

附件

免于临床试验的体外诊断试剂 临床评价技术指导原则

一、编制目的

对免于临床试验的体外诊断试剂，申请人应采用非临床试验的方式确认产品满足使用要求或者预期用途，包括将待评价试剂与境内已上市产品进行同品种方法学比对研究，证明待评价试剂与已上市产品实质等同；或采用待评价试剂与参考测量程序或诊断准确度标准进行比较研究的方式，考察待评价试剂与参考测量程序或诊断准确度标准的符合率/一致性。

本指导原则旨在为申请人对免于临床试验的体外诊断试剂临床评价提供技术指导，同时为药品监督管理部门对该部分资料的技术审评提供参考依据。

二、适用范围

本指导原则适用于列入免于临床试验体外诊断试剂目录（以下简称“目录”）产品的临床评价。

对免于临床试验的体外诊断试剂，申请人应提交申报产品与“目录”对应项目的对比资料，该资料应能证明申报产品与“目录”所述的产品具有等同性。

对免于临床试验的体外诊断试剂，申请人可依据本指导原则的要求进行临床评价，也可依据《体外诊断试剂临床试验技术指

导原则》的要求进行临床试验。

三、基本原则

对免于临床试验的体外诊断试剂，申请人可将待评价试剂与境内已上市产品进行比对，证明待评价试剂与已上市产品实质等同，或与参考测量程序/诊断准确度标准检测结果具有良好的一致性。此处所述实质性等同，指预期用途相同，且具有相同的安全性及有效性。

申请人应采用最终定型的试剂进行临床评价。进行临床评价之前，应确定产品的基本性能，通常包括适用的样本类型、特异性、精密度、检出限和/或定量限、测量区间、阳性判断值、参考区间等，以便为待评价试剂进行临床评价提供依据。

如通过临床评价无法证明待评价试剂与境内已上市产品实质等同，或与参考测量程序/诊断准确度标准检测结果具有良好的一致性，应通过临床试验的方式对申报试剂进行评价。

四、具体要求

（一）同品种方法学比对研究

1. 对比试剂的选择

免于临床试验的体外诊断试剂，如采用本指导原则所述方式进行方法学对比，应通过比对分析确定与待评价试剂相适宜的境内已上市产品作为对比试剂。

申请人应首先对待评价试剂与境内已上市产品的预期用途进行比对分析。预期用途是指体外诊断试剂的一般用途或功能，包括样本类型、被测物和适应证等。适应证是指体外诊断试剂诊

断、预防、预测、治疗监测或预后观察的疾病或状态，包括适用人群。如待评价试剂的预期用途在境内已上市产品的预期用途范畴内，则认为二者具有相同的预期用途。如待评价试剂与境内已上市产品的预期用途存在功能差异（如：分别用于辅助诊断和治疗监测）、适用人群差异（如：分别用于成人和儿童）等，则认为二者具有不同的预期用途。

预期用途相同时，申请人需继续对二者的基本原理、性能指标、阳性判断值、参考区间等进行比对分析。如存在差异，申请人还需进一步评估差异是否会对人体样本的检测性能产生明显差异。如二者间的差异不会对人体样本的检测性能产生明显影响，则认为具有可比性。如二者的基本原理（方法学）差异较大或性能指标差异较大等，应详述差异对人体样本的检测性能的影响。

综上，对比试剂需满足以下条件：

- （1）境内已经取得上市许可；
- （2）与待评价试剂具有相同的预期用途；
- （3）对于定量试剂，对比试剂与待评价试剂的检测结果应具有相同的计量单位或二者之间的计量单位可相互转化；
- （4）优先选择与待评价试剂检测结果偏差较小的试剂，不建议选择性能劣于待评价试剂的作为对比试剂。

2. 试验地点

试验过程由申请人进行管理并负责试验数据的真实性、合规性及完整性。境外申请人如需进行中国境内开展的方法学比对研究，应通过其在中国的代理人在中国境内开展试验。

申请人也可以提交在境外完成的试验资料，资料内容应满足本指导原则的技术要求。申请人还应充分考虑流行病学差异、受试人群的差异、使用的环境与条件差异等对研究结果产生的可能影响。必要时，申请人应针对差异因素在我国境内进行补充研究。

3. 试验人员

试验操作人员应为具有相应试验能力的专业技术人员，并熟悉待评价试剂和对比试剂的检测流程。

4. 试验样本

应采用与预期用途声称样本类型一致的人体样本进行试验，样本背景信息应清晰，样本来源应可溯源。样本背景信息包括但不限于：样本来源、唯一且可追溯的编号、样本类型和其他与该检测相关的背景信息（如年龄、性别、干扰因素等）；对于试剂检测结果有明确疾病指向的产品，其纳入的样本所来源的病例应有明确的临床诊断信息。

样本采集方式及稳定性应符合待评价试剂和对比试剂说明书有关要求。原则上，应采用临床真实样本进行研究。当遇真实样本浓度无法覆盖检测范围时，应充分阐述无法获得的依据，酌情采用从类似病史患者获取的混合样本，一般而言，混合样本不应超过总样本量的 20%。如特定的临床样本无法获得，可适当采用稀释或者经处理去除分析物获得低浓度样本、添加分析物获得高浓度样本等人工制备样本。人工制备样本应充分考虑样本的背景信息，基质效应等影响因素。

5. 试验时间

试验应考虑到临床实际使用情况和日间精密度影响,在保证符合样本稳定性的条件下,设定合理的持续时间,如3~20天,应对所有样本进行编盲,采用待评价试剂和对比试剂分别进行随机检测,整个试验应有质量控制。

6. 试验方法

试验方法可参考境内外有关方法学比对的技术指导文件,并重点关注以下内容。

6.1 定量产品

6.1.1 样本要求

应选择浓度覆盖线性/测量区间的预期适用人群样本和干扰样本进行研究。研究应纳入一定数量样本,一般不少于100例,并注意医学决定水平和测量区间内的各个浓度水平均应包含一定数量的样本。

试验前应预先设定可接受偏倚的限值,如比对研究试验结果无法达到预设标准,可适当扩大样本量进行评价。

如果待评价试剂与比对方法对不同人群亚组表现出不同的性能,则应对不同的人群亚组进行分层统计,每个人群应分别纳入至少100例样本。

如果待评价试剂对不同的人群具有不同的参考区间,则应对不同的人群进行分层统计,每个人群应分别纳入至少100例样本。注意:此处不包含已知的生理变化(如女性生理周期、性别、年龄等),产生的不同参考区间的情形。

申请人还应根据产品的预期用途,适用人群、临床适应证、

不同亚群基因型分配、临床使用中的可接受标准等综合情况充分考虑各种影响因素，采用合理的方法确定样本数量。

6.1.2 数据剔除

在数据收集过程中，应对所有数据及时记录并检查。如确定某些异常结果由可解释、可接受的原因引起，应记录原因并将其剔除出数据分析。如不能确定原因，须将原始结果保留在数据集中。

6.1.3 统计分析

申请人应综合考虑数据特征等因素，确定具体的统计分析方法，并提供其选择依据。定量产品的统计分析通常采用如下顺序进行。

6.1.3.1 数据作图与审查

完成数据收集后，应先绘制散点图和差异图，对数据进行分析 and 审查，观察数据是否覆盖了线性/测量区间以及是否存在离群值，并初步了解待评价试剂和对比试剂测量值之间的潜在变异特征，确定如何更好地表征这些差异。

散点图和差异图可显示待评价试剂和对比试剂的比较结果。其中，散点图应显示所有数据，其 **x** 轴表示对比试剂的测量结果，**y** 轴表示待评价试剂的测量结果，**x** 轴和 **y** 轴应使用相同的数值范围和间距。差异图的 **x** 轴表示测量浓度（如：待评价试剂和对比试剂的测量平均值），**y** 轴表示待评价试剂和对比试剂测量值的差值（如：二者的绝对差值或百分比差值）。申请人可根据具体情况选择特定类型的差异图，具体可参考国内外相关指南文件。

6.1.3.2 计算相关系数或决定系数

6.1.3.3 回归分析

应根据散点图和差异图判断数据是否满足相应的假设前提，并据此确定最佳的回归分析方法。常见的回归分析方法包括 Deming 回归、Passing-Bablok 回归分析和最小二乘回归等。

回归分析一般包括：获得线性回归方程 $\hat{y} = a + bx$ ， b 是回归系数， a 是截距， x 代表对比试剂的检测结果， y 代表待评价试剂的检测结果，并计算回归系数及截距的 95% 可信区间。

亦可进行假设检验评价二者偏倚。一般对回归方程中的截距 a 和回归系数 b 进行假设检验。

6.1.3.4 评价待评价试剂与对比试剂的偏倚

应通过区间估计评价待评价试剂与对比试剂的偏倚。一般将医学决定水平处的预期偏倚及其 95% 可信区间与申请人声称的可接受偏倚的限值进行比较。可接受偏倚的限值由申请人咨询临床机构后根据临床需求设定，或者参考相关的国内外标准等设定。如果预期偏倚的 95% 可信区间未超出申请人声称的可接受偏倚的限值，说明待评价试剂与对比试剂的检测结果符合预期偏倚标准。如果预期偏倚的 95% 可信区间超出申请人声称的可接受偏倚的限值，但包含可接受偏倚的限值，此时申请人应继续通过扩大样本量和/或额外对样本进行分析等方式，评价待评价试剂与对比试剂的检测结果的偏倚是否可接受。

6.1.3.5 定量产品统计分析中应注意的问题

(1) 如出现离群值，应对离群值产生原因进行分析，并进

行两次统计分析（包括/不包括离群值），如两次统计分析的结果不一致，应进行合理分析。

（2）相关系数仅能表示待评价试剂与对比试剂线性相关关系的密切程度，而非一致性，因此，不能仅采用相关系数评价二者的一致性。

（3）应在散点图上显示两条直线，一条为 $y=x$ 的直线，另一条为拟合回归直线。

6.2 定性产品

6.2.1 样本要求

样本所来源的病例应能代表预期适用人群。研究应纳入一定数量的阳性样本和阴性样本，一般各不少于 50 例，并注意包含一定数量的阳性判断值附近的样本和干扰样本。研究样本数量将影响可信区间宽度，扩大样本量，将获得更窄的可信区间。

申请人应根据产品的预期用途、适用人群、临床适应证、不同亚群基因型分配、临床可接受标准等综合情况充分考虑各种影响因素，采用合理的方法确定样本数量。

6.2.2 数据收集期间检查

在数据收集过程中，应对所有数据及时记录并检查。如确定某些异常结果由可解释的原因引起，应记录原因并将其剔除出数据分析。如不能确定原因，须将原始结果保留在数据集中。

6.2.3 统计分析

申请人应综合考虑数据特征等因素，确定具体的统计分析方法，并提供其选择依据。评价待评价试剂与对比试剂的一致性。

以 $R \times C$ 表的形式总结待评价试剂和对比试剂的检测结果，

并据此计算二者的阳性符合率、阴性符合率和总符合率。同时，计算阳性符合率、阴性符合率和总符合率的双侧 95%可信区间。

以下表 1 给出建议采用的 2×2 表样例：

表 1 待评价试剂和对比试剂的 2×2 表

	对比试剂		总计
	阳性	阴性	
待评价 阳性	a	b	a+b
试剂 阴性	c	d	c+d
总 计	a+c	b+d	a+b+c+d

阳性符合率=100%× a/（a+c）

阴性符合率=100% ×d/（d+b）

总符合率=100% ×（a+d）/（a+b+c+d）

还可同时进行假设检验评价二者一致性， 如进行 Kappa 检验，并计算 Kappa 值的双侧 95%可信区间等。

6.2.4 不一致结果的处理

待评价试剂与对比试剂的不一致可能源于待评价试剂或对比试剂的误差，不一致结果可通过“诊断准确度标准”或其他合理方法进行确认并分析原因，确认结果不应纳入统计分析。

6.3 半定量产品

本指导原则所述半定量产品一般指结果报告为多个等级（如：阴性、+、++或+++）的试剂。该类产品的样本要求、试验时间和顺序等，可参考定量或定性产品的相关要求，具体视产品特点而定。

半定量产品所采用的统计分析方法也应视情况而定，一般可采用如下方法评价待评价试剂与对比试剂的一致性。首先，采用 $R \times C$ 表示待评价试剂与对比试剂的检测结果。在 $R \times C$ 表中，待评价试剂与对比试剂的等级数为多个等级，且二者的等级数相等，如“阴性，+，++”等。然后，计算二者在各等级的单项符合率、总符合率及 95%可信区间，还可同时进行假设检验，如 Kappa 检验，并计算 Kappa 值的双侧 95%可信区间等。

6.4 不同样本类型

如果待评价试剂包括几种样本类型，应进行样本适用性研究。对于研究中显示可比的样本（如血清和血浆），可仅选择一种样本类型进行方法学比对研究。对于研究中显示不可比的样本（如：血液和尿液样本），应根据本指导原则，各选择不少于 100 例样本，对每种样本类型分别进行方法学比对研究。

此处所述可比样本，一般指：性能指标相同、阳性判断值相同、预期人群一致、临床意义相同。反之，则应视为不可比样本。

6.5 统计分析中应注意的问题

6.5.1 样本剔除应有合理理由，不可随意剔除。如确需剔除，应说明剔除数量及理由。

6.5.2 应参考权威技术指南等，根据产品适用的统计分析模型，计算斜率、截距、医学决定水平的预期偏倚、阳性符合率、阴性符合率、总符合率及其 95%可信区间等。

6.5.3 如比对研究采用稀释或者经处理去除分析物获得低浓度样本、添加分析物获得高浓度样本等人工制备样本时，应进行

两次统计分析（包括/不包括人工制备样本），并对统计结果进行合理分析。

（二）与参考测量程序或诊断准确度标准进行比较研究

如有参考测量程序或诊断准确度标准，也可参照本指导原则“（一）同品种方法学比对研究”章节相关要求，采用待评价试剂与参考测量程序或诊断准确度标准进行比对研究。参考测量程序指被接受作为提供适合其预期用途的测量结果的测量程序。诊断准确度标准指使用一种方法或联合多种方法，包括实验室检测、影像学检测、病理和随访信息在内的临床信息，来界定状况、事件和关注特征有无的标准。

（三）变更事项相关的方法学比对研究

对于变更事项相关的方法学比对研究，一般采用变更后产品与变更前产品进行比对研究。如变更后产品与变更前产品不具有可比性，应采用变更后产品与其他已上市产品进行方法学比对研究，具体要求可参考本指导原则。

五、临床评价报告要求

报告应至少包括如下内容：待评价试剂与对比试剂/参考测量程序或诊断准确度标准的描述性比对分析以及比对性能数据。报告应由试验人员签字，并由申请人和/或代理人签章。

（一）描述性比对分析

如采用同品种方法学比对研究，申请人应提交待评价试剂与对比试剂的详细比对分析资料，应详述二者的异同，并提交异同来源的支持资料。如二者存在差异，申请人还应提交差异是否会

对人体样本的检测性能产生明显影响的评估资料。比对分析项目包括但不限于附件所列项目，若存在不适用项目，应说明不适用的理由。

如采用参考测量程序或诊断准确度标准作为对比方法，申请人应详细说明选择参考测量程序或诊断准确度标准作为对比方法的理由，并提交所选参考测量程序或诊断准确度标准的详细资料，包括操作方法、判定标准等。

（二）比对性能数据

比对性能数据应对试验设计、试验实施和统计分析等情况进行清晰的描述，至少包括如下内容。

1.基本信息。包括但不限于：申请人名称、试验人员、试验时间及地点、待评价试剂名称、对比试剂/参考测量程序或诊断准确度标准、配套使用的其他试剂/仪器的产品名称、生产企业、规格/型号和批号等。

2.试验设计。详细说明样本入组、排除和剔除标准、样本量、统计分析方法的选择等内容。

3.试验实施情况，具体包括：

3.1 样本选择情况，包括例数和样本分布等。

3.2 试验过程描述。

4.试验管理，包括参加人员、质量控制情况、数据管理、出现的问题及处理措施等。

5.统计分析及评价结果总结。根据确定的统计方法对检测结果进行统计分析，对产品的性能进行合理评价。应详细说明申请

人声称的可接受偏倚的限值或可接受一致性的限值的确定依据，并详细说明统计分析中各计算公式及其来源、以及选择的理由。如涉及统计软件，应明确所采用的统计软件及软件版本号。

6.数据汇总表

应以表格形式对入组样本进行汇总描述，至少包括以下内容：可溯源样本编号、样本类型、待评价试剂和对比试剂或参考测量程序/诊断准确度标准的检测结果、样本背景信息（如适用）等。数据汇总表应由操作者签字，并可溯源。申请人应保存所有原始试验数据备查。

六、其他资料

除以上评价报告外，有关待评价试剂对人体样本检测性能评价的相关文献，可作为补充资料提交。

附件：待评价试剂与对比试剂的比对分析项目

附件

待评价试剂与对比试剂的比对分析项目

比对项目	对比试剂	待评价试剂	异同
1.预期用途			
(1) 适用人群			
(2) 样本类型			
(3) 被测物			
(4) 适用的疾病/状态			
(5) 使用环境			
2.基本原理			
3.主要组成成分			
4.阳性判断值或参考区间			
5.检验方法			
6.性能要求：如			
(1) 检出限和/或定量限			
(2) 线性/测量区间			
(3) 精密度			
(4) 干扰和交叉			
(5) HOOK 效应（如有）			
7.各种对照/质控			
8.结果报告形式			

注：比对项目包括但不限于上述项目，可根据实际情况增加。