

Clinical Trial Endpoints and Their Clinical Meaningfulness in Early Stages of Alzheimer's Disease

S. Cohen¹, J. Cummings², S. Knox³, M. Potashman⁴, J. Harrison⁵⁻⁷

1. Toronto Memory Program, Toronto, ON, Canada; 2. Chambers-Grundy Center for Transformative Neuroscience, Department of Brain Health, School of Integrated Health Sciences, University of Nevada, Las Vegas (UNLV), NV, USA; 3. Biogen International GmbH, Baar, Switzerland; 4. Biogen Inc., Cambridge, MA, USA; 5. Metis Cognition Ltd, Wiltshire, UK; 6. Alzheimer Center, AU Medical Center, Amsterdam, the Netherlands; 7. Institute of Psychiatry, Psychology & Neuroscience, King's College London, UK

Corresponding Author: Sharon Cohen, MD. Toronto Memory Program, Toronto, ON, Canada. Phone (+1) 416-386-9761; Fax (+1) 416-386-0458; Email: cohen@memorydisorders.ca

Abstract

As the focus of Alzheimer's disease (AD) therapeutic development shifts to the early stages of the disease, the clinical endpoints used in drug trials, and how these might translate into clinical practice, are of increasing importance. The clinical meaningfulness of trial outcome measures is often unclear, with a lack of conclusive evidence as to how these measures correlate to changes in disease progression and treatment response. Clarifying this would benefit all, including patients, care partners, primary care providers, regulators, and payers, and would enhance our understanding of the relationship between clinical trial endpoints and assessments used in everyday practice. At present, there is a wide range of assessment tools used in clinical trials for AD and substantial variability in measures selected as endpoints across these trials. The aim of this review is to summarize the most commonly used assessment tools for early stages of AD, describe their use in clinical trials and clinical practice, and discuss what might constitute clinically meaningful change in these measures in relation to disease progression and treatment response.

Key words: Early Alzheimer's disease, clinical meaningfulness, clinical trials, endpoints, assessment tools.

Abbreviations: 1^o: primary; 2^o: secondary; A β : amyloid beta; AD: Alzheimer's disease; ADAS-Cog: Alzheimer's Disease Assessment Scale—Cognitive Subscale; AD PACE: Alzheimer's Disease Patient and Caregiver Engagement; ADCOMS: Alzheimer's Disease Composite Score; ADCS-ADL: Alzheimer's Disease Cooperative Study—Activities of Daily Living Scale; ADCS-ADL-MCI: Alzheimer's Disease Cooperative Study—Activities of Daily Living Scale—Mild Cognitive Impairment; ADL: activities of daily living; A-IADL-Q: Amsterdam Instrumental Activities of Daily Living Questionnaire; APCC: Alzheimer's Prevention Initiative Preclinical Composite Cognitive Test; ApoE: apolipoprotein E; BADL: basic activities of daily living; BEHAVE-AD: Behavioral Pathology in Alzheimer's Disease Rating Scale; CANTAB: Cambridge Neuropsychological Test Automated Battery; CDR-SB: Clinical Dementia Rating—Sum of Boxes; CFC: Cognitive-Functional Composite; ChEI: cholinesterase inhibitor; CIBIC Plus: Clinician's Interview Based Impression of Change with caregiver input; DAD-6: Disability Assessment for Dementia; DMT: disease-modifying therapy; EMA: European Medicines Agency; EQ-5D: EuroQol 5-Dimension Questionnaire; FAQ: Functional Activities Questionnaire; FDA: Food and Drug Administration; GAS: Goal Attainment Setting; GDS: Geriatric Depression Scale; GPCOG: General Practitioner Assessment of Cognition; IADL: instrumental

activities of daily living; iADRS: integrated Alzheimer's Disease Rating Scale; iMAP: Insights to Model Alzheimer's Progression in Real Life; MCI: mild cognitive impairment; MCID: minimal clinically important difference; MID: minimal important difference; MIS: Memory Impairment Screen; MMSE: Mini-Mental State Examination; MoCA: Montreal Cognitive Assessment; NPI-C: Neuropsychiatric Inventory Questionnaire—Clinician version; NPI-Q: Neuropsychiatric Inventory Questionnaire; NTB: Neuropsychological Test Battery; QOL: quality of life; QOL-AD: Quality of Life in Alzheimer's Disease Scale; RBANS: Repeatable Battery for the Assessment of Neuropsychological Status.

Introduction

Alzheimer's disease (AD) is a neurodegenerative disorder accounting for ~60–80% of late-onset dementia cases worldwide (1). It represents an enormous disease burden for patients, care partners, healthcare institutions and society at large (1). The main clinical features of AD are progressive impairments of cognition and function and changes in behavior (2, 3). It has been postulated that the optimal opportunity to meaningfully impact the course of AD, reduce disease burden and preserve function, may be prior to substantial clinical symptom appearance when the underlying pathology is limited (4).

The current treatment development landscape in AD includes therapeutic candidates that target the underlying pathophysiology of AD in the hope of significantly altering the course of the disease (5). Many of these are being investigated as potential treatments for the early stages of AD; 12/29 Phase 3 trials (41.4%) and 37/73 Phase 2 trials (50.7%) involve patients with mild cognitive impairment (MCI) due to AD or prodromal-to-mild AD dementia (5). In these trials, certain clinical measures such as Clinical Dementia Rating—Sum of Boxes (CDR-SB), Alzheimer's Disease Assessment Scale—Cognitive Subscale (ADAS-Cog), and Alzheimer's Disease Cooperative Study—Activities of Daily Living Scale—Mild Cognitive Impairment (ADCS-ADL) are often used as

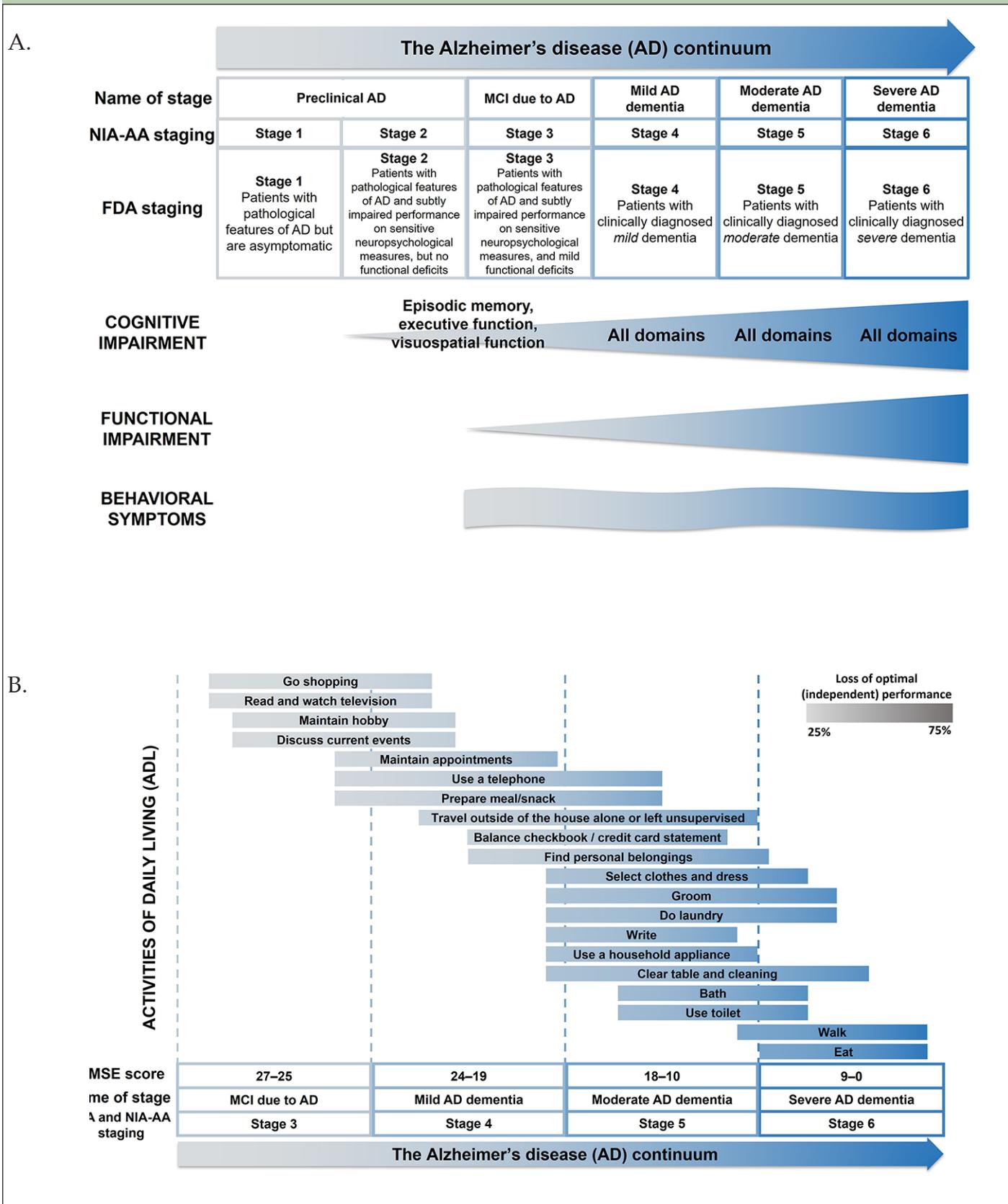
Table 1. Clinical features of AD: cognition, function, and behavior

Clinical feature	Domain	Description	Clinical presentation of domain in Alzheimer's disease
Core cognitive memory domains	Episodic memory	Recollection of past events and personal experiences (Harvey 2019)	<ul style="list-style-type: none"> • Forgetting recent personal events (Alzheimer's Association 2020) • Losing items around the home (Alzheimer's Association 2020) • Forgetting recent conversations (Alzheimer's Association 2020)
	Semantic memory	Knowledge of the world (facts, ideas, meanings, and constructs) and long-term storage of ideas and concepts (Harvey 2019)	<ul style="list-style-type: none"> • Difficulty naming an object (Alzheimer's Association 2020) • Difficulty recognizing an object (agnosia) (Alzheimer's Association 2020)
	Working memory*	A short-term storage module in which concepts are operated upon to solve problems (Harvey 2019)	<ul style="list-style-type: none"> • Impaired problem-solving skills (Alzheimer's Association 2020) • Losing "train of thought" (Alzheimer's Association 2020)
Other core cognitive domains	Executive functions	A set of processes used to control, regulate, and manage thoughts and behavior (Harvey 2019)	<ul style="list-style-type: none"> • Struggling to follow a series of steps (Alzheimer's Association 2020) • Difficulty performing a familiar task (Alzheimer's Association 2020) • Impaired judgment (Alzheimer's Association 2020)
	Language	Verbal expression, language interpretation, and communication (Harvey 2019)	<ul style="list-style-type: none"> • Difficulty finding words (anomic aphasia) (Alzheimer's Association 2020) • Struggling to follow a conversation (Alzheimer's Association 2020)
	Visuospatial functions	The integration of space and visual form (Harvey 2019)	<ul style="list-style-type: none"> • Impaired perception of objects and faces (Alzheimer's Association 2020) • Inability to navigate familiar surroundings (Alzheimer's Association 2020)
Other cognitive domains	Attention*	The capacity to focus and sustain concentration (Harvey 2019)	<ul style="list-style-type: none"> • Impaired concentration (Alzheimer's Association 2020) • Losing track of thoughts (Alzheimer's Association 2020)
	Praxis	Monitor and process data in the surrounding environment to plan a motor action (Ward et al. 2015)	<ul style="list-style-type: none"> • Inability to perform a movement with a body part (Ward et al. 2015) • Inability to imitate a hand gesture (Ward et al. 2015)
Function	Basic activities of daily living	Ability to perform basic activities that are strongly correlated with motor functioning and coordination (Mlinac and Feng 2016)	<ul style="list-style-type: none"> • Reduced ability to carry out self-maintenance such as bathing, dressing, and grooming (Alzheimer's Association 2020) • Reduced ability to carry out activities independently, such as walking and eating (Alzheimer's Association 2020)
	Instrumental activities of daily living	Ability to perform complex activities, requiring greater cognitive organization than basic activities of daily living (Mlinac and Feng 2016)	<ul style="list-style-type: none"> • Reduced ability to manage finances or medications (Alzheimer's Association 2020) • Reduced ability to shop, cook, conduct housework, or drive (Alzheimer's Association 2020)
Behavior		Social behavior and psychological status (Kales et al. 2015; Alzheimer's Association 2020)	<ul style="list-style-type: none"> • Neuropsychiatric symptoms such as aggression, anxiety, apathy, sleep disturbance, and psychosis (Kales et al. 2015; Alzheimer's Association 2020)

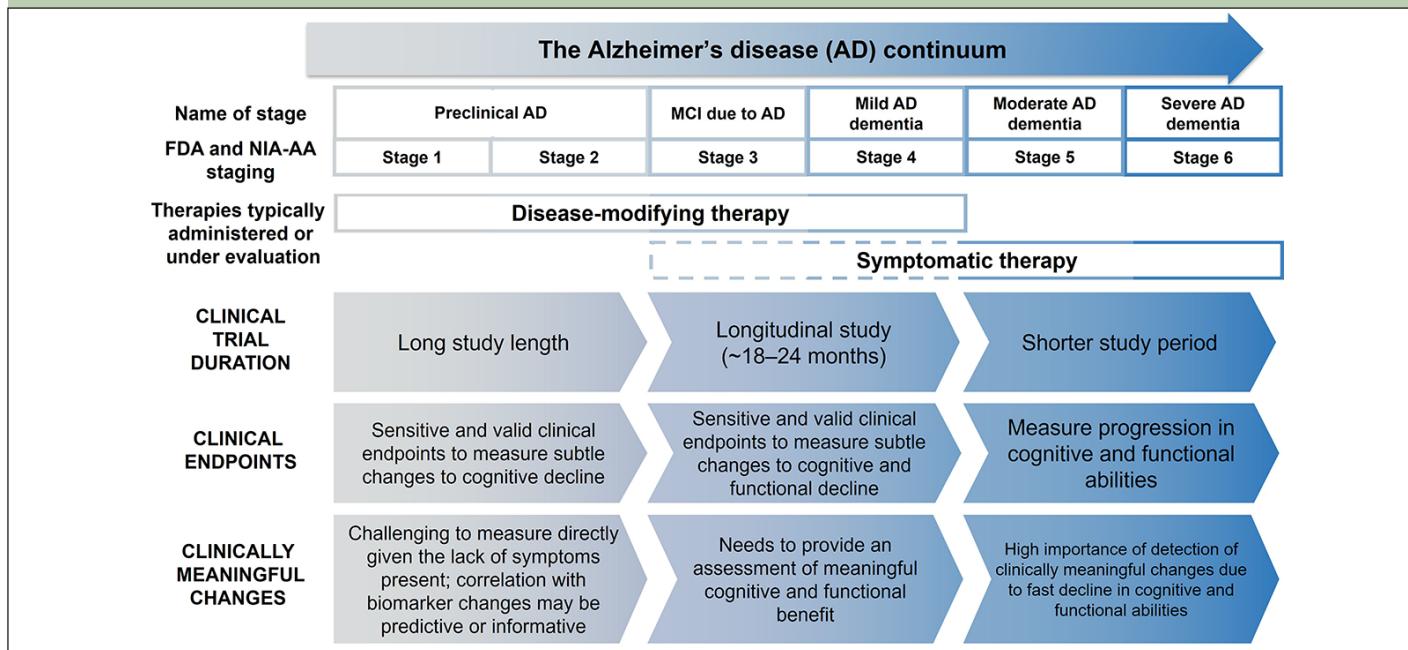
*Working memory and attention are executive functions

trial outcomes (5), but a variety of factors (including practical limitations such as long administration times and limited sensitivity to change in early stages of the disease) limit their use outside clinical trials in the diagnosis and disease monitoring of patients with AD. As new therapeutic options emerge for the treatment of AD, such as aducanumab, which was recently granted accelerated approval for the treatment of AD by the US Food and Drug Administration (FDA), it is important that the clinical meaningfulness of the endpoints used to evaluate these therapeutics in clinical trials is understood. Firstly, there is a need to evaluate the totality of evidence

from both clinical and biomarker effects to make an accurate estimate of effect on disease progression (6). It is critical to have a scientifically justified approach to assess the significance and persistence of the potential impact of a therapeutic along the disease course (6). Secondly, there needs to be a pragmatic focus on endpoint selection in clinical trials, considering the overlap of cognitive, functional, and behavioral changes, and the differences in sensitivity of these endpoints in detecting changes along the disease course, especially in the early disease stages (7-9). Finally, there is a need for translation of clinical trial outcome data to inform decision making in clinical

Figure 1. Cognitive, functional, and behavioral impairment across the AD continuum

A. There is progressive cognitive and functional decline, initially observed in Stages 2 and 3 of AD, as classified by the FDA in 2018 (U.S. Food and Drug Administration 2018). Behavioral impairment differs across the continuum, with variable appearance and resolution of behavioral symptoms. B. Different ADL are impaired differentially across the AD continuum, with each ADL being initially affected during different stages of AD. Adapted from Galasko 1998 with permission (Galasko 1998). Abbreviations: AD = Alzheimer's disease; ADL = activities of daily living; FDA = Food and Drug Administration; MCI = mild cognitive impairment; MMSE, Mini-Mental State Examination; NIA-AA = National Institute of Aging-Alzheimer's Association

Figure 2. Selection of clinical endpoints depends on the stage of Alzheimer's disease

Clinical trials for AD vary in length based on the population of patients they involve, as classified here by the FDA staging (U.S. Food and Drug Administration 2018). Selection of clinical endpoints for clinical trials differs dependent upon the clinical stage of AD, with measurement of clinically meaningful changes also having varying challenges across the AD continuum. Abbreviations: AD = Alzheimer's disease; FDA = Food and Drug Administration; MCI = mild cognitive impairment; NIA-AA = National Institute of Aging–Alzheimer's Association

practice, from the perspectives of healthcare practitioners, payers, patients, and care partners.

Here we review commonly used clinical trial endpoints and their applicability in clinical practice. We also examine clinically meaningful changes in the outcome measures in early stages of AD, focusing on those used in MCI due to AD and mild AD dementia.

Key clinical features assessed in AD clinical trials

The three main clinical features commonly assessed in AD clinical trials are cognition, function, and behavior (3). These can be assessed separately or through composite or global assessments (2, 3). Clinical presentation of AD decline can vary among individuals (Table 1) (10).

In the assessment of cognition, multiple domains are evaluated, including attention, memory, executive function, language, visuospatial processing, and praxis (Table 1), but there is limited consensus on the discreteness of these functions (11). Cognitive decline impacts all cognitive domains to a variable extent and is associated with increasing functional impairment over time (Table 1; Figure 1A, 1B) (11-14).

The assessment of function describes a patients' ability to perform activities normally associated with everyday life. Function can be assessed through the performance of basic activities of daily living (BADL), which involve activities required for personal self-maintenance (Figure 1B) (8), and instrumental activities of daily living (IADL),

which involve higher-level activities, such as making a meal or cleaning a room. IADL place greater demand on cognitive resources and therefore represent a more sensitive outcome to assess early functional loss than BADL (Figure 1B) (8). Notably, decline in IADL can be masked by compensatory mechanisms in very early stages of AD, presenting a challenge for accurate and complete clinical assessment (15).

Behavior in AD is assessed through presence and severity of neuropsychiatric symptoms (NPS) (Figure 1; Table 1) (7, 16, 17). NPS worsen with cognitive impairment (7) and the presence of NPS may predict further cognitive decline, decreased functioning, increased care partner burden, increased healthcare utilization and costs, earlier need for institutionalization, and increased risk of death (16-19). However, patient-to-patient variability exists in the occurrence and severity of NPS (17), so cognition can serve only as a rough proxy for progressive decline in function and behavior in individuals with AD. Evaluation of all three clinical features is key for comprehensive clinical assessment of disease stage, progression, and potential treatment response.

Key clinical outcomes of interest across the AD continuum

The updated National Institute of Aging–Alzheimer's Association (NIA-AA) research framework proposed to classify individuals along the continuum of AD, based

Table 2A. A selection of assessment tools commonly used only in clinical trials or in both clinical trials and clinical practice

Assessment type	Assessment tool	Sensitivity and specificity for detecting AD dementia (unless stated otherwise)	Key advantages	Key limitations	Use across AD continuum				Setting	
					Stage 1/2	Stage 3	Stage 4/5	Stage 6	Clinical trials (1°/2°)	Clinical practice
					Pre-clinical AD	MCI due to AD	Mild-to-moderate AD dementia	Severe AD dementia		
Cognitive	ADAS-Cog-13/14 (Black et al. 2009; Sheehan 2012)	Not available	• Covers all cognitive areas in dementia	• Ceiling effects • Length of assessment makes unsuitable for clinical practice					Y - 1°	N
	CANTAB (De Roeck et al. 2019)	Sensitivity: 68% Specificity: 98%	• Automated test delivery and data capture • Language independent	• Does not include language-based measures • Fee for use					Y - 1°/2°	N
	CogState (De Roeck et al. 2019)	Sensitivity: 100% Specificity: 85%	• Automated test delivery and data capture	-					Y - 1°	N
	MMSE (Sheehan 2012; Cordell et al. 2013; Patnode et al. 2020)	Sensitivity: 89% Specificity: 89%	• Minimal training requirements	• Floor and ceiling effects • Fee for use					Y - 2°	Y
	MoCA (Sheehan 2012; Patnode et al. 2020)	Sensitivity: 78–100% Specificity: 65–94%	• Minimal training requirements • Sensitive to early stages of AD	• Floor effects at moderate AD					Y - 2°	Y
	NTB (Black et al. 2009; Schneider and Goldberg 2020)	Not available	• High test-retest reliability • Uses pre-baseline exposure to tests to induce stability	-					Y - 1°	N
	RBANS (Duff et al. 2008; Black et al. 2009)	Sensitivity: 98% Specificity: 82%	• Minimal practice effects • Translated and validated in multiple languages	-					Y - 1°/2°	Y
Functional	ADCS-ADL (Black et al. 2009)	Not available	• Graded responses allow gradual transitions to be detected	• Reliant on reliable informant					Y - 1°/2°	Y
	ADCS-ADL-MCI (Pedrosa et al. 2010; Marshall et al. 2012)	Sensitivity for MCI: 87% Specificity for MCI: 87%	• Suitable for MCI	• Reliant on reliable informant					Y - 1°/2°	Y
	A-IADL-Q (Sikkens et al. 2013; Koster et al. 2015)	Sensitivity: 74% Specificity: 65%	• Sensitive to changes in IADL functioning over time	• Reliant on reliable informant					Y - 1°/2°	Y
	FAQ (Marshall et al. 2012; Patnode et al. 2020)	Sensitivity: 90% Specificity: 83%	• High reliability	• Reliant on reliable informant					Y - 2°	Y
Behavioral	BEHAVE-AD / Empirical-BEHAVE-AD (Sheehan 2012; Ismail et al. 2013)	Sensitivity: 79% Specificity: 73%	• Empirical-BEHAVE-AD created to avoid bias	• Length of assessment makes less suitable for clinical practice					Y	N
	NPI-Q (Sheehan 2012; Ismail et al. 2013)	Sensitivity: 86% Specificity: 76%	• Short • Reliable across different languages and cultures	• Reliant on reliable informant					Y - 2°	Y
Composite / Global	ADCOMS (Wang et al. 2016; Schneider and Goldberg 2020)	Sensitivity: 82%	• May enable small sample sizes to show drug-placebo difference	• No alternate forms • Psychometrics not yet known					Y - 1°	N
	iADRS (Schneider and Goldberg 2020)	Not available	• Sensitive to tracking disease progression	• Minimal use as prospective clinical endpoint					Y - 1°	N
	CFC (Jutten et al. 2020)	Not available	• Moderate-to-high test-retest reliability • Minimal practice effects	• No current evidence for utility in subjective cognitive decline					Y	N
	CDR-SB (Black et al. 2009; O'Bryant et al. 2010; Marshall et al. 2012; Sheehan 2012)	Sensitivity: 74% Specificity: 81%	• Short • Good inter-rater reliability	• Length of assessment makes less suitable for clinical practice • Ceiling effects in early stages of AD					Y - 1°/2°	N
	CIBIC Plus (Sheehan 2012)	Not available	• Comprehensive	• Training required • Reliant on reliable informant					Y - 1°	N

Table 2A (Continued). A selection of assessment tools commonly used only in clinical trials or in both clinical trials and clinical practice

Assessment type	Assessment tool	Sensitivity and specificity for detecting AD dementia (unless stated otherwise)	Key advantages	Key limitations	Use across AD continuum				Setting	
					Stage 1/2	Stage 3	Stage 4/5	Stage 6	Clinical trials (1°/2°)	Clinical practice
					Pre-clinical AD	MCI due to AD	Mild-to-moderate AD dementia	Severe AD dementia		
Patient-reported outcome	GAS (Black et al. 2009)	Not available	<ul style="list-style-type: none"> Individualized for patient; Can detect longitudinal changes; High importance to patients and caregivers 	Not widely favored for use in clinical practice	–	–	–	–	Y - 2°	N
Patient and caregiver QOL/burden	EQ-5D (Sheehan 2012)	–	<ul style="list-style-type: none"> Short Completed by patient or informant 	<ul style="list-style-type: none"> Reliant on reliable informant 	–	–	–	–	Y - 2°	Y
	QOL-AD (Sheehan 2012)	–	<ul style="list-style-type: none"> Extensively validated Disease specific Completed by patient and caregiver 	<ul style="list-style-type: none"> Informant version reliant on reliable informant 	–	–	–	–	Y - 2°	Y
	Zarit Burden Interview (Sheehan 2012)	–	<ul style="list-style-type: none"> Disease specific 	<ul style="list-style-type: none"> Length of assessment is restrictive 	–	–	–	–	Y - 2°	Y

Abbreviations: 1° = primary, 2° = secondary, AD = Alzheimer's disease, ADAS-Cog = Alzheimer's Disease Assessment Scale—Cognitive Subscale, ADCOMS = Alzheimer's Disease Composite Score, ADCS-ADL = Alzheimer's Disease Cooperative Study—Activities of Daily Living Scale, ADCS-ADL-MCI = Alzheimer's Disease Cooperative Study—Activities of Daily Living Scale—Mild Cognitive Impairment, A-IADL-Q = Amsterdam Instrumental Activities of Daily Living Questionnaire, BEHAVE-AD = Behavioral Pathology in Alzheimer's Disease, Rating Scale, CANTAB = Cambridge Neuropsychological Test Automated Battery, CDR-SB = Clinical Dementia Rating—Sum of Boxes, CFC = Cognitive-Functional Composite, CIBIC Plus = Clinician's Interview Based Impression of Change with caregiver input, EQ-5D = EuroQol 5-Dimension Questionnaire, FAQ = Functional Activities Questionnaire, GAS = Goal Attainment Setting, IADL = instrumental activities of daily living, iADRS = integrated Alzheimer's Disease Rating Scale, MCI = mild cognitive impairment, MMSE = Mini-Mental State Examination, MoCA = Montreal Cognitive Assessment, NPI-Q = Neuropsychiatric Inventory Questionnaire, NTB = Neuropsychological Test Battery, QOL = quality of life, QOL-AD = Quality of Life in Alzheimer's Disease Scale, RBANS = Repeatable Battery for the Assessment of Neuropsychological Status.

on the presence of biomarkers of AD neuropathology (4). Cognition, function, and behavior are affected differentially across the AD continuum (Figure 1A, 1B) (20). According to the latest guidance from the FDA, during Stage 1, there is no objective evidence of cognitive decline or functional impairment (20), although some individuals report subjective cognitive decline (20). During Stage 2, there is objective evidence of impairment observed in select cognitive domains, with preservation of IADL (13). Stages 3 and 4 represent MCI due to AD and mild AD dementia, respectively. During Stage 3, one or more of the cognitive domains are affected and there are difficulties with word-finding (a component of the language domain) (11, 13). During Stage 4, multiple cognitive domains are affected and there is increased cognitive and functional decline compared with Stage 3 (4). Notably, cognition and function are often correlated in Stage 4, as cognitive decline generally precedes functional decline (12, 14). However, this is not always the case, as many environmental variables, compensatory strategies, and assistive devices may ameliorate functional decline, without any accompanying improvement in cognition. While executive function may be the most relevant cognitive domain in determining function (21, 22), reliance on specific cognitive domains will vary depending on the demands of the IADL in question.

The focus of drug development is shifting to therapeutic interventions targeting early stages of AD. During MCI and mild AD dementia, patients tend to exhibit slow and variable progression, which can complicate disease assessment (9, 23). Patients may remain in Stage 3 for ~3–5 years and in Stage 4 for ~3

years, with the individual rate of decline influenced by age, sex, apolipoprotein E (APOE) genotype, and potentially other factors yet to be identified (24). The variable length of time during which an individual may remain in the early stages of disease poses challenges for detecting treatment response.

As the disease progresses to Stages 5 and 6, where moderate and severe AD dementia occur, all cognitive domains are increasingly impaired, with greater functional declines in IADL, and progressive deterioration in BADL (Figure 1B) (4, 8). While cognition and function decline continuously, behavioral symptoms vary throughout the disease, with apathy being among the most common symptoms in early stages of AD, followed by aggression and agitation in later stages (Figure 1A) (17).

Clinical trial endpoints for early stages of AD

Historically, clinical trial endpoints have been designed to measure symptom changes in patients with mild-to-moderate or severe AD dementia (25, 26). With the shift of focus toward developing AD therapeutics targeting MCI due to AD and mild AD dementia, it has become necessary to design and validate clinical trial endpoints sensitive to changes that occur during these stages (Figure 2) (5, 25, 26). Changes in MCI due to AD and mild AD dementia are challenging to capture due to high variability in baseline functioning and cognitive ability, and the floor and/or ceiling effects of many of the clinical assessment tools that evaluate AD, which may limit their utility across the continuum of disease (9). These

Table 2B. A selection of assessment tools commonly used only in clinical practice

Assessment type	Assessment tool	Sensitivity and specificity for detecting AD dementia (unless stated otherwise)	Key advantages	Key limitations	Use across AD continuum			
					Stage 1/2	Stage 3	Stage 4/5	Stage 6
					Pre-clinical AD	MCI due to AD	Mild-to-moderate AD dementia	Severe AD dementia
Cognitive	Clock-drawing test (Sheehan 2012)	Sensitivity: 86% Specificity: 96%	<ul style="list-style-type: none"> • Fast, simple • No training requirements • Minimal language requirement 	<ul style="list-style-type: none"> • Assesses narrow part of cognitive dysfunction 				
	GPCOG (Cordell et al. 2013; Patnode et al. 2020)	Sensitivity: 82% Specificity: 83%	<ul style="list-style-type: none"> • Minimal training requirements • Brief 	<ul style="list-style-type: none"> • Informant version reliant on reliable informant 				
	Mini-Cog (Cordell et al. 2013; Patnode et al. 2020)	Sensitivity: 88% Specificity: 71%	<ul style="list-style-type: none"> • Easy to interpret • No training requirements • Brief 	<ul style="list-style-type: none"> • Relatively low sensitivity and specificity 				
	MIS (Cordell et al. 2013; Patnode et al. 2020)	Sensitivity: 73% Specificity: 93%	<ul style="list-style-type: none"> • Free to use • Minimal training requirements 	<ul style="list-style-type: none"> • Lacks information on executive functions or visuospatial functions 				
Functional	DAD-6 (de Rotrou et al. 2012)	Sensitivity for MCI: 83% Sensitivity for MCI: 84%	<ul style="list-style-type: none"> • High discriminative ability to distinguish MCI from mild dementia 	<ul style="list-style-type: none"> • Reliant on reliable informant 				
Behavioral	GDS (Sheehan 2012)	Not available	<ul style="list-style-type: none"> • Reliable 	<ul style="list-style-type: none"> • Validated in mild dementia but not moderate-to-severe dementia 				

Abbreviations: AD = Alzheimer's disease, DAD-6 = Disability Assessment for Dementia Scale, GDS = Global Depression Scale, GPCOG = General Practitioner Assessment of Cognition, MCI = mild cognitive impairment, MIS = Memory Impairment Scale.

changes may be difficult to notice, depending on patient's lifestyle (9) and potential compensatory mechanisms (27), but can be meaningful to the individuals experiencing them. Measures used in clinical trials for MCI due to AD and mild AD dementia must be sensitive enough to differentiate continuing decline in the placebo group from slowed progression in response to active treatment (25).

A variety of tools assessing clinical features are used as primary or secondary endpoints in clinical trials for MCI due to AD and mild AD dementia. These include measures of cognition, function, and behavior, as well as global change, quality of life (QOL), and care partner burden (Table 2A) (2, 3, 5, 28-42). Both the European Medicines Agency (EMA) and FDA have provided guidance on endpoints for AD trials, dependent on disease stage (20, 43). When selecting appropriate assessment tools for clinical trials, several factors should be considered: baseline cognitive severity, test validity, test sensitivity and specificity, test completion time, test-retest reliability, inter-rater reliability, global multi-site setting application, and ability to show clinically meaningful changes (3, 28). These measures for patients with MCI due to AD and mild AD dementia have differing levels of psychometric performance data (Table 2A). Patients with early AD have subtle cognitive deficits and do not present with functional impairment; patients who are closer to the onset of dementia may have noticeable functional deficits that progress slowly, creating sensitivity issues with currently available scales. Furthermore, the extent to which an individual compensates for cognitive deficits and adjusts for daily activities is very variable. It is thus

challenging to establish a clinically meaningful effect during a trial of reasonable duration. Both guidelines suggest an alternative time-to-event approach (e.g., time to the occurrence of a clinically meaningful event during disease progression) to evaluate beneficial effects in trials in early AD. While independent assessment of daily function and cognitive effects may be acceptable, the FDA guideline states that measurable cognitive benefit should not allow for an overall finding of efficacy in the absence of meaningful functional benefit, and vice versa. The EMA guideline states that measures of cognition and function may be included as secondary endpoints to contribute to the overall assessment of efficacy. Both guidelines agree that the use of multiple individual tests may increase the persuasiveness of an overall finding of efficacy. There is a need to construct more sensitive rating scales involving domains that have been shown to be impaired consistently in early AD.

Complementary to clinical measures, imaging and fluid biomarkers quantify measures of brain characteristics or peripherally available markers of pathological changes due to AD, even at early stages when clinical measures might have limited sensitivity. Robust evidence supports the utility of biomarkers to predict the likely trajectory of clinical progression (2), and as such, should be evaluated in conjunction with clinical features to provide a comprehensive estimate of ongoing disease progression (4).

Ability of clinical endpoints to detect changes in clinical features in early stages of AD

Detection of changes in cognition in early AD

As cognition includes a broad spectrum of processes, capturing its full breadth with any one assessment tool is challenging. The EMA guidance on designing AD trials stipulates that specific components of currently available tools are sensitive to detection of disease progression in MCI and mild AD dementia (43). Episodic memory, executive function, and visuospatial function, which are known to be most affected during MCI due to AD (Figure 1A), using tools that focus on these functions may therefore be most appropriate for use in early stages of AD (9, 13, 31). Appropriate cognitive assessment tools that can measure longitudinal changes, especially subtle shifts as observed in MCI, are critical to understanding disease progression and determining treatment effects in longitudinal studies (44).

The Mini-Mental State Exam (MMSE) and the Montreal Cognitive Assessment (MoCA) are similar cognitive assessments tools; both are brief and examine orientation, immediate and delayed recall, language abilities, attention, and visuospatial ability (5). MoCA has more measures of executive function. Due to the floor and ceiling effects of the MMSE, low sensitivity in early stages of AD, and test items weighted heavily toward the memory domain, with minimal assessment of other domains (language, visuospatial function, and executive function), the MoCA may be more appropriate for use in early stages of AD (45). However, although commonly used in clinical practice, the MoCA is not widely used in clinical trials, whereas the MMSE is used in both (Table 2A) (5). The MMSE may be the preferred tool in longitudinal studies as the MoCA may be a more challenging test (that includes executive function, complex visuospatial processing, and higher-level language) for patients with late-stage AD, where cognitive function is more severely impaired (46).

Another tool used widely in clinical trials for mild-to-moderate AD dementia is the ADAS-Cog, which assesses the cognitive domains of memory, language, and praxis (45). The standard ADAS-Cog-11 suffers from ceiling effects at mild stages of AD (47, 48), whereas the modified ADAS-Cog-13, through the addition of a delayed word recall task and a number cancellation task, has an increased sensitivity in mild AD dementia compared with the ADAS-Cog-11. However, sensitivity to change in MCI due to AD remains limited (Table 2A) (48). Additional tools used in AD clinical trials or clinical practice include the Neuropsychological Test Battery (NTB) (32) and the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) (Table 2A) (28, 30). These tools, and a variety of others, have been translated and validated in multiple languages (28, 49, 50). However, for broad application, global interpretation

of results and sensitive detection of changes in early AD, assessment tools should be validated in populations of diverse educational and cultural backgrounds (9, 28).

The cognitive scales discussed above are composite measures of multiple cognitive domains comprising a global cognitive measure. However, as new compounds may have differential effects across cognitive domains, this approach may represent an oversimplification (2). A detailed discussion is beyond the remit of this paper, but interested readers are directed to a recent review on this topic (2).

Detection of changes in function in early AD

Assuming that functional changes in AD are driven by cognitive decline, functional assessments are crucial to evaluating the impact of cognitive decline on the ability to carry out everyday activities. Both EMA and FDA guidelines emphasize the importance of outcome measures that assess meaningful cognitive function (with combined or separate assessment of cognitive and functional measures appropriate for use as primary endpoints in early AD), but neither guideline has provided the definition of “meaningful” (20, 43).

Functional abilities in IADL decline in early stages of AD due to the high level of cognitive ability associated with these activities, whereas BADL are affected later in the disease continuum (Figure 1B) (8); therefore, assessment of IADL is critical to detect any possible change in function in early stages of AD. Generally, tools have tended to focus on lower-level functioning or BADL. One exception is the ADCS-ADL-MCI, modified from the ADCS-ADL in order to emphasize IADL (Table 2A) (39). Likewise, the Amsterdam Instrumental Activities of Daily Living Questionnaire (A-IADL-Q) is sensitive to longitudinal change of IADL in patients with subjective memory complaints, MCI due to AD, and AD dementia (Table 2A) (35). Both tools align with the EMA recommendation to preferentially assess IADL over BADL (43). However, functional assessments are generally informant-based (8), which may not be the most appropriate approach in early stages of AD; in MCI due to AD, some degree of patient insight is generally preserved. Nonetheless, informant-based functional measures may be useful for longitudinal studies that extend beyond the point where patients can respond independently.

Executive function—a repertoire of cognitive processes used to control, regulate, and manage thoughts and behavior—is highly correlated with functional performance (51). A significant body of research has established that executive function assessment tools serve as a reliable proxy measure of functional skills (52, 53). For example, significant correlations between the NTB, which measures executive function and memory, and the Disability Assessment for Dementia (DAD), which measures functional abilities, ($r=0.55$) were observed in studies of lecuzotan. Cognitive assessments that do not

include robust measures of executive function, such as the widely used ADAS-Cog 11, do not tend to correlate highly with functional abilities (32).

Detection of changes in behavior in early AD

Currently, there are limited tools available for assessment of behavior across the AD continuum. One tool, used widely as a clinical trial endpoint, is the Neuropsychiatric Inventory Questionnaire (NPI-Q) (3, 54). This is a self-administered, informant-based questionnaire, developed to assess NPS in clinical practice. Derived from the original NPI-10 (54) the NPI-Q is widely used in both trials and practice, and assesses 12 neuropsychiatric domains, as well as informant distress levels (Table 2A). Although not designed specifically for early stages of AD, it has been used to detect NPS in patients early in the AD continuum (55). Patients who exhibited accelerated decline on the NPI-Q were more likely to develop dementia; thus, this tool may be useful in detecting changes in behavior earlier on in the course of disease.

Another informant-based assessment tool measuring NPS, extensively used in both research and clinical centers, is the Behavioral Pathology in Alzheimer's Disease Rating Scale (BEHAVE-AD) (Table 2A) (54). Like all such measures, the BEHAVE-AD can be affected by informant status, such as recall issues or interpretation bias. The modified Empirical-BEHAVE-AD as well as the NPI-Clinician version (NPI-C) mitigate some of these problems by including direct observation of the patient's behavior (54). However, in early stages of AD, patients themselves may be more insightful and knowledgeable than their informants regarding their own neuropsychiatric symptoms, including anxiety and apathy.

Detection of changes in early AD by composite or global scales

The FDA recognized the value of composite assessments for the evaluation of patients with MCI due to AD (54). Composite scales that include both cognition and function within a single scale may more effectively assess longitudinal changes than single measures, by enabling enhanced measurement of individual variability in clinical decline (34). In addition, composite assessments that include a semi-structured approach allow clinicians the flexibility to evaluate patients more comprehensively. Composite scores have increased power to detect change and may enable shorter or smaller trials. Although FDA guidance no longer recommends specific endpoints, it continues to emphasize the relevance of a link between cognition and function (20).

One measure explicitly identified in the FDA's guidance was the CDR, which has served as a primary or secondary endpoint in multiple clinical trials (56).

The CDR is a semi-structured interview that comprises 75 items relevant to cognition and function, takes ~25 minutes to administer, and is conducted sequentially with the informant and the patient (36). The CDR global score yields an overall rating of disease severity, while the CDR-SB score yields more detailed information on performance across six categories (three functional domains and three cognitive domains) (37). However, the use of CDR-SB in clinical practice may be limited by the subscales (three out of six) involving aspects of daily function (e.g., personal care) not usually impaired in early stages of the disease (57), the lack of behavioral items, and administration requirements (36).

A more practical alternative to the CDR-SB is the Cognitive-Functional Composite (CFC) that is comprised of seven cognitive tests of memory and executive function, plus the functional component represented by a shortened version of the A-IADL-Q (35). The CFC takes only ~25 minutes to administer. The CFC was recently shown to be sensitive to clinical progression in AD, specifically in early stages (34). Other composite scales that have shown sensitivity to changes in early stages of AD are the Alzheimer's Disease Composite Score (ADCOMS), which incorporates items from the MMSE, CDR, and ADAS-Cog-12 (42), and the integrated Alzheimer's Disease Rating Scale (iADRS), which combines scores from ADAS-Cog-14 and ADCS-IADL (58). For ADCOMS, individual measures are administered normally and select items are summed and weighted to generate the ADCOM score; the iADRS score is generated from a sum of the two component tools (42, 58).

Additional considerations for detection of change in clinical features in early stages of AD

Shifts in cognition and function are subtle and decline is generally slow in the early stages of AD (25). To effectively assess change, assessment tools must be sensitive to subtle change and be conducted repeatedly and over a sufficient period of time (25, 59). This enables differentiation among periods of decline, stability, or improvement, to assess a therapeutic effect. While test frequency will depend on trial duration, as well as disease stage and measurement tool sensitivity, a challenge to conducting frequent assessments is the increased vulnerability to practice effects, which may result in improved scores over time masking treatment benefit. A key consideration for selecting endpoints such as the NTB and the CFC is that they are not prone to exhibit practice effects (60, 61).

Both patient and care partner must be considered when selecting clinical endpoints for early AD (62). The role of the care partner shifts across the AD continuum; in the early stages when the patient remains relatively independent, the care partner provides emotional support, facilitates tasks if needed, and assists in planning for the future in anticipation of disease progression; with

greater patient decline the care partner provides more hands-on assistance (63). The What Matters Most study, part of the Alzheimer's Disease Patient and Caregiver Engagement (AD PACE) initiative, found that while traditional cognitive assessments address a range of meaningful outcomes, they do not address outcomes like decreased socialization and mood-related symptoms highly valued by patients and care partners (27). Many of these changes, among others deemed important by patients and informants (64), are reflected in current measures (Table 2A, 2B) (65); nevertheless, a range of tools are required to capture the full spectrum of changes that are considered important by patients and informants.

As AD has marked impact on patients and their families, assessing QOL may enable an improved holistic understanding of the impact of AD and of potential therapeutic benefit. Indeed, the EMA recommends including secondary endpoints such as health-related QOL scales in clinical trials (43). In addition to the insight they provide researchers, these assessment tools provide key information to population-health decision makers, such as payers, policymakers, and advocacy organizations, regarding clinical meaningfulness of changes in disease severity (66, 67). In early AD, particularly during MCI due to AD, there is less impact on QOL and burden compared with AD dementia, due to a lesser impairment on cognition and function (11). Therefore, these assessment tools may be less sensitive to change in the disease's early stages, compared with more advanced disease. QOL assessments must be interpreted carefully, as they measure many influences on QOL, including financial wellbeing, the strength of personal relationships, and the extent of a patient's supportive social network, all pre-disease aspects of an individual's life that are not directly affected by drug treatment assessed in trials. Nevertheless, using QOL measures and evaluating care partner burden in clinical trials, such as with the QOL in AD Scale (QOL-AD) or Zarit Burden Interview, may provide important insight into both patient and care partner (Table 2A).

Beyond measures of QOL, the pharmacoeconomic impact of therapy is often important to decisions of patient access. Additionally, functional and global outcomes—which are linked to costs, resource utilization, and care-partner burden—are valued by payers in assessment of a treatment's cost effectiveness (67). Cost effectiveness is linked to overall health-system burden, which similarly influences the decisions of payers, regulators, and policymakers. The perspective of payers and other health decision makers, in addition to patients and their care partners, should be considered carefully when selecting endpoints for clinical trials.

Are clinical trial endpoints applicable to clinical practice?

It is important that clinical trial endpoints capture data that are meaningful to clinicians, patients, and others AD stakeholders. Measuring the appropriate outcomes can instill clinicians with the confidence that an approved therapeutic for AD is beneficial for their individual patients. However, what works in clinical trials, both in terms of the tool and frequency of administration, may not necessarily work in practice (Table 2A, 2B) (68, 69). Ultimately, the choice of tool should be guided by researchers' and clinicians' differing needs and constraints. Nevertheless, some tools lend themselves more readily than others for use in both settings—the MMSE, MoCA, FAQ, and NPI-Q (Table 2A), for example—and leveraging such instruments in both contexts may provide a valuable continuum of insight from trial to clinic.

Despite the practical limitations such as cost, training requirements, and administration time, the tools used in clinical practice must measure cognition, function, and behavior to provide a comprehensive view into the disease and accurately assess and monitor patient status. Ideally, assessment tools used in clinical practice should be sensitive to the disease stage in which they are used. They should be brief and free of charge, should not require specialist personnel or extensive training, and, if possible, administered digitally (9). These characteristics are reflected in the various tools widely used in clinical practice today (Table 2A, 2B). A number of different approaches have been used to overcome some of the constraints with using structured instruments in clinical practice, such as shortening or translating tools, as seen with the NPI-Q (shortened from the NPI), which is suitable for both clinical trials and clinical practice (70). However, a battery specific to the restrictions of clinical practice, and mindful of patient, care partner, and system burden, would benefit the assessment of AD in this setting, as well as the evaluation of potential future therapeutics.

Clinical meaningfulness

Clarifying the clinical meaningfulness of trial outcomes would promote understanding of the totality of benefits of treatment. Clinical meaningfulness is the practical importance of a treatment effect regarding its impact on the patient and/or family; this differs from statistical significance, which describes the probability of trial outcomes being due to random chance (71). Any clinically meaningful difference determined between groups must be statistically significant; however, alone, statistical significance between groups does not necessarily reflect clinically meaningful changes for patients (71-73). Clinical meaningfulness comprises two key elements: the relevance of the domains measured and the magnitude of treatment effect (74), and can be described either

qualitatively or quantitatively. Qualitative clinical meaningfulness approaches, often assessed through semi-structured interviews or focus groups (75), involve gaining the opinion and experiences of patients, care partners, or clinicians (76). In contrast, quantitative clinical meaningfulness involves examination of thresholds of the minimal clinically important difference (MCID), also called minimal important difference (MID) (76, 77). Time to event analyses (described above) provide relevance and magnitude data to assess meaningfulness.

The MCID is the smallest change in an outcome measure score that leads to meaningful change for patients, which could be determined by patients, care partners, or clinicians. A difference between drug and placebo that is smaller than the MCID would not be considered clinically meaningful even if it were statistically significant, whereas a statistically significant difference that is at least as large as the MCID would be considered clinically significant. MCID values are highly sensitive and are dependent on the clinical trial population (and disease severity) to which they are derived and applied (76-78). MCIDs can also be used to examine individual-level score changes. For example, an individual patient whose score change is at least as large as the MCID can be categorized as a responder (65).

Several methods exist to calculate MCIDs – the two main methods involve a distribution-based approach or an anchor-based approach (74, 78). In the distribution-based approach, the MCID is determined entirely quantitatively, based on the effect size statistic, and is commonly defined as $\frac{1}{2}$ standard deviation (or standard error of the mean) of the baseline outcome measure in the assessed population (73, 74, 78). The anchor-based approach involves anchoring a change in an outcome to a known meaningful change that may be based on previous research or opinion (73, 74, 77-79).

If the anchor-based approach is taken, the anchor must be carefully selected to ensure it is sensitive enough to detect differential levels of change, is relevant to the respondent's perspective of clinically meaningful changes, and appropriate for disease stage. In earlier stages of AD, patients can report their symptomatic changes; in later stages, there is increased reliance on care partner's and clinician's report due to the patient's lack of awareness that may compromise their ability to accurately report their symptoms (78). While several good anchors exist in AD, consensus is currently lacking (73, 77, 79). This is particularly challenging when considering whether to employ patient or care partner or clinical perspectives. A clear interpretation of the anchor value is needed to clarify the clinical meaningfulness scores (76). These challenges underscore the need to understand concepts of meaningfulness qualitatively and to apply these learnings when assessing clinically meaningful change. Examining concepts, behaviors, and activities that patients or care partners value, and considering impacts of intervention on the ability could add insights to an assessment (27, 64, 65).

Defining clinically meaningful changes in early stages of AD

Defining clinically meaningful changes in early stages of AD is challenging due to the slow progression at this stage, as well as high levels of patient variability in symptom initiation and presentation (9, 23, 25). With disease-modifying therapies (DMTs), point difference and effect size may increase over time; thus, it is important to consider the duration of treatment necessary to achieve an effect. Given that clinical trials in early stages of AD have varying durations and use different scales, a consensus regarding treatment duration and point differences or effect sizes may be difficult to achieve (67). Furthermore, this should be considered along with changes in amyloid biomarkers or downstream biomarkers of intracellular tau pathology (phosphorylated tau) and neurodegeneration (total tau or neurofilament light) to establish the agent's biological effect (8, 67). Considering such challenges, it is not surprising that currently there is no consensus on clinically meaningful changes in early stages of AD, with little quantitative research having been conducted on tools used across the AD continuum (74, 80-82).

Are existing definitions of clinically meaningful changes in AD applicable to early stages of disease?

Many MCIDs have been defined based on symptomatic changes, as with those reported in cholinesterase inhibitors (ChEI) trials (81, 83-85). One such study reported a 4-point change in ADAS-Cog as clinically meaningful at 6 months, on a group level, in patients with mild-to-moderate AD dementia (81). However, novel therapeutics may slow disease progression without evidence of symptomatic improvement (81). The previously published thresholds are mostly defined in late-stage patient populations, and as MCIDs vary for each instrument as well as according to disease stage or severity (80, 81, 83-86), this further limits their applicability to trial outcomes for therapeutics that potentially alter underlying disease pathology. One such therapy that has recently received accelerated approval by the US FDA is aducanumab, which selectively targets aggregated forms of amyloid beta (A β) comprising amyloid plaques (87, 88). Unlike other treatments for AD (i.e., ChEI and N-methyl-D-aspartate receptor antagonists), which are approved for mild or moderate-to-severe AD, the FDA-approved label specifies that treatment with aducanumab should be initiated in patients with the mild cognitive impairment or mild dementia stage of disease (87), for which more subtle effects are consistent with maintaining cognitive integrity, a key goal of early intervention (9, 23, 25). There is a need to modify existing thresholds that are based on later stages of AD and develop new measures to assess the clinical outcomes and meaningful effects of treatments approved for early stages of AD. This highlights the need

Table 3. Recommended requirements of an ideal assessment tool for use in early stages of AD in clinical trials and / or clinical practice

Ideal properties of an assessment tool	
1	Be sensitive to subtle changes in early stages of AD
2	Be able to track impairment longitudinally and be sensitive to changes over time
3	Be able to show clinically meaningful changes regarding progression of AD
4	Have well-established psychometric and clinical validity
5	Have minimal or no range restrictions such as floor and ceiling effects
6	Have minimal practice effects
7	Be coherent with biomarker changes
8	Be cross-culturally applicable or have alternate forms, each with well-established psychometric and clinical validity
9	Be able to provide valuable information to the patient, family, caregiver, primary care provider, and payer

Abbreviation: AD = Alzheimer's disease.

for more stage-specific approaches for assessing and defining clinically meaningful changes in early AD.

The Insights to Model Alzheimer's Progression in Real Life (iMAP) is one study aiming to address some of these gaps, by assessing the ability of the RBANS and Alzheimer's Prevention Initiative Preclinical Composite Cognitive Test (APCC) to predict clinically meaningful outcomes, including a diagnosis of MCI due to AD or mild AD dementia, and changes in the CDR-SB (89). Research to examine MCIDs for differentially declining groups, based on natural history, have been reported for three assessment tools commonly used in clinical trials: the MMSE; CDR-SB; and FAQ (80). On average, meaningful changes were observed to be a 1–3-point decrease in MMSE, 1–2-point increase in CDR-SB, and 3–5-point increase in FAQ. The values reflect the range of meaningful change estimates across MCI due to AD; mild, moderate and severe AD dementia. These results should be interpreted with caution as The National Alzheimer's Coordinating Center database used in the study may not be representative of the broader AD population, due to differences in patient characteristics and enrollment procedures (90). In addition, the diagnostic criteria for MCI due to AD did not reflect that of current clinical practice and a substantial population was not biomarker-verified. Further, study methods included several limitations. Primarily, the study used a binary anchor to assess meaningful changes in patients, and the extent to which this reflects a "minimum" is unknown. This study also relied only on clinician opinion, without patient or care partner viewpoint inclusion, potentially leading to a narrow view of disease progression.

Assessing trial outcomes via cognitive, functional, and behavioral changes, plus disease progression biomarkers, will allow better interpretation of an AD treatment's clinical benefit (78). An example of a composite endpoint is the Preclinical Alzheimer Cognitive Composite, which detected subtle cognitive decline and an increased risk for progression to MCI or decline in CDR in clinically

normal older adults with elevated A β biomarkers over three years (91). However, longer and larger trials may be required to detect clinically meaningful changes in cognition related to amyloid status, especially if cognition is the sole domain measured (92). Composite or global assessment tools such as the CDR may capture adequate cognitive and functional data to provide evidence of clinical meaningfulness. The FDA's dual outcome criteria for trials is based on the core clinical features of cognition, plus a global or functional measure, to provide meaningful information (20). Therefore, using tools that are sensitive to early stages of AD for both cognition and function in clinical trials could facilitate conclusions regarding clinical meaningfulness in AD.

Additional challenges of defining clinically meaningful changes in early stages of AD

Additional challenges exist in quantitatively defining clinically meaningful changes in longitudinal data (76). Clinical meaningfulness definitions involve heavy reliance on patient, care partner, or clinician perception, and can involve retrospective recall, introducing of the potential for recall bias (76). Choosing a common mean from a wide distribution of change across the clinical trial population may mask patient variability, leading to possible reduced applicability of the final score to numerous patients (76). It can be informative to use the distribution-based approach to examine the proportion of patients who experience meaningful change, defined as at least $\frac{1}{2}$ standard deviation (or standard error of the mean) of the outcome measure at baseline (73).

Variability of patients' symptoms, in addition to differences in patients' priorities in early stages of AD (Figure 2), complicate current attempts to use an MCID. For example, it may be meaningful for some patients to retain the ability to perform a task with reasonable success or efficiency, whereas for others, it is the ability to simply complete a task, even if they need to complete

it differently from previously. Certain aspects of clinical features, including decline in memory and planning, and changes in behavior, were deemed highly important to patients and care partners in a recent AD PACE study (27). Therefore, defining clinical meaningfulness in these aspects is key to address the viewpoint of patients and care partners, and for investigation of the efficacy of future therapeutics.

The need for consensus

The FDA states that clinical meaningfulness can be detected through changes in cognition, and it emphasizes the link between cognition and function (20). Both the EMA and FDA recommend that endpoints show clinically meaningful changes and note that novel assessment tools of AD are needed to provide evidence of this; neither agency provides a precise definition (20, 43). The FDA acknowledges that in early stages of AD (Stage 2 of their continuum [Figure 2]), it may be difficult to establish a clinically meaningful effect on subtle cognitive deficits within a clinical trial's timeframe (20); additional longitudinal, real-world data collection, using assessment tools appropriate for clinical practice, may be beneficial.

Currently, there are differences in the interpretation of clinical meaningfulness from clinical trial data and real-world practice. Clear definitions, more comprehensive clinical meaningfulness discussions, and correlations between clinically meaningful differences for outcomes in AD clinical trials and clinical practice would provide key information regarding how a therapeutic may impact the everyday lives of the patients, their families, and care partners. Subsequently, this would also increase a clinician's confidence in prescribing a therapeutic, add value to the payer for reimbursement, and help policymakers make informed decisions on the application of clinical trial outcomes in healthcare systems.

Future perspectives

A number of assessment tools are currently in development or being validated that may provide increased sensitivity in the early stages of AD than current options (34). Several recommendations exist for an ideal endpoint for early stages of AD (Table 3), with a focus on longitudinal assessment of change with well-established psychometric and clinical validity (25, 68). Analysis approaches such as evaluating the drug-placebo difference over time ("area under the curve") or time-to-event analysis may facilitate a quantitative assessment of the meaningful preservation of cognition and function.

Digital biomarkers hold promise as potential future endpoints for trial and clinical practice use, enabling passive, objective, large-scale data collection, with very low patient and researcher burden (59, 93, 94). They may capture data on potentially clinically meaningful changes in a patient's day-to-day life (93), which are currently not

captured in traditional measures (25, 94). These potential future biomarkers are subject to patient acceptance and ethical concerns; comprehensively evaluating and validating these is essential prior to clinical trial or clinical practice use (59, 94). Further use of current or future clinical trial endpoints, or potential digital biomarkers, could aid progress toward consensus on full, simple, and clear definitions of clinically meaningful changes. This would advance the field, given the potential emergence of AD therapeutics that alter underlying pathology.

Conclusion

With the emergence of novel therapeutics for AD, it is increasingly important to ensure that assessment tools measure stage-specific, clinically meaningful changes in relation to AD progression and treatment response, focusing particularly on early AD. Progress in these efforts will improve patients' expectations and knowledge of a prescribed therapeutic's potential benefits to their everyday lives, and provide clearer understanding of an approved therapeutic's socioeconomic relevance. Use of an appropriate battery that captures all aspects of the disease in clinical practice, while ensuring low burden on patients, care partners, and the health care system, will also aid in the assessment of the potential therapeutic benefit. Regulatory guidance is needed on outcome selection for AD clinical trials, especially those assessing therapeutics that potentially alter underlying disease pathology. This guidance should consider all stakeholders' viewpoints to better generate meaningful data from trials. Support from real-world data and post-approval observations of effects of a therapeutic may provide key information on the applicability of trial results to a wider population.

Funding and financial disclosures: Editorial support was funded by Biogen provided by Helios Medical Communications and MediTech Media, Ltd.

Conflict of interest: SC: Provided consultation in the past 2 years to Alnylam, Biogen, Cassava Sciences, Cogstate, Eli Lilly, INMune Bio, ProMIS Neurosciences, RetiSpec, and Roche (no personal fees) and receives research support (paid to institution) from AgeneBio, Alector, Anavex, Biogen, Cassava Sciences, Eisai, Eli Lilly, Genentech, Green Valley, Janssen, Novo Nordisk, RetiSpec, Roche, and Vielight. JC: provided consultation to AB Science, Acadia, Alkahest, AlphaCognition, AriBio, Avanir, Axsome, Behren Therapeutics, Biogen, Biohaven, Cassava, Cerecin, Cortexyme, Diadem, EIP Pharma, Eisai, GemVax, Genentech, Green Valley, Grifols, Janssen, LSP, Merck, NervGen, Novo Nordisk, Oligomerix, Ono, Otsuka, PRODEO, ReMYND, Renew, Resverlogix, Roche, Signant Health, Suven, United Neuroscience, and Unlearn AI pharmaceutical, assessment, and investment companies. JH: receipt of personal fees in the past 2 years from Actinogen, AlzeCure, Aptinyx, AstraZeneca, Athira Therapeutics, Axon Neuroscience, Axovant, Bial Biotech, Biogen Idec, BlackThornRx, Boehringer Ingelheim, Brands2Life, Cerecin, Cognito, Cognition Therapeutics, Compass Pathways, Corlieve, Curasen, EIP Pharma, Eisai, G4X Discovery, GfHEU, Heptares, Ki Elements, Lundbeck, Lysosome Therapeutics, MyCognition, Neurocentria, Neurocog, Neurodyn Inc, Neurotrack, the NHS, Novartis, Novo Nordisk, Nutricia, Probiodrug, Prothema, Recognify, Regeneron, reMYND, Rodin Therapeutics, Samumed, Sanofi, Signant, Syndesi Therapeutics, Takeda, Vivoryon Therapeutics and Winterlight Labs. Additionally, he holds stock options in Neurotrack Inc. and is a joint holder of patents with My Cognition Ltd. SK is an employee and shareholder of Biogen. MP was an employee of Biogen at the time of the development of this manuscript.

Open Access: This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, duplication, adaptation, distribution and

reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

References

1. 2020 Alzheimer's disease facts and figures. *Alzheimers Dement*. Published online March 10, 2020; doi:10.1002/alz.12068
2. Schneider LS, Goldberg TE. Composite cognitive and functional measures for early stage Alzheimer's disease trials. *Alzheimers Dement (Amst)* 2020;12(1):e12017; doi:10.1002/dad2.12017
3. Sheehan B. Assessment scales in dementia. *Ther Adv Neurol Disord* 2012;5(6):349-58; doi:10.1177/1756285612455733
4. Jack CR, Jr., Bennett DA, Blennow K, et al. NIA-AA Research Framework: toward a biological definition of Alzheimer's disease. *Alzheimers Dement* 2018;14(4):535-62; doi:10.1016/j.jalz.2018.02.018
5. Cummings J, Lee G, Ritter A, Sabbagh M, Zhong K. Alzheimer's disease drug development pipeline: 2020. *Alzheimers Dement (N Y)* 2020;6(1):e12050; doi:10.1002/trc2.12050
6. Cummings J, Fox N. Defining disease modifying therapy for Alzheimer's disease. *J Prev Alzheimers Dis* 2017;4(2):109-15; doi:10.14283/jpad.2017.12
7. Ismail Z, Smith EE, Geda Y, et al. Neuropsychiatric symptoms as early manifestations of emergent dementia: provisional diagnostic criteria for mild behavioral impairment. *Alzheimers Dement* 2016;12(2):195-202; doi:10.1016/j.jalz.2015.05.017
8. Mlinac ME, Feng MC. Assessment of activities of daily living, self-care, and independence. *Arch Clin Neuropsychol* 2016;31(6):506-16; doi:10.1093/arclin/acw049
9. Sabbagh MN, Boada M, Borson S, et al. Early detection of mild cognitive impairment (MCI) in primary care. *J Prev Alzheimers Dis* 2020;7(3):165-70; doi:10.14283/jpad.2020.21
10. Harada CN, Natelson Love MC, Triebel KL. Normal cognitive aging. *Clin Geriatr Med* 2013;29(4):737-52; doi:10.1016/j.cger.2013.07.002
11. Albert MS, DeKosky ST, Dickson D, 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* 2011;7(3):270-9; doi:10.1016/j.jalz.2011.03.008
12. Liu-Seifert H, Siemers E, Sundell K, et al. Cognitive and functional decline and their relationship in patients with mild Alzheimer's dementia. *J Alzheimers Dis* 2015;43(3):949-55; doi:10.3233/jad-140792
13. Roberts R, Knopman DS. Classification and epidemiology of MCI. *Clin Geriatr Med* 2013;29(4):753-72; doi:10.1016/j.cger.2013.07.003
14. Wajman JR, Schultz RR, de Medeiros Correia Marin S, Bertolucci PHF. Correlation and adaptation among functional and cognitive instruments for staging and monitoring Alzheimer's disease in advanced stages. *Rev Psychiatr Clin* 2014;41:5-8
15. Amieva H, Mokri H, Le Goff M, et al. Compensatory mechanisms in higher-educated subjects with Alzheimer's disease: a study of 20 years of cognitive decline. *Brain* 2014;137(Pt 4):1167-75; doi:10.1093/brain/awu035
16. Fernández M, Gobart AL, Balañá M. Behavioural symptoms in patients with Alzheimer's disease and their association with cognitive impairment. *BMC Neurol* 2010;10:87; doi:10.1186/1471-2377-10-87
17. Kales HC, Gitlin LN, Lyketsos CG. Assessment and management of behavioral and psychological symptoms of dementia. *BMJ* 2015;350:h369; doi:10.1136/bmj.h369
18. Cerejeira J, Lagarto L, Mukaeleva-Ladinska EB. Behavioral and psychological symptoms of dementia. *Front Neurol* 2012;3:73; doi:10.3389/fneur.2012.00073
19. Cloutier M, Gauthier-Loiselle M, Gagnon-Sanschagrin P, et al. Institutionalization risk and costs associated with agitation in Alzheimer's disease. *Alzheimers Dement (N Y)* 2019;5:851-61; doi:10.1016/j.trci.2019.10.004
20. U.S. Food and Drug Administration (FDA). Alzheimer's disease: developing drugs for treatment guidance for industry. 2018. <https://www.fda.gov/files/drugs/published/Alzheimer%2E%20%99s-Disease---Developing-Drugs-for-Treatment-Guidance-for-Industry.pdf>
21. Harvey PD. Domains of cognition and their assessment dialogues. *Clin Neurosci* 2019;21(3):227-37; doi:10.31887/DCNS.2019.21.3/pharvey
22. Razani J, Casas R, Wong JT, Lu P, Alessi C, Josephson K. Relationship between executive functioning and activities of daily living in patients with relatively mild dementia. *Appl Neuropsychol* 2007;14(3):208-14; doi:10.1080/09084280701509125
23. Ganguli M, Jia Y, Hughes TF, et al. Mild cognitive impairment that does not progress to dementia: a population-based study. *J Am Geriatr Soc* 2019;67(2):232-8; doi:10.1111/jgs.15642
24. Vermunt L, Sikkes SAM, van den Hout A, et al. Duration of preclinical, prodromal, and dementia stages of Alzheimer's disease in relation to age, sex, and APOE genotype. *Alzheimers Dement* 2019;15(7):888-98; doi:10.1016/j.jalz.2019.04.001
25. Harvey PD, Cosentino S, Curiel R, et al. Performance-based and observational assessments in clinical trials across the Alzheimer's disease spectrum. *Innov Clin Neurosci* 2017;14(1-2):30-9
26. Schmitt FA, Ashford W, Ernesto C, et al. The severe impairment battery: concurrent validity and the assessment of longitudinal change in Alzheimer's disease. *The Alzheimer's Disease Cooperative Study*. *Alzheimer Dis Assoc Disord* 1997;11 Suppl 2:S51-6
27. DiBenedetti DB, Slota C, Wronski SL, et al. Assessing what matters most to patients with or at risk for Alzheimer's and care partners: a qualitative study evaluating symptoms, impacts, and outcomes. *Alzheimers Res Ther* 2020;12(1):90; doi:10.1186/s13195-020-00659-6
28. Black R, Greenberg B, Ryan JM, et al. Scales as outcome measures for Alzheimer's disease. *Alzheimers Dement* 2009;5(4):324-39; doi:10.1016/j.jalz.2009.05.667
29. De Roeck EE, De Deyn PP, Dierckx E, Engelborghs S. Brief cognitive screening instruments for early detection of Alzheimer's disease: a systematic review. *Alzheimers Res Ther* 2019;11(1):21; doi:10.1186/s13195-019-0474-3
30. Duff K, Humphreys Clark JD, O'Bryant SE, Mold JW, Schiffer RB, Sutker PB. Utility of the RBANS in detecting cognitive impairment associated with Alzheimer's disease: sensitivity, specificity, and positive and negative predictive powers. *Arch Clin Neuropsychol* 2008;23(5):603-12; doi:10.1016/j.acn.2008.06.004
31. Harrison JE, Hendrix S. The assessment of cognition in translational medicine: a contrast between the approaches used in Alzheimer's disease and major depressive disorder. *Behav Neurosci* 2019;29:297-308; doi:10.1016/B978-0-12-803161-2.00021-7
32. Harrison JE, Rentz DM, Brashears HR, Arrighi HM, Ropacki MT, Liu E. Psychometric evaluation of the neuropsychological test battery in individuals with normal cognition, mild cognitive impairment, or mild to moderate Alzheimer's disease: results from a longitudinal study. *J Prev Alzheimers Dis* 2018;5(4):236-44; doi:10.14283/jpad.2018.31
33. Ismail Z, Emeremni CA, Houck PR, et al. A comparison of the E-BEHAVE-AD, NBRS, and NPI in quantifying clinical improvement in the treatment of agitation and psychosis associated with dementia. *Am J Geriatr Psychiatry* 2013;21(1):78-87; doi:10.1016/j.jagp.2012.10.013
34. Jutten RJ, Harrison JE, Brunner AJ, et al. The Cognitive-Functional Composite is sensitive to clinical progression in early dementia: longitudinal findings from the Catch-Cog study cohort. *Alzheimers Dement (N Y)* 2020;6(1):e12020-e; doi:10.1002/trc2.12020
35. Koster N, Knol DL, Uitdehaag BM, Scheltens P, Sikkes SA. The sensitivity to change over time of the Amsterdam IADL Questionnaire(C). *Alzheimers Dement* 2015;11(10):1231-40; doi:10.1016/j.jalz.2014.10.006
36. Marshall GA, Amariglio RE, Sperling RA, Rentz DM. Activities of daily living: where do they fit in the diagnosis of Alzheimer's disease? *Neurodegener Dis Manag* 2012;2(5):483-91; doi:10.2217/nmt.12.55
37. O'Bryant SE, Lacritz LH, Hall J, et al. Validation of the new interpretive guidelines for the clinical dementia rating scale sum of boxes score in the national Alzheimer's coordinating center database. *Arch Neurol* 2010;67(6):746-9; doi:10.1001/archneurol.2010.115
38. Patnode CD, Perdue LA, Rossom RC, et al. Screening for cognitive impairment in older adults: updated evidence report and systematic review for the US Preventive Services Task Force. *JAMA* 2020;323(8):764-85; doi:10.1001/jama.2019.22258
39. Pedrosa H, De Sa A, Guerreiro M, et al. Functional evaluation distinguishes MCI patients from healthy elderly people—the ADCS/MCI/ADL scale. *J Nutr Health Aging* 2010;14(8):703-9; doi:10.1007/s12603-010-0102-1
40. Roalf DR, Moore TM, Wolk DA, et al. Defining and validating a short form Montreal Cognitive Assessment (s-MoCA) for use in neurodegenerative disease. *J Neurol Neurosurg Psychiatry* 2016;87(12):1303-10; doi:10.1136/jnnp-2015-312723
41. Sikkes SA, Pijnenburg YA, Knol DL, de Lange-de Klerk ES, Scheltens P, Uitdehaag BM. Assessment of instrumental activities of daily living in dementia: diagnostic value of the Amsterdam Instrumental Activities of Daily Living Questionnaire. *J Geriatr Psychiatry Neurol* 2013;26(4):244-50; doi:10.1177/0891988713509139
42. Wang J, Logovinsky V, Hendrix SB, et al. ADCOMS: a composite clinical outcome for prodromal Alzheimer's disease trials. *J Neurol Neurosurg Psychiatry* 2016;87(9):993-9; doi:10.1136/jnnp-2015-312383
43. European Medicines Agency. Guideline on the clinical investigation of medicines for the treatment of Alzheimer's disease. 2018. https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-clinical-investigation-medicines-treatment-alzheimers-disease-revision-2_en.pdf
44. Li F, Takechi H, Kokuryu A, Takahashi R. Longitudinal changes in performance on cognitive screening tests in patients with mild cognitive impairment and Alzheimer disease. *Dement Geriatr Cogn Dis Extra* 2017;7(3):366-73; doi:10.1159/000481910
45. Pinto TCC, Machado L, Bulgakov TM, et al. Is the Montreal Cognitive Assessment (MoCA) screening superior to the Mini-Mental State Examination (MMSE) in the detection of mild cognitive impairment (MCI) and Alzheimer's disease (AD) in the elderly? *Int Psychogeriatr* 2019;31(4):491-504; doi:10.1017/

- s1041610218001370
46. Trzepacz PT, Hochstetler H, Wang S, Walker B, Saykin AJ, for the Alzheimer's Disease Neuroimaging Initiative. Relationship between the Montreal Cognitive Assessment and Mini-Mental State Examination for assessment of mild cognitive impairment in older adults. *BMC Geriatr* 2015;15(1):107; doi:10.1186/s12877-015-0103-3
 47. Doraiswamy PM, Kaiser L, Bieber F, Garman RL. The Alzheimer's Disease Assessment Scale: evaluation of psychometric properties and patterns of cognitive decline in multicenter clinical trials of mild to moderate Alzheimer's disease. *Alzheimer Dis Assoc Disord* 2001;15(4):174-83; doi:10.1097/00002093-200110000-00003
 48. Podhorna J, Krahne T, Shear M, Harrison JE. Alzheimer's Disease Assessment Scale-Cognitive subscale variants in mild cognitive impairment and mild Alzheimer's disease: change over time and the effect of enrichment strategies. *Alzheimers Res Ther* 2016;8:8; doi:10.1186/s13195-016-0170-5
 49. Muntal S, Doval E, Badenes D, et al. New normative data from the Spanish-language version of the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS), form A. *Neurologia (Engl Ed)* 2020;35(5):303-10; doi:10.1016/j.nrl.2017.09.001
 50. Shen JH, Shen Q, Yu H, et al. Validation of an Alzheimer's disease assessment battery in Asian participants with mild to moderate Alzheimer's disease. *Am J Neurodegener Dis* 2014;3(3):158-69
 51. McAlister C, Schmitter-Edgecombe M. Executive function subcomponents and their relations to everyday functioning in healthy older adults. *J Clin Exp Neuropsychol* 2016;38(8):925-40; doi:10.1080/13803395.2016.1177490
 52. Roy M, Molnar F. Systematic review of the evidence for Trails B cut-off scores in assessing fitness-to-drive. *Can Geriatr J* 2013;16(3):120-42; doi:10.5770/cgj.16.76
 53. Szlyk JP, Myers L, Zhang Y, Wetzel L, Shapiro R. Development and assessment of a neuropsychological battery to aid in predicting driving performance. *J Rehabil Res Dev* 2002;39(4):483-96
 54. Stella F. Assessment of neuropsychiatric symptoms in dementia: toward improving accuracy. *Dement Neuropsychol* 2013;7(3):244-51; doi:10.1590/S1980-57642013DN7030003
 55. Nunes PV, Schwarzer MC, Leite REP, et al. Neuropsychiatric inventory in community-dwelling older adults with mild cognitive impairment and dementia. *J Alzheimers Dis* 2019;68(2):669-78; doi:10.3233/JAD-180641
 56. U.S. Food and Drug Administration (FDA). Alzheimer's disease: developing drugs for the treatment of early stage disease. 2013. https://isctm.org/public-access/FDAGuidance_AD_Developing_Drugs_Early_Stage_Treatment.pdf. Accessed March 15, 2022
 57. Samtani MN, Raghavan N, Novak G, Nandy P, Narayan VA. Disease progression model for Clinical Dementia Rating-Sum of Boxes in mild cognitive impairment and Alzheimer's subjects from the Alzheimer's Disease Neuroimaging Initiative. *Neuropsychiatr Dis Treat* 2014;10:929-52; doi:10.2147/ndt.S62323
 58. Wessels AM, Siemers ER, Yu P, et al. A combined measure of cognition and function for clinical trials: the integrated Alzheimer's Disease Rating Scale (iADRS). *J Prev Alzheimers Dis* 2015;2(4):227-41; doi:10.14283/jpad.2015.82
 59. Dorsey ER, Papapetropoulos S, Xiong M, Kieburtz K. The first frontier: digital biomarkers for neurodegenerative disorders. *Digit Biomark* 2017;1(1):6-13; doi:10.1159/000477383
 60. Jutten RJ, Harrison J, Lee Meeuw Kjoe PR, et al. A novel cognitive-functional composite measure to detect changes in early Alzheimer's disease: test-retest reliability and feasibility. *Alzheimers Dement (Amst)* 2018;10:153-60; doi:10.1016/j.dadm.2017.12.002
 61. Harrison J, Rentz DM, McLaughlin T, et al. Cognition in MCI and Alzheimer's disease: baseline data from a longitudinal study of the NTB. *Clin Neuropsychol* 2014;28(2):252-68; doi:10.1080/13854046.2013.875595
 62. Bouwens SF, van Heugten CM, Verhey FR. Review of goal attainment scaling as a useful outcome measure in psychogeriatric patients with cognitive disorders. *Dement Geriatr Cogn Disord* 2008;26(6):528-40; doi:10.1159/000178757
 63. Dean K, Wilcock G. Living with mild cognitive impairment: the patient's and carer's experience. *Int Psychogeriatr* 2012;24(6):871-81; doi:10.1017/s104161021100264x
 64. Tochel C, Smith M, Baldwin H, et al. What outcomes are important to patients with mild cognitive impairment or Alzheimer's disease, their caregivers, and health-care professionals? A systematic review. *Alzheimers Dement (Amst)* 2019;11:231-47; doi:10.1016/j.dadm.2018.12.003
 65. Ropacki MT, Hannisdottir K, Hendrix S, et al. Clinically meaningful outcomes in early Alzheimer disease: a consortia-driven approach to identifying what matters to patients. *Ther Innov Regul Sci* 2017;51(3):380-90; doi:10.1177/2168479016689712
 66. Kahle-Wroblekski K, Ye W, Henley D, et al. Assessing quality of life in Alzheimer's disease: implications for clinical trials. *Alzheimers Dement (Amst)* 2017;6:82-90; doi:10.1016/j.dadm.2016.11.004
 67. Bradley P, Akehurst R, Ballard C, et al. Taking stock: a multistakeholder perspective on improving the delivery of care and the development of treatments for Alzheimer's disease. *Alzheimers Dement* 2015;11(4):455-61; doi:10.1016/j.jalz.2014.01.007
 68. Cordell CB, Borson S, Boustani M, et al. Alzheimer's Association recommendations for operationalizing the detection of cognitive impairment during the Medicare Annual Wellness Visit in a primary care setting. *Alzheimers Dement* 2013;9(2):141-50; doi:10.1016/j.jalz.2012.09.011
 69. National Institute for Health and Care Excellence (NICE). *Dementia: assessment, management and support for people living with dementia and their carers*. 2018. <https://www.ncbi.nlm.nih.gov/books/NBK513207/>. Accessed 15 March 2022
 70. Franzen S, van den Berg E, Goudsmit M, et al. A systematic review of neuropsychological tests for the assessment of dementia in non-Western, low-educated or illiterate populations. *J Int Neuropsychol Soc* 2020;26(3):331-51; doi:10.1017/s1355617719000894
 71. Ranganathan P, Pramesh CS, Buyse M. Common pitfalls in statistical analysis: clinical versus statistical significance. *Perspect Clin Res* 2015;6(3):169-70; doi:10.4103/2229-3485.159943
 72. U.S. Food and Drug Administration. Patient-reported outcome measures: use in medical product development to support labeling claims (final). 2009. <https://www.fda.gov/media/77832/>. Accessed 15 March 2022
 73. McLeod LD, Cappelleri JC, Hays RD. Best (but oft-forgotten) practices: expressing and interpreting associations and effect sizes in clinical outcome assessments. *Am J Clin Nutr* 2016;103(3):685-93; doi:10.3945/ajcn.115.120378
 74. Edgar CJ, Vradenburg G, Hassenstab J. The 2018 Revised FDA Guidance for Early Alzheimer's disease: establishing the meaningfulness of treatment effects. *J Prev Alzheimers Dis* 2019;6(4):223-7; doi:10.14283/jpad.2019.30
 75. Staunton H, Willgoss T, Nelsen L, et al. An overview of using qualitative techniques to explore and define estimates of clinically important change on clinical outcome assessments. *J Patient Rep Outcomes* 2019;3(1):16-; doi:10.1186/s41687-019-0100-y
 76. Cook CE. Clinimetrics corner: the minimal clinically important change score (MCID): a necessary pretense. *J Man Manip Ther* 2008;16(4):E82-E3; doi:10.1179/jmt.2008.16.4.82E
 77. Devji T, Carrasco-Labra A, Qasim A, et al. Evaluating the credibility of anchor based estimates of minimal important differences for patient reported outcomes: instrument development and reliability study. *BMJ* 2020;369:m1714; doi:10.1136/bmj.m1714
 78. Rentz DM, Wessels AM, Annapragada AV, et al. Building clinically relevant outcomes across the Alzheimer's disease spectrum. *Alzheimers Dement (N Y)* 2021;7(1):e12181; doi:10.1002/trc.212181
 79. Wyrych KW, Norquist JM, Lenderking WR, Acaster S. Methods for interpreting change over time in patient-reported outcome measures. *Qual Life Res* 2013;22(3):475-83; doi:10.1007/s11136-012-0175-x
 80. Andrews JS, Desai U, Kirson NY, Zichlin ML, Ball DE, Matthews BR. Disease severity and minimal clinically important differences in clinical outcome assessments for Alzheimer's disease clinical trials. *Alzheimers Dement (N Y)* 2019;5:354-63; doi:10.1016/j.trci.2019.06.005
 81. Rockwood K, Fay S, Gorman M. The ADAS-cog and clinically meaningful change in the VISTA clinical trial of galantamine for Alzheimer's disease. *Int J Geriatr Psychiatry* 2010;25(2):191-201; doi:10.1002/gps.2319
 82. Schrag A, Schott JM. What is the clinically relevant change on the ADAS-Cog? *J Neurol Neurosurg Psychiatry* 2012;83(2):171-3; doi:10.1136/jnnp-2011-300881
 83. Burback D, Molnar FJ, St John P, Man-Son-Hing M. Key methodological features of randomized controlled trials of Alzheimer's disease therapy. Minimal clinically important difference, sample size and trial duration. *Dement Geriatr Cogn Disord* 1999;10(6):534-40; doi:10.1159/000017201
 84. Hensel A, Angermeyer MC, Riedel-Heller SG. Measuring cognitive change in older adults: reliable change indices for the Mini-Mental State Examination. *J Neurol Neurosurg Psychiatry* 2007;78(12):1298-303; doi:10.1136/jnnp.2006.109074
 85. Howard R, Phillips P, Johnson T, et al. Determining the minimum clinically important differences for outcomes in the DOMINO trial. *Int J Geriatr Psychiatry* 2011;26(8):812-7; doi:10.1002/gps.2607
 86. Feeney J, Savva GM, O'Regan C, King-Kallimanis B, Cronin H, Kenny RA. Measurement error, reliability, and minimum detectable change in the Mini-Mental State Examination, Montreal Cognitive Assessment, and Color Trails Test among community living middle-aged and older adults. *J Alzheimers Dis* 2016;53(3):1107-14; doi:10.3233/jad-160248
 87. Aduhelm. Prescribing information. Biogen, Inc; 2021.
 88. Budd Haeberlein S, Aisen PS, Barkhof F, et al. Two Randomized Phase 3 Studies of Aducanumab in Early Alzheimer's Disease. *J Prev Alzheimers Dis* (2022). <https://doi.org/10.14283/jpad.2022.30>
 89. Graf A, Risson V, Gustavsson A, et al. Assessment of clinical meaningfulness of endpoints in the generation program by the insights to model Alzheimer's progression in real life (iMAP) study. *J Prev Alzheimers Dis* 2019;6(2):85-9; doi:10.14283/jpad.2018.49
 90. National Alzheimer's Coordinating Center. <http://www.naccdata.org>
 91. Papp KV, Buckley R, Mormino E, et al. Clinical meaningfulness of subtle cognitive decline on longitudinal testing in preclinical AD. *Alzheimers*

- Dement 2020;16(3):552-60; doi:10.1016/j.jalz.2019.09.074
92. Insel PS, Weiner M, Mackin RS, et al. Determining clinically meaningful decline in preclinical Alzheimer disease. Neurology 2019;93(4):e322; doi:10.1212/WNL.0000000000007831
93. Gold M, Amatniek J, Carrillo MC, et al. Digital technologies as biomarkers, clinical outcomes assessment, and recruitment tools in Alzheimer's disease clinical trials. Alzheimers Dement (N Y) 2018;4:234-42; doi:10.1016/j.jADI.2018.04.003
94. Kourtis LC, Regele OB, Wright JM, Jones GB. Digital biomarkers for Alzheimer's disease: the mobile/wearable devices

How to cite this article: S. Cohen, J. Cummings, S. Knox, et al. Clinical Trial Endpoints and Their Clinical Meaningfulness in Early Stages of Alzheimer's Disease. *J Prev Alz Dis* 2022;3(9):507-522; <http://dx.doi.org/10.14283/jpad.2022.41>