

生物标志物免疫分析方法学验证难点:稳定性

满素勤^{1,2}, 张雪^{1,2}, 原月梦^{1,2}, 朱晰^{1,2}, 刘丹^{1,2*}

1. 上海益诺思生物技术股份有限公司, 上海 201203

2. 国家上海新药安全评价研究中心, 上海 201203

摘要: 在生物标志物分析方法学验证项中, 稳定性验证是难点。与药动力学分析不同的是, 生物标志物分析待测物为内源性物质, 其稳定性与标准品(一般为重组蛋白)有很大不同。所以, 生物标志物稳定性验证样品应采用内源性真实样本。但是受分析方法变异影响, 内源性真实样本的准确浓度无法得知, 而且难以获取新鲜的未冻存过的样本, 最终导致稳定性的接受标准难以界定。运用趋势控制图和计算机分析等手段可以更准确地判断生物标志物的稳定性。

关键词: 生物标志物; 验证; 稳定性; 趋势分析

中图分类号: R967 文献标志码: A 文章编号: 1674-6376(2020)11-2340-05

DOI: 10.7501/j.issn.1674-6376.2020.11.033

Stability: challenge in biomarker immunoassay method validation

MAN Suqin^{1,2}, ZHANG Xue^{1,2}, YUAN Yuemeng^{1,2}, ZHU Xi^{1,2}, LIU Dan^{1,2}

1. Shanghai InnoStar Bio-Tech Co. Ltd., Shanghai 201203, China

2. National Shanghai Center for New Drug Safety Evaluation and Research, Shanghai 201203, China

Abstract: In biomarker bioanalytical method validation, one of the most challenges is stability. Compared with pharmacokinetic (PK) assays, the analytes of biomarker assay is endogenous. The stability of endogenous analytes is differed with its corresponding recombinant reference standard. To that end, biomarker stability evaluation should be performed using endogenous incurred sample. However, the accurate concentration of endogenous sample is difficult to assess due to the assay variation. In addition, it is hard to acquire fresh unfrozen sample. All above factors lead to the difficulty of defining acceptance criteria for stability evaluation. Trending chart and computational analysis could be used to determine biomarker stability more accurately.

Key words: biomarker; validation; stability; trending analysis

生物标志物在药物研发周期中起着重要作用, 它可以被用来评估药物的作用机制、筛选病人、评估药物的安全性和药效、选择药物剂量和药物监测。生物标志物的检测数据可用于探索性研究、内部决策和监管审查。是否能有效地将生物标志物应用于上述目的取决于生物标志物检测数据的质量, 而数据的质量则依赖于生物标志物分析方法的完全验证^[1]。

2018年版美国食品药品监督管理局(FDA)生物分析方法验证指导原则指出: 生物标志物分析方法的验证应阐述和药物分析方法验证相同的问题, 准确性、精密度、灵敏度、选择性、平行性、定量范

围、重复性和稳定性是界定方法的重要特征, 但是FDA也承认生物标志物方法验证的特殊性^[2]。在所列验证项中, 稳定性验证是生物标志物与药物分析方法学验证差异较大的部分, 也是生物标志物验证中的难点。主要原因是: 首先, 生物标志物分析待测物为内源性物质, 其稳定性与标准品(对于大分子而言一般为重组蛋白)有很大不同; 其次, 受分析方法变异影响, 内源性真实样本的准确浓度难以测定; 另外, 难以获取新鲜的未冻存过的样本作为“0”时间点的 T_0 样本(基线水平, 名义上的零时间点样本, 未经稳定性条件处理); 最后, 如何界定稳定性的接受标准需要仔细考量。

收稿日期: 2019-12-23

基金项目: “十三五”重大专项(2018ZX09201017-008)

第一作者: 满素勤(1989—), 女, 博士, 主要从事药动力学及药效动力学研究。Tel: (021)60211999 E-mail: sqman@ncdser.com

*通信作者: 刘丹, 主要从事药代动力学及药理学研究。E-mail: dliu@ncdser.com

本文以配体结合试验(ligand binding assays, LBA)为基础,探讨了生物标志物分析稳定性验证的独特性及挑战,综合分析相关文献中的稳定性验证方法,阐述如何科学的进行生物标志物分析稳定性验证。特殊类型的生物标志物(基于流式的生物标志物, RNA等)的稳定性验证更为复杂,不在本文讨论范围。

1 生物标志物分析与药动学(PK)生物分析稳定性验证项的差异

影响生物标志物稳定性的因素较多,如时间、温度、湿度、是否有降解酶的存在、天然半衰期、保存条件、基质、光照和容器吸附等^[3]。因此生物标志物分析中稳定性验证对于研究样本的准确测定至关重要。尽管生物标志物分析稳定性验证可参照PK生物分析,但是二者又有显著不同,其不同之处主要体现在验证样品不同、标准品是否具有代表性、接受标准差异。详细总结于表1。

2 内源性蛋白与重组蛋白不同

生物标志物免疫分析所用的标准品一般为重组蛋白,它和内源待测物并非一致,其不同之处主要表现在蛋白序列、折叠、糖基化修饰、纯度等方面。内源性生物标志物不仅可能存在多种异构体形式^[4],而且可能受疾病状态的遗传学、环境、基质等影响而产生独特的转录后修饰,使得其与标准品之间的差别进一步增大^[5]。最后,标准品和内源待测物在免疫分析中与试剂(捕获抗体、检测抗体等)的结合特性可能有所不同^[4]。

基于此,在生物标志物分析中,如果标准品不能充分代表内源性待测物会带来一定的风险。根据分析结果使用目的不同,风险可以是对数据解读的影响,也可以是影响决策、项目进度和资源分配。

在严重的情形下,测得的结果甚至可能不是目的待测物^[4],如CUZD1的ELISA试剂盒测的其实是CA125^[6]。因此在为生物标志物分析方法选择标准品时,不仅要尽量选择经过充分表征的标准品,还应考虑如下因素:内源性待测物的理化特性(如蛋白是以单体还是二聚体亦或三聚体存在);生物标志物的异构体形式、转录后修饰、剪切等;重组的标准品的表达系统(真核还是原核);蛋白的错误折叠和失活;批次之间的差异;在稳定性方面重组的和内源的蛋白是否不同^[4]。

2015年,辉瑞的科学家比较了重组的和内源待测物的冻融稳定性和长期冻存稳定性。冻融稳定性结果显示,经过1次冻融,重组转化生长因子- β 1(TGF- β 1)回收率在87%~110%,而内源性TGF- β 1回收率在5%~96%,内源性TGF- β 1对冻融的耐受较差。在另外一项研究中,长期冻存稳定性结果显示,重组白细胞介素(IL)-13可保持4个月稳定,而内源(病人样本)IL-13稳定性长达15个月。由此作者得出结论,重组纯化的蛋白不能反映内源蛋白的稳定性,在生物标志物稳定性验证中,应尽可能地采用真实的病人样本^[7]。

美国FDA发起的Critical Path Institute(C-Path)专家组共识里也指出:在可行的情况下,无论何时,应使用天然的基质和内源待测物制备稳定性样本,因为重组的蛋白会得出错误的结果^[3]。

总之,生物标志物的稳定性验证应采用内源性真实样本而非重组的标准品加样样品。尽管重组蛋白的稳定性评估不能代表内源生物标志物的稳定性,但是对重组蛋白的稳定性评估也有一定的价值,如研究项目持续时间较长,批次之间的样品分析产生不一致性,此时重组蛋白稳定性数据有助于

表1 比较生物标志物分析与PK生物分析稳定性验证

Table 1 Stability verification of comparative biomarker analysis and PK Bioanalysis

比较研究	PK生物分析	生物标志物分析
验证样品	真实基质中加入标准品,即加样样品	真实基质中加入标准品(含内源待测物和加样蛋白)和真实基质(内源待测物),以后者作为主要验证样品
标准品代表性	标准品与待测物一致,能充分代表待测物的稳定性	标准品稳定性与内源待测物稳定性不同,标准品不能代表或只能部分代表待测物稳定性
验证内容	采样过程、试验台、冻融、短期冻存、长期冻存	采样过程、试验台、冻融、短期冻存、长期冻存
接受标准	各稳定性条件下测定值与理论值(名义值)相比较 与理论值相对误差(RE%)在 $\pm 20\%$ 之间 $RE\% = (\text{测得值} - \text{名义值}) / \text{名义值}$	因同时含有内源待测物和加样样品[加样质控(QC)]或者只含内源待测物(内源QC),故理论值未知。受方法变异影响,基线水平(T_0 样本)无法准确测定 未经冻存的 T_0 样本难以获取 接受标准难以界定,后文将详述

分析和解释相关原因^[4]。

3 稳定性验证方法和接受标准的演变与“迭代”

在早期,生物标志物稳定性验证主要参照PK的做法,如2011年的文章^[1],采用个体基质和加样样品进行试验台、短期和长期稳定性评估,采用的接受标准为平均回收率75%~125%,精密度(RSD)≤25%,4/6的基质满足上述标准。在2012年的文章中^[8],用真实基质配制标准品加样QC,使用RE[RE=测得值/(加样理论值+内源值)]评估稳定性。

随着人们逐渐意识到重组蛋白和内源蛋白稳定性的巨大差异^[7],在2016年的文章中指出^[9],如果可以,应使用3个水平的内源性样品进行稳定性验证。在同年发表的1篇白皮书中明确说明^[10],没有必要同时检测加样分析QCs和内源样品QCs的稳定性,只需证明内源样品QCs的稳定性即可;此外,使用真实样本支持生物标志物的稳定性更为准确,真实样本应采用个体样本而非“大量混合”(mega pool)基质。

使用未加样的内源真实样本评估稳定性会带来新的挑战。首先,由于生物标志物分析实验室和临床研究机构常常不在一起,意味着从样本采集到运输最后再到分析实验室检测,样本至少经过1次冻融且经过一定时间的冻存,所以 T_0 的真实样本难以获取。在2017年发布的白皮书中提到^[11],为评估生物标志物的稳定性,理想的情况是通过内部捐献等手段尽可能获取新鲜样本,并在方法开发早期就进行稳定性评估;如果不可行,则尽可能采集最新的样本,如一次冻融的样本作为 T_0 样本。其次,难以筛选内源性真实样本,使得待测物浓度合理的分布于标准曲线的定量范围,即难以获得代表高、中、低浓度的内源QC样本。最后,也是最重要的一点,因为后续的稳定性样品测得值都会与 T_0 的值比较,所以要求 T_0 的值应尽可能测得准确。但是,除非能证明重组的标准品和内源待测物完全一致,否则测得的值只是相对的准确值^[12],并且方法本身的变异导致准确的测定 T_0 更为困难。基于此,在2018年发表的文章中,Allinson^[13]提出了1个尽量消除方法变异影响的接受标准:如果测得的稳定性样品的值偏移 T_0 值超过 $\pm 3 \times \text{RSD}$ (由方法验证时的数据计算而来),则可认为不稳定,笔者认为这个标准可以达到99.7%的置信水平。另外一种得到普遍认同的方法是使用趋势分析^[3,14-15]。

在2019年发表的C-Path专家组共识中提到^[3],通过检测与基线水平样本的偏移来判断样品的稳

定性,同时监测随时间变化的稳定性评估的趋势和运用控制图(control chart)有助于鉴别任何可能与样品稳定性有关的超出界限的表现。在2019年的1个案例中,Kar等^[16]用趋势分析评估血清内源性TGF- $\beta 1$ 的稳定性。具体做法为,汇集20个月的稳定性样本测得值,计算均值(mean)和标准偏差(SD, standard deviation),以月份为横坐标,测得值为纵坐标做趋势分析图。基于临床生化检测所采用的Westgard规则^[17],参照Bruijnsvoort等所用的规则^[18],作者设定的警戒标准为:

- (1) 1个点的值超过 $\text{mean} \pm 3 \text{ SD}$
- (2) 2个点的值超过 $\text{mean} \pm 2 \text{ SD}$
- (3) 4个点的值超过 $\text{mean} \pm 1 \text{ SD}$
- (4) 4个连续的点位于均值的同侧

出现上述任何一种异常结果,则触发稳定性样品调查程序。结果显示,从第11个月开始,连续8个月的稳定性测得值都小于均值,此结果提示从第11个月开始,内源性样品开始变得不稳定。同时应当注意,如果只有1个数值超出警戒限,但随后的值又回到了警戒限内,则很可能是方法变异,跟稳定性无关^[15]。

4 稳定性验证方法的选择和接受标准的界定

首先可以明确的是,应采用内源真实样本进行生物标志物的稳定性验证。如果在方法验证时,真实的病人样本不可获取,则可采用购买或内部捐献的办法获得健康人的内源性样本。但如果健康人和病人体内待测生物标志物表达水平差异较大,如健康人内源水平极低而病人较高,则不得不在稳定性验证时使用加样样品,但后续高浓度的内源样品可用时,应补充加样样品和内源样品稳定性的比较。

目前看来,广为接受且相对较为准确的生物标志物稳定性评估方法是运用控制图的趋势分析。但该方法需要的稳定性时间点较多,适用于持续时间长、样本量大的临床II、III期研究。但如果稳定性时间点较少,则需要尽可能的测得准确的 T_0 值,后续的稳定性样本测得值与 T_0 值进行比较,为减少方法本身变异对测得值的影响,应采用 $T_{0\text{测得值}} \pm 3 \times \text{RSD}$ (方法验证的批内RSD)的接受标准。但如果方法验证时的批内RSD较大(如>20%),则该标准可能不适用。

最后,应采用“符合目的”的原则,选择合适的稳定性验证内容。如仅用于内部决策或探索性目的(数据不提交监管审查)的生物标志物分析方法,

则只需部分验证,稳定性验证可以只进行短期冻存和冻融2项,但如果数据用于监管审查,则方法需要完全验证,稳定性验证需包含所有的验证内容,即采样过程、试验台、冻融、短期冻存、长期冻存等^[19]。

总之,应结合方法验证结果和实际的临床应用综合考量,采用“符合目的”的原则制定稳定性验证方法和接受标准。

5 其他关键考量点

在样本采集、运输和保存过程中,内源性蛋白质生物标志物容易降解和失活^[15]。因此在稳定性评估中应考虑到预期的正式研究中的样品处理和保存条件,包括所有临床研究场点的冰箱温度范围,样品采集、运输和处理流程等^[15]。可以使用严格的样品采集处理手册或统一培训并结合操作视频教学的方法,尽可能的保证正式研究的样本处理过程和验证样品处理过程的一致性以及多个临床研究场点之间的样本采集处理的一致性。新型样品采集管如BD的P800采集管和加入DPP-4抑制剂的K₂EDTA管可以减少样品处理过程带来的稳定性问题。此外,处理完的样品在冻存之前也可以考虑加入样品稳定剂以利于样品的长期冻存稳定性^[15]。对于一些含低丰度蛋白的样品(如脑脊液、唾液、尿液等),在稳定性考察时可以考虑在样品中加入抗吸附剂和使用低吸附管保存样本(如在脑脊液中加入聚山梨酯-20有助于A β 42的稳定)^[11]。在检测唾液中内源性精氨酸酶活性时,加入二硫苏糖醇(DTT)和蛋白酶抑制剂能明显提升精氨酸酶的稳定性^[16]。

尽管在方法开发时加速稳定性(如保存于37℃)可以作为有用的工具来了解待测物的稳定性特点,但是在正式验证时不能用加速稳定性来替代真实冻存的稳定性评估^[20]。对于冻存稳定性,如果已证明在较高温度冻存下的稳定性(如-20℃),那么是否意味着样品在更低温度(如-70℃)的冻存是稳定的?答案是否定的。根据Arrhenius反应动力学定律^[20],从化学上而言分子在更低的保存温度下更稳定,但是蛋白质在更低的温度冻存或许会增加变性和部分丢失其三维结构,进而在配体结合试验中会影响抗体对其识别。

在稳定性评估中需要的样品的量较大,如果个体样本的量无法满足要求,也可以使用混合样本,混合样本会掩盖个体样本之间的稳定性差异^[15],所以应考虑评估混合样本的这个特点对临床研究的影响。如果已知受试药物可以和待测生物标志物

相互作用,则应建立不同的稳定性样本,模拟给药前/安慰剂样品和给药后真实样本,并分别评估其稳定性^[15]。

如果稳定性研究持续时间较长,在整个稳定性研究中采用了不同批次的试剂盒,则在稳定性评估时可以考虑使用计算机工具计算校正因子,对来自不同批次试剂盒的标准曲线进行桥接,这样得出的结果更为公正和准确^[16,21]。但是需注意,使用简单的、非计算机统计计算而来的校正因子进行桥接可能会带来错误的结果,因为这种简单的校正因子是基于两种错误的假定,一是不同分析批之间的方法变异是恒定的,二是标准曲线上不同的浓度带来的变异也是恒定的^[16]。

6 结语

生物标志物稳定性验证受重组蛋白和内源性蛋白稳定性差异、完全新鲜的内源性T₀样本难以获取、方法变异导致T₀值难以准确测定等因素的影响,使得稳定性验证成为生物标志物验证中具有挑战性的验证项。但运用趋势分析和计算机统计等手段,可以相对准确的进行生物标志物稳定性评估。

参考文献

- [1] Valentin M A, Ma S, Zhao A, et al. Validation of immunoassay for protein biomarkers: bioanalytical study plan implementation to support pre-clinical and clinical studies [J]. J Pharm Biomed Anal, 2011, 55(5): 869-877.
- [2] FDA. Bioanalytical Method Validation Guidance [EB/OL]. (2018-05-25) [2019-11-23]. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/bioanalytical-method-validation-guidance-industry>.
- [3] Piccoli S P, Michael Sauer J. Points to consider document: scientific and regulatory considerations for the analytical validation of assays used in the qualification of biomarkers in biological matrices. biomarker assay collaborative evidentiary considerations writing group, critical P [EB/OL]. (2019-06-11) [2019-11-23]. <https://c-path.org/wp-content/uploads/2019/06/evidconsid-whitepaper-analyticalsectionv2019.pdf>.
- [4] Cowan K J, Amaravadi L, Cameron M J, et al. Recommendations for selection and characterization of protein biomarker assay calibrator material [J]. AAPS J, 2017, 19(6): 1550-1563.
- [5] Walsh C T, Garneau-Tsodikova S, Gatto G J. Protein posttranslational modifications: The chemistry of proteome diversifications [J]. Angew Chemie - Int Ed, 2005, 44(45): 7342-7372.

- [6] Prassas I, Brinc D, Farkona S, et al. False biomarker discovery due to reactivity of a commercial ELISA for CUZD1 with cancer antigen CA125 [J]. *Clin Chem*, 2014, 60(2): 381-388.
- [7] Fraser S, Fleener C, Ogborne K, et al. When close is not close enough: a comparison of endogenous and recombinant biomarker stability samples [J]. *Bioanalysis*, 2015, 7(11): 1355-1360.
- [8] Houghton R, Gouty D, Allinson J, et al. Recommendations on biomarker bioanalytical method validation by GCC [J]. *Bioanalysis*, 2012, 4(20): 2439-2446.
- [9] Jani D, Allinson J, Berisha F, et al. Recommendations for use and fit-for-purpose validation of biomarker multiplex ligand binding assays in drug development [J]. *AAPS J*, 2016, 18(1): 1-14.
- [10] Richards S, Amaravadi L, Pillutla R, et al. 2016 White Paper on recent issues in bioanalysis: Focus on biomarker assay validation (BAV): (Part 3 - LBA, biomarkers and immunogenicity) [J]. *Bioanalysis*, 2016, 8(23): 2475-2496.
- [11] Safavi A, Beaver C, Zeng J, et al. 2017 White Paper on recent issues in bioanalysis: a global perspective on immunogenicity guidelines & biomarker assay performance (Part 3 - LBA: immunogenicity, biomarkers and PK assays) [J]. *Bioanalysis*, 2017, 9(24): 1967-1996.
- [12] Cowan K, Gao X, Parab V, et al. Fit-for-purpose biomarker immunoassay qualification and validation: Three case studies [J]. *Bioanalysis*, 2016, 8(22): 2329-2340.
- [13] Allinson J L. Clinical biomarker validation [J]. *Bioanalysis*, 2018, 10(12): 957-968.
- [14] Bruijnsvoort M Van, Meijer J, Den Beld C Van. The application of control charts in regulated bioanalysis for monitoring long-term reproducibility [J]. *Bioanalysis*, 2017, 9(24): 1955-1965.
- [15] Dakappagari N, Zhang H, Stephen L, et al. Recommendations for clinical biomarker specimen preservation and stability assessments [J]. *Bioanalysis*, 2017, 9(8): 643-653.
- [16] Kar S, Islam C. New approaches for biomarker stability determination in regulated bioanalysis: trending, bridging and incurred samples [J]. *Bioanalysis*, 2019, 11: 1837-1844.
- [17] Westgard J O. The need for a system of quality standards for modern quality management [J]. *Scand J Clin Lab Invest*, 1999, 59(7): 483-486.
- [18] Whitcomb B W, Perkins N J, Albert P S, et al. Treatment of batch in the detection, calibration, and quantification of immunoassays in large-scale epidemiologic studies [J]. *Epidemiology*, 2010, 21(Suppl 4): S44-50.
- [19] Islam R, Kar S, Islam C, et al. Adaptation of commercial biomarker kits and proposal for "drug development kits" to support bioanalysis: call for action [J]. *Bioanalysis*, 2018, 10(12): 945-955.
- [20] Van De Merbel N, Savoie N, Yadav M, et al. Stability: recommendation for best practices and harmonization from the global bioanalysis consortium harmonization team [J]. *AAPS J*, 2014, 16(3): 392-399.
- [21] Feng F, Thompson M P, Thomas B E, et al. A computational solution to improve biomarker reproducibility during long-term projects [J]. *PLoS One*, 2019, 14(4): 1-15.