

阿尔茨海默病药物临床试验技术指导原则（试行）

国家药品监督管理局药品审评中心

2025 年 5 月

目 录

一、概述	1
(一) 疾病特点及治疗现状	1
(二) 目的及适用范围	2
二、总体考虑	2
(一) 整体研发策略	2
(二) 研发目标	3
三、关键技术考虑	4
(一) 研究人群	4
(二) 生物标志物的作用和类型	7
(三) 有效性评估工具	9
四、临床药理学研究	14
(一) 药代动力学	14
(二) 药效学	14
(三) 药物相互作用	15
五、探索性临床试验	15
六、确证性临床试验	16
(一) 不同研发目标药物的试验设计	16
(二) 有效性指标	23
七、统计学考量	28
八、安全性评估	30
九、与监管机构沟通	32
十、参考文献	33

阿尔茨海默病药物临床试验技术指导原则（试行）

一、概述

（一）疾病特点及治疗现状

阿尔茨海默病（Alzheimer's disease, AD）是一种中枢神经系统退行性疾病，以渐进性记忆障碍及认知功能减退伴日常生活能力下降和行为改变为特征，是全球需应对的公共卫生领域重大疾病之一。根据国际阿尔茨海默病协会发布的《世界阿尔茨海默病 2021 年报告》，全世界约有 4100 万阿尔茨海默病患者。有文献报道我国 60 岁及以上老年人群中约有 983 万阿尔茨海默病痴呆患者。随着人口老龄化进程加速，阿尔茨海默病发病率和患病率将逐年增加。

阿尔茨海默病的发病机理尚未完全明确，相关的假说包括： β -淀粉样蛋白（ β -amyloid, $A\beta$ ）级联假说、tau 蛋白过度磷酸化假说、脑血管异常假说、神经炎症假说和氧化应激假说等，其中 $A\beta$ 级联假说在药物研发中一直处于较为重要的地位。阿尔茨海默病的主要病理特征为 $A\beta$ 沉积形成淀粉样斑块、tau 蛋白异常磷酸化形成的神经原纤维缠结以及神经元缺失和胶质细胞增生等。

传统的阿尔茨海默病治疗药物主要为症状治疗药物，包括胆碱酯酶抑制剂和谷氨酸受体拮抗剂等。近年来批准上市了针对 β -淀粉样蛋白等新靶点的药物，但临床仍存在迫切需

求。

（二）目的及适用范围

本指导原则旨在为阿尔茨海默病不同临床分期、不同研发目标的药物临床试验提供技术建议，适用于化学药品和治疗用生物制品的药物研发。

本指导原则主要讨论临床试验设计所需关注的重点内容。应用本指导原则时，还需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）和其他国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、总体考虑

（一）整体研发策略

阿尔茨海默病是一个连续的疾病谱已成为共识，大致可分为三期，包括临床前 AD/认知未受损 AD、前驱期 AD/AD 源性轻度认知障碍（mild cognitive impairment, MCI）、AD 痴呆三个分期。既往的阿尔茨海默病治疗药物临床试验主要针对已表现出典型痴呆症状的时期，目前认为阿尔茨海默病具有漫长的症状前时期，在症状发生前 10-20 年就可出现阿尔茨海默病特征性的神经病理学改变，故可能有机会在阿尔茨海默病疾病过程的早期进行干预。

基于不同的阿尔茨海默病临床分期、预期的治疗效应和产品开发目标，有效性和安全性评价技术要求有所区别，因此，通常应结合药物的作用机制特点和临床定位制定合理的研发策略。

鼓励在早期研发阶段基于药物特点和研发目标开展必要的患者体验数据研究，并基于获得的患者体验数据完善临床研发计划及试验方案。

此外，临床研发策略还需要考虑是否作为单药治疗进行开发，是否计划与目前已批准的标准治疗药物（如胆碱酯酶抑制剂、美金刚、抗 $A\beta$ 抗体等）联合使用，或者是否计划与靶向相似或不同的阿尔茨海默病病理生理机制的新药联合使用等。

（二）研发目标

不同作用机制的阿尔茨海默病药物，在临床研发计划和临床试验设计中可能有不同的目标，其中主要的研发目标包括：

1.疾病预防：在临床前 AD 期通过干预可能的致病机制来预防疾病症状发生。

2.疾病修饰治疗：具有延缓或阻止临床症状进展的作用，且具有改变神经病理学进程的证据。

3.症状治疗：不改变疾病总体病程的治疗效应，可包括增强认知和功能改善。

4.痴呆的精神行为症状（behavioural and psychiatric symptoms of dementia, BPSD）的对症治疗。

三、关键技术考虑

（一）研究人群

1.常染色体显性阿尔茨海默病

常染色体显性阿尔茨海默病是由几种已知的致病基因相关突变（如 PSEN1、PSEN2、APP 等）导致的，所占比例不足 1%。影响常染色体显性阿尔茨海默病患者症状发生和疾病进展的因素尚不完全了解。常染色体显性阿尔茨海默病是开发新药以及验证评价工具的重要模型，但该类型阿尔茨海默病的病理生理学特征及其对治疗的反应与散发性阿尔茨海默病的相似程度仍有待确定。

2.散发性阿尔茨海默病

常见的阿尔茨海默病类型为散发性阿尔茨海默病，约占 99%。散发性阿尔茨海默病是一种高度复杂的多因素疾病，其发病机理和病理生理学过程尚未完全明确。已知阿尔茨海默病的病理过程在临床症状出现前 10-20 年开始，但该神经病理变化过程与症状进展和特定的测量结局之间的关系有待进一步研究确定。

需要具有高敏感性和特异性的诊断标准来选择同质化的研究人群。选择研究人群时通常需考虑临床症状和神经病理生理生物标志物状态等不同的维度。随着阿尔茨海默病领

域研究的发展，对于痴呆前的临床分期可能会更加细化，同时更加重视生物标志物的应用，以便在早期进行干预。国际上已经制定了多套阿尔茨海默病诊断标准，其中在药物临床研发中广泛应用的主要包括美国国家老龄化研究所-阿尔茨海默病协会（NIA-AA）标准和国际工作组（IWG）标准，两者均可用于阿尔茨海默病药物临床试验的诊断标准和研究人群选择。在目前阶段，NIA-AA 和 IWG 标准仍未得到充分验证并仍在不断完善。

通常需根据阿尔茨海默病的病理生理特点、试验药物的作用机制和研发目标，考虑研究人群的适宜临床分期。基于试验药物的临床定位，可在同一个临床研究中纳入不同临床分期的目标受试者，例如前驱期 AD/AD 源性 MCI 和轻度 AD 痴呆患者。

美国阿尔茨海默病协会工作组 2024 年发布的诊断和分期标准中包括两种类型的临床分期：综合征分类性认知分期和数字临床分期。其中，综合征分类性认知分期包括认知未受损、轻度认知障碍和痴呆，其中痴呆进一步细分为轻度、中度和重度。数字临床分期分为 7 期，0 期具有相关的基因突变，但无阿尔茨海默病特征性病理生理变化和临床症状。1~3 期代表痴呆前期，4~6 期代表痴呆期。综合征分类性认知分期和数字临床分期，两种分期系统可解决药物临床研发中的不同需求，两者的分期划分有一定的重叠交叉。数字临

床分期 1 期和 2 期大致对应认知未受损期；数字临床分期 3 期大致对应轻度认知障碍期，但轻度认知障碍期也适用于数字临床分期 2 期中的某些个体；数字临床分期 4、5 和 6 期分别大致对应轻度、中度和重度痴呆期。

此外，采用神经心理学测评可以进一步区分认知功能障碍患者的疾病严重程度，有利于针对性地纳入受试者。常用的评估量表包括：临床痴呆评定量表（clinical dementia rating, CDR）、简易精神状态检查（mini mental state examination, MMSE）和总体衰退量表（global deteriorate scale, GDS）等。对于早期的阿尔茨海默病患者，可能需使用更为敏感的神经心理学测试量表，但需要经过充分的验证，并提供支持性证据。

3. 混合型痴呆和混合型阿尔茨海默病

混合型阿尔茨海默病通常具备共存疾病的临床表现和病理学证据。据文献报道，尸检时混合型阿尔茨海默病至少占有所有阿尔茨海默病病例的 50%。根据 IWG 诊断标准，混合型阿尔茨海默病必须与具有非典型临床表现的非典型阿尔茨海默病区分开，例如脑后部变异型、少词变异型和额叶变异型等。

阿尔茨海默病痴呆和血管性痴呆经常同时存在，患者同时具有神经退行性变化和血管变化，但其他疾病也可能促使混合型痴呆患者的认知衰退，例如正常颅压性脑积水、海马

硬化以及路易体痴呆、额颞叶痴呆和亨廷顿病等其他痴呆。

根据 IWG 诊断标准,对于混合型阿尔茨海默病的诊断,必须有典型或非典型阿尔茨海默病的证据(基于临床表型),且同时有至少 1 个阿尔茨海默病病理的证据;同时,还应具备共存疾病的临床表现以及神经影像学或生化证据。

4.关于特殊人群的考虑

临床试验中纳入不同年龄组的受试者是有必要的,例如老年与更高年龄的阿尔茨海默病患者。确证性临床试验通常应纳入合理数量的老年患者(包括>65 岁、>75 岁和>85 岁),其中 75 岁及以上受试者的样本量需足够评估该人群的有效性和安全性。研究人群应能体现目标人群的年龄和合并症情况。

(二)生物标志物的作用和类型

阿尔茨海默病临床试验中的生物标志物可根据其潜在用途进行分类:①诊断:用于明确诊断;②富集:用于选择目标人群;③预后:用于确定疾病病程;④预测:用于预测未来对治疗的临床应答和安全性评估;⑤药效学:用于确定预期或非预期的药效学活性。

在阿尔茨海默病药物的临床研发中,用于上述特定目的的大部分生物标志物尚需进行验证。目前,脑脊液(cerebrospinal fluid, CSF)标志物以及 MRI 和 PET 影像标志物可用于研究人群的富集,但这些生物标志物在临床前

AD 的使用场景可能仍需要进行验证。随着检测技术的发展和经验积累，预期将有更多的生物标志物可用于阿尔茨海默病药物临床试验。鼓励开发和验证可靠、灵敏的新生物标志物。

淀粉样蛋白 PET 和 CSF A β 多肽亚型可测量淀粉样蛋白生物学的不同方面：PET 用于检测 A β 斑块负荷；CSF A β 42 用于检测可溶性 A β 42 单体水平，与斑块间接相关。对于试验富集，CSF 与 PET 淀粉样蛋白生物标志物是相关的，但尚不清楚相关程度，这取决于检测类型和阈值。这些检测方法在多大程度上能够作为可相互替换的富集措施检测出不同的潜在生物阶段，应通过数据来证明，以确保选择的是一个同质化人群。尽管 CSF A β 42 检测的性能已显著提高，但建议不局限于测量 A β 42，A β 42 与其他生物标志物的比值（如 A β 42/A β 40、A β 42/p-tau181、A β 42/t-tau 等）相较于单独使用 A β 42 可能具有更好的预测性。与 A β 相关生物标志物类似，还可测量总 Tau 蛋白（t-Tau）或磷酸化 Tau 蛋白（p-Tau）水平、Tau 蛋白 PET 等。

APOE 是与阿尔茨海默病发病风险相关的易感基因。APOE ϵ 4 状态可作为临床试验人群富集的指标之一，但如果仅纳入具有该特定基因型的患者而不纳入任何非携带者数据，则必须证明该药物普遍适用是合理的（除非该药物仅特定针对 APOE ϵ 4 基因型的阿尔茨海默病人群）。

脑区域结构和代谢变化的下游局部生物标志物（例如，通过 MRI 评估的海马萎缩，通过 FDG PET 评估的皮质低代谢等），虽然病理学特异性不足，但对于检测和定量分析疾病进展可能有一定的价值。

替代终点在特定的早期阿尔茨海默病药物开发项目中的可接受性取决于疾病的分期、试验入组的人群、药物的作用机制、以及当前治疗的可及性等因素。通常情况下，即使拟使用替代终点支持上市申请，也应在临床试验中考察临床结局评估指标，以评估早期的临床变化，这些临床变化可能为观察到的生物标志物变化提供潜在的支持。目前尚无可靠的药效学指标作为替代终点，鼓励申请人在研发中探索和充分验证潜在的替代终点。如拟使用替代终点支持上市申请，需提供充分的支持性依据，并与监管机构进行沟通。

在识别疾病进展可能更快的患者方面，目前尚无一种特定的生物标志物始终优于其他可选择标志物。认知功能的衰退过程可能会受到认知储备、合并疾病、生活方式和认知训练等因素的影响，应予以关注，并提高临床试验质量控制。

（三）有效性评估工具

1. 一般考虑

鼓励开发和验证可靠、灵敏的工具来测量认知、日常功能等，特别是在疾病早期。可考虑利用现有的文献和数据资源，开展适当的患者体验数据研究，结合研究目的、目标适

应症人群的疾病特点、药物作用机理及临床定位等因素，开发和验证适宜的有效性评估工具。

认知、日常功能和总体评估指标涵盖了评估阿尔茨海默病患者有效性的关键领域，其他还包括与健康有关的生活质量评估工具（包括通用和疾病特异性的评价工具）、痴呆的精神行为症状评估工具等。生物标志物也可作为有效性评估工具之一。

量表是阿尔茨海默病领域有效性评价的常用工具之一。由于目前没有理想的单独量表可全面评估药物的治疗效应，申请人可能需要使用多种量表来评估阿尔茨海默病药物的有效性。虽然已经出现了许多评价认知和日常功能变化的评价工具，但目前仍没有可满足上述不同临床分期所有要求的标准评价工具，因此，对评估工具的选择仍保持开放态度，但需提供合理的依据。

在使用评估工具时，建议由对治疗分配设盲的评估者执行。如果存在可使研究者知晓盲底的特殊情况（如严重不良事件等），则应尽可能使所有临床疗效评估者保持盲态。评估者最好不在其他方面参与临床试验的实施。应提前对评估者进行培训，从而最大程度的减少评估变异性，以及最大程度提高评估者之间的可靠性。理想情况下，不同维度的评估者培训应标准化，以减少评估变异性。

目前在重度痴呆患者中进行的研究相对较少，在该类患

者中通常需要调整评估工具，以便对日常功能和总体领域进行全面评估，且应更加侧重于日常功能以及痴呆的精神行为症状。

可以通过选择或删除现有量表中的单个条目，或使用单个子测试的特定加权，或同时进行这两种操作，开发更敏感的量表用于早期的阿尔茨海默病患者。如果开发同时评价认知和日常功能两个维度的综合评估量表，应量化单个条目和维度。应用此类方法开发新评估工具时，需考虑临床治疗目标并通过拟定的工具充分体现这些目标。无论采用哪种方法，新开发的评估工具必须证明能够测量临床获益。

日常生活活动能力的条目（例如处理财务、遵守约定、任务准确性和技术技能等）可能是痴呆较早期的敏感指标，而基本的自理能力或工具性活动（例如购物、洗衣服或烹饪等），可能在后续进一步发展阶段更敏感。现有评估量表中的某些组成部分，特别是认知评估量表，对检测阿尔茨海默病较早期的疾病进展较敏感（例如，单词记忆、单词识别、执行功能等）。开发新的评估工具应进行验证（例如，在独立人群中进行前瞻性研究），建议必要时与监管机构进行沟通。

2. 有效性评估工具示例

已有许多不同维度的有效性评价工具在既往临床试验中使用，以下仅列出其中的部分示例。目前正在开发更多的有效性评估工具。

认知功能测试

阿尔茨海默病评价量表的认知子量表（Alzheimer's disease assessment scale cognitive subscale, ADAS-cog）已广泛应用于轻度至中度痴呆的阿尔茨海默病研究。由于其天花板效应和地板效应，原始版本对疾病早期和晚期的变化的灵敏度有限，后续对该量表的条目进行了部分改良。

日常功能

已有几种评价日常功能的量表可测试与躯体活动相关的基本日常生活能力（例如，如厕、活动、穿衣、洗澡等）和工具性日常生活能力（例如购物、烹饪、洗衣服、处理财务、使用交通工具、驾驶和打电话等）。但是，这些量表更侧重于一般自理能力或家庭生活能力，而忽视了现代生活中可能更为适用的其他能力（如使用科技产品的能力），导致一些测试量表的敏感性较低。其中，阿尔茨海默病协同研究 ADL 量表（Alzheimer's disease cooperative study-activities of daily living, ADCS-ADL）是一种广泛应用于轻度至中度阿尔茨海默病患者临床试验的日常功能评价量表，也可结合药物的预期治疗效应，考虑是否应用于其他临床分期患者。目前尚没有一种工具能够比其他工具更好地评价日常功能的轻微改变，可考虑在特定的试验人群中验证现有工具，或开发新的日常功能评价工具，聚焦于在认知衰退初期患者中也能反映出变化的条目。

总体疗效评价

总体疗效评价是指由具有经验的临床医师对患者整体状况作出的主观独立等级评定。虽然临床医师的总体疗效评价具有一定局限性，但可以作为评价药物治疗晚期痴呆患者临床获益的一种有用手段。

总体疗效评价量表允许临床医师对患者的症状和行为做出单一的主观综合判断，这不同于通过涵盖不同维度的综合量表或成套测试来评价不同维度的功能。临床医师访谈印象变化量表（Clinician's interview based impression of change plus, CIBIC-plus）可以对照基线对痴呆患者的总体临床状况进行评价，根据对患者及其护理者的半结构化访谈中获得的信息来评价，但没有考虑任何来源的认知表现。阿尔茨海默病协同研究组织的医生临床变化总体印象量表（Alzheimer's disease cooperative study unit clinician's global impression of change, ADCS-CGIC）是另一种根据患者及其护理者的总体信息的半结构化访视。

认知和日常功能复合评价

临床痴呆评定量表-总分（clinical dementia rating-sum of box, CDR-SB）是一项复合评价量表，由3项认知领域和3项日常功能领域的评分组成，可评估认知和日常功能两个维度的变化。

痴呆的精神行为症状

阿尔茨海默病病理行为评定量表（behavioural pathology in Alzheimer's disease rating scale, BEHAVE-AD）、痴呆行为评定量表（behavioural rating scale for dementia, BRSD）、神经精神症状问卷（neuropsychiatric inventory, NPI）等可作为阿尔茨海默病痴呆的精神行为症状的评价工具。Cohen-Mansfield 激越问卷主要用于评价患者的激越。

四、临床药理学研究

（一）药代动力学

应结合相关指导原则开展全面的药代动力学研究。

除旨在确定药物的吸收、分布、代谢和消除的常规药代动力学研究外，通常应考察老年人群的药代动力学特征。群体药代动力学模型有助于模拟以老年人为主的患者人群中的药代动力学特征。对于某些作用机制的假设，药物穿过血脑屏障并在脑中进行靶点结合的信息对于解读药物作用机制的临床效应是重要的。必要时，还应在肝脏和/或肾脏功能不全患者中进行药代动力学研究。

（二）药效学

在阿尔茨海默病药物开发早期，确定药物可能具有治疗活性的药理学机制很重要。表征药物的主要药效学活性（例如，对于受体/神经递质途径的活性，对淀粉样蛋白级联的活性，对 Tau 聚集的活性，对神经炎症的活性等）将影响后续的临床研发计划。鼓励探索采用新技术新方法考察药物的药

效学，例如通过模型模拟等方法。健康受试者中，不良反应和药理学活性的可能替代标志物（如果可用且相关）有助于估计适当的剂量范围。

（三）药物相互作用

药代动力学相互作用研究应结合相关指导原则开展。

若试验药物为特定代谢酶或转运体的底物或诱导剂、抑制剂时，应选用特定的指针药物进行相互作用研究。

对预期在临床实践中会同时给予的试验药物、其他抗痴呆药物和其他药品，通常应进行药代动力学相互作用研究，除非有明确的基于机制的证据表明，预期不会发生相互作用和/或试验药物的给药途径会限制与其他药品的相互作用。

另外，应考虑到阿尔茨海默病患者大多数为老年这种特征，特别是该人群通常可能对与试验药物同时给予的某些药物（包括精神疾病药物、抗血小板药物和脂质代谢药物）的药效学更为敏感。

五、探索性临床试验

随着阿尔茨海默病研究领域的快速发展，有多款新靶点、新机制创新药物研发，但其中某些药物的探索性试验可能没有为确证性临床试验提供可靠的“概念验证”信息，导致大型确证性临床试验失败。因此，在开展确证性临床试验之前，强烈建议在特征明确的患者人群中进行充分的探索性试验。

探索性临床试验可能有以下目标：证明靶点结合、从临

床和实验室角度评估短期不良反应、确定药代动力学特征、确定最大耐受剂量、确定 PK/PD 关系、确定剂量-效应关系、有效性的初步评估、概念验证、确定能够从治疗中获益的患者亚群以及确证性临床试验的人群选择等。

此类试验的持续时间将取决于预期可测量的应答时间或待评估的参数。某些生物标志物的研究已取得显著进展，可考虑作为作用机制验证研究的主要观察指标。建议在探索性临床试验中同时收集多方面的临床有效性和安全性数据，为后续开展确证性临床试验提供更多参考。

六、确证性临床试验

（一）不同研发目标药物的试验设计

1. 症状治疗

对于症状治疗药物，应通过随机、双盲、平行对照试验证明在认知、日常功能或临床总体评估方面的治疗效应，该治疗效应通常需基于较基线的变化来证明，也可通过达到有临床意义获益的患者比例（应答率）来证明。如拟证明应答率方面的治疗效应，应预先明确定义“应答”，例如，可定义为认知指标改善至预先指定的程度，同时在其他两个领域（日常功能和总体评估）至少没有恶化。考虑到试验过程中疾病的自然进程，需要谨慎制定应答的标准。

不同阿尔茨海默病痴呆药物试验之间的安慰剂组人群可能存在差异，且无法排除存在未经治疗的症状改善，为了

确定新药在症状治疗药物中的定位以及研究分析的敏感性，推荐的试验设计是试验药、阳性对照药和安慰剂相比的三臂试验。为最大程度减少使用安慰剂的伦理问题，可以接受不平衡比例的随机化。考虑到轻度至重度痴呆的阿尔茨海默病患者目前已有多种症状治疗药物可供选择，在该类患者中进行长期安慰剂对照的可行性可能会受到限制，故也可考虑采用阳性对照、优效性设计的双臂试验。由于症状治疗的疗效评价指标通常为主观指标，对试验敏感性存在担忧等因素，一般不能接受仅与阳性对照药物对比的非劣效性设计作为有效性的关键证据。

如果新药计划在标准对症治疗（例如胆碱酯酶抑制剂、谷氨酸受体拮抗剂等）的基础上使用，则可采用双臂、安慰剂对照、加载（add-on）设计。

对于痴呆前的早期阿尔茨海默病患者（前驱期 AD/AD 源性 MCI 等），在保障受试者安全的前提下，可考虑选择安慰剂作为对照。

研究持续时间通常应足以评估药物的治疗效应。阿尔茨海默病症状治疗药物的研究持续时间依赖于研究人群的临床分期。轻度至中度阿尔茨海默病患者的对照临床试验通常需持续至少 6 个月。

建议在双盲期后进行至少 6 个月的长期安全性研究，可考虑采用开放标签设计。应根据药品的预期起效速度和试验

持续时间，定期进行阶段性的有效性和安全性评价。在治疗结束后，应在适当时限内对患者进行随访，以观察与停药相关的可能不良事件。

2. 疾病修饰治疗

当药物治疗延缓了疾病的病理学或病理生理学进展，同时证明可延缓症状进展，可被认为是疾病修饰的治疗药物。疾病修饰作用，需证明受试者临床体征和症状衰退速度的延缓，且这些结果与经过充分验证的生物标志物的显著变化是相关的。此类生物标志物应基于合理的疾病模型，反映出潜在疾病过程的关键病理生理学特征。

在保障受试者安全的前提下，推荐采用随机、双盲、安慰剂对照、优效性试验确定有效性。在有公认的阳性对照时，也可采用试验药、阳性对照药和安慰剂相比的三臂试验设计。

胆碱酯酶抑制剂或美金刚等常作为标准的对症治疗药物，特别是对于轻度至中度痴呆的阿尔茨海默病，可在随机化时对这些药物的使用进行分层。

研究持续时间与研究人群和预期的治疗效应等相关，最短持续时间取决于预期的疾病进展速度和药物作用活性强度。例如，对于痴呆前的早期阿尔茨海默病患者（前驱期 AD/MCI 源性 MCI、数字临床分期的第 3 期等）和轻度至中度阿尔茨海默病患者，通常需持续至少 18 个月，某些情况下可能需要进行更长时间的研究。

根据产品的作用机制，治疗干预的时机可能对结局至关重要。如果仅在前驱期 AD/AD 源性 MCI 等早期患者中证实了药物的延缓疾病进展治疗作用，其有效性难以外推至病程后期（中度或重度痴呆）。

对于疾病修饰治疗药物，通过对各治疗组之间临床症状随时间变化的速度进行统计比较（例如斜率分析等）是合适的。但应考虑到，尽管已知阿尔茨海默病的自然病程可能随时间推移呈近似线性模型，但尚不清楚该线性特点在具有干预治疗效应（潜在的疾病修饰作用）的临床试验中是否成立，以及治疗效应在治疗过程中是否保持不变。此外，随时间推移而增加的药理学可逆效应也可能影响线性特点。因此，应定期测量各平行组的治疗效应。通常期望在基于模型的分析中纳入多个时间点数据，以更好表征预设指标的下降速率是否随时间而减慢。使用的模型和检验假设应是合理的。此类研究可通过延迟启动治疗的设计来加强，目的是证实延迟启动治疗患者与更早期启动治疗患者之间在临床指标上的差异在整个研究期间可维持。使用延迟开始治疗的设计，随访时间长度和分析参数非常重要，如果随访时间过短，可能难以显示出预期的疗效变化曲线。

也可通过时间-事件统计分析方法来评估可能的疾病修饰效应。在疾病较早期，症状轻微，疗效随时间的变化可能难以评估，此时推荐的指标是衰退至一个预设的有临床意义

终点的时间。预设事件必须具有明确的临床重要性（例如，若预设事件为“发生痴呆”，则相应的评估指标可为“至发生痴呆的时间”），通常应事先基于可接受的主要疗效终点评估工具进行定义，而不是简单的根据某个维度量表评分的下降来定义（例如，ADAS-cog 下降 2 分）。预期患者发生该事件之前的时间必须足够长。

以上描述的确定疾病修饰效应的方法有其缺点，并且可能进一步受到以下因素的干扰：安慰剂治疗患者的可能改善、总体脱落率和缺失数据差异、治疗依从性较差、治疗效应随病程的变化、终点指标随时间推移的敏感性等。因此，应在研究方案中论证主要分析方法的选择、统计模型、基本假设和要求等的科学性。

减缓或延迟临床衰退的证据，应与生物标志物显示的延缓大脑神经退行性变进展的证据一同提供。在缺乏经充分验证的生物标志物作为结局指标的情况下，难以证明与延缓潜在神经病理学过程相关的疾病修饰效应，故通过上述创新试验设计和相应的合理分析支持的病程变化证据可能是可接受的替代研发目标。

疾病修饰治疗的联合用药：阿尔茨海默病的病理生理学可能涉及多种通路（可能是多因素的），因此，联合使用具有互补作用机制的疾病修饰治疗药物可能具有重要的治疗作用。如果拟联合使用两种疾病修饰治疗药物进行研究，通常

需要显示每种药物对目标作用机制和临床有效性的贡献。例如，开展一项将联合治疗与两个单药治疗组进行比较的试验，并在必要时与安慰剂也进行比较。但是，由于在长期研究期间每组所需的样本量大，因此疾病修饰治疗可能难以实现完全的析因设计。

3. 疾病预防药物的开发

预防的总体目标是降低目标人群中的发病率，或防止处于临床前状态的分期进展到后期症状更加明显的分期。可根据基因标记（例如 APOE $\epsilon 4$ 状态、常染色体显性突变等）、生物标志物（例如 CSF 或外周血 A β 和 tau 水平、PET 中淀粉样蛋白或 tau 示踪剂的滞留等）或环境风险因素（例如血管或代谢因素）等，富集预防试验的人群。

阿尔茨海默病是一种多因素疾病，但尚未确定每种风险因素对疾病发生的相对贡献，难以在个体水平上解释群体的风险。

针对症状前临床分期的阿尔茨海默病潜在病理生理机制的药物干预可能是预防策略的合理方法。应在富集人群中进行随机、双盲、安慰剂对照试验。但是，如此早期的疾病模型仍有待验证，同时需考虑个体间变异性和其他风险因素对疾病进展速度的影响。阿尔茨海默病病理累积和临床症状发生的时间进程尚未完全确定，大脑对结构变化的反映和适应能力在个体间存在很大差异（认知储备），甚至每个特定

患者在每天也会有变化。由于上述原因，高风险人群的主要干预目标仍然是预防认知受损，因为目前还没有生物标志物可以被视为有效的替代终点。

预防试验患者在基线时没有可评估的临床损害，且患者症状的出现时间具有不同的潜伏期，因此在典型持续时间（如 2 年或更短）的临床试验中很难证明具有临床意义的获益，通常需要大样本量和长期随访（通常至少 3 年）。目前仍然缺少相关的科学信息来为预防试验设计提供可靠的支撑，如申请人拟开展预防试验，建议与监管机构进行沟通。

4.痴呆的精神行为症状的对症治疗

总体上，阿尔茨海默病的对症治疗还包括针对痴呆的精神行为症状的治疗，例如激越、攻击行为、情感淡漠、精神病性症状（妄想和幻觉）、抑郁症状、焦虑和睡眠障碍等。虽然痴呆的精神行为症状未纳入正式的阿尔茨海默病诊断分类中，但其在阿尔茨海默病患者人群中十分常见，是晚期痴呆患者临床恶化的重要原因，可导致疾病负担和压力增加，特别是对于家庭成员或照护者。痴呆的精神行为症状是多变的，随着病程波动，应考虑“假特异性”问题。激越和攻击等行为症状在痴呆晚期更常见，而抑郁和情感淡漠等情绪症状在痴呆早期更常见。症状群是经验性的还是具有生物学支持证据尚未确定，因此，治疗 **BPSD** 单一症状或症状群的可能性必须有强有力的理论依据，并取决于药物的作用机制。

对于痴呆的精神行为症状的对症治疗，通常应选择安慰剂、平行对照试验。尽管利培酮因为在老年人群中存在特定的安全性问题，仅被批准用于持续性恶化的短期（6 周）治疗，但一种新的研究药物可以在标准治疗（非药物治疗和药物治疗）的基础上进行评价。此外，患者的周围环境对治疗结局可能有较大影响。不同研究中心的标准治疗差异很大，在临床试验中应尽可能减少这种差异。

对于痴呆的精神行为症状的对症治疗，建议持续 8 至 12 周，但也取决于症状及其波动情况，并应说明理由。在临床实践中可能会延长治疗时间，因此需要采集更长期的数据为有效性的维持、反跳作用、撤药现象和安全性等提供证据。如果该弱势人群中出现严重的安全性问题，则不能只依赖开放扩展期的数据，需要采用平行对照设计。

（二）有效性指标

有效性指标的选择需关注患者感受、功能状态和生活质量等，可采用临床结局评估（clinical outcome assessment, COA）工具进行有效性评价。

创新药物临床研发时通常需考虑药物的作用机制和研发目标，设计相应的适宜疗效指标去证实是否能够达到预期的研发目标，但对于阿尔茨海默病药物，不同临床分期患者的临床特征和治疗学需求存在较大差异，因此，在考虑药物作用机制和研发目标的基础上，还应结合研究人群的不同临

床分期考虑合理的有效性评价指标。

1.阿尔茨海默病痴呆期

阿尔茨海默病痴呆期患者的主要临床表现为认知功能障碍伴日常生活能力下降等，对于痴呆期患者，可基于以下三个维度的指标评估患者的临床获益：

(1) 认知终点：主要通过相关临床量表评估的认知功能；

(2) 日常功能终点：日常生活能力；

(3) 总体终点(总体临床疗效)：通过总体评价来体现。

上述三个维度中的每个维度应设置合理的有效性变量。

考虑到不同痴呆程度的阿尔茨海默病患者临床特征和治疗学需求存在一定差异，故相应的有效性评价指标的侧重点可能有所区别。在轻度至中度阿尔茨海默病痴呆患者中，通常需证明对认知的影响具有临床意义，同时将日常功能指标或临床总体评价指标作为共同主要终点来确认疗效的临床意义。在重度阿尔茨海默病痴呆患者中，认知表现变化的临床获益相关性可能较低，且更加难以量化，因此，日常功能和临床总体评价维度指标可能更适合作为共同主要终点来确定该人群中具有临床意义的改善。

阿尔茨海默病痴呆期关注的次要疗效终点可包括健康相关生活质量量表以及行为和精神病学症状量表等。在痴呆晚期，精神行为问题对患者和照护者有较大影响。

2.前驱期 AD/AD 源性 MCI

在痴呆前的早期阿尔茨海默病患者(例如前驱期 AD/AD 源性 MCI, 或数字临床分期第 3 期等), 由于当前可用量表的局限性, 使用两个共同主要终点进行认知和日常功能或临床总体评价可能是困难的, 但仍有必要证明试验结果的临床意义。这也适用于在一项研究中同时研究痴呆前早期阿尔茨海默病患者和轻度痴呆阿尔茨海默病患者。

目前使用的认知量表具有天花板效应, 其敏感度可能不足以检测认知的微小变化, 而复杂的神经心理组合测试可能也难以在大型临床试验中实施。此外, 接近痴呆期的患者在日常功能方面存在微小但已经可以观察到的损害, 但每个患者能够补偿其认知损害和调整其日常活动的程度各不相同, 且日常功能损害的进展可能非常缓慢, 使用现有量表可能会产生可行性问题(研究的样本量和把握度)。因此, 针对早期阿尔茨海默病患者构建更敏感的评分条目, 或筛选那些已证明在早期阿尔茨海默病患者中持续受损的维度, 并开发出可靠、灵敏的工具, 可能是今后的发展方向。

对于痴呆前的早期阿尔茨海默病患者研究或同时纳入痴呆前早期阿尔茨海默病患者和轻度痴呆阿尔茨海默病患者的研究, 采用同时评价认知和日常功能两个维度的复合测量指标作为主要疗效指标是可接受的。该复合测量指标应能够充分且有意义的评估认知和日常功能的独立影响, 且应将认知和日常功能指标分别单独作为次要疗效指标。此外, 可

考虑将工具性活动能力、执行功能和健康相关生活质量指标等也作为次要疗效指标，以支持有效性的全面评估。

3.临床前 AD 期

对于临床前 AD 期患者，干预目标多为预防或延缓疾病症状发生。已在预防试验中使用的认知终点包括痴呆的诊断（基于阈值评分）、认知功能显著衰退、基于某些测试纵向表现的认知功能改变等。时间-事件分析可能作为一种支持临床相关性的补充指标，但目前还存在可行性方面的问题，包括试验时长和脱落率等。时间-事件分析中的“事件”应具有明确的临床意义，例如发生认知受损等。在某个生物标志物被确定为治疗效应的可靠替代指标之前，通常应对患者进行足够时间的随访，以采集相关的认知变化。

某些药物可能定位于疾病治疗。在阿尔茨海默病中尽早干预是非常有必要的，因此具有阿尔茨海默病特征性病理生理变化但没有功能损害或敏感神经心理评估中可检测到异常的患者（例如数字临床分期第 1 期）是临床试验的重要目标人群。在这一疾病阶段，在典型持续时间（如 2 年或更短）的临床试验中很难证明具有临床意义的获益，因为这些患者在基线时没有可评估的临床损害，且患者症状的出现时间具有不同的潜伏期。对此，通过各种生物标志物检测所明确的对阿尔茨海默病特征性病理生理变化的效果，预计可能是适当的评估方式。与使用神经心理测试评估一样，通过多个单

独生物标志物检测得到的治疗效果将增加假定效果的说服力。如拟使用替代终点支持上市申请，通常应在临床试验中考察临床疗效指标的变化，这些临床变化可能为观察到的生物标志物变化及其临床意义提供潜在的支持。考虑到目前尚无可靠的药效学指标作为替代终点，此类情形建议与监管机构沟通。

对于数字临床分期的第2期，仅在敏感的神经心理测试评估中可检测到轻微认知损害且没有功能损害的证据，可能需要很长的时间来确定在这些患者中有临床意义的治疗效果，可行性方面存在较大挑战。对于在敏感的神经心理测试评估中检测到的微小变化，其临床意义的解释可能具有挑战性，但是，认知功能对于阿尔茨海默病患者的日常功能具有重要影响，非常显著的认知变化可能可代表具有明显临床意义的变化。在多个单独的神经心理测试中展示出具有获益疗效的一致趋势可增强研究结果的说服力。基于上述认识，通过敏感的神经心理测试评估对认知具有明显临床意义的疗效，同时这种认知维度的疗效可以得到阿尔茨海默病特征性病理生理显著变化的支持，其总体临床获益可能是可接受的。此类情形建议与监管机构沟通。

4. 痴呆的精神行为症状

痴呆的精神行为症状通常被视为单独的适应症，需开展单独的临床试验。这些临床试验通常需要可靠且经过验证的

测量工具来评估药物的疗效。既往临床试验中已使用了多种量表，应根据目标症状和研究人群进行选择。鼓励开发评估早期痴呆的精神行为症状的敏感工具。针对阿尔茨海默病痴呆的精神行为症状的临床试验中，应将认知和日常功能作为次要疗效终点，以排除药物可能对这些症状的负面影响。在以认知和日常功能为主要疗效终点的临床试验中，也可将痴呆的精神行为症状作为次要疗效终点，但在这种情况下不能将研究结果外推成独立的适应症。

七、统计学考量

阿尔茨海默病治疗药物临床试验的统计设计与分析，应符合临床试验生物统计学的基本原则。需注意的是在确证性试验阶段统计分析方法的选择（包括缺失数据的处理策略）应与预先指明的估计目标对应，能反映试验目的与拟验证的研究假设，方法选择应有充分的依据。针对估计目标和亚组分析有如下考虑。

阿尔茨海默病药物确证性临床试验的估计目标框架应符合 ICH E9 (R1) 的要求。需根据试验目的确定估计目标，包括人群、治疗、变量（终点）、伴发事件及其处理策略和群体层面汇总五个属性，并在试验方案和统计分析计划中明确描述，可参见 ICH E9 (R1)。

阿尔茨海默病不同临床分期、不同研发目标的确证性临床试验的主要估计目标可能会不同，通常应反映与产品研发

目标相对应的干预效应。鼓励在临床研发的早期阶段收集可能影响疗效评价的伴发事件，为后续构建估计目标提供依据。

对伴发事件的明确定义及相应处理策略的适宜选择，在阿尔茨海默病药物临床试验中至关重要。伴发事件可能包括但不限于终止治疗、使用非研究药物（包括方案规定的“补救”药物）、发生与阿尔茨海默病或者治疗无关的事件和死亡等。对于伴发事件适合的首选处理策略，可能因关注的临床问题（例如，证明症状治疗效应或疾病修饰的治疗效应等）而有所不同。例如，对于症状治疗，因安全性或耐受性原因“终止治疗”，可考虑采用疗法策略；若治疗目的是延缓或阻止临床症状进展，因疗效不佳而“终止治疗”的受试者可视为治疗失败而采用复合策略，如果是为了估计纯粹的药物疗效也可考虑采用假想策略。对于阿尔茨海默病痴呆相对晚期的长期试验中，可能会有一定比例的患者死亡，此时对死亡的处理可能会依赖于预设的研究假设（如药物治疗效应是否对生存期存在影响）。应特别注意，伴发事件的预期发生率及其对结局变量的潜在影响，在不同研发目标的试验中可能有差异，通常需具体问题具体分析，必要时可与监管机构沟通交流。

阿尔茨海默病药物的临床试验涉及疗程长、安慰剂对照，以及主观性强的量表工具等。为了保证主估计方法得到的统计推断结果的稳健性，建议针对同一估计目标采用一种或多种形式的敏感性分析，包括对缺失数据的不同处理方法进行

敏感性分析等。

阿尔茨海默病的发病机理目前尚未完全明确，有多种相关的假说，即便是同一疾病分期的患者，往往都具有不同程度的异质性。因此，针对目标适应症全人群的确证性临床试验，应进行必要的亚组分析（例如，不同年龄组、不同临床分期、生物标志物水平等），以便能更好的提供支持性证据评价药物的治疗效应与安全性。对于确证性亚组分析，需要在临床试验方案和/或统计分析计划中预先定义亚组 and 规定多重性调整方法（如涉及）。具体可参见《药物临床试验亚组分析指导原则（试行）》。

八、安全性评估

阿尔茨海默病药物通常应遵循新药临床安全性评价的一般要求（可参考《新药临床安全性评价技术指导原则》等），以及 ICH E1 对药物暴露量和暴露时间的要求。

对于不良事件，通常需描述治疗持续时间、药物使用剂量、严重程度、恢复情况、恢复时间等，尤其是不同年龄组（如老年组和超高龄组患者）以及其他相关变量的特征。应特别注意药物作用机制或作用靶点相关的潜在安全性风险，例如胆碱酯酶抑制剂的拟胆碱作用等；对于 β -淀粉样蛋白抗体药物，通常需要通过 MRI 监测淀粉样蛋白相关性影像异常（amyloid related imaging abnormalities, ARIA）。

因阿尔茨海默病患者有认知功能障碍，常不能准确和客

观地表达不适症状，因此需要临床研究者和照料者进行细致观察，必要时应增加随访频率，以便及时发现和处理。在发现 AE 时，应注意不要引入偏倚，询问 AE 的首选方法是对受试者进行开放性和非引导性口头提问。

在短期临床试验之后，建议进行至少 6 个月的治疗随访。可采用开放标签扩展设计。这不仅可以充分回应伦理问题，还可以积累药物的中长期安全性数据和估计对症状影响的最长持续时间。

需要特别关注神经系统不良事件的发生或加重，特别是脑血管事件、锥体外系症状、定向障碍、步态进一步受损、癫痫发作等。根据作用机制和作用靶点信息，对于可能会发生的特定神经系统不良事件，需要进行特殊的监测。靶向作用于 β -淀粉样蛋白的抗体药物治疗已显示出可引起不同程度和发生频率的 ARIA，具体取决于药物的活性、靶点、剂量和患者特征（APOE ϵ 4 基因携带情况、基线脑微出血数量或其他）等。这类不良事件的临床影响尚需进一步研究，在临床试验期间应予以特别关注，定期进行 MRI 监测，收集相关的安全性信息，并制定风险控制措施。此外，通常还需监测停药的影响。

应特别关注幻觉的发生，以及情感障碍或精神障碍的其他体征和症状。根据试验药物的药效学特征，可能需记录神经-行为异常，特别是定向力障碍、激越和攻击行为等。

药物过量：根据作用机制、预期治疗方案，可能需要对药物过量的风险和影响进行研究。

自杀想法和行为：通常应使用经验证的量表对试验药物引发自杀想法和行为的可能性进行系统的评估，例如哥伦比亚自杀严重程度评定量表（columbia suicidality severity rating scale, C-SSRS）或其他经验证的量表。应报告自杀事件（包括自杀意念、自杀未遂和自杀身亡）的发生率，并分析相关因素的影响，包括剂量、治疗持续时间及其他促成因素等。应提供自杀患者陈述或行为的叙述性总结。

根据药物的药效学特征，监测其对心血管系统的影响，例如直立性低血压的发生、诱发心律失常的可能性或增加心肌梗死的风险等。

考虑到阿尔茨海默病患者需要长期药物治疗，故通常需获得大样本且具有代表性的患者在足够长周期内的数据（可参考 ICH E1 相关要求）。不同年龄组（年轻、年老和高龄）的长期安全性特征可能不同，应予以考虑。早期疾病阶段的患者人群应特别关注长期安全性，这些患者可能在无症状期接受治疗多年，某些不良反应可能是明显的。

应长期监测对死亡率的影响，特别是对于无症状期的患者人群。上市后风险管理计划中也应关注死亡事件的监测。

九、与监管机构沟通

鼓励申请人在药物研发的关键阶段，基于药物特点和前

期研究数据，就临床研发计划、临床试验方案设计的关键问题、本指导原则未涵盖的问题等与药品监督管理机构进行沟通交流。

十、参考文献

[1] EMA, Committee For Medicinal Products For Human Use (CHMP). Clinical investigation of medicines for the treatment of Alzheimer's disease[EB/OL]. 2018. <https://www.ema.europa.eu/en/clinical-investigation-medicines-treatment-alzheimers-disease> [published online ahead of print].

[2] FDA. Early Alzheimer's disease: developing drugs for treatment, guidance for industry[EB/OL](2024-03-12). <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/early-alzheimers-disease-developing-drugs-treatment>.

[3] McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the national institute on aging - Alzheimer's association workgroups on diagnostic guidelines for Alzheimer's disease[J]. *Alzheimers Dement*, 2011, 7(3): 263-269.

[4] Jack CR Jr, Bennett DA, Blennow K, et al. NIA - AA research framework: toward a biological definition of Alzheimer's disease [J]. *Alzheimers Dement*, 2018, 14(4): 535-562.

[5] Jack CR Jr, Andrews JS, Beach TG, et al. Revised criteria for diagnosis and staging of Alzheimer's disease: Alzheimer's

Association Workgroup[J]. *Alzheimers Dement*. 2024 Jun 27. doi: 10.1002/alz.13859. Online ahead of print.

[6] Dubois B, Feldman HH, Jacova C, et al. Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria[J]. *Lancet Neurol*. 2014 Jun;13(6):614-29. doi: 10.1016/S1474-4422(14)70090-0.

[7] 田金洲, 解恒革, 王鲁宁, 等. 中国阿尔茨海默病痴呆诊疗指南(2020 年版). *中华老年医学杂志*[J], 2021, 40(3): 269-283.

[8] 中华医学会神经病学分会痴呆与认知障碍学组. 阿尔茨海默病源性轻度认知障碍诊疗中国专家共识 2024. *中华神经科杂志*[J], 2024, 57(7): 715-737.

[9] 中国医师协会神经内科医师分会, 阿尔茨海默病药物临床试验写作组. 阿尔茨海默病药物临床试验中国专家共识. *中华脑科疾病与康复杂志(电子版)* [J], 2022, 12(1): 9-20.

[10] Gauthier S, Rosa-Neto P, Morais JA, et al. World Alzheimer Report 2021: Journey through the diagnosis of dementia. London: Alzheimer's Disease International, 2021. <https://www.alzint.org/u/World-Alzheimer-Report-2021.pdf>.