

9101 分析方法验证指导原则

分析方法验证 (Analytical procedure validation) 是证明采用的分析方法适合于其预期的目的。分析方法验证是分析方法生命周期的一部分, 在建立兽药质量标准、变更兽药生产工艺或配方、修订原分析方法、或采用法定分析方法作为新开发兽药的分析方法时, 需对分析方法需经进行验证; 本指导原则讨论符合兽药典要求的分析方法验证中需要考虑的要素, 提供关于如何实施和评价分析方法各种验证试验的指导和建议, 适用于化学、生物/生物技术的原辅料药物及其制剂新建或修订的分析方法, 也可适用于基于风险控制策略的其他分析方法。生化药物质量控制中采用的方法包括理化分析方法和生物学测定方法, 相对于理化分析方法而言, 生物学测定方法存在更多的影响因素, 在依据本指导原则进行具体验证时, 还需结合生物制品的特点考虑。本指导原则尽可能与ICH Q2分析方法验证指导原则保持一致。

分析方法验证研究

需验证的待测量的质量属性有: 鉴别试验、杂质(纯度)或其他定量测量(包括限度控制或定量测量)、含量/效价或其他定量测量(如药品溶出度、释放度和物理常数测定)等。

需验证指标的性能特征有专属性/选择性、准确度、精密度和范围。由于待测量的质量属性具有各自的特点, 随分析技术和分析对象的复杂与否而有不同的要求, 应根据分析方法的预期用途和所采用的具体技术, 选择一组适当的性能特征及其标准进行验证。推荐使用预定义的性能特征及其标准来证明分析方法对其预期用途的适用性。表1中列出待测量的质量属性和相应的性能特征供验证时选择。

表1 待测量的质量属性的典型性能特征和相应的验证试验

待测量的质量属性 分析方法性 能特征(2)	鉴别	杂质(纯度) 其他定量测量(1)		含量/效价 其他定量测量(1)
		定量测量	限度控制	
专属性(3) 专属性试验	+	+	+	+
范围 响应 (校正模型) 范围下限	-	+	-	+
	-	QL [†]	DL	-
准确度(4) 准确度试验	-	+	-	+
精密度(4) 重复性试验	-	+	-	+
中间精密度试验	-	+(5)	-	+(5)

-表示该性能特征通常不需被评估;

+表示该性能特征通常应被评估；

†表示通常不需评估检测限，但在某些特别或复杂的情况下是被推荐的。

QL,DL: 分别代表定量限，检测限

(1)其他定量测量的范围限如接近技术的检测限或定量限，可遵循杂质检测方案，否则建议采用含量分析方案。

(2)在某些分析方法用于物理化学性质情况下，某些性能特征可以用技术固有合理性来代替。

(3)某一分析方法不够专属，应用一种或多种其他辅助分析方法予以补充，除非有合理的证明。

(4)准确度和精密度可以分别评估，也可以使用联合验证的方式评估。

(5)精密度包括重复性、中间精密度和重现性。如已有重现性试验数据，由重现性数据集可得出中间精密度，不需另行中间精密度独立研究。

在分析方法验证中，如需要，应使用标准物质或其他经适当表征符合预期目的的物质进行试验，这些物质均应有鉴别、纯度或任何其他必要特性的证明文件。标准物质所需的纯度取决于预期的用途。对标准物质的定义、标定、管理和使用均应符合国家相关规定。

分析方法验证不是一个孤立的过程，是与分析方法开发过程相联系的不可分割的一个整体。分析方法开发研究中获得的合适数据可作为替代的验证数据；如果有科学依据，可简化已建立的分析方法的验证工作。在实践中，为适当地评价验证试验结果，可以设计验证实验，以提供分析方法性能的完整信息，如专属性/选择性、范围、准确度和精密度等。在实施验证研究前，可将耐用性作为分析方法开发的一部分进行评估。除本指导原则所述的验证方法外，在有适当科学依据的情况下，其他方法也可适用并被接受。分析方法验证应选择与其产品质量属性相对应的最适宜的验证方法和方案。

在实施验证研究之前，应制定验证方案。方案应包括分析方法的预期目的、需要验证的性能特征及其标准、验证实验设计等相关信息。如使用不属于验证方案范围和利用先验知识(例如来源于开发或前期验证)的其他研究数据，应该提供充分的合理性证明。

验证研究的结果应全面总结并形成验证报告。方法验证的理由、过程、数据和图表等，包括那些支持验证结果而未列入验证方案的试验，均应文件化呈现。

验证研究的实验设计应反映常规分析中使用的平行试验次数以产生可报告结果。如合理，可使用不同的平行次数实施一些验证试验，或者根据验证期间生成的数据调整分析方法中的平行试验次数。

1 分析方法生命周期内验证

在分析方法整个生命周期内，随着对产品认知的深入、技术水平的发展和监管要求的提高，分析方法可能需要变更。变更后的分析方法可能需要部分或全部重新验证。已给定的性能特征是否需要重新验证，需应用科学和基于风险的评估原则予以证明；重新验证的程度取决于分析性能特征受变更影响的程度。

如适用，通过使用来自多个实验室的数据共同验证，可证明分析方法符合预定义的性能标准，并满足分析方法在不同实验室转移的要求。在符合分析方法生命周期变化的背景下，应考虑经验证的分析方法转移。当分析方法在不同实验室间转移时，通常会执行验证实验的子集。

2 可报告范围

可报告范围通常来自质量标准，并取决于方法的预期用途。通过证明分析方法在应用于待测量的质量属性所规定的范围或在其限度内的样品时，提供的结果具有可接受的响应、准确度和精密度，来确认可报告范围。如适用，可报告范围应涵盖质量标准限度的上限和下限或报告限度。

表2举例说明了某些待测量的质量属性推荐的可报告范围。如合理，其他范围也可以接受，例如对于高纯化学原料药，可报告范围的上下限可以更窄。在某些情况下，例如含量较低时，更宽的上限可能更为实际。

表 2 某些待测量的质量属性推荐的可报告范围

待测量的质量属性	可报告范围下限	可报告范围上限
原料药和制剂含量	标示量80%或限度下限 80%	标示量120%或限度上限 120%
效价	限度下限-20%	限度上限+20%
含量均匀度	标示量70%	标示量130%
溶出试验 速释制剂 单点指标 多点指标 调释制剂	Q值-45% 可报告范围下限(根据质量标准来论证)或QL, 如适用 可报告范围下限(根据质量标准来论证)或QL, 如适用	最高规格标示量的130%
杂质检查(1)	报告阈值	限度的 120%
纯度检测(以面积%计)	限度下限的80%	限度上限或100%

(1) 当含量和杂质检测采用同一试验且仅使用一个标准时，线性验证应考察杂质报告水平至含量指标可接受标准的 120%。

在中药分析中，可报告范围应根据分析方法的具体应用和校正关系、准确度、精密度结果及要求进行确定。对于有毒的、具特殊功效或药理作用的成分，其验证范围应大于被限定含量的区间。

3 稳定性指示特性的证明

如果一个经过验证的定量分析方法，可以检测原料药、辅料、制剂或其他药品在存储过程中的相关质量属性的变化，则被认为是一个稳定性指示特性方法。为了证明稳定性指示特性方法的专属性/选择性，应研究含有相关降解产物的样品，包括使用加标分析物和含所有已知干扰物的样品，已暴露于各种物理和化学强制条件下的样品，或为已过期或在强制条件下存储的实际样品。

4 多变量分析方法的考虑

多变量分析方法又称多元分析方法，通过使用多个输入变量(例如，具有多个波长变量的光谱)的多变量校正模型来确定结果。多变量校正模型将输入数据与所关注的属性值(即模型输出)联系起来。

多变量分析方法的成功验证应考虑校正、内部测试和验证；通常分开发和验证两个阶段进行。

- **第一阶段** 模型开发包括校正和内部测试

校正数据用于创建校正模型,测试数据用于内部测试和模型优化。测试数据可以是一组单独的测试数据,也可以是校正数据集的一部分(例如,交叉验证方法)。内部测试步骤用于获得对模型性能的评估和微调算法的参数(如偏最小二乘法(PLS)潜在变量数),以在给定的数据集内选择合适的模型。

• 第二阶段 模型验证

使用独立样本组成的独立验证数据集对模型进行验证。对于鉴别库,验证包括分析库中未能代表的样本(即挑战样本),以证明库模型的判别能力。

每个用于验证定量或定性多变量方法的样品均需确定其量值或其类别,它们通常由经过验证的方法或药典中的参考方法获得。当使用参考分析方法时,其性能应达到或超过多变量分析方法的预期性能。为确保样品和测量的稳定性,应在合理的时间内尽可能用参考方法对同一样品进行分析和多变量数据收集。在某些情况下,可能需要相关性分析(correlation)或转换来提供相同的度量单位。任何假设或计算都应加以解释。

验证试验、方法学 and 评价

以下描述了评价分析方法性能的实验方法学,它们是根据分析方法设计所确定的主要性能特征进行分类的。然而,有关其他性能特征的信息可能来自同一数据集。如合理,可以使用不同方法来证明分析方法符合预期目的和相关的性能标准。

1 专属性/选择性

专属性(又称特异性)和选择性均用于描述在其它物质存在下分析方法测定某一物质不受干扰的程度。其他物质可能包括杂质、降解产物、有关物质、基质或操作环境中存在的其他组分。专属性方法是具有完全选择性的方法,通常用于描述最终状态,明确可以对目标分析物进行检测。选择性是一个相对术语,用于描述混合物或基质中特定被分析物可被检测且不受具有类似行为的其他组分干扰的程度。无论采用何种分析方法,应用于何种待测量的质量属性,均应考察其专属性/选择性。当分析方法不具专属性时,可以证明其选择性。然而,在存在潜在干扰的情况下,应使鉴别或定量测试的干扰最小化,并证明该测试符合预期目的。某些分析方法专属性的缺乏可以由其他辅助分析方法来补充,如果一种方法不能提供足够的区分能力,建议采用两种或两种以上的分析方法,以达到必要的专属性/选择性。

分析方法的专属性或选择性可采用以下方法之一或它们的组合来证明,有些实验可与准确度研究相结合。

1.1 无干扰

专属性/选择性可通过证明分析物的鉴别和/或定量不受其他共存物质如杂质、降解产物、有关物质、基质或存在于操作环境中的其他成分的干扰影响来验证。如采用阴性试样(除去含待测成分药材或不含待测成分的模拟处方试样)试验,取不含被测成分的阴性试样与含被测成分的供试品在同一条件试验并比较,以确认是否存在干扰。

1.2 与正交方法比较

专属性/选择性可通过将拟采用方法的测量结果与另一个良好表征的分析方法的测量结果相比较来验证,理想情况下,另一个良好表征的分析方法应为基于不同测量原理的方法(即正交方法)。

1.3 技术固有合理性

某些情况下,分析技术的专属性可通过技术参数如质谱中同位素的分辨率、核磁共振信号的化学位移等来确保和预测,如果被证明是合理的,则不需要实验研究。

1.4 数据要求

鉴别

对于鉴别试验,关键是证明其基于分子结构特征和/或其他特性,能区分样品中所关注的成分与其他化合物的能力,该能力可通过与已知标准物质比较,从含有被分析物的样品中得到的阳性结果和从不含被分析物样品中得到的阴性结果来证明。另外,鉴别试验可用于鉴别与分析物结构相似或密切相关的物质,以确认不致产生阳性结果。选择这些潜在干扰物质时,应基于科学判断并考虑到任何可能产生的干扰。

含量、纯度和杂质检测

应证明分析方法的专属性/选择性,以满足测定样品中分析物的含量或效价的准确度要求。

应使用代表性数据如色谱图、电泳图、光谱图、生物反应来证明专属性,如合适,图中的每个成分应适当加以标记。过高或过低地要求方法的专属性既不切实际也不科学。

对于分离技术,应在适当的水平研究合理的区分,例如,在色谱的关键分离中,专属性可用两个最接近的洗脱组分的分离度来证明。或者,可通过比较不同组分的光谱来评估可能的干扰。

对于非分离技术,例如生物测定、ELISA 法、qPCR 法等,专属性可通过使用标准物质或其他适当表征的物质以确认对分析物无干扰来证明。如分析物是与某种工艺相关的杂质,还必须确认供试品及其他成分无干扰来证明专属性。

如某一方法不具专属性或没有充分的选择性,应使用其他方法来确保足够的区分。例如,在用滴定法测定原料药含量时,可结合使用合适的杂质检查方法。

• 如杂质或有关物质可获得:

对于含量或效价测定,应证明分析物在杂质和/或辅料或其他成分存在时能被区分。实际上,可通过在原料药或制剂中加入适当水平的杂质和/或辅料,与未添加杂质或辅料的样品检测结果相比较,以证明分析物的检测结果不受共存物质的影响。或者,可通过设计的强制降解原料药或制剂样品制备含有适量杂质的样品。

对于纯度或杂质检测,可通过降解原料药或制剂,也可以在原料药或制剂中加入杂质以使杂质或有关物质达到适当的水平,并证明这些分析物在单独存在和/或与样品基质中的其他成分共存时均能被准确测量,以此来建立区分。

• 如杂质或有关物质不可获得:

如果杂质、有关物质或降解产物不能通过制备和分离获得,专属性可通过将含有典型杂质、有关物质或降解产物的样品的检测结果与另一种已被良好表征的方法(如药典方法或其他经过验证的正交分析方法)相比较来证明。所采用的方法应是合理的。

化学药含量测定应对两种方法的结果,杂质检查应对检出的杂质个数,必要时可采用光二极管阵列检测和/或质谱检测,进行峰纯度检查。

对于中药分析检测,专属性/选择性除可通过与另一方法的测定结果比较或用阴性试样试验来证明外,必要时,还应评价色谱相邻洗脱组分峰的分离度和峰纯度。

2 范围

分析方法的范围通常是指分析方法能达到适当水平的响应、精密度和准确度,具有良好校正关系的最高和最低结果的区间。范围可以使用适当的校正模型(如线性、非线性或多变量)通过对可报告结果(产生可报告范围)的直接评估来验证。校正模型范围如线性范围应覆盖准确度和精密度的验证范围。

在某些情况下,根据样品制备(如稀释)和所选择的分析方法,可使用一个或多个适当的工作范围来确定可报告范围。

通常,工作范围对应于呈现在分析仪器上的最低和最高样品浓度或纯度水平,在该范围内,分析方法可提供可靠的结果。通常需要数学计算来生成可报告的结果。可报告范围和工作范围可以相同。

如果无法获得足够纯(或含有相当量杂质)的物质来验证整个范围(例如,100%纯度),则可适当地外推可报告范围,并提供合理性证明。

2.1 响应

2.1.1 线性响应

分析物浓度和响应之间的线性关系应在分析方法的整个范围内进行评估,以确认分析方法对预期用途的适用性。可采用拟定的方法,如用对照品或直接用原料药制成的标准贮备液经精密稀释,或分别精密称取对照品或精密称取混合对照品,制备系列溶液来证明测量响应与浓度呈线性关系。

以响应信号作为分析物浓度或含量的函数作图评价线性关系,并应证明分析方法在给定的范围内具备获得与真实样品值(已知值或理论量)成比例的数值的能力。应采用适当的统计方法(例如,用最小二乘法计算回归曲线)对实验结果进行评价。

由回归曲线得到的数据有助于提供线性关系的数学估计。应提供数据图、相关系数或其平方值 (coefficient of determination, 判定系数)、y轴截距和回归曲线斜率。分析实测数据点与回归曲线的偏差可能有助于评价线性,例如,对于线性响应,应评估回归分析的残差图中任何非随机模式的影响。

为建立线性关系,建议至少设计5个浓度水平并适当地分布在范围内。

为获得线性关系,必要时可对测量响应数据使用如对数函数等进行数学转换。若采用其他方法评价线性,应证明评价线性关系的其他方法的合理性。

2.1.2 非线性响应

有些分析方法可能显示非线性响应。在这些情况下,有必要构建一个模型/函数来描述分析方法响应与活性/浓度之间的关系。应通过非线性回归分析(例如,判定系数)来评估模型的适用性。

例如,免疫分析或细胞分析可能显示S-型响应。当浓度范围足够宽,响应受到上、下渐近线的约束时,就会出现S-型试验曲线。在这种情况下使用的常见模型是四参数或五参数逻辑函数,不过也存在其他可接受的模型。对于这些分析方法,线性的评价与浓度-响应曲线形状的考虑是分开的。因此,浓度-响应的线性关系不是必需的;而应评价分析方法在给定范围内获得值与已知或理论的样品真值成比例关系的能力。

2.1.3 相对响应

色谱定量分析基于进入检测器中各组分的量与检测器的色谱响应成比例关系。同一色谱条件下,待测物质与参比物质可有不同的色谱响应,例如不同的紫外吸收系数。校正因子定义为单位质量参比物质(包括内标)的色谱响应与单位质量待测物的色谱响应的比值。校正因子法,通常以标准物质的色谱响应校正待测物质的色谱响应实现待测物质的定量分析法,常用于化学药中有关物质、中药材及其复方制剂中多指标成分的测定。在方法开发或方法验证期间,应确定使用适当的校正因子,并以文件化呈现。

校正因子用于有关物质检测时,通常以主成分为参比,也可以供试品中存在的已知有关物质或加入的另一成分为参比;当校正因子近似等于1(待测物质与标准物质的相对响应因子为0.8~1.2)或有关物质的量已被高估时,可不使用校正因子计算;否则,应使用校正因子计算。

2.1.4 多变量校正

用于构建多变量校正模型的算法可以是线性的,也可以是非线性的,只要模型适合于建立分析信号与待测物的质量属性之间的关系。多变量方法的准确度取决于多种因素,如校正样品在校正范围内的分布和参考方法的误差。

在多变量分析中,测量数据通常通过导数或归一化进行预处理。

除了对参考结果和预测结果进行比较外,线性评估还应包括方法误差(残差)在校正范围内如何变化的信息。残差分布图可用于评估整个工作范围内模型预测的残差。

2.2 范围下限的验证

如待测量的质量属性要求分析方法范围接近该方法的范围下限,使用以下方法估计检测限和定量限。检测限(Detection limit, DL或Limit of detection, LOD)系指试样中被测物能被检测出的最低量。药品的杂质检查方法,应通过测试来确定方法的检测限。检测限仅作为限度试验指标和定性鉴别的依据,没有定量意义。定量限(Quantitation limit, QL或Limit of quantitation, LOQ)系指试样中被测物能被定量测定的最低量,其确定结果应符合一定的准确度和精密度要求。对微量或痕量药物或药物杂质和降解产物进行定量测定时应确定方法的定量限。

2.2.1 基于直观评价

直观评价既可用于非仪器分析方法,也可用于仪器分析方法。

通过分析含已知浓度待测物的样品,能被可靠地分辨、检出的待测物的最小量即为检测限,能以可接受的准确度和精密度定量检测的待测物的最小量即为定量限。

2.2.2 信噪比法

这种方法适用于具有基线噪音的分析方法。通过将已知浓度样品的测量信号与空白样品的测量信号比较来确定信噪比,或者可使用适当基线区域内的信号代替空白样品的信号,以建立分析物能被可靠检测或定量的最低浓度。对于检测限,信噪比为3:1通常被认为是可以接受的;对于定量限,信噪比应不小于10:1。

信噪比应在一个预定义的区域内确定,如可能,应对称地分布于待测峰两侧。

2.2.3 基于线性响应的标准差和斜率

检测限(DL)可以表示为:

$$DL = \frac{3.3 \sigma}{S}$$

定量限 (QL)可以表示为:

$$QL = \frac{10 \sigma}{S}$$

斜率S可以从分析物的标准曲线中估算出来。标准偏差 σ 的估算可以采用多种方法, 例如:

(1) 根据空白的标准偏差

通过分析适当数量的空白样本的背景响应值的大小, 计算其标准偏差。

(2) 根据标准曲线

使用含有分析物的样品, 在DL和QL范围内评价特定的标准曲线。回归曲线的剩余标准差(即均方根误差/偏差)或回归曲线y轴截距的标准偏差可作为标准偏差。

2.2.4 基于范围下限的准确度和精密度

除使用上述方法估算值, 定量限可通过准确度和精密度测量直接验证。

2.2.5 数据要求

(1) 检测限

应报告检测限和用于确定检测限的方法。如果DL是基于视觉评价或基于信噪比确定的, 应呈现相关数据、图表和所用方法。在通过计算或外推获得DL估算值时, 该估算值随之可通过分析浓度接近或等于检测限水平的适当数量的样品来验证。

应注意仪器检测限与方法检测限的区别, 如所述的检测限是基于某种仪器的检测响应, 应关注不同仪器检测限的差别, 若将由一台仪器获得的仪器检测限作为方法检测限时应谨慎。

(2) 定量限

应报告定量限和用于确定定量限的方法。对于杂质和微量或痕量成分的检测, 分析方法的定量限应不得高于报告阈值。

不论采用何种方法得到QL估算值, 都应将其视为初始值, 随之应通过分析浓度接近或等于定量限水平的适当数量的样品来确认。如QL远远低于报告限度(例如, QL约低于报告限度的1/10以下时), 可合理地省略前述的确认验证。

3 准确度和精密度

准确度和精密度可以分别进行评价, 它们各有预定义的可接受标准。将这两个性能特征联合验证是评价分析方法适用性的另一种方法。

3.1 准确度

准确度系指用该建立方法测定量的结果与真实值或参考值接近的程度, 一般用回收率(%)表示。准确度应在规定的范围内测定。分析方法的报告范围应建立在, 在常规测试条件下(如存在样品基质和使用描述的样品制备步骤)得到证明。

通常, 准确度可通过下述的研究之一来确认。在某些情况下, 如精密度、范围内的响应和专属性已经确定, 可以推论方法的准确度。

3.1.1 与标准物质比较

用分析方法测定已知纯度的分析物如标准物质、良好表征的杂质或有关物质，将测定结果与理论预期结果比较进行评价。

3.1.2 加样回收试验

在不含待测成分的所有基质中添加已知量的待测成分，如无法获得模拟所有样品成分的基质试样，可将已知量的待测成分添加入或富集在待测试样中；分别测定已添加或富集待测成分的试样和未添加待测成分的试样中的待测成分量，将两者的测定结果进行比较来评价回收率。在加样回收试验中须注意添加的待测成分量与供试品中待测成分原含有量之和必须在校正模型范围之内；添加的量要适当，过小则引起较大的相对误差，过大则干扰成分相对减少，真实性差。

3.1.3 与正交方法比较

将拟采用的分析方法的结果与基于不同测量原理的另一良好表征的方法(正交方法)的结果进行比较。应报告另一方法的准确度。在无法获得所有药品成分的样品来模拟加样回收研究所需的基质时，正交方法可与定量杂质测量一起用于确认主要(primary)的测量值。

3.1.4 数据要求

在可报告范围内，使用适当数量的平行样品，在适当的浓度水平评价准确度，如设计至少3种不同浓度，每种浓度分别制备至少3份供试品进行测定，用至少9份测定结果进行评价，且浓度的设定应考虑样品的浓度范围。

准确度试验结果应报告为在试样中已知添加量分析物的平均回收率，或报告为平均值与可接受真值之间的差值，同时提供合理的 $100(1-\alpha)\%$ 置信区间(或其他合理的统计区间)。除另有合理的证明，所提供的置信区间应与对应的准确度可接受标准相匹配。对于杂质检测，应描述与主成分对应的单杂或总杂的测定方法(例如，重量/重量或与主成分面积百分比)。

对于多变量方法的定量应用，应使用合适的度量指标，如均方根预测误差(Root Mean-Square Error of Prediction, RMSEP)。如果RMSEP与可接受的均方根校正误差(Root Mean-Squared Error of Calibration, RMSEC)相当，则表明在使用独立的测试集进行测试时模型足够准确。对于分类等定性应用，可以使用误分类率或阳性预测率来表征方法的准确度。

3.2 精密度

精密度系指在规定的测定条件下，同一个均匀供试品，经多次取样测定所得结果之间的接近程度。

在相同条件下，由同一个分析人员测定所得结果的精密度称为重复性；在同一个实验室，不同时间由不同分析人员用不同设备测定结果之间的精密度，称为中间精密度；在不同实验室由不同分析人员测定结果之间的精密度，称为重现性。

含量测定、其他定量测定和杂质或纯度的定量测定应验证方法的精密度。

研究方法的精密度应使用均匀的、真实的样品，或在无法获得这样的样品时，可使用人工制备的样品(例如，在基质混合物中或在不含待测成分的样品中添加相应数量的分析物)。

3.2.1 重复性

在可报告范围内，取同一浓度(相当于100%浓度水平)的供试品，用至少平行6份的测定结果进行评价；或设计至少3个不同浓度水平，每个浓度水平分别平行制备至少3份供试品溶

液进行测定,用至少9个测定结果进行评价。评价重复性,浓度水平的选择和设计,应根据供试品中待测成分的含量或浓度可能的变化范围即可报告范围来确定,以保证重复性评价结果的可靠性。

3.2.2 中间精密度

中间精密度应达到的程度取决于所使用方法预期的目的,应确定随机事件对分析方法精密度的影响。应考察随机变动因素如不同日期、不同环境、不同分析人员、不同仪器对精密度的影响。理想情况下,选择的变动因素应基于并通过对分析方法开发和风险评估的理解予以证明,没有必要单独研究这些影响。鼓励使用实验设计研究中间精密度。

3.2.3 重现性

通过实验室间试验评估重现性。并不是每一次申报都要求提供重现性资料。但在分析方法标准化的情况下应该考虑重现性,例如,拟在国家药品质量标准中收载的或将在多个实验室使用的分析方法,应通过在不同实验室的协同检验获得重现性试验结果,提交重现性试验资料。协同检验的目的、过程和重现性结果均应记载并作为附件提交。应注意重现性试验用样品质量的一致性和贮存运输中的环境对该一致性的影响,以免影响重现性结果。

3.2.4 数据要求

所有的精密度试验都应报告标准偏差、相对标准偏差(变异系数),和适当的 $100(1-\alpha)\%$ 置信区间或其他合理的统计区间。除另有合理的证明,置信区间应与对应的精密度可接受标准相匹配。

对于多变量分析方法,常规指标预测均方根误差(RMSEP)包含了准确度和精密度。

3.3 准确度和精密度的联合验证

评价准确度和精密度的可选方法是通过建立一个综合性能标准来考虑它们的总体影响。合并的标准可提供方法产生可接受的总体变化结果的更全面证明,也能反映已建立的准确度和精密度的各自标准。

在方法开发过程中生成的数据有助于确定最佳的方法,并完善与合并的准确度和精密度相比较的适当的性能标准。

可采用预测区间、容忍区间或置信区间来评价联合的准确度和精密度。预测区间也称为期望容忍区间,可以被解释为下一个或几个可报告值以一定概率水平落在给定的置信区间;容忍区间也称为样本容忍区间,可以被解释为所有未来可报告值以一定概率水平落在给定的置信区间。也可采用其他合理的统计学方法。

数据要求

如果选择了综合性能标准,结果可作为综合值报告,以提供分析方法适用性的适当总体认知。如证明与分析方法适用性有相关性,准确度和精密度的各自结果应作为补充信息报告。应描述所使用的方法。

4 耐用性

分析方法的耐用性是指分析方法参数微小但刻意变化时,测量结果不受影响的能力,可用于说明方法正常使用时的可靠性,为所建立的方法用于常规检验提供依据。

传统上,耐用性并不是严格意义上需验证的性能特征。耐用性研究通常贯穿整个方法开发、验证过程,即方法全生命周期之中,属于风险评估的范畴。在方法开发阶段,就应考察其耐用性,根据所研究的方法类型对方法在预期操作环境中的适用性进行评价。

分析方法在开发过程中和验证前至少进行了部分优化,且进行了耐用性研究,通常,在分析方法验证中无需重复耐用性研究的实验过程。然而,耐用性试验有助于发现影响方法的变量,必要时,在方法验证中确认或完善耐用性评价仍是有意义的。耐用性评价资料应作为分析方法开发数据的一部分。

如果测试条件要求苛刻,则应在标准中列出的方法中写明,并注明可以接受变动的范围。可以先采用均匀设计确定主要影响因素,再通过单因素分析等确定变动范围;也可以采用试验设计(Design of Experiment, DoE)进行因素考察,尤其是存在多种因素可能有交互影响的情况下。

不同分析方法影响耐用性变动因素可能不同,典型的变动因素有:被测溶液的稳定性、样品的提取次数、时间等。液相色谱法典型的变动因素有:流动相的组成和 pH 值、不同品牌或不同批号的同类型色谱柱、柱温、流速等。气相色谱法典型的变动因素有:不同品牌或批号的色谱柱、固定相、不同类型的担体、载气流速、柱温、进样口和检测器温度等。

5 系统适用性试验

系统适用性试验 (System Suitability Testing, SST) 是分析方法的一个组成部分,被定义为对系统和方法性能特征的检查,通常在方法开发过程中设置,在方法验证中确认,在方法日常使用中用于确保其符合预期用途。

系统适用性试验的建立基于对方法开发数据、风险评估、耐用性以及先验知识等的理解,它是分析方法开发、验证,特别是耐用性试验的产物,是方法验证与方法应用相连接的桥梁。

作为重要的分离技术,色谱法有更多的变量,已有明确可设置的系统适用性试验参数。然而,如有必要且可能,其他分析方法也应设置系统适用性试验。设与不设,或设置何种特定的系统适用性试验参数取决于方法的类型和耐用性实验结果。方法越复杂、受影响的因素越多,需要设置的系统适用性试验参数越多,以保证在后续的方法转移和日常使用中,经过验证的分析方法始终具备通过验证时的性能。

系统适用性要求应在样品分析之前和/或期间得到满足。在方法运行时不符合系统适用性要求将导致由分析方法所获得的结果不可信、不能用;在继续分析之前,对不符合系统适用性要求的原因进行分析或调查,必要时采取纠正措施。

6 统计学考虑

统计学方法是评价分析方法验证结果的有用工具,方法验证中涉及很多方面,其中准确度和精密度的评价是最为重要的内容之一。

有多种统计方法可用于评价准确度和精密度。对于定量分析方法,只有在真值或可接受的参比值已获得时才能评估其准确度。在某些情况下,有必要评估其相对准确度。在许多分析方法中,即使不能直接评估准确度,也应评估精密度。

若高、中、低浓度水平的准确度所有可报告值是独立的,且在各浓度水平是相近的,可将各浓度水平的准确度所有报告值合并起来评价;如果上述条件不成立,就需要采用方差模型进行分析,分别验证每个浓度水平的准确度。若高、中、低浓度水平的标准差所有可报告

值数据是独立的，且在各浓度水平是相近的，可将各浓度水平的标准差所有报告值合并起来评价；如果不满足上述条件，数据转换也许仍然允许合并所有数据；如果转换不成功，则分别验证每个浓度水平的精密度。

一种统计学方法不一定适用于所有情况。如适用，其他统计学方法可用于方法验证结果的评价。

基于概率评价的统计学方法也有一定的风险。分析方法验证的另一目的是评估分析方法可能的风险点，是对由分析方法误差而导致试验结果判断错误的概率是否在允许范围之内进行的评价。

分析方法应具有确定的目标，评价分析方法准确度和/或精密度是否符合要求的最终标准是分析方法是否满足预期的目的，即分析方法是否满足质量控制目的。

备注：该指导原则进行较全面的修订，所以为采用修订模式公示。