

阿尔茨海默病治疗药物临床试验技术指导原则

(征求意见稿)

2024年9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、临床研发的整体考虑	2
(一) 治疗目标	3
(二) 研究人群	3
(三) 生物标志物的作用和类型	6
三、临床药理学研究	8
四、探索性临床试验	9
五、确证性临床试验	10
(一) 不同治疗目标药物的试验设计	10
(二) 有效性评价指标	15
(三) 有效性评估的工具	19
六、安全性评估	23
(一) 神经系统不良事件	24
(二) 精神系统不良事件	25
(三) 心血管系统不良事件	25
(四) 长期安全性	25
七、与监管机构沟通	26
八、参考文献	26

阿尔茨海默病治疗药物临床试验技术指导原则

(征求意见稿)

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丧失伴日常生活能力下降和行为改变为特征, 是全球需应对的重大公共卫生领域疾病之一。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 AD 患者。有文献报道我国 60 岁及以上老年人群中约有 983 万 AD 患者。随着人口老龄化进程加速, AD 发病率和患病率可能将逐年增加。

AD 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、 $A\beta$ 和脑血管异常的相互作用、神经炎症假说、氧化应激假说及遗传相关假说等, 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。AD 的主要病理特征为 $A\beta$ 沉积形成神经炎性斑块、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 AD 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, 近年来获批上市的新靶点、新机制药物包括仑卡奈单抗和甘露特钠等, 目前仍存在迫切

23 临床需求。

24 (二) 目的及适用范围

25 本指导原则旨在为 AD 不同疾病阶段的治疗药物临床试
26 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
27 研发。AD 的精神行为症状通常被视为单独的适应症，必须
28 进行单独的试验，本指导原则不适用该适应症的药物研发。

29 本指导原则主要讨论临床试验设计的重点关注内容。应
30 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
31 (GCP)、国际人用药品注册技术协调会(ICH)和其他国内
32 外已发布的相关指导原则。

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

36 二、临床研发的整体考虑

37 AD 是一个连续的疾病谱已成为共识，大致可包括临床
38 前 AD/认知未受损 AD 阶段、AD 痴呆阶段，以及介于两者
39 之间的中间阶段(具体参见“临床分期”章节)。既往的 AD 治
40 疗药物临床试验主要针对已表现出典型痴呆症状的阶段，目
41 前认为 AD 具有漫长的症状前阶段，在症状发生前 10-20 年
42 就可出现 AD 特征性的神经病理学改变，故可能有机会在 AD
43 疾病过程的非常早期进行干预。

44 基于不同的 AD 临床分期、预期的治疗效应和产品开发

45 目标，有效性和安全性评价技术要求有所区别，因此，通常
46 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
47 略。

48 此外，临床研发策略还需要考虑是否计划与目前的标准
49 治疗（如胆碱酯酶抑制剂、美金刚等）联合使用，是否作为
50 单药治疗进行开发，或者是否计划与靶向相似或不同 AD 病
51 理生理机制的新药联合使用等。

52 （一）治疗目标

53 不同作用机制的 AD 治疗药物，在临床研发计划和临床
54 试验设计中可能有不同的目标，其中主要的治疗目标包括：

55 1. 症状治疗：不改变疾病总体病程的治疗效应，可包括
56 认知和日常功能等维度的治疗效应。

57 2. 延缓疾病进展：延缓或阻止症状进展，并具有延缓神
58 经病理学过程的证据。

59 3. 疾病预防：在临床前阶段通过干预可能的致病机制来
60 预防症状发生。

61 （二）研究人群

62 1. 患者类型

63 1.1 常染色体显性 AD

64 常染色体显性 AD 是由几种已知的淀粉样蛋白基因相关
65 突变（如 PSEN1、PSEN2、APP 等）导致的，所占比例不足
66 1%。影响常染色体显性 AD 患者症状发生和疾病进展的因素

67 尚不完全了解。常染色体显性 AD 是开发新药以及验证评价
68 工具的重要模型，但该类型 AD 的病理生理学特征及其对治
69 疗的反应与散发性 AD 的重叠程度仍有待确定。

70 1.2 散发性 AD

71 常见的 AD 类型为散发性 AD，约占 99%。散发性 AD
72 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
73 程尚未完全明确。已知 AD 的病理过程在临床症状出现前数
74 十年开始，但该神经病理变化过程与症状进展和测量的特定
75 结局之间的关系有待进一步研究确定。

76 1.3 混合型痴呆和混合型 AD

77 混合型 AD 通常具备共存疾病的临床表现和病理学证据。
78 据文献报道，尸检时混合型 AD 至少占有所有 AD 病例的 50%。

79 AD 痴呆和血管性痴呆经常同时存在，患者同时具有神
80 经退行性变化和血管变化，但其他病理也可能促使混合型痴
81 呆患者的认知衰退，例如正常颅压性脑积水、海马硬化以及
82 路易体痴呆、额颞叶痴呆和亨廷顿病等其他痴呆。

83 2. 诊断

84 通常需要具有高敏感性和特异性的诊断标准来选择同
85 质化的研究人群。国际上已经制定了多套 AD 诊断标准，其
86 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
87 研究所-阿尔茨海默病协会（NIA-AA）标准和国际工作组
88 （IWG）标准，两者均可用于 AD 药物临床试验的诊断标准

89 和研究人群富集。NIA-AA 标准最早由美国神经病学、语言
90 障碍和卒中-阿尔茨海默病与相关疾病协会（NINCDS-
91 ADRDA）于 1984 年制定的标准，美国国家老龄化研究所-阿
92 尔茨海默病协会分别于 2011 年、2018 年和 2024 年对此标准
93 进行修订。IWG 标准是由 AD 研究诊断标准国际工作组（IWG）
94 于 2007 年制定，2014 年进行了修改。NIA-AA 和 IWG 两个
95 标准目前仍在不断的修订和完善。

96 3.临床分期

97 通常需根据试验药物的作用机制和治疗目标，考虑研究
98 人群的适宜临床分期阶段。NIA-AA 标准和 IWG 标准的临床
99 分期有相似之处，但也存在一定的差异。

100 在 NIA-AA 标准 2011 年版中，包括临床前 AD、AD 源
101 性轻度认知障碍（MCI）和 AD 痴呆阶段。在 NIA-AA 标准
102 2018 年版中概述了两种类型的分期：综合征分类性认知分期
103 和数字临床分期。其中，综合征分类性认知分期在很大程度上
104 保留了 2011 年版的临床分期方法，具体包括认知未受损
105 （CU）、轻度认知障碍（MCI）和痴呆阶段，其中痴呆进一步
106 细分为轻度、中度和重度。数字临床分期包括 6 个阶段，其
107 中 1~3 阶段代表痴呆前期，4~6 阶段代表痴呆期。综合征分
108 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
109 床研发中的不同需求，两者的阶段划分有一定的重叠交叉。
110 数字临床分期 1 阶段和 2 阶段大致对应 CU 阶段；数字临床

111 分期 3 阶段大致对应 MCI 阶段，但 MCI 阶段也适用于数字
112 临床分期 2 阶段中的某些个体；数字临床分期 4、5 和 6 阶
113 段分别大致对应轻度、中度和重度痴呆阶段。NIA-AA 标准
114 2024 年版基本保留了 2018 年版的综合征分类性认知分期和
115 数字临床分期，同时在 6 阶段数字临床分期的基础上增加了
116 0 阶段，即具有相关的基因突变，但无 AD 特征性病理生理
117 变化和临床症状。

118 IWG 标准的临床症状分期与 NIA-AA 标准 2011 年版较
119 为相似，包括临床前 AD、前驱期 AD 和 AD 痴呆阶段。

120 此外，采用神经心理学测评可以进一步区分认知功能障
121 碍患者的疾病严重程度，有利于针对性地纳入受试者。常用
122 的评估量表包括：临床痴呆评定量表（clinical dementia rating,
123 CDR）、简易精神状态检查（mini mental state examination,
124 MMSE）和总体衰退量表（global deteriorate scale, GDS）等。

125 （三）生物标志物的作用和类型

126 AD 临床试验中的生物标志物可根据其潜在用途进行分
127 类：①诊断：用于明确诊断；②富集：用于选择人群；③预
128 后：用于确定病程；④预测：用于预测未来对治疗的临床应
129 答和安全性评估；⑤药效学：用于确定预期或非预期的药效
130 学活性。

131 在 AD 治疗药物的临床研发中，用于上述特定目的的大
132 部分生物标志物尚需进行验证。鼓励开发和验证可靠、灵敏

133 的新生物标志物。

134 目前，脑脊液（CSF）标志物以及 MRI 和 PET 影像标
135 志物已经可以用于研究人群的富集。淀粉样蛋白 PET 和 CSF
136 A β 42 可测量淀粉样蛋白生物学的不同方面：PET 用于检测
137 A β 纤维聚集体；CSF A β 42 用于检测可溶性 A β 42 单体水平，
138 仅与斑块间接相关。对于试验富集，CSF 与 PET 淀粉样蛋白
139 生物标志物是高度相关的，但尚不清楚相关程度，这取决于
140 检测类型和阈值。这些检测不同潜在生物学阶段的方法在多
141 大程度上能够作为可相互替换的富集措施，应通过数据来证
142 明，以确保选择的是一个同质化人群。尽管 CSF A β 42 检测
143 的性能已显著提高，但建议不局限于测量 A β 42。A β 42 和 Tau
144 比值相较于单独使用 A β 42 可能具有更好的预测性。与 A β 相
145 关生物标志物类似，还可测量总 Tau 蛋白（t-Tau）或磷酸化
146 Tau 蛋白（p-Tau）水平、Tau 蛋白 PET 等。

147 APOE 是与 AD 发病风险相关的易感基因，约 58% 的 AD
148 患者携带至少 1 个 APOE ϵ 4 等位基因。APOE ϵ 4 状态可作为
149 临床试验人群富集指标之一，但如果仅纳入具有该特定基
150 因型的患者而不包括任何非携带者数据，则必须证明该药物
151 普遍适用是合理的。此外，应注意 APOE ϵ 4 对 AD 诊断不具
152 有足够特异性，携带 APOE ϵ 4 不一定会患病。

153 脑区域结构和代谢变化的下游局部生物标志物（例如，
154 通过 MRI 评估的海马萎缩，通过 FDG PET 评估的皮质低代

155 谢等), 虽然病理学特异性不足, 但对于检测和定量分析疾病
156 进展可能有一定的价值。

157 在识别疾病进展可能更快的患者方面, 目前尚无一种特
158 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
159 退过程可能会受到认知储备、合并疾病、生活方式和认知训
160 练等因素的影响, 应予以关注, 并提高临床试验质量控制。

161 三、临床药理学研究

162 (一) 药代动力学

163 应结合相关指导原则开展全面的药代动力学研究。

164 除旨在确定药物的吸收、分布、代谢和消除的常规药代
165 动力学研究外, 群体药代动力学模型也可能有助于模拟以老
166 年人为主的人群中的药物浓度。对于某些作用机制的假设,
167 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
168 试验结局是重要的。必要时, 还应在肝脏和/或肾脏功能不全
169 患者中进行药代动力学研究。

170 (二) 药效学

171 在 AD 药物开发早期, 确定药物可能具有治疗活性的药
172 理学机制很重要。表征药物的主要药效动力学活性(例如,
173 对于受体/神经递质途径的活性, 对淀粉样蛋白级联的活性,
174 对 Tau 聚集的活性, 对神经炎症的活性等)将影响后续的临床
175 研发计划。健康受试者中, 不良反应和药理学活性的可能
176 替代标志物(如果可用且相关)有助于估计适当的剂量范围。

177 (三) 药物相互作用

178 药代动力学相互作用研究应结合相关指导原则开展。

179 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，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，否则应进行药代动力学相互作用研究。

183 另外，应考虑到 AD 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。

187 四、探索性临床试验

188 随着 AD 治疗药物研究领域的快速发展，出现了大量的新靶点、新机制创新药物，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。

194 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。

199 此类试验的持续时间将取决于预期可测量的应答时间
200 或待评估的参数。某些生物标志物的研究已取得显著进展，
201 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
202 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，
203 为后续开展确证性临床试验提供更多参考。

204 五、确证性临床试验

205 （一）不同治疗目标药物的试验设计

206 1. 症状治疗

207 对于症状治疗药物，应通过随机、双盲、平行对照试验
208 证明在认知、日常功能或临床总体评估方面的治疗效应，该
209 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
210 临床意义获益的患者比例（应答率）来证明。如拟证明应答
211 率方面的治疗效应，应预先明确定义“应答”，例如，可定义
212 为认知指标改善至预先指定的程度，同时在其他两个领域
213 （日常功能和总体评估）至少没有恶化。考虑到试验过程中
214 疾病的自然进程，需要谨慎制定应答的标准。

215 轻度至重度痴呆 AD 患者目前已有多种症状治疗药物可
216 供选择，故推荐采用阳性对照、优效性设计的双臂试验。如
217 果不定位于优于阳性对照药，为了研究的敏感性，以及确定
218 新药在症状治疗药物中的定位，推荐的试验设计是试验药、
219 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
220 安慰剂的伦理问题，可以接受不均衡比例的随机化。由于对

221 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
222 物对比的非劣效性设计作为有效性的关键证据。

223 如果新药计划在标准对症治疗（例如胆碱酯酶抑制剂等）
224 的基础上使用，则双臂、安慰剂对照、加载（add-on）设计是
225 可以接受的。

226 对于痴呆前的早期阶段 AD 患者（前驱期 AD/AD 源性
227 MCI 等），在保障受试者安全的前提下，可考虑选择安慰剂
228 作为对照。

229 研究持续时间通常应足以评估药物的治疗效应。AD 症
230 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
231 度至中度 AD 患者的对照临床试验通常需持续至少 6 个月。

232 建议在双盲期后进行至少 6 个月的长期安全性随访，可
233 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
234 持续时间，定期进行阶段性的有效性和安全性评价。在治疗
235 结束后，应在适当时限内对患者进行随访，以观察与停药相
236 关的可能不良事件。

237 2. 延缓疾病进展治疗

238 当药物治疗延缓了疾病的病理学或病理生理学进展时，
239 可被认为是延缓疾病进展的治疗药物。延缓疾病进展的作用，
240 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
241 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
242 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

243 键病理生理学特征。

244 首选随机、双盲、对照、优效性试验确定有效性，在保
245 障受试者安全的前提下，对照药可以选择安慰剂或延缓疾病
246 进展治疗一线药物。胆碱酯酶抑制剂或美金刚等常作为标准
247 的对症治疗药物，特别是对于轻度至中度 AD，应在随机化
248 时对这些药物的使用进行分层。

249 研究持续时间与研究人群和预期的治疗效应等相关，最
250 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
251 度。例如，对于轻度至中度 AD 患者和痴呆前的早期阶段 AD
252 患者（前驱期 AD/AD 源性 MCI、数字临床分期的第 3 阶段
253 等），通常需持续至少 18 个月，某些情况下可能需要进行更
254 长时间的研究。根据产品的作用机制，治疗干预的时机可能
255 对结局至关重要。如果在前驱期 AD/AD 源性 MCI 等早期阶
256 段患者中证实了药物的延缓疾病进展治疗作用，其有效性难
257 以外推至病程后期（中度或重度痴呆）。

258 对于延缓疾病进展治疗药物，通过对各治疗组之间临床
259 症状随时间变化的速度进行统计比较（斜率分析）是合适的。
260 但应考虑到，尽管已知 AD 的自然病程可能随时间推移呈近
261 似线性模型，但尚不清楚该线性特点在具有干预治疗效应
262 （潜在的延缓疾病进展作用）的临床试验中是否成立，以及
263 治疗效应在治疗过程中是否保持不变。此外，随时间推移而
264 增加的药理学可逆效应也可能影响线性特点。因此，应定期

265 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
266 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
267 时间而减慢。使用的模型和检验假设应是合理的。此类研究
268 可通过延迟开始治疗的设计来加强，目的是证实延迟开始治
269 疗患者与更早期开始治疗患者之间在临床指标上的差异在
270 整个研究期间可维持。使用延迟开始治疗的设计，随访时间
271 长度和分析参数非常重要，如果随访时间过短，可能难以显
272 示出预期的疗效变化曲线。

273 也可通过时间-事件统计分析方法来评估可能的延缓疾
274 病进展效应。在疾病较早阶段，症状轻微，疗效随时间的变
275 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
276 床意义终点的时间。预设事件必须具有明确的临床重要性
277 （例如，至痴呆时间），通常是基于可接受的主要疗效终点评
278 估工具，而不是简单的根据某个维度量表评分的下降来定义
279 （例如，ADAS-cog 下降 2 分）。预期患者发生该事件之前的
280 时间必须足够长。

281 以上描述的确确定延缓疾病进展效应的方法有其缺点，并
282 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
283 改善、总体脱落率和缺失数据差异、治疗依从性较差、治疗
284 效应随病程的变化、终点指标随时间推移的敏感性等。因此，
285 应在研究方案中论证主要分析选择、统计模型、基本假设和
286 要求等的科学性。

287 减缓或延迟临床衰退的证据，应与生物标志物显示的延
288 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
289 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
290 在神经病理学过程相关的延缓疾病进展效应，故通过上述创
291 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
292 可接受的替代研发目标。

293 延缓疾病进展治疗的联合用药：AD 的病理生理学可能
294 涉及多种通路（可能是多因素的），因此，联合使用具有互补
295 作用机制的延缓疾病进展治疗药物可能具有重要的治疗作
296 用。如果拟联合使用两种延缓疾病进展治疗药物进行研究，
297 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
298 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
299 试验，并在必要时与安慰剂也进行比较。但是，由于在长期
300 研究期间每组所需的样本量大，因此延缓疾病进展治疗可能
301 难以实现完全的析因设计。不纳入单药治疗组需要有科学依
302 据。此类情形应具体问题具体分析，建议与监管机构进行沟
303 通。

304 3. 疾病预防的开发

305 痴呆一级预防的总体目标是降低目标人群中的发病率，
306 二级预防的目标是防止处于临床前状态的阶段进展到后期
307 症状更加明显的阶段。

308 可根据基因标记（例如 APOE ϵ 4 状态、常染色体显性突

309 变等)、生物标志物(例如 CSF A β 和 tau 水平、PET 中淀粉
310 样蛋白或 tau 示踪剂的滞留等)或环境风险因素(例如血管
311 或代谢因素)等,富集预防试验的人群。

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

314 针对症状前阶段 AD 潜在病理生理机制的药物干预可能
315 是预防策略的合理方法。应在富集人群中进行随机、双盲、
316 安慰剂对照试验。但是,如此早期阶段的疾病模型仍有待验
317 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
318 度的影响。AD 病理累积和临床症状发生的时间进程尚未确
319 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
320 异(认知储备),甚至每个特定患者在每天也会有变化。由
321 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
322 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。

323 预防试验患者在基线时没有可评估的临床损害,且患者
324 症状的出现时间具有不同的潜伏期,因此在典型持续时间
325 (如 2 年或更短)的临床试验中很难证明具有临床意义的获
326 益,通常需要大样本量和长期随访(通常至少 3 年)。目前
327 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
328 撑,如申请人拟开展预防试验,建议与监管机构进行沟通。

329 (二) 有效性评价指标

330 有效性指标的选择需关注患者(或看护者)功能状态和

331 生活质量等，可采用临床结局评估（ clinical outcome
332 assessment, COA）工具进行有效性评价。

333 创新药物临床研发时通常需考虑药物的作用机制和治
334 疗目标，设计相应的适宜疗效指标去证实是否能够达到预期
335 的治疗目标，但对于 AD 治疗药物，不同临床分期患者的临
336 床特征和治疗学需求存在较大差异，因此，在考虑药物作用
337 机制和治疗目标的基础上，应基于研究人群的不同临床分期
338 考虑合理的有效性评价指标。

339 1.AD 痴呆阶段

340 AD 痴呆阶段患者的主要临床表现为认知功能障碍伴日
341 常生活能力下降等，对于此阶段患者，可基于以下三个维度
342 的指标评估患者的临床获益：

343 （1）认知终点：通过客观测试进行测量；

344 （2）日常功能终点：日常生活能力；

345 （3）总体终点（总体临床疗效）：通过总体评价来体现。

346 上述三个维度中的每个维度应设置合理的有效性变量。

347 考虑到不同痴呆程度的 AD 患者临床特征和治疗学需求
348 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
349 区别。在轻度至中度 AD 痴呆患者中，对认知的影响应具有
350 临床意义，同时应将日常功能指标或临床总体评价指标作为
351 共同主要终点来确认疗效的临床意义。在重度 AD 痴呆患者
352 中，认知表现变化的临床获益相关性可能较低，且更加难以

353 量化，因此，日常功能和临床总体评价维度指标可能更适合
354 作为共同主要终点来确定该人群中具有临床意义的改善。

355 AD 痴呆阶段关注的次要疗效终点可包括健康相关生命
356 质量量表以及行为和精神病学症状量表等。在痴呆晚期，行
357 为问题对患者和照护者有较大影响。

358 2. 介于 AD 痴呆和临床前 AD 的中间阶段

359 在痴呆前的早期阶段 AD 患者（例如前驱期 AD/AD 源
360 性 MCI、数字临床分期第 3 阶段等），由于当前可用量表的
361 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
362 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
363 义。这也适用于在一项研究中同时研究痴呆前早期阶段 AD
364 患者和轻度痴呆 AD 患者。

365 目前使用的认知量表具有天花板效应，其敏感度可能不
366 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
367 也难以在大型临床试验中实施。此外，接近痴呆阶段的患者
368 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
369 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，
370 且日常功能损害的进展可能非常缓慢，使用现有量表可能
371 会产生可行性问题（研究的样本量和把握度）。因此，针
372 对早期阶段 AD 患者构建更敏感的评分条目，或筛选那些已
373 证明在早期阶段 AD 患者中持续受损的维度，并开发出可靠、
374 灵敏的工具，可能是今后的发展方向。

375 对于痴呆前的早期阶段 AD 患者，采用同时评价认知和
376 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
377 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
378 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
379 作为次要疗效指标。此外，可考虑将工具性活动能力、执行
380 功能和健康相关生命质量指标等也作为次要疗效指标，以支
381 持有效性的全面评估。

382 3.临床前 AD 阶段

383 对于临床前 AD 阶段患者，干预目标多为疾病预防。已
384 在一级预防和二级预防试验中使用的认知终点包括痴呆的
385 诊断（基于阈值评分）、认知功能显著衰退、基于某些测试
386 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
387 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，
388 包括试验时长和脱落率等。时间-事件分析中的“事件”应具有
389 明确的临床意义，例如发生认知受损等。在某个生物标志物
390 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
391 足够时间的随访，以采集相关的认知变化。

392 某些药物可能定位于疾病治疗，但目前尚无评估临床前
393 AD 患者治疗效应的“金标准”指标。对于特别早期阶段 AD 患
394 者，例如数字临床分期的第 2 阶段，仅在敏感的神经心理测
395 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
396 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

397 效果，可行性方面存在较大挑战。对于在敏感的神经心理测
398 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
399 战性，但是，认知功能对于 AD 患者的日常功能具有重要影
400 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
401 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
402 致趋势可增强研究结果的说服力。基于上述认识，通过敏感
403 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
404 这种认知维度的疗效可以得到 AD 特征性病理生理显著变化
405 的支持，其总体临床获益可能是可接受的。此类情形建议与
406 监管机构沟通。

407 (三) 有效性评估的工具

408 1. 一般考虑

409 鼓励开发和验证可靠、灵敏的工具来测量认知、日常功
410 能等，特别是在疾病早期阶段。可考虑利用现有的文献和数
411 据资源，开展适当的患者体验数据研究，结合研究目的、目
412 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，
413 开发和验证事宜的有效性评估工具。

414 认知、日常功能和总体评估指标涵盖了评估 AD 患者有
415 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
416 具(包括通用和疾病特异性的评价工具)、痴呆的行为和精神
417 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，
418 具体参见“生物标志物的作用和类型”章节。

419 由于目前没有理想的单独量表工具可全面评估药物的
420 治疗效应，申请人可能需要使用多种量表工具来评估 AD 治
421 疗药物的有效性。虽然已经出现了许多评价认知和日常功能
422 变化的评价工具，但目前仍没有可满足上述不同临床分期阶
423 段所有要求的标准评价工具，因此，对评估工具的选择仍保
424 持开放态度，但需提供合理的依据。

425 在使用评估工具时，建议由对治疗分配设盲的评估者执
426 行。如果存在可使研究者知晓盲底的特殊情况（如严重不良
427 事件等），则应尽可能使所有临床疗效评估者保持盲态。评估
428 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
429 进行培训，从而最大程度的减少评估变异性，以及最大程度
430 提高评估者之间的可靠性。理想情况下，不同维度的评估者
431 培训应标准化，以减少评估变异性。

432 目前在重度痴呆患者中进行的研究相对较少，在该类患
433 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
434 行全面评估，且应更加侧重于日常功能。

435 可以通过选择或删除现有量表中的单个条目，或使用单
436 个子测试的特定加权，或同时进行这两种操作，开发更敏感
437 的量表用于早期阶段 AD 患者。如果开发同时评价认知和日
438 常功能两个维度的综合评估量表，应量化单个条目和维度。
439 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
440 过拟定的工具充分体现这些目标。无论采用哪种方法，新开

441 发的评估工具必须证明能够测量临床获益。

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

450 2.有效性评估工具示例

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

454 认知功能测试

455 阿尔茨海默病评价量表的认知子量表（ADAS-cog）已广
456 泛应用于轻度至中度 AD 研究，但由于天花板效应和地板效
457 应，原始版本对疾病早期和晚期的变化的灵敏度有限，后续
458 对该量表的条目进行了部分改良。该量表有 12 项条目
459 （ADAS-Cog12），覆盖记忆、语言、定向、视空间结构、
460 观念运用、注意力等评测。通过对项目的增减先后形成了
461 ADAS-Cog11（删除了注意力条目）、ADAS-Cog13（ADAS
462 -Cog11 基础上增加了词语延迟回忆和数字划削测验）、

463 ADAS-Cog14(ADAS-Cog13 版本基础上增加了迷宫测验),
464 不同版本在 AD 治疗药物临床试验中均有应用。成套神经心理
465 理学测试 (neuropsychological test battery, NTB) 由 9 个认
466 知测验组成, 涉及记忆、注意、执行及语言功能等, 在轻度
467 到中度 AD 人群中可能具有较好的测量性能。

468 日常功能

469 已有几种评价日常功能的量表, 可测试与躯体活动相关
470 的基本日常生活能力 (例如, 如厕、活动、穿衣、洗澡等),
471 以及工具性日常生活能力 (例如购物、烹饪、洗衣服、处理
472 财务、使用交通工具、驾驶和打电话等)。但是, 这些量表
473 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
474 中可能更为适用的其他能力 (如使用科技产品的能力), 导
475 致一些测试量表的敏感性较低。阿尔茨海默病协同研究 ADL
476 量表 (ADCS-ADL) 是一种广泛应用于轻度至中度 AD 患者
477 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
478 应, 考虑是否应用于其他临床分期患者。目前尚没有一种工
479 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
480 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
481 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
482 目。

483 总体疗效评价

484 总体疗效评价是指由具有经验的临床医师对患者整体

485 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
486 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
487 临床获益的一种有用手段。

488 总体疗效评价量表允许临床医师对患者的症状和行为
489 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
490 合量表或成套测试来评价不同维度的功能。临床医师访视印
491 象变化量表(CIBIC-plus)可以对照基线对痴呆患者的总体临
492 床状况进行评价，根据对患者及其护理者的半结构化访视中
493 获得的信息来评价，但没有考虑任何来源的认知表现。阿尔
494 茨海默病协同研究组织的医生临床变化总体印象量表
495 (ADCS-CGIC)是另一种根据患者及其护理者的总体信息的
496 半结构化访视。与CIBIC-plus相比较,ADCS-CGIC更详细，
497 关注于包括认知、行为、社会 and 日常能力在内的15个方面。

498 认知和日常功能复合评价

499 临床痴呆评定量表-总分(clinical dementia rating-sum of
500 box, CDR-SB)是一项复合评价量表，由3项认知领域和3
501 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
502 度的变化。

503 六、安全性评估

504 AD治疗药物通常应遵循新药临床安全性评价的一般要
505 求(可参考《新药临床安全性评价技术指导原则》等)，以
506 及ICH E1对药物暴露量和暴露时间的要求。

507 对于不良事件，通常需描述治疗持续时间、药物使用剂
508 量、恢复时间，尤其是不同年龄组（如老年组和超高龄组患
509 者）以及其他相关变量的特征。应特别注意药物作用机制或
510 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
511 胆碱作用等；对于 β -淀粉样蛋白裂解酶抑制剂，通常需要通过
512 MRI 监测淀粉样蛋白相关性影像异常（ARIA），例如伴
513 有水肿的 ARIA（ARIA-E）和伴有含铁血黄素沉积的 ARIA
514 （ARIA-H）。

515 因 AD 患者有认知功能障碍，常不能准确和客观地表达
516 不适症状，因此需要临床研究者和照料者进行细致观察，必
517 要时应增加随访频率，以便及时发现和处理。在发现 AE 时，
518 应注意不要引入偏倚，询问 AE 的首选方法是对受试者进行
519 开放性和非引导性口头提问。

520 在短期临床试验之后，建议进行至少 6 个月的治疗随访。
521 可采用开放标签设计，考察长期安全性以及疗效的维持情况。

522 （一）神经系统不良事件

523 需要特别关注神经系统不良事件的发生或加重，特别是
524 脑血管事件、锥体外系症状、定向障碍、步态进一步受损、
525 癫痫发作等。根据作用机制和作用靶点信息，对于可能会发
526 生的特定神经系统不良事件，需要进行特殊的监测。靶向作
527 用于 β -淀粉样蛋白的单克隆抗体治疗已显示出可引起不同
528 程度和发生频率的 ARIA，具体取决于药物的活性、靶点、

529 剂量和患者特征（APOE ε4 状态或其他）等。这些不良事件
530 的临床意义尚需进一步研究，在临床试验期间应予以特别关
531 注，定期进行 MRI 监测，并收集相关的安全性信息。此外，
532 通常还需监测停药的影响。

533 （二）精神系统不良事件

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

537 药物过量和自杀：根据作用机制、预期治疗方案，可能
538 需要对药物过量的风险和影响进行研究。应使用经验证的量
539 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
540 估，例如 InterSePT 自杀想法量表、哥伦比亚自杀严重程度评
541 定量表（C-SSRS）或其他经验证的量表等。应报告自杀事件
542 的发生率（包括自杀意念到自杀身亡），并分析相关因素的
543 影响，包括剂量、治疗持续时间及其他促成因素等。应提供
544 自杀患者陈述或行为的叙述性总结。

545 （三）心血管系统不良事件

546 根据药物的药效学特征，监测其对心血管系统的影响，
547 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
548 肌梗死的风险等。

549 （四）长期安全性

550 考虑到阿尔茨海默病患者需要长期药物治疗，故通常需

551 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
552 参考 ICHE1 相关要求）。不同年龄组（年轻、年老和高龄）
553 的长期安全性特征可能不同，应予以考虑。早期疾病阶段的
554 患者人群应特别关注长期安全性，这些患者可能在无症状阶
555 段接受治疗多年，某些不良反应可能是明显的。

556 应长期监测对死亡率的影响，特别是对于无症状阶段的
557 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。

558 七、与监管机构沟通

559 鼓励申请人在药物研发的关键阶段，基于药物特点和前
560 期研究数据，就临床研发计划、临床试验方案设计的关键问
561 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
562 通交流。

563 八、参考文献

564 [1] EMA, Committee For Medicinal Products For Human
565 Use (CHMP). Clinical investigation of medicines for the
566 treatment of Alzheimer’s disease[EB/OL]. 2018. [https://www.
567 ema.europa.eu/en/clinical-investigation-medicines-treatment-
568 alzheimers-disease](https://www.ema.europa.eu/en/clinical-investigation-medicines-treatment-alzheimers-disease) [published online ahead of print].

569 [2] FDA. Early Alzheimer’s disease: developing drugs for
570 treatment, guidance for industry[EB/OL](2024-03-12). [https://
571 www.fda.gov/regulatory-information/search-fda-guidance-
572 documents/early-alzheimers-disease-developing-drugs-treatment.](https://www.fda.gov/regulatory-information/search-fda-guidance-documents/early-alzheimers-disease-developing-drugs-treatment)

573 [3] McKhann GM, Knopman DS, Chertkow H, et al. The

574 diagnosis of dementia due to Alzheimer's disease:
575 recommendations from the national institute on aging- Alzheimer'
576 s association workgroups on diagnostic guidelines for
577 Alzheimer's disease[J]. *Alzheimers Dement*, 2011, 7(3): 263-269.

578 [4] Jack CR Jr, Bennett DA, Blennow K, et al. NIA - AA
579 research framework: toward a biological definition of
580 Alzheimer's disease [J]. *Alzheimers Dement*, 2018, 14(4): 535-
581 562.

582 [5] Jack CR Jr, Andrews JS, Beach TG, et al. Revised criteria
583 for diagnosis and staging of Alzheimer's disease: Alzheimer's
584 Association Workgroup[J]. *Alzheimers Dement*. 2024 Jun 27. doi:
585 10.1002/alz.13859. Online ahead of print.

586 [6] Dubois B, Feldman HH, Jacova C, et al. Advancing
587 research diagnostic criteria for Alzheimer's disease: the IWG-2
588 criteria[J]. *Lancet Neurol*. 2014 Jun;13(6):614-29. doi:
589 10.1016/S1474-4422(14)70090-0.

590 [7] 田金洲, 解恒革, 王鲁宁, 等. 中国阿尔茨海默病痴
591 呆诊疗指南(2020 年版). *中华老年医学杂志*[J], 2021, 40(3):
592 269-283.

593 [8] 中华医学会神经病学分会痴呆与认知障碍学组. 阿
594 尔茨海默病源性轻度认知障碍诊疗中国专家共识 2024. *中华
595 神经科杂志*[J], 2024, 57(7): 715-737.

596 [9] 中国医师协会神经内科医师分会, 阿尔茨海默病药物
597 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.

598 中华脑科疾病与康复杂志(电子版) [J], 2022, 12(1): 9-20.

599 [10] Gauthier S, Rosa-Neto P, Morais JA, et al. World
600 Alzheimer Report 2021: Journey through the diagnosis of
601 dementia. London: Alzheimer's Disease International, 2021.
602 <https://www.alzint.org/u/World-Alzheimer-Report-2021.pdf>.