

汉佩尔辨识法在酶联免疫吸附试验室内质控中的应用

周月兰, 苏武锦(广西壮族自治区南宁市中心血站 530003)

【摘要】 目的 探讨汉佩尔辨识法(Hampel identifier)在室内质控图中的应用,评估其对质控离群值的检出能力。**方法** 使用弱阳性 HIV 质控品和常规标本一起检测,连续测定 20 次,收集原始数据,并在其基础上衍生模拟数据;然后对原始数据和模拟数据进行分析并绘制 Z-分数图、Hampel identifier (Zi-分数图)与箱式图,比较 3 种质控图对离群值的检出能力。**结果** 对于非正态数据,Hampel identifier 与箱式图的离群值检出能力一致,均高于 Zi-分数图。**结论** 使用多种离群值检测方法的功能、应用条件及其限制,防止离群值的漏判。

【关键词】 Zi-分数; 汉佩尔辨识法; 箱式图; 质量控制图; 离群值
DOI:10.3969/j.issn.1672-9455.2015.11.024 文献标志码:A 文章编号:1672-9455(2015)11-1555-02

The application of Hampel identifier in internal quality control about ELISA ZHOU Yue-lan, SU Wu-jin (Nan-ning Blood Center, Nanning, Guangxi 530003, China)

【Abstract】 Objective To study the value of Hampel identifier in internal quality control and to assess the detection ability of quality control outliers. **Methods** The weak positive quality product of HIV together with routine specimens were detected for 20 times, raw data and simulation data derived on the basis of the raw data were collected and analyzed, Z-scores figure, Hampel identifier (Zi-scores figure), box plot correspondingly were drawn, and the ability of outlier detection about the three kinds of quality control charts were compared. **Results** Hampel identifier was consistent with the box plot on outlier detection ability for abnormal distribution data, both were superior to Z-score chart. **Conclusion** Medical laboratory managers should master functions, application conditions and limitations of outlier detection methods to prevent the missing outliers.

【Key words】 Z-scores; Hampel identifier; box plot; quality control charts; outliers

目前,基于数据正态性分布特性的 Zi-分数图已广泛应用于我国医学实验室质量控制^[1-2],但实际应用时,由于受某些外源性 & 内源性因素的影响,酶联免疫吸附试验(ELISA)检测的质控数据不一定完全符合正态分布,当数据中出现离群值时常被忽略^[3-4]。有学者认为箱式图对非正态分布数据离群值的检出能力更强^[5-6]。Wilcoxon^[7]推荐使用汉佩尔辨识法(Hampel identifier)的质控方法在国内却鲜有报道。现通过实例及模拟数据,探讨该法的应用条件及优缺点,报道如下。

1 资料与方法

1.1 一般资料 弱阳性(4 NCU/mL) HIV 质控品(北京康彻斯坦生物技术有限公司,生产批号:201403001)及常规标本。

1.2 仪器与试剂 全自动酶免分析系统(STAR 全自动加样器及 FAME 全自动酶免分析仪);伯乐公司 HIV 诊断试剂盒(ELISA),是国家批准检验合格产品,于有效期内使用。

1.3 方法

1.3.1 数据收集 每批样品均平行检测常规标本、试剂盒内阴性对照品和弱阳性质控品(1 次),并连续测定 20 次,收集质控数据 20 例(A 组),同时从原始数据中衍生模拟出其他 4 组数据:B 组,怀疑存在 1 个高值为离群值;C 组,怀疑存在 1 个以上的高值离群值;D 组,怀疑存在 1 个以上的高值离群值(个别离群值很高);E 组,怀疑存在 1 个高值和 1 个低值离群值。

1.3.2 质控图绘制 使用 Z-分数图、Hampel identifier (Zi-分数图)及箱式图分别对各组数据进行绘图。(1)Z-分数图: $Z = \frac{X_i - \bar{X}}{S}$, X_i 为单个检测值, $\bar{X}$ 为均值, S 为标准差。Z-分数图平

均数为 0,界限为 ± 1 、 ± 2 、 ± 3 ,刻度至少从 $-4 \sim +4$ 。若 $|Z| \geq 3$ 则为离群值。(2) Zi-分数图: $Z_i = \frac{0.6745(X_i - X_m)}{MAD}$, X_m 是该批检测数据 X_i 的中位数, MAD 是

该批检测数据(X_i)减去该批数据中位数差值之绝对值($|X_i - X_m|$)的中位数^[7-9]。Zi-分数图平均数为 0,界限为 ± 1 、 ± 2 、 ± 3 ,刻度至少从 $-4 \sim +4$ 。若 $|Z_i| \geq 3$ 则为离群值。(3)箱式图:P2.5、P25.0、P50.0、P75.0 及 P97.5 为描述定量变量的 5 个百分位点。箱式图的 5 条线分别表示该 5 个百分位点,由 P2.5~P25.0 和 P75.0~P97.5 构成两条“丝”,为两端 45% 的数据;由 P25.0~P75.0 构成“箱”,为中间 50% 的数据。离群值为距箱上缘或下缘 1.5~3 倍四分位数间距的值,用“O”表示;极值为超出距箱上缘或下缘 3 倍四分位数间距的值,用“<”表示。

1.3.3 数据分析 使用 Z-分数图、Zi-分数图及箱式图分别对各组数据进行分析,比较 3 种质控图对质控离群值的检出能力。

1.4 统计学处理 采用 SPSS 17.0 软件计算统计量;夏皮洛-威尔克(Shapiro-Wilk)检验进行正态性检验;MedCalc11.4.2.0 英文版绘制箱式图;Microsoft Office Excel 2003 绘制 Z-分数图 & Zi-分数图。 $P < 0.05$ 为差异有统计学意义。

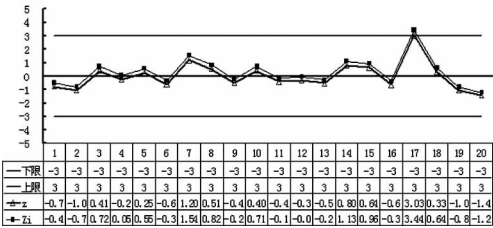
2 结果

2.1 数据统计 原始数据(A 组)和衍生模拟数据(B、C、D、E 组)经 Shapiro-Wilk 正态性检验,均 $P < 0.05$ 。

2.2 Z-分数图与 Zi-分数图的比较 A 组中,例 17 的 Z 分数为 3.03, > 3 , Zi 分数为 3.44, > 3 , Z-分数图 & Zi-分数图均检

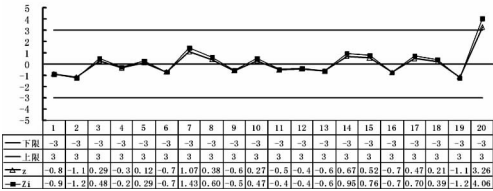
出例 17 为离群值。B 组中,计算获得的 20 个 Z 分数绘制成的 Z-分数图,检出例 20($Z=3.26$)为离群值,Zi-分数图也检出例 20($Z_i=4.06$)为离群值。C 组中,Z-分数图未检出离群值,Zi-分数图则检出例 17($Z_i=3.53$)、例 20($Z_i=3.98$)为离群值。D 组中,Z-分数图仅检出例 20($Z=4.05$)为离群值,Zi-分数图则检出例 17($Z_i=3.53$)、例 20($Z_i=15.71$)为离群值。E 组中,Z-分数图未检出离群值,Zi-分数图则检出例 17($Z_i=3.44$)、例 20($Z_i=-3.23$)为离群值,见图 1~5。

2.3 箱式图分析 A 组中,例 17 为离群值。B 组中,例 20 为离群值。C 组中,例 17、例 20 为离群值。D 组中,例 17 为离群值、例 20 为极值。E 组中,例 17、例 20 为离群值。见图 6。



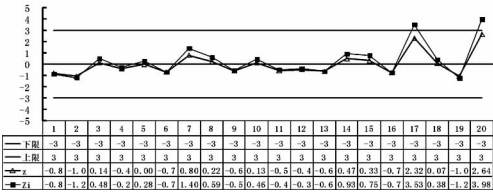
注:Z-分数图与 Zi-分数图比较。

图 1 A 组分数图



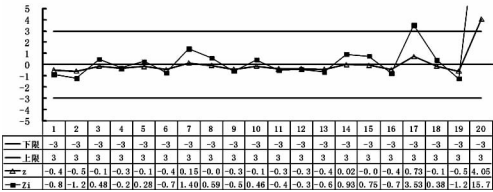
注:Z-分数图与 Zi-分数图比较。

图 2 B 组分数图



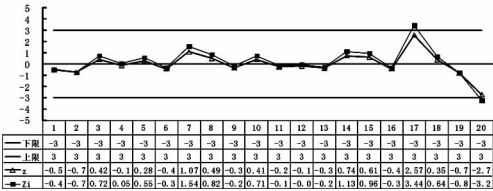
注:Z-分数图与 Zi-分数图比较。

图 3 C 组分数图



注:Z-分数图与 Zi-分数图比较。

图 4 D 组分数图



注:Z-分数图与 Zi-分数图比较。

图 5 E 组分数图

2.4 Zi-分数图与箱式图的比较 2 种质控图检出离群值的总

数、例数相一致。

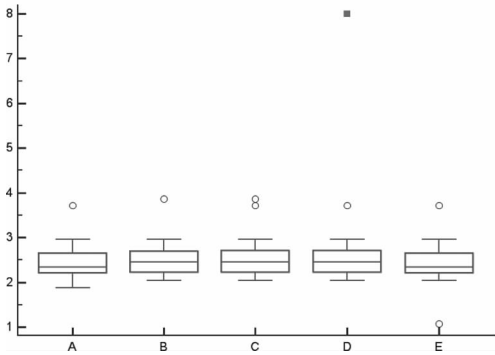


图 6 箱式图

3 讨论

离群值是指样本中的 1 个或多个观测值,离开其他观测值较远,提示其可能来自不同的总体^[10]。本研究利用 Z-分数图、Zi-分数图及箱式图,分别对实例及模拟数据的离群值进行分析,本组结果显示:(1)当数据中仅存在 1 个离群值时(A 组、B 组),3 种方法均能检出,但 Z-分数图法不如 Zi-分数图法敏感(见图 1、图 2),A 组中第 17 个检测值的 Z 分数为 3.03 时,Zi 分数为 3.44;B 组中第 20 个检测值的 Z 分数为 3.26 时,Zi 分数为 4.06。(2)当数据中存在 2 个同一方向离群值时(C 组),Zi-分数图、箱式图均能检出,而 Z-分数图则不能(见图 3),除非存在很大的离群值时(D 组),最大的数值方能被检出(见图 4),但此时该组的 CV 值已高达 46.48%,远大于 ELISA 实验批间差异的 20%最高限。(3)当数据中存在 2 个不同一方向(即一高一低)离群值时(E 组),Zi-分数图、箱式图均能检出,而 Z-分数图则不能(见图 5),但第 20 个检测值($S/CO=1.07$)已经非常接近临界值($Cut\ Off\ 值=1$),若再小于 1 时,表明该实验处于失控状态了,而 Z-分数图再检出离群值也无实际意义。

Z 分数,即单个观察值减去总体均值,得到的差值除以总体标准获得的无因次度量。其用于判断离群值操作简便,但较适于符合正态分布的变量。ELISA 受试剂、设备、温度、操作各个环节等多种因素影响,检测数据不一定完全符合正态分布,若出现离群值,其标准差变大,Z 分数相应变小其绝对值可能会小于 3,离群值出现漏检。当存在多个离群值时,还会出现某个离群值会屏蔽另一个离群值被检出的情况,甚至导致无任何离群值被检出(见图 3~5)。

国外学者汉佩尔(Hampel)建议以中位数和中位数绝对偏差(MAD)作为位置和宽度的稳健估计,已有相关报道证明该方法的有效性^[7-9]。箱式图的箱子两端分别是上四分位数和下四分位数,中间横线是中位数,两端连线分别是除离群值外的最小值和最大值,另外标记的可能是离群值,其更适于非正态分布资料^[9-11]。本研究结果表明,Zi-分数图与箱式图对非正态资料的离群值检出能力是一致的,两者均高于 Z-分数图。

本组认为,判断数据是否存在离群值,须考虑数据中是否存在单个或多个离群值,有些检出离群值的方法只用来检出 1 个离群值,如 Z-分数图。这是否可以称之为“屏蔽效应”,还需做进一步研究,但目前至少可以确定某些离群值的存在会掩盖另一个离群值的检出,甚至导致所有离群值都无法被检出的可能。因此,医学实验室的管理人员,应该掌握多种离群值检测功能、应用条件及其限制,防止离群值的漏判。

参考文献

[1] 陈雪,李书平,赵欣,等.血液核酸筛查(下转第 1559 页)

3 讨 论

脑卒中在我国已成为中老年病死的主要因素之一,其不仅发病率高,且并发症及后遗症多,具有较高的伤残率和病死率,同时还有病程较长、治愈困难等特点^[7]。脑卒中发病常为突发事件,使其家属面临巨大压力,常使患者家属处于一种身心应激状态,产生焦虑、恐惧等情绪反应。同时患者常会伴有不同程度的神经功能损伤,生活难以自理,所以后期的康复也是一长期复杂的过程。在患者的急救、治疗和康复过程中,家属往往面临患者病情变化、经济、情感、照顾与工作学习的矛盾等压力,导致各种心理问题。

本研究结果显示,大部分家属需要承担多重角色,正常的工作生活受到严重影响,使其身心负担加重,出现一系列的不适应证,影响心理健康。本组调查结果表明,脑卒中患者家属的心理健康水平较低,其中躯体化、抑郁、焦虑、敌对等 SCL-90 阳性项目分数均高于常模,差异有统计学意义($P<0.05$ 或 $P<0.01$),与张玲云等^[8]的研究结果基本相符。对影响患者家属心理健康状况因素的统计分析结果提示,经济状况、与患者的关系、患者神经缺失程度、患者意识水平都对其心理健康状况影响显著。原因如下:(1)与患者的关系:子女偏执得分显著高于配偶。在家庭中,子女对父母往往存在不同程度的心理依赖,对患者康复的愿望表现更为强烈,因此在心理上表现更为偏执。王俊涛^[9]研究认为脑卒中患者的家属照顾者中配偶的焦虑程度高,但本研究显示配偶与子女的焦虑程度相似,可能与近年来子女对老年患者照顾的参与程度的增加有关。(2)经济状况:低收入家属在支付患者治疗费用问题上存在较多担忧,SCL-90 的抑郁和焦虑因子得分上差异显著,而在应对方式上,经济能力中等者表现更为积极,与实际情况有所差异。医师应一方面在制订治疗方案的过程中尽量考虑患者的经济状况,另一方面也应充分与家属沟通,使其对治疗费用有一定的心理准备,减少不良情绪的产生。(3)患者神经缺失程度:患者的神经缺失程度直接影响家属照顾患者的程度,神经功能缺失越严重,自理能力越差,越依赖家属的治疗,家属的心理压力也会越大。(4)患者意识状态:脑卒中患者在 2 级意识水平下抑郁、焦虑、SAI、TAI、抑郁自评分数较高,意识水平过低或相对正常时,家属的心理健康状况较 2 级意识水平情况下更好,可

能与心理期待有关,2 级意识水平下是否为一个期待的关键转折点,到目前为止尚未有研究证实。患者家属的心理压力与患者意识状态程度密切相关,随着患者意识状态的改善,家属的心理压力也得以减轻,因此患者的疗效是家属心理压力改善的最根本途径。

综上所述,本研究对急性脑卒中患者家属心理健康状况的分析结果提示,医务护理人员需从多方面对家属进行积极的心理干预,增强脑卒中患者的社会支持程度,有利于患者的尽早康复。

参考文献

- [1] 高强, 关敏, 吴志刚, 等. 急性期脑卒中患者家属焦虑状况调查分析[J]. 中国康复医学杂志, 2010, 25(9): 903-905.
- [2] Chou KR, Lin JC, Chu H. The reliability and validity of the Chinese version of the caregiver burden inventory[J]. Nurs Res, 2002, 51(5): 324-331.
- [3] 张作记. 行为医学量表手册[M]. 北京: 中华医学电子音像出版社, 2005: 212, 223.
- [4] 汪向东, 王希林, 马弘. 心理卫生评定量表手册[M]. 北京: 中国心理卫生出版社, 1999: 109, 113.
- [5] 中华神经科学会. 美国国立卫生院神经功能缺损评分(NIHSS)[J]. 中华神经科杂志, 1996, 29(6): 381-383.
- [6] 王忠诚. Glasgow 昏迷评分(GCS)对脑卒中患者的评价[J]. 神经外科学, 1998, 11(6): 112-115.
- [7] 石滴坚. 微创碎吸术手术时机对高血压脑出血患者血清炎性细胞因子水平的影响[J]. 疑难病杂志, 2010, 9(2): 92-95.
- [8] 张玲云, 许春艳. 社区脑卒中患者家庭家属照顾者心理健康状况调查[J]. 中华现代护理杂志, 2009, 15(34): 3641-3642.
- [9] 王俊涛. 脑卒中患者主要照顾亲属心理健康状况的影响因素分析[J]. 中国实用护理杂志, 2012, 28(32): 69-70.

(收稿日期: 2014-12-25 修回日期: 2015-02-10)

(上接第 1556 页)

- 利用室内质控品 Ct 值建立质控图方法初探[J]. 中国输血杂志, 2014, 27(11): 1154-1155.
- [2] 罗伟, 罗智敏, 刘建兵, 等. 用 Excel 绘制两种实验室常用室内质量控制图方法[J]. 检验医学与临床, 2013, 10(5): 620-622.
 - [3] 缪希莉. 临床免疫检验质量影响因素分析[J]. 检验医学与临床, 2014, 11(3): 395-397.
 - [4] 苏武锦, 钟周琳, 邱昌文, 等. 四分位极差在血站 ELISA 质控中的应用研究[J]. 国际检验医学杂志, 2014, 35(6): 735-737.
 - [5] David C, Hoaglin-Frederick M, John WT, 等. 探索性数据分析[M]. 北京: 中国统计出版社, 1998: 62-101.
 - [6] 官生平. SPC 统计过程管制[M]. 厦门: 厦门大学出版社, 2004: 104-106.

- [7] Wilcox R. Applying contemporary statistical techniques [M]. Amsterdam: Academic Press, 2003: 55-79.
- [8] Perarson RK. Outliers in process modeling and identification[J]. IEEE Transactions on Control Systems Technology, 2002, 37(10): 55-63.
- [9] Liu H, Shah S, Jiang W. On-line outlier detection and data cleaning[J]. Computers and Chemical Engineering, 2004, 67(28): 1635-1647.
- [10] 国家质量监督检验检疫总局. GB/T4883-2008 数据的统计处理 and 解释正态样本离群值的判断和处理[S]. 北京: 中国标准出版社, 2008.
- [11] 孙振球. 医学统计学[M]. 3 版. 北京: 人民卫生出版社, 2010: 18-19.

(收稿日期: 2014-12-25 修回日期: 2015-02-10)