

## 【 一致性评价专栏 】

## 两阶段生物等效性研究的试验设计及关注要点

孙 华, 梁大虎, 谢海棠\*

皖南医学院弋矶山医院临床药学部, 安徽 芜湖 241000

**摘 要:** 两阶段生物等效性研究目前已得到多个国家生物等效性指南的认可,但在实施两阶段生物等效性研究时如何在控制 I 类错误的基础上保证目标把握度是很大的挑战。对目前国内外文献发表的两阶段设计生物等效性研究方法加以综述,详细介绍了文献方法的研究策略、检验水准的校正方法、样本量再估算等,为国内药品申办者在开展两阶段生物等效性研究提供参考。

**关键词:** 生物等效性; 两阶段设计; 适应性设计; 样本含量再评估; 把握度; 检验水准

**中图分类号:** R945      **文献标志码:** A      **文章编号:** 1674-6376 (2017) 05-0593-07

**DOI:** 10.7501/j.issn.1674-6376.2017.05.003

## Two-stage bioequivalence designs and problems need to focus on

SUN Hua, LIANG Da-hu, XIE Hai-tang

Department of Clinical Pharmacy, Yijishan Hospital of Wannan Medical College, Wuhu 241000, China

**Abstract:** Two-stage designs for the assessment of bioequivalence have been recently accepted in various regulatory authorities. However, controlling type I error rates around 5% at targeted power is still a great challenge for applying two-stage method. This paper reviewed the feature of present designs of the two-stage bioequivalence. The decision tree, nominal significance level, and sample size recalculation in previously published methods were also introduced in detail, which would be referential for domestic sponsors in the study of two-stage design bioequivalence.

**Key words:** bioequivalence study; two-stage design; adaptive design; sample size recalculation; power; size of test

在生物等效性研究中,样本量是非常关键的因素,生物等效性试验样本量估算取决于检验水准( $\alpha$ )、检验效能( $1-\beta$ )、关键药动学参数(AUC、 $C_{\max}$ )的个体内变异系数(within-subject coefficient of variation, CV)、试验制剂和参比制剂待评价药动学参数指标的实际差异(T/R)和设定的等效界限<sup>[1-2]</sup>。但由于个体内变异及参比制剂与受试制剂间差异的真值往往在试验前不可知,在实施中,有可能由于对其错误的估计,造成纳入的样本例数过少或过多。样本例数过少,检验效能不足,犯 II 类错误的可能性加大,可能将等效的制剂误判为不等效,增加申办者的风险;而样本例数过多,会使研究成本增加,纳入过多的受试者暴露于试验药物也不符

合伦理原则。如何兼顾统计效率与经济效益的矛盾双方,无论在理论上还是实践中都有重要的意义。

两阶段设计(two-stage design, TSD)生物等效性研究由于其在样本量估算方面的适应性、灵活性,近年来已被各国指南所认可(FDA、EMA、加拿大、澳大利亚、日本、新西兰、WHO等)<sup>[2-8]</sup>。虽然在2015年版《中国药典》9011 药物制剂人体生物利用度和生物等效性试验指导原则中明确指出:“在证明生物等效性时,可以接受两阶段试验方法,最初一组受试者给药并分析数据,如果不能证明生物等效,则可以增加招募一组受试者,在最终分析中合并两组的结 果”<sup>[9]</sup>,但具体如何实施在指南中没有明确规定。本文将综述国内外相关文献,

收稿日期: 2017-03-06

基金项目: 国家自然科学基金面上项目(81173134); 安徽省卫生厅医学科研重点项目(2010A013)

作者简介: 孙 华(1977—),女,副主任药师,主要从事临床药代动力学研究。E-mail: sunhua9406@163.com

\*通信作者: 谢海棠,研究员/教授,硕士生导师。Tel: (0553)5739810 E-mail: xiehaitang@sina.com

对两阶段设计生物等效性研究的设计特点、决策树、中期分析、样本量再估算、检验水准的校正方法、结果判定等做详细介绍,为两阶段生物等效性研究提供可采纳的方法,以促进国内新药临床试验的发展。

### 1 两阶段生物等效性方法的提出

在以药动学参数为终点评价指标的生物等效性试验中,一般采用双周期、双交叉的设计方法,根据选定的药动学参数和预设的接受限,对两者的平均生物等效性做出判定<sup>[10]</sup>。血药浓度-时间曲线下面积(AUC)反映暴露的程度,峰浓度( $C_{\max}$ )和达峰时间( $T_{\max}$ )是受到吸收速度影响的参数。当关键药动学参数(AUC、 $C_{\max}$ )的几何均数比(geometric mean ratio, GMR)的90%可信区间(CI)落在预设的接受范围内时(通常要求为80.00%~125.00%),则被认为生物等效。在进行两周期双交叉平均生物等效性评价时,推荐的统计方法为:对关键药动学参数(AUC、 $C_{\max}$ )进行对数转换后,采用多因素方差分析,考虑可以合理假定对相应变量有影响的方差来源,如序列、序列内受试者、周期和制剂,得到误差的均方(MS),带入以下公式,来计算关键药动学参数(AUC、 $C_{\max}$ )的几何均数比的CI值。

$$100(1-2\alpha)\%CI = \ln^{-1}(\mu_T - \mu_R \pm t_{\alpha, df} \cdot Se) \times 100\%$$

$$Se = s \times \sqrt{2/n}$$

其中 $\mu_T$ 、 $\mu_R$ 分别代表受试制剂与参比制剂关键药动学参数的均数(经对数转换后), $Se$ 为标准误差, $s$ 为MS的平方根,来自方差分析。从公式中可以发现,当 $s$ 较大,说明研究药物个体内变异较大,CV与 $s$ 满足以下关系式: $CV = [\exp(s^2) - 1]^{1/2}$ ,即使GMR接近于1( $\mu_T - \mu_R$ 接近于零),当样本例数( $n$ )不够时,其GMR的90%CI亦有可能超出预设的等效范围,特别对于高变异药物( $CV > 30\%$ )的生物等效性判定时,失败的案例有可能是由于样本例数不够,把握度过低,而将本为等效的两制剂误判为不等效。采用经典的单阶段固定样本量的生物等效性设计往往会出现上述问题,一旦失败,过去只能重新组织更大规模的生物等效性研究,造成人力物力资源的浪费。因此早在上世纪90年代,即有国内外学者提出可以在生物等效性研究中沿用随机对照临床研究(random control trial, RCT)中的适应性设计的思路,在生物等效性研究中提出成组序贯设计(group sequential design)、追加样本设计

(add-on design)等<sup>[11-13]</sup>,其中两阶段研究是成组序贯研究的特例,与以疗效指标为主要评价指标的RCT研究相比,由于以药动学参数为终点评价指标的生物等效性研究的规模较小,两阶段研究比较适用<sup>[4]</sup>。

两阶段设计是将生物等效性研究分成两个阶段,当第一阶段实验完成后进行期中分析(interim analysis),如拒绝零假设则可提前得出有效结论结束试验,否则进入下一阶段试验,重新评估样本量,招募第二阶段的受试者进行研究。一般来说,由实际试验数据估算的样本含量往往要比不确切信息下估算的更加准确,因此,根据期中分析的结果对本含量进行调整,可以保证试验达到预定的把握度。另外对于不能拒绝无效假设的,也可提前结束以避免不必要的浪费,提高试验效率,更加符合伦理学的要求。但由于研究分为两个阶段,有可能使得检验水准增加( $\alpha$  inflation),增加发生I类错误的概率(type I error, TIE)<sup>[14]</sup>,因此,有学者针对在两阶段设计中如何控制TIE进行了专门的研究。

目前两阶段生物等效性研究已在很多国家与地区生物等效性指南中得到认可<sup>[2-8]</sup>。2015版《中国药典》的9011指南也明确指出两阶段设计可以被接受<sup>[9]</sup>。各国指南均要求调整后的显著性水平应在方案中预先规定,指南还对实施过程提出了一些建议,如日本指南中就提出第二阶段招募的受试者例数( $n_2$ )应不少于第一阶段受试者例数的1/2( $n_2 \geq n_1/2$ )<sup>[6]</sup>。欧盟、WHO、中国指南中要求在分析两阶段合并的数据时,在方差分析模型中应包括阶段项<sup>[3,8-9]</sup>。加拿大指南中对两阶段设计中显著性水平调整方法给出建议,推荐采用Pocock  $\alpha$  消耗函数方法,两阶段调整后的名义检验水准 $\alpha'$ (nominal alpha)均采用0.029 4<sup>[4]</sup>。

### 2 两阶段生物等效性研究设计方案

根据假设检验数据统计的原理,重复检验若不对检验水准进行调整,会造成TIE的增大。重复实施 $k$ 次假设检验,每次的检验水准 $\alpha$ 均为0.05,最终TIE可用下述公式估计: $1 - (1 - \alpha)^k$ 。在实施两阶段生物等效性研究时,若不对两阶段 $\alpha$ 进行调整,TIE可能上升为9.75%,如果不对累积TIE进行控制,可能会使不等效的药物获得上市许可,增加患者风险。因此,必须在设计两阶段生物等效性时,对累积TIE进行控制。在生物等效性研究中控制累积TIE有以下6种方法。

### 2.1 Bonferroni 校正法

Bonferroni 校正法其每阶段的名义检验水准调整为： $\alpha/k$ ， $k$  为阶段数，当采用两阶段设计时，第一阶段与第二阶段的  $\alpha$  调整为 0.025，则总的 TIE 基本可被控制在 5% 以下。Bonferroni 校正法在控制 TIE 方面相对保守，每阶段的置信区间调整为 95% ( $CI=1-2\alpha'$ )，纳入的总的样本量一般比固定样本设计的样本量要大。目前认为本法应用于生物等效性研究时由于样本例数要求较大，不符合伦理原则，一般不被推荐。

### 2.2 Pocock 校正法

Pocock 消耗函数的公式为： $\alpha(t)=\alpha\ln[1+(e-1)t]$ ， $t$  为试验时间，采用 Pocock 校正法，各阶段名义检验水准基本相等，当采用两阶段设计时，每阶段的名义检验水准 ( $\alpha'$ ) 调整为 0.0294<sup>[14-15]</sup>，在加拿大的生物等效性指南中推荐采用本校正方法<sup>[4]</sup>，在欧盟的指南中亦参考本方法，认为在第一阶段与第二阶段均采用 94.12% ( $CI=1-2\alpha'$ ) 的置信区间可被接受<sup>[8]</sup>。

但 Pocock 校正法是基于平行组优效性假设检验的序贯设计的，数据一般要求已知变异度并符合正态分布，中期分析的时间点要求是在  $N/2$  ( $N$  为总的样本量) 处。生物等效性一般为交叉设计，其药动力学参数需进行对数转换后才符合正态分布，在试验前个体内的变异度为未知，进行中期分析的时间点并不一定在  $N/2$  时，因此，生物等效性研究不符合上述 Pocock 校正法应用的前提条件，在两阶段样本量不同的情况下，仍采用相同的名义检验水准理论依据不充分，可能会造成 TIE 的增加，需要采用蒙特卡罗拟合的方法对试验结果的 TIE 进行验证，模拟过程可由统计软件 (如 C、R、SAS 等) 实现<sup>[16-18]</sup>。

### 2.3 Potvin 适应性两阶段设计 (方法 B、C、D)

Potvin 曾提出了 4 种 (方法 A、B、C、D) 双交叉生物等效性研究的两阶段适应性设计方案，针对不同 T/R 的估计值与个体内变异的案例进行了模拟。方法 A 由于未对  $\alpha$  进行调整，会造成 TIE 增大，不被推荐。而方法 B、C、D 的研究路径见图 1 和 2，Potvin 团队认为在受试制剂与参比制剂  $T/R=0.95$  时应用方法 B、C，在  $T/R=0.90$  时应用方法 D，均可以在控制 TIE 的前提下保证把握度<sup>[19-20]</sup>。

根据对第一阶段数据分析时  $\alpha$  是否固定，可将适应性两阶段设计主要分成两种类型，类型 I 为无

论中期分析时的结论是中止研究 (不拒绝无效假设)，还是继续进入至第二阶段，在两阶段均采用同一调整后的名义检验水准；类型 II 为根据中期分析的把握度，对第一阶段数据进行统计分析时，先采用  $\alpha=0.05$ ，根据计算得到的置信区间与把握度判断是否等效后再决定是否需要对  $\alpha$  进行调整。这两种类型的设计方法的决策树示意图见图 1、2。

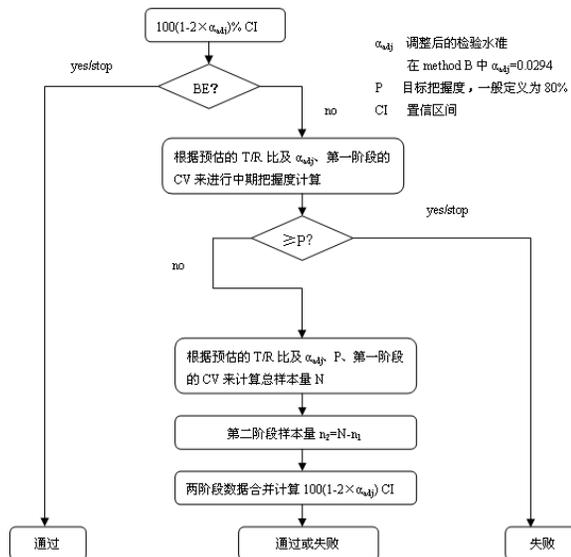


图 1 类型 I 两阶段设计决策树示意图 (方法 B)

Fig.1 Graphical description of type I two-stage bioequivalence design (decision tree of method B)

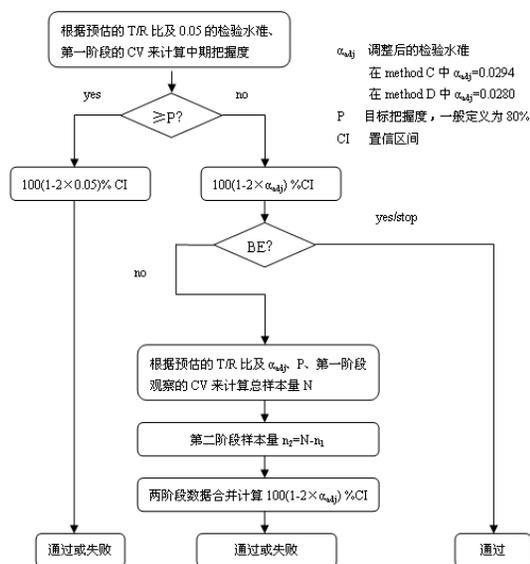


图 2 类型 II 两阶段设计决策树示意图 (方法 C、D)

Fig.2 Graphical description of type II two-stage bioequivalence design (decision tree of method C/D)

类型 I 的设计方案主要是以 Pocock 校正法为基础，Potvin 提出的方法 B 即属于此类，两阶段的名义检验水准均调整为 0.029 4；而类型 II 的设计方案是基于传统的生物等效性理论，在未进入第二阶段前（此时  $\alpha$  并未被消耗，仍可采用  $\alpha=0.05$ ）来进行生物等效性的判定。若决定继续第二阶段的研究，再对两阶段的  $\alpha$  进行调整，Potvin 设计的方法 C 与方法 D 均属此类，方法 C 与方法 D 研究路径一致，只是名义检验水准不同，在方法 C 中  $\alpha_{adj}$  调整为 0.029 4，而方法 D 中  $\alpha_{adj}$  调整为 0.028 0<sup>[19-20]</sup>，适用于不同的 T/R 估计值。

Potvin 提出的两阶段适应性设计是目前两阶段生物等效性研究的基础，后来的学者基本是在他的方法基础上加以改良。

**2.4 Fuglsang 提出的适应性两阶段设计（方法 E）**

对于试验前对个体内变异度估计不足，而根据第一阶段数据得出的实际 CV 较大时，往往在第二阶段纳入的样本量  $n_1 \ll N$ ，这种情况下，需要在第二阶段纳入较多的受试者才能保证把握度，此时又可能造成 TIE 的增加，有必要引入无效规则的概念（futility rule）<sup>[21-22]</sup>。Fuglsang 在 Potvin 方法 C 的基础上提出了方法 E<sup>[26]</sup>，对总样本例数设置了无效规则，定义总样本量的上限，规定  $N_{max}$  为  $2n_1$ 、 $3n_1$ 、 $4n_1$  或直接控制其例数，如规定  $N_{max} \leq 60$  或  $N_{max} \leq 80$  等，当通过中期分析时计算得到的总样本例数超出样本量上限时，试验终止，不再往下进行。此外方法 E 要求无论第一阶段的数据结果，人为强制规定都必须进入到第二阶段，还在方案中对第二阶段纳入的受试者数目最小值的限制规定为  $n_{2min} = n_1/2$ 。决策树示意图见图 3。值得注意的是，对受试者例数上限进行控制可能会减小研究的把握度。

Fuglsang 认为与 Potvin 提出方法 C 相比较，对于高变异低初始样本量的研究来说，方法 E 所纳入的总受试者例数较少，更加符合伦理学的要求。此后，Fuglsang 还把两阶段设计推广至平行设计生物等效性研究<sup>[27-28]</sup>。

**2.5 Karalis 等提出的总体样本量控制的适应性两阶段设计**

Potvin 提出的两阶段设计，在期中分析进行把握度计算，所应用的 T/R 采用的是假设值（一般假设为 0.90 或 0.95），Fuglsang 提出的方法 E，在期中分析时也是以 T/R 为 0.95 的假设进行运算。Potvin 未设计无效规则，而 Fuglsang 提出的方法 E 中的无

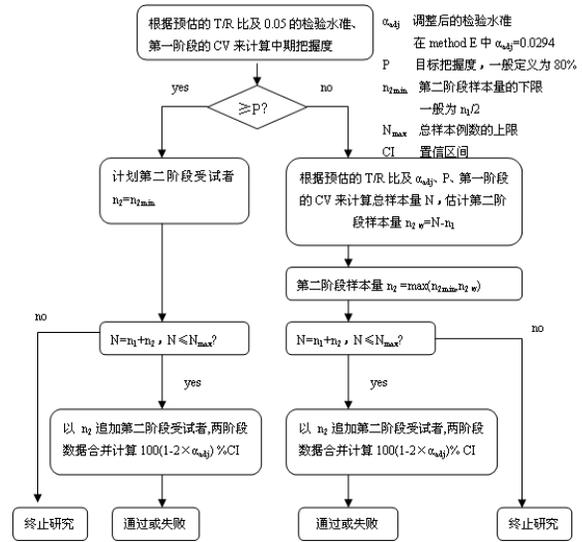


图 3 Fuglsang 提出的适应性两阶段设计决策树示意图（方法 E）

Fig.3 Graphical description of adaptive two-stage bioequivalence design invented by Fuglsang (decision tree of method E)

效规则中针对样本总量，未对第一阶段 GMR 的点估计值（point estimate）提出要求。对此，Karalis 设置了下列无效规则：若第一阶段的 GMR 超出接受范围（0.80~1.25），则终止研究；设置样本量上限（upper sample size limit, UL），若经中期分析， $n_1 + n_2 > UL$ ，试验终止，不再往下进行。Karalis 等在 Potvin 方法的基础上提出两种总样本量控制的适应性两周期设计 STD-1 与 STD-2，与之前不同的是，Karalis 的方法在中期分析做样本量再估计采用的是第一阶段的 GMR（而不是 T/R 的估计值）。Karalis 认为有样本总量的限制，可以控制 TIE 的上升，UL 设置得越低，结果越保守，Karalis 在 STD-1 中名义检验水准设置为 0.028 0，STD-2 中名义检验水准设置为 0.029 4，在其论文中模拟了  $UL=62$  与  $UL=150$  两种情况，没有发现 TIE 的明显上升<sup>[29-30]</sup>。STD-1 与 STD-2 两种方法的决策树示意图见图 4、5。

**2.6 Fuglsang 提出在更高把握度要求下的 TIE 控制方法**

大多文献在数据模拟时，把 80% 作为目标把握度，在 Fuglsang 2013 年的论文中<sup>[31]</sup>，针对有申办者出于商业的考虑，希望将把握度增大至 90% 时，提出了名义检验水准的调整方案，以 90% 为目标把握度，采用 Potvin 方法 B（假设 T/R=0.95）， $\alpha_{adj}$  调整

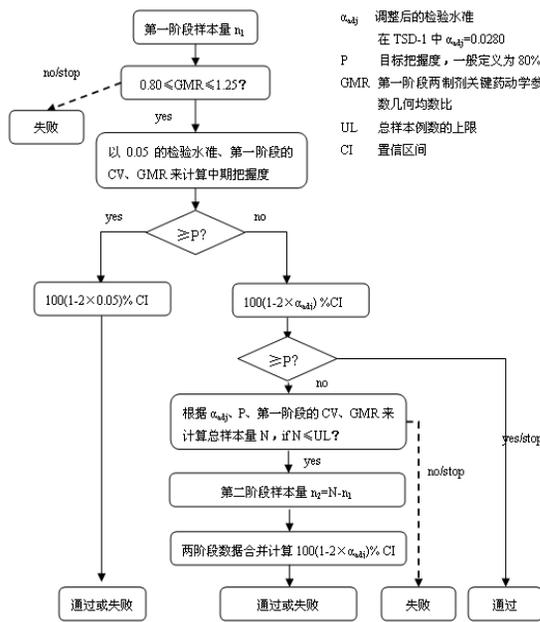


图 4 Karalis 提出的两阶段设计决策树示意图 (方法 TSD-1)

Fig.4 Graphical description of adaptive two-stage bioequivalence design invented by Karalis et al (decision tree of method TSD-1)

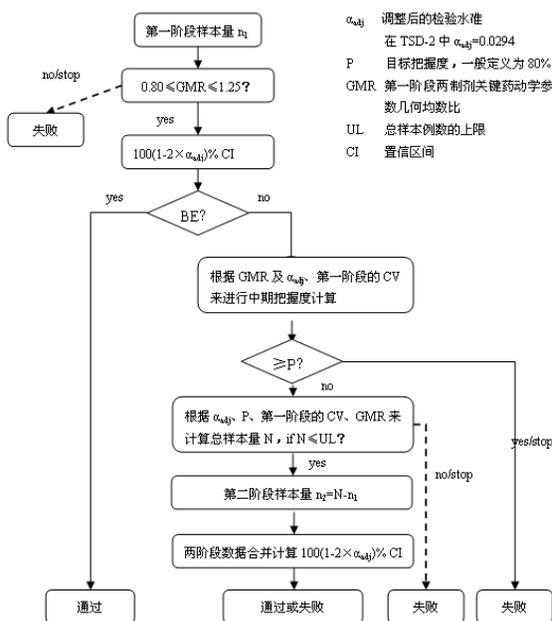


图 5 Karalis 提出的适应性两阶段设计决策树示意图 (方法 TSD-2)

Fig.5 Graphical description of adaptive two-stage bioequivalence design invented by Karalis et al (decision tree of method TSD-2)

至 0.028 4, 采用 Potvin 方法 C (假设 T/R=0.95),  $\alpha_{adj}$  调整至 0.027 4, 用 Potvin 方法 D (假设 T/R=0.90),  $\alpha_{adj}$  调整至 0.026 9, 可将总的 TIE 控制在 5% 以下。

### 3 实施两阶段生物等效性的关注要点

#### 3.1 两阶段生物等效性研究与预试验的差别

在传统的固定样本生物等效性研究正式开展前, 往往会先进行预试验, 预试验的主要目的是考查采血时间点是否合理、了解体内浓度分布范围、评估两周期给药间的清洗时长及分析生物样本分析方法的可行性, 也有申办者通过组织预试验来估计 T/R 差异以及药动学参数个体内变异情况。预试验一般不作假设检验, 其结果不进行等效性的判定, 且预试验的数据不可与正式研究数据合并。预试验与正式研究的临床研究中心、监护项目、入试者入排标准、采血点、清洗期、生物样本分析方法等允许不相同。而两阶段生物等效性中的第一阶段与第二阶段都属于正式研究, 两阶段的试验应按同一方案执行, 在同一临床研究中心实施, 两阶段的药物批号、临床研究条件和生物样本分析方法均应保持一致。

#### 3.2 两阶段生物等效性研究的方案设计要点

要求在开展前对两阶段生物等效性研究前做出详细的研究方案, 明确两阶段生物等效性研究的方法类型、名义检验水准的调整方法、第一阶段的样本例数和中期分析样本再评估的策略。推荐根据试验药物的变异度、预估的 T/R 以及可能的脱落率来设计第一阶段样本量, 双交叉生物等效性研究的样本量可用下面的公式<sup>[1,2]</sup>来估计:

$$n_e \geq 2 \times [t_{(\alpha)} + t_{(\beta)}]^2 \times [CV / (v - \delta)]^2$$

公式中,  $n_e$  代表估计的样本量;  $v$  代表生物等效限 (以对数值表示),  $\delta$  代表试验制剂与参比制剂的差异 (对数转换后), 只有 CV 值在被低估把握度不足才进入到第二阶段。那种认为既然准备进行两阶段生物等效性研究, 而将第一阶段样本例数设置得较低的做法不被鼓励。第一阶段的例数过少, 有可能会造成总的样本量增加。

在实施两阶段生物等效性研究时, 对第一阶段数据的中期分析非常关键<sup>[32]</sup>, 需要统计第一阶段结果的 GMR, 并根据第一阶段方差分析的误差均方 ( $s^2$ ) 来计算出 CV ( $CV = [\exp(s^2) - 1] / 2$ ), 按方案规定的决策树方法对总样本例数进行再评估。根据申办者对把握度的期望值、研究的成本预算、研究药

物的 CV 估计等来综合考虑是否设置 UL 及其数值。

### 3.3 数据合并时应关注的内容

采用单阶段固定样本量两周期双交叉生物等效性设计时,在对关键药动学参数(AUC、 $C_{\max}$ 对数转换后)进行多因素方差分析(ANOVA)的因素项一般包括:序列、嵌套在序列间的个体、周期、制剂,而在进行两阶段生物等效性研究中由于增加了阶段项这一误差来源,在欧盟指南、中国药典 9011 指南中要求在进行多因素方差分析时加入阶段项,在欧盟 PKWP 问答中推荐在 ANOVA 模型中将阶段、序列、序列×阶段、个体(序列×阶段)、周期(阶段)、制剂作为因素来源进行分析,而制剂×阶段的交互作用没有实际意义,不适用于两阶段生物等效性研究。另外也有文献认为虽然欧盟药典专门提出了序列×阶段的因素项,但在实际应用中大都无意义。应注意采取措施保证两阶段研究的受试者人群、入排标准、研究实施方案、生物样本分析方法一致,以减小阶段间差异对结果的影响。由于序列与阶段均属于个体间效应,无论是否纳入方差分析因素,方差分析的残差(residual error)结果不受影响,不会影响最终等效性判定时 CI 的计算。

鼓励申办者在试验完成后对研究的 TIE 及把握度进行验算,前文中介绍的 Potvin、Karalis、Fuglsang 等作者发表的论文方法已采用蒙特卡罗拟合法对试验结果进行了验证,能在控制 TIE 的基础上保证目标把握度,申办者在实施时可直接引用,但需注意引用的前提是变异度和 T/R 估计值不能超过文献验证时预设的范围。

### 3.4 伦理学的考虑

一个良好设计的生物等效性研究应符合伦理与科学两大原则,尽量减少纳入的受试者例数。新西兰的指南中指出<sup>[7]</sup>:在生物等效性研究中纳入不超过 40 名受试者是比较合适的。对于高变异的药物来说,申办者应综合伦理与科学的原则采用合适的研究方法:如选择参比制剂校正的平均生物等效性(reference-scaled average bioequivalence, RSABE)<sup>[23-25]</sup>,或是同位素标记制剂的方法<sup>[33]</sup>,这两种方法所介绍需样本量相对较少,更符合伦理的原则。对于高变异的药物来说,若选择采用两阶段生物等效性,设计中建议尽量引入总量控制。

## 4 结论

本文从统计分析角度综述了国外关于两阶段生物等效性研究的相关指南和文献,并阐述了实施两

阶段生物等效性的注意事项。生物等效性试验是新药研发过程中评价药品质量和桥接安全性、有效性数据的重要手段。生物等效性试验评价中样本量的估计是一个比较复杂的问题,特别对于高变异的药物。两阶段生物等效性研究由于其在样本量估算方面的适应性、灵活性,近年来已被各国指南所认可,但关于具体的两阶段生物等效性方法药监部门没有相应的评价细则,如何在两阶段生物等效性研究中在控制 TIE 的基础上保证目标把握度,需大量试验基础与理论相结合。本文详细介绍了 Potvin、Karalis、Fuglsang 等作者发表的两阶段生物等效性论文的研究策略、检验水准的调整方案、样本量的再估算方法,希望能给国内药品申办者和临床研究单位在实施两阶段生物等效性研究时提供参考。

### 参考文献

- [1] Phani B R B, Someswara R K, Sanketh K C, et al. Sample size estimation for highly variable drugs using reference scaled average bioequivalence criteria [J]. *Int J Rec Sci Res*, 2015, 6(7): 5040-5045.
- [2] Food and Drug Administration, Center for Veterinary Medicine. Supplemental examples for illustrating statistical concepts described in the VICH *in vivo* bioequivalence guidance GL52 [EB/OL]. (2016-12-16) [2017-1-24]. <http://www.fda.gov/downloads/AnimalVeterinary/GuidanceComplianceEnforcement/GuidanceforIndustry/UCM415701.pdf>.
- [3] European Medicines Agency, Committee for medicinal products for human use(CHMP). Guideline on the investigation of bioequivalence [EB/OL]. (2010-1-20) [2017-1-24]. [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Scientific\\_guideline/2010/01/WC500070039.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2010/01/WC500070039.pdf)
- [4] Health Canada. Therapeutic products directorate guidance document: Conduct and analysis of comparative bioavailability studies [EB/OL]. (2012-5-22) [2017-1-24]. [http://www.hc-sc.gc.ca/dhp-mps/prodpharma/applic-demanded/guide-ld/bio/gd\\_standards\\_ld\\_normes-eng.php](http://www.hc-sc.gc.ca/dhp-mps/prodpharma/applic-demanded/guide-ld/bio/gd_standards_ld_normes-eng.php)
- [5] Australian Department of Health, Therapeutic Goods Administration. Guideline on the investigation of bioequivalence [EB/OL]. (2012-5-22) [2017-1-24]. <https://www.tga.gov.au/clinical-efficacy-and-safety-guidelines/clinical>
- [6] Ministry of Health, Labour and Welfare of Japan. Guideline for bioequivalence studies of generic products [EB/OL]. (2012-2-29) [2017-1-24]. <http://www.nihs.go.jp/d>

- rug/be-guide(e)/Generic/GL-E\_120229\_BE.pdf
- [7] New Zealand Medicines and Medical Devices Safety Authority. New Zealand regulatory guidelines for medicines. Section 7. Bioequivalence testing of oral medicines [EB/OL]. (2014-9-?) [2017-1-24]. <http://www.medsafe.govt.nz/regulatory/Guideline/Full%20-%20NZ%20Regulatory%20Guidelines%20for%20Medicines.pdf>
- [8] World Health Organization Expert Committee on Specifications for Pharmaceutical Preparations. Multisource (generic) pharmaceutical products: guidelines on registration requirements to establish interchangeability Annex 7 [EB/OL]. [2017-1-24]. [http://www.who.int/medicines/areas/quality\\_safety/quality\\_assurance/Annex7-TRS992.pdf?ua=1](http://www.who.int/medicines/areas/quality_safety/quality_assurance/Annex7-TRS992.pdf?ua=1)
- [9] 中国药典 [S]. 四部. 2015. 药物制剂人体生物利用度和生物等效性试验指导原则, Ch.P(2015) Vol IV, 2015: 356-362.
- [10] Schuirmann D J. A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability [J]. J Pharmacokinet Biopharm, 1987, 15(6): 657-680.
- [11] Gould A L. Group sequential extensions of a standard bioequivalence testing procedure [J]. J Pharmacokinet Biopharm, 1995, 23(1): 57-86.
- [12] Hauck W W, Preston P E, Bois F Y. A group sequential approach to crossover trials for average bioequivalence [J]. J Biopharm Stat, 1997, 7(1): 87-96.
- [13] 韩可勤. 两阶段抽样方法在生物等效性评价中的应用 [J]. 数理医药学杂志, 1996, 9(2): 143-144.
- [14] Fenta H M. Determination of sample size for two stage sequential designs in bioequivalence studies under  $2 \times 2$  crossover design [J]. Sci J Clin Med, 2014, 3(5): 82-90.
- [15] Pocock S J. Group sequential methods in the design and analysis of clinical trials [J]. Biometrika, 1977, 64(8): 191-199.
- [16] Schütz H. Two-stage designs in bioequivalence trials [J]. Eur J Clin Pharmacol, 2015, 71(3): 271-281.
- [17] Kieser M, Rauch G. Two-stage designs for cross-over bioequivalence trials [J]. Stat Med, 2015, 34(16): 2403-2416.
- [18] Pocock S J. Interim analyses for randomized clinical trials: the group sequential approach [J]. Biometrics, 1982, 38: 153-162.
- [19] Potvin D, DiLiberti C E, Hauck W W, et al. Sequential design approaches for bioequivalence studies with crossover designs [J]. Pharm Stat, 2008, 7(4): 245-262.
- [20] Montague T H, Potvin D, Diliberti C E, et al. Additional results for 'Sequential design approaches for bioequivalence studies with crossover designs' [J]. Pharm Stat, 2012, 11(1): 8-13.
- [21] Fuglsang A. Futility rules in bioequivalence trials with sequential designs [J]. AAPS J, 2014, 16(1): 79-82.
- [22] Karalis V, Macheras P. On the statistical model of the two-stage designs in bioequivalence assessment [J]. J Pharm Pharmacol, 2014, 66(1): 48-52.
- [23] Haidar S H, Davit B, Chen M L, et al. Bioequivalence approaches for highly variable drugs and drug products [J]. Pharm Res, 2008, 25(1): 237-241.
- [24] Wonnemann M, Frömke C, Koch A. Inflation of the Type I Error: Investigations on regulatory recommendations for bioequivalence of highly variable drugs [J]. Pharm Res, 2015, 32(1): 135-143.
- [25] 何春远, 孙 华, 谢海棠. 高变异药物生物等效性试验及量化评价 [J]. 中国临床药理学与治疗学, 2016, 21(7): 721-730.
- [26] Fuglsang A. A sequential bioequivalence design with a potential ethical advantage [J]. AAPS J, 2014, 16(4): 843-846.
- [27] Fuglsang A. Sequential bioequivalence approaches for parallel designs [J]. AAPS J, 2014, 16(3): 373-378.
- [28] Fuglsang A, Schütz H, Labes D. Reference datasets for bioequivalence trials in a two-group parallel design [J]. AAPS J, 2015, 17(2): 400-404.
- [29] Karalis V, Macheras P. An insight into the properties of a two-stage design in bioequivalence studies [J]. Pharmaceut Res, 2013, 30(7): 1824-1835.
- [30] Karalis V. The role of the upper sample size limit in two-stage bioequivalence designs [J]. Int J Pharmaceut, 2013, 456(1): 87-94.
- [31] Fuglsang A. Sequential bioequivalence trial designs with increased power and controlled type I error rates [J]. AAPS J, 2013, 15(3): 659-661.
- [32] Fuglsang A. Controlling type I errors for two-stage bioequivalence study designs [J]. Clin Res Regul Aff, 2011, 28(4): 100-105.
- [33] Parr A, Gupta M, Montague T H, et al. Re-introduction of a novel approach to the use of stable isotopes in pharmacokinetic studies [J]. AAPS J, 2012, 14(3): 639-645.