

体外诊断试剂能力验证应用策略探讨

李丽莉, 卢大伟* (中国食品药品检定研究院, 北京 100050)

摘要 目的: 探讨体外诊断检测领域能力验证策略, 为能力验证提供者组织能力验证活动提供参考, 推动实验室质量体系持续改进。方法: 通过对组织体外诊断检测领域能力验证活动的工作进行总结, 从方案设计到结果分析所涉及的关键环节进行梳理剖析, 只有将这些关键环节进行有效的质量控制, 才能成功地开展能力验证活动。结果与结论: 能力验证提供者持续改进质量管理体系建设, 为实验室提供质量评价保障的同时, 推动了实验室间结果的一致性和标准化。

关键词: 能力验证; 正确度验证; 认可准则; 质量控制

中图分类号: R95 文献标识码: A 文章编号: 1002-7777(2018)10-1400-06

doi:10.16153/j.1002-7777.2018.10.014

Discussion on Proficiency Testing Strategy for *in vitro* Diagnostic Reagents

Li Lili, Lu Dawei* (National Institutes for Food and Drug Control, Beijing 100050, China)

Abstract Objective: To explore the proficiency testing strategy in the field of *in vitro* diagnostic testing, to provide references for proficiency testing providers and to promote the continuous improvement of the laboratory quality system. **Methods:** The work of the proficiency testing activities in the field of *in vitro* diagnostic testing was summarized. The key points from the project design to the results were analyzed. If only the quality control of the key points were achieved, the proficiency testing activities could be successfully carried out. **Results and Conclusion:** The proficiency testing providers should continuously improve the quality control system, provide the quality evaluation for the laboratory and to promote consistency and standardization of inter-laboratory results.

Keywords: proficiency testing; trueness-based proficiency testing; accreditation criteria; quality control

能力验证 (proficiency testing, PT) 是指利用实验室间比对来确定实验室的检测能力或检查机构的检测能力的活动^[1]。实验室间比对是指按照预先制定的准则, 组织两个或多个实验室对相同或类似的样品进行检测或测量, 要求其在规定时间内报告测定结果, 并且对结果进行评定的活动^[2]。能力验证因其能够有效检验实验室检测能力而备受国际组织的重视。近年来, 我国实验室能力验证体系已经逐步成型, 但体外诊断检测领域还有许多亟待深入研究和不断完善的方面。随着体外诊断技术的发

展, 商品化体外诊断产品广泛应用于疾病诊治, 对使用商品化体外诊断产品的临床检测实验室及从事体外诊断产品质量评价的实验室都提出了更高的要求。笔者就组织体外诊断检测能力验证应关注的关键环节进行探讨。

1 能力验证的计划方案

方案设计是能力验证活动的基础和关键^[1]。应该成立专家组针对所要开展的能力验证项目和活动实施过程进行详细设计。能力验证方案包含以下内容: 本次能力验证活动的具体名称, 实施本次验证

活动的机构, 活动组织成员及其职责分工, 能力验证活动实施范围, 能力验证受试参数, 受试样品性状、数量、制备、运输、分组及编号原则, 上报表格及操作规程, 时间安排和费用等^[3]。

能力验证计划方案经相关专家组反复讨论通过后, 具体方案要求文件化, 并以此作为整个能力验证活动的组织和运作的指导性文件, 然后将相关通知、报名表等由组织者向实施范围内的检验机构邮寄发放。

2 能力验证的样品

受试样品是能力验证活动的核心, 应根据能力验证的目的配制, 确定样品的性状特征、受试参数浓度、包装形式, 以及需要加入的背景干扰物等。如果是常规检验项目, 主要目的是为检验参与活动的实验室日常检验水平, 受试样品要尽量与实验室日常检测样品的浓度、性状等保持一致, 这样才能检验参与实验室的常规操作水平^[2]。

能力验证所提供测试样品的均匀性和稳定性是保证能力验证结果可信度的关键。为了确保能力验证结果的公平、可靠、真实, 必须要保证每个参加实验室收到的受试样品是相同的, 每个样品受试参数的检测量不应存在统计差别, 在能力验证中出现的任何离群或可疑结果都不应该来源于受试样品之间的差异^[4]。

2.1 能力验证样品的制备

严格控制能力验证样品的制备过程, 保证能力验证受试样品的一致性和可操作性对顺利开展能力验证是至关重要的^[5]。能力验证样品的稳定性、均匀性和互通性(即基质与患者样品的一致程度)直接影响到参加者的检测结果。能力验证中使用具有互换性的受试样品对不同的检测系统(或检测方法)的标准化或一致性更具有意义^[6], 这对能力验证提供者在调查品的准备和评价方面提出了更高的要求。同时, 还要考虑到各参与实验室上报结果的客观性和真实性, 防止可能出现的因实验室间受试参数检出值的相互串通, 而导致评价结果的统计偏离和公正性缺失^[6]。

体外诊断试剂能力验证样品还要考虑正确度方面的验证, 应注意以下几个方面: 1) 能力验证样品多采用人源样品, 避免了加工处理过的样品引入的基质效应对结果造成的偏移; 2) 能力验证样

品要尽量与实验室日常检测样品的浓度、性状等保持一致, 这样才能检验参与实验室常规操作水平; 3) 靶值确定通常采用国际公认参考方法赋值, 用正确赋值来考评实验室比使用方法组中位数或公认值方法更为科学; 4) 用同一个标准直接比较各实验室检测结果的差异, 推动实验室质量的改进, 也契合了国家和行业对结果互认的要求。

2.2 能力验证样品的均匀性

用于能力验证的受试样品, 必须进行样品均匀性检验^[4]。受试样品的均匀性是指存在于样品中的一种或多种特定指标具有相同的结构或数量的一种状态。一般采用以下方法检验受试样品的均匀性: 随机抽取 10 个或以上的样品, 选择有代表性的和对不均匀性敏感的特定指标进行测试, 并在相同条件下至少重复检验 2 次。均匀性检验所用的测试方法, 其精密度和灵敏度不应低于所要开展的能力验证指定检验方法的精密度和灵敏度; 并且均匀性检验所用的受试样品取样量应小于能力验证指定检验方法的取样量。同时要满足在限定时间内, 由同一实验室的相同人员采用相同方法和设备, 完成受试样品的均匀性检验。按照 CNAS-GL03 指南文件^[7]对合格受试样品的评价给予规定, 均匀性检验通常采取单因子方差分析法对均匀性的检验结果进行判定。该分析法表明, 如果受试样品的结果之间无显著性差异, 则表明样品是均匀的; 对需要定性的受试样品, 按定性检验的要求确定受试样品的均匀性。以中国食品药品检定研究院(以下简称中检院) 2017 年体外诊断试剂能力验证项目血清中尿酸、尿素测定(NIFDC-PT-104) 为例, 5 份能力验证样本, 制备方式一致, 因此随机抽取其中一份样本, 进行均匀性分析。采用精密度良好的日立 7180 全自动生化分析仪对抽取 10 份样本进行测定, 每个样本测定 3 次。利用单因素方差分析法对所得数据进行数据统计分析, 若 $F < F_{\alpha}(f_1, f_2)$ 及给定显著性水平 $\alpha(0.05)$ 的临界值 $F_{\alpha}(f_1, f_2)$, 则表明样品内和样品间无显著性差异, 样品是均匀的。尿酸、尿素均匀性检测记录见表 1 和表 2, 尿酸 F 值为 1.339864, F 临界值为 1.355166; 尿素 F 值为 1.071361, F 临界值为 1.355166。两个项目的 F 值均小于 F 临界值, 说明组间无显著性差异, 所制备的样本是均匀的。

表1 UA (尿酸) 的均匀性检测结果

差异源	平方和	自由度	均方	<i>F</i>	<i>P</i> 值	$\mu\text{mol} \cdot \text{L}^{-1}$ <i>F</i> 临界值
组间	513.6077	83	6.188045	1.339864	0.056666	1.355166
组内	775.8933	168	4.618413			
总计	1289.501	251				

表2 UREA (尿素) 的均匀性检测结果

差异源	平方和	自由度	均方	<i>F</i>	<i>P</i> 值	$\mu\text{mol} \cdot \text{L}^{-1}$ <i>F</i> 临界值
组间	0.634458	83	0.007644	1.071361	0.349989	1.355166
组内	1.198667	168	0.007135			
总计	1.833125	251				

2.3 能力验证样品的稳定性

由于参与能力验证的机构往往分布在全国, 所以如何保证受试样品从发出到实验人员手中时仍能保持性能性状的稳定一致, 是至关重要的。受试样品的稳定性是指在规定的时间内, 经过相应的贮存、运输后, 使其特定的指标值仍能保持在规定范围内。

体外诊断试剂能力验证样品多为性质较为活跃的生物制品, 验证其稳定性是必不可少的。一般对于样品稳定性的检验方法可根据样品的性质和计划要求设计。参与稳定性检验的样品应从需要发给参加能力验证活动的实验室样品中抽取, 进行稳定性检验所需要的时间跨度必须能够维持整个能力验证过程, 即从发放样品开始一直到预定的允许参加实验室最后递交检验结果的时间为止的时间跨度。例如, 中检院2017年体外诊断试剂能力验证

项目设计的时间跨度为30天, 实验室将若干个受试样品分别放置在 $-20\text{ }^{\circ}\text{C}$ 、 $4\text{ }^{\circ}\text{C}$ 及 $25\text{ }^{\circ}\text{C}$ 的环境下保存, 进行稳定性测定^[4]。按照CNAS-GL03指南文件采用*t*检验法对样品稳定性的检验结果进行判定。当对受试样品的测定检验结果的偏差不超过预定检验方法要求的精密度时, 就可以认定该样品在设计的时间范围内是稳定的。例如, 甲型流感病毒核酸检测(NIFDC-PT-105), 随即抽取10份冻干的流感样品, 以实时荧光定量PCR法各检测1次, *Ct*值差异 <1.0 为稳定。按照变异系数及*Ct*值差异判定是否满足要求。能力验证样品P44、P53和P80分别经 $4\text{ }^{\circ}\text{C}$ 放置5天、 $4\text{ }^{\circ}\text{C}$ 放置2天+室温($25\text{ }^{\circ}\text{C}$)放置3天, 以及复溶后反复冻融2次等3种处理后, 与 $-20\text{ }^{\circ}\text{C}$ 保存的对照相比, 其*Ct*值差值的绝对值均 <1.00 。具体结果见表3。

表3 甲型流感病毒核酸检测(NIFDC-PT-105)稳定性检验结果

样品编号	<i>Ct</i>	处理
P44	27.49	$-20\text{ }^{\circ}\text{C}$ 对照
	27.15	$4\text{ }^{\circ}\text{C}$ 5天
	27.18	$4\text{ }^{\circ}\text{C}$ 2天 + $25\text{ }^{\circ}\text{C}$ 3天
	27.23	复溶后反复冻融2次

续表3

样品编号	Ct	处理
P53	17.91	-20℃对照
	17.92	4℃ 5天
	18.32	4℃ 2天 + 25℃ 3天
	18.17	复溶后反复冻融 2次
P80	22.32	-20℃对照
	22.30	4℃ 5天
	22.27	4℃ 2天 + 25℃ 3天
	21.91	复溶后反复冻融 2次

注：检验方法 Cut-off 值为 35，以上处理后 Ct 值差异均 < 1.0，对定性判定无影响。

2.4 能力验证样品的编组

为避免出现实验室间互通窜改结果和伪造检验数据的情况发生，能力验证组织实施机构应该对受试样品和参与实验室进行科学编号。常用的能力验证样品设计是“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^[8-9]。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数据串通问题。以中检院2017年组织的血清中尿酸、尿素测定项目（NIFDC-PT-104）为例，受试样品包含5份不同尿素、尿酸水平的冰冻血清样本，5份样品随机编号，一个水平的样品用于准确性评价，另一水平的样品用于精密度评价，精密度和准确性结果均在规定范围内，判断为满意；任一条件不符合，均判定为不满意。通过编号随机和测量水平分割，降低能力验证者串通数据的概率。

能力验证计划组织者应发放给每个参加者仅供平行样测试用的适量样品，除非确认原样品破损，否则绝不随意提供备用样品，以防实验室拿备用样品进行比对串通。还应尽量缩短能力验证结果上报的时间。能力验证计划组织者应将样品邮寄时间及大约到达时间提前通知参加者，让参加者提前准备实验条件，以最快的速度、最短的时间完成测试及结果上报。

2.5 能力验证样品的运输

为保证提供给分布在全国各地的参与实验室均匀一致的受试样品，能力验证组织者还应对样品采取适当的存储条件、适宜的运输方式和安全快捷的分发途径。应该按照受试样品的特性采用相应的包装和运输方式：对需要低温冷藏的样品，应低温冷藏包装；对需要避光保存的样品应采用避光材料；对需要防震防碰的样品，则要考虑采用海绵、泡沫等防震材料和防碰结构包装；对活跃的化学品和微生物样品应该用真空包装或密封包装以避免空气污染和样品外溢。

3 能力验证检验结果的评价

CNAS-GL02 文件^[10]对一个重要指标——靶值的确定方法根据不确定度由小到大进行了排列：（1）已知值；（2）有证参考值；（3）参考值；（4）由各专家实验室获得的公议值；（5）从参加实验室获得的公议值。方法“（5）”是目前PT提供者最常采用的方法，通常为参加者结果的中位数或稳健均数，但得到的靶值不确定度是最大的，并且受限于参加实验室数量、样品的互通性、方法分组的同质性等因素，使统计靶值与真值产生较大偏移。中检院2017年组织的血清中尿酸、尿素测定项目（NIFDC-PT-104）和甲型流感病毒核酸检测（NIFDC-PT-105），其中NIFDC-PT-104为定量测定项目，NIFDC-PT-105为定性测定项目。两个项目均采用方法（3）参考值，根据能力验证受试

样品与可溯源至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、测量或比对来确定参考值的方法,确定参考值,为体外诊断试剂能

力验证实施正确度评价/调查进行了有益地尝试。两个能力验证项目结果见表4、表5。

表4 NIFDC-PT-104 能力验证结果汇总

实验室 代码	尿酸 ($\mu\text{mol} \cdot \text{L}^{-1}$)					尿素 ($\text{mmol} \cdot \text{L}^{-1}$)				
	样本1	样本2	样本3	样本4	样本5	样本1	样本2	样本3	样本4	样本5
619	145	278	350	436	494	3.85	5.78	7.83	6.68	18.33
732	152	282	352	431	486	3.91	5.94	7.96	6.71	18.78
003	161	296	367	448	504	3.91	5.97	7.98	6.72	18.56
213	146	278	349	432	491	3.71	5.65	7.53	6.40	18.05
625	145	271	327	404	443	3.58	5.76	8.00	6.48	20.46
298	148	280	350	434	495	3.89	5.96	7.95	6.70	18.95
530	145	282	358	441	503	3.93	5.97	7.93	6.74	18.55
312	153	288	355	434	498	3.70	5.70	7.66	6.44	18.22
388	143	279	351	437	495	3.80	5.75	7.76	6.53	18.04
250	144	277	350	430	486	3.70	5.60	7.56	6.38	17.92
478	148	276	348	429	485	3.89	5.88	7.95	6.66	18.65
011	149	295	372	470	520	3.89	5.92	7.91	6.77	18.75
941	145	278	348	435	497	3.76	5.73	7.70	6.47	17.97
939	140	266 [*]	335	415	473	3.64	5.54	7.46	6.18	17.70
808	146	285	352	438	495	3.77	5.79	7.75	6.59	18.16
408	-	-	-	-	-	4.05	5.95	8.12	6.84	19.28
均值	147	281	351	434	491	3.81	5.81	7.82	6.57	18.54
SD	5.3	8.4	11.2	14.7	17.5	0.13	0.14	0.20	0.18	0.69
CV/%	3.6	3.0	3.2	3.4	3.6	3.4	2.5	2.5	2.8	3.7

注：“-”该实验室未参加此项目。

表 5 NIFDC-PT-105 能力验证结果汇总

实验室代码	样品编号	预期结果	报告结果	结果判断
213	FI01050005	P/N/P/N	P/N/P/N	满意
158	FI01050009	N/P/P/P	N/P/P/P	满意
383	FI01050006	P/P/N/P	P/P/N/P	满意
625	FI01050003	N/N/P/P	N/N/P/P	满意
498	FI01050010	P/P/N/P	P/P/N/P	满意
976	FI01050002	N/P/P/P	N/P/P/P	满意

注：P 代表阳性，N 代表阴性。

4 结语

举办能力验证活动要考虑得非常周全，从开始的方案设计到最后的分析涉及很多关键的环节：首先，要保证方案的可操作性和严谨性；其次，既要考虑受试样品的性能、状态、均匀性、稳定性，又要考虑包装、贮存、运输对受试样品的影响；再次，检测结果报出值的有效数字要求与统计方法的适应性等方面对结果统计的影响；最后，还应关注参加者之间可能进行的数据串通，以免引起统计结果的偏离与不客观。只有将这些关键环节进行有效的质量控制，才可能组织好能力验证活动。

参考文献：

- [1] 能力验证提供者认可准则[S]. CNAS-CL03: 2010.
 [2] 能力验证提供者认可指南[S]. CNAS-GL31: 2011.
 [3] 能力验证计划设计方案[S]. CNAS-WI21-01/01-A/1: 2012.
 [4] 能力验证样品均匀性和稳定性评价指南[S]. CNAS-

GL03: 2006.

- [5] 郝曼. 能力验证活动对现代实验室建设的重要性[J]. 现代测量与实验室管理, 2014, (3): 46-47.
 [6] MILLER WG, JONES GR, HOROWITZ GL, et al. Proficiency Testing /external Quality Assessment: Current Challenges and Future Directions[J]. Clin Chem, 2011, 57(12): 1670-1680.
 [7] 能力验证样品均匀性和稳定性评价指南[S]. CNAS-GL03: 2006
 [8] 国家计量局. [87]量局法字第188号中华人民共和国强制检定的工作计量器具目录[S]. 1987.
 [9] 王飞. 可燃气体检测报警器计量管理与检定的策略研究[J]. 科技风, 2016, (5): 94-97.
 [10] 能力验证结果的统计处理和评价指南[S]. CNAS-GL02: 2006.

(收稿日期 2018年4月11日 编辑 郑丽娥)