

确证性临床试验期中分析案例分析与思考

Summary and reflections on the interim analysis of confirmatory clinical trials

衡明莉¹, 钟思睿², 张晨博¹,
孙凤宇¹

(1. 国家药品监督管理局药品审评中心; 2. 国家
药品监督管理局药品审评检查大湾区分中心)

HENG Ming-li¹, ZHONG Si-ru²,
ZHANG Chen-bo¹, SUN Feng-yu¹

(1. Center for Drug Evaluation, National
Medical Products Administration, Beijing
100076, China; 2. Guangdong - Hong
Kong - Macao Greater Bay Area Center
for Drug Evaluation and Inspection,
National Medical Products
Administration, Shenzhen 518000,
Guangdong Province, China)

作者简介: 衡明莉 (1980 -), 女, 副研究员, 主
要从事生物统计学审评工作

通信作者: 衡明莉, 副研究员

Tel: (010) 87965772

E-mail: mingli_heng@163.com

摘要: 随着国内外临床试验适应性设计指导原则的发布, 在临床试验中设计期中分析, 对临床试验的成功与否, 具有很大的指导意义, 它不但有助于提高研发效率, 而且, 对一些临床试验选择了不正确的统计方法, 从而导致 I 类错误膨胀、引入偏倚等严重问题, 可能引起临床试验失败有重要防范意义。本文结合目前临床试验中出现的 3 个案例, 对确证性临床试验期中分析相关统计学问题进行了分析与总结, 促进了对期中分析的深入理解, 为确证性临床试验进行高质量地设计、实施、分析和结果解释提供参考。

关键词: 临床试验; 期中分析; 统计

DOI: 10.13699/j.cnki.1001-6821.2025.19.022

中图分类号: R969 **文献标志码:** C

文章编号: 1001-6821(2025)19-2841-03

Abstract: Objective With the release of guidance for adaptive design in clinical trials both domestically and internationally, designing interim analyses in clinical trials has great guiding significance for the success of clinical trials. It not only helps to improve research and development efficiency, but also has important preventive significance for clinical trials that have chosen incorrect statistical methods, resulting in serious issues such as type I error inflation and introduction of bias, which may lead to the failure of clinical trials. This article combines three cases that have occurred in current clinical trials to analyze and summarize the statistical issues related to confirmatory clinical trial interim analyses, promoting a deeper understanding of interim analyses and providing references for the high-quality design, implementation, analysis, and interpretation of results of confirmatory clinical trials.

Key words: clinical trial; interim analysis; statistics

美国食品药品监督管理局 (Food and Drug Administration, FDA) 于 2018 年 9 月发布了《药物和生物制品临床试验适应性设计》^[1] 的指导原则草案, 随之 2021 年 1 月国家药品监督管理局 (National Medical Products Administration, NMPA) 发布了《药物临床试验适应性设计指导原则 (试行)》^[2], 在确证性临床试验中, 越来越多的申办者或者研究者在设计临床试验方案时, 考虑设计了期中分析, 同时在创新药物研发中也得到广泛应用。但是, 作者发现, 在确证性临床试验中的期中分析实施中, 发现了诸多问题, 本文结合目前案例, 对期中分析存在的相关统计学问题进行剖析, 希望能够起到抛砖引玉的作用, 促申办者或者研究者进行深层次的思考, 为临床试验高质量设计、实施和分析提供参考。

1 案例分析

1.1 案例1

1.1.1 设计及实施

设计:A药拟基于一项III期随机对照双盲临床试验观察其用于儿童及成年人某种皮肤病的局部外用疗效。随机分配方法是采用分层区组随机化,分层因素为病情程度和年龄,方案中没有明确关于样本量重估和疗效的期中分析,基于儿科人群安全性成立数据监查委员会(data monitoring committee,DMC),DMC章程描述如下:“研究设立数据监查委员会,定期审查试验进行中累积数据,特别关注低年龄人群有安全风险,就研究的继续、修改或终止提供建议,若儿童发生SAE则立即召开临时会议”。

临床试验实施:第一次期中分析时,儿科人群占比35.8%。DMC建议的其中一条为:建议增加样本量,增加儿科人群受试者比例。申请人根据DMC建议,增加儿科受试者样本量,儿科人群占比提高至52.2%。

1.1.2 统计学考量

①涉及儿童用药,本试验成立DMC的目的主要是是监查安全性,方案、统计分析计划(statistical analysis plan,SAP)及DMC章程均未涉及样本量重估以及疗效的期中分析。

②第一次期中分析时入组人群各年段构成比例与自然人群就诊患者相似;试验结果显示,成人疗效较儿童差,增加儿童比例更有利于做出阳性结果。

综上,本研究目的是评估儿童及成人患者的有效性和安全性,设置DMC的目的是在临床试验过程中监测入组儿童患者的安全性,不涉及样本量重估和疗效的期中分析,本研究中DMC职责超出了方案,在试验过程中,DMC进行了关于疗效的期中分析,期中分析结果显示,儿童疗效较成人好,DMC根据期中分析结果提出增大样本量,增加儿科受试者比例的建议,申请人未关注DMC职责与方案不一致,接受了DMC建议,在试验过程中通过增加儿童患者例数增加儿童患者比例,成年人比例减少,导致总人群疗效因为儿童比例增加而引入偏倚,本试验无法估计总人群疗效,没有反应出真正临床试验的疗效。

1.2 案例2

1.2.1 试验设计及结果

B药拟通过一项随机、双盲、安慰剂平行对照、多中心III期临床试验证明其在某风湿疾病中的有效性和安全性。主要终点指标为第24周时根据美国风湿病学会类风湿关节炎疾病活动评估核心标准达到

20%改善(American college of rheumatology 20,ACR20)的受试者比例。申请人设置了DMC,进行安全性审核。对于主要终点,结果显示,试验组和安慰剂组应答率分别为46.3%和33.9%, $P<0.05$ 。

1.2.2 统计学考量

①期中分析前后结果不能自洽:将DMC会议前后入组的受试者分别进行分析,结果显示:DMC会议前入组的受试者,试验组和对照组应答率分别为43.81%和40.63%。DMC会议后入组的受试者,试验组和对照组应答率分别为49.37%和25.93%。第二阶段对照组应答率比第一阶段低15%,两阶段安慰剂组应答率存在较大差异。

②各分项结果不能自洽:ACR20的各分项指标,仅患者评价的两项指标,试验组有优于对照组的趋势,其余指标均未显示试验组有优于对照组的趋势。其中对于疾病活动度的评价,虽然患者评价显示试验组有好与对照组的趋势,但医生评价的疾病活动度未显示该趋势。

本研究临床试验方案没有设计关于疗效的期中分析,仅仅对药物的安全性设计了期中分析,但是在此期中分析前后的分析结果表明,期中分析前组间疗效近似,期中分析后组间差异变大,两阶段组间疗效差异的不一致主要与两阶段安慰剂组应答率存在较大差异有关;另外,ACR20各分项指标间不能自洽,综上,本研究内部一致性存疑,可能是期中分析操作不当引入了偏倚。

1.3 案例3

1.3.1 设计及实施

C药拟通过一项随机对照试验支持其新增适应症,患者按1:1随机分配至试验组和对照组,对照组为常规治疗,试验组为试验药物加常规治疗。在方案V1.0版本未预设期中分析,方案V1.3计划A时点进行安全性期中分析,方案V1.4修改期中分析时间节点,计划B时点(B时点晚于A时点)进行安全性期中分析。

1.3.2 统计学考量

①方案V1.4版本修订了期中分析时间节点,存在先进行期中分析后进行方案修订行为。

②申请人方案中预设1次期中分析,但实际进行了两次期中分析。

③申请人方案预设期中分析仅涉及安全性,但实际也进行了疗效的期中分析。

④两次期中分析均未成立DMC,均由申请人自己进行分析。

综上,临床试验在实施过程涉及破盲,可能为安全性和有效性评价带来偏倚。

2 总结与思考

药物临床试验中常见期中分析目的为样本量重估和有效性分析,其次为以无效性期中分析和 II/III 期无缝设计等为目的的期中分析,无论是哪种期中分析,均需考虑设置期中分析的适用性、合理性及完整性等统计学关键问题^[3]。

对于样本量重估,需要注意问题包括但不限于:

①如果期中分析增大了样本量,在统计学上,有较成熟的 I 类错误控制方法和计算组间比较 P 值的方法^[4,5];但两阶段数据进行统计分析,各组疗效点估计和区间估计以及组间差异点估计和区间估计存在困难^[6]。②若临床试验预设确有证性亚组,对确证性亚组进行以样本量重估为主要目的的期中分析,根据亚组在全人群中比例放大全人群样本量,此时会增大全人群 I 类错误。③非劣效设计的临床试验中考虑样本量重估,可能增大 I 类错误^[9]。

在有效性期中分析中,需要注意问题包括但不限于:①期中分析时间节点,为保持期中分析时结果的稳健性,建议选择较保守的时间节点^[7](例如 2/3 以上信息量)。②在以事件为主要终点的试验中事件发生速度,如果事件发生速度较快,可能期中分析还未结束或刚完成期中分析不久,预设的主要分析所有事件数就能得到收集,此种情形下,考虑期中分析并不合理。③期中分析达到阳性,是否继续试验直至收集到方案预设主要终点事件数,以验证结果稳健性。④疗效比较的期中分析,是否设置 DMC,DMC 设置是否符合《药物临床试验数据监查委员会指导原则》^[10],DMC 实际操作中是否存在不规范问题(例如 DMC 成员作为研究者参与临床试验、DMC 成员参与盲态审核会、DMC 建议超出预设等)。⑤多区域临床试验(multi-regional clinical trials, MRCT)在进行期中分析时,中国样本量、事件数是否足够进行一致性评价^[11]。

在 II/III 期操作无缝设计中,可能会面临两阶段数据差异较大,尤其是期中分析后,试验组疗效与期中分析前相似,但安慰剂组疗效与期中分析前相比明显下降,导致整体安慰剂疗效降低的情况,在这种情形下,结果的可解释性可能受到影响,进而疗效估计

可能受到质疑。另外,在一个临床试验中,先进行无效期中分析,再进行有效性期中分析,逻辑存在矛盾,无效期中分析基于对药物疗效没有信心,而有效性期中分析基于对药物疗效信心较大,虽然能够控制 I 类错误,但存在逻辑上的矛盾。

总之,期中分析应用得当,能够加速药物研发;应用不当,可能引入偏倚^[8]、统计分析面临挑战、结果难以解释等问题,试验设计时应充分考虑并予以重视。

参考文献:

- [1] FOOD AND DRUG ADMINISTRATION (FDA). Adaptive designs for clinical trials of drugs and biologics guidance for industry [EB/OL]. Washington (America): FDA, 2018-09 [2025-07-30]. <https://www.fda.gov/ucm/groups/fdagov-public/fdagov-drugs-gen/documents/document/ucm201790.pdf>.
- [2] 国家药品监督管理局药品审评中心. 药物临床试验适应性设计指导原则(试行) [EB/OL]. 2021-01-29 [2025-07-30]. <https://www.cde.org.cn/zdlyz/domesticinfopage?zdyzIdCODE=4409e51a403a911757af6caf3ecef129>.
- [3] ICH. E9: Statistical Principles for Clinical Trials [EB/OL]. 1998-02-05 [2025-07-30]. <https://www.cde.org.cn/ichWeb/guideIch/toGuideIch/3/0>.
- [4] WANG P, CHOW S C. Sample size re-estimation in clinical trials [J]. *Stat Med*, 2021, 40(27): 6133-6149.
- [5] CHANG M. Adaptive design method based on sum of p-values [J]. *Stat Med*, 2007, 26(14): 2772-2784.
- [6] ROBERTSON D S, CHOUDARI - OSKOOEI B, DIMAIRO M, et al. Point estimation for adaptive trial designs I: A methodological review [J]. *Stat Med*, 2023, 42(2): 122-145.
- [7] BRIEL M, BASSLER D, WANG A T, et al. The dangers of stopping a trial too early [J]. *J Bone Joint Surg Am*, 2012, 94 (Suppl 1): 56-60.
- [8] SCHOU I M, MARSCHNER I C. Meta-analysis of clinical trials with early stopping: an investigation of potential bias [J]. *Stat Med*, 2013, 32(28): 4859-4874.
- [9] FRIEDE T, KIESER M. Blinded sample size re-estimation in superiority and noninferiority trials: Bias versus variance in variance estimation [J]. *Pharm Stat*, 2013, 12(3): 141-146.
- [10] 国家药品监督管理局药品审评中心. 药物临床试验数据监查委员会指导原则(试行) [EB/OL]. 2020-09-23 [2025-07-30]. <https://www.cde.org.cn/zdlyz/domesticinfopage?zdyzIdCODE=7b29e6094564d9d962dc4a3af13862de>.
- [11] ICH. E17: General principle on planning and designing multi-regional clinical trials [EB/OL]. 2019-11-12 [2025-07-30]. <https://www.cde.org.cn/ichWeb/guideIch/toGuideIch/3/0>.

(收稿日期 2025-08-23)