

## 研究报告

## 实验室技术能力综合评价方法初步探讨

周爽,岳兵,赵馨,马兰,赵云峰

(国家食品安全风险评估中心 卫生部食品安全风险评估重点实验室,北京 100021)

**摘要:**目的 探讨实验室技术能力综合量化分析评价方法。方法 利用国家食品安全风险评估中心组织的3次食品中重金属分析考核结果,比较了Z比分数方法、点分数法和组合分数法对考核的综合量化分析结果。结果 Z比分数法得到的5个考核目标物的满意率分别为78.1%、93.8%、81.2%、90.6%和93.8%;点分数法评价3次考核的满意率分别为75%、71.9%、90.6%;组合分数法评价3次考核各实验室综合表现满意率为81.2%。结论 3种不同评价方法各有侧重,适合于不同目的的结果分析。Z比分数法适用于单一因素分析能力评价;点分数法适用于一次多个因素分析能力综合性评价;组合分数法适用于实验室历次考核结果综合能力分析。

**关键词:**食品污染物;重金属;能力评价;Z比分数法;点分数法;组合分数法

中图分类号:R194.2 文献标识码:A 文章编号:1004-8456(2013)03-0233-05

### Preliminary discussion on comprehensive statistical treatments of proficiency test results

Zhou Shuang, Yue Bing, Zhao Xin, Ma Lan, Zhao Yunfeng

(Key Laboratory of Food Safety Risk Assessment of Ministry of Health, China National Center for Food Safety Risk Assessment, Beijing 100021, China)

**Abstract: Objective** To discuss the comprehensive assessment methods for evaluation of proficiency test results.

**Methods** Statistical methods of Z-score, point-score-system and combination scores were used to evaluate the laboratory test results from three times and provide a general assessment. **Results** The pass rates that analyzed by Z-score method for the five targets were 78.1%, 93.8%, 81.2%, 90.6% and 93.8%, respectively. Using point-score-system, the pass rates of the three tests were 75%, 71.9% and 90.6%. And with combination scores, 81.2% of the laboratories showed satisfying performances. **Conclusion** The three evaluation methods showed advantages in different aspects and fit for the comprehensive evaluation of proficiency tests. Z-score fit for single tested substance; point-score-system was suitable for each test with multi-target substances; combination scores were competent for comprehensive assessment of each laboratory during several tests.

**Key words:** Food contaminants; heavy metal; laboratory test evaluation; Z-score; point-score-system; combination scores

利用实验室间比对或能力验证评价实验室检验能力的活动,是不同检验机构结果一致性比较的有力手段。铅、镉、汞、砷作为食品安全评价的主要污染物,其检测结果倍受关注,且重金属性能稳定,检验方法易于统一,结果更具有可比性,因此,对重金属的检测成为食品检验能力分析最经常选择的项目。考核结果分析是评价实验室能力的重要环节,采用有效的评价方法,将有助于对参加考核实验室能力和水平的了解。Z比分数法为进行实验室间比对或能力验证时通常使用的统计分析和评价

方法<sup>[1-3]</sup>,其适合于单次考核中单一项目结果的评价,随着实验室比对或能力验证工作的推广,考核项目和考核次数的增加,需要引入综合的评价方法。近年来,欧盟在实验室能力验证考核中,尝试使用了点分数法<sup>[4-5]</sup>和组合分数法<sup>[6-7]</sup>对实验室能力进行综合性评价。本文比较和探讨了欧盟检验能力综合评价方法,对国家食品安全风险监测网组织的3次食品中重金属检验质控考核结果进行分析,探讨了利用多次考核结果对实验室进行检验能力综合性评价的方法。

收稿日期:2013-03-05

基金项目:“十二五”科技支撑计划(2011BAK10B07)

作者简介:周爽 女 助理研究员 研究方向为分析化学

E-mail:szhoupk@gmail.com

通信作者:赵云峰 男 研究员 研究方向为食品安全

E-mail:zhaoyf703@126.com

## 1 材料与amp;方法

### 1.1 数据来源

2006、2008、2011年国家食品安全风险评估中心(中国疾控中心营养与食品安全所)组织的3次食品中重金属分析质控考核结果,3次考核参与机构

范围相同,均为新疆生产建设兵团和31个省级疾病预防控制中心。

## 1.2 考核内容

2006年为桃树叶与奶粉混合基质中铅、砷分析<sup>[8]</sup>;2008年为豆奶粉中镉分析;2011年为豆奶粉中镉、汞分析<sup>[9]</sup>。

## 1.3 结果分析评价方法

### 1.3.1 Z比分数法

用Z比分数法对考核单一目标物正态分布结果进行评价<sup>[10]</sup>。Z比分数的计算公式为:

$$Z \text{ 值} = \frac{\text{测定值} - \text{指定值}}{\text{标准偏差}}$$

结果判定:  $|Z| \leq 2$  为满意结果;  $2 < |Z| < 3$  为可疑结果,建议实验室关注;  $|Z| \geq 3$  或未提交结果为不满意结果。

### 1.3.2 点分数法

点分数法是近年来欧盟参比实验室在组织一次能力验证包括多个考核项目时对实验室能力的综合评价方法。

其方法原理为:首先根据考核项目性质,分成违禁物质、有限值要求和定性分析三类,以Z值为基础,给不同考核项目类别的Z值赋以点分数<sup>[5]</sup>,具体如表1所示,然后将考核项目点分数累加。如果该实验室此次能力验证活动的总点分数大于该次考核满分分数的2/3,则认为该实验室此次分析能力测试结果满意<sup>[5]</sup>。该方法通常规定,对未及时上报结果,其Z值按3.0处理。

表1 不同考核项目点分数赋分方法

Table 1 Definition of the points for tested substances

项目或内容	结果范围	点分数
违禁物质分析	$ Z  < 2$	+1.5
	$ Z  > 2$	+1.0
有限量要求指标分析	$ Z  < 1$	+2.0
	$1 <  Z  < 2$	+1.0
	$ Z  > 2$	+0.5
定性分析假阴性	未检出	0
定性分析假阳性	1~2个假阳性结果	-1.0
	>2个假阳性结果	-2.0

### 1.3.3 组合分数法

组合分数法可以用于一个实验室历年考核结果的综合评价。组合分数法先计算每个项目的Z值,分别按下式计算调整Z比分数之和(Rescaled sum of Z-Scores, RSZ)和实验室相对能力(relative laboratory performance, RLP)两个参数:

$$RSZ = \frac{\sum Z}{\sqrt{n}} \quad RLP = \sqrt{\frac{1}{n} \sum Z^2}$$

式中, $n$ 为历次考核中考核项目总数。同时利用两个参数综合评价历年考核结果,如果  $-2 < RSZ$

$< 2$ ,且  $RLP < 1.5$ ,则该实验室能力满意<sup>[6-7]</sup>。该方法通常规定,对未及时上报结果,其Z值按3.0处理。

## 2 结果

### 2.1 Z比分数法评价结果

按Z比分数法计算Z值,表2为各实验室3次考核的Z值结果。

表2 3次质控考核结果的Z值表

Table 2 Z-scores of each tested substances

实验室编号	2006年铅	2006年砷	2008年镉	2011年镉	2011年汞
LC1	-0.41	0.49	-0.76	-0.13	-0.18
LC2	0.33	-1.14	-0.68	0.24	-0.54
LC3	-0.41	-0.16	-0.24	-0.81	0.63
LC4	-0.41	0.16	0.17	-0.61	0.71
LC5	-0.74	0.16	-0.15	0.44	0.60
LC6	-0.08	-0.81	1.31	0.62	0.24
LC7	2.37	-1.47	0.65	0.31	0.09
LC8	0.98	1.14	0.13	0.31	-0.07
LC9	-0.25	-1.14	0.65	0.25	-0.07
LC10	-0.9	-1.47	0.02	-0.69	0.82
LC11	-0.16	-0.16	-2.07	0.50	0
LC12	2.45	1.14	0.64	2.46	-0.01
LC13	-0.25	0.49	-0.98	0.46	-0.07
LC14	-0.25	1.3	-0.57	-0.44	0.74
LC15	0.25	1.3	-3.6	0.15	0.05
LC16	-0.33	0.49	-0.18	0.60	0.14
LC17	-0.57	0.65	-0.42	-0.67	-0.10
LC18	0.58	0.49	0.76	-0.73	0.63
LC19	2.45	-1.14	0.11	0.32	-0.02
LC20	-0.33	-0.16	-0.37	-0.77	-0.43
LC21	1.06	1.14	-0.49	-0.64	-0.08
LC22	0.16	-0.88	-0.41	-1.12	-0.28
LC23	0	1.47	3.4	0.58	-0.21
LC24	1.39	-0.88	1.44	-0.39	0.04
LC25	—	—	—	—	-3.64
LC26	0	-1.14	0.98	0.41	-0.14
LC27	6.535	-0.88	-1.4	0.40	-0.02
LC28	0.41	1.79	-0.39	0.65	0.25
LC29	0.08	-0.49	-2.2	0.28	-0.11
LC30	2.37	0.49	0.37	0.56	-0.16
LC31	0	-1.14	0.72	-0.46	0.47
LC32	—	—	—	—	—

注:“—”为未提供数据

2006年进行铅、砷考核,30个实验室提交铅元素分析有效数据,总体满意率78.1%,12.5%结果可疑,6.25%结果不满意。30个实验室提交了砷元素分析有效数据,总体满意率93.8%,6.2%结果不满意。

2008年镉元素考核,共30个实验室提交有效数据,总体满意率81.2%,6.25%结果可疑,12.5%结果不满意。

2011年进行了镉和汞考核。共30个实验室提

交镉元素分析有效数据,总体满意率 90.6%,9.4% 结果可疑。共 31 个实验室提交汞元素分析有效数据,总体满意率 93.8%,6.2% 结果不满意。

2.2 点分数系统法评价结果。

点分数法不仅适用于定量结果准确性评价,同

时兼顾了定性筛查结果中假阳性、假阴性的量化评判。本文按有限量值要求指标赋分,综合分析评价了各次食品中重金属实验室分析能力。

根据表 2 的 Z 值结果,查找其对应的点数,并累加,得到每次考核中各实验室的总点分数,结果见表 3。

表 3 点分数法评价结果

Table 3 Test results of heavy metals by point-score-system

实验室编号	2006 年(满分 4.0)		2008 年(满分 2.0)		2011 年(满分 4.0)		3 次考核(满分 10)	
	点分数	评价结果	点分数	评价结果	点分数	评价结果	点分数	评价结果
LC1	4.0	满意	2.0	满意	4.0	满意	10.0	满意
LC2	3.0	满意	2.0	满意	4.0	满意	9.0	满意
LC3	4.0	满意	2.0	满意	4.0	满意	10.0	满意
LC4	4.0	满意	2.0	满意	4.0	满意	10.0	满意
LC5	4.0	满意	2.0	满意	4.0	满意	10.0	满意
LC6	4.0	满意	1.0	不满意	4.0	满意	9.0	满意
LC7	1.5	不满意	2.0	满意	4.0	满意	7.5	满意
LC8	3.0	满意	2.0	满意	4.0	满意	9.0	满意
LC9	3.0	满意	2.0	满意	4.0	满意	9.0	满意
LC10	3.0	满意	2.0	满意	4.0	满意	9.0	满意
LC11	4.0	满意	0.5	不满意	4.0	满意	8.5	满意
LC12	1.5	不满意	2.0	满意	2.5	不满意	6.0	不满意
LC13	4.0	满意	2.0	满意	4.0	满意	10.0	满意
LC14	3.0	满意	2.0	满意	4.0	满意	9.0	满意
LC15	3.0	满意	0.5	不满意	4.0	满意	7.5	满意
LC16	4.0	满意	2.0	满意	4.0	满意	10.0	满意
LC17	4.0	满意	2.0	满意	4.0	满意	10.0	满意
LC18	4.0	满意	2.0	满意	4.0	满意	10.0	满意
LC19	1.5	不满意	2.0	满意	4.0	满意	7.5	满意
LC20	4.0	满意	2.0	满意	4.0	满意	10.0	满意
LC21	2.0	不满意	2.0	满意	4.0	满意	8.0	满意
LC22	4.0	满意	2.0	满意	3.0	满意	9.0	满意
LC23	3.0	满意	0.5	不满意	4.0	满意	7.5	满意
LC24	3.0	满意	1.0	不满意	4.0	满意	8.0	满意
LC25	1.0	不满意	0.5	不满意	1.0	不满意	2.5	不满意
LC26	3.0	满意	2.0	满意	4.0	满意	9.0	满意
LC27	2.5	不满意	1.0	不满意	4.0	满意	7.5	满意
LC28	3.0	满意	2.0	满意	4.0	满意	9.0	满意
LC29	4.0	满意	0.5	不满意	4.0	满意	8.5	满意
LC30	2.5	不满意	2.0	满意	4.0	满意	8.5	满意
LC31	3.0	满意	2.0	满意	4.0	满意	9.0	满意
LC32	1.0	不满意	0.5	不满意	1.0	不满意	2.5	不满意
满意度	75%		71.9%		90.6%		90.6%	

2006 年进行了铅、砷两个项目考核,满分为 4.0 分,点分数 >2.7 的为满意单位,共有 24 家实验室,满意度 75%。在 8 个不满意实验室中,除未上报数据的 2 家实验室外,6 家实验室的综合评价分数较低。

2011 年有镉、汞两个考核项目,满分为 4.0 分,点数 >2.7 的为满意单位,共 29 家,占 90.6%;其余 3 家为不满意单位,其中 2 家未上报数据,1 家单位测定结果在  $1 < |Z| < 2$  之间。该评价结果与 Z 比分数法满意度相当。

此外,对 3 次考核的 5 个项目进行总体评价,则满分为 10.0 分,点数 >6.7 的为满意单位,共

29 家,占 90.6%,与逐年评价相比差异明显。可见参评项目数量对于点分数法的评价结果影响显著,项目数越多,每个项目对评价结果影响越小,评价结果愈趋于稳定。因此,点分数法作为综合评价体系,不适用于单一测定值的评价,而对于多项目考核则更易于得到稳定、客观的评价结论。

2.3 组合分数法综合评价

根据表 2 的 Z 值结果,计算参数 RSZ 和 RLP,结果见表 4。该方法通常规定,对未及时上报结果,其 Z 值按 3.0 处理。通过 RSZ 和 RLP 参数值评价每个实验室在多次能力验证中的总体表现。

表4 组合分数法评价结果

Table 4 Proficiency test results of heavy metals by combination scores

实验室编号	RSZ	RPL	评价结果
LC1	-0.44	0.45	满意
LC2	-0.80	0.67	满意
LC3	-0.44	0.51	满意
LC4	0.01	0.47	满意
LC5	0.14	0.48	满意
LC6	0.57	0.75	满意
LC7	0.87	1.29	满意
LC8	1.11	0.69	满意
LC9	-0.25	0.61	满意
LC10	-0.99	0.91	满意
LC11	-0.85	0.96	满意
LC12	2.99	1.66	不满意
LC13	-0.16	0.54	满意
LC14	0.35	0.75	满意
LC15	-0.83	1.72	不满意
LC16	0.32	0.39	满意
LC17	-0.50	0.53	满意
LC18	0.77	0.65	满意
LC19	0.77	1.22	满意
LC20	-0.92	0.46	满意
LC21	0.44	0.78	满意
LC22	-1.13	0.68	满意
LC23	2.34	1.68	不满意
LC24	0.72	0.99	满意
LC25	3.74	3.14	不满意
LC26	0.05	0.70	满意
LC27	2 921	2 923	不满意
LC28	1.21	0.90	满意
LC29	-1.09	1.02	满意
LC30	1.62	1.13	满意
LC31	-0.18	0.67	满意
LC32	6.71	3.00	不满意

分别以 RSZ 和 RLP 为横、纵坐标,将 2006—2011 年 3 次考核中 32 个实验室的能力验证结果绘制在图中(见图 1),其中阴影部分为  $-2 < \text{RSZ} < 2$  且  $\text{RLP} < 1.5$  的实验室,即综合能力满意实验室,共有 26 个,占 81.2%。共 6 个实验室为不满意单位,其中 LC27 实验室由于 2006 年铅测定值偏离定值太远,Z 值为 6 535,导致 RLP 和 RSZ 值异常高,没有显示在图 1 中,其余 5 个实验室 LC25 和 LC32 多次未上报数据,其他 3 个实验室 LC12, LC15, LC23 均在满意范围附近。

对典型实验室进行对比分析,结合表 2 中 Z 比分数法的结果,LC12 实验室在 2006 年铅考核及 2011 年镉考核中,Z 值分别为 2.45 和 2.46,以 Z 比分数法判断均为可疑结果,与组合分数法评价为不满意实验室结果相符。LC15 和 LC23 实验室在 5 项考核结果中,虽然都只有 1 项不满意,但其 Z 值分别为 -3.6 和 3.4,  $|Z| \geq 3$ ,与参考值偏离很大,因此组合分数法仍评价为不满意实验室。而对于 LC7、

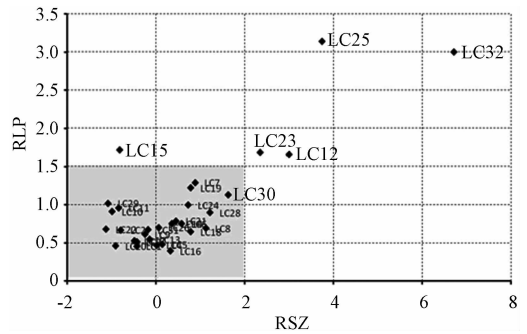


图1 2006—2011年实验室总体评价图

Figure 1 Comprehensive evaluations of laboratories by combination scores

LC11、LC19、LC29 和 LC30 这 5 个实验室,在 5 项考核中,也各有 1 项结果可疑,对应的 Z 值分别为 2.37、-2.07、2.45、-2.2 和 2.37,与 LC15 和 LC23 实验室相比,表现相对较好,组合分数法综合评价这 5 个实验室为满意实验室。

### 3 讨论

本文采用三种方法评价了 3 次考核 5 种目标物的能力验证结果,其中 Z 比分数法得到 5 种考核目标物的满意率分别为 78.1%、93.8%、81.2%、90.6% 和 93.8%;点分数系统法得到 3 次考核的满意率分别为 75%、71.9% 和 90.6%;组合分数法得到 2006—2011 年各实验室综合表现满意率为 81.2%。可见三种评价方法在满意率方面具有一定的可比性,纵观 3 次考核结果,各项目间存在差异。如在 2006 年考核中,对铅的检测满意率较低(参考值为 0.68 mg/kg),究其原因,可能来自实验室背景污染对考核结果产生的影响。而在 2008 年镉元素考核中,低浓度样品的考核结果满意率低(参考值为 52.7  $\mu\text{g}/\text{kg}$ ),表明实验室监测低含量镉的能力还需要进一步加强。

通过三种评价方法的应用,可以看出,三种评价方法各有侧重。其中 Z 比分数法是基础,其他两种方法的参数均由 Z 值衍生而来。Z 比分数法计算的是每个考核项目的 Z 值,更适用于对单次考核中每个项目的单独分析,该方法侧重细节评价,更容易发现个别异常值并对其进行跟踪和纠正。点分数系统法以各考核项目 Z 值为基础,计算得到一个实验室在单次考核中的总点分数,更适用于对单次考核中每个实验室对多个考核项目的综合表现做出评价;同时,该方法的评价体系还可以同时对定性分析中假阳性和假阴性结果进行量化赋分,结果更加全面,信息量丰富;但该方法对定量结果的判定将可疑和不满意列入一个等级,其适用范围还需进一步深入探讨。组合分数法以多次考核中所有

目标物的 Z 值为基础,计算组合参数,更适用于评价一个实验室在多次能力验证中的总体表现,该方法能够对实验室检验能力进行持续跟踪,避免了因一次失误而引起的不公正评价,为更加客观、综合衡量实验室水平提供了途径。

综上,三种方法在数据范围、评价目标方面具有良好的互补性,综合使用能够实现有能力验证结果从细节到整体的全面评价。

## 参考文献

- [ 1 ] 中国合格评定国家认可委员会. CNAS—GL02 能力验证结果的统计处理和评价指南[S]. 2006-06-01.
- [ 2 ] ISO. ISO/IEC 17043: 2010 Conformity assessment—General requirements for proficiency testing[S]. ISO, Switzerland, 2010. <http://www.iso.org/iso/home/standards.htm>
- [ 3 ] ISO. ISO 13528:2005 Statistical methods for use in proficiency testing by interlaboratory comparisons[S]. ISO, Switzerland, 2005. <http://www.iso.org/iso/home/standards.htm>
- [ 4 ] Federal Office of Consumer Protection and Food Safety (BVL).

- NSAIDs in Milk, Interlaboratory Study NSAID, Report on Results[R]. Berlin: BVL, 2010, 21-29.
- [ 5 ] Manfred Stoyke, Wolfgang Radeck, Petra Gowik. Anthelmintics in bovine milk and muscle: interlaboratory studies among EU National Reference Laboratories[J]. Accred Qual Assur, 2012, 17(4): 405-412. DOI 10. 1007/s00769-012-0919-x.
- [ 6 ] Von Holst C, Alder L. In: Fernandez-Alba A R (ed) Comprehensive analytical chemistry XLIII [M]. chap 10. Elsevier, Amsterdam, 2003.
- [ 7 ] Uhlig S, Lischer P. Statistically-based performance characteristics in laboratory performance studies [J]. Analyst 1998, 123: 167-172.
- [ 8 ] 赵馨,李业鹏,常迪,等. 2006 年全国省级疾控中心食品中铅、砷含量实验室间比对结果分析[J]. 中国卫生检验杂志, 2008, 18(1): 10-11.
- [ 9 ] 马兰,赵馨,周爽,等. 2011 年省级疾控中心食品中金属污染物质控考核结果分析[J]. 中国食品卫生杂志, 2012, 24(4): 361-363.
- [ 10 ] FAPAS. Protocol for Proficiency Testing Schemes [M]. 2nd. FAPAS, 2010: 10-12.

## · 法规文件 ·

# 关于印发《2013 年食品安全国家标准项目计划》的通知

卫办监督函〔2013〕359 号

各有关单位:

根据《食品安全法》和《食品安全国家标准管理办法》规定,我委在向社会公开征求意见的基础上制定了《2013 年食品安全国家标准项目计划》,现印发给你们,请认真组织落实。有关工作要求如下:

### 一、填报项目委托协议书,及时落实食品安全国家标准项目计划

2013 年食品安全国家标准计划项目承担单位应当填写《2013 年食品安全国家标准制(修)订项目委托协议书》(可从卫生计生委网站 <http://www.moh.gov.cn> 下载),打印后由承担单位负责人签字并加盖单位公章(一式五份),于 2013 年 5 月 20 日前报送食品安全国家标准审评委员会秘书处(以下简称秘书处)。逾期未提交协议书的,视为自动放弃标准起草单位和起草人资格。秘书处对协议书进行审核后,于 2013 年 5 月 31 日前报送我委。

### 二、加强日常管理,确保食品安全国家标准项目及相关经费按时保质执行

(一)项目承担单位和项目负责人要加强食品安全国家标准制定、修订工作的管理,保证项目质量和进度,请于 2013 年 12 月 30 日前向秘书处提交工作中期进展报告和经费使用情况报告,于 2014 年 6 月 30 日前完成任务,向秘书处提交送审材料和经费决算报告。经费决算报告由财务负责人和单位负责人签字并加盖公章。

(二)未按期完成任务提交送审材料的,项目承担单位和项目负责人应当提交说明,并附经费使用情况报告,加盖单位公章后报秘书处。我委将视情况予以通报批评,并根据国家有关财经法规制度,对已拨付的项目经费采取追回等必要的处理措施。

(三)相关省(区、市)卫生厅(局、卫生计生委)、有关单位要支持并督促下属单位承担的项目工作,秘书处要督促检查项目执行情况,确保项目计划整体进度。

附件:食品安全国家标准制(修)订项目委托协议书.docx(略)

国家卫生和计划生育委员会办公厅

二〇一三年五月二日