

Evidence-based Practice Center Systematic Review Protocol

Project Title: Diagnosis and Treatment of Clinical Alzheimer's-type Dementia (CATD)

Initial publication date if applicable: TBD

Amendment Date(s) if applicable:

(Amendments Details—see Section VII)

I. Background and Objectives for the Systematic Review

Dementia is a clinical syndrome affecting about 10% of older U.S. adults^{1, 2} in which an acquired cognitive deficit interferes with a person's independence in daily activities.³ It adversely affects patient quality of life, burdens caregivers, increases institutionalization, and is costly to families and society.⁴ Agitation, aggression and other behavioral and psychological symptoms in dementia (BPSD) are common,⁵ especially late in the disease course. These symptoms may imperil the safety of the patient and others, and often are highly distressing to caregivers. Most individuals with dementia have Alzheimer's disease (AD) as at least part of their underlying disease process.^{6, 7}

Historically, premortem AD diagnosis has been operationalized clinically, as by the Diagnostic and Statistical Manual of Mental Disorders (DSM) and the National Institute of Neurological and Communicative Disorders and Stroke and Alzheimer's Disease and Related Disorders (NINCDS-ADRDA) criteria.^{8, 9} These criteria require acquired, persistent impairment in memory and another cognitive domain with associated functional disability not attributable to another disorder. However, this approach is limited. First, the details of these clinical diagnostic criteria differ (e.g., DSM, ICD, NINCDS-ADRDA) and have evolved over time.^{3, 10} Second, these criteria dichotomously frame AD as present or absent. This is counter to current understanding of AD as an insidiously progressive disease in which neuropathological changes accumulate over decades, symptoms may first be detectable relatively late in the course, and multiple neuropathologies may contribute to clinical AD. Third, in clinical settings, even when the etiology of dementia is thought to be AD, it may not be possible to differentiate between isolated AD, dementia due to a combination of AD plus another etiology (e.g., cerebrovascular disease), or dementia due to a non-AD neurodegenerative disease. Many individuals clinically diagnosed with AD do not meet neuropathological AD criteria at autopsy and/or have additional pathological changes (e.g., microinfarcts or Lewy bodies).¹¹ For these reasons, patients with probable AD based on a clinical evaluation may be labeled more tentatively as having clinical Alzheimer's-type dementia (CATD).

Neuropsychological testing may quantify the severity of cognitive impairment and the pattern of cognitive performance across multiple domains, helping to clinically diagnose dementia and distinguish between different dementia subtypes. However, access to comprehensive neuropsychological testing is limited in many clinical settings. This has heightened interest in identifying which individual cognitive tests or combinations of cognitive tests are most accurate for clinically diagnosing CATD in patients in whom this condition is

suspected. Currently there are no evidence-based guidelines about the merits of either brief cognitive testing or comprehensive neuropsychological testing in this patient population. Nevertheless, identification of brief tests that are sensitive and specific for CATD could increase the efficiency of the diagnostic evaluation of patients in clinical settings, especially in primary care.

Limitations in the accuracy of the clinical diagnosis of AD as the underlying cause of CATD, even after a full clinical evaluation including neuropsychological testing, have spurred efforts to identify biomarkers specific for AD. Brain imaging and cerebrospinal fluid (CSF) biomarkers now exist that may reflect manifestations of AD pathology, including localized neuronal hypometabolism, localized neuronal loss, cortical amyloid deposition, abnormal β -amyloid metabolism, and accumulation of tau pathology.¹² Blood tests are earlier in development. However, precision of existing biomarker assays may be limited, and definitions of normal thresholds for individual biomarkers and their combinations are not fully standardized. Further, positive and negative predictive values likely vary depending on patient age, education, level of cognitive or functional impairment or disease stage, risk for non-AD pathology, and time between biomarker collection and neuropathological diagnosis. The most recent systematic review on the diagnostic sensitivity and specificity of CSF and brain imaging biomarkers for autopsy confirmed AD in symptomatic patients only includes studies published through 2011.¹¹ Due to the publication of many relevant biomarker studies since that time and the development of new imaging methods and assays, an updated review is needed to synthesize the most current evidence on the accuracy and harms of these tests for diagnosing AD. Identification of biomarkers that are sensitive and specific for AD, are detectable early in the disease course, and are associated with minimal burden to patients would be useful for clinical decision making.

There are many interventions that may be considered for treatment of CATD, with the goal of improving, stabilizing or slowing the decline in cognition, function, quality of life, and BPSD. These include nonpharmacological interventions, prescription pharmacological interventions, and nonprescription pharmacological interventions (e.g., over-the-counter drugs, vitamins or supplements, herbal remedies).

A recent Agency for Healthcare Research and Quality (AHRQ) report examined the effects of nonpharmacological interventions, including cognitive training, physical activity, and diet, on prevention or slowing of cognitive decline in adults with normal cognition or MCI.¹³ Moderate-strength evidence showed that cognitive training could improve the cognitive domain trained in patients with normal cognition, most physical activity interventions showed no benefit, and evidence about diet interventions was insufficient to draw conclusions about their benefits and harms. We are not aware of a recent review on the effect of nonpharmacological interventions for treatment of cognition, function and quality of life in patients with established CATD. Though nonpharmacological interventions are recommended as first line treatments for BPSD,¹⁴ a recent AHRQ report found that patient-level and care delivery-level interventions were not superior to usual care for managing agitation and aggression, and that evidence was insufficient

to draw conclusions about the effectiveness of most caregiver-level interventions.¹⁵ While these interventions generally are presumed safe, trials rarely have reported information about harms.

A 2008 American Academy of Family Physicians (AAFP)/American College of Physicians (ACP) guideline focused on pharmacological treatment of CATD.¹⁶ It reported that evidence from mostly short-term randomized controlled trials (RCTs) showed that cholinesterase inhibitors and memantine statistically significantly improved cognition, but that the mean differences in cognitive scores between active treatment and control groups were not clinically important. Some studies reported that more patients assigned cholinesterase inhibitors had clinically important improvements in cognition than did those assigned placebo, suggesting a possible subpopulation benefit. However, these studies did not report formal test results for whether the proportions with clinically important improvements significantly differed between treatment groups. Data reported on function was limited and these treatments did not improve behavioral symptoms. The guideline stated that evidence was insufficient to compare the effectiveness of different pharmacological agents for treatment of dementia. The guideline recommended that decisions to initiate one of these therapies should be individualized to the patient, should consider issues of adverse effects, ease of use, and cost, and that further research on the clinical effectiveness of pharmacological treatments for dementia was urgently needed. Nonprescription pharmacological treatments were included in the evidence review but were not addressed in the guideline. Though no new medications have been approved for treatment of CATD by the U.S. Food and Drug Administration (FDA) since before the 2008 AAFP/ACP guideline, numerous new trials of existing agents have been published. A recent nonsystematic review reported that antipsychotics and mood stabilizers for treatment of BPSD in patients with dementia did not improve behavioral symptoms more than placebo, but had a substantially increased risk of harms.¹⁷ Results for selective serotonin reuptake inhibitor antidepressants were mixed. Nonprescription pharmacological treatments were not addressed.

With respect to nonprescription pharmacological treatments, claims abound on the internet and elsewhere about the cognitive benefits of dozens of over-the-counter drugs, supplements, and herbal remedies on cognition and function in patients with CATD. Anecdotally, patient and caregiver questions to primary care providers about the potential benefits of these agents are common. The efficacy of some older nonprescription agents has been evaluated in RCTs and systematic reviews, and, for many of these, it is unlikely that new trials exist to warrant a fresh review. For old agents with new trials, new agents, or agents that otherwise have received increased public interest (e.g., cannabinoids, ginseng, omega 3, ginkgo, huperzine A), a new comprehensive systematic review that examines the effects of these agents, including not just cognition, but also function, quality of life, BPSD and harms, would have clinical value.

Primary care providers routinely provide dementia care, and require current, evidence-based guidance to optimize their clinical practices for diagnosing and treating dementia. To address this need, the AAFP nominated this topic to update their 2008 AAFP/ACP guideline on prescription and nonprescription pharmacological treatment of CATD,¹⁶ and broaden it by adding questions about the efficacy and harms of nonpharmacological CATD treatment, and the accuracy and

harms of diagnostic testing of adults with suspected cognitive impairment. Because a separate ongoing AHRQ review is focused on the efficacy and harms of nonpharmacological treatments for patients with CATD, these interventions will not be included in the present review except when included as a control group for a pharmacological intervention. Therefore, the scope of the present review will be limited to cognitive and biomarker diagnostic testing for CATD and AD, and prescription and nonprescription pharmacological treatment of patients with CATD.

Because primary care providers must make clinical decisions in individual patients, average results on diagnostic accuracy and treatment efficacy and harms may have limited applicability. Identification of patient characteristics that are associated with diagnostic test accuracy and harms, and with pharmacological treatment efficacy and harms may help physicians, patients and caregivers make more informed individualized decisions about how to test, whether to treat, with what treatment and when, and when to stop treatment. Therefore, this review also will examine whether factors such as age, sex, race/ethnicity, education, depression, pre-treatment cognitive or functional level/CATD stage, pre-treatment BPSD severity, and living setting modify the accuracy and comparative accuracy of diagnostic tests and the efficacy and comparative effectiveness of pharmacologic treatments.

II. The Key Questions

KQ 1: In adults with CATD, what are the efficacy and harms of prescription pharmacological interventions versus placebo/inactive control for treatment of cognition, function, and quality of life?

KQ 1a: In adults with CATD, does the efficacy of prescription pharmacological interventions versus placebo/inactive control vary as a function of patient characteristics (i.e., age, sex, race/ethnicity, depression, pre-treatment cognitive or functional level/CATD stage, living setting)?

KQ 2: In adults with CATD, what are the efficacy and harms of nonprescription pharmacological interventions versus placebo/inactive control for treatment of cognition, function, and quality of life?

KQ 2a: In adults with CATD, does the efficacy of nonprescription pharmacological interventions versus placebo/inactive control vary as a function of patient characteristics (i.e., age, sex, race/ethnicity, depression, pre-treatment cognitive or functional level/CATD stage, living setting)?

KQ 3: In adults with CATD, what are the comparative effectiveness and harms of prescription pharmacological interventions versus other active interventions for treatment of cognition, function, and quality of life?

KQ 3a: In adults with CATD, what are the comparative effectiveness and harms of prescription pharmacological interventions versus other prescription pharmacological interventions for treatment of cognition, function, and quality of life?

- KQ3b: In adults with CATD, what are the comparative effectiveness and harms of prescription pharmacological interventions versus nonprescription pharmacological interventions for treatment of cognition, function, and quality of life?
- KQ3c: In adults with CATD, what are the comparative effectiveness and harms of prescription pharmacological interventions versus nonpharmacological interventions for treatment of cognition, function, and quality of life?
- KQ 3d: In adults with CATD, does the comparative effectiveness of prescription pharmacological interventions versus other active interventions for treatment of cognition, function, and quality of life vary as a function of patient characteristics (i.e., age, sex, race/ethnicity, depression, pre-treatment cognitive or functional level/CATD stage, living setting)?
- KQ 4: In adults with CATD and behavioral and psychological symptoms of dementia (BPSD), what are the efficacy and harms of prescription pharmacological interventions versus placebo/inactive control for treatment of BPSD?
- KQ 4a: In adults with CATD and BPSD, what are the efficacy and harms of prescription pharmacological interventions versus placebo/inactive control for reducing frequency and severity of future BPSD?
- KQ 4b: In adults with CATD and BPSD, does the efficacy of prescription pharmacological interventions versus placebo/inactive control for reducing frequency and severity of future BPSD vary as a function of patient characteristics (i.e., age, sex, race/ethnicity, depression, pre-treatment cognitive or functional level/CATD stage, pre-treatment BPSD severity, living setting)?
- KQ 4c: In adults with CATD and BPSD, what are the efficacy and harms of prescription pharmacological interventions versus placebo/inactive control for acute treatment of BPSD?
- KQ 4d: In adults with CATD and BPSD, does the efficacy of prescription pharmacological interventions versus placebo/inactive control for acute treatment of BPSD vary as a function of patient characteristics (i.e., age, sex, race/ethnicity, depression, pre-treatment cognitive or functional level/CATD stage, pre-treatment BPSD severity, living setting)?
- KQ 5: In adults with CATD and BPSD, what are the efficacy and harms of nonprescription pharmacological interventions versus placebo/inactive control for treatment of BPSD in adults with CATD and BPSD?
- KQ 5a: In adults with CATD and BPSD, what are the efficacy and harms of nonprescription pharmacological interventions versus placebo/inactive control for reducing frequency and severity of future BPSD?
- KQ 5b: In adults with CATD and BPSD, does the efficacy of nonprescription pharmacological interventions versus placebo/inactive control for reducing

frequency and severity of future BPSD vary as a function of patient characteristics (i.e., age, sex, race/ethnicity, depression, pre-treatment cognitive or functional level/CATD stage, pre-treatment BPSD severity, living setting)?

- KQ 5c: In adults with CATD and BPSD, what are the efficacy and harms of nonprescription pharmacological interventions versus placebo/inactive control for acute treatment of BPSD?
- KQ 5d: In adults with CATD and BPSD, does the efficacy of nonprescription pharmacological interventions versus placebo/inactive control for acute treatment of BPSD vary as a function of patient characteristics (i.e., age, sex, race/ethnicity, depression, pre-treatment cognitive or functional level/CATD stage, pre-treatment BPSD severity, living setting)?
- KQ 6: In adults with CATD and BPSD, what are the comparative effectiveness and harms of prescription pharmacological interventions versus other active interventions for treatment of BPSD?
- KQ 6a: In adults with CATD and BPSD, what are the comparative effectiveness and harms of prescription pharmacological interventions versus other prescription pharmacological interventions for reducing frequency and severity of future BPSD?
- KQ 6b: In adults with CATD and BPSD, what are the comparative effectiveness and harms of prescription pharmacological interventions versus nonprescription pharmacological interventions for reducing frequency and severity of future BPSD?
- KQ 6c: In adults with CATD and BPSD, what are the comparative effectiveness and harms of prescription pharmacological interventions versus nonpharmacological interventions for reducing frequency and severity of future BPSD?
- KQ 6d: In adults with CATD and BPSD, does the comparative effectiveness of prescription pharmacological interventions versus other active interventions for reducing frequency and severity of future BPSD vary as a function of patient characteristics (i.e., age, sex, race/ethnicity, depression, pre-treatment cognitive or functional level/CATD stage, pre-treatment BPSD severity, living setting)?
- KQ 6e: In adults with CATD and BPSD, what are the comparative effectiveness and harms of prescription pharmacological interventions versus other prescription pharmacological interventions for acute treatment of BPSD?
- KQ 6f: In adults with CATD and BPSD, what are the comparative effectiveness and harms of prescription pharmacological interventions versus nonprescription pharmacological interventions for acute treatment of BPSD?
- KQ 6g: In adults with CATD and BPSD, what are the comparative effectiveness and harms of prescription pharmacological interventions versus nonpharmacological interventions for acute treatment of BPSD?

- KQ 6h: In adults with CATD and BPSD, does the comparative effectiveness of prescription pharmacological interventions versus other active interventions for acute treatment of BPSD vary as a function of patient characteristics (i.e., age, sex, race/ethnicity, depression, pre-treatment cognitive or functional level/CATD stage, pre-treatment BPSD severity, living setting)?
- KQ 7: In adults with suspected CATD, what are the accuracy, comparative accuracy, and harms of different individual cognitive diagnostic tests and their combinations for making the diagnosis of CATD as defined by full clinical evaluation and/or neuropsychological testing with explicit diagnostic criteria?
- KQ 7a: Do the accuracy and comparative accuracy of cognitive tests for making the diagnosis of CATD as defined by full clinical evaluation and/or neuropsychological testing with explicit diagnostic criteria vary as a function of patient characteristics (i.e., age, sex, race/ethnicity, education, pre-testing cognitive or functional level CATD stage)?
- KQ 8: In adults with a clinical diagnosis of CATD, what are the accuracy, comparative accuracy, and harms of brain imaging, CSF, and blood tests for diagnosing pathologically confirmed Alzheimer's disease as the underlying etiology?
- KQ 8a: Do the accuracy and comparative accuracy of brain imaging, CSF, and blood tests for pathologically confirmed Alzheimer's disease as the underlying etiology of CATD vary as a function of patient characteristics (i.e., age, sex, race/ethnicity, depression, education, pre-testing cognitive or functional level CATD stage)?

Table 1. PICOTS (Populations, Interventions, Comparators, Outcomes, Timing, Settings/Study Design)

KQ	Population	Intervention	Treatment Comparator or Diagnostic Reference Standard	Health Outcomes & Harms	Timing	Setting	Study Design
<p>KQ 1-3: Drug treatment efficacy, comparative effectiveness & harms on cognition, function & quality of life</p>	<p>Adults with CATD ≥50 years of age <u>Patient characteristics to be assessed as possible treatment effect modifiers</u> Age Sex Race/ethnicity Depression Pre-treatment cognitive or functional level/CATD stage Living setting</p>	<p><u>Prescription pharmacologic (drug) treatment</u> Cholinesterase inhibitors NMDA antagonists <u>Nonprescription pharmacologic (drug) treatment</u> OTC supplements Vitamins Herbals</p>	<p><u>For efficacy comparisons</u> Placebo Other inactive control <u>For comparative effectiveness comparisons</u> Prescription drug treatment Nonprescription drug treatment Nondrug treatment</p>	<p><u>Efficacy and comparative effectiveness:</u> Change in patient cognition (global screen, multidomain, memory, executive function, language, attention), function, or QOL on validated test Change in disease stage based on validated test Change in patient “at home” IADL or ADL function Change in patient residence to different level of independence <u>Harms:</u> <i>General</i> FDA defined SAEs Withdrawals due to AEs <i>Psychiatric</i> Somnolence Confusion/Delirium <i>Nonpsychiatric</i> Falls Extrapyramidal symptoms Stroke Mortality (all-cause, CVD, non-CVD)</p>	<p>≥24 weeks</p>	<p><u>Cognitive outcomes:</u> Community-dwelling Assisted living <u>Functional & QOL</u> <u>outcomes:</u> Community-dwelling Assisted living Nursing home</p>	<p><u>Efficacy and comparative effectiveness:</u> RCT, CCT, systematic review of RCTs or CCTs <u>Harms:</u> RCT, CCT, controlled prospective cohort studies with ≥1000 participants, systematic review of any of these study designs</p>

KQ	Population	Intervention	Treatment Comparator or Diagnostic Reference Standard	Health Outcomes & Harms	Timing	Setting	Study Design
<p>KQ 4-6: Drug treatment efficacy, comparative effectiveness & harms on BPSD</p>	<p>Adults with CATD ≥50 years of age with BPSD (studies specified BPSD inclusion criterion) <u>Patient characteristics to be assessed as possible treatment effect modifiers</u> Age Sex Race/ethnicity Pre-treatment cognitive or functional level/CATD stage Pre-treatment BPSD severity Living setting</p>	<p><u>Prescription pharmacologic treatment</u> Cholinesterase inhibitors NMDA antagonists Antipsychotics, second generation (any) and first generation (only haloperidol) Antidepressants Anti-seizure/mood stabilizers Anxiolytics, benzodiazepine Anxiolytics, other Hormonal agents (Disinhibited sexual behavior only) Cannabinoids Combinations <u>Nonprescription pharmacologic treatment</u> OTC supplements Vitamins Herbals</p>	<p><u>Efficacy comparisons</u> Placebo Other inactive control <u>Comparative effectiveness comparisons</u> Prescription drug treatment Nonprescription drug treatment Nondrug treatment</p>	<p><u>Efficacy and comparative effectiveness:</u> <i>Primary</i> Change in the frequency and/or severity of patient BPSD* on validated tests Agitation/ aggression Psychosis Depression Anxiety Disinhibited sexual behavior Change in patient QoL on validated test Change in validated general behavior scale <i>Secondary</i> Change in caregiver/staff outcomes on validated tests Depression Global stress/distress QOL Burden <u>Harms:</u> <i>General</i> FDA defined composite SAE outcome Withdrawals due to AE <i>Psychiatric</i> Somnolence Confusion/Delirium <i>Nonpsychiatric</i> Falls Extrapyramidal symptoms Stroke Mortality (all-cause, CVD, non-CVD)</p>	<p><u>Agitation, aggression, psychosis or Disinhibited sexual behavior outcomes:</u> ≥2 weeks <u>Depression or anxiety outcomes:</u> ≥24 weeks</p>	<p>Community-dwelling Assisted living Nursing home</p>	<p><u>Efficacy and comparative effectiveness:</u> RCT, CCT, systematic review of RCTs or CCTs <u>Harms:</u> RCT, CCT, controlled prospective cohort studies ≥1000 participants, systematic review of any of these study designs</p>

KQ	Population	Intervention	Treatment Comparator or Diagnostic Reference Standard	Health Outcomes & Harms	Timing	Setting	Study Design
KQ 7-8: Diagnostic test accuracy & harms (also see Table 2 below)	<p><u>Cognitive tests:</u> Adults ≥ 50 years of age with suspected CATD</p> <p><u>Biomarker tests only:</u> Adults ≥ 50 years of age with clinical syndrome of CATD</p> <p><u>Patient characteristics to be assessed as possible effect modifiers of diagnostic test accuracy</u> Age Sex Race/ethnicity Education Depression</p> <p>Pre-test cognitive or functional level/ CATD stage</p>	<p><u>Brief, validated cognitive tests:</u> Global (brief screens, multi-domain batteries)</p> <p>Single domain tests (memory, executive, language, attention)</p> <p><u>Biomarker tests:</u> <i>Brain imaging</i> CT/MRI</p> <p>Medial temporal atrophy/hippocampal volume Cortical thickness DTI indices PET ^{18}F-FDG PET Amyloid PET ^{11}C-PiB and fluorinated tracers (e.g. florbetapir, flutemetamol, florbetaben)</p> <p>Tau PET fMRI: resting state and task specific activation SPECT: resting state cerebral perfusion <i>CSF tests</i> Aβ42 Aβ42/Aβ40 ratio t-tau p-tau t-tau/Aβ42 ratio p-tau/Aβ42 ratio neurofilament light protein <i>Blood tests</i> Aβ42 Aβ42/Aβ40 ratio APP <i>Combinations</i></p>	<p><u>Cognitive tests:</u> Full clinical evaluation and/or neuropsychological testing with explicit diagnostic criteria</p> <p><u>Biomarker tests:</u> Postmortem neuropathological confirmation of AD</p>	<p><u>Accuracy and comparative accuracy (e.g., TP, FP, TN, FN, sensitivity, specificity, PPV, NPV)</u> Of cognitive tests for confirming clinical syndrome of CATD Of biomarker tests for confirming that etiology of CATD is AD</p> <p><u>Harms:</u> <i>Psychological or behavioral</i> True positive: Labeling stigma False positive: Incorrect diagnosis Labeling stigma Side effects of unneeded interventions (e.g., restrictions on independence)</p> <p>False negative: Unexplained symptoms Failure to make appropriate interventions (e.g., safety precautions, future planning)</p> <p>Any test result: Patient or caregiver mental distress</p> <p><i>Physical</i> Directly from diagnostic tests: Pain Infection Headache Radiation</p>	Any	Community-dwelling Assisted living	<p><u>Accuracy and comparative accuracy:</u> Controlled observational studies (i.e., cross-sectional, retrospective cohort, case control); systematic review of controlled observational studies</p> <p><u>Harms:</u> Controlled observational studies (i.e., cross-sectional, retrospective cohort, case control, prospective cohort); systematic review of controlled observational studies</p>

*For this report, two psychological symptoms that are components of BPSD have been excluded due to their coverage in recent, high quality systematic reviews – apathy and sleep disturbances.^{18, 19} In addition, wandering was also eliminated, as this symptom is usually treated with nonpharmacologic interventions, which are not covered as interventions in this review.

†Strength of evidence (SOE) will be evaluated for the 1-2 most commonly reported validated treatment efficacy outcomes for each of the following test categories: disease stage, global cognitive screening tests, global multidomain cognitive tests, memory, executive functioning, language, attention, function, quality of life, BPSD agitation/aggression, and the harms outcome of serious adverse events. Additional treatment outcomes will be considered for SOE grading when available data allow. For diagnostic tests, SOE will be graded for the 1-2 most commonly reported validated tests for each of the following categories: global cognitive screening tests, global multidomain cognitive tests, memory, MRI, PET, and CSF tests. Additional diagnostic testing outcomes will be considered for SOE grading when available data allow.

A β = beta amyloid, AD = Alzheimer's dementia, ADL = activities of daily living, AE = adverse events, APOE = apolipoprotein E, APP = amyloid precursor protein, BPSD = behavioral and psychological symptoms of dementia, CATD = clinical Alzheimer's-type dementia, CCT = controlled clinical trial, CSF = cerebrospinal fluid, CT = computed tomography, CVD = cardiovascular disease, DTI = diffusion tensor imaging, FDG = fluorodeoxyglucose, fMRI = functional magnetic resonance imaging, FN = false negative, FP = false positive, IADL = instrumental activities of daily living, MCI = mild cognitive impairment, MRI = magnetic resonance imaging, NMDA = N-methyl-D-aspartate, NPV = negative predictive value, OTC = over-the-counter, PET = positron emission tomography, PPV = positive predictive value, p-tau = abnormally phosphorylated tau, QOL = quality of life, RCT = randomized clinical trial, ROC = receiver operating characteristic, SAE = serious adverse events, SPECT = single-photon emission computed tomography, TN = true negative, TP = true positive, t-tau = total tau

Table 2. Prescription Drugs Used for Treatment of CATD Cognition, Function, Quality of Life or BPSD

Class of drug	Drug name(s)
Cholinesterase inhibitor	Donepezil*, rivastigmine*, galantamine*
NMDA receptor antagonist	Memantine*
Cholinesterase inhibitor/NMDA receptor antagonist combination	Donepezil/ Memantine*
1 st generation (typical) antipsychotic	only Haloperidol
2 nd generation (atypical) antipsychotic	e.g., Risperidone, quetiapine, olanzapine, aripiprazole, clozapine
Anti-depressant, selective serotonin-reuptake inhibitor (SSRI)	e.g., Citalopram, escitalopram, sertraline, fluoxetine, fluvoxamine, paroxetine
Anti-depressant, serotonin-norepinephrine reuptake inhibitor (SNRI)	e.g., Duloxetine, venlafaxine
Anti-depressant, other†	e.g., Trazodone, bupropion, mirtazapine
Anti-seizure/mood stabilizer	e.g., Valproate, gabapentin, carbamazepine, lamotrigine
Anti-anxiety, benzodiazepine	e.g., Clonazepam, diazepam, lorazepam, temazepam, alprazolam
Anti-anxiety, other	Buspirone
Mixed	Dextromethorpan/ Quinidine
Hormones (antiandrogens, estrogens, gonadotropin-releasing hormone analogues)	e.g., medroxyprogesterone acetate, cyproterone acetate, leuprolide
Cannabinoids	e.g., medical marijuana

*US FDA approved indication for Alzheimer's dementia

†Excludes MAO-inhibitor, tricyclic and tetracyclic antidepressants.

BPSD = behavioral and psychological symptoms of dementia, CATD = clinical Alzheimer's-type dementia, NMDA = N-methyl-D-aspartate, SSRI = selective serotonin reuptake inhibitor, SNRI = selective norepinephrine reuptake inhibitor

Table 3. Cognitive Tests to be Assessed for Diagnostic Accuracy

Cognitive Test Categories	Cognitive Test Names	Cognitive Domains Evaluated	Approximate Administration Time
Screening and Brief (< 30min) Global Instruments	Cognitive Abilities Screening Instrument (CASI)	Global	15-20 min
	Mini-Cog	Global	< 5 min
	Mini-Mental State Exam (MMSE)	Global	5-10 min
	Montreal Cognitive Assessment (MoCA, also MoCA-Blind version)	Global	10-15 min
	St. Louis University Mental Status (SLUMS)	Global	5-10 min
	Short Test of Mental Status (STMS)	Global	5-10 min
	Telephone Interview for Cognitive Status (TICS & TICS-M)	Global	10 min
	Clock Drawing	Global	<5 min
Short Batteries	Alzheimer's Disease Assessment Scale-Cognition (ADAS-Cog)	Global	30-40 min
	CERAD Battery	Global	30-40 min
	Mattis Dementia Rating Scale (DRS & DRS-2)	Global	30-40 min (both)
	Repeatable Battery for the Assessment of Neuropsychological Status (RBANS, multiple versions)	Global	30-40 min
	Computer administered (e.g., CogState, CANS-MCI)	Global	Varies by test
Types of Individual Tests Administered as Part of Longer Battery	Trail making tests (e.g., TMT part B, DKEFS)	Executive Function	Varies by test
	Coding tasks (e.g., Digit symbol [WAIS], symbol digit)	Executive Function	Varies by test
	Design and figure fluency tasks (e.g., DKEFS)	Executive Function	Varies by test
	Concept formation switching and rule attainment (e.g., Wisconsin Card Sort)	Executive Function	Varies by test
	Figure recall tasks (e.g., BVRT, RCFT, Taylor)	Nonverbal memory	Varies by test
	List-learning tests (e.g., CVLT, Buschke, Hopkins, RAVLT)	Verbal memory	Varies by test
	Prose/paragraph recall (e.g., Boston story, Logical Memory)	Verbal memory	Varies by test
	Confrontation naming (e.g., BNT)	Language	Varies by test
	Verbal fluency-letter/phonemic (e.g., FAS, CFL, includes COWAT)	Language, Executive Function	Varies by test
	Verbal fluency-category/semantic (e.g., names, animals)	Language, Executive Function	Varies by test

Abbreviations: ADAS-Cog = Alzheimer's Disease Assessment Scale-Cognition, BNT = Boston Naming Test, BVRT = Benton Visual Retention Test, CASI = Cognitive Abilities Screening Instrument, CERAD = Consortium to Establish a Registry for Alzheimer's Disease, COWAT = Controlled Oral Word Association Test, CPT = Continuous Performance Test, CVLT = California Verbal Learning Test, DKEFS = Delis-Kaplan Executive Function System, DRS = Dementia Rating Scale, MMSE = Mini-Mental State Exam, MoCA = Montreal Cognitive Assessment, RAVLT = Rey Auditory Verbal Learning Test, RBANS = Repeatable Battery for the Assessment of Neuropsychological Status, RCFT = Rey-Oosterrieth Complex Figure Test, SLUMS = St. Louis University Mental Status, STMS = Short Test of Mental Status, TICS = Telephone Interview for Cognitive Status, TMT = Trail Making Test, TOVA = Tests of Variables of Attention, WAIS = Wechsler Adult Intelligence Scale

III. Analytic Framework

Figure 1. Analytic Framework for Key Questions 1-3: Efficacy, comparative effectiveness and harms of pharmacological treatment for treatment of cognition, function and quality of life in patients with CATD. This figure depicts key questions 1-3 within the context of the PICOTS described in the previous section. In general, the figure illustrates how prescription or nonprescription drug treatment versus control may result in final health outcomes such as changes in cognition, function and quality of life. It also illustrates how adverse events may occur. Finally, it illustrates how the effect of drug treatments versus control on cognitive, functional, quality of life, and harms outcomes may vary as a function of different patient characteristics (possible effect modifiers).

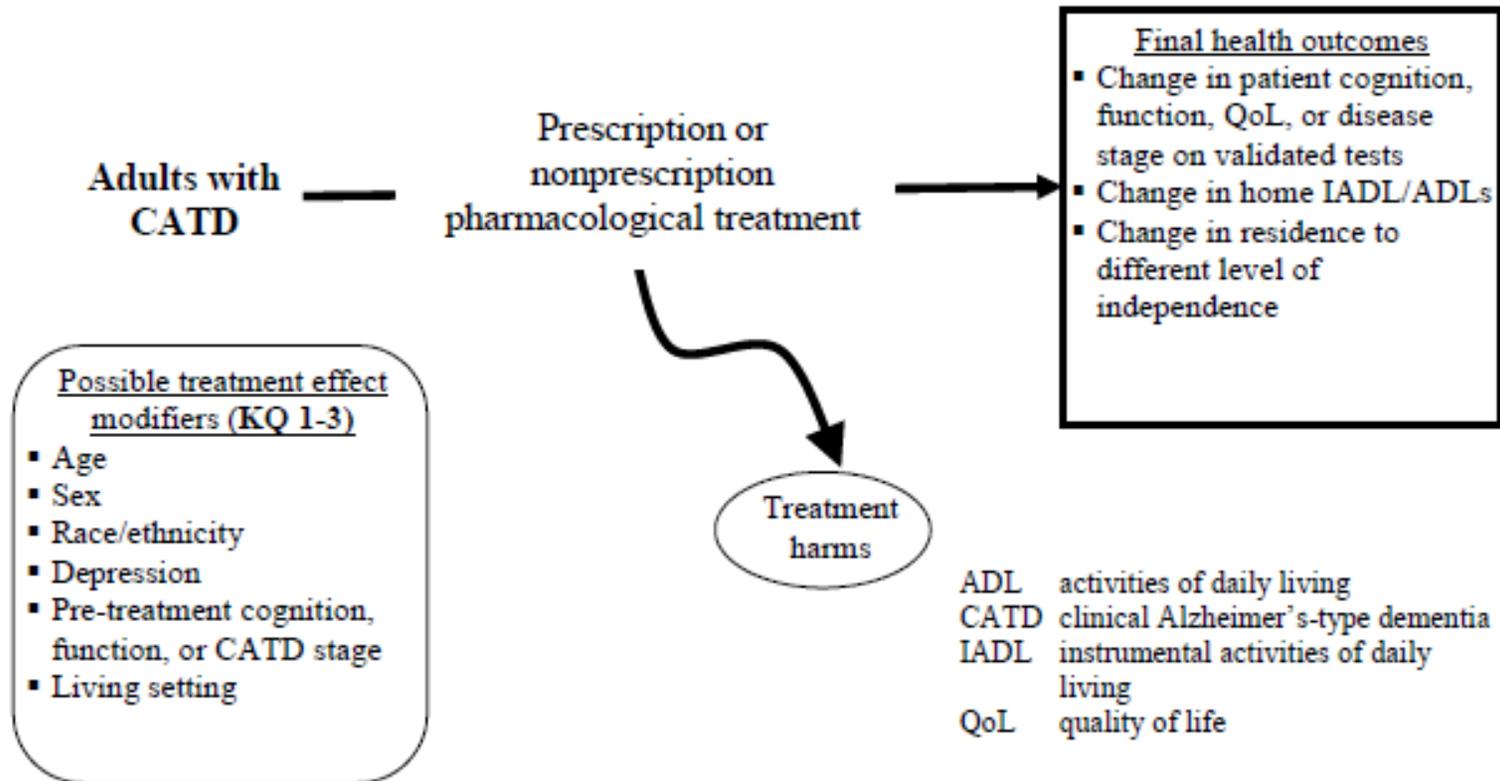


Figure 2. Analytic Framework for Key Questions 4-6: Efficacy, comparative effectiveness and harms of pharmacological treatment for behavioral and psychological symptoms of dementia (BPSD) in patients with CATD who have BPSD. This figure depicts key questions 4-6 within the context of the PICOTS described in the previous section. In general, the figure illustrates how prescription and nonprescription drug treatment versus control may result in final health outcomes such as changes in BPSD, patient quality of life, and caregiver outcomes. It also illustrates how adverse events may occur. Finally, it illustrates how the effect of drug treatments versus control on BPSD, patient quality of life, and caregiver outcomes may vary as a function of different patient characteristics (possible effect modifiers).

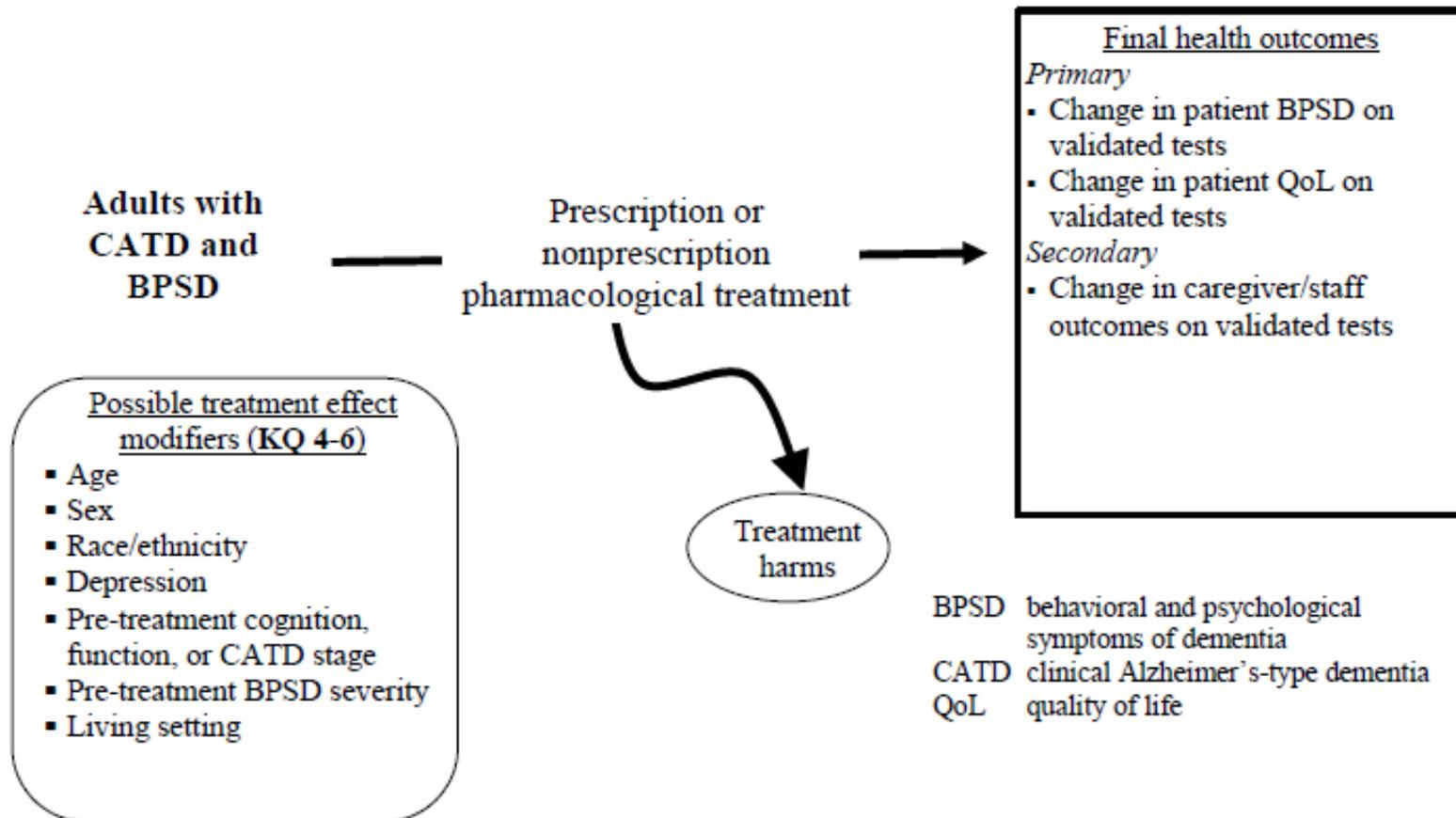
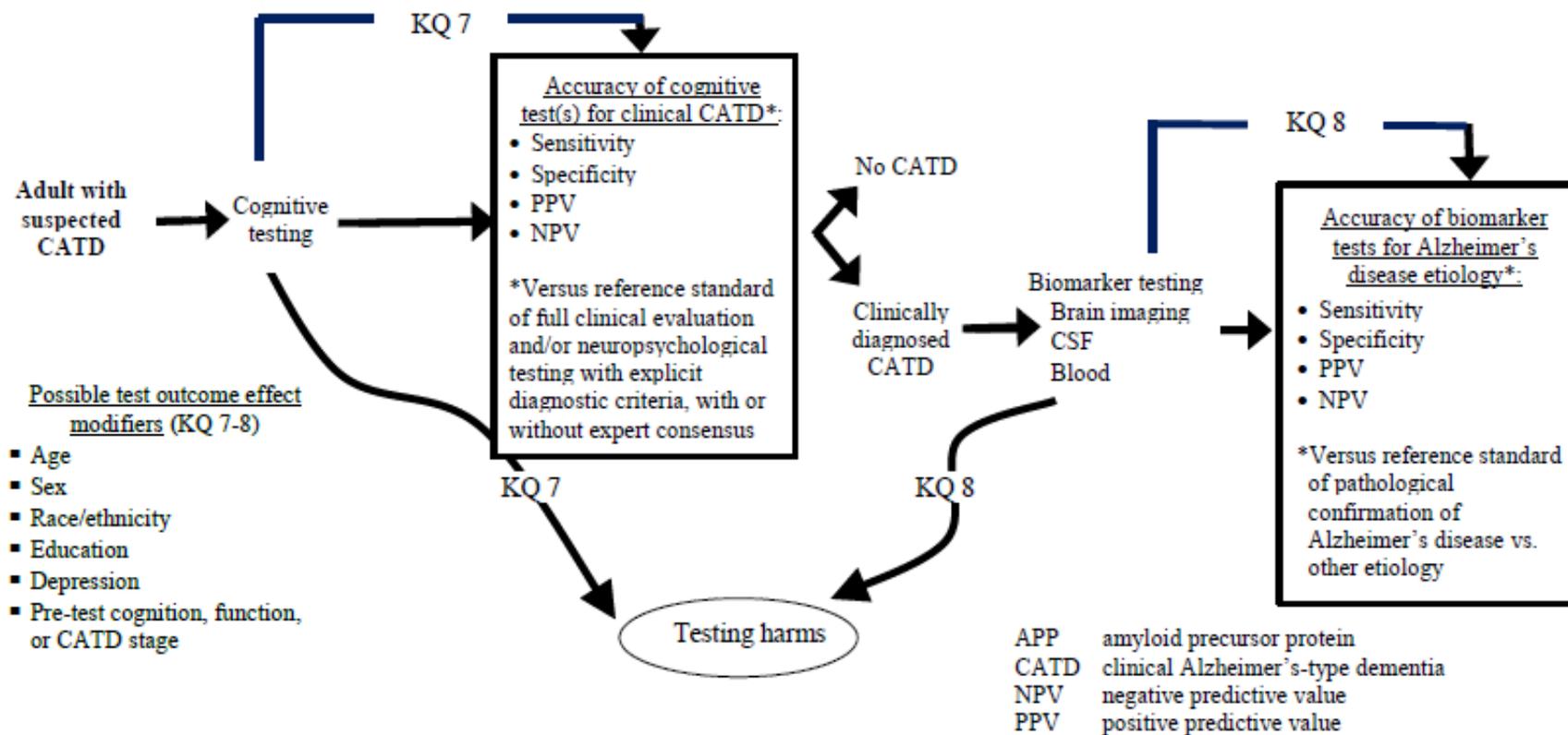


Figure 3. Analytic Framework for Key Questions 7-8: Accuracy, comparative accuracy and harms of diagnostic testing in patients with suspected CATD. This figure depicts key questions 7-8 within the context of the PICOTS described in the previous section. In general, the figure illustrates how cognitive tests may identify patients with clinically diagnosed CATD and how, in patients with clinically diagnosed CATD, biomarker tests may identify those with pathologically confirmed AD. It also illustrates how adverse events may occur. Finally, it illustrates how the accuracy of diagnostic testing may vary as a function of different patient characteristics (possible effect modifiers).



IV. Methods

A. Criteria for Inclusion/Exclusion of Studies in the Review

Table 4. Study Inclusion Criteria

Category	Entry Criteria
Study Population	Adults aged ≥ 50 years KQ 1-3: CATD KQ 4-6: CATD with BPSD (study specified BPSD inclusion criterion) KQ 7: Suspected CATD KQ 8: Clinically diagnosed CATD Exclude: Normal cognition, MCI, or dementia known to be secondary solely to TBI, FTD, PD, LBD, stroke or other non-AD etiology
Study Objectives	KQ 1-3: Evaluate efficacy, comparative effectiveness, and harms of pharmacologic treatment for CATD targeted for symptoms of cognition, function, and quality of life KQ 4-6: Evaluate efficacy, comparative effectiveness, and harms of pharmacologic treatment for CATD targeted for symptoms of BPSD KQ 7: Evaluate diagnostic accuracy, comparative accuracy and harms of cognitive tests for the reference standard of the clinical diagnosis of CATD KQ 8: Evaluate diagnostic accuracy, comparative accuracy and harms of biomarker tests for the reference standard of pathologically confirmed AD to determine whether AD is the underlying etiology of clinically diagnosed CATD KQ 1a, 2a, 3d, 4a, 5a, 6d, 7a, 8a: Evaluate possible effect modifiers of CATD pharmacological treatment efficacy, and comparative efficacy; and of cognitive and biomarker diagnostic test accuracy
Study Design	KQ 1-6: Treatment efficacy and comparative effectiveness: RCT or CCT, systematic review of RCTs or CCTs that assessed ROB of included studies using validated tools. Treatment harms: RCT, CCT, , controlled prospective cohort studies of ≥ 1000 participants (will consider smaller cohort studies if evidence from larger cohort studies is insufficient); systematic review of RCTs, CCTs, or large, controlled prospective cohort studies that assesses ROB of included studies using validated tools. KQ 7-8: Controlled observational studies (i.e., cross-sectional, retrospective cohort, case control) with ≥ 25 participants*, systematic review of these study designs that assessed ROB of included studies using validated tools. For these diagnostic key questions, we will consider including studies with fewer than 25 participants if data from larger studies are unavailable or provide insufficient evidence to draw conclusions about their test performance.

Category	Entry Criteria
Interventions	<p>KQ 1-3: For targeting cognitive, functional and quality of life outcomes: cholinesterase inhibitors, NMDA antagonists,</p> <p>KQ 4-6: For targeting BPSD and quality of life outcomes: Prescription drugs: cholinesterase inhibitors, NMDA antagonists, antipsychotics, antidepressants, anxiolytics, antiseizure/mood stabilizer, hormones (disinhibited sexual behavior only), cannabinoids, combinations Orally ingested over-the-counter supplements, vitamins, herbal medications</p> <p>KQ 7: Diagnostic tests: cognitive tests that are English-language and available in the U.S. (brief global screens, short multi-domain batteries, individual domain tests [memory, executive, language])</p> <p>KQ 8: Brain imaging tests in contemporary use (CT, MRI, PET, fMRI, SPECT), CSF tests, blood tests</p>
Comparisons	<p>KQ 1-6: Placebo, other inactive control, prescription pharmacological treatment, nonprescription pharmacological treatment, nonpharmacological treatment</p> <p>KQ 7: Cognitive tests to diagnose clinical CATD, with clinical diagnosis based on full clinical evaluation and/or neuropsychological testing with explicit diagnostic criteria (e.g., DSM-IV, DSM-5, ICD, or NINCDS-ADRDA), with or without expert consensus</p> <p>KQ 8: “Normal level” on biomarker test, other biomarker tests</p>
Outcomes	<p>KQ 1-3: Patient-related outcomes: Change in cognition (global, memory, executive function, language, attention), function, quality of life, and disease stage on validated tests; harms (FDA defined composite SAEs, withdrawals due to AEs, somnolence, confusion/delirium, falls, extrapyramidal symptoms, stroke, mortality)</p> <p>KQ 4-6: Patient-related outcomes: Change in BPSD and quality of life on validated tests; harms as listed for KQ 1-3</p> <p>Caregiver/staff outcomes: depression, QoL, global stress/distress/burnout, burden</p> <p>KQ 7-8: Sensitivity, specificity, positive predictive value, and negative predictive value of specific diagnostic test cut-off values, or data which enable their calculation</p>
Possible treatment/test outcome modifiers	<p>KQ 1-8: pretreatment age, race/ethnicity, sex, depression, pretreatment/pretesting cognitive or functional level/CATD stage</p> <p>KQ 1-6: living setting</p> <p>KQ 4-6 only: pre-treatment BPSD severity, living setting</p> <p>KQ 7-8 only: education</p> <p>Indicate whether reported subgroup analyses or tests of interaction were planned <i>a priori</i> versus post hoc, as post hoc analyses are at greater risk for false positive findings.</p>

Category	Entry Criteria
Timing	KQ 1-3: Cognitive, functional, quality of life and harms outcomes: ≥ 24 weeks KQ 4-6: BPSD and harms outcomes: ≥ 2 weeks for outcomes of agitation/aggression, psychosis, or disinhibited sexual behavior; otherwise ≥ 24 weeks KQ 4-6: Quality of life and harms outcomes: ≥ 24 weeks KQ 7: Cognitive diagnostic testing: ≤ 6 months between cognitive test and clinical diagnosis of CATD KQ 8: Any, including pre- or post-mortem, and any interval between biomarker and pathological assessment. For CSF or blood biomarkers with limited post-mortem stability, inclusion may need to be restricted to a maximum duration post-mortem.
Setting	KQ 1-3: For cognitive outcomes: community-dwelling, assisted living KQ 1-6: For functional, quality of life, and BPSD outcomes: community-dwelling, assisted living, nursing home KQ 7-8: For diagnostic testing: community-dwelling, assisted living
Publication type	Published in full text in peer reviewed journals
Language of Publication	English only, due to resource limitations

*We will exclude controlled observational studies with $N < 25$, since these small observational studies are often lower in quality, inadequately powered on their own, and inappropriate to pool. Regarding study quality, the quality of the evidence is low since statistical adjustment is not possible with very small sample sizes because models become unstable when the number of cases is not much larger than the number of covariates (e.g. 10 to 15-fold). Regarding statistical power, without pooling, studies with 12 participants per arm cannot reject null hypotheses even when true associations are large (i.e. Cohen's $D = 1.2$ for $N=24$ at 80% power). Regarding appropriateness for pooling, small studies are prone to overestimate the magnitude of an association, potentially exaggerating the accuracy and harms of diagnostic testing, and biasing the pooled estimates.²⁰

AD = Alzheimer's disease, APOE = apolipoprotein E, BPSD = behavioral and psychological symptoms of dementia, CATD = clinical Alzheimer's-type dementia, CCT = controlled clinical trial, CSF = cerebrospinal fluid, CT = computed tomography, DSM = Diagnostic and Statistical Manual of Mental Disorders, fMRI = functional magnetic resonance imaging, FTD = frontotemporal dementia, ICD = International Classification of Disease, LBD = Lewy body dementia, MRI = magnetic resonance imaging, NINDCS-ADRDA = National Institute of Neurological and Communicative Disorders and Stroke and Alzheimer's Disease and Related Disorders, NMDA = N-methyl-D-aspartate, NPV = negative predictive value, PD = Parkinson's disease, PET = positron emission tomography, PPV = positive predictive value, RCT = randomized controlled trial, ROC = receiver operating characteristics, SAE = serious adverse events, SPECT = single-photon emission computed tomography, TBI = traumatic brain injury.

B. Searching for the Evidence: Literature Search Strategies for Identification of Relevant Studies to Answer the Key Questions

Electronic database search: We will search Ovid Medline, Ovid Embase, PsycINFO, and the Cochrane Central Register of Controlled Trials (CENTRAL) to identify randomized controlled trials, nonrandomized controlled clinical trials, observational studies, and systematic reviews published and indexed in these bibliographic databases. The search algorithm will include relevant controlled vocabulary and natural language terms for the concepts of Alzheimer's disease, mild cognitive impairment, dementia, drug treatment, cognitive tests, biomarkers, and diagnostic accuracy, and will be combined with validated filters to select study designs (Appendix 2). We will supplement our electronic database searching with backward citation searches of included studies and of highly relevant recent systematic reviews.

Grey literature search: We will search ClinicalTrials.gov to identify relevant completed studies that did not report outcomes and analyses in the published literature to help assess publication and reporting bias, and to identify and track ongoing studies that may contribute information to address the key questions in the future. To solicit Pharmaceutical Manufacturer protocols with additional information about published or unpublished drug studies, AHRQ will open a Supplemental Evidence and Data for Systematic Reviews (SEADS) portal and send out a notification through its listserv. Recommendations of additional potentially eligible references will be sought from the Technical Expert Panel.

We will update both the electronic database and grey literature searches while the draft report is under peer/public review.

C. Study Selection

We will review bibliographic database search results for individual studies and systematic reviews relevant to our PICOTS framework and study-specific entry criteria (Table 4). References identified from these electronic databases and from citation searches of systematic reviews, peer and public review or through the SEADs portal will be pooled and deduplicated in EndNote (EndNote X7 and X8, Clarivate Analytics, Philadelphia, PA). Search results then will be downloaded into Distiller (DistillerSR, Evidence Partners, Ottawa, Canada) where they will be further deduplicated.

Titles and abstracts will be reviewed by two independent investigators to identify studies meeting PICOTS framework and inclusion/exclusion criteria. Studies considered ineligible by both investigators will be excluded from the review, while those considered potentially eligible by at least one of these investigators will be forwarded for full text screening. All studies forwarded for full-text screening will be independently evaluated by two investigators to

determine if inclusion criteria are met and, if excluded, to determine the reason(s) for exclusion. Reasons for exclusion at this stage will be reported. Differences in screening decisions will be resolved by consultation between investigators, and, if necessary, consultation with a third investigator. Before and throughout screening, team members will meet regularly to discuss study entry criteria, the screening process and issues as they arise to ensure consistency within and between investigators.

D. Assessment of Methodological Risk of Bias of Individual Studies and Quality of Systematic Reviews

Based upon AHRQ guidance, we will assess each eligible individual study for risk of bias in its design, analysis and reporting.^{21, 22} Two investigators will independently assess each study for bias in several different domains, and then, considering these assessments, also rate its overall risk of bias.

For individual CATD treatment studies, for each outcome of interest (i.e., stage, global cognition and cognition domains, function, quality of life, and BPSD symptoms), risk of bias for each of the following domains will be rated as high, medium or low using a risk of bias tool (Appendix 4):

- Selection bias: adequacy of randomization method (RCT), accounting for imbalance in prognostic variables (observational studies)
- Attrition bias: loss to follow-up, overall and differentially between treatment or diagnostic testing groups
- Detection bias: outcome measurement quality, outcome assessor masking
- Performance bias: intention to treat/test analysis, adjustment for potential confounding variables, participant masking to treatment assignment
- Reporting bias: selective reporting of outcomes

For individual diagnostic test studies, for each test of interest (e.g., brief cognitive tests, biomarkers), risk of bias for each of the following domains will be rated as high, low or unclear using the QUADAS-2 tool:²³

- Patient selection: consecutive or random sample enrolled, avoided case-control design, avoided improper exclusions
- Index test: index test interpreted without knowledge of the reference standard, any index test threshold prespecified
- Reference standard: reference standard likely to correctly classify target condition (i.e., AD), reference standard results interpreted without knowledge of results of index test
- Flow and timing: appropriate interval between index test and reference standard, all patients received same reference standard, all patients included in analysis

Considering the different domain risk of bias ratings, each investigator will independently rate overall study risk of bias for each individual study as high, medium or low. Investigators will consult to reconcile any discrepancies in risk of bias ratings for both individual domains and overall.

E. Systematic reviews that directly address a question in our review and assessed risk of bias for included individual studies using appropriate validated tools will be assessed for quality. We will use AMSTAR 2 criteria for systematic reviews of CATD treatment studies,²⁴ and modified AMSTAR 2 criteria for systematic reviews of diagnostic test studies. **Data Abstraction and Data Management**

For all eligible studies, one investigator will extract selected data and a second reviewer will check the accuracy of extracted data.

Studies determined to be high risk of bias will have only limited data extracted. Information extracted from both treatment and diagnostic studies will include author, year of publication, population description and number enrolled, study design, and funding source. Information extracted only from treatment studies will include intervention, comparator, and types of treatment efficacy and harms outcomes. Information extracted only from diagnostic studies will include diagnostic test, reference test, and measures of diagnostic performance assessed.

Studies judged to have low to moderate risk of bias will undergo additional data extraction. Fields that will be extracted from all studies will include participant eligibility criteria, setting, and participant baseline characteristics (age, race/ethnicity, sex, depression, and pretreatment/pretesting cognitive and functional level/CATD stage).

Additional fields that will be extracted only from treatment studies will include intervention details (drug class, name, dose and delivery route), control intervention details, follow-up duration, living setting, and which validated efficacy, comparative effectiveness, and harms outcomes were reported. Based on the frequency with which they are reported, we will decide which specific stage, cognitive (global screen, global multidomain, memory, executive, language, attention), functional, quality of life, and behavioral outcome measures to extract. We anticipate extracting detailed efficacy results for the 1-2 most common measures for each outcome category, but will consider extracting data for additional outcome measures if data allow.

Additional fields that will be extracted only from diagnostic studies will include prevalence of reference condition in tested population, index test (e.g., specific cognitive test, brain imaging, CSF or blood test) and cut-off values used to categorize participants, specific reference standard (e.g., DSM, full clinical evaluation with pathologic confirmation) and methods of participant sampling and recruitment, time interval between measurements of index and reference test, and

sensitivity, specificity, true positives, true negatives, false positives, and false negatives at each combination of index and reference test threshold.

In addition, pretreatment BPSD severity will be extracted for behavioral treatment studies, and education will be extracted for diagnostic studies.

Systematic reviews determined to be high quality may be used to replace de novo data extraction processes for specific population/treatment/outcome comparisons that are sufficiently relevant. Individual studies in included systematic reviews will be tracked for contribution to unique population/treatment/outcome comparisons to avoid double-counting study results. Any reviews used to replace de novo data extraction will be supplemented by data extraction from eligible studies published after the search date of the review.

F. Data Synthesis

Results will be organized first by key question. Then, for key questions 1-3, results will be organized by treatment comparison, and then by targeted treatment outcome (disease stage, cognition [global screen, global multidomain, memory, executive, language, attention], function, quality of life) and harms. Similarly, for key questions 4-6, results will be organized by treatment comparison, and then by targeted treatment outcome (agitation/aggression, psychosis, depression, anxiety, disinhibited sexual behavior, general behavior) and harms. For key questions 7-8, results will be organized by diagnostic test category (cognitive, brain imaging, CSF, blood), within each diagnostic test category by specific test, and then by diagnostic accuracy outcomes and harms. For studies with low and moderate risk of bias, we first will describe the results in evidence tables.

When a comparison is adequately addressed by a previous high quality systematic review and no new studies are available, we will reiterate the conclusions drawn from that review. When new eligible trials were published since the search date of the prior review, previous systematic review data will be synthesized with data from these new trials.

For treatment studies, we will prioritize analyses of outcomes framed as responders, or improved or stable versus declined, or meeting an a priori established threshold for a clinically meaningful improvement where available. For these binary outcomes, we will calculate risk ratios (RR) and absolute risk differences (RD) with the corresponding 95 percent confidence intervals (CI) for binary outcomes. For continuous outcomes, we will calculate weighted mean differences (WMD) and/or standardized mean differences (SMD) with the corresponding 95 percent CIs. For diagnostic test studies, we will report/calculate sensitivity, specificity, positive and negative likelihood ratios and their 95 percent confidence intervals for each population and combination of index and reference test thresholds.

We will assess individual study clinical and methodological heterogeneity to determine appropriateness of pooling data.²⁵ For treatment studies, we will evaluate clinical heterogeneity by whether the populations, treatment interventions and controls, and outcomes are comparable. For diagnostic studies, we will evaluate clinical heterogeneity by whether the populations, index test thresholds, reference test thresholds, and measures of test performance are comparable. When we judge that data are appropriate for pooling, we will synthesize data using a generalized linear mixed model approach.²⁶ If the analyses yield substantial heterogeneity, we will stratify the results to assess treatment outcomes and diagnostic accuracy, respectively, based on patient or study characteristics and/or explore sensitivity analysis.

When data allow, we will perform stratified analyses to evaluate a priori selected possible effect modifiers of CATD treatment efficacy, comparative effectiveness, and cognitive and biomarker testing diagnostic accuracy, comparative accuracy and harms. For all key questions, we will examine age, sex, race/ethnicity, depression, and pretreatment cognitive or functional status/CATD stage. For KQ 4-6 only, we will examine pretreatment BPSD severity. For KQ 1-6 only, we will examine living setting. For KQ 7-8 only, we will examine education. We will record whether the possible effect modifiers were identified a priori. In addition, we will examine if treatment efficacy differs as a function of drug dose, treatment duration, and treatment follow-up duration, and if diagnostic accuracy differs as a function of the time interval between diagnostic test measurement and the determination of the reference diagnosis.

G. Grading the Strength of Evidence (SOE) for Major Comparisons and Outcomes

For the key questions on the benefits and harms of CATD treatment, SOE will be graded²⁷ for the direction of the treatment effect (i.e., whether benefits or harms are greater or are not different between one treatment and another). SOE will be evaluated for the 1-2 most commonly reported validated treatment efficacy outcomes for each of the following test categories: stage, global cognitive screening tests, global multidomain cognitive tests, memory, executive functioning, language, attention, function, quality of life, BPSD agitation/aggression, and serious adverse events. Additional treatment outcomes may be considered for SOE grading based on available data.

For the key questions on the accuracy of diagnostic tests, SOE will be graded²⁸ for the magnitude of the test sensitivity, specificity, positive likelihood ratio, and negative likelihood ratio of the 1-2 most commonly reported validated tests for each of the following categories: global cognitive screening tests, global multidomain cognitive tests, memory, MRI, PET imaging, and CSF tests. Additional diagnostic testing outcomes will be considered for SOE grading based on available data.

Two investigators will independently assess five required domains (listed below) and other possible factors to grade the strength of evidence for each treatment comparison and diagnostic test for included studies. Differences in individual domain ratings and overall strength of evidence (SOE) grades will be resolved by consultation between investigators, and, if necessary, consultation with a third investigator. The five required strength of evidence domains will be: (1) study limitations; (2) directness; (3) consistency; (4) precision; and (5) reporting bias. When considered appropriate to a body of evidence, we also will consider dose-response association across or within studies, unmeasured confounders that would decrease an effect, and strength of association.²⁹

Study limitations will be rated as low, medium, or high based on the design and risk of bias of the aggregated individual studies within an evidence base. Directness will be rated as either direct or indirect based on whether the evidence directly links the intervention to the primary outcome of interest for the review. Because patients with suspected or confirmed CATD may not be capable of reliable self-reporting of outcomes, results reported by caregivers will not be downgraded for indirectness. Because our primary question about the diagnostic tests is about their accuracy and not whether they impact clinical outcomes, results for sensitivity, specificity, positive predictive value and negative predictive value also will not be downgraded for indirectness. Consistency within an evidence base will be rated as consistent or inconsistent based on whether treatment effects or diagnostic test performance from multiple studies are similar. For treatment effects, we will assess consistency in direction (effect estimates on same side of no effect or of a minimally important difference, if one is available), while for diagnostic test performance, we will assess consistency of the magnitude of test results (range of estimates). An evidence base will not be rated inconsistent if differences in results may be accounted for by heterogeneity in study characteristics. When evidence is based on a single study, regardless of its size or the number of participating study centers, consistency will be rated as unknown. Precision is the degree of certainty around an outcome effect estimate based on the sufficiency of the total sample size and/or number of events. Precision will be rated as precise or imprecise based on the degree of certainty surrounding each effect estimate. An imprecise estimate is one for which the confidence interval is wide enough to include clinically distinct conclusions regarding the direction of the effect (for treatment benefits or harms) or magnitude of the effect (for measures of diagnostic test performance) based upon established minimal detectable differences when available.

For treatment comparisons, the starting SOE grade for an evidence base derived from RCTs will be high, while the starting SOE grade for an evidence based derived from observational studies will be low. For diagnostic tests, the starting SOE grade for an evidence base will be high.²⁸ The final SOE grades then will be upgraded or downgraded based on the ratings for the individual SOE domains. Based on these elements, we will assess the overall SOE²⁹ for each comparison and outcome as:

- **High:** Very confident that estimate of effect lies close to true effect. Few or no deficiencies in body of evidence. Findings believed to be stable.
- **Moderate:** Moderately confident that estimate of effect lies close to true effect. Some deficiencies in body of evidence. Findings likely to be stable, but some doubt.
- **Low:** Limited confidence that estimate of effect lies close to true effect. Major or numerous deficiencies in body of evidence. Additional evidence necessary before concluding that findings are stable or that estimate of effect is close to true effect.
- **Insufficient:** No evidence, unable to estimate an effect, or no confidence in estimate of effect. No evidence is available or the body of evidence precludes judgment.

An overall rating of high SOE will be assigned when included studies were RCTs with a low risk of bias, and the results were consistent, direct, and precise. If SOE for a treatment-outcome or testing-outcome comparison is rated insufficient based on assessment of only low to moderate risk of bias studies, we will consider evaluating eligible high risk of bias studies that address the same treatment-outcome or testing-outcome comparison.

H. Assessing Applicability

Applicability of studies will be determined according to the PICOTS framework. Factors that may affect applicability include when studies have narrow eligibility criteria or when study population characteristics (e.g., age, race, sex, presence or lack of comorbidities, living setting, country of residence) differ from those in population studies of individuals with undiagnosed cognitive impairment or with clinically diagnosed CATD or AD. This limitation in applicability may be magnified if these population characteristics are associated with diagnostic test accuracy or treatment response. In addition, applicability of study findings may be limited if the studied diagnostic tests or treatments are not easily available in typical clinical settings.³⁰

V. References (See Appendix 1)

VI. Definition of Terms

AAFP	American Academy of Family Physicians
A β	beta amyloid
ACP	American College of Physicians
AD	Alzheimer's dementia
ADAS-Cog	Alzheimer's Disease Assessment Scale-Cognition
ADCS-ADL	Alzheimer's Disease Cooperative Study Activities of Daily Living Scale
ADL	activities of daily living
AE	adverse events
AHRQ	Agency for Healthcare Research & Quality

APOE	apolipoprotein E
APP	amyloid precursor protein
BNT	Boston Naming Test
BPSD	behavioral and psychological symptoms of dementia
BVRT	Benton Visual Retention Test
CASI	Cognitive Abilities Screening Instrument
CATD	clinical Alzheimer's-type dementia
CCT	controlled clinical trial
CERAD	Consortium to Establish a Registry for Alzheimer's Disease
CI	confidence intervals
COWAT	Controlled Oral Word Association Test
CPT	Continuous Performance Test
CSF	cerebrospinal fluid
CT	computed tomography
CVD	cardiovascular disease
CVLT	California Verbal Learning Test
DKEFS	Delis-Kaplan Executive Function System
DRS	Dementia Rating Scale
DTI	diffusion tensor imaging
DSM	Diagnostic and Statistical Manual of Mental Disorders
EPC	Evidence-based Practice Center
FDG	fluorodeoxyglucose
fMRI	functional magnetic resonance imaging
FN	false negative
FP	false positive
IADL	instrumental activities of daily living
ICD	International Classification of Disease
KI	key informant
KQ	key question
MCI	mild cognitive impairment
MMSE	Mini-Mental State Exam
MoCA	Montreal Cognitive Assessment
MRI	magnetic resonance imaging
NA	not applicable
NINCDS-ADRDA	National Institute of Neurological and Communicative Disorders and Stroke and Alzheimer's Disease and Related Disorders
NMDA	N-methyl-D-aspartate
NPI	Neuropsychiatric Inventory
NPI-Q	Neuropsychiatric Inventory Questionnaire
NPV	negative predictive value

OTC	over-the-counter
PET	positron emission tomography
PICOTS	Populations, interventions, comparators, outcomes, timing, and settings
PPV	positive predictive value
p-tau	abnormally phosphorylated tau
QOL	quality of life
QOL-AD	Quality of Life in Alzheimer's Disease
RAVLT	Rey Auditory Verbal Learning Test
RBANS	Repeatable Battery for the Assessment of Neuropsychological Status
RCFT	Rey-Oosterrieth Complex Figure Test
RCT	randomized clinical trial
RD	risk difference
ROC	receiver operating characteristic
RR	risk ratio
SLUMS	St. Louis University Mental Status
SMD	standardized mean difference
SPECT	single-photon emission computed tomography
SR	systematic review
STMS	Short Test of Mental Status
TEP	technical expert panel
TICS	Telephone Interview for Cognitive Status
TMT	Trail Making Test
TN	true negative
TOVA	Tests of Variables of Attention
TP	true positive
TR	topic refinement
t-tau	total tau
US FDA	United States Food and Drug Administration
WAIS	Wechsler Adult Intelligence Scale
WMD	weighted mean difference

VII. Summary of Protocol Revisions

Table 5. Changes Between Draft and Final Protocols

Section	Original Protocol	Revised Protocol	Rationale
General	Scope included evaluation of diagnostic accuracy of cognitive tests for clinical diagnosis of MCI, and of biomarkers to identify underlying Alzheimer's disease pathology in patients with clinical MCI.	Eliminated MCI.	Narrowed the scope to focus on CATD, as understanding the accuracy of diagnosing CATD is a higher clinical priority than is diagnosing MCI. The AHRQ posting document described the burden of dementia on the patient, family and society, and didn't discuss anything about MCI. Also, clinicians believe that risk of safety issues and consideration of available interventions is less important at MCI stage than at CATD stage, so that recognizing CATD currently is a higher priority than recognizing MCI.
General	Included nonpharmacologic interventions as a primary intervention of interest.	Excluded nonpharmacologic interventions for the Intervention category, but retain for as a comparison	Nonpharmacologic interventions will not be examined as an intervention in this review because they are being examined as an intervention in the NIA sponsored dementia care interventions review. However, nonpharmacologic interventions will be retained as a comparator to pharmacologic interventions in this review, as this comparator reflects the real-world decision-making of a clinician considering pharmacological treatment. If the NIA report scope expands to completely overlap this comparison, it will be removed from the current project via a protocol modification.

Section	Original Protocol	Revised Protocol	Rationale
Introduction	Methods language was in the Introduction.	Removed all methods language from introduction. Added it to the methods if it wasn't already there. If it was already there, just deleted from the introduction.	Methods language was inappropriate to include in the Introduction section.
Key Questions	Subquestions combined efficacy and comparative effectiveness.	Separated out efficacy and comparative effectiveness subquestions.	To clarify which comparisons were being evaluated for which treatments.
Key Questions	Wording stated that we'd look at efficacy of pharmacological treatments to prevent or respond to BPSD.	Changed prevent to reduced the frequency and severity, changed respond to acute treatment, and broke these concepts into separate questions.	Sought to clarify distinct clinical problems and treatment goals.

Section	Original Protocol	Revised Protocol	Rationale
Key Questions (also carried this change through Table 1, analytic framework figures 1-3, Table 4 Inclusion Criteria, section IV.D. Data abstraction, and section IV.F. Data synthesis)	Did not include comorbidities as possible effect modifiers.	Added depression as possible effect modifier.	Depression may be important in interpreting results of treatment and diagnostic studies.
Key Questions (also carried through rest of protocol)	Pre-treatment cognitive or functional level	Pre-treatment cognitive or functional level/CATD stage	There was some TEP confusion about whether the protocol would look at CATD stage as a potential effect modifier. We wanted to clarify that our prior wording addressed the same concept.
II. Key Questions	Treatment setting and testing setting	Living setting	We sought to clarify that we were evaluating whether results were different as a function of where the participant lived.
Table 1	Treatment comparisons were presented in a single list.	Broke out the comparisons into effectiveness and comparative effectiveness categories.	To clarify which comparisons were being compared to which treatment interventions.

Section	Original Protocol	Revised Protocol	Rationale
Table 1	Cognitive domains to be evaluated as outcomes were not fully specified.	Specified that treatment outcome cognitive domains will include global cognitive screening, global multidomain, memory, executive function, language, attention	Clarification/transparency.
Table 1	All harms outcomes were in one list.	Categorized harms outcomes as general, psychiatric, and nonpsychiatric.	This change to categorize harms outcomes was made for clarification.
Table 1	Confusion listed as a treatment harm.	Changed to confusion/delirium	A TEP member asked how confusion would distinguish between delirium and cognitive decline that was part of the natural history of a participant's underlying dementia. We changed to a broader confusion/delirium term with the intent to extract how studies defined the outcome, since a brief review suggested that studies didn't strictly use the term delirium to refer to this concept.
Table 1	Mortality listed as a treatment harm.	Changed to mortality (all-cause, CVD, non-CVD).	Because of the concern that antipsychotics or other treatments may increase CVD mortality, but not other causes of mortality, breaking out results by cause-specific mortality may prevent missing an important risk.

Section	Original Protocol	Revised Protocol	Rationale
Table 1	Caregiver harms for BPSD drug treatments listed as depression, anxiety, QOL, injury	Caregiver harms for BPSD drug treatments changed to depression, QOL, global stress/distress/burnout, burden	The revised list better captures the most important caregiver outcomes in clinical settings.
Table 1	Brief cognitive batteries	Brief multidomain cognitive batteries.	We changed the wording to help readers understand the concept being addressed.
Table 1	Details not provided for which cognitive domains will be evaluated as treatment outcomes.	Stated that the following cognitive domains will be evaluated as treatment outcomes: global cognitive screening, multidomain cognitive tests, memory, executive function, language, attention	More specificity to CATD. These also were considered the highest priority cognitive domains by TEP members.
Table 1 (also carried through this change in Table 4, IV.F. Data synthesis)	Hypersexuality	Sexually disinhibited behavior	The revised wording more accurately describes what is happening with patients. In most patients, their libido isn't increased.
Table 1	SAEs	FDA defined composite outcome of SAEs	We changed the wording to help the reader know this was a very specific definition, and not the review team deciding which individual harms outcomes are serious adverse events.

Section	Original Protocol	Revised Protocol	Rationale
Table 1 (also carried through this change in Table 2)	Antipsychotics listed as a category of prescription pharmacologic treatments.	Changed to include only haloperidol from first generation antipsychotics and any second generation antipsychotic.	We believe that use of first generation antipsychotics is increasingly uncommon, especially so by primary care providers. Among first generation antipsychotics, haloperidol may still be used some.
Table 1	Alpha blockers listed as a category of prescription pharmacologic treatments.	Eliminated	This class of medications is not generally used for this indication.
Table 1 (also carried through this change in Table 2, Table 4)	CNS stimulants listed as a category of prescription pharmacologic treatments.	Eliminated	We won't be extracting information on apathy, and CNS stimulant use in patients with CATD is almost entirely for apathy.
Table 1	Brain imaging tests included whole brain/grey matter volume.	Eliminated these tests of from review.	These whole brain/total brain/global volume measures are associated with head size and are nonspecific for AD. To our knowledge, clinicians do not look at global atrophy to help determine if the underlying etiology is AD. We kept focal volume measures in the review (e.g. hippocampal volume).
Table 1		Added CSF tests t-tau/Aβ42 ratio, p-tau/Aβ42 ratio, and neurofilament light protein.	A relatively recent systematic review on CSF tests for diagnosis of Alzheimer's disease reported that these tests may be useful.

Section	Original Protocol	Revised Protocol	Rationale
Table 1		Added mental distress due to correct diagnosis and radiation as potential diagnostic test harms.	The absence of this outcome was an oversight in the prior protocol.
Table 1 (also carried through this change in Table 2)	List of prescription drug categories included sleep medications.	Eliminated sleep medications for list of prescription drug categories.	Sleep disturbances as a symptom target and treatment outcome were previously eliminated from scope due to lower priority than other included BPSD. Since this outcome eliminated, drugs targeted to these symptoms should be eliminated.
Table 2 (also carried through this change in Table 4)	No minimum size limit on RCTs and CCTs	Added language that we may limit to RCTs and CCTs ≥ 10 or ≥ 25 participants per study arm if data from larger studies provide sufficient evidence for a treatment comparison.	This change was instituted in an attempt to efficiently eliminate small, often single-center and low quality trials that are highly unlikely to contribute to strength of evidence. Small studies also are likely to have low statistical power, more likely to be low quality, will not be able to adjust for potentially confounding variables, and are prone to small sample bias. We included language stating that we would consider smaller trials if evidence without them was insufficient.
Table 2	Antidepressants, other was listed as an included drug class for treatment of BPSD.	Specified that the category of “antidepressant, other” excluded MAO-inhibitors, tricyclic and tetracyclic antidepressants.	These types of antidepressants are not used by primary care providers. They are very infrequently used by specialists in patients of any age, let alone in older patients with dementia, in whom risk of their side effects is even greater.

Section	Original Protocol	Revised Protocol	Rationale
Table 3	List of cognitive test categories indicated as being evaluated for diagnostic accuracy included measures from the attention domain.	Eliminated cognitive diagnostic tests that evaluate attention but that do not evaluate other cognitive domains.	Attention is a nonspecific cognitive domain for diagnosis of CATD and underlying Alzheimer’s disease. Decreased attention may be more likely to pick up medical causes of cognitive impairment, such as causes of delirium (medications, intoxication, infection, etc.)
Table 4		Added that we will consider studies with fewer than 25 participants total for diagnostic studies if data from studies with ≥ 25 participants are insufficient.	We found in a preliminary review of biomarker studies with pathologic reference group that many studies had arms with slightly fewer than 25 participants. We also provided justification for why we would initially exclude diagnostic studies with fewer than 25 participants, including the higher likelihood that these would be low quality studies, their low statistical power, their inability to adjust for potentially confounding variables, and their small sample bias.
Table 4	No delivery route specified for nonprescription OTC supplements, vitamins, herbal medicines.	Limited included pharmacologic interventions to those that are orally ingested.	This change was implemented to eliminate aroma therapy.

Section	Original Protocol	Revised Protocol	Rationale
IV.B.Search (also carried through this change in IV.C, IV.D, IV.E, and IV.F)	No information about searching for systematic reviews, selecting them, rating their quality, extracting or synthesizing their data.	Added information about all these activities.	In an attempt to improve the efficiency of the review, when appropriate, we plan to use data from systematic reviews to replace de novo data extraction. This required additional methods language for how we'll deal with the systematic review.
IV.E. Data abstraction	Caregiver/social support, family history of CATD, and APOE-e4 status were listed as possible effect modifiers.	All these were eliminated as possible effect modifiers.	Family history of CATD is not expected to affect treatment response or diagnostic testing accuracy. APOE-e4 is not expected to affect response of treatments targeted to behavior and results of drug treatments to prevent CATD have suggested the effect of cholinesterase inhibitors on cognition are not modified by APOE-e4 levels. Although caregiver/social support possibly could impact function, QOL, and BPSD, it would not be expected to modify the effects of drug treatment on these outcomes.

Section	Original Protocol	Revised Protocol	Rationale
IV.E. Data abstraction	In section describing approach for possibly using prior systematic reviews to replace de novo data extraction, we included language stating plan to assess systematic reviews for author conflict of interest, and, when judged to be present, extracting a sample of the primary studies.	Eliminated the plan to try to formally assess systematic reviews for author conflict of interest.	We found no validated method to judge when author conflict of interest was likely. Further, the proposed approach seemed likely to add substantial work, while possibly introducing its own bias.
IV.F. Data synthesis	Described plan to possibly calculate effect size without stating how we would calculate it.	Eliminated this language, leaving prior language on plan to evaluate standardized mean differences.	For clarification.

Section	Original Protocol	Revised Protocol	Rationale
IV.F. Data synthesis	List of behavioral outcomes included apathy, wandering and sleep disturbance.	All these were eliminated as possible behavioral outcomes.	These factors all were supposed to have been eliminated in an earlier revision but were overlooked. The rationale at that time was that recent systematic reviews had been published on drug treatment for apathy and sleep disturbance in patients with dementia and that wandering was primarily treated with nonpharmacologic treatment.
IV.G. Strength of evidence	Stated that we'd assess strength of evidence at nonspecific domain level.	Specified that we would determine which validated stage, cognitive, functional, quality of life, behavioral measures, SAE, brain imaging, and CSF measures were most commonly reported across studies, and pick the 1-2 most common in each domain for strength of evidence grading. We stated we'd consider grading SOE for other measures depending on available data.	To provide more specificity of strength of evidence grading methods a priori.

Amendments After Final Protocol

If we need to amend this protocol, we will give the date of each amendment, describe the change and give the rationale in this section. Changes will not be incorporated into the protocol. Example table below:

Table 6.

Date	Section	Original Protocol	Revised Protocol	Rationale
<i>Effective date of the change in protocol</i>	<i>Location of change in the protocol</i>	<i>Language of the original protocol.</i>	<i>Description of changes in protocol.</i>	<i>Justification of why the change will improve the report. If necessary, describe why the change does not introduce bias. Do not use justification as “because the AE/TOO/TEP/Peer reviewer told us to” but explain what the change hopes to accomplish.</i>

VIII. Review of Key Questions

Key questions were refined by the Evidence-based Practice Center (EPC), then reviewed by AHRQ staff and AAFP—the topic nominator and nonsponsoring partner—to assure that they addressed the clinical questions that drove the nomination of this topic. These reviews also aimed to make the key questions more explicit about the populations, interventions, comparisons, outcomes, treatment duration, settings and study designs being considered.

IX. Key Informants

Key Informants constituted a group of patients, practicing primary and specialty care clinicians, representatives from relevant professional organizations (American Psychiatric Association, American Psychological Association), representatives from the nonsponsoring partner (AAFP), and content experts from relevant federal government agencies (NIA, NINDS, VA). They provided verbal and written feedback to help the EPC refine the key questions, PICOTS, analytic framework and project scope during the topic refinement stage.

X. Technical Experts

Technical Experts constitute a multi-disciplinary group of clinical, content, and methodological experts identified to provide input in defining populations, interventions, comparisons, or outcomes and possibly to identify studies or databases to search. For the present project, Technical Experts will be targeted to provide broad expertise and diverse perspectives pertinent to the review, including: primary care; geriatrics; geropsychiatry; psychology; neurology; pharmacological treatment of CATD; neuropsychology; use of brain imaging, CSF and blood biomarkers in diagnosis of MCI and AD; epidemiology; systematic reviews; clinical guidelines; and complex medical patients/multimorbidity. Divergent and conflicting opinions are common and perceived as healthy scientific discourse that results in a thoughtful, relevant systematic review. Therefore, study questions, design, and methodological approaches do not necessarily represent the views of individual technical and content experts.

Technical Experts will provide information to AHRQ and the EPC on the important clinical and research issues pertinent to CATD drug treatment, MCI, CATD and AD diagnosis, and on proposed key questions and PICOTS. Technical Experts will be given the opportunity to review the draft report during the peer/public review comment period. Technical Experts will not perform analysis of any kind or contribute to the writing of the report.

Technical Experts must disclose any financial conflicts of interest greater than \$5,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals are invited to serve as Technical Experts and those who present

with potential conflicts may be retained. The AHRQ TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

XI. Peer Reviewers

Peer reviewers will be invited to provide written comments on the draft report based on their clinical, content, or methodological expertise. The EPC will consider all peer review comments on the draft report in preparation of the final report. Peer reviewers will not participate in writing or editing the final report or other products. The final report will not necessarily represent the views of individual reviewers. The EPC will complete a disposition of all peer review comments. The disposition of comments will be published three months after the publication of the evidence report.

Potential Peer Reviewers must disclose any financial conflicts of interest greater than \$5,000 and any other relevant business or professional conflicts of interest. Invited Peer Reviewers may not have any financial conflict of interest greater than \$5,000. Peer reviewers who disclose potential business or professional conflicts of interest may submit comments on draft reports through the public comment mechanism.

XII. EPC Team Disclosures

EPC core team members are required to disclose any financial conflicts of interest greater than \$1,000 and any other relevant business or professional conflicts of interest. Related financial conflicts of interest that cumulatively total greater than \$1,000 will usually disqualify EPC core team investigators.

XIII. Role of the Funder

This project was funded under Contract No. HHS29032005T / HHS290201500008I TO #5 from the Agency for Healthcare Research and Quality (AHRQ), U.S. Department of Health and Human Services. AHRQ staff assisted in developing the scope and key questions, but had no role in study selection, quality assessment, or synthesis. The AHRQ Task Order Officer reviewed contract deliverables for adherence to contract requirements and quality. The authors of this report are responsible for its content. Statements in the report should not be construed as endorsement by the Agency for Healthcare Research and Quality or the U.S. Department of Health and Human Services.

XIV. Registration

This protocol will be registered in the international prospective register of systematic reviews (PROSPERO).

Appendix 1: Bibliography

1. Plassman BL, Langa KM, Fisher GG, et al. Prevalence of dementia in the United States: the aging, demographics, and memory study. *Neuroepidemiology*. 2007;29(1-2):125-32. doi: 10.1159/000109998. PMID: 17975326.
2. Langa KM, Larson EB, Crimmins EM, et al. A Comparison of the Prevalence of Dementia in the United States in 2000 and 2012. *JAMA Intern Med*. 2017 Jan 01;177(1):51-8. doi: 10.1001/jamainternmed.2016.6807. PMID: 27893041.
3. McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011 May;7(3):263-9. doi: 10.1016/j.jalz.2011.03.005. PMID: 21514250.
4. Hurd MD, Martorell P, Delavande A, et al. Monetary costs of dementia in the United States. *N Engl J Med*. 2013 Apr 4;368(14):1326-34. doi: 10.1056/NEJMsa1204629. PMID: 23550670.
5. Zhao QF, Tan L, Wang HF, et al. The prevalence of neuropsychiatric symptoms in Alzheimer's disease: Systematic review and meta-analysis. *J Affect Disord*. 2016 Jan 15;190:264-71. doi: 10.1016/j.jad.2015.09.069. PMID: 26540080.
6. What is Alzheimer's? Alzheimer's Association. www.alz.org/alzheimers-dementia/what-is-alzheimers. Accessed on November 2, 2018.
7. Alzheimer's Disease. Centers for Disease Control and Prevention. www.cdc.gov/aging/aginginfo/alzheimers.htm. Accessed on November 2, 2018.
8. Association AP. Diagnostic and Statistical Manual of Mental Disorders, Third Edition, Revised (DSM-III-R); 1987.
9. McKhann G, Drachman D, Folstein M, et al. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*. 1984 Jul;34(7):939-44. PMID: 6610841.
10. Association AP. Diagnostic and Statistical Manual of Mental Disorders: DSM-5; 2013.
11. Cure S, Abrams K, Belger M, et al. Systematic literature review and meta-analysis of diagnostic test accuracy in Alzheimer's disease and other dementia using autopsy as standard of truth. *J Alzheimers Dis*. 2014;42(1):169-82. doi: 10.3233/JAD-131559. PMID: 24840572.
12. Frisoni GB, Boccardi M, Barkhof F, et al. Strategic roadmap for an early diagnosis of Alzheimer's disease based on biomarkers. *Lancet Neurol*. 2017 Aug;16(8):661-76. doi: 10.1016/S1474-4422(17)30159-X. PMID: 28721928.
13. Kane RLB, M.; Fink, H.A.; Brasure, M.; Davila, H.; Desai, P.; Jutkowitz, E.; McCreedy, E.; Nelson, V.A.; McCarten, J.R.; Calvert, C.; Ratner, E.; Hemmy, L.S.; Barclay, T. Interventions To Prevent Age-Related Cognitive Decline, Mild Cognitive Impairment, and Clinical Alzheimer's-Type Dementia. AHRQ Publication No. 17-EHC008-EF. Rockville, MD: Quality AfHRa; February 2017 2017. www.effectivehealthcare.ahrq.gov/reports/final.cfm

14. Reus VI, Fochtmann LJ, Eyer AE, et al. The American Psychiatric Association Practice Guideline on the Use of Antipsychotics to Treat Agitation or Psychosis in Patients With Dementia. *Am J Psychiatry*. 2016 May 1;173(5):543-6. doi: 10.1176/appi.ajp.2015.173501. PMID: 27133416.
15. Brasure M, Jutkowitz E, Fuchs E, et al. *Nonpharmacologic Interventions for Agitation and Aggression in Dementia*. Rockville (MD); 2016.
16. Qaseem A, Snow V, Cross JT, Jr., et al. Current pharmacologic treatment of dementia: a clinical practice guideline from the American College of Physicians and the American Academy of Family Physicians. *Ann Intern Med*. 2008 Mar 4;148(5):370-8. PMID: 18316755.
17. Kales HC, Gitlin LN, Lyketsos CG. Assessment and management of behavioral and psychological symptoms of dementia. *BMJ*. 2015 Mar 2;350:h369. doi: 10.1136/bmj.h369. PMID: 25731881.
18. Ruthirakuhan MT, Herrmann N, Abraham EH, et al. Pharmacological interventions for apathy in Alzheimer's disease. *Cochrane Database Syst Rev*. 2018 May 4;5:CD012197. doi: 10.1002/14651858.CD012197.pub2. PMID: 29727467.
19. McCleery J, Cohen DA, Sharpley AL. Pharmacotherapies for sleep disturbances in dementia. *Cochrane Database Syst Rev*. 2016 Nov 16;11:CD009178. doi: 10.1002/14651858.CD009178.pub3. PMID: 27851868.
20. Hackshaw A. Small studies: strengths and limitations. *Eur Respir J*. 2008 Nov;32(5):1141-3. doi: 10.1183/09031936.00136408. PMID: 18978131.
21. Viswanathan M, Ansari M, Berkman N, et al. *Assessing the Risk of Bias of Individual Studies in Systematic Reviews of Health Care Interventions AHRQ*. 2012.
22. *Methods Guide for Medical Test Reviews*. AHRQ Publication No. 12-EC017. Rockville, MD: Quality AfHRA; June 2012.
www.effectivehealthcare.ahrq.gov/reports/final.cfm.
23. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011 Oct 18;155(8):529-36. doi: 10.7326/0003-4819-155-8-201110180-00009. PMID: 22007046.
24. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ*. 2017 Sep 21;358:j4008. doi: 10.1136/bmj.j4008. PMID: 28935701.
25. Fu R, Gartlehner G, Grant M, et al. Conducting quantitative synthesis when comparing medical interventions: AHRQ and the Effective Health Care Program. *J Clin Epidemiol*. 2011 Nov;64(11):1187-97. doi: 10.1016/j.jclinepi.2010.08.010. PMID: 21477993.
26. Takwoingi Y. *Meta-analysis of test accuracy studies in Stata: a bivariate model approach*. Version 1.1.
27. Owens DK, Lohr KN, Atkins D, et al. AHRQ series paper 5: grading the strength of a body of evidence when comparing medical interventions--agency for healthcare research and quality and the effective health-care program. *J Clin Epidemiol*. 2010 May;63(5):513-23. doi: S0895-4356(09)00093-6 [pii]
10.1016/j.jclinepi.2009.03.009. PMID: 19595577.

28. Singh S, Chang SM, Matchar DB, et al. Chapter 7: grading a body of evidence on diagnostic tests. *J Gen Intern Med.* 2012 Jun;27 Suppl 1:S47-55. doi: 10.1007/s11606-012-2021-9. PMID: 22648675.
29. Berkman ND, Lohr KN, Ansari M, et al. Grading the Strength of a Body of Evidence When Assessing Health Care Interventions for the Effective Health Care Program of the Agency for Healthcare Research and Quality: An Update. *Methods Guide for Effectiveness and Comparative Effectiveness Reviews.* Rockville (MD); 2008.
30. Atkins D, Chang S, Gartlehner G, et al. Assessing the Applicability of Studies When Comparing Medical Interventions. *Methods Guide for Effectiveness and Comparative Effectiveness Reviews.* Rockville (MD); 2008.

Appendix 2:

Electronic Literature Search Strategies for Key Questions 1 & 2--CATD Treatment

Database: Ovid MEDLINE(R) Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily, Ovid MEDLINE and Versions(R) <1946 to June 06, 2018>

Search Strategy:

-
- 1 exp Alzheimer Disease/ (83289)
 - 2 Dementia/ (45164)
 - 3 (dementia or alzheimer*).ti. (97725)
 - 4 1 or 2 or 3 (139036)
 - 5 limit 4 to "therapy (best balance of sensitivity and specificity)" (6136)
 - 6 limit 5 to english language (5766)
 - 7 limit 6 to (addresses or autobiography or bibliography or biography or case reports or clinical conference or comment or comparative study or congresses or consensus development conference or consensus development conference, nih or dataset or dictionary or directory or editorial or evaluation studies or "expression of concern" or festschrift or government publications or guideline or historical article or interactive tutorial or interview or lectures or legal cases or legislation or letter or news or newspaper article or observational study or patient education handout or periodical index or personal narratives or portraits or "review" or "scientific integrity review" or validation studies or video-audio media or webcasts) (2140)
 - 8 limit 7 to (adaptive clinical trial or clinical study or clinical trial, all or clinical trial, phase i or clinical trial, phase ii or clinical trial, phase iii or clinical trial, phase iv or clinical trial or controlled clinical trial or randomized controlled trial) (587)
 - 9 6 not 7 (3626)
 - 10 8 or 9 (4213)
 - 11 limit 10 to ("all child (0 to 18 years)" or "adult (19 to 44 years)") (372)
 - 12 limit 11 to ("middle age (45 to 64 years)" or "middle aged (45 plus years)" or "all aged (65 and over)" or "aged (80 and over)") (337)
 - 13 10 not 11 (3841)
 - 14 12 or 13 (4178)

Database: Embase Classic+Embase <1947 to 2018 June 08>

Search Strategy:

-
- 1 exp *Alzheimer disease/ (103967)
 - 2 *dementia/ (50978)
 - 3 (alzheimer* or dementia*).ti. (136009)
 - 4 1 or 2 or 3 (165774)
 - 5 limit 4 to english language (149043)
 - 6 limit 5 to "therapy (best balance of sensitivity and specificity)" (11249)
 - 7 limit 6 to "reviews (best balance of sensitivity and specificity)" (2742)
 - 8 6 not 7 (8507)

- 9 limit 8 to (embryo <first trimester> or infant <to one year> or child <unspecified age> or preschool child <1 to 6 years> or school child <7 to 12 years> or adolescent <13 to 17 years>) (58)
- 10 limit 9 to (adult <18 to 64 years> or aged <65+ years>) (38)
- 11 8 not 9 (8449)
- 12 10 or 11 (8487)
- 13 limit 12 to (book or book series or conference proceeding or trade journal) (94)
- 14 12 not 13 (8393)
- 15 limit 14 to conference abstracts (2173)
- 16 14 not 15 (6220)
- 17 limit 16 to (abstract report or books or "book review" or chapter or conference abstract or "conference review" or editorial or letter or note or patent or reports or "review" or short survey or tombstone) (472)
- 18 16 not 17 (5748)
- 19 limit 18 to (amphibia or ape or bird or cat or cattle or chicken or dog or "ducks and geese" or fish or "frogs and toads" or goat or guinea pig or "hamsters and gerbils" or horse or monkey or mouse or "pigeons and doves" or "rabbits and hares" or rat or reptile or sheep or swine) (355)
- 20 18 not 19 (5393)

Database: PsycINFO <1806 to June Week 1 2018>

Search Strategy:

-
- 1 exp *ALZHEIMER'S DISEASE/ (37834)
 - 2 *dementia/ (26693)
 - 3 (dementia* or alzheimer*).ti. (51716)
 - 4 1 or 2 or 3 (64340)
 - 5 limit 4 to "therapy (best balance of sensitivity and specificity)" (8415)
 - 6 limit 5 to (childhood <birth to 12 years> or adolescence <13 to 17 years>) (56)
 - 7 limit 6 to adulthood <18+ years> (46)
 - 8 5 not 6 (8359)
 - 9 7 or 8 (8405)
 - 10 limit 9 to animal (765)
 - 11 9 not 10 (7640)
 - 12 limit 11 to (100 childhood <birth to age 12 yrs> or 120 neonatal <birth to age 1 mo> or 140 infancy <2 to 23 mo> or 160 preschool age <age 2 to 5 yrs> or 180 school age <age 6 to 12 yrs> or 200 adolescence <age 13 to 17 yrs> or 320 young adulthood <age 18 to 29 yrs> or 340 thirties <age 30 to 39 yrs>) (377)
 - 13 limit 12 to (360 middle age <age 40 to 64 yrs> or "380 aged <age 65 yrs and older>" or "390 very old <age 85 yrs and older>") (345)
 - 14 11 not 12 (7263)
 - 15 13 or 14 (7608)
 - 16 limit 15 to (abstract collection or bibliography or chapter or clarification or "column/opinion" or "comment/reply" or dissertation or editorial or encyclopedia entry or interview or letter or obituary or poetry or publication information or review-book or review-media or review-software & other or reviews) (661)
 - 17 15 not 16 (6947)

- 18 limit 17 to ("0200 book" or "0240 authored book" or "0280 edited book" or "0300 encyclopedia" or "0400 dissertation abstract") (13)
- 19 17 not 18 (6934)
- 20 limit 19 to english language (6617)
- 21 limit 20 to "therapy (maximizes specificity)" (1068)

Appendix 3:

Electronic Literature Search Strategies for Key Question 3-- CATD Diagnosis

Cognitive Testing for Diagnosing Clinical CATD

Database: Ovid MEDLINE(R) Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily, Ovid MEDLINE and Versions(R) <1946 to June 06, 2018>

Search Strategy:

-
- 1 *Alzheimer Disease/ (66402)
 - 2 alzheimer*.ti. (61779)
 - 3 mild cognitive impairment.ti. (5911)
 - 4 MCI.ti. (983)
 - 5 or/1-4 (80974)
 - 6 exp Neuropsychological Tests/ (163514)
 - 7 screen*.ti. (154444)
 - 8 test*.ti. (384720)
 - 9 detect*.ti. (326398)
 - 10 battery.ti. (5744)
 - 11 assess*.ti. (324127)
 - 12 validat*.ti. (63099)
 - 13 tool*.ti. (73368)
 - 14 instrument*.ti. (36613)
 - 15 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 (1429090)
 - 16 5 and 15 (14872)
 - 17 15 and 16 (14872)
 - 18 limit 17 to "diagnosis (best balance of sensitivity and specificity)" (3339)
 - 19 limit 18 to english language (3222)
 - 20 limit 19 to (addresses or autobiography or bibliography or biography or case reports or clinical conference or comment or congresses or dataset or dictionary or directory or editorial or "expression of concern" or festschrift or interactive tutorial or interview or lectures or legal cases or legislation or letter or news or newspaper article or patient education handout or periodical index or personal narratives or portraits or video-audio media or webcasts) (80)
 - 21 19 not 20 (3142)

Database: Embase Classic+Embase <1947 to 2018 June 08>

Search Strategy:

-
- 1 alzheimer*.ti. (86006)
 - 2 mild cognitive impairment.ti. (9199)
 - 3 MCI.ti. (2187)
 - 4 or/1-3 (93844)
 - 5 exp Neuropsychological Tests/ (91842)
 - 6 screen*.ti. (203834)
 - 7 test*.ti. (483228)
 - 8 detect*.ti. (388678)

- 9 battery.ti. (5642)
- 10 assess*.ti. (435259)
- 11 validat*.ti. (88547)
- 12 tool*.ti. (94680)
- 13 instrument*.ti. (43294)
- 14 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 (1710921)
- 15 4 and 14 (10676)
- 16 limit 15 to "diagnosis (best balance of sensitivity and specificity)" (3110)
- 17 limit 16 to english language (2954)
- 18 limit 17 to conference abstracts (715)
- 19 17 not 18 (2239)
- 20 limit 19 to (book or book series or trade journal) (23)
- 21 19 not 20 (2216)
- 22 limit 21 to (books or "book review" or chapter or conference abstract or "conference review" or editorial or letter or note or patent or short survey or tombstone) (39)
- 23 21 not 22 (2177)
- 24 limit 23 to (embryo <first trimester> or infant <to one year> or child <unspecified age> or preschool child <1 to 6 years> or school child <7 to 12 years> or adolescent <13 to 17 years>) (10)
- 25 23 not 24 (2167)

Database: PsycINFO <1806 to June Week 1 2018>

Search Strategy:

-
- 1 alzheimer*.ti. (28851)
 - 2 mild cognitive impairment.ti. (3998)
 - 3 MCI.ti. (452)
 - 4 or/1-3 (31933)
 - 5 screen*.ti. (18161)
 - 6 test*.ti. (99582)
 - 7 detect*.ti. (17678)
 - 8 battery.ti. (3009)
 - 9 assess*.ti. (92838)
 - 10 validat*.ti. (22170)
 - 11 tool*.ti. (14533)
 - 12 instrument*.ti. (13145)
 - 13 exp neuropsychological assessment/ (16377)
 - 14 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 (267118)
 - 15 4 and 14 (2610)
 - 16 diagnos*.ti. (39731)
 - 17 sensitivity.ti,ab. (83414)
 - 18 specificity.ti,ab. (33176)
 - 19 exp diagnosis/ (164560)
 - 20 16 or 17 or 18 or 19 (273286)
 - 21 15 and 20 (987)
 - 22 limit 21 to english language (923)

- 23 limit 22 to (abstract collection or bibliography or chapter or "column/opinion" or "comment/reply" or dissertation or editorial or encyclopedia entry or interview or letter or obituary or poetry or publication information or review-book or review-media or review-software & other or reviews) (97)
- 24 22 not 23 (826)
- 25 limit 24 to (100 childhood <birth to age 12 yrs> or 120 neonatal <birth to age 1 mo> or 140 infancy <2 to 23 mo> or 160 preschool age <age 2 to 5 yrs> or 180 school age <age 6 to 12 yrs> or 200 adolescence <age 13 to 17 yrs>) (3)
- 26 24 not 25 (823)

Biomarker Testing for Determining Whether Clinical CATD is Attributable to AD

Database: Ovid MEDLINE(R) and Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Daily and Versions(R) <1946 to August 29, 2018>

Search Strategy:

-
- 1 Alzheimer Disease/ (84019)
- 2 alzheimer*.ti,ab. (125917)
- 3 Dementia/ (45557)
- 4 Cognition Disorders/ (61418)
- 5 Cognitive Dysfunction/ (9488)
- 6 MCI.ti,ab. (14712)
- 7 mild cognitive impairment.ti,ab. (12950)
- 8 or/1-7 (232479)
- 9 BIOMARKERS/ (234077)
- 10 Neuroimaging/ (8717)
- 11 exp Hematologic Tests/ (238907)
- 12 Cerebrospinal Fluid/ (17957)
- 13 exp Magnetic Resonance Imaging/ (398799)
- 14 exp Positron-Emission Tomography/ (50761)
- 15 exp Tomography, Emission-Computed, Single-Photon/ (29951)
- 16 CT.ti,ab. (294643)
- 17 computed tomography.ti,ab. (210479)
- 18 PET.ti,ab. (83676)
- 19 positron emission.ti,ab. (51751)
- 20 imag*.ti,ab. (1027090)
- 21 neuroima*.ti,ab. (41199)
- 22 single photon.ti,ab. (23622)
- 23 SPECT.ti,ab. (26030)
- 24 magnetic resonance.ti,ab. (297013)
- 25 MRI.ti,ab. (205862)
- 26 (blood or plasma or serum).ti,ab. (2992475)
- 27 or/9-26 (4644755)
- 28 Alzheimer Disease/dg [Diagnostic Imaging] (3589)
- 29 exp Alzheimer Disease/di [Diagnosis] (13604)
- 30 exp Cognitive Dysfunction/di [Diagnostic Imaging] (2600)

- 31 exp Cognitive Dysfunction/di [Diagnosis] (2600)
- 32 Cognition Disorders/dg [Diagnostic Imaging] (932)
- 33 exp Cognition Disorders/di [Diagnosis] (25169)
- 34 or/28-33 (39723)
- 35 8 and 27 (53962)
- 36 34 or 35 (80966)
- 37 exp AUTOPSY/ (40362)
- 38 autops*.ti,ab. (70532)
- 39 neuropath*.ti,ab. (118139)
- 40 histopath*.ti,ab. (200564)
- 41 postmortem.ti,ab. (32649)
- 42 Braak.ti,ab. (982)
- 43 37 or 38 or 39 or 40 or 41 or 42 (427946)
- 44 36 and 43 (5553)
- 45 limit 44 to "diagnosis (best balance of sensitivity and specificity)" (916)
- 46 limit 45 to english language (879)
- 47 limit 46 to (addresses or autobiography or bibliography or biography or case reports or dataset or dictionary or directory or interactive tutorial or interview or lectures or legal cases or legislation or letter or news or newspaper article or patient education handout or periodical index or personal narratives or portraits or video-audio media or webcasts) (30)
- 48 46 not 47 (849)
- 49 limit 48 to yr="2012 -Current" (318)
- 50 11 or 26 (3121413)
- 51 48 and 50 (150)
- 52 49 or 51 (408)

Database: Embase Classic+Embase <1947 to 2018 August 15>

Search Strategy:

-
- 1 Alzheimer Disease/ (164879)
 - 2 alzheimer*.ti,ab. (162410)
 - 3 Dementia/ (102680)
 - 4 Cognition Disorders/ (26235)
 - 5 Cognitive Dysfunction/ (76211)
 - 6 MCI.ti,ab. (26533)
 - 7 mild cognitive impairment.ti,ab. (19772)
 - 8 or/1-7 (347165)
 - 9 BIOMARKERS/ (164164)
 - 10 Neuroimaging/ (95977)
 - 11 exp Hematologic Tests/ (236604)
 - 12 Cerebrospinal Fluid/ (116263)
 - 13 exp Magnetic Resonance Imaging/ (767997)
 - 14 exp Positron-Emission Tomography/ (123609)
 - 15 exp Tomography, Emission-Computed, Single-Photon/ (6109)
 - 16 CT.ti,ab. (468224)
 - 17 computed tomography.ti,ab. (244121)

- 18 PET.ti,ab. (136571)
- 19 positron emission.ti,ab. (63108)
- 20 imag*.ti,ab. (1302984)
- 21 neuroima*.ti,ab. (55752)
- 22 single photon.ti,ab. (24397)
- 23 SPECT.ti,ab. (41018)
- 24 magnetic resonance.ti,ab. (342354)
- 25 MRI.ti,ab. (332862)
- 26 (blood or plasma or serum).ti,ab. (4061097)
- 27 or/9-26 (6165659)
- 28 exp Cognitive Dysfunction/di [Diagnosis] (45351)
- 29 8 and 27 (92640)
- 30 29 or 28 (125049)
- 31 exp AUTOPSY/ (179900)
- 32 autops*.ti,ab. (115539)
- 33 neuropath*.ti,ab. (165200)
- 34 histopath*.ti,ab. (268113)
- 35 postmortem.ti,ab. (42076)
- 36 Braak.ti,ab. (1803)
- 37 31 or 32 or 33 or 34 or 35 or 36 (652613)
- 38 30 and 37 (11978)
- 39 limit 28 to "diagnosis (maximizes specificity)" (2848)
- 40 limit 39 to embase (2437)
- 41 limit 40 to (embryo <first trimester> or infant <to one year> or child <unspecified age> or preschool child <1 to 6 years> or school child <7 to 12 years> or adolescent <13 to 17 years>) (57)
- 42 40 not 41 (2380)
- 43 limit 42 to english language (2207)
- 44 limit 43 to (books or chapter or conference abstract or "conference review" or editorial or letter or note or "review" or short survey or tombstone) (256)
- 45 43 not 44 (1951)
- 46 limit 45 to (book or book series or trade journal) (14)
- 47 45 not 46 (1937)
- 48 limit 47 to yr="2012" (101)
- 49 11 or 26 (4149515)
- 50 47 and 49 (205)
- 51 48 or 50 (295)

Database: Embase Classic+Embase <1947 to 2018 August 15>
 Search Strategy:

-
- 1 Alzheimer Disease/ (164879)
 - 2 alzheimer*.ti,ab. (162410)
 - 3 Dementia/ (102680)
 - 4 Cognition Disorders/ (26235)
 - 5 Cognitive Dysfunction/ (76211)

6 MCI.ti,ab. (26533)
7 mild cognitive impairment.ti,ab. (19772)
8 or/1-7 (347165)
9 BIOMARKERS/ (164164)
10 Neuroimaging/ (95977)
11 exp Hematologic Tests/ (236604)
12 Cerebrospinal Fluid/ (116263)
13 exp Magnetic Resonance Imaging/ (767997)
14 exp Positron-Emission Tomography/ (123609)
15 exp Tomography, Emission-Computed, Single-Photon/ (6109)
16 CT.ti,ab. (468224)
17 computed tomography.ti,ab. (244121)
18 PET.ti,ab. (136571)
19 positron emission.ti,ab. (63108)
20 imag*.ti,ab. (1302984)
21 neuroima*.ti,ab. (55752)
22 single photon.ti,ab. (24397)
23 SPECT.ti,ab. (41018)
24 magnetic resonance.ti,ab. (342354)
25 MRI.ti,ab. (332862)
26 (blood or plasma or serum).ti,ab. (4061097)
27 or/9-26 (6165659)
28 [Alzheimer Disease/dg [Diagnostic Imaging]] (0)
29 exp Alzheimer Disease/di [Diagnosis] (17185)
30 exp Cognitive Dysfunction/di [Diagnostic Imaging] (45351)
31 exp Cognitive Dysfunction/di [Diagnosis] (45351)
32 [Cognition Disorders/dg [Diagnostic Imaging]] (0)
33 exp Cognition Disorders/di [Diagnosis] (45351)
34 or/28-33 (45351)
35 8 and 27 (92640)
36 34 or 35 (125049)
37 exp AUTOPSY/ (179900)
38 autops*.ti,ab. (115539)
39 neuropath*.ti,ab. (165200)
40 histopath*.ti,ab. (268113)
41 postmortem.ti,ab. (42076)
42 Braak.ti,ab. (1803)
43 37 or 38 or 39 or 40 or 41 or 42 (652613)
44 36 and 43 (11978)
45 limit 44 to "diagnosis (best balance of sensitivity and specificity)" (2692)
46 limit 45 to english language (2512)
47 limit 46 to (addresses or autobiography or bibliography or biography or case reports or
dataset or dictionary or directory or interactive tutorial or interview or lectures or legal cases or
legislation or letter or news or newspaper article or patient education handout or periodical index
or personal narratives or portraits or video-audio media or webcasts) [Limit not valid in Embase;
records were retained] (20)

- 48 46 not 47 (2492)
- 49 limit 48 to yr="2012 -Current" (1033)
- 50 11 or 26 (4149515)
- 51 48 and 50 (423)
- 52 49 or 51 (1253)

Appendix 4

Risk of Bias Assessment Decision Aid

Selection Bias

DEFINITION

Systematic differences between baseline characteristics of the groups that arise from self-selection of treatments, physician-directed selection of treatments, or association of treatment assignments with demographic, clinical, or social characteristics. Good randomization produces study groups that are likely comparable for known and unknown risk factors, removes investigator bias in allocation, and allow the most valid statistical inference in comparing outcomes between groups. In randomized studies, whether there is bias in allocation of study participants to treatment groups is a function both of whether the methods of randomization are good AND whether randomization successfully achieved a balance between treatment groups in risk factors or prognostic covariates.

ASSESSMENT GUIDANCE

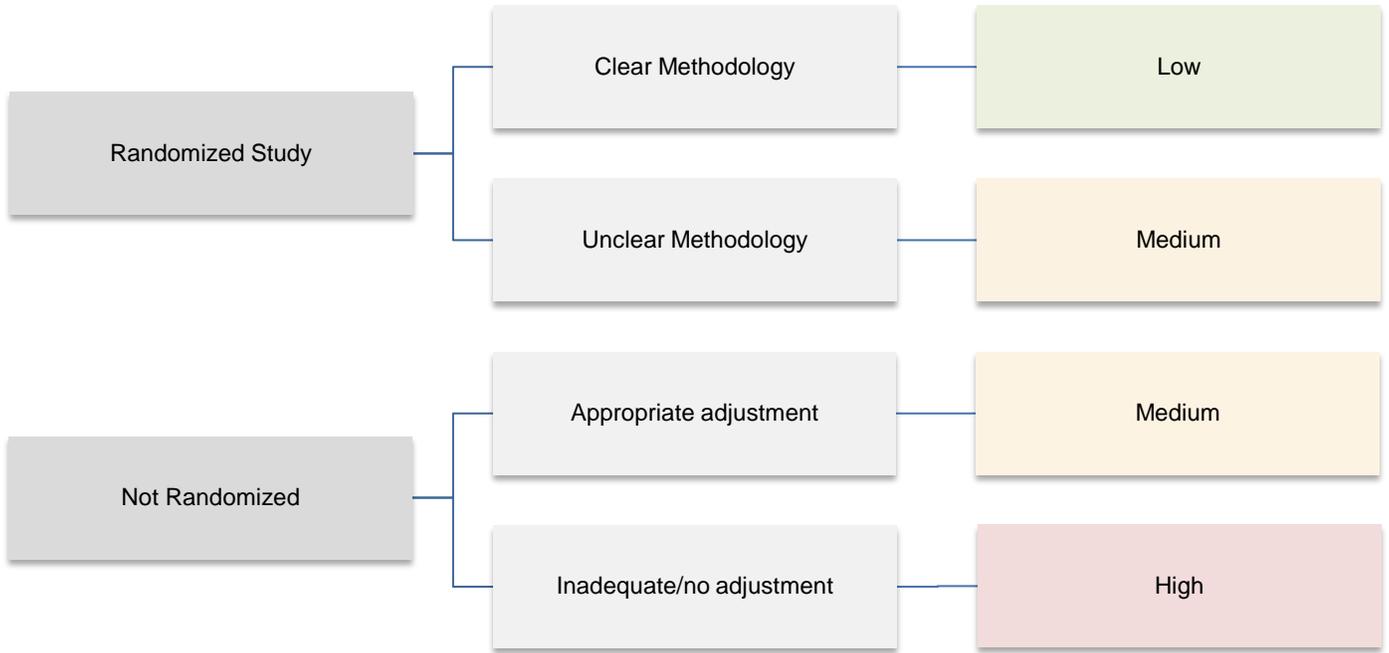
OPTION 1: The study reports that it was randomized.

- **Clear Methodology:** The study used a randomization method such as random numbers table, computer-generated random number producing algorithm, blocked randomization, stratified randomization, adaptive randomization (e.g., minimization).
- **Unclear Methodology:** Study reports that allocation/assignment was randomized but gives no further detail.

OPTION 2: Study is not randomized (for treatment efficacy outcomes, CCTs are the only eligible nonrandomized study design)

- Study it uses systematic allocation of treatment by investigator: Systematic and predictable investigator allocation of treatment assignment (e.g., alternation, based on day of week, based on the month of birthday)
- Study should use an appropriate statistical adjustment (propensity score, instrumental variable, multivariate).

Figure 4. Selection Bias Assessment Guidance



Attrition

DEFINITION

Loss of participants from the study, potential systematic differences in that loss to follow-up, and how losses were accounted for in the results (e.g., incomplete follow-up, differential attrition). Those who drop out of the study or who are lost to follow-up may be systematically different from those who remain in the study. Attrition bias can potentially change the collective (group) characteristics of the relevant groups and their observed outcomes in ways that affect study results by confounding and spurious associations. **Overall attrition** refers to attrition in all groups combined for a given outcome comparison and timepoint. **Differential attrition** refers to the absolute difference between groups in attrition for a given outcome comparison and timepoint.

ASSESSMENT GUIDANCE

***Studies that have long-term outcomes that are 5 years and longer should be assessed on a case-by-case basis.**

OPTION 1: Study has low overall attrition (<10%). Reasons for incomplete/missing data should be adequately explained.

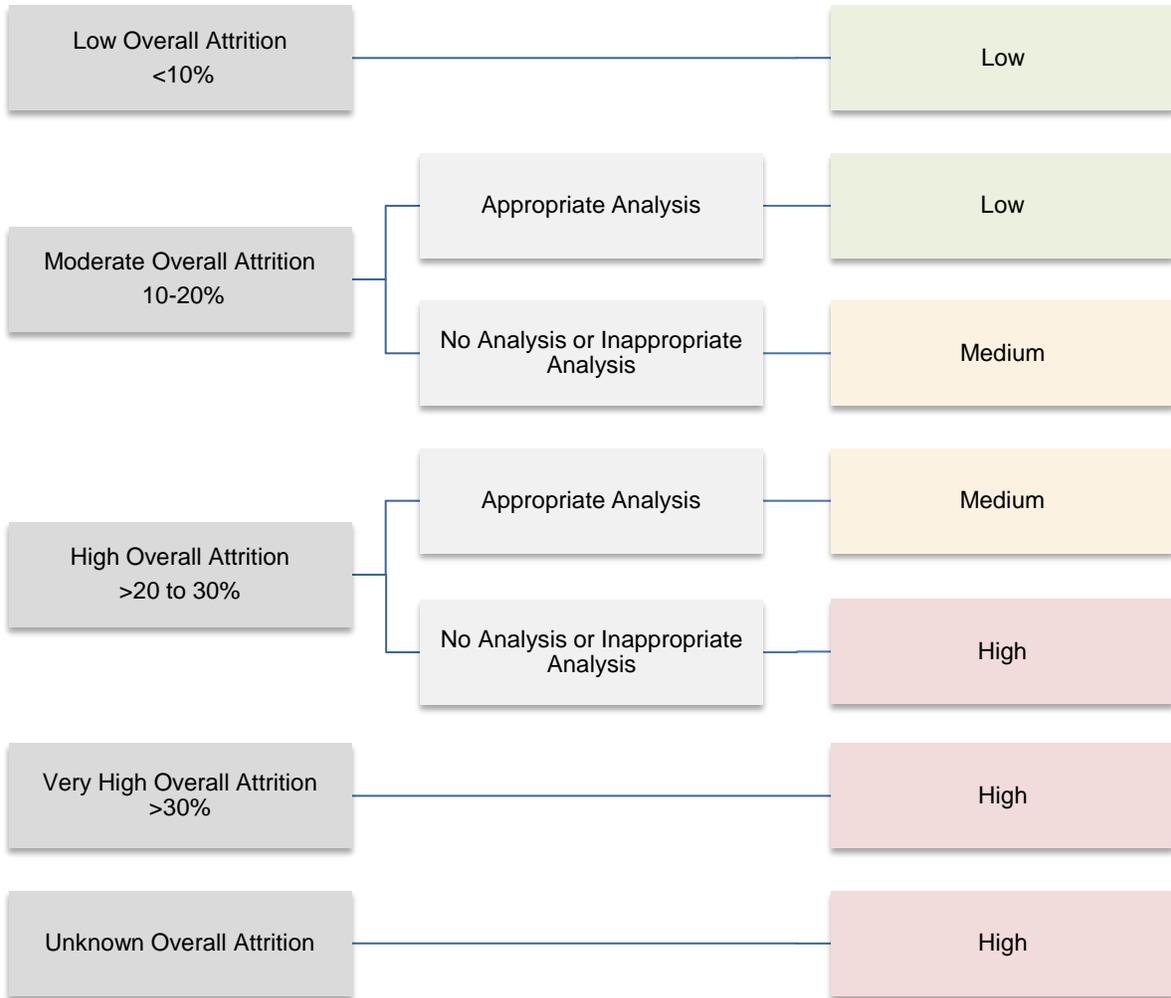
OPTION 2: Study has moderate overall attrition (10 to 20%). Reasons for incomplete/missing data should be adequately explained and authors should attempt to address attrition in their analysis. Analysis should be done with appropriate method, noting that this may help explain the size and direction of the potential bias, but they don't eliminate the bias. Last valued carried forward is not an appropriate adjustment. Some imputation methods might be appropriate (to be evaluated on a case-by-case basis).

OPTION 3: Study has high overall attrition (>20 to 30%) Reasons for incomplete/missing data should be adequately explained and authors should to address attrition in their analysis with an appropriate method. This reduces, but does not eliminate, the risk of attrition bias. Last valued carried forward is not an appropriate adjustment. Some imputation methods might be appropriate (to be evaluated on a case-by-case basis).

OPTION 4: Study has very high overall attrition (>30%) Authors may attempt to address attrition in their analysis, but the risk of attrition bias is high.

OPTION 5: Reporting of attrition by study arm is inadequate. It is unclear how many participants have been lost in each group. Risk of attrition bias is high.

Figure 5. Attrition Assessment Guidance



Performance Bias

DEFINITION

Systematic differences in the care provided to participants and protocol deviation. Examples include contamination of the control group with the exposure or intervention, unbalanced provision of additional interventions or co-interventions, difference in co-interventions, and inadequate blinding of providers and participants. **Intention-to-Treat Principle (ITT)** is when the study counts events in all randomized participants according to their treatment assignment, regardless of whether they received assigned treatment. It does not exclude participants from analysis for nonadherence, protocol deviations, withdrawal, or anything else that happens after randomization. To exclude such participants undercuts the benefit of randomization in minimizing selection bias. **Modified ITT (mITT)** is where analyses exclude randomized participants who did not receive any of their assigned treatment. This is not strictly ITT, but is accepted as such by the FDA in evaluating drug trials for approval.

ASSESSMENT GUIDANCE

Domain	Assessment	Options
1. ITT/Adjustment of Known Confounders	OPTION 1A: Study is a RCT. Check if study uses ITT or modified ITT.	-Yes -No -Unclear/Not Reported
	OPTION 1B: Study is a CCT. Check for adjustment of known confounders. Adequate adjustment includes adjustment for at least age, sex, and baseline cognition. Partially adequate adjustment adjusts for 1 or 2 of these potential confounder categories. Inadequate adjustment does not adjust for any of these potential confounder categories.	-Adequate -Partially Adequate -Inadequate
2. Participant Blinding	For all studies, check to see if participant blinding is described in text.	Yes
		No
		Unclear

Overall Performance Rating	Low =ITT or adequate adjustment of confounders. Participants are blinded.	Medium =Unclear ITT or partially adequate adjustment of confounders. Participant blinding is unclear or not described.	High = No ITT or inadequate adjustment of known confounders.
-----------------------------------	--	---	---

Detection Bias

DEFINITION

Systematic differences in outcomes assessment among groups being compared, including systematic misclassification of the exposure or intervention, covariates, or outcomes because of variable definitions and timings, diagnostic thresholds, recall from memory, inadequate assessor blinding, and faulty measurement techniques. Erroneous statistical analysis might also affect the validity of effect estimates.

ASSESSMENT GUIDANCE

Assessment	Options
1. Check if outcome assessors were blinded to treatment assignment.	-Yes -No -Unclear
2. Check if studies used validated, reliable, outcomes measure and that the groups assessed using comparable outcome measures. Please flag the test for the team if you are unsure if a test is validated or think that the measure is based on unconfirmed self-report.	-Yes -No -Unclear

Overall Performance Rating	Low = 2 Yes OR 1 Yes, 1 Unclear	Medium = All unclear	High = At least 1 No
-----------------------------------	--	-----------------------------	-----------------------------

Reporting Bias

DEFINITION

Systematic differences between reported and unreported findings (e.g., differential reporting of outcomes or harms, incomplete reporting of study findings). Reporting bias includes selective analysis (e.g., study combines intervention groups or adjusts planned analysis without explanation).

ASSESSMENT GUIDANCE

- Check if all outcomes reported in the methods section reported in the result section and vice versa.
- If study indicates that additional information is available in a separate protocol, protocol papers should be checked to ensure no relevant information is missed.

Assessment	Options	Rating
Check if all outcomes are reported without selective analysis?	Yes	Low
	No	High
	Unclear	Medium

Overall Risk of Bias Assessment

ASSESSMENT GUIDANCE: Overall risk of bias is determined by reviewer, or team, consensus. The flow chart below provides a guide for how to rate overall risk of bias, based on the assessment of each individual domain. Reviewers should use this guide when making judgements about overall risk of bias. However, there may be cases where deviation from this guide is necessary and appropriate. For clarification and transparency, reviewers should provide a brief written justification for these deviations.

Figure 6. Overall Risk of Bias Assessment Guidance

