JJF

**中华人民共和国国家计量技术规范**

JJF XXXX—202X

检验医学定量检测项目基于患者数据的质量控制算法溯源方法

A Traceable Method of Patient-based Real-time Quality Control Algorithms for Quantitative Testing Results in Laboratory Medicine

（征求意见稿）

XXXX－XX－XX 发布 XXXX－XX－XX实施

国家市场监督管理总局发布

检验医学定量检测项目基于

JJF XXXX-202X

患者数据的质量控制算法溯

源方法

**A Traceable Method of Patient-based Real-time Quality Control Algorithms for Quantitative Testing Results in Laboratory Medicine**

**归口单位：**全国法制计量管理计量技术委员会

**主要起草单位:**

本规范委托全国法制计量管理计量技术委员会负责解释

本规范主要起草人：

参加起草人：

目 录

[引言 II](#_Toc4440)

[1 范围 1](#_Toc15543)

[2 引用文件 1](#_Toc29105)

[3 术语和定义 1](#_Toc27504)

[4 概述 3](#_Toc14945)

[5 计量特性 4](#_Toc18040)

[5.1 假性失控率 4](#_Toc28452)

[5.2 误差检出率 4](#_Toc10044)

[5.3 错误报告患者检测结果数量中位值/平均值 4](#_Toc24017)

[5.4 算法评价值 4](#_Toc32233)

[6 数字量具特性 4](#_Toc28874)

[6.1 质量在控数字量具 5](#_Toc671)

[6.2 质量失控数字量具 5](#_Toc8682)

[7 溯源要求 7](#_Toc18715)

[7.1 溯源方法 7](#_Toc31234)

[7.2 溯源周期 8](#_Toc4810)

[附 录 A 10](#_Toc10534)

[临床检验分析性能质量规范要求 10](#_Toc23144)

[附 录 B 11](#_Toc26188)

[表B.1 假性失控记录表 11](#_Toc11668)

[表B.2 误差检出记录表 11](#_Toc30996)

[表B.3 错误报告患者检测结果数量记录表 11](#_Toc19453)

# 引言

JJF 1001-2011《通用计量术语及定义》、JJF 1071-2010《国家计量校准规范编写规则》共同构成本规范编制工作的基础性系列规范。

结合当前临床检验定量测定室内质量控制算法研究与发展现状，对算法软件的评价至关重要。本规范是为了使利用检验医学定量检测项目基于患者数据的质量控制算法能够满足临床的需求。本规范主要参考GB/T 19000-2016《质量管理体系 基础和术语》、GB/T 3358.2-1993《统计学术语 第二部分 统计质量控制术语》、WS/T 403-2012《临床生物化学检验常规项目分析质量指标》、WS/T 415-2013《无室间质量评价时实验室检测评估方法》、CLSI C24 4th ed《Statistical Quality Control for Quantitative Measurement Procedures: Principles and Definitions》进行制定。

本规范为首次发布。

检验医学定量检测项目基于患者数据的质量控制算法溯源方法

# 1 范围

本规范适用于实验室或生产厂商基于患者数据的质量控制算法软件的评价。

# 2 引用文件

本规范引用下列文件：

JJF 1001-2011 通用计量术语及定义

JJF 1071-2010 国家计量校准规范编写规则

GB/T 2887-2011 计算机场地通用规范

GB/T 3358.2-1993 统计学术语 第二部分 统计质量控制术语

GB/T 19000-2016 质量管理体系 基础和术语

WS/T 403-2012 临床生物化学检验常规项目分析质量指标

WS/T 415-2013 无室间质量评价时实验室检测评估方法

凡是注明日期的引用文件，仅该日期对应的版本适用于本规范；不注明日期的引用文件，其最新版本（包括所有的修改单）适用于本规范。

# 3 术语和定义

下列术语和定义适用于本规范。

3.1质量控制 quality control

质量管理的一部分，致力于满足质量要求。

[GB/T 19000-2016，术语 3.3.7]

3.2 质量控制算法 quality control algorithm

质量管理所用的算法。

3.3 失控状态 out-of-control condition

某过程或过程中的某部分操作处于非稳定状态。

注: 对于定量测量程序，失控状态通常表现为测量程序偏离稳定均值或者稳定精密度增大。

[Statistical Quality Control for Quantitative Measurement Procedures: Principles and Definitions，CLSI C24 4th ed]

3.4 在控状态 in-control status of QC

某过程或过程中的某部分操作处于非失控状态的稳定状态。

3.5 控制限 control limits

控制图上用于判断过程是否处于在控状态的界限。

注：控制限分上控制限（upper control limit）和下控制限（lower control limit）,当过程处于统计控制状态时，所考虑的统计量将以很高的概率位于上/下控制限之间。

[GB/T 3358.2-1993，过程量度术语5.39]

3.6 不精密度 imprecision

检验结果精密度指标，指同一实验室用同种方法在多次独立检验中分析同一样品所得结果的离散程度。

注1：

不精密度有多种，如“批内”、“批间”、“室内”、“室间”不精密度等，计量学领域分重复性、中间、再现性不精密度等。本标准中的不精密度指室内（或中间）不精密度。

注2：

对于临床实验室，不精密度一般来自室内质控数据；在方法评价中，本标准不精密度包括批内和批间不精密度。

注3：

不精密度用标准差（SD）或变异系数（CV）表示。

[WS/T 403-2012，术语2.2]

3.7 允许总误差 allowable total error

TEa

指分析过程需要达到的质量目标，即单次测量或单次检测结果内不精密度（随机误差）和偏倚（系统误差）可允许的最大误差范围。

注1：

在质量控制（QC）计划中，假定无标本特异性影响，因为标本特异性属方法整体性能的一部分，非QC监测范围。

注2：

有些出版物将允许总误差称为“ATE”。

[CLSI C24 4th ed，1.4.2 Definitions]

3.8 特异性 specificity

在出现干扰现象（影响量）时，试验或检测程序能够正确地识别或定量确定某一实体物质的能力。

注：在质量控制中，是指当特定原因变异确实不存在时，质量控制系统能够指出特定原因变异不存在的概率。

[WS/T 415-2013，术语2.8]

3.9 敏感度 sensitivity

测量系统的示值变化除以相应的被测量值变化所得的商。

注1:

在质量控制中，是指质量控制系统检测误差的能力。

注2:

在定性试验中， 是指试验方法获得阳性结果的能力。

[WS/T 415-2013，术语2.7]

# 4 概述

基于患者数据的质量控制算法是通过患者数据动态监控实验室检测系统质量的算法软件，按照原理可分为统计学和机器学习两类。

在构建方法阶段，如图1所示，两类算法均利用患者历史检测数据，其中统计学方法通过数据清洗、分析处理、浮动特征值计算后，给出用于检测应用阶段的最优控制限；机器学习方法通过数据清洗、分析处理、模型训练后，构建出用于检测应用阶段的最优模型。

|  |
| --- |
| 历史检测数据  统计学方法  最优模型  分析处理  浮动特征值计算  数据清洗  机器学习方法  数据清洗  模型训练  分析处理  最优控制限 |
| 图1 基于患者数据的质量控制算法构建流程 |

在检测应用阶段，如图2所示，两类算法均以患者检测数据为输入，统计学方法通过构建方法阶段给出的最优控制限，判定检测系统的质量；机器学习方法通过构建方法阶段给出的最优模型，判定检测系统的质量。

|  |
| --- |
| 患者检测数据  在控状态/  失控状态  统计学方法  控制限判断  浮动特征值计算  分析处理  在控状态/  失控状态  模型预测  分析处理  机器学习方法 |
| 图2 基于患者数据的质量控制算法应用流程 |

# 5 计量特性

检验医学定量检测项目基于患者数据的质量控制算法的计量特性包括：假性失控率、误差检出率、错误报告患者检测结果数量中位值/平均值、算法评价值四个评价指标。

## 5.1 假性失控率

假性失控率（*PFR*）表征被测试对象的特异性，反映算法对于预期在控状态的误判概率。

## 5.2 误差检出率

误差检出率（*PED*）表征被测试对象的敏感度，反映算法对于预期失控状态的识别概率。

## 5.3 错误报告患者检测结果数量中位值/平均值

错误报告患者检测结果数量（number of patient results affected before error detection, Nped）表征被测试对象的识别速度，反映算法对预期失控状态的识别中所延误的结果数量。其定义为：在某次质量控制检测中，以引入失控状态数据的位置为起点，以算法初次识别为失控状态的位置为终点，统计两点之间的结果个数。应计算多个Nped的中位值（median of Nped, MNped）/平均值（average of Nped, ANPed)，每次质量控制检测均采用相同的检测结果数量。

## 5.4 算法评价值

算法评价值反映算法对于假性失控率、误差检出率、错误报告患者检测结果数量中位值/平均值的综合特性，其算法评价值（）表征被测试对象的准确性和可靠性。

# 6 数字量具特性

本规范所使用的数字量具包括两类参考数据集，分别为：质量在控数字量具（I）、质量失控数字量具（II）。数字量具应涵盖临床常规检验项目。

## 6.1 质量在控数字量具

质量在控数字量具是依据临床实际在控检测结果而来。每个检验项目的临床排除标准为：

剔除原始患者数据中血液肿瘤科化疗前后患者和儿科患者的数据部分；

b）剔除已知明确的错误患者结果数据和非患者标本结果数据。

每个检验项目的临床纳入标准为：

a）按照实际测试量，提取时间不少于6个月的患者数据（包含足够的节假日及季节引起的变异因素）；

b）提取的患者数据至少包括两个不同批号定标品和试剂改变。

每个检验项目的在控数据量不少于20条数据/天/维度。每条在控检测结果应保留小数点后2位，其相对标准测量不确定度不大于15%。

## 6.2 质量失控数字量具

质量失控数字量具是依据临床在控检测结果计算得来，其临床排除标准和临床纳入标准遵循6.1。不同核查场景下的质量失控数字量具计量特性不同。对于首次核查，每个检验项目的失控数据量不少于20/天/维度，质量失控数字量具包括三种类型：比例偏倚数字量具、恒定偏倚数字量具、随机失控数字量具；对于期间核查和关键点核查，每个检验项目的失控数据量不少于20/天/维度，质量失控数字量具至少包含比例偏倚数字量具类型。每条在控检测结果应保留小数点后2位，其相对标准测量不确定度不大于20%。

### 6.2.1 系统偏倚数字量具

系统偏倚数字量具包含比例偏倚数字量具和恒定偏倚数字量具两种类型。

### 6.2.1.1 比例偏倚数字量具

比例偏倚数字量具中的每条检测结果满足本检测结果在±0.5范围内（含±0.5），以0.02为间隔的偏倚，依据公式（4）：

 （4）

式中：

——比例偏倚数字量具；

*x*——质量在控数字量具；

*n*——偏倚因子，取值为±0.5、±0.48、±0.46、...、±0.06、±0.04、±0.02。

### 6.2.1.2 恒定偏倚数字量具

恒定偏倚数字量具中的每条检测结果满足±1/6、±1/4、±1/2、±1、±1.5倍TEa的偏倚，依据公式（5）：

 （5）

式中：

——恒定偏倚数字量具；

*x*——质量在控数字量具；

*n*——偏倚因子，取值为±1/6、±1/4、±1/2、±1、±1.5；

TEa——项目的允许总误差；

——所有质量在控检验结果的平均值。

### 6.2.2 随机失控数字量具

随机失控数字量具中的所有检测结果满足一个噪音分布的偏倚，对应质量失控中发生的均值和方差均波动的随机失控，依据公式（6）：

 （6）

式中：

——随机失控数字量具；

*x*——质量在控数字量具；

*e*——引入的噪音成分。

本数字量具按照所分布的噪音类型可分为含高斯分布、U形分布、均匀分布三类。

高斯分布随机失控数字量具中的高斯分布噪音均值满足±0.3、±0.5、0倍TEa，标准差满足1/3、1/2、1.5、3倍的质量在控数字量具的标准差；

U形分布随机失控数字量具中的U形分布噪音均值满足±0.3、±0.5、0倍TEa，标准差满足1/3、1/2、1.5、3倍的质量在控数字量具的标准差；

均匀分布随机失控数字量具中的均匀分布噪音上下限范围满足±0.3、±0.5、0倍TEa或1/3、1/2、1.5、3倍的质量在控数字量具的标准差。

# 7 溯源要求

## 7.1 溯源方法

溯源方法采用本规范所使用的数字量具计算假性失控概率、误差检出概率、错误报告患者检测结果数量中位值/平均值、算法评价值。每一项核查指标的计算记录格式见附录B。对于初始开机时核查，应满足：*PED*>95%，*PFR*<5%；对于过程核查，应满足：*PED*>90%，*PFR*<0.5%。

### 7.1.1 假性失控率溯源方法

采用数字量具（I）输入算法，依据公式（7）计算*PFR*。

（7）

式中：

——数字量具个数；

——第个实际判定结果，取值为0（判定在控状态）或者1（判定失控状态）。

### 7.1.2 误差检出率溯源方法

采用数字量具（II）输入算法，依据公式（8）计算*PED*。

（8）

式中：

——数字量具个数；

——第个实际判定结果，取值为0（判定在控状态）或者1（判定失控状态）。

### 7.1.3 错误报告患者检测结果数量中位值/平均值溯源方法

为了使结果更可靠，设置多个虚拟日。每个虚拟日含有相同数量的数字量具（II），数量通常为1000。在数字量具（II）的初始位置，引入一部分数字量具（I），以满足算法在计算时的数据量条件。每个虚拟日下，以数字量具（II）出现位置为起点，算法初次判定为失控状态的位置为终点，统计两点之间的检测结果数量。如果未检测出，结果数量记作数字量具（II）个数的110%，计算多个虚拟日的结果数量的中位值/平均值。检测过程应满足~~以下要求：~~数据过滤比例不超过全部数据的5%，计算步长不超过40。

### 7.1.4 算法评价值溯源方法

基于假性失控率、误差检出率、错误报告患者检测结果数量中位值/平均值，可依据公式（9）和测量不确定度评定计算得出。

（9）

式中：

，——由专家经验评估的权重系数，默认取值为2/3，1/3；

——每个虚拟日下数字量具（II）的数量。

## 7.2 溯源周期

### 7.2.1 首次核查

算法首次部署开机核查时，应实施全部检查项目：假性失控率，误差检出率，错误报告患者检测结果数量中位值/平均值，算法评价值。其中质量失控数字量具至少5万/维度，且需核查全部偏倚类型：比例偏倚、恒定偏倚、随机失控。

### 7.2.2 周期核查

算法常规使用期间，可利用日常工作产生的检验和质控数据，定期对算法的计量特性进行核查。核查周期至少1次/年。只针对假性失控率，误差检出率。其中质量失控数字量具每维度数量为建方法所用数据集的1/4倍到1倍之间，只需核查比例偏倚类型。

### 7.2.3 关键点核查

现有算法的任一要素（软件、仪器、试剂）变更影响质量控制算法计量特性的情况发生后，应在算法重新启用前对可能受影响的计量特性进行核查。关键点核查时只针对假性失控率，误差检出率。其中质量失控数字量具至少1万/维度，偏倚类型至少但不限于比例偏倚类型。影响算法计量特性的情况包括但不限于：

检测程序所在硬件更换升级；

检测程序所在操作系统升级；

检测程序所在软件环境进行其他软件的安装和删除；

检测程序本身安装升级；

人工远程自动校准；

仪器发生维修；

关键零部件更换；

关键试剂更换；

仪器搬迁；

设施和环境的严重失控。

# 附 录 A

# 临床检验分析性能质量规范要求

临床检验分析性能质量规范要求由高到低依此分为三个等级，用TEa表示，用于方法性能评价。其中

等级1 基于直接或间接大样本临床结局研究确定的分析性能质量规范

等级2 基于被检测量人群生物学变异确定的分析性能质量规范

等级3 基于方法学原理确定的分析性能质量规范

临床实验室检测项目分析性能质量规范可参考国家卫生健康委临床检验中心（National Center for Clinical Laboratories, NCCL），北京市卫生健康委临床检验中心（Beijing Center for Clinical Laboratories, BCCL）医学检验能力验证和/或质量评价计划评价标准及临床实验室改进1988修正案（Clinical Laboratory Improvement Amendments of 1988，CLIA) 能力验证计划评价标准。

# 附 录 B

溯源记录格式

B.1 假性失控记录表

假性失控记录表见表B.1。

# 表B.1 假性失控记录表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 样本类型 | 项目 | 样本序号 | 检测结果值（10^9/L） | 预期状态 | 算法判定结果 |
| 全血 | 白细胞(WBC) | 1 | 5.35 | 0 | 0 |
| 全血 | 白细胞(WBC) | 2 | 5 | 0 | 0 |
| 全血 | 白细胞(WBC) | 3 | 6.05 | 0 | 0 |
| 全血 | 白细胞(WBC) | 4 | 4.84 | 0 | 0 |
| 全血 | 白细胞(WBC) | 5 | 6.25 | 0 | 1 |
| 全血 | 白细胞(WBC) | 6 | 8.51 | 0 | 0 |
| ... | ... | ... | ... | ... | ... |

注：0指代在控状态，1指代失控状态。

B.2 误差检出记录表

误差检出记录表见表B.2。

# 表B.2 误差检出记录表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 样本类型 | 项目 | 样本序号 | 检测结果值（10^9/L） | 预期状态 | 算法判定结果 |
| 全血 | 白细胞(WBC) | 1 | 11.6 | 1 | 1 |
| 全血 | 白细胞(WBC) | 2 | 15.8 | 1 | 1 |
| 全血 | 白细胞(WBC) | 3 | 13.8 | 1 | 1 |
| 全血 | 白细胞(WBC) | 4 | 13.1 | 1 | 0 |
| 全血 | 白细胞(WBC) | 5 | 11 | 1 | 1 |
| 全血 | 白细胞(WBC) | 6 | 10.8 | 1 | 0 |
| ... | ... | ... | ... | ... | ... |

注：0指代在控状态，1指代失控状态。

B.3 错误报告患者检测结果数量记录表

可按表B.2记录每个虚拟日下的检出结果，统计出每个虚拟日的错误报告患者检测结果数量之后，再次按表B.3记录，以计算MNped。

# 表B.3 错误报告患者检测结果数量记录表

|  |  |  |  |
| --- | --- | --- | --- |
| 样本类型 | 项目 | 虚拟日编号 | NPed |
| 全血 | 白细胞(WBC) | 1 | 60 |
| 全血 | 白细胞(WBC) | 2 | 53 |
| 全血 | 白细胞(WBC) | 3 | 46 |
| 全血 | 白细胞(WBC) | 4 | 59 |
| 全血 | 白细胞(WBC) | 5 | 71 |
| ... | ... | ... | ... |