

附件 1

药物真实世界研究设计与方案框架

指导原则（试行）

2023 年 02 月

目 录

一、概述.....	1
二、真实世界研究设计的主要类型.....	1
(一) 观察性研究设计.....	2
(二) 实用临床试验设计.....	4
(三) 单臂研究设计.....	6
三、真实世界研究方案的主体框架.....	8
(一) 方案摘要.....	9
(二) 研究背景.....	9
(三) 研究目的.....	9
(四) 研究假设.....	10
(五) 整体设计.....	10
(六) 研究人群.....	10
(七) 治疗或干预.....	11
(八) 研究终点.....	12
(九) 基线变量及重要协变量.....	13
(十) 观察期/随访期与观测/随访时间点.....	14
(十一) 数据治理/数据管理计划.....	14
(十二) 偏倚考虑.....	15
(十三) 统计分析计划.....	15
(十四) 质量控制.....	20
(十五) 伦理.....	20
(十六) 注册登记.....	21
(十七) 方案修订.....	21
(十八) 组织实施.....	21
四、真实世界研究设计的其它考虑.....	21
(一) 真实世界研究路径的可行性.....	21
(二) 目标人群的代表性.....	21

(三) 混合型研究设计	22
(四) 估计目标	23
(五) 模仿目标临床试验	25
五、与药审中心的沟通	25
参考文献	26
附录1 名词解释	28
附录2 中英文词汇对照	32

药物真实世界研究设计与方案框架

指导原则（试行）

一、概述

已发布的《真实世界证据支持药物研发与审评的指导原则（试行）》和《用于产生真实世界证据的真实世界数据指导原则（试行）》为真实世界证据支持药物研发与审评奠定了基础。采用真实世界研究支持儿科药物、罕见疾病药物等监管决策的指导原则也相继发布。

为了指导申办者科学合理的设计真实世界研究，明确真实世界研究方案撰写的技术要求，本指导原则将重点阐述药物研发及评价中真实世界研究设计以及研究方案制订的基本考虑，为药物研发中开展真实世界研究提供指导意见。

本指导原则适用于通过真实世界研究获得药物评价的临床证据。真实世界证据支持药物研发和监管决策的适用情形参见《真实世界证据支持药物研发与审评的指导原则（试行）》。本指导原则也可供以非注册研究为目的的真实世界研究参考。

二、真实世界研究设计的主要类型

真实世界研究设计包括观察性（或非干预性）研究设

计和干预性研究设计（如实用型临床试验）。单臂研究设计是一种特殊的设计形式，其研究组可以是干预性的，也可以是观察性的，其外部对照通常基于真实世界数据而设定。

（一）观察性研究设计

观察性研究可分为队列研究、病例对照研究和横断面研究等。以因果推断为目的的观察性研究通常采用队列研究设计。本指导原则后文若无特别说明，所述观察性研究均是指队列研究。

根据研究方案中定义的真实世界研究起始时间和结局发生的时间，队列研究可分为回顾性、前瞻性和回顾前瞻性队列研究。回顾性队列研究收集的是历史数据，即研究开始前生成的数据；前瞻性队列研究收集的是研究开始后的数据；回顾前瞻性队列研究既收集已有的历史数据，也收集研究开始后的数据。

队列研究设计主要考虑目标人群队列、因果推断和质量控制三个方面。其它方面的考虑在下文的研究方案部分有具体阐述。

1. 目标人群队列

目标人群队列根据临床所关心的问题而定，具体以数据体现，即目标人群从研究的治疗开始到观察期结束所形

成的纵向观测数据。目标人群队列的具体定义应基于研究目的、入排标准（包括本研究用治疗的初治者（本指导原则在实操上定义为：纳入研究队列之前在充分的洗脱期内未使用研究用治疗的病例）或非初治者）、数据来源和数据治理/管理计划综合考虑。鉴于数据来源的多样性，需充分评估研究人群的代表性及研究结论的外推性。目标人群所收集的重要变量包括治疗（含研究队列和对照队列）、基线、协变量（如基线协变量、时依变量）和结局变量等。观察性研究的样本量应在充分考虑混杂因素、缺失数据等因素的基础上满足统计假设的要求，通常不设上限，特别是回顾性研究。队列起始时间、观察期/随访期的长短和观测时间点/访视点的确定应符合所研究疾病的特征、临床实践和临床评价要求。

2. 因果推断

观察性研究由于变量间因果关系的不确定性和复杂性使得因果推断具有挑战。不同分析模型的选择往往会导致分析结果不同，因此，为了避免结果驱动的偏倚，需要在设计阶段明确主分析将要采用的分析数据集、分析模型及其相对应的统计假设。为使研究结果更为准确和稳健，应考虑混杂偏倚、选择偏倚、信息偏倚等重要偏倚识别及控制方法，以及缺失数据的处理策略及其基于的假设；还应

针对可能影响研究结果的各种因素，如模型假设背离或各类潜在偏倚来源，充分考虑敏感性分析及定量偏倚分析计划和策略。

3. 质量控制

质量控制的主要目的是保证获得高质量的分析数据。一方面，需要事先制定数据治理计划（针对历史数据）或数据管理计划（针对前瞻性收集数据），保证所产生的数据能够满足适用性要求（参见《用于产生真实世界证据的真实世界数据指导原则（试行）》）；另一方面，应制定具体措施保障观测变量值的准确性，例如在保障测量工具、度量单位和评价方法的一致性方面的具体措施。

（二）实用临床试验设计

实用临床试验(pragmatic clinical trial, PCT)又称实操临床试验或实效临床试验，是指尽可能接近真实世界临床实践的临床试验，是介于传统的随机对照试验（randomized controlled trial, RCT）和观察性研究之间的一种研究类型，属于干预性研究。与 RCT 不同的是：PCT 的干预既可以是标准的，也可以是非标准的；既可以采用随机分组方式，也可以自然选择入组；受试病例的入选标准可以相对较宽泛；对干预结局的评价不局限于临床有效性和安全性；PCT 更多地使用临床终点，而很少使用传统 RCT 中可能使用的

替代终点；可以同时考虑多个治疗组，以反映临床实践中不同的标准治疗，或设置多个剂量组达到剂量探索目的；一般不设安慰剂对照；如果因难以实施而不采用盲法，应考虑如何估计和控制由此产生的偏倚；数据的收集通常依赖于患者日常诊疗记录，但也可以设置固定的随访时间点，其时间窗通常较 RCT 更宽。与观察性研究不同的是，PCT是干预性研究，尽管其干预的设计具有相当的灵活性。

PCT设计应重点考虑以下因素：①收集到的数据是否适用于支持产生真实世界证据；②治疗领域和干预措施等是否符合各种形式的常规临床实践；③是否具有足够的可以用于评价的病例数（特别是临床结局罕见的情况）；④参与 PCT 的各试验中心甚至不同的数据库之间对终点的评价和报告方法是否一致；⑤是否采用随机化方法控制偏倚；⑥当盲法不可行时，应考虑非盲对结局变量，特别是患者报告的结局，可能产生的影响，可使用不受治疗分组影响的客观终点（如中风、死亡等），以减少非盲可能带来的偏倚；⑦分析方法的考虑可参照观察性研究的分析方法。

对于实用随机临床试验，还需要特别阐明治疗策略的选择（如单次治疗策略或持续治疗策略）和有效性的主分析所基于的数据集。由于P-RCT在随机化之后出现的治疗策略更改、剂量改变、停药、转组、数据缺失等情况较RCT更

为普遍，因此，相较于RCT通常基于ITT/mITT（调整ITT）进行主分析，P-RCT则需要考虑基于符合方案数据集是否更为合理的问题，或者考虑更加合适的数据集定义，并在样本量计算时予以充分考虑。

（三）单臂研究设计

采用单臂研究首先要考虑的问题是其前提条件是否充分，例如，采用RCT难以实施或具有重大伦理风险，属于危及生命、复发难治、无药可治或甚为罕见的疾病。单臂研究组如果是干预性的，为单臂试验；如果是非干预性的，为单臂观察性研究。无论是干预或非干预的，单臂研究设计通常应设置外部对照，外部对照采用的形式有基于疾病自然史队列数据或其他外部数据的历史对照或平行对照，或者目标值对照。为了减少偏倚，采用外部对照需考虑其目标人群特征（人口学、基线水平和临床特征等）、诊断和治疗标准、伴随治疗、结局的测量和评价标准等对结局（预后）有潜在影响的各种因素与研究组是否足够相似，以保证与研究组有较好的可比性。此外，单臂研究设计至少还应考虑以下内容。

1. 研究组设置

研究组的设置主要分干预性和非干预性，前者更为常用。对于干预性设计，研究组需要定义标准干预，且在研究实

施过程中严格执行所规定的干预措施；对于非干预性设计，研究的治疗通常没有统一标准，且在治疗过程中患者可能会同时接受其它治疗，使得治疗模式较为复杂多样，对此可通过设置合理的入选和排除标准来定义较明确的目标治疗。

2. 对照设置

(1) 历史对照

以既往获得的疾病自然史队列或其它外部真实世界数据作为对照，应考虑人群异质性及不同历史时期对疾病的定义、诊断、分类、自然史和可用的治疗手段等对疗效可比性和一致性的影响。

(2) 平行外部对照

收集与研究组同期的疾病自然史队列或其它外部真实世界数据作为对照。

(3) 目标值对照

目标值的确定应有充分依据，优先依次考虑国家标准、行业标准和专家共识，否则，需要根据已有的相关信息，包括但不限于公开发表的文献、研究报告、相关研究的原始数据等，通过综合分析确定目标值。

(4) 混合对照

将既往及研究同期获得的外部数据混合在一起形成对照臂。这些外部数据可以是日常的病例记录，也可以是过

去开展不同的临床研究（观察性或干预性的）所获得的数据。研究开始前需评估外部数据的适用性、代表性和预先设定不同部分数据合成时的权重系数，建议预先设置敏感性分析评估混杂因素、不同权重系数等对研究结论的影响。

3. 其它考虑

采用外部对照的单臂研究由于混杂因素、人群异质性和各种可能偏倚的影响，因果推断结论具有较大的不确定性。为克服或减少这些局限，除上述考虑外，还应注意：

①主要终点采用客观指标，如肿瘤临床研究的客观缓解；②明确并严格把握入组人群的入排标准及筛选过程；③要确保所采集的数据符合真实世界数据的适用性要求；④较之于历史对照，更鼓励采用平行外部对照；⑤事先恰当地定义主分析的统计分析方法，如合理利用多因素模型、倾向评分方法，虚拟对照方法、工具变量方法等；⑥若对照组选择或主分析模型采用基于匹配的方法，应在方案中事先明确匹配标准；⑦要充分使用敏感性分析和偏倚的定量分析来考察未知或未测量的混杂因素、效应异质性、模型假设不成立以及其它各种可能偏倚对分析结果的影响。

三、真实世界研究方案的主体框架

不同设计类型的真实世界研究方案的主体框架基本相同，个别不同之处将在主体框架下相应内容中分别阐述。

以下是真实世界研究方案的建议框架，但并不排除个别研究项目的某些特殊考虑。

（一）方案摘要

以表格形式摘录研究方案的主要内容，突出重点，力求简洁。主要内容包括：标题、研究背景、研究目的、研究假设、整体设计、研究人群（含诊断标准、入选标准、排除标准、剔除标准等）、治疗或干预（定义研究组和对照组）、研究终点、基线变量与重要协变量、安全性指标、观察期与观测时间点、数据来源、数据治理或数据管理、样本量及其确定依据、统计分析、偏倚控制等。

（二）研究背景

简要介绍研究背景，包括国内外研究的现状和意义，本研究的前期基础等。应充分论述选择真实世界研究路径的主要依据（如其它研究类型的不可行性和伦理风险等），以及本研究的定位，例如用于注册的关键支持证据、辅助支持证据、或基于探索目的的研究等。

（三）研究目的

根据目标人群、治疗（含对照）和结局，简要阐述研究目的，即本研究计划回答的临床科学问题，明确主要目的和次要目的（如果有），也可包括探索性目的。

（四）研究假设

根据研究目的提出研究假设。

（五）整体设计

简述研究的整体设计，包括多中心或单中心、观察性或干预性、单臂或双臂/多臂等要素。若是观察性研究，应说明是回顾性的还是前瞻性的，或回顾前瞻性的。

干预性研究应说明是否采用随机化分配，如果采用，应详细说明具体的随机化分配方法及实施过程；是否采用盲法，如果采用盲法（单盲或双盲），应说明具体实施办法；如果采用开放设计，需说明终点事件是否采用盲评，若采用盲评，如何实施。

单臂研究应说明研究组是干预性还是观察性的，以及采用何种形式的外部对照。

（六）研究人群

1. 诊断标准

如果所研究的疾病有不同诊断标准，应该说明本研究所采用的诊断标准及其出处，并给出所采用诊断标准的具体内容，若内容较多可以在附件中呈现。还应标明疾病代码（如 ICD9/ICD10 等）。

2. 入选/排除标准

入排标准的制定应能代表研究的目标人群。一般而言，

观察性研究的入排标准较干预性研究宽松。应注意入排标准可能导致的恒定时间偏倚或选择偏倚，必要时对重要的入排标准做出解释，并评估其对分析结果的影响。

（七）治疗或干预

本指导原则中，关于采用研究药物或治疗策略的人群队列，对于观察性研究，称之为“治疗组”或“治疗队列”；对于干预性研究（如 PCT），称之为“试验组”。关于采用非研究药物或治疗策略的人群队列，称之为“对照组”或“对照队列”。

1. 治疗组/试验组

对于治疗组的定义，应阐明具体治疗方法，如药物治疗的剂量、频次、给药途径、疗程等，以及药物的商品名和生产厂家；如果是物理治疗（如放疗或激光治疗），应给出具体的治疗参数。观察性研究中治疗策略和治疗模式由临床实践所决定，因此具有多样性，在数据收集、因果推断和结果解释时应予以考虑。

对于试验组的定义，与观察性研究不同的是，治疗方法通常应固定下来，形成相对标准的治疗策略。

2. 对照组

真实世界研究通常选择阳性对照或标准治疗对照，阳性对照应是目前或数据采集的起止期（如回顾性研究）临

床实践中公认的疗效明确的治疗方法或治疗策略。对照组应像治疗组或试验组一样描述具体的治疗方法或策略。

除单臂研究的历史对照外，观察性研究对照的选择应与治疗组同时期。对于回顾性研究，为了避免病例选择偏倚，原则上应选择研究所定义的数据采集的起止期内所有的治疗组和对照组的病例，或者采用严格的随机抽样方法选择病例（例如因为已有数据量太大而无法承受治理和分析全部数据的情况）。对于前瞻性研究，对照组的选择标准，特别是与治疗组的匹配方法，应明确定义。

干预性研究对照组的选择与 RCT 类似。

单臂研究对照的选择见上述第三章的“（三）单臂研究设计”中的“对照设置”一节。

3. 伴随治疗

真实世界研究中，伴随治疗的情况较为常见，应在方案中尽可能地阐述清楚可能出现的伴随治疗，对于未能预见的伴随治疗，也需要在分析过程中予以充分考虑。

（八）研究终点

1. 有效性终点

应定义主要终点和次要终点，如必要还需定义关键次要终点。有效性终点应完整定义，包括终点的名称、观测的时间点或时间段、测量方法与工具、计算方法、评价方

法等。必要时，可设置独立第三方终点事件判定委员会，并描述实施办法，例如所执行的标准操作程序（SOP）。应注意，真实世界研究的主要终点通常不采用替代终点，如采用需充分说明理由。

2. 安全性终点

根据研究目的，安全性终点可以是主要终点、（关键）次要终点、或者探索性终点。除了明确定义终点事件外，还应考虑对终点事件的编码（如用 MedDRA 编码）、分级（如用 CTCAE 对安全性事件的严重程度分级）、发生时间、发生频率（如反复多次发生，如何计算发生率）等加以说明。需要指出的是，与有效性终点不同，安全性终点在大多数情况下无法预先确定具体终点事件及其发生的时间和严重程度，因此其不确定性给统计分析和结果的解释带来挑战，具体考虑可参照相关临床试验指南。

3. 探索性终点（如果有）

研究如有需要，也可以设置某个/些探索性终点，例如药物经济学终点等。

（九）基线变量及重要协变量

研究方案中应明确基线变量和重要的协变量，以及它们的度量单位和观测时间。这些变量的确定依据主要来自对目标研究人群现有的研究成果，例如指南、专家共识、

公开发表的文献、会议报告等提及的影响疗效的变量/因素，也有来自项目组专家的认识。重要协变量的确定应具备合理性，可结合各因素间的因果路径图确定，并综合考虑前期数据评估结果。在已确定的重要协变量中，建议在方案中明确协变量的属性，如效应修正因素、危险因素、混杂因素（包括时依混杂因素）、中间变量、碰撞变量、工具变量等。

（十）观察期/随访期与观测/随访时间点

应明确研究的观察期或随访期，以及观测对象的观测或随访的起始时间、时间间隔和时间点，合理定义窗口期。

（十一）数据治理/数据管理计划

在真实世界研究中，应准确理解数据治理和数据管理的概念。对于既往数据，无论是病历记录的原始数据，还是开展不同临床研究所获得的数据，都应经过统一的数据治理使其满足分析的要求。对于前瞻性收集的数据，应通过严格和规范的数据管理，为研究提供高质量的用于分析的数据。数据治理计划通常应与研究方案同步完成。

应明确研究数据来源，包括所来自的研究中心、收集数据的起止时间、数据存储的系统和记录形式。如果是来源于既往的研究，应描述原始数据的记录和存储形式，以保证研究数据可追溯。

数据治理和数据安全的具体要求可参见《用于产生真实世界证据的真实世界数据指导原则（试行）》。

（十二）偏倚考虑

偏倚是真实世界研究特别需要考虑的问题，在方案中应充分考虑各种潜在偏倚及其影响，并制定控制偏倚的有效措施。常见的偏倚包括：因测量、数据收集或评价方法的不准确或不一致导致的信息偏倚，因选择性地入选和/或排除数据或失访、退出、剔除、记录缺失等导致的选择偏倚，因人群变化、治疗变化、研究背景变化等原因导致的疗效异质性，因分析中未能充分控制混杂因素导致的混杂偏倚，因未事先确定主分析方法而选择采用不同分析方法中最有利的结果导致的结果驱动等偏倚。此外，不同的研究中还可能发生其它具体的信息偏倚，例如，在记录生存时间时可能产生的恒定时间偏倚或领先时间/起点时间偏倚，基于文献的 meta 分析可能存在的发表偏倚，回顾性研究中回忆以往事件可能产生的回忆偏倚，因入选非初治病例而产生的幸存者偏倚等。

（十三）统计分析计划

为了避免结果驱动偏倚和保证研究过程的透明性，真实世界研究特别强调，至少主分析计划应该与研究方案同步确定，这与 RCT 中规定统计分析计划可以在数据库锁定

之前完成有很大不同。如果主分析计划篇幅较大，可以附件形式呈现。独立的主分析计划除了摘录方案中的一些关键要素外，如研究目的、目标人群、终点指标及其定义等，在统计分析方面至少应该包含以下内容。

1. 样本量估计

临床试验的样本量估计通常要考虑的因素有：研究类型、比较类型（优效性或非劣效性）、统计分析方法、结局变量预期的效应量或参数、统计分布、检验水准、单双侧检验、检验效能、分配比例、多重性、脱落剔除率、依从性等。真实世界研究应采用主分析所对应的样本量估计方法，在估计时，除了需要考虑上述因素外，还需考虑混杂因素等的调整对样本量的影响。

需注意，对于采用外部对照组的单臂试验研究，对照组的样本量通常应不少于治疗组的样本量，或可以数倍于治疗组。另外，观察性研究（特别是回顾性研究）的数据缺失率较高，设计时应留有充分余地。

2. 数据集定义

真实世界研究的数据来源及其质量有很大差异，而且不同分析回答的问题各异，应根据不同的分析定义不同的数据集，如有效性数据集和安全性数据集。如果涉及随机分配，应基于随机分组定义数据集。如果分析的目标人群

是数据集的一个子集，应将子集标记为对应的目标人群。

3. 缺失数据处理

真实世界研究中，数据缺失较为普遍，甚至缺失比例较大。在数据治理及数据管理过程中，应尽可能追踪捕捉遗漏的记录，使数据质量有所改善。尽管如此，主分析或敏感性分析中仍面临缺失数据的处理问题，在主分析计划中和敏感性分析计划中（如适用）应阐述缺失数据的处理方法，并说明其理由。

4. 描述性分析

描述性分析能够刻画变量（特别是基线变量）的主要特征。所有指标/变量，包括终点变量，均应进行描述性分析，所采用的描述统计量应根据变量的分布特征合理选择。

5. 异质性分析

应事先考虑可能的异质性因素，如研究中心、年龄、性别、病情程度等，为亚组分析或分层分析打好基础。同时阐述异质性的评估方法，如采用何种分析模型；以及异质性的判断标准，例如以 0.10 的检验水准判断分组与潜在异质性因素是否存在交互作用，但应注意异质性判断标准的确定应结合研究目的和临床意义综合考虑。

6. 主分析

主分析围绕主要终点进行统计分析，是研究结论的最主要依据，应给予详尽和严谨的阐述，包括但不限于：统

计假设；非调整分析和调整分析所采用的模型及其所基于的假设；拟纳入调整分析的协变量的初步考虑以及明确在分析过程中根据观测数据筛选各类变量的明确规则，包括混杂因素/时依混杂因素、风险因子、中间变量和潜在异质性因素的识别；如果采用倾向性评分匹配方法，应定义匹配比例、匹配方法及其具体参数设置（如卡钳值），以及匹配的均衡性验证方法；对于生存结局的分析需考虑是否存在竞争风险问题。此外应对模型假设进行必要的验证，如非线性关系、非等比例风险等。需要指出，即使 PCT 设计中采用随机分配策略，其主分析对协变量的考虑仍建议与观察性研究相同，因为在研究的实施过程中，PCT（特别是群随机设计）对基线均衡性的控制远不如 RCT 严格。有关因果推断的具体方法可参阅《真实世界证据支持药物研发与审评的指导原则（试行）》的附录部分和其它相关专业文献。

7. 亚组分析

应根据现有的研究结论和认知、可能的异质性因素等明确定义需要进行亚组分析的因素，此外也可考虑主要协变量中与分组变量交互显著的因素进行亚组分析。有关亚组分析的具体方法可参阅《药物临床试验亚组分析指导原则（试行）》。

8. 敏感性分析

由于真实世界研究中因果推断结论具有不确定性，因此结论的稳健性尤其重要，充分的敏感性分析有助于判断结论的稳健性。敏感性分析应根据不同的假设情景展开，这些情景包括但不限于：缺失数据的不同处理机制；不同的数据集定义；不同的分析方法；模型中不同的协变量组合；未知或不可测的混杂因素影响等。

9. 定量偏倚分析

偏倚对研究结论的影响是因果推断中需要特别考虑的。鼓励对于各种可能的偏倚，明确判断其是否存在，或设置相关偏倚参数值或分布的方法，并基于偏倚的定量分析考察其对结果的影响，例如，根据剔除标准，分别对剔除病例的数据集和不剔除病例的数据集进行分析，比较其结果的差异以判断是否存在选择偏倚；根据混合型研究中内部数据与外部数据间的疗效差异，判断是否存在效应异质性偏倚，并根据结果设置异质性偏倚参数分布进行校正分析。偏倚参数的分布反映偏倚的大小和不确定性，利用临界点分析考察各种可能偏倚的影响也可视为定量偏倚分析的一种方法。另外，敏感性分析与定量偏倚分析亦可合并描述。

10. 安全性分析

真实世界研究，特别是回顾性研究，对安全性事件的

主动监测存在较明显的不足，可能需要提供某些外部证据以弥补其不足，例如研究药物在其它研究中的安全性信息和不良反应监测系统报告的信息。如果研究假设要回答研究药物比对照药物具有更好的安全性，还应提供充分的对照药物的安全性信息。对于主要研究目的是回答安全性问题的情况，可参阅相关指导原则或文献。

需要指出，上述与研究方案同步的主分析计划主要是呈现事先确定的将要做的各种分析以及这些分析的假设和条件，而与分析对应的结果的具体输出形式（统计图表）可以在正式的数据分析之前确定。

（十四）质量控制

一般而言，真实世界研究的质量控制的目标与 RCT 类似，但需要特别关注数据治理过程的质量控制，具体可参见《用于产生真实世界证据的真实世界数据指导原则（试行）》。

（十五）伦理

真实世界研究的伦理要求可参照国家卫生健康委员会《涉及人的生物医学伦理审查办法》等管理要求执行，回顾性观察性研究经伦理委员会审查批准后可采用泛知情同意等形式。

（十六）注册登记

应描述本研究在公共网站注册登记的情况。

（十七）方案修订

在真实世界研究的实施过程中，如果更改数据治理计划或统计分析计划中的主分析计划，属于方案的实质性变更，需将修订方案与药审中心充分沟通以达成一致。

（十八）组织实施

可参照一般的临床研究方案，并根据实施项目的特点，制定实施计划。

四、真实世界研究设计的其它考虑

（一）真实世界研究路径的可行性

在进行研究设计前，应首先对采用真实世界研究路径的可行性进行评估，包括但不限于以下考虑：①传统 RCT 是否不可行；②是否有比 RCT 更好或可替代的研究路径；③真实世界数据是否足以支持将要开展的研究。无论是数据的质量还是数量（样本量）应均能支持统计分析并产生真实世界证据；④该项真实世界研究在药物研发中的定位，明确该研究所形成的证据在整个证据链中的作用。

（二）目标人群的代表性

观察性研究中目标人群的代表性非常重要，确定研究人群的理想方法是采用严格的随机抽样。然而，由于临床

研究的实际情况，研究人群通常采用的是便利抽样的方法确定的，因此，应充分评估研究人群与目标人群的特征是否存在异质性，及其对研究结论外推性（即外部效度）的潜在影响。

（三）混合型研究设计

本指导原则中，混合型研究是指同时基于真实世界数据和真实世界数据外的研究数据的研究。含有实用元素的随机对照试验和利用真实世界数据混合形成研究臂和/或对照臂（以下简称混合臂）而开展的研究是其中较为典型的应用。

混合臂研究设计的关键是将内部数据与外部数据的合并应基于合理的统计模型和方法，以保证内部人群与外部人群特征相一致为原则，将外部数据根据个体水平或整体水平、匹配或赋权等方式与内部数据进行融合，并尽可能进行充分的敏感性分析和定量偏倚分析。若采用基于贝叶斯理论的方法，还应配合充分的针对先验分布和其它相关参数设置的模拟分析。由于内外部人群特征重叠程度及效应一致性程度都会影响外部数据所能够提供的有效样本量，因此，混合臂研究估算所需样本量时应确保当前试验纳入足够数量的受试者，使分析结果达到稳健和可靠。

（四）估计目标

ICH E9 (R1) 将临床试验中构建估计目标归纳为五个重要属性，即目标人群、治疗、终点、伴发事件和汇总统计量。真实世界研究中，如何构建估计目标目前仍处于探索阶段，但与传统的RCT相比，还需考虑一些更复杂的问题，例如，研究人群的异质性、治疗方法的灵活性、伴发事件的多样性、终点选择的特殊性、敏感性分析的复杂性等。现阶段，本指南对真实世界研究中围绕估计目标的主线进行设计并无特定要求，但也鼓励真实世界研究设计中积极探索估计目标实施的可行性。以下是真实世界研究设计中估计目标需要特别考虑的问题。

1. 研究人群的异质性

由于真实世界研究中的入选标准较为宽松，且不采用随机化或不能严格按照随机分配方案实施，故人群异质性是普遍存在的，其来源除了人口学特征、临床特征、地域和研究中心等外，还可能包括不愿意参加 RCT 的患者或 RCT 中代表性不足的患者群体（例如少数民族、老年人和居住在偏远地区的人）。因此，在疗效估计中应考虑所适用的目标人群。

2. 治疗方法的灵活性

真实世界研究中，对患者的治疗策略通常有较多的选择，如不同的治疗剂量、合并治疗、疗程等。患者的治疗

依从性以及治疗选择的偏好等，都是定义治疗时需要考虑的。

3. 伴发事件的多样性

真实世界研究中，除了需要考虑RCT中常见的因安全性、不耐受性、缺乏疗效等引起的伴发事件和终点伴发事件（如死亡）外，还应考虑因患者行为因素（如患者对治疗的偏好、治疗的便捷性、医患关系等）和非行为因素（如医保政策的调整影响现在接受的治疗药物、健康状况改善等）导致治疗中断的伴发事件，它们对疗效的影响是不同的。

4. 研究终点选择的特殊性

真实世界研究一般选择临床终点而非替代终点，最好选择单一测量、易观测的临床结果（例如死亡或住院）；复合临床终点的使用应确保构成复合事件的每一组成部分能够获得有效记录，否则要慎重考虑其可行性，例如使用量表工具时，某一维度或条目记录的严重缺失，可能导致整个量表的评价结果失效。

5. 敏感性分析的复杂性

真实世界研究由于需要控制混杂和偏倚而使得因果推断具有复杂性，同时为使估计量更为准确和可靠（稳健），也对敏感性分析提出了更高的要求。

此外，真实世界研究中还有许多面临挑战的问题，如

数据融合问题，观察性研究中个体生存时间记录的区间缺失等，有待深入研究和探索。

（五）模仿目标临床试验

模仿目标临床试验是一种真实世界研究方法，即基于现有的真实世界数据，模仿一个良好的 RCT 设计（明确的入排标准、治疗策略、治疗分配方法、随访期限和时间点、终点事件评价和统计分析计划等），产生一个真实世界研究的子集，通过因果推断得出研究结论。该法有助于思考、识别和避免不必要的偏倚，如恒定时间偏倚、非初治者偏倚等，并提供合理的方法来阐明在观察性研究中可能需要做出的权衡。模仿目标临床试验要考虑使用的场景，该法的前提是需要拥有非常大样本量的适用真实世界数据，并且具有较高的 RCT 的仿真度，目前在应用方面有待获得共识，但不失为一种值得探索的方法。

五、与药审中心的沟通

就真实世界研究相关技术问题与药审中心的沟通交流可参照《真实世界证据支持药物注册申请的沟通交流指导原则（试行）》。

参考文献

- [1] 国家药品监督管理局. 真实世界证据支持药物研发与审评的指导原则(试行). 2020.
- [2] 国家药品监督管理局. 用于产生真实世界证据的真实世界数据指导原则(试行). 2021.
- [3] 国家药品监督管理局. 真实世界研究支持儿童药物研发与审评的技术指导原则(试行). 2020.
- [4] 国家药品监督管理局. 基于人用经验的中药复方制剂新药临床研发指导原则(试行). 2022.
- [5] 国家药品监督管理局. 罕见疾病药物临床研发技术指导原则(试行). 2022.
- [6] 国家药品监督管理局. 罕见疾病药物临床研究统计学指导原则(试行). 2022.
- [7] 国家药品监督管理局. 单臂试验支持上市的抗肿瘤药上市许可申请前临床方面沟通交流技术指导原则(试行). 2020.
- [8] 国家药品监督管理局. 单臂试验支持上市的抗肿瘤药进入关键试验前临床方面沟通交流技术指导原则(试行). 2020.
- [9] 国家药品监督管理局. 药物临床试验亚组分析指导原则(试行). 2020.
- [10] Duke Margolis Center for Health Policy. A frame work for regulatory use of real-world evidence. 2017.
- [11] Boslaugh S . Encyclopedia of epidemiology. [M] SAGE Publications. 2008.
- [12] Chen EY, Raghunathan V, Prasad V. An overview of cancer drugs approved by the us food and drug administration based on the surrogate end point of response rate[J]. JAMA Intern Med .2019;179(7):915-921
- [13] Hernán MA, Robins JM. Per-protocol analyses of pragmatic trials[J]. N Engl J Med, 2017, 377(14): 1391-1398.
- [14] Hernán MA, Robins JM. Causal Inference[M]. Boca Raton: Chapman & Hall/CRC 2019.
- [15] ICH E10: Choice of Control Group and Related Issues in Clinical Trials. 2000
- [16] James S. Importance of post-approval real-word evidence[J]. Eur Heart J Cardiovasc Pharmacother, 2018; 4(1):10-11.
- [17] Last JM. A Dictionary of Epidemiology[M]. 4th Edit. Oxford University Press. 2001.
- [18] Lash TL, Fox MP, Fink AK. Applying quantitative bias analysis to epidemiologic data[M]. Springer Science & Business Media, 2011.
- [19] Lash TL, Fox MP, MacLehose RF, et al. Good practices for quantitative bias analysis[J], Int. J. Epidemiol., 2014, 43(6): 1969–1985.

- [20] Roland M, Torgerson DJ. Understanding controlled trials: What are pragmatic trials?[J]. *BMJ*, 1998, 316(7127): 285.
- [21] Sherman RE, Anderson SA, Dal Pan GJ, et al. Real-world evidence—what is it and what can it tell us[J]. *N Engl J Med*, 2016, 375(23): 2293-2297.
- [22] Strayhorn JM. Virtual controls as an alternative to randomized controlled trials for assessing efficacy of interventions [J]. *BMC Med. Res. Methodol.* 2021; 21(3): 1-14.
- [23] Sugarman J, Califf RM. Ethics and regulatory complexities for pragmatic clinical trials[J]. *JAMA*, 2014, 311(23): 2381-2382.
- [24] US FDA. Framework for FDA’s real-world evidence program. December 2018. 2019.
- [25] US FDA. FDA Submitting Documents Using Real-World Data and Real-World Evidence to FDA for Drug and Biological Products. 2022.
- [26] Velentgas P, Dreyer NA, Nourjah P, et al. Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide. Rockville (MD): Agency for Healthcare Research and Quality (US), 2013.
- [27] Von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE)statement: guidelines for reporting observational studies[J]. *Ann. Intern. Med.* 2007, 147(8): 573-577.

附录 1 名词解释

1. 初治者 (New User) : 纳入研究队列之前在充分的洗脱期内未使用研究用治疗的病例。

2. 单臂临床研究 (Single-arm/One-arm Study) : 一种只设置试验组或治疗组的非随机临床研究, 通常采用外部对照, 如历史对照、平行对照或目标值对照等。

3. 定量偏倚分析 (Quantitative Bias Analysis, QBA) : 可用于评估研究结果对各种可能的系统误差来源 (如错误分类、不受控制的混杂和选择偏倚等) 敏感性的一类方法, 可基于设置的各类偏倚大小及其分布考察各类偏倚对效应估计结果的影响方向和程度, 亦可用于提供进一步校正各种可能偏倚后的分析结果。

4. 估计目标 (Estimand) : 对治疗效应的精确描述, 反映了针对临床试验目的提出的临床问题。它在群体水平上汇总比较相同患者在不同治疗条件下的结局。

5. 观察性研究 (Observational Study) : 又称非干预性研究, 根据特定研究问题, 不施加主动干预的、以目标人群为对象的、探索治疗与结局因果关系的研究。

6. 回顾性观察性研究 (Retrospective Observational Study) : 在研究开始时确定目标人群、并根据历史数据 (研究开始前生成的数据) 开展的观察性研究。

7. 模仿目标临床试验 (Target Trial Emulation) : 是一种

真实世界研究方法，即基于现有的真实世界数据，模仿一个良好的 RCT 设计，产生一个真实世界研究的子集，通过因果推断得出研究结论。

8. 偏倚 (Bias)：任何在研究设计、数据收集、分析、结果解释、报告中系统性地导致估计量偏离真值的倾向。

9. 前瞻性观察性研究 (Prospective Observational Study)：在研究开始时确定目标人群、并在研究开始前确定将要收集的治疗和结局数据的观察性研究。

10. 实用临床试验 (Pragmatic Clinical Trial/Pragmatic Trial, PCT)：又称实操/实效临床试验，指尽可能接近临床真实世界环境的临床试验，是介于RCT和观察性研究之间的一种研究类型。

11. 数据治理 (Data Curation)：指针对特定临床研究问题，为适用于统计分析而对原始数据所进行的治理，其内容至少包括数据采集（可包含多个数据源）、数据安全性处理、数据清洗（逻辑判断及异常数据处理、数据完整性处理等）、数据导入和结构化（通用数据模型、归一化、自然语言处理、医学编码、衍生点位等）、数据传输等若干环节。

12. 外部对照 (External Control)：在临床试验中，以研究对象以外的数据为对照，以评价所研究的治疗或干预

效果。外部对照可以是历史数据，也可以是平行观测所获得的数据，还可以是目标值。

13. 虚拟对照方法（Method of Virtual Control）：基于反事实论证思路，首先根据单臂设计所考虑的重要变量，利用现有真实世界数据建立不用试验药物的预后预测模型；然后将单臂试验所获得的协变量数据代入预测模型，计算得到假设不用试验药物的预后预测结果作为对照（虚拟对照）；最后将试验药物实际观测到的结果与虚拟对照进行比较，从而做出是否有效的推断。

14. 因果推断（Causal Inference）：基于真实世界数据，刻画干预或治疗与临床结局或健康结局的因果关系路径，充分考虑各种协变量和已测或未测混杂因素的影响，并控制可能的偏倚，采用恰当的统计模型和分析方法，做出干预或治疗与临床结局或健康结局的因果关系的推断结论。

15. 真实世界数据（Real-World Data, RWD）：来源于日常所收集的各种与患者健康状况和/或诊疗及保健有关的数据。并非所有的真实世界数据经分析后都能成为真实世界证据，只有满足适用性的真实世界数据才有可能产生真实世界证据。

16. 真实世界研究（Real-World Research/Study, RWR/RWS）：指针对预设的临床问题，在真实世界环境下收集与研究对象健康状况和/或诊疗及保健有关的数据（真

实世界数据)或基于这些数据衍生的汇总数据,通过分析,获得药物的使用情况及潜在获益-风险的临床证据(真实世界证据)的研究过程。

17. 真实世界证据(Real-World Evidence, RWE):指通过对适用的真实世界数据进行恰当和充分的分析所获得的关于药物的使用情况和潜在获益-风险的临床证据。

18. 中间变量(Intermediate Variable):指处于因果关系链中间、既受药物治疗影响、同时又影响结局的变量,或与结局有关联的变量;前者又称中介变量(mediator)。

19. 主分析(Primary Analysis):指针对主要终点采用的统计分析方法。

20. 主分析计划(Primary Analysis Plan):指针对主要终点采用的统计分析方法以及相应的增加结果稳健性的敏感性分析计划。在真实世界研究中,特别强调主分析计划应在研究方案中事先确定下来,如果在研究过程中修订主分析计划,属于方案的重大调整。

附录 2 中英文词汇对照

中英文词汇对照表

中文	英文
标准操作程序	Standard Operation Procedure, SOP
病例对照研究	Case-control Study
初治者	New User
定量偏倚分析	Quantitative Bias Analysis, QBA
队列研究	Cohort Study
发表偏倚	Publication Bias
工具变量	Instrumental Variable
估计目标	Estimand
观察性研究	Observational Study
恒定时间偏倚	Immortal-time Bias
横断面研究	Cross-sectional Study
患者报告结局	Patient Reported Outcome, PRO
回顾性研究	Retrospective Study
回忆偏倚	Recall Bias
混杂因素	Confounder
可追溯性	Traceability
每个协变量阳性事件数	Events per Variable, EPV
碰撞变量	Collider Variable
领先时间偏倚	Lead-time Bias

中文	英文
模仿目标临床试验	Target Trial Emulation
目标临床试验	Target Trial
起点时间偏倚	Zero-time Shift Bias
前瞻性研究	Prospective Study
倾向评分	Propensity Scores, PS
时依变量	Time-varying Variable
实用临床试验	Pragmatic Clinical Trial, PCT
数据管理	Data Management
数据治理	Data Curation
实用随机临床试验	Pragmatic Randomized Clinical Trial, P-RCT
统计分析计划	Statistical Analysis Plan, SAP
幸存者偏倚	Survivor Bias
虚拟对照	Virtual Control
衍生变量	Derived Variable
因果推断	Causal Inference
真实世界数据	Real World Data, RWD
真实世界研究	Real World Research/Study, RWR/RWS
真实世界证据	Real World Evidence, RWE
中间变量	Intermediate Variable
主分析	Primary Analysis
主分析计划	Primary Analysis Plan