



CNAS-CL03

能力验证提供者认可准则
Accreditation Criteria for Proficiency Testing
Providers
(ISO/IEC 17043: 2010)

中国合格评定国家认可委员会

目 次

目 次	1
前 言	2
ISO引言	3
1 范围	4
2 规范性引用文件	4
3 术语和定义	4
4 技术要求	7
4.1 总则	7
4.2 人员	7
4.3 设备、设施和环境	8
4.4 能力验证计划的设计	9
4.5 方法或程序的选择	12
4.6 能力验证计划的运作	12
4.7 数据分析和能力验证计划结果的评价	14
4.8 报告	15
4.9 与参加者的沟通	16
4.10 保密	16
5 管理要求	17
5.1 组织	17
5.2 管理体系	18
5.3 文件控制	18
5.4 要求、标书和合同的评审	19
5.5 分包服务	20
5.6 服务和供应品的采购	20
5.7 服务客户	20
5.8 投诉和申诉	21
5.9 不符合工作的控制	21
5.10 改进	21
5.11 纠正措施	21
5.12 预防措施	22
5.13 记录的控制	22
5.14 内部审核	23
5.15 管理评审	23
附录A（资料性附录）能力验证计划的类型	25
附录B（资料性附录）能力验证的统计方法	29
附录C（资料性附录）能力验证的选择和利用	36
参考文献	39

前 言

本准则规定了能力验证提供者应满足的质量管理和技术要求，是 CNAS 对能力验证提供者能力进行认可的基本依据。

本准则等同采用 ISO/IEC 17043:2010 “Conformity assessment – General requirements for proficiency testing”。

本准则的附录是信息性资料，不是要求，旨在帮助理解和实施本准则。这些附录，在能力验证计划的分类、统计方法，以及各利益相关方选择和利用能力验证等方面提供了信息。这些资料虽然不是要求，但对能力验证提供者的技术运作及服务具有积极的作用。

对于能力验证样品制备和结果评价活动，除本准则外，CNAS 还制定了 CNAS-GL02 《能力验证结果的统计处理和评价指南》和 CNAS-GL03 《能力验证样品均匀性和稳定性评价指南》，供能力验证提供者参考。

本准则代替：CNAS-CL03: 2006 《能力验证计划提供者认可准则》。

ISO引言

实验室间比对广泛地用于诸多目的，其应用在国际上日益增长，实验室间比对的代表性目的有：

- a) 评定实验室从事特定检测或测量的能力及监视实验室的持续能力；
- b) 识别实验室存在的问题并启动改进措施，这些问题可能与诸如不适当的检测或测量程序、人员培训和监督的有效性、设备校准等因素有关；
- c) 建立检测或测量方法的有效性和可比性；
- d) 增强实验室客户的信心；
- e) 识别实验室间的差异；
- f) 根据比对的结果，帮助参加实验室提高能力；
- g) 确认声称的不确定度；
- h) 评估某种方法的性能特征——通常称为协作试验；
- i) 用于标准物质/标准样品的赋值及评定其在特定检测或测量程序中使用的适用性；
- j) 支持由国际计量局（BIPM）及其相关区域计量组织，通过“关键比对”及补充比对所达成的国家计量院间测量等效性的声明。

能力验证包含了利用实验室间比对确定实验室的能力，如 a)至 g)所列，但能力验证通常不从事 h)、i)和 j)活动，因为在这些比对中实验室的能力已被设定，但这些应用可以为实验室的能力提供独立的证明。因此，本准则的要求也可用于 h)、i)和 j)中的许多技术策划和运作。

对实验室能力的持续信任，不仅对实验室及其客户至关重要，而且对其他利益相关方也极其重要。这些相关方包括例如管理部门、实验室认可机构，以及其他对实验室持有特定要求的组织。GB/T 27011 要求认可机构考虑实验室参加能力验证的情况及结果。其他的合格评定活动，诸如检查或产品认证，对能力验证的需求也在不断增长。适当时，本准则中的大部分要求也适用于上述领域，尤其是在管理、策划和设计、人员、质量的保证、保密性和其他方面。

能力验证提供者认可准则

1 范围

本准则规定了能力验证提供者所需能力和建立、运作能力验证计划的通用要求。这些要求适用于所有类型的能力验证计划,并可作为某个特殊应用领域特定技术要求的基础。

2 规范性引用文件

下列引用文件对本准则的应用是必不可少的。对注明日期的引用文件,只采用所引用的版本;对未注明日期的引用文件,采用其最新版本(包括任何修订)。

GB/T 27000-2006, 合格评定 词汇和通用原则 (ISO/IEC 17000: 2004, IDT)

ISO/IEC Guide 99: 2007, International Vocabulary of metrology—Basic and general concepts and associated terms (VIM)

3 术语和定义

GB/T 27000-2006、ISO/IEC Guide 99: 2007 中的术语和定义适用于本准则,此外,本准则还采用下列术语和定义。

3.1 指定值 assigned value

对能力验证物品的特定性质赋予的值。

注:在某些定性或半定量计划中,能力验证物品的特性不是以量值来表示。

3.2 协调者 coordinator

负责组织和管理能力验证计划运作中所有活动的一人或多人。

3.3 客户 customer

通过合同性协议获得能力验证计划的组织或个人。

3.4 实验室间比对 interlaboratory comparison

按照预先规定的条件,由两个或多个实验室对相同或类似的物品进行测量或检测的组织、实施和评价。

3.5 离群值 outlier

一组数据中被认为与该组其他数据不一致的观测值。

注:离群值可能来源于不同的总体,或由于不正确的记录或其他粗大误差的结果。

3.6 参加者 participant

接受能力验证物品并提交结果以供能力验证提供者评价的实验室、组织或个人。

注:在某些情况下,参加者可以是检查机构。

3.7 能力验证 proficiency testing

利用实验室间比对, 按照预先制定的准则评价参加者的能力。

注 1: 在本准则中, 术语“能力验证”具有极为广泛的含义, 包括但不限于以下类型:

- a) 定量计划 quantitative scheme
该类计划是确定能力验证物品的一个或多个被测量的量。
- b) 定性计划 qualitative scheme
该类计划是对能力验证物品的一个或多个特性进行鉴别或描述。
- c) 顺序计划 sequential scheme
该类计划是将检测或测量的一个或多个能力验证物品按顺序分发, 并按期返回能力验证提供者。
- d) 同步计划 simultaneous scheme
该类计划中, 分发能力验证物品, 在规定期限内同时进行检测或测量。
- e) 单次计划 single occasion exercise
该类计划中, 为单个需求提供能力验证物品。
- f) 连续计划 continuous scheme
该类计划中, 按规定间隔提供能力验证物品。
- g) 抽样 sampling
该类计划中, 为后续的分析抽取样品。
- h) 数据转换和解释 data transformation and interpretation
该类计划中, 提供成组的数据或其他信息, 要求对信息进行处理以给出解释(或其他结论)。

注 2: 在医学领域的某些能力验证提供者, 利用术语“外部质量评价(EQA)”表示其能力验证计划和/或更广义的计划(见附录 A)。但本准则中的要求只适用于符合能力验证定义的 EQA 活动。

3.8 能力验证物品 proficiency testing item

用于能力验证的样品、产品、人工制品、标准物质/标准样品、设备部件、测量标准、数据组或其他信息。

3.9 能力验证提供者 proficiency testing provider

对能力验证计划建立和运作中所有任务承担责任的组织。

3.10 能力验证轮次 proficiency testing round

向参加者发放能力验证物品、评价和报告结果的单个完整流程。

3.11 能力验证计划 proficiency testing scheme

在检测、测量、校准或检查的某个特定领域, 设计和运作的一轮或多轮能力验证。

注: 一项能力验证计划可以包含一种或多种特定类型的检测、校准或检查。

3.12 稳健统计方法 robust statistical method

对给定概率模型假定条件的微小偏离不敏感的统计方法。

3.13 能力评定标准差 standard deviation for proficiency assessment

基于可获得的信息，评价能力验证结果所用的分散性度量。

注 1：标准差只适用于比例尺度和定距尺度的结果。

注 2：并非所有的能力验证计划都根据结果的分散性进行评价。

3.14 分包方 subcontractor

能力验证提供者所用的，从事本准则中规定活动并影响能力验证计划质量的组织或个人。

注：术语“分包方”包括许多能力验证提供者所称的合作方。

3.15 计量溯源性 metrological traceability

通过一条形成文件的、具有规定测量不确定度及不间断的校准链，使测量结果与参考标准联系起来特性。

注 1：对于该定义，参考标准可以是通过实际复现的某个测量单位的定义，或是包括非序量测量单位的一个测量程序，或是一个测量标准；

注 2：计量溯源性要求建立校准等级图；

注 3：参考标准的技术要求中必须包括利用参考标准建立校准等级图的时间、以及有关该参考标准的其他计量信息，例如，在校准等级图中首次进行校准的时间；

注 4：对于测量模型中含有一个以上输入量的测量，其每个输入量值在计量上应能溯源，相关的校准等级图可以形成支岔结构或网络状。每个输入量值在建立计量溯源中的工作量，应与该量值对测量结果的贡献相对应；

注 5：测量结果的计量溯源性不能保证测量的不确定度适用于某个特定的目的，或不存在差错；

注 6：如果将比对用于核查，则两个测量标准之间的比对可以视为校准，必要时，可对测量标准之一的量值和测量不确定度进行修正；

注 7：ILAC 认为，确认计量溯源性的要素有：与国际测量标准或国家测量标准相联系的一条不间断的计量溯源链；形成文件的测量不确定度；形成文件的测量程序；认可的技术能力；对 SI 的计量溯源性；以及校准的间隔（见 ILAC P-10: 2002）；

注 8：简称的术语“溯源性”，既可表示“计量溯源性”，有时也可表示其他概念，诸如“样品溯源性”、“文件溯源性”、“材料溯源性”，这些都代表了一件物品的历史（溯源）。所以，如有产生混淆的风险时，最好采用术语的全称“计量溯源性”。

[ISO/IEC Guide 99: 2007, 定义 2.4.1]

3.16 测量不确定度 measurement uncertainty/ uncertainty of measurement

不确定度 uncertainty

根据所用到的信息，表征赋予被测量的量值分散性的非负参数。

注 1: 测量不确定度包括源自系统影响的不确定度分量, 诸如与修正值和测量标准指定量值相伴随的不确定度分量, 以及定义的不确定度。有时, 并不修正估计的系统影响, 取而代之是将系统影响与伴随的测量不确定度结合在一起;

注 2: 参数可以是称为标准测量不确定度的标准差 (或其特定的倍数), 或是具有规定包含因子概率的半宽区间;

注 3: 测量不确定度通常包含多个分量, 其中一些分量可根据系列测量的量值的统计分布, 采用 A 类不确定度评定进行评估, 并用标准差表征。其他一些分量则采用 B 类不确定度评定, 由基于经验或其他信息的概率密度函数进行评估, 也可用标准差表征;

注 4: 通常, 对于给出的一组信息, 测量不确定度应伴随被测量所宣称的量值, 该值的改变也会造成与其相伴的不确定度的改变。

[ISO/IEC Guide 99: 2007, 定义 2.26]

4 技术要求

4.1 总则

能力验证计划的建立和运作应由具有能力实施实验室间比对, 并能获得特定类型能力验证物品的专业技术的能力验证提供者承担。能力验证提供者或其分包方还应具备对于待测特性的测量能力。

注: 符合 GB/T 27025 或 GB/T 22576 可证明能力验证提供者实验室的能力, 或承担能力验证计划相关检测或测量工作的分包方实验室的能力。符合 GB/T 15000.7 可证明提供能力验证物品的标准物质/标准样品生产者的能力。

4.2 人员

4.2.1 能力验证提供者应有管理和技术人员, 他们具有履行其职责所需权限、资源和技术能力。

4.2.2 能力验证提供者管理层应规定组织中关键岗位所需资格和经验的最低要求, 并确保人员满足要求。

4.2.3 能力验证提供者应使用长期雇用人员或签约人员。在使用签约人员及其他的技术人员及关键支持人员时, 能力验证提供者应确保这些人员是胜任的且受到监督, 并按照能力验证提供者管理体系要求工作。

注: 对用于某专项工作的或作为顾问组或指导小组 (见 4.4.1.4) 成员的技术专家, 通过诸如小组职责范围描述或其它方式签有正式协议, 可视为满足本要求。

4.2.4 能力验证提供者应授权专门人员从事如下工作:

- a) 选择适当的能力验证物品;
- b) 策划能力验证计划;
- c) 进行特定类型的抽样;

- d) 操作特定的仪器设备;
- e) 进行测量以确定能力验证物品被测量的稳定性和均匀性、指定值及其不确定度;
- f) 制备、处置并分发能力验证物品;
- g) 操作数据处理系统;
- h) 进行统计分析;
- i) 评价能力验证参加者的能力;
- j) 提供意见和解释;
- k) 批准发布能力验证报告。

4.2.5 能力验证提供者应保留所有技术人员（包括签约人员）的相关授权、能力、教育和专业资格、培训、技能和经验的最新记录，这些信息要易于获取，并应包括其承担相应工作的能力评价和确认日期。

4.2.6 能力验证提供者应为运作能力验证计划的每个人员制定教育、培训和技能的目标。能力验证提供者应有确定培训需求和提供人员培训的政策和程序。培训计划应与能力验证提供者当前和预期的任务相适应。

注：建议考虑人员定期再培训需求。人员培训政策宜考虑技术更新和证明持续能力的需求，并以持续提高技能为目的。

4.2.7 能力验证提供者应确保工作人员获得必要的培训，以确保人员能够胜任测量、仪器设备操作以及其它影响能力验证计划质量的工作，应评价培训活动的有效性。

注：可使用客观尺度来评定人员所达到的能力。

4.3 设备、设施和环境

4.3.1 能力验证提供者应确保具有与能力验证计划运作相适应的设施，主要包括用于能力验证物品的制造、处置、校准、检测、储存和分发，以及数据处理、通讯、资料记录和检索的设施和设备。

4.3.2 能力验证提供者应确保环境条件不会影响能力验证计划或所要求的运作质量。当能力验证计划是在能力验证提供者固定设施以外的场所或委托给分包方运作时，应予以特别注意。应将影响能力验证计划的设施和环境条件的技术要求制定成文件。

4.3.3 应对影响能力验证计划质量的区域的进入和使用加以控制。能力验证提供者应根据其特定情况确定控制的程度。

4.3.4 能力验证提供者应识别可能显著影响能力验证物品、检测和校准质量的环境条件，包括规范和测量程序所要求的条件。提供者应对这些条件加以控制和监测，并应记录所有相关监控活动。当环境条件危及到能力验证计划的质量或运作时，应停止相关能力验证活动。

注：环境条件包括诸如生物消毒、灰尘、电磁干扰、辐射、湿度、供电、温度、声级、振级，使其适应于相关的技术活动。

4.3.5 应将不相容活动的相邻区域进行有效隔离。应采取措施以防止交叉污染。

4.3.6 能力验证提供者应确保用于确定能力验证物品的含量、均匀性和稳定性的实验室方法和设备的性能特征得到适当地确认和维持。

4.4 能力验证计划的设计

4.4.1 策划

4.4.1.1 能力验证提供者应识别和策划直接影响能力验证计划质量的过程，并确保这些过程依据既定程序进行。

注：在建立一项计划及相关信息时也可考虑相关方的利益。

4.4.1.2 能力验证提供者不应将能力验证计划的策划工作分包（见 5.5.2）。

注：能力验证提供者可向顾问、专家或指导小组寻求建议和帮助（见 4.4.1.4）。

4.4.1.3 能力验证提供者应在能力验证计划开始之前制定文件化的方案，说明本次能力验证计划的目标、目的以及基本设计情况，并提供以下信息，必要时应说明其取舍理由：

- a) 能力验证提供者的名称和地址；
- b) 协调者以及其他参与能力验证计划设计和运作人员的姓名、地址和联系方式；
- c) 分包的活动以及参与能力验证计划运作分包方的名称和地址；
- d) 参加计划应满足的条件；
- e) 能力验证计划预期的参加者数量和类型；
- f) 所选定的被测量或特性，包括参加者需要鉴别、测量或检测的有关信息；
- g) 对能力验证物品预期的量值范围和/或特性的描述；
- h) 所提供能力验证领域中涉及的潜在的主要误差来源；
- i) 对能力验证物品生产、质量控制、存储、分发的要求；
- j) 防止参加者串通或伪造结果的合理预警措施，以及当怀疑串通或伪造时可执行的程序；
- k) 将提供给参加者的信息描述，以及能力验证计划各阶段时间表；
- l) 对于连续能力验证计划，给参加者分发能力验证物品的频次或日期，参加者返回结果的截止日期，若有必要，参加者进行检测或测量的日期；
- m) 参加者准备检测材料以及进行检测或测量所使用的方法或程序的有关信息；
- n) 用于能力验证物品均匀性和稳定性检验的检测或测量方法的程序，必要时确定其生物活性；
- o) 为参加者准备其使用的所有标准化报告格式；
- p) 所使用统计分析的详细描述；
- q) 所有指定值的来源、计量溯源性和测量不确定度；
- r) 参加者能力评价的准则；
- s) 返回给参加者的数据、中期报告或信息的描述；
- t) 参加者结果和根据能力验证计划结果所做结论的公布范围描述；
- u) 能力验证物品丢失或损坏时应采取的措施。

4.4.1.4 能力验证提供者应能获得所需的专业技术和检测、校准、抽样、检查以及统计分析等相关领域的经验。如有必要，可通过成立顾问组（可适当命名）来实现。

4.4.1.5 适当时，应利用专业技术确定以下几方面的事宜：

- a) 4.4.1.3中所列的策划要求；
- b) 识别和解决在制备和保持均匀的能力验证物品方面，或在为能力验证物品提供稳定的指定值方面可预料到的困难；
- c) 为参加者准备详细的指导书；
- d) 对前几轮能力验证中参加者所提出的所有技术难题或意见进行评议；
- e) 为评定参加者的能力提供建议；
- f) 对参加者结果和能力进行整体评述，适当时，分组或逐一评议；
- g) 为参加者分别提供或在报告中提供建议（在保密许可范围内）；
- h) 回复参加者的反馈；
- i) 策划或参加与参加者的技术会议。

4.4.2 能力验证物品的制备

4.4.2.1 能力验证提供者应建立并执行程序，以确保能力验证物品制备按照 4.4.1 描述的方案进行。

注：建议能力验证提供者充分考虑制备足够数量的能力验证物品，以便替代分发期间丢失或损坏的能力验证物品，或者在完成结果评价之后，用于对参加者的培训或者用作标准物质/标准样品等。

4.4.2.2 能力验证提供者应建立并执行程序，确保所有能力验证物品以恰当的方式获得、收集、制备、处置、储存，需要时包括废弃物的处理。该程序应确保制备能力验证物品所用材料的获取符合相关法规和伦理道德要求。

4.4.2.3 能力验证物品的基体、被测量和浓度等应尽可能地与日常检测或校准物品和材料的类型相似。

4.4.2.4 当能力验证计划中需要参加者将制备和/或处理后的能力验证物品再提交给能力验证提供者时，能力验证提供者应提供能力验证物品的制备、包装及运输说明书。

4.4.3 均匀性和稳定性

4.4.3.1 应基于不均匀性和不稳定性对参加者能力评定可能产生的影响，建立合适的均匀性和稳定性判定准则。

注 1：本条要求旨在确保每个参加者都收到具有可比性的能力验证物品，并且这些能力验证物品在整个能力验证过程中保持稳定。为此，需要周密的策划、制备和运输，且通常需要通过检验来确认。

注 2：某些情况下，无法对能力验证物品开展均匀性和稳定性检验，例如可用于制备能力验证物品的材料非常有限。

注 3: 某些情况下, 仅可获得不完全均匀或稳定的材料, 如果在指定值的不确定度中或者结果评价时考虑到这一点, 这些材料仍可用作能力验证物品。(见 B. 3. 1. 3 以及 ISO 13528: 2005 附录 B)

注 4: 均匀性和稳定性在 GB/T 15000. 7、GB/T 15000. 3 和 ISO 13528 中有更进一步的讨论。

4. 4. 3. 2 应将均匀性和稳定性的评定程序形成文件并执行, 只要可行, 应根据适用的统计设计进行评定。如可能, 能力验证提供者应根据统计方法, 从整批检测材料中随机抽取一定数量、具有代表性的能力验证物品来评定材料的均匀性。

注: 某些情况下, 从整批能力验证物品中采用随机分层抽样或系统抽样更为合适。

4. 4. 3. 3 均匀性评定通常应在能力验证物品被包装成最终形式之后、分发给参加者之前进行, 除非不可行, 例如稳定性研究表明必须以散装的方式来保存。

注 1: 当预计包装不会影响均匀性时, 可在样品包装之前进行均匀性检验。

注 2: 某些情况下, 由于实际的、技术的或后勤保障的原因, 均匀性检验在样品分发前无法实施。

4. 4. 3. 4 应证实能力验证物品足够稳定, 确保这些材料在能力验证实施过程(如储存和运输条件)中不会发生明显变化。如果无法保证, 则应将稳定性量化, 并考虑作为能力验证物品指定值测量不确定度的补充分量, 必要时, 在能力评定准则中予以考虑。

4. 4. 3. 5 若使用先前轮次中保留的能力验证物品, 在样品分发前, 能力验证提供者应确认本次能力验证计划中待测的特性值。

4. 4. 3. 6 当无法进行均匀性和稳定性检验时, 能力验证提供者应证明能力验证物品的收集、制备、包装和分发程序可以充分满足能力验证要求。

4. 4. 4 统计设计

4. 4. 4. 1 应根据数据的特性(定量或定性, 包括排序和分类)、统计假定、误差的性质以及预期的结果数量, 制定和建立符合能力验证目标的统计设计。

注 1: 统计设计涵盖了能力验证计划的数据获取方案的策划、数据的收集、分析和报告等过程。统计设计通常基于能力验证计划的既定目标, 比如确定已规定了大小的误差, 或确定带有规定了测量不确定度的指定值。

注 2: 数据分析方法可能非常简单(如描述统计学), 也可能由于利用带有概率假设或不同能力验证物品结果组合的统计模型而非常复杂。

注 3: 当能力验证计划根据委托人(如客户、监管机构或认可机构)给定的规范设计时, 可直接使用规范中的统计设计和数据分析方法。

注 4: 在统计设计缺少可靠信息时, 可利用先期的实验室间比对来获得。

4. 4. 4. 2 能力验证提供者应将用来确定指定值和评价参加者测量结果的统计设计和数据分析方法形成文件, 并对选用的原因和假定进行说明。能力验证提供者应能证实统计假定合理并确保统计分析按照规定程序进行。

4. 4. 4. 3 在设计统计分析时, 能力验证提供者应仔细考虑下列事项:

- a) 能力验证中每个被测量或特性所要求或期望的准确度（正确度和精密度）以及测量不确定度；
- b) 达到统计设计目标所需的最少参加者数量；当参加者数量不足以达到目标或不能对结果进行有意义的统计分析时，提供者应将评定参加者能力的替代方法细则形成文件，并提供给参加者；
- c) 有效数字和/或小数位数与所报告结果的相关性；
- d) 需要检测或测量的能力验证物品数量，以及对每个能力验证物品或每项测定的检测、校准或测量的重复次数；
- e) 用于确定能力评定标准差或其它评定准则的程序；
- f) 用于识别和（或）处理离群值的程序；
- g) 只要适用，对统计分析中剔除值的评价程序；
- h) 只要适当，与设计相符的目标和能力验证轮次的频率。

4.4.5 指定值

4.4.5.1 能力验证提供者应将确定能力验证计划被测量或特性指定值的程序形成文件。该程序应考虑所需的计量溯源性和测量不确定度，以证明能力验证计划符合其目标。

注：计量溯源性并不是在所有情况都可行或适用。

4.4.5.2 校准领域的能力验证计划的指定值应具有计量溯源性和测量不确定度。

4.4.5.3 对于校准领域以外的能力验证计划，应考虑参加者或其他利益相关方的特定需求，或根据能力验证计划的设计来确定指定值的计量溯源性和测量不确定度的需求、关联性和可行性。

注：所需的计量溯源链可根据能力验证物品类型、被测量或特性以及可溯源校准和标准物质/标准样品可获得性而有所不同。

4.4.5.4 若将公议值作为指定值，能力验证提供者应将原因形成文件，并应根据能力验证计划方案来评估指定值不确定度。

4.4.5.5 能力验证提供者应有关于指定值披露的政策。该政策应确保参加者不能从早期的披露中获益。

4.5 方法或程序的选择

4.5.1 通常期望参加者使用他们自己选择的检测方法、校准或测量程序，这些方法或程序应与其日常所使用的一致。能力验证提供者可依据能力验证计划的设计要求参加者使用指定的方法。

4.5.2 若允许参加者使用他们自己选择的方法，能力验证提供者应：

- a) 制定政策并按程序对不同检测或测量方法得到的结果进行比对；
- b) 了解每一被测量所用不同检测或测量方法的技术等效性，并采取措施评价不同方法的参加者结果。

4.6 能力验证计划的运作

4.6.1 给参加者的指导

4.6.1.1 能力验证提供者应在发放能力验证物品前及早告知参加者能力验证物品可能到达或将要分发的日期，除非能力验证计划的设计不适于这样做。

4.6.1.2 能力验证提供者应向所有参加者提供详细的文件化的指导书。指导书应包括：

- a) 要求参加者按照大多数日常检测样品的处理方式处理能力验证物品（除非能力验证计划有不同于该原则的特定要求）；
- b) 能力验证物品检测或校准影响因素的详细说明，例如：能力验证物品的性质、存储条件、是否限定检测方法，以及检测或测量的时间要求；
- c) 进行检测或校准之前，能力验证物品的准备和/或状态调节的详细程序；
- d) 处置能力验证物品的适当指导，包括安全要求；
- e) 参加者检测和/或校准时特定的环境条件，如适用，要求参加者报告测量期间相关环境条件；
- f) 检测或测量结果及其不确定度记录和报告方式的明确和详细的说明。如果指导书要求报告结果的测量不确定度，应包括包含因子和置信概率（只要可行）；
注：指导书通常包括测量单位、有效数字或小数位数、报告的依据（如按千基重量计或按“收到基”计）等参数。
- g) 提供者接收用于分析的能力验证结果的截止日期；
- h) 能力验证提供者用于问询的详细联络信息；
- i) 适用时，返回能力验证物品的说明。

4.6.2 能力验证物品的处置和存储

4.6.2.1 能力验证提供者应确保能力验证物品从制备到分发的过程中得到妥善标识、隔离和免受污染或降解。

4.6.2.2 能力验证提供者应确保提供安全的存储区域和/或存储室，防止能力验证物品在制备和分发期间受损或变质。应制定授权从这些区域收发能力验证物品的合适程序。

4.6.2.3 适当时，为发现存储期间可能的变质，应定期检查存储或存放能力验证物品、化学品和材料的条件。

4.6.2.4 如果使用有潜在危险的能力验证物品、化学品和材料，应有设施保证其安全处置、去污染和废弃处理。

4.6.3 能力验证物品的包装、标识和分发

4.6.3.1 能力验证提供者应控制包装和标记过程，以确保符合有关国家、地区或国际的安全和运输要求。

注：对于某些类型的材料而言，正确分发能力验证物品可以避免严重问题的发生。例如，要求持续冷藏或不能暴露于 X 射线下、碰撞或振动等。大部分的化学材料采用

气密包装，以避免受到空气中污染物的污染，例如，在运输中会接触到燃料蒸汽或发动机排放的废气。

4.6.3.2 能力验证提供者应规定能力验证物品运输相关的环境条件。必要时，提供者应监视运输过程中的这些环境条件，评估其对能力验证物品的影响。

4.6.3.3 在能力验证计划中，若要求参加者将能力验证物品传递给其他参加者，对于该传递，应向参加者提供文件化的指导书。

4.6.3.4 能力验证提供者应确保每个能力验证物品的包装上牢固地贴有标识，设计时应确保标识在整个能力验证期间保持清晰和完整。

4.6.3.5 能力验证提供者应按照程序来确认能力验证物品的送达。

注：参加者如果没有按期收到能力验证物品，可以按照 4.6.1.1，告知能力验证提供者。

4.7 数据分析和能力验证计划结果的评价

4.7.1 数据分析和记录

4.7.1.1 所有的数据处理设备和软件在投入使用前，应依据程序确认。计算机系统维护应包括备份操作和系统恢复方案，并应记录维护和运作检查的结果。

4.7.1.2 应运用适当的方法记录和分析参加者提交的结果。应建立和执行程序以检查数据输入、传输、统计分析和报告的有效性。

4.7.1.3 数据分析应给出与能力验证计划统计设计相符的总计统计量、能力统计量以及有关信息。

4.7.1.4 应使用稳健统计方法或检出统计离群值的适当方法，使离群值对总计统计量的影响降至最低。

4.7.1.5 能力验证提供者应有文件化的准则和程序，以处理不适合统计评价的检测结果，例如：计算错误、转换错误和其他粗大误差。

4.7.1.6 能力验证提供者应有文件化的准则和程序，以识别和处理在分发后才发现的不适合能力评定的能力验证物品，例如，能力验证物品不均匀、不稳定、损坏或被污染。

4.7.2 能力评定

4.7.2.1 能力验证提供者应运用符合能力验证计划目标的有效评定方法。该方法应形成文件并包含对评定依据的描述。能力评定不应分包(见 5.5.2)。

4.7.2.2 适当时，为达到能力验证计划的目的，能力验证提供者应对参加者的能力提供以下方面的专家评议：

- a) 将整体表现与预期比较，并考虑测量不确定度；
- b) 室内及室间的变异，及其与先前能力验证轮次、类似能力验证计划、或已公布的精密度数据的比较；
- c) 方法或程序间的差异；
- d) 误差的可能来源（针对离群值）和提高能力的建议；

- e) 给参加者的建议和指导性反馈, 作为参加者持续改进程序的一部分;
- f) 由于非正常的因素, 无法进行结果评价和能力评定的情况;
- g) 其他的建议、意见或总体评述;
- h) 结论。

注: 在某个特定的能力验证计划实施过程中或完成之后, 定期向参加者提供单独的汇总表是有用的。汇总表可包括各参加者在连续的能力验证计划的各轮次中能力的最新汇总。如需要, 可对这些汇总做进一步分析并强调其趋势。

4.8 报告

4.8.1 能力验证报告应清晰、全面, 包含所有参加者结果的资料, 并指出每个参加者的能力。最终报告的批准不应分包(见 5.5.2)。

注: 如果不能为参加者提供所有原始资料, 那么可以提供一个结果汇总, 例如, 以表格或图形的形式。

4.8.2 除非不适用或能力验证提供者有正当理由, 否则报告应包括以下内容:

- a) 提供者的名称和详细联系信息;
- b) 协调者的姓名和详细联系信息;
- c) 报告批准人的姓名、职位、签名或等效标识;
- d) 提供者分包活动的说明;
- e) 报告发布日期和状态(如初期的、中期的或最终的);
- f) 报告的页码和清晰的结束标记;
- g) 结果保密程度的声明;
- h) 能力验证计划报告的编号和清晰标识;
- i) 对能力验证物品的清晰描述, 包括能力验证物品制备、均匀性和稳定性评定的必要细节;
- j) 参加者的结果;
- k) 统计数据及总计统计量, 包括指定值、可接受结果的范围和图形表示;
- l) 用于确定指定值的程序;
- m) 指定值的计量溯源性和测量不确定度的详细信息;
- n) 用于确定能力评定标准差或其它评定准则的程序;
- o) 对应每组参加者使用的检测方法/程序的指定值和总计统计量(如果不同组的参加者使用了不同的方法);
- p) 能力验证提供者和技术顾问对参加者的能力评述;
- q) 能力验证计划设计和实施的信息;
- r) 数据统计分析的程序;
- s) 对统计分析解释的建议;
- t) 基于本轮能力验证结果的评述或建议。

注：对于连续能力验证计划，提供较简单的报告即可，上述很多内容在常规报告中可以省略，但应包含在参加者可获得的能力验证计划协议或阶段性的汇总报告中。

4.8.3 应使参加者在规定时间内获得报告。对于运作周期可能很长或涉及易腐坏材料的计划，可在最终结果发布之前提供初步或预期结果。

注：这样可以及早调查可能的错误。

4.8.4 能力验证提供者应有个人和组织使用报告的政策。

4.8.5 当有必要发布一个能力验证计划新的或修正版的报告时，应包括以下内容：

- a) 唯一性的标识；
- b) 所要替换或修改的原始报告的引用；
- c) 修改和重新发布原因的说明。

4.9 与参加者的沟通

4.9.1 能力验证提供者应提供有关能力验证计划的详细信息。包括如下内容：

- a) 能力验证计划范围的详细信息；
- b) 参加者应支付的各项费用；
- c) 参加条件的文件化规定；
- d) 保密措施；
- e) 申请细则。

4.9.2 能力验证提供者应迅速将能力验证计划设计或运作中的所有变化通知参加者。

4.9.3 应制定文件化的程序，使参加者能够对能力验证计划中有关其能力评定提出申诉，应将申诉的途径告知能力验证计划参加者。

4.9.4 适当时，与参加者交流的相关记录应保留并归档。

4.9.5 若能力验证提供者发布参加声明或能力声明，应包含足够的信息，以免引起误导。

4.10 保密

4.10.1 除非参加者宣布放弃保密，否则应对能力验证计划参加者的身份保密；其身份应仅为参与能力验证计划运作的相关人员所知。

4.10.2 参加者向能力验证提供者提供的所有信息都应视为保密信息。

注：参加者可以选择在能力验证计划范围内放弃保密，以便能够互相交流和帮助，例如为了提高能力。出于管理或认可的目的，参加者也可以放弃保密。在大多数情况下，参加者可以自己向相关权威机构提供能力验证结果。

4.10.3 若某个利益相关方要求能力验证提供者直接向其提供能力验证结果，应将该安排在计划开始前告知参加者。

4.10.4 例外情况下，若管理机构要求能力验证提供者直接向其提供能力验证结果，提供者应书面通知受影响的参加者。

5 管理要求

5.1 组织

5.1.1 能力验证提供者或其所在组织，应是一个具有法律地位和能够承担法律责任的实体。

5.1.2 能力验证提供者有责任确保所提供能力验证活动符合本准则要求，并满足参加者、法定管理机构和对其提供承认的组织的需求。

5.1.3 能力验证提供者管理体系应覆盖其在固定设施内、离开其固定设施的场所和在相关临时设施中进行的工作。

5.1.4 如果能力验证提供者所在的组织还从事其它活动，为识别潜在利益冲突，能力验证提供者应确定该组织中涉及能力验证活动，或对能力验证活动有影响的关键人员的职责。当识别出潜在利益冲突时，能力验证提供者应实施相关程序以确保其所有活动的公正性。

5.1.5 能力验证提供者应：

- a) 有管理人员和技术人员，不论他们的其他责任，他们应具有所需的权力和资源来履行包括实施、保持和改进管理体系的职责，识别对于管理体系或提供能力验证计划的各项程序的偏离，以及采取措施预防或减少这些偏离；
- b) 有措施确保其管理层和员工不受任何对工作质量有不良影响的、来自内外部的不正当的商业、财务和其他压力；
- c) 有保护客户的机密信息和所有权的政策和程序，包括保护电子存储和传输结果的程序；
- d) 有政策和程序以避免卷入任何会降低其在能力、公正性、判断力或运作诚实性方面的可信度的活动；
- e) 确定组织和管理结构、在母体组织中的地位，以及质量管理、技术运作和支持服务之间的关系；
- f) 规定对能力验证计划运作质量有影响的所有管理、操作和核查人员的职责、权力、相互关系和所需的能力；
- g) 确保全体人员理解他们活动的相关性和重要性，以及如何为实现管理体系目标做出贡献；
- h) 由熟悉各项活动程序的人员，对技术人员包括在培员工，进行充分地监督；
- i) 有技术管理者，全面负责技术运作和提供确保能力验证计划运作质量所需的资源，包括提供在相关检测、校准或检查领域以及4.4.1.4中统计方面的必要专业知识和经验方面的支持；
- j) 指定一名员工作为质量主管（适当的称谓），不论其他职责，应赋予其在任何时候都能确保管理体系得到实施和遵循的责任和权力。质量主管应有直接渠道接触决定能力验证提供者政策或资源的最高管理者；

k) 指定关键管理人员的代理人。

注：在能力验证提供者人员数量少时，一个人可能有多项职能，对所有主要职能都指定代理人可能不现实。

5.1.6 最高管理者应确保在组织内部建立适宜的沟通机制，并就管理体系有效性事宜进行沟通。

5.2 管理体系

5.2.1 能力验证提供者应建立、实施和保持与其活动范围相适应的管理体系，这些活动包括所提供能力验证的类型、范围和数量。

5.2.2 能力验证提供者对其政策、计划、程序和指导书的规定和文件化程度，应能满足保证能力验证各方面质量的需要。体系文件应传达至有关人员，并被其理解、获取和执行。

注：这些方面可包括但不限于：能力验证物品质量（如均匀性和稳定性）、特性（如设备校准和方法确认）、特性值的指定（如使用适当的统计程序）、参加者能力评价、能力验证物品的分发、储存和运输程序、检测结果统计处理以及报告。

5.2.3 能力验证提供者管理体系与质量有关的各项政策，包括质量方针声明，应在质量手册（适当的称谓）中阐明。应制定总体目标并在管理评审时加以评审。质量方针声明应由最高管理者授权发布，至少包括下列内容：

- a) 管理者向参加者及其它客户提供能力验证服务的质量的承诺；
- b) 管理者关于服务标准的声明；
- c) 与质量有关的管理体系的目的；
- d) 要求与能力验证活动有关的所有人员熟悉质量文件，并在工作中执行这些政策和程序；
- e) 管理者对遵守本准则及持续改进管理体系有效性的承诺。

5.2.4 最高管理者应提供建立和实施管理体系以及持续改进其有效性承诺的证据。

5.2.5 最高管理者应将满足客户要求和法定要求的重要性传达到组织。

5.2.6 质量手册应包括或指明含技术程序在内的支持性程序，并概述管理体系中所用文件的架构。

5.2.7 质量手册中应规定技术管理者和质量主管的作用和责任，包括确保遵循本准则的责任。

5.2.8 当策划和实施管理体系的变更时，最高管理者应确保保持管理体系的完整性。

5.3 文件控制

5.3.1 总则

能力验证提供者应建立和保持程序来控制构成其管理体系的所有文件（内部制订或来自外部的），诸如法规、标准、其他规范化文件、能力验证计划协议、检测或校准方法，以及图纸、软件、规范、指导书和手册。

5.3.2 文件的批准和发布

5.3.2.1 凡作为管理体系组成部分发布的所有文件,在发布之前应由授权人员审查并批准使用;提供者应建立用于识别管理体系文件的当前修订状态和分发的清单,或等效的文件控制程序,并使之易于获取,以防止使用无效和(或)作废的文件。

5.3.2.2 文件控制程序应确保:

- a) 在对能力验证计划有效运作起重要作用的所有作业场所都能得到相应文件的授权版本;
- b) 定期审查文件,必要时进行更新,以确保其持续适用并满足使用要求;
- c) 及时地从所有使用或发布处撤除无效或作废文件,或用其他方法保证防止误用;
- d) 出于法律或知识保存目的而保留的作废文件,应有适当的标记。

5.3.2.3 能力验证提供者制订的管理体系文件应有唯一性标识。该标识应包括发布日期和(或)修订标识、页码、总页数或表示文件结束的标记和发布机构。

5.3.3 文件变更

5.3.3.1 除非另有特别指定,文件的变更应由原审查和批准责任人进行审查和批准。被指定的人员应获得进行审查和批准所依据的有关背景资料。

5.3.3.2 若可行,更改的或新的内容应在文件或适当的附件中标明。

5.3.3.3 如果能力验证提供者的文件控制系统允许在文件再版之前对文件进行手写修改,则应确定修改的程序和权限。修改之处应有清晰的标注、签名缩写并注明日期。修订的文件应尽快发布。

5.3.3.4 应制订程序来描述如何更改和控制保存在计算机系统文件中的文件。

5.4 要求、标书和合同的评审

5.4.1 能力验证提供者应建立和保持对要求、标书和合同的评审政策和程序。这些评审应确保:

- a) 对包括所用的检测和校准方法、测量设备及能力验证物品在内的要求应予以充分规定,形成文件,并易于理解;
- b) 能力验证提供者有能力和资源满足这些要求;
- c) 能力验证计划在技术上是可行的。

注 1: 当客户要求为特定目的而开展能力验证计划,或与通常提供的计划在不同水平、有不同参加频次要求时,该评审显得尤为重要。

注 2: 当能力验证计划在目录或其他公告中已有详尽描述,且参加者加入例行计划时,该评审可简化。

5.4.2 应保存包括任何变化在内的评审记录。在执行合同期间,就客户的要求或工作结果与客户进行讨论的有关记录,也应予以保存。

5.4.3 评审的内容应覆盖要求的所有方面,包括被能力验证提供者分包出去的任何工作。

5.4.4 适用时,对合同或商定的能力验证计划方案的任何偏离均应通知参加者和其他客户。

5.4.5 能力验证计划开始后如果修改要求或合同,应重复进行同样的合同评审过程,并将所有修改内容通知所有受到影响的人员。

5.5 分包服务

5.5.1 当能力验证提供者需将工作分包时,应能证明分包方的经验和技術能力能够胜任指定的任务,并符合本准则和其它适用标准的相关条款。

5.5.2 能力验证提供者不得分包能力验证计划的策划(见 4.4.1.2)、能力评价(见 4.7.2.1)或最终报告的批准(见 4.8.1)。

注:这并不排除能力验证提供者使用任何来自顾问、专家或指导小组的建议和协助。

5.5.3 能力验证提供者应将分包或可能分包的服务提前以书面的形式通知参加者。

注:例如,该通知可以在能力验证计划文件中采取以下形式的声明:“能力验证计划的各个方面可能有时被分包。分包时,会分包给有能力的分包方,并且能力验证提供者对该工作负责”。

5.5.4 能力验证提供者应就分包工作对参加者和其他客户负责,由法定管理机构指定的分包方除外。

5.5.5 能力验证提供者应保存能力验证计划运行中使用的所有分包方的登记表,包括分包的范围,并保存有关工作按照本准则和其他适用标准进行的能力评价记录。

5.6 服务和供应品的采购

5.6.1 能力验证提供者应有选择其所使用的、影响能力验证计划质量的服务和供应品的政策和程序。该程序应包含试剂、能力验证物品、标准物质/标准样品,以及与能力验证计划有关的其他消耗材料的采购、接收和存储。

5.6.2 能力验证提供者应确保所采购的、影响能力验证计划质量的供应品、设备和消耗材料,只有在经检查或以其他方式验证了符合规范或要求之后才投入使用。应保存所采取的符合性检查活动的记录。

5.6.3 影响能力验证计划质量的物品的采购文件,应包含描述所购服务和供应品的信息。这些采购文件在发出之前,其技术内容应经过审查和批准。

5.6.4 能力验证提供者应对影响能力验证计划质量的重要供应品和服务的供应商进行评价,并保存这些评价的记录和获批准的供应商名单。

注:可以理解,某些能力验证提供者可能要根据其母公司或依托机构规定的政策执行其采购程序。

5.7 服务客户

5.7.1 在确保为参加者保密的前提下,能力验证提供者应在澄清客户疑问和允许客户监视其相关工作表现方面积极与参加者和其他客户合作。

5.7.2 能力验证提供者应向客户征求反馈，无论是正面的还是负面的；应分析和利用这些反馈，以改进管理体系、能力验证计划及客户服务。

注：反馈的类型示例包括：客户满意度调查、与客户一起复查能力验证报告。

5.8 投诉和申诉

能力验证提供者应有政策并按照程序处理来自参加者、客户或其他方面的投诉；应保存所有投诉和申诉的记录以及能力验证提供者所开展调查和纠正措施的记录。

5.9 不符合工作的控制

5.9.1 当工作的任何方面不符合其程序或与客户达成一致的要求时，能力验证提供者应实施既定的政策和程序。该政策和程序应确保：

- a) 确定对不符合工作进行管理的责任和权力，规定当识别出不符合工作时所采取的措施（包括必要时暂停实施中的计划工作、扣发报告）；
对不符合工作的严重性进行评价；
- c) 立即对需要采取的措施和时限作出决定，同时对不符合工作的可接受性作出决定；
- d) 适合时，通知能力验证计划的参与者和其他客户，召回或取消已经发送给参加者的不合格能力验证物品或报告；
- e) 规定批准恢复工作的职责。

注：对管理体系或技术活动的不符合工作或问题的识别，可能发生在管理体系和技术运作的各个环节，例如参加者的投诉、管理评审和内部或外部审核、质量控制、能力验证物品的制备、均匀性和稳定性检验、数据分析、给参加者的指导书，及材料的处理和存储。

5.9.2 当评价表明不符合工作可能再度发生，或对能力验证提供者或分包方对政策和程序的符合性产生怀疑时，应立即执行 5.11 中规定的纠正措施程序。

5.10 改进

能力验证提供者应通过利用质量方针、各项质量目标、审核结果、数据分析、纠正措施、预防措施和管理评审来持续改进管理体系的有效性。

5.11 纠正措施

5.11.1 总则

能力验证提供者应制定政策和程序并指定适当的人员，以在识别出不符合工作和对管理体系或技术运作的政策和程序的偏离后实施纠正措施。

注：见 5.9.1 的注。

5.11.2 原因分析

纠正措施程序应从确定问题根本原因的调查开始。

注：原因分析是纠正措施程序中最关键有时也是最困难的部分。根本原因通常并不明显，因此需要仔细分析产生问题的所有潜在原因。潜在原因可包括：客户要求、能力验证物品及其特性、方法和程序、员工技能和培训、消耗性供应品、能力验证物

品的制备、均匀性和稳定性检验、统计设计、给参加者的指导书，以及材料的处理和储存。

5.11.3 纠正措施的选择和实施

5.11.3.1 需要采取纠正措施时，能力验证提供者应识别出各项可能的纠正措施，并选择和实施最可能消除问题和防止问题再次发生的措施。

5.11.3.2 纠正措施应与问题的严重程度和风险大小相适应。

5.11.3.3 能力验证提供者应将纠正措施调查所引起的任何变更制定成文件并加以实施。

5.11.4 纠正措施的监控

能力验证提供者应对纠正措施的结果进行监控，以确保采取的纠正措施有效。

5.11.5 附加审核

当对不符合或偏离的识别，导致对能力验证提供者符合其政策和程序，或符合本准则产生怀疑时，能力验证提供者应尽快依据 5.14 条的规定对相关活动区域进行审核。

注：附加审核常在纠正措施实施后进行，以确定纠正措施的有效性。仅在识别出问题严重或对能力验证计划有危害时，才有必要进行附加审核。

5.12 预防措施

5.12.1 应识别技术方面和管理体系方面所需的改进和潜在不符合的原因。当识别出改进机会或需采取预防措施时，应制定、执行和监控这些措施计划，以减少类似不符合情况发生的可能性并借机改进。

5.12.2 预防措施程序应包括措施的启动和控制，以确保其有效性。

5.13 记录的控制

5.13.1 总则

5.13.1.1 能力验证提供者应建立和保持识别、收集、索引、存取、存档、存放、维护和清理质量记录和技术记录的程序。质量记录应包括内部审核报告和管理评审报告以及纠正措施和预防措施的记录。

5.13.1.2 所有记录应清晰明了，并以便于存取的方式存放和保存于具有防止损坏、变质、丢失的适宜环境的设施中。应规定记录的保存期。

注：记录可以存于任何形式的媒体上，例如硬拷贝或电子媒体。

5.13.1.3 所有记录应予安全保护和保密，并符合相关的法规要求。

5.13.1.4 能力验证提供者应执行程序来保护和备份以电子形式存储的记录，并防止未经授权的侵入或修改。

5.13.2 技术记录

5.13.2.1 能力验证提供者应按照规定的时间，保存每项能力验证的所有技术数据记录。这些记录包括（但不限于）以下内容：

a) 均匀性和稳定性检验结果；

- b) 给参加者的指导书;
- c) 参加者的原始反馈;
- d) 统计分析的校核数据;
- e) 报告所需的信息 (见4.8);
- f) 最终报告 (汇总或单独的, 或两者兼而有之)。

注 1: 建议保存充足的信息, 为各轮次能力验证结果的处理建立一个审核路径。

注 2: 技术记录是开展能力验证活动所得数据和信息的累积。技术记录可包括表格、合同、工作单、工作手册、核查表、工作笔记、分包方报告和参加者反馈。

5.13.2.2 数据输入、检查和计算应在产生的当时进行记录, 并能按照特定任务和责任人分类识别。

5.13.2.3 当记录中出现错误并进行修改时, 应采取以下措施:

- a) 标示更改内容和修改日期;
- b) 避免原始数据的丢失;
- c) 标示更改人。

5.14 内部审核

5.14.1 能力验证提供者应根据预定的日程表和程序, 定期地对其活动进行内部审核, 以验证其运作持续符合管理体系和本准则的要求。内部审核计划应涉及管理体系的全部要素, 包括技术程序和能力验证物品制备、存储和分发, 以及报告能力验证计划运作的活动。质量主管负责按照日程表的要求和管理层的需要策划和组织内部审核。审核应由经过培训和具备资格的人员来执行, 只要资源允许, 审核人员应独立于被审核的活动。

注: 建议管理体系内部审核计划每十二个月完成一次。

5.14.2 当审核中发现的情况导致对运作的有效性, 包括对能力验证物品的适宜性和正确性、程序、统计评价和数据表达产生怀疑时, 能力验证提供者应及时采取纠正措施, 并通知在能力验证计划中可能受影响的客户和/或参加者。

5.14.3 审核活动的领域、审核发现的情况和因此采取的任何纠正措施, 应予以记录。

5.14.4 跟踪审核活动应验证和记录纠正措施的实施情况及有效性。

5.15 管理评审

5.15.1 能力验证提供者的最高管理者应根据预定的日程表和程序, 定期地对其管理体系和能力验证活动进行评审, 以确保其持续适用和有效, 并进行必要的变更或改进。评审应考虑到:

- a) 政策和程序的适用性;
- b) 管理和监督人员的报告;
- c) 近期内部审核的结果;
- d) 纠正措施和预防措施;
- e) 由外部机构进行的评审;

- f) 工作量和类型的工作的变化；
- g) 客户、咨询组或参加者的反馈；
- h) 投诉和申诉；
- i) 改进的建议；
- j) 其他相关因素，如资源以及员工培训。

注 1：管理评审的典型周期为每 12 个月一次。

注 2：评审结果应当输入能力验证提供者的策划系统，并包括目标和行动计划。

注 3：管理评审包括对日常管理会议中有关议题的研究。

注 4：能力验证提供者是较大组织的一部分时，宜召开覆盖能力验证活动的单独的评审会议。

5.15.2 应记录管理评审中的发现和为其采取的各项措施。管理者应确保这些措施在适当和约定的时限内得到执行。

附录A （资料性附录）

能力验证计划的类型

A.1 总则

能力验证已经成为检测、校准、检查各领域实验室活动的一项重要内容。根据使用方的需求、能力验证物品的性质、所用方法及参加者的数量，能力验证计划会有所不同。但是，大部分能力验证计划具有的共同特征，即将一个实验室所得的结果与一个或多个不同实验室所得的结果进行比较。

能力验证计划中的检测或测量类型决定了进行能力比较的方法。实验室活动有三种基本类型：定量的、定性的以及解释性的。

——定量测量的结果是数值型的，并用定距或比例尺度表示。定量测量检测的精密度、正确度、分析灵敏度以及特异性可能有所差异。在定量能力验证计划中，对数值结果通常进行统计分析。

——定性检测的结果是描述性的，并以分类或顺序尺度表示，如微生物的鉴定，或识别出存在某种特定的被测量（如某种药物或某种特性等级）。用统计分析评定能力可能不适用于定性检测。

——对于解释性计划，“能力验证物品”是与参加者能力的解释性特征相关的一个检测结果（如描述性的形态学说明）、一套数据（如确定校准曲线）或其他一组信息（如案例研究）。

其他能力验证计划依其目的不同而有另外的特点，见定义第 3.7 注 1 的 a) 至 h)。以下讨论一些常用的能力验证类型，并在图 A.1 中说明。这些计划可能是“单次的”即只做一次，或是“连续的”即按一定间隔实施。

A.2 顺序参加的计划

顺序参加的计划（有时被称作测量比对计划）是将能力验证物品连续地从一个人参加者传送到下一个参加者（即按顺序参加），有时需要传送回能力验证提供者进行再次核查。图 A.1 中的模式 1 简述了这类计划的设计，其主要特点概述如下：

a) 使用参考实验室，其能为能力验证物品提供可靠的、具有计量溯源性的指定值，且该指定值具有足够小的测量不确定度。对于分类或顺序的特性，指定值应由专家公议或其他权威公议来确定。在能力验证计划实施过程中，可能有必要在特定阶段对能力验证物品进行核查，以确保指定值没有明显变化。

b) 各个测量结果与参考实验室确定的指定值进行比较。协调者应考虑各参加者声称的测量不确定度，或声称的专业水平。按组比较结果可能有困难，因为测量能力彼此接近的参加者可能相对较少。

c) 完成顺序参加能力验证计划需要较长时间（有时需若干年），由此造成一些困难。例如：

- 确保物品的稳定性，
- 严格监控物品在参加者间的传递及各参加者允许的测量时间，
- 在计划实施过程中需向参加者单独反馈结果，而不是等到计划结束。

d) 用于该类能力验证的能力验证物品（测量制品）可包括，如测量参考标准（如电阻器、千分尺和频率计），医学计划中有确诊结果的组织切片。

e) 据此设计但仅限于单个参加者独立进行检测的计划通常称为“测量审核”。

f) 某些情况下，在所有参加者（或部分参加者）完成测量比对后，能力验证物品的指定值由公议确定。

A.3 同步参加的计划

A.3.1 总则

同步参加的计划是从材料源中随机抽取子样，同时分发给参加者共同进行测试。有些计划中，要求参加者自己抽取样品作为分析用能力验证物品。完成检测后，将结果返回能力验证提供者与指定值比对，以表明单个参加者的能力和一组参加者整体的能力。该类计划所用的能力验证物品可包括食品、体液、农产品、水、土壤、矿物和其他环境材料，某些情况下，也可是以前建立的标准物质/标准样品的一部分。建议或有教育意义的评论是能力验证提供者反馈给参加者的报告中的典型部分，目的在于促进（参加者）能力的提升。图 A.1 中的模式 2 描述了这类能力验证计划的典型例子，通常针对检测实验室。模式 3 介绍了一种经常与同步计划结合使用的计划类型，用于监督或教育目的。

如附录 B 所述，可以有多种方式确定这类能力验证计划的指定值。然而，能力的评定既可以基于参加者（全部参加者或部分“专家”参加者）的公议值，也可基于独立确定的指定值。

已知值计划使用独立于参加者的指定值，且制备的能力验证物品具有已知的被测量或特性。有证标准物质/标准样品也可用于该类计划，可以直接利用其标准值和测量不确定度。在该类计划中，也可以在重复性条件下将能力验证物品和有证标准物质/标准样品进行直接比较，但应确保有证标准物质/标准样品与能力验证物品非常接近。以前能力验证所用的能力验证物品如可证明是稳定的，也可用于该类计划。

能力验证的一种特殊应用，常被称作“盲样”能力验证，是指能力验证物品与实验室收到日常客户的物品或样品无法区别。该类能力验证可能是困难的，因为这要求与实验室的日常客户协作。另外，由于独特的包装和运输的需要，批量处理通常难以实现，且均匀性检验也有困难。

A.3.2 分割水平设计

一种常用的能力验证设计是“分割水平”设计，其中两个能力验证物品具有类似（但不相同）水平的被测量。该设计用于评估参加者在某个特定的被测量水平下的

精密度，它避免了用同一能力验证物品做重复测量，或者在同一轮能力验证中使用两个完全相同的能力验证物品带来的问题。

A. 3.3 分割样品检测计划

“分割样品”设计是经常被参加者的客户以及某些管理机构采用的能力验证特殊类型。

注：不要将该类设计与 A. 3.2 中讨论的分割水平设计相混淆。

分割样品能力验证通常用于少量参加者（通常只有两个参加者）数据的比较。在该类能力验证计划中，某种产品或材料的样品被分成两份或多份，每个参加者检测其中的一份（见图 A.1 模式 5）。该类计划的用途包括识别不好的准确度、描述持续偏倚以及验证纠正措施的有效性。该设计可用于评价作为检测服务提供方的一个或两个参加者，或用于参加者数量太少而无法进行适当结果评价的情况。

该类计划中，其中的一个参加者由于使用了参考方法和更先进的设备等，或通过参加承认的实验室间比对计划获得满意结果而证实了其自身的能力，可认为其测量具有较高的计量水平（即较小的测量不确定度）。该参加者的结果可用作该类比对的指定值，该参加者可作为顾问实验室或指导实验室，其他参加者的结果与之比对。

A. 3.4 部分过程计划

能力验证的一种特殊类型，用于评定参加者完成检测或测量全过程中若干部分的能力。例如，现有的某些能力验证计划是评定参加者转换和报告一套数据的能力（而不是进行实际的测试或测量），或基于一套数据或能力验证物品（如用于诊断的染色血液涂片）做出解释的能力，或根据规范抽取及制备样品或试样的能力。

A. 4 外部质量评价（EQA）计划

EQA 计划（例如为医学检验实验室提供的计划）根据传统的能力验证模型提供多种实验室间比对计划，但比在 A. 2、A. 3 和图 A. 1 中所描述的应用模式更宽泛。许多 EQA 计划设计的目的是对实验室完整的工作流程进行深入了解，而不是仅针对检测过程。多数 EQA 计划是包括长期实验室能力跟踪的连续性计划。EQA 计划的一个典型特征是向参加者提供教育机会并促进质量改进，为实现该目的，EQA 计划反馈给参加者的报告中包括了咨询和教育性的评议。

某些 EQA 计划在评价分析阶段的同时，也评定检测的分析前阶段和分析后阶段的能力。在这类 EQA 计划中，能力验证物品的性质可能与传统能力验证计划中所用的有很大差异，这些“能力验证物品”可能是一个调查表或案例分析（见图 A.1 模式 3），由 EQA 提供者发放给每个参加者并要求其反馈特定的答案。另一种情况是，能力验证物品可能带有一些分析前信息，要求参加者选择适当的方法进行检测或结果解释，而不仅是实施检测。在“样品复查”计划中，可能要求参加者给 EQA 提供者提交“能力验证物品”（见图 A.1 模式 4）。该类能力验证物品可能是处理过的试样或样品（如染色玻片或固定的组织）、实验室数据（如测试结果、实验室报告或质量保证/质量控制记录）或文件（如程序或方法确认准则）。



注：根据指定值确定的方式，指定值在能力验证物品分发之前或在参加者结果反馈之后确定。

图 A.1 常见能力验证计划类型示例

附录B （资料性附录）

能力验证的统计方法

B.1 总则

能力验证的结果可以多种形式出现，包含不同的数据类型且以不同的统计分布为基础。分析这些结果时，应根据不同情况选择适用的统计方法，由于统计方法种类繁多，本准则中不可能一一规定。ISO 13528 中详细描述了各种情况下优先使用的具体方法，对于其他方法，只要具有统计依据并向参加者进行了详细描述，也可使用。ISO13528 中使用的某些方法，特别是均匀性检验和稳定性检验方法，在 IUPAC 技术报告《分析化学实验室能力验证国际协议》中已略有改动。这些文件还给出了设计和直观数据分析的指南。对于特定类型能力验证计划也可参考其他资料，例如，用于校准的测量比对计划。

本附录和参考文献中给出的方法，涵盖了几乎所有类型能力验证计划所共有的基本步骤，例如，

- a) 指定值的确定，
- b) 能力统计量的计算，
- c) 能力评定，
- d) 能力验证物品均匀性和稳定性的初步判定。

对新开展的能力验证计划，由于新的问题、新的形式、人造测试物品、检测或测量方法的不一致性或测量程序的不同等原因，导致初始阶段参加者结果的一致性往往较差。协调者可能不得不使用相对能力稳健指标（如百分位数），直到一致性得到改进。一旦参加者结果的一致性得到改进且能力验证计划也得到完善时，所用的统计方法也要进一步改善。

本附录中只考虑了能力验证数据处理的统计方法而没有考虑分析研究的统计方法。若将实验室间比对数据用于引言中提到的其它目的时，可能需要使用不同的统计方法。

B.2 指定值及其不确定度的确定

B.2.1 指定值的确定有多种方法，以下列出最常用的方法。在大多数情况下，按照以下次序，指定值的不确定度逐渐增大。

- a) 已知值 —— 根据特定能力验证物品配方（如制造或稀释）确定的结果；
- b) 有证参考值 —— 根据定义的检测或测量方法确定（针对定量检测）；
- c) 参考值 —— 根据对能力验证物品和可溯源到国家标准或国际标准的标准物质/标准样品或参考标准的并行分析、测量或比对来确定；
- d) 由专家参加者确定的公议值 —— 专家参加者（某些情况下可能是参考实验

室)应当具有可证实的测定被测量的能力,并使用已确认的、有较高准确度的方法,且该方法与常用方法有可比性;

e) 由参加者确定的公议值 —— 使用 ISO 13528 和 IUPAC 国际协议中的统计方法,并考虑离群值的影响。

B.2.2 指定值的确定应确保公平地评价参加者,并尽量使检测或测量方法间吻合一致。只要可能,应通过选择共同的比对小组以及使用共同的指定值达到这一目的。

B.2.3 对上述每类指定值,ISO 13528 和 IUPAC 国际协议中详细地描述了确定其不确定度的程序。此外,ISO/IEC 导则 98-3 中给出了确定不确定度的其它信息。

B.2.4 对定性数据(也称为“分类值 categorical”或“名义值 nominal”)或半定量值(也称为“顺序值 ordinal”)确定其指定值的统计方法并没有在 ISO 13528 和 IUPAC 国际协议提及。通常,这些指定值需要由专家进行判断或由制造过程确定。某些情况下,能力验证提供者可使用大多数参加者的结果(预先确定的比例,如 80%或更高)来确定公议值。然而,这个比例应基于能力验证计划的目标和参加者的能力和经验水平来确定。

B.2.5 离群值可按下列方法进行统计处理:

a) 明显错误的结果,如单位错误、小数点错误、或者错报为其他能力验证物品的结果,应从数据集中剔除,单独处理。这些结果不再计入离群值检验或稳健统计分析。

b) 当使用参加者的结果确定指定值时,应使用适当的统计方法使离群值的影响降到最低,即可以使用稳健统计方法或计算前剔除离群值。在较大的或常规的能力验证计划中,如存在有效的客观判据,则可自动筛除离群值。

c) 如果某结果作为离群值被剔除,则仅在计算总计统计量时剔除该值。但这些结果仍应当在能力验证计划中予以评价,并进行适当能力评定。

注:ISO 13528 给出了一种确定公议均值和标准差的稳健方法,该方法不需要剔除离群值。

B.2.6 需考虑的其他事项

a) 理想情况下,如果指定值由参加者公议确定,能力验证提供者应当有确定该指定值正确度和检验数据分布的程序。

b) 能力验证提供者应当有依据不确定度来判断指定值是否可接受的准则。在 ISO 13528 和 IUPAC 国际协议中给出了该准则,该准则是基于限定指定值不确定度对能力评定的影响而建立的,即,准则限定了由于指定值的不确定度而使参加者得到一个不可接受的评估结果的可能性。

B.3 能力统计量的计算

B.3.1 定量结果

B.3.1.1 能力验证结果通常需要转化为能力统计量,以便进行解释和与其他确定的目标作比较。其目的是依据能力评定准则来度量与指定值的偏离。所用统计方法可能从

不做任何处理到复杂的统计变换。

B.3.1.2 能力统计量对参加者应是有意义的。因此，统计量应适合于相关检测，并在某特定领域得到认同或被视为惯例。

B.3.1.3 按照对参加者结果转化由简至繁的顺序，定量结果的常用统计量如下：

a) 差值 D ，由 (B.1) 式计算

$$D = x - X \quad (\text{B.1})$$

其中

x 为参加者结果；

X 为指定值。

b) 百分比相对差 $D_{\%}$ ，由 (B.2) 式计算：按照原文符号和公式

$$D_{\%} = \frac{(x - X)}{X} \times 100 \quad (\text{B.2})$$

c) z -比分数由 (B.3) 式计算：

$$z = \frac{x - X}{\hat{\sigma}} \quad (\text{B.3})$$

其中 $\hat{\sigma}$ 是能力评定标准差。

ISO 13528 中， $\hat{\sigma}$ 可由以下方法计算：

- 与能力评价的目标和目的相符，由专家判定或法规规定（规定值）；
- 根据以前轮次的能力验证得到的估计值或由经验得到的预期值（经验值）；
- 由统计模式得到的估计值（通用模式）；
- 由精密度试验得到的结果；
- 由参加者结果得到的传统标准差或稳健标准差。

d) ζ 值由式 (B.4) 计算，除了使用标准不确定度代替扩展不确定度外，计算与

E_n 值非常类似。其解释与 z 比分数相同。

$$\zeta = \frac{x - X}{\sqrt{u_{lab}^2 + u_{av}^2}} \quad (\text{B.4})$$

其中

u_{lab} 是参加者结果的合成标准不确定度；

u_{av} 是指定值的标准不确定度。

e) E_n 值由 (B. 5) 式计算:

$$E_n = \frac{x - X}{\sqrt{U_{lab}^2 + U_{ref}^2}}, \quad (\text{B. 5})$$

其中

U_{lab} 为参加者结果的扩展不确定度;

U_{ref} 为参考实验室指定值的扩展不确定度。

注 1: 仅当 x 和 X 独立时, 式 (B. 4) 和 (B. 5) 才是正确的。

注 2: 其他的统计方法参见 ISO 13528 和 IUPAC 国际协议。

B. 3. 1. 4 需要考虑的其它事项。

a) 通过参加者结果与指定值之差完全可以确定实验室能力, 对于参加者也是最容易理解的。量 $(x - X)$ 在 GB/T 6379. 4 和 ISO 13528 中称为“实验室偏倚的估计值”。

b) 百分比相对差不依赖于指定值的大小, 参加者也很容易理解。

c) 对于高度分散或者偏态的结果, 次序响应量, 或是不同响应的数量有限时, 百分位数和秩是有效的。但该方法仍应慎用。

d) 根据检测的特性优先或需要使用变换结果。例如, 稀释的结果呈现几何尺度, 需做对数变换。

e) 如果 $\hat{\sigma}$ 由公议确定, 变异度的估计值应可靠, 即, 基于足够多次的观测以降低离群值的影响, 并得到足够小的不确定度。

f) 如果根据参加者报告的测量不确定度的估计值来确定能力统计量 (例如 E_n 值和 ζ 值), 只有当所有参加者采用一致的方法 (比如按照 ISO/IEC 指南 98-3 的原则) 评估不确定度时, 该方法才有意义。

B. 3. 2 定性结果和半定量结果

B. 3. 2. 1 对于定性结果和半定量结果, 如果应用统计方法, 必须与结果的特性相适应。对定性数据 (也称之为“分类”数据), 可采用直接将参加者结果与指定值进行比较的技术。如果两者相同, 则结果是可接受的; 如果不相同, 可根据专家判断参加者结果是否满足预期用途。某些情况下, 能力验证提供者可审查参加者的结果, 并确定该能力验证物品不适于评估, 或者指定值不正确。这些决定应当是能力验证计划的一部分, 并在计划开始前使参加者了解。

B. 3. 2. 2 对半定量结果 (也称为“顺序”结果), 用于定性数据的技术也适用。顺序结果包括很多类型, 例如, 响应为等或级、感官评价, 或化学反应强度 (如 1+, 2+, 3+, 等)。有时, 这些响应结果由数字表示, 如, 1=差, 2=不满意, 3=满意, 4=良好,

5=优秀。

B. 3. 2. 3 对顺序数据进行常规的总计统计量的计算是不合适的，即使结果以数值表示。因为这些数值并不是区间尺度，也就是说，客观意义上，1 和 2 间的差可能与 3 和 4 间的差并不相同，因而不能解释其平均值和标准差的意义。因此，对半定量结果使用诸如 z 比分数的统计量是不合适的。特定的统计量，如秩或顺序统计量，对顺序数据是可以使用的。

B. 3. 2. 4 描述出（或作图表示）所有参加者结果的分布，以及每一类结果的数量或百分比，并给出总计统计量（如众数和极差）是适当的。根据与指定值的接近程度评价结果的可接受性也是适当的，例如，结果落在指定值的上或下一个数值范围内即为可接受的。某些情况下，利用百分位数评估能力也是合适的，如，可以规定距离众数或指定值最远的 5%的结果是不可接受的。这些规则应根据能力验证计划的目的来确定，并事先告知参加者。

B. 3. 3 合成的能力比分数

当对一个特定被测量使用了一个以上能力验证物品或有一组相关被测量时，可根据一轮能力验证计划中两个或两个以上的结果评定参加者的能力。这样可以对参加者能力进行全面评定。

图方法，如 Youden 图或曼德尔 h 统计量图，是解释参加者能力的有效工具（见 ISO 13528）。

总之，不鼓励使用能力比分数的平均值，因为这将掩盖对一个或多个能力验证物品的较差的检测或测量能力，而这正是需要调查的。最常用的合成的能力比分数是可接受结果的数量（或百分比）。

B. 4 能力评定

B. 4. 1 初始能力

B. 4. 1. 1 应根据能力度量所涉及的特性建立能力评定准则，用于能力评定的特性如下：

a) 专家公议，由顾问组或其他有资格的专家直接确定报告结果是否与预期目标相符合；专家达成一致是评估定性测试结果的典型方法。

b) 与目标的符合性，根据方法性能指标和参加者的操作水平等预先确定准则。

c) 用统计方法确定比分数（值），其准则应当适用于每个比分数（值）；比分数的常用例子如下：

1) z 比分数和 ξ 比分数（简单起见，示例中仅给出了 z 比分数，对 ξ 也适用）；

—— $|z| \leq 2.0$ 表明能力“满意”，无需采取进一步措施；

—— $2.0 < |z| < 3.0$ 表明能力“有问题”，产生警戒信号；

—— $|z| \geq 3.0$ 表明能力“不满意”，产生措施信号。

2) 对 E_n 值:

—— $|E_n| \leq 1.0$ 表明能力“满意”，无需采取进一步措施;

—— $|E_n| > 1.0$ 表明能力“不满意”，产生措施信号。

B. 4. 1. 2 分割样品设计的目的可能是为了识别结果中是否有不适当的校准或较大的随机波动。在这种情况下，评价需要建立在有足够多结果，并且测试结果的范围要足够宽的基础上。应用图形方法可以有效地识别并描述这些问题，在 ISO 13528 中有说明。由于坐标问题，这些图中纵轴表示结果之差，而不是一个参加者的结果对另一个参加者的结果作图。需要着重考虑的是一个参加者得到的结果是否有，或预期有较小的测量不确定度。这种情况下，该结果就是被测量实际水平的最佳估计值。如果两个参加者有几乎相同的测量不确定度，则两者结果的平均值就是实际水平的优先估计值。

B. 4. 1. 3 只要可能，应当使用 ISO 13528 和 IUPAC 国际协议所述的图形来显示实验室能力（如直方图，误差条形图，次序 z 比分数图）。这些图可用来显示：

- a) 参加者结果的分布;
- b) 多个能力验证物品结果间的关系;
- c) 不同方法所得结果分布的比较。

B. 4. 2 长期监测能力

B. 4. 2. 1 能力验证计划可包含长期监测能力的程序。该程序应当使参加者能观测到其能力的变动，是否呈现趋势性的变化或不一致的结果，以及随机变化。

B. 4. 2. 2 图形方法有助于更多读者理解数据分析结果。传统的“休哈特”控制图是有用的，特别是出于自我改进的目的。数据列表和总计统计量可以提供更详细信息。用来评定能力的标准化能力比分数，如 z 比分数，可用于绘制这些图和表。ISO 13528 给出了其它示例和图形工具。

B. 4. 2. 3 当标准差公议值作为能力验证标准差时，由于参加者的变化及其对比分数的未知影响，长期监测能力时应当谨慎。通常，由于参加者逐渐熟悉能力验证计划或者方法得到改进，实验室间标准差会随时间而减小。即便参加者本身的能力没有变化时，也会导致 z 比分数的明显变大。

B. 5 能力验证物品均匀性和稳定性

B. 5. 1 本准则要求通过有效的统计方法（包括随机抽取代表性数量的样品）证明能力验证物品“充分均匀”。ISO 13528 和 IUPAC 国际协议详述了具体程序。这两个文件中定义的“充分均匀”是相对于能力验证计划的评估区间而言，因而其推荐的方法是由评估区间确定不均匀性所带来的不确定度的允许量。ISO 13528 规定了不均匀性和

不稳定性的严格限制，以减小不确定度的影响，进而减小对能力评估的影响。相对 ISO 13528 中相应准则，IUPAC 国际协议扩展了允许不均匀性和不稳定性估计值统计检验的准则。

B. 5. 2 由于需求不同，本准则与 GB/T15000. 7 和 GB/T15000. 3 中的规定有所不同，后者用于确定有证标准物质/标准样品的参考值及其不确定度。GB/T15000. 3 使用方差分析法评估“瓶间”变异和“瓶内”变异（适当时），然后用这些方差作为指定值不确定度的分量。由于需要准确评估有证标准物质/标准样品的不确定度分量，因此，随机抽取样品数量可能超过能力验证所需的样品数量，而后者只是用来检查一批能力验证物品的不一致性。

B. 5. 3 稳定性检验通常用来确保被测量在一轮能力验证中不发生变化。如 ISO 13528，IUPAC 国际协议和 GB/T15000. 3 所述，能力验证物品的稳定性检验应考虑能力验证计划正常运作中的各种条件，例如，分发给参加者的运输和处置条件。尽管通常只做较少的检测或测量，但可接受的不稳定性的判定准则与 ISO 13528 中不均匀性的判定准则相同。

附录C （资料性附录）

能力验证的选择和利用

C.1 总则

本附录制定了参加者和其他利益相关方选择和利用能力验证计划的原则，同时也意在促进利益相关方（如认可机构、管理机构或参加者的客户）协调使用能力验证计划。

由于能力验证计划的结果可能被用于评定参加者的能力，所以利益相关方和参加者对能力验证计划的建立和运作具有信心非常重要。

对参加者来说，清楚地了解利益相关方关于参加能力验证计划的政策，用于判断能力验证计划满意表现的准则，以及对每轮能力验证所有不满意结果的跟踪政策和程序也很重要。然而，除了管理机构的特定要求外，选择适当的能力验证计划并正确评价结果也是参加者自身的责任。

但应认识到，除能力验证计划外，利益相关方也考虑由其它活动产生的检测数据的适用性，例如参加者使用控制样的自身内部质量控制程序的结果、与其他参加者的分割样品数据的比对，以及对有证标准物质/标准样品检测的结果。因此，选择能力验证计划时，参加者应考虑到可利用的或已开展的其他质量控制活动。

C.2 能力验证计划的选择

C.2.1 实验室（以及其他类型的参加者）应选择适合于其检测或校准范围的能力验证计划。所选能力验证计划应符合本准则要求。

C.2.2 选择能力验证计划时，应考虑以下因素：

- a) 涉及的检测、测量或校准应与参加者所开展的检测、测量或校准类型相匹配；
- b) 利益相关方对计划设计的细节、确定指定值的程序、给参加者的指导书、数据统计处理以及最终总结报告的可获得性；
- c) 能力验证计划运作的频次；
- d) 与有意向参加者相关的能力验证计划组织保障方面（如时间、地点、样品稳定性考虑、样品发送安排）的适宜性；
- e) 接受准则（即用于判定能力验证中的满意表现）的适宜性；
- f) 成本；
- g) 能力验证提供者对参加者保密的政策；
- h) 报告结果和分析数据的时间表；
- i) 确信能力验证物品适宜性的特性（如均匀性、稳定性，以及适当时对国家或国际标准的计量溯源性）；
- j) 与本准则的符合性。

注：有些能力验证计划中的检测可能与参加者所开展的检测不完全匹配（如对同一个测定使用不同的国家标准），但如果数据处理考虑了检测方法的重大差异或其他因素，参加能力验证计划在技术上仍然是可行的。

C.3 参加能力验证计划的政策

C.3.1 需要时，利益相关方应将参加能力验证计划的政策制定成文件；该类文件化的政策应向实验室和其他利益相关方公布。

C.3.2 参加特定能力验证计划的政策中应说明的事项包括：

- a) 参加特定能力验证计划是强制性的或自愿的；
- b) 参加的频次；
- c) 利益相关方用于判定满意或不满意表现的准则；
- d) 若参加者的表现被判定为不满意，是否要求其参加后续的能力验证计划；
- e) 在能力评定和随后的决定中如何使用能力验证的结果；
- f) 利益相关方为参加者保密的政策细节。

C.4 参加者对能力验证的利用

C.4.1 参加者应从能力验证计划组织和设计的评价中得出有关自身能力的结论，其间应考虑能力验证计划和参加者客户的需求间的关系。考虑的信息应包括：

- a) 能力验证物品的来源和特性；
- b) 所使用的检测和测量方法，可能时，特定检测或测量方法的指定值；
- c) 能力验证计划的组织（如统计设计、重复测试次数、被测量、实施方式）；
- d) 能力验证提供者用于评定参加者能力的准则；
- e) 相关的监管、认可或其他要求。

C.4.2 参加者应保留参加能力验证的记录，包括所有不满意结果的调查结论及随后的纠正或预防措施。

C.5 利益相关方对结果的利用

C.5.1 认可机构

C.5.1.1 GB/T 27011-2005第7.15条规定了认可机构利用能力验证的要求。

注：ILAC P-9中进一步规定了认可机构符合ILAC多边承认协议要求的能力验证政策。

C.5.1.2 能力验证结果对于参加者和认可机构都有用。然而，利用这些结果确定参加者能力有其一定的限制。某一次能力验证计划的满意表现可以代表这一次的能力，但不能反映出持续的能力；同样，在一次能力验证计划中的不满意表现，也许反映的是参加者偶然地偏离了正常的能力状态。正因如此，在认可过程中，认可机构不能使用能力验证作为唯一的手段。

C.5.1.3 对报告了不满意结果的参加者，认可机构应有政策以：

- a) 确保参加者在规定时间内开展调查和评议其能力，并采取适当的纠正措施；
- b) （必要时）确保参加者进行后续的能力验证以确认所采取的纠正措施有效；

c) (必要时) 确保由合适的技术评审员对参加者进行现场评价, 以确定纠正措施是有效的。

C. 5. 1. 4 认可机构应对能力验证计划表现不满意可能采取的后续措施告知获准认可机构。这些措施可能包括在规定时间内确认有效纠正措施后维持认可、对相关检测暂停认可 (视纠正措施情况而定), 直至撤销相关检测的认可。

注: 通常认可机构所选择的措施将基于参加者能力的长期表现及最近现场评审的情况。

C. 5. 1. 5 认可机构应有政策, 从获准认可机构处反馈其对能力验证结果所采取的措施, 尤其是对不满意结果的措施。

C. 5. 2 其他利益相关方

C. 5. 2. 1 参加者可能需要向其他利益相关方证明其能力, 例如客户或分包协议中的相关方。能力验证结果以及其他质量控制活动可用于证明能力, 尽管这不是唯一的方式。

注: 利用能力验证数据证实能力时通常与其他证据联合使用, 例如认可。见 C. 5. 1. 2。

C. 5. 2. 2 参加者有责任确保向评价其能力的利益相关方提供所有相关信息。

C. 6 监管机构对能力验证的利用

C. 6. 1 对需要评价所辖参加者能力的监管机构, 能力验证计划的结果是有用的。

C. 6. 2 监管机构运作能力验证计划时, 应符合本准则要求。

C. 6. 3 使用独立的能力验证提供者的监管机构应:

a) 在承认能力验证计划之前, 寻求其符合本准则的文件化证据;

b) 与参加者讨论能力验证计划的范围和运作参数, 以便按相关规定恰当地判定参加者的能力。

参考文献

- [1] ISO/IEC Guide 98-3, Uncertainty of measurement — Part 3: Guide to the expression of uncertainty in measurement (GUM:1995)
- [2] GB /T 27011-2005 合格评定 认可机构通用要求
- [3] GB/T 27025 检测和校准实验室能力的通用要求
- [4] ISO 3534-1, Statistics — Vocabulary and symbols — Part 1: General statistical terms and terms used in probability
- [5] GB/T 6379.1 测量方法与结果的准确度（正确度与精密度）第1部分：总则与定义
- [6] GB/T 6379.2 测量方法与结果的准确度（正确度与精密度）第2部分：确定标准测量方法重复性与再现性的基本方法
- [7] GB/T 6379.4 测量方法与结果的准确度（正确度与精密度）第4部分：确定标准测量方法正确度的基本方法
- [8] ISO 13528:2005, Statistical methods for use in proficiency testing by interlaboratory comparisons
- [9] GB/T 22576 医学实验室 质量和能力的专用要求
- [10] GB/T 15000.7 标准样品工作导则(7) 标准样品生产者能力的通用要求
- [11] GB/T 15000.3 标准样品工作导则(3) 标准样品 定值的一般原则和统计方法
- [12] ISO/TS 21748, Guide to the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation
- [13] EN 14136, Use of external quality assessment schemes in the assessment of the performance of in vitro diagnostic examination procedures
- [14] ASTM E1301-95, Standard Guide for Proficiency Testing by Interlaboratory Comparisons
- [15] Standards for EQA schemes in laboratory medicine. Version 4.02, December 2004. Clinical Pathology Accreditation (UK) Ltd. Sheffield, UK
- [16] National Occupational Standards for External Quality Assessment, HCS-EQA1 to HCS-EQA12. Competence Framework for Healthcare Science. (www.skillsforhealth.org.uk/)
- [17] EURACHEM/CITAC Guide CG4, Quantifying Uncertainty in Analytical Measurement, 2nd edition, 2000
- [18] THOMPSON M., ELLISON S. L. R., WOOD R., “The International Harmonized Protocol for the proficiency testing of analytical chemistry laboratories” (IUPAC Technical Report), in Pure and Applied Chemistry, Vol. 78, No. 1, pp. 145-196, 2006
- [19] ILAC P-9:2005, ILAC Policy for Participation in National and International Proficiency Testing Activities
- [20] ILAC P-10:2002, ILAC Policy on Traceability of Measurement Results