



中华人民共和国医药行业标准

YY/T ××××.2-20××

# 体外诊断检验系统 性能评价方法 第2部分：正确度

In vitro diagnostic test systems—Performance evaluation method  
—Part 2: Trueness

(征求意见稿)

××××-××-××发布

××××-××-××实施

国家药品监督管理局 发布

## 前 言

本部分按照 GB/T 1.1-2009 《标准化工作导则第 1 部分：标准的结构和编写》给出的规则起草。应注意，本部分中的某些条款可能涉及专利权的问题，起草单位并不负责在任何专利权的鉴别。

YY/T ×××× 《体外诊断检验系统 性能评价方法》，由下列部分组成：

- 体外诊断检验系统 性能评价方法 第 1 部分：精密度
- 体外诊断检验系统 性能评价方法 第 2 部分：正确度
- 体外诊断检验系统 性能评价方法 第 3 部分：检出限与定量限
- 体外诊断检验系统 性能评价方法 第 4 部分：线性区间与可报告区间
- 体外诊断检验系统 性能评价方法 第 5 部分：分析特异性
- 体外诊断检验系统 性能评价方法 第 6 部分：定性试剂的重复性、诊断敏感性和诊断特异性

本部分为 YY/T XXXX 的第 2 部分。

本部分由全国医用临床检验试验室和体外诊断系统标准化技术委员会（SAC/TC136）归口。

本部分起草单位：XXXXX。

本部分主要起草人：XXXXX。

## 引 言

在对体外诊断医疗器械产品进行性能评价时，体外诊断仪器、试剂、校准品等共同参与，反映的是仪器、试剂、校准品等组成的测量系统的性能，因此本系列标准采用系统的概念进行描述。

分析性能的评价是指对测量系统检测患者样品时可能具有的误差大小进行的估计。设计各种试验，对测量系统稳定状态（固有状态）误差的各种类型进行估计。体外诊断检验系统的分析性能包括精密度、正确度、准确度、检出限与定量限、线性区间与可报告区间、分析特异性等。

测量正确度是评价均匀样品的一系列测量结果中系统测量误差的性能特征。正确度是一个定性概念，但测量偏倚可以被评估。评估偏倚需有可用于确定被测量参考量值的适当参考物质或参考测量程序。正确度要与“准确度”进行区分。本文件提供了三种评价方法，第一种方法是分析具有被测量参考量值的适当参考物质，第二种方法是基于用患者样品进行方法学比对并估计偏倚，第三种方法是采用回收试验。



# 体外诊断检验系统 性能评价方法

## 第 2 部分：正确度

### 1. 范围

本部分规定了体外诊断检验系统的正确度性能评价方法。

本部分适用于制造商对定量测定的体外诊断检验系统进行性能评价。本部分不适用于结果报告为名义标度和序数标度的体外诊断检验系统，例如用于血细胞鉴定、微生物分型、核酸序列鉴定、尿液颗粒鉴定，结果报告为阴性、阳性或1+、2+、3+的体外诊断检验系统的性能评价。本部分也不适用于基于定量测量并通过阈值判断结果的定性体外诊断检验系统（例如酶联免疫吸附法的病原微生物抗原或抗体检测试剂盒）的性能评价。

本部分不适用于医学实验室正确度性能验证，也不适用于产品型式检验。

### 2. 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件，仅注日期的版本适用于本文件。凡是不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

YY/T 1441 体外诊断医疗器械性能评估通用要求

### 3. 术语和定义

下列术语和定义适用于本文件。

#### 3.1

**测量偏倚** measurements bias

**偏倚** bias

系统测量误差的估计值。

注1：偏倚反相关于正确度。

注2：偏倚的估计是一系列测量值的平均值减去参考量值。

[GB/T 29791.1—2013, 定义 A. 3. 25]

#### 3.2

**测量正确度** measurement trueness

**正确度** trueness

无穷多次重复测量所得量值的平均值与一个参考量值间的一致程度。

注1：测量正确度不是一个量，且因而不能以数字来表达。一致程度的量度在 GB/T 6379.3/ISO 5725-3 中给出。

注2：测量正确度与系统测量误差反相关，但与随机测量误差不相关。

注3：术语测量准确度不应用于测量正确度，且反之亦然。

[GB/T 29791.1—2013, 定义 A. 3. 34]

#### 3.3

**参考测量程序** reference measurement procedure

被接受作为提供适合其下列预期用途的测量结果的测量程序，预期用途包括评价测量同类量的其他测量程序测得量值的测量正确度、校准或参考物质赋值。

注 1：体外诊断医疗器械参考测量程序的要求在 GB/T19702/ISO 15193 中描述；

注 2：GB/T 21415/ISO 17511 和 YY/T 0638/ISO 18153 中给出用参考测量程序为体外诊断医疗器械校准物赋值的举例；

注 3：无需与同类量的测量标准相关而得到测量结果的测量程序被称为原级参考测量程序。见 ISO/IEC 指南 99:2007，定义 2.8。

[GB/T 29791.1—2013, 定义 3.59]

### 3.4

#### 参考物质 reference material

一种或多种指定特性足够均匀和稳定，已被证明适合在测量过程中或名义特性检验中预期应用的物质。

注 1：具有或没有指定量值的参考物质可用于测量精密度控制，而只有具有指定量值的参考物质可用于校准或测量正确度控制。

注 2：在给定测量中，给定参考物质只能用于校准或质量保证之一。

注 3：参考物质由含量以及名义特性的物质组成。

包含量的参考物质例子

示例1：标明纯度的水、其动态粘度用于校准粘度计；

示例2：用作校准物的含有标明质量分数的葡萄糖的血浆；

示例3：对于内在胆固醇浓度没有指定量值的人血清，只用作测量精密度控制的材料。

包含特性的参考物质例子：

示例4：指示一种或多种特定颜色的色图；

示例5：含有特定核酸序列的 DNA 化合物；

示例6：含有 19-雄烯二酮的尿液。

注4：参考物质有时候会整合到一个体外诊断医疗器械中。

示例1：在三相点容器中已知三相点的物质；

示例2：在透射滤光片支架上已知光密度的玻璃片；

示例3：固定在显微镜载玻片上均一尺寸的微球。

注5：带有权威机构发布的证书，并指明用于获得带有相关不确定度和溯源性的指定特性值的有效程序的参考物质被称为有证参考物质。

对胆固醇浓度有指定量值和相关测量不确定度的人血清，用作校准物或测量正确度控制物质。

注6：有些参考物质的量值在计量学上溯源到一个单位系统之外的测量单位。这些物质包括由世界卫生组织指定国际单位（IU）的生物来源测量标准。

注7：参考物质的性能指标包括其材料的溯源性，说明其来源和处理过程。体外诊断医疗器械参考物质的性能指标要求在 GB/T 19703/ISO 15194<sup>[12]</sup>有描述。

注8：参考物质的用途可以包括测量系统的校准、测量程序的评价、为其它材料赋值及质量控制。。

注9：名义特性的检验给出了名义特性的值和相关的不确定度。此不确定度不是测量不确定度。

注10：由 ISO/ REMCO（参考物质委员会）有个类似的定义，但使用术语测量过程来表示检验，涵盖测量和名义特性检验两种含义。

[GB/T 29791.1—2013, 定义 3.58]

### 3.5 正确度控制物质

用于评价特定测量系统的特定量测量偏倚的参考物质。

注 1：正确度控制物质的基质通常与预期检测的临床样品近似；

注 2：应评价并建立正确度控制物质与临床样品的互换性；

注 3：正确度控制物质可被其制造者作为有证参考物质。

[ISO/DIS 17511, 定义 3.46]

## 4 总则

### 4.1 总体要求

制造商在对体外诊断医疗器械进行性能评价时，其计划、实施、评价和文件化等相关过程应符合 YY/T 1441 的规定。制造商应规定所有管理和实施体外诊断医疗器械性能评估相关人员的责任和相互关系，并确保具备充足的资源。制造商设计评价方案，并进行测试，做好相关记录，所有文件和记录作为该产品技术文件的一部分。性能评价的负责人应对性能评价结果最终评定和审查，并形成评价报告。

### 4.2 正确度评价方法

测量正确度是评价均匀样品的一系列测量结果中系统测量误差的性能特征。正确度是一个定性概念，但测量偏倚可以被评估，它是一个与正确度反义的可测量的量。本文件提供了三种评价方法，第一种方法是分析具有被测量参考量值的适当参考物质，第二种方法是基于用患者样品进行方法学比对并估计偏倚，第三种方法是采用回收试验。

### 4.3 待评价产品

待评价产品一般是体外诊断试剂，该试剂与体外诊断仪器、校准品等组成特定的测量系统，评价时，应对包括待评价产品在内的测量系统的名称、型号、批号等基本信息进行记录并报告。

### 4.4 试验注意事项

正确度评价试验的研究者应能正确、熟练操作待评价的产品和（或）比较方法，以及相关的校准程序、质控程序、维护程序等。在试验开始时，应对待评价的产品和（或）比较方法进行校准，在试验过程中的校准频率应依照待评价的产品和（或）比较方法的使用说明。试验中应运行质控程序，一旦待评价的产品和（或）比较方法出现失控，应重新测定。

应实时检测数据的完整性和有效性。如果因质控原因或其他已识别和确认出的错误来源，影响到数据的真实性时，则剔除错误数据，并及时重复测试以补充数据。若剔除数据较多时，应考虑测量系统性能的稳定性及此时进行性能评价的适宜性。

## 5 使用参考物质的正确度评价

### 5.1 参考物质要求

推荐的参考物质包括：

- 具有互换性的有证参考物质或国家标准品、参考品；
- 具有互换性的正确度控制物质；
- 具有互换性的能力验证（PT）物质或室间质量评价（EQA）物质；
- 参考方法赋值的临床样品。

所选用参考物质的不确定度应与该产品校准品的不确定度相当。参考物质应至少 3 个浓度，覆盖待评价产品的测量区间。

## 5.2 测试时间和测试次数

同一水平中的一组测试应在重复性条件下进行，即在短暂的时间间隔内，由同一操作人员测试。不同水平的测试可以不在同一日进行。同一水平的重复测试次数应不小于 6。

## 5.3 数据分析

### 5.3.1 偏倚估计

偏倚的估计值  $B$  由下式给出：

$$B = \bar{X} - \mu \quad \dots\dots\dots (1)$$

式中：

$\bar{X}$ ：测试结果的算术平均值；

$\mu$ ：参考物质的被测特性的接受参考值。

### 5.3.2 计算置信区间

$$B - \sqrt{U_{lab}^2 + U_{ref}^2} \leq B \leq B + \sqrt{U_{lab}^2 + U_{ref}^2} \quad \dots\dots\dots (2)$$

式中：

$U_{lab}$ ：实验室测量不确定度；

$U_{ref}$ ：参考物质的不确定度。

### 5.3.3 示例

数据分析示例可参见附录 A。

## 6 使用患者样品的正确度评价

### 6.1 样品要求

应尽可能使用未经处理的样品，样品不应对比方法和待评价产品产生已知干扰，一般情况下样品数量不少于 100 份，分析物浓度应尽可能在测量区间内均匀分布。应记录样品相关信息（例如是否存在溶血、黄疸、脂血、浊度）。若样品量不够用于重复测定或复测，或测量区间的某些样品难以获得时，可通过混合、添加、稀释等办法进行处理。处理过程可能会引起样品互换性问题，因此使用处理过的样品应进行标注和记录，并且不能超过样品总数的 20%。

### 6.2 比较方法的选择

对比较方法有以下要求：

- 具有与待评价产品相同的单位或能够转换成相同的单位；
- 比待评价产品更低的不确定度；
- 优先选择参考测量程序或标准方法。

### 6.3 测试时间和测试次数

应尽可能使用当天采集的样品同时使用比较方法和待评价产品进行测试。若使用贮存的样品，应保证样品的稳定性。可在一段时间期限内完成所有样品的测试。无论是比较方法还是待评价产品，每个样品可进行一次测试，也可重复多次测试，重复测试的目的是减小测量的随机误差。

### 6.4 数据分析

#### 6.4.1 散点图或偏差图绘制

试验结束后，应目测检查数据，其目的是评估样品是否覆盖了预期的测量区间、检测是否存在离群值，并且可以对两种方法的差异有初步了解并决定如何去描述两种方法的变异特点。

可绘制散点图和偏差图进行目测检查。散点图用 x 轴表示比较方法测量结果，y 轴表示待评价产品测量结果。偏差图用 x 轴表示被测物浓度，y 轴表示待评价产品与比较方法间的差值。如果比较方法是参考测量程序，则偏差图的 x 轴一般用比较方法的测量结果。如果比较方法不是参考测量程序，则可用待评价产品与比较方法的均值作为 x 轴。x 轴也可为样品按浓度排序的顺序编号（即顺序偏差图）。为了判断两个方法间的差异是否一致或与浓度成比例，偏差图的 y 轴可以用两个方法的测量结果的差值，也可用差值的百分比。

根据差值的特点，偏差图有以下几种情况：

- a) 差值恒定变化（SD 恒定），即两个方法间的差值是恒定的；
- b) 差值成比例变化（CV 恒定），即两个方法间的差值与浓度成比例；
- c) 差值混合变化（SD 和 CV），即低浓度时差值恒定而高浓度时差值与浓度成比例；
- d) 偏倚随浓度改变，即两个方法间的偏倚以线性方式改变；
- e) 偏倚非线性，即两个方法间的偏倚呈非线性。

#### 6.4.2 离群值检验

可用广义极端学生化偏差 (Extreme Studentized deviate, ESD) 方法检验离群值，离群值的数量不应超过 5%，具体步骤为：

a) 根据 100 份样品的测定结果计算  $(Y_i - X_i)/X_i$  的平均值 ( $\bar{d}$ ) 和标准差 ( $s$ )，并根据  $\bar{d}$  和  $s$  计算最大偏倚 ( $ESD_i$ )， $ESD_i = \max(|d_i - \bar{d}|/s)$ ；

b) 根据检验统计的样品数量（本试验中为 100）计算临界值 ( $\lambda_i$ )；

$$\lambda_i = \frac{t_{v,p}(n-i)}{\sqrt{(n-i+1)(v+t_{v,p}^2)}} \dots\dots\dots (3)$$

式中：

$n$ ：原始数据组中的样品数量；

$i$ ：1, 2... $h$ ；

$v = n - i - 1$ ；

$$p = 1 - \frac{\alpha}{2(n-i+1)}$$

$t_{v, p}$ : t 分布中自由度为  $v$ 、概率为  $p$  下对应的  $t$  值;

$\alpha$ : 显著性水平, 一般选 0.05 或 0.01。

c) 若  $ESD_1 > \lambda_1$ , 则认为该观察值为离群值, 排除后进行第二个 ( $i=2$ ) 最大偏倚比较;

d) 离群值的个数是通过找到最大的  $i$  使得  $ESD_i > \lambda_i$ ;

e) 排除离群值后补充数据, 达到至少 100 例样品的数量要求。

### 6.4.3 偏倚初步估计

对于标准偏差恒定 (SD 恒定) 或者差值成比例变化 (CV 恒定) 的偏差图, 如果两种程序间的差值成正态分布, 则利用差值平均值作为估算的偏倚; 如果非正态分布, 则用中位数作为方法间估算的偏倚。

### 6.4.4 回归分析

回归分析类型通常包括线性回归方法(ORL)、加权最小二乘法(WLS)、Deming 回归和 Passing-Bablok 回归分析方法等。应根据方法间差值的变化特性 (SD 恒定、CV 恒定、差值混合变化) 选择适当的模型拟合。

### 6.4.5 医学决定水平处偏倚

医学决定水平处偏倚及其 95%CI 计算公式:

a) 对于 OLS 回归, 计算公式如下:

$$\hat{B}_{md} = a + (b - 1)X_{md} \quad \dots\dots\dots (4)$$

$$[\hat{B}_{md,L}, \hat{B}_{md,U}] = \hat{B}_{md} \pm t_{(n-2, 1-\frac{\alpha}{2})} \cdot S_{y \cdot x} \cdot \sqrt{\frac{1}{n} + \frac{(X_{md} - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}} \quad \dots\dots\dots (5)$$

式中:

$\hat{B}_{md}$ : 医学决定水平处偏倚估计值;

a: 回归方程的截距;

b: 回归方程的斜率;

$X_{md}$ : 为医学决定水平;

$\hat{B}_{md,L}, \hat{B}_{md,U}$ : 分别为医学决定水平处偏倚 95% 置信区间下限、上限估计值;

$t_{(n-2, 1-\frac{\alpha}{2})}$ : t 分布 95% 分位数;

$S_{y \cdot x}$  为回归标准误, 计算公式如下:

$$S_{y \cdot x} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-2}} \quad \dots\dots\dots (6)$$

b) 对于 WLS 回归, 计算公式如下:

$$\hat{B}_{md} = a + (b - 1)X_{md} \quad \dots\dots\dots (7)$$

$$[\hat{B}_{md,L}, \hat{B}_{md,U}] = \hat{B}_{md} \pm t_{(n-2, 1-\frac{\alpha}{2})} \cdot \sqrt{\frac{1}{\sum_{i=1}^n w_i} + \frac{(X_{md} - \bar{x}_w)^2}{\sum_{i=1}^n w_i x_i^2 - \frac{(\sum_{i=1}^n w_i x_i)^2}{\sum_{i=1}^n w_i}}} \quad \dots\dots\dots (8)$$

式中:

$\hat{B}_{md}$ : 医学决定水平处偏倚估计值;

a: 回归方程的截距;

b: 回归方程的斜率;

$X_{md}$ : 医学决定水平;

$\bar{x}_w$ : 加权平均值;

$w_i$ : 加权值;

$\hat{B}_{md,L}, \hat{B}_{md,U}$ : 分别为医学决定水平处偏倚 95%置信区间下限、上限估计值;

$t_{(n-2, 1-\frac{\alpha}{2})}$ : t 分布 95%分位数 (a 取值为 0.05 时);

c) 对于 Deming 回归, 计算公式如下:

$$\hat{B}_{md} = a + (b - 1)X_{md} \quad \dots\dots\dots (9)$$

$$[\hat{B}_{md,L}, \hat{B}_{md,U}] = \hat{B}_{md} \pm t_{(n-2, 1-\frac{\alpha}{2})} \cdot \sqrt{\hat{\sigma}_a^2 + x_{md}^2 \hat{\sigma}_b^2 + 2x_c \hat{\sigma}_{ab}} \quad \dots\dots (10)$$

式中:

$\hat{B}_{md}$ : 医学决定水平处偏倚估计值;

a: 回归方程的截距;

b: 回归方程的斜率;

$X_{md}$ : 医学决定水平;

$\hat{\sigma}_a^2$ : 截距方差估计值;

$\hat{\sigma}_b^2$ : 斜率方差估计值;

$\hat{\sigma}_{ab}$ : 截距斜率协方差估计值;

$\hat{B}_{md,L}, \hat{B}_{md,U}$ : 分别为医学决定水平处偏倚 95%置信区间下限、上限估计值;

$t_{(n-2, 1-\frac{\alpha}{2})}$ : t 分布 95%分位数 (a 取值为 0.05 时)。

#### 6.4.6

数据分析示例可参见附录 B。

### 7 回收试验

#### 7.1 适用条件

在对被测量定义清楚, 且有与被测量一致的纯品或 5.1 所规定的高浓度参考物质条件下, 可采用回收试验进行正确度评价。

#### 7.2 试验方法

预先准备标准溶液作为 A 液。A 液可以直接选用合适的高浓度参考物质, 若无参考物质, 也可用与被测量一致的纯品进行配制, 配制时应采用重量法, 以减小配制过程中的不确定度。选择合适浓度的临床样品, 作为 B 液。将不同体积量的标准溶液 A 加入样品 B 中, 配制成至少 3 个不同浓度的回收样品, 覆盖待评价产品的测量区间。每个浓度重复检测 3 次。需注意的是:

- 为保证样品的基质皆可能接近临床样品, 加入的标准溶液体积一般不超过总体积的 10%;
- 因为标准物溶液加入的体积不到 10%, 为保证得到不同浓度的回收样品, 标准溶液的浓度应足够高;
- 保证加样准确度, 可采用重量法;
- 至少有一份回收样品的被测物浓度在医学决定水平附近。

#### 7.3 数据分析

根据下列公式计算回收率  $R$ 。

$$R = \frac{C \times (V_0 + V) - C_0 \times V_0}{V \times C_s} \times 100\% \quad (11)$$

式中:

- $R$ : 回收率, %;
- $V$ : 标准溶液 A 的体积;
- $V_0$ : 样品 B 的体积;
- $C$ : 将标准溶液 A 加入样品 B 后的检测浓度均值;
- $C_0$ : 样品 B 的浓度;
- $C_s$ : 样品 A 的浓度。

### 8 再评价

当体外诊断医疗器械出现设计或生产程序变更时, 应根据需要重新进行正确度评价研究, 以确保符合产品的预期用途和性能声明。

## 附录 A

## (资料性附录)

## 使用参考物质的正确度评价数据分析示例

## A.1 评价方法

使用 3 个水平的血清胆固醇、总甘油、游离甘油和甘油三酯标准物质进行胆固醇检测系统正确度评价，通过测量均值和接受参考值及其不确定度计算计算偏倚和置信区间。

## A.2 测量结果

使用仪器配套校准物进行系统校准，校准通过后使用配套质控物质进行质控，当质控物质测量结果落在其范围内时，进行系统正确度评价试验。试验在重复性测量条件下进行，即在短暂的时间间隔内，由同一操作人员对每一水平标准物质重复测量 6 次，测量结果见表 A.1：

表 A.1 正确度评价结果

测量单位：mg/dL

测量项目	胆固醇		
	水平 1	水平 2	水平 3
1	193.2	149.7	120.1
2	192.9	149.5	118.9
3	192.1	150	119.2
4	191.6	150.2	119.5
5	192.4	149.3	119.8
6	191.9	150.1	119.5

## A.3 评价结果

## A.3.1 国家标准物质信息

国家标准物质赋值及其不确定度见表 A.2：

表 A.2 国家标准物质赋值及其不确定度

单位：mg/dL

测量项目	胆固醇		
	水平 1	水平 2	水平 3
赋值	197.6	152.4	122.1

不确定度	2.5	1.9	1.6
------	-----	-----	-----

### A.3.2 计算实验室测量均值及测量不确定度

测量系统校准物质赋值结果为 200.0 mg/dL，不确定度为 5.0 mg/dL，假定校准物质的不确定度在（100~200）mg/dL 区间内呈线性分布，因此在国家标准物质的 3 个测量水平上校准物带来的不确定度估计值、每一水平国家标准物质测量结果的均值、标准差、均值的标准误以及测量不确定度，计算过程中比原始数据多保留一位有效数字，见表 A.3：

表 A.3 测量均值及不确定度

单位：mg/dL

测量项目	胆固醇		
	水平 1	水平 2	水平 3
测量均值	192.35	149.80	119.50
标准差	0.61	0.36	0.42
测量不确定度	0.25	0.15	0.17

### A.3.3 计算偏倚及其置信区间

实验室测量偏倚及其置信区间的计算结果见表 A.4：

表 A.4 测量偏倚计算结果

单位：mg/dL

测量项目	胆固醇		
	水平 1	水平 2	水平 3
赋值	197.6	152.4	122.1
测量均值	192.4	149.8	119.5
偏倚	-5.2	-2.6	-2.6
$U_{lab}^2$	0.3	0.1	0.1
$U_{ref}^2$	6.3	3.6	2.6
CI <sub>下限</sub>	-7.7	-4.5	-4.2
CI <sub>上限</sub>	-2.7	-0.7	-1.0

注：正确度评价时，所选用测量标准的不确定度与校准物不确定度相当，一般情况下两个不确定度平方比（大/小）不超过 10 即可认为相当。

## 附录 B

## (资料性附录)

## 使用患者样品的正确度评价的数据分析示例

## B.1 评价方法

假定评价试剂盒和对比试剂盒线性范围和医学决定水平平均相同，线性区间（10~300）mg/dL，医学决定水平 125mg/dL。

使用分析物浓度在线性区间内均匀分布的至少 100 份人血清样本，在对比系统和评价系统上同时对每份样本测量 1 次。记录样品相关信息（例如是否存在溶血、黄疸、脂血、浊度）。对测量结果进行离群值分析、剔除统计离群值后进行总体偏差估计、回归分析、以及医学决定水平处偏差的估计。

## B.2 测量结果

分别使用对比系统和评价系统配套校准物进行各自系统校准，校准通过后使用配套质控物质进行质控，当质控物质测量结果落在其范围内时，进行系统正确度评价试验。试验在重复性测量条件下进行，即在短暂的时间间隔内，由同一操作人员对每一样品测量 1 次，测量结果见表 B.1：

表 B.1 正确度评价结果

测量单位：mg/dL

序号	对照	考核	序号	对照	考核	序号	对照	考核
1	21	23	41	98	95	81	105	107
2	128	130	42	168	172	82	124	127
3	43	40	43	243	250	83	109	112
4	125	128	44	265	272	84	79	83
5	165	163	45	134	138	85	92	95
6	223	228	46	123	128	86	84	90
7	228	242	47	78	81	87	28	24
8	19	20	48	25	21	88	29	26
9	24	21	49	34	35	89	89	94
10	68	72	50	48	52	90	73	75
11	75	76	51	198	204	91	57	56
12	129	125	52	196	203	92	189	194
13	142	137	53	45	54	93	194	199
14	54	51	54	65	69	94	199	209
15	37	35	55	37	32	95	78	83
16	88	94	56	89	84	96	42	44
17	94	87	57	90	95	97	78	82
18	107	113	58	91	95	98	89	94
19	123	128	59	186	195	99	95	100
20	43	41	60	54	57	100	104	109
21	55	52	61	28	27	101	189	192
22	65	69	62	95	99	102	28	25
23	78	75	63	132	138	103	147	150
24	91	87	64	139	142	104	128	132
25	12	13	65	256	260	105	18	15

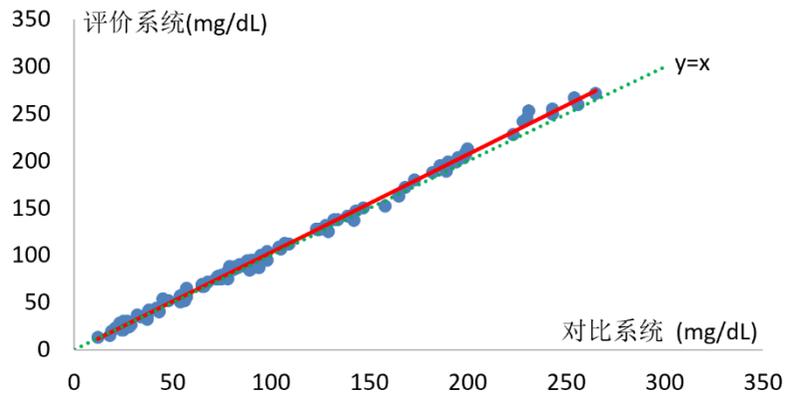
26	23	28	66	231	253	106	25	23
27	25	30	67	78	83	107	189	194
28	82	88	68	92	95	108	254	267
29	199	207	69	37	39	109	173	180
30	230	245	70	75	79	110	66	69
31	158	152	71	56	52	111	73	78
32	143	147	72	98	104	112	85	89
33	125	128	73	57	65	113	81	85
34	79	88	74	54	56	114	182	188
35	72	75	75	38	42	115	195	204
36	38	42	76	32	37	116	65	67
37	66	67	77	56	59	117	75	75
38	54	56	78	189	189	118	57	59
39	27	30	79	200	213	119	190	199
40	83	87	80	243	255	120	82	86

B.3 数据统计分析

B.3.1 图示分析

B.3.1.1 散点图

以对比系统测量结果为 x 轴，评价系统测量结果为 y 轴绘制散点图，见图 1。

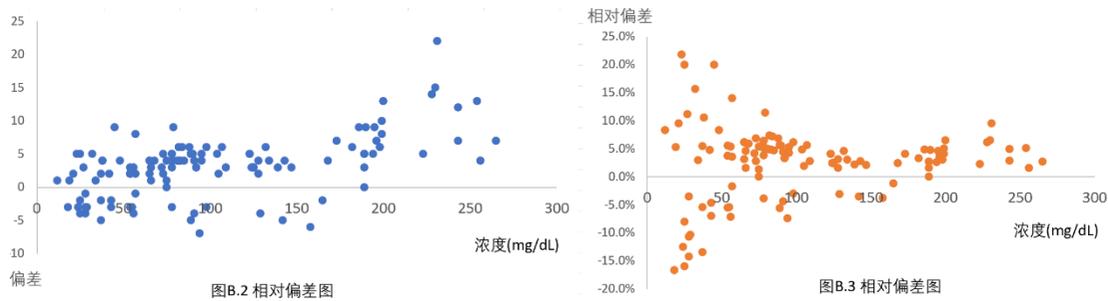


图B.1 散点图

散点图显示对比系统测量结果和评价系统测量结果紧密分布在  $y=x$  的周围，差异与浓度不成比例。

B.3.1.2 偏差图

以对比系统和评价系统测量结果的平均值作为 x 轴，以评价系统与对比系统测量结果之差及其差值占对比系统测量结果的百分比分别作为 y 轴绘制偏差图和相对偏差图。



偏差图显示对比系统和评价系统测量结果的偏差在低浓度时偏差基本恒定,在高浓度时相对偏差基本恒定,同时可以看出 100 mg/dL 可以作为偏差分组的临界值。

**B. 3. 2 离群值检验**

根据 B. 3. 1. 2 偏差分布图得知, x 轴数据小于 100mg/dL 时偏差基本恒定, 大于等于 100mg/dL 时相对偏差基本恒定, 因此, 将以上数据分为两组, 其中小于 100mg/dL 的有 75 个, 大于等于 100mg/dL 的有 45 个, 分别使用 ESD 法进行偏差和相对偏差离群值检验, 检验结果见表 B. 2。

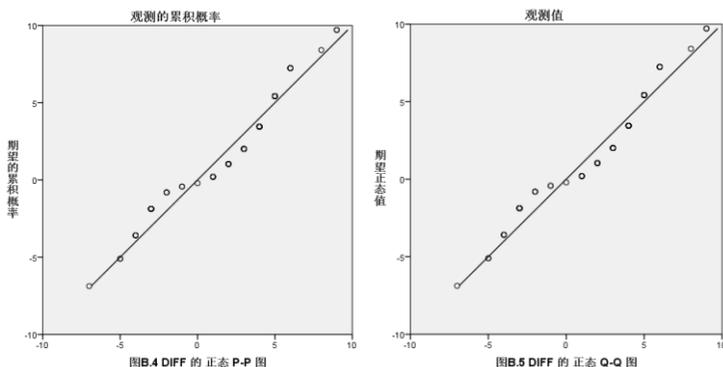
**表 B. 2 离群值检验结果**

偏差类型	偏差			相对偏差		
	1	2	3	4	5	6
序号	1	2	3	4	5	6
平均值	1. 827	1. 730	1. 630	3. 00%	2. 85%	2. 77%
SD	3. 633	3. 559	3. 478	2. 52%	2. 35%	2. 30%
ESDi	1. 974	2. 043	1. 831	2. 585	1. 564	1. 620
$\lambda_i$	3. 648	3. 638	3. 627	3. 435	3. 415	3. 392
偏差	9. 000	9. 000	8. 000	9. 52%	6. 52%	6. 50%
编号	53	34	73	66	30	79
结果判定	非离群值	非离群值	非离群值	非离群值	非离群值	非离群值

**B. 3. 3 偏倚估计**

**B. 3. 3. 1 低值区间 (<100mg/dL)**

x 轴数据小于 100mg/dL 的 75 个数据之间的偏差基本恒定,其偏差分布 P-P 图、Q-Q 图分别见图 B. 4、图 B. 5:



从 P-P 图、Q-Q 图可以看出数据紧邻对角线, 未见明显偏离正态情形; 同时进行偏度和峰度检验,

检验结果见表 B. 3。

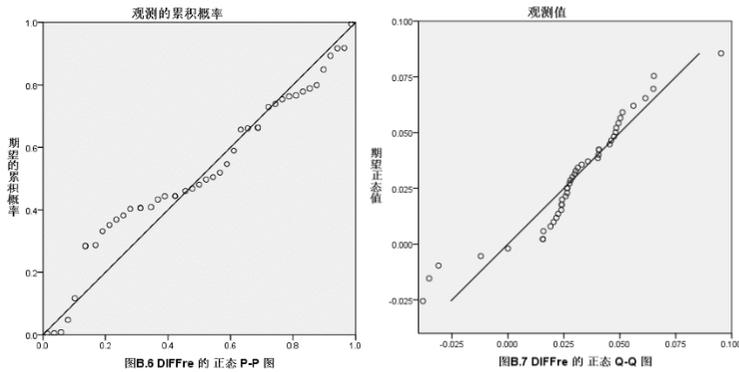
表 B. 3 正态性检验结果

描述统计量									
	N	极小值	极大值	均值	标准差	偏度		峰度	
	统计量	统计量	统计量	统计量	统计量	统计量	标准误	统计量	标准误
DIFF	75	-7.00	9.00	1.8267	3.63308	-.463	.277	-.600	.548
有效的 N (列表状态)	75								

偏度和峰度的 u 检验结果均小于 1.96,  $P > 0.05$ , 说明该组数据在偏度和峰度上均未偏离正态分布。既然数据分布未偏离正态, 则用偏差的平均值作为该区间内偏差的估计值, 偏倚及其 95%CI 为: 偏差的平均值为 1.8, 其 95%CI 为 (1.0, 2.7)。

B. 3. 3. 2 高值区间 (A.100mg/dL)

x 轴数据大于等于 100mg/dL 的 45 个数据之间的相对偏差基本恒定, 其相对偏差的分布的 P-P 图、Q-Q 图分别见图 B. 6、B. 7:



从 P-P 图、Q-Q 图可以看出数据不紧邻对角线, 呈一定的离散状态, 怀疑数据存在偏离正态情形; 使用偏度和峰度检验进行进一步确定, 检验结果见表 B. 4。

表 B. 4 正态性检验结果

描述统计量									
	N	极小值	极大值	均值	标准差	偏度		峰度	
	统计量	统计量	统计量	统计量	统计量	统计量	标准误	统计量	标准误
DIFFre	45	-.04	.10	.0300	.02523	-.692	.354	1.994	.695
有效的 N (列表状态)	45								

偏度和峰度的 u 检验结果分别为 1.95、2.87, 峰度检验的  $P < 0.01$ , 说明该组数据在峰度上偏离正态分布。既然数据分布偏离正态, 则用相对偏差的中位值作为该区间内相对偏差的估计值, 相对偏差及其 95%CI 为:

相对偏差的中位值为 2.9%, 其 95%CI 为 (2.6%, 4.0%)。

B. 3. 4 一元一次回归分析

根据偏差的分布选择适合的拟合方式获得回归方程, 本例中不考虑偏差分布情况, 分别就以下几种回归分别举例。以下公式中 i 表示参与回归样本数量,  $i=1 \dots n$ ; j 表示重复次数,  $j=1 \dots r$ 。

B. 3. 4. 1 最小二乘法 (LSD/OLS/OLR)

a) 方程及参数

方程:  $y = a + bx$

均值:  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ ;  $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ ;  $x_i = \frac{1}{r} \sum_{j=1}^r x_{ij}$ ;  $y_i = \frac{1}{r} \sum_{j=1}^r y_{ij}$

斜率:  $b = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$ ;  $\sigma_b = \frac{S_{y \cdot x}}{\sum_{i=1}^n (x_i - \bar{x})^2}$

截距:  $a = \bar{y} - b\bar{x}$ ;  $\sigma_a = S_{y \cdot x} \cdot \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}$

回归标准误:  $S_{y \cdot x} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-2}}$

系数95%CI:  $\bar{y} \pm t_{(n-2, 1-a/2)} \cdot \sigma$

b) 参数计算

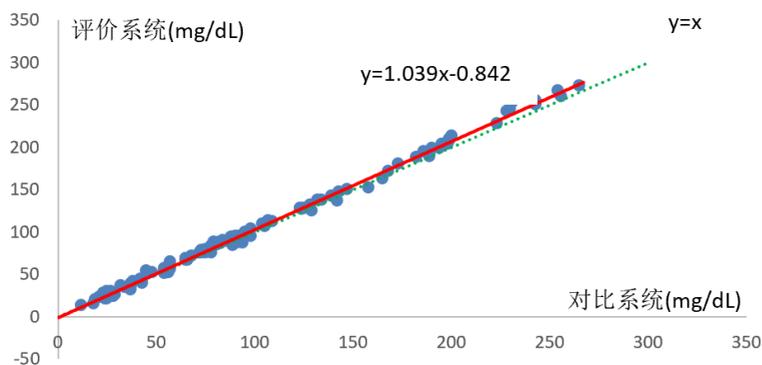
按照以上公式计算结果见表B. 5:

表 B. 5 OLS 回归结果

模型	系数	标准误差	t	Sig.	95% CI	
					下限	上限
(常量)	-0.842	0.663	-1.269	0.207	-2.154	0.471
斜率	1.039	0.005	190.100	0.000	1.029	1.050

c) 回归方程

使用OLS得到的回归方程为  $y = -0.842 + 1.039x$ 。OLS回归散点图见图B. 8:



图B.8 OLS回归

B. 3. 4. 2 加权的最小二乘法 (WLS)

a) 方程及参数

方程:  $y = a_w + b_w x$

权重:  $\hat{w}_i = \frac{1}{\hat{\sigma}_i^2}$ ;  $\hat{\sigma}_i = a_\sigma + b_\sigma x_i$ ;  $\sigma_i = |e_i|$ ;  $e_i = y_i - \hat{y}_i = y_i - (a + b x_i)$

加权均值:  $\bar{x}_w = \frac{\sum_{i=1}^n x_i w_i}{\sum_{i=1}^n w_i}$ ;  $\bar{y}_w = \frac{\sum_{i=1}^n y_i w_i}{\sum_{i=1}^n w_i}$

斜率:  $b_w = \frac{\sum_{i=1}^n w_i x_i y_i - \frac{\sum_{i=1}^n w_i x_i \sum_{i=1}^n w_i y_i}{\sum_{i=1}^n w_i}}{SSx_w}$ ;  $\sigma_b = \frac{s_{y \cdot x}}{SSx_w}$ 。其中  $SSx_w = \sum_{i=1}^n w_i x_i^2 - \frac{(\sum_{i=1}^n w_i x_i)^2}{\sum_{i=1}^n w_i}$

截距:  $a_w = \bar{y}_w - b_w \bar{x}_w$ ;  $\sigma_a = s_{y \cdot x} \sqrt{\frac{1}{\sum_{i=1}^n w_i} + \frac{\bar{x}_w^2}{SSx_w}}$

回归标准误:  $s_{y \cdot x} = \sqrt{\frac{\sum_{i=1}^n e_i^2 w_i}{n-2}}$ 。其中  $e_i = y_i - \hat{y}_i = y_i - (a_w + b_w x_i)$

系数95%区间:  $\bar{y} \pm t_{(n-2, 1-a/2)} \cdot \sigma$

b) 参数计算

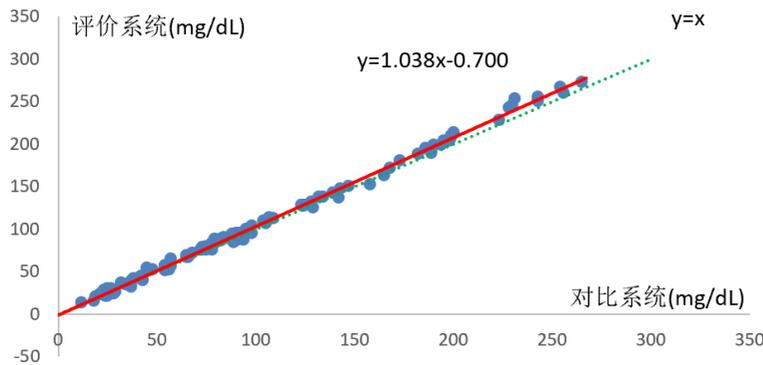
按照以上公式计算结果见表B. 6:

表 B. 6WLS 回归结果

模型	系数	标准误差	t	Sig.	95% CI	
					下限	上限
(常量)	-0.700	0.620	-1.129	0.261	-1.927	0.528
斜率	1.038	0.006	179.599	0.000	1.027	1.050

c) 回归方程

使用WLS得到的回归方程为  $y = -0.700 + 1.038x$ 。WLS回归散点图见图B. 9:



图B.9 WLS回归

### B. 3. 4. 3 Deming回归

a) 方程及参数

方程:  $y = a + bx$

均值:  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ ;  $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ ;  $x_i = \frac{1}{r} \sum_{j=1}^r x_{ij}$ ;  $y_i = \frac{1}{r} \sum_{j=1}^r y_{ij}$

斜率:  $b = \frac{\hat{\sigma}_y^2 - \lambda \hat{\sigma}_x^2 + \sqrt{(\hat{\sigma}_y^2 - \lambda \hat{\sigma}_x^2)^2 + 4\lambda \hat{\sigma}_{xy}^2}}{2\hat{\sigma}_{xy}}$ ;  $\hat{\sigma}_b^2 = \frac{b^2}{n\hat{\sigma}_{xy}^2} (\hat{\sigma}_x^2 \hat{\sigma}_y^2 - \hat{\sigma}_{xy}^2)$ 。其中  $\hat{\sigma}_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$ ;  $\hat{\sigma}_y^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2$ ;

$\hat{\sigma}_{xy} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$ ;  $\lambda = \frac{\sum_{i=1}^n \sum_{j=1}^r (x_{ij} - x_i)^2}{\sum_{i=1}^n \sum_{j=1}^r (y_{ij} - y_i)^2}$ , 当  $r=1$ 时假定预期值为1;  $\hat{\sigma}_{ab} = -\frac{\bar{x}b^2}{n\hat{\sigma}_{xy}^2} (\hat{\sigma}_x^2 \hat{\sigma}_y^2 - \hat{\sigma}_{xy}^2)$

截距:  $a = \bar{y} - b\bar{x}$ ;  $\hat{\sigma}_a^2 = \frac{1}{n} [\hat{\sigma}_y^2 - 2b\hat{\sigma}_{xy} + b^2\hat{\sigma}_x^2 + \frac{\bar{x}^2 b^2}{\hat{\sigma}_{xy}^2} (\hat{\sigma}_x^2 \hat{\sigma}_y^2 - \hat{\sigma}_{xy}^2)]$

系数95%区间:  $\bar{y} \pm t_{(n-2, 1-a/2)} \cdot \sigma$

b) 参数计算

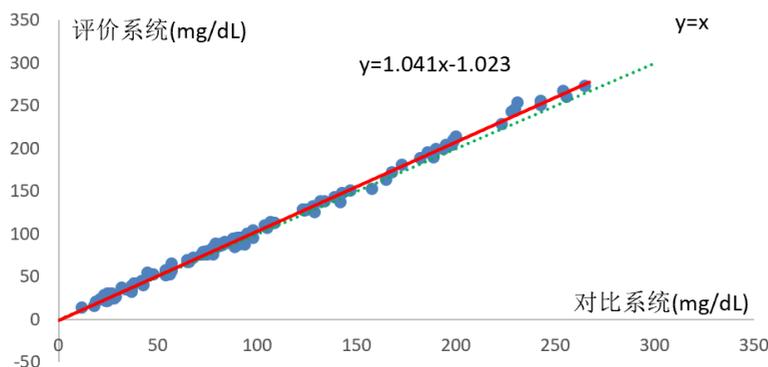
按照以上公式计算结果见表B.7:

表 B.7 Deming 回归结果

模型	系数	标准误差	t	Sig.	95% CI	
					下限	上限
(常量)	-1.023	0.658	-1.554	0.123	-2.327	0.281
斜率	1.041	0.005	191.704	0.000	1.030	1.052

c) 回归方程

使用Deming得到的回归方程为  $y = -1.023 + 1.041x$ 。Deming回归散点图见图B.10:



图B.10 Deming回归

### B.3.4.4 Passing-Bablok回归

a) 方程及参数

方程:  $y = a + bx$

斜率: 当n为奇数时,  $b = S_{(\frac{n+1}{2}+k)}$ ; 当n为偶数时,  $b = \frac{1}{2} (S_{(\frac{n}{2}+k)} + S_{(\frac{n}{2}+1+k)})$ 。其中n组排序的测

量数据对中  $S_{ij}$  为任意两点的斜率,  $S_{ij} = \frac{(y_i - y_j)}{(x_i - x_j)}$ ; k为  $S_{ij}$  小于-1的数量

截距:  $a = \text{中位值}\{y_i - bx_i\}$

系数95%区间:

$$S_{(m_1+k)} \leq b \leq S_{(m_2+k)} ; \text{其中 } m_1 = \frac{n-C_\gamma}{2}, m_2 = n - m_1 + 1, C_\gamma = Z_{(1-\frac{\gamma}{2})} \sqrt{\frac{n(n-1)(2n-5)}{18}}, Z_{(1-\frac{\gamma}{2})}$$

为标准正态分布的分位数,  $\gamma$ 为显著性水平

$$a_L = \text{中位数}\{y_i - b_U x_i\} \quad a_U = \text{中位数}\{y_i - b_L x_i\}$$

b) 参数计算

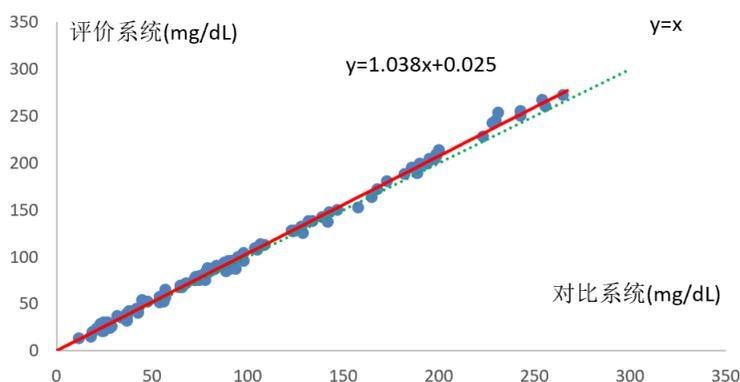
按照以上公式计算结果见表B. 8:

表 B. 8 Passing-Bablok 回归结果

模型	系数	95% CI	
		下限	上限
(常量)	0.025	-0.879	1.012
斜率	1.038	1.025	1.049

c) 回归方程

使用Passing-Bablok得到的回归方程为  $y = 0.025 + 1.038x$ 。Passing-Bablok回归散点图见图B. 11:



图B.11 Passing-Bablok回归

### B. 3. 5 医学决定水平处偏差

将医学决定水平浓度代入回归方程，按照公式  $\hat{B}_{md} = a + (b - 1)X_{md}$  计算其偏差估计值。

对于偏差 95%CI，不同回归方式对应的计算公式各不相同。以下列举几种常见的计算公式。

OLS 回归医学决定水平处偏差 95%CI 计算公式为：

$$[\hat{B}_{md,L}, \hat{B}_{md,U}] = \hat{B}_{md} \pm t_{(n-2, 1-\frac{\alpha}{2})} \cdot \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-2}} \cdot \sqrt{\frac{1}{n} + \frac{(X_{md} - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}$$

WLS 回归医学决定水平处偏差 95%CI 计算公式为：

$$[\hat{B}_{md,L}, \hat{B}_{md,U}] = \hat{B}_{md} \pm t_{(n-2, 1-\frac{\alpha}{2})} \cdot \sqrt{\frac{1}{\sum_{i=1}^n w_i} + \frac{(X_{md} - \bar{x}_w)^2}{\sum_{i=1}^n w_i x_i^2 - \frac{(\sum_{i=1}^n w_i x_i)^2}{\sum_{i=1}^n w_i}}}$$

Deming 回归医学决定水平处偏差 95%CI 计算公式为：

$$[\hat{B}_{md,L}, \hat{B}_{md,U}] = \hat{B}_{md} \pm t_{(n-2, 1-\frac{\alpha}{2})} \cdot \sqrt{\hat{\sigma}_a^2 + x_{md}^2 \hat{\sigma}_b^2 + 2x_c \hat{\sigma}_{ab}}$$

OLS、WLS、Deming 回归医学决定水平处 (125 mg/dL) 偏差 95%CI 计算结果分别见表 B. 9、B. 10、B. 11:

表 B. 9 125 mg/dL 处偏差及其 95%CI (OLS)

类型	偏倚	下限	上限	单位
----	----	----	----	----

偏差	4.1	3.4	4.8	mg/dL
相对偏差	3.3%	2.7%	3.9%	/

表 B.10 125 mg/dL 处偏差及其 95%CI (WLS)

类型	偏倚	下限	上限	单位
偏差	4.1	3.5	4.7	mg/dL
相对偏差	3.2%	2.8%	3.7%	/

表 B.11 125 mg/dL 处偏差及其 95%CI (Deming)

类型	偏倚	下限	上限	单位
偏差	4.1	3.4	4.9	mg/dL
相对偏差	3.3%	2.7%	3.9%	/

### 参考文献

- [1] GB/T 3358.1-2009 统计学词汇及符号 第1部分：一般统计术语与用于概率的术语
- [2] GB/T 3358.2-2009 统计学词汇及符号 第2部分：应用统计
- [3] GB/T 3358.3-2009 统计学词汇及符号 第3部分：试验设计
- [4] GB/T 6379.1-2004 测量方法与结果的准确度（正确度与精密度） 第1部分：总则与定义
- [5] GB/T 6379.2-2004 测量方法与结果的准确度（正确度与精密度） 第2部分：确定标准测量方法重复性与再现性的基本方法
- [6] GB/T 6379.3-2012 测量方法与结果的准确度(正确度与精密度) 第3部分：标准测量方法精密度的中间度量
- [7] GB/T 6379.4-2006 测量方法与结果的准确度(正确度与精密度) 第4部分：确定标准测量方法正确度的基本方法
- [8] GB/T 6379.5-2006 测量方法与结果的准确度(正确度与精密度) 第5部分：确定标准测量方法精密度的可替代方法
- [9] GB/T 6379.6-2009 测量方法与结果的准确度（正确度与精密度） 第6部分：准确度值的实际应用
- [10] GB/T 26124-2011 临床化学体外诊断试剂盒
- [11] GB/T 29791.1-2013 体外诊断医疗器械 制造商提供的信息（标示）第1部分：术语、定义和通用要求
- [12] WS/T 492-2016 临床检验定量测定项目精密度与正确度性能验证
- [13] 体外诊断试剂分析性能评估（准确度-方法学比对）技术审查指导原则
- [14] 体外诊断试剂分析性能评估（准确度-回收试验）技术审查指导原则
- [15] JCGM 200:2012 International vocabulary of metrology – Basic and general concepts and associated terms (VIM)
- [16] EP09c Measurement Procedure Comparison and Bias Estimation Using Patient Samples（用患者样品进行测量方法比对及偏倚估计，2018年）
- [17] EP15-A3 User Verification of Precision and Estimation of Bias; Approved Guideline--Third Edition（用户对精密度和准确度的性能验证，2014年）
- [18] EP10-A3-AMD Preliminary Evaluation of Quantitative Clinical Laboratory Measurement Procedures（临床试验室定量测定方法的预评估，2014年）