RESEARCH ARTICLE

Alzheimer's & Dementia
Translational Research
& Clinical Interventions

Development of a novel cognitive composite outcome to assess therapeutic effects of exercise in the EXERT trial for adults with MCI: The ADAS-Cog-Exec

Diane M. Jacobs[1,2,3]　│　Ronald G. Thomas[2,4]　│　David P. Salmon[1,2,3]　│　Shelia Jin[2,4]　│　Howard H. Feldman[1,2,3]　│　Carl W. Cotman[5]　│　Laura D. Baker[6]　│　for the Alzheimer's Disease Cooperative Study EXERT Study Group[1,#]　│　for the Alzheimer's Disease Neuroimaging Initiative[1,**]

[1] Department of Neurosciences, University of California San Diego, La Jolla, California, USA

[2] Alzheimer's Disease Cooperative Study, University of California San Diego, La Jolla, California

[3] Shiley-Marcos Alzheimer's Disease Research Center, University of California San Diego, La Jolla, California, USA

[4] Division of Biostatistics, Department of Family Medicine & Public Health, University of California San Diego, La Jolla, California, USA

[5] Institute for Memory Impairments and Neurological Disorders, University of California Irvine, Irvine, California, USA

[6] Department of Internal Medicine–Geriatrics, Wake Forest School of Medicine, Winston-Salem, North Carolina, USA

Correspondence
Diane M. Jacobs, Department of Neurosciences, University of California San Diego, 9500 Gilman Drive, #0949, La Jolla, CA 92093-0949, USA.
E-mail: djacobs@ucsd.edu

[#] Members of the Alzheimer's Disease Cooperative Study (ADCS) EXERT Study Group are listed in the Acknowledgments

[**] Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

Abstract

Introduction: Use of cognitive composites as primary outcome measures is increasingly common in clinical trials of preclinical and prodromal Alzheimer's disease (AD). Composite outcomes can decrease intra-individual variability, resulting in improved sensitivity to detect longitudinal change and increased statistical power. We developed a novel composite outcome, the ADAS-Cog-Exec, for use in the EXERT trial—a Phase 3 randomized, controlled, 12-month exercise intervention in mild cognitive impairment (MCI).

Methods: Three combinations of cognitive measures selected from the Alzheimer's Disease Assessment Scale-Cognitive Subscale version 13 (ADAS-Cog13), tests of executive function, and the Clinical Dementia Rating (CDR) were created based on previously documented sensitivity to longitudinal change in MCI and to the effects of exercise. Optimally weighted composites of each combination were modeled using data from the ADNI-1 MCI cohort. Ten-fold cross-validation was performed to obtain a bias-corrected mean to standard deviation ratio (MSDR). The cognitive composites were assessed for their sensitivity to detect 12-month change in MCI.

Results: The MSDR of 12-month change for each of the composite outcomes tested exceeded that of the ADAS-Cog13 total score. The composite with the highest MSDR (MSDR = 0.48) and associated statistical power included scores on ADAS-Cog13 Word

Recall, Delayed Word Recall, Orientation, and Number Cancellation subtests; Trail-Making Tests A & B, Digit Symbol Substitution and Category Fluency; and cognitive components of the CDR (Memory, Orientation, Judgement & Problem Solving).

Discussion: An optimally weighted cognitive composite measure was identified and validated for use in EXERT. This composite contained selected subtests from the ADAS-Cog13, additional measures of executive function, and box scores for cognitive components of the CDR. Because this composite score demonstrated high sensitivity to longitudinal change in MCI it will be used as the primary outcome measure for the EXERT trial.

KEYWORDS
aerobic exercise, Alzheimer's disease, cognition, composite outcome, mild cognitive impairment, randomized controlled trial

1 | INTRODUCTION

Cognitive outcome measures that are commonly used in clinical trials of Alzheimer's disease (AD), such as the Alzheimer's Disease Assessment Scale-Cognitive Subscale (ADAS-Cog),[1] are relatively insensitive to the changes that are typical of the early stages of the disease, and they are limited in their ability to detect subtle longitudinal decline in preclinical and prodromal AD. As a result, it is difficult to achieve acceptable statistical power for early-stage trials without unwieldy sample sizes that are both logistically challenging and unduly expensive. In response to this challenge, trialists are increasingly turning to composite measures wherein selected sensitive test items or subscales are combined, often with optimal weighting. By lessening the influence of atypical performance on a single test on longitudinal trajectory, composite outcomes can decrease intra-individual variability and more accurately reflect change over time in cognitive function. Such linear combinations of multiple cognitive outcome measures have demonstrated improved efficiency and statistical power in clinical trials of preclinical[2,3] and early-stage[4-10] AD. Here we describe the development and validation of a novel cognitive composite measure for use as the primary outcome in the EXERT (Exercise in Adults with Mild Memory Problems) trial (NCT02814526), a Phase 3, multicenter, randomized controlled study to examine the effects of a 12-month structured aerobic exercise intervention on cognition and other measures of brain function in 300 adults with amnestic mild cognitive impairment (MCI).[11]

The primary outcome of EXERT was pre-specified as an optimized cognitive composite that is maximally sensitive both to change over time in MCI and to beneficial effects of aerobic exercise on cognition. This composite, referred to as the ADAS-Cog-Exec, was to include subtests of the ADAS-Cog, version 13 (with Delayed Word Recall and Number Cancellation; ADAS-Cog13)[12] and additional tests of executive function, including the Trail-Making Test, Digit Symbol Substitution Test, and Category and Letter Word Fluency. It was hypothesized that including additional measures of executive function in the ADAS-Cog-Exec would increase sensitivity of the cognitive outcome to treatment effects given prior evidence of beneficial exercise effects on this cognitive domain.[13-19]

2 | METHODS

2.1 | EXERT cognitive outcome measures

Cognitive outcome measures in the EXERT trial protocol include the ADAS-Cog13 (includes Delayed Word Recall, Number Cancellation);[12] Trail-Making Test, Parts A & B;[20] the Digit Symbol Substitution Test from the Wechsler Adult Intelligence Scale-Revised;[21] Category Fluency (Animals and Vegetables); and Letter Word Fluency (Letters F and L). These tests are administered to all participants at baseline and every 6 months thereafter for 18 months. Primary efficacy of the exercise intervention will be assessed at the 12-month time point when the supervised phase of the intervention is completed. Unsupervised exercise will continue for another 6 months to test sustainability of the intervention given the extensive support received in the first 12 months, with a final assessment at month 18. The primary outcome for the trial was pre-specified to be the 12-month change in performance on a supplemented and optimized version of the ADAS-Cog13 that includes subtests shown to be maximally sensitive to change over time in MCI using the latest available scientific data, plus measures of executive function previously shown to improve with increased aerobic exercise in clinical trials of older adults.[13-19]

2.2 | Modeling of potential cognitive composites

Data used for modeling potential cognitive composite outcomes were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). ADNI-1[22] was determined to be the longitudinal cohort study best suited for modeling the ADAS-Cog-Exec composite because of similarities in the cohort characteristics (eg, age, severity of cognitive impairment) and neuropsychological test batteries. Letter Fluency, however, was not administered in ADNI-1 so

this measure from the EXERT battery could not be included in the modeling of the composite outcome. Test data from N = 390 participants with MCI from the ADNI-1 cohort were used to model the EXERT ADAS-Cog-Exec composite; data from seven participants were excluded because their baseline ADAS-Cog13 scores were higher (ie, more impaired) than the highest baseline score observed in the EXERT cohort.

2.3 | Cognitive composites tested

Three combinations of cognitive composite measures were selected and tested based on a priori knowledge and hypotheses about their sensitivity to longitudinal change in MCI and sensitivity to the effects of exercise (Table 1). Composite 1 combined the ADAS-Cog13 total score with scores on the supplemental executive function measures (Trail-Making Test A & B, time to completion; Digit Symbol Substitution Test, total correct; and Category Fluency, mean of Animals and Vegetables). Composite 2 included select subscales of the ADAS-Cog13 plus the supplemental executive function measures. For this composite, ADAS-Cog13 Word Recall, Delayed Word Recall, and Orientation subtests were included based upon prior work showing their high sensitivity to longitudinal decline in adults with MCI.[4] Number Cancellation was included as an additional measure of executive function in Composite 2. Composite 3 included all Composite 2 components plus box scores for the cognitive components of the Clinical Dementia Rating (CDR; ie, Memory, Orientation, Judgement & Problem Solving) that have demonstrated utility in improving sensitivity of a cognitive composite outcome in adults with MCI.[4]

2.4 | Deriving optimal weights

Scores on tests for which larger values indicate more impairment (ie, ADAS-Cog subtests; CDR box scores; Trail-Making time to completion) were transformed so that higher scores indicate better performance. Standard scores (z-scores) for intra-individual longitudinal change were then calculated by subtracting each participant's baseline score from the 12-month follow-up score, obtaining the usual change score measure, then dividing each change score by the baseline standard deviation for that measure. Optimal weights for each of the three composites were calculated using previously described methods[5] wherein weights are derived for each variable that maximize the ratio of the mean to the standard deviation (SD) of composite score change over time. This method for deriving the weights used classic linear algebraic methods to determine the eigenvector corresponding to the largest eigenvalue of the ratio.[5]

2.5 | Statistical analyses

Sensitivities of the three composites to longitudinal decline in MCI were measured using the mean to SD ratio (MSDR) of change from

RESEARCH IN CONTEXT

1. **Systematic review**: The authors reviewed the literature on novel cognitive composite outcome measures and modified versions of the Alzheimer's Disease Assessment Scale-Cognitive Subscale (ADAS-Cog) using electronic databases (eg, PubMed) and search engines (eg, Google Scholar). While previously described measures have demonstrated improved sensitivity to detect cognitive change in mild cognitive impairment (MCI) compared to the ADAS-Cog, the composite described herein was developed specifically to detect the effects of an aerobic exercise intervention.
2. **Interpretation**: We developed a novel cognitive composite outcome measure, called the ADAS-Cog-Exec, that was theoretically derived to be sensitive to the effects of aerobic exercise and optimally weighted to detect change over time in MCI. The sensitivity of this novel composite to detect 12-month change and its associated statistical power far exceeded that of the ADAS-Cog13 total score alone.
3. **Future directions**: Results demonstrate sensitivity of the ADAS-Cog-Exec to detect subtle cognitive decline in MCI; however, evidence of its ability to detect response to an aerobic exercise intervention awaits completion of the EXERT trial.

baseline to 12 months. Because the MSDR of a composite score that is calculated from the same data used to develop the composite score may be inflated, K-fold cross-validation (k = 10) was performed to obtain a bias-corrected MSDR. Sensitivities of the composites to change over time were evaluated by comparing their bias-corrected MSDRs with one another and with the MSDR for the ADAS-Cog total score, wherein a larger MSDR indicates greater sensitivity (ie, greater effect size and/or less noise). Additionally, we performed power analyses to determine the requisite sample size associated with each of these composite measures. All analyses were performed using custom programs from the R statistical computing platform, version 3.5.[23]

3 | RESULTS

The ADNI-1 MCI cohort did not differ significantly from the EXERT cohort recruited up to the time of these analyses in terms of age, education, or baseline scores on the Mini-Mental State Examination (MMSE; Table 2). There were significantly more men in the ADNI-1 cohort than in the current EXERT cohort.

The bias-corrected MSDR of 12-month change from baseline for each of the proposed ADAS-Cog-Exec composites exceeded that of the ADAS-Cog13 total score (Table 3). Of note, using only selected

TABLE 1 Cognitive measures included in composites tested

Composites tested	Composite 1	Composite 2	Composite 3
	• ADAS-Cog13 • Executive Function Tests	• Select ADAS-Cog13 Subtests • Executive Function Tests	• Select ADAS-Cog13 Subtests • Executive Function Tests • CDR Box Scores, Cognitive Components
ADAS-Cog13 Subtests Included	All subtests	Word Recall Delayed Word Recall Orientation Number Cancellation	Word Recall Delayed Word Recall Orientation Number Cancellation
Executive Function Tests Included	Trail-Making A Trail-Making B Digit Symbol Substitution Category Fluency	Trail-Making A Trail-Making B Digit Symbol Substitution Category Fluency	Trail-Making A Trail-Making B Digit Symbol Substitution Category Fluency
Clinical Dementia Rating Box Scores Included			Memory Orientation Judgement & Problem Solving

Abbreviations: ADAS-Cog13, Alzheimer's Disease Assessment Scale-Cognitive Subscale version 13; CDR, Clinical Dementia Rating.

TABLE 2 Baseline characteristics of ADNI-1 and EXERT cohorts

	ADNI-1[a]	EXERT[b]	P-value
N	390	176	
Sex (% female)	35	54	<.001
Age	74.9 ± 7.4	74.6 ± 6.0	.98
Education	15.7 ± 3.0	16.3 ± 2.3	.97
MMSE	27.0 ± 1.8	27.7 ± 2.1	.92
ADAS-Cog13	18.5 ± 6.0	14.9 ± 6.7	.68
CDR-sum of boxes	1.6 ± 0.9	1.5 ± 0.9	.93

[a] N = 7 participants with baseline ADAS-Cog13 scores outside the range observed in EXERT (ie, total score > 33) were excluded.
[b] Characteristics of EXERT cohort recruited as of May 8, 2019.
Abbreviations: ADAS-Cog13, Alzheimer's Disease Assessment Scale-Cognitive Subscale version 13; ADNI, Alzheimer's Disease Neuroimaging Initiative; CDR, Clinical Dementia Rating; MMSE, Mini-Mental State Examination.

TABLE 3 Mean to SD ratio (MSDR) for 12-month change from baseline

	MSDR	95% CI
ADAS-Cog13	0.2730	0.1444, 0.3238
Composite 1[a]	0.3132	0.3104, 0.3159
Composite 2[a]	0.3710	0.3677, 0.3742
Composite 3[a]	0.4780	0.4753, 0.4806

[a] Bias-corrected MSDR using k-fold validation (k = 10).
Abbreviations: ADAS-Cog13, Alzheimer's Disease Assessment Scale-Cognitive Subscale version 13; MSDR, mean to standard deviation ratio; SD, standard deviation.

subtests of the ADAS-Cog13 plus the executive function tasks (Composite 2) rather than the ADAS-Cog13 total score plus the executive function tasks (Composite 1) improved the bias-corrected MSDR from 0.31 to 0.37. For Composite 3, adding box scores for the CDR cognitive components (Memory, Orientation, Judgement & Problem Solving) to Composite 2 further improved the bias-corrected MSDR

to 0.48. Power curves comparing requisite sample sizes for each of these composite measures for effect sizes from 0.3 to 0.5 are shown in Figure 1. Derived weights for Composite 3 are shown in Table 4.

4 | DISCUSSION

We developed and validated an optimally weighted cognitive composite measure, referred to as the ADAS-Cog-Exec, for use as the primary outcome in the EXERT trial (NCT02814526), a Phase 3, multicenter, randomized controlled study to examine the effects of a 12-month structured aerobic exercise intervention on cognition and other measures of brain function in 300 adults with amnestic MCI. Of the three theoretically derived combinations of cognitive scores examined, the composite with the greatest sensitivity to detect 12-month change included ADAS-Cog13 subtests that assess episodic memory (Immediate and Delayed Word Recall), orientation (Orientation) and executive function (Number Cancellation); supplemental measures of executive function (Trail-Making A & B, Digit Symbol Substitution, Category Fluency); and box scores for cognitive components of the CDR (Memory, Orientation, Judgement & Problem Solving). The sensitivity of Composite 3 to detect 12-month change and its associated statistical power far exceeded that of the ADAS-Cog13 total score alone.

Our approach in developing the ADAS-Cog-Exec was guided by previous efforts by others to improve the sensitivity of the ADAS-Cog for use in adults with MCI. The initial attempt to improve sensitivity of the original 11-item ADAS-Cog was to add Delayed Word Recall and Number Cancellation (or Mazes), creating the ADAS-Cog13.[12] As the focus of AD research has shifted toward earlier disease stages, further modifications to the ADAS-Cog are needed to prevent ceiling effects and to provide adequate range of scores to detect change, particularly among adults with early MCI or preclinical AD. Common approaches for improving the utility of the ADAS-Cog for pre-dementia populations include modifying the scoring criteria and adding additional test items. A narrative review of the ADAS-Cog[24] found 31 modified versions and concluded that adding tests of memory, executive function,
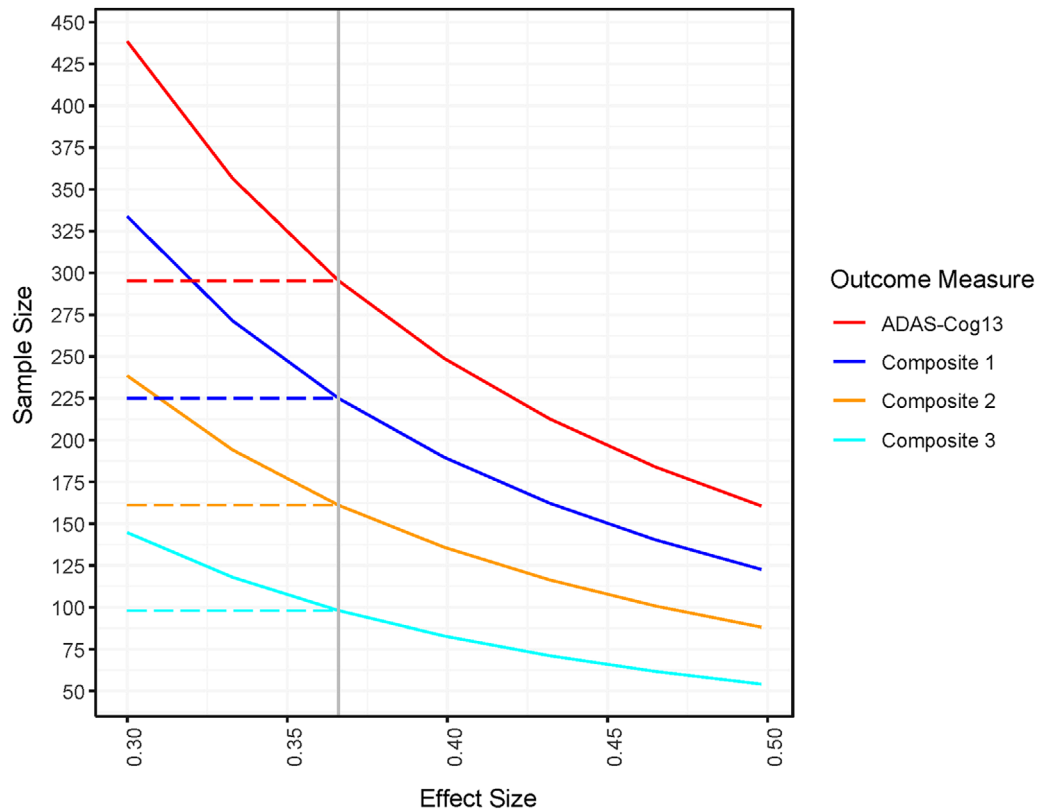
FIGURE 1 Requisite sample size to detect 12-month change (assumes 80% power, 20% attrition, equal group allocation, 2-sided alpha = 0.05; reference effect size 0.366)

TABLE 4 Derived weights for Composite 3

Z-score measure	Estimated weight
ADAS-Cog13 immediate word recall	0.2330
ADAS-Cog13 delayed recall	0.0735
ADAS-Cog13 orientation	0.1088
ADAS-Cog13 number cancellation	−0.2436
Trail-Making A (time to completion)	0.0586
Trail-Making B (time to completion)	0.1080
Digit symbol substitution (number correct)	−0.0577
Category fluency (mean of Animals and Vegetables)	0.1602
CDR memory box score	0.1043
CDR orientation box score	0.3012
CDR Judgement & Problem Solving box score	0.1030

Abbreviation: ADAS-Cog13, ADAS-Cog13, Alzheimer's Disease Assessment Scale-Cognitive Subscale version 13; CDR, Clinical Dementia Rating

and/or daily functioning is one of the best methods for improving the sensitivity of the ADAS-Cog for use in pre-dementia cohorts.

A complementary approach to improve test sensitivity is to remove subtests that have a restricted range or poor power to discriminate level of cognitive impairment. Previous detailed psychometric analyses of the ADAS-Cog-11 revealed that 7 of 11 component subtests had substantial ceiling effects and were notably skewed, even among dementia patients with mild-to-moderate AD; the only subtests that had no significant ceiling effects or notable skew were Word Recall, Word Recognition, and Orientation.[25,26] Raghavan et al.[4] similarly identified Word Recall, Delayed Word Recall, and Orientation as the most informative measures from the ADAS-Cog-13 in MCI and early dementia due to AD. Incorporating box scores from the CDR cognitive items (Memory, Orientation, Judgement & Problem Solving) with these subtests from the ADAS-Cog further enhanced the sensitivity of this novel composite, resulting in improved efficiency in MCI and early AD trials.[4] These earlier studies informed our selection of specific ADAS-Cog and CDR measures to be included in Composites 2 and 3.

The ADAS-Cog-Exec was designed not only to be optimally sensitive to cognitive decline in MCI, but also to detect the hypothesized cognitive benefit from the aerobic exercise intervention. Previous randomized exercise intervention trials have found tests of executive function to be among the most sensitive to the beneficial cognitive effects of aerobic exercise in younger adults,[19] older adults,[14,16-18] and individuals with MCI[13,27] or mild AD.[27,28] Results from these trials informed our decision to include additional outcome measures of executive function in EXERT and our a priori decision to include the Trail-Making Test, Digit-Symbol Substitution Test, and Category Fluency in each of the three potential composite outcomes.

Our results are consistent with previous reports showing that composite measures can improve the statistical power of efficacy

analyses.[2-7,9,29] Cognitive composite measures are increasingly common in AD research, particularly in clinical trials assessing the preclinical and prodromal phases of AD. During these very early phases, cognitive impairment is often subtle, rate of cognitive decline typically is slow, and longitudinal intra-individual variability on single test scores can make detecting change difficult. Composite measures have the potential to track time trends more reliably by reducing the impact of variability attributable to poor performance on a single test measure at a single test session.

Several approaches have been used for creating composite measures for AD trials[8] that include (1) a data-driven approach in which data-reduction techniques such as factor analysis[30] or partial least squares regression[3,9] are used to combine optimally weighted items that depend on their relative contributions in detecting clinical decline; (2) a theory-driven approach in which tests are selected and combined based upon a priori knowledge of their characteristics and sensitivities;[10,31,32] and (3) a global approach in which tests from multiple cognitive domains are combined in an unweighted fashion to provide an aggregate measure of overall performance.[2] The ADAS-Cog-Exec was developed using a combination of theory-driven and data-driven approaches wherein a priori selected measures were optimally weighted using a previously described multivariate mixed model on repeated measures approach.[5]

4.1 | Limitations

The investigation described here supports our initial prediction that the ADAS-Cog-Exec will outperform the ADAS-Cog13 with regard to sensitivity to detect longitudinal decline in participants with MCI. Validation of this composite measure for use in EXERT relied on change observed in the ADNI-1 MCI cohort, which may differ somewhat from the EXERT cohort. Although the cohorts are similar with regard to age, education, and overall cognitive functioning, there are significantly more men in the ADNI-1 cohort than in EXERT (see Table 2). Additionally, there may be important differences related to characteristics of individuals who self-select to participate in an exercise trial, or to cohort-specific inclusion/exclusion criteria that are not captured in this investigation. EXERT is ongoing, which limits the extent to which the Composite 3 ADAS-Cog-Exec can be further validated.

4.2 | Conclusion and future directions

The results of this investigation describe the development and validation of an optimal composite measure, the ADAS-Cog-Exec, that will be used as the primary outcome to evaluate efficacy in the EXERT randomized controlled exercise intervention trial in adults with MCI. Sensitivity to change in cognition and associated statistical power was improved relative to the ADAS-Cog13 total score by a cognitive composite measure composed of ADAS-Cog13 subtests of memory, orientation, and executive function; supplemental tests of executive function; and CDR subscales of memory, orientation, and judge-

ment/problem solving. These results provide preliminary but promising evidence that the new ADAS-Cog-Exec composite may be a valuable and sensitive tool for detecting subtle cognitive changes linked to disease progression—and therefore response to intervention—in adults with MCI. Full validation of the ability of this composite to accomplish these goals awaits completion of the EXERT trial.

CONFLICTS OF INTEREST

REFERENCES

1. Mohs RC, Rosen WG, Davis KL. The Alzheimer's disease assessment scale: an instrument for assessing treatment efficacy. *Psychopharmacol Bull*. 1983;19:448-450.
2. Donohue MC, Sperling RA, Salmon DP, et al. The preclinical Alzheimer cognitive composite: measuring amyloid-related decline. *JAMA Neurol*. 2014;71:961-970.
3. Langbaum JB, Hendrix SB, Ayutyanont N, et al. An empirically derived composite cognitive test score with improved power to track and evaluate treatments for preclinical Alzheimer's disease. *Alzheimers Dement*. 2014;10:666-674.
4. Raghavan N, Samtani MN, Farnum M, et al. The ADAS-Cog revisited: novel composite scales based on ADAS-Cog to improve efficiency in MCI and early AD trials. *Alzheimers Dement*. 2013;9:S21-31.
5. Xiong C, Luo J, Morris JC, Bateman R. Linear combinations of multiple outcome measures to improve the power of efficacy analysis—application to clinical trials on early stage Alzheimer disease. *Biostat Epidemiol*. 2017;1:36-58.
6. Ard MC, Raghavan N, Edland SD. Optimal composite scores for longitudinal clinical trials under the linear mixed effects model. *Pharm Stat*. 2015;14:418-426.
7. Shifflett B, Jacobs DM, Salmon DP, et al. Composite endpoints for Alzheimer's disease clinical trials: improved performance via optimal weighting of component measures. *Alzheimer's & Dementia*. 2018;14.
8. Jonaitis EM, Koscik RL, Clark LR, et al. Measuring longitudinal cognition: individual tests versus composites. *Alzheimers Dement*. 2019;11:74-84.
9. Wang J, Logovinsky V, Hendrix SB, et al. ADCOMS: a composite clinical outcome for prodromal Alzheimer's disease trials. *J Neurol Neurosurg Psychiatry*. 2016;87:993-999.
10. Jutten RJ, Harrison JE, Lee Meeuw Kjoe PR, et al. Assessing cognition and daily function in early dementia using the cognitive-functional composite: findings from the Catch-Cog study cohort. *Alzheimers Res Ther*. 2019;11:45.
11. Baker LD, Cotman C, Morrison RH, et al. EXERT: a phase 3 multi-site randomized controlled trial of aerobic exercise in MCI - study design and methods. *Alzheimer's & Dementia*. 2017;13:P613.
12. Mohs RC, Knopman D, Petersen RC, et al. Development of cognitive instruments for use in clinical trials of antidementia drugs: additions to the Alzheimer's Disease Assessment Scale that broaden its scope. *Alzheimer Dis Assoc Disord*. 1997;11(suppl 2):S13-21.
13. Baker LD, Frank LL, Foster-Schubert K, et al. Effects of aerobic exercise on mild cognitive impairment: a controlled trial. *Arch Neurol*. 2010;67:71-79.
14. Baker LD, Frank LL, Foster-Schubert K, et al. Aerobic exercise improves cognition for older adults with glucose intolerance, a risk factor for Alzheimer's disease. *J Alzheimers Dis*. 2010;22:569-579.
15. Colcombe S, Kramer AF. Fitness effects on the cognitive function of older adults: a meta-analytic study. *Psychol Sci*. 2003;14:125-130.
16. Kramer AF, Hahn S, Cohen NJ, et al. Ageing, fitness and neurocognitive function. *Nature*. 1999;400:418-419.
17. Kramer AF, Erickson KI, Colcombe SJ. Exercise, cognition, and the aging brain. *J Appl Physiol*. 2006;101:1237-1242.
18. Sink KM, Espeland MA, Castro CM, et al. Effect of a 24-month physical activity intervention vs health education on cognitive outcomes in sedentary older adults: the LIFE randomized trial. *JAMA*. 2015;314:781-790.
19. Stern Y, MacKay-Brandt A, Lee S, et al. Effect of aerobic exercise on cognition in younger adults: a randomized clinical trial. *Neurology*. 2019;92:e905-e16.
20. Army Individual Test Battery. *Manual of Directions and Scoring*. Washington, DC: War Department, Adjutant General's office; 1944.
21. Wechsler D. Wechsler Adult Intelligence Scale-Revised: The Psychological Corporation; 1981.
22. Aisen PS, Petersen RC, Donohue MC, et al. Clinical core of the Alzheimer's Disease Neuroimaging Initiative: progress and plans. *Alzheimers Dement*. 2010;6:239-246.
23. R Core Team. *R: A language and environment for statistical computing*. Vienna, Austria; 2018.
24. Kueper JK, Speechley M, Montero-Odasso M. The Alzheimer's Disease Assessment Scale-Cognitive Subscale (ADAS-Cog): modifications and responsiveness in pre-dementia populations. *A Narrative Review J Alzheimers Dis*. 2018;63:423-444.
25. Cano SJ, Posner HB, Moline ML, et al. The ADAS-cog in Alzheimer's disease clinical trials: psychometric evaluation of the sum and its parts. *J Neurol Neurosurg Psychiatry*. 2010;81:1363-1368.
26. Hobart J, Cano S, Posner H, et al. Putting the Alzheimer's cognitive test to the test II: Rasch Measurement Theory. *Alzheimers Dement*. 2013;9:S10-20.
27. Ohman H, Savikko N, Strandberg TE, Pitkala KH. Effect of physical exercise on cognitive performance in older adults with mild cognitive impairment or dementia: a systematic review. *Dement Geriatr Cogn Disord*. 2014;38:347-365.
28. Hoffmann K, Sobol NA, Frederiksen KS, et al. Moderate-to-high intensity physical exercise in patients with Alzheimer's disease: a randomized controlled trial. *J Alzheimers Dis*. 2016;50:443-453.
29. Burnham SC, Raghavan N, Wilson W, et al. Novel statistically-derived composite measures for assessing the efficacy of disease-modifying therapies in prodromal Alzheimer's disease trials: an AIBL Study. *J Alzheimers Dis*. 2015;46:1079-1089.
30. Dowling NM, Hermann B, La Rue A, Sager MA. Latent structure and factorial invariance of a neuropsychological test battery for the study of preclinical Alzheimer's disease. *Neuropsychology*. 2010;24:742-756.
31. Clark LR, Racine AM, Koscik RL, et al. Beta-amyloid and cognitive decline in late middle age: findings from the Wisconsin Registry for Alzheimer's Prevention study. *Alzheimers Dement*. 2016;12:805-814.
32. Skinner J, Carvalho JO, Potter GG, et al. The Alzheimer's Disease Assessment Scale-Cognitive-Plus (ADAS-Cog-Plus): an expansion of the ADAS-Cog to improve responsiveness in MCI. *Brain Imaging Behav*. 2012;6:489-501.