

# 药物临床试验中心化监查统计指导原则 (试行)

2021年12月

# 目 录

一、 引言 .....	1
二、 临床试验风险管理体系.....	2
(一) 关键数据和关键流程.....	3
(二) 风险评估 .....	3
(三) 风险控制 .....	5
(四) 风险沟通与报告.....	6
三、 中心化监查中的统计应用.....	7
(一) 中心化监查的常用统计指标.....	8
(二) 中心化监查中常用的统计学方法.....	9
1. 试验层面的统计学方法.....	9
2. 研究中心层面的统计学方法.....	10
3. 试验数据的中心化监查.....	12
四、 中心化监查计划和报告.....	13
五、 其他考虑 .....	14
六、 参考文献 .....	16
附录: 中英文对照表.....	18

# 药物临床试验中心化监查统计指导原则（试行）

## 一、引言

为了保证临床试验受试者的权益与安全，临床试验数据的质量与完整，申办者需要对临床试验各个阶段进行监查。ICH E8（R1）指出应在研究计划期间确定研究的关键质量因素，并在研究实施过程中对这些因素进行风险管理。临床试验的数据量和复杂性的增加给临床试验监查带来了更大的挑战。ICH E6（R2）和《药物临床试验质量管理规范》均明确提出申办者应采用系统的、基于风险的方式对临床试验进行监查，不同情况下可考虑采取不同的监查策略，并强调了中心化监查的优势。电子数据采集系统的普遍使用以及统计学评估方法的引入，为中心化监查的实施提供了条件，其与传统现场监查相结合可提高临床试验的质量和效率。

现场监查是指在开展临床试验的过程中，由申办者或其代表在研究中心对试验质量进行实地评估。现场监查能够通过比对原始记录和病例报告表来发现数据录入错误以及缺失数据，保证研究文件的真实完整性；评估研究中心的人员对研究方案和相关流程的熟悉程度；评估对方案的依从性及完成研究药品清点等。上述监查活动，部分或者全部也可以通过信息化的方式在临床试验机构以外的地点完成，其本质也属于现场监查。

中心化监查是指由申办者或其代表使用累积的数据及时地对试验质量进行远程评估。中心化监查作为现场监查的补充，还能帮助调整现场监查的频率和协助识别潜在问题数据，从而提示现场监查的重点。

目前制药、生物技术公司临床试验监查的主要方法仍旧是定期、频繁地访问临床试验现场。与此同时，国内外临床试验关于中心化监查的一系列的实践证实中心化监查方法可以加快质量审核流程，使临床试验质量核查更加高效。通过中心化监查核查数据，申办者可以更早地监测到临床运营质量问题，并进行针对性的快速整改，尤其是涉及样本量较大的多中心临床试验。

目前国内对中心化监查的认知、技术研究与应用还处在早期阶段。因此，明确基于风险的监查策略，并制订相应的实施准则十分必要。本指导原则主要针对中心化监查中的统计学问题，重点关注中心化监查的适用范围和使用时需考虑的因素，以及监查中风险管理措施可能采用的统计学方法，为申办者实践与应用中心化监查在方法的选择和实施方面提供技术指导。本指导原则仅代表当前的观点和认识，随着研究和认识的深入将予以修订完善。

## **二、临床试验风险管理体系**

申办者应在临床试验的整个运营过程中建立风险管理体系，保证临床试验数据的可靠性，提升整个临床试验的

质量管理，从而保护受试者权益与安全。临床试验质量管理体系的建立主要围绕以下几个环节展开：

### （一）关键数据和关键流程

ICH E8（R1）指出申办者应前瞻性地确定临床研究的关键质量因素（Critical to Quality Factors），其中关键数据和关键流程是关键质量因素的重要组成部分。临床试验中数据与流程的重要性存在差异，非关键数据和流程的偶发错误一般不会对药物的安全性和有效性结论有太大影响。而关键数据与流程的错误将损害到受试者的权益或研究结果的可靠性和完整性。

关键数据与关键流程通常是指（包括但不限于）：

- 知情同意是否恰当获得；
- 方案入排标准在招募时的执行情况，尤其是保证受试者权益的标准；
- 研究药物记录和管理流程体系；
- 与临床试验有效性终点相关或方案特定要求的安全性终点相关（例如，严重不良事件、死亡、脱落等）的评估流程体系；
- 与临床试验的可靠性、完整性相关的流程体系（例如，方案违背管理、盲态保持管理等）。

### （二）风险评估

风险评估通常包括识别风险、分析风险，从而为实施风险控制提供依据。早期识别关键数据和关键流程，其核心是为了落实临床试验中的风险评估及管理。

在确定关键数据和关键流程后，申办者应进行风险评估，以确定可能影响关键数据收集或关键流程实际实施的风险性质、来源和潜在原因，从而形成监查中的风险指标。风险评估首先应识别重要的风险，确立风险的优先次序，再对风险进行分析，其中包括对风险的定量估计和风险范围的定性描述。在风险评估期间，着重考虑防止或减少关键数据及关键流程在临床实施、数据收集和最后报告方面可能的重要错误。以监查为目的的风险识别通常应考虑要收集的数据类型、收集这些数据所需的特定手段，以及临床监查中固有的保护受试者权益相关的问题。

风险评估应当考虑以下三个方面：风险发生的可能性；能检测到风险的限度；风险对于受试者权益保护的影响和对试验结果可靠性的影响。申办者应依据风险评估结果制定监查计划（例如，决定哪些风险可以通过监查得到解决），确定最适合应对这些风险的监查活动的类型和强度。

申办者还可以确定哪些风险可以通过监查之外的手段更好地管理（例如，修改方案以消除风险来源）。申办者需要定期评估新发生的风险，并决定监查方式是否需要调整以更有效地进行风险管理。

### （三）风险控制

申办者应决定哪些风险需要降低和哪些风险可以接受。用于将风险降低到可接受水平的方法应与风险的重要性相符。减少风险的措施可纳入方案设计、实施规划、监查计划中；各协作方之间应确定角色和责任、确保遵守标准操作程序，保障对相应措施以及流程的培训。风险控制中应考虑到变量的医学和统计特征以及试验的统计学设计，预先设定质量风险的容忍度（Quality Tolerance Limit, QTL），以识别可能影响临床试验受试者安全性或试验结果可靠性的系统性问题。QTL 反映的是试验层面风险指标可接受的执行偏差的变异程度，可以通过统计学方法来定义。当监测到风险指标数值超出预先设定的 QTL 时应触发风险评估，以确定是否需要采取风险控制措施。QTL 应在试验开始前尽早设定，一是有助于及时纠正或改进流程，以保障临床试验的实施；二是有助于指导与研究目标至关重要参数的管控监查，并有助于设计更多基于风险的监查策略。

设定变异程度范围或 QTL 的风险指标可以考虑以下几点（包括但不限于）：

#### 1. 试验研究数据

基于关键数据的识别和相应的风险评估，注意力应集中在那些重要的风险指标数值超过 QTL 的情况。现在越来越多的数据直接来源于电子源数据，在预设 QTL 范围后，实

现数据的及时测量、跟踪、报告及必要时采取相应措施会更加容易。

## 2. 试验方案流程

应建立有效的机制，及时发现违背方案或《药物临床试验质量管理规范》的情况，并评估它们对研究目标和受试者权益的影响。设置风险指标及 QTL，以及时监测到关键问题，并触发必要的监查升级（例如，额外的现场访问、额外的方案培训等）。

## 3. 试验管理流程

临床试验管理过程中应定义可以进行中心化监查的风险指标，以便有针对性地设定监查活动，设置触发机制。可能触发监查升级的情况包括在 eCRF 系统录入数据出现过度延迟，或延迟汇报严重不良事件等。数据缺乏变异性也会引发进一步的监查，例如降压药试验中对血压测量的某特定一位数的偏好。

可利用信息化技术整合各种来源的试验数据，开发可视化的中心化监查系统，以监查试验进程和数据的质量，并定期汇总监查指标报告以记录和展示正在进行的监查活动是否按预定的监查策略和程序执行，以提升临床运营质量。

### （四）风险沟通与报告

各职能部门应相互协作定期对收集的信息进行分析总



结，包括试验数据本身及其相关时间采集点的变异性评估、对超过 QTL 或方案偏离的评估、缺失数据的评估等。还可以通过对单个或多个指标进行中心内部和各中心间的深入统计分析以获取更多的信息。任何趋势性分析应与整体的科学价值和数据可用性，以及相应的优先级和风险等级相关，同时应结合来自于现场监查报告和数据管理报告中的信息。整个临床试验中各方需及时沟通，保证信息的公开性和透明性，所有重要的质量管理措施都应记录归档，从而更好地支持风险评估与控制。

在临床试验结束后，应定性和定量汇总试验在 QTL 范围内的实施情况，此类信息可汇总在临床研究报告（数据质量保障章节）中。所有跟 QTL 有关的风险指标都应呈现在报告中，无论指标数值是否超出 QTL。如果风险指标数值超出设定的 QTL，临床研究报告还应报告其对受试者安全和试验数据可靠性的潜在影响，并进行原因分析及描述所采取的措施。

### **三、中心化监查中的统计应用**

基于风险的质量管理是一个结合了现场和中心化监查的动态临床试验管理过程，可以持续提高试验实施的质量。在整个基于风险的监查过程中，合理实施中心化监查，尤其是合理地运用统计学方法，能够进一步提高现场监查的效率。中心化监查中可以使用统计学方法和模型来检查及

管理试验流程和采集的临床数据，用来识别非典型的数据模式或者异常趋势，以实现以下目的：

- 在试验层面监测试验整体质量，发现可能影响受试者安全性或试验结果可靠性的系统性问题；
- 在研究中心层面，进行中心之间的比较，发现潜在高危中心，为现场监查确定目标和程度；
- 在试验数据层面评估数据质量，检查试验数据一致性和完整性，确定数据收集过程中是否存在系统性问题、明显错误或者数据真实性问题。

#### （一）中心化监查的常用统计指标

ICH E6（R2）建议从试验层面设定风险指标及其 QTL。试验层面的风险指标及相应的 QTL 可以针对某一试验进行专门设置，也有一些风险指标可能适用于申办者所有的临床试验，并在申办者的质量体系中定义。试验层面风险指标的数量不宜过多，应选取与受试者权益和试验数据完整性密切相关的指标。常用的指标包括入组的受试者不符合入排标准的比例、试验药物相关的严重方案违背发生率、无法观测到主要或关键次要终点的受试者比例等。

类似地，对于关键数据和关键流程，也可以以研究中心为单位设定中心层面的风险指标，使用从单个中心层面计算的统计量来判断一个中心的表现。试验层面的风险指标一般也可以用于中心层面。常见的中心层面的风险指标

包括入组速率、筛选失败率、病例报告表完成时间、数据质疑率、质疑解决时间和活跃的质疑数量、（严重）不良事件数量、缺失或者延迟的随访数量、方案违背数量等。中心层面的风险指标也需要匹配定义相应的阈值。单个中心的指标数值超出阈值时触发相应的措施（例如，加强数据审查、增加现场监查等）。阈值的选择会受到很多因素的影响，包括试验目的、试验设计、指标的类型、不同地区/中心的差异等。因此，统计师需要和其他团队一起根据试验具体情况确定阈值的定义方式。

基于试验数据质量和真实性的监查，可以针对上述设定的试验层面或者中心层面的风险指标进行开展，也可以使用非关键数据（例如，基线特征、伴随治疗等）作为定期中心化监查的内容，判断是否存在异常的数据模式。

## （二）中心化监查中常用的统计学方法

统计学方法的选择应服务于具体的监查目的。以下介绍的是中心化监查中可参考的统计学方法，申办者可以根据试验设计和数据的特点选择合适的方法。

### 1. 试验层面的统计学方法

以监测试验整体质量、发现可能影响受试者安全性或者试验结果可靠性的系统问题为目的时，可使用选定的风险指标在试验层面的汇总值，与试验开始前预先设定的QTL进行比较。QTL的设定可参照历史数据，并考虑指标

的医学和统计特点，以及试验的统计学设计。QTL 可以设定下限或上限，或两者兼有，此外还可以设置更窄的预警阈值，预警阈值的上限或下限应在 QTL 范围内，起到预警和及早干预的作用。

风险指标的数值与阈值（QTL 或预警阈值）的比较可以是简单的是否大于或小于上、下限的判断，也可以使用预设的统计推断的方法，比如假设检验、区间估计、贝叶斯方法等。

## 2. 研究中心层面的统计学方法

以发现高风险研究中心，为现场监查确定重点和程度为目的时，前述试验层面比较风险指标的数值与阈值的方法可以运用于中心层面，也可以使用离群值检测的统计方法进行中心间比较，找出潜在高风险研究中心。这类方法假设参加同一研究的中心遵从同一研究方案进行受试者筛选、治疗、评估，因此具有很大程度上的相似性，表现为中心层面的风险指标应服从同一分布。当某个中心与其他中心的表现明显不一致，其风险指标的数值成为离群值时，则预示该中心可能有潜在的质量风险。阈值的设置与比较可以是一个中心里受试者不良反应的发生率和所有中心相比的差异是否在某预设的百分比范围外；或者在一定统计模型下，一个中心的指标数值标准化后在所有中心的分布中是否在某预设的分位数之外；以及检验本中心指标与其

余中心指标差异的  $P$  值是否小于某预设的检验水准等。

使用中心层面的风险指标汇总值进行中心间比较时，需要考虑因中心入组患者数或随访时间不一样而带来的不同程度的不确定性。即使在同质化风险的假设下，患者数量较少的中心和较多的中心相比，风险估计的变异程度会更大。如果不考虑样本量，直接将单中心和整体进行比较，不论是使用差异的绝对值还是相对比例，关注点可能均会集中在样本量较小的中心。当风险指标与随访时间相关时，例如不良事件发生数量与患者暴露人年相关，随访时间也会带来风险指标变异的不同。常用处理方法包括漏斗图，以标准差的倍数作为阈值对各中心进行评估；也可以通过假设检验，对某一中心与其他中心进行比较，较小的  $P$  值标识潜在离群中心；还可以考虑贝叶斯方法，使用贝叶斯收缩估计量处理中心大小的不同。另外，在使用多个风险指标进行监查时，应关注多重性问题，选择合适的方法控制错误发现率。这些方法都可以在中心层面识别出各个风险指标潜在的离群值，作为现场监查的指导，提示监查员应更加关注相应的异常指标。

有时也需要能够结合多个风险指标，得到反映中心质量风险的综合评分。这时可以基于预先设定的权重计算所有指标的加权平均值，或者根据指标之间的相关性矩阵决定权重，还可以使用其他多变量统计方法（例如，马氏距

离、在主成分分析后采用基于部分主成分的欧氏距离等)。这个综合评分(距离)衡量每个中心与研究整体数据分布的差异度。评分越高(距离越远)则提示更大程度的离群。因此,这样的评分可以作为中心质量风险的整体衡量,用于确定现场监查的频率和紧急性。

### 3. 试验数据的中心化监查

为了发现数据的系统性问题或者错误,包括系统的测量误差(例如,使用了未校正的仪器),或是数据填写错误、缺失,或者真实性问题等,除了可以针对关键数据和关键流程进行统计分析,也可以针对一些非关键数据开展分析,发现其异常模式。统计方法的选择和前述发现高风险研究中心的方法类似。此外,对于单变量数据,可以核查单个中心变量的均值、方差、峰度和偏度等描述性统计指标,通过比较其是否处在其余各中心相对应指标的标准差倍数以外来发现离群点;也可配合可视化图形(例如,直方图、茎叶图、箱线图等)发现异常值。对于多个连续型变量,可以检验这些变量在中心层面的相关系数矩阵和试验层面的相关系数矩阵的一致性来判定是否存在异常数据。另外,可以通过分析结果中端点数字的分布或检验其随机性(例如,应用火山图或者 Benford 定律等方法)来识别是否存在数据真实性的问题。

对于基于累加数据的动态连续监测,可以考虑过程控

制的统计工具（例如，控制图），通过对变量随着累加数据的变化趋势来区分随机波动和异常波动，以判断系统的稳定性并及时采取措施消除系统异常。

#### **四、中心化监查计划和报告**

临床试验中基于风险的质量管理通常围绕以下环节展开：通过收集信息，确定关键数据与流程进行风险评估；实施风险控制，包括设定试验层面的风险指标及 QTL、中心层面风险指标及阈值等；结合前面步骤所收集的信息、试验期间出现的新信息，及风险管理工具，对发现与风险相关的结果和数据及时沟通，确定所需采取的措施。

所有临床试验都可能存在数据完整性、患者安全性和方案依从性及合规性风险。依据质量源于设计的理念，临床运营管理中应系统性地按照基于风险的方法进行监查，包括现场监查和中心化监查。中心化监查计划是进行中心化监查的基础，为定期审阅数据提供系统性计划。

中心化监查计划需由跨部门构成的中心化监查小组共同撰写与执行。统计方面应侧重考虑但不限于监查计划的制定，关键数据与流程的确定；中心化监查的方法选择；根据试验特点决定的监查时间、类型、频次和程度；触发监查活动调整的标准等。监查计划可以针对实际情况进行调整。

中心化监查计划应首先依据方案明确风险指标，风险

指标的确定通常与风险管理计划同步，都应在首例受试者入组前完成。风险指标确定的同时也需设定相应的阈值。同时要注意从什么时间节点开始进行风险评估，因为试验早期可能因为入组人群较少而没有充分的代表性。此外，也要事先规定好数据审阅的频率。

在设定好风险指标后，可依据这些信息制定相对应的图表模板，并依据预设的时间频率产生相应的报表以便中心化监查团队定期审阅，从而实现对临床试验运营质量的及时有效监控。审阅结果需记录在中心化监查报告中并归档。试验层面风险指标的审阅尤为重要，如果数值超过QTL的阈值，则需要详细记录，同时与质量稽查部门开始合作调查，必要时落实纠正与预防措施。

## **五、其他考虑**

基于临床数据整体状况，由跨部门协作进行的中心化监查能够及时地发现系统性问题，提高试验结果的可信度。虽然使用的统计学方法可能会比较复杂，但由于其并不特别依赖于主观指标的选择或者阈值的设定，因此，具备一定的普适性。在相似的结构化数据下，统计程序或系统设定可以在不同试验中重复使用，类似研究的监查结果也可以进行比较。需要明确的是，中心化监查中数据审阅不同于期中分析，中心化监查计划中图表的设计服务于监测临床试验的总体风险信号，而不是对安全性和/或有效性数据



的统计分析，尤其要避免破盲的发生。当监测到特定风险信号时，有可能需要对数据进行深入分析，以确定临床运营中是否存在系统性风险，甚至数据真实性问题。必要时，建议与监管部门进行沟通。

## 六、参考文献

- [1] ICH. ICH E6(R2): Good Clinical Practice: Integrated Addendum to ICH E6(R1) (2017).
- [2] ICH. ICH E8(R1): General Considerations for Clinical Studies (2021).
- [3] ICH. ICH Q9: Quality Risk Management (2006).
- [4] FDA. Oversight of Clinical Investigations-A Risk-Based Approach to Monitoring (2013).
- [5] FDA. A Risk-Based Approach to Monitoring of Clinical Investigations Questions and Answers Guidance for Industry (2019).
- [6] EMA. Reflection paper risk based quality management clinical trials (2013).
- [7] 国家药监局《药物临床试验质量管理规范》(2020).
- [8] ICH. E9(R1): Addendum on Estimands and Sensitivity Analysis in Clinical Trials to the Guideline on Statistical Principles for Clinical Trials (2019) .
- [9] Bhagat R, Bojarski L, Chevalier S, et al. Quality Tolerance Limits: Framework for Successful Implementation in Clinical Development. *Therapeutic Innovation & Regulatory Science*, 2021, 55(2): 251–261.
- [10] Zink RC, Dmitrienko A, Dmitrienko A. Rethinking the Clinically Based Thresholds of TransCelerate BioPharma for Risk-Based Monitoring. *Therapeutic Innovation & Regulatory Science*, 2018, 52(5): 560–571.
- [11] Trotta L, Kabeya Y, Buyse M. Detection of Atypical Data in Multicenter Clinical Trials Using Unsupervised Statistical Monitoring. *Clinical Trials*, 2019, 16(5): 512–522.
- [12] Bottle A, Aylin P. *Statistical Methods for Healthcare Performance Monitoring*. CRC Press, 2016.
- [13] 胡劲, 徐炎, 周高超, 等. 通过基于风险评估统计模型的中心化监查优化抗肿瘤药物临床试验质量管理. *中国食品药品监管*, 2021.
- [14] Shwartz M, Ren J. Estimating a composite measure of hospital quality from the Hospital Compare database: differences when using a Bayesian hierarchical latent variable model versus denominator-based weights. *Medical care*, 2008, 46(8): 778–785.

[15] Spiegelhalter D, Sherlaw-Johnson C, Bardsley M. Statistical Methods for Healthcare Regulation: Rating, Screening and Surveillance: Statistical Methods for Healthcare Regulation. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 2012, 175(1): 1–47.

[16] Benford F. The law of anomalous numbers. *Proceedings of the American philosophical society*, 1938, 551–572.

## 附录：中英文对照表

中文	英文
贝叶斯收缩估计量	Bayesian Shrinkage Estimator
错误发现率	False Discovery Rate, FDR
电子病历报告表	Electronic Case Report Form, eCRF
关键质量因素	Critical to Quality Factors
纠正与预防措施	Corrective and Preventive Action, CAPA
控制图	Control Chart
马氏距离	Mahalanobis Distance
欧氏距离	Euclidean Distance
现场监查	On-site Monitoring
预警阈值	Secondary Limit
中心化监查	Centralized Monitoring
质量风险的容忍度	Quality Tolerance Limit, QTL