

# 药物临床试验协变量校正指导原则

2020年12月

# 目 录

一、 概述 .....	1
二、 试验设计中有关协变量的考虑.....	2
(一) 常见的重要协变量 .....	3
(二) 随机化的分层因素 .....	4
(三) 对协变量数量的控制 .....	5
三、 校正协变量的统计分析方法 .....	5
四、 结果的报告和解读 .....	6
(一) 基线变量的特征分析 .....	7
(二) 分析方法对结果解读的影响.....	7
(三) 校正与未校正协变量的分析.....	7
(四) 协变量与处理因素交互作用的探查 .....	8
五、 参考文献.....	9
附录 1：词汇表 .....	10
附录 2：中英文对照.....	11

# 药物临床试验协变量校正指导原则

## 一、概述

在随机对照临床试验中，除处理因素以外还存在其他协变量，如果在试验设计时不进行有效控制，或在统计分析时不进行合理的校正，则可能使检验效能降低，或使疗效估计产生偏倚。因此，在随机对照临床试验中对于协变量的处理应予以慎重考虑。

本指导原则中，协变量是指在干预之前（通常是在随机化之前）观测到的，并且预期与主要研究结果有关联的变量。校正协变量的意义是使得对于任意一个受试者，随机分组到试验组或对照组的预期疗效差异与协变量的观测值无关。由于随机分组的原因，随机对照试验中的各个协变量的取值在试验组与在对照组的概率分布是相同的，而任何观测到的分布不均衡都应归结于随机抽样误差。因此，随机对照试验中协变量校正的主要目的是减少终点变量中与处理因素无关的冗余变异从而使疗效估计更加精确。协变量可以是连续型

的、有序分类的或无序分类的。人口统计学指标（如年龄或体重）、疾病特征（如病程或严重程度）、预后因素、病理学结果、生理学因素、遗传因素、社会学因素（如经济状况、职业、教育水平），以及研究中心或研究者等都可能是协变量。同时，主要疗效指标的基线值也可能是非常重要的协变量。

在临床试验中，为了保证入组受试者对于目标人群的代表性，试验受试者的协变量通常对应一定的取值范围。当存在对终点变量影响较大的协变量时，终点变量的变异度会增加，导致疗效估计误差增大、相关假设检验的效能降低。因此，如何识别并控制潜在的协变量，更科学合理地分析处理因素与终点变量间的效应关系是临床试验中的关键问题。

本指导原则旨在阐明确证性随机对照临床试验中协变量的处理原则，并为试验设计、统计分析、临床试验报告中如何处理和解读重要的协变量提供建议。

## **二、 试验设计中有关协变量的考虑**

在临床试验中，有关协变量控制和校正的考虑起始于试

验设计阶段，并需要在研究方案中事先确定。实际临床试验中可能有很多协变量与主要研究结果有关，因此在试验设计时需要识别重要的、具有生物学意义和临床意义的协变量，并在随机分组时加以控制，在统计分析时加以校正。

### （一）常见的重要协变量

#### 1. 与终点指标关联性较强的协变量

如果协变量与主要终点指标有较强的关联性，协变量的变异以及抽样误差更有可能影响终点变量，造成疗效估计的误差增大以及相应的统计学检验效能降低。因此，通常需要将该协变量引入疗效分析的统计学模型中以提高疗效估计的精度。例如，某病情评估指标属于反映受试者的病情严重程度的连续型变量，并且在基线和干预后均有观测。无论疗效的评估是基于该指标在治疗终点时的实际取值，还是治疗终点时较基线的取值变化，评估结果均与基线取值有较强的关联性。此时，该疾病评估指标的基线取值应纳入统计分析模型中，以在疗效估计时进行相应的协变量校正。

当协变量的变异度较大时，可能导致疗效估计的精度下

降，也可能使疗效估计产生偏倚。因此，可以事先考虑把预期变异较大的协变量引入疗效估计的统计分析模型中，以校正协变量对疗效估计带来的潜在影响。

## 2. 中心因素

在多中心随机对照临床试验中，各研究中心在临床实践、试验条件、受试者基线特征等方面可能存在不同程度的差异，而这些因素可能与终点指标相关，故而在多中心临床试验中通常会选择中心因素作为需要校正的协变量。特别是在国际多区域临床试验中，不同区域的受试者可能存在种族、文化、饮食习惯、临床实践等方面的差异。区域因素通常综合性地包含这些特征和信息，可以考虑以国家或区域分类作为中心因素进行校正。当试验中心数量较多时，单个中心预期入组病人数量可能非常有限，此时以中心为协变量进行校正通常会带来模型估计和结果解读方面的挑战。此时，可以考虑不对中心因素进行协变量校正，或按预先定义的方式合并中心（或国家/地区）后进行校正。

### （二）随机化的分层因素

在随机对照临床试验中，针对与终点变量有较强关联性的协变量，采用分层随机的方法将受试者分配到不同治疗组中，以进一步降低组间协变量的不均衡和控制偏倚。分层因素建议不宜过多，并且通常需要在统计分析模型中加以校正。

### (三) 对协变量数量的控制

如果在统计分析模型中纳入过多的协变量，特别是与终点变量关联性不大或者相互之间相关性很强的协变量，可能导致协变量某些取值组合情况下的样本量很少。这种情况下，经过协变量校正的疗效估计可能产生偏倚，检验效能出现下降，甚至可能导致模型的过度拟合、模型信息矩阵奇异等问题，给统计分析结果的科学性、可靠性和可解释性提出挑战。因此应在试验设计阶段尽可能地选取具有临床意义、与试验终点变量相关性强的关键协变量，以控制纳入统计分析模型的协变量个数。事实上，在随机对照临床试验中，除了分层因素等常规需要校正的协变量外，纳入统计分析模型的协变量数量建议尽可能少。

## 三、校正协变量的统计分析方法

在随机对照临床试验中，通常基于终点变量的类型选择

不同的校正协变量的统计分析方法。例如，对于连续型疗效终点变量，协变量校正可以采用线性模型；对于时间-事件（time to event）型终点变量，协变量的校正可以采用 Cox 比例风险模型；对于二分类疗效终点变量（如有效/无效），每个组别的汇总统计量可以为率（如有效率），评估两组差异的统计量（以下简称为评估统计量）可以是处理组之间的率差、率比或率的优势比（Odds Ratio, OR），不同类型的评估统计量需用不同的协变量校正的统计学模型，例如 logistic 回归模型可以用于评估统计量为 OR 的协变量调整。

校正协变量的统计模型通常基于一系列的假设，因此需要关注模型的适用性要求，还需对模型的假设是否成立进行预先判断。例如，协方差分析模型需要进行残差分析和方差齐性评估，而 Cox 模型需要考虑比例风险模型的假定是否满足等。如果所选分析模型的假设不成立，可能导致对治疗效果的错误估计。

#### 四、结果的报告和解读

除了在研究设计时采用分层随机控制协变量，在分析时用合适的方法对协变量进行校正以外，在研究总结报告中还



要注意正确解释协变量对主要分析结果的影响，评价主要结论的稳健性，并且在试验报告中进行讨论。

### (一) 基线变量的特征分析

在随机对照试验中，一般需要报告各处理组的基线变量特征。由于是随机分组，基线变量的组间差异来自于随机误差。因此基线数据的分析和报告通常基于各处理组的描述性统计，而不进行假设检验或统计推断。

如果出现非预期的基线变量在各处理组间明显不均衡的情况，这可能影响疗效的估计。此时，可以考虑在统计分析计划内容以外增加补充分析对该协变量进行校正，进一步评估主要分析结果的稳健性。

### (二) 分析方法对结果解读的影响

协变量校正会基于特定的统计模型，因此分析结果的解读应结合模型假设的合理性。若发现对模型假定有较大的偏离，需在研究总结报告中予以描述，同时采用其它模型进行补充分析以支持主要分析结果的稳健性。

### (三) 校正与未校正协变量的分析

在随机对照临床试验中，通常根据研究目的和协变量特征确定将校正协变量的分析方法预设为主要分析方法，未采用校正协变量的方法作为敏感性分析。当校正和不校正的结论不一致时，则需要进一步进行深入探讨。

#### (四) 协变量与处理因素交互作用的探查

一般情况下，确证性临床试验的主要目的是衡量处理在目标人群中的整体效应。主要分析中通常不会纳入协变量与处理因素的交互作用项，而在敏感性分析中可以考虑对协变量与处理因素的交互作用进行统计分析。

事实上，除非在设计时有针对性的考虑，临床试验往往都不具有足够的检验效能对协变量和处理因素的交互作用进行检验。因此对于交互作用的检验即使没有统计学意义也不足以完全证明分层亚组间疗效的一致性。如果协变量和处理因素存在具有统计学和临床意义的交互作用，这说明疗效可能在分层人群中有所不同，这种情况下，需从临床角度探讨交互作用的潜在来源、对主要分析结果的影响，并对基于主要分析的结论予以谨慎解释。

## 五、参考文献

- [1] Altman D, Dore C. Randomization and baseline comparisons in clinical trials. *The Lancet*, 1990, 335(8682):149-53.
- [2] Beach M L, Meier P. Choosing covariates in the analysis of clinical trials. *Controlled Clinical Trials*, 1989, 10(4):161-175.
- [3] Committee for Proprietary Medicinal Products (CPMP), Points to consider on adjustment for baseline covariates. *Statistics in Medicine*, 2004, 23:701-709.
- [4] D. Tu, K. Shalay, and J. Pater. Adjustment of treatment effect for covariates in clinical trials: statistical and regulatory issues. *Drug Information Journal*, 2000, 34:511-523.
- [5] EMA. Guideline on adjustment for baseline covariates in clinical trials. 2015
- [6] FDA. Adjusting for Covariates in Randomized Clinical Trials for Drugs and Biologics with Continuous Outcomes. Guidance for Industry (DRAFT GUIDANCE). 2019
- [7] G. Raab, S. Day, and J. Sales, How to select covariates to include in the analysis of a clinical trial. *Controlled Clinical Trials*, 2000, 21:330-342.
- [8] S. Assmann, S. Pocock, L. Enos, and L. Kasten, Subgroup analysis and other (mis)uses of baseline data in clinical trials.

*The Lancet*, 2000, 255:1064-1069.

[9] Senn S. Covariate imbalance and random allocation in clinical trials. *Statistics in Medicine*, 1989, 8(4):67–75.

[10] Tukey J W. Use of Many Covariates in Clinical Trials. *International Statistical Review*, 1991, 59(2):123-137.

[11] 赵耐青, 陈峰. 基线与协变量. 临床试验统计学 (第十三章). 人民卫生出版社. 2018. 202-210.

## 附录 1：词汇表

**协变量 (Covariate):** 在干预之前 (通常是在随机化之前) 观测到的, 并且预期与主要研究结果有关联的变量。

**多区域临床试验 (Multi-regional clinical trial, MRCT):** 一项按照单个方案在多个地区实施的临床试验。

**分层随机 (Stratified randomization):** 依据关键因素 (如年龄、性别、种族、疾病状态等) 对研究对象进行分组 (层), 然后在每层内分别进行的随机化。分层随机可以有效地提高在关键因素或者特别关注的研究对象亚组中分布的平衡性。用于定义分层的因素称为分层因素。

**过度拟合 (Over-fitting):** 在数据分析 (如建模) 中过于精确的契合或匹配某一数据集, 导致分析结果与额外的观测数据不匹配或者无法可靠地预测未来的观测结果。

**交互作用 (Interaction):** 当某一个因素 (协变量) 对于结局变量的影响随另一因素变化而变化时, 则称这两个因素之间存在交互作用。

## 附录 2: 中英文对照

### 中英文词汇对照

中文	英文
过度拟合	Over-fitting
交互作用	Interaction
分层随机	Stratified randomization
协变量	Covariate
抽样误差	Sampling error
偏倚	Bias
敏感性分析	Sensitivity analysis