

方法验证性能参数的获取和评价方式再探讨*

朱容蝶^{1,2}, 耿颖¹, 谭德讲^{1**}, 杨化新¹, 何兰¹, 刘万卉²

(1. 中国食品药品检定研究院, 北京 102629; 2. 烟台大学, 烟台 264003)

摘要: 综述了目前国内外方法验证的法规和/或指导原则中对于方法性能参数进行验证和评价的统计学方法; 并对如何通过风险分析、实验设计、样本量和结果计算以获得更可靠性能参数等方面的问题进行了探讨。

关键词: 方法验证; 准确度; 精密度; 关键因素; 风险分析; 样本量; 实验设计; 统计分析

中图分类号: R 917

文献标识码: A

文章编号: 0254-1793(2019)02-0202-05

doi: 10.16155/j.0254-1793.2019.02.03

Further discussion about the acquisition and evaluation of performance parameters in procedure validation*

ZHU Rong-die^{1,2}, GENG Ying¹, TAN De-jiang^{1**},
YANG Hua-xin¹, HE Lan¹, LIU Wan-hui²

(1. National Institutes for Food and Drug Control, Beijing 102629, China; 2. Yantai University, Yantai 264003, China)

Abstract: The statistical methods for validation and evaluation of performance parameters of analytical procedures in domestic and foreign regulations and/or guidelines were summarized. The issues to obtain more reliable performance parameters by risk analysis, experimental design, sample size calculation, and results calculation evaluation were also discussed.

Keywords: procedure validation; accuracy; precision; critical factors; risk analysis; sample size; design of experiment; statistical analysis

方法验证是根据所检测药品质量属性的要求, 预先设定一定的验证内容和限度标准, 通过设计合理的实验来验证所采用的分析方法是否科学、合理, 并适用它所期望的应用目的^[1]。该内容在药品质量控制领域不可或缺, 故各国药典^[2-4]均制定了相关的指导原则或指南, 且由美国、日本和欧洲协调一致, 形成了具有广泛影响力的 ICH Q2(R1)^[1]。

ICH Q2(R1)的最终版本于2005年成型, 至今未作修改。各国药典基于 ICH 指导原则, 结合实际应用中遇到的问题, 进行了许多新的探索和拓展(参见本专栏《方法验证法规发展历程及存在问题》一文)。近年来, 对方法验证的研究主要集中于以下3个问题: (1) 如何科学地确定方法满足用途的标准或依据; (2) 如何获得可靠的性能参数; (3) 如何全面地

* 国家“重大新药创制”科技重大专项(2015ZX09303001)

** 通信作者 Tel: (010) 53851581; E-mail: tandj@nifdc.org.cn

第一作者 Tel: 18811714736; E-mail: rongdiezhu@163.com



描述方法验证性能参数的稳健性或变异性。

有关第 1 个问题,作者已在本专栏《定量理化分析方法满足预期用途的判断标准探讨》一文中做了专门论述。本文主要结合 ICH Q2(R1) 和最新版 USP 所刊载的 <1210>^[5],对上述的问题(2)和(3)进行比较,进而提出对验证试验的实验设计和参数描述的个人观点以供参考。

1 关于描述方法变异性和可靠性的性能参数探讨

ICH Q2(R1)^[1]与 USP<1225>^[4]除在方法分类(USP 将溶出度/释放度作为单独一类方法,而 ICH 和《中华人民共和国药典》将其与主成分定量测定合并为一类)上有差异外,对各类方法验证时所需的性能参数要求基本一致,如各类方法都要求提供专属性;对主成分定量测定的方法还要求提供准确度、精密度、线性和范围;对杂质检测方法还要求提供定量下限(limit of quantitation, LOQ)(杂质分析测定方法的定量限度)或检测下限(limit of detection, LOD)(杂质分析测定方法的检测限度)。本文主要对影响 I 类和 II 类分析方法的关键性能参数——准确度、精密度、LOD 与 LOQ 的获取和评价进行探讨,对线性的评价新增回归线标准误,范围的确定参照准确度、精密度和线性。

为了更好地描述方法的变异性和性能参数的可靠性,相比于 ICH Q2(R1),USP<1210>^[5]规定了更加明确的计算方式和新的统计评价指标,如(1)对准确度的评价增加了偏倚(bias)的置信区间;(2)对精密度的评价增加了置信上限;(3)对线性的评价增加了回归线标准误(standard error of regression line);(4)对 LOD 的计算明确规定采用线性 95% 的预测区间;(5)对方法的评价还增加了方法变异度(variability),即采用准确度和精密度联合(合成)的方式评价方法,同时还增加了用方法变异度计算的容忍区间和预测区间来总体评价方法的适用范围。

上述各统计评价指标的具体含义如下:偏倚的置信区间指在一定置信水平下,真实偏倚有一定概率落在测量结果周围的程度;精密度的置信上限指在一定置信水平下,精密度可能的最大离散程度;回归线的标准误即回归线的残差标准偏差;方法变异是指分析方法造成的变异,是准确度和精密度的合成;预测区间是指用验证试验的样本结果来预计在某一置信水平下,下一次或下几次实验时,该方法所测结果的所在范围;容忍区间是指用验证试验的样本结果来预计该方法在给定的置信水平下,预测该方法所有

检测结果中有期望概率值(desired probability value, P)所在的范围。LOD 和 LOQ 的确定,采用线性 95% 的预测区间,既充分考虑了由空白信号导致的 I 类错误与 II 类错误,同时也考虑了回归线的不确定性以及与未来观测相关的变异性^[5]。

上述改变是对 ICH Q2(R1)指南应用中所遇问题的解答,也是更加科学准确、严谨规范地保障今后方法验证研究的有力手段,是 USP 专家近 8 年的探索和实践的成果。

2 关于获取方法验证性能参数的探讨

ICH Q2(R1)对所规定的性能参数提供了多种计算方式,如对 LOD 运算有 5 种方式:基于目测、基于信噪比、基于响应和斜率的标准偏差、基于空白的标准偏差和基于校准曲线,但哪些方式更优,更能科学地保证所获参数的准确和可靠,却没有更多论述。

USP 专家近年来除补充和完善了方法验证的性能参数,也对如何可靠地获得这些参数进行了比较研究,最终从科学性和经济性的角度出发,提出了一套相对合理的实验设计和新增了准确度-精密度联合评定的方式^[5]。

但 USP<1210>提供的实例中,其实验方案和计算方式仍有不足。如主成分定量测定类方法的准确度和精密度评价时,采用 3 个浓度点,9 个实验条件的实验方案,且仅用报告值进行参数计算,有下列不足:

(1)仅可给出所选浓度范围内相对可靠的总体验证参数,若对每个浓度点的准确度、重复性、精密度和中间精密度进行计算时,因样本量小(仅 3 次重复结果),难以保证其稳健性和可靠性。

(2)验证试验中仅对 3 个浓度点进行准确度和精密度确定,如某个浓度点(如标示量的 50%)的准确度或精密度过低,则无法确定该方法的适用范围是否真正适合其评价某一质量标准的产品(如 90%~110%)。所以,选择 5 个或更多的浓度点更有利于判断不同范围内,方法能否满足其预期应用。

上述 2 点缺陷是该实验设计的先天缺陷,要解决上述缺陷,需要更为严谨的实验设计。

3 关于可靠的方法验证试验设计的探讨

针对 USP<1210>中实验设计的不足,作者从方法验证的重要性、每个浓度点准确度和精密度计算的可靠性及工作实际出发,提出方法验证试验通用的实验设计原则,见图 1。

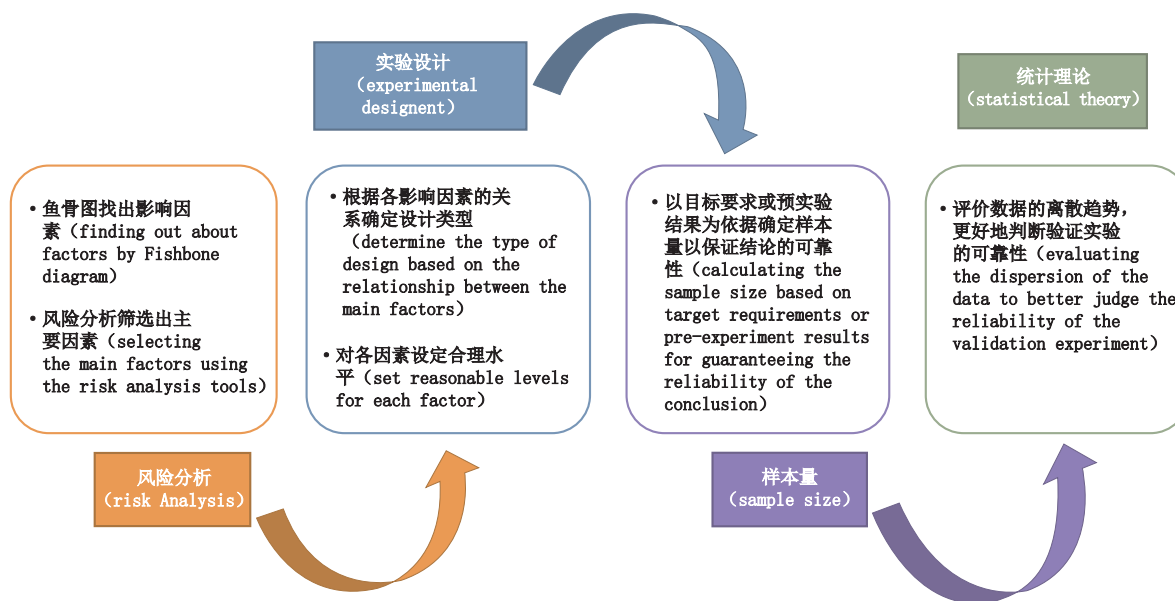


图 1 方法验证的实验设计和性能参数计算流程图

Fig. 1 Experimental design for procedure validation and flow chart for performance parameter calculation

3.1 用风险分析确定方法验证时的关键影响因素

风险分析对于方法验证中各影响因素的分析和评估非常有益^[6-7]。在方法验证试验设计中,首先通过风险分析工具,从实验的各影响因素中筛选出对准确度和精密度有重要影响的因素。具体操作

如下:

(1) 因果图(也叫鱼骨图)列举出影响分析方法的所有因素,见图 2。该鱼骨图不仅为方法验证试验确定影响因素,同时也为方法确认(verification)和转移时影响因素的筛选提供重要依据。

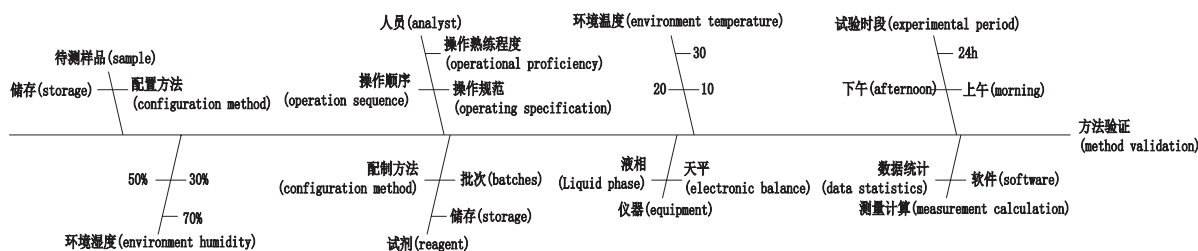


图 2 方法验证中各影响因素的鱼骨图

Fig. 2 Fishbone diagram of each influencing factor in procedure validation

(2) 按照对准确度与精密度影响的严重性和发生频率,对上述因素进行风险评估^[8],找出主要的影响因素。风险评估阶段,应选择具有丰富经验的试验技术人员,结合方法开发、验证、关键操作步骤等进行综合评估。

(3) 结合鱼骨图对影响因素进行再评价,最终确定各个因素的风险等级,选定需要进行评价的主要因素。一般在方法验证中,其主要影响因素是试验人员、主要检测仪器和关键检测试剂,其次是环境温湿度、实验时间段等。

3.2 析因设计确保性能参数的可靠性

在方法验证阶段,一般使用多因素的析因设计^[9]。根据影响因素的关系,确定用正交法^[5]还是嵌套法^[10]。图 3 是假定方法验证过程主要考察的影响因素为仪器、人员,对于其他影响因素,如关键试剂、实验环境的温湿度等,通过将其均匀分布到各影响因素中,以降低或抵消这些因素对实验的影响。将仪器和人员各设置为 2 水平,每个水平进行 2 次重复的实验设计方案。

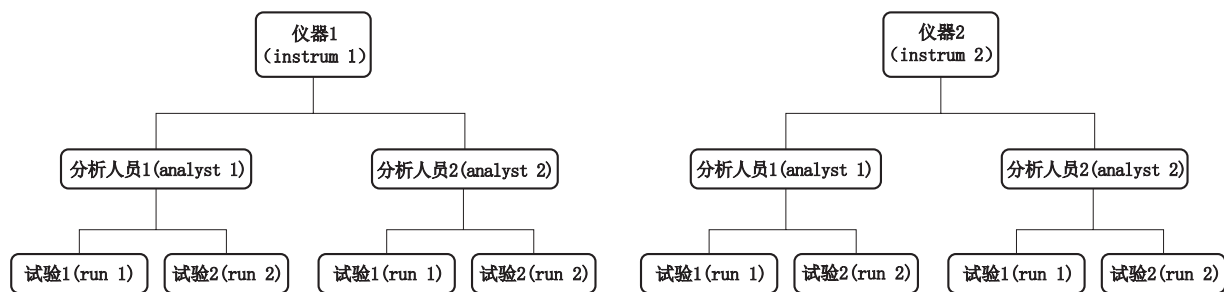


图3 方法验证的实验设计

Fig. 3 Experimental design for procedure validation

为了考察方法的适用范围,在实验设计中,也需考虑方法验证时的浓度选择。

对于含量测定类方法,若所检测的产品质量标准为 90.0%~110.0%,可用以标示量为中心的 5 个浓度点,如 80%、90%、100%、110%、120%。若所考察的方法同时为溶出度服务,则需在标示量 1%~120% 的浓度范围内选择合适的浓度点,以便对溶出曲线进行评价。

对于杂质定量测定类试验的验证,浓度范围的选择需要考虑 LOD 和产品质量限度标准;如假定 LOD 为 $0.5 \mu\text{g} \cdot \text{mL}^{-1}$,质量标准规定的产品限度值为 $6.0 \mu\text{g} \cdot \text{mL}^{-1}$,则至少需要从 $0.5 \sim 8.0 \mu\text{g} \cdot \text{mL}^{-1}$ 的范围内设置不少于 6 个浓度,如 0.5、1.0、2.0、4.0、6.0、8.0 $\mu\text{g} \cdot \text{mL}^{-1}$ 。

3.3 计算样本量以保证所获结果的科学性和可靠性

方法验证试验中,足够的样本量不仅可以提供准确和可靠的方法性能参数,而且可以确保试验结果的把握度和完整性^[11],对于判定各因素对分析方法的影响大小也很重要。但样本量过大,则浪费时间、人力和成本^[12]。

要获得足够的样本量,需要结合前期的预实验数据,通过科学的假设进行统计计算。一般可采用基于效能确定样本量法。通常,选择在 5% 的显著性水平下,90% 或 80% 的效能。由此固定的样本量大小可以有效地避免 I 类错误,同时又可以减少 II 类错误,而得到的准确度和精密度更加可靠,可更好地保证此方法的检测目的。具体的计算参考如下公式^[10]:

$$n \geq \frac{(t_{\alpha, df} + t_{\beta/2, df})^2 \sigma_{ip}^2}{\theta^2}$$

$t_{\alpha, df}$ 和 $t_{\beta/2, df}$: 是指学生 t 分布中的分布点值; α 和 β : 分别为单侧 I 型和 II 型错误发生率,代表了在验证研究中做出错误结论的风险大小; df : 验证

中实验设计的自由度,通常为 $n-1$; σ_{ip}^2 : 是精密度(中间精密度)的初步估值; θ : 是可接受的偏差值(acceptable deviation)。

在实际应用中,一般都使用统计软件(JMP、Minitab 等)进行计算。使用者输入相应的 α 、效能 $(1-\beta)$ 、 θ 和 σ_{ip} 值,选择相应的应用目的,软件可根据使用者的诉求快速计算出所需的样本量。

4 讨论和结论

本文结合 ICH Q2(R1) 和 USP 的最新研究进展,主要讨论了在方法验证中,哪些方法性能参数可以更好地描述方法的适用性,如何科学可靠地获得这些参数和如何对其科学地计算。从中可看到:

(1) ICH Q2(R1) 需要进一步修订完善。其列举的实验设计过于简单,未从整体考虑;给出的统计评价参数不能从整体把握方法的变异性。

(2) USP<1210> 列举的实例中,实验设计虽简便高效,但仅选择 3 个浓度点,易产生误判风险,也难以给出各浓度点可靠的准确度和精密度。建议采用 USP<1033> 的实验设计方案。

(3) 为保证方法验证研究的科学性和可靠性,建议在方法验证报告中提供如下参数:①准确度及其变异,即偏倚(或相对偏倚)及其置信区间;②各精密度及其置信上限;③线性描述参数中的斜率、截距、相关系数(r)或决定系数(R^2)和回归线的标准误;④回归线 95% 的预测区间计算的 LOD。

(4) 在性能参数限度标准已确定的前提下,为有充足的把握表明试验结果的可靠性,需确定试验的样本量,进而确定条件数(c)与重复数(r)。一般增加实验条件数比仅增加重复数更能获得可靠的性能参数^[13]。

(5) 为从整体判断把握方法是否满足预期用途,建议在方法验证报告中提供新的统计评价指标,如

容忍区间、预测区间或方法能力指数(见本栏中《定量理化分析方法满足预期用途的判断标准探讨》一文)。这些参数的提供,对使用者和监管者今后更好地利用和评价该方法都非常重要,故也建议能在行业中尽快推广。

参考文献

- [1] ICH Q2(R1) Harmonised Tripartite Guideline Validation of Analytical Procedure: Text and Methodology [S]. 2005
- [2] 中华人民共和国药典 2015 年版. 四部 [S]. 2015: 374 (9101) ChP 2015. Vol IV [S]. 2015: 374 (9101)
- [3] OMCL Network of the Council of Europe Quality Assurance Document [EB/OL]. (2014) [2018-12-8]. <https://www.edqm.eu>
- [4] USP 41-NF 36 [S]. 2018: 7665 (<1225>Validation of Compendial Procedures)
- [5] USP 41-NF 36 [S]. 2018: 7623 (<1210>Statistical Tools for Procedure Validation)
- [6] ICH Q9 Harmonised Tripartite Guideline Quality Risk Management [S]. 2005
- [7] International Organization for Standardization. ISO/DIS 31000 Risk Management — Guidelines [S]. 2017
- [8] YYT 0316-2008/ISO14971 中华人民共和国医药行业标准. 医疗器械 风险管理对医疗器械的应用 [S]. 2009
YYT 0316-2008/ISO14971 Pharmaceutical Industry Standards of People's Republic of China. Medical Devices— Application of Risk Management to Medical Devices [S]. 2009
- [9] 谭德讲, 马双成. 药品监督与检定中的统计学应用 [M]. 北京: 中国科学技术出版社, 2011: 58
TAN DJ, MA SC. Statistical Applications in Biopharmaceutical Regulatory Science [M]. Beijing: China Science and Technology Press, 2011: 58
- [10] USP 41-NF 36 [S]. 2018: 6803 (<1033>Biological Assay Validation)
- [11] CHOW SC, SHAO J, WANG HS. Sample Size Calculations in Clinical Research Second Edition [M]. London: Chapman & Hall / Crc. 2008: 142
- [12] NOORDZIJ M, DEKKER FW, ZOCCALI C. Sample size calculations [J]. Nephron Clin Pract, 2011, 118(4): 319
- [13] <1210> Validation of Compendial Procedures [J/OL]. USP Pharmacopeial Forum, 2014, 40(5) [2018-12-01]. <http://www.usppf.com/pf/pub/index.html>

(本文于 2018 年 11 月 25 日收到)

《药物分析杂志》编辑部声明

本刊采用在线投稿系统,作者稿件一经本刊审核通过,确定录用,可优先数字出版,同时被中国学术期刊网络出版总库等数据库收录,进入因特网提供信息服务,并通过本刊在线系统等实现全文查询。本刊所付稿酬包含刊物内容上网服务报酬,不再另付。

本刊未委托其他任何机构或个人代理征收稿件,所有稿件须登录本刊网站 (<http://www.ywfxzz.cn>) 在线投稿,并须提交加盖公章的单位介绍信。

本刊未委托其他任何机构或个人代收任何费用,所有收费按本刊缴费通知办理。