



中华人民共和国国家标准

GB/T 15000.8—2023/ISO Guide 33:2015

代替 GB/T 15000.8—2003, GB/T 15000.9—2004

标准样品工作导则 第 8 部分：标准样品的使用

Directives for the work of reference materials—
Part 8: Good practice in using reference materials

(ISO Guide 33:2015, Reference materials—
Good practice in using reference materials, IDT)

2023-09-07 发布

2024-01-01 实施

国家市场监督管理总局
国家标准化管理委员会 发布

目 次

前言	III
引言	V
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 符号	2
5 约定	3
6 RM 及其在测量中的作用	3
6.1 RM 的一般应用	3
6.2 特性值	4
6.3 不确定度陈述	5
6.4 溯源性陈述	6
7 RM 和 CRM 的处置	6
8 精密度的评估	7
8.1 概述	7
8.2 重复测量次数	7
8.3 对 RM 的要求	8
8.4 测量	8
8.5 数据处理	8
8.6 精密度的计算和评估	8
9 偏倚的评估	9
9.1 概述	9
9.2 偏倚核查方法	9
9.3 偏倚数据应用	10
10 校准	10
10.1 概述	10
10.2 建立计量溯源性	10
10.3 校准模型	11
11 为其他材料赋值	11
11.1 概述	11
11.2 纯物质	12
11.3 称量法和容量法	12
12 约定标尺	13

12.1 概述 13

12.2 pH 标尺 13

12.3 辛烷值 13

13 CRM 和 RM 的选择 14

13.1 概述 14

13.2 CRM 的选择 14

13.3 RM 的选择 16

13.4 与测量系统的相关性 16

附录 A (资料性) 标准样品常规应用中的关键特性 17

附录 B (资料性) 校准模型和相关联不确定度模型 18

附录 C (资料性) 决策错误 20

参考文献 21

前 言

本文件按照 GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

本文件是 GB/T 15000《标准样品工作导则》的第8部分。GB/T 15000 已经发布了以下部分：

- 第1部分：在技术标准中陈述标准样品的一般规定；
- 第2部分：常用术语及定义；
- 第3部分：标准样品 定值和均匀性与稳定性评估；
- 第4部分：证书、标签和附带文件的内容；
- 第5部分：质量控制样品的内部研制；
- 第6部分：标准样品包装通则；
- 第7部分：标准样品生产者能力的通用要求；
- 第8部分：标准样品的使用。

本文件代替 GB/T 15000.8—2003《标准样品工作导则(8) 有证标准样品的使用》和 GB/T 15000.9—2004《标准样品工作导则(9) 分析化学中的校准和有证标准样品的使用》，与 GB/T 15000.8—2003 和 GB/T 15000.9—2004 相比，除结构调整和编辑性改动外，主要技术变化如下：

- 更改了标准的范围(见第1章, GB/T 15000.8—2003 第1章, GB/T 15000.9—2004 第1章)；
- 增加了术语及定义“特性值”“标准值”“信息值”“校准样品”和“质量控制样品”(见 3.3~3.7)；
- 删除了术语及定义“测量过程”“影响量”“精密度”“重复性”“重复性条件”“重复性标准差”“重复性限 r ”“再现性(复现性)”“再现性条件”“再现性限 R ”“偏倚”“准确度”“正确度”“不确定度”“估计”“估计值”“估计量”和“原假设”(见 GB/T 15000.8—2003 的 2.1、2.2、2.5~2.20)；
- 将“统计原理”更改为“约定”(见第5章, GB/T 15000.8—2003 第4章)；
- 将“RM 及其在测量中的作用”更改为“标准样品及其在测量中的作用”(见第6章, GB/T 15000.8—2003 第5章)；
- 增加了“RM 和 CRM 的处置”(见第7章)；
- 将“测量过程的评定”更改为“精密度的评估”和“偏倚的评估”(见第8章、第9章, GB/T 15000.8—2003 第6章)；
- 增加了“校准”“为其他材料赋值”和“约定标尺”(见第10章、第11章、第12章)。

本文件等同采用 ISO Guide 33:2015《标准样品 标准样品的使用指南》。文件类型由 ISO 的指南文件调整为我国的国家标准。

本文件做了下列最小限度的编辑性改动：

- 为与现有标准保持一致，将标准名称修改为《标准样品工作导则 第8部分：标准样品的使用》；
- 结合国际标准的情况及国内标准情况增加了译者注。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由全国标准样品技术委员会(SAC/TC 118)提出并归口。

本文件起草单位：中国标准化协会、广东美的厨房电器制造有限公司、天津东方绿色科技发展有限公司、建筑材料工业技术监督研究中心、山东省分析测试中心、中标能效科技(青岛)有限公司、北京华塑晨光科技有限责任公司。

本文件主要起草人：徐大军、吴忠祥、郑江、胡晓燕、邢小茹、唐本玲、何平、王向红、安子怡、陈志永、

王桓、曹际娟、石雨婷、王晓、陈剑、陈宏愿。

本文件及其所代替文件的历次版本发布情况为：

- GB/T 15000.8, 2003 年首次发布；
- GB/T 15000.9, 2004 年首次发布；
- 本次为第一次修订。

引 言

GB/T 15000《标准样品工作导则》是指导我国标准样品工作的基础性和通用性标准。GB/T 15000旨在确立普遍适用于标准样品生产者制备、定值、包装储存、分发和应用工作的准则,由8个部分构成。

- 第1部分:在技术标准中陈述标准样品的一般规定。目的在于规定技术标准制定中陈述标准样品的一般要求、标准样品与技术标准的关系、选择标准样品的技术准则等。
- 第2部分:常用术语及定义。目的在于规定与标准样品有关的术语和定义,特别是用于标准样品产品说明书、证书和相应的研制报告中的术语。
- 第3部分:标准样品 定值和均匀性与稳定性评估。目的在于就技术问题提供更具体的指导,并提供标准样品均匀性、稳定性评估和定值过程的基本模型。
- 第4部分:证书、标签和附带文件的内容。目的在于确立适用于编写标准样品证书、标签和附带文件内容时需要遵守的相关要求。
- 第5部分:质量控制样品的内部研制。目的在于规范和指导质量控制样品的研制和使用。
- 第6部分:标准样品包装通则。目的在于确立适用于标准样品包装和储存运输方面的基本规则。
- 第7部分:标准样品生产者能力的通用要求。目的在于规定有关组织管理和运行要求,以确保标准样品生产者能够生产出符合要求的标准样品。
- 第8部分:标准样品的使用。目的在于为标准样品的正确使用提供指导。

本文件以实例形式给出使用建议,在某种意义上也反映了有关RM的复杂程度。其详细程度有助于任何在实验室中对质量管理负有责任的人员(如规程、作业指导书、标准操作程序及类似文件的起草者、审核者、管理者和评审者等)对本文件的理解和标准样品的使用。

标准样品工作导则

第 8 部分：标准样品的使用

1 范围

1.1 本文件描述了在测量过程中标准样品(RM),特别是有证标准样品(CRM)的规范使用,主要包括测量方法精密度和正确度评估、质量控制、给其他材料赋值、校准和建立约定标尺。本文件也涉及不同类型 RM 关键特性的各种应用。

1.2 CRM 的特性值具有计量溯源性,可溯源至国际标尺或其他测量标准。CRM 之外的 RM 特性值虽然通常未建立这种溯源性,但此类 RM 仍可用于测量程序的部分评估,包括评估不同水平的精密度。

1.3 RM 的主要应用包括精密度控制(第 8 章)、偏倚评估(第 9 章)、校准(第 10 章)、校准 RM 制备(第 11 章)和维持约定标尺(第 12 章)。

注：不是所有类型的 RM 都能够用于上述所有应用。

1.4 ISO Guide 34¹⁾和 ISO Guide 35^[2]也涉及校准标准样品的制备。本文件仅适用于诸如实验室设备校准用少量 RM 的制备和赋值,不适用于以分发为目的的规模化 RM 的生产,后者已包含在 ISO Guide 34 和 ISO Guide 35 中。

1.5 本文件未涉及诸如天然气分析、临床化学和制药行业工作标准的研制。此类活动已包含在 ISO Guide 34 和 ISO Guide 35 中。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中,注日期的引用文件,仅该日期对应的版本适用于本文件;不注日期的引用文件,其最新版本(包括所有的修改单)适用于本文件。

ISO Guide 30 标准样品 常用术语及定义(Reference materials—Selected terms and definitions)

注：GB/T 15000.2—2019 标准样品工作导则 第 2 部分：常用术语及定义(ISO Guide 30:2015, IDT)

ISO/IEC Guide 98-3:2008 测量的不确定度 第 3 部分：不确定度表述(GUM)[Uncertainty of measurement—Part 3: Guide to the expression of uncertainty in measurement (GUM)]

注：GB/T 27418—2017 测量不确定度评定和表示(ISO/IEC Guide 98-3:2008, MOD)

ISO/IEC Guide 99 国际计量学词汇 基础概念和相关术语(VIM)[International vocabulary of metrology—Basic and general concepts and associated terms (VIM)]

注：本文件中“测量不确定度表述指南”简称“GUM”,“国际计量学词汇 基础概念和相关术语”简称“VIM”。

3 术语和定义

ISO/IEC Guide 98-3, ISO/IEC Guide 99 和 ISO Guide 30 中界定的以及下列术语和定义适用于本文件。

注：更多术语可通过 ISO 在线浏览平台 <https://www.iso.org/OBP/ui/> 获得。

1) ISO Guide 34:2000 被 ISO 17034:2016 替代,原指南文件废止。ISO 17034:2016 已等同转化为 GB/T 15000.7—2021。

3.1

标准样品 reference material; RM

具有一种或多种规定特性足够均匀且稳定的材料,已被确定其符合测量过程的预期用途。

注 1: 标准样品²⁾是一个通用术语。

注 2: 特性可以是定量的或定性的(例如:物质或物种的特征)。

注 3: 用途可包括测量系统的校准、测量程序的评估、给其他材料赋值和质量控制。

注 4: ISO/IEC Guide 99:2007 有类似的定义(5.13),但限定“测量”术语仅用于定量的值。然而,ISO/IEC Guide 99:2007(5.13)的注 3 中明确包括定性特性,称作“标称特性”。

[来源:ISO Guide 30:2015,2.1.1]

3.2

有证标准样品 certified reference material; CRM

采用计量学上有效程序测定的一种或多种规定特性的标准样品,并附有证书提供规定特性值及其不确定度和计量溯源性的陈述。

注 1: 值的概念包括标称特性或定性属性,如特征或序列,该特性的不确定度可用概率或置信水平表示。

注 2: 标准样品生产和认定所采用的计量学上有效程序已在 ISO Guide 34 和 ISO Guide 35 中给出。

注 3: ISO Guide 31^[17]给出了证书内容的编写要求。

注 4: ISO/IEC Guide 99:2007(5.14)有类似的定义。

[来源:ISO Guide 30:2015,2.1.2,有修改]

3.3

特性值 property value

〈标准样品的特性值〉表示标准样品的物理、化学或生物所对应特性量的值。

[来源 ISO Guide 30:2015,2.2.1]

3.4

标准值 certified value

赋予标准样品特性的值,其具有不确定度和计量溯源性陈述,并在标准样品证书中标明。

[来源:ISO Guide 30:2015,2.2.3]

3.5

信息值 indicative value; information value; informative value

标准样品的量或特性的值,其仅用于提供信息。

注: 信息值不能作为计量溯源链的参考标准。

[来源 ISO Guide 30:2015,2.2.4]

3.6

校准样品 calibrant

用于设备或测量程序校准的标准样品。

[来源:ISO Guide 30:2015,2.1.21]

3.7

质量控制样品 quality control material

用于测量质量控制的标准样品。

[来源:ISO Guide 30:2015,2.1.22]

4 符号

下列符号适用于本文件。

α 第一种类型错误(第 I 类错误)的风险

2) reference material 也译为参考物质、标准物质。

β	第二种类型错误(第Ⅱ类错误)的风险
χ^2	卡方
d	测量偏倚
k	包含因子
s_w	由重复观测计算获得的标准差
$\sigma_{w(i)}$	实验室内标准差所要求的值
$u(\)$	括号中参数的标准不确定度
$U(\)$	括号中参数的扩展不确定度
u_{CRM}	有证标准样品特性值的标准不确定度
u_{meas}	测量有证标准样品所得值的标准不确定度
u_{prep}	制备校准样品所得值的不确定度
x_{CRM}	有证标准样品中规定特性的值
x_{meas}	测量有证标准样品所得值
x_{prep}	制备校准样品所得值
$\bar{x}$	重复观测的平均值

5 约定

本文件遵守以下约定。

5.1 被测量具有唯一但不可知的“真值”。

5.2 本文件中的所有统计学方法均基于如下假定：

- a) 标准值是 CRM 特性真值的最佳估计值。
- b) 与材料(如均匀性)或测量过程相关的所有变量是随机的或遵循正态分布。本文件中规定的概率值假设为正态分布。当不是正态分布时可能导致概率不同。

5.3 本文件使用的“有证标准样品”(CRM)的概念还包括特性值具有计量溯源性或测量不确定度陈述的 RM。这些特性值假设通过 ISO Guide 34 和 ISO Guide 35 中规定的定值方式获得。

5.4 本文件使用术语 RM 时,指能够用于所描述目的的任何 RM。使用 CRM 是一种选择,但通常不是最经济的那种。大多数实际应用中,在不需要特性值、不确定度和溯源性陈述时可使用 RM。

5.5 任何没有赋予计量溯源性和测量不确定度陈述的指示值、信息值、参考值或其他形式的赋值,均不能用于要求被测量赋值的计量学应用,如校准或给其他材料赋值。尽管这些量值可以核查 RM 是否适合精密控制或其他不要求特性值的应用。

5.6 本文件使用了不确定度传播律。其他不确定度传播的方法也能够使用。在有些情况下,这些替代方法是应用环境所要求的。有关这些事项进一步的指导参见 GUM 及其附件。

6 RM 及其在测量中的作用

6.1 RM 的一般应用

6.1.1 RM 尤其是 CRM 广泛应用于如下目的：

- 设备或测量程序的校准(第 10 章)；
- 建立计量溯源性(第 9 章、第 10 章、第 11 章)；
- 方法确认(第 8 章、第 9 章)；
- 为其他材料赋值(第 11 章)；
- 测量或测量程序的质量控制(第 8 章、第 9 章)；
- 维持约定标尺(第 12 章)。

图 1 给出了包括抽样和样品制备的某项测量的流程框图,揭示了 CRM 可能起到的作用。

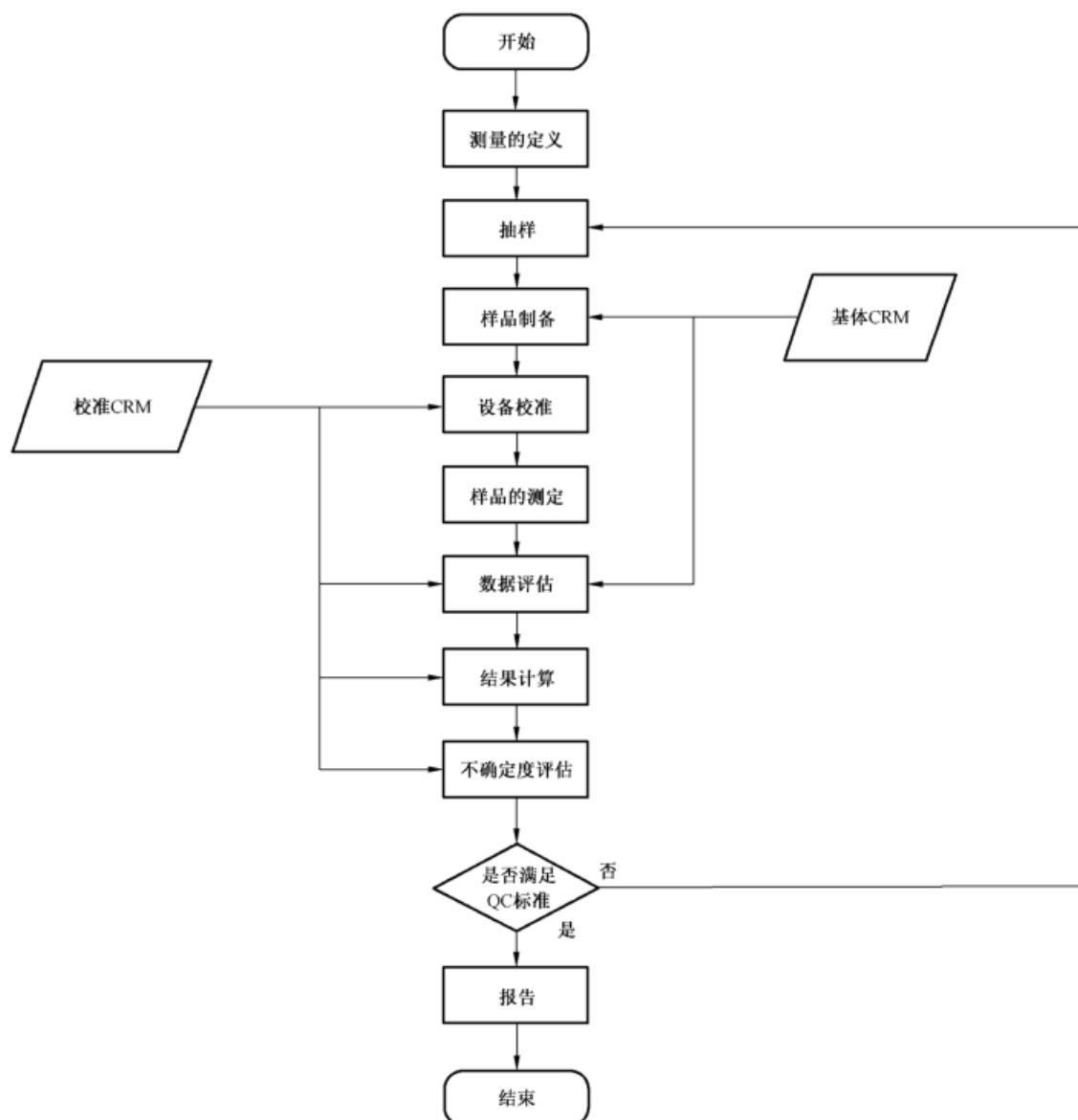


图 1 某项测量以及 CRM 在其中作用的框图

6.1.2 诸如 ISO/IEC 17025^[4]和 ISO 15189^[5]等关于实验室质量体系的标准要求测量结果具有计量溯源性,且测量设备要校准。计量溯源性是得到可比、兼容测量结果的先决条件。

示例:一种酒精体积分数为 12%的葡萄酒能够很好地与另外一种酒精体积分数为 13.5%的葡萄酒进行比较。

6.1.3 通常,大家理所当然地认为测量结果以合适的单位表示就是可比的。为了达到对测量数据的这种期望,实验室宜确保所有设备适当校准,并使用最终可以溯源到相关单位的测量标准。很多情况下,这个单位是国际单位制(SI)的一部分。

6.1.4 本文件附录 A 中给出了 RM 的关键特性和一般应用的对照表。

6.2 特性值

6.2.1 概述

6.2.1.1 CRM 对一个或多个特性进行了定值,这些特性值伴有:

- a) 相关特性的清晰说明；
- b) 不确定度陈述；
- c) 计量溯源性陈述；
- d) 证书的有效期。

使用者宜核实所有这些信息并以明确的形式表述。

6.2.1.2 信息值不宜用于本文件中描述的 CRM 的任何用途。

注：实际使用的信息值术语与本文件不总是一致。

6.2.2 特性的界定

6.2.2.1 相关特性的清晰界定对于确定 CRM 是否适用于预期用途有极大的帮助。CRM 的使用者有责任评估标准样品对预期用途的适用性。

示例：对于土壤中的痕量元素，重要的是界定是否是总量、不完全消解（如王水溶解）得到的含量、可浸出量或是包含痕量元素的某个特定成分。

6.2.2.2 特性值宜以合适的单位、最好是以 SI 单位给出。特性值宜以适当的数字位数给出，一方面避免准确度的不必要损失，另一方面避免给出过于准确的错误表述。

注：GUM(ISO/IEC Guide 98-3:2008, 第 7 章)给出了测量结果和相关不确定度的取整指导。

6.3 不确定度陈述

6.3.1 不确定度陈述宜易于理解，其中包括提供将所陈述的不确定度转换为标准不确定度的所有必要信息。若给出扩展不确定度，则需注明相应的包含因子。

示例：一份混合气体的校准证书中陈述一氧化碳的质量分数如下：

$$x_{\text{CO}} = (41\,122 \pm 28) \mu\text{mol/mol} (k = 2)$$

扩展不确定度为 $28 \mu\text{mol/mol}$ 。标准不确定度可通过计算得到：

$$u = \frac{U}{k} = \frac{28}{2} \mu\text{mol/mol} = 14 \mu\text{mol/mol}$$

6.3.2 若给出包含区间，则宜规定特性值的（假定）概率密度函数，包括给定区间的包含概率（如 95%）。这样的区间可以是不对称的。有时，有必要做附加的相关假定，例如相应的包含因子。在这些情况下，宜遵循 GUM(ISO/IEC Guide 98-3:2008, 6.3) 的规定。

示例：烟煤中的碳含量为 760.1 mg/g ，不确定度为 2.1 mg/g 。在不确定度陈述中附有如下标注：“不确定度以 95% 置信区间表示，在该标准样品用于校准目的时使用。”

从研制报告中很清楚地得知特性值是由实验室间协作得到的，因此假设正态分布比较合理。95% 置信区间的半宽度为标准差的 1.96 倍。因此用（包含）因子 2（相当于正态分布 95.45% 的置信水平）无实际差异。

标准不确定度通过 $u = \frac{U}{k} = \frac{2.1}{2} \text{ mg/g} = 1.05 \text{ mg/g}$ 得到。

6.3.3 陈述的不确定度宜与特性值的单位一致，或者表达为特性值的分数（例如，相对扩展不确定度）。如果使用分数，宜确认能否实现以一种明确的方式转化为绝对标准不确定度。

注：这种分数包括百分比、千分比和百万分比（ppm），然而这些都不建议用来表示不确定度，因为其不明确。

6.3.4 CRM 附有证书，其中标明规定特性及其值和相关不确定度（见 6.2.1.1）。描述如何得到特性值的不确定度超出了本文件的范围，但是知道其潜在的主要贡献很重要。

注：涉及建立特性值不确定评定的详细信息在 ISO Guide 35 中给出。

CRM 特性值不确定度的主要贡献包括^[2]：

- 定值的不确定度；
- 长期稳定性引入的不确定度；
- 短期稳定性（运输条件下标准样品的稳定性）引入的不确定度；
- 瓶间变化引入的不确定度。

6.3.5 在某些情况下,了解不确定度评定中不同部分的详细信息是很有帮助的,尤其某个贡献是最大者时。这些信息可以从 CRM 生产者获得。

6.3.6 20 世纪 90 年代末以前生产的 CRM 证书中,部分特性值附有的不确定度未包含批不均匀性和不稳定性的影响。证书中的不确定度宜包含可能引起特性值分散在不同批次间和随时间的影响。最终,陈述的不确定度宜适用于将在测量过程中使用的单个包装^[2]。

注:若陈述的不确定度过小,扩展不确定度具有比陈述的更低的置信水平。

6.3.7 没有特性值的 RM 宜给出诸如特性的(瓶间)均匀性和长期稳定性,以便此特性能够应用。使用者宜确认此信息是否以一种能够评估 RM 适用性的形式提供。这种评估可包括在随后的不确定度计算中使用相关均匀性和稳定性的信息。

6.4 溯源性陈述

6.4.1 计量溯源性是测量结果的一个特性。由于通过 RM 定值过程得到的值(特性值)是测量结果,特性值也能够同样具有这个特性。CRM 特性值的一个关键特性是其已很好地建立了溯源性。

6.4.2 CRM 的使用者宜确认特性值附有其计量溯源性的陈述。该陈述宜告知使用者特性值溯源到的测量标尺,以便使用者能够判断 CRM 是否适用于预期用途。

注:在多数情况下,测量标尺是 SI 单位。

6.4.3 为了解释计量溯源性陈述,以下几点是必要的,这也是 CRM 证书或其他文件中宜陈述的:

- a) 被测量的说明;
- b) 特性值溯源到的单位;
- c) 用于样品处理/转化的方法和用于定值的测量程序/技术;
- d) 定值方式(如单一方法、两种方法、多个实验室等)。

注:CRM 附带的文件能够通过包括网址、电子邮件或发表的公开刊物等不同渠道获取。

6.4.4 测量结果的提供者有责任支持结果或值的计量溯源性陈述。如果是 CRM,则 RM 生产者承担此责任。审核计量溯源性对其预期用途的适宜性是使用者的责任。

6.4.5 为评估计量溯源性陈述,使用者可能需要得到证书以外的更多信息。计量溯源性陈述由 6.4.3 中的条款支持。

RM 使用者宜仔细核查这些信息并确保特定标准样品预期用途的适用性。RM 使用者宜核查本条款所描述信息的可获得性以帮助评估。若这些信息的必要部分不可获得,则 RM 可能不适用于某个用途。

7 RM 和 CRM 的处置

7.1 宜遵守使用和储存说明,它们是保证标准样品特性值及其不确定度有效的要素。宜始终避免在测量程序的实施过程中不正确地使用 RM 和 CRM。

7.2 宜遵守证书上的有效期限,不宜使用超过有效期的 CRM。

7.3 尤其对于可多次使用的 CRM,确保其包装容器有适合的密封性,并以恰当的方式储存非常重要。某些情况下,有必要对剩余的部分重新包装,否则给出的特性值可能无效,CRM 无法使用或不可靠。使用者宜遵循生产者在这方面提供的使用说明。

7.4 宜遵守最小取样量。小于最小取样量可能没有代表性。

7.5 CRM 的取样,宜以能够代表整个包装样品特性的方式完成,否则,长此以往,剩余的 CRM 不再对已生产和定值的包装批有代表性,因此证书上给出的值和不确定度不再有效。

注 1:CRM 取样前再次均匀化通常是必要的。这些说明通常在 CRM 附带的文件中给出。

注 2:所谓“单剂量”CRM 是按单次使用的剂量设计的,通常每个单元仅含足够 1 次或 2 次测量所需的样品,但是,当单剂量标准样品设计成单次使用的量时,不宜再分取。

8 精密度的评估

8.1 概述

8.1.1 实验室进行测量程序精密度的核查需要将重复性条件(或其他指定条件)下实验室内标准差与该标准差所要求的值进行比较。

注：重复性或再现性条件下的标准差是精密度的度量。

8.1.2 精密度评估可以是实验室开发或确认方法活动的一部分。理想情况下,进行此类实验宜使用能够覆盖方法所适用的各种基体(或某个基体内的差异)和不同特性值水平的 $RM^{[6]}$ 。评估也可在多个实验室进行。更多有关通过实验室间研究评估精密度的指导在 ISO 5725^[7-12] 中已给出。

8.1.3 测量程序定期核查得到的结果可记录在质量控制图中,极差控制图可用于此目的^[13]。

8.2 重复测量次数

8.2.1 为了在所要求的置信区间内对精密度进行可靠评估,重复测量所必需的次数能够用 χ^2 检验估计。所要求的重复测量次数(n)主要取决于精密度评估所选择的 α 和 β 的值以及备择假设。统计假设检验中与第 I 类和第 II 类风险有关的概率 α 和 β 的讨论在附录 C 中给出。

8.2.2 表 1 所示为 $\alpha=0.05$ 时不同 β 值所对应的自由度 ν (此时, $\nu=n-1$) 与测量过程的实验室内标准差(s_w)和实验室内标准差所要求的值(σ_{w0})之比率之间关系。

示例:当 $n=10$ 时,如测量过程的实验室内标准差(σ_{w0})等于或大于 2.85 倍 σ_{w0} 所要求的值,则在 $\alpha=0.05$ 时测量结果的方差通过适当的检验(见 8.6)的概率不超过 1%。

表 1 测量过程的标准差与标准差要求值之间的比率

(对应 $\alpha=0.05$ 时不同的 β 值和自由度 ν)

ν	$\alpha=0.05$			
	$\beta=0.01$	$\beta=0.05$	$\beta=0.1$	$\beta=0.5$
1	159.5	31.3	15.6	2.73
2	17.3	7.64	5.33	2.08
3	6.25	4.71	3.66	1.82
4	5.65	3.65	2.99	1.68
5	4.47	3.11	2.62	1.59
6	3.80	2.77	2.39	1.53
7	3.37	2.55	2.23	1.49
8	3.07	2.38	2.11	1.45
9	2.85	2.26	2.01	1.42
10	2.67	2.15	1.94	1.40
12	2.43	2.01	1.83	1.36
15	2.19	1.85	1.71	1.32
20	1.95	1.70	1.59	1.27
24	1.83	1.62	1.52	1.25
30	1.71	1.54	1.46	1.22
40	1.59	1.45	1.38	1.19
60	1.45	1.35	1.30	1.15
120	1.30	1.24	1.21	1.11

8.3 对 RM 的要求

8.3.1 为了评估测量精密度,要求 RM 具有足够的均匀性和稳定性。至少在进行精密度核查测量期间,RM 中所涉及的所有特性宜是足够稳定的。必要时,宜采取特别措施对所用 RM 的稳定性进行监测。这些措施可包含用其他手段证明所研究测量过程的稳定性,如使用 CRM 或者使用稳定性已被证明的其他测量过程。

8.3.2 使用 RM 绘制控制图时,异常结果可能是由 RM 的稳定性引起的,而不属于测量系统的问题。因此,RM 使用者宜了解这种可能性并将之包含在根本原因分析中。

8.3.3 尽管用于精密度评估的 RM 不一定要具有计量可溯源的特性值,但由于精密度测量可能依赖于被测量的名义值,因此,一般需要了解有关参数的名义值,以便对所选用于核查的 RM 适宜性进行评估。

8.3.4 有关 RM 适宜性的更通用的指导见第 13 章。

8.4 测量

8.4.1 使用者宜进行独立的重复测量。“独立”的实际意义是指某个重复结果不受先前测量的影响。进行重复测量意味着重复整个程序。例如,在对固体材料进行化学分析时,从试样称量到最终读数或结果计算的程序均宜重复进行。

示例:如果土壤 RM 中铅的测量包括取样、试样消解和随后测量等分试样,那么对于涉及的取样、消解和测量,结果是独立的。反之,如果只是重复测量一个等分试样,那么所得到的标准差也只是测量某个等分试样的重复性影响。

在开始实验工作之前,核查标准差 σ_{wO} 对应过程的哪一部分是非常重要的。许多检测文字标准中的重复性对应着整个标准检测方法。因此,评估这些检测方法重复性的测量宜是完整测量,即对每个(子)样品重复使用该检测方法。

8.4.2 独立重复测量可通过各种方式实现,这取决于过程的性质。不过,由于在程序的任何一步出错都会影响所有重复测量,因此不总是推荐平行重复测量。而且,一些步骤,如测量设备的校准,也需要包含在重复过程中。

示例:在进行铁矿石分析时,分析步骤的重复要在不同的时间进行,并且要包含合适的校准。此时,由重复测量计算得到的标准差也包含来自每日间变异和校准的贡献。

8.5 数据处理

8.5.1 首先宜对所获得的数据仔细核查任何异常情况。被识别为技术无效的数据,无论它们是否表现符合(假定的)完整数据集的概率分布,均宜剔除。技术无效结果通常是指在样品处理和/或测量过程中发生了异常情况。

8.5.2 第二种类型的异常情况是指在数据集中一些观测值与其他观测值看上去相差很远。通常但不总是能够找到某个技术原因,然后把这些数据剔除。如果不能找到技术解释,则需采用 ISO 5725-2^[8]或 ISO 16269-4^[14]所描述的方法对这些数据进行可疑值核查,离群值宜被剔除,极个别情况下(如计算错误)也可用修正的数据替换。只要可能,宜依据一个以上离群值检验的结果剔除离群值,但通常歧离值宜保留在数据集中。

注 1:过多的可疑离群值表示测量过程有问题。

注 2:剔除离群值,特别是剔除歧离值,将降低数据集的分散性并可导致所研究测量过程的标准差的值过小。

注 3:大多数离群值检验的应用均要求有一个与数据概率密度函数的(期望的)形状相对应的假定。如果这个假定与数据的性质不一致,那么不能应用这种离群值检验。

8.6 精密度的计算和评估

8.6.1 测量过程的精密度可通过比较在重复性条件下实验室内标准差与实验室内标准差所要求的值(σ_{wO})进行评估。

计算平均值($\bar{x}$)和标准差(s_w):

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad \dots\dots\dots (1)$$

$$s_w = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \quad \dots\dots\dots (2)$$

式中:

x_i ——单个结果;

n ——剔除离群值后结果的个数。

8.6.2 计算下列比值:

$$\chi_c^2 = \frac{s_w^2}{\sigma_{WO}^2} \quad \dots\dots\dots (3)$$

式中, σ_{WO} 为实验室内标准差所要求的值。

$$\chi_{table}^2 = \frac{\chi_{(n-1)}^2; 0.95}{n-1} \quad \dots\dots\dots (4)$$

表示 $(n-1)$ 自由度的 χ^2 分布被自由度 $(n-1)$ 分割的第 0.95 分位点。

χ^2 的解释:

当 $\chi_c^2 \leq \chi_{table}^2$ 时, 没有证据表明测量过程未达到所要求的精密度。

当 $\chi_c^2 > \chi_{table}^2$ 时, 有证据表明测量过程未达到所要求的精密度。

注: χ^2 值可由查表或软件计算得到, 在包括参考文献^[15]在内的许多地方均能找到它们。

9 偏倚的评估

9.1 概述

9.1.1 偏倚核查是实验室中的一项重要应用。它能够作为测量结果质量保证和/或方法确认的一部分来完成。偏倚核查的必要条件是其所依据的参考标准是可靠的并具备计量溯源性。

9.1.2 如果材料的类型和特性与预期用途相关, 则采用 CRM 实施偏倚评估是适当的。偏倚核查前, 使用者宜确认 CRM 的适用性。

9.1.3 质量控制样品 (QCM) 及其他未定值的 RM 能够用于精密度评估 (见第 8 章), 但由于缺少具备计量溯源性的特性值, 它们不能用于评估测量偏倚。

注: ISO Guide 80^[3] 涵盖了质量控制样品的研制。

9.1.4 本章给出偏倚评估的指南, 本文件第 8 章给出了评估精密度的指南。

9.2 偏倚核查方法

9.2.1 使用 CRM 偏倚核查构成测量结果的计量学基础。它是测量程序确认的必要环节。如果使用相同的 CRM, 针对同一被测量的测量结果具有计量溯源性, 那么这些结果无显著性偏倚。这种情况下, 该测量程序得到的结果可溯源至同一参考标准。

9.2.2 测量值与证书上标明的特性值之差宜小于与该差值相关联的标准不确定度, 即

$$x_{meas} - x_{CRM} \leq k \sqrt{u_{meas}^2 + u_{CRM}^2} \quad \dots\dots\dots (5)$$

注: 选择包含因子需要用 (假定的) 概率密度函数和包含概率。通常使用 95% 作为包含概率, 但该选择可取决于具体应用。

9.2.3 如果式 (5) 成立, 则测量值与特性值在其各自的不确定度范围内是相互一致的。由于 CRM 的特性值可计量溯源至某些陈述的参考标准 (理想情况下计量溯源至 SI), 因此在该条件下得到的 CRM 特性值结果确认了由该测量程序所得结果的计量溯源性。

9.2.4 使用者宜独立于偏倚核查所用 CRM 来校准仪器设备。实验室随后通过比较测量结果与陈述特性值之差来评估测量方法所有步骤的正确性。若实验室在日常样品测量时能够使用相同的测量程

序,则可证明其测量结果可溯源至 CRM 特性值所陈述的参考标准。

注 1: 实践中,要将测量程序用于 CRM 时,可能有必要修订测量程序的某些部分。为测量 CRM 而对日常测量程序的必要修订可能会降低评估的有效性。必要的修订越多(或说越严重),评估的用处越小。

注 2: 对于定义被测量的测量程序,例如酶的测定,改变测量程序将改变被测量的定义。

9.2.5 本章所概述的方法偏倚评估不限于使用单一 CRM。实际上,如果存在一个以上 CRM,则建议使用多个 CRM 来核查方法,以便覆盖与方法适用范围相关的被测量值范围。

9.3 偏倚数据应用

9.3.1 由用于校准 CRM 得到的偏倚估计值可直接用于修正,这种修正可以使用加法、乘法或其组合。有必要了解校准模型及其细节以决定如何应用修正。

9.3.2 对于测试,偏倚修正更加复杂,因为 CRM 的属性可能不能完全反映日常样品的属性。在很多情况下,建议改进方法以便消除偏倚,而不是尝试修正。有些标准测试方法给出了可接受偏倚的准则。

注: 如果 CRM 的属性是不充分的且测量程序不能被改进,那么该 CRM 不适用于评估所考察测量程序的偏倚。

9.3.3 偏倚用公式(6)表示:

$$d = x_{\text{meas}} - x_{\text{CRM}} \quad \dots\dots\dots (6)$$

相关联的标准不确定度用公式(7)评估:

$$u(d) = \sqrt{u_{\text{meas}}^2 + u_{\text{CRM}}^2} \quad \dots\dots\dots (7)$$

9.3.4 如果偏倚是显著的,即 $|d| > U(d)$,通常尝试找出产生偏倚的原因并予以降低或消除。

注: $U(d) = k \cdot u(d)$,式中 k 表示适当选择的包含因子。

9.3.5 如果不可能充分降低或彻底消除偏倚,宜使用偏倚修正测量结果且宜将与偏倚相关联的不确定度包含在不确定度评定中。修正可能是加和的或是乘积的,这取决于它们是否依赖于被修正的量值。

9.3.6 如果测得的偏倚没有被修正且显著,该偏倚宜包含在不确定度估算中。粗略的近似算法是将偏倚的平方(即 d^2)加入到不确定度估算中来计算未修正的显著偏倚。

9.3.7 如果在被测量值的某个范围内评估了偏倚,能够计算平均偏倚及其相关联的不确定度。GUM (ISO/IEC Guide 98-3:2008,F.2.4.5)中给出了通用方法。

10 校准

10.1 概述

10.1.1 对于校准,需使用 CRM,CRM 宜在以下几个方面适用于设备的校准:

- a) 物理形态;
- b) 合适的认定特性;
- c) 特性值的范围及其与测量范围的相关性;
- d) 反映常规样品属性的适用性(互换性)。

10.1.2 有必要使用一组 CRM 进行设备校准,特别是在测量设备示值与特性值之间可能偏离比例关系的情况下。

10.1.3 与特性值相关联的不确定度应用于评定由校准引入的测量不确定度,为此,可使用 GUM 不确定度传播律,或其他任何概率分布/不确定度的传播机理。

10.2 建立计量溯源性

10.2.1 使用 CRM 进行器具校准是建立由该器具所获得校准函数的计量溯源性的便捷途径。通常,CRM 的特性值用于校准模型。

10.2.2 在某些情况下,仅能获得纯物质来作为 CRM,而校准方法需要另一种物理形态。这种情况下,由校准样品配制所得到的值及其不确定度宜用于测量过程。

10.2.3 使用 CRM 校准设备时,宜在一定程度上建立质量保证措施。至少,宜使用适当的 QCM、先前使用的校准样品或其他方法来核查校准,保证前后校准结果的一致性。

注:可结合其他保证测量结果有效性的质量保证措施核查校准的一致性。

10.3 校准模型

10.3.1 从建立计量溯源性和评定测量不确定度的角度而言,基于校准目的使用 CRM 是非常清晰的。将校准特性值输入所测样品的赋值公式中,本文件的附录 B 给出了三种常用的方法,即

——单点校准;

——双点校准;

——多点校准。

10.3.2 单点校准是最简单的方法,一个校准样品(本文件中的 CRM)用于校准测量设备,然后用于给被测样品赋值。

10.3.3 双点校准需要两个校准样品,其中之一的特性值要比样品的特性值大,另一个要比样品的特性值小,通过在两个校准样品之间的线性插值给其他样品赋值。

10.3.4 多点校准广泛用于校准测量设备,尤其在分析化学领域。测量一系列校准样品,然后根据测得的响应,通常使用曲线回归的方法建立测得响应与待测量之间的关系。

注:一种简单的曲线关系就是直线。

10.3.5 附录 B 给出了基于这三种主流方法赋值的公式及不确定度的评定方法。

11 为其他材料赋值³⁾

11.1 概述

11.1.1 尤其在设备校准时,CRM 常通过混合、稀释或其他方法用于制备其他的 RM。新制备的 RM 的特性值在一定程度上取决于用以制备 RM 的 CRM 的特性值。这些用途通常被冠以“为其他材料赋值”的标题。制备方法包括称量法和容量法。

11.1.2 CRM 的这种用途被广泛使用。事实上,分析化学中常用的校准大多数是基于 CRM 的这种作用。纯物质通常被用于制备混合物或溶液,它们随后被用于校准设备。有时,这些混合物或溶液在使用前要进一步稀释。浓度、物质的量分数或其他成分数值的计算基于纯度数据和制备数据。

如果制备过程使用的设备已被适当校准、相应的环境条件被监测,就可以获得可溯源至 SI 单位的特性值。

注 1:环境条件在称量法的准确度中起到主导作用,尤其是称量大体积目标物时,例如钢瓶,空气浮力可发挥重要的作用。控制环境条件的需求根据不确定度评定的严格程度和预期的准确度水平而定。

注 2:除其他因素以外,浓度测量也取决于温度。与现场实验室其他不确定度分量相比,这些影响可能比较小,但是对于容量法处理过程,相对于其他来源的不确定度,它们并不小。

11.1.3 宜核查这些校准样品量值的一致性。这种核查可按下列方式进行:

——将新校准样品与之前的、经确认的校准样品比较;

——评价质量控制中使用新校准样品的影响,例如测定一个 QCM。

11.1.4 如果校准对整个测量程序的性能有重要影响,则有必要将新校准样品与之前的、经确认的校准样品进行比对。从新校准样品的制备计算得出的量值(x_{prep})宜与使用之前的校准样品得出的值(x_{meas})进行比对。之前经确认的校准样品用于校准比对试验所用的仪器。若满足公式(8),新校准样品为有效:

$$|x_{\text{prep}} - x_{\text{meas}}| \leq k \sqrt{u_{\text{prep}}^2 + u_{\text{meas}}^2} \quad \dots\dots\dots (8)$$

式中, k 表示置信水平为 95% 时选择的包含因子。在大多数情况下, $k=2$ 是合适的选择(请参考 ISO/IEC Guide 98-3:2008 第 7 章,获得选择包含因子的进一步指南)。

3) 此处也可译为“为其他样品赋值”。

除使用上述准则外,如果观察到的差值($x_{\text{prep}} - x_{\text{mess}}$)与测量或测试方法要求的不确定度相比很小,也可认为新校准样品是有效的。

11.2 纯物质

11.2.1 在很多不仅限于化学分析的测量领域,纯物质在建立计量溯源性中发挥着重要作用。

注:“纯”是一种理想化的概念,就像均匀或稳定一样。实际上,没有材料是绝对纯的。

11.2.2 对于成分测量,纯物质通常是所谓的“溯源链”的基础,即纯物质和待测物质成分测量之间的第一次链接。用于此目的的所有材料宜已对杂质定值,且这些与预期用途相关杂质需要进行鉴别和量化。

示例 1:氮气(6.0 级)通常含有几个 nmol/mol 的苯。这种杂质在制备合成天然气时不是很重要(此处氮气的物质的量一般为 0.5 cmol/mol~20 cmol/mol,而苯,如果有的话,为 1 μmol/mol~10 μmol/mol),但是在制备空气质量标准时,这种杂质却高度相关(此处苯的目标含量为 5 nmol/mol~50 nmol/mol)。

示例 2:在用于温标固定点的材料中,杂质的存在可导致偏离,如凝固点的下降。

11.2.3 许多化学品或其他纯物质都附有相关杂质的数据。在计量情境中,这些信息只有规定了以下内容时才有用:

——测量单位(如 mol/mol,若以物质的量的分数表示时);

——指定值的不确定度。

示例 1:对于土壤中痕量污染物分析的检测实验室,通常用新的纯化学物质制备的校准样品与之前的校准样品比对就足够了,由此可以确定无显著的偏离。通过诸如 QCM、PTM(能力验证样品)的重复分析或使用 CRM 可以获得这种确定。

示例 2:生产检测实验室用的校准样品的生产商宜提供具有计量溯源性的成分数据,并对纯物质(包括基体)中的杂质进行鉴别和量化。

11.2.4 纯度分析不限于溶解、稀释或以其他方式制成适合进行化学成分测量的纯物质,溶剂、稀释气体等也宜进行纯度分析,因为它可能含有可定量检测到的杂质,这些杂质可影响测量结果。

在测量和检测中,这些纯度分析可作为试剂空白核查进行,即核查杂质的水平是否足够低,不会对测量程序的后续步骤产生负面影响。

11.3 称量法和容量法

11.3.1 称量法和容量法广泛用于校准样品的制备技术。

11.3.2 校准样品特性值是基于制备过程计算得到的。

11.3.3 与特性值相关的不确定度能够通过不确定度传播律和 10.3.1 所述的模型获得。

11.3.4 许多校准样品的一种或多种特性随时间是不稳定的。为了在校准仪器期间获得有效的结果,较早赋予校准样品的特性值宜在其相应的不确定度内仍然有效。开展某些类型的稳定性核查是必要的。

11.3.5 校准样品的稳定性核查可采用许多种方法完成。下面列出一些方法:

- 开展稳定性检验;
- 比较使用新旧校准样品对同一样品测量获得的测量结果,如 PTM 留样或 QCM;
- 使用旧校准样品校准仪器,再测量新校准样品,然后将测得值与校准样品的指定值进行比较。
- 使用独立 CRM 核查整个测量程序。

稳定性检验可能比较费力,但若有异常,则是必要的。ISO Guide 34 和 ISO Guide 35 中有关于稳定性检验的内容。

11.3.6 赋予校准样品的值宜在整个有效期限内有效。实验室宜按此为它们的校准样品设置寿命,这样就无须再做专门核查,并依赖好的质量控制来发现校准样品存在的问题。

如果要确定其寿命,宜定期核查校准样品的质量,以保证由此获得的结果的计量溯源性,校准样品长期使用尤其如此。

12 约定标尺

12.1 概述

12.1.1 自早期文明以来,人们已经使用过很多种测量标尺,最初的测量标尺几乎全部是约定的、无关的和 inaccurate 的。科学技术的进步以及国际贸易的发展建立了一个唯一的、合理的、自治的国际单位制(SI),且已在全世界广泛使用。然而,SI 对某些测量类型不适用,对于这些测量需要建立、维持和使用某些不属于 SI 范围的约定单位。另外,有些与被测量有关的单位虽然处在 SI 框架内,但按照定义复现该单位时在技术上却是非常困难和昂贵的。因此使用将材料特性赋予参照值而建立的实用标尺实现测量则显得更为方便。虽然一个参照值标尺和单纯的约定标尺在理论上彼此不同,但就标准样品使用而言它们是类似的,因此都作为约定标尺一起讨论。

12.1.2 约定标尺是以标准样品赋值为基础,赋值由标准规范、国际建议或其他参考文件规定,因此复现约定标尺的一个固定点的标准样品在全世界应具有相同的质量,这种类型的 CRM 按特性值定值,亦即在标准设备上用参考方法测量。

12.1.3 显然 CRM 只能保证一个测量标尺的固定点。标尺的测量或是要求一个固定点和通过此点的一个数学函数,或是要求两个或多个固定点,在各点之间应用规定的内插方法。

注:存在特殊的不连续的标尺,例如在地质实验中用于测量硬度的莫氏标尺,该标尺是以赋予 10 个硬度级的 10 种矿物为基础,每种较硬的矿物都会将较不硬的矿物划伤。

12.1.4 一个约定标尺有两个基本支撑:一个是用于复现固定点的有证标准样品,另一个是给出测量程序的标准规范(或类似文件)。它们都宜被严格定义以确保在该约定标尺上的测量一致性。标准规范提供建立和使用以赋值为基础的标尺所需要的详细信息,或者提供依据假定的测量中所使用的实验和计算的程序。宜在详述测量方法的同一标准规范中规定对于 CRM 的要求。借助于必需的 CRM 和相应的标准规范,使用者能创建测量标尺,并利用这一标尺测量其样品或校准仪器。

12.1.5 在评定标尺的测量不确定度时,使用者宜考虑创建标尺的不确定度以及用 CRM 确定固定点时的不确定度。有时使用者的最终使用中所需要的不确定度水平比由 CRM 所确定的固定点的不确定度要低(如血液 pH 测量)。他们需要意识到在标尺上测量的不确定度必定大于固定点的不确定度。另外,CRM 的重复测量和标尺的建立(点的合理选择、内插方法的特征及重复性等)也影响总不确定度。

12.1.6 确定标尺固定点的 CRM 的选择宜取决于最终使用要求的不确定度水平。为使标尺上测量值的不确定度减到最小,使用者宜使用以标尺单位定值的 CRM。显然,使用者理应通晓有关复现标尺的测量程序以及正确使用该 CRM 的说明书的全部信息。

在某些情况下,以标尺单位定值的 CRM 如果得不到或价格昂贵,或者在测量的不确定度水平下不必使用这种 CRM,则使用者能够应用纯的化学物质来确定固定点。如果选择这种程序,使用者宜知道物质纯度和所依据标尺的特性之间的关系,而且测量的不确定度只能粗略估计。

约定标尺有多种多样,确定它们时 CRM 的使用方法差别很大。为了说明一些约定标尺的不同特性,12.2 和 12.3 中给出了两个例子。

12.2 pH 标尺

12.2.1 由于绝对单离子活度不能用实验方法测量,因而 pH 被认为是一个不严格的物理量。为了使被测的 pH 尽可能有意义,采用了一个约定的 pH 标尺,它由赋予 pH 的参比溶液来定义。这些溶液的 pH 是通过测量无迁移的氢-银/氯化银电池的电动势,并根据约定用给定的计算方法计算得到的。

12.2.2 很多国家标准规范中描述了参比溶液的制备以及 pH 的赋值方法,这些参比溶液标准值的不确定度限定在千分之几的 pH 单位之内。

12.3 辛烷值

12.3.1 辛烷值标尺是由美国材料与试验协会(ASTM)和英国石油学会(IP)联合制定的标准规范定义

的。国际标准以及许多国家标准都参考了这些文件。ASTM D 2699-95a/IP273^[16]和 ASTM D 2700-95a/IP236^[17]分别描述了按研究方法和按马达法作为发动机燃料爆震特性的检验方法。在这两个标准中,燃料的辛烷值通过比较燃料的爆震趋势与在标准操作条件下已知辛烷值的 ASTM 标准燃料的混和物爆震特性而测定。标准样品和混合的附属设备在这两个标准的附录中给出。

12.3.2 ASTM 标准参考了美国国家标准技术研究院(NIST)的标准样品 SRM No.1816a(异辛烷,纯度 99.987%)和 SRM No.1815a(正庚烷,纯度 99.987%)。这些样品主要用于商业生产的 ASTM 爆震检验标准燃料的定值,在该标准中给出了这些标准燃料的规范,也列出了供应商,并由供应商承担标准样品符合规范要求的要求。ASTM 根据样品的物理特性进行定值,供应商需要检验待定值的标准样品并同时检验相应的 SRM 以确保生产可接受的标准样品的溯源性。ASTM 给供应商发证书,授权他们保证发运的标准样品是按规定测试的并提供测试结果。

13 CRM 和 RM 的选择

13.1 概述

13.1.1 标准样品,特别是有证标准样品,能够用于测量过程中各种用途。这些用途包括:

- 校准;
- 建立计量溯源性;
- 方法确认;
- 给其他材料的特性赋值;
- 质量控制。

13.1.2 符合 ISO Guide 34 和 ISO Guide 35 生产的 RM 和 CRM 是最佳选择。理想状态下,此类符合性在所附带的文件中已明确说明。使用者宜核查这些信息,如未附说明,请咨询 RM 的生产者。

13.1.3 不是每个 RM 都能用于每种用途。在一个特定的测量过程中,RM 仅能用于单一用途。

示例:浓度经确定的天然气混合物(甲烷、乙烷、丙烷、异丁烷、正丁烷、氮气、二氧化碳)能够用来校准气相色谱(GC),但在同一测量过程中,同一 CRM 不能同时用来核查 GC 的正确度。然而,同一 CRM 能够用来评价其他方面性能,如保留时间、峰间距及精密度。

13.1.4 标准样品可有不同的形态^[18]。常见的形态包括:

- a) 以化学纯度和其他特性(如熔点、黏度、燃烧焓)定值的纯物质;
- b) 通常由纯物质用称量法制备得到的标准溶液和气体混合物;
- c) 以选定特性组成(如特定化学成分)定值的基体标准样品。此类标准样品可由天然或人工合成获得;
- d) 以诸如熔点、黏度、辛烷值、闪点、硬度和吸光率等特性定值的物理化学标准样品;
- e) 以诸如味道、香气等感官特性定值的参比实物或人工制品。此类标准样品也包括从纤维种类到微生物标本的样品。

是否选择 CRM 或其他 RM 用作某种用途取决于诸多因素,如特定的应用和可获得性。

13.1.5 标准样品可从互联网各种数据库中查询到,如有证标准样品数据库(COMAR)、各国计量院关键比对数据库(KCDB)、RM 生产者网站。

13.2 CRM 的选择

13.2.1 本文件中所提 CRM 至少附带下列信息的文件:

- 对证书负责的生产者或认定机构情况;
- 对标准样品的描述,包括所来自的动植物物种的明确说明;
- CRM 的预期用途;
- 对于每个特性的特性值和相应(扩展)不确定度;
- 标准值的计量溯源性;

- 证书的失效日期(或有效期);
- 使用说明,包括限制说明;
- 适宜的存储条件。

注:对于定性测量用 CRM,特性(如颜色)可用数字表示,不确定度可用概率表示。

13.2.2 预期用途阐明 CRM 用于测量过程的目的。

13.2.3 实验室宜能解释和证明选择 CRM 的理由,也包括某特定 CRM 不适用的理由。使用者宜对 RM 进行正式的适用性评价,除非有证据表明所选 RM 不会显著影响测量结果。图 2 中给出了如何进行 CRM 适用性评价。13.3 给出了在评价中需要包含的各个方面。

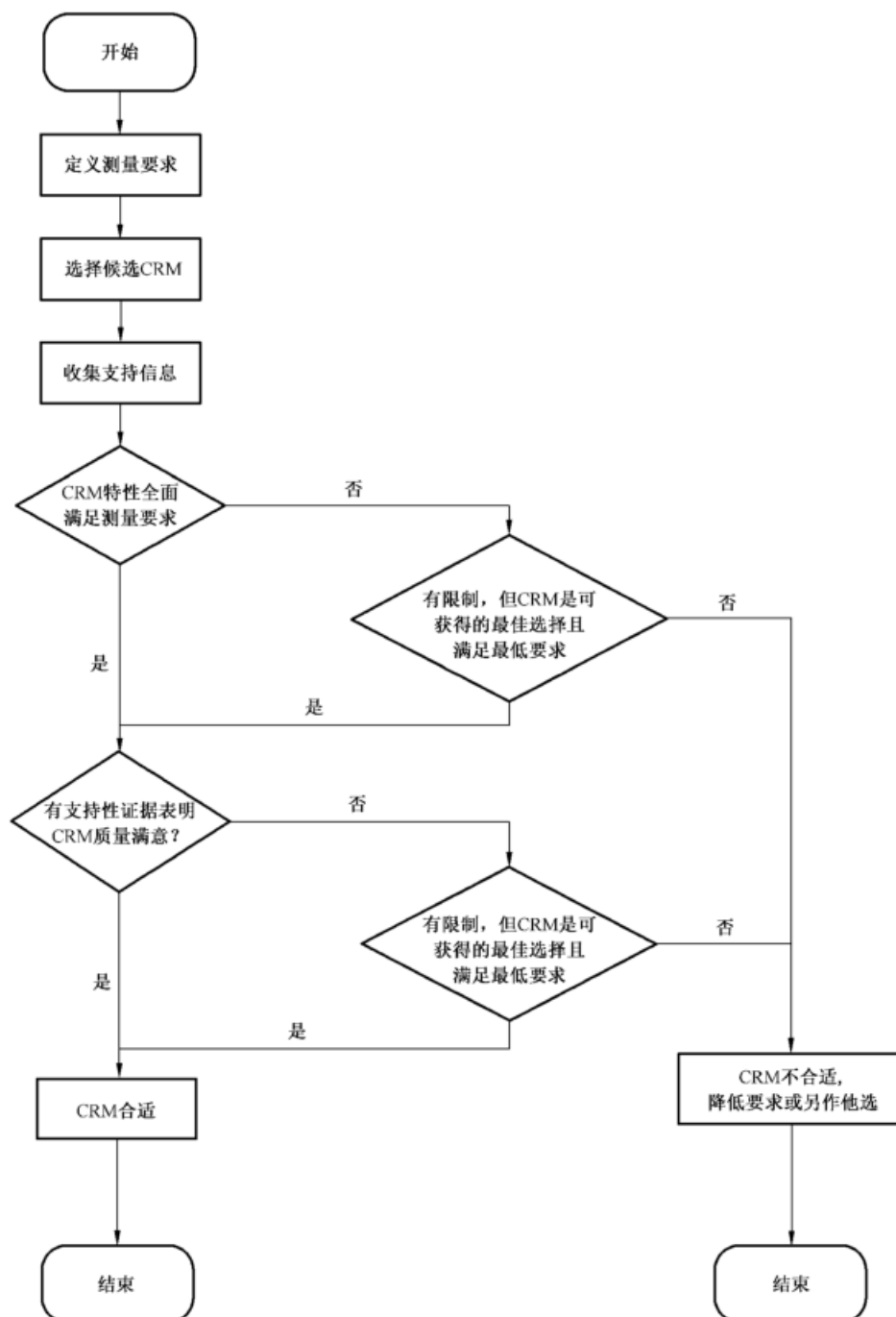


图 2 CRM 适用性评价(源自^[18],有修改)

13.3 RM 的选择

13.3.1 为讨论 RM 的选择(CRM 除外),有两点需要考虑:

- a) 已定值的 RM;
- b) 未定值的 RM,即没有给出特性值的 RM。

具有声称特性值的 RM 宜满足 ISO Guide 35 的要求,因此,特性值宜具有计量溯源性(最好溯源到 SI)。为满足相关用途,这些 RM 宜有与 RM 证书中信息量至少相同的文件信息包。对于此类 RM,适用性要求等同于 CRM(见 13.2)的选择。

13.3.2 许多 RM 没有特性值。大多数的应用须知道标称值或特性值的预期范围。此类 RM 经常用来做各种精密度控制,如日常的实验室内部质量控制。

此外,为保障使用,此类 RM 宜按照 ISO Guide 35 进行均匀性和稳定性的核查。

注:有一些情况下,即使 RM 没有诸如稳定性数据也能在测量过程中使用。这些情况仅限于在测量过程中此类 RM 的稳定性能够用另一种标准间接证明,如在测量过程中使用的已证明稳定性的 RM(CRM)。

13.4 与测量系统的相关性

CRM 的使用者宜根据证书上提供的认定(定值)的方法、预期用途的声明、CRM 正确使用的说明等因素,考虑 CRM 的何种特性与测量程序相关。

- a) 水平:CRM 宜具有与测量过程预期使用水平相适应的特性值,如浓度。
- b) 基体:CRM 宜与用于检测过程的对象(待测物)具有尽量相近的基体,如低合金钢中的碳或不锈钢中的碳。
- c) 形态:其可能是试片或加工样或粉末,也有可能需要进一步的制备。最好使用与待测样品形态相同(如固体或气体)的 CRM。
- d) 最小取样量:如果 CRM 证书上规定了能代表总体的最小取样量,使用者宜遵循。
- e) 包装量:CRM 的单元包装量对于整个实验项目宜是充足的,必要时还需要有所预留。避免在一个特定的测量过程中还需开启其他新的 CRM 单元,除非 CRM 是仅供单次使用。
- f) 稳定性:只要可能,CRM 宜在整个试验阶段中具有稳定的特性。具体有如下三种情况。
 - 1) 特性稳定,无须采取预防措施。
 - 2) 标准值可能受储存条件的影响,比如,在开瓶前后均需按照证书中的条件进行储存。
 - 3) CRM 证书已确定在规定时间内(以已知速率发生变化)的特性。使用者宜遵循在证书或其他相关文件中的使用说明。特性值及其不确定度仅在这些条件下是有效的。
- g) 标准值可接受的不确定度:标准值的不确定度需要与 1.1 中描述的精密度和正确度要求相匹配。
- h) 互换性:只要相关,使用者需要评估 CRM 对于预期用途是否具有互换性^[20-21]。可能会有来自 CRM 生产者的评价数据帮助使用者评价 CRM 的互换性。

在这种情况下,基质效应和样品处理的效应都宜评价。

附录 A

(资料性)

标准样品常规应用中的关键特性

标准样品的关键特性及其在常规应用中的相关性见表 A.1。

表 A.1 标准样品的关键特性及其在常规应用中的相关性

	精密度控制	偏倚控制	校准/约定标尺	给其他材料赋值
有关特性的描述	需要	需要	需要	需要
特性值		需要	需要	需要
陈述的不确定度		需要	需要	需要
均匀性的规定水平	需要	⁹	⁹	⁹
稳定性的规定水平	需要	⁹	⁹	⁹
计量溯源性陈述		需要	需要	需要
使用说明书	需要	需要	需要	需要
证书的有效期		需要	需要	需要
⁹ 给出的特性值相关的不确定度已包括此贡献。				

附录 B

(资料性)

校准模型和相关联不确定度模型

B.1 单点校准

B.1.1 假定 y_{CRM} 和 y_{sample} , 分别是测量 CRM 和样品得到的响应值, x_{CRM} 是 CRM 的特性值, 样品的测量结果计算如下:

$$x_{\text{sample}} = x_{\text{CRM}} \frac{y_{\text{sample}}}{y_{\text{CRM}}} \quad \dots\dots\dots (\text{B.1})$$

x_{sample} 相关联的不确定度可以计算如下:

$$u^2(x_{\text{sample}}) = x_{\text{sample}}^2 \left\{ \frac{u^2(x_{\text{CRM}})}{x_{\text{CRM}}^2} + \frac{u^2(y_{\text{CRM}})}{y_{\text{CRM}}^2} + \frac{u^2(y_{\text{sample}})}{y_{\text{sample}}^2} \right\} \quad \dots\dots\dots (\text{B.2})$$

不确定度 $u(x_{\text{sample}})$ 仅包含了测量重复性及与 CRM 特性值相关联不确定度的效应; 通常, 该不确定度估算宜再加入由诸如抽样、再抽样和样品转化等带来的不确定度。

B.1.2 这一模型的应用及其相关联的不确定度评定限定用于以下情况: x 和 y 变量间的关系能够使用下式完全覆盖所感兴趣的范围

$$y = a_1 x$$

式中, a_1 表示斜率。如果存在显著的零点偏离或者需要考虑非线性, 那么单点校准模型可能是无效的, 或者需在标准不确定度表达式中加入其他因子。

B.2 双点校准

B.2.1 双点校准需要两个 CRM, 其一的特性值要高于未知样品的值, 另一个要低于未知样品的值。两个 CRM 的结果用于线性内插。因此, 宜选择足够小的区间以保证检测器的非线性 (如果有的话) 不会导致未知样品赋值偏倚。因此, 区间的最大值取决于所用测量系统的非线性。

B.2.2 双点校准模型如下:

$$x_{\text{sample}} = \frac{x_2 - x_1}{y_2 - y_1} (y_{\text{sample}} - y_1) + x_1 \quad \dots\dots\dots (\text{B.3})$$

式中, y 表示响应值, x 表示待测量 (如浓度)。CRM 的特性值用 x_1 和 x_2 表示, 它们的响应分别为 y_1 和 y_2 , 样品的响应值为 y_{sample} , 样品值为 x_{sample} 。 x_{sample} 的合成标准不确定度能通过使用不确定度传播律以右边各变量的不确定度来表示。

x_{sample} 的合成标准不确定度计算如下:

$$u^2(x_{\text{sample}}) = \left[\frac{\partial x_{\text{sample}}}{\partial x_1} \right]^2 u^2(x_1) + \left[\frac{\partial x_{\text{sample}}}{\partial x_2} \right]^2 u^2(x_2) + \left[\frac{\partial x_{\text{sample}}}{\partial y_1} \right]^2 u^2(y_1) + \left[\frac{\partial x_{\text{sample}}}{\partial y_2} \right]^2 u^2(y_2) + \left[\frac{\partial x_{\text{sample}}}{\partial y_{\text{sample}}} \right]^2 u^2(y_{\text{sample}}) \quad \dots\dots\dots (\text{B.4})$$

式中, 各灵敏系数可以用 x_{sample} 表达式的微分给出。灵敏系数的表达式如下:

$$\begin{aligned} \frac{\partial x_{\text{sample}}}{\partial x_1} &= -\frac{y_{\text{sample}} - y_1}{y_2 - y_1} + 1 \\ \frac{\partial x_{\text{sample}}}{\partial x_2} &= \frac{y_{\text{sample}} - y_1}{y_2 - y_1} \\ \frac{\partial x_{\text{sample}}}{\partial y_1} &= \frac{x_2 - x_1}{(y_2 - y_1)^2} (y_{\text{sample}} - y_1) - \frac{x_2 - x_1}{y_2 - y_1} \end{aligned}$$

$$\frac{\partial x_{\text{sample}}}{\partial y_2} = -\frac{x_2 - x_1}{(y_2 - y_1)^2}(y_{\text{sample}} - y_1)$$

$$\frac{\partial x_{\text{sample}}}{\partial y_{\text{sample}}} = \frac{x_2 - x_1}{y_2 - y_1}$$

不确定度 $u(x_{\text{sample}})$ 仅包含了测量重复性及与 CRM 特性值相关联不确定度的效应；通常，该不确定度估算宜再加入由诸如抽样、二次抽样和样品转化等带来的不确定度。

B.3 多点校准

B.3.1 多点校准是应用广泛的校准技术。它不需要如无零偏差(单点校准)、仪器响应值 y 和特性值 x 关系的线性(双点校准和单点校准)之类的假定。

B.3.2 为了获得一个未知样品的测量值，只要给定这个样品的测量响应值，多点校准的数据通常需要使用一个拟合程序以及一个适当的模型来获得。在大多数情况下，宜使用最小二乘法得到校准模型的参数。

B.3.3 用于曲线拟合的方法有很多。为了传播 x 和 y 的不确定度，考虑到各自的不确定度，宜采用一种在两个方向上使平方差最小的方法。在气体分析领域，这种方法已写入 ISO 标准^[22]中。

B.3.4 作为最小二乘拟合的次优备选，传统的普通最小二乘法也可以使用。参考文献^[25]给出了实例。

B.4 标准加入

在分析化学文献中，有一种特别的多点校准方法叫作标准加入法。不同于使用独立的校准样品，校准标准被添加到(已转化的)样品中。测量值通过外推法得到。

附录 C

(资料性)

决策错误

根据精密度和正确度评估测量过程总有可能出现错误结论,这是因为:

- a) 测量结果的不确定度;
- b) 通常进行的重复测定次数是有限的。

虽然增加测量的次数将减少做出不正确结论的机会,但在很多情况下,由于增加测量次数要增加测量费用,因而要将增加测量次数的经济性与做出错误结论的风险相权衡。因此,在制定评价测量程序的准则时,必须考虑最终使用所要求的精密度和正确度水平。

基于本文件的目的,使用了术语“原假设”。

这种情况下,原假设为测量过程的偏倚不超过实验者所选择的限值而方差不超过预定值,备择假设是与原假设相反的假设(同时见 ISO 3534-1^[26])。

在接受或拒绝原假设时可能会出现两种类型的错误:

- a) 第 I 类错误:当实际上原假设为真时,拒绝原假设犯的错误。
 - 1) I 型风险:犯第 I 类错误的概率,其值随实际情况而异。
 - 2) 显著性水平:一个给定值,通常记为 α ,它是犯第 I 错误的概率界限。
- b) 第 II 类错误:当实际上备择假设为真时,没有拒绝原假设的错误。
 - 1) II 型风险:犯第 II 类错误的概率通常以 β 表示,其值随实际情况而异,且仅在备择假设明确规定时才能计算;
 - 2) 检验的功效:未犯第 II 类错误的概率通常指定为 $1 - \beta$ 。它是当实际上备择假设为真时,拒绝原假设的概率。

α 和 β 值的选择通常是基于经济上的考虑,它取决于决定后果的重要程度。这些值连同备择假设宜在测量过程开始之前予以选择。

参 考 文 献

- [1] ISO Guide 34 General requirements for the competence of reference material producers
- [2] ISO Guide 35 Reference materials—General and statistical principles for certification
- [3] ISO Guide 80 Guidance for the in-house preparation of quality control materials (QCMs)
- [4] ISO/IEC 17025 General requirements for the competence of testing and calibration laboratories
- [5] ISO 15189 Medical laboratories—Requirements for quality and competence
- [6] NEN 7779:2008 Environment—Measurement Uncertainty
- [7] ISO 5725-1 Accuracy (trueness and precision) of measurement methods and results—Part 1: General principles and definitions
- [8] ISO 5725-2 Accuracy (trueness and precision) of measurement methods and results—Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method
- [9] ISO 5725-3 Accuracy (trueness and precision) of measurement methods and results—Part 3: Intermediate measures of the precision of a standard measurement method
- [10] ISO 5725-4 Accuracy (trueness and precision) of measurement methods and results—Part 4: Basic methods for the determination of the trueness of a standard measurement method
- [11] ISO 5725-5 Accuracy (trueness and precision) of measurement methods and results—Part 5: Alternative methods for the determination of the precision of a standard measurement method
- [12] ISO 5725-6 Accuracy (trueness and precision) of measurement methods and results—Part 6: Use in practice of accuracy values
- [13] ISO 7870-5 Control charts—Part 5: Specialized control charts
- [14] ISO 16269-4 Statistics interpretation of data—Part 4: Detection and treatment of outliers
- [15] Snedecor G.W., & Cochran W.G. Statistical Methods (8th edition). Ames, Iowa, Blackwell Publishing Professional
- [16] ASTM D 2699-1995a Test Method for Knock Characteristics of Motor Fuels by the Research Method
- [17] ASTM D 2700-1995a Standard Test Method for Motor Octane Number of Spark-Ignition Engine Fuel
- [18] International Laboratory Accreditation Cooperation, ILAC G9:2005, Guidelines for the Selection and Use of Reference Materials, <http://www.ilac.org>
- [19] ISO Guide 31 Reference materials—Contents of certificates and labels
- [20] ISO 15194 In vitro diagnostic medical devices—Measurement of quantities in samples of biological origin—Requirements for certified reference materials and the content of supporting documentation
- [21] Vesper H.W., Miller W.G., Myers G.L. Reference Materials and Commutability. Clin. Biochem. Rev. 2007, 28(4), pp.139-147
- [22] ISO 6143 Gas analysis—Comparison methods for determining and checking the composition of calibration gas mixtures

[23] Analytical Methods Committee (AMC) Technical brief No.10.Fitting a linear functional relationship to data with error on both variables.Royal Society of Chemistry 2002,UK

[24] Ripley B.D., & Thompson M.Regression techniques for analytical bias,Analyst (Lond.),1987,112 pp.377-383

[25] Ellison S.L.R., Rosslein M., Williams A,eds.Eurachem/Citac Guide—Quantifying uncertainty in analytical measurement,LGC,London,Third Edition,2012

[26] ISO 3534-1 Statistics—Vocabulary and symbols—Part 1;General statistical terms and terms used in probability
