

标准研讨

四参数模型在生物效价计算中的注意事项^{*}谭德讲¹, 黄维金¹, 朱容蝶¹, 王萌¹, 杨化新^{1**}, 周航宇², 马莉¹

(1. 中国食品药品检定研究院, 北京 102629; 2. 北京中医药大学东直门医院检验科, 北京 100700)

摘要: 通过调研目前国内法规和使用四参数模型计算效价的实验室数据, 对生物效价测定在使用四参数模型计算中常见问题总结如下: (1) 《中华人民共和国药典》虽然收录了多个品种的四参数模型计算要求, 但对使用该模型进行效价计算时缺乏具体要求。(2) 国内实验室在使用该模型计算生物效价时, 在实验设计、系统适用性和样本适用性等方面均存在评价不足, 没有提供充分的数据分析及结果, 从而导致结果的可信程度下降。(3) 由于差异性检验的自身缺陷, 建议今后使用该模型时, 对样品适用性的统计检验逐步过渡到使用等效性检验, 以保证计算的准确性和科学性。

关键词: 生物效价; 四参数模型; 模型选择; 系统适用性; 样本适用性

中图分类号: R 917

文献标识码: A

文章编号: 0254-1793(2018)07-1238-05

doi: 10.16155/j.0254-1793.2018.07.20

Some attention issues in potency calculating using 4-parameter model^{*}TAN De-jiang¹, HUANG Wei-jin¹, ZHU Rong-die¹, WANG Meng¹,
YANG Hua-xin^{1**}, ZHOU Hang-yu², MA Li¹

(1. National Institutes for Food and Drug Control, Beijing 102629, China; 2. Dongzhimen Hospital, Beijing University of Chinese Medicine, Beijing 100700, China)

Abstract: After investigating the current domestic regulations and the original data & calculation using four-parameter model from labs, some common problems were summarized as follows: (1) Although some kinds of products' monographs have been included in Pharmacopoeia of the People's Republic of China (ChP) which uses four parameter model for potency calculation, the monograph in ChP lacks specific requirements for the calculation. (2) There were some insufficiency assessment in experiment design, system suitability and sample suitability in some domestic labs when using the model to calculate, which led to the decrease of the credibility of the results. (3) Owing to the disadvantages of difference test, it suggested that the statistical test should be gradually transitioned to the equivalent test for assessment of the suitability of the sample to promote the accuracy and scientificity of the calculation, when the model was used for calculation of potency.

Keywords: potency; 4-parameter model; model selection; system suitability; sample suitability

* 药品医疗器械审评审批制度改革专项课题(ZG2016-4-05); 国家十二五“重大新药创制”科技重大专项(2015ZX09303001)

** 通信作者 Tel:(010)53851655; E-mail: yanghx@nifdc.org.cn

第一作者 Tel:(010)53851581; E-mail: tandj@nifdc.org.cn

体内外效价测定是生物制品活性评价的重要指标。随着现代生物技术和国际社会对实验动物的 3R 原则的实施,在效价测定中,体外生物学活性测定法越来越受到重视^[1-3],并逐步成为日常检验的主要手段,避免了使用整体动物产生的巨大个体差异,同时可以更好地反映活性的全反应域,从而更好地了解其活性特点。

由于生物活性在全反应域中,剂量-效应关系多呈 S 曲线,故在描述和计算效价时,需要使用描述 S 曲线的统计拟合模型。描述 S 曲线的统计模型有很多^[4-5],如 Logistic 模型(包括四参数、五参数和三参数形式)、Weibull 模型、四参数 Gompertz 模型等。要获得最佳的拟合曲线,需要有严格的实验设计和便于使用的统计分析工具。目前美国药典(USP)和欧洲药典(EP)都选择使用四参数方法进行拟合并计算的。典型的四参数模型公式,如 Rodbard 公式为 $y=D+\frac{A-D}{1+(\frac{x}{C})^B}$ (USP 收录),其公式中的 4 个参数:A、D 分别表示 S 曲线的上、下渐近线上的值,即最高和最低生物反应量值,C 表示半数反应剂量,B 表示在半数反应量附近描述曲线陡峭程度的斜率。

国内在使用四参数模型进行实验分析和效价测定中,常常没有考虑到前期的实验设计具体问题,也没有对实验数据是否符合该拟合模型以及是否与标准品曲线平行等问题进行比较。本文结合实验室在日常使用中的这些问题,对使用该类模型进行生物效价测定中的实验设计要求、统计要求、拟合模型的选择和评价及其常见问题进行分析,希望对大家今后的准确把握有益。

1 关于满足生物效价数据分析的三大前提假设问题

实验数据的分析,首先要求做到数据具有独立性(independence)、正态性(normality)、方差恒定(constant variance)。只有满足上述三大前提^[6],所计算的结果方可保证具有可靠性。

1.1 数据的独立性

数据的独立性是统计分析的基本要求,一般通过实验设计和具体操作要求进行保证。在细胞类生物效价测定中,常见问题是(1)没有评价样品是否具有板效应和位置效应等影响;(2)各样品不同剂量组缺少重复样本;(3)重复均出自同一个样本。

避免板效应和位置效应的问题,要求实验者必须进行相应的实验评估。只有样品测定在培养板中位

置效应不明显,方可按顺序排放样品,否则,应按照均衡分配各样品在培养板中的不同位置,以达到减小系统操作误差的目的。

关于缺少重复样本,由于生物实验的变异大,当只有 1 个数据时,对模型的判断会因 1 个点的数据偏离而导致模型失效;另外,不能很好地给出各浓度内的精密度评估。

对重复而言,由于该类实验多通过不同倍比的稀释等方式测试多个浓度,故对每个浓度系列而言,都无法避免其操作系统误差;这时,使用 2 个以上的高浓度样本进行独立的系列稀释,可相应地弥补其不足。

1.2 数据的正态性和方差恒定

对于生物效价实验,不仅变异较一般理化实验的大,而且其结果往往呈对数正态分布或其他类型的分布,这时需要对原始数据进行相应的转换,使之符合正态分布方可进行分析,如对正偏态的数据使用对数转换,对于泊森分布的数据使用平方根转换等;如果转换后,还达不到要求,需再采用 Dixon 法^[7]等进行必要的离群值(outlier)检查和剔除,使其达到正态性和方差恒定的要求。

2 关于方法的系统适用性评价问题

对生物活性测定方法的统计分析,首先需要根据所检测产品的剂量-反应曲线选择合适的统计分析模型。所谓系统适用性^[6]评价,就是对数据所拟合的曲线是否具有很好模型符合性的评价,或者在初期确定最佳曲线模型的问题。

2.1 模型选择问题

严格意义上,在方法建立的初期,应该首先使用标准物质或典型产品的检测进行模型筛选,并经过多次试验的结果进行验证或联合评价,这样才能获得最准确、可靠的模拟计算模型。由于该过程需要复杂的计算,故多需要相应的统计软件协助完成,在 SoftMax、JMP、PLA 等统计软件,均有多种曲线模型供比较和选择。

采用四参数模型可满足解释生物反应的要求:(1)生物实验的反应区间有一定范围,四参数法恰恰具有上下渐近线,能很好地表达这种反应;(2)其 4 个参数具有现实且比较易于理解的意义,除上、下渐近线表示更低剂量和更高剂量的生物反应趋于稳定外,半数反应量可反映出相应的生物效应强弱,而其斜率反映了在半数反应量周围的生物效应特征;

(3) 在等效范围内,选择四参数模型替代如五参数等更复杂的模型形式,虽会失去一些效能,但基本可以正确反映出生物反应的特征。

实际应用中,四参数模型的公式也很多,如 Rodbard 四参数模型和 Logistic 四参数模型(两者具有本质一致性,仅因 1 种是剂量不进行对数转换,1 种采用对数转换,所计算的参数值不同而已)、Richard 四参数和 Gompertz 模型(生长模型),还有 Weibull 模型等。这些模型均可模拟曲线特征,但对特定数据组类型有效能差异或拟合优劣的一些差异,故 USP 和 EP 均对所选用的方程进行了限定;EP 选择严格意义的四参数 Logistic 模型,而 USP 选择了不需对剂量进行对数转换的 Rodbard 四参数模型。

2.2 四参数模型拟合中的浓度设置问题

四参数模型进行系统性评价中,国内常见的问题是所用检测浓度点设置问题,包括浓度点的分布和浓度点的间距选择 2 个方面。

2.2.1 浓度点的分布问题 参数模型一般要求浓度点至少比参数多 1 个点,如对四参数模型而言,至少要求有 5 个浓度点。然而,要想得到稳健可靠的拟合,该模型至少使用 7 个浓度测定点,且其中上下渐近线上至少各有 2 个点,在拐点周围要求有至少 3 个点。使用过少的点,得出的模型参数可靠性或稳健性相对较差。但通常人们都推荐使用 8 个以上的浓度点。

2.2.2 浓度点的间距选择问题 有时,会发现实验者使用了很多个浓度点,浪费了很多资源,然而,还是没有获得稳定可靠的参数,为何呢?通过曲线拟合可以发现,所用浓度没有很好地达到全反应域的剂量(分布),给模型的拟合造成了损失所致。故在实验方法建立初期,一定要用标准品或典型样品找出其全反应域的剂量区间,然后调整所用剂量使之符合浓度分布点的最低要求。有时固定使用 1:1 的倍比稀释并非全适合,开始可选用 1:2 或 1:3 的倍数稀释。待确定出全反应域后,根据实际操作的简便性,可调整到 7~8 个浓度。

2.3 判断模型拟合的优劣(拟合优度)的指标问题

验证模型的有效性应进行统计检验,目前广泛使用^[8]的有模型的确定系数、失拟(lack-of-fit, LOF) F 检验等。

2.3.1 决定系数(R^2) 使用皮尔森相关系数在用于评价模型是否达到较优的拟合中并非完全恰当,这是

因为(1)精密度很低的数据样本和精密度很高的数据样本可获得相同的相关系数;(2)要求相关系数具有一定的统计显著性也没有很大意义,因为即使 1 个很弱的相关系数(如 0.2),当样本量足够大时,也可具有统计显著性;而且不可能对相关系数确定这样 1 个统计显著性的阈值。目前主要的应用软件均提供相关系数的平方(R^2),即决定系数来判断,如当 R^2 为 0.8 时,可判断模型可解释大约 80% 的数据。

2.3.2 失拟(LOF)的 F 检验 另 1 个常见描述拟合优度的指标是对 LOF(lack of fit)进行检验,也存在一些问题,如(1)LOF 检验要求在校准曲线的多个浓度水平上具有独立的重复检验;这与要求操作简便、经济相矛盾;(2)一些无关紧要的数据偏离即可导致假定的模型产生显著的失拟,而一些重要的偏离也可能因样本量不足而无法发现其统计显著性。

2.3.3 等效性检验(TOST)和校正赤迟信息准则(AICc) 最新研究建议^[8],应使用 TOST 和 AICc(corrected akaike information criterion)进行拟合优度的分析。这 2 种判断方法,可以很好地避免上述在拟合优度判断上的缺点,更科学地保证所得模型的拟合性能。

3 关于方法的样品适用性评价问题

生物活性测定方法的样品适用性问题是正确计算效价的重要组成部分,它包括对样品的相似性(similarity)或平行性评价和浓度范围 2 个方面的评价。这里主要谈目前存在问题较多的相似性或平行性问题,即是判断检验样品与标准品是否具有同源性的问题,是否可计算其相对效价的问题。对于生物大分子而言,很多结构往往会有差异,且会是多种同类物质的混合物。故这里的同源性,主要是指来源一致,而非分子结构完全相同。生物检定的模型,要求其生物活性物质基础必须同源才能计算其相对效价;非同源的物质,两者肯定不具有平行性,但应牢记:统计检验有平行性,仅说明两者具有“相同的剂量-反应模式”,未必就能保证完全同源。只有从溯源和物质基础上保证后,再获得平行性,方可获得其可靠的相对效价。关于统计的平行性分析问题,应注意平行线比较的参数和平行线检验的方式问题。

3.1 平行性所比较的参数

对四参数 Logistic 回归模型而言,常用平行性比较的参数有(1)使用四参数中的斜率和上下渐近线进行比较;(2)使用非平行性方差进行统计比较,如

PLA 2.0;(3)其他参数:使用通过额外部分的平方和与独立部分的平方和做参数进行比较,如 SoftMax;使用残差和置信区间,如 Statlia。使用不同的软件,计算结果不同(主要是所用参数不同所致),无法相互进行比较,故在日常应用中,应选择同款软件(参数不会变化),以便今后的相互比较。

3.2 平行性检验的方式

按照传统统计方法,一般使用差异性检验中的 F 检验和卡方检验。目前,EP 使用 F 检验,认为 F 检验适于各种情况,也容易进行分析;而卡方检验对权重非常敏感,所得结论在整个方法的生命周期中难以稳健。差异性检验(F 检验和卡方检验)的问题是确定出可靠的显著性水平。然而,USP 经过组织国际专家的论证,认为使用等效性检验更合适^[9],故其更推荐使用等效性检验。

一般在方法建立的初期,由于数据较少,难以确定出相应的等效区间,故建议先使用差异性检验,待获得足够的的数据后,逐步确定出科学的等效区间(见“4”项)后,再使用等效性检验方法。

目前 USP 和 EP 已经收录了等效性检验的方法,国内还没有法规等倡导使用,故应用较少见,但随着监管的要求逐步提高,最终会要求使用者应用该方法的。

在进行平行性的等效性检验中,不同的软件采用的参数及对参数的比较方式各不相同,如 XYMP 软件,采用相对差值法,即

$$\text{斜率 } B \text{ 采用 } \Delta B = 100(B_{\text{Test}} - B_{\text{Ref}}) / B_{\text{Ref}} \times 100\%$$

$$\text{下渐近线 } D \text{ 采用 } \Delta D = (D_{\text{Test}} - D_{\text{Ref}}) / D_{\text{Ref}} \times 100\%$$

$$\text{反应窗使用上渐近线 } \Delta A = (A_{\text{Test}} - A_{\text{Ref}}) / A_{\text{Ref}} \times 100\%$$

而杨鹤飞^[10]则使用一些参数转换后的比值法:

上渐近线的比值(ratio of upper asymptotes):

$$r_1 = A_{\text{test}} / A_{\text{Ref}}$$

$$\text{反应窗的比值(ratio of effective windows): } r_2 = (A - D)_{\text{test}} / (A - D)_{\text{Ref}}$$

$$\text{在半数反应量周围的斜率比值(ratio of slopes at EC}_{50}\text{): } r_3 = [(A - D) \times B]_{\text{test}} / [(A - D) \times B]_{\text{Ref}}$$

4 等效性检验的等效区间计算问题

采用等效性检验的关键是确定出科学、可靠的等效区间。这是目前等效性检验难以实施的 1 个瓶颈问题。

从文献[6, 11]看,主要有 4 种获得等效区间的方式:(1)利用所获参数历史数据的容忍区间来确定

其等效区间;(2)利用所获参数历史数据的广义置信区间来计算其等效区间;(3)利用 ROC 曲线将标准品与失败样品的历史数据进行分析,找出 2 种样品的分界点(cut-off point),作为其等效区间。(4)根据已有经验进行初步确定。

5 关于四参数模型中要求输出的统计指标问题

上述所论,生物效价计算中的四参数 Logistic 模型最终以等效性检验为最佳选择,而方法的开始阶段因结果数据较少,难以实施,故建议从方法应用的开始阶段即能输出如下指标,以便实验者能有效判断结果的科学可靠性:

每次实验的各剂量组的描述性统计指标:反应均值(mean)、标准差(SD)和 RSD。

每次实验的拟合分析指标:包括使用共同斜率和渐近线的限制性模型分析结果和非限定性模型的分析结果,即各计算方式的上下渐近线及斜率的计算值、均方误差或 RMSE、决定系数(R^2)、AICc 值或 TOST 结果等。

每次实验的方差分析指标:包括不同处理间的总方差及 F 检验结果,各不同区组处理间(包括制品间)、回归、非线性、非平行性方差及其 F 检验结果、残差。

对系统适用性的评价结论:标准品在模型中的拟合优度(LOF)结果、 R^2 结果和 AICc 结果,及其各自相对应的结论。

对样品适用性的评价结论:样品在模型中的 LOF 结果、 R^2 结果和相对于标准品的非平行性结果,以及各自相对应的结论。

等效性评价指标的输出:根据拟采用的等效性评价指标,将每次的结果值一同输出以供记录和今后比较。

相对效价的输出:需要输出待测样品相对于标准品的相对效价比(potency ratio)、95% 的置信区间和相对置信区间;并按照稀释关系,求出待测样品原液中的相对效价、95% 的置信区间;最后给出相应的 EC_{50} 值。

6 讨论

本文对生物制品中新兴的一种生物效价分析—使用四参数 Logistic 回归模型进行计算时,从实验设计、模型选择和参数比较等一些常见问题进行了汇总、分析和阐述。

在实验设计方面,需在实验前充分考虑数据分析

的基本要求,均衡各随机误差项等,从而保证分析的可靠性。

在模型选择方面,由于生物活性产品本身的复杂性,其产生的生物反应本身也很复杂,在选择计算模型时,只能是近似模型,无法达到完美的一致,且任何模型都会存在限制和陷阱^[12],故在进行模型选择时,应充分查阅国内外相关文献,通过实验室数据的长期积累,以及经过多种统计模型比较后,最终确定1个最接近的模型;这正是各国药典等法规文件或方法研发单位规定使用同一模型进行某种方法评价的原因所在。本文所讨论的四参数模型,是一类适于常见生物反应的评价方法,实验者应清楚,有些生物反应,使用五参数回归更适合,但因计算上的复杂性和四参数回归的损失不影响其根本判断,故一些法规仍推荐使用四参数回归。

在所比较的参数方面,还未见使用不同参数导致的结果差异的文献,且目前国际上的商用和内部使用统计软件(本文介绍)都被认可。但从实验室的长期比较和观察考虑,应固定使用同一软件。待积累较多数据后,可进行各种参数比较的分析。

等效性检验在国内还没有被推广应用,但由于其结论的科学性、可靠性,将很快会被监管部门重视,所以,建议使用四参数回归模型进行分析的实验室,尽量早掌握。

参考文献

[1] SCHWARZ TF, SPACZYNSKI M, SCHNEIDER A, *et al.* Immunogenicity and tolerability of an HPV-16/18 AS04-adjuvanted prophylactic cervical cancer vaccine in women aged 15-55 years[J]. *Vaccine*, 2009, 27(4): 581

- [2] COUCH RB, DECKER WK, UTAMA B, *et al.* Evaluations for *in vitro* correlates of immunogenicity of inactivated influenza A H5, H7 and H9 vaccines in humans[J]. *PLoS One*, 2012, 7(12): e50830
- [3] CUMBERLAND WN, FONG Y, YU X, *et al.* Nonlinear calibration model choice between the four and five parameter logistic models[J]. *J Biopharm Stat*, 2015, 25(5): 972
- [4] Molecular Devices Corporation. *Softmax Pro User's Manual* [EB/OL]. [2017-10-15]. <http://www.docin.com/p-241442998.html>
- [5] LEE J, BOTTJE WG, KONG BW. Genome-wide host responses against infectious laryngotracheitis virus vaccine infection in chicken embryo lung cells[J]. *BMC Genomics*, 2012, 24(13): 143
- [6] USP 40-NF 35[S]. 2017: General Information <1032>. Design and development of biological assay
- [7] USP 40-NF 35[S]. 2017: General Chapter <111>. Design and analysis of biological assay
- [8] USP in Process Revision <1210>. Statistical Tools for Procedure Validation [J/OL]. *Pharmacoepial Forum* 2014, 40(5) [2017-10-15]. <http://www.usppf.com/pf/pub/index.html>
- [9] 谭德讲,冯国双,朱容蝶,等. 一致性评价所用统计分析方法辨析[J]. *中国新药杂志*, 2017, 26(24): 2881
- TAN DJ, FENG GS, ZHU RD, *et al.* Analysis of the meanings of equivalence and the required statistical evaluation methods[J]. *Chin J New Drugs*, 2017, 26(24): 2881
- [10] YANG H, KIM HJ, ZHANG L, *et al.* Implementation of parallelism testing for four-parameter logistic model in bioassays[J]. *PDA J Pharm Sci Technol*, 2012, 66(3): 262
- [11] 冯国双,朱容蝶,谭德讲,等. 药品检验中等效区间的确定方法与应用[J]. *中华医学杂志*, 2017, 97(48): 3835
- FENG GS, ZHU RD, TAN DJ, *et al.* The determination of equivalence limits and application in pharmaceuticals assays[J]. *Natl Med J China*, 2017, 97(48): 3835
- [12] CARDOT JM, DAVIT BM. *In vitro-in vivo* correlations: tricks and traps[J]. *AAPS J*, 2012, 14(3): 491

(本文于2018年6月8日修改回)