

## 附件 1：不确定度评定指导原则公示稿（第一次）

# 不确定度评定指导原则

本指导原则简述测量不确定度 (Measurement Uncertainty) 评定的基本原理、定义、术语、作用和评定方法，作为在药品质量控制中的应用指导。本指导原则并不旨在将测量不确定度的评定作为分析检测的常规任务和纯粹的数学工作。

### 1. 基本概念

测量不确定度用于表征合理地赋予被测量之值的分散性，是与测量结果相联系的非负参数。它是对被测量客观值在某一量值范围内的评估，是对测量结果质量的定量表征。

任何测量结果都是估计值，都有一定的不确定度。对测量不确定度进行评定的目的是增加对测量过程的理解，减小测量不确定度并降低测量结果应用于决策的风险。

### 2. 适用范围

在开发、修改和应用分析方法时，除应按要求进行分析方法验证、确认和转移外，结合测量不确定度评定结果，考虑不确定度的允许量，可以更科学、客观地评价分析方法是否符合预期的目的，并为限度制订提供合理的依据。

判定测量结果与标准中规定限值的符合性时，特别是对限值边缘的检测结果进行符合性判定时，考虑一定置信水平下的扩展不确定度，有利于正确解释测量结果，减小误判的风险。同时，在合理评定测量不确定度的基础上，针对影响测量结果的主要因素采取改进措施，如增加重复试验次数，使用准确度更高的仪器和标准物质等，有助于提高检测结果的准确性。

在标准物质的标定、测量仪器的校准或检定时，或其他对检测或校准结果有特殊要求等情况下，如需要，可按相应要求评定和表示测量结果的不确定度。

### 3. 评定方法和程序

实验室应制定与目的相适应的测量不确定度评定程序，对于不同的检测项目和检测对象，可以采用不同的评定方法。

测量结果不确定度评定的原理和方法基于不确定度传播律，在数学模型为线性时，有效应用不确定度传播律无需任何条件。但对于数学模型为非线性时，应

用不确定度传播律是有条件的。在数学模型为明显非线性，泰勒级数近似中高阶项涉及的输入量不相互独立，且其概率密度函数（PDF）不为高斯分布时，用不确定度传播律可能得到不可靠的评定结果，这时可采用一些近似或假设的方法处理，或考虑采用蒙特卡洛法（MCM）或者用其他合适的替代方法评定测量不确定度。

由于某些检测方法的性质，决定了无法从计量学和统计学角度对测量不确定度进行有效而严格的评定，这时应基于对相关理论原理的理解或使用该检测方法实践经验进行分析，列出各主要的不确定度分量，并做出合理的评定。同时应确保测量结果的报告形式不会造成对所给测量不确定度的误解。如果检测结果不是用数值表示或者不是建立在数值基础上（如合格/不合格，阴性/阳性，或基于视觉和触觉等的定性检测），则宜采用其他方法评定测量不确定度，例如假阳性或假阴性的概率。

对于某些通用检测方法，如果该方法规定了测量不确定度主要来源的极限值和计算结果的表示形式时，实验室只要按照该检测方法的要求操作，出具测量结果报告，即被认为符合不确定度评定要求。

测量不确定度不仅和测量仪器、测量方法、测量条件、测量程序、数据处理过程以及操作人员的水平等因素有关，还和对测量过程的把握程度及对不确定度来源的识别和量化水平等因素有关。通过对测量过程的分析，充分识别各种影响因素，可以提高不确定度评定的可靠性。测量不确定度一般可按以下基本步骤进行评定：

### 3.1 识别不确定度来源

测量不确定度来源的识别应从分析测量过程入手，从影响测量结果的因素考虑，即对测量方法、测量系统和测量程序作详细研究，为此应尽可能画出测量系统原理或测量方法的方框图和测量流程图如鱼骨图（又称石川图）。

测量结果的不确定度有很多来源，包括仪器、计算方法和分析人员等可能引入的不确定度分量，通常有但不限于：被测量的定义不完善；被测量定义的重现性不理想；取样的代表性不够；对测量过程受环境影响的认识不足或对环境条件的测量与控制不完善；对模拟仪器的读数存在人为偏移；测量仪器的计量性能存在局限性（如分辨率或鉴别力不够）；计量标准或标准物质提供的标准值不准确；

引用的常数或其他参数值不准确；测量方法和测量程序中的近似和假设；在相同条件下，被测量重复观测值的变化。

### 3.2 建立被测量数学模型

规定被测量，如：含量、浓度、相对密度等，明确被测量与其所依赖的输入量（例如被测数量、常数、校准标准值等）的关系，还应包括对已知系统影响量的修正，建立数学模型。

$$Y = f(X_1, X_2, \dots, X_n)$$

其中  $Y$  为被测量（输出量）； $X$  为影响量（输入量）。

在建立模型时，有一些潜在的不确定度来源不能明显地呈现在上述函数关系中，它们对测量结果本身有影响，但由于缺乏必要的信息，无法写出它们与被测量的函数关系。因此，在具体测量时无法定量地计算出它们对测量结果影响的大小，在计算公式中只能将其忽略，而在模型中应包括这些来源。这些来源在数学模型中可以将其作为被测量与输入量之间函数关系的修正因子（其最佳值为0），或修正系数（其最佳值为1）处理。

此外，有些特殊不确定度来源，如取（抽）样、预处理、方法偏离、测试条件的变化以及样品类型的改变等也应考虑在模型中。

每一个来源的不确定度对合成不确定度的贡献量即为一个不确定度分量。在识别不确定度来源后，有必要对不确定度的各个分量进行预估算。对小于最大分量的三分之一的分量一般可以忽略（除非这种分量数目较多），通常只需对其估计一个上限即可。重点应识别并仔细评估那些重要的分量，特别是占支配地位的分量。对难以用数学模型表示的检测量，必要时，仍可对各个分量作出预估算。

### 3.3 评定标准不确定度

以标准偏差表示的测量结果  $x_i$  的不确定度为标准不确定度，标准不确定度  $u(x_i)$  的评定可分为 A 类和 B 类。

为了简化评定过程，在样品量较大情况下，在逐项计算标准不确定度时，一般可忽略自由度；在样品量较小情况下，应根据具体情况考虑自由度。

#### 3.3.1 A 类评定

A 类评定系指对在规定条件下测得的量值用统计分析方法进行测量不确定度分量的评定。

对被测量  $X$  进行  $n$  次独立的等精度测量，得到测量结果为  $x_1, x_2, \dots, x_n$ ，

被测量  $X$  的最佳估计值是  $n$  个独立测得值的算术平均值  $\bar{x}$ ，
$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

单次测量结果的实验标准差为：
$$u(x_i) = s(x_i) = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$$

观测列平均值即估计值的标准不确定度：
$$u(\bar{x}) = s(\bar{x}) = \frac{s(x_i)}{\sqrt{n}}$$

为了使 A 类不确定度评定结果可靠，要求重复测量次数足够多。另外，在 A 类评定时应尽可能考虑随机效应的来源，使其反映到被测量中。

对被测量  $X_i$  在重复条件或复现性条件下，均进行了  $n$  次独立测量，有  $x_{i1}, x_{i2}, \dots, x_{in}$ ，其平均值为  $\bar{x}_i$ ，如有  $m$  组这样的被测量，则合并样本标准差为：

$$s_p = \sqrt{\frac{1}{m} \sum_{i=1}^m s_i^2} = \sqrt{\frac{1}{m(n-1)} \sum_{i=1}^m \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2}$$

单次测量结果的标准不确定度为：
$$u(x_i) = \sqrt{\frac{\sum_{i=1}^m s_i^2}{m}}$$

$n$  次测量平均值的测量结果标准不确定度为：
$$u(\bar{x}) = \sqrt{\frac{\sum_{i=1}^m s_i^2}{nm}}$$

在重复性条件或复现性条件下，对  $X$  进行  $n$  次重复测量，计算结果中的最大值与最小值之差  $R$ （称为极差）。在  $X$  可以估计接近正态分布的前提下，单次测量结果  $x_i$  的实验标准差  $s(x_i)$  可以近似地评估：
$$s(x_i) = \frac{R}{C} = u(x_i)$$
，式中极差系数  $C$  由表 1 给出，其值与测量次数  $n$  有关。一般在测量次数较少时采用极差法。

表 1 极差系数  $C$

$n$	2	3	4	5	6	7	8	9	10	15	20
$C$	1.13	1.69	2.06	2.33	2.53	2.70	2.85	2.97	3.08	3.47	3.73

### 3.3.2 B 类评定

B 类评定方法不同于 A 类评定，系指当被测量的估计值  $x_i$  不是由重复观测

得到，其标准不确定度  $u(x_i)$  可用  $x_i$  的可能变化的有关信息或经验来进行的一类评定。

常用检测项目及方法的 B 类不确定度主要包括：样品、对照品称量引入的分量；对照品纯度引入的分量；测试溶液制备所用的容量器具引入的分量；仪器性能引入的分量；测定方法的偏差等。评定 B 类标准不确定度时常用的概率分布有正态分布、矩形（均匀）分布及三角分布等（见表 2）。

表 2 常用分布与  $k$ 、 $u(x_i)$

分布类型	$p(\%)$	$k$	$u(x_i)$
正态	99.73	3	$a/3$
	95	1.96	$a/1.96$
矩形（均匀）	100	$\sqrt{3}$	$a/\sqrt{3}$
三角	100	$\sqrt{6}$	$a/\sqrt{6}$

其中： $\alpha$  为区间半宽度， $k$  为包含因子。

区间半宽度  $\alpha$  一般根据以下信息确定：

- 先前测量的数据；
- 对有关技术资料 and 测量仪器特性的了解和经验；
- 生产厂提供的技术说明书；
- 校准证书、检定证书或其他文件提供的数据；
- 手册或某些资料给出的参考数据；
- 检定规程、校准规范或测试标准中给出的数据；
- 其他有用的信息。

若已知置信水平  $p$  的区间半宽度（用  $\pm\alpha$  表示），则将  $\alpha$  值除以该置信水平对应的包含因子  $k_p$ ，即可得到标准不确定度  $u(x) = \frac{\alpha}{k_p}$ 。正态分布的置信水平  $p$  与包含因子  $k_p$  之间的关系见表 3。

表 3 正态分布情况下置信水平  $p$  与包含因子  $k_p$  间的关系

$p(\%)$	50	68.27	90	95	95.45	99	99.73
$k_p$	0.67	1	1.645	1.960	2	2.576	3

如果  $\alpha$  值给出时没有给定置信水平，有理由认为可能是极限值，通常假定其为矩形分布，标准不确定度为  $u(x) = \alpha/\sqrt{3}$ ；

如果  $\alpha$  值给出时没有给定置信水平，但是有理由认为不可能是极限值，通常假定其为三角分布，标准不确定度为  $u(x) = \alpha/\sqrt{6}$ 。

如估计值  $x_i$  来源于说明书、校准证书、手册或其它资料，其中同时还明确给出了扩展不确定度  $U(x_i)$  和包含因子  $k$  的取值，则标准不确定度  $u(x_i) = \frac{U(x_i)}{k}$ 。

### 3.3.3 相对标准不确定度

标准不确定度  $u(x_i)$  与测量结果  $x_i$  的比值  $u_{rep}(x_i)$  为相对标准不确定度。

### 3.4 合成标准不确定度 $u_c(y)$

当测量结果是由若干个分量求得时，按各分量的方差或（和）协方差算得的标准不确定度为合成标准不确定度。

$$u_c(y) = \sqrt{\sum_{i=1}^N u_i^2(y) + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N u_i(y) u_j(y) r(x_i, x_j)}$$

式中， $u_i(y) = c_i u(x_i)$ ， $c_i$  称为灵敏系数，可通过以各分量为变量求偏导数获得。

当各影响量  $X_i$  独立无关时，相关系数  $r(x_i, x_j) = 0$ ，则

$$u_c(y) = \sqrt{\sum_{i=1}^N u_i^2(y)}$$

在实际工作中，只要无明显证据证明某几个分量有强相关时，均可按不相关处理。若发现分量间存在强相关，如采用相同仪器测量的分量，则尽可能改用不同仪器分别测量这些量，使其不相关。

实际工作中，为消除计量单位的影响，一般采用相对标准不确定度计算得到合成相对标准不确定度。

$$u_{c,rep}(y) = \sqrt{\sum_{i=1}^N u_{rep}^2(x_i)}$$

以合成相对标准不确定度乘以测量结果平均值 $\bar{Y}$ ，即得到合成标准不确定度，一个带有计量单位的量值。

$$u_c(y) = u_{c,rep}(y) \cdot \bar{Y}$$

### 3.5 计算扩展不确定度 $U$ 或 $U_p$

扩展不确定度是一个区间，包含被测量值分散性的主要区域，根据输出量（被测量）的分布情况，求出所要求的置信水平 $p$ 下的包含因子 $k$ ，则 $U_p = k_p u_c(y)$ 。

如果 $Y$ 接近于正态分布，多数情况下取 $p = 95\%$ ，此时， $k_p$ 为1.960（见表3），可近似取 $k_p = 2$ 。

若不能判断 $Y$ 的分布，则取 $k = 2$ 或 $3$ （一般取 $k = 2$ ）：

$$U = k u_c(y)$$

### 3.6 报告测量不确定度

报告测量不确定度应使用扩展不确定度，同时给出特定的置信水平或包含因子。测量不确定度评定报告应尽可能详细，以便使用者可以正确地利用测量结果，可由以下内容组成：被测量的测量模型；不确定度来源；A类和B类标准不确定度评定方法和评定过程；合成标准不确定度及其计算过程；扩展不确定度及包含因子；报告测量结果，包括被测量的估计值及其测量不确定度。

通常最终报告的扩展不确定度根据需要取一位或两位有效数字，一般采用常规的修约规则将数据修约到需要的有效数字，也可将最末位后面的数字进位而不是舍去。

报告测量结果时，通常，在相同计量单位下，被测量的估计值应修约到其末位与不确定度的末位一致。