

• 临床实验室质量管理 •

DOI: 10.13602/j.cnki.jcls.2015.08.17

检测结果一致性评价方法的选择

李云飞, 马建锋, 陈晓婷(南京医科大学第一附属医院检验学部, 南京 210029)

摘要: 目的 比较不同评价方法对检测结果一致性评价的适用性。方法 选择多份临床样本制备的混合血清 20 份作为浓度集中组; 另选择覆盖检测系统不同浓度的样本 20 例作为系列浓度组。2 组分 2 批次测定胱抑素 C, 同一组 2 批结果分别用配对 t 检验、简单相关性分析、比对及偏倚评估和 Bland-Altman 法 4 种方法进行分析, 比较 4 种方法对检验结果一致性评价的优劣。结果 对于浓度集中组的 2 批检测结果间的比较: 配对 t 检验分析 2 批数据差异有统计学意义 ($P < 0.05$); 简单相关性分析 2 组数据有相关性 ($r = 0.58$, $P < 0.05$); 用患者样本进行方法学比对及偏倚评估法分析, 其拟合回归方程 $Y = 0.535X + 0.496$, 在 0.63、1.15 mg/L (依据参考区间: 0.63 ~ 1.15 mg/L, 下同) 处偏倚分别为 32.2% 和 3.4%, 差异有统计学意义 ($P < 0.05$); 用 Bland-Altman 法分析作差值 - 均值散点图, 界限的最大允许误差为 7%; 作比值 - 均值散点图, 界限的最大允许误差为 6.4%, 均表明 2 批次结果一致性良好。对于系列浓度的 2 批检测结果间的比较: 配对 t 检验分析 2 组数据间差异无统计学意义 ($P > 0.05$); 用简单相关性分析 $r = 0.999$, 对相关系数进行 t 检验 $P < 0.05$; 用患者样本进行方法学比对及偏倚评估法分析, 拟合回归方程 $Y = 1.015X - 0.035$, 在 0.63、1.15 mg/L 处偏倚分别为 4.01% 和 1.54%, 差异无统计学意义 ($P > 0.05$); 用 Bland-Altman 法分析作差值 - 均值散点图, 界限的最大允许误差为 8%; 作比值 - 均值散点图, 界限的最大允许误差为 5.6%, 表明 2 批次结果一致性良好。结论 浓度集中组建议用 Bland-Altman 法分析其差异可接受性; 浓度分布较宽的 2 组数据可以用配对 t 检验、患者样本进行方法学比对及偏倚评估和 Bland-Altman 法来分析, 推荐后两种方法, 简单相关性分析不宜单独分析。

关键词: 一致性; 配对 t 检验; 简单相关系数分析; 方法学比对及偏倚评估; Bland-Altman 法

中图分类号: R446.1

文献标志码: A

国家临床实验室质量管理规范要求检验项目在用于临床检测前需评价其检测性能, 且在不同检测系统上检测的相同项目要进行一致性评价, 以确定同一临床样本在不同仪器上能得到一致的结果。常用的方法有配对 t 检验、简单相关系数分析、比对及偏倚评估和 Bland-Altman 法。用不同的统计学方法对同一组数据分析, 可能得到不同的结果。本研究以胱抑素 C 测定为例, 观察在不同情况下、不同统计学方法的分析结果, 以期选择最为合适的统计学方法。

1 材料和方法

1.1 研究对象 收集南京医科大学第一附属医院检验学部 2013 年 2 月 18 日临床样本检测后的血清样本。一组为浓度集中组, 由多份样本制备的混合血清, 分成 20 份; 另一组为系列浓度组, 选择覆盖检测系统不同浓度的血清样本 20 份。

1.2 仪器和试剂 AU5400 全自动生化分析仪(贝克曼库尔特公司); 胱抑素 C 试剂(批号: 211122K, 校正品批号: 212191L, 北京利德曼公司)。

1.3 方法 2 组样本分别进行 2 个批次测定, 以第 1 次测定为目标结果。比较 2 批次结果的一致性。

1.4 统计学分析 采用 SPSS 13.0 统计软件进行。数据以 $\bar{x} \pm s$ 表示, 计量资料用配对 t 检验分析和相关性分析, 以 $P < 0.05$ 为有统计学意义; 用患者样本进行方法学比对及偏倚评估时先求出拟合的回归方程, 计算在有临床意义浓度处的结果误差是否超过国家实验室管理规范或临床可接受误差; 用 Bland-Altman 法计算出 2 批测量一致性的界限, 观察 2 次结果差值或比值与 2 批结果均值作散点图, 结合临床实际, 判断 2 批测量方法结果是否一致。

2 结果

2.1 浓度集中组 4 种方法分析结果 浓度集中组 20 份样本分 2 个批次测定, 结果见表 1。

2.1.1 配对 t 检验 经配对 t 检验分析 2 组数据差异有统计学意义 ($t = 5.064$, $P < 0.05$), 表明 2 个批次测定结果不一致。

2.1.2 相关性分析 其相关系数 (r) = 0.58, 进行 t 检验 $t = 3.02$, $P < 0.05$, 说明 2 组数据间存在相关性。

2.1.3 患者样本进行方法学比对及偏倚评估 浓度集中组 2 批测定结果散点图, 以第 1 次测定结果为横坐标 (X), 第 2 次测定结果为纵坐标 (Y), 其拟

作者简介: 李云飞, 1976 年生, 女, 主管技师, 大学本科, 主要从事临床生化检验。

通信作者: 陈晓婷, 主任技师, E-mail: 13851522791@163.com。

合回归方程为 $Y = 0.535X + 0.496$ 。该项目参考区间为 0.63 ~ 1.15 mg/L 在 $X = 0.63$ 和 1.15 mg/L 处的偏倚分别为 32.2% 和 3.4% ,不能满足临床要求。见图 1。

表 1 浓度集中组胱抑素 C 的测定结果

编号	第 1 批结果 (mg/L)	第 2 批结果 (mg/L)	差值 (mg/L)	比值	均值 (mg/L)
1	1.100	1.090	0.010	1.009	1.095
2	1.060	1.040	0.020	1.019	1.050
3	1.110	1.070	0.040	1.037	1.090
4	1.080	1.090	-0.01	0.991	1.085
5	1.130	1.090	0.040	1.037	1.110
6	1.130	1.090	0.040	1.037	1.110
7	1.130	1.080	0.050	1.046	1.105
8	1.110	1.080	0.030	1.028	1.095
9	1.140	1.100	0.040	1.036	1.120
10	1.100	1.060	0.040	1.038	1.080
11	1.140	1.100	0.040	1.036	1.120
12	1.110	1.110	0.000	1.000	1.110
13	1.150	1.110	0.040	1.036	1.130
14	1.110	1.090	0.020	1.018	1.100
15	1.150	1.110	0.040	1.036	1.130
16	1.130	1.110	0.020	1.018	1.120
17	1.160	1.120	0.040	1.036	1.140
18	1.120	1.140	-0.020	0.982	1.130
19	1.100	1.120	-0.020	0.982	1.110
20	1.130	1.100	0.030	1.027	1.115
\bar{x}	1.120	1.095			
s	0.025	0.023			

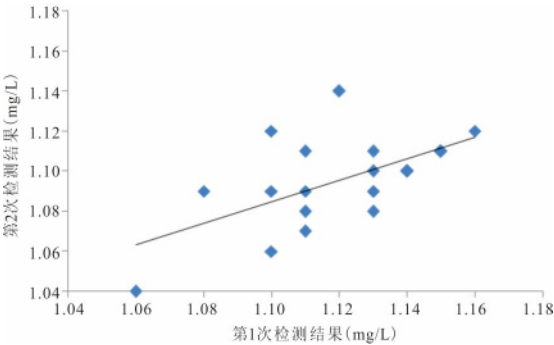


图 1 浓度集中组 2 批测定结果散点图

2.1.4 Bland-Altman 法分析 计算各样本 2 批结果差值,以 $\bar{x} \pm 2s$ 为允许范围,作差值 - 均值散点图,结果表明,所有点均在允许范围内,界限的最大误差为 0.070 5,即最大允许误差 7%,说明 2 批次结果一致性好,见图 2。计算 2 次结果比值并作比值 - 均值散点图,结果表明,界限的最大允许误差为 6.4%,发现有 2 个点位于界限范围外,超过 95% 数据在规定误差范围内,位于次外围点的误差有 2.8%,为临床可以接受的误差,说明 2 批次结果一致性良好,见图 3。

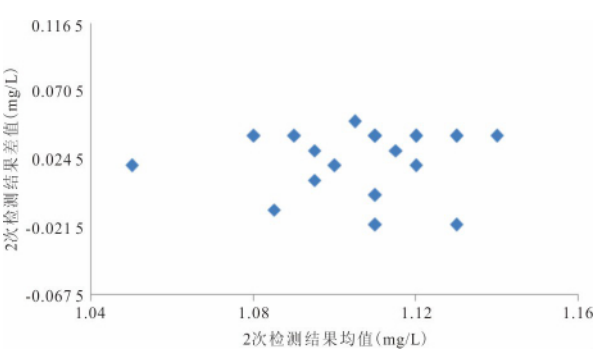


图 2 Bland-Altman 法分析浓度集中组差值 - 均值散点图

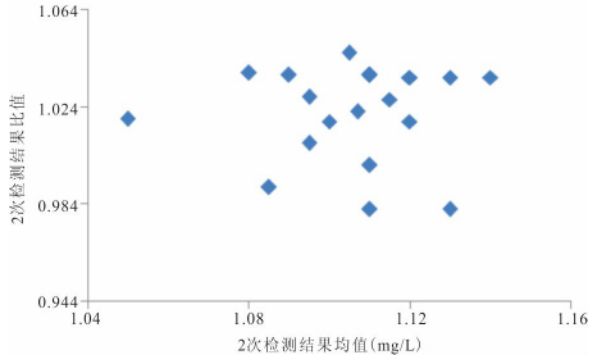


图 3 Bland-Altman 法分析浓度集中组比值 - 均值散点图

2.2 系列浓度组 选择覆盖检测系统不同浓度的样本 20 例,2 个批次测定结果见表 2。

表 2 系列浓度组胱抑素 C 测定结果

编号	第 1 批结果 (mg/L)	第 2 批结果 (mg/L)	差值 (mg/L)	比值	均值 (mg/L)
1	0.380	0.370	0.010	1.027	0.375
2	0.420	0.400	0.020	1.050	0.410
3	0.530	0.540	-0.010	0.981	0.535
4	0.520	0.490	0.030	1.061	0.505
5	0.905	0.885	0.020	1.023	0.895
6	0.710	0.725	-0.015	0.979	0.718
7	0.835	0.820	0.015	1.018	0.828
8	0.730	0.705	0.025	1.035	0.718
9	0.940	0.920	0.020	1.022	0.930
10	0.865	0.860	0.005	1.006	0.863
11	1.280	1.260	0.020	1.016	1.270
12	1.700	1.705	-0.005	0.997	1.703
13	1.560	1.525	0.035	1.023	1.543
14	1.990	1.960	0.030	1.015	1.975
15	2.295	2.290	0.005	1.002	2.293
16	2.655	2.635	0.020	1.008	2.645
17	2.210	2.165	0.045	1.021	2.188
18	2.695	2.675	0.020	1.007	2.685
19	3.940	3.935	0.005	1.001	3.938
20	9.080	9.215	-0.135	0.985	9.148
\bar{x}	1.812	1.804			
s	1.957	1.985			

2.2.1 配对 t 检验 经配对 t 检验分析,2 组数据差异无统计学意义($t = 0.967$, $P > 0.05$),说明 2 个

批次测定结果可替代。

2.2.2 相关性分析 其 $r=0.999$, 进行 t 检验 $t=94.80$, $P<0.05$, 说明 2 组数据存在相关性。

2.2.3 患者样本进行方法学比对及偏倚评估 以第 1 次测定结果为横坐标 (X), 第 2 次测定结果为纵坐标 (Y) 作散点图, 其拟合回归方程为 $Y=1.015X-0.035$ 。该项目参考区间为 $0.63 \sim 1.15 \text{ mg/L}$, 在 X 分别为 0.63 和 1.15 mg/L 处的偏倚分别为 4.01% 和 1.54% , 能满足临床要求。见图 4。

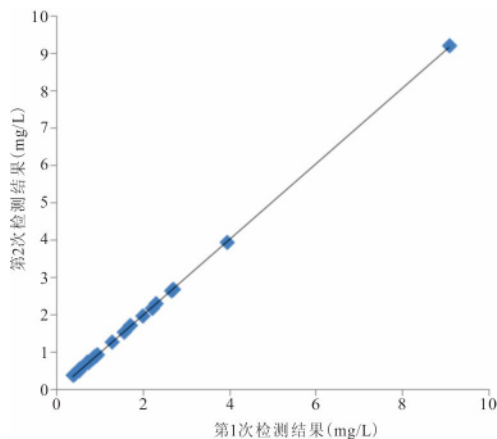


图4 系列浓度组 2 批测定结果散点图

2.2.4 Bland-Altman 法分析 计算系列浓度各样本 2 批结果的差值, 以 $\bar{x} \pm 2s$ 为允许范围, 作差值 - 均值散点图, 结果只有 1 个点在允许范围外, 即 95% 的数据在允许范围内, 界限的最大允许误差为 8% , 说明 2 批次结果一致性良好, 见图 5。计算 2 次结果比值, 以 $\bar{x} \pm 2s$ 为允许范围, 作比值 - 均值散点图, 同样只有 1 个点位于界限范围外, 符合 95% 数据在范围内的规定, 其界限的最大允许误差为 5.6% , 说明 2 批次结果一致性好, 见图 6。

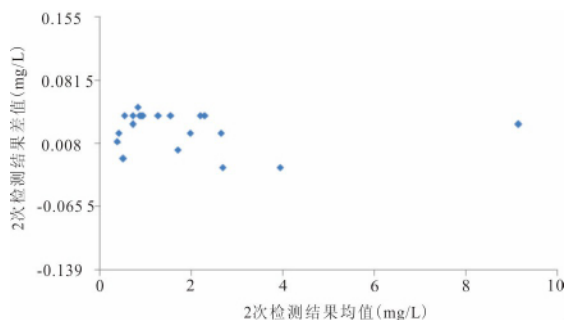


图5 Bland-Altman 法分析系列浓度组差值 - 均值散点图

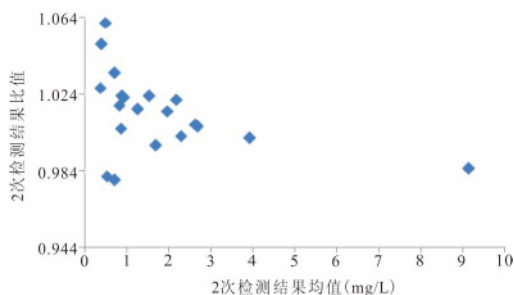


图6 Bland-Altman 法分析系列浓度组比值均值散点图

3 讨论

分析性能的准确性一般通过分析与国家/国际参考品的偏差、回收试验、方法学比对等方法来实现。然而, 对于许多项目来说, 并没有国家或国际的参考品, 回收试验也需要明确测量值的标准物质, 有时还要冒基质效应的风险, 因此临床上常通过与参考检测系统测量结果的一致性验证, 来证明 2 种检测系统检测结果是否等价或可交换, 从而评价检测系统与参考检测系统检测结果的差异。

本研究中用配对 t 检验比对 2 组数据差值所代表的总体均数与 0 之间的差异是否有统计学意义, 发现浓度集中组中 2 组结果不一致, 但 2 批次间最大的误差不到 5% , 平均误差为 0.8% 。对免疫比浊法检测胱抑素 C 来说, 其误差已非常小, 说明统计结果与临床实际并不相符, 分析产生偏差原因, 发现该方法的批内精密度较高, 其 CV 值在 1% 左右, 且可能与配对 t 检验的值与差值的均数、差值标准差和标本量有关。系列浓度组分析结果表明, 高浓度下的误差绝对值大, 增大了差值的标准差, 结果间差异无统计学意义。如样本浓度差异越大, 则其 t 值会越小。

简单的相关性分析对系统误差的识别能力较差, 因为其可以人为地将 1 组数据按比例升高或下降, 而相关系数不变, 因此无法检出结果间的差异, 故不推荐单独用该方法检查 2 组结果的差异。用患者样本进行方法学比对及偏倚评估是一种较为科学的评价方法。本研究中浓度集中组的结果显示不等价(参考区间下限存在约 32% 的误差, 超过临床质量允许要求), 原因可能为我们选择的浓度远离该点, 此时与评估浓度相近的参考区间上限误差仅为 3.4% , 而在系列浓度组分析中结果显示可等价, 这是因为实验浓度范围包含重要医学水平值的缘故。

Bland-Altman 法分析可以任意选择 2 批结果的差值或比值来进行统计, 先求出允许的分布宽度, 再相对于均值作散点图, 以 95% 的数据在范围内作为

一致性的判断基础 ,可以同时检出可能存在的系统误差或随机误差。本研究中 2 组数据的判断与实际均相符。对该分析方法研究后结果表明 ,应用 2 次测量结果的均值作为 X 轴 ,以差值作为 Y 轴分析更为合理。然而 ,无论用哪种检测结果作为 X 轴 ,均可能观察到错误的测量间差值的散布趋势 ,从而掩盖测量间差值真实的散布趋势。

综上所述 ,配对 t 检验对随机误差检出不敏感 ,浓度集中组数据间差异有统计学意义的结果存在假阳性 ,浓度分布宽的数据间可以用配对 t 检验 ,但也不能兼顾检出随机误差 ,不建议单独使用 ;简单相关系数分析不能检出系统误差 ,不建议单独使用 ;用患者样本进行方法学比对及偏倚评估适用于系列浓度组数据分析 ,特别关注了临床上有特殊意义点的正确性 ,可作为建议的统计方法 ,但其不适用于浓度集中的数据间显著性分析 ;Bland-Altman 法分析适合不同情况下的显著性检验 ,同时兼顾了系统误差和

随机误差的检出 ,是值得推荐的统计方法 ,但在使用时应注意 X 轴须用 2 种方法测量结果均值 ,而不是任何一种方法的结果。

4 参考文献

- [1]陈卉. Bland-Altman 分析在临床测量方法一致性评价中的应用[J]. 中国卫生统计 2007 24(3):308-315.
- [2]李云飞,马建峰. 降钙素原定量检测系统的性能验证[J]. 临床检验杂志 2014 32(11):876-878.
- [3]王杨,陈涛,徐涛,等. Bland-Altman 分析结果散点分布趋势与测量方法变异程度间的关联度分析[J]. 中华疾病控制杂志, 2012 16(6):535-538.
- [4]梅金平,周欣,姬文婕,等. 流式细胞术与光学比浊法检测血小板聚集率的比较[J]. 临床检验杂志 2015 33(2):81-84.

(收稿日期:2015-06-01)

(本文编辑:许晓蒙)

• 读者 • 作者 • 编者 •

《临床检验杂志》可直接使用缩略形式的常用词汇

对于以下医学检验工作者比较熟悉的常用词汇 ,本刊允许在论文撰写中直接使用其缩略语 ,可以不标注中文。

磷酸盐缓冲液(PBS)	白细胞介素(IL)	乙型肝炎表面抗原(HBsAg)
核糖核酸(RNA)	肿瘤坏死因子(TNF)	乙型肝炎 e 抗原(HBeAg)
脱氧核糖核酸(DNA)	干扰素(IFN)	抗 HBsAg 抗体(抗 HBs)
聚合酶链反应(PCR)	人类白细胞抗原(HLA)	抗 HBeAg 抗体(抗 HBe)
酶联免疫吸附试验(ELISA)	系统性红斑狼疮(SLE)	抗 HBcAg 抗体(抗 HBc)
免疫球蛋白 G(IgG)	类风湿关节炎(RA)	严重急性呼吸综合征(SARS)
免疫球蛋白 A(IgA)	人类免疫缺陷病毒(HIV)	红细胞(RBC)
免疫球蛋白 M(IgM)	甲型肝炎病毒(HAV)	白细胞(WBC)
免疫球蛋白 D(IgD)	乙型肝炎病毒(HBV)	血红蛋白(Hb)
免疫球蛋白 E(IgE)	丙型肝炎病毒(HCV)	