

罕见疾病药物开发中疾病自然史研究  
指导原则

2023年7月

# 目录

一、背景 .....	1
二、疾病自然史研究定义 .....	2
三、疾病自然史研究的应用 .....	3
(一) 识别特定患者人群 .....	3
(二) 助力研究设计 .....	3
(三) 识别和开发生物标志物 .....	4
(四) 评估和建立临床结局评估工具 .....	4
(五) 作为外部对照数据 .....	5
(六) 其他应用场景 .....	7
四、疾病自然史研究类型 .....	8
(一) 回顾性队列研究 .....	9
(二) 前瞻性队列研究 .....	11
(三) 双向队列研究 .....	12
(四) 横断面研究 .....	13
五、疾病自然史研究的实施 .....	14
(一) 及早规划和实施 .....	14
(二) 鼓励患者群体的参与 .....	14
(三) 数据采集 .....	15
(四) 研究对象保护 .....	16
六、参考文献 .....	16

## 一、背景

疾病自然史是指在未经干预的情况下，疾病从发生到进展，直至稳定、痊愈或者恶化并导致个体死亡或永久性功能丧失的预后和转归过程。对疾病自然史的深入研究和全面了解，是人类认识疾病，并对疾病进行诊断、治疗以及开展药物研发的基础。罕见疾病由于其发病率和患病率极低，临床医生对疾病的认知普遍不足，开展罕见疾病的疾病自然史研究存在诸多困难，因此，我国很多种类的罕见疾病缺乏可靠的疾病自然史研究数据。近年来，随着医药行业对罕见疾病治疗药物研发热情不断增加，罕见疾病的疾病自然史研究就显得尤为重要。在我国，虽然目前已建立一些患者登记数据库，获得了部分罕见疾病的病史信息，但是由于现有罕见疾病患者登记数据库自身的局限性，可用且可靠的疾病自然史信息仍然十分有限。

为了鼓励罕见疾病药物研发，美国 FDA 于 2019 年发布了“Rare diseases: Natural History Studies for Drug Development”指南征求意见稿<sup>[1]</sup>，而我国目前尚无相关指导原则，为推动和规范我国罕见疾病的疾病自然史研究，特制定本指导原则。本指导原则旨在结合我国罕见疾病研究现状，提出符合我国研发实践的罕见疾病药物开发中疾病自然史研究的考虑要点。

本指导原则仅代表药品监管部门当前的观点和认知；随

着科学研究的进展，相关内容将不断完善与更新。应用本指导原则时，还请同时参考药物临床试验质量管理规范（good clinical practice, GCP）、人用药品技术要求国际协调理事会（International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use, ICH）和其他国内外已发布的相关指导原则。

## 二、疾病自然史研究定义

疾病自然史研究是指预先设计、旨在增加对疾病了解（包括疾病发展和转归）的观察性研究。其目的是识别人口统计学、遗传、环境和其他因素（例如，治疗方式、用药等）与疾病的发展和结局的关系。疾病自然史研究中，患者可能会接受当前的标准治疗和/或针对并发症的治疗，对于可能改变患者的一些临床表现甚至影响疾病进程的干预措施，应该作为研究要素纳入研究内容中。

疾病自然史研究强调预先的计划性，需要针对性地收集患者疾病过程的相关数据，以便系统地提供疾病发展和转归过程中的重要信息。

疾病登记系统是获得疾病自然史研究数据的常用平台，但两者并不完全等同，周密设计的疾病自然史研究可以通过可靠的疾病登记系统予以有效实施，反之，良好实施的疾病自然史研究也有助于提升登记系统数据的数据质量。

### 三、疾病自然史研究的应用

疾病自然史研究在医疗实践中广泛应用，并为制定和优化临床诊疗方案等提供重要依据。在药物研发方面，疾病自然史研究有助于全面和深入理解疾病发展的全过程，准确识别疾病发展中敏感且重要的变化因素，在新药研发各个阶段和环节发挥重要作用。

#### （一）识别特定患者人群

对于诊断为同一类罕见疾病的患者人群，可能具有显著的基因型和/或表型异质性。例如，不同表型可表现为不同系统的器官受累，或具有不同的严重程度或恶化速度。疾病自然史研究通过在广泛的患者群体中采集数据，可以帮助发现一些能够预测疾病进展或本身具有重要临床意义的特定事件或可检测的实验室检查指标（包括影像学等）的变化，获得对疾病不同基因型和表型的深入认识，以及对患者结局的影响因素的认知。因此，精巧设计的疾病自然史研究有助于精准定位特定的潜在获益人群，同时，可以按照疾病进展、恶化程度以及最终预后判断临床治疗需求的紧迫性，规划研究人群的优先级，制定科学合理的整体研发策略。

#### （二）助力研究设计

在确定潜在获益患者人群的基础上，进一步确定其症状、体征以及实验室检查（包括影像学等）的变化情况，了解其进展模式（受累器官、受累数量和程度分级等），上述信息

可作为重要依据指导临床研究方案中入排标准、干预阶段、试验持续时间、研究终点和随访数据采集时间点的合理选择。

同时，通过发现一些提示疾病进展的重要预测因素，作为探索治疗反应的预测因子，用于提高早期有效性评价和剂量选择的敏感性。

### （三）识别和开发生物标志物

生物标志物通常是指能被客观测量和评价，反映生理或病理过程，以及对暴露或治疗干预措施产生生物学效应的指标<sup>[2]</sup>。

疾病自然史研究有助于识别或指导开发生物标志物，用于药物开发过程的疾病诊断、病程监测、治疗反应预测，或有助于指导患者的用药选择和药物剂量选择。此外，通过疾病自然史研究而发现的一些生物标志物，经验证后，还可能作为临床试验的终点或替代终点。

### （四）评估和建立临床结局评估工具

临床结局评估（Clinical outcome assessment, COA）是描述或反映个体感受、功能、疾病进展或生存状况的重要工具。临床结局评估可在试验期间用于评估药物的临床获益（包括疗效和安全性）。通常，临床结局评估包括以下四种类型：

- 医生报告结局
- 观察者报告结局（例如，由照护者报告的临床结局）
- 患者报告结局

●功能测试结局（例如，记忆测试或行动能力测试的结果）

疾病自然史研究有助于开发新的临床疗效评估工具，或评价现有评估工具的准确性。

#### （五）作为外部对照数据

为确证药物用于特定患者人群的获益大于风险，开展周密设计并实施的随机对照临床试验仍然是金标准。然而罕见疾病药物研发的难点之一在于受试者少，难以开展大样本量的对照研究；此外，对于一些尚无有效治疗手段的罕见疾病，临床试验过程中，研究者希望受试者有更多接受潜在有效治疗的机会，开展与安慰剂对照的研究可行性较低。因此，在罕见疾病药物研发中，疾病自然史研究数据可作为外部对照数据用于评价药物的疗效<sup>[117]</sup>。

外部对照数据可分为以下两种类型：

（1）非平行外部对照数据，也称为历史对照数据，可以从既往临床试验的对照药物组（如安慰剂组）或自然史研究获得。

（2）平行外部对照数据，即临床试验中与试验组同时收集的对照组数据。

在一些情况下，历史对照可能作为一种外部对照数据使用；在某些情形下，历史对照中的患者可以进入到药物临床研究，这类患者的数据可以作为自身对照的部分，提供更具

有可比性的信息。但是，由于历史对照可能存在无法控制的偏倚，从而限制了其作为外部对照数据证明试验药物疗效的可靠性。然而，在病程可预测的一些情况下，疾病自然史研究可以描绘疾病发生发展过程特征（如某些罕见疾病在未经干预的情况下不会自发改善/缓解），从而为研究设计中的基线数据或疾病随时间恶化的相关数据提供依据。此外，在有些情况下，疾病自然史研究结果可提供未经治疗的疾病发生发展数据，也可作为外部对照用以证明试验药物的疗效。

如拟采用外部对照用于评价药物有效性，应充分考虑：治疗带来获益的改善程度是否显著超过潜在偏倚和已知的进展变异；有效性评估是否可以客观测量，不受患者或研究者的主观因素影响，也不受到研究对象选择的影响；治疗与有效性之间是否存在较强的时间相关性；药物药效学机制是否已在非临床研究中明确。如符合上述条件可用于作为研究的外部对照时，还应关注外部对照组与试验组是否具备影响预后判断的关键要素的可比性，包括疾病严重程度、疾病持续时间、既往治疗和其他任何可能影响结局和结局发生时间的因素。

由于缺乏随机和未能充分识别基线差异可能导致对预后的影响，选择偏倚是使用外部对照时需要重点关注的问题。选择偏倚可能包括（但不限于）以下方面：患者关键特征未被评估，或因评估方法的演变，导致评估结果缺少可比性（如

疾病定义、诊断技术和安全性监测方法等已经发生显著改变)；治疗标准可能发生了变化；数据收集间隔和质量可能缺乏一致性和可比性。如果对照组结局的评估方法不明确或不可靠，其数据不适用于支持药物的开发以及支持监管机构决策。

当使用疾病自然史研究作为外部对照研究时，其研究设计可参考真实世界研究的相关指导原则<sup>[3,4]</sup>。

#### (六) 其他应用场景

计划、组织和实施疾病自然史研究的作用远不止用于药物开发。通过实施规范的研究，疾病自然史研究还可以有助于提高医护团队对疾病的认知；帮助医护人员或研究者与患者建立良好沟通，辅助罕见疾病的诊断和加快新药临床试验入组；促进患者对当前诊疗标准和临床实践的理解，促进医护人员或药物研发人员了解患者需求，并不断改善和提高临床管理能力。

此外，疾病自然史研究还可提供人口统计学、疾病患病率等流行病学数据、疾病特征信息，辅助监测疾病进展，以及量化疾病的负担（包括直接和/或间接医疗成本），为药物开发成功后纳入医疗保障体系提供衡量依据，以提高广大患者对创新药可及性。

当计划利用罕见疾病自然史数据支持药物研发时，鼓励申请人就研发策略进行沟通交流。

#### 四、疾病自然史研究类型

观察性研究大致可分为队列研究、病例对照研究和横断面研究。队列研究通常收集纵向数据，分析干预或暴露因素以及其他因素对结局的影响，是一个由因到果的推断过程，是疾病自然史研究中的主要研究类型。病例对照研究是一个由果到因的推断过程，由于疾病自然史研究中的对照不好选择，故鲜有应用。

队列研究将某一特定人群按是否暴露于某种因素或依据暴露程度分为不同的亚组，追踪观察两组或多组人群疾病结局的情况，比较各组之间结局发生率的差异，从而判定影响因素与结局之间有无因果关联及关联强度。

相较横断面研究，队列研究在数个时间点收集队列中所有患者的数据。队列疾病自然史研究通常可提供更丰富、更全面的与疾病发生和进展以及转归相关的信息，是一种更有价值的疾病自然史研究。队列研究可以提供患者在未经治疗的情况下，疾病相关指标随时间的变化信息；此外，队列研究可以更好描绘或区分疾病表型和亚组的多样性，也有助于识别罕见疾病的预后因素。

队列研究的主要局限性在于与横断面研究相比，通常需要更长的研究时间，和更集中、更丰富的资源；另外，队列研究也需要考虑患者无法获益从而依从性差/患者脱落所可能导致的数据缺失。

队列研究又分为回顾性队列研究、前瞻性队列研究和回顾前瞻性队列研究。

### （一）回顾性队列研究

回顾性队列研究是研究者根据明确的研究目的，设计周密的研究方案，采集已有的临床纵向观测数据（这些数据一般来源于电子病历系统、医保数据库、患者登记系统等），通过统计分析最终得出结论的临床研究类型。

回顾性研究可通过已发表的科学文献、疾病诊疗专家的意见和经验，以及其他来源的信息获取疾病信息，也可以直接从患者处收集数据。

需注意的是，回顾性疾病自然史研究可能会有以下局限：

- 评估方法随时间变更，而导致历史数据与当前患者结果间缺乏可比性；

- 诊断标准随时间变化，或诊断准确性随医疗技术发展而不断提高，导致不同时期的诊断结果存在偏倚；

- 未收集或记录研究所需的数据；

- 在不同时间点或采用不同路径收集数据，所记录的数据可能会因为评估方法或数据质量不同，以及观察指标本身发生了变化，而与当前患者接受治疗的情况缺乏可比性；

- 医学术语改变，或者不同医疗机构间所使用的术语不一致，所导致的评估或分析错误；

- 特定医疗机构提供的历史数据可能存在患者选择偏

倚（例如，疾病自然史研究中仅包括疾病严重程度最重的患者）；

- 患者的病历或历史病程信息可能不足以识别和了解疾病的发生发展；

- 如果使用某一数据库的现有患者数据，在数据库中登记时间最长的患者信息可能会被过度使用，而那些刚刚被记录的患者数据可能未被捕捉，即存在病程选择偏倚<sup>[5]</sup>；

- 医学文献中发表的疾病自然史研究可能会侧重于罕见疾病中的某一类人群，如疾病程度更严重或治疗结局更成功的患者情况，因此文献研究通常不能反映疾病的全部特征，从而不能充分替代疾病自然史研究；如将医学文献发表的疾病自然史研究数据作为药物研发的参考信息，需要充分评估其质量及潜在偏倚；

- 回顾性疾病自然史研究可能因患者入选标准，或研究选择的开始日期和截止日期而产生偏倚。

上述因素均可能导致疾病信息收集不完整，并可能影响所获得的信息的代表性和可靠性。

虽然回顾性研究存在局限性，但是由于回顾性研究使用的是已有数据，因此其实施和推进可能更快。在我国对大多数罕见疾病缺乏充分认知的情况下，回顾性研究有利于用较短的时间弥补当前对于罕见疾病认识的不足。此外，回顾性研究还有助于确定哪些缺失的信息可能需要通过后续前瞻

性研究进行收集、分析和弥补。

## （二）前瞻性队列研究

前瞻性队列研究是研究者根据明确的研究目的，设计周密的研究方案，确定研究对象入排标准以及合理的观察指标，并前瞻性地详细记录临床纵向观测数据，通过统计分析最终得出结论的临床研究类型。

在前瞻性研究中，对患者评估是基于预先的研究方案设计。方案设计决定了收集数据的标准。前瞻性研究可解决回顾性研究的诸多局限，如：

- 描述医疗情况和治疗方式时，使用标准、统一、最新的定义。收集这些数据时，采用统一的、事先在研究方案中明确医学术语。

- 制定持续、有计划的医学随访时间表，包括诊室随访和远程随访等。

- 明确标准操作规程或统一的评估标准，以及尽量采用最新的诊断标准，或及时跟进诊疗现状的变化，确保信息收集的一致性（例如，使用相同的临床结局评价方法、采用可比性的指标）。

- 收集可能与疾病发生发展有关的数据，以及患者的伴随治疗信息。

前瞻性疾病自然史研究的不足之处主要在于通常需要更长的研究时间和研究准备时间。为了缓解这类挑战，建议

利用真实世界证据和/或最大限度地使用结构化的电子病历数据，并鼓励工业界、学术界和患者组织的合作，如果可能，以患者负担最小但科学价值最大的方式生成疾病自然史数据。

### （三）双向队列研究

双向队列研究又称回顾前瞻性队列研究，是指根据前瞻性队列研究进行设计，同时收集回顾性队列数据所开展的研究。该设计的主要优点是可以充分利用已有的队列数据，以及补足回顾队列中未达到所设定的最短观察期的数据，并与前瞻设计的队列数据合并，形成样本量更大、观察期更长、纵向数据更完整的观察性研究数据。

双向队列研究也需要回顾现有的医疗记录、已发表的科学文献、疾病诊疗专家的意见和经验，以及其他来源的信息获取疾病信息。通过回顾已有的资料，结合规范的前瞻性跟踪观察，可以收集疾病在不同时间点的临床特征数据，记录患者在不同时间点的病情和治疗情况。因此，双向队列研究可以模拟和预测疾病的发展进程。

前瞻性研究和回顾性的结合可以优势互补，双向队列研究兼具前瞻性研究和回顾性研究的优点，如：

- 有利于用更快速、高效地对疾病，尤其是一些罕见疾病有初步地认知，而缺失的信息可以进行前瞻性观察。

- 充分收集可能与疾病发生发展有关的数据，模拟预测

疾病进程，找到最优的治疗方案，以做出更精准的治疗决策。

- 统一收集数据、操作规程和评估标准，确保信息收集、诊断和治疗的一致性。

当然，双向队列研究对于过去资料的留存也有一定要求。对回顾性收集的信息的准确性需客观鉴别，也需要与前瞻性跟踪得到的信息尽可能统一标准。

#### （四）横断面研究

横断面研究是对特定时间点和特定范围内人群疾病或健康状况及有关因素的分布状况的一种研究。横断面研究所展现的数据是有关患者疾病自然史全病程的纵向数据中一个时间点或较短时期内的横向数据。

横断面研究对罕见疾病药物研发的价值主要体现在以下方面：

- 可以根据同一时间所测定的、处于不同疾病进展阶段的患者情况，从而推测疾病进展的一般过程。

- 可以描述当疾病快速进展时，治疗可能带来的即刻疗效情况。

横断面研究为研究现有患者群体情况提供了一种快速有效的方法，但在较短时期内收集的数据可能无法充分表征疾病的发展过程。例如，一些快速进展的疾病亚型，可能会因为病程偏倚，而无法对其发生发展特点进行充分表征。

虽然横断面数据不适合作为临床试验中的外部对照，但

可为后续研究设计提供依据。

上述几种研究类型各自存在有局限性和实施上的挑战，出于实效性考虑，可以考虑整合不同来源数据，如将可比的横断面研究数据加入已存在的纵向研究数据，或将回顾性研究数据整合入前瞻性研究等。

## 五、疾病自然史研究的实施

### （一）及早规划和实施

疾病自然史研究的起点是收集、组织和分析所有当前可用数据；应及早规划和实施疾病自然史研究。

疾病自然史研究可以由药物研发企业发起，也可以由研究者或者患者组织等发起。对于一些罕见疾病，早期开始疾病自然史研究（甚至在试验性治疗药物被开发前即开始），将有助于通过更长时间、在更广泛患者群体中采集数据，获得对疾病更深入的认识。

不应该为了等待疾病自然史研究的结果而延缓药物开发，因为疾病自然史研究经常面临一系列未知因素的影响。但仍需重视疾病自然史对于罕见疾病药物开发的重要价值。来自疾病自然史研究的数据可作为药物注册研究的有益补充。

建议研发单位及早与药物监管部门进行研究目的和数据源可行性分析，对数据源的选择进行合理的论证，避免引入选择偏倚和发表偏倚。

### （二）鼓励患者群体的参与

在疾病自然史研究的实施过程中，鼓励寻求特定疾病患者组织的协助，鼓励患者群体的参与。计划周密的罕见疾病自然史研究有助于确定未来加入试验的患者群体，加快后续招募。患者持续参与研究可确保随访数据的稳健性，并可以提供来自患者视角的观点（例如，如何减少患者和家庭的负担、提升研究可接受性等）。此外，对退出研究或选择不完全参与的研究对象进行访谈，也将有助于减少数据缺失，提高研究质量。得益于独特视角，患者的参与还有助于研究的设计和执行<sup>[6]</sup>。

### （三）数据采集

对于罕见疾病，寻找患者加入疾病自然史研究可能存在一定挑战，因此通常需要多地区、多研究中心的参与。鼓励在多地区多人群中开展疾病自然史研究，特别是对于计划开展国际多中心研究的药物，通过对多地区人群的疾病自然史研究，将有助于评估不同地区间潜在差异的影响。在启动多中心研究时，鼓励采取措施加强各研究中心之间的一致性。

疾病自然史研究数据可以通过不同方式获取，包括一些新兴的数据采集方式，如用于采集临床评估结局（**clinical outcome assessment, COAs**）的数字监测设备（可与其他数据采集方法结合使用）。数据可以来自于医生或医疗机构，也可以采集自患者本人或其家庭。提升便利性可以促进患者参与研究。应在研究中关注参与者的便利性，尽量减少额外负

担。因此对于地理位置分散、偏远以及存在其他不便利因素（如疾病晚期的患者）的患者，可以考虑运用远程方式采集样本、收集数据和/或医疗判断来提升罕见疾病自然史研究中患者人群的代表性和多样性。

当需要采集实验室检测数据（包括影像学等）时，每个实验室均应具有相应的检测资质。为了降低变异性，不同实验室间应进行一致性比较，也可以考虑设置独立的中心实验室分析样本。

用于中国患者的药物，在其开发过程中，如计划开展自然病史研究，通常情况下建议在中国患者中开展，以满足了解中国患者自然病史特征的需求；是否可参考境外患者的自然病史研究数据，将取决于具体需要解决的临床和/或监管问题。

#### （四）研究对象保护

疾病自然史研究应该有相应措施充分保护患者隐私并确保数据的保密性。对于数据的应用目的、使用范围、使用数据的相关方、研究对象权利等，应向研究对象或家属（监护人）予以充分告知并获得知情同意或获得泛知情同意或豁免（如回顾性研究），研究方案应获得伦理委员会批准，相应数据的收集与使用应符合相关法律法规的要求。

## 六、参考文献

[1]. US. Food and Drug Administration. Rare diseases: Natural

History Studies for Drug Development (Draft Guidance for Industry). <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/rare-diseases-natural-history-studies-drug-development>.

[2]. Biomarkers Definitions Working Group. Bethesda, Md. Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. *Clinical Pharmacology and Therapeutics*, 2001,69: 89-95.

[3]. 国家药监局药审中心关于发布《用于产生真实世界证据的真实世界数据指导原则（试行）》的通告（2021年第27号）。

<https://www.cde.org.cn/main/news/viewInfoCommon/2a1c437ed54e7b838a7e86f4ac21c539>.

[4]. 关于《药物真实世界研究设计与方案框架指导原则（征求意见稿）》公开征求意见的通知。

<https://www.cde.org.cn/main/news/viewInfoCommon/ea778658adc3d1ae3ffe3f1cc0522e5e>.

[5]. Delgado-Rodriguez M and Llorca J, 2004, Bias, *J Epidemiol Community Health* 58(8):635–641.

[6]. 关于公开征求《组织患者参与药物研发的一般考虑指导原则（征求意见稿）》意见的通知。

<https://www.cde.org.cn/main/news/viewInfoCommon/2981a587>

ba89aa3368130ac1565fb04f.

[7] EMA Guideline on clinical trials for small populations (2006) .

<https://www.ema.europa.eu/en/documents/scientific->

[guideline/guideline-clinical-trials-small-populations\\_en.pdf](https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-clinical-trials-small-populations_en.pdf)