A composite measure of cognitive and functional progression in Alzheimer's disease: Design of the Capturing Changes in Cognition study

ELSEVIER

Alzheimer's & Dementia: Translational Research & Clinical Interventions ■ (2017) 1-9

Alzheimer's & Dementia

Featured Article

A composite measure of cognitive and functional progression in Alzheimer's disease: Design of the Capturing Changes in Cognition study

Roos J. Jutten[a,*], John Harrison[a,b], Frank Jan de Jong[c], André Aleman[d], Craig W. Ritchie[e], Philip Scheltens[a], Sietske A. M. Sikkes[a,f]

[a]*Alzheimer Center, Department of Neurology, VU University Medical Center, Amsterdam Neuroscience, Amsterdam, The Netherlands*
[b]*Metis Cognition Ltd, Park House, Wiltshire, United Kingdom*
[c]*Department of Neurology, Erasmus Medical Center, Rotterdam, The Netherlands*
[d]*Department of Neurosciences, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands*
[e]*Centre for Dementia Prevention, University of Edinburgh, Edinburgh, United Kingdom*
[f]*Department of Epidemiology & Biostatistics, VU University Medical Center, Amsterdam, The Netherlands*

Abstract

Introduction: Cognitive testing in Alzheimer's disease (AD) is essential for establishing diagnosis, monitoring progression, and evaluating treatments. Assessments should ideally be brief, reliable, valid, and reflect clinically meaningful changes. There is a lack of instruments that meet all these criteria. In the Capturing Changes in Cognition (Catch-Cog) study, we seek to correct these deficiencies through the development and validation of a composite measure combining cognition and function: the cognitive-functional composite (CFC). We expect that the CFC is able to detect clinically relevant changes over time in early dementia stages of AD.

Methods/Design: We will include patients (n = 350) with mild cognitive impairment or mild dementia due to AD from memory clinics in the Netherlands and the United Kingdom. We will include cognitively healthy volunteers (n = 30) as a control group. The CFC is based on the "cognitive composite" and the Amsterdam instrumental activities of daily living questionnaire. We will investigate test–retest reliability with baseline and 2- to 3-week follow-up assessments (n = 50 patients and n = 30 healthy controls). We will involve experts and participants to evaluate the initial feasibility and refine the CFC if needed. Subsequently, we will perform a longitudinal construct validation study in a prospective cohort (n = 300) with baseline, 3-, 6-, and 12-month follow-up assessments. The main outcome is cognitive and functional progression measured by the CFC. Reference measures for progression include traditional cognitive and functional tests, disease burden measures, and brain imaging methods. Using linear mixed modeling, we will investigate longitudinal changes on the CFC and relate these to the reference measures. Using linear regression analyses, we will evaluate the influence of possible confounders such as age, gender, and education on the CFC.

Conclusion: By performing an independent longitudinal construct validation, the Catch-Cog study of the novel CFC will contribute to the improvement of disease monitoring and treatment evaluation in mild AD.

Keywords: Alzheimer's disease; Cognition; Composite measure; Daily function; Longitudinal construct validation; Mild cognitive impairment; Prospective cohort

*Corresponding author. Tel.: +31 20 4448527; Fax: ■■■.
E-mail address: r.jutten@vumc.nl

http://dx.doi.org/10.1016/j.trci.2017.01.004

1. Background

Assessing cognition in Alzheimer's disease (AD) is essential for establishing diagnosis, monitoring progression, and evaluating treatments [1,2]. Commonly used cognitive tests have shown adequate quality for diagnostic use [3,4]. However, the quality of these tests for the measurement of changes over time remains questionable [5].

One limitation is the duration of cognitive assessment, which can take up to several hours. This can be burdensome for patients and result in fatigue and loss of concentration. These factors add to measurement error and may be a reason for patients to abort the testing procedure [6]. A European Task Force suggested that measuring progression in mild AD should focus on the domains that are vulnerable for decline, specifically episodic memory (EM), working memory (WM), and executive functioning (EF) [7]. A benefit of this specificity is more concise testing.

A variety of tests are available for the previously specified domains [8]. However, most of these are unable to detect changes over time in mild cognitive impairment (MCI) and mild AD [9]. For example, mixed results are found for the cognitive part of the Alzheimer's Disease Assessment Scale (ADAS-Cog), a test battery frequently used to evaluate therapies in AD [10]. Previous studies have demonstrated that most ADAS-Cog subtests suffer from either floor or ceiling effects in MCI and mild AD, which strongly limits their sensitivity to changes over time [11–13]. However, there is also evidence that some parts show good responsiveness in these disease stages [14,15]. Potentially sensitive tests for EF originate from the Neuropsychological Test Battery (NTB) [16]. Based on existing data on the ADAS-Cog and NTB, Harrison et al. selected three EM tests and two EF tests with a total administration time of 20 minutes. First results showed this "cognitive composite" (CC) to be a concise and reliable measure in mild AD [17].

Although cognitive performance is an important predictor of everyday life performance, test scores only explain part of the variance in functional status, which limits their clinical relevance [18]. Informant reports measuring "instrumental activities of daily living" (IADL) may complement cognitive assessments to provide a clinically meaningful change [19]. IADL are cognitively complex everyday activities, such as cooking and managing finances [20]. Unfortunately, the psychometric quality of most existing IADL instruments is questionable or unknown [21,22]. Recent promising developments include the Amsterdam IADL questionnaire (A-IADL-Q): an informant-based measure with good psychometric properties regarding reliability, validity, responsiveness, and diagnostic accuracy in early dementia [23–26]. The A-IADL-Q is now incorporated in the European Prevention of Alzheimer's Dementia study given its potential capacity to measure functional changes in preclinical and prodromal AD [27].

Combining sensitive cognitive and functional tests into a single composite measure may yield a useful tool to detect clinically relevant changes over time in MCI and mild AD [28]. This is highly relevant for symptomatic and disease-modifying trials, in which treatments are tested that aim to improve cognition and function [7]. Previous studies have proposed composite measures as end points for longitudinal changes. Most of these involve cognitive tests only [29–31] or address global function without focusing on specific activities of daily living [32], which hampers their clinical relevance. Furthermore, they are designed using retrospective data sets and thus need further validation in independent cohorts. An independently validated measure to detect clinically meaningful changes over time in MCI and mild AD is thus still lacking. Therefore, the "Capturing Changes in Cognition" (Catch-Cog) study has been designed. We aim to develop and validate a short composite measure combining cognition and function: the cognitive-functional composite (CFC). The CFC is based on preparatory work on the CC and A-IADL-Q. We expect that the CFC is able to detect changes over time in MCI and mild AD and that these changes relate to clinical and biological measures associated with disease progression.

2. Methods and design

2.1. Study participants

We will include patients ($n = 350$) with MCI or mild AD. They will be recruited via outpatient memory clinics from the (1) VU University Medical Center (VUmc) Alzheimer Center, Amsterdam, The Netherlands ($n = 140$); (2) the Alzheimer Center Rotterdam, The Netherlands ($n = 50$); (3) the University Medical Center Groningen, The Netherlands ($n = 60$); and (4) the Brain Health Clinic at the University of Edinburgh, United Kingdom ($n = 100$). Before inclusion, participants have undergone a dementia assessment in their center, including medical history, neurological and neuropsychological examination, and brain imaging. Diagnoses are made according to the National Institution on Aging criteria [1,33], in a multidisciplinary diagnostic meeting including at least a neurologist or psychiatrist with neuropsychology input. To ensure mild AD, we will include people with a Mini–Mental State Examination (MMSE) score $\geq$18 [34]. Other inclusion criteria include age $\geq$50; sufficient proficiency of the study language; and availability of a study partner. Exclusion criteria address potential confounders for cognitive and functional decline, specifically presence of another significant neurological or psychiatric disorder; Geriatric Depression Scale score $\geq$6 [35]; and current abuse of alcohol or drugs. We will also exclude people who participate in a clinical trial within our follow-up time frame, to avoid potential practice effects due to repeated cognitive testing.

In the VUmc Alzheimer Center, we will additionally include cognitively healthy participants ($N = 30$) as a control group. They will be recruited from an existing database containing healthy volunteers. Before enrollment, all

participants have undergone a neuropsychological screening to ensure cognitive performance within the range of age- and education-adjusted norms; age ≥50; and availability of a study partner. The Medical-Ethical Committee of the VUmc approved the study for all Dutch centers. The South East Scotland Research Ethic Committee approved the study for the UK site.

2.2. Study design

We will use a mixed-methods design to develop the CFC (see Fig. 1). Based on preparatory work on the CC and A-IADL-Q, we will design a first version of the CFC in our working group (consisting of R.J.J., J.H., F.J., A.A., C.W.R., P.S., and S.A.M.S.). We will pilot test this version in patients (n = 50) and healthy controls (n = 30) to investigate test–retest reliability (baseline and 2- to 3-week follow-up assessments) (A). During the test–retest study, we will evaluate the initial feasibility by interviewing a subsample of patients (n = 15) (B). Additionally, we will investigate experts' needs and wishes for a measure of clinical progression, using an online survey that we will distribute among various professional dementia networks (C). Furthermore, we will involve an advisory board consisting of health care professionals and potential future end users of the CFC (D). We will use input from these experts to establish content validity. Finally, output from all four steps (A–D) will be integrated, discussed in the working group, and used to determine the final version of the CFC.

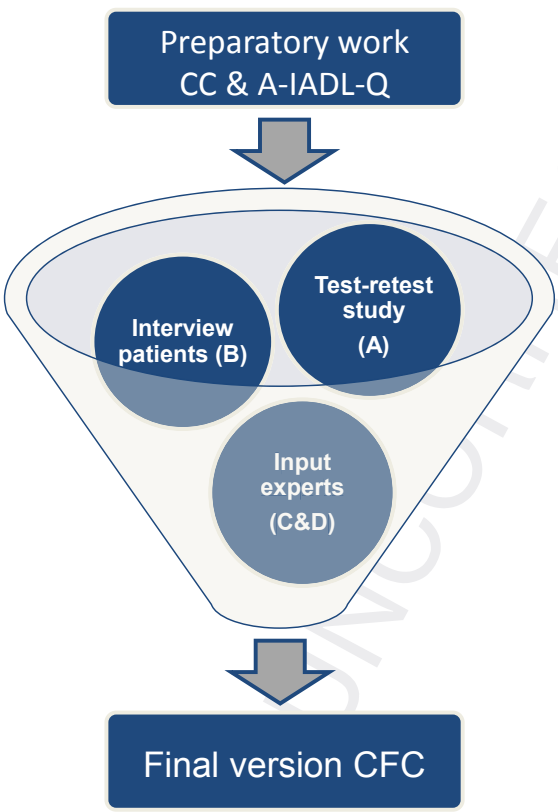

Fig. 1. Development procedure of the cognitive-functional composite. The first version of the CFC is based on the CC and A-IADL-Q. Output from the test–retest study (A), participant interviews (B), expert survey (C), and advisory board (D) will be integrated to determine the final version of the CFC. Abbreviations: A-IADL-Q, Amsterdam IADL questionnaire; CC, cognitive composite; CFC, cognitive-functional composite.

Subsequently, we will perform a longitudinal construct validation study in a prospective cohort with baseline, 3-, 6-, and 12-month follow-up assessments (n = 300). A construct validation approach is chosen [36] because a gold standard for "clinical progression" is lacking. That is, we will include measures that assess different aspects of disease progression, such as subjective perceived decline, disease burden, and brain atrophy. We will also include traditional cognitive and functional tests to compare the CFC with. As shown in Fig. 2, the CFC and reference test of cognition, function, and subjective perceived decline will be assessed at each time point. Disease burden measures will be repeated at 6- and 12-month follow-up. Apathy evaluation and brain imaging will be repeated at 12-month follow-up. For a subgroup (n = 100), the 3-month follow-up will be discarded, to examine potential practice effects that may result from repeated testing within the 3-month time frame [37]. We will compare their trajectory of decline with the subgroup for which the 3-month assessment was retained.

2.3. Outcome parameters

Main outcome parameter is progression in cognition and function measured by the CFC. Reference measures consist of traditional cognitive and functional tests, subjective perceived decline, disease burden measures, and structural brain imaging.

2.3.1. The cognitive-functional composite

The cognitive part of the CFC is based on the CC, which includes (1) ADAS-Cog Word Recognition; (2) ADAS-Cog Orientation; (3) ADAS-Cog Word Recall; (4) Controlled Oral Word Association Test; and (5) Category Fluency Test (see Table 1). Previous work on the CC demonstrated good internal consistency (Cronbach's alpha = 0.80) and test–retest reliability at 4 (r = 0.89), 12 (r = 0.85), 18 (r = 0.84), and 24 weeks (r = 0.84) in mild AD [17]. To cover the EF and WM domains more broadly, we complemented the CC with the Digit Span Backward Task. This test has also been a feature of the NTB [38]. In addition, we included the Digit Symbol Substitution Test. This measure has performed as being sensitive to changes in recently reported clinical drug trials of cognitively enhancing compounds [39]. It has also been listed in recent guidance for dementia drug development as a measure of timed EF, as well as having been selected as the EF component of recently proposed theoretically and empirically driven composite measures for preclinical AD [30,40].

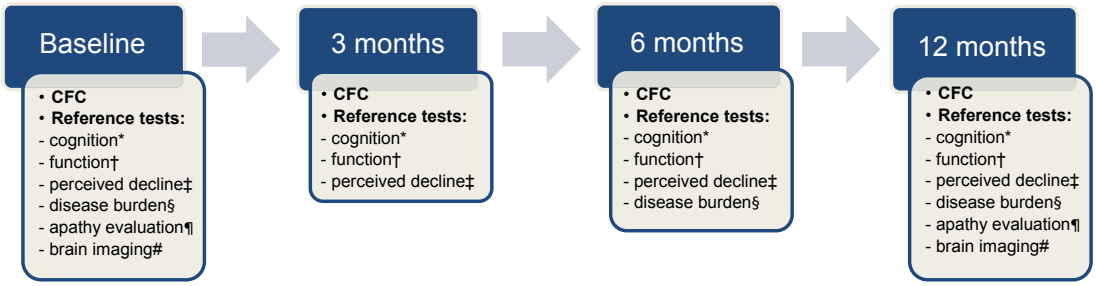

Fig. 2. Schematic overview of the longitudinal construct validation study design. Reference tests including corresponding symbols: *Mini–mental State Examination, Clinical dementia rating scale, and Alzheimer's Disease Assessment Scale–Cognitive subscale. †Alzheimer's Disease Cooperative Study–Activities of Daily Living inventory and Cognitive Function Index. ‡Visual analogue scales for subjective perceived decline in cognitive functioning, everyday functioning, and social functioning. §Zarit Burden Inventory-12 item version and Quality of Life in Alzheimer's disease scale. ¶Apathy Evaluation Scale. #MRI scan including at least 3D-weighted T1, T2 and 3D FLAIR imaging. Abbreviations: CFC, cognitive functional composite; MRI, magnetic resonance imaging; FLAIR, fluid-attenuated inversion recovery.

The functional part of the CFC is based on the A-IADL-Q: an informant-based, computerized questionnaire covering a broad range of IADL activities. For each activity, difficulty in performance is rated on a five-point Likert scale (ranging from "no difficulty in performing this task" to "no longer able to perform this task"). Good psychometric properties have been demonstrated previously: factor analysis supported unidimensionality, high internal consistency (reliability coefficient: 0.97) and good test–retest reliability (κ values > 0.60 for 87.9% of the items) [23]. A construct validation study showed in accordance with prior hypotheses medium to high correlations with traditional measures of everyday and cognitive functioning, suggesting good construct validity [24]. Furthermore, a recent longitudinal validation study demonstrated that the A-IADL-Q was able to measure changes in IADL functioning, in particular in patients with dementia [26]. In the present study, we will use a short version of the A-IADL-Q containing 32 items, which was recently developed and showed good psychometric quality [41].

The ultimate CFC score will be based on the combination of both components. We will explore both theoretically or empirically driven weighting of the subcomponents, to determine what provides most optimal weighting for the score (e.g., use equal weights for all components or differential weights for different components).

2.3.2. Cognitive reference tests

Reference measures for cognition include the MMSE, Clinical Dementia Rating (CDR) scale, and the ADAS-Cog-13. These tests are widely used in both clinical practice and research. The MMSE was originally designed as screening test for the grading of dementia severity [34]. It consists of 30 items all briefly screening different aspects of cognition (e.g., memory, attention and visuospatial skills). Total scores range from 0 to 30, with lower scores reflecting more severe impairment.

The CDR was developed for the staging of dementia severity [42]. Based on an interview with both the study partner and participant, the clinician rates the participant's cognitive and functional performance in six areas: memory, orientation, judgment and problem solving, community affairs, home and hobbies, and personal care. Each area is rated as 0 ("healthy"), 0.5 ("questionable dementia"), 1

Table 1
Overview of the cognitive-functional composite

Content aspects		Administration aspects		
Test	Domain	Completed by	Modality	Duration
Cognitive part				
ADAS-Cog Word Recognition	EM	Participant	On paper	20–25 minutes
ADAS-Cog Orientation	EM			
ADAS-Cog Word Recall	EM			
Digit Span Backward Task	WM			
COWAT	EF			
CFT	EF			
DSST	EF			
Functional part				
A-IADL-Q-SV	IADL	Study partner	Electronical	10–15 minutes

Abbreviations: ADAS-Cog, Alzheimer's Disease Assessment Scale–Cognitive subscale; EM, episodic memory; WM, working memory; COWAT, Controlled Oral Word Association Test; CFT, Category Fluency Test; EF, executive functioning; DSST, Digit Symbol Substitution Test; A-IADL-Q-SV, Amsterdam IADL Questionnaire–short version; IADL, instrumental activities of daily living.

("mild dementia"), 2 ("moderate dementia"), or 3 ("severe dementia"). Adding the rating of all boxes results in a total score ranging from 0 to 18, with higher scores reflecting more severe dementia [43].

The ADAS-Cog-13 is a cognitive test battery that measures cognitive performance by combining ratings of 13 subtests (e.g., constructional praxis, object, and finger naming) [10]. Because three ADAS-Cog-13 subtests are incorporated in the CFC, we will assess the remaining subtests after assessing the CFC. Performance on the CFC ADAS-Cog tests will be included in the scoring. Total scores range from 0 to 85, with higher scores indicating more severe impairment.

2.3.3. Functional reference tests

Reference measures for daily function include the Alzheimer's Disease Cooperative Study—Activities of Daily Living inventory (ADCS-ADL) and the Cognitive Function Instrument (CFI). The ADCS-ADL was designed to assess functional abilities affected in AD and is still widely used in clinical trials [44]. It was developed for a mild-to-moderate AD population and contains both basic and instrumental activities. For 23 different activities, the levels of performance and independency during the past 4 weeks are rated by the study partner. Total scores range from 0 (nonperformance or need for extensive help) to 78 (independent performance).

The CFI was originally developed to detect early clinical changes in individuals at the preclinical stages of AD [45]. The questionnaire includes 14 items that ask about decline in day-to-day cognitive and functional abilities, compared with 1 year ago. Response options include "yes" (0), "no" (1), or "maybe" (0.5), with total scores ranging from 0 to 14. There is a version for the participant and for the study partner with the same questions. In the present study, we will only include the study partner version, as patients are already in the clinical phase of the disease and insight in functioning is likely to be comprised.

2.3.4. Subjective perceived decline

Subjective perceived decline will be measured using visual analogue scales (VAS), ranging from 0 ("no decline") to 100 ("severe decline"). Participants and study partners are independently asked to rate severity of decline in (1) cognitive functioning; (2) everyday functioning; and (3) social functioning, compared to 3 months ago.

2.3.5. Disease burden measures

Caregiver burden will be measured using the short version of the Zarit Burden Inventory (ZBI-12). The ZBI is one of the most commonly used instruments for assessing burden experienced by the caregivers of dementia patients. To minimize respondent burden, we selected the ZBI-12, which was found to produce comparable results to the original version with equal psychometric quality [46]. Each item can be rated on a five-point scale, with total scores ranging from 0 to 48. Higher scores suggest greater caregiver burden.

Quality of life will be measured using the Quality of Life in Alzheimer's Disease scale (QoL-AD) [47]. The QoL-AD was found to be a reliable measure for quality of life in AD patients with an MMSE >10 [48]. We will assess the self-report version for the participant and the informant-based version for the study partner. Both consist of 13 items, rated on a four-point scale. Total scores range from 13 to 52, with higher scores reflecting better quality of life.

Finally, we will include an apathy measure, as apathy can be a predictor of disease severity in AD [49]. We will use the informant-based version of the Apathy Evaluation Scale [50], which consists of 18 statements about the participant's thoughts, feelings, and activity. Each item is rated on a four-point scale. Total scores range from 0 to 72, with higher scores indicating more severe apathy.

2.3.6. Brain atrophy

Brain atrophy will be measured using magnetic resonance imaging (MRI). For each participant, an MRI without contrast will be acquired at baseline and 12-month follow-up. Scans will be performed on 3 Tesla scanners. Sequences include 3D T1-weighted imaging, T2-weighted imaging, and 3D fluid-attenuated inversion recovery (FLAIR). To explore changes in brain activity and functional and structural connectivity in relation to the CFC, a resting state scan (4D T2-weighted imaging) and diffusion tensor imaging will be additionally performed in the research center Groningen. Scans will be analyzed using visual rating and quantitative volumetric imaging tools.

2.3.7. Secondary study parameters

Age, gender, education, cultural background, and disease severity at baseline are secondary study parameters. We will investigate their influence on the CFC and provide norms if necessary. Additionally, we will record whether patients receive any cognitive enhancing treatment during the study period, to ensure that we can account for this afterward.

2.4. Procedures

Eligible participants will receive written and oral information. After 1–2 weeks, the research team contacts the potential participant and study partner to determine whether they are interested to join the study and to answer any further questions. When both are willing to participate, baseline and follow-up visit(s) will be scheduled. At the beginning of the first visit, both the participant and study partner sign the informed consent form in presence of the rater.

Visits take place at either the participants' home or the hospital, depending on the participant's preference, with the requirement that this should be consistent for each study visit. In case of testing at home, separate visits

for the MRI scan will be scheduled nearby the baseline and 12-month follow-up. Study visits are conducted by raters with a background in neuropsychology. To ensure high quality and consistent application, we will organize annual rater meetings which include training in all involved tests.

Each study visit includes a cognitive assessment for the participant, which consists of the cognitive part of the CFC followed by the cognitive reference tests. In the meantime, the study partner completes the functional part of the CFC and the visit-related questionnaires independently on an iPad. Following this, the participant completes the VAS and visit-related disease burden measures on the iPad with assistance from the rater. Finally, the rater completes the remaining interview-based measures with the study partner.

2.5. Sample size

For the test–retest study, we will use the minimal recommended sample size of 50 patients and 30 controls [36]. The sample size for the longitudinal study is based on the objective of investigating the ability of the CFC to detect changes over time. Therefore, sample size formulas for linear models of longitudinal correlated observational data were used [51]. Assuming a power $(1 - \beta)$ of 0.80 and a significance level (α) of 0.05 (two sided), a sample size of 240 patients is sufficient. As we expect a maximum dropout of 20%, we will initially include 300 participants.

2.6. Statistical analyses

We will investigate test–retest reliability of the CFC using intraclass correlations and apply the Bland-Altman method to explore systematic bias such as practice effects. Using baseline data of the longitudinal study, we will investigate the factor structure of the CFC by confirmatory factor analysis. The number of factors will be based on preliminary findings on the CC and A-IADL-Q [17,24]. We will investigate whether the CFC data meet the criteria for item response theory (IRT) or bifactor modeling. Subsequently, we will investigate internal consistency using Cronbach's alpha or IRT reliability coefficients when appropriate.

We will relate longitudinal changes on the CFC (as dependent variable) to changes on the reference measures of disease progression (as independent variables) using linear mixed models with random effects. We will calculate confidence intervals of repeated-measures effect sizes for the CFC and traditional tests. We expect that changes on the CFC moderately relate to changes on the traditional tests but that effect sizes for the CFC are higher than for the traditional tests. We will investigate the clinical relevance of changes by linking actual changes to subjective feelings of change as measured by the VAS. When the data fit an IRT model, we also use anchor-based bookmarking methods to determine the minimal important change [36]. Using linear regression analyses, we will evaluate the influence of possible confounders such as age, gender, and education and investigate whether norms are necessary. When the data fit an IRT model, we will use differential item functioning to explore the influence of possible confounders per item.

3. Discussion

Our aim in the Catch-Cog study is to develop and validate a composite measure combining cognition and function: the CFC. We expect that the CFC is able to detect clinically relevant changes over time in MCI and mild AD. We will investigate this with a test–retest study followed by a longitudinal construct validation in a multicenter, prospective cohort. The CFC is based on preparatory work on the CC and A-IADL-Q. The reliability and validity of these measures have already been demonstrated in existing cohort data [17,23,24,26]. The present study goes a step beyond by performing an independent validation, which is necessary to determine whether the CFC is suitable for implementation in future cohorts and clinical trials [28].

Other composite measures are described in the literature. Recently designed composite measures for detecting cognitive changes in preclinical AD include the theoretical based ADCS preclinical Alzheimer cognitive composite [40] and the empirically derived Alzheimer's Prevention Initiative composites [30,52]. Although some subtests will be able to detect decline in later disease stages (i.e., MCI and mild AD) as well, others will probably show floor effects in these stages. Existing composite measures for MCI and mild AD contain tests that have shown to be sensitive in these stages, and some have also included a functional component [29–32]. However, they do not focus on specific IADL functions. We expect the Catch-Cog study to contribute to this field by designing a composite measure that integrates (1) sensitive cognitive tests and (2) a measure focusing on specific daily skills that are vulnerable for decline in AD. Although there is evidence that cognitive impairment precedes functional impairment in mild AD [53], we do not expect that decline on the CFC will be primarily driven by changes on the cognitive tests. In contrast, we believe that combining our selected cognitive and functional measures may improve statistical power to detect changes and aid the measurement of clinical progression in early dementia stages. The Food and Drug Administration encourages the use of assessment tools that combine cognitive and functional end points, if they are properly validated and have the potential to detect clinically meaningful changes [54].

An important strength of Catch-Cog is the mixed-methods approach for developing and validating the CFC, including the use of input from different stakeholders (e.g.,

patients and experts). This will advance the clinical relevance and acceptability for patients to ease future implementation of the CFC. Another strength includes the international, multicenter character of the study, which enables us to cross-culturally validate the CFC.

A main challenge for this study is the absence of a gold standard for "clinical progression." Furthermore, included reference tests may show limited sensitivity to changes, which could be a potential limitation. We aim to obviate this with a construct validation approach, by involving different clinical and biological measures related to disease progression that are less likely to suffer from ceiling effects, such as hippocampal volume. Second, it could be argued that a follow-up period of 1 year is relatively short for expecting progression in MCI and mild AD. However, both the A-IADL-Q and subtests of the CC have shown to be able to capture changes within the 1-year time frame. We therefore expect the CFC to detect decline after 1 year as well. We also aim to set up future research projects that will address a longer follow-up period for the CFC.

To conclude, we expect Catch-Cog to contribute to the improvement of longitudinal measurement in mild AD. A short and concise composite measure combining cognition and function will advance the monitoring of clinical progression as well as the evaluation of treatment effects.

RESEARCH IN CONTEXT

1. Systematic review: We searched PubMed for publications on measurement instruments for clinically relevant changes over time in mild dementia due to Alzheimer's disease.
2. Interpretation: There is an urgent need for a brief, reliable, valid, and clinically relevant measure, which is able to detect changes over time in mild Alzheimer's disease (AD). In the Catch-Cog study, our aim is to design and validate a composite measure combining sensitive cognitive and functional tests: the cognitive-functional composite (CFC). The CFC is developed based on preparatory work, input from patients and experts, and test–retest analyses. We will investigate its sensitivity over time by performing a longitudinal construct validation study in a multicenter, prospective cohort consisting of subjects with mild cognitive impairment and mild AD.
3. Future directions: By performing an independent longitudinal validation, we expect the novel CFC to contribute to the improvement of disease monitoring and treatment evaluation.

Acknowledgments

The present study is supported by a grant from ZonMw Memorabel (733050205). Research of the VUmc Alzheimer Center is part of the neurodegeneration research program of the Amsterdam Neuroscience Institute. The VUmc Alzheimer Center is supported by Alzheimer Nederland and Stichting VUmc Fonds.

The Amsterdam IADL questionnaire is free for use in all public health and not-for-profit agencies and can be obtained from the authors following a simple registration.

R.J.J., F.J., A.A., and C.W.R., and P.S. report no relevant conflicts of interest. In the past two years, J.H. has received honoraria and paid consultancy from Abbvie, A2Q, Amgen, Anavex, AstraZeneca, Avraham, Axon, Axovant, Biogen Idec, Boehringer Ingelheim, Bracket, Catenion, CRF Health, DeNDRoN, EnVivo Pharma, Enzymotec, ePharmaSolutions, Eisai, Eli Lilly, Forum Pharma, Fresh Forward, GfHEu, Heptares, Janssen AI, Johnson & Johnson, Kaasa Health, Kyowa Hakko Kirin, Lundbeck, MedAvante, Merck, MyCognition, Mind Agilis, Neurocog, Neurim, Neuroscios, Neurotrack, Novartis, Nutricia, Orion Pharma, Pharmanet/i3, Pfizer, Prana Biotech, PriceSpective, Probiodrug, Prophase, Prostrakan, Regeneron, Reviva, Roche, Sanofi, Servier, Shire, Takeda, TCG, TransTech Pharma & Velacor. S.A.M.S. is supported by grants from JPND and ZonMw and has provided consultancy services in the past 2 years for Nutricia and Takeda. All funds were paid to her institution.

References

[1] McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR Jr, Kawas CH, 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;7:263–9.

[2] Snyder PJ, Kahle-Wrobleski K, Brannan S, Miller DS, Schindler RJ, DeSanti S, et al. Assessing cognition and function in Alzheimer's disease clinical trials: do we have the right tools? Alzheimers Dement 2014;10:853–60.

[3] Mathews M, Abner E, Kryscio R, Jicha G, Cooper G, Smith C, et al. Diagnostic accuracy and practice effects in the National Alzheimer's Coordinating Center Uniform Data Set neuropsychological battery. Alzheimers Dement 2014;10:675–83.

[4] Schmand B, Eikelenboom P, van Gool WA, Alzheimer's Disease Neuroimaging I. Value of neuropsychological tests, neuroimaging, and biomarkers for diagnosing Alzheimer's disease in younger and older age cohorts. J Am Geriatr Soc 2011;59:1705–10.

[5] Bossers WJ, van der Woude LH, Boersma F, Scherder EJ, van Heuvelen MJ. Recommended measures for the assessment of cognitive and physical performance in older patients with dementia: a systematic review. Dement Geriatr Cogn Dis Extra 2012;2:589–609.

[6] Narhi V, Laaksonen S, Hietala R, Ahonen T, Lyyti H. Treating missing data in a clinical neuropsychological dataset–data imputation. Clin Neuropsychol 2001;15:380–92.

[7] Vellas B, Andrieu S, Sampaio C, Coley N, Wilcock G. Endpoints for trials in Alzheimer's disease: a European task force consensus. Lancet Neurol 2008;7:436–50.

[8] Schmand B, Rienstra A, Tamminga H, Richard E, van Gool WA, Caan MW, et al. Responsiveness of magnetic resonance imaging and

neuropsychological assessment in memory clinic patients. J Alzheimers Dis 2014;40:409–18.

[9] Mura T, Proust-Lima C, Jacqmin-Gadda H, Akbaraly TN, Touchon J, Dubois B, et al. Measuring cognitive change in subjects with prodromal Alzheimer's disease. J Neurol Neurosurg Psychiatry 2014; 85:363–70.

[10] Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. Am J Psychiatry 1984;141:1356–64.

[11] Karin A, Hannesdottir K, Jaeger J, Annas P, Segerdahl M, Karlsson P, et al. Psychometric evaluation of ADAS-Cog and NTB for measuring drug response. Acta Neurol Scand 2014;129:114–22.

[12] Cano SJ, Posner HB, Moline ML, Hurt SW, Swartz J, Hsu T, et al. The ADAS-cog in Alzheimer's disease clinical trials: psychometric evaluation of the sum and its parts. J Neurol Neurosurg Psychiatry 2010; 81:1363–8.

[13] Podhorna J, Krahnke T, Shear M, E Harrison J, for the Alzheimer's Disease Neuroimaging I. Alzheimer's Disease Assessment Scale–Cognitive subscale variants in mild cognitive impairment and mild Alzheimer's disease: change over time and the effect of enrichment strategies. Alzheimer's Res Ther 2016;8:8.

[14] Raghavan N, Samtani MN, Farnum M, Yang E, Novak G, Grundman 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.

[15] Hendrix S, Ellison N, Stanworth S, Tierney L, Mattner F, Schmidt W, et al. Methodological aspects of the phase II study AFF006 evaluating amyloid-beta -targeting vaccine AFFITOPE(R) AD02 in early Alzheimer's disease—prospective use of novel composite scales. J Prev Alzheimers Dis 2015;2:91–102.

[16] Harrison J, Minassian SL, Jenkins L, Black RS, Koller M, Grundman M. A neuropsychological test battery for use in Alzheimer disease clinical trials. Arch Neurol 2007;64:1323–9.

[17] Harrison J, Dgetluck N, Gawryl M, Moebius H, Hilt D. Validation of a novel cognitive composite assessment for mild and prodromal Alzheimer's disease. Alzheimer's Demen 2013;9:P661.

[18] Royall DR, Lauterbach EC, Kaufer D, Malloy P, Coburn KL, Black KJ. The cognitive correlates of functional status: a review from the Committee on Research of the American Neuropsychiatric Association. J Neuropsychiatry Clin Neurosci 2007; 19:249–65.

[19] Rabin LA, Wang C, Katz MJ, Derby CA, Buschke H, Lipton RB. Predicting Alzheimer's disease: neuropsychological tests, self-reports, and informant reports of cognitive difficulties. J Am Geriatr Soc 2012;60:1128–34.

[20] Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist 1969; 9:179–86.

[21] Sikkes SA, de Lange-de Klerk ES, Pijnenburg YA, Scheltens P, Uitdehaag BM. A systematic review of Instrumental Activities of Daily Living scales in dementia: room for improvement. J Neurol Neurosurg Psychiatry 2009;80:7–12.

[22] Kaur N, Belchior P, Gelinas I, Bier N. Critical appraisal of questionnaires to assess functional impairment in individuals with mild cognitive impairment. Int Psychogeriatr 2016;28:1425–39.

[23] Sikkes SA, de Lange-de Klerk ES, Pijnenburg YA, Gillissen F, Romkes R, Knol DL, et al. A new informant-based questionnaire for instrumental activities of daily living in dementia. Alzheimers Dement 2012;8:536–43.

[24] Sikkes SA, Knol DL, Pijnenburg YA, de Lange-de Klerk ES, Uitdehaag BM, Scheltens P. Validation of the Amsterdam IADL Questionnaire(c), a new tool to measure instrumental activities of daily living in dementia. Neuroepidemiology 2013;41:35–41.

[25] Sikkes SA, Pijnenburg YA, Knol DL, de Lange-de Klerk ES, Scheltens P, Uitdehaag BM. Assessment of instrumental activities of daily living in dementia: diagnostic value of the Amsterdam Instrumental Activities of Daily Living Questionnaire. J Geriatr Psychiatry Neurol 2013;26:244–50.

[26] Koster N, Knol DL, Uitdehaag BM, Scheltens P, Sikkes SA. The sensitivity to change over time of the Amsterdam IADL Questionnaire. Alzheimers Dement 2015;11:1231–40.

[27] Ritchie CW, Molinuevo JL, Truyen L, Satlin A, Van der Geyten S, Lovestone S. Development of interventions for the secondary prevention of Alzheimer's dementia: the European Prevention of Alzheimer's Dementia (EPAD) project. Lancet Psychiatry 2016;3:179–86.

[28] Hendrix SB. Measuring clinical progression in MCI and pre-MCI populations: enrichment and optimizing clinical outcomes over time. Alzheimer's Res Ther 2012;4:24.

[29] Burnham SC, Raghavan N, Wilson W, Baker D, Ropacki MT, Novak G, 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; 46:1079–89.

[30] Langbaum JB, Hendrix SB, Ayutyanont N, Chen K, Fleisher AS, Shah RC, 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;10:666–74.

[31] Coley N, Gallini A, Ousset PJ, Vellas B, Andrieu S. Evaluating the clinical relevance of a cognitive composite outcome measure: an analysis of 1414 participants from the 5-year GuidAge Alzheimer's prevention trial. Alzheimers Dement 2016.

[32] Wang J, Logovinsky V, Hendrix SB, Stanworth SH, Perdomo C, Xu L, et al. ADCOMS: a composite clinical outcome for prodromal Alzheimer's disease trials. J Neurol Neurosurg Psychiatry 2016;87:993–9.

[33] Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011;7:270–9.

[34] Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12:189–98.

[35] Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, et al. Development and validation of a geriatric depression screening scale: a preliminary report. J Psychiatr Res 1983;17:37–49.

[36] de Vet HCW, Terwee CB, Mokkink LB, Knol DL. Measurement in medicine. New York: Cambridge University Press; 2011.

[37] Goldberg TE, Harvey PD, Wesnes KA, Snyder PJ, Schneider LS. Practice effects due to serial cognitive assessment: implications for preclinical Alzheimer's disease randomized controlled trials. Alzheimers Dement 2015;1:103–11.

[38] Harrison J, Rentz DM, McLaughlin T, Niecko T, Gregg KM, Black RS, et al. Cognition in MCI and Alzheimer's disease: baseline data from a longitudinal study of the NTB. Clin Neuropsychol 2014;28:252–68.

[39] McIntyre RS, Harrison J, Loft H, Jacobson W, Olsen CK. The effects of vortioxetine on cognitive function in patients with major depressive disorder: a meta-analysis of three randomized controlled trials. Int J Neuropsychopharmacol 2016.

[40] Donohue MC, Sperling RA, Salmon DP, Rentz DM, Raman R, Thomas RG, et al. The preclinical Alzheimer cognitive composite: measuring amyloid-related decline. JAMA Neurol 2014;71:961–70.

[41] Jutten RJ, Peeters CFW, Leijdesdorff SMJ, Jelle Visser P, Maier AB, Terwee CB, et al. Detecting functional decline from normal ageing to dementia: development and validation of a short version of the Amsterdam IADL Questionnaire. ArXiv e-prints; 2016. Q6

[42] Hughes CP, Berg L, Danziger WL, Coben LA, Martin RL. A new clinical scale for the staging of dementia. Br J Psychiatry 1982; 140:566–72.

[43] Williams MM, Storandt M, Roe CM, Morris JC. Progression of Alzheimer's disease as measured by Clinical Dementia Rating Sum of Boxes scores. Alzheimers Dement 2013;9:S39–44.

[44] Galasko D, Bennett D, Sano M, Ernesto C, Thomas R, Grundman M, et al. An inventory to assess activities of daily living for clinical trials in Alzheimer's disease. the Alzheimer's Disease Cooperative Study. Alzheimer Dis Assoc Disord 1997;11:S33–9.

[45] Amariglio RE, Donohue MC, Marshall GA, Rentz DM, Salmon DP, Ferris SH, et al. Tracking early decline in cognitive function in older individuals at risk for Alzheimer disease dementia: the Alzheimer's disease cooperative study cognitive function instrument. JAMA Neurol 2015;72:446–54.

[46] Bedard M, Molloy DW, Squire L, Dubois S, Lever JA, O'Donnell M. The Zarit Burden Interview: a new short version and screening version. Gerontologist 2001;41:652–7.

[47] Logsdon RG, Gibbons LE, McCurry SM, Teri L. Assessing quality of life in older adults with cognitive impairment. Psychosom Med 2002; 64:510–9.

[48] Thorgrimsen L, Selwood A, Spector A, Royan L, de Madariaga LM, Woods RT, et al. Whose quality of life is it anyway? The validity and reliability of the Quality of Life-Alzheimer's Disease (QoL-AD) scale. Alzheimer Dis Assoc Disord 2003;17:201–8.

[49] Godefroy O, Bakchine S, Verny M, Delabrousse-Mayoux JP, Roussel M, Pere JJ. Characteristics of Alzheimer's disease patients with severe executive disorders. J Alzheimers Dis 2016; 51:815–25.

[50] Marin RS, Biedrzycki RC, Firinciogullari S. Reliability and validity of the Apathy Evaluation Scale. Psychiatry Res 1991;38:143–62.

[51] Liu G, Liang KY. Sample size calculations for studies with correlated observations. Biometrics 1997;53:937–47.

[52] Ayutyanont N, Langbaum JB, Hendrix SB, Chen K, Fleisher AS, Friesenhahn M, et al. The Alzheimer's prevention initiative composite cognitive test score: sample size estimates for the evaluation of preclinical Alzheimer's disease treatments in presenilin 1 E280A mutation carriers. J Clin Psychiatry 2014;75:652–60.

[53] Liu-Seifert H, Siemers E, Price K, Han B, Selzler KJ, Henley D, et al. Cognitive impairment precedes and predicts functional impairment in mild Alzheimer's disease. J Alzheimers Dis 2015;47:205–14.

[54] Food and Drug Administration. Draft guidance for industry. Alzheimer's disease: developing drugs for the treatment of early stage disease. Washington, DC: Center for Drug Evaluation and Research; 2013.