

白蛋白临床试验设计中的统计学考量

孔媛媛¹, 姚晨², 贾继东³

1 首都医科大学附属北京友谊医院临床流行病学与循证医学研究室, 国家消化系统疾病临床医学研究中心, 北京 100050

2 北京大学第一医院医学统计室, 北京 100034

3 首都医科大学附属北京友谊医院肝病中心, 北京 100050

通信作者: 孔媛媛, kongyy@ccmu.edu.cn (ORCID: 0000-0002-2586-1443)

摘要: 白蛋白在临床中的应用广泛, 其临床试验设计的合理性直接影响研究结果的可靠性和临床应用价值。本文综述了白蛋白临床试验设计中的关键统计学考量, 包括对主要终点的选择、统计学假设及非劣效界值的设定、腹水改善的临床评价标准、样本量的估计、中期分析等问题进行探讨, 旨在为临床研究者提供方法学参考, 以优化临床试验设计、提高其科学性和可行性。

关键词: 白蛋白类; 临床试验; 统计学(主题)

基金项目: 国家重点研发计划(2023YFC2306902, 2023YFC2306900); 北京市卫生健康委高层次公共卫生技术人才建设项目(学科骨干-010-018)

Statistical considerations in the design of albumin clinical trials

KONG Yuanyuan¹, YAO Chen², JIA Jidong³

1. Department of Clinical Epidemiology and Evidence-Based Medicine, Beijing Friendship Hospital, Capital Medical University, National Clinical Research Center for Digestive Diseases, Beijing 100050, China; 2. Department of Medical Statistics, Peking University First Hospital, Beijing 100034, China; 3. Liver Research Center, Beijing Friendship Hospital, Capital Medical University, Beijing 100050, China

Corresponding author: KONG Yuanyuan, kongyy@ccmu.edu.cn (ORCID: 0000-0002-2586-1443)

Abstract: Albumin is widely used in clinical practice, and the rationality of trial design directly affects the reliability of research findings and clinical application value. This article reviews the key statistical considerations in the design of albumin clinical trials, including the selection of primary endpoints, the establishment of statistical hypotheses and non-inferiority margins, clinical evaluation criteria for ascites improvement, sample size, and interim analyses, in order to provide methodological guidance for clinical researchers to optimize clinical trial design and enhance its scientific rigor and feasibility.

Key words: Albumins; Clinical Trials; Statistics as Topic

Research funding: National Key Research and Development Program (2023YFC2306902, 2023YFC2306900); Beijing Municipal Health Commission High level Public Health Technical Talent Construction Project (Subject-backbones-010-018)

白蛋白作为一种重要的血浆蛋白, 在多种临床病症的治疗中发挥着重要作用^[1-3]。但人血白蛋白逐渐进入临床应用是国内外长期临床实践的结果, 大多并未经过

设计严谨的关键临床试验。此外, 因血液来源不足, 故需要大量进口, 这也是一项潜在的安全性问题。目前, 采用基因工程技术制备的人血白蛋白已经进入临床试

验阶段,为了验证其疗效和安全性,设计合理的临床试验至关重要。但目前评估白蛋白治疗的注册临床试验的终点选择存在较大异质性。此外,已发表的在失代偿期肝硬化患者中评估白蛋白治疗效果的临床研究中,终点选择也不尽相同,如 ATTIRE 研究采用治疗失败率(包括感染、肝性脑病、肾衰竭、多器官衰竭或死亡)^[4], ANSWER 研究采用 18 个月生存率^[5], MATCH 研究采用并发症发生率(如感染、肝肾综合征等), MACHT 研究采用住院时间^[6-7]。

试验设计其他方面的差异也值得注意,包括入组标准、给药方案、随访时间、统计假设、非劣效界值等。以上因素影响了不同临床试验结果的可比性和临床应用的普适性。本文将主要从统计学角度探讨白蛋白注册临床试验设计中的若干关键问题,包括主要终点的选择、统计学假设、非劣效界值、样本量计算以及腹水改善标准的确定等,以期对白蛋白临床试验设计提供参考。

1 主要终点的选择

研究终点是临床试验设计的核心要素之一,不仅能决定试验的成败,还直接影响研究样本量的确定。通常,终点的选择应基于临床重要性进行排序,同时考虑终点事件的预期频率和药物作用的预期效果。

对于研究终点排序的考量,研究者需要权衡临床重要性与统计学效力之间的关系。当临床重要性高的终点发生频率较低时,可能导致研究效力不足。在这种情况下,可以考虑选择临床重要性相对次要、但统计效力更高的终点作为主要终点,以提高研究的统计学效力,增加研究的可行性。

在临床试验中,主要终点的数量可以是单一指标或多指标。多终点设计包括复合终点、共同终点和多终点。对于旨在证明肝硬化腹水患者临床获益的白蛋白试验,可考虑双主要终点设计,即将血浆白蛋白水平改善和腹水改善作为双主要终点,或将白蛋白水平改善作为主要终点,腹水改善作为关键次要终点,以全面评估药物的疗效。

2 统计学假设及非劣效界值的设定

在白蛋白治疗肝硬化腹水合并低白蛋白血症患者的临床试验中,对于白蛋白水平改善可以考虑等效性假设,即对试验药物与阳性对照药物进行生物等效性统计学检验;对于临床获益如腹水改善,则可以考虑非劣效假设,即对试验药物的疗效终点不劣于阳性对照药物进

行统计学检验^[8]。对于非劣效检验,要合理设定非劣效界值,通常需要综合权衡临床意义和统计学效力^[9]。在白蛋白临床试验中,可考虑有效率的非劣效界值设定为 10%~15%。具体设定时,需要结合阳性对照药物的腹水改善标准的定义及有效率、临床专家意见及前期临床试验数据综合考虑。

3 腹水改善标准对有效率及样本量的影响

腹水改善标准的设定直接影响样本量的估算。在白蛋白临床试验中,腹水改善的评价指标并不统一,例如采用体质量、腹围或腹水深度的变化等。此外,有关具有临床意义的腹水改善的定义及有效率的具体标准亦存在争议。

以腹水量相对基线的变化作为腹水改善的评价指标为例,腹水改善的标准为腹水程度较基线时下降一级(基线腹水 1 级者腹水消失),或腹水深度测量值较基线下降达到一定标准,一般被认为有临床意义。腹水改善标准的不同定义会影响试验药物与对照药物间有效率差值的方向和样本量的大小。

假设将腹水改善标准分别定义为腹水深度较基线下降 $\leq 20\%$ 、 25% 、 30% 、 40% 或 50% ,阳性对照药有效率的估计值为 $45\% \sim 70\%$,有效率的差值(试验组-对照组)设为 0 或 -5% ,非劣效界值分别采用 -10% 、 -15% ,把握度(power)= 0.8 ,单侧 $\alpha=0.025$ 时,对试验所需的样本量估算模拟结果详见表 1、2。由此可见,若率差= 0 ,非劣效界值= -10% , $n=734 \sim 874$;非劣效界值= -15% , $n=328 \sim 390$ 。若率差= -5% ,非劣效界值= -10% , $n=3\ 054 \sim 3\ 472$;非劣效界值= -15% , $n=766 \sim 870$ 。

腹水改善标准直接影响疗效的高低及率差方向。模拟结果显示,当腹水下降标准定义为 $<30\%$ 时,组间疗效相近,率差= 0 概率显著增加;当腹水下降标准定义为 $\geq 30\%$ 时,腹水改善率降低,试验组疗效可能低于对照组,率差 <0 概率显著增加(表 3)。

4 中期分析样本量重估的考量

在临床试验中,可基于中期分析的效应量或变异性估计值对本量进行重新估计^[10-12]。但应注意,由于非劣效设计的目标是证明试验组的疗效不显著低于对照组(效应量 $\geq -\Delta$),因此样本量重估更容易导致非劣效设计的 I 类错误膨胀^[11,13-15],主要是由于以下原因:(1)非劣效设计的假设检验边界(Δ)是人为设定的,样本量调整可能改变试验对 Δ 的敏感性;(2)增加样本量会导致置

表1 腹水有效率、率差及非劣效界值与样本量的关系(率差=0)

Table 1 Impact of ascites treatment efficacy, rate difference and non-inferiority margin on sample size (rate difference=0)

对照组有效率估计值	有效率差值	非劣效界值	样本量	扩充后样本量(设脱落率10%)
45%	0	-10%	778(389/组)	866(433/组)
50%	0	-10%	786(393/组)	874(437/组)
55%	0	-10%	778(389/组)	866(433/组)
60%	0	-10%	754(377/组)	838(419/组)
65%	0	-10%	716(358/组)	796(398/组)
70%	0	-10%	660(330/组)	734(367/组)
45%	0	-15%	346(173/组)	386(193/组)
50%	0	-15%	350(175/组)	390(195/组)
55%	0	-15%	346(173/组)	386(193/组)
60%	0	-15%	336(168/组)	374(187/组)
65%	0	-15%	318(159/组)	354(177/组)
70%	0	-15%	294(147/组)	328(164/组)

注: power=0.8, 单侧 $\alpha=0.025$ 。

表2 腹水有效率、率差及非劣效界值与样本量的关系(率差=-5%)

Table 2 Impact of ascites treatment efficacy, rate difference and non-inferiority margin on sample size (rate difference=-5%)

对照组有效率估计值	有效率差值	非劣效界值	样本量	扩充后样本量(设脱落率10%)
45%	-5%	-10%	3 062(1 531/组)	3 404(1 702/组)
50%	-5%	-10%	3 124(1 562/组)	3 472(1 736/组)
55%	-5%	-10%	3 124(1 562/组)	3 472(1 736/组)
60%	-5%	-10%	3 062(1 531/组)	3 404(1 702/组)
65%	-5%	-10%	2 936(1 468/组)	3 264(1 632/组)
70%	-5%	-10%	2 748(1 374/组)	3 054(1 527/组)
45%	-5%	-15%	766(383/组)	852(426/组)
50%	-5%	-15%	782(391/组)	870(435/组)
55%	-5%	-15%	782(391/组)	870(435/组)
60%	-5%	-15%	766(383/组)	852(426/组)
65%	-5%	-15%	734(367/组)	816(408/组)
70%	-5%	-15%	688(344/组)	766(383/组)

注: power=0.8, 单侧 $\alpha=0.025$ 。

表3 腹水不同改善标准与有效率差值关系的模拟结果

Table 3 Simulation analysis of the relationship between different improvement criteria for ascites and efficacy rate difference

腹水改善标准 (腹水程度下降一级 或深度较基线下降率 达到以下标准)	平均有效率 估计值	有效率差值 ≥ 0 的 概率(试验组-对 照组)
$\leq 20\%$	54.5% ~ 69.7%	84.0%
25%	54.5% ~ 65.4%	71.1%
30%	48.0% ~ 54.5%	18.8%
40%	43.8% ~ 49.9%	21.2%
50%	35.0% ~ 45.3%	14.7%

注: 非真实数据, 参考某临床试验 II 期数据模拟, 重复抽样 1 000 次。

信区间变窄, 更容易满足非劣效标准(即置信区间的上限小于 Δ), 从而增加错误得出非劣效结论的可能性; (3) 如果中期分析的效应量估计值存在偏差, 样本量调整可能进一步放大这种偏差, 导致 I 类错误膨胀。

但对于优效假设来说, 在满足相关条件时, 采用希望区域法对样本量重估能够有效控制 I 类错误^[11,16]。主要是由于以下原因: 本方法的核心思想是基于中期分析结果, 将试验分为“希望区域”和其他区域, 并在希望区域内调整样本量以提高试验成功的可能性^[12,17-18]。希望区域法通过中期分析将试验分为以下几种情况: (1) 不可能区域: 中期数据表明试验成功的可能性极低, 此时通常停止试验。(2) 希望区域: 中期数据表明试验有一定成功的可能性, 但统计效能不足, 此时可以通过增

加样本量来提高效能。(3)成功区域:中期数据表明试验已经非常接近成功,此时无须调整样本量。样本量的调整仅在希望区域内进行,而在其他区域进行。在优越假设下,假设检验的边界(如效应量为零)是固定的,不依赖于人为设定的参数(如非劣效界值 Δ),样本量调整不会改变假设检验的边界,因此不会引入额外的偏倚。

鉴于白蛋白临床试验的主要终点为等效设计,关键次要终点为非劣效设计,针对非劣效假设的样本量重估即使采用希望区域法也可能会导致 I 类错误膨胀,故在预先决定是否进行期中分析对样本量进行重估时,应仔细权衡上述影响因素。

5 小结

白蛋白临床试验设计中仍存在诸多挑战,在坚持科学性和严谨性的同时,也应充分考虑临床试验的具体情况和可实施性。目前亟需建立标准化的疗效评价终点,特别是临床改善的客观评价指标,并合理设定样本量估计参数。可通过增加上市后的长期随访,以进一步确证血浆白蛋白水平提升与远期临床获益的关联。未来的研究应继续探索这些因素的优化,通过方法学创新和跨学科合作,以推动白蛋白临床试验的健康发展,从而为广大患者提供来源充足、质量可靠、疗效稳定的白蛋白制品。

利益冲突声明: 本文不存在任何利益冲突。

作者贡献声明: 孔媛媛负责数据分析、文章撰写;姚晨负责统计内容指导、论文审阅;贾继东负责临床内容指导、论文审阅。

参考文献:

- [1] Chinese Society of Hepatology, Chinese Medical Association. Guidelines on the management of ascites and complications in cirrhosis[J]. *J Clin Hepatol*, 2017, 33(10): 1847-1863. DOI: 10.3969/j.issn.1001-5256.2017.10.003.
中华医学会肝病学分会. 肝硬化腹水及相关并发症的诊疗指南[J]. *临床肝胆病杂志*, 2017, 33(10): 1847-1863. DOI: 10.3969/j.issn.1001-5256.2017.10.003.
- [2] European Association for the Study of the Liver. EASL Clinical Practice Guidelines for the management of patients with decompensated cirrhosis[J]. *J Hepatol*, 2018, 69(2): 406-460. DOI: 10.1016/j.jhep.2018.03.024.

- [3] RUNYON BA, AASLD PRACTICE GUIDELINES COMMITTEE. Management of adult patients with ascites due to cirrhosis: An update[J]. *Hepatology*, 2009, 49(6): 2087-2107. DOI: 10.1002/hep.22853.
- [4] CHINA L, FREEMANTLE N, FORREST E, et al. A randomized trial of albumin infusions in hospitalized patients with cirrhosis[J]. *N Engl J Med*, 2021, 384(9): 808-817. DOI: 10.1056/nejmoa2022166.
- [5] CARACENI P, RIGGIO O, ANGELI P, et al. Long-term albumin administration in decompensated cirrhosis (ANSWER): An open-label randomised trial[J]. *Lancet*, 2018, 391(10138): 2417-2429. DOI: 10.1016/S0140-6736(18)30840-7.
- [6] FERNÁNDEZ J, CLÀRIA J, AMORÓS A, et al. Effects of albumin treatment on systemic and portal hemodynamics and systemic inflammation in patients with decompensated cirrhosis[J]. *Gastroenterology*, 2019, 157(1): 149-162. DOI: 10.1053/j.gastro.2019.03.021.
- [7] THÉVENOT T, BUREAU C, OBERTI F, et al. Effect of albumin in cirrhotic patients with infection other than spontaneous bacterial peritonitis. A randomized trial[J]. *J Hepatol*, 2015, 62(4): 822-830. DOI: 10.1016/j.jhep.2014.11.017.
- [8] ICH Harmonised Tripartite Guideline. Statistical principles for clinical trials. International Conference on Harmonisation E9 Expert Working Group[J]. *Stat Med*, 1999, 18(15): 1905-1942.
- [9] Committee for Medicinal Products for Human Use (CHMP). Guideline on the choice of the non-inferiority margin[M]. London: European Medicines Agency, 2005.
- [10] GOULD AL. Sample size re-estimation: Recent developments and practical considerations[J]. *Stat Med*, 2001, 20(17-18): 2625-2643. DOI: 10.1002/sim.733.
- [11] PROSCHAN MA, HUNSBERGER SA. Designed extension of studies based on conditional power[J]. *Biometrics*, 1995, 51(4): 1315-1324.
- [12] MEHTA CR, POCOCK SJ. Adaptive increase in sample size when interim results are promising: A practical guide with examples[J]. *Stat Med*, 2011, 30(28): 3267-3284. DOI: 10.1002/sim.4102.
- [13] FRIEDE T, KIESER M. Sample size recalculation in internal pilot study designs: A review[J]. *Biom J*, 2006, 48(4): 537-555. DOI: 10.1002/bimj.200510238.
- [14] SCHUMI J, WITTES JT. Through the looking glass: understanding non-inferiority[J]. *Trials*, 2011, 12: 106. DOI: 10.1186/1745-6215-12-106.
- [15] SNAPINN SM. Noninferiority trials[J]. *Curr Control Trials Cardiovasc Med*, 2000, 1(1): 19-21. DOI: 10.1186/cvrm-1-1-019.
- [16] DENNE JS. Sample size recalculation using conditional power[J]. *Stat Med*, 2001, 20(17-18): 2645-2660. DOI: 10.1002/sim.734.
- [17] JOSHUA CHEN YH, DEMETS DL, GORDON LAN KK. Increasing the sample size when the unblinded interim result is promising[J]. *Stat Med*, 2004, 23(7): 1023-1038. DOI: 10.1002/sim.1688.
- [18] CUI L, HUNG HM, WANG SJ. Modification of sample size in group sequential clinical trials[J]. *Biometrics*, 1999, 55(3): 853-857. DOI: 10.1111/j.0006-341x.1999.00853.x.

收稿日期: 2024-12-23; 录用日期: 2024-12-30

本文编辑: 王莹

引证本文: KONG YY, YAO C, JIA JD. Statistical considerations in the design of albumin clinical trials[J]. *J Clin Hepatol*, 2025, 41(3): 420-423.

孔媛媛, 姚晨, 贾继东. 白蛋白临床试验设计中的统计学考量[J]. *临床肝胆病杂志*, 2025, 41(3): 420-423.