2003, 38(3): 207-208.
[5] 刘风华, 王李洁. 血标本放置方法及时间对血清葡萄糖浓度检测结果的影响[J]. 检验医学与临床, 2010, 7(14): 1501-1502.
[6] 王晋, 吴夏枫, 朱旋. 血标本放置时间对血糖检验结果的影响[J]. 中华全科医师杂志, 2007, 6(6): 376-377.

(收稿日期: 2010-12-23)

·质控与标规·

血液分析仪的方法确认和性能验证

胡丽涛[1], 王　薇[2], 王治国[2△]
(1. 中国医学科学院/北京协和医学院研究生院　100730; 2. 卫生部北京医院/卫生部临床检验中心　100730)

摘　要: 目的　阐述血液分析仪的方法确认和性能验证。**方法**　主要参考美国临床和实验室标准化研究所(CLSI) H26-A 文件相关内容, 并结合具体实际进行总结归纳。**结果**　血液分析仪的方法确认内容包括精密度、线性、空白限、检测下限、定量下限、携带污染率等; 血液分析仪的性能验证与方法确认有相似之处但又有别于确认。**结论**　对血液分析仪进行方法确认和性能验证是检测结果准确可比的重要保证。

关键词: 方法;　血液分析仪;　方法确认;　携带污染;　检出下限;　验证
DOI: 10.3969/j.issn.1673-4130.2011.13.045　**文献标识码:** A　**文章编号:** 1673-4130(2011)13-1497-03

美国食品和药品监督管理局(FDA)规定: 确认是通过提供客观证据对预期用途或应用要求已得到满足的认定。血液分析仪确认的目的有: 从医学角度评估安全性和临床效力; 确定性能信息; 验证系统操作性能特征的适用度; 获得监管机构的审批。血液分析仪确认具体评价以下的性能特征: 空白限(本底, LoB)、检出下限(LLoD)和定量限(LoQ)、携带污染率、精密度、分析测量区间(AMI)、可比性等。本文参考了美国临床和实验室标准化研究所(CLSI) H26-A[1] 文件及有关文献[2], 具体内容如下。

1　方法确认研究(厂商的工作)

1.1　方法确认准备工作

1.1.1　标本的选择　所有用于确认试验的血标本都应该是临床检测之后剩余的, 否则涉及到知情同意。对用于性能评估的样本量没有明确要求, 但是进行综合性能评估需要在分析测量区间(AMI)内各浓度水平大量取样, 且充足的标本量更有能力区分个别的离群值和不同方法间的显著差异。CLSI EP09 提到: "足量的标本量可以提高方法学比较统计估计的可信度, 增加包括干扰物质效应的机会, 使得研究更全面"。厂商的性能声明应该建立在实际的标本检测上, 不应该有超出观测值的推断。每个试验中心的试管应该统一, 不能既用 EDTA-K_2 又用 EDTA-K_3 试管, 国际血液学标准化委员会(ICSH)推荐使用 EDTA-K_2 试管, 使用采集后 8 h 内的标本以减少标本老化对被测物的影响。应该排除溶血、凝血标本、血量少于 1 mL 和采集后超过 8 h 的标本。

1.2　医学允许误差　CLSI EP21 提到: "临床医生作为检测数据的最初使用者, 考虑的是总的分析误差, 而不是随机误差或系统误差。很久以来我们只估计随机误差和系统误差而没有把两者结合起来。"换句话说, 只是使用了不精密度和偏倚来表示仪器的性能, 没有计算总的分析误差 TE, 这样是不完整的。实验室检测的医学目标如下: 检测结果准确反映患者的状况; 检测结果的变异反映的是患者的变化而不是仪器的变异(偏倚或不精密度); 偏倚不会导致患者诊断结果的改变; 同一患者的系列标本的改变能够与生物变异区分开来。分析目标只与患者诊疗有关。表 1 是基于医学要求的性能目标。

1.3　仪器校准和质量控制　方法确认的主要目的是比较试验的自动分析仪(TAA)和比对的自动分析仪(CAA)的患者结果。因为血液学商业物质的基质效应已知, 所以可用新鲜的全血标本来完成校准。这是证实校准稳定性的主要机制。用可溯源的校准物按厂商的说明书对 CAA 进行校准, 然后用 CAA 和正常的新鲜全血标本为 TAA 校准, 这样能够确保新鲜血标本结果交叉检测平台的可比性。分别用商业产品进行校准可能达不到目的。确认研究中, 每天都应该用全血标本交叉核对(至少 10 份正常标本)进行校准, 去掉不合理的结果。此外, 在进行血液分析仪确认前必须做好质量控制。血液分析仪的室内质量控制除了传统的商品质控物外还可以使用患者结果, 形态学是血液学检测的特点之一, 手工显微镜镜检也是血液分析仪质量控制的一部分。

表 1　基于医学要求的性能目标

项目	最大偏倚(%)	最大不精密度(%)	最大总误差(%)
WBC	9.8	8.0	12.7
RBC	3.4	2.5	4.2
Hb	3.1	2.0	3.7
HCT	2.8	2.2	3.6
PLT	14.7	9.9	17.7

2　方法确认研究的内容

2.1　LoB、LLoD 和定量下限(LLoQ)　血液学中的 LoB 通常被称为"本底", 本底是由于试剂或电子噪音所致, 表现为检测出假性的标本成分。自动血液分析仪确认的核心问题是准确定量极低浓度的 WBC 和 PLT, 这个极低值必须与仪器的本底区分开来。LLoD 指在一定概率下标本可被检测出来的最低浓度, 在血液学中, 指可与本底区分开来的最低的血细胞浓度值。LLoQ 指标本中能够准确定量的最低浓度值, 且定量结果在可接受的精密度和准确度范围内, 即满足分析性能目标的最低的 WBC 和 PLT 浓度。图 1 展示了检测结果的分布及报告

△　通讯作者, E-mail: zhiguo_w@hotmail.com。

形式。

准确定量低浓度 WBC 在某些临床情况下非常重要，包括决定是否化疗和判断骨髓移植恢复信息；准确定量低值 PLT 对预测出血和决定是否输注 PLT 非常重要。LLoD 和 LLoQ 的确定可根据 CLSI EP17 文件。LLoD 值是本底的均值加上一个常数倍 s，若本底值服从正态分布，则这个常数为 1.64；若不服从正态分布则常数为 $1.654/(1-1/[4(n-k)]$，n 为总的重复检测次数，k 为标本个数。CLSI EP17 推荐使用不同浓度的样本至少测定 60 次。可以选用 6 个不同浓度的样本每个测定 10 次，以发现样本与样本之间的变异。其他被测物如 RBC、Hb 和 HCT 则不需要确定 LLoD 和 LLoQ，因为这些项目接近 0 时人已经不能存活了。

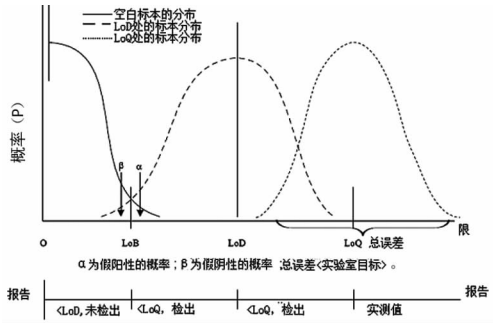

图 1　检测结果分布图及报告形式

2.2　携带污染率　携带污染率是指一个标本对接下来标本的影响百分率。对于血液学仪器，如果前一标本的浓度很高，则携带污染通常对后续标本产生正向偏倚，表现后续结果的假性升高；反之则使后续标本的值假性降低。临床上在检测患者标本和质控标本前分别预吸入全血标本或质控标本来解决携带污染的问题。携带污染研究必须至少在 3 台试验的自动分析仪上进行至少 3 次分析。先检测 1 份高浓度(HTV)新鲜全血标本，然后检测低浓度(LTV)的全血标本，重复 3 次(如表 2)。不能用质控物或吸入空气代替，用于携带污染研究的被测物及浓度可参照表 2。

表 2　用于携带污染研究的被测物及浓度*

检测项目	高浓度	低浓度(不包括边界值)
WBC(10^9/L)	>90	0～3
RBC(10^{12}/L)	>6.20	0～1.5
Hb(g/L)	220	0～50
PLT(10^9/L)	>900	0～30

*：假定这些浓度在分析测量范围内，若厂家声明的分析测量范围更窄则需要适当调整。

这样做的目的是要较大地拉开高浓度和低浓度分析物的浓度差距。从健康的献血者采血后需要人为地升高和降低血标本中被测物的浓度，使之符合表中所列的浓度要求。但是高浓度和低浓度标本都应该含有一定量的被测物，不能是没有血细胞的血浆。低浓度标本不用简单的空气或稀释液组成，少量的全血可以加入较大体积的没有血细胞的血浆中以提供适当的基质，基质中的细胞浓度不能被检测出来。由于质控物中细胞的组成与全血不同，存在基质效应，没有上表中要求的人体内真正高浓度和低浓度的被测物，因此质控物质不能代替新鲜的全血标本。每个被测物可计算 95%、97.5% 和 99% 的置信区间。

2.3　不精密度(重复性和重现性)　重复性是指一系列重复检测结果的标准差 *s* 和变异系数 *CV*；重现性包括了所有来源的变异。在校准不变的情况下 *CV* 值的确定可以得出单个检测的预期变异，例如 WBC、RBC 和 PLT 在低浓度水平的不精密度明显升高，因为在固定的稀释和计数模式下血液分析仪检测到的细胞要少。短期不精密度和长期不精密度对证明系统稳定性和设置校准点都是有用的。

正常标本不精密度研究用的是其值在实验室参考区间内的新鲜全血。将同一份血液标本连续吸取 30 份；样本量太少得出的置信区间太宽。至少要在 TAA 上做 4 份不精密度研究，共 12 个批次。

ICSH 建议不精密度的研究应该包括病理范围浓度的标本。由于贫血和出血的人可能涉及到输血的问题，因此不精密度研究还应该包括下面 2 种病理标本：Hb 为(60～100)g/L 和 PLT 为$(0\sim50)\times10^9$/L。此外严重白血病的患者还涉及到治疗，不精密度研究还应该包括 WBC 为$(0\sim2)\times10^9$/L 浓度的标本。至少应该在 3 台 TAA 上做 3 个不精密度研究(31 次重复测定)，总共至少检测 9 个批次。贫血标本只要分析 RBC、Hb 和 HCT，其他不属于这些被测物的可疑或无效数据可以忽略；PLT 减少的标本只检测 PLT，无关的无效或可疑数据可以忽略。同样，白血病标本只需检测 WBC 浓度和分化。

表 3　方法比较的目标范围

项目	范围	分布(%)
WBC($\times10^9$/L)	<2.0	10
	2.0～5.0	10
	5.1～11.0	45
	11.1～50.0	25
	>50.1	10
RBC($\times10^{12}$/L)	<3.00	5
	3.00～4.00	15
	4.01～5.00	55
	5.01～6.00	20
	>6.01	5
Hb($\times$g/L)	<100	10
	100～120	15
	121～160	60
	161～180	10
	>181	5
PLT(10^9/L)	<40	10
	40～125	20
	126～300	40
	301～500	20
	500～600	5
	>601	5

2.4　可比性(相关)　可比性研究需要将系统线性范围内的结果与适当的参考方法或对比方法进行比较，研究人群应包括仪器的目标人群。已上市仪器可以作为新血液分析仪的比较方

法。方法比较的目标范围如表 3，表中的数据基本上与 ICSH 和 CLSI EP09 推荐的目标范围一致，覆盖了正常参考区间、低于和高于正常参考区间的范围。对于确定分析结果范围，参考方法可以帮助确保仪器的准确性。超过临床重要范围的准确度评价都要选用替代的参考方法，表 4 列出了推荐的参考区间。

表 4　推荐的参考方法及使用的标准

项目	小于	大于	参考方法
Hb(g/L)	—	—	根据相关算法计算、氰化高铁血红蛋白法
HCT(%)	30	60	微量离心法*
MCV(fL)	75	115	红细胞比容/红细胞计数、显微镜法#
PLT(10^9/L)	25	n/a	单克隆抗体、相差显微镜
Ret(%)	0.001	5.0	流式细胞术

*：对于贫血或正常的 HCT 离心时间是 5 min，对于较高的 HCT 需要离心更长的时间；#：当 MCV>115 fl 时要用瑞氏染色涂片显微镜检判断是否是由于 RBC 聚集或叠连造成的假性升高；—：无数据。

2.5　分析测量区间(线性)　线性的概念不适合用于血液学细胞计数(如 RBC、WBC 和 PLT)，只有血红蛋白浓度与比色仪的吸光度呈一定关系。对于粒子计数测量(WBC、RBC 和 PLT)和相关测量(Hb、HCT)回收是一个更恰当的分析术语，通过对已知高浓度的标本进行一系列稀释建立分析测量范围，可作为临床可报告范围(CRI)的子集。

建立分析测量范围用于说明书中的声明的方法有 2 种：一种是对试验的自动分析仪和可比较的自动分析仪进行新鲜血液的相关分析，这种方法最具有临床相关性；另一种方法是至少从 2 个不同的厂家购买商品的线性试剂盒并按照厂家的说明进行评估。这些数据得出的可分析测量范围是基于商品试剂盒材料，并不代表使用临床新鲜血液标本检测的性能。本实验至少要用不小于 3 台的试验的自动分析仪。

3　性能验证(实验室的工作)

ISO 提出：仪器应该满足所需的性能要求并与检测相关的规定一致。一般用户实验室应该按照厂家确认的原则和程序进行验证。目的是要验证厂家所述的性能对于实验室具体的仪器是正确的。验证与确认有区别也有联系，具体内容包括：准确度、精密度、检测结果的可报告范围和正常参考区间。

3.1　LoB、LLoD 和 LLoQ　LoB 的验证起始程序与确认部分程序相同。一旦通过验证，每天都要进行本底核查以保证反应试剂微粒和电极噪音没有变化，如果出现问题需要重新研究 LoB。LLoD 和 LLoQ 的验证同确认部分。

3.2　携带污染　携带污染的验证起始程序与确认部分程序相同。当检测血细胞减少的患者标本，应保证其前面检测的 WBC、RBC 和 PLT 的浓度不超过患者标本的浓度值，因为即使是很小的携带污染都可能假性提高血细胞减少患者标本的血细胞浓度。反之，在患者标本或质控血标本之前吸入了不含血细胞的液体，由于其预稀释作用使得血细胞的浓度值假性下降。

3.3　不精密度(重现性)　用户实验室验证应该强调在医学决定水平的短期和长期的重现性。适宜使用实用的方法和所有实验室都可获得的标本。短期不精密度：正常标本、医学决定水平；长期不精密度：重复性和仪器。

3.4　分析测量区间　见确认部分。

3.5　可比性(相关)　根据具体实验室的患者人群对厂家的数据进行修正是很重要的。比如，如果实验室很多标本都是来自早产儿，则有比必要对有核红细胞正确定量以及对 WBC 浓度进行修正。如果实验室要换一台新的仪器，将其与实验室现有的仪器做比对是很重要的，以保证患者结果不受仪器更换的影响。厂家的确认研究可能不包括实验室现有仪器的模式。

3.6　参考区间　具体检测方法的参考区间根据目标人群个体内和个体间的生物学变异，样本的数量和将分成亚组非常重要。分析前的各方面应该严格控制，以保证从一般人群中获得标本的重现性。分析过程的标准化可使不同人群的数据具有可比性，并且最终的结果都能用于不同地方、不同时间的人群。参考限的计算方法可以影响获得的结果，如使用不适当的统计方法或不正确地排除了离群值。欧洲关于医疗体外诊断装置的指导说明 98/79(5.5.5 条款)提出：参考区间应该周期性修订，特别是每一次改变了分析或分析前程序。认可管理组织也要求实验室评价其参考区间的适合性。

3.7　CRI　为满足临床需要，建立临床可报告范围是实验室根据检测技术做出的一个临床判断。如果实验室验证的 AMI 和 CRI 都窄，则没有必要采取进一步行动。如果患者的结果超出了 AMI 则需要根据具体实验室进行确认。CRI 一般在方法验证开始建立，只有在方法学改变时才需要重新评价。例如报告的结果超出了 AMI 的上限，则需要对标本进行稀释后重新检测，用稀释倍数来计算最后的检测结果。厂家一般规定了超出 AMI 的标本的稀释和浓缩程序。方法确认期间验证了 CRI 的低限一般是 AMI 的低限。低于 CRI 下限的值一般报告为小于低限。

参考文献

[1] Clinical and Laboratory Standards Institute. H26-A2. Validation, verification and quality qssurance of automated hematology analyzers; approved standard-second edition[S]. Wayne, PA: CLSI, 2010.

[2] 王治国. 临床检验方法确认与性能验证[M]. 北京：人民卫生出版社，2009.

(收稿日期：2010-12-17)

参数与统计量

描述总体特征的数值为参数，通常是未知的，一般用希腊字母表示，如 μ、σ、π 等。描述样本特征的数值为统计量，是已知的或可计算获得的，用英文字母表述，如 S、P 等。从总体中随机抽样可获得样本，以样本为基础、通过统计推断(参数估计、假设检验)可获得对总体的认识。