

C O D E X A L I M E N T A R I U S

国际食品标准



联合国粮食
及农业组织



世界卫生组织

E-mail: codex@fao.org - www.codexalimentarius.org

国际食品法典
结果不确定度评估导则
CAC/GL 59-2006
2011 年修订

1 引言

应 ISO/IEC 17025 的要求，实验室要对分析结果的不确定度进行评估。食品检测实验室按残留分析 GLP 导则修订版(CAC/GL 40-1993, Rev. 1- 2003)操作可得到充足的数据，它们来自方法研究、实验室间的验证实验以及实验室内质量检测，这些数据可以用来估计不确定度，尤其是实验室经常进行的实验。此导则的制定还考虑了 CCMAS 的总体建议。

1.1 不确定度的概念和组成

所谓测量的不确定度是指由测量过程产生的数据的“不确定度”。在分析化学中，它通常被定义为实验过程的不确定度，也包括与采样有关的不确定因子的不确定度。

所报告的或实验结果的不确定度的“估计”描述了一个范围，其真值能期望以约定的概率水平位于其中。它与测量误差的概念不同，后者用来描述一个单独的结果与真值之间的差。不确定度则提供了一个所报告结果在较高置信水平上面的准确性。

数据不确定度的来源是很多的，详见表 1 与 2。不确定度的评估要很好地理解与估计有关的测量过程的各个步骤对不确定度的影响。

2 不确定度来源的确定

测量的不确定度由很多因素形成，包括样本采集在内的每一个测试步骤都会影响评估结果。一个分析结果的不确定度，由该测定的三个主要方面所影响。

- 外部操作：抽样 (S_s)、包装、运送和储藏¹；
- 测试样品的制备：分样、样品的制备与加工 (S_{sp})；
- 分析 (S_A)：提取、净化、蒸发、衍生、仪器测定²

合成不确定度 (S_{Res})和相对不确定度(CV_{Res})可以根据误差传播定律计算：

$$S_{\text{Res}} = \sqrt{S_s^2 + (S_{sp}^2 + S_A^2)} ; S_{\text{Res}} = \sqrt{S_s^2 + S_L^2} \quad (1)$$

如果整个样本都进行了分析，且平均残留保持不变，等式可写为：

$$CV_{\text{Res}} = \sqrt{CV_s^2 + CV_L^2} \quad \text{and} \quad CV_L = \sqrt{CV_{sp}^2 + CV_A^2} \quad (2)$$

CV_L 是实验室检测状态的相对不确定度，可能来源于分样、样品的制备与加工以及分析阶段。

如果采样不属于实验室管理范畴，通常只要求实验室估计它所能掌控的那些因素所引起的不确定度，即仅限于实验室内的过程。

2.1 分析测量中的误差

在一般测量中，误差可分为 3 类：过失误差、随机误差和系统误差。

过失误差 分析过程中产生的无意/非预期误差，这种类型的误差使测定结果失效。实验室质量保证程序要尽量减少过失误差，在不确定度的估计中计入过失误差或统计过失误差是不可能的。本文件不作进一步讨论。

随机误差 它出现在整个测量过程中，使重复测出的结果落在均值的两边。测量的随机误差是不可抵消的，只能通过增加测量次数与培训分析人员来减少其影响。

系统误差 出现在很多实验中，而它们的作用是很不同的。在一个实验中，所有系统误差之和称为偏离 (bias)。多次测量系统误差不会减少系统误差，因此，通过重复分析不能直接检出个别的系统误差。除非采取合适的预防措施，否则发现不了系统误差。实际测试中，分析中的系统误差可以通过下述方法识别——分析技术中使用参照物质，样品由另一个分析人员或由另一个实验室分析，或采用另外的分析方法再分析。然而，仅当参照物按分析、基质、浓缩完全匹配时才能发现系统误差。一个方法的偏离也可由回收率研究来得到。不过，回收率研究只评估分析的影响 (S_A)，而不能用于实际样品分析中的偏离或者这种偏离在分析步骤开始之前已经

¹ 包装，运输，储存和实验室样本准备可能对残留检测产生显著影响，但基于目前的信息，通常无法量化这些步骤对不确定度的影响。这些误差有采样地点的选择，采样时间，错误标注的分析物分解或样本污染。

² 如果结果根据回收率而做了更正，那么此项相关的不确定度也应随之更正。

存在在农药分析中，除非当平均回收率显著偏离 100%时，结果通常不进行回收率校正。如果结果已用回收率校正，那么，与回收率相关的不确定度应该并入测定的不确定度估计中。

在表 1 和表 2 中列出了几种误差源的例子。要注意的是，在不确定度估计中，不是所有提及的误差源都必须被评估，一些误差源已并入了总的不确定度中，而另一些则忽略而不予考虑。但是，在淘汰前要识别与评估所有的误差源。进一步的信息可从已发表的文献获得^{3,4}。

表 1: 实验测试样品制备过程中的误差来源

	系统误差来源	随机误差来源
样本制备	样品分析部分的不正确的选择	分析样品与其他样品接触造成的污染 淋洗、冲刷程度不同所造成的去除程度不同
样本处理 (S _{Sp})	样品制作过程中分析物的分解，样本的交叉污染	分析样品间的不均匀性 分析样品混匀和粉碎的不均匀性 样品均质过程中温度的变化 植物的成熟度对均质效果的影响

表 2: 分析中误差的来源 (S_A):

	系统误差来源	随机误差来源
提取/净化	分析物回收不完全 共同提取物的干扰（吸附剂的过载）	样品组成（如：水、脂肪及糖分的含量）变化 样品基质及溶剂的组成和温度
定量分析	共同提取化合物的干扰 分析用标准品纯度不准 天平的精确性与线性 质量/体积测量偏差 操作人员读取仪器、设备的偏差 被测物不来源于样本（比如，被包装材料污染） 测定物由于残留定义的不同而不同 偏差校正	衍生化反应的不完整和可变性 分析过程中实验室环境条件的改变 进样、色谱及检测器条件的变化（基质效应、系统惰性、检测器响应、信噪比等） 操作者影响（疏忽） 校准

3 测量不确定度的评估方法

实验室评估测量不确定度的程序有很多种选择，最常用的两个程序是“自下向上（bottom up）”方法与“自上向下（top down）”方法。

自下向上法:

“自下向上（bottom up）”方法或称逐个组分（component-by-component）法，组成以行为作基础的过程。分析者拆分所有的分析操作至各初级行为，从而估计这些行为的贡献进而得到测量过程的组合不确定度值。这“bottom up”方法是相当费力并要求对整个分析过程具有详细的了解。对分析者的好处是该方法提供了对所测量的不确定度有明显贡献的分析行为的透彻了解。因此，在该方法的进一步应用中可安排一些关键的控制点来减少或管理所测量的不确定度。

自上向下法:

该方法是基于分析方法的建立及从实验室检测样本、已发表的文献数据以及实验室间的协作试验等得来的长期的精确的数据。基于实验室间研究项目所得的不确定度估计也需要考虑数据的实验室间变异性以及提供一个方法性能与结合它应用的不确定度的可靠的估计。不过要知道，协作研究项目是设计来评估特定方法与所参与实验室的性能的，通常，它们并不评估样本在较高均质性的制备或分析过程中的不精确性。

³ EURACHEM 对分析法不确定度定量的指导，第 2 版，1999, <http://www.measurementuncertainty.org>

⁴ Ambrus A. 残留数据的可靠性, Accred. Qual. Assur. 9, pp. 288-304. 2004.

农药残留分析实验室通常要在众多的日常样本中寻找大量残留样本中 200 个以上残留值，用以代表无限多个组合情况。因此建议，对于多个残留的关于不确定度的评估程序，实验室要使用一个适当选择的分析物与样本基质的范围，使它们能代表这些残留物与被分析的样品（依据农药残留分析 GLP 规范的选择具有代表性物理、化学性质的成分），而不是对每个方法/分析物/基质组合来建立不确定度。这个关于分析物与基质的代表性范围的选择所提供的评估，应有验证数据和选择基质/分析物的研究报告做依据。

总之，实验室应使用它们自己的长期精确数据或以实际操作做基础的程序（逐个组分的计算）来建立与精炼不确定度数据。

在某种情形下，利用样本的变异性来评估不确定度也是合适的。这要求对样本批次内的分析物变异性的了解，而不是实验室或分析物已有的变异性的了解。从 8 500 多个残留数据（表 4）的统计分析所得的值提供了当今的最佳评估（参见上页注 1）⁵。这些评估可用来结合到组合不确定度值中。

同样，如果想要知道分析员间与实验室间及分析物的变异性，还必须考虑到样本储存与分析过程中分析物的稳定性。

3.1 含多种成分的分析结果的不确定度评估

对于多种成分残留结果的场合，包括结构与旋光异构体、代谢物与其他的分解产物等多种成分，其结果的不确定度的评估应采用不同的方法，特别当各组分的残留的全体或部分已建立了 MRL 的情况。基于多个峰的测定所得到结果的随机误差与系统误差的评估已在近期文献中详细说明⁶。

4 可接受的不确定度指导值

由单一实验室的一系列实验的标准偏差来作为标准不确定度的一个测定，要求大量实验数据，是不易办到的。而对于较少的数据量，标准偏差的真值可按下面公式来估计：

由观察数（ n ）有关的标准偏差的真值（ σ ），计算的标准偏差（ S ），95%置信度的均值的预期范围（ X ）在表 3 中给出了说明。乘积因子 f ，作为测定数的函数，提供了估计值与真值间的函数。

表 3 为计算标准偏差与均值的预期范围的 f 值

N	$S_{\min}=f_1\sigma$	$S_{\max}=f_2\sigma$	$\bar{X} = \pm f_3S$
	f_1	f_2	f_3
5	0.35	1.67	1.24
7	0.45	1.55	0.92
15	0.63	1.37	0.55
31	0.75	1.25	0.37
61	0.82	1.18	0.26
121	0.87	1.13	0.18

例：实验室操作的可重复性 CV_L ，由 5 个试样来测定，它们来自包含有残留物的均质性样本，平均残留是 0.75mg/kg，标准偏差是 0.2mg/kg。该样本的真实残留值期望在 $0.75 \pm 1.24 \times 0.2 = 0.75 \pm 0.248 \text{mg/kg}$ 之间；而测定结果的真实不确定度，按 95% 的置信度在 0.069 6（ 0.2×0.35 ）与 0.334（ 0.2×1.67 ）之间。

在表 4 中给出，标准不确定度的指导值，它们是基于大量数据并可用于评判实验室内所估的不确定度的合理性，用以避免不合理的偏高与偏低的值。

⁵ Ambrus A and Soboleva E. (Contribution of sampling to the variability of residue data), JAOAC. 87, 1368-1379, 2004.

⁶ Soboleva E., Ambrus A., Jarju O. (Estimation of uncertainty of analytical results based on multiple peaks), J. Chromatogr. A. 1029. 2004, 161-166

表 4. 在农药残留采样与分析的主要阶段中有代表性的期待不确定度

阶段	相对不确定度	注释
植物源产品的采样 从批中随机取出的份样对在集样过程中平均残留的变异性的影响，未将后续步骤的误差计入。	中等和较小的样品. (样本大小 ≥ 10) ^a : 26-30% ^b 较小的样品(样本大小 ≥ 5) ^a : 36~40% ^b	对依从 MRLs 试验，抽样不确定度限定为 0，就如同对散料样本平均残留 MRLs 所限定的那样。
动物产品的采样	n 个超标样品至少有一个被检出的概率 (β_t)，单个超标样品被检出的概率 (β_p)，它们之间的关系式 a: $1-\beta_t = (1-\beta_p)^n$	初级样本应从整批中随机抽取。
样本处理 包括为使分析样本和再抽样所得的均质化而进行的物理操作，但不包括分析物的分解与蒸发。	样本的基质与仪器设备会引起较大的变化，不能给出代表性的值。分析者应试着保持它 c 低于 8%~10%。	它可能受用来碎化/均质化样品的设备与样本基质的影响，但它跟分析物无关。
分析 包括从添加开始对试样进行分析的全过程	实验室内的可重复性：浓度为 $1\mu\text{g/kg}$ 至 1 mg/kg ^c ，16-53%。 实验室间平均可重复性在： $0.001\sim 10\text{ mg/kg}$: 25% ^d 之间。	不同日期测定多种多样的农药—产品组合的回收率实验中，可容易测定代表性 CV _A 值。

注:

- (a) 符合 MRL 的农药残留检测推荐采样方法 (CAC/GL 38-1999).
- (b) Ambrus A. Soboleva E. 采样对残留数据变异的影响, JAOAC, 87, 1368-1379, 2004;
- (c) 残留分析实验室良好习惯做法的原则 (CAC/GL 40-1993, Rev. 1-2003)
- (d) Alder L., Korth W., Patey A., van der Schee and Schoeneweis S., 农药残留分析中的不确定度评估, J. AOAC International, 84, 1569-1578, 2001

此外，对由单个实验室评估得出的不确定度，基于实验室间的再现性值，主管部门与其他的风险管理者可以确定一个默认的扩展不确定度值，来判断是否符合基于实验室间的重复性的依从 MRLs（见第 5 节）。比如，对 CV_L，一个 50% 扩展的不确定度，可考虑作为一个合理的默认值。

5 不确定度信息的使用

如有必要，报告结果的同时应给出扩展不确定度，U，如下

$$\text{结果} = x \pm U (\text{单位})$$

扩展不确定度 U，如 EURACHEM 推荐，可以通过标准组合不确定度(S_{Res})以及一个包含因子 2，或者通过有效自由度小于 20，符合要求的置信水平（通常为 95%）上的司徒顿（student）t 值计算得到，同时。扩展不确定度可按如下公式分别计算

$$U = 2S_{\text{Res}} \quad \text{或} \quad U = t_{v,0.95}S_{\text{Res}} \quad (3)$$

所报告结果的具体数值，应按最后 1 位数字为不确定值的一般规则处理，有效数字的修约只在最终结果时进行，因为在计算阶段就进行修约，会引起计算结果的不必要的误差。

为说明问题，可将对一个样本的残留的最佳估计的有关内容作出报告，至于对这结果如何解释将依赖于检验的目的。比如，检验是否符合国家 MRL 标准，或者对出口商品来讲是否符合 Codex MRL 标准。

5.1 符合 MRL 的测试

图 1 表示残留测定值， 不确定度与 MRL 的关系。

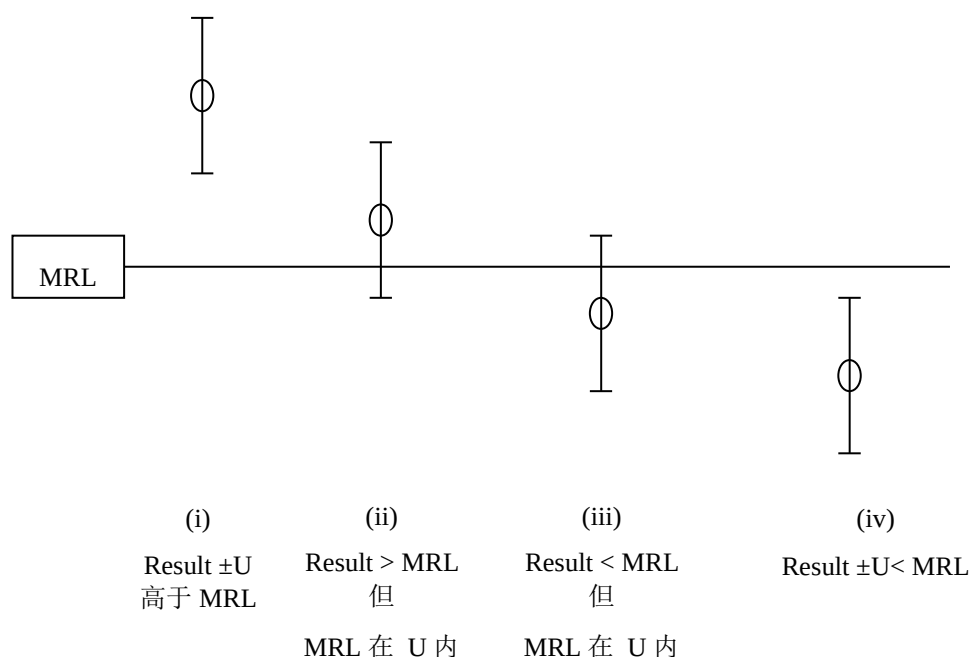


图 1. 测量值、期望不确定度和 MRL 之间的关系

情况 (i)

分析结果以扩展不确定度为限，大于 MRL。结果表示取样批次的残留大于 MRL。

情况(ii)

分析结果大于 MRL，测量不确定度下限值小于 MRL。

情况(iii)

分析结果小于 MRL，测量不确定度上限值大于 MRL。情况(iv)

情况(iv)

分析结果以扩展不确定度为限，小于 MRL。

5.2 决定环境

图 1 中解释的情况是有关植物源样品的。对于畜产品的残留与 MRLs 的符合度， 应由基于分布自由统计量的抽样方案来判定， 相应的例子见于文件《测定农药残留与 MRLs 符合度的推荐采样方法》(CAC/GL 33-1999) 中。

因为每个样本中的残留都应符合 MRL， 当然每个样本的最小样本量与样本的量应按《采样方法导则》(Codex Sampling Procedure) 规定实行， 所以， 扩展的不确定度应采用 $U=KSL$ ， 其中 $SL=CV_L \times \text{残留}$ 。

对情况 (i) 该做何决策是清楚的， 在对本地产品或进口产品检验其与国家 MRL 的符合性时， 为避免分析报告中对不确定度的过多解释， 实验室可报告结果如下： 样本中残留量 “不低于 ‘X-U’”。这满足 MRL 被超出的要求， 除非对有关测定不确定度有任何合理的怀疑。

对情况 (iv) 显然该样本符合 MRL。

对情况 (ii) 与 (iii) 不能下结论说样品残留超过了 MRL 或者毫无疑问地符合 MRL 了， 决策者的行动需作进一步的考虑， 讨论如下： 情况 (ii) 与 (iii) 的含意依赖于国家的实际状况以及可能对商贸货物的接受性有多大的影响。对检验结果如情况 (ii) 与 (iii) 的产品， 不论在国内市场还是国际贸易中都须谨慎处置。例如， 对出口产品的认证， 遇到残留结果如情况 (ii) 与 (iii) 的， 将不适宜出口； 对国家的进口产品， 若残留水平如情况 (ii)， 那么， 在一适宜的置信水平下， 认为它不符合 MRL； 而对情况 (iii)， 进口方通常就放行。

正文使用的术语表^a

空白 (样本, 试剂)	(i) 物质 (样品或样品一部分或样品的提取物) 已知不含有可检出的被检测物, 也已知为一基质空白。 (ii) 在不使用任何样品的情况下, 仅用溶剂与试剂进行的完全分析 (用水可替代样品, 来实现该分析), 也可作为已知试剂空白或程序空白。
组合标准不确定度	对测定结果 y , 总不确定度 $U_c(y)$ 是一估计的标准偏差, 它等于总方差的正的平方根, 这总方差是组合所有不确定度分量按不确定度传递 (误差传递定律) 而获得。
污染	在抽样或分析任一步骤中, 通过任一途径, 其他分析物进入样品、提取物、标准液等。
残留定义	指农药及它的代谢物、衍生物及相关化合物的总称, 用于 MRL 的实施或对食物暴露量评估。
测定系统	用来检出与测定分析物的浓度或含量的任一系统, 例如: GC-FPD, LC-MS/MS, LC 一柱后衍生, ELISA, TLC-比色法或生物测定法。
水平	在此文件中, 指浓度 (如 mg/kg, µg/ml) 或质量 (如 ng, pg)。
批次	同时提交的一定量的食品, 并由抽样方已知或认定, 它们具有统一特点, 如产地, 生产者, 品种, 包装, 包装类型, 标志物, 代销商等。
基质效应	在分析物浓度或含量的测定中, 样品中的一个或多个未检出成份的影响, 某些测定系统 (如 GC, LCMS, ELISA) 对某分析物的响应, 可能受样品中的附属提取物的影响。
程序空白	见空白
试剂空白	见空白
响应值	当分析物出现时, 从检测仪器发出的绝对或相对输出信号。
添加	为测定回收率, 分析物的注入或标准品的注入。
标准不确定度	表示为某一不确定组分的标准偏差。
单元 (作为样品的一部分)	单个水果、蔬菜、动物、谷物籽粒、罐头等, 例如。一个苹果、一牛排、麦粒、一番茄罐头。
残留超标	残留超出了 MRL 或由任何其他理由来认定为不合法的。

注(a).这里给出的定义皆基于参考文献^{7,8,9,10}。附加的定义是由残留分析 GLP 上 GLs 的修正结论给出。

⁷ EURACHEM (2000) EURACHEM/CITAC 分析法不确定度定量化指导 第2版。 <http://www.measurementuncertainty.org>

⁸ 国际食品法典。符合 MRL 的农药残留测定推荐采样方法 ftp://ftp.fao.org/codex/standard/en/cxg_033e.pdf

⁹ Willetts P, Wood R (1998) Accred Qual Assur 3: 231-236.

¹⁰ 度量衡中基础和一般术语的国际词汇表, 日内瓦, 1993

附件

指导性解释

在指南文件 CAC/GL 59-2006 中提到, 通过 ISO/IEC 17025 认证的实验室, 以及满足良好实验室规范 (GLP) 的残留分析试验室中都要求对分析数据的测量不确定度进行评估 (MU)。对于符合规定的食品中化学残留和污染, 不管是国内还是国际标准化中, 对于分析测试的物品或货物的测试结果报告中都需要考虑到相关的测量不确定度。

尽管实验室之间采用非常相似的分析测试方法, 但报告的测量不确定度值却会有很大的不同, 这种情况也很常见。这说明测量不确定度的评估在一些食品分析试验室中依然是一个发展中的科学概念。本附件的目的是描述一些实验室在测量不确定性的评估, 特别是内部使用方法验证、质量控制和多残留分析方法中长期精度数据时可能采用的一些选择。还可以在农药残留分析结果的测量不确定度评估时预见更协调的方法, 从而减少在法规决策和 MRL 制定时可能产生的纠纷。

有两种实验室中常用的测量不确定度的评估方法: GUM (测量不确定度指南) 法或“自下而上法”, 和分析精密度和误差来评估的“自下而上法”。

GUM 法对分析过程的每个单元都进行严格的分析, 并且对每一步骤中的系统误差和随机误差都进行评估。这一过程最初相当费力, 需要分析者有或者开发一个详细的了解分析步骤的过程和识别关键控制点的方法。除非所有这一过程中所有步骤都考虑进去, 才有可能评估测量不确定度。另一方面来讲, 如果操作者的误差由于忽略而被取消, 会提供一个高估的不确定度。一般来讲, 自下而上法比起分析化学来更适合于物理计量学, 尤其适用于更复杂的农药多残留分析方法。

自上而下法的支持者指出, 从内部实验室收集的数据验证, 长期精密度和分析质量控制(QC)可为测量不确定度的评估提供更可靠的信息。某些时候, 能力测试 (PT) 的数据也可以用来评估测量不确定度, 且作为唯一依据估计或经常结合内部数据。能力测试研究的内部实验室间重现性数据也可以为单一实验室评估提供一个有用的“基准”。

测量不确定度的评估应该考虑所有的选项。最初的目标应该是使用现有的信息获得最好的估计。最初的实验室评估应该通过与替代方法、文献报告、以及能力测试研究的比较中来验证。此外, 专业的判断也是评估和确证测量不确定度的重要因素。评估应该检查更精密的可用数据, 比如在项目过程定期分析中生成的内部批次质量控制数据。

本附件主要关注于测量不确定度评估的“自上而下法”, 以不同来源获得的数据为基础。

应用一个在食品中农药残留分析中测量不确定度评估的默认值

欧盟成员国对于进入欧盟的食品中农药残留采用了测量不确定的默认值为 $\pm 50\%$ 。默认值是基于一系列欧盟成员之间能力验证研究涉及的主管残留实验室参与的对水果和蔬菜中农药多残留研究的统计结果。这些研究报告的平均相对标准偏差介于 20%~25%之间, 提供了一个接近 50%的测量不确定度值。

在缺乏其他统计数据时, 符合欧盟 MRL 规定的实验室检测食品可以采用默认值 50%, 可以通过参与欧盟或类似的能力验证研究来建立其分析能力, 或者也可以用可以接受的相关的长期精度和测量结果误差来证明。然而, 从长远来看, 现任的实验室验证默认值的采用应该由内部精密度和验证数据来独立评估测量不确定度。

使用由 Horwitz 关系来源的精密数据

如果特定的方法中没有实验室间数据和重现性标准偏差, 则测量不确定度可以用 Horwitz 方程来表示, 其与分析物浓度标准偏差相关。Horwitz 关系变异系数(CV)和分析物浓度的结果是基于大量的食品协作研究的文献报道。Horwitz 方程在比较实验室内部测量不确定度评估和实验室间出版研究的期望值时也是一个有用的工具。

来源于实验室间研究的精密数据 (协作研究和能力验证研究)

实验室间研究结果受到不精密度和偏差的共同制约。如果这些研究涉及到足够数量的实验室和旨在涵盖实际测试条件(分析物和基质的范围), 获得的重复性标准偏差则反映了在实践中可能会遇到的典型的错误。因此, 能力验证研究的数据将被用来提供合理的评估测量不确定度。

协作研究方法通常对分析步骤有明确的文件说明, 通常只涉及专业的及经验丰富的残留分析实验室。在这些条件下, 当方法应用在重复性条件下时, 方差分析可能是最好的可行的应用方法, 尤其是当误差来源于样本本身且可以忽略时。提供一个有能力可以达到特定的协作研究实验室的分析性能的实验室, 研究获得的重复性标准偏差将为评估测量不确定度提供一个很好的基础。然而, 有能力的实验室在实验室间重复性条件下应该提高实验室间方法的精密度, 从而降低测量不确定度。

如果有证参考物质(CRMs)应用在协作研究中, 研究报告应该提供一个对方法参考值的偏差估计, 评估测量不确定度时也应同时考虑。

在能力验证研究中, 通常情况下实验室分析采用自己的方法, 可以是一个标准方法、或者优化的标准方法、或者开发经过实验室确证的方法。此外, 在这种情况下参与协作研究的实验室的分析性能有很大不同。考虑到这些因素, 能力验证研究的重复性标准偏差将有可能比参与方法协作研究的要大。基于此基础上的测量不确定度也要比多个参与实验室报告的要大。不过, 基于采用不同方法的一系列参与的专业实验室而评估的测量不确定度会比在国际贸易中判断食品中农药残留是否符合规定而更实际和有用。欧盟成员国采用的测量不确定度 50% 的默认值即是基于一系列农药和食品基质的能力验证数据中得来的。

不管实验室评估测量不确定度时使不使用能力验证的数据, 在比较和确证像实验室内确证或质量控制实验中, 这些数据都可以提供有用的信息。

由实验室内部确证和质量控制数据派生的测量不确定度

化学计量人员的共识是分析过程中不确定性数据的最佳来源的来源是实验室的方法确认和/或验证研究和长期质量控制数据。这是基于假设实验室对合适的质量控制(QC)样本、有证参考物质(CRMs)或基质添加均进行了确认和/或验证研究, 并且有足够的经验来建立长期误差和重复性数据。

有证参考物质在食品基质中农药残留中有限的可用性通常需要实验室专注于添加样品或其他适当的特征样本作为内部质量控制。基质匹配质控样本如产生的残留样品、能力验证的剩余样品、或添加样品的使用就可以为实验室提供在监测和控制方法(和分析物)时同时收集偏差和精密度的能力。控制图在评估长期精密度和监测统计分析过程中是一个很好的工具。

在评估测量不确定度时, 显著的偏差和不确定度偏差也应该考虑进去, 如 5.4 小节中的示例讨论。

偏差最好是由有证标准物质来测定。然而考虑到食品中的农药以及多残留监测中大量的农药时有证标准物质的缺乏, 通常需要依靠基质添加样本的回收率来提供方法偏差的信息。

能力验证研究实验室的性能可以由参与的独立实验室对公认值的偏差指示来进一步提供, 在某些情况下, 可以是能力验证样品的添加水平。然而, 偏差在应用到测量不确定度的评估时, 应该由一定数量的能力验证实验室验证或者以此为基础。

工作示例

下列的工作示例采用的评估测量不确定度的步骤就建立在内部实验室数据的不同组合、内部实验室数据精密度和实验室间数据的基础上。Horwitz 方程和能力验证研究的结果为实验室内部不确定度评估进一步提供了有效的基准。

下面的示例采用典型农药残留毒死蜱作为假设数据, 大量的样本见 Eurolab Technical report No 1/2007 和 the Nordtest Report TR537。

1 用 Horwitz 方程评估测量不确定度

Horwitz 方程表示的是分析物浓度的重复性标准偏差。

$$u' = 2^{1-0.5 \log c}$$

注: u' = 重复性相对标准偏差

c = 分析物浓度(in g/g)

相对扩展不确定度, U' (95% 置信水平) 由下式计算: $U' = 2u'$

由于 Horwitz 方程是一个分析物浓度的函数, 它将提供一系列取决于农药浓度的不确定度值, 如下表所示:

浓度 (mg/kg)	u' (%)	U' (%)
1.0	16	32
0.1	22.6	45
0.01	32	64

示例 1:

毒死蜱在实验室番茄样本中的浓度为 0.40 mg/kg, 按 Horwitz 方程计算浓度为 0.40 mg/kg 时, 重复性相对标准偏差为 18.4%。即 $u' = 18.4\%$, $U' = 2u' = 37\%$, 因此实验室报告的结果即为 0.40 ± 0.15 mg/kg。

实验室报告应该指出不确定度为置信水平为 95%，包含因子为 2 时的扩展不确定度值。除非另有说明，这是扩展不确定度报告的普遍结果。

假如缺少数据支持，Horwitz 方程 仅可以作为可能的测试指标和结果不确定度时应谨慎使用。分析方法的进步，特别是仪器技术可以在非常低的检测定量限时提供比 Horwitz 方程更小的不确定度值。Thompson 和 Lowthian 报道，实验室往往在低浓度时不采用 Horwitz 函数。然而值得注意的是 Thompson 的概念限制了浓度低于 0.1mg/kg 时， u' 的最大值为 22%。

2 利用不确定度默认值 50%来评估测量不确定度

在采用测量不确定度默认值之前，实验室应该保证在日常检测中不确定度不高于此值。

示例 2:

实验室测定番茄样本中毒死蜱含量为 0.40 mg/kg。采用公认的不确定度默认值 $\pm 50\%$ 来计算结果。因此，实验室报告结果应为 0.40 ± 0.20 mg/kg。

3 基于实验室内质量控制和能力验证研究数据的评估测量不确定度

1) 使用能力验证研究中的定义值或公认值

$$U' = 2u' \quad \text{公式 1}$$

$$u' = \sqrt{u'(R_w)^2 + u'(\text{bias})^2} \quad \text{公式 2}$$

此处 U' = 相对扩展不确定度

u' = 相对合成标准不确定度

$u'(R_w)$ = 实验室间相对标准不确定度(实验室间重复性相对标准偏差)

$u'(\text{bias})$ = 由偏差引起的相对标准不确定度

示例 3:

在此例中， $u'(R_w)$ 由实验室内部质量控制数据得到，最好是长期质控数据。 $u'(\text{bias})$ 由能力验证数据估算。

番茄中毒死蜱的实验室结果 = 0.40 mg/kg，毒死蜱添加浓度为 0.5 mg/kg 的质控样本批次间相对标准偏差为(每周一个添加样本，连续添加 3 个月) = 15%。该实验室参与了六项能力验证研究，分析物包括了不同蔬菜和水果基质中毒死蜱检测。对于这些研究，实验室结果与定义值之间的相对偏差为-15%, 5%, -2%, 7%, -20% 和-12%，是能力验证研究的 16 个参与实验室的平均值。在这 6 项研究中毒死蜱的平均重复性相对标准偏差(SR)为 25%。

$$u'(\text{bias}) = \sqrt{\text{RMS}'\text{bias}^2 + u'(\text{C ref})^2} \quad \text{公式 3}$$

其中 $\text{RMS}'\text{bias}$ = 相对标准偏差的均方根

$u'(\text{C ref})$ = 毒死蜱在六个研究中定义值的平均相对不确定度

$$\text{RMS}'\text{bias} = \sqrt{\frac{\sum (\text{bias})^2}{n}} \quad (n = \text{能力验证研究的数目}) \quad \text{公式 4}$$

$$= \sqrt{\frac{(-15)^2 + (5)^2 + (-2)^2 + (7)^2 + (-20)^2 + (-12)^2}{6}}$$

$$= 11.9\%$$

$$u'(\text{C ref}) = \frac{S_R}{\sqrt{m}} \quad \text{公式 5}$$

其中 S_R = 毒死蜱在六个实验中的平均相对标准偏差

$$m = \frac{25}{\sqrt{16}} = 6.3\%$$

$$\text{所以, } u'(\text{bias}) = \sqrt{(11.9)^2 + (6.3)^2} = 13.5\%$$

$$\text{从公示 2 可得, } u' = \sqrt{(15)^2 + (13.5)^2} = 20\%$$

从公示 1 可得, 相对扩展不确定度(置信区间 95%) = 40%。

实验室报告结果应为 $0.40 \pm 0.16 \text{ mg/kg}$ 。

注1 RMS' bias 值包括偏差和不确定度偏差。

注2 计算测量不确定的最佳估计是由于不同基质和不同毒死蜱浓度的能力验证数据。

注3 如果可能的话, 测量不确定度的计算应该建立在达到或接近最临界浓度上, 比如说 Codex 的限量值 MRL。

2) 有证参考物质的能力验证研究(CRMs)

如果一个合适的有证参考物质含有毒死蜱, 在能力验证研究样本中分布均匀, 那么就不需要计算 $u'(C_{\text{ref}})$ 。在这种情况下, $u'(C_{\text{ref}})$ 应该是参考物质表明的不确定度转换成相对标准偏差。

例如, 如果在置信区间为 95%时, 对于毒死蜱在有证参考物质的参考值为 $0.489 \pm 0.031 \text{ mg/kg}$, 那么:

$$u(C_{\text{ref}}) \text{ (标准偏差)} = \frac{0.031}{2} = 0.0155 \text{ mg/kg}$$

$$u'(C_{\text{ref}}) \text{ (相对标准偏差)} = \frac{0.0155 \times 100}{0.489} = 3.17\%$$

在某些不太可能的情况下, 含有毒死蜱的有证参考物质分布在不同的能力验证研究中, 那么 $u(C_{\text{ref}})$ 值就用来计算 U。在两种情况下, RMS' bias 都需要用公式 4 计算。

示例 4:

Study No.	CRM	相对偏差	$u'(C_{\text{ref}})$
1	A	-12%	2.3%
2	B	-15%	1.7%
3	C	-3%	2.0%
4	C	5%	2.0%
5	C	-20%	2.0%
6	A	0%	2.3%

平均 $u'(C_{\text{ref}}) = 2.05\%$

由公式 4 得到 $\text{RMS}'_{\text{bias}} = 11.6\%$

由公式 3 得到 $u'(\text{bias}) = 11.8\%$

注 4 与有证参考物质相关的相对不确定性可能小于与定义值或者公认值。如果由分析不精密度造成的实验室的相对标准不确定度 $u'(\text{RW})$ 是相同的, 如 15%, 那么从公式 1 和 2 可以得到

$$u' = 19\%$$

$$U' = 38\%$$

实验室报告的结果应该是 $0.40 \pm 0.15 \text{ mg/kg}$ 。

4 使用实验室间质控数据来评估测量不确定度

示例 5:

- 番茄中毒死蜱的实验室结果为 0.40 mg/kg。
- 制备标准添加溶液的校正液纯度为 $95 \pm 2\%$ (分析纯)。
- 3 个月内毒死蜱添加浓度为 0.5mg/kg 的批次质控样本的 14 个回收率 (%) 值为: 90, 100, 87, 89, 91, 79, 75, 65, 80, 82, 115, 110, 65, 73, 平均回收率为 86%, 相对标准偏差为 15%。

假设参考物质的不确定度为扩展不确定度 U (置信区间为 95%)

$$u'(C_{\text{ref}}) = \frac{2}{2} = 1\%$$

注5 假设添加溶液和添加的番茄样本的不确定度差异是不显著的, 很有可能是这种情况, 如果不是的话,

$u'(C_{\text{ref}})$ 对整体不确定度的贡献将是非常小的。

$u'(RW) = 15\%$ (实验室间重复性相对标准偏差)

由公式 4, 考虑到回收率 100 - % 的偏差,

$$RMS'_{\text{bias}} = 20\%$$

由公式 3 可得, $u'(\text{bias}) = 20\%$

由公式 2 可得, $u' = 25\%$

由公式 1 可得, $U' = 50\%$

实验室报告结果应为 $0.40 \pm 0.20 \text{ mg/kg}$ 。

注6 这个结果的不确定度没有校正回收率。如果 3 个月内分析的平均回收率的结果在分析步骤的最后做了校正, 则 $u'(\text{bias})$ 值只反应了平均回收率的不确定度。那么 $u'(\text{bias})$ 值可以用回收因子的相对标准不确定度 (平均回收率的相对不确定度) 和添加浓度的相对标准不确定度 $u'(C_{\text{ref}})$ 表示。

$$\text{平均回收率的相对标准标准不确定度 } u'_{\text{Rec}} = \frac{u'(R_w)}{\sqrt{n}} \quad \text{公式 6}$$

其中 n = 平均回收率计算时的重复次数

$$u'_{\text{Rec}} = \frac{15}{\sqrt{14}} = 4\%$$

$$u'(\text{bias}) = \sqrt{u'(\text{Rec})^2 + u'(C_{\text{ref}})^2} \quad \text{公式 7}$$

$$\text{因此 } u'(\text{bias}) = \sqrt{(4)^2 + (1)^2} = 4.1\%$$

这样, 从公式 2 和 1, $u'(RW)$ 值为 15% 时计算

$$u' = 15.5\%, \quad U' = 31\%$$

如果用回收率结果校正, 则结果报告应为

$$0.40 \pm 0.12 \text{ mg/kg}$$

注7 该例表明, 如果对回收率的校正是基于分析过程中的 9 个或多个重复试验, 并且使用的有证参考物质的纯度有很高的不确定度, 则测量不确定度的合理评估仅仅应从实验室间重复性标准偏差来计算。