

革新的医薬品・医療機器・再生医療製品実用化促進事業（厚生労働省）
アルツハイマー病治療薬の臨床評価基準策定のためのレギュラトリーサイエンス研究

アルツハイマー病治療薬の臨床評価及び臨床開発における留意点と課題について

東京大学医学部附属病院

平成 29 年 3 月

[趣旨]

平成 24 年度から厚生労働省の「革新的医薬品・医療機器・再生医療製品実用化促進事業」が開始された。本事業において、東京大学医学部附属病院は、「アルツハイマー病（AD）治療薬の臨床評価方法に関するガイドライン」の策定を目標とし、医薬品医療機器総合機構（PMDA）との連携の下、バイオマーカーを活用した AD 治療薬の臨床評価基準策定のための調査研究及びモデリング・シミュレーションを活用した AD の臨床効果予測モデルの構築のための研究を進めている。

本書は、本事業の一環として、今般、PMDA の協力を得て、現在開発が増加している疾患修飾を目的としたアルツハイマー病治療薬の臨床評価及び臨床開発に関して、現時点での考え方・留意すべき事項と今後の課題をとりまとめたものである。

目次

I. 緒言	3
II. 臨床試験の選択基準.....	4
1. AD dementia	4
2. MCI due to AD/ Prodromal AD	5
3. バイオマーカーの利用について.....	6
III. 臨床試験に用いる有効性評価項目	7
1. AD dementia	8
2. MCI due to AD/ Prodromal AD	9
3. バイオマーカーの利用について.....	10
IV. 各開発段階において実施する臨床試験に関する留意事項.....	10
1. 第 I 相試験	10
2. 第 II 相試験	11
3. 第 III 相試験（検証的試験）	11
4. 第 III 相試験（長期投与試験）	12
5. 臨床データパッケージ.....	12
V. 今後の検討課題	14
参考文献	15

I. 緒言

アルツハイマー病（以下、AD）は、緩徐に認知機能障害が進行し自立した日常生活が困難となるため、患者本人はもとより介護者にも著しい負担を来し、少子高齢化社会を迎えたわが国においては、社会経済学的な観点からも極めて深刻かつ喫緊の対策を要する疾患となっている。医師を対象とした各種疾患の治療満足度と治療における薬剤の貢献度を1994年から定期的に調査したヒューマンサイエンス振興財団による報告によると、ADについては一貫して、満足度、貢献度ともに下位に位置している。ADに対して現在、コリンエステラーゼ阻害剤や *N*-methyl-D-aspartate (NMDA) 受容体阻害剤等の症状改善薬¹が臨床使用されているが、疾患修飾薬²は国内外いずれにおいても承認されておらず、世界各国で開発が進められている。ADの病態生理変化は認知症を発症する何年も前から始まっているという推定に基づき、疾患修飾薬の開発においては、認知症患者のみならず、病態生理変化は始まっているが認知症が発症していない段階からの早期治療介入が提唱されている。これらの患者を対象とした臨床試験を実施するにあたっては、病態進行の各ステージに適した被験者選択基準及び評価項目をどのように設定するかが重要な課題であるが、特に臨床症状が軽度である早期段階においては、臨床症状を中心とした従来の方法では適切な被験者選択や有効性評価が困難であるといった課題があり、ADの病態生理変化を反映するバイオマーカーの活用の期待が高まっている。また、臨床試験の大規模化、長期化に伴い国際共同治験を利用した開発も増加していることも踏まえ、適切な臨床評価及び臨床開発のための基準の整備が必要となっている。

このような状況を踏まえ、本報告書においてADの疾患修飾を目的とした治療薬の臨床試験を実施する上で留意すべき事項及び解決すべき課題を取りまとめた。バイオマーカーの利用や認知症発症前段階における有効性評価方法に関しては、さらなるエビデンスの集積が必要であり、現時点では、臨床試験の実施に際して未だ多くの課題が残されている。これらの課題に対し、今後集積されるエビデンスをもとに検討を重ねた上で、AD治療薬の臨床評価方法を確立する必要がある。

本文書における定義

¹ 症状改善薬：ADの臨床症状を緩和する効果はあるが、病態の進行は抑制しない薬剤

² 疾患修飾薬：ADの病態機序に作用することで神経変性や神経細胞死を遅延させ、その結果として臨床症状の進行を抑制する薬剤

II. 臨床試験の選択基準

疾患修飾薬の臨床試験においては、有効性及び安全性評価に適した共通の病態を有する均質な集団を選択できるよう、適切な選択基準を設定する必要がある。現在疾患修飾薬の開発が主に行われている AD の病態生理変化を背景とする認知症（AD dementia）及び軽度認知障害（Mild Cognitive Impairment（以下、MCI） due to AD / prodromal AD）について、選択基準の設定における留意事項を以下に示す。

1. AD dementia

AD の臨床診断基準としては、米国精神医学会による精神疾患の診断・統計マニュアル改訂第 4 版（Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision : DSM-IV-TR）及び米国立神経疾患・脳卒中研究所とアルツハイマー病・関連障害協会（National Institute of Neurological and Communicative Disorders and Stroke AD and Related Disorders Association : NINCDS-ADRDA）による診断基準が広く用いられてきた。しかし、これらの診断基準は他の原因による認知症との鑑別があいまいであること、疾患の根底にある病態生理変化を考慮していないこと等を踏まえ、NINCDS-ADRDA による診断基準は 2011 年に米国立老化研究所（National Institute of Aging : NIA）とアルツハイマー病協会（Alzheimer's Association : AA）（以下、NIA-AA）により改訂がなされ、中核臨床基準に加え、AD の病態生理変化を反映するバイオマーカーを取り入れた AD dementia の診断基準が作成された。当該診断基準において、AD に特徴的なバイオマーカーとして、Amyloid beta（以下、A β ）の脳内への沈着を反映する脳脊髄液（cerebrospinal fluid、以下 CSF）中の A β ₄₂ の低下、陽電子放出断層撮影（positron emission tomography、以下 PET）によるアミロイドイメージング、神経変性を反映するバイオマーカーとして、CSF 中のタウやリン酸化タウの上昇、MRI (magnetic resonance imaging) における内側側頭葉・海馬の萎縮、FDG-PET (fluorodeoxy glucose-positron emission tomography) における側頭頭頂葉・楔前部の代謝低下等が示されている。また、2014 年に公表された Dubois らによる International Working Group (IWG)-2 criteria では、AD の診断には CSF 中の A β ₄₂ 低下およびタウまたはリン酸化タウの上昇、PET によるアミロイドイメージング、または家族性 AD の原因となる遺伝子変異（PSEN1、PSEN2 または APP）のいずれかが必要とされている。

A β やタウ等の AD の病態生理変化に関与する分子を標的とした疾患修飾薬の臨床試験を実施するには、AD 以外の原因による認知症患者を適切に除外する必要がある。そのためには AD の病態生理変化を反映するバイオマーカーを利用することが重要である。現時点ではいずれの診断基準・バイオマーカーを用いるべきかコンセンサスは得られていないものの、上記の A β 沈着や神経変性を反映するバイオマーカーを選択基準に取り入れることを検討する必要がある。ただし、現時点ではバイオマーカーの選択基準への利用に際しては、「II.3. バイオマーカーの利用について」で示すように様々な課題がある。

治験に組み入れる被験者の重症度は、複数の検査に基づき統合的に評価すべきである。具

体的な検査・評価指標としては、例えば Mini-Mental State Exam (MMSE)、Alzheimer's Disease Assessment Scale - Cognitive Subscale (ADAS-cog)、Clinical Dementia Rating (CDR) 等が挙げられるが、より適切と考えられる指標があれば他の指標を用いることも可能である。

2. MCI due to AD/ Prodromal AD

疾患修飾薬の開発においては、認知症発症前の早期段階からの治療介入が有用である可能性が提唱されており、日常生活は自立しており認知症の基準は満たさないものの、認知機能の持続的な低下を有する軽度認知障害の被験者を対象とした臨床試験が複数の薬剤で実施されている。認知症発症前の AD を表す用語として「MCI due to AD」や「Prodromal AD」が用いられており、それぞれ NIA-AA、Dubois らにより AD の病態生理変化を有する患者を特定できるようバイオマーカーを取り入れた診断基準が公表されている。「MCI due to AD」と「Prodromal AD」のいずれの基準を用いるべきかについてコンセンサスは得られておらず、現時点でいずれかの用語・基準の使用を推奨するものではない。MCI due to AD/ Prodromal AD (以下、MCI/pAD) を対象とした薬剤の評価にあたっては、AD の病態の有無及び認知症発症リスクを考慮して薬剤を投与すべき患者を適切に診断できること、臨床試験成績から薬剤の臨床的意義が十分に示されることが必要と考えられるが、以下に記載するように診断基準及び有効性の評価方法について更なる検討が必要である。

疾患修飾薬の臨床試験においては、AD の病態を有する患者を適切に選択して組み入れる必要があるが、臨床症状のみに基づく診断では、AD 以外の原因による認知機能障害を適切に除外することは困難である。また、認知症へ進行するリスクは患者毎に異なり、 $A\beta$ の沈着等の AD に特徴的な病態生理変化を有しても、生涯認知症を発症しない患者も存在する。薬剤投与には必ずリスクが伴うことを考慮すると、AD を原因とする認知機能障害があり、かつ高い確率で認知症を発症すると予測される患者のみを被験者とするのが適切であり、このような患者を組み入れるために、適切な認知機能検査、 $A\beta$ の沈着及び神経変性を反映するバイオマーカー等を利用して選択基準を設定する必要がある。

しかしながら、AD への進行を正確に予測するバイオマーカーや認知機能検査の組合せ、バイオマーカーの陽性・陰性の判断基準、カットオフ値等については現在研究が進められているところであり、現時点では利用に際して様々な課題があることから (II.3. バイオマーカーの利用について)、バイオマーカーを利用した臨床試験デザインの立案に際しては、最新の知見を踏まえて検討する必要がある。

MCI/pAD の被験者の重症度評価において、従来 AD dementia に対して用いられてきた MMSE や ADAS-cog 等は天井効果により認知機能障害が検出できない可能性があり、MCI/pAD の重症度評価に適した指標を検討する必要がある。

また、現時点では、MCI/pAD の臨床診断基準が十分に確立されていない状況にあるため、

臨床試験において設定された選択基準が将来的な診断基準と異なる可能性がある。その場合には、実施された臨床試験成績に基づき、審査時点における最新の診断基準に合致する患者、または想定される投与対象における有効性・安全性が十分に説明できる必要があり、将来的な診断基準との違いの程度によっては、当該試験成績が承認申請の主要な根拠としては不十分と判断される可能性もある。

3. バイオマーカーの利用について

AD dementia 又は MCI/pAD の選択基準におけるバイオマーカーの利用においては、以下の点に留意する必要がある。

- ・ CSF バイオマーカー（CSF 中の A β 、タウ等、以下同様）及び画像バイオマーカー（アミロイドイメージング、MRI、FDG-PET 等）について、測定方法や撮像機器の違いによる測定結果の違いや、同一の方法を用いた場合にも測定施設間・評価者間で結果のばらつきがあることが指摘されており、標準的な評価法が確立していない。現時点では、評価のばらつきを抑えるために、治験に用いるバイオマーカーの採取や保存、撮像等は施設間で共通の手順を定め、中央測定・中央評価を行うことが望ましい。また、CSF バイオマーカーは将来標準的な測定方法が確立した場合に再評価できるよう、検体の一部を保管しておくことが有用と考える。現時点では ADNI 等の大規模研究で用いられている測定方法に準拠することを推奨するが、適切な根拠が示されれば、他の測定方法を用いることも可能である。
- ・ CSF バイオマーカーとアミロイドイメージングを併用する等、複数のバイオマーカーを組み合わせて用いる場合には、各バイオマーカー間の相関関係や差異に基づき、一方のバイオマーカーが陽性で、もう一方が陰性となった症例をどのように取り扱うか検討する必要がある。複数のバイオマーカーのうち、全てに適合した場合のみ臨床試験に組み入れるのではなく、いずれかでも陽性であれば組み入れる選択基準とする場合は、異なるバイオマーカーの基準に基づき組み入れられた集団で患者背景・治験薬の有効性及び安全性に差異が認められないか検討できるよう、それぞれのバイオマーカーの陽性被験者について一定の症例数を確保することが望ましい。また、選択基準に用いないバイオマーカーについても、事後的な部分集団解析等にも利用できるよう、組入れ時から可能な限り情報を得ておくことが有用と考えられる。患者背景としてのバイオマーカーの情報は、有効性及び安全性の個人差及び民族差を考察する際にも、重要な情報となり得る。
- ・ バイオマーカーの測定値にカットオフ値を設けて選択基準とする場合は、臨床試験の対象（病期・人種等）や測定方法を考慮し、類似した臨床試験や研究報告の情報に基づき妥当な値を検討する必要がある。一方で、日本人を対象とした研究報告は限られており、バイオマーカーに関する国内外差の有無は明らかではない。外国人を中心としたバイオマーカー研究に基づいて計画された国際共同治験に日本から参画する場合は、利

用予定のバイオマーカーの民族差の有無について可能な限り事前に検討する必要がある。

- ・ 臨床試験の被験者選択にバイオマーカーを使用する場合、製造販売後も適切な投与対象を選択するために当該バイオマーカーによる事前の確認が必要となる可能性があり、臨床試験で用いたバイオマーカーが実臨床においても同様に利用できるかを検討する必要がある。例えばアミロイドメージングを選択基準に用いた場合、薬剤承認時に PET リガンドやその合成装置も承認されていなければ、臨床現場において薬剤の適切な投与対象が選択できないことが懸念される。したがって、バイオマーカーを治験の選択基準に用いる薬剤の開発に際しては、当該バイオマーカーの測定方法の開発状況も考慮する必要がある。また、アミロイドメージングは利用可能な施設が限られ、検査費用が高額となる可能性があること、CSF バイオマーカーは実施に侵襲性を伴うこと等が試験の実施可能性及び製造販売後の適切な投与対象の選択に影響する可能性も考慮しておく必要がある。ただし、疾患の科学的理解を深め、より有効な治療法の検討のため、現状では臨床現場では利用し難いバイオマーカーを治験の中で測定することは、被験者の負担が許容できるものであれば否定されるものではない。
- ・ 特定のバイオマーカーの陽性例のみを実臨床での投与対象と想定する場合、陽性例に投与を限定することの妥当性、又は当該バイオマーカーのカットオフ値の妥当性等を検討するために、陰性例も探索的に臨床試験に組み入れ、有効性・安全性を評価することが望ましい。ただし、バイオマーカー陰性例に対して臨床試験において当該医薬品の有効性が示される可能性が極めて低いと考えられる場合や安全性の懸念がある場合はこの限りではない。
- ・ 治験の選択基準に用いたバイオマーカーの確認が、実臨床での投与対象の選択においても必須とされるかは、バイオマーカーの用途（疾患の診断、予後の予測、治療効果の予測等）、薬剤のリスク/ベネフィット及び審査時の医療環境に応じて判断される。特定のバイオマーカーの陰性例では有効性が期待できず、かつ安全性のリスクが高い薬剤においては、投与前のバイオマーカーの確認が必須となる可能性もある。
- ・ APOE ε4 は AD の発症及び進行のリスク遺伝子型として知られており、薬剤によっては APOE ε4 保因者において有害事象の発現リスクが非保因者に比べて増加することも示唆されているため、部分集団解析等に利用できるよう、APOE ε2 も含め APOE の遺伝子型の情報を得ておくことが望ましい。ただし、被験者への情報開示の有無を含め、遺伝情報やバイオマーカーに関する情報の取扱いには十分留意する必要がある。

以上の点を踏まえ、臨床試験においていずれのバイオマーカーを用いるか検討する必要がある。

III. 臨床試験に用いる有効性評価項目

AD dementia 及び MCI/pAD を対象とした臨床試験における、有効性評価項目の設定に関する留意事項を以下に示す。

1. AD dementia

AD dementia を対象とした薬剤の有効性評価方法として、米国の認知症治療薬の臨床評価方法に関するガイドライン案（Guidelines for the clinical evaluation of antidementia drugs, first draft. November 8, 1990. FDA）においては、AD の中核症状である認知機能及び AD の全般臨床症状の 2 つを有効性の主要評価項目として評価することが必要とされている。また、欧州の認知症治療薬の臨床評価方法に関するガイドライン案（Draft guideline on the clinical investigation of medicines for the treatment of Alzheimer's disease and other dementias. EMEA Guideline London 28 Jan 2016 EMA/CHMP/539931/2014）では、認知機能及び日常生活動作を主要評価項目とし、全般臨床症状は副次評価項目として評価することが必要とされている。

AD dementia に対する有効性を評価するために、認知機能評価に加え、日常生活動作又は全般臨床症状の評価を行うことは世界的なコンセンサスとなっており、本邦においても、AD dementia を対象とした薬剤の検証的試験については、認知機能に加え、日常生活動作又は全般臨床評価の 2 つの主要評価項目を設定し、有効性を示す必要がある。なお、探索的な目的で、限られた試験規模・試験期間で実施する試験においては、認知機能評価、日常生活動作又は全般臨床評価のいずれか 1 つを主要評価項目とし、その他を副次評価項目とすることも可能と考えられる。

それぞれの評価項目で用いる評価指標は、妥当性及び信頼性が確認され、AD dementia の臨床症状・重症度を反映し、薬効を検出する上で十分な感度を有するものを用いる。具体的には、海外で作成された評価指標の日本語版を作成して用いる場合には、言語や文化の差異を考慮し、検査内容や難易度の等価性に十分配慮する必要がある。AD dementia の治験では、これまで一般に、ADAS-cog・Severe Impairment Battery (SIB)（認知機能）、Alzheimer's Disease Cooperative Study - Activities of Daily Living Inventory (ADCS-ADL)・Disability Assessment for Dementia (DAD)（日常生活動作）、the Clinician's Interview-Based Impression of Change plus caregiver input (CIBIC-plus)・Alzheimer's disease cooperative study- clinical global impression of change (ADCS-CGIC)・Clinical Dementia Rating – Sum of Boxes (CDR-SB)（全般臨床評価）等の指標が用いられてきた。一方で、軽度 AD dementia の患者ではこれらの評価指標では天井効果が生じる可能性があり、対象患者の重症度に応じて、適切な主要評価項目を検討する必要がある。

評価者の習熟度や情報提供者が有する情報量の不足が評価に影響を及ぼす可能性が考えられるため、事前に十分なトレーニングを行ったり、情報提供者が被験者の情報を十分に得られるような基準を設ける等の対策が必要である。

また、副次評価項目として可能な限り複数の評価指標を用いて試験を行い、薬剤の有効性を様々な観点から評価できるようにしておくことが望ましい。

2. MCI due to AD/ prodromal AD

現時点では、MCI/pAD における有効性評価方法として、確立した評価項目はない。以下に、現時点で利用可能と考えられる評価項目を示すが、いずれの評価項目を用いる場合でも、認知症発症前の段階から早期に治療介入を行うことの意義が明確に示される必要がある。

MCI/pAD を対象とした臨床試験において、「AD dementia の発症」（発症までの期間又は発症割合）を主要評価項目として用いることは適切である。しかしながら、AD において認知機能や日常生活の障害は緩徐に進行するため、AD dementia の発症時点の判断には評価者間でばらつきが生じることが懸念される。AD dementia の発症を判定する評価者に対しては事前に十分なトレーニングを行う等、評価者間で均質な判断がなされる方策をとる必要がある。また、症状の経過の詳細、神経心理検査の結果（採点に評価者の主観が含まれる場合は、判断の根拠となった情報・説明を含む）等の情報を収集した上で、複数の専門家により中央評価を行い、判定の妥当性を確認することが望ましい。さらに、AD dementia の発症までの期間を有効性評価項目とする場合には、評価・検査頻度の規定が有効性の結果に影響を及ぼす可能性があるため、評価・検査頻度に関する適切な基準を設定する必要がある。

また、MCI/pAD における有効性評価方法として、米国の早期 AD に対する臨床評価ガイドライン案（Guidance for Industry, Alzheimer's Disease: Developing Drugs for the Treatment of Early Stage Disease, Draft Guidance, February 2013, FDA）では、CDR-SB のような認知機能及び日常生活の障害の双方を包括した指標を単独の主要評価項目として用いることが適切であるとしている。また、欧州の臨床評価ガイドライン案（Draft guideline on the clinical trial investigation of medicines for the treatment of Alzheimer's disease and other dementias. EMEA Guideline London 28 Jan 2016 EMA/CHMP/539931/2014）においても、認知機能と日常生活に対する影響をあわせて評価できるような複合評価指標を単独の主要評価項目として用いることの可能性について言及されている。MCI/pAD においては、日常生活や全般臨床症状の障害は軽度であり、これらを AD dementia を対象とした臨床試験と同様に主要評価項目とすることは困難であることから、本邦においても、認知機能及び日常生活の障害を包括的に評価可能な指標を用いて MCI/pAD における単独の主要評価項目として設定することは可能と考える。また、CDR-SB のみならず、MCI/pAD の評価に適した新規の評価項目を新たに開発して用いることも可能と考えるが、AD の病態進行に伴うスコアの推移や既存の評価項目及び AD dementia の発症との関係等を踏まえてスコアの変化の臨床的意義が説明可能でなければならない。臨床的意義を説明する際には、適切な縦断的研究のデータ等も踏まえて、AD の病態進行に伴うスコアの変化を検討することが有用と考える。また、副次評価項目として、AD dementia の発症までの期間・割合や既存の評価項目を評価し、これらの指標でも有効性が示されることが望ましい。

MCI/pAD の段階では認知機能及び日常生活の障害が軽度であることから、これらの臨床

的な検査指標においてプラセボに対し統計的に有意な変化が示されるのみでなく、副次評価項目として神経変性を反映すると考えられるバイオマーカーを評価すること等により、疾患の進行に確かに影響していることを示すことが、薬効を説明する上で有用な情報になると考える（「Ⅲ.3. バイオマーカーの利用について」参照）。

3. バイオマーカーの利用について

ADNI等の観察研究において、自然経過におけるADの臨床症状とバイオマーカー変化の関係が検討されており、バイオマーカーの変化を有効性評価に用いることで、臨床症状のみを評価するよりも小規模な試験で治療効果を評価できる可能性が報告されている。臨床試験においては、探索的試験や検証的試験における副次評価項目として可能な限りバイオマーカーの評価を行い、薬剤が目的とする標的に作用していることを確認するとともに、薬剤の臨床効果とバイオマーカー変化の関係について検討することが望ましい。また、国際的な開発戦略を取る場合は、国内外いずれにおいても開発の早期段階からバイオマーカーのデータを取得し、有効性、安全性及び用法・用量における民族差の検討に利用することが望ましい。

一方で、薬剤の介入に基づくバイオマーカー変化と臨床効果の関係は未だ明確にされておらず、薬剤介入によってどのバイオマーカーがどの程度変化すれば、ADの病態進行を確かに抑制し、臨床症状の改善につながるかは不明である。したがって、臨床試験においていずれのバイオマーカーを有効性の指標とすれば、ADの病態進行を抑制していることを適切に示していると言えるのか現時点では明確ではなく、いずれのバイオマーカーについても、検証的試験において臨床症状の評価に代わるサロゲートエンドポイントとして用いることは現状では適切とは言えない。有効性の指標としてのバイオマーカーの利用については、ADNI等の観察研究や薬剤の介入試験のデータを踏まえ、更なる検討が必要である。以上より、現時点では探索的な評価として、可能な限りバイオマーカーを評価することが推奨される。

IV. 各開発段階において実施する臨床試験に関する留意事項

1. 第I相試験

第I相試験は、非臨床試験で得られた情報をもとに、治験薬をヒトに投与する臨床開発の最初の段階であり、忍容性及び安全性、並びに薬物動態を検討する。また、バイオマーカーを用いた薬力学的検討も行うことが望ましい。

原則として、健康成人を対象とし、AD患者は高齢者が多いため、高齢者における忍容性及び安全性、並びに薬物動態の検討を行うことも考慮する。ただし、Aβ、タウ等を標的とした免疫療法における第I相試験においては、各被験薬の安全性プロファイルも踏まえ被験者の安全を十分に確保した上で、適応対象とするAD患者で実施することも可能である。

2. 第Ⅱ相試験

非臨床試験及び第Ⅰ相試験の結果を評価して、患者を対象とした第Ⅱ相試験を開始する。第Ⅱ相試験は、患者における治験薬の有効性・安全性を探索的に検討する前期第Ⅱ相試験と、第Ⅲ相試験の実施に際して用量反応関係を明らかにし、治験薬の用法・用量を決定するための後期第Ⅱ相試験に分けることもできる。

用量反応関係を検討する後期第Ⅱ相試験の試験デザインは、通常、プラセボ対照無作為化二重盲検並行群間比較試験とし、プラセボに加え、少なくとも治験薬を2用量以上で行うことが推奨される。

第Ⅱ相試験における主要評価項目は、「Ⅲ.有効性評価項目」を参考に設定する。なお、前期第Ⅱ相試験として用量を探索的に検討する目的で、バイオマーカーを指標とした薬力学的効果と薬物動態の関係を検討することは可能と考える。ただし、現時点では臨床症状の評価に代わるサロゲートエンドポイントとして用いることができるバイオマーカーはないため、バイオマーカー変化のみに基づき適切な用量を設定することは困難である。したがって、第Ⅲ相試験に用いる用量を設定するための後期第Ⅱ相試験においては、原則として、臨床症状に対する用量反応性の検討が必要である。

なお、海外で第Ⅱ相試験が先行して実施され、臨床症状に対する用量反応性の検討がなされている場合も、現時点では臨床症状への影響を適切に推測できるバイオマーカーが確立しておらず、薬物動態やバイオマーカーの変化に基づく薬力学的効果の国内外の比較により日本人における臨床的有効性を推測することは困難と考えるため、原則として、日本人を対象とした（国内又は国際共同）第Ⅱ相試験（用量設定試験）を実施し、国内外の臨床症状に基づく用量反応性の異同を検討した後に第Ⅲ相試験（検証的試験）に進むべきである。

3. 第Ⅲ相試験（検証的試験）

第Ⅱ相試験によって安全性と期待される有効性が確認され、臨床的な推奨用量が推定された薬物について、治験薬の有効性を検証することを目的として第Ⅲ相試験（検証的試験）を行う。試験デザインは、プラセボ対照二重盲検無作為化並行群間比較試験とすることが適切である。なお、第Ⅱ相試験の結果から臨床推奨用量を1つの用法・用量に絞ることができない場合等、必要に応じて第Ⅲ相試験において複数の用法・用量を設定することも可能である。主要評価項目は、「Ⅲ.有効性評価項目」を参考に設定する。

疾患修飾作用について検討を行う場合、無作為化並行群間比較試験において臨床症状の改善効果を示すことに加え、バイオマーカーを用いてADの病態生理学的な進行の抑制を評価することが適切である。一方、FDAによる早期ADに対する臨床評価ガイドライン（案）では、臨床試験デザインとして randomized-start design や randomized-withdrawal design を用いることで、疾患修飾作用を示す方策も提案されている。ただし、当該試験デザインを用いる場合には、評価方法や評価期間等について個別に検討が必要であるため、事前にPMDAと相談することが推奨される。

国際共同治験として実施する場合の日本人において必要な成績や臨床データパッケージについては「IV. 5. 臨床データパッケージ」参照。

4. 第Ⅲ相試験（長期投与試験）

AD 治療薬は一般に長期間投与されるため、治験薬を臨床推奨用量で 1 年間以上投与し、長期投与時の安全性を検討する試験が必要である。症例数は「致命的でない疾患に対し長期間の投与が想定される新医薬品の治験段階において安全性を評価するために必要な症例数と投与期間について」（平成 7 年 5 月 24 日付、薬審第 592 号）を参考に計画すべきである。なお、投与期間が 1 年以上に及び、臨床推奨用法・用量で投与された症例が十分に存在する臨床試験がある場合は、別途長期安全性を検討するための試験は必須ではない。

AD dementia 及び MCI/pAD をいずれも対象とする場合、基本的にはそれぞれの病期で、長期投与時の安全性を十分に評価できる症例数を設定することが望ましい。両者を合わせて長期投与時の安全性評価に必要なデータを充足させる計画とする場合には、MCI/pAD と AD dementia の臨床試験のデータを相互補完的に利用することの妥当性の説明が求められる。（「5. (2) MCI/pAD」参照）

長期投与試験の主たる目的は、臨床推奨用法・用量で長期間投与したときの安全性及び有効性を検討することであるため、長期投与試験は臨床推奨用法・用量が推定できた段階で、当該用法・用量を用いて実施することが適切である。

5. 臨床データパッケージ

(1) AD dementia

原則として、複数の無作為化並行群間比較試験成績に基づき、日本人において再現性を以て治験薬の有効性を示す必要がある。また、AD dementia の臨床評価の検証的試験においては、「Ⅲ. 有効性評価項目」で述べたように 2 つの主要評価項目を設定して有効性を示す必要がある。国内単独開発、ブリッジング戦略、及び国際共同治験への参画等の複数の方法がある状況で、上記の条件をどのように満たすかを整理することは、開発戦略を検討する上で重要となるため、代表的な開発手法における臨床データパッケージの構築の事例を以下にまとめた。ただし、必要なデータパッケージについては最新の知見の集積状況、薬剤の特徴等も踏まえて、個々に検討されるべきである。

1) 国内単独開発

第Ⅱ相試験で設定された主要評価項目において治験薬のプラセボに対する優越性が示された上で、第Ⅲ相試験において「Ⅲ. 有効性評価項目」で示した 2 つの主要評価項目のいずれにおいても治験薬のプラセボに対する優越性が示される必要がある。

2) ブリッジング戦略

海外で有効性の検証試験が別途実施されており、海外検証試験成績の外挿を目的に日本でブリッジング開発を行う場合は、適切なブリッジング対象試験と比較可能なブリッジング試験を行い、国内外における有効性、安全性及び用量反応性の類似性を示す必要がある。民族間の類似性は、バイオマーカーの解析結果を含めて多角的に検討することが有用であるが、現時点では、原則として臨床症状の評価による用量反応性の類似性を示す必要がある。海外におけるブリッジング対象試験は、複数用量で用量反応性が検討され、「Ⅲ. 有効性評価項目」で示した 2 つの主要評価項目においてプラセボに対する優越性が示された試験とする。日本でのブリッジング試験においても、ブリッジング対象試験と同じ 2 つの主要評価項目においてプラセボに対する優越性が示され、用量反応性の類似性が確認される等、予め設定したブリッジングの成立条件が満たされる必要がある。

3) 国際共同治験への参画

臨床開発が世界的に進行している場合、早期の開発段階から日本が国際共同治験に参加し、その結果を踏まえて、第Ⅲ相試験を国際共同治験として実施することの妥当性を確認した上で、国際共同第Ⅲ相試験を実施することが望ましい。国際共同治験を利用した開発において重要な考え方を以下に示す。

- ① 国際共同第Ⅲ相試験の全体集団の成績で、「Ⅲ. 有効性評価項目」で示した 2 つの主要評価項目のいずれにおいてもプラセボに対する治験薬の優越性が示された上で、2 つの主要評価項目のいずれにおいても日本人集団と全体集団の成績の一貫性が示される必要がある。また、バイオマーカーについても可能な限り副次評価項目として評価し、全体集団の成績との一貫性を検討する。
- ② 上記①が、日本が参加する 2 つ以上の国際共同第Ⅲ相試験により再現性を以て示される必要がある。ただし、国際共同第Ⅲ相試験が 1 つであっても、日本が参加する国際共同第Ⅱ相試験や国内臨床試験において明確な有効性が示されている（適切な主要評価項目において治験薬のプラセボに対する優越性が示される等）場合には、それを再現性の根拠とすることができる可能性もある。
- ③ 日本が 1 つの国際共同第Ⅲ相試験にのみ参加した場合、当該試験において上記①が示されるのみでなく、各主要評価項目の治験薬とプラセボとの群間差等を踏まえても明らかな臨床的意義が示されていると判断でき、かつ当該国際共同第Ⅲ相試験において、日本人集団でも 1 つ以上の主要評価項目において治験薬のプラセボに対する優越性が示される等、治験薬の有効性が明確かつ頑健に示された場合、再現性の根拠として別途日本人を対象とした臨床試験の実施が必要となるか否かは検討の余地があると考えらる。

(2) MCI due to AD/ Prodromal AD

原則として AD dementia と同様に複数の無作為化並行群間比較試験成績に基づき再現性を以て治験薬の有効性を示す必要があるが、MCI/ pAD を対象とした臨床試験以外に、AD

dementia を対象とした臨床試験を計画している場合、MCI/ pAD と AD dementia の臨床試験のデータを相互補完的に利用することの妥当性（各試験の対象集団が共に AD の病態生理変化を有し、MCI/pAD として組み入れられた被験者が高い確率で AD dementia に進行することが十分な根拠に基づき説明できるか、また臨床推奨用量、治験薬の安全性のプロファイル、及びリスク/ベネフィットのバランスが各試験の対象集団で同様と判断できるか等）が説明可能であれば、MCI/pAD を対象とした臨床試験成績に加えて、AD dementia を対象とした検証的試験等の成績を治験薬の有効性の再現性を支持するデータとして利用できる可能性がある。

MCI/pAD のみを対象とした臨床開発の場合、検証的試験が大規模かつ長期間となることが想定されることも考慮すると、例えば、複数の用量群が設定された 1 つの検証的試験で臨床的に意義のある頑健な有効性（AD dementia の発症抑制等）が示され、かつ用量反応性が示されれば、必ずしも複数試験において有効性を検証する必要がない場合も想定される。

V. 今後の検討課題

AD の病態生理変化を有するが認知機能が正常、あるいはごく軽度の認知機能低下のみが存在する段階を Preclinical AD として、早期治療介入の対象とする考え方が提唱されている。世界的に Preclinical AD を対象とした臨床試験も開始されており、神経細胞の障害が進行する前に治療介入することで、より効果的に病態の進行を抑制できる可能性がある。一方で、現時点では臨床で使用し得る Preclinical AD の診断基準は定まっておらず、MCI/pAD や AD dementia への進行リスクに関して更なるデータの蓄積が必要である。

Preclinical AD に対する薬剤の評価にあたっては、疾患の進行リスクが高い患者を適切に選択できることが重要であり、長期投与時の安全性についても慎重に検討する必要がある。また、認知症の発症抑制効果を示すデータが必要であるが、臨床症状の変化が小さいこと、薬効評価に長期間を要すること等、治験を実施するにあたっての課題は多い。わずかな認知機能変化を検出できる評価指標と、臨床的な治療効果を予測可能なバイオマーカーの確立が望まれる。

参考文献

アルツハイマー病に対する疾患修飾薬の臨床評価に際して参考とすべき文献など

Albert MS et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*, 7:270-9, 2011

Dubois B et al. Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. *Lancet Neurol*, 13:614-629, 2014

Dubois B et al. Revising the definition of Alzheimer's disease: a new lexicon. *Lancet Neuro*, 9:1118-27, 2010

Dubois B et al. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *Lancet Neurol*, 6:734-46, 2007

McKhann GM 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. *Alzheimers Dement*, 7:263-9, 2011

Sperling R et al. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 7:280-92, 2011

Sperling R et al. Amyloid-related imaging abnormalities in patients with Alzheimer's disease treated with bapineuzumab: a retrospective analysis. *Lancet Neurol*, 11:241-9, 2012

Weiner MW et al. The Alzheimer's Disease Neuroimaging Initiative: a review of papers published since its inception. *Alzheimers Dement*, 8:S1-68,2012

「平成 22 年度国内基盤技術調査報告書－2020 年の医療ニーズの展望－」（財）ヒューマンサイエンス振興財団

ガイドライン：

<欧米>

Committee for Medicinal Products for Human Use (CHMP), European Medicines Agency

Draft guideline on the clinical investigation of medicines for the treatment of Alzheimer's disease and other dementias. EMEA Guideline London 28 Jan 2016 EMA/CHMP/539931/2014

http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2016/02/WC500200830.pdf

U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER)

Guidance for Industry Alzheimer's Disease: Developing Drugs for the Treatment of Early Stage Disease. February 2013

<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM338287.pdf>

<本邦>

ICHガイドライン (International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use、日米EU医薬品規制調和会議)

E1 : The Extent of Population Exposure to Assess Clinical Safety for Drugs Intended for Long-Term Treatment of Non-Life-Threatening Conditions

致命的でない疾患に対し長期間の投与が想定される新医薬品の治験段階において安全性を評価するために必要な症例数と投与期間について (平成 7 年 5 月 24 日付薬審第 592 号厚生省薬務局審査課長通知)

E5 : Ethnic Factors in the Acceptability of Foreign Clinical Data

外国で実施された医薬品の臨床試験データの取扱いについて (平成 10 年 8 月 11 日付医薬発第 739 号厚生省医薬安全局長通知)

外国臨床データを受け入れる際に考慮すべき民族的要因について (平成 10 年 8 月 11 日付医薬審第 672 号厚生省医薬安全局審査管理課長通知)

E5 Ethnic Factors : Questions and Answers

「外国臨床データを受け入れる際に考慮すべき民族的要因についての指針」に関する Q&A について (平成 16 年 2 月 25 日付厚生労働省医薬食品局審査管理課事務連絡)

「外国臨床データを受け入れる際に考慮すべき民族的要因についての指針」に関する Q&A について (その 2) (平成 18 年 10 月 5 日付厚生労働省医薬食品局審査管理課事務連絡)

E7 : Studies in Support of Special Populations: Geriatrics

高齢者に使用される医薬品の臨床評価法に関するガイドラインについて (平成5年12

月2日薬新薬第104号)

E7 Studies in Support of Special Populations: Geriatrics Questions & Answers

「高齢者に使用される医薬品の臨床評価法に関するガイドライン」に関する質疑応答集 (Q&A) (平成22年9月17日厚生労働省医薬食品局審査管理課事務連絡)

その他

国際共同治験に関する基本的考え方について (平成19年9月28日付薬食審査発第0928010号厚生労働省医薬食品局審査管理課長通知)

国際共同治験に関する基本的考え方 (参考事例) について (平成24年9月5日付厚生労働省医薬食品局審査管理課事務連絡)