

溃疡性结肠炎治疗药物临床试验技术指导原则

（征求意见稿）

2021 年 6 月

目 录

一、适用范围	1
二、概述	1
(一) 适应症特点	1
(二) 临床治疗学现状及需求	2
三、总体考虑	2
(一) 受试者	2
(二) 疗效指标评价	3
(三) 方法学考虑	4
四、临床药理学研究	5
五、探索性临床试验	6
六、确证性临床试验	6
(一) 总体设计	6
(二) 对照组	6
(三) 受试者	7
(四) 给药剂量和治疗持续时间	7
(五) 基础治疗	8
(六) 补救治疗	8
(七) 疗效指标	8
1. 主要疗效指标	8

2. 次要疗效指标	9
(八) 统计分析考虑	10
七、安全性评价	11
八、儿童溃疡性结肠炎临床试验	12
九、参考文献	14

溃疡性结肠炎治疗药物临床试验技术指导原则

一、适用范围

本指导原则旨在为溃疡性结肠炎（Ulcerative Colitis, UC）治疗药物的研发提供技术指导。

本指导原则适用于化学药品和治疗用生物制品的药物研发，仅作为推荐性建议。应用本指导原则时，还应同时参考药物临床试验质量管理规范（GCP）、国际人用药品注册技术协调会（ICH）和其他境内外已发布的相关指导原则。

本指导原则仅代表药品监管部门当前的观点和认识，不具有强制性的法律约束力。随着科学研究的进展，本指导原则中的相关内容将不断完善与更新。

二、概述

本指导原则主要讨论 UC 治疗药物研发中临床试验设计的重点关注内容。关于临床试验设计或统计学分析的一般性问题可参考相关指导原则。

（一）适应症特点

UC 是一种病因尚不十分清楚的慢性非特异性肠道炎症性疾病，最常发生于青壮年期。临床表现为持续或反复发作的腹泻、黏液脓血便伴腹痛、里急后重和不同程度的全身症状。可有皮肤、黏膜、关节、眼、肝胆等肠外表现。并发症包括中毒性巨结肠、肠穿孔、下消化道大出血、上皮内瘤变以及癌变。

我国 UC 缺少确切的发病率数据，基于区域性的流行病学调查提示我国的 UC 发病率为 0.42-2.22/10 万。虽然与西方国家相比发病率仍然较低，然而与 20 年前相比，呈明显上升趋势。

（二）临床治疗学现状及需求

UC 的治疗目标在于诱导并维持缓解以及黏膜愈合，防治并发症，改善生命质量。

主要根据疾病活动严重程度、病变累及范围、治疗反应等选择治疗方案。

传统治疗药物主要包括氨基水杨酸制剂、糖皮质激素、免疫抑制剂等。对于传统治疗无效、不能耐受以及糖皮质激素依赖的中重度活动性 UC 患者，可选择肿瘤坏死因子 α (TNF α) 拮抗剂。TNF α 拮抗剂存在严重感染和恶性肿瘤等风险。对传统治疗或 TNF α 拮抗剂应答不足、失应答或无法耐受的中重度活动性 UC 患者，可选择 $\alpha 4 \beta 7$ 整合素拮抗剂。

三、总体考虑

药物研发临床试验的设计基于临床试验目的而定。

（一）受试者

活动期 UC：根据临床症状及入组前 1 个月内的内镜检查和组织学检查结果，提供疾病活动性的证据。病变范围推荐采用蒙特利尔分型确定：E1：局限于直肠，未达乙状结肠；E2：累及左半结肠（脾曲以远）；E3：广泛病变累及脾曲以近乃至全结肠。严重程度分型常用 Mayo 评分确定，包括轻、中、重度。对内镜和组织学检查结果的评价，应建立并严格遵循 SOP（Standard Operation Procedure），推荐采用中心阅片。

如果目标人群是经过特定单药或多药联合充分治疗后仍被确定存在活动性炎症的患者，应预先明确“充分治疗”的定义，包括剂量和治疗持续时间。在入组时应全面评估患者诊断、用药依从性和治疗充分性。

如果目标人群是对某种治疗不耐受的患者，应预先明确“不耐受”的定义，包括导致患者无法继续治疗的不良反应类型，如：硫嘌呤类药物引起骨髓抑制。

如果目标人群是对某种治疗依赖的患者，应预先明确“依赖”的定义。对于糖皮质激素依赖，可以接受以下定义：①虽能维持缓解，但激素治疗 3 个月后，泼尼松仍不能减量至 10mg/d；②在停用激素 3 个月内复发。

缓解期 UC：如果不存在或存在非常轻微的症状和体征，且内镜证实无肉眼可见活动性炎症，可将患者归类为缓解期。确切的定义取决于评价黏膜炎症和临床症状的工具。

（二）疗效指标评价

UC 治疗药物临床试验中，主要评价临床症状和黏膜炎症两个方面。

对于临床症状的评价，推荐采用患者报告结局（Patient Reported Outcome, PRO）量表，该量表应包括 UC 的重要症状，如排便次数和便血。目前可接受 Mayo 评分中排便次数（Stool Frequency, SF）和便血（Rectal Bleeding, RB）单项评分，鼓励申办者进一步研发和验证 PRO 量表。临床症状评价指标包括症状缓解、症状应答等。可以接受以下症状缓解定义：RB 和 SF 单项评分均为 0；或 RB 单项评分为 0，SF 单项评分 ≤ 1 且较基线降低至少 1 分。

对于黏膜炎症的评价，目前主要采用内镜检查评估量表，包括：Mayo 评分中内镜单项评分、UC 内镜严重程度指数（UC Endoscopic Index of Severity, UCEIS）。黏膜炎症内镜评价指标包括内镜缓解、内镜应答等。可以接受以下内镜缓解定义：UCEIS

≤1 分；或 Mayo 评分中内镜单项评分为 0 或 1 分。为减少偏倚，推荐采用中心阅片，应由两名及以上经过培训的消化内镜专家进行盲态下评价，图像存储备查。

单独内镜检查仅能提供对黏膜视觉外观的评估。鼓励进行黏膜炎症的组织学评价，以了解缓解的深度。目前常用的组织学组织病理学评分系统包括 Nancy 指数、Geboes 评分和 Robarts 组织病理学指数。为保证组织学样本的处理质量，要求严格遵循病理样本 SOP。为减少组织病理学评价的差异，病理读片应采用中心阅片，建议由两名及以上经过培训的病理专家进行双盲读片。

其他有效性评价还包括炎症生物标志物、健康相关生命质量评价等。

另外，UC 病程改变试验的最终治疗目标包括预防疾病对患者生命质量（健康相关生命质量等）、中期并发症（肠道损伤、UC 相关手术和住院等）和长期并发症（胃肠和肠外异型增生或恶性肿瘤、死亡）的影响，因此，在整体研发计划中，还需要考虑评价药物对相关终点指标的影响。

（三）方法学考虑

既接受传统的临床研发设计，也接受适应性设计等新颖设计。如采用新颖设计，建议与药品监管部门事先沟通。

在多区域临床试验中，应关注临床药理和临床实践方面存在的差异可能对种族敏感性的影响。建议在早期阶段加入全球研发，以保证受试者能够充分代表中国人群。

应明确定义试验旨在解决的科学问题，确定估计目标。试验计划、设计、实施、分析和解释必须与估计目标相一致，参见 ICH E9 (R1)。主要估计目标应反映内镜和症状缓解实现（诱导缓解）

或维持（维持缓解）的治疗效应。在估计目标的定义中，需要考虑伴发事件的影响，伴发事件包括停止治疗（对试验药物不耐受、缺乏疗效或疾病进展等）和其他药物使用的变化（如：采用补救治疗、改变基础治疗、糖皮质激素未能按照规定方案减量或使用禁用药物等）。

处理伴发事件“停止治疗”的最合适策略取决于治疗目的。若治疗目的是在短期内诱导缓解，则主要关注对内镜和症状缓解的治疗效应，此时建议采用疗法策略。若治疗目的是维持缓解或证明长期治疗的效果，则建议采用复合变量策略，将停止治疗的受试者定义为治疗失败，因为 UC 是一种需要长期治疗的疾病。

若采用了补救治疗，尤其是糖皮质激素（不论试验方案是否预先规定），应视为治疗失败，建议采用复合变量策略。将逐渐减量的糖皮质激素作为基础治疗时，若未能按照方案规定逐渐减量，则应视为补救治疗的一种特殊情况。

四、临床药理学研究

通常，临床药理学研究包括人体耐受性试验、药代动力学（pharmacokinetics, PK）、药效学（pharmacodynamics, PD）和 PK/PD 试验等。

根据药物特点（如：局部作用药物、蛋白质和单克隆抗体等），参考相关指导原则，充分研究药物的 PK 特征，关注特殊人群（如：肝肾功能损害、老年人、儿童等），为后续临床试验中合理的给药方案及剂量调整等提供支持。

早期研发阶段应针对药物处置过程关键环节，如：关键代谢酶或转运体的底物、诱导剂或抑制剂等，参考相关指导原则开展药物相互作用研究。基于群体 PK 分析也有助于描述已知或新确

定的相互作用的临床影响，并提供剂量调整的建议。UC 患者可能合并使用糖皮质激素、免疫抑制剂等，应关注这些常见合并用药对临床安全性和有效性的影响。

五、探索性临床试验

建议采用随机、双盲、安慰剂对照设计。

可基于前期 PK/PD 评估结果设置多个剂量组，充分评价药物的量效关系，为后续给药方案的选择提供依据。

研究时间取决于药物的作用机制、安全性和对于所选疗效指标的预期作用，一般不应短于 6~8 周。

六、确证性临床试验

（一）总体设计

随机、双盲、安慰剂或阳性对照设计。随机分组时可以考虑采用分层随机化的方法，如：考虑既往治疗、病变范围等因素。

可以分别开展诱导研究和维持研究。

也可以在一项研究中同时评价诱导缓解治疗和维持缓解治疗的有效性，如采用随机撤药设计，或持续治疗设计。

随机撤药设计是指将诱导缓解研究中达到应答或缓解的受试者再次随机分配至安慰剂组或试验药组，这种策略通常需要招募更多受试者，因为研究必须具有进行最终评估的充分把握度。

持续治疗设计是指在诱导缓解研究基础上继续进行维持缓解研究，保持设盲和随机继续进行维持缓解研究。如果采用持续治疗设计，对于无应答的受试者，应慎重评估继续治疗的获益和风险，设置合理的退出机制，以保证受试者利益。

（二）对照组

对照的选择应有充分依据，并符合伦理。对于活动性 UC 治

疗的一线适应症，需要与标准一线治疗进行直接比较。对于二线适应症（在标准一线治疗失败或不耐受之后），可选择安慰剂或阳性对照。

在采用阳性对照时，应考虑该阳性药对于目标适应症是否具有充分的临床有效性证据。阳性对照研究的目的既可以是证明研究药物与阳性药相比的优效性，也可以是证明其非劣效性。非劣效界值应遵循相关指导原则预先确定。不同临床试验入选人群基线严重程度、病变范围以及评价有效的终点指标可能不同，在确定非劣效界值时需要关注。

（三）受试者

根据适应症定位制定具体的入排标准，尽可能接近预期的目标治疗人群。关注疾病活动性的严重程度、病变范围、既往治疗的疗效等。

入组活动性疾病治疗试验的受试者应该具有活动性疾病的证据。同时规定症状和黏膜炎症的最低水平。如：Mayo 评分 6~12 分，且内镜单项评分 ≥ 2 分，或 UCEIS 评分 ≥ 5 分。

应排除克罗恩病、肠结核、其他慢性肠道感染性疾病、肠道恶性肿瘤等。

如果目标人群同时包括既往传统治疗失败和生物制剂治疗失败的中重度活动性 UC，建议预先规定不同人群在总人群中所占比例。

（四）给药剂量和治疗持续时间

给药方案应有前期临床试验数据等科学支持，基于量效关系和整体获益风险评估合理选择。

为证明诱导缓解治疗的有效性，研究的持续时间通常为 6~

12 周，但取决于试验药物的作用机制、预期作用程度和时间过程。

为证明有效性的维持，研究总持续时间（包括诱导阶段）应至少为 12 个月。

（五）基础治疗

对于入组时正在接受 UC 治疗的受试者，一般允许在试验期间以基础治疗的方式继续使用这些治疗，并维持稳定剂量（糖皮质激素除外）。对于入组时正在使用糖皮质激素的受试者，在试验的早期阶段应持续使用固定剂量的糖皮质激素，之后按照规定方案逐渐减量。

应预先规定入组前基础治疗的剂量和持续时间，以确保受试者在随机分组前接受了充分治疗。对于硫嘌呤类药物，需要接受稳定剂量至少 3 个月。

如果不允许使用基础治疗，应根据药物的 PK 和 PD 特点等，规定足够的洗脱期。

在广泛病变中，合并使用局部用药可能影响结肠镜检查的结果，因此，如果主要目的是评估全身作用药物的有效性，建议停用局部用药。

（六）补救治疗

针对可能发生的疾病恶化风险，在方案中事先规定补救治疗的标准和补救治疗药物，以确保临床试验受试者安全。

（七）疗效指标

1. 主要疗效指标

主要终点应包括症状缓解和内镜缓解两个方面。

可采用以下单一主要终点：达到临床缓解（即同时达到症状缓解和内镜缓解）的受试者比例。

也可采用以下共同主要终点：达到症状缓解的受试者比例，和达到内镜缓解的受试者比例。在共同主要终点分析中症状缓解和内镜缓解均应达到统计学意义。

在一项研究中同时评价诱导缓解治疗和维持缓解治疗的有效性时，通常希望将两个时间点都作为主要分析。由于内镜缓解所需时间通常长于症状缓解所需时间，因此评估内镜终点的时间点可能与临床症状终点的时间点不同。

举例 1：一项在中重度活动性 UC 受试者中评估某药物有效性和安全性的随机、双盲、安慰剂对照、III 期临床试验，采用随机撤药设计，包括 12 周诱导阶段和 40 周维持阶段。受试者随机分配接受试验药或安慰剂治疗 12 周，第 12 周达到临床应答的受试者将被重新随机分配，接受试验药或安慰剂治疗 40 周。诱导阶段主要终点是第 12 周达到临床缓解的受试者比例，维持阶段主要终点是第 52 周达到临床缓解的受试者比例。

举例 2：一项在中重度活动性 UC 受试者中评估某药物有效性和安全性的随机、双盲、安慰剂和阳性对照、III 期临床试验，采用持续治疗设计。受试者随机分配接受试验药或安慰剂治疗，第 12 周达到临床应答的受试者继续进行双盲治疗，直至第 52 周。诱导阶段共同主要终点是第 12 周达到症状缓解的受试者比例，和第 12 周达到内镜缓解的受试者比例。维持阶段共同主要终点是第 52 周达到症状缓解的受试者比例，和第 52 周达到内镜缓解的受试者比例。

如果选择其他主要终点，应有充分的科学依据，并事先与药品监管部门进行沟通。

2.次要疗效指标

- 达到内镜缓解的受试者比例（适用于使用单一主要终点

时)。

- 达到症状缓解的受试者比例（适用于使用单一主要终点时）。

- 同时达到内镜缓解和症状缓解的受试者比例（适用于使用共同主要终点时）。

- 达到内镜单项评分为 0 分的受试者比例。

- 对于在基线时正在使用糖皮质激素的受试者，达到撤离激素的缓解的受试者比例。通常定义为在特定时间点达到缓解，且在此前一定时间窗内无激素暴露（如：第 52 周达到缓解，且在第 40 到第 52 周期间未使用激素）。

- 临床症状和内镜检查的单项评分或总评分较基线的变化。

- 黏膜炎症的组织学评价，如：达到组织学正常的受试者比例。

- 炎症生物标志物评价，如：粪便钙卫蛋白正常的受试者比例等。

- 达到缓解或应答的时间（仅症状评分和生物标志物）。

- 维持缓解的受试者比例。定义为在入组时（或诱导研究结束时）达到缓解的受试者在至 52 周研究结束时持续处于缓解状态。

- 健康相关生命质量评价，使用经验证的生命质量评估量表，如：炎症性肠病调问卷（IBDQ）相对基线的变化等。

- UC 相关外科手术或住院的受试者比例。

（八）统计分析考虑

统计分析(包括缺失数据的处理)应与既定的估计目标对应，且必须在研究方案中预先定义并阐明其合理性。

收集与既定估计目标相关的必要数据至关重要。对于主要估

计目标，若采用复合变量策略将特定伴发事件定义为治疗失败，分析时无需纳入伴发事件之后的临床结局变量数据。但这些数据可能会被用于估计反映其他科学问题的估计目标，例如：采用连续型变量的临床症状指标来估计治疗效应时，可使用疗法策略来反映额外药物的使用情况。

区分伴发事件和缺失数据也很重要。特别是对于一年内进行多达 3 次内镜检查的试验，拒绝接受反复内镜检查可能会导致数据缺失，但其并非是伴发事件，且根据估计目标，需要在统计分析中处理这些受试者的缺失数据。

在一项研究中同时评价诱导缓解和维持缓解的有效性时，将两者作为共同主要终点，不涉及多重性问题。若在一项研究中评价诱导缓解或维持缓解的有效性时，则可能会产生多重性问题，应制定相应多重性调整策略。

七、安全性评价

原则上遵循安全性评价的共性标准。

UC 通常需要长期治疗，因此需要有足够的暴露量和暴露时间进行安全性观察，参考 ICH E1。建议在长期试验中设立独立的科学委员会。

UC 临床试验中应特别关注以下安全性评价内容：

1. UC 治疗药物通常具有免疫调节剂的作用，应特别关注严重感染、自身免疫疾病和肿瘤等风险的可能性。

2. 对于生物制剂，关注治疗期间是否产生抗体，以及抗体是否影响长期有效性和安全性。合并使用免疫抑制剂可能增加机会性感染和恶性肿瘤等严重不良反应的风险，并降低检测免疫原性的能力，在评价安全性时需要特别关注合并用药情况。

八、儿童溃疡性结肠炎临床试验

在儿童患者中，UC 的总体疾病病理、进展和治疗目标与成人相似。然而，儿童炎症性肠病（Inflammatory Bowel Disease, IBD）特点是病程更复杂，炎症反应可能更高，对糖皮质激素和免疫抑制剂治疗需求也更高。与成人 IBD 相比，儿童 IBD 的癌变风险更高、疾病持续时间更长、疾病严重或扩展程度更为严重。此外，儿童 UC 患者生长发育迟缓和骨量减少的风险增加。

1. 总体考虑

临床研发项目应尽量纳入 2 岁及以上的儿科受试者，除非存在明显安全性隐患（如：幼龄动物或成人发生重大不良事件，或其他免疫缺陷），提示需要排除某些年龄组，或有证据表明某些年龄组不太可能获益。

当有足够的证据支持研究药物对儿科 UC 患者的安全性和预期获益后，应尽早开展儿科研究。

所有入组的儿科受试者都应具备明确的 UC 诊断，并经内镜和组织学检查结果确认活动性炎症。6 岁以下儿童的疾病表型可能不同于青少年或成人，需进行原发性免疫缺陷病的基因检测，并根据检测结果入选或排除这些患者。

2. 临床药理学和剂量探索研究

在选择各年龄或体重组的起始剂量和最终剂量时，应考虑成人和/或儿童中所有 PK、PD 或其他数据。可在所有数据的基础上应用群体 PK 建模，因为这种方法可进行广泛的协变量分析，量化体重、年龄和其他协变量的影响。在此之前，需要获得所有年龄组的 PK 数据。若研究目的为达到某个相同的暴露量（如：成人暴露量），则可采用这些协变量分析的结果，以确定是否需

要在各年龄组采用不同剂量。

3. 有效性研究

儿童 UC 的治疗目标为诱导并维持临床缓解及黏膜愈合，促进生长发育，改善患儿生命质量。

推荐的主要疗效指标与成人相似，包括内镜和症状缓解两个方面。对于临床症状的评价，可采用儿童溃疡性结肠炎活动指数（pediatric UC activity index, PUCAI），PUCAI<10 视作症状缓解，PUCAI 较基线下降 ≥ 20 分视作症状应答。对于黏膜炎症的评价，目前主要采用成人的内镜检查评估量表。

次要疗效指标可参考成人，但需包括：生长参数（包括身高和体重）和营养状况改善、肠外表现方面的疗效等。有关性成熟的观察时间将取决于受试者入组时的年龄。

4. 安全性研究

长期安全性评估需要在更长时间内收集更多患者数据，数据收集可能会持续至成年期，可在开放性扩展研究和上市后观察性研究中评估。

如果担忧药物对免疫系统有影响，且通过临床前研究或成人临床试验无法解决，应开展相应的儿童临床试验（如：对免疫系统、疫苗接种的影响等）。

5. 数据外推

基于成人和儿童在整体疾病病理和进展以及治疗目标等方面的某些相似性，可参考相关指导原则，通过数据外推，对儿童用药的有效性和安全性进行评价。如果考虑采用数据外推，建议与药品监管部门沟通。

九、参考文献

[1] EMA, Guideline on the development of new medicinal products for the treatment of Ulcerative Colitis, 28 June 2018.

[2] FDA, Ulcerative Colitis: Clinical Trial Endpoints Guidance for Industry (DRAFT GUIDANCE), August 2016.

[3]中华医学会消化病学分会炎症性肠病学组.炎症性肠病诊断与治疗的共识意见(2018年•北京).中华消化杂志, 2018,38(5): 292-311.

[4] 中华医学会儿科学分会消化学组, 中华医学会儿科学分会临床营养学组. 儿童炎症性肠病诊断和治疗专家共识. 中华儿科杂志, 2019,7(57): 501-507.

[5]中华医学会消化病学分会炎症性肠病学组.中国消化内镜技术诊断与治疗炎症性肠病的专家指导意见. 中华炎症性肠病杂志, 2020, 4(4): 283-291.

[6]NMPA.药物临床试验的一般考虑指导原则.2017年1月

[7] Turner D, Ricciuto A, Lewis A, et al. STRIDE-II: an update on the Selecting Therapeutic Targets in Inflammatory Bowel Disease (STRIDE) initiative of the International Organization for the Study of IBD (IOIBD): determining therapeutic goals for treat-to-target strategies in IBD [J]. Gastroenterology, 2021;160(5):1570-1583.

[8]Berre C L, Peyrin-Biroulet L, Sandborn W J, et al. Selecting Endpoints for Disease-Modification Trials in Inflammatory Bowel Disease: the SPIRIT consensus from the IOIBD[J]. Gastroenterology, 2021;160(5):1452-1460.

[9]Ma C, Panaccione R, Fedorak R N., et al. Heterogeneity in

definitions of endpoints for clinical trials of ulcerative colitis: a systematic review for development of core outcome set[J]. Clinical Gastroenterology and Hepatology, 2018;16 (9) :637-647.

[10]Xue Li, Peige Song, Jun Li, et al. The disease burden and clinical characteristics of inflammatory bowel disease in the Chinese population: a systematic review and meta-analysis [J]. Int J Environ Res Public Health, 2017; 14 (3): 238.