器械临床试验机构审核签章后交申办者。临床试验报告应当由协调研究者签名、注明日期，经组长单位医疗器械临床试验机构审核签章后交申办者。临床试验所有资料均应由申办者签章。

### 三、临床试验设计

体外诊断试剂临床试验设计与产品预期用途、适应证、适用人群（目标人群）、被测物特点、检测样本类型、产品使用方法（如使用者）和检测结果报告方式（如定性、定量）等直接相关。临床试验结论应能够证明产品的临床性能满足预期用途的要求，并支持说明书中所描述的相关内容。

根据产品特点和产品性能评价需要，体外诊断试剂临床试验可能包括不同的临床试验目的，有必要针对各个临床试验目的，分别进行科学的临床试验设计，包括选择适当的临床试验设计类型，确定适合的对比方法、受试者入组/排除标准和临床评价指

标等，并进行科学的样本量估算。

同时，临床试验是根据抽样得到的有限的受试者样本得出研究结果，对未来具有类似情况的受试者总体（目标人群）做出统计学推断的过程。因此在临床试验设计时，需要根据统计学原理，对试验相关因素作出合理、有效的安排，并对试验结果进行科学合理的分析。

对于全新的体外诊断试剂或相比已上市同类产品有重大差异的产品，在正式开展临床试验之前，可考虑设计一个小样本的探索性试验。探索性试验的开展有助于减少非预期的结果导致临床试验中需要改变关键设计的可能性。一般而言，为了做出科学上有效的确证推理，探索性试验数据不应与临床试验阶段的研究数据合并。本文不针对探索性试验的设计和管理提出具体要求，但探索性试验实施过程中可借鉴本文相关内容。

### （一）设计类型

根据在临床试验过程中试验体外诊断试剂检测结果对受试者的影响，体外诊断试剂临床试验主要包括两种设计类型：观察性研究和干预性研究。

观察性研究中，采用试验体外诊断试剂对样本进行检测的同时，受试者还会接受常规临床诊断和实验室检测，试验体外诊断试剂检测结果不用于患者的管理，不影响临床决策；临床试验中通过评价该检测结果与确定受试者目标状态的临床参考标准（或其他方法）判定结果的一致性，确认产品临床性能。



干预性研究中，试验体外诊断试剂检测结果将用于患者管理或指导治疗，通过评价治疗效果或患者受益，为支持体外诊断试剂安全有效性的判定提供证据。

临床试验设计时应根据体外诊断试剂的特点和预期用途，选择适当的设计类型。

本指导原则的内容主要基于观察性研究设计进行讨论并提出要求，其他临床试验设计的情形可依据具体的情况参照执行。

### 1. 观察性研究中的横断面研究和纵向研究

一般的，体外诊断试剂的观察性研究主要涉及横断面研究，即评价单一时间点采集样本的检测结果与临床参考标准（或其他方法）判定结果的一致性。除此以外，有些产品需要进行纵向数据研究，即需要多个时间点采集样本的检测结果才能评价产品临床性能。例如，某些用于治疗监测的体外诊断试剂，在临床试验中应对受试者及其样本中的被测物进行治疗前后多个时间点的观测，以证明被测物检测结果的变化与病情发展、治疗效果的相关性。临床试验方案中应根据被测物特点、疾病进程等明确受试者观测时间、临床评价指标等。

### 2. 观察性研究中对比方法的选择

一般情形下，观察性研究中，采用试验体外诊断试剂与临床参考标准进行比较研究，评价试验体外诊断试剂检测结果与受试者目标状态的相关性，临床评价指标一般包括临床灵敏度和临床特异度等。临床参考标准是指现有条件下临床上可获得的能够用来确定受试者目标状态的最佳方法，通常来自临床和实验室的医

学实践，包括：现有条件下公认的、可靠的、权威的疾病诊断标准（如组织病理学检查、影像学检查、病原体分离培养鉴定、长期随访所得的结论等），疾病诊疗指南中明确的疾病诊断方法，行业内专家共识推荐的或临床上公认的、合理的参考方法等。临床参考标准可能是一种方法，也可能是多种方法相结合。

对于境内已有同类产品上市的体外诊断试剂，临床试验亦可采用试验体外诊断试剂与已上市同类产品（对比试剂）进行比较研究的方法，评价两种方法检测结果的一致性，评价指标通常包括阳性符合率、阴性符合率等。对比试剂在预期用途、适用人群、样本类型、检测方法学、检测性能等方面应与试验体外诊断试剂具有较好的可比性。

为了更加全面地评价体外诊断试剂的临床性能，临床试验中有时需要将与临床参考标准的比较研究和与境内已上市同类产品的比较研究相结合，对产品的临床性能进行综合评价，从而支持有关预期用途的所有声称内容。

对于目前临床上不存在或无法获得适当的临床参考标准、或临床参考标准尚不能全面评价产品临床性能，同时境内亦无同类产品上市的情况，设计临床试验方案时，应在证明产品临床意义的同时，依据现有临床实践和理论基础，选择目前公认、合理的方法，进行比较研究，进一步确认产品临床性能。

临床试验对比方法的选择应根据产品预期用途、样本类型、检测结果报告方式以及临床参考标准和对比试剂的可获得性等

因素进行综合考虑，临床试验结论应能够支持预期用途声称的内容。临床试验方案中应描述对比方法的选择依据。

体外诊断试剂变更注册的临床试验可采用变更后产品与变更前产品进行比较研究；对于变更前后产品性能发生显著变化或增加临床适应证等情形，亦可采用变更后产品与临床参考标准或境内已上市同类产品进行比较研究，证明变更后产品的临床性能符合要求。

### 3. 观察性研究中的特殊情形

对于某些体外诊断试剂，临床试验设计中可能遇到需要特殊考虑的情形，例如：某些情况下，试验体外诊断试剂与对比试剂由于样本采集、处理、保存等差异导致不能使用同一份样本进行检测（例如：适用样本为拭子样本，但两种方法适用的拭子材质和保存液不同的情况），此时可针对每位受试者进行两次样本采集，并分别进行试验体外诊断试剂和对比试剂的检测，两次采集样本的顺序应遵循随机原则。需要注意的是，一般仅在一次样本采集不会影响下一次样本采集时才考虑采用此种试验方法。

#### （二）受试者选择和样本收集

临床试验方案中应根据试验体外诊断试剂的预期用途、目标人群和检测要求等合理确定临床试验受试者入组/排除标准、受试者分层入组要求和样本收集方法等。

##### 1. 临床试验的受试人群

临床试验受试者应来自产品预期用途所声称的适用人群（目

标人群），如具有某种症状、体征、生理、病理状态或某种流行病学背景等情况的人。非目标人群入组可能引入受试者选择偏倚，导致临床试验结果不能反映产品的真实情况。

受试人群应能够代表目标人群的特征，包括人口学特征（性别、年龄）、症状、体征、合并症以及疾病的阶段、部位和严重程度等；同时受试者应排除不适合该临床试验的生理或病理特征。

根据以上要求合理设定受试者入组/排除标准，并在临床试验过程中采取适当的措施确保只有符合该标准的人方能入组。

此外，受试者入组还需根据产品特点考虑其他可能的影响因素，如不同民族、不同种族、不同地域的影响等。

举例来说，用于疾病辅助诊断、鉴别诊断的产品，受试者应来自具有疑似症状或有相关流行病学背景的人，包括具有目标疾病状态的受试者和不具有目标疾病状态的受试者。具有目标疾病状态的受试者应能够尽量覆盖疾病状态的全部特征，包括症状典型和非典型、疾病的分型、分期、病程的长短、病情的轻重等，以评价产品的临床灵敏度；不具有目标疾病状态的受试者需包括具有相同或相似的症状、易与目标疾病状态相混淆的其他疾病病例等，以评价产品的临床特异度；此外还应考虑纳入可能对检测产生干扰的样本等。

## 2. 受试者分层入组

当体外诊断试剂临床性能预期在不同亚组的人群中有差异，且对某些重要亚组的临床性能需得到准确评价时，建议采用分层

入组的方式，且亚组的样本量应满足统计学要求。

分层入组是将目标人群划分为预先指定的非重叠的不同亚组，针对每个亚组分别入组受试者。例如，根据需要按性别（男性、女性）和年龄组（低于或高于特定年龄）对目标人群进行分层。分层入组方式不仅确保了对重要亚组的充分评价，还有利于获得更准确的性能结果。

### 3. 样本收集

体外诊断试剂临床试验中应按照临床试验方案规定的入组/排除标准、受试者招募方式、样本采集方式等规定进行受试者入组并采集样本。

特定情况下，某些受试者样本可以来自既往的、其他研究的样本集或无特定用途的样本集。此种情形下则需特别注意避免引入偏倚，例如：①试验过程中，经过编盲后，试验操作者和结果判读者应不能区分来自既往样本集的样本；②既往样本集样本其储存、处理等应符合要求；③既往样本集样本的纳入可能导致具有罕见状况的受试者在临床试验受试人群中的比例显著高于在自然状态下目标人群总体中的比例，这个潜在偏倚应在统计分析中加以考虑。

如果既往样本纳入较多，需格外注意充分论证可能的选择偏倚等问题，例如：①受试人群是否能够代表目标人群的各种特征（而不仅仅是最典型的特征）；②样本是否来自足够大的样本集从而在一定程度上实现抽样的随机性要求；样本集中的样本是否

存在非随机的样本剔除；③样本是否具有充分的受试者临床信息；④定量检测的临床试验中，样本是否能够覆盖整个线性范围；⑤样本是否为产品适用的样本类型，且保存条件和时间满足被测物稳定性要求等。

### （三）临床试验机构数量和要求

#### 1. 临床试验机构数量

对于需要进行临床试验的体外诊断试剂，第二类产品应选择不少于2家（含2家）符合要求的临床试验机构、第三类或新研制产品应选择不少于3家（含3家）符合要求的临床试验机构开展临床试验。需进行变更注册临床试验的，一般可选择不少于2家（含2家）符合要求的临床试验机构开展临床试验。

#### 2. 临床试验机构要求

申办者应根据产品特点及其预期用途，综合不同地区人群差异、流行病学背景、病原微生物的特性等因素选择具有代表性的机构开展临床试验，包括受试人群的代表性、临床条件（预期使用环境和使用者的代表性）的代表性等。

#### 3. 中心效应

整个临床试验过程应严格按照伦理委员会同意的临床试验方案、有关标准操作规程等文件开展，其中试验操作、结果判读等应与体外诊断试剂产品说明书中的相关规定一致；临床试验开始前，申办者应当负责组织与该临床试验相关的培训，包括试验体外诊断试剂的储存、操作、管理等，以确保临床试验操作的一

致性，最大限度地控制试验误差；试验实施过程中要有必要的监查及质控措施；各临床试验机构之间试验样本量应尽量均衡；应当考虑各家临床机构开展临床试验的时间不同可能带来的影响，尽量同期开展临床试验。

应该注意，尽管使用了同一研究方案，各临床试验机构的临床数据进行合并统计时，仍有可能出现中心效应，即各中心试验结果出现显著差异。中心效应可能反映了各中心之间受试人群、临床试验操作等方面的差异，中心效应的产生可能影响研究结果的解释。临床试验方案设计时应考虑如何避免由于中心效应所带来的潜在偏倚。

#### （四）临床评价指标

临床评价指标应在临床试验的设计阶段确定，并在临床试验方案中予以明确。

定性检测的体外诊断试剂临床试验评价指标通常包括诊断准确性（灵敏度、特异度、预测值、似然比、ROC 曲线下面积等）或检测一致性（阳性/阴性符合率、总符合率、Kappa 值等）。

半定量检测的体外诊断试剂临床试验评价指标通常包括：各等级符合率、阴/阳性符合率及 Kappa 值等。

定量检测的体外诊断试剂临床试验评价指标通常包括回归分析的回归系数、截距、相关系数和决定系数等。

#### （五）临床试验的统计学分析

临床试验结果的统计分析应建立在正确、完整的数据基础上，选择适当的临床评价指标来评价体外诊断试剂的临床性能，并采用适当的统计模型对数据进行分析。

### 1. 统计学分析的基本考虑

体外诊断试剂的统计分析一般包括评价指标的参数估计（含置信区间估计）和假设检验。参数估计是在保证评价指标满足期望精度水平（置信区间的宽度一定）的前提下，确认灵敏度、特异度、（回归方程的）回归系数和截距等评价指标的水平。假设检验则需对统计学指标提出无效假设及备择假设，通过假设检验进行相关统计学推断。如有必要，需在统计分析之前考虑对样本的分布假设进行验证，从而合理选择统计方法。

对于有确定的临床可接受标准的研究，应在临床试验方案中明确接受标准，并经过临床试验数据的统计学分析证明临床评价指标的评价结果（区间估计）满足可接受标准要求。临床可接受标准应为行业广泛认可的结果，一般依据相关检测试剂的性能水平、风险判定和临床需求等因素进行设定。

#### 1.1 定性检测的统计学分析

定性检测临床试验一般以  $2 \times 2$  表的形式总结两种分析方法的检测结果，并据此计算灵敏度（阳性符合率）、特异度（阴性符合率）、总符合率、Kappa 值等指标及其置信区间。

除此之外，还可同时进行假设检验评价两种分析方法的一致性。



## 1.2 半定量检测的统计学分析

半定量检测的体外诊断试剂通常是指：检测结果报告为几个等级值（例如：阴性、+、2+、3+）、或者报告为终点稀释度的试剂等。临床试验可采用  $R \times C$  表形式总结检测结果，并据此计算各等级的符合率、阴/阳性符合率及 **Kappa** 值等指标及其置信区间。

## 1.3 定量检测的统计学分析

1.3.1 应根据临床试验数据绘制散点图，并进行相关性分析。

1.3.2 采用 **Bland-Altman** 法，计算一致性限度，评价两种检测结果的一致性。一致性限度应在临床所能接受的界值范围内。

1.3.3 采用回归分析对两种检测方法的一致性进行评价。应根据数据分布特点等因素选择适用的回归分析方法，如 **Passing-Bablok** 回归、**Deming** 回归和最小二乘回归等。回归分析应重点观察回归方程的回归系数和截距等指标，计算回归系数和截距的置信区间。亦可同时对相关评价指标进行假设检验。

1.3.4 应特别针对医学决定水平附近的检测结果进行分析。

## 1.4 ROC 分析

对于试验体外诊断试剂检测结果为定量或半定量数据，临床参考标准判断结果为定性结果的统计学分析，也可采用受试者工作特征（**ROC**）曲线的方法，以 **ROC** 曲线下面积反映试验体外诊断试剂检测的诊断价值，或同时比较两种试剂的诊断价值。对于体外诊断试剂的临床试验，采用 **ROC** 分析方式进行数据统计

时仍应进一步以推荐的阳性判断值进行灵敏度、特异度等指标（及其置信区间）的评价。

## 2. 定性检测的不一致样本分析

在定性检测试剂临床试验中，如有试验体外诊断试剂与对比方法检测结果不一致的情况，应对不一致结果进行综合分析，说明是否影响对产品临床性能的判定；对检测结果不一致的样本可采用临床参考标准或其他恰当的方法进行分析，但该分析结果不应纳入原有统计分析。

3. 如有必要，应对获得的数据集进行分层、分段统计。

4. 纳入临床试验的样本不应随意剔除，应在临床试验方案中设定样本剔除标准；如有任何剔除，应在临床试验小结和报告中详细列出所有样本剔除的情形，并说明理由。

5. 在体外诊断试剂临床试验的全过程中，生物统计学分析的结论非常重要，同时也要充分考虑到临床诊疗对于体外诊断试剂临床性能的要求；在产品安全有效性的评价中，应综合考虑评价结果的统计学意义及其所代表的临床意义。

## （六）样本量要求

适当的样本量是保证体外诊断试剂临床性能得到准确评价的必要条件。体外诊断试剂临床试验的样本量与多种因素相关，包括评价指标、检测的可重复性、干扰因素、亚组间的差异性、以及被测物特点等。临床试验方案中应对临床试验需要的最低样本量进行估算，并说明依据。

## 1. 关于样本量要求的基本考虑

1.1 临床试验样本量应满足统计学要求，应采用适当的统计学方法进行估算。

1.2 临床试验样本量的确定应考虑临床性能的各种影响因素，保证对临床性能的充分评价，如：受试人群应能够代表目标人群的各种特征，考虑到不同亚组中检测性能的评价需要，以及多种被测物（或多种亚型等）检测性能评价的要求，应在估算最低总样本量的基础上，保证各种组别/类型样本的例数满足要求。当体外诊断试剂临床性能预期在不同亚组的人群中有差异，且对某些重要亚组的临床性能需得到准确评价时，应对亚组样本量单独进行统计学估算。

1.3 如果试验体外诊断试剂适用于不同的样本类型，则还需考虑不同样本类型的样本量要求。

采用不同样本类型检测相同被测物时，可能因其样本基质差异、被测物来源差异、被测物浓度水平差异、干扰因素差异以及采样部位差异等因素导致产品分析性能、临床性能以及适用人群、适应证、参考区间等方面的差异。针对不同样本类型应考虑上述差异造成的影响，确定合理的样本量要求。

如不同样本类型在临床性能、适用人群、适应证、参考区间等方面存在显著差异，需针对不同的样本类型分别进行临床试验设计，包括分别进行样本量估算和统计学分析。

如通过临床前研究证明不同样本类型之间存在分析性能差

异，但不对临床性能造成显著影响，例如某些免疫学检测中的血清与全血样本，则临床试验中应以一种样本类型为主进行临床试验，样本例数满足统计学要求；其他样本类型适当纳入一定数量阳性和阴性样本，通过同源样本对比试验或与对比方法的比较研究进行临床性能的评价。

如不同样本类型在样本基质、被测物来源、被测物浓度水平、干扰因素以及采样部位等方面几乎没有差异，且经临床前研究证明分析性能没有差异，例如某些检测中的血清和血浆样本，则临床试验中不同样本类型可进行汇总统计。

1.4 如果临床试验包含不同的临床试验目的，则需分别进行临床试验设计，包括分别进行相应的样本量估算。例如：临床试验目的包括：①评价试验体外诊断试剂与对比试剂的检测一致性研究；②评价试验体外诊断试剂用于疾病鉴别诊断的灵敏度和特异性，则针对上述两个临床试验目的，应分别进行临床试验设计，并估算最低样本量。

1.5 对于定量检测试剂，在线性范围内的各个浓度水平均应有一定量的样本例数，并着重考虑对医学决定水平的检测性能进行充分验证；对于定性检测试剂，临床试验样本则应包含一定数量的阳性判断值附近样本（如涉及）。

## 2. 采用统计学方法进行样本量估算

采用统计学公式进行样本量估算的相关要素一般包括评价指标的预期值、评价指标的可接受标准（如适用）、I类和II类错

误概率以及预期的受试者脱落剔除率等。

应在方案中明确用于样本量估算的评价指标。评价指标的预期值根据已有临床数据（基于目标人群样本）、小样本探索性试验（如有）的结果或其他适当的评价数据来设定，应在临床试验方案中明确这些参数的确定依据。

一般情况下，I类错误概率 $\alpha$ 设定为双侧 0.05 或单侧 0.025，II类错误概率 $\beta$ 设定为不大于 0.2。

本文附件中举例说明几种常见的样本量估算方法，供参考。

3. 应注意，样本量估算仅仅是基于统计学要求的最低样本量估计，临床试验样本量应以临床性能得到充分评价为目标，保证目标人群的各种特征均有充分数量的代表性受试者入组，从而使临床试验结果全面、真实地反映目标人群的情况。

### （七）偏倚的控制

偏倚是指在临床试验方案设计、实施及结果分析时，有关影响因素所致的系统误差，致使对试验体外诊断试剂安全有效性的评价偏离真值。偏倚干扰临床试验得出正确的结论，在临床试验的全过程中均需防范其发生。

1.盲法：体外诊断试剂的比较研究试验中应对试验体外诊断试剂和对比方法（如有）的试验操作者和结果评价者设盲，使其在试验过程中不知晓受试者的疾病诊断或其他相关检测结果等信息，从而避免引入偏倚。

2.试验体外诊断试剂检测应与临床参考标准的判断或对比

方法检测尽量同步进行，避免因疾病进程不同或样本保存时间差异较大而造成临床试验结论偏离真值。

3.在临床试验机构选择、受试者选择、试验过程、统计学分析等各个阶段均需进行偏倚的控制。例如：受试人群应尽可能代表目标人群的特征，以避免选择偏倚；不同临床试验机构在临床试验中应统一试验操作和判读标准等，以避免中心效应的发生。

#### **（八）临床试验设计中还需考虑的其他因素**

体外诊断试剂的一些固有特征可能影响临床试验设计中的关键要素，在进行临床试验设计时，应加以考虑，包括体外诊断试剂的检验原理、使用方法（包括对技能水平的要求等）、使用条件（使用者和使用环境）、使用中人为因素的影响等。

例如：

1.预期供消费者个人自行使用的体外诊断试剂，临床试验中，除需评价试验体外诊断试剂临床性能以外，还需评价无医学背景的消费者对产品说明书的认知能力，并证明无医学背景消费者与专业检验人员检测结果的一致性。

2. 如果体外诊断试剂的检测操作和数据处理、解读等对使用者有专门的技能要求，则临床试验中对试验人员的技能水平、培训要求等应能够代表该产品上市后对预期使用者的要求。

#### **四、临床试验质量管理**

体外诊断试剂临床试验应符合《医疗器械临床试验质量管理规范》的相关要求，维护临床试验过程中受试者权益和安全，保

证临床试验过程规范，结果真实、准确、完整和可追溯。临床试验质量管理应涵盖临床试验的全过程，包括临床试验的方案设计、实施、监查、稽查、检查，数据的采集、记录、保存、分析、总结和报告等。以下针对体外诊断试剂临床试验质量管理中需要特别关注的问题进行说明：

### （一）临床试验前管理

1. 临床试验前，应确保产品设计已定型，完成试验体外诊断试剂的分析性能评估、阳性判断值或参考区间研究、质量检验以及风险受益分析等，且结果应能够支持该项临床试验。

2. 临床试验中使用的试验体外诊断试剂按照医疗器械生产质量管理规范的相关要求生产且质量合格。

### （二）临床试验开展

各临床试验机构原则上应当尽量同期开展临床试验，如在时间阶段上有较大的差异，应有合理的解释，确认采用了同一临床试验方案，并进行偏倚和中心效应分析。

### （三）数据与记录

1. 在临床试验中，主要研究者应当确保将任何观察与发现均正确完整地予以记录。临床试验的源数据至少应当包括：

1.1 所使用的试剂和仪器的信息，包括名称、规格/型号、批号/序列号、数量、接收日期、储存条件、使用情况及剩余试剂的处理等。

1.2 受试者筛选入选记录、受试者基本信息（如性别、年龄、

入组时间等)、临床诊疗信息、样本检验记录以及不良事件等。

1.3 临床试验用样本来源、编号、保存、使用、留存、销毁等各环节的完整记录。

1.4 记录者的签名及日期。

2. 临床试验源数据不得随意更改;确需作更改时应当说明理由,签名并注明日期。

3. 样本管理及溯源:临床试验样本应由开展试验的临床试验机构提供,应具有唯一的可溯源编号,每一份样本应可溯源至唯一受试者(如有特殊情况应在方案和报告中说明)。

4. 临床试验数据应具有可追溯性,临床试验报告、病例报告表、临床试验数据表等文件中的数据均应一致且可以追溯至源文件。

#### (四) 临床试验所需试剂管理

临床试验中试验体外诊断试剂、对比试剂及其配套使用的其他试剂(例如:核酸提取试剂等)和仪器、设备等的运输、使用、储存等,均应符合相关要求。

### 五、其他

1. 对于特定的体外诊断试剂,具体的临床试验方法、统计方法、样本量估算等可能有特定的要求,申办者应根据具体情况进行科学的选择和设计。如有相关产品的技术审查指导原则,应参考其中的相关要求。

2. 部分临床试验采用试验体外诊断试剂与核酸序列测定、



GC-MS/MS 等实验室检测参考方法进行比较研究，而这些方法并非临床常规检测技术，需要专门的设备仪器和试验条件，且大部分临床试验机构不具备相关检测条件。对于此类情况，申请人应尽可能选择具备相应条件的临床试验机构开展临床试验，确无检测条件的部分临床试验机构可将此部分测试委托给专门的测序机构、或具备一定检测资质的实验室进行检测，并对检测结果进行认可。产品注册申报时应提交临床试验机构与受委托机构的委托证明文件、对比方法的方法学研究资料和质控数据等。但不得委托申办者的实验室进行相关检测。由申办者直接委托的检测结果不得作为临床试验数据提交。

附件：样本量估算方法举例

## 附件

# 样本量估算方法举例

申办者应结合产品的具体特点、统计学分析模型等因素选择适当的样本量估算方法，并在方案中明确样本量确定的依据。同时应充分考虑可能的脱落剔除率等情况，合理设定样本量要求。

这里举例说明几种样本量估算方法。

### 一、定性检测的样本量估算举例

1.评价指标有确定的临床可接受标准时，需证明产品评价指标满足可接受标准要求。此时可采用单组目标值法样本量公式估算最低样本量。

$$n = \frac{\left[ Z_{1-\alpha/2} \sqrt{P_0(1-P_0)} + Z_{1-\beta} \sqrt{P_T(1-P_T)} \right]^2}{(P_T - P_0)^2}$$

公式中， $n$  为样本量； $Z_{1-\alpha/2}$ 、 $Z_{1-\beta}$  为显著性水平和把握度的标准正态分布的分数位， $P_0$  为评价指标的临床可接受标准， $P_T$  为试验体外诊断试剂评价指标预期值。

一般的，体外诊断试剂临床试验中，与已上市同类产品的对比试验均可根据临床需要设定适当的临床可接受标准，并采用上述公式进行最低样本量估算。

例：定性检测试剂临床试验，采用试验体外诊断试剂与已上市同类产品进行比较研究的方法，根据临床需求，阳性、阴性符

合率应分别达到 85%和 90%，根据探索性试验结果，试验体外诊断试剂与对比试剂阳性、阴性符合率预期分别可达到 90%和 94%。

阳性组 (n+) 和阴性组 (n-) 最低样本量估计分别为：

$$n_{+} = \frac{[1.96\sqrt{0.85(1-0.85)} + 0.84\sqrt{0.90(1-0.90)}]^2}{(0.90-0.85)^2} = 362$$
$$n_{-} = \frac{[1.96\sqrt{0.90(1-0.90)} + 0.84\sqrt{0.94(1-0.94)}]^2}{(0.94-0.90)^2} = 388$$

根据以上估算，总样本例数预计为 750 例。

按照脱落剔除率 10%，则应至少入组 834 例受试者，实际入组受试人群中，阳性组和阴性组样本例数应分别满足上述 n+和 n-的最低要求。

应注意，临床试验样本量除需满足上述统计学估算的最低样本量要求外，还应保证入组病例覆盖受试者的各种特征；如涉及不同样本类型，还需考虑不同样本类型的例数要求等。

2.对于临床试验的参数估计中只保证评价指标满足期望精度水平（置信区间的宽度一定），而不设定临床可接受标准的情况，可采用如下公式：

$$n = \frac{[Z_{1-\alpha/2}]^2 P(1-P)}{\Delta^2}$$

公式中 n 为样本量， $Z_{1-\alpha/2}$  为置信度标准正态分布的分位数，P 为评价指标预期值， $\Delta$  为 P 的允许误差大小。

应注意，P 和  $\Delta$  的取值应有充分依据，除非有特殊理由，否则不建议设置  $\Delta > 0.05$ ，当预期值更高时还应考虑更优的精度。

采用上述公式，可根据灵敏度或特异度的预期值分别估算具有目标疾病状态的受试者（阳性）或不具有目标疾病状态的受试者（阴性）的样本量。

例如：某项标志物检测试剂用于相关疾病的辅助诊断，通过对已有资料进行分析得知，该检测试剂的灵敏度预期为 85%，特异度预期为 90%，临床试验采用试验体外诊断试剂与临床参考标准进行比较研究的方法，评价试验体外诊断试剂的临床性能。允许误差 $\Delta$ 取值 0.05，则具有目标疾病状态的受试者（阳性）最低样本量（ $n_+$ ）估计为：

$$n_+ = \frac{1.96^2 \times 0.85 \times (1 - 0.85)}{0.05^2} = 196$$

不具有目标疾病状态的受试者（阴性）最低样本量（ $n_-$ ）估计为：

$$n_- = \frac{1.96^2 \times 0.9 \times (1 - 0.9)}{0.05^2} = 138$$

根据以上估算，总样本例数预计为 334 例。

按照脱落剔除率 10%，则应至少入组 371 例受试者，实际入组受试人群中，具有目标疾病状态的受试者（阳性）和不具有目标疾病状态的受试者（阴性）样本例数应分别满足上述  $n_+$  和  $n_-$  的最低要求。

该临床试验中，除需满足最低样本量要求外，具有目标疾病状态的受试者（阳性）还应确保覆盖疾病的各种类型、不同疾病进程、不同疾病严重程度的病例；不具有目标疾病状态的受试者（阴性）则需涵盖临床特异度评价所需的各种受试者样本等。如

果不同亚组人群预期灵敏度或特异度不同,还可能需要进行分层入组,并分别估算亚组样本量。

3. 需要注意的是,当评价指标 P 接近 100%时,上述两种样本量估算方法可能不适用,应考虑选择更加适宜的方法进行样本量估算和统计学分析,如精确概率法等。

## 二、定量检测的样本量估算

对于定量检测,亦可针对适当的评价指标,选择适宜的统计学方法,进行样本量估算。临床试验方案中建议同时给出与所选定评价指标对应的临床可接受标准,并提供确定依据。