

临床检验质量控制程序的性能验证*

张诗诗, 王 薇, 王治国 (北京医院 国家老年医学中心 卫生部临床检验中心, 北京 100730)

摘要:随着实验室质量控制(QC)计划的发展和风险管理的引入,许多不同类型的QC可用于操控分析过程的性能并及时准确地检出可能发生的任意误差。因此,进一步验证这些QC程序的有效性对于提升实验室管理水平是至关重要的。针对给定的质量控制策略,实验室需要确认其合理性并定义性能度量。通过指定一个区分结果可接受或不可接受的质量规范,结合允许总误差(TEa)表明分析过程的固有精密度;运用西格玛(σ)指标评价分析系统在控操作期间产生不可靠结果的概率并判断分析过程对于失控条件的耐受程度;用 ΔP_E 反映失控和在控状态产生不可靠结果的概率差;以及用 P_{ed} 表示质控规则的误差检出概率;实验室可以评估因失控而产生的不可靠患者结果预期数($E[N_u]$),验证质量控制程序的性能,并判断产生及报告不可靠患者结果的风险是否可以接受。对于符合实验室风险标准的质量控制策略,在将失控状态判断为失控之前,先于和后于最终可接受的质控活动所报告的不可靠患者结果的预期数可以用作设计标准,分别用 $E(N_{uf})$ 和 $E(N_{uc})$ 表示。通过计算最大的 $E(N_{uf})$ 和 $E(N_{uc})$ 值以及确定其是否被实验室接受,进一步验证临床实验室的质控策略是否合理。

关键词:质量控制;失控状态;质量控制性能评估;质控策略

中图分类号:R446 文献标志码:A 文章编号:1671-7414(2016)06-150-05

doi:10.3969/j.issn.1671-7414.2016.06.045

Performance Verification of Quality Control Procedures in Clinical Laboratory Medicine

ZHANG Shi-shi, WANG Wei, WANG Zhi-guo (National Center for Clinical Laboratories,
Beijing Hospital, National Center of Gerontology, Beijing 100730, China)

Abstract: With the development of laboratory Quality Control (QC) plans and the introduction of risk management, numerous different types of quality control measures were applied to handle the performance of analytical processes and to timely and accurately detect any errors that may occur. Therefore, it is of critical importance to improve the level of laboratory management by further verifying the effectiveness of quality control procedures in the clinical laboratories. For a given QC strategy, the clinical laboratory should validate its rationality and define the performance indicators. By designating a quality specification which can distinguish acceptable and unacceptable results, combining the allowable total error (TEa) to demonstrate the inherent imprecision of analytical processes, applying the sigma indicator to evaluate the probability of producing unreliable results in-control state by the analytical system and determine the degree of tolerance for out-of-control conditions in analysis, using ΔP_E to reflect the discrepancy of probabilities about producing unreliable results between out-of-control and in-control states, and expressing with P_{ed} to indicate the probability of error detection capability by the QC rules, the clinical laboratory can estimate the expected number of unreliable patient results ($E[N_u]$) produced because of out-of-control conditions, verify the performance of QC procedures, and decide whether the risk of producing and reporting unreliable patient results can be accepted. For the QC strategies up to the laboratory risk standards, before the determination of out-of-control conditions, the expected number of unreliable patient results produced before and after the last accepted QC estimation can be used as the design criteria, abbreviated as $E(N_{uf})$ and $E(N_{uc})$ respectively. Then the clinical laboratory can further validate the rationality of a QC strategy by computing the maximum $E(N_{uf})$ and $E(N_{uc})$ values and determining if they are acceptable to the laboratory.

Keywords: quality control; out-of-control condition; QC performance evaluation; QC strategy

1 质量控制性能验证的概述 随着实验室质量控制(QC,简称质控)计划的发展和风险管理的引入,许多不同类型的QC可以用于操控分析过程的性能并及时准确地检出可能发生的任意误差。因此,

进一步验证这些QC程序的有效性对于提升实验室管理水平是至关重要的。

首先,验证实验室是否有一个合理的质控策略。传统上,验证质控策略集中于评价其是否有足

* 基金项目:北京市自然科学基金(基金编号:7143182),北京医院资助课题(BJ-2015-025)。

作者简介:张诗诗(1991—),女,在读硕士研究生,研究方向:实验室质量管理,E-mail:Athena0915@163.com。

通讯作者:王治国,硕士,研究员,主要从事临床检验质量控制方法研究和室内质量评价工作,E-mail:zgwang@nccl.org.cn。

够的功效去检出任何“关键的”失控状态(指 $\geq 5\%$ 的检验结果处于不可接受的状态)^[1,2]。

其次,定义性能度量。在风险管理的时代,要求用不同的性能度量来判定 QC 策略的功效。关于特定的 QC 策略在失控时是如何影响患者标本的检验,即为探讨其由于失控条件而产生的不可靠患者结果预期数 $E(N_u)$ 。因为不知道即将发生的失控状态达到何种程度以及各程度的失控状态产生的不同 $E(N_u)$,所以 QC 验证问题变成“任何程度的失控状态产生的 $E(N_u)$ 是否可以接受?”^[3]。

本文重点在于,针对给定的质控策略,如何计算 $E(N_u)$ 以及如何使用它来验证质控策略的可接受性^[4]。质控策略包括每次质控中评估的质控样本数、质控频率(或质控活动之间被评估的患者样本平均数)和质控规则,目的是用于检出在控分析系统的改变。根据系统失控时产生额外的不可靠结果数高于在控时产生的不可靠结果数,可以测量质控策略的功效。

2 质量规范

2.1 允许总误差(TEa) 虽然指定一个可将结果分为接受或不可接受的质量规范需要投入大量的工作^[5,6];但实际的重点在于可运用它来验证质控程序。本文中我们使用表达为百分数的 TEa。TEa 通常是指定的单一值或在不同浓度下分别给出的不同值,但关键是 TEa 需要反映出所要求的质量。

当标本检测结果的报告值与其真值有差别并超过 TEa 时,该结果被认为是不可靠的。虽然分析过程始终存在一些固有分析误差使得标本结果的真值难以确定,但在理想的情况下,当分析系统在控时,固有分析误差的大小与 TEa 相比是很小的。

2.2 西格玛(σ)度量 σ 是利用偏倚调整 TEa 之后与稳定的在控分析不精密度的比值(如公式)。 σ 可用于评价分析系统在控操作期间产生不可靠结果的几率有多大^[7],因为即使分析系统处于在控状态,也有一定的产生不可靠结果的概率。 σ 也是判断分析过程在失控状态下有多大耐受程度的一种简便方法。见图 1,若 1 000 000 个患者标本在失控条件出现时采用 6σ 过程进行评估,且失控条件会在评估过程中产生 2σ 偏倚,则预期有 32 个结果是不可靠的;若 1 000 000 个患者标本在失控条件出现时采用 2σ 过程进行评估,且失控条件会在评估过程中产生 2σ 偏倚,则预期有 454 531 个结果是不可靠的。

$$\sigma = \frac{TEa - |\text{bias}|}{CV}$$

σ 越低,对于给定的失控条件将会有更高的 $E(N_u)$ 。通常,低 σ 过程比高 σ 过程要求更多的质控资源来产生相等的 $E(N_u)$ 。

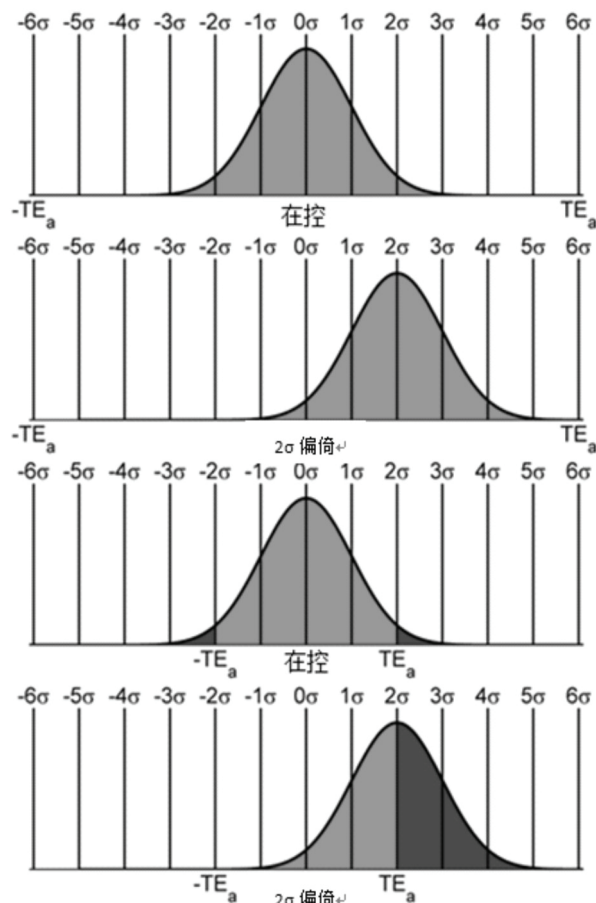


图 1 6σ 过程和 2σ 过程

顶部的 2 张测量误差分布图显示了采用允许 TEa 为 $\pm 6\sigma$ 的一个在控过程和一个带有 2σ 偏倚的失控过程。底部的 2 张测量误差分布图显示了采用 TEa 为 $\pm 2\sigma$ 的一个在控过程和一个带有 2σ 偏倚的失控过程。浅色阴影面积表示产生结果符合明确质量要求的概率,深色阴影面积表示产生结果不符合明确质量要求的概率。

3 失控检出

3.1 失控和在控状态产生不可靠结果的概率差(ΔP_E) 一般认为,失控严重的状态产生不可靠结果的可能性高,而失控微小的状态对产生不可靠的结果几乎没有影响。当分析系统处于在控状态时可以计算产生不可靠结果的概率,当失控条件的大小被给出时也可以计算产生不可靠结果的概率,且二者的概率差可以表示为 ΔP_E 。

见图 2,当分析过程在控时, $SE=0$, $\Delta P_E=0$;对于足够严重的失控状态, ΔP_E 会接近 1。计算失控状态的 ΔP_E ,若失控范围从 0(在控)到 1.5 倍 TEa ($1.5 \times 4\sigma = 6\sigma$),将产生一个从 0 过渡到接近

于1的 ΔP_E 曲线作为评估失控状态严重程度的函数。

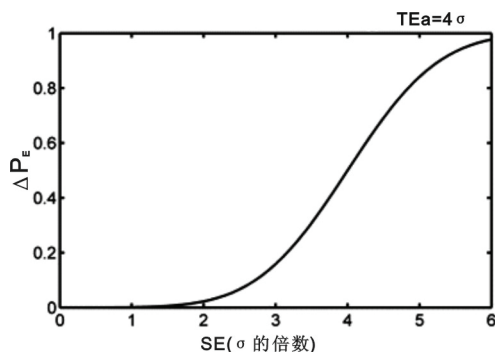


图2 ΔP_E 曲线

SE代表失控状态的系统误差,以过程稳定的分析不精密度(σ)的倍数表示。 ΔP_E 表示产生不符合规定TEa的患者结果的概率变化。

3.2 误差检出概率(P_{ed}) 针对已明确严重程度的失控状态,运用质控规则判断其为失控的概率称为该质控规则的误差检出功效,即误差检出概率,表示为 P_{ed} 。

检出已明确严重程度的失控状态需要进行的质控活动预期(或平均)数与 P_{ed} 有关,呈反比关系;即 P_{ed} 越高,检失控状态所要求的质控活动平均数越低。对于任意给定的质控规则和失控状态严重程度,检出失控的质控活动平均数是 $1/P_{ed}$ 。

不同的质控规则有不同的功效曲线^[8~10]。如图3,对于 2σ 偏倚的失控状态,质控规则判断失控的概率大约是40%($P_{ed}=0.4$),因此,平均要用2.5次的质控活动($1/P_{ed}=1/0.4$)来检出分析检测系统中的 2σ 失控偏倚。

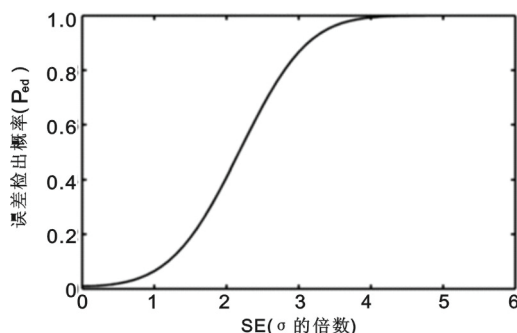


图3 误差检出概率(P_{ed})

具有2个浓度水平质控品 $1_{3s}/2_{2s}/R_{4s}$ 多规则质控的功效曲线^[8],横轴SE代表失控状态的系统误差,以 σ 的倍数表示。SE=0意味着分析过程处于在控状态。纵轴代表质控规则判断失控的概率,作为评价失控状态严重程度的函数。

3.3 受影响患者的预期数($ANP_{affected}$) 将 $1/P_{ed}$ 与质控活动之间检查患者样品的平均数相结合,在

识别和解决持续失控状态之前,可以预测有多少患者样本受影响,即受影响患者的预期数 $ANP_{affected}$ ^[11]与质控活动之间的受检患者样本平均数呈正比,与质控频率呈反比。

当分析系统在控时(SE=0), $ANP_{affected}$ 很大,反映了质控规则在假失控(在无失控状态时质控规则判断为失控)之间所评估的患者样品数。见图4,质控规则假失控率大约是0.01,期望平均每100次质控活动发生1次质控规则假失控,且在质控活动之间平均检测100个患者样本,则在控系统在该质控规则假失控之间的 $ANP_{affected}$ 是 $100 \times 100 = 10000$ 。

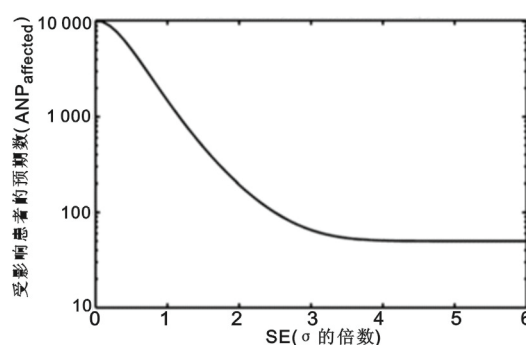


图4 受影响患者的预期数($ANP_{affected}$)

失控程度越严重,质控规则的 P_{ed} 越大且 $1/P_{ed}$ 越小,导致 $ANP_{affected}$ 也越小。对于足够严重的失控状态,在出现失控状态后的第一次评估,将确保质控规则判断为失控。假设在检测过程中的任意时刻失控状态有相同的发生概率,当极严重的失控状态在其发生后的第一次质控活动中被检出时, $ANP_{affected}$ 将是质控活动间所评估的患者样本平均数的一半。

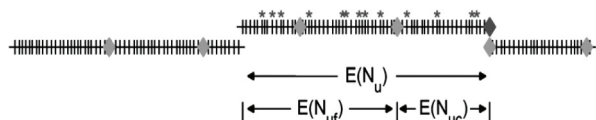
$ANP_{affected}$ 可作为评价失控状态严重程度的函数^[11]。横轴SE代表失控状态的系统误差,以 σ 的倍数表示。SE=0意味着分析过程处于在控状态。纵轴表示 $ANP_{affected}$,以对数刻度表示。

4 $E(N_u)$ 的划分和运用

4.1 $E(N_u)$ 的划分 $E(N_u)$ 是由于失控状态而产生的不可靠患者结果预期数,它是 ΔP_E 和 $ANP_{affected}$ 的积^[12,13]。对于严重的失控状态,一般经验规则是 $E(N_u)$ 为质控活动之间患者数量的一半^[14]。

$E(N_u)$ 提供了一种度量,直接描述产生患者结果的可靠性以及质控策略受给定严重程度的失控状态影响的性能之间的关系,且可以进一步分成两部分。在失控状态发生和最后可接受的质控之前所产生的不可靠结果被认为是不能识别和纠正的“最终的”结果,其预期数缩写为 $E(N_{uf})$ 。在最后

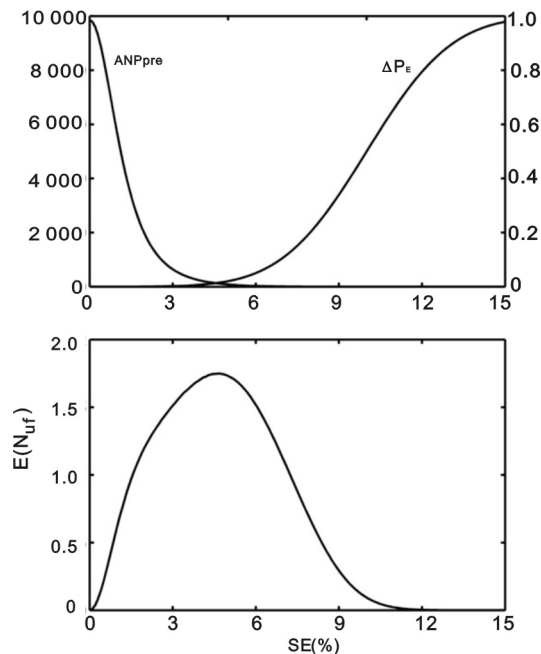
可接受的质控之后所产生的不可靠结果由于大部分是近期产生的则被认为是“可纠正的”结果,其预期数缩写为 $E(N_{uc})$ 。图5中 $E(N_{uf})=9$ (从发生失控状态到最后的可接受质控之间有9个红色星号)及 $E(N_{uc})=4$ (最后的可接受质控和质控判断为失控之间有4个红色星号)。



向上移动的线代表失控状态,水平线代表随时间变化的分析过程;每条垂直线代表一次患者样本检测。每个菱形代表一次质控活动;绿色菱形代表质控在控,红色菱形代表质控失控。红色星号代表不可靠的患者样本结果。

图5 由于失控状态而产生的不可靠结果预期数 $E(N_u)$

4.2 运用 $E(N_{uf})$ 是 ANP_{pre} (发生失控状态和最后的可接受质控活动之间受影响患者的预期数)和 ΔP_E 之积(见图6)。 $E(N_{uf})$ 对于非常微小和非常严重的失控状态均很小,并且在非常微小和非常严重的失控状态之间的某处获得最大值。若最大的 $E(N_{uf})$ 值被认为是可接受的,则对于任意的失控情况,质控策略可以保证实验室质量;若最大的 $E(N_{uf})$ 值被认为是不可接受的,则应该检查和调整质控策略,直到实现可接受的最大 $E(N_{uf})$ 。 SE 代表失控状态系统误差的大小,以百分数表示。

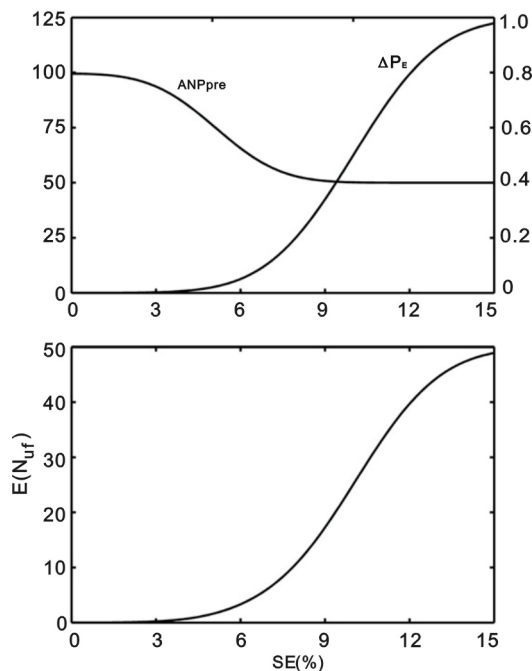


SE 代表失控状态系统误差的大小,以百分数表示。

图6 ANP_{pre} , ΔP_E 和 $E(N_{uf})$ 之间关系的实例

$E(N_{uc})$ 是 ANP_{post} (最后的可接受质控活动和质控活动判断为失控之间受影响患者的预期数)与 ΔP_E 之积(见图7)。 ANP_{post} 依赖于失控状态发生

后的第一次质控活动能否检出失控情况。如果失控状态在其发生后的第一次质控活动中未被检测出,那么在最后的可接受质控活动和质控活动判断为失控之间所有的患者样本将受失控状态的影响。如果失控状态在其发生后的第一次质控活动中被检测出,那么如上文所述,我们预期在质控活动之间被检测的患者样本平均数有一半会受到影响^[11]。 SE 代表失控状态系统误差的大小,以百分数表示。



SE 代表失控状态系统误差的大小,以百分数表示。

图7 ANP_{post} , ΔP_E 和 $E(N_{uc})$ 之间关系的实例

最大的 $E(N_{uc})$ 值将总是接近于质控活动之间被检测患者样本平均数的一半。实验室在失控状态被识别和解决后,需要及时纠正最大的 $E(N_{uc})$ 值,以阻止不可靠的患者结果在发生严重的失控状态时被不合理地运用。若实验室有到位的程序可以及时识别、重复和纠正最大的 $E(N_{uc})$ 且使其达到质控活动之间被检测的患者样本平均数的一半,那么最大的 $E(N_{uc})$ 值是可以接受的。否则,应该采取措施,如改变质控频率来减少最大的 $E(N_{uc})$ 值。

5 总结 质控策略是由质控过程中所评估的质控物样本数,质控规则以及质控频率(质控活动之间被评估的患者样本平均数)所组成。通过对 $E(N_u)$ 的划分并计算最大的 $E(N_{uf})$ 和 $E(N_{uc})$ 值以及确定它们是否被实验室接受,进一步验证临床实验室质控策略的合理性。

若最大的 $E(N_{uf})$ 值太高,可以通过增加质控过程中被评估的质控物样本数,使用有更高 P_{ed} 的质控规则,或者减少质控活动之间 (下转 157 页)

- Wang NJ, Yao WJ, Wei YH. Biochemical analysis of urine for diagnosis of kidney disease[J]. Lab Med Clin, 2010, 7(6): 519-521.
- [2] 李熙建, 谭同均, 朱国文, 等. 胱抑素 C 与尿微量蛋白及尿酶联合检测在不同程度肾病中的应用价值[J]. 现代检验医学杂志, 2008, 23(6): 70-73.
- Li XJ, Tan TJ, Zhu GW, et al. Application of serum cystatin C, urinary microprotein and urinary enzyme in the degree of nephropathy[J]. J Mod Lab Med, 2008, 23(6): 70-73.
- [3] 赵海舰, 张传宝, 王 薇, 等. 脂类检验项目室内质控变异系数分析[J]. 中华检验医学杂志, 2012, 35(12): 1172-1175.
- Zhao HJ, Zhang CB, Wang W, et al. Analysis of the co-efficient of variation of internal quality control of lipid testing[J]. Clin J Lab Med, 2012, 35(12): 1172-1175.
- [4] 刘晓春, 周微雅, 何 毅, 等. 广西三级医院临床化学互认检验项目室内质控变异系数调查与分析[J]. 现代检验医学杂志, 2013, 28(5): 147-150, 154.
- Liu XC, Zhou WY, He Y, et al. Investigation and analysis of clinical chemistry internal quality control coefficient of variation of the mutual recognition of test items in Guangxi tertiary hospital laboratories[J]. J Mod Lab Med, 2013, 28(5): 147-150, 154.
- [5] 白 玉, 王治国, 王 薇, 等. 全国常规化学检验项目室内质控变异系数的分析[J]. 检验医学, 2011, 26(3): 207-209.
- Bai Y, Wang ZG, Wang W, et al. Analysis of coefficient of variation of internal quality control of routine chemistry items[J]. Lab Med, 2011, 26(3): 207-209.
- [6] Skiek M. Acceptability limits based on biology goals in haematology EQAS[J]. Accred Qual Assur, 2005, 10(3): 112-115.
- [7] Klee GG. Establishment of outcome-related analytic performance goals[J]. Clin Chem, 2010, 56(5): 714-722.
- [8] Fraser CG. General strategies to set quality specifications for reliability performance characteristics[J]. Scand J Clin Lab Invest, 1999, 59(7): 487-490.
- [9] 王治国. 临床检验质量控制技术[M]. 3版. 北京: 人民卫生出版社, 2014: 124.
- Wang ZG. Quality control technology in clinical laboratory[M]. 3th Ed. Beijing: People's Medical Publishing House, 2014: 124.

收稿日期: 2015-12-10

修回日期: 2016-09-10

(上接 153 页) 被评估的患者样本数三者相结合, 由此改变质控策略来减少 $E(N_{uf})$ 值; 也可以通过减少分析系统不精密度或增加 TE_a 规范(任选一个)来减少其值。

最大的 $E(N_{uc})$ 值依赖于质控活动之间被测的患者样本平均数。如果最大的 $E(N_{uc})$ 值太高, 就要降低质控活动之间被评估的患者样本数。

参考文献:

- [1] Westgard JO. How to establish control limits. In: Assuring the right quality right[C]. Madison (WI): Westgard QC, Inc, 2007: 245-259.
- [2] Brooks ZC. Choosing our own rules. In: Performance-driven quality control[C]. Washington, DC: AACC Press, 2001: 89-114.
- [3] Parvin CA. Quality-control (QC) performance measures and the QC planning process[J]. Clin Chem, 1997, 43(4): 602-607.
- [4] Parvin CA. Statistical topics in the laboratory sciences[J]. Methods in Molecular Biology, 2007(404): 353-375.
- [5] Kenny D, Fraser CG, Hytloft Petersen P, et al. Consensus agreement[J]. Scand J Clin Lab Invest, 1999, 59(7): 585.
- [6] Fraser CG. Quality specifications. In: Biological variation: from principles to practice[C]. Washington, DC: AACC Press, 2001: 29-66.
- [7] Westgard JO. How to use sigma-metrics QC selection tool. In: Assuring the right quality right[J]. Madison (WI): Westgard QC, Inc, 2007: 125-146.
- [8] Parvin CA. Assessing the impact of the frequency of quality control testing on the quality of reported patient results[J]. Clin Chem, 2008, 54(12): 2049-2054.
- [9] Yundt-Pacheco JC, Parvin CA. The impact of QC frequency on patient results[J]. MLO Med Lab Obs, 2008, 40(9): 24, 26-27.
- [10] Cembrowski GS, Carey RN. Quality control procedures. In: Laboratory quality management[J]. Chicago: ASCP Press, 1989: 59-79.
- [11] Parvin CA. Comparing the power of quality-control rules to detect persistent systematic error[J]. Clin Chem, 1992, 38(3): 358-363.
- [12] Parvin CA. Comparing the power of quality-control rules to detect persistent increases in random error[J]. Clin Chem, 1992, 38(3): 364-369.
- [13] Parvin CA, Kuchipudi LS, Yundt-Pacheco JC. Should I repeat my 1:2s QC rejection? [J]. Clin Chem, 2012, 58(5): 925-929.
- [14] Parvin CA. New insight into the comparative power of quality-control rules that use control observations within a single analytical run[J]. Clin Chem, 1993, 39(3): 440-447.

收稿日期: 2016-01-04

修回日期: 2016-05-28