

# 中华人民共和国卫生行业标准

WS/T 494—2017

---

## 临床定性免疫检验重要常规项目分析 质量要求

Guideline for performance characteristics of immunological qualitative test

2017-09-06 发布

2018-03-01 实施

---

## 目 次

前言 .....	III
引言 .....	IV
1 范围 .....	1
2 术语和定义 .....	1
3 临床定性免疫测定的方法 .....	4
3.1 筛查试验 .....	4
3.2 诊断试验 .....	4
3.3 确认试验 .....	4
4 临床定性免疫检验重要常规项目分析质量控制指标 .....	4
4.1 内容 .....	4
4.2 定性测定精密度 .....	5
4.3 以 COI 或 S/CO 比值表示结果的定性免疫测定的精密度 .....	5
4.4 准确度 .....	7
4.5 分析敏感性(即最低检出限) .....	16
4.6 对转化血清盘的检测能力 .....	16
4.7 分析特异性(交叉反应) .....	16
4.8 干扰因素 .....	16
参考文献 .....	18

## 前 言

本标准按照 GB/T 1.1—2009 给出的规则起草。

本标准起草单位：北京医院、首都医科大学附属北京朝阳医院、中国医学科学院北京协和医院、四川省医学科学院四川省人民医院。

本标准主要起草人：李金明、张瑞、王露楠、张括、谢洁红、王清涛、徐英春、黄文芳、袁红。

## 引 言

临床定性免疫检验有许多重要的常规检验项目,临床意义明确,应用广泛。临床定性免疫检验结果准确可比是医疗卫生工作的基本需要,提高和保证检验结果的准确性和可比性是临床检验质量管理及改进工作的重要内容。质量管理必须有明确的质量控制指标,临床定性免疫检验重要常规检验项目的分析质量控制指标主要包括:精密度、准确度、分析敏感性(即最低检出下限)、对转化血清盘的检测能力、分析特异性(交叉反应)和干扰因素等。

现有的临床定性免疫检验重要常规项目的检测,其检测性能差距较大,甚至有些商品试剂盒在其说明书上没有给出其质量控制指标等。此外,目前在某些项目上对这些试剂的重要性能指标应达到的最低要求缺乏相应的标准和规定。因此,临床实验室很难完整地评价所使用的试剂性能,也使得检验结果的准确性和可比性难以得到提高。

本标准将对开展临床定性免疫测定的实验室所需了解的检测系统或试剂(包括商品化的以及临床实验室自制的试剂)分析质量控制指标的解释和最低要求进行规定,从而保证临床实验室可以正确评价相关的定性免疫试剂。

# 临床定性免疫检验重要常规项目分析 质量要求

## 1 范围

本标准规定了临床定性免疫测定的方法和临床定性免疫检验重要常规项目质量控制指标。本标准适用于开展各种临床定性免疫测定的医学实验室。

## 2 术语和定义

下列术语和定义适用于本文件。

### 2.1

#### 准确度 accuracy

分析物的测定结果与真实结果之间的接近程度。准确度的定义在定性测定,是指样本阳性或阴性测定结果与其真实结果的一致性程度。通常通过方法学比较来实现。

### 2.2

#### 分析敏感性 analytical sensitivity

#### 测定下限 detection limits

可重复检测出待测物质的最低浓度水平。不同类型的标本,测定下限可能会有所不同。

### 2.3

#### 分析特异性 analytical specificity

一种检测方法仅对样本中的待测物质反应,而与其他物质不发生反应的能力。

### 2.4

#### 分析物 analyte

实验室试验所检测的物质或成分。

注:包括任意元素、离子、混合物、物质、因子、传染性物质、细胞、细胞器、(酶、激素或免疫原)活性、特性、存在与否、浓度、活性、强度、或者其他需要确定的特征。

### 2.5

#### 偏倚 bias

检测结果和真值之间的差异。

### 2.6

#### 5%的检测浓度 5% of detection concentrat

#### $C_5$

检测浓度为  $C_5$  的分析物时将产生 5% 的阳性结果。用浓度小于  $C_5$  的样本进行重复性检测时,将持续得到阴性结果。

### 2.7

#### 临界值浓度 critical concentrat

#### $C_{50}$

处于或接近临界值的分析物浓度,多次重复检测此浓度的单一样本时将获得 50% 的阳性结果和 50% 的阴性结果。

## 2.8

95%的检测浓度 95% of detection concentrat

$C_{95}$

检测浓度为  $C_{95}$  的分析物时将分别产生 95% 的阳性结果。用浓度大于  $C_{95}$  的样本进行重复性检测时,将持续得到阳性结果。

## 2.9

临床敏感性 clinical sensitivity

当特定疾病存在时,患者标本检测结果为阳性或超过正常值范围的比率。

## 2.10

临床特异性 clinical specificity

当特定疾病不存在时,患者标本检测结果为阴性或者在正常值范围内的比率。

## 2.11

诊断准确度 diagnostic accuracy

待评价试剂的检测信息和诊断标准的符合程度。

注 1: 诊断准确度可以多种形式表达,包括敏感性-特异性配对、似然比配对和接受者操作特性曲线(ROC 曲线)下的面积。

注 2: 诊断准确度需在关注状况下,结合特定标准与使用的方法进行阐述。

注 3: 诊断准确度不等同于准确度。

## 2.12

符合率 efficiency

一检测试剂或方法给出正确结果(包括阳性结果和阴性结果)的百分比。

## 2.13

假阳性结果 false-positive results

分析物为阴性的样本或患者,检测结果为阳性。

## 2.14

假阴性结果 false-negative results

分析物为阳性的样本或患者,检测结果为阴性。

## 2.15

不精密度 imprecision

特定条件下得到的独立测定结果的分散程度。数值上用标准差和变异系数表示。

## 2.16

阴性预测值 negative predictive value

临床上,阴性预测值为一检测试剂或方法测定为阴性的标本实际上为阴性的可能性。

注: 计算公式见表 1。

表 1 敏感性、特异性和阳性预测值、阴性预测值的计算

试剂的检测结果	分析物阳性	分析物阴性	合计
阳性结果数	TP	FP	TP+FP
阴性结果数	FN	TN	FN+TN
合计	TP+FN	FP+TN	TP+FP+FN+TN
敏感性	$\frac{TP}{TP+FN} \times 100\%$		

表 1 (续)

试剂的检测结果	分析物阳性	分析物阴性	合计
特异性		$\frac{TN}{TN+FP} \times 100\%$	
阳性预测值		$\frac{TP}{TP+FP} \times 100\%$	
阴性预测值		$\frac{TN}{TN+FN} \times 100\%$	
符合率		$\frac{TN+TP}{TP+FP+TN+FN} \times 100\%$	
式中： TP —— 分析物存在时检测结果为阳性的数量； FP —— 分析物不存在时结果为阳性的数量； TN —— 分析物不存在时检测结果为阴性的数量； FN —— 分析物存在时检测结果为阴性的数量。			

## 2.17

**精密度 precision**

在定性测定中,精密度的概念是一个阳性或阴性样本,重复多次测定得到阳性或阴性结果的比率。在评价定性测定时,不能使用强阳性或阴性样本,只能使用接近临界浓度的样本。在评价化学发光免疫试验(chemiluminescence immunoassay, CLIA)和酶联免疫吸附试验(enzyme-linked immunosorbent assay, ELISA)等可将结果以临界值(cut-off Index, COI 或 S/CO)方式表示的试验中,精密度的定义与定量测定的定义相同。精密度无法用数字来表示,只能通过不精密度如标准差和变异系数来评估。

## 2.18

**阳性预测值 positive predictive value**

临床上,阳性预测值为一检测试剂或方法测定为阳性的标本实际上为阳性的可能性。

注:计算公式见表 1。

## 2.19

**流行率 prevalence**

与特定的人群中总的成员数相比,患病或受特定健康状况影响的人数所占的比例。

## 2.20

**定性检测 qualitative tests**

只提供两种反应结果的检测方法(即阳性/阴性或者是/否或者有反应的/无反应的)。

## 2.21

**敏感性 sensitivity**

当分析物存在时,其检测结果呈阳性或者超过正常值范围的比率。

## 2.22

**特异性 specificity**

当分析物不存在时,其检测结果呈阴性或者在正常值范围内的比率。

## 2.23

**转化血清盘 seroconversion panels**

由分析物从无到有的过程中不同阶段样本所构成的血清盘。

## 2.24

**真阴性结果 true-negative results**

分析物为阴性的样本或患者,检测结果为阴性。

## 2.25

**真阳性结果 true-positive results**

分析物为阳性的样本或患者,检测结果为阳性。

## 3 临床定性免疫测定的方法

## 3.1 筛查试验

筛查试验用于检测整个人群或部分人群中分析物的存在情况。一般而言,筛查试验应具有较高的临床敏感性(即临床检出率大于95%),对其特异性和阳性预测值的要求则取决于对各种因素的综合考虑,例如假阳性结果是否会对被检测人经济上或心理上产生严重的不良影响、对误诊病例的治疗是否会产生严重的后果(例如,由于风疹抗体水平较高,而怀疑妊娠期感染,为避免感染造成的不良后果,需要进行流产手术以终止妊娠)、是否有可以对阳性筛查结果进行确认的方法、确认是否易于执行、确认方法是否价格昂贵等。通常情况下,筛查试验的阳性预测值都会低于诊断试验和确认试验。也就是说,筛查试验结果阴性提示被检测人分析物阴性的可能性很高,而筛查试验结果阳性仅提示阳性结果的可能,但需要进一步的确认。

## 3.2 诊断试验

诊断试验用于检测临床上已怀疑某种分析物(例如感染性病原体抗原或相关抗体)的存在情况。如果待测的感染性病原体抗原或相关抗体对于治疗以及判断预后有重要意义,则该诊断试验应当具有足够高的敏感性。如果试验的结果可以很容易地通过确认试验进行确认,且确认试验的准确度高,那么对诊断试验的特异性要求可适当降低。

## 3.3 确认试验

确认试验用于验证筛查试验或者诊断试验结果。对于确认试验,特异性和阳性预测值比敏感性和阴性预测值更为重要,确认试验的特异性应大于98%。确认试验可以是免疫测定,也可以是培养或核酸检测。用于确认试验的免疫测定方法有免疫印迹试验(WB)、重组免疫印迹(RIBA)、抗体中和试验等,在大多数情况下,一种试验方法如果不属于筛查试验或诊断试验,那么就应当属于确认试验。如果筛查试验或者诊断试验本身的特异性和阳性预测值就很高,那么就没有必要再使用确认试验进行结果的确认。

## 4 临床定性免疫检验重要常规项目分析质量控制指标

## 4.1 内容

临床定性免疫重要常规项目分析质量控制指标的内容包括:

- a) 精密度:包括定性测定的精密度和可以COI或S/CO比值报告结果的精密度评价两类;
- b) 准确度:特定的定性免疫测定系统或试剂或试验的准确度以与临床诊断、金标准方法和经过验证确认血清样本盘的检测比较来评价,常以敏感性(包括临床敏感性)/特异性(包括临床特异性)和阳性符合率/阴性符合率来表示;
- c) 分析敏感性(即最低检出下限);



- d) 对转化血清盘的检测能力；
- e) 分析特异性(交叉反应)；
- f) 干扰因素。

## 4.2 定性测定精密度

### 4.2.1 不精密度的来源

不精密度的来源包括：

- a) 样本状态及处理；
- b) 样本、试剂的运输和贮存条件；
- c) 操作人员；
- d) 环境条件；
- e) 仪器、试剂或检测系统；
- f) 检测程序(加样、温育、洗涤、结果判读时间等)。

### 4.2.2 50%的检测浓度

试剂生产厂家根据检测目的及敏感性和特异性建立  $C_{50}$ ，此临界值一旦确立，用户不可随意更改，若检测结果低于临界值则判定为阴性或无反应，高于临界值则判定为阳性或有反应。

在理想条件下对恰好为  $C_{50}$  的样本进行一系列重复性检测，将产生 50% 的阴性结果和 50% 的阳性结果。实际操作中，理想条件不易达到，因此接近临界值的分析物浓度，即出现 50/50 分界点的浓度，每个实验室会存在差异。尽管如此，每个实验室的此分析物浓度均称为  $C_{50}$ 。如果在  $C_{50}$  的浓度基础上，逐步增加待测物的浓度，并对浓度梯度进行检测，理应获得相应逐步增大的阳性结果百分比和更小的阴性结果百分比。同理，逐步降低待测物浓度，则应得到相反的结果。如果待测物的浓度接近  $C_{50}$ ，测定结果将具有不确定性，同一份样本的多次检测结果(阳性或阴性、有或无、有反应或无反应)不可能保持一致。

### 4.2.3 5%~95%的检测浓度区间

类似于  $C_{50}$  的定义，检测浓度为  $C_5$  和  $C_{95}$  的分析物时将分别产生 5% 和 95% 的阳性结果。用浓度小于  $C_5$  的样本进行重复性检测时，将持续得到阴性结果，用浓度大于  $C_{95}$  的样本进行重复性检测时，将持续得到阳性结果。

分析物浓度位于  $C_5 \sim C_{95}$  区间之外(小于  $C_5$  或大于  $C_{95}$ )时，候选方法对同一样本的重复性检测将得到相同结果。 $C_5 \sim C_{95}$  区间越窄，检测方法精密度越好。 $C_5 \sim C_{95}$  区间( $\geq C_5$  且  $\leq C_{95}$ )和同一样本重复检测可获得一致结果时的浓度范围，在使用相同的分析物但采用不同方法检测时可能存在差异。由于  $C_5 \sim C_{95}$  区间反映了重复检测可能获得不完全一致结果的浓度范围，因此  $C_5 \sim C_{95}$  区间的宽度可表示定性检测的精密度。

对于定性免疫检测而言，实验室了解  $C_5 \sim C_{95}$  区间的意义在于，明确同一样本重复检测可获得一致结果时的浓度范围。 $C_5 \sim C_{95}$  区间为作为检测系统或试剂的“灰区”。

注：从  $C_5 \sim C_{95}$  之间的浓度范围称为方法的“95%区间”，切勿将该术语和 95% 置信区间混淆。

### 4.2.4 定性测定精密度的要求

定性免疫测定精密度分析质量指标为  $C_5 \sim C_{95}$  浓度区间(亦即“灰区”)，这一区间范围  $\leq C_{50} \pm 20\%$ 。

## 4.3 以 COI 或 S/CO 比值表示结果的定性免疫测定的精密度

### 4.3.1 概述

有些定性免疫测定，如酶联免疫吸附试验(ELISA)、化学发光免疫分析试验(CLIA)等，样本的测定

结果可以 COI 或 S/CO 比值来表示,检测系统或试剂厂家会在其试剂说明书中给出试剂的批内和批间变异,这对于对实验室评价试剂的精密度以及评估其室内质控的批间变异结果有参考价值。

批内精密度是指严格的相似条件下,所得到的最佳的精密度;批间精密度指在同一实验室,由同一(组)操作员在同一仪器上,使用同一方法和同种、同一批号试剂,在一段时间内(一般为一个月或 20 个工作日)对同一检测样本(常用质控品)测量结果的精密度。

#### 4.3.2 批内和批间精密度评估的基本原则

4.3.2.1 操作者应熟悉检测系统、试剂方法和仪器工作原理,了解并掌握仪器的操作步骤和各项注意事项,应在评估阶段维持仪器的可靠和稳定。

4.3.2.2 用于评估试验的样本一般常采用临床实验室收集的血清(浆)库;当实验室收集的样本不稳定或不易得到时,也可考虑使用稳定的、以蛋白质为基质的商品物质,如校准品或质控品。

4.3.2.3 评估精密度时,应至少评估两个浓度水平样本的精密度。当两个浓度的精密度有显著差异时,宜增加为三个浓度。所选样本浓度应在测量范围内有医学意义,即至少有一个浓度在医学决定水平左右,在定性测定,即为接近临界水平的浓度。具体可参考试剂说明书中在评价精密度时所用的检测样本的浓度水平,甚至阴性样本。

#### 4.3.3 批内精密度的评估

4.3.3.1 试剂和校准品。可使用不同批号的试剂和校准物。

4.3.3.2 评估方法。三个不同浓度(参考试剂盒说明书)的样本,在一个测试批内重复进行至少 20 个检测,计算所得 S/CO 值的均值和 *SD*,计算批内 *CV*。

4.3.3.3 质量控制。检验时应同时至少测一个质控品。当质控品结果超出规定的失控限,不论实验结果是否满意都应弃去不用,重新进行试验以取得实验数据。要保存所有的质控数据和失控处理记录。

#### 4.3.4 批内精密度的要求

批内变异系数应 $<10\%$ ,同时应不大于试剂盒说明书指出的批内 *CV* 的  $10\%$ 。

#### 4.3.5 批间精密度的评估

4.3.5.1 试剂和校准品。可使用不同批号的试剂和校准物。

4.3.5.2 评估方法。3 个不同浓度(参考试剂盒说明书)的样本,在 10 d 以上时间内单次(孔或管)重复进行至少 20 批检测,计算所得 S/CO 值的均值和 *SD*,计算批间 *CV*。

4.3.5.3 质量控制。检验时应同时至少测一个质控品。当质控品结果超出规定的失控限,不论实验结果是否满意都应弃去不用,重新进行试验以取得实验数据。要保存所有的质控数据和失控处理记录。

#### 4.3.6 批间精密度的要求

批间变异系数应小于  $15\%$ ,同时应不大于试剂盒说明书指出的批间 *CV* 的  $20\%$ 。

#### 4.3.7 报告内容

精密度的评估影响因素较多,在报告精密度时,应同时说明下列各项:

- a) 批内标准差;
- b) 批内变异系数;
- c) 批间标准差;
- d) 批间变异系数;
- e) 实验进行的工作天数;

- f) 检验批次数；
- g) 每个批次重复检测数和总检测数。

## 4.4 准确度

### 4.4.1 概述

特定的定性免疫测定系统或试剂的准确度以与临床诊断、金标准方法和经过验证确认血清样本盘的检测比较来评价,常以敏感性(包括临床敏感性)/特异性(包括临床特异性)和阳性符合率/阴性符合率来表示。

### 4.4.2 诊断准确度和诊断标准

一种定性免疫测定试剂的准确度,取决于其检测结果和诊断标准的一致程度。所谓诊断标准,是指确认分析物是否存在的最好的检测方法。诊断标准只给出两种结果:分析物存在或不存在。诊断标准可以是单个检测方法,例如已经过验证的确认试验;也可以是不同方法和技术的结合,包括追踪随访、公认的诊断指南等;如果诊断标准是不同方法的结合,需说明是如何将不同的方法相结合,最终给出阳性/阴性的结果,包括:选择方法的原因,两种检测方法使用的顺序,第一种方法中结果为何种情况时使用第二种方法进行检测,判定规则等。例如,红斑狼疮、类风湿关节炎采用美国风湿学会的指南,幽门螺杆菌感染的诊断则是结合了培养、组织学检测和尿素酶实验进行的。在实际应用时,除了直接比较检测结果和诊断标准的结果外,常见的方法还有使用已经过诊断标准确认的参考血清盘进行评价,将试剂的检测结果与参考血清盘的预期结果相比较。

判断诊断的方法是否为目前最好的方法或者能否作为诊断标准,取决于临床的、实验室的以及相关的政策法规等很多方面。有时存在多种待选的方法可以作为标准的检测方法,有时可能没有一致的公认的诊断标准,又或者存在某种诊断标准,但是这种方法在某一特定人群中是不适用的,而这一人群又存在一定的比例。

诊断准确度可用多种方式来表述,包括敏感性(及临床敏感性)和特异性(及临床特异性)、阳性似然比和阴性似然比、置信区间的 ROC 曲线分析。本标准中仅对试剂说明书中最常见的敏感性和特异性进行说明和规定。

### 4.4.3 敏感性和特异性

#### 4.4.3.1 概述

敏感性和特异性只能是一个估计值。这是因为对敏感性和特异性的计算只能基于一组人群,而不可能是所有的人群。即使是相同的一组人群但时间不同,其敏感性和特异性的数值都会有所不同。置信区间水平可对由于样本选择的不同产生的不确定度进行一定估计。研究人群的数量越多,其不确定度越小。

敏感性和特异性是试剂性能的重要指标,提示了试剂检测分析物是否存在的能力。需要说明的是一种敏感性等于 $(1 - \text{特异性})$ 的试剂不具有诊断价值,即当分析物存在时的试剂的阳性率与分析物不存在时试剂的阳性率完全相同时,该试剂则不具有诊断价值。而敏感性和特异性均接近 1 的试剂具有良好的诊断能力。

#### 4.4.3.2 敏感性和特异性的计算

从统计学的角度来说,最理想的方式就是将待评价试剂与诊断标准同时检测具有代表性的人群(包括分析物存在的人群以及分析物不存在的人群),并将结果进行比较。在这种情况下,可计算敏感性和特异性两种指标,计算方法见表 1。

### 4.4.3.3 敏感性和特异性的置信区间

敏感性和特异性的置信限有多种计算方法,数据呈正态分布时,常用的较为简单的方法是正态近似二项式分布。当样本量较少或比率接近“0%”或“100%”的情况下,这种近似正态的假设通常是不成立的。

敏感性和特异性的精确置信限(Clopper-Pearson 方法)可以通过二项分布、多种统计软件包、使用 F-表手工计算或者从已经发表的表中获得。有能力计算精确置信限的用户可使用上述方法。除此以外,一种由 Wilson 创建的计分置信区间的直接计算方法可以应用于所有例子中。本文件建议使用计分置信区间,介绍如下。

敏感性的 95% 计分置信区间计算公式见式(1):

$$[100 \times (Q_{1,se} - Q_{2,se})/Q_{3,se}, 100 \times (Q_{1,se} + Q_{2,se})/Q_{3,se}] \dots\dots\dots(1)$$

其中  $Q_{1,se}$ ,  $Q_{2,se}$  和  $Q_{3,se}$  的数据可通过式(2)~式(4)计算:

$$Q_{1,se} = 2 \times TP + 1.96^2 = 2 \times TP + 3.84 \dots\dots\dots(2)$$

$$Q_{2,se} = 1.96 \sqrt{1.96^2 + 4 \times TP \times FN / (TP + FN)} = 1.96 \sqrt{3.84 + 4 \times TP \times FN / (TP + FN)} \dots\dots\dots(3)$$

$$Q_{3,se} = 2(TP + FN + 1.96^2) = 2(TP + FN) + 7.68 \dots\dots\dots(4)$$

特异性的 95% 计分置信区间计算公式见式(5):

$$[100 \times (Q_{1,sp} - Q_{2,sp})/Q_{3,sp}, 100 \times (Q_{1,sp} + Q_{2,sp})/Q_{3,sp}] \dots\dots\dots(5)$$

其中  $Q_{1,sp}$ ,  $Q_{2,sp}$  和  $Q_{3,sp}$  的数据可从表 1 得到并通过式(6)~式(8)计算:

$$Q_{1,sp} = 2 \times TN + 1.96^2 = 2 \times TN + 3.84 \dots\dots\dots(6)$$

$$Q_{2,sp} = 1.96 \sqrt{1.96^2 + 4 \times FP \times TN / (FP + TN)} = 1.96 \sqrt{3.84 + 4 \times FP \times TN / (FP + TN)} \dots\dots\dots(7)$$

$$Q_{3,sp} = 2(FP + TN + 1.96^2) = 2(FP + TN) + 7.68 \dots\dots\dots(8)$$

式(1)~式(8)中,1.96 是从标准正态分布对应的 95% 置信区间得来的。若置信水平改变,则 1.96 也应被相应的百分位点所替代。

### 4.4.3.4 敏感性、特异性的计算举例

对敏感性、特异性计算方法举例见表 2。

表 2 待评价试剂与诊断标准相比较进行敏感性、特异性的计算举例

待评价试剂的检测结果	诊断标准		总计
	分析物存在	分析物不存在	
阳性结果数	88	2	90
阴性结果数	14	336	350
共计	102	338	440

敏感性 =  $88/102 \times 100\% = 86.3\%$

特异性 =  $336/338 \times 100\% = 99.4\%$

敏感性和特异性的 95% 置信区间为敏感性(74.3%, 93.2%)和特异性(96.7%, 99.9%)。

### 4.4.3.5 阳性预测值和阴性预测值

对于临床医生而言,免疫测定方法的临床预测值是最为重要的参数。因此,确立有效准确的预测值

并说明检测的临床意义是非常重要的。阳性预测值的确立应当建立在大样本的检测结果之上,此外分析物在某一人群中的流行率对于检测结果的预测值会产生非常大的影响,见表3~表5,95%敏感性和特异性试验在流行率为10%的人群中,其阳性预测值为67.9%,而在流行率为1%的人群中,其阳性预测值仅为16%。这里流行率表示单位人群(通常单位人群指100 000人)中的病例总数,其不同于发病率,发病率指每年单位人群中新发病例的数量。因此,如果临床上对怀疑某种病原体感染(有临床表现,或有其他相关指标提示可能感染)的患者进行检测,则阳性预测值将大为提高;相反,如果对所有就诊者常规筛查,则阳性预测值低。

阳性预测值是指一检测试剂或方法测定为阳性的标本实际上为阳性的可能性。阴性预测值是指一检测试剂或方法测定为阴性的标本实际上为阴性的可能性。这些值可帮助正确地解释实验结果。临床实验室可根据试剂厂家提供的临床敏感性和临床特异性的结果,在本实验室检测人群分析物流行率的基础上,来计算检测结果的阳性预测值和阴性预测值。

#### 4.4.3.6 定性免疫检验项目敏感性、特异性要求

作为筛查试验,检测的敏感性应大于95%。作为诊断试验,检测敏感性和特异性均应大于95%。作为确认试验特异性应大于98%。

#### 4.4.4 阳性符合率和阴性符合率

##### 4.4.4.1 概述

如果没有可用的诊断标准,那么厂家需提供待评价试剂与某一已验证方法进行比较的结果,在这种情况下,则不能够计算敏感性和特异性,而是计算阳性符合率和阴性符合率。这种计算反映的并非方法的准确度,而是待评价试剂和已验证方法的一致性程度。此外,由于是两种方法的相互比较,而并不知道样本的真实状态,包括阳性预测值、阴性预测值、阳性似然比和阴性似然比这样的指标均不能够进行计算。尽管如此,能够给出待评价试剂与已验证方法的一致性也还是对实验室有参考意义的。

这种情况下,实验室需要了解以下数据:

- 待评价试剂与已验证方法相比较的 $2 \times 2$ 表;
- 对已验证方法的描述,以及其具体操作程序;
- 阳性符合率和阴性符合率,以及其对应的置信区间。

表3 流行率对于90%敏感性和特异性试验阳性预测值和阴性预测值的影响

结 果	80%流行率		50%流行率		10%流行率		1%流行率		0.5%流行率	
	检出 (+)	未检出 (-)	检出 (+)	未检出 (-)	检出 (+)	未检出 (-)	检出 (+)	未检出 (-)	检出 (+)	未检出 (-)
+	72 000	8 000	45 000	5 000	9 000	1 000	900	100	450	50
-	2 000	18 000	5 000	45 000	9 000	81 000	9 900	89 100	9 950	89 550
敏感性 [TP/(TP+ FN)]	90%		90%		90%		90%		90%	
特异性 [TN/(TN+ FP)]	90%		90%		90%		90%		90%	

表 3 (续)

结 果	80%流行率		50%流行率		10%流行率		1%流行率		0.5%流行率	
	检出 (+)	未检出 (-)	检出 (+)	未检出 (-)	检出 (+)	未检出 (-)	检出 (+)	未检出 (-)	检出 (+)	未检出 (-)
阳性预示值 TP/(TP+ FP)	79 200/(79 200+ 200)=97.3%		45 000/(45 000+ 5 000)=90.0%		9 000/(9 000+ 9 000)=50%		900/(900+ 9 900)=8.33%		450/(450+ 9 950)=4.3%	
阴性预示值 TN/(TN+ FN)	18 000/(18 000+ 8 000)=69.2%		45 000/(45 000+ 5 000)=90.0%		81 000/(81 000+ 100)=99.88%		89 100/(89 100+ 100)=99.89%		89 550/(89 550+ 50)=99.94%	

表 4 流行率对于 95%敏感性和特异性试验阳性预测值和阴性预测值的影响

结 果	80%流行率		50%流行率		10%流行率		1%流行率		0.5%流行率	
	检出 (+)	未检出 (-)	检出 (+)	未检出 (-)	检出 (+)	未检出 (-)	检出 (+)	未检出 (-)	检出 (+)	未检出 (-)
+	76 000	4 000	47 500	2 500	9 500	500	950	50	475	25
-	1 000	19 000	2 500	47 500	4 500	85 500	4 950	94 050	4975	94 525
敏感性 [TP/(TP+ FN)]	95%		95%		95%		95%		95%	
特异性 [TN/(TN+ FP)]	95%		95%		95%		95%		95%	
阳性预示值 TP/(TP+ FP)	76 000/(76 000+ 1 000)=98.7%		47 500/(47 500+ 2 500)=95.0%		9 500/(9 500+ 4 500)=67.9%		950/(950+ 4 950)=16%		475/(475+ 4 975)=8.7%	
阴性预示值 TN/(TN+ FN)	19 000/(19 000+ 4 000)=82.6%		47 500/(47 500+ 2 500)=95.0%		85 500/(85 500+ 500)=99.4%		94 050/(94 050+ 50)=99.9%		94 525/(94 525+ 25)=99.97%	

表 5 流行率对于 99%敏感性和特异性试验阳性预测值和阴性预测值的影响

结 果	80%流行率		50%流行率		10%流行率		1%流行率		0.5%流行率	
	检出 (+)	未检出 (-)	检出 (+)	未检出 (-)	检出 (+)	未检出 (-)	检出 (+)	未检出 (-)	检出 (+)	未检出 (-)
+	79 200	800	49 500	500	9 900	100	990	10	495	5
-	200	19 800	500	49 500	900	89 100	990	98 010	995	98 505
敏感性 [TP/(TP+ FN)]	99%		99%		99%		99%		99%	

表 5 (续)

结 果	80%流行率		50%流行率		10%流行率		1%流行率		0.5%流行率	
	检出 (+)	未检出 (-)	检出 (+)	未检出 (-)	检出 (+)	未检出 (-)	检出 (+)	未检出 (-)	检出 (+)	未检出 (-)
特异性 [TN/(TN+FP)]	99%		99%		99%		99%		99%	
阳性预示值 TP/(TP+FP)	79 200/(79 200+200)=99.7%		49 500/(49 500+500)=99.0%		9 900/(9 900+900)=91.7%		990/(990+990)=50%		495/(495+995)=33.2%	
阴性预示值 TN/(TN+FN)	19 800/(19 800+800)=96.1%		49 500/(49 500+500)=99.0%		89 100/(89 100+100)=99.9%		94 050/(94 050+50)=99.99%		94 525/(94 525+25)=99.99%	

## 4.4.4.2 阳性符合率和阴性符合率的计算

所有样本均分别进行待评价试剂和已验证方法的检测,得出待评价试剂与已验证方法相比较的 $2 \times 2$ 表,见表 6。

表 6 待评价试剂与已验证方法相比较进行阳性符合率和阴性符合率的计算

待评价试剂的检测结果	已验证方法	
	+	-
+	<i>a</i>	<i>b</i>
-	<i>c</i>	<i>d</i>
共计	<i>a+c</i>	<i>b+d</i>

表 6 和表 2 的区别在于,表 6 中的阳性结果不代表分析物存在。因此  $a$ 、 $b$ 、 $c$ 、 $d$  也不代表表 2 中的  $TP$ 、 $FP$ 、 $FN$ 、 $TN$ 。表 2 中的数据表示待评价试剂结果正确的可能性,表 6 中的数据则表示待评价试剂与已验证方法符合的可能性。 $a$  代表已验证方法和待评价试剂的检测结果均为阳性; $b$  代表验证方法为阴性,待评价试剂的检测结果为阳性; $c$  代表已验证方法为阳性,待评价试剂的检测结果为阴性; $d$  代表待评价试剂与已验证方法均为阴性。

$$\text{总符合率} = (a + d) / (a + b + c + d) \times 100\%$$

但是,总符合率不能足够地反映两种方法的一致程度。如  $b+c$  的和可以相同,但是  $b$  和  $c$  均为完全不同的结果。因此,还需提供以下两组数据:

$$\text{阳性符合率} = a / (a + c) \times 100\%$$

$$\text{阴性符合率} = d / (b + d) \times 100\%$$

阳性符合率可以解释为已验证方法检测结果为阳性的样本中待评价试剂的阳性率,或待评价试剂检测结果为阳性的样本中已验证方法的阳性率,但是在本标准中,仅代表已验证方法检测结果为阳性的样本中待评价试剂的阳性率。

## 4.4.4.3 阳性符合率和阴性符合率的置信区间

符合率的 95% 计分置信区间的计算方法见式(9):

$$[100\%(Q_1 - Q_2)/Q_3, 100\%(Q_1 + Q_2)/Q_3] \dots\dots\dots(9)$$

其中  $Q_1$ 、 $Q_2$  和  $Q_3$  的值可运用下面公式得到：

$$Q_1 = 2(a + d) + 1.96^2 = 2(a + d) + 3.84 \dots\dots\dots(10)$$

$$Q_2 = 1.96\sqrt{1.96^2 + 4(a + d)(b + c)/n} = 1.96\sqrt{3.84 + 4(a + c)(b + d)/n} \dots\dots(11)$$

$$Q_3 = 2(n + 1.96^2) = 2n + 7.68 \dots\dots\dots(12)$$

式(10)~式(12)中,1.96 是从标准正态分布对应的 95%置信区间得来的。

阳性符合性(positive percent agreement, PPA)95%计分置信区间的计算方法见式(13)：

$$[100 \times (Q_{1,ppa} - Q_{2,ppa})/Q_{3,ppa}, 100 \times (Q_{1,ppa} + Q_{2,ppa})/Q_{3,ppa}] \dots\dots\dots(13)$$

$Q_{1,ppa}$ 、 $Q_{2,ppa}$  和  $Q_{3,ppa}$  的值可运用式(14)~式(16)得到：

$$Q_{1,ppa} = 2a + 1.96^2 = 2a + 3.84 \dots\dots\dots(14)$$

$$Q_{2,ppa} = 1.96\sqrt{1.96^2 + 4ac/(a + c)} = 1.96\sqrt{3.84 + 4ac/(a + c)} \dots\dots\dots(15)$$

$$Q_{3,ppa} = 2(a + c + 1.96^2) = 2(a + c) + 7.68 \dots\dots\dots(16)$$

阴性符合率(negative percent agreement, NPA)95%计分置信区间的计算方法见式(17)：

$$[100 \times (Q_{1,npa} - Q_{2,npa})/Q_{3,npa}, 100 \times (Q_{1,npa} + Q_{2,npa})/Q_{3,npa}] \dots\dots\dots(17)$$

$Q_{1,npa}$ 、 $Q_{2,npa}$  和  $Q_{3,npa}$  的值可运用式(18)~式(20)得到：

$$Q_{1,npa} = 2d + 1.96^2 = 2d + 3.84 \dots\dots\dots(18)$$

$$Q_{2,npa} = 1.96\sqrt{1.96^2 + 4bd/(b + d)} = 1.96\sqrt{3.84 + 4bd/(b + d)} \dots\dots\dots(19)$$

$$Q_{3,npa} = 2(b + d + 1.96^2) = 2(b + d) + 7.68 \dots\dots\dots(20)$$

#### 4.4.4.4 阳性符合率和阴性符合率的计算举例

阳性符合率和阴性符合率计算举例见表 7。

表 7 待评价试剂与已验证方法相比较进行阳性符合率和阴性符合率的计算举例

待评价试剂的检测结果	已验证方法		总计
	+	-	
+	80	10	90
-	8	342	350
共计	88	352	440

阳性符合率 =  $80/88 \times 100\% = 90.9\%$ 。

阴性符合率 =  $342/352 \times 100\% = 97.2\%$

总符合率 =  $(80 + 342)/440 \times 100\% = 95.9\%$

95%置信区间为：敏感性(78.8%, 96.4%)和特异性(93.5%, 98.8%)。

#### 4.4.4.5 总符合率的意义

两种试剂的总符合率相同,不代表两种试剂的检测性能相同。以表 8 和表 9 为例,两组试剂与已验证方法的总符合率相同,但是两种试剂的性能可以有很大差别,试剂 A 的阳性符合率(67.8%)远低于试剂 B 的阳性符合率(97.6%)。



表 8 试剂 A 与已验证方法相比较的结果

试剂 A	已验证方法		总计
	+	-	
+	80	2	82
-	38	1 024	1 062
共计	118	1 026	1 144
阳性符合率	$80/118 \times 100\% = 67.8\%$		
阴性符合率	$1\ 024/1\ 026 \times 100\% = 99.8\%$		
总符合率	$(80+1\ 024)/1\ 144 \times 100\% = 96.5\%$		

表 9 试剂 B 与已验证方法相比较的结果

试剂 B	已验证方法		总计
	+	-	
+	80	38	118
-	2	1 024	1 026
共计	82	1 062	1 144
阳性符合率	$80/82 \times 100\% = 97.6\%$		
阴性符合率	$1\ 024/1\ 062 \times 100\% = 96.4\%$		
总符合率	$(80+1\ 024)/1\ 144 \times 100\% = 96.5\%$		

#### 4.4.4.6 阳性符合率和阴性符合率的意义

待评价试剂与已验证方法相比较而得出的阳性符合率和阴性符合率,与敏感性和特异性相比,具有两个主要的缺陷:

- “符合”不等于“正确”。两种检测方法可能高度符合但是敏感性和特异性都很低。相反,两种检测方法不符合也并不意味着待评价试剂是错误的而比较方法是正确的。
- 符合率受检测人群的分析物的阳性率影响。

通常,两种试剂在分析物阴性的人群中的符合率与分析物阳性人群中的符合率是不同的,因此,即使其他条件不变,只要改变研究人群中分析物阳性和阴性的比例,两种试剂的符合率就会产生改变(可能是比较大的改变)。即两种试剂的符合率受检测人群的分析物的阳性率影响。

现举例(见表 10)如下:

表 10 待评价试剂和已验证方法比较结果——分析物流行率高

待评价试剂	已验证方法	总计	诊断标准	
			+	-
+	+	80	78	2
+	-	10	10	0

表 10 (续)

待评价试剂	已验证方法	总计	诊断标准	
			+	-
-	+	8	2	6
-	-	342	12	330
合计		440	102	338

研究人群共计 220 例,分析物的阳性率为 23.2%(102/440),分析物阳性人群中总符合率为 88.2%[(78+12)/102],分析物阴性人群中总符合率为 98.2%[(2+330)/338],所有人群总符合率为 95.9%[(78+12+2+330)/440],阳性符合率为 90.9%[(80)/(80+8)],阴性符合率为 97.2%[(342)/(342+10)]。

如果保持分析物阳性人群中总符合率不变,分析物阴性人群中总符合率不变,只是改变人群中分析物的阳性和阴性的比例,如将分析物阴性的人数增加 4 倍,即 1 352 人,则结果见表 11。

表 11 待评价试剂和已验证方法比较结果——分析物流行率低

待评价试剂	已验证方法	总计	诊断标准	
			+	-
+	+	86	78	8
+	-	10	10	0
-	+	26	2	24
-	-	1 332	12	1 320
合计		1 454	102	1 352

研究人群共计 1 454 例,分析物阳性人群中总符合率为 88.2%[(78+12)/102],分析物阴性人群中总符合率为 98.2%[(8+1 320)/1 352]。而不同的是所有人群总符合率为 97.5%[(78+12+8+1 320)/1 454],高于 95.9%。而差别更大的是,阳性符合率为[86/(86+26)],远低于 90.9%。阴性符合率为 99.2%[(1 332)/(1 332+10)],略高于 97.2%。

研究样本的阳性率越低,两种试剂相比较得出的阳性符合率也越低,而阴性符合率则越高。因此,如果试剂厂家仅提供表 6 的结果,则其得出的阳性符合率和阴性符合率仅在其研究人群中适用。因为,表 6 中并没有研究人群分析物阳性率的信息。

#### 4.4.4.7 对不符合结果进一步检测的意义

即使采用诊断标准对不符合的结果进行阳性和阴性的确认,仍然不能够将两组结果进行合并,因为待评价试剂与已验证方法相符合的结果不能表示分析物就是阳性或阴性的。

表 12 待评价试剂和已验证方法比较结果——用另一种已验证方法对不符合结果进行检测(错误地将两组结果合并)

待评价试剂	已验证方法	总计	另一种已验证方法		总计(合并后)
			+	-	
+	+	80	不适用	不适用	90
+	-	10	10 ↑	0	0

表 12 (续)

待评价试剂	已验证方法	总计	另一种已验证方法		总计(合并后)
			+	-	
-	+	8	2↓	6	2
-	-	342	N/A	N/A	348
合计		440	N/A	N/A	440

如果试剂说明书中一定要提供第三种已验证方法的结果,宜采用表 13 格式,以供用户参考。

表 13 待评价试剂和已验证方法比较结果——用另一种已验证方法  
对不符合结果进行检测(推荐形式)

待评价试剂	已验证方法	总计	另一种已验证方法	
			+	-
+	+	80	不适用	不适用
+	-	10	10	0
-	+	8	2	6
-	-	342	N/A	N/A
合计		440	N/A	N/A

#### 4.4.5 敏感性/阳性符合率和特异性/阴性符合率相关的信息

##### 4.4.5.1 概述

除了敏感性和特异性以外,还有许多其他的因素临床实验室需要了解。如果不了解其研究人群本身的构成和特征以及其他的因素,仅敏感性和特异性本身不能说明该试剂的性能。

##### 4.4.5.2 研究人群

应符合以下要求:

- 涵盖疾病的各个阶段;
- 涵盖一些不典型的、较难诊断的患病人群;
- 各组不同特征的人群均需有一定的数量。

##### 4.4.5.3 外部有效性

外部有效性是指结果是在足够真实地反映实际应用的环境下得出的。实际应用的环境包括操作人员、仪器设备、检测人群等。因此实际应用的环境最好符合以下要求:

- 检测系统或试剂评价的检测过程完全按照临床实验室使用的试剂说明书进行,如果诊断试剂还要求使用配套的仪器设备,那么评价时使用的也是厂家最终提供给用户的仪器设备;
- 使用多台仪器设备进行检测;
- 多个操作人员进行操作,操作人员受过一定培训,有着不同的专业经验;
- 预期用途和操作环境均与实际用户应用时相近。

#### 4.4.5.4 关于敏感性/阳性符合率和特异性/阴性符合率所应了解的内容

在评价敏感性/阳性符合率和特异性/阴性符合率时,应对以下因素有充分了解:

- 该检测系统或试剂预期使用的人群;
- 该检测系统或试剂敏感性/阳性符合率和特异性/阴性符合率研究所涵盖的人群特征;
- 分析物以及分析物的检测原理;
- 比较方法是诊断标准还是已验证的方法;
- 待评价试剂与诊断标准/已验证方法比较结果的数据,最好是 $2 \times 2$ 表格;
- 如与诊断标准相比较,了解敏感性和特异性,或阳性似然比和阴性似然比的结果;如与已验证方法比较,了解总符合率、阳性符合率和阴性符合率的结果。无论哪种指标,都需给出95%的置信区间。

#### 4.5 分析敏感性(即最低检出限)

检测限是指检测方法可检测出的最低被测量浓度,也称检测低限或最小检出浓度,有时也称为分析灵敏度。

分析物浓度位于 $C_{5} \sim C_{95}$ 区间之外(小于 $C_{5}$ 或大于 $C_{95}$ )时,候选方法对同一样本的重复性检测将得到相同结果。因此, $C_{95}$ 代表了某一试剂可以测出的最低被测量浓度。

评估试剂分析敏感性所使用的样本,如检测项目有国家参考品,则可使用国家参考品或经国家参考品标化的参考品进行检测,如没有国家参考品,则使用可以溯源或量化的样本,如国际标准物质,或与国际标准物质溯源的样本,是更有意义的。

对于感染性疾病用于诊断感染的抗原和抗体的定性免疫测定,在不影响测定特异性的情况下,最低检出限越低越好,对于有国际标准物质的检测项目,较容易得到这个浓度,例如HBsAg化学发光免疫测定的最低检出限应 $< 0.1$  IU/mL,ELISA的最低检出限应 $< 0.2$  IU/mL。抗-HBs的最低检出限应 $< 10$  mIU/mL。对于商品试剂盒,某个特定指标的最低检出限可参考试剂盒说明书。

#### 4.6 对转化血清盘的检测能力

大部分抗原、抗体(尤其是抗体)定性检测指标没有国际标准物质,评估了解分析敏感性只能是通过较强阳性样本的系列(如倍比)稀释进行,但由于机体针对特定病原体抗原产生的抗体是多克隆的,含量不均一,系列稀释后,不能反映抗体天然产生的状况。此时,评估一个定性测定系统或试剂的分析敏感性的最佳方法,是采用转化血清盘。

对转化血清盘的检测能力的分析质量指标可定为不同试剂检测同一个阳转血清盘,越早检出者越好,具体对特定转化血清盘的检测能力可参考特定试剂盒的说明书。

#### 4.7 分析特异性(交叉反应)

对于感染性疾病的特异抗原和抗体,分析特异性评估质量指标是指用无特定病原体感染和特定病原体外的其他病原体感染者样本检测时,不应出现阳性结果。例如,HBsAg检测系统或试剂,当用其检测甲型肝炎病毒(HAV)、丙型肝炎病毒(HCV)、人类免疫缺陷病毒(HIV)、梅毒螺旋体、风疹病毒、麻疹病毒、弓形虫等病原体感染者或正常人样本检测时,不应出现阳性。

对于特定的自身抗体等,分析特异性则是指试剂盒所用抗原与特定自身抗体以外的其他抗体(包括其他自身抗体)的交叉反应程度。理想的分析质量指标是应无交叉反应性,如果有交叉反应,应当在试剂说明书中予以说明。

#### 4.8 干扰因素

4.8.1 干扰物质是体外诊断试剂使用过程中造成测量误差的一个主要原因,针对体外诊断试剂进行的

干扰实验是指通过实验查找出对体外诊断试剂测量结果产生影响的物质的过程。

4.8.2 干扰物质来源:干扰物质可能来自内源或外源物质。可疑干扰物质的来源通常有:

- 常见的异常标本,例如溶血、黄疸及脂血;
- 类风湿因子、嗜异性抗体;
- 普通的处方药及非处方药;
- 患者群体中异常的生化代谢物;
- 患者群体中常见的治疗药物。如由于试剂盒研制所用成分原因可能导致的干扰,如国外化学发光免疫试验常采用的链霉亲和素-生物素包被系统,则要考虑患者服用生物素治疗所产生的干扰作用;
- 干扰测量程序的药物(包括代谢物);

已报道干扰相似测量程序的物质:

- 标本处理过程中的添加物,例如抗凝剂、防腐剂;
- 采集及处理过程中接触标本的物质,例如血清分离设备、导管、标本收集容器及塞子;
- 影响某些实验的膳食物质,例如咖啡因, $\beta$ -胡萝卜素,罂粟籽。

4.8.3 临床实验室需了解干扰因素的相关信息有:

- 干扰因素为何种物质;
- 干扰因素对检测试剂造成怎样的影响(假阴性/假阳性);

可以造成影响的浓度范围。对于干扰因素的具体的分析质量控制指标是在常见的一定浓度范围内,不出现干扰,如:“类风湿因子 $<2\ 000\ \text{U/mL}$ 不出现干扰结果”;如为药物治疗,可说明血药浓度或治疗剂量和采样时间的限制等,例如:“使用生物素 $5\ \text{mg/d}$ 的剂量时,需在至少用药 $8\ \text{h}$ 后方能采样”。

参 考 文 献

- [1] 李金明,临床酶免疫测定技术.北京:人民军医出版社,2008
- [2] I/LA18-A2 Vol.21 No.15 Specifications for Immunological Testing for Infectious Diseases; Approved Guideline—Second Edition
- [3] EP15-A Vol.21 No.25 User Demonstration of Performance for Precision and Accuracy; Approved Guideline
- [4] EP5-A2 Vol.24 No.25 Evaluation of Precision Performance of Quantitative Measurement Methods; Approved Guideline—Second Edition
- [5] EP7-A Vol.22 No.27 Interference Testing in Clinical Chemistry; Approved Guideline
- [6] EP12-A Vol.22 No.14 User Protocol for Evaluation of Qualitative Test Performance; Approved Guideline
- [7] EP17-A Vol.24 No.34 Protocol for Determination of Limits of Detection and Limits of Quantitation; Approved Guideline
- [8] Food and Drug Administration Center for Devices and Radiological Health.2007.Statistical Guidance on Reporting Results from Studies Evaluating Diagnostic Tests
-