

引言

很多重要的决策都是建立在化学定量分析的结果基础上，例如，化学定量分析的结果可以用于估计收益、判定某些材料是否符合特定规范或法定限量，或估计货币价值。当我们使用分析结果来作为决策依据的时候，很重要的一点是必须对这些结果的质量有所了解，换句话说，就是必须知道用于所需目的时，这些结果在多大程度上是可靠的。化学分析结果的用户，特别是涉及国际贸易领域时，正在受到越来越大的压力减少取得化学分析结果的重复劳动。达到这个目的的前提是必须建立对由非用户自身机构所得数据的信心。在化学分析的某些领域，现在已经有一个正式的（经常是法定的）要求，就是要求实验室引进质量保证措施来确保其能够并且正在提供所需质量的数据。这些质量保证措施包括：使用经确认的分析方法、使用规定的内部质量控制程序、参加水平测试项目、通过根据 ISO 17025 [H.1] 进行的实验室认可和建立测量结果的溯源性。

在分析化学中，过去曾经把重点放在通过特定方法获得的结果的精密度，而不是他们对所定义的标准或 SI 单位的溯源性。这种思路导致使用“官方方法”来满足法定要求和贸易要求。但是，因为现在正式要求建立结果的可信度，所以必须要求测量结果可以溯源至所定义的标准，如 SI 单位、标准物质或（如果适用）所定义的方法或经验方法（参见 5.2 节）。内部质量控制程序、水平测试和实验室认可可以作为辅助方法来证明与给定标准的溯源性。

上述要求的结果是：化学家们所从事的分析工作，正受到越来越大的压力，要求其证明结果的质量，特别是通过度量结果的可信度来证明结果的适宜性。这一般包括期望某个结果与其他结果相吻合的程度，通常与所使用的分析方法无关。度量该项内容的一个有用的方法就是测量不确定度。

虽然化学家们认识测量不确定度的概念已经有很多年了，但是直到 1993 年 ISO 才联合 BIPM、IEC、IFCC、IUPAC、IUPAP 和 OIML 出版了《测量不确定度表述指南》[H.2]，该指南正式确定了适用于广泛测量领域的评估和表达测量不确定度的通用原则。本指南文件说明了 ISO 指南的概念如何运用到化学测量中。它首先引入了不确定度的概念及不确定度和误差的区别，然后描述了评估不确定度的步骤，并在附录 A 中给出了评估过程的实际例子。

评估不确定度时，要求分析人员密切注意产生不确定度的所有可能来源。虽然对不确定度来源的详尽研究需要付出相当多的工作，但是，所付出的努力与所分析对象的复杂程度应相适宜。实际上，初步的分析就可快速确定不确定度最重要的来源。正如实例中所显示的那样，合成不确定度的数值几乎完全取决于那些重要的不确定度分量。评估不确定度时，正确的做法应该是集中精力分析最大的不确定度分量。此外，对于某特定实验室中指定方法（即：特定的测量程序）完成不确定度评估后，经过有关质量控制数据验证后，这一不确定度估计值能可靠地适用于以后该实验室使用该方法所得到的结果中。只要测量过程本身或所使用的设备未变化，就不需要再进一步进行不确定度评估了。在测量过程本身或所使用的设备发生变化时，需要重新审查不确定度评估结果，并将这项工作作为通常进行的方法再确认的一部分。

EURACHEM 第一版“分析测量中的不确定度评估”指南 [H.3] 是根据 ISO 指南于 1995 年出版的。

编制 EURACHEM 指南的第二版时，依据了化学实验室测量不确定度评估的实践经验，并更清楚地认识到实验室引入正式质量保证程序的必要性。第二版强调的是实验室引入测量不确定度评估程序时，应该与其现有的质量保证措施结合起来，因为上述质量保证措施通常提供了评估测量不确定度所需要的很多信息。所以本指南中明确给出了完全按 ISO 指南中的原则使用确认和有关的数据来进行不确定度评估。本方法也与 ISO17025:1999 [H.1] 的要求相吻合。

注：附录 A 中给出了实例。附录 B 中给出了定义列表。按惯例将文中第一次出现的有定义的术语用黑体字打印，并在方括号中给出了其在附录 B 中的索引号。定义主要摘自国际计量学基本和通用标准术语词汇表 (VIM) [H.4]、指南 [H.2] 和 ISO 3534 (统计学 - 词汇表和符号) [H.5]。附录 C 中用通用术语列出了产生测量结果的化学分析的总体结构。附录 D 描述了用于识别不确定度分量和筹划下一步所需试验的通用程序。附录 E 中列出了分析化学中测量不确定度评估所使用的某些统计方法。附录 F 中讨论了接近检出限时测量不确定度的处理方法。附录 G 列出了很多常见的不确定度来源和评估不确定度数值的方法。附录 H 是参考文献。

1 应用范围和领域

1.1 本指南给出了定量化学分析中评估和表述不确定度的详细指导。它是基于“ISO 测量不确定度表述指南” [H.2] 中所采用的方法,适用于各种准确度和所有领域——从日常分析到基础研究、经验方法和合理方法(参见第 5.3 节)。需要化学测量并可以使用本指南原理的一些常见领域有:

- (1) 制造业中的质量控制和质量保证;
- (2) 判定是否符合法定要求的测试;
- (3) 使用公认方法的测试;
- (4) 标准和设备的校准;
- (5) 与标准物质研制和认证有关的测量活动;
- (6) 研究和开发活动。

1.2 请注意在某些情况下可能还需要额外的指南。特别是,本指南中未包括如何使用公认方法(包括多重测量方法)给标准物质赋值,在符合性声明中如何使用不确定度估计值和和低浓度如何使用和表述不确定度,对此可能还需要额外的指南。与取样操作有关的不确定度也未在本指南中明确涉及。

1.3 由于一些领域的实验室已采用了正式的质量保证措施,所以本指南说明了应该如何使用从下列过程获得的数据进行测量不确定度评估:

- (1) 实验室作为规定测量程序 [B.8] 使用某种方法,对该方法所得分析结果的已识别来源的不确定度影响的评价;
- (2) 实验室中规定的内部质量控制程序的结果;
- (3) 为了确认分析方法而在一些有能力的实验室间进行的协同试验的结果;
- (4) 用于评价实验室分析能力的水平测试项目的结果。

1.4 无论是进行测试还是评估测量程序的操作性能,本指南始终认为有效的质量保证和控制措施,可以确保过程稳定并受控。这些措施通常包括诸如合格的工作人员、对设备和试剂的正确维护和校准、使用适当的参考标准、文件化的测量程序、使用适当的核查标准和控制图。参考文献 [H.6] 中有关于分析质量保证程序的更进一步的信息。

注:本段的意思是本指南认为所有分析方法的实施均是按已充分文件化的程序来进行。因此对分析方法的任何引用都意味着其有这类的程序。严格意义上讲,测量不确定度仅适用于这类程序的结果,而不是通常的测量方法 [B.9]。

2 不 确 定 度

2.1 不确定度的定义

2.1.1 本文所使用的（测量）不确定度的术语定义取自现行版本的《国际计量学基本和通用术语词汇表》[H.4]。定义如下：

测量不确定度：表征合理地赋予被测量之值的分散性，与测量结果相联系的参数。

注 1：这个参数可能是，如标准偏差 [B.23]（或其指定倍数）或置信区间宽度。

注 2：测量不确定度一般包括很多分量。其中一些分量是由测量序列结果的统计学分布得出的，可表示为标准偏差。另一些分量是由根据经验和其他信息确定的概率分布得出的，也可以用标准偏差表示。在 ISO 指南中将这不同种类的分量分别划分为 A 类评定和 B 类评定。

2.1.2 在化学分析的很多情况中，被测量 [B.6] 是指被分析物的浓度^①。然而，化学分析也可用于测量其他量，例如颜色、pH 值等，所以本文中使用了“被测量”这一通用术语。

2.1.3 上述不确定度的定义主要考虑了分析人员确信被测量可以被合理地赋值的数值范围。

2.1.4 通常意义上，不确定度这一词汇与怀疑一词的概念接近。在本指南中，如未加限定词，不确定度一词可能指上述定义中的有关参数，或是指对于一个特定量的有限知识。测量不确定度一词没有对测量有效性怀疑的意思，正相反，对不确定度的了解表明对测量结果有效性的信心增加了。

2.2 不确定度的来源

2.2.1 在实际工作中，结果的不确定度可能有很多来源，例如定义不完整、取样、基体效应和干扰、环境条件、质量和容量仪器的不确定度、参考值、测量方法和程序中的估计和假定以及随机变化等（在第 6.7 节中给出了不确定度来源更完整的说明）。

2.3 不确定度的分量

2.3.1 在评估总不确定度时，可能有必要分析不确定度的每一个来源并分别处理，以确定其对总不确定度的贡献。每一个贡献量即为一个不确定度分量。当用标准偏差表示时，测量不确定度分量称为标准不确定度 [B.13]。如果各分量间存在相关性，在确定协方差时必须

^① 在本指南中，未加限定词的术语“浓度”适用于任何具体的量，如质量浓度、数量浓度、数字浓度或体积浓度，除非引用了单位（例如用 $\text{mg}\cdot\text{l}^{-1}$ 表示的浓度明显就是质量浓度）。也应注意，用来表示成份的其他量，如质量分数、物质含量和摩尔分数，能直接表示浓度。

加以考虑。但是，通常可以评价几个分量的综合效应，这可以减少评估不确定度的总工作量，并且如果综合考虑的几个不确定度分量是相关的，也无需再另外考虑其相关性了。

2.3.2 对于测量结果 y ，其总不确定度称为合成标准不确定度 [B.14]，记做 $u_c(y)$ 是一个标准偏差估计值，它等于运用不确定度传播律将所有测量不确定度分量（无论是如何评价的）合成为总体方差的正平方根（请参考第 8 节）。

2.3.3 在分析化学中，很多情况下要用到扩展不确定度 U [B.15]。扩展不确定度是指被测量的值以一个较高的置信水平存在的区间宽度。 U 是由合成标准不确定度 $u_c(y)$ 乘以包含因子 k [B.16]。选择包含因子 k 时应根据所需要的置信水平。对于大约 95% 的置信水平， k 值为 2。

注：对于包含因子 k 应加以说明，因为只有如此才能复原被测量值的合成标准不确定度，以备在可能需要用该量进行其他测量结果的合成不确定度计算时使用。

2.4 误差和不确定度

2.4.1 区分误差和不确定度很重要。误差 [B.19] 定义为被测量的单个结果和真值 [B.3] 之差。所以，误差是一个单个数值。原则上已知误差的数值可以用来修正结果。

注：误差是一个理想的概念，不可能被确切地知道。

2.4.2 另一方面，不确定度是以一个区间的形式表示，如果是为一个分析过程和所规定样品类型做评估时，可适用于其所描述的所有测量值。一般不能用不确定度数值来修正测量结果。

2.4.3 此外，误差和不确定度的差别还表现在：修正后的分析结果可能非常接近于被测量的数值，因此误差可以忽略。但是，不确定度可能还是很大，因为分析人员对于测量结果的接近程度没有把握。

2.4.4 测量结果的不确定度并不可以解释为代表了误差本身或经修正后的残余误差。

2.4.5 通常认为误差含有两个分量，分别称为随机分量和系统分量。

2.4.6 随机误差 [B.20] 通常产生于影响量的不可预测的变化。这些随机效应使得被测量的重复观察的结果产生变化。分析结果的随机误差不可消除，但是通常可以通过增加观察的次数加以减少。

注：虽然在一些不确定度的出版物中是这样说的，但是，实际上算术平均值 [B.22] 或一系列观察值的平均值的实验标准差不是平均值的随机误差。它是由一些随机效应产生的平均值不确定度的度量。由这些随机效应产生的平均值的随机误差的准确值是不可知的。

2.4.7 系统误差 [B.21] 定义为在对于同一被测量的大量分析过程中保持不变或以可以预测的方式变化的误差分量。它是独立于测量次数的，因此不能在相同的测量条件下通过增加分析次数的办法使之减小。

2.4.8 恒定的系统误差，例如定量分析中没有考虑到试剂空白，或多点设备校准中的不准确性，在给定的测量值水平上是恒定的，但是也可能随着不同测量值的水平而发生变化。

2.4.9 在一系列分析中，影响因素在量上发生了系统的变化，例如由于试验条件控制得不充分所引起的，会产生不恒定的系统误差。

例子：

(1) 在进行化学分析时，一组样品的温度在逐渐升高，可能会导致结果的渐变。

(2) 在整个试验的过程中，传感器和探针可能存在老化影响，也可能引入不恒定的系统误差。

2.4.10 测量结果的所有已识别的显著的系统影响都应修正。

注：测量仪器和系统通常需要使用测量标准或标准物质来调节或校准，以修正系统影响。与这些测量标准或标准物质有关的不确定度及修正过程中存在的不确定度必须加以考虑。

2.4.11 误差的另一个形式是假误差或过错误差。这种类型的误差使测量无效，它通常由人为失误或仪器失效产生。记录数据时数字进位、光谱仪流通池中存在的气泡或试样之间偶然的交叉污染等原因是这类误差的常见例子。

2.4.12 有此类误差的测量是不可接受的，不可将此类误差合成进统计分析中。然而，因数字进位产生的误差可进行修正（准确），特别是当这种误差发生在首位数字时。

2.4.13 假误差并不总是很明显的。当重复测量的次数足够多时，通常应采用异常值检验的方法检查这组数据中是否存在可疑的数据。所有异常值检验中的阳性结果都应该小心对待，可能时，应向实验者核实。通常情况下，不能仅根据统计结果就剔除某一数值。

2.4.14 使用本指南获得的不确定度并没有考虑出现假误差或过错误差的可能性。

3 分析测量和不确定度

3.1 方法确认

3.1.1 在实践中，日常检测使用的分析方法的适用性通常是通过方法确认研究 [H.7] 加以评价的。这种研究对于整体性能和个别影响因子分别产生数据，并可以在正常使用中利用这些数据来评估该方法产生的测量结果的测量不确定度。

3.1.2 方法确认研究依赖于整体方法性能参数的确定。这些参数是在方法开发和实验室间研究或按照实验室内确认方案获得的。单个的误差或不确定度来源通常仅当其与所使用的总体精密度测量相比较时比较显著才加以研究。重点主要放在识别和消除（而不是修正）显著影响。这就形成了大多数潜在的显著影响因素已被识别，并已通过与整体精密度相比较来检查其显著性，并表明可以忽略的局面。在这种情况下，分析人员可以得到的数据主要包括总体性能数据及很多影响是非显著的证据和残余的某些显著影响的测量值。

3.1.3 定量分析方法的确认研究通常确定了下列部分或全部参数：

(1) 精密度：主要精密度测量包括重复性标准偏差 s_r 、复现性标准偏差 s_R ，（ISO 3534-1）和中间精密度，有时表示为 s_{zi} ，其中 i 表示变化因素的数目（ISO 5725-3: 1994）。重复性 s_r 表示在短时间内由同一操作人员操作同一设备等在同一间实验室内观察到的变异性。 s_r 可以在一个实验室内评估，也可以通过实验室间研究来评估。特定方法的实验室间的复现性标准偏差 s_R 可能只能用实验室间研究的方法直接评估；它表示不同实验室分析同一样品的差异性。当实验室内一个或几个因素（例如时间、设备和操作人员）发生变化时，中间精密度与所观察到的结果的变化有关；所得到的数值可能有所不同，这取决于哪些因素是保持恒定的。中间精密度评估通常是在实验室内获得的，但是也可以由实验室间研究取得。无论是通过独立方差合成的方法得到的，还是通过研究完整的操作方法获得的，分析过程所观察到的精密度是总体不确定度的基本分量。

(2) 偏差：分析方法的偏差通常是通过研究相关标准物质或通过加料研究而确定的。使用适当的标准数值来确定总体偏差，在建立对公认的标准溯源性 [B.12] 时很重要（见 3.2 节）。偏差可以表示为分析回收率（观察值除以期望值）。偏差应小到可忽略或应加以修正，但不论何种情况，与偏差确定有关的不确定度始终是总不确定度的一个基本分量。

(3) 线性：线性是分析方法在测量一个区间内的浓度时的重要特性。对于纯标准或实际样品响应能否成线性是可以确定的。通常并不将线性量化，而是通过检验或使用非线性显著性检测来检查线性。显著的非线性通常使用非线性校准函数加以修正，或通过选择更严格的操作区间加以消除。残余的线性偏差通常是在计算覆盖几个浓度的总体精密度估计值或与校准有关的不确定度中充分考虑（附录 E.3）。

(4) 检出限：在方法确认中，通常确定检出限只是为了建立一个方法实际操作区间的下限。虽然接近检出限的不确定度在评估时需要仔细考虑和特别处理（附录 F），检出限（无

论它是如何确定的)与测量不确定度的评估没有直接关系。

(5) 稳健性：很多方法开发和方法确认方案中需要直接研究特定参数的灵敏度。通常通过初步的“稳健性测试”来得到，即观察一个和多个参数变化的影响。如果影响显著（与稳健性测试的精密度相比较），就需要进一步详细研究来确定影响的大小，并因此选择一个所允许的操作区间因此稳健性测试的数据也可提供关于重要参数影响的信息。

(6) 选择性/特定性：虽然定义不是很严格，但是这两个术语均指一个分析方法对所要求的被分析物的专一响应程度。典型的选择性研究中需要研究可能的干扰物的影响，通常采用在空白和被分析物加强型的样品中加入潜在的干扰物并观察其响应值。结果通常用于证明实际影响并不显著。然而，因为这种研究是直接测量响应值的变化，所以当干扰物的浓度范围已知时，有可能使用这些数据来评估与潜在的干扰物有关的不确定度。

3.2 方法性能的实验研究

3.2.1 方法确认和方法性能研究的详细设计和实施在别处有详细的讨论 [H.7]，在此不再重复。然而，研究中影响不确定度评估的主要原则是很关键的，因此在下面加以讨论。

3.2.2 代表性是最重要的。也就是说，研究应尽可能反映方法正常使用中影响因素数量和范围的实际情况，并涵盖方法范畴内的浓度范围和样品类型。例如，在精密度试验中如果一个因子已代表性地变化，那么这个因子的影响就可以直接出现在所观察到的变化中，而不需要做额外研究，除非方法需要进一步的优化。

3.2.3 本文中，代表性变化是指影响参数必须具有与该参数的不确定度相适应的数值分布。对于连续参数，这可以是允许的范围或给出的不确定度；对于样品基体等非连续因子，这个范围与方法正常使用中所允许或者所遇到的不同类型相一致。注意代表性不只扩展到了数值的范围，而且包括其分布。

3.2.4 在选择变化因子时，必须保证只要可能就改变那些较大的影响量。例如，日与日之间的变化（可能产生于重新校准的影响）比重复性更大时，5天中的每一天进行2次测量就比2天中的每一天进行5次测量要给出更好的中间精密度估计值。在充分的控制下，在不同的天数里进行10次单独的测量可能会更好，虽然它对于日内重复性没有提供额外的信息。

3.2.5 通常处理随机选取的数据比处理从系统变化获取的数据更简单。例如：在足够长的时间内进行随机次数的试验通常包括代表性的环境温度影响，而按照24 h间隔系统完成的试验可能由于工作日内常规的环境温度变化而产生偏差。前者只需要评估总体标准偏差，后者还需要环境温度的系统变化，并随后按照温度的实际分布进行调整。然而随机变化有效性降低了。小量的系统研究就可以快速确定一个影响量的大小，但是通常需要测量30多次才能给出高于约20%相对准确度的不确定度分量。只要可能，系统地研究少量的主要影响因素是更好地选择。

3.2.6 当因素已知或可能产生交互作用时，必须保证对交互作用的影响加以考虑。这可以通过保证随机选取不同水平的交互作用参数或通过以获得方差和协方差信息为目的的缜密的系统设计来达到。

3.2.7 在对总体偏差研究时，标准物质和数值应与日常检测的物质相关。

3.2.8 为调查和确定某个影响量的显著性而进行的研究，必须有足够的力量在这些影响量

变得实际显著之前确定这些影响量。

3.3 溯源性

3.3.1 能够可信地比较来自不同实验室的结果或同一实验室不同时期的结果是重要的。通过保证所有的实验室均使用同样的测量尺度或同样的“参考点”，就可达到这一点。许多情况下是通过建立能够到达国家或国际基准，理想的情况下是测量国际单位制 SI（为了长期的一致）的校准链。熟悉的例子是分析天平。每个天平用标准砝码来校准，而后者（最终）跟国家基准核对，如此直至千克基准。这种可到达已知参考值的不间断链的比较提供了对共同“参考点”的溯源性，确保不同的操作者使用同一测量单位。在日常测量中，对用来获得或控制某个测量结果的所有的中间测量，均建立溯源性，可极大地帮助达到一个实验室（一个时期）和另一个实验室（另一个时期）的测量结果的一致性。因此在所有测量领域中溯源性是一个重要的概念。

3.3.2 溯源性的正式定义 [H.4] 是：“通过一条具有规定不确定度的不间断的比较链，使测量结果或测量标准的值能够与规定的参考标准，通常是与国家测量标准或国际测量标准联系起来的特性。”

需要提及不确定度是由于实验室间的一致性在一定程度上受到每个实验室的溯源性链所带来的不确定度的限制。溯源性因此与不确定度紧密联系。溯源性提供了一种将所有有关的测量放在同一测量尺度上的方法，而不确定度则表征了校准链链环的“强度”以及从事同类测量的实验室间所期望的一致性。

3.3.3 通常，某个可溯源至特定参考标准的结果的不确定度，将由该标准的不确定度与对照该标准所进行的测量的不确定度组成。

3.3.4 完整的分析过程的结果的溯源性应通过下列步骤的综合使用来建立：

- (1) 使用可溯源标准来校准测量仪器；
- (2) 通过使用基准方法或与基准方法的结果比较；
- (3) 使用纯物质的标准物质 RM；
- (4) 使用含有合适基体的有证标准物质 CRM；
- (5) 使用公认的、规定严谨的程序。

下面依次讨论每个步骤。

3.3.5 测量仪器的校准

在任何情况下，所使用的测量仪器的校准必须可溯源到适当的标准。分析过程的定量阶段通常使用其值可溯源至 SI 的纯物质的标准物质来进行校准。这种做法为这部分过程结果提供了至 SI 的溯源性。然而，有必要使用额外的程序，为诸如萃取和样品净化等属于定量阶段之前的操作结果建立溯源性。

3.3.6 使用基准方法来进行测量

下面是目前对基准方法的描述：

“测量的基准方法是一种具有最高计量学特性，其操作可用 SI 单位进行完整地描述并被理解，其结果无需参考相同量的标准即可接受的方法。”

基准方法的结果通常直接可溯源至 SI 单位，并且相对于该参考标准具有所能获得的最

小不确定度。基准方法通常只由国家测量机构来实施，很少用于日常测试或校准。如有可能，通过直接比较基准方法和测试或校准方法的测量结果来达到对基准方法结果的溯源性。

3.3.7 使用纯物质标准物质（RM）

通过测量含有或由已知量纯物质组成的样品可证明溯源性。例如，可通过加料或标准加样来达到。然而，评估测量系统对所使用的标准和所测试的样品的不同响应总是必要的。可惜的是，对于许多化学分析，在加料或标准加样的具体例子中，对不同响应的修正和其不确定度可能较大。因此，虽然原则上该结果对 SI 单位的溯源性得以建立，但实际上，除最简单例子外，所有的例子，其结果的不确定度可能大得不可接受或甚至无法量化。假如其不确定度不能量化，则其溯源性并未建立起来。

3.3.8 使用有证标准物质（CRM）进行测量

对有证基体的 CRM 进行测量，并将测量结果与其有证数值比较可证明溯源性。当有合适基体的 CRM 时，该程序与使用纯物质 RM 相比能够降低不确定度。假如 CRM 的值可溯源至 SI，则这些测量也可溯源至 SI 单位，使用标准物质的不确定度评估在 7.5 节讨论。然而，即使在这种情况下，假如样品和标准物质的成分之间没有很好匹配，其结果的不确定度可能大得无法接受甚至无法量化。

3.3.9 使用公认程序

通过使用规定严谨并且普遍接受的程序可达到适当的可比性。该程序通常用输入参数加以规定，例如一组规定了萃取时间，颗粒大小等等。当这些输入参数的值按正常方式可溯源至所规定的参考标准时，使用这类程序产生的结果被认为是可溯源的。其结果的不确定度来自所规定的输入参数的不确定度和规范不完整的影响以及执行过程中的变化（见 7.8.1 节）。当估计另一种方法或程序的结果可与这类公认程序的结果相比较时，则可通过比较两者的结果来建立对该公认值的溯源性。

4 测量不确定度的评估过程

4.1 不确定度的评估在原理上很简单。下述段落概述了为获取测量结果的不确定度估计值所要进行的工作。随后的章节提供了用于不同情况下的附加指南，特别是关于使用方法确认研究的数据和使用正式的不确定度传播律。这些步骤包括：

(1) 第一步 规定被测量

清楚地写明需要测量什么，包括被测量和被测量所依赖的输入量（例如被测数量、常数、校准标准值等）的关系。只要可能，还应该包括对已知系统影响量的修正。该技术规定资料应在有关的标准操作程序（SOP）或其他方法描述中给出。

(2) 第二步 识别不确定度的来源

列出不确定度的可能来源。包括第一步所规定的关系式中所含参数的不确定度来源，但是也可以有其他的来源必须包括那些由化学假设所产生的不确定度来源。附录 D 给出了用结构图表示的一般步骤。

(3) 第三步 不确定度分量的量化

测量或估计与所识别的每一个潜在的不确定度来源相关的不确定度分量的大小。通常可能评估或确定与大量独立来源有关的不确定度的单个分量。还有一点很重要的是要考虑数据是否足以反映所有的不确定度来源，计划其他的实验和研究来保证所有的不确定度来源都得到充分的考虑。

(4) 第四步 计算合成不确定度

在第三步中得到的信息，是总不确定度的一些量化分量，它们可能与单个来源有关，也可能与几个不确定度来源的合成影响有关。这些分量必须以标准偏差的形式表示，并根据有关规则进行合成，以得到合成标准不确定度。应使用适当的包含因子来给出扩展不确定度。

图 1 用图示方法表示了该过程。

4.2 下面的章节对执行上述步骤提供了指南，并展示了如何根据所获得的有关多个不确定度来源的合成影响信息来简化这些步骤

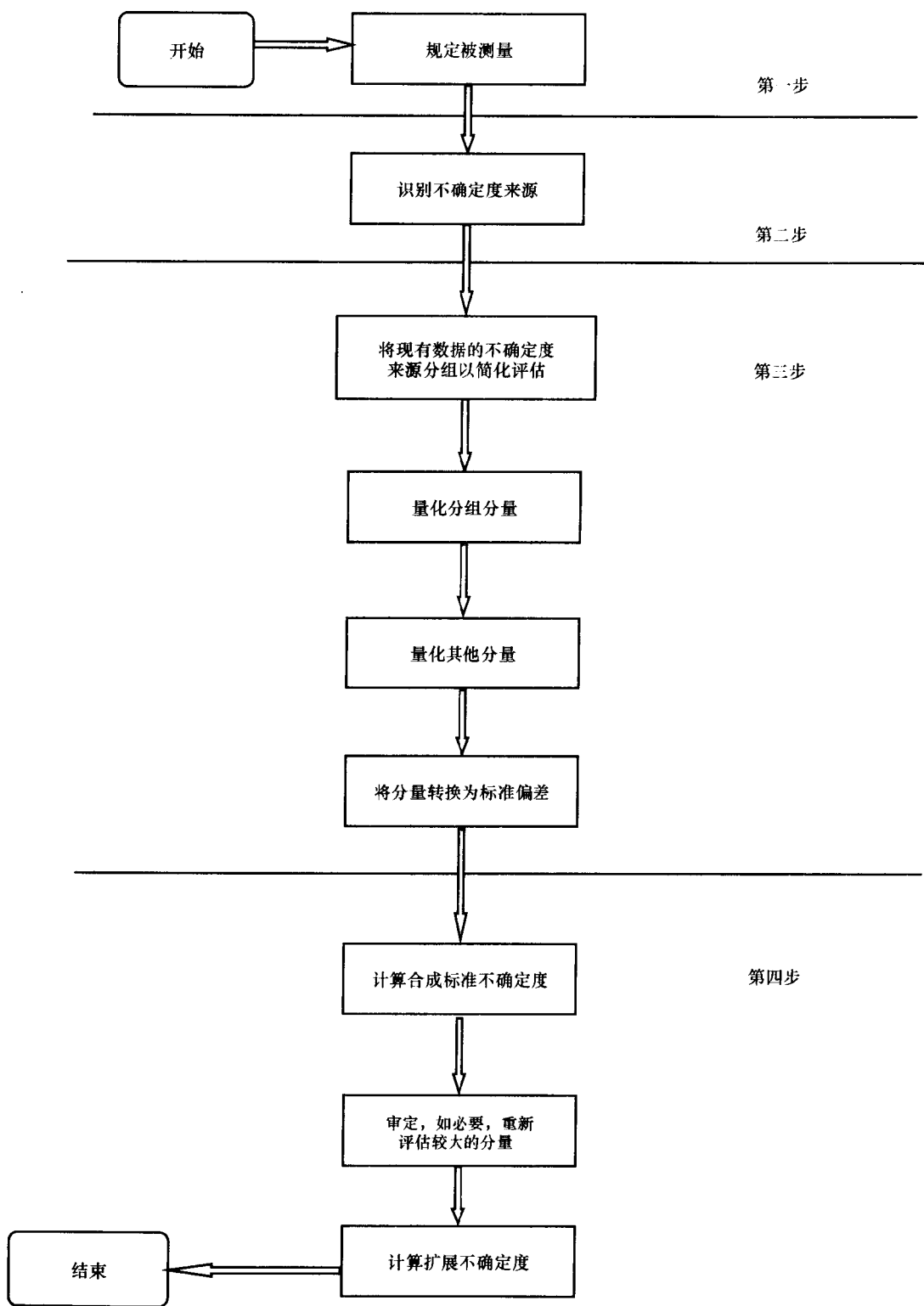


图 1 不确定度评估过程

5 第一步 被测量的技术规定

5.1 在本不确定度评估的指南中，“被测量的技术规定”要求清楚明确地说明正在测量什么，并定量表述被测量的值与其所依赖的参数之间的关系。这些参数可能是其他被测量、不能直接测量的量或者常数，并应明确过程中是否包括取样步骤。如果有，与取样过程有关的不确定度评估必须考虑。所有这类信息应在标准操作程序（SOP）中写明。

5.2 在分析测量中，特别重要的是要区别其测量结果独立于所使用的方法的测量和其测量结果依赖于所使用的方法的测量。后者通常称作经验方法。下面的例子将进一步澄清这一点。

例子：

(1) 测定合金中镍含量的方法通常会产生相同的结果，该结果以相同单位表示，一般表示为质量或摩尔分数。原则上，需要对方法偏差或基体产生的任何系统影响进行修正，虽然更为常用的做法是确保任何此类影响是较小的。除了作为参考信息外，结果通常无需引述所使用的特定方法。该方法不是经验方法。

(2) “可萃取脂肪”的测定可能相差很大，它取决于所规定的萃取条件。由于“可萃取脂肪”完全取决于条件的选择，所以所使用的方法就是经验方法。考虑对该方法的内在偏差进行修正是无意义的，因为被测量是由所使用的方法所定义的。结果的报告通常提到所使用的方法，而不对该方法的内在偏差进行修正。该方法被认为是经验方法。

(3) 当基质或基体的变化具有很大的和不可预测的影响时，通常开发程序的惟一目的是达到测量同一材料的实验室之间具有可比性。然后可将该程序采用为地方，国家或国际标准方法，凭此进行贸易或做出决定，而无需绝对度量被分析物和实际含量。按常规可不对方法偏差或基体影响进行修正（而不管在方法开发过程中它们是否已被最小化）。通常报告结果时没有对基体或方法偏差进行修正。该方法被认为是经验方法。

5.3 经验方法和非经验方法（有时称为合理方法）的区别很重要，因为它影响了不确定度的评估。在 5.2 例子（2）和（3）中，因为采用了习惯做法，与一些较大的影响量相关联的不确定度在日常使用中并不相关。所以，应适当考虑结果是独立于还是依赖于所使用的方法，并且只有那些与所报告的结果有关的影响量才应包括在不确定度评估过程中。

6 第二步 识别不确定度来源

6.1 应列出不确定度有关来源的完整清单。在本步骤，无需考虑单个分量的量化问题本步骤的目的是明确应考虑什么。在第三步中，将考虑处理每一个来源的最佳方法。

6.2 在列出不确定度来源的清单时，通常方便的办法是从那些根据中间数值计算被测量的基本表达式开始。这个表达式中的所有参数可能都有一个与其数值相关的不确定度，因此都是不确定度的潜在来源。此外，也可能有其他参数并没有明显地出现在用于计算被测量数值的表达式中，但却影响该测量结果，例如：萃取时间或温度。它们也是不确定度的潜在来源。所有这些不同的来源都应包括在内。在附录 C（分析过程的不确定度）中给出了其他信息。

6.3 附录 D 中的因果图是列出不确定度来源的一种非常方便的方法，它表示了他们之间的相互关系，以及他们对结果的不确定度的影响，也有助于避免重复计算不确定度来源。虽然还有其他方法可以列出不确定度的来源，本书的后面章节和附录 A 的例子中均采用因果图的方法。其他信息参见附录 D（分析不确定度来源）。

6.4 列出不确定度的来源后，他们对于测量结果的影响原则上可以用一个正式的测量模型来表述，其中每一个影响量都与公式中的一个参数有关或在公式中是变量。然后该公式形成了用影响该测量结果的所有独立因素表示的测量过程的完整模型。该函数可能非常复杂，不能明确地写出，但是只要可能，就应该这样做。因为表达式的方式通常决定了合成各个不确定度分量的方法。

6.5 还有一个很有用的技巧是将测量过程考虑成一系列独立的操作（有时称为单元操作），每一个操作可以单独评价以得到与之相关的不确定度估计值。当类似的测量过程共享相同的单元操作时，这是一个很有用的方法。每个操作各自的不确定度构成了总不确定度的分量

6.6 实际上，在分析测量中，更普遍的做法是考虑与整体方法性能的要害有关的不确定度，例如可观察的精密度和相对于合适的标准物质所测得的偏差。这些构成了不确定度评估中的主要分量，最好在模型中表示为影响结果的独立因素。然后，有必要对其他可能的分量进行评估，只需要检查他们是否显著，并只量化那些显著的量。本方法的进一步指南，见第 7.2.1 节，其更适用于使用方法确认数据的情况。

6.7 典型的不确定度来源包括：

(1) 取样

当内部或外部取样是规定程序的组成部分时，例如不同样品间的随机变化以及取样程序存在的潜在偏差等影响因素构成了影响最终结果的不确定度分量。

(2) 存储条件

当测试样品在分析前要储存一段时间，则存储条件可能影响结果。存储时间以及存储条件因此也被认为是不确定度来源。

(3) 仪器的影响

仪器影响可包括，如对分析天平校准的准确度限制；保持平均温度的控温器偏离（在规

范范围内)其设定的指示点;受进位影响的自动分析仪。

(4) 试剂纯度

即使母材料已经化验过,因为化验过程中存在着某些不确定度,其滴定溶液浓度将不能准确知道。例如许多有机染料,不是 100% 的纯度,可能含有异构体和无机盐。对于这类物质的纯度,制造商通常只标明不低于规定值。关于纯度水平的假设将会引进一个不确定度分量。

(5) 假设的化学反应定量关系

当假定分析过程按照特定的化学反应定量关系进行的,可能有必要考虑偏离所预期的化学反应定量关系,或反应的不完全或副反应。

(6) 测量条件

例如,容量玻璃仪器可能在与校准温度不同的环境温度下使用。总的温度影响应加以修正,但是液体和玻璃温度的不确定度应加以考虑。同样,当材料对湿度的可能变化敏感时,湿度也是重要的。

(7) 样品的影响

复杂基体的被分析物的回收率或仪器的响应可能受基体成份的影响。被分析物的物种会使这一影响变得更复杂。

由于改变的热力情况或光分解影响,样品/被分析物的稳定性在分析过程中可能会发生变化。

当用“加料样品”用来估计回收率时,样品中的被分析物的回收率可能与加料样品的回收率不同,因而引进了需要加以考虑的不确定度。

(8) 计算影响

选择校准模型,例如对曲线的响应用直线校准,会导致较差的拟合,因此引入较大的不确定度。

修约能导致最终结果的不准确。因为这些是很少能预知的,有必要考虑不确定度。

(9) 空白修正

空白修正的值和适宜性都会有不确定度。在痕量分析中尤为重要。

(10) 操作人员的影响

可能总是将仪表或刻度的读数读高或低。

可能对方法作出稍微不同的解释。

(11) 随机影响

在所有测量中都有随机影响产生的不确定度。该项应做为一个不确定度来源包括在列表中。

注:这些来源不一定是独立的。

7 第三步 量化不确定度

7.1 引言

7.1.1 已按第二步（第 6 章）的说明识别不确定度来源后，下一个步骤就是要量化这些不确定度来源所产生的不确定度。可以通过以下方法：

(1) 评价每个不确定度来源的不确定度，然后将其按照第 8 章所述的方法合成。例子 A1~A3 说明了这种方法的使用；或者

(2) 使用方法性能数据，直接确定来自一些或所有不确定度来源的结果的不确定度合成分量。例子 A4~A6 给出了该过程的使用。

实践中，上述方法的组合使用通常是必要和便利的。

7.1.2 不管使用哪种方法，评价不确定度最需要的大部分信息可从确认研究的结果、质量保证/质量控制（QA/QC）的数据和其他为核查方法性能进行的实验工作中得到。然而，不是所有来源的不确定度都有评估其不确定度所需的数据，可能有必要进行 7.10~7.14 节所描述的其他工作。

7.2 不确定度的评估程序

7.2.1 用来评价总不确定度的程序取决于可提供的方法性能数据。制定这一程序的步骤包括：

(1) 解决现有数据与信息需求之间的矛盾

首先，检查不确定度来源表，看哪些不确定度来源可用现有的数据计算，而不管现有的数据是来自对具体分量的明确研究，还是来自整个方法实验过程中的隐含变化。应对照步骤二所列的表来核查这些来源，并列出任何余下的来源，以提供哪些不确定度分量已包括在内的可审查记录。

(2) 策划获取所需的其他数据

对于没有被现有数据适当地包括的不确定度来源，可以从文献或现有资料（证书、仪器规格等）中获取其他信息，或策划实验以获取所需其他数据。附加的实验可采取具体研究单个不确定度分量的方式，或采用常用的方法性能研究以保证重要因素有代表性的变化。

7.2.2 很重要的是认识到不是所有的分量都会对合成不确定度构成显著的贡献。实际上，只有很少的分量才会有显著影响。除非数量很多，比最大的分量小三分之一的那些分量无需深入评估。对于每一个分量或合成分量的贡献进行初步评估，去掉那些不重要的分量。

7.2.3 下面的章节对将要采取的程序提供了指南，这取决于现有的数据和所需的其他信息。第 7.3 节给出了使用以前的实验研究数据（包括确认数据）的要求。第 7.4 节简述了单纯从单个不确定度来源评估不确定度的方法。对于所识别的所有或很少不确定度来源来说这可能

是必要的，这取决于现有的数据，因此也在稍后的章节中加以考虑。7.5~7.9 节描述了在不同情况下评估不确定度的方法。7.5 节适用于使用极为匹配的标准物质的情况。7.6 节适用于使用协同研究数据，7.7 节适用于使用实验室内部确认数据。7.8 节描述了经验方法的特殊处理。7.9 节讨论了特别方法 (ad-hoc method)。量化单个不确定度分量的方法，包括实验研究、文献和其他数据、建立模型和专业判断，详见 7.10~7.14 节。7.15 节讨论了在不确定度评估中对于已知偏差的处理。

7.3 以前研究的相关性

7.3.1 当不确定度评估至少部分是基于以前的方法性能研究时，有必要证明使用以前研究结果的有效性。典型的包括：

(1) 证明以前所得到的精密度具有相当水平。

(2) 特别是通过用相关标准物质偏差的确定（例见 ISO Guide 33 [H.8]），通过适当的加料研究或通过有关水平测试项目或其他实验室比对的满意表现来证明使用以前所得到的偏差数据的合理性。

(3) 通过定期质量控制样品结果表明持续运行在统计控制之内以及实施有效的分析质量保证程序。

7.3.2 满足上述条件，并且方法在其使用范围和领域内应用时，通常可以使用该实验室以前的研究（包括确认研究）数据来直接评估该实验室的不确定度。

7.4 量化单个分量来评估不确定度

7.4.1 有些情况下，特别是没有或有很少方法性能数据时，最适合的程序可能是分别评估每个不确定度分量。

7.4.2 合成单个分量的一般程序是准备一个详细的试验过程的定量模型（参考第 5 节和第 6 节，尤其是第 6.4 节），评估与单个输入参数相关的标准不确定度，并使用不确定度传播律对其进行合成，参见第 8 章。

7.4.3 为了更清楚地说明问题，通过试验方法和其他方法评估单个分量的详细指南推迟到第 7.10~7.14 节中给出。附录 A 中的例子 A1 和 A3 给出了该程序的详细描述。使用该程序的详细指南也参见 ISO 指南 [H.2]。

7.5 极匹配的有证标准物质

对有证标准物质的测量通常作为方法确认或重新确认的组成部分，有效地达到了对整个测量程序用可追溯的标准物质所进行的校准。因为这一程序提供了许多不确定度潜在来源的综合影响的信息，所以它为不确定度评估提供了很好的数据。详见 7.7.4 节。

注：ISO 指南 33 [H.8] 对于使用标准物质检查方法性能方面给出了有用的描述。