

ICS 点击此处添加 ICS 号

CCS 点击此处添加 CCS 号

WS

中华人民共和国卫生行业标准

XX/T XXXXX—XXXX

代替 WS/T 356-2011

参考物质互通性评估指南

Guideline for evaluation of commutability of reference materials

(点击此处添加与国际标准一致性程度的标识)

(征求意见稿)

在提交反馈意见时，请将您知道的相关专利连同支持性文件一并附上。

XXXX - XX - XX 发布

XXXX - XX - XX 实施

发布

目 次

前 言	3
参考物质互通性评估指南.....	4
1 范围	4
2 术语和定义	4
3 评估原理	5
4 实验材料	5
4.1 测量程序和参考物质	5
4.2 临床样本.....	5
5 测量程序	6
5.1 测量程序的选择	6
5.2 评价方法.....	6
6 数据分析	7
6.1 回归方法的选择	7
6.2 数据分布的线性评价	8
6.3 OLS 回归分析	8
6.4 Deming 回归分析.....	9
附录 A.....	14
附录 B.....	17
附录 C.....	21
附录 D.....	22

前 言

本标准按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

本标准由国家卫生健康委临床检验标准专业委员会负责技术审查和技术咨询，由国家卫生健康委医疗管理服务指导中心负责协调性和格式审查，由国家卫生健康委员会医政医管局负责业务管理、法规司负责统筹管理。

本标准替代WS/T 356-2011《基质效应与互通性评估指南》，本标准主要技术变化如下：

- 修改了标准名称
- 修正了原标准中部分文字叙述和公式上面的不当之处
- 增加了Deming回归分析的内容
- 附件增加了IFCC评估方案的内容

本标准起草单位：北京医院、广东省中医院、首都医科大学附属北京天坛医院、中国计量科学研究院、北京市医疗器械检验研究院、解放军总医院第七医学中心、山东大学齐鲁医院、广东省人民医院、贵州省人民医院。

本标准主要起草人：张传宝、张天娇、黄宪章、张国军、武利庆、李胜民、刘杰、张义、侯铁英、聂瑛洁。

参考物质互通性评估指南

1 范围

本标准规定了临床检验参考物质互通性评价的实验设计和数据分析方法。

本标准适用于体外诊断产品生产商、临床实验室、室间质量评价/能力验证组织者和标准物质生产者在制备和使用参考物质过程中的互通性评价。

2 术语和定义

2.1

参考物质 reference material

在一项或多项特性上具有足够的均匀性和稳定性，用于校准、给其他物质赋值或提供质量保证的物质。

[来源：ISO/IEC 导则 99-2007，定义 5.1.3]

2.2

制备物 processed samples

本标准中，是指除临床样本以外的经过加工处理步骤得到的参考物质，包括：标准物质、校准品、质评/质控品等。

2.3

基质 matrix

一个物质系统中除待测物之外的所有成分。

2.4

基质效应 matrix effect

独立于被分析物存在的对测量和测量所得量值产生影响的样品特性。

注：对质谱分析技术，基质效应指标本中除待测物以外的其他成分对分析方法测量能力的干扰。

[来源：改写ISO/IEC导则99-2007，定义5.14]

2.5

互通性 commutability

参考物质的特性，对于给定的参考物质的规定量，由两个给定测量程序所得测量结果之间关系与另一个指定物质所得测量结果之间关系的一致程度。

注1: 互通性, 有时亦作互换性, 均为“commutability”的中译文。

注2: 临床检验领域参考物质的互通性, 通常指用不同测量程序测量该物质时, 各测量程序测量结果之间的数字关系, 与用这些测量程序测量实际临床样本时测量结果的数字关系的一致程度, 即该物质理化性质与实际临床样本的接近程度。

[来源: ISO/IEC导则99-2007, 定义5.15]

2.6

计量溯源性 metrological traceability

溯源性 traceability

通过文件规定的不间断的校准链, 测量结果与参照对象联系起来特性, 校准链中的每项校准均会引入测定不确定度。

[来源: JJF 1001-2011 4.14]

3 评估原理

在临床实验室常规测量中, 待测物的检测信号与其活性或浓度的关系常常受到环境条件(如温度或基质)的影响, 因此两种测量程序之间的关系很大程度上依赖于所选择的待测样本的性质。由于临床检验分析的对象是临床样本, 故在本标准中亦使用一批具有代表性的临床样本作为比对标准。

评估方式如下: 用两种测量程序同时对选定的一系列具有代表性的临床样本和参考物质进行分析, 利用两种测量程序得到临床样本的结果建立数学关系。参考物质测定的结果偏离这一数学关系的程度即反映其基质效应的大小。一般来说, 参考物质与临床样品的性质差异越大, 数据的偏离程度将越大, 该物质的互通性越差。

测量程序的不精密度和非特异性会影响测量数据的离散程度, 可以通过多次重复测定并采用均值进行数据分析的方式降低不精密度, 从而减少随机误差对互通性结果判读的影响。

4 实验材料

4.1 测量程序和参考物质

评价互通性的测量程序包含试剂、校准物和仪器等, 应对其进行说明。实验过程中, 所使用的试剂批号不宜更换。

评价的参考物质一般包括标准物质、校准品、室间质评样本等。如参考物质浓度较高, 需对其进行稀释以得到合适浓度水平的样本, 通过稀释制备得到样本都进行互通性评价。

4.2 临床样本

4.2.1 单人份临床样本

优先使用不同来源的单人份新鲜临床样本，单人份临床样本可来自健康人或患者，其浓度应大致均匀地分布，并包含待评价的参考物质的浓度水平，样本体积应以满足所有测量程序分析所需为宜，应避免使用含有已知干扰或交叉反应物质的样本。

4.2.2 混合临床样本

由于部分样本来源有限，如新生儿或罕见病相关项目，如不能收集足够体积的单人份临床样本，可选择混合临床样本，尽量避免使用过多单人份样本进行混合。

4.2.3 临床样本的添加

应尽量避免在临床样本中添加纯品待测物或其他添加剂，以避免影响样本的基质。实际操作中，在无法收集到高浓度的样本而必须进行添加时，应标记添加过的样本，并说明这些样本对互通性结果判定可能造成的影响。

4.2.4 临床样本的浓度和数量

临床样本的浓度水平应均匀分布在测量程序范围内，并涵盖参考物质的浓度范围。应选择具有代表性的临床样本（如健康人和患者的样本），避免使用含有已知干扰物的样品。若明确冰冻样本不影响测定，亦可采用新鲜冰冻样本。

样本数量可根据测量程序的性能特征而有所不同，应 ≥ 20 个不同浓度水平的样本。

注：为便于实验，可根据实验批次计算样本份数，将实验样本分装后进行保存。

4.2.5 样本稳定性

需考察临床样本中被测量的稳定性，以满足实验要求；如样本需要冰冻储存，应避免反复冻融。

5 测量程序及实验方法

5.1 测量程序的选择

用于互通性评价的测量程序应具备如下特性：a) 足够精密，精密度不佳会带来较大的不确定度，影响互通性的评价；b) 足够特异，使用特异性不足的测量程序可能会得到错误的互通性评价结果。若实验过程中使用了参考测量程序，则参考测量程序为比对程序，常规测量程序为实验程序；若不含参考测量程序，则将常规测量程序进行两两比较。

注：理想的比对程序应为无基质效应的参考测量程序或指定比对程序。在实际工作中也可选用常规测量程序作为比对程序，但是在这种情况下，比较难以判断基质效应是来自于比对程序拟或待评价程序。

5.2 评价方法

5.2.1 分别使用实验程序与比对程序测定所有样本，重复测定 ≥ 3 批，实验批内每个样品测定1次，每批测定都需对项目进行重新校准。临床样本数量 ≥ 20 个，参考物质随机穿插在临床样本中，同批次测定。

5.2.2 实验完成后，将样本在适宜条件下保存以备再次测定。如在数据分析过程中发现问题，有必要选用其它比对程序（如参考测量程序）对样本进行重新测定。

注:参考物质可能比临床样更均匀,因此它们的重复性可能更好。如果发现以上情况且可量化,宜适当增加患者样本重复测定的次数,或减少参考物质重复测定次数,以确保两种样本测定结果的方差相似。同样原因,两种程序也需具有不同的重复性,如果程序间的重复性差别已知且可量化,宜适当调整各测量程序的重复次数。

5.2.3 实验程序与比对程序宜同步进行,若不能实现同步测定,应在适宜的条件下储存样本,并提供相关信息,以证明储存条件不会影响测量结果。但应注意储存条件如冰冻,可能导致结合蛋白改变而引入基质效应,对样本互通性产生影响。

6 数据分析

数据分析方法的选择应考虑以下因素:参考物质与临床样本数量、测量程序的数量、是否包含参考测量程序。首先根据数据的分布特征选取合适的回归模型,通常使用戴明(Deming)回归方法、普通最小二乘法(ordinary least square, OLS)、Passing-Bablok (PB)回归法。IFCC互通性评价工作组推荐采用偏倚/偏差差值法来评价参考物质的互通性,需要预先设定一个合理的评价标准,强调该标准应是一个固定且与医学预期用途相关的标准,该评价标准可克服回归分析法的不足,评价案例列于附录D,仅供参考。本指南主要介绍普通最小二乘法和Deming回归评价互通性的方案。

6.1 回归方法的选择

6.1.1 回归分析方法是假定一组变量(X , Y)间具有线性关系并估计线性模型参数和标准差。常用的回归方法有:普通最小二乘法(OLS)、Deming回归法和PB(Passing-Bablok)回归法。回归方法的选择和正确使用在很大程度上取决于数据的特性,即 X 和 Y 的标准差以及参考物质和临床样本的重复检测次数。在建立数学模型前,应确定数据是否符合正态分布,符合正态分布或经过对数转换后符合正态分布,采用OLS和Deming回归,不符合正态分布的数据采用PB回归进行分析。

6.1.2 OLS和Deming回归:

均采用最小二乘法拟合,但两者对数据的假设不同。OLS回归假设变量 Y 服从标准差恒定的正态分布,自变量 X 无随机误差,主要适用于比对程序为参考测量程序的情况。

Deming回归假定两个测量程序所得的结果都含有随机误差,根据观察点到回归线垂直距离的平方和最小来估计回归模型斜率和截距,适用于两种常规测量程序间的互通性评价。Deming回归在参数估计时更为严谨,估计的参数与实际情况的一致性更好,在互通性研究中的适用范围更广。

a) 当 X 和 Y 数据标准差在测量范围内都近似恒定,可以对测量结果的均值使用常规的Deming回归方法。

b) 若标准差与浓度成比例变化(变异系数恒定)时,推荐使用加权Deming回归分析或者将结果进行对数转换后使用非加权Deming回归分析。

6.1.3 PB回归:

是一种非参数程序,对误差分布类型没有规定。PB回归可用于标准差恒定或变异系数恒定的情况,唯一的限制是假设标准差的比例等于斜率的平方。PB回归使用任意两个数据点之间的斜率来估计回归线

的斜率；对截距进行估计时应使一半的数据点落在回归线的上方，一半落在下方。PB回归稳健性好，离群值对结果的影响小，但会造成的较宽的预测区间。

6.1.4 多项式回归：

采用合适的统计分析软件（如SPSS）进行多项式回归分析，以比对方法测定每个临床样本的均值为自变量（ x 轴），测量方法测定每个样本的均值为因变量（ y 轴），求出最佳拟合回归方程，以期得到最小的预测区间（见下文），提高检出基质效应的能力。若最佳拟合为二项式，则回归方程为： $y = a_0 + a_1x + a_2x^2$ ，对该二项式回归方程中的回归系数 a_2 进行统计分析，若 a_2 与0有显著性差异（如 t 检验结果 $p < 0.05$ ），则采用此二项式回归模式；若 a_2 与0无显著性差异（如 $p > 0.05$ ），则使用线性回归分析。

注1：若两种方法测定结果之间不呈线性，很可能是因为20个临床样本的浓度范围分布较窄，建议增加样本以得到更宽的 x 值范围，不能用20个样本的数据来查找非线性的原因。

6.2 数据分布的线性评价

利用临床样本及参考物质重复测定结果的均值（使用不同符号）作散点图， y 轴为实验程序结果， x 轴为比对程序结果，采用OLS回归进行初步的线性回归分析。通过计算回归线的截距和斜率（回归线不一定要通过原点）进行统计学分析，检查数据是否适合进行线性回归分析，具体可参见WS/T XXX-XXXX《定量检验程序分析性能验证指南》。

如果数据适合线性回归，检查是否存在异常值。从散点图中可明显观察到离群值，需采用统计方法进行离群值判断，可采用附录B中描述的离群值检测方法。发现潜在异常值后，需剔除并记录原因（例如，仪器或操作人员失误等），并补充样本进行数据替换。

6.3 OLS 回归分析

6.3.1 比对程序的要求和选择

使用本方法时，要求比对程序为参考测量程序，对于参考物质没有或只有极轻微的基质效应，具有极小的随机误差。比对程序选择顺序如下：

- a) 一级参考测量程序，如同位素稀释质谱法测定血清胆固醇；
- b) 二级参考测量程序，如美国CDC改良的Abell-Kendall法测定胆固醇；
- c) 指定的比对方法，如美国CDC的HDL-C指定比对方法；

6.3.2 数据分析

a) 利用临床样本及制备样本重复测定结果的均值（使用不同符号）作散点图， y 轴为待评价程序结果， x 轴为比对程序结果。

b) 线性回归分析

——目视，评估方法和比对方法测定结果呈线性关系，无明显弯曲；在实验浓度范围内，临床样本的评估方法测定值（回归线的 y 轴）呈均匀分布。

——检查数据是否适合回归分析（参考最新版CLSI/NCCLS文件EP6-定量测定方法的线性评估：统计方法）。

——将评估方法测定临床样品结果的均值作为 y 值，比对方法测定临床样品的均值作为 x 值，进行线性回归分析（见附录A）。

c) 用以下公式计算给定 x 值下（重复测量均值），新鲜临床样品预测值 \hat{y} 的双侧95%置信区间。

$$\hat{y} \pm t(0.975, n-2)S_{y \cdot x} \sqrt{1 + \frac{1}{n} + \frac{(\bar{x}_i - \bar{x})^2}{\sum(\bar{x}_i - \bar{x})^2}}$$

$$S_{y \cdot x} = \sqrt{\sum(\hat{y} - \bar{y}_i)^2 / (n - g)}$$

\hat{y} : 用回归方程计算各已知浓度 (x) 下的预测值;

n : 临床样本的数量;

g : 常数项, 线性回归时为 2, 二次回归时为 3;

$S_{y \cdot x}$: 直线回归的剩余标准差;

\bar{x}_i : x 轴上第 i 个值 (比对程序测定某样本的均值);

\bar{y}_i : y 轴上第 i 个值 (待评价程序测定某样本的均值);

\bar{x} : 所有样本比对方法测定均值的整体均值;

对于参考物质, 利用以上公式, 计算出比对程序测定均值 (x), 对应的预测值 (\hat{y}) 的 95% 置信区间, 如果待评价程序的测定均值 (y) 落在该区间内, 说明该参考物质在两种程序间具有互通性。需注意的是, 若两种测定方法间的特异性差异较大, 测定结果间的相关性将较受到影响, 从而导致计算出来的置信区间偏大, 以至于无法检出不太显著的基质效应, 影响最终的结论。

d) 在散点图上, 可将一系列临床样本的比对方法测定均值 (x) 与对应的 y 值的 95% 预测区间在回归线两边标记出来 (见附录 A)。若制备样品的点落在预测区间线条之外则说明存在基质效应 (见附录 B)。若制备样品的点在置信区间内, 则表明不存在基质效应 (见附录 A)。

注: 若回归曲线附近的数据点分布和待测物浓度的关系既非比例关系亦非恒定关系, 宜将数据分成几组较小的浓度区间, 在每个区间内单独进行线性回归分析。区间组的划分标准为: 每组内新鲜临床样品结果的分布保持大体一致, 每组应至少含有 10 份新鲜临床样品的数据, 且包含制备样品。

6.4 Deming 回归分析

6.4.1 Deming 回归流程图 (图 1):

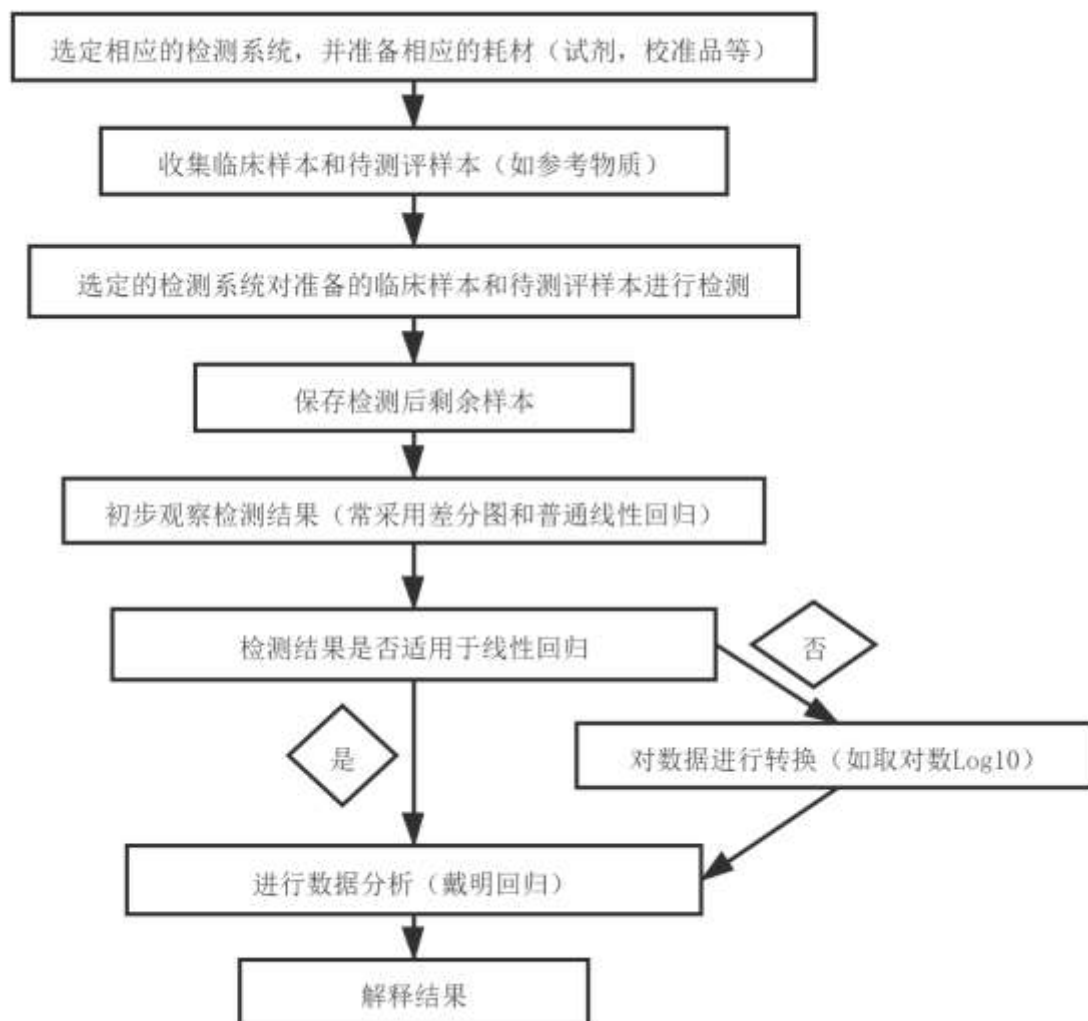


图 1 Deming 回归评估待测评样本互通性流程图

6.4.2 绘制偏差图：

通过偏差图检查两测量程序（X、Y）获得的临床样品结果的均值分布，偏差图的X轴为两测量程序所得结果的均值（若测量程序包括参考测量程序则X轴为参考测量程序的均值），Y轴为两测量程序所得结果的差值，如图2。

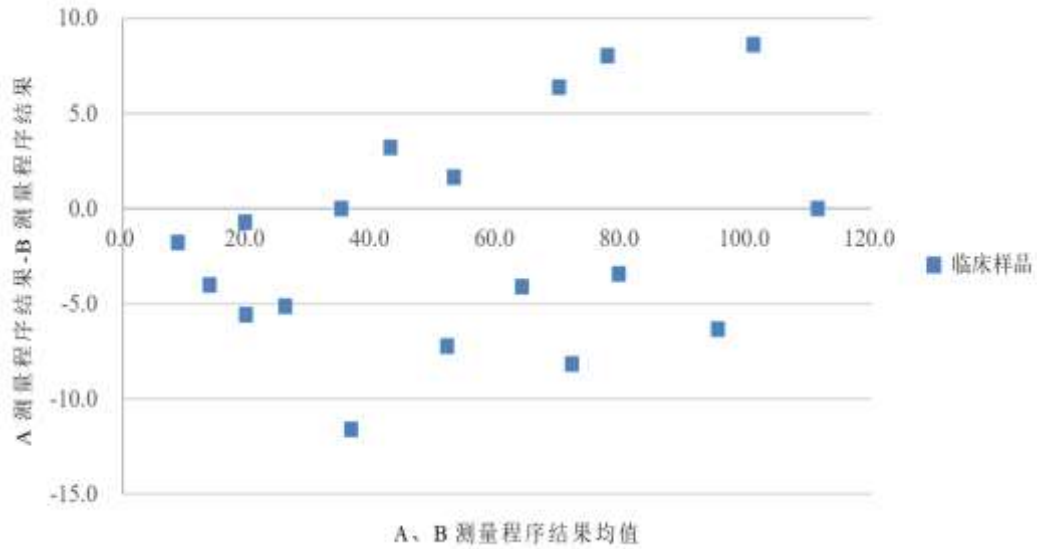


图 2 测量程序的偏差图

6.4.3 数据转化:

若偏差/偏倚是恒定的,则可直接进行Deming回归。若偏差/偏倚随着被测量浓度的增加而增加,则需将数据进行 \log_{10} 转换,若转换后的偏差/偏倚恒定则将转换后的数据进行Deming回归。

6.4.4 Deming 回归及互通性评估:

以两测量程序测得临床样本结果的均值(或 \log_{10} 均值)为相应的X、Y轴,并绘制Deming回归图。计算临床样品对于相应的X值,其Y值的预测值及双侧95%的预测区间,并绘制在回归图上,如图3。临床样品和参考物质采用不同的符号绘制在同一图表上。若参考物质的检测结果在区间内,则认为参考物质是互通的,若在区间外,则认为参考物质不互通。

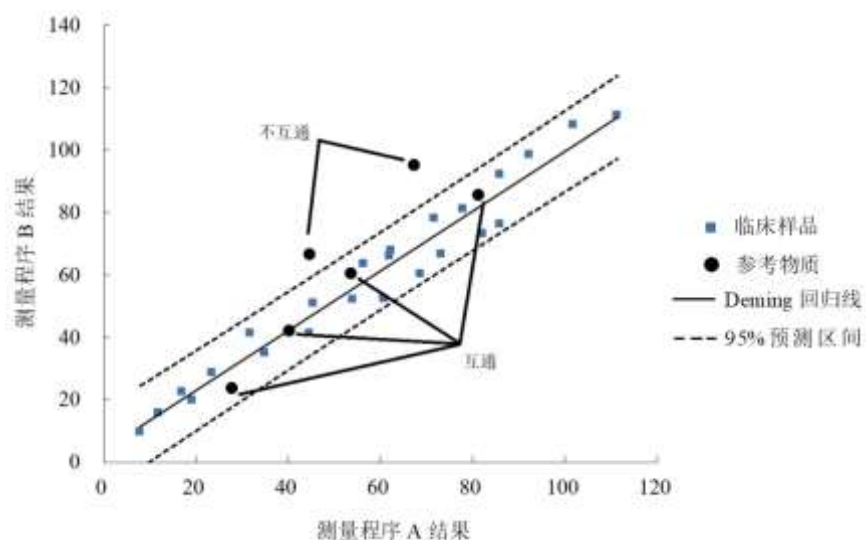


图 3 Deming 回归方法评估参考物质互通性的评价结果

6.4.5 计算过程:

采用两种不同的检测系统 X、Y 检测临床样本和待测评样本时，往往认为两种检测系统的结果之间存在线性关系，临床样本在两系统上的结果之间的关系可以表示为：

$$Y = \alpha_H + \beta_H X \quad (1a)$$

$$Y = \alpha_{PC} + \beta_{PC} X \quad (1b)$$

α_H , α_{PC} ：分别为临床样本和参考物质在两个检测系统上线性关系的截距

β_H , β_{PC} ：分别为临床样本和参考物质在两个检测系统上线性关系的斜率

对 X、Y 两检测系统结果做 Deming 回归的详细计算过程如下：

$$\hat{\beta}_H = \frac{\hat{\sigma}_Y^2 - \lambda \hat{\sigma}_X^2 + \sqrt{(\hat{\sigma}_Y^2 - \lambda \hat{\sigma}_X^2)^2 + 4\lambda \hat{\sigma}_{XY}^2}}{2\hat{\sigma}_{XY}} \quad (2)$$

$$\hat{\alpha}_H = \bar{Y} - \hat{\beta}_H \bar{X} \quad (3)$$

$$\hat{\sigma}_X^2 = \frac{1}{n} \sum_{i=1}^n (\bar{X}_i - \bar{X})^2 \quad (4)$$

$$\hat{\sigma}_Y^2 = \frac{1}{n} \sum_{i=1}^n (\bar{Y}_i - \bar{Y})^2 \quad (5)$$

$$\hat{\sigma}_{XY} = \frac{1}{n} \sum_{i=1}^n (\bar{X}_i - \bar{X})(\bar{Y}_i - \bar{Y}) \quad (6)$$

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n \bar{X}_i \quad (7)$$

$$\bar{Y} = \frac{1}{n} \sum_{i=1}^n \bar{Y}_i \quad (8)$$

$$\lambda = \hat{\sigma}^2(\varepsilon_Y) / \hat{\sigma}^2(\varepsilon_X) \quad (9)$$

n：临床样本数量

\bar{X} , \bar{Y} ：分为临床样本在 X、Y 检测系统上的总均值，及所有临床样本的检测均值

\bar{Y}_i , \bar{X}_i ，别为第 i 个临床样本重复检测的均值

$$\hat{\sigma}^2(\varepsilon_X) = \frac{1}{n(N_H - 1)} \sum_{i=1}^n \sum_{j=1}^{N_H} (X_{ij} - \bar{X}_i)^2 \quad (10)$$

$$\hat{\sigma}^2(\varepsilon_Y) = \frac{1}{n(N_H - 1)} \sum_{i=1}^n \sum_{k=1}^{N_H} (Y_{ik} - \bar{Y}_i)^2 \quad (11)$$

N_H ：为每个样本重复检测的次数

j：在 X 检测系统上样本重复检测次数

k：在 Y 检测系统上样本的重复检测次数

从以上公式便可以计算出 $\hat{\alpha}_H$, $\hat{\beta}_H$ 从而确定回归线方程 $Y = \hat{\alpha}_H + \hat{\beta}_H X$ ，将该方程绘制为一条直线，并计算该直线 95% 置信度的预测区间，95% 置信度的区间计算方法如下：

$$[L, U] = \bar{Y}_{PC_pred} \pm t(1 - \frac{\gamma}{2}, n(N_H - 1)) * \hat{\sigma}(\bar{Y}_{PC_pred}) \quad (10)$$

$$\bar{Y}_{PC_pred} = \hat{\alpha}_H + \hat{\beta}_H \bar{X}$$

$$\hat{\sigma}(\bar{Y}_{PC_pred}) = \sqrt{(\bar{X}_{PC} - \bar{X})^2 \hat{\sigma}_{\beta_H}^2 + \frac{1}{N_{PC}} (\hat{\beta}_H^2 \hat{\sigma}^2(\varepsilon_X) + \hat{\sigma}^2(\varepsilon_Y)) (1 + \frac{1}{n})} \quad (11)$$

$$\hat{\sigma}_{\beta_H}^2 = \frac{\hat{\beta}_H^2}{n \hat{\sigma}_{XY}^2} (\hat{\sigma}_X^2 \hat{\sigma}_Y^2 - \hat{\sigma}_{XY}^2) \quad (12)$$

[L, U]：分别为检测区间的下限和上限： γ 为所选置信度，如计算 95% 的预测区间则 $\gamma=0.05$ ，如计算 99% 的预测区间 $\gamma=0.01$ 。通常将该预测区间的上限和下限在 Deming 回归图上分别做一条直线，形成一

个区间，通过待测物质在 X、Y 检测系统上的结果在相同的坐标系点出来，如果检测结果落在该区间内则认为互通，如果超过该区间则认为不互通（图 3），计算案例见附录 B。

附录 A
资料性附录
数据分析-肌酐测定互通性-OLS 回归

A.1 利用参考方法作为比对程序，某酶法试剂盒作为待评价程序，评价 6 个参考物质在两个程序间的互通性。本实验测定 20 个临床样 (S1~S20) 本，5 个参考物质 (P1~P5)，每个样本测定 3 次，各样本的均值如下表。

表 A.1 各样本测定结果

样本	比对程序测定均值(\bar{x}_i)	评价程序测定均值(\bar{y}_i)
S1	129.3	116.3
S2	881.9	820
S3	163.3	147.7
S4	886.4	827.3
S5	298.3	269
S6	112.5	103.3
S7	322.9	292.7
S8	293.3	263.7
S9	599.9	541.5
S10	730.6	658.5
S11	250.3	230
S12	516.7	466
S13	160.9	149
S14	70.8	63.5
S15	43.7	39
S16	343.5	307.5
S17	230.8	215
S18	366.2	331
S19	328.7	302.5
S20	340.9	312
P1	200	220
P2	80	70
P3	283	285
P4	390	330
P5	515.2	520

A.2 用以下公式计算给定 x 值下（重复测量均值），新鲜临床样品预测值 \hat{y} 的两侧 95%置信区间。

$$\hat{y} \pm t(0.975, n-2)S_{y \cdot x} \sqrt{1 + \frac{1}{n} + \frac{(\bar{x}_i - \bar{x})^2}{\sum(\bar{x}_i - \bar{x})^2}}$$

$$S_{y \cdot x} = \sqrt{\sum(\hat{y} - \bar{y}_i)^2 / (n - g)}$$

\hat{y} : 用回归方程计算各已知浓度 (x) 下的预测值;

n: 临床样本的数量;

g : 常数项, 线性回归时为2, 二次回归时为3;

$S_{y \cdot x}$: 直线回归的剩余标准差;

\bar{x}_i : x 轴上第 i 个值(比对程序测定某样本的均值);

\bar{y}_i : y 轴上第 i 个值(待评价程序测定某样本的均值);

\bar{x} : 所有样本比对方法测定均值的整体均值, 本实验为353.55

A.3 计算结果汇总

附表A.2 计算结果汇总表

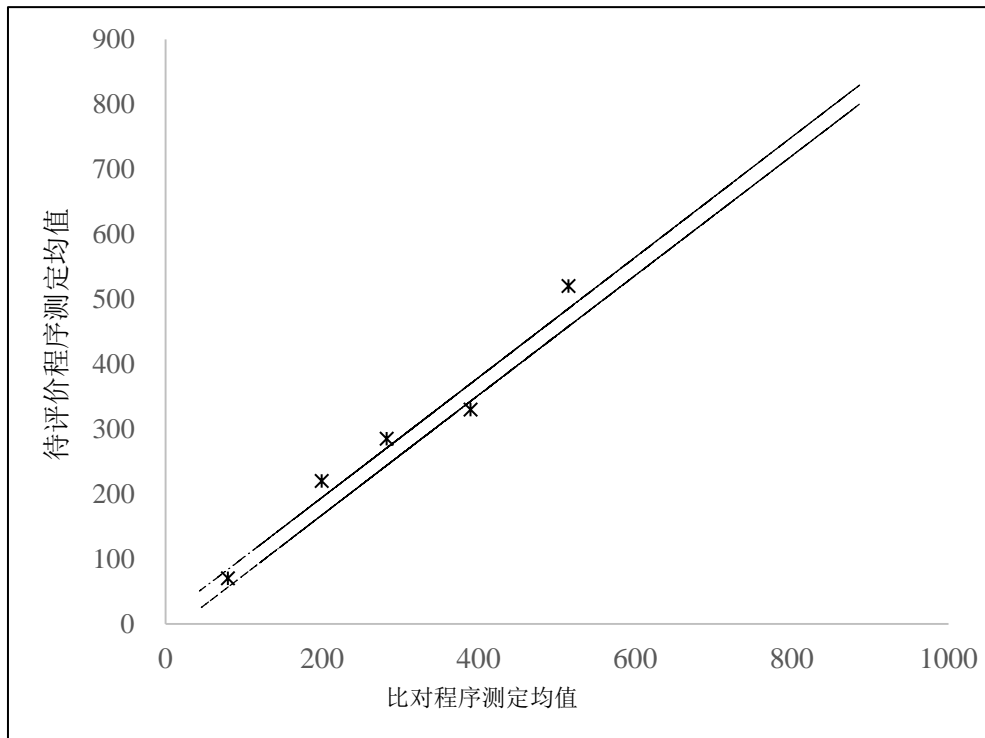
样品	比对程序		待评价程序				
	\bar{x}_i	$(\bar{x}_i - \bar{x})^2$	\bar{y}_i	\hat{y}	$(\hat{y} - \bar{y}_i)^2$	95%上限	95%下限
S1	129.3	50286	116.3	115.7	0.33	129.2	102.2
S2	881.9	279159	820	810.6	88.05	825.3	796.0
S3	163.3	36193	147.7	147.1	0.34	160.6	133.7
S4	886.4	283934	827.3	814.8	156.96	829.5	800.1
S5	298.3	3052	269	271.8	7.65	285.0	258.5
S6	112.5	58103	103.3	100.2	9.53	113.8	86.7
S7	322.9	939	292.7	294.5	3.17	307.7	281.2
S8	293.3	3629	263.7	267.1	11.90	280.4	253.9
S9	599.9	60691	541.5	550.2	76.39	563.8	536.7
S10	730.6	142170	658.5	670.9	154.21	684.9	656.9
S11	250.3	10660	230	227.4	6.52	240.8	214.1
S12	516.7	26620	466	473.4	55.05	486.8	460.0
S13	160.9	37112	149	144.9	16.80	158.3	131.5
S14	70.8	79945	63.5	61.7	3.20	75.4	48.0
S15	43.7	96004	39	36.7	5.34	50.4	22.9
S16	343.5	101	307.5	313.5	36.00	326.7	300.3
S17	230.8	15066	215	209.4	30.89	222.8	196.1
S18	366.2	160	331	334.5	11.97	347.7	321.2
S19	328.7	617	302.5	299.8	7.10	313.1	286.6
S20	340.9	160	312	311.1	0.81	324.3	297.9
P1	200	23576	220	181	1521.00	194.4	167.6
P2	80	74827	70	70.2	0.04	83.8	56.6
P3	283	4977	285	257.6	750.8	270.9	244.3
P4	390	1329	330	356.4	697.0	369.7	343.1
P5	515.2	26132	520	472.1	2294.4	485.5	458.7

A.4 参考物质互通性结果判断

表A.3 参考物质互通性结果判断

样本	\bar{x}_i	\bar{y}_i	\hat{y}	95%上限	95%下限	判断
P1	200.0	220.0	181.0	194.4	167.6	正基质效应
P2	80.0	70.0	70.2	83.8	56.6	无基质效应
P3	283.0	285.0	257.6	270.9	244.3	正基质效应

P4	390.0	330.0	356.4	369.7	343.1	负基质效应
P5	515.2	520.0	472.1	485.5	458.7	正基质效应



图A.1 两个方法测定结果散点图

图中虚线为其预测值Y的95%置信区间；“*”代表制备物，凡是落在虚线范围外的“*”，均判断为有基质效应。

附录 B

(资料性附录)

数据分析-胆固醇两种测定方法互通性评估-Deming 回归

B.1 X、Y检测系统的对临床样本和待测评样本的原始检测值

样本编号	X 检测系统原始结果			Y 检测系统原始结果		
	RUN 1	RUN 2	RUN 3	RUN 1	RUN 2	RUN 3
S1	206.27	213.13	204.53	217.43	208.78	219.86
S2	143.71	146.45	144.76	161.57	161.95	163.32
S3	118.59	117.81	126.21	127.43	133.14	129.01
S4	224.56	231.18	222.17	246.47	236.57	245.42
S5	249.09	248.05	247.22	266.61	274.72	261.04
S6	206.03	205.6	202.75	217.59	215.63	222.02
S7	220.74	224.57	217.09	236.31	241.65	240.33
S8	175.12	173.99	181.63	182.33	191.41	195.36
S9	242.66	245.04	245.72	269.62	272.5	265.39
S10	162.58	158.43	162.55	177.67	169.33	173.15
S11	131.18	137.38	128.5	137.68	131.66	144.67
S12	242.84	247.71	250.34	264.58	270.42	268.85
S13	133.99	140.23	138.34	155.64	153.43	150.03
S14	226.87	209.17	216.33	238.32	241.54	242.92
S15	259.22	257.44	254.87	285.09	283.05	282.53
S16	180.03	178.13	182.35	188.97	200.60	196.30
S17	99.39	106.08	103.99	110.16	108.02	111.61
S18	276.38	279.19	276.34	284.93	298.46	304.16
S19	126.17	132.06	123.88	135.89	138.1	138.66
S20	181.95	173.08	181.81	192.45	201.37	193.65
Ps 1	107.17	106.39	113.02	117.60	107.32	105.53
Ps 2	122.86	124.04	122.47	131.54	140.24	139.05
Ps 3	150.00	143.63	147.22	182.82	170.44	172.81
Ps 4	198.39	200.01	196.92	215.74	221.86	210.35
Ps 5	221.84	221.94	223.17	250.75	241.86	239.59

根据上表数据可以得出：

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n \bar{x}_i = 191$$

$$\bar{Y} = \frac{1}{n} \sum_{i=1}^n \bar{y}_i = 206$$

$$\hat{\sigma}_{\bar{X}}^2 = \frac{1}{n} \sum_{i=1}^n (\bar{X}_i - \bar{X})^2 = 2541$$

$$\hat{\sigma}_{\bar{Y}}^2 = \frac{1}{n} \sum_{i=1}^n (\bar{Y}_i - \bar{Y})^2 = 3011$$

$$\sigma_{\bar{X}\bar{Y}} = \frac{1}{n} \sum_{i=1}^n (\bar{X}_i - \bar{X})(\bar{Y}_i - \bar{Y}) = 2760$$

$$\hat{\sigma}^2(\varepsilon_X) = \frac{1}{n(N_H - 1)} \sum_{i=1}^n \sum_{j=1}^{N_H} (X_{ij} - \bar{X}_i)^2 = 15.06$$

$$\hat{\sigma}^2(\varepsilon_Y) = \frac{1}{n(N_H - 1)} \sum_{i=1}^n \sum_{k=1}^{N_H} (Y_{ik} - \bar{Y}_i)^2 = 22.08$$

$$\text{所以: } \hat{\lambda} = \frac{\hat{\sigma}^2(\varepsilon_Y)}{\hat{\sigma}^2(\varepsilon_X)} = 1.466$$

$$\text{又得: } \hat{\beta}_H = \frac{\hat{\sigma}_Y^2 - \hat{\lambda} \hat{\sigma}_X^2 + \sqrt{(\hat{\sigma}_Y^2 - \hat{\lambda} \hat{\sigma}_X^2)^2 + 4 \hat{\lambda} \hat{\sigma}_{\bar{X}\bar{Y}}^2}}{2 \hat{\sigma}_{\bar{X}\bar{Y}}} = 1.088$$

$$\text{所以: } \hat{\alpha}_H = \bar{Y} - \hat{\beta}_H \bar{X} = -1.282$$

又因为:

$$\hat{\sigma}_{\beta_H}^2 = \frac{\hat{\beta}_H^2}{n \hat{\sigma}_{\bar{X}\bar{Y}}^2} (\hat{\sigma}_{\bar{X}}^2 \hat{\sigma}_{\bar{Y}}^2 - \hat{\sigma}_{\bar{X}\bar{Y}}^2) = 2.9 \times 10^{-4}$$

$$\text{所以: } \text{当 } \bar{X}_{PC} = 145.0, \bar{Y}_{PC_{pred}} = 156.5$$

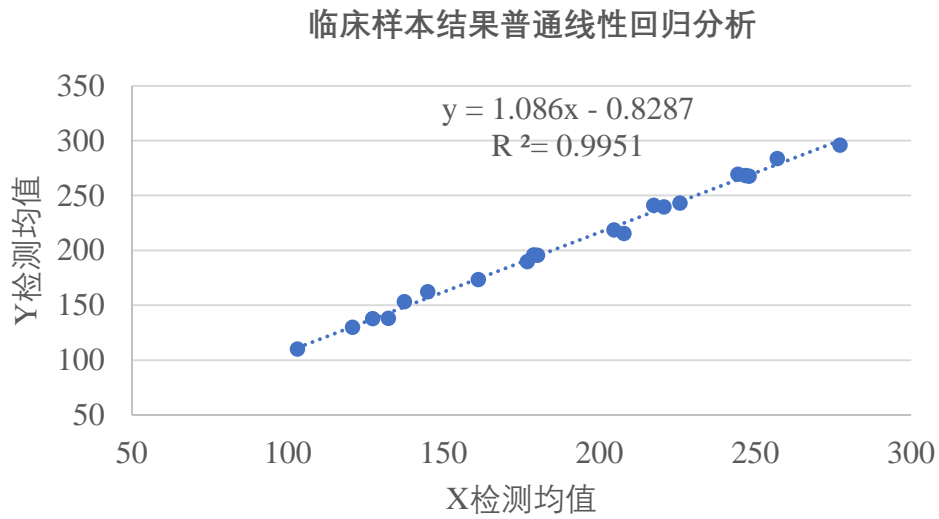
$$\hat{\sigma}(\bar{Y}_{PC_{pred}}) = \sqrt{(\bar{X}_{PC} - \bar{X})^2 \hat{\sigma}_{\beta_H}^2 + \frac{1}{N_{PC}} (\hat{\beta}_H^2 \hat{\sigma}^2(\varepsilon_X) + \hat{\sigma}^2(\varepsilon_Y)) (1 + \frac{1}{n})} = 3.82$$

所以: 当 $\gamma = 0.05$ 时

$$[L, U] = \bar{Y}_{PC_{pred}} \pm t\left(1 - \frac{\gamma}{2}, n(N_H - 1)\right) * \hat{\sigma}(\bar{Y}_{PC_{pred}}) = 148.8 - 164.2$$

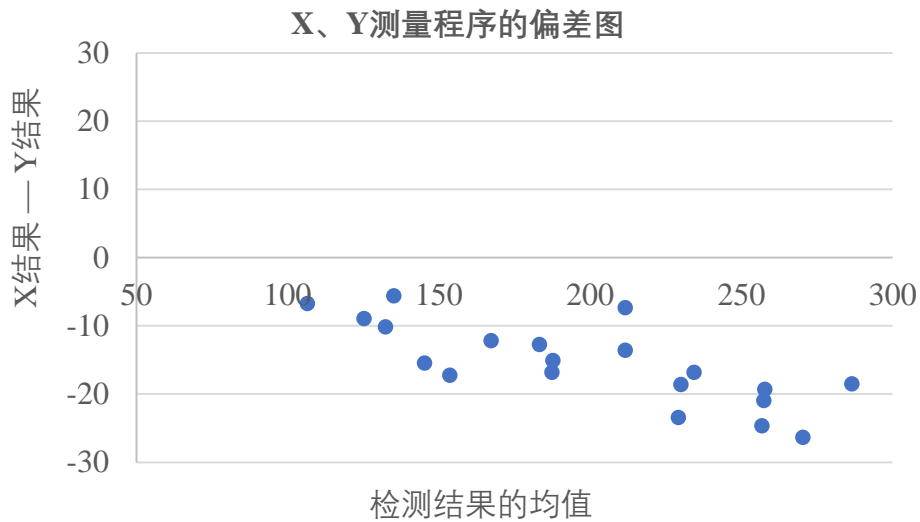
由以上计算可以得出Deming回归方程: $Y = \hat{\alpha}_H + \hat{\beta}_H X$, 并计算出其95%的预测区间, 当参考物质落在这个区间内, 则认为互通, 反之则认为不互通(图B.3) 其中Ps 3号样本在两种系统上不具有互通性, 呈现出正基质效应。

B.2 对两个测量程序的均值绘制散点图，评价是否适合线性回归：



图B.1 X、Y检测系统结果间普通线性回归结果

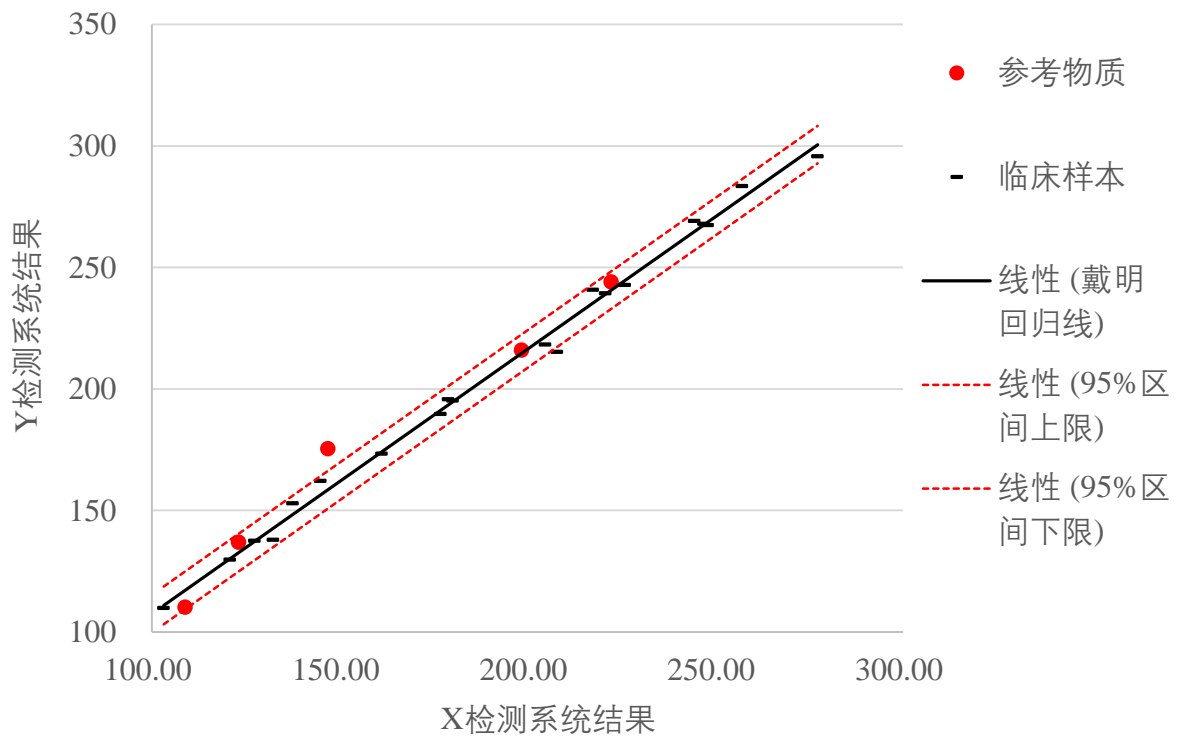
B.3 对适合线性回归的数据绘制偏差图：



图B.2 X、Y检测系统结果间偏差图

B.4 对数据进行计算，绘制互通性评价图：

待测评样本在X、Y检测系统间的互通性



图B. 3 Deming回归互通性评价结果图

附录 C

(资料性附录)

Deming 回归评价互通性的离群值判断方法

对附录B中的数据做以下处理。

1. 收集两个检测系统对于临床样本和待测评样本的检测结果，其中每个样本至少重复检测3次，临床样本的数量要多于20例，且需要覆盖待测评物的浓度区间

2. 计算两个系统间对样本重复检测的均值，公式为： $\bar{X}_i = \sum_{j=1}^{N_H} X_{ij}$ ； $\bar{Y}_i = \sum_{k=1}^{N_H} Y_{ik}$ 。并且绘制散点图，看是否适用于回归分析（图4）。

3. 计算检测系统重复检测样本的标准差，并以样本的标准差为纵坐标，样本的浓度为横坐标（通常为两个检测系统的均值）做散点图（图5），观察散点图，确定检测系统的标准差是否和样本浓度相关，如随着浓度增大而增大，减小而减小。如果标准差与样本浓度无关，则可以往下进行戴明回归，如果不是则需要对原始检测结果进行数据转换如去对数Log10，直到标准差不具有浓度变化趋势。

4. 剔除数据中的离群值。通过上文中的公式11、12来计算X、Y检测系统的随机误差， $\sigma(\varepsilon_X) = \sqrt{\sigma^2(\varepsilon_X)}$ ， $\sigma(\varepsilon_Y) = \sqrt{\sigma^2(\varepsilon_Y)}$ 。满足以下公式则可以判定为离群值：

$$X_{i, \max} - X_{i, \min} > q \sigma(\varepsilon_X), Y_{i, \max} - Y_{i, \min} > q \sigma(\varepsilon_Y)$$

根据样本数量和每个样本的重复检测次数可以导出在99%置信度对应q值，如下表

样本数量	样本重复检测次数	自由度	99%区间对应的q值
20	3	40	4.37
	4	60	4.59
	5	80	4.78
	6	100	4.91
30	3	60	4.28
	4	90	4.55
	5	120	4.71
	6	150	4.85
40	3	80	4.25
	4	120	4.50
	5	160	4.68
	6	200	4.81

注：按照上述公式通常一组数据里被判定为离群值的概率大概为2%，通常来讲剔除的离群值不能超总体数据的5%，并且剔除利群之后用于分析的临床样本数量不应该少于20例

附录 D

(资料性附录)

数据分析-胆固醇两种测定方法互通性评估-IFCC 方法

D. 1. IFCC 偏倚/偏差差值法

IFCC推荐偏差差值法 (difference in bias, 简称差值法) 来评价物质的互通性, 该方法是将待评价物质系统间偏差与临床样本系统间偏差均值的差值的分布范围与临床允许范围做比较, 来判断参考物质是否具有互通性。此方法的评价标准固定且与临床需求密切相关, 克服了回归分析法的不足。同时该方法提出的误差模型可以估计方法间偏倚的变异来源, 从而更好的解释互通性评价结果。

D. 2. 评估原理

差值法的评价原理源于误差模型理论, 该模型认为不同测量系统对同一样本检测结果的差异主要来源于三方面, 1、方法间常见偏差, 常为恒定常数或者和浓度相关的连续函数, 2、样本特异性产生的偏差, 常指不同方法的特异性不同而造成的对样本中的检测物质和干扰物质反应不一致, 从而造成检测结果的偏差。3、不同方法批内变异组分, 及不同方法的不同精密度在检测结果上造成的偏差, 往往可以通过重复多次测量计算的标准差进行量化。

差值法在评价互通性时, 检测系统两两配对, 基于以上的误差模型理论按照下列公式 (1-3) 计算带测样本在两系统间的相对偏差及其不确定度:

$$d_{RM} = B_{RM} - B_{CS}(1)$$

$$u(d_{RM}) = \sqrt{\frac{(S_x^2 + S_y^2)}{p} + \frac{S_B^2}{n}} (2)$$

$$U(d_{RM}) = k \times u(d_{RM})(3)$$

其中, B_{RM} 为参考物质在两系统间检测结果的偏差, B_{CS} 为临床样本在两系统间检测结果的偏差均值, d_{RM} 为二者的差值; S_x^2 、 S_y^2 分别表示待评价物质在X、Y两个系统的标准差, p 为检测次数, S_B 为临床样本检测结果系统间偏差的标准差, n 为临床样本例数, $u(d_{RM})$ 是 d_{RM} 的合成不确定度; k 为扩展因子, 取 $k=2$, $U(d_{RM})$ 为 d_{RM} 的扩展不确定度。为了评价待测物质的互通性, 需根据指定偏倚 d_{RM} 允许范围, 及指定的评价标准 C , 改标准常根据待测物质的临床应用相关。

当 $d_{RM} \pm U(d_{RM})$ 在 $0 \pm C$ 范围内, 则参考物质具有互通性; 若 $d_{RM} \pm U(d_{RM})$ 超出 $0 \pm C$ 范围, 则参考物质不具有互通性; 若 $d_{RM} \pm U(d_{RM})$ 与 $0 \pm C$ 范围存在重叠, 则参考物质的互通性难以确定。

D. 3. 样本收集和检测流程

D. 3. 1 临床样本

收集 n 个临床样本, $n \geq 30$, 且收集的样本的浓度要均匀分布并覆盖参考物质及相邻浓度范围。其中临床样本可以是单人份样本或者混合样本, 1) 单人份临床样本, 如单人份血清或者血浆等, 是互通性评估中最理想的检测材料。但某些情况下收集足够量的单人份临床样本较难实现, 这时可以选择将多份单人份临床样本混合来满足评估所要求样本量, 在这样的情况下, 样本特异性将会受到影响, 因此, 这时不能评估样本特异性造成的变异进行评估。

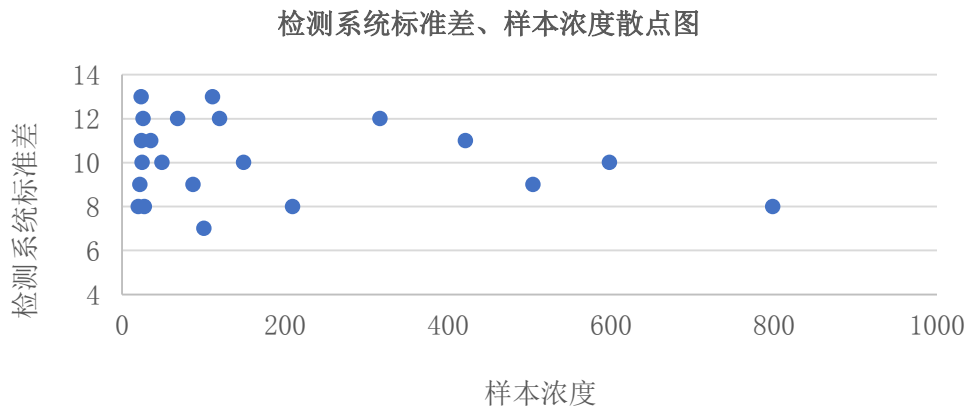
D. 3. 2 检测流程

收集的 n 个临床样本以及相应的带测评物质应在一个批次内测定完成, 测定时临床样本和参考物质随机分布不需要按照浓度进行排序, 每个样本最少重复测定 $k (\geq 2)$ 次, 但考虑到需要计算精密度和剔除离群值, 因此一般进行3次重复检测。如需要考虑位置效应带来的变异, 则需要将带测评物质分成至少 $p (p \geq 5)$ 组, 代表 p 个不同的位置, 穿插在整个临床样本的测定中, 并且在每个位置都重复检测3次

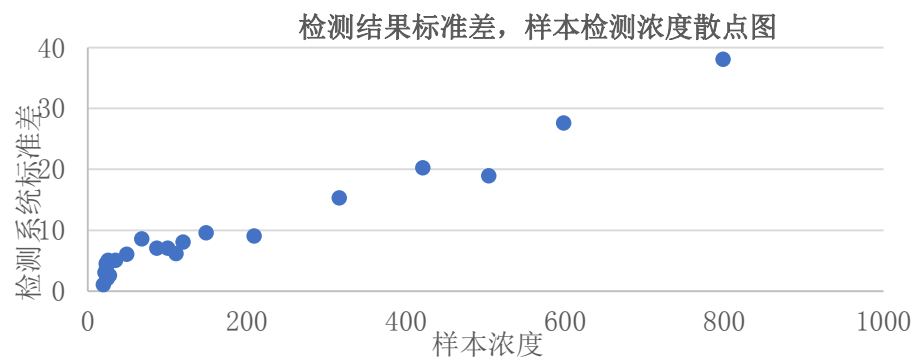
D.4. 数据处理

D.4.1 明确是否需要进行数据转换

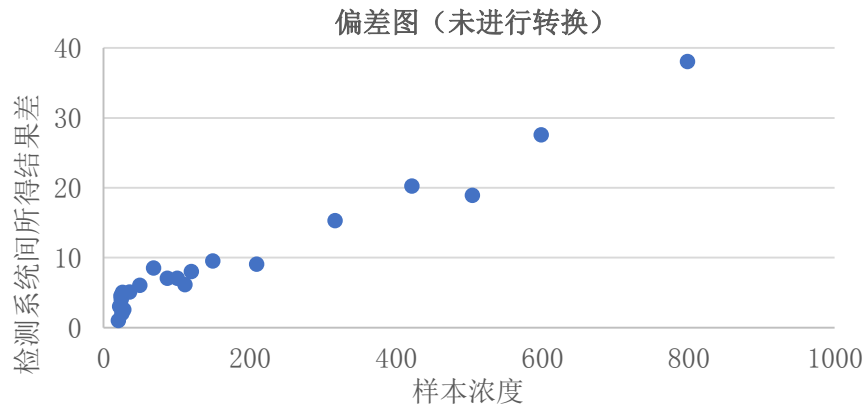
在进行互通性评价之前，需要对检测所得结果进行精密度检查，计算每个临床样本重复测量的标准差，并以样本的标准差为纵坐标，样本浓度为横坐标作图，则通常出现两种情况。如果如图D.1所示，其标准差和浓度之间不存在连续的变化趋势（如标准差随着浓度增高而逐渐增加或随着浓度降低逐渐降低），这时则可以合并所有临床样本的标准差作为方法重复测量标准差的估计值记作 S_x 或 S_y （常指取均值）。如果如图D.2所示，标准差与浓度之间存在明显的正向或负向相关性，则需要对数据进行转换，通常对检测结果取对数 \ln 。正常情况下取对数后的偏差图（纵坐标为临床样本在两个检测系统上检测结果取对数后的差值，横坐标为样本浓度，常为两个检测系统取均值）如图D.4所示，这时取对数后的差值和浓度不具有明显的正相关或者负相关性，这时则对两个检测系统的结果进行 \ln 转换后再进行后续的计算。



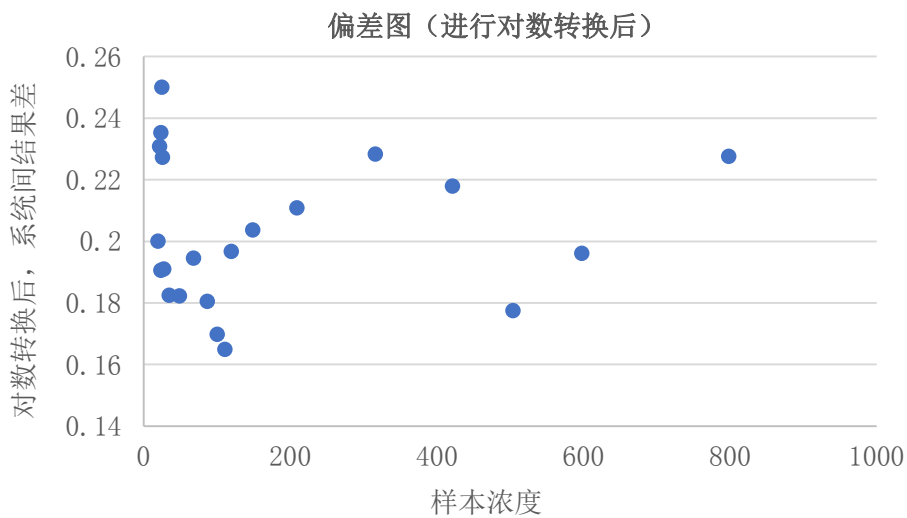
图D.1 检测系统呈现出恒定标准差



图D.2 检测系统精密度成浓度趋势性变化



图D.3 取对数前的偏差图



图D.4 取对数后的偏差图

D.4.2 评价浓度趋势性变异

差值法进行互通性评价时,使用临床样本的在不同方法间的偏差的均值来代表两方法间临床样本的常见偏差,该平均偏差能准确的代表临床样本的真实常见偏差的前提是临床样本间的常见偏差无浓度趋势性变异。因此在进行互通性评价之前,要先对临床样本浓度趋势性变异进行评估和量化。评估方法如下:

$$B_i = y_i - x_i \quad (4)$$

$$S_{MSSD} = \sqrt{\frac{1}{2(n-1)} \sum_{i=1}^{n-1} (B_{i+1} - B_i)^2} \quad (5)$$

$$S_B = \sqrt{\frac{1}{(n-1)} \sum_{i=1}^{n-1} (B_i - B_{CS})^2} \quad (6)$$

y_i : 为临床样本i方法y的均值

x_i : 为临床样本i方法x的均值

n : 为临床样本数量

B_i : 为临床样本i方法间的偏差

B_{CS} : 为所有临床样本方法间偏差的均值

S_{MSSD} 是基于均方递差的变异估计值，该值排除了浓度趋势性变异组分。 S_B 是临床样本在方法房间偏差的变异，其包含了浓度趋势性变异组分。因此如果 S_{MSSD} 和 S_B 之间有较大差异则说明存在明显的浓度趋势性变异，反之则不存在。

D.4.3 数据统计分析进行互通性评价

根据临床样本有无浓度趋势性变异，通常分为三种情况来计算 $u(d_{RM})$

- a、临床样本在整个浓度区间不存在明显的浓度趋势性变异，且覆盖了参考物质的浓度，这种情况下用如下公式评估互通性结果

$$u(d_{RM}) = \sqrt{\frac{S_x^2 + S_y^2}{p} + \frac{S_B^2}{n}} \quad (7)$$

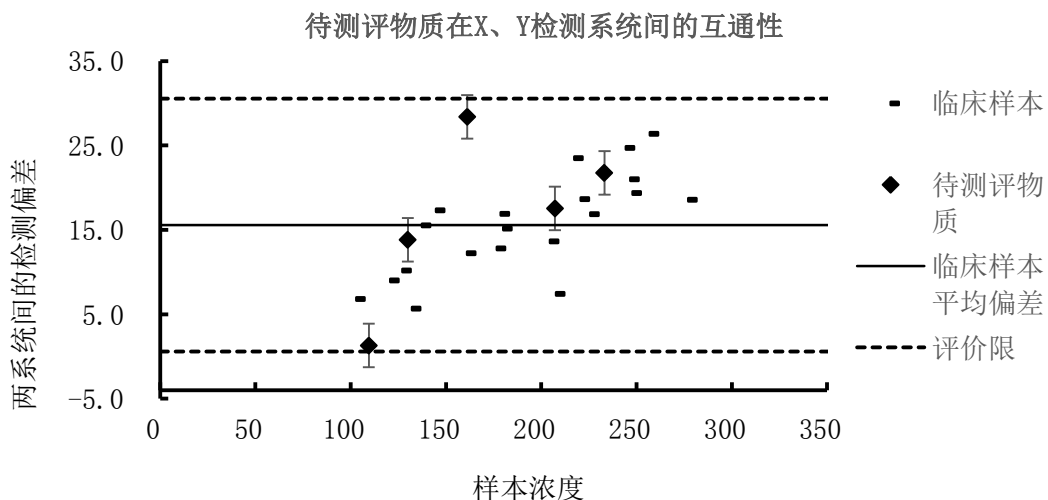
- b、在包含参考物质的浓度区间近似没有浓度趋势性变异，在整个浓度区间存在浓度趋势性变异，这时需要确定一个包含参考物质浓度的区间，这个区间不存在明显的浓度趋势性变异，这个区间的样本数量为 q ， q 的数量必须足够，这时采用如下公式评估互通性。

$$u(d_{RM}) = \sqrt{\frac{S_x^2 + S_y^2}{p} + \frac{S_B^2}{q}} \quad (8)$$

- c、整个区间存在浓度趋势性变异，且无法找到一个包含参考物质的区间不存在明显的浓度趋势性变异时，通过选取参考物质浓度前后相等数量的临床样本来形成一个区间（共 q 个样本，前后各 $q/2$ 个），这个区间相对于a、b两种情况所用的区间通常会小一些，所以其浓度趋势性变异的会更小，但 q 必须为偶数，且 $q \geq 6$ ，再次计算其浓度趋势性变异的大小，如不存在明显的浓度趋势性变异，按照如下公式评估互通性。

$$(d_{RM}) = \sqrt{\frac{S_x^2 + S_y^2}{p} + \frac{S_{MSSD}^2}{q}}$$

如果选取了最小的浓度区间仍存在明显的浓度趋势变异，则无法进行互通性评价。利用IFCC方法计算附录C中的数据可以得到如下结果（图D.5），可以看到Ps 1、Ps 3号样本的互通性难以确认。



图D.5 IFCC方法评价互通性结果图

注：IFCC方法要求确认原始结果是否需要转换，并且最后通过确认这些样本浓度趋势变异的大小从而选取不同的计算方法来评估互通性，本案例旨在提供一个简单的计算结果，及假设不需要进行数据转换和样本没有浓度趋势性变异时能得到的结果。

参考文献

- [1] EP30-A. Characterization and qualification of commutable reference materials for laboratory medicine; Approved guideline[S]. CLSI, 2010.
- [2] EPI4-A3. Evaluation of commutability of processed samples, Approved guideline—Third edition[S]. CLSI, 2014.
- [3] Greg M W, Heinz S, Robert R, et al. IFCC Working Group Recommendations for Assessing Commutability Part 1: General Experimental Design[J]. Clinical Chemistry, 2018,64(3):447-454.
- [4] Nilsson Göran, Budd J R, Neil G, et al. IFCC Working Group Recommendations for Assessing Commutability Part 2: Using the Difference in Bias between a Reference Material and Clinical Samples[J]. Clinical Chemistry, 2018, 64(3):455-464.
- [5] IFCC Working Group Recommendations for Assessing Commutability Part 3: Using the Calibration Effectiveness of a Reference Material[J]. Clinical Chemistry: Journal of the American Association for Clinical Chemists, 2018.
- [6] Fasce C F, Robert R, Copeland W H, et al. A Discussion of Enzyme Reference Materials: Applications and Specifications[J]. Clinical Chemistry, 1973,19(1):5-9.
- [7] 陈文祥, 申子瑜, 郭健, 等. 临床检验量值溯源中的重要术语与概念及有关问题.中华检验医学杂志, 2005, 28(2):142-146.
- [8] Federica, Braga, Mauro, et al. Commutability of reference and control materials: an essential factor for assuring the quality of measurements in Laboratory Medicine.[J]. Clinical chemistry and laboratory medicine, 2019,57(7):967-973.
- [9] NCCLS. Preparation and validation of commutable frozen human serum pools as secondary reference materials for cholesterol measurement procedures ; Approved guideline[J]. NCCLS document, 1999.
- [10] 周伟燕, 刘庆香, 刘珍妮, 邓宇航, 张传宝. 计量溯源性和临床检验参考系统.临床检验杂志, 2020, 38(10): 721-728.
- [11] 刘庆香, 周伟燕, 张天娇, 张传宝. 参考物质互通性评价方案.检验医学, 2021, 36(3): 245-249.