Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring 1 (2015) 521-532

Cognitive & Behavioral Assessment

Five-year biomarker progression variability for Alzheimer's disease dementia prediction: Can a complex instrumental activities of daily living marker fill in the gaps?

Ioannis Tarnanas[a,b,*], Anthoula Tsolaki[c], Mark Wiederhold[d], Brenda Wiederhold[e], Magda Tsolaki[c]

[a]Health-IS Lab, Department of Management, Technology and Economics, ETH Zurich, Zurich, Switzerland
[b]Piaget Research Foundation, Nürnberg, Germany
[c]3rd Department of Neurology, Medical School, Aristotle University of Thessaloniki, Thessaloniki, Greece
[d]Division of Cognitive and Restorative Neurology, Virtual Reality Medical Center, San Diego, CA, USA
[e]Virtual Reality Medical Institute, Clos Chapelle aux Champs, Brussels, Belgium

Abstract

Introduction: Biomarker progressions explain higher variability in cognitive decline than baseline values alone. This study examines progressions of established biomarkers along with a novel marker in a longitudinal cognitive decline.

Methods: A total of 215 subjects were used with a diagnosis of normal, mild cognitive impairment (MCI) or Alzheimer's disease (AD) at baseline. We calculated standardized biomarker progression rates and used them as predictors of outcome within 5 years.

Results: Early cognitive declines were more strongly explained by fluorodeoxyglucose-positron emission tomography, precuneus and medial temporal cortical thickness, and the complex instrumental activities of daily living (iADL) marker progressions. Using Cox proportional hazards model, we found that these progressions were a significant risk factor for conversion from both MCI to AD (adjusted hazard ratio 1.45; 95% confidence interval 1.20–1.93; $P = 1.23 \times 10^{-5}$) and cognitively normal to MCI (adjusted hazard ratio 1.76; 95% confidence interval 1.32–2.34; $P = 1.55 \times 10^{-5}$).

Discussion: Compared with standard biological biomarkers, complex functional iADL markers could also provide predictive information for cognitive decline during the presymptomatic stage. This has important implications for clinical trials focusing on prevention in asymptomatic individuals.

Keywords: Computerized cognitive assessment; Rate of progression; Diagnostics; Early detection; Biomarker; Biomarker progressions; Cognitive declines; MCI; MRI; PET; Alzheimer's disease

1. Introduction

Accurate and early Alzheimer's disease (AD) staging and differential diagnosis possess a pressing modern challenge, partly fueled by recent AD disease-modifying treatment paradigms that only work if applied during the presymptomatic phase [1]. Accurate and earlier diagnosis of patient states is difficult, partly because, despite the popularity of the AD cascade model [2], amyloid and tau-based, pathologic progressions, such as neuritic plaques and neurofibrillary pathology, are interacting in a much more complex way than previously thought [3]. The complexity of the AD pathologic events is now accepted to occur years before symptomatic onset and it challenges current knowledge of the underlying

The authors state that no conflicts of interest exist for this study.

*Corresponding author. Tel.: +41-71-224-72-44; Fax: +41-632-97-75.
E-mail address: itarnanas@ethz.ch

http://dx.doi.org/10.1016/j.dadm.2015.10.005

pathologic pathways [4]. Determining new diagnostic criteria that incorporate biomarkers to construct models of disease progression enabled the mechanism to stage and stratify patients during the presymptomatic phase [5]. For example, the revised National Institute on Aging -Alzheimer's Association (NIA-AA) criteria [6] helped reduce heterogeneity in trial groups, monitor treatment outcomes, and match persons to presumptive treatments. However, despite the deeper understanding and availability of AD in vivo biomarkers, the evidence base for this is relatively limited [7]. A major challenge is to construct models of disease progression that estimate biomarker ordering and dynamics directly from real-world data sets enabling quantitative evaluation of the disease since its earliest stages [8]. At the presymptomatic stage, this would mean to allow the capturing of healthy individuals at risk of developing AD.

Hypothetical models of AD progression have been proposed that describe presymptomatic sequences in which different biomarkers become abnormal [9]. The most well validated of these models generally propose that cerebrospinal fluid (CSF) amyloid pathology and amyloid positron emission tomography (PET) abnormalities precede CSF phosphorylated and total tau (t-tau), fluorodeoxyglucose-positron emission tomography (FDG-PET) hypometabolism, and measures of brain metabolism precede regional neurodegeneration, e.g., volume and atrophy rate markers derived from structural magnetic resonance (MRI), which all occur before a significant clinical change in cognitive performance test scores [10]. When attempting to validate the ordering of these biomarkers, e.g., Brickman et al. [11], CSF brain amyloidosis, neuronal degeneration, namely elevated CSF tau protein, decreased cortical FDG-PET, and medial temporal atrophy on MRI, the results are always dependent on defining abnormal biomarker levels and choosing cut points, which are not easy to establish. Others are also attempting to determine biomarker ordering using a priori staging based on clinical diagnosis and not informed directly by measured data sets [12]. Such attempts can only provide ordering of a small number of biomarkers and limit the temporal resolution of such models to crude stages (e.g., normal, early mild cognitive impairment [MCI], late MCI, or AD). For instance, empirically derived MCI stages or subtypes demonstrate heterogeneity that is not captured by conventional criteria in MCI cognitive profiles. Conventional profiles are susceptible to false-positive errors, which implicates the result of prior MCI studies and may be diluting important biomarker relationships [13]. Moreover, because the way a biomarker is measured can make a difference in diagnostic accuracy, harmonized protocols are still needed [14–16].

In the context mentioned previously, a recently introduced, probabilistic, event-based model (EBM) provided a generative model of AD progression, as a sequence of events, at which individual biomarkers become abnormal. Recent work [17] demonstrated the EBM's consistent ability to learn normal and abnormal distributions of presymptomatic AD biomarker values from data, without requiring any a priori staging or cut points. Researchers might be using such an approach to stage subjects retrospectively and follow a large elderly cohort over a long period of time. For example, Rembach et al. [18] showed such an analysis in plasma amyloid beta and Lim et al. [19] estimated the rate of change of prodromal AD biomarkers and obtained an average cognitive trajectory over time. Similarly, Tarnanas et al. [20] showed a 2-year rate of change but with the introduction of a novel computer-based marker along with MRI and event-related potential biomarkers in subjects with MCI. However, although a promising approach, one issue not systematically examined previously is whether biomarker changes from baseline value to end point or biomarker changes over all the intermediate time points (referred in this study as biomarker progressions) were more strongly associated with cognitive declines. A recent study [21] examined the relative ability of baseline values versus biomarker progressions at each stage of AD in predicting cognitive declines and proved that progressions explained higher variability in cognitive declines than values at the baseline. This finding provides an improved model of the longitudinal, nonlinear association between biomarker and regional atrophy progressions and shows that future clinical trials would benefit by identifying such biomarker progressions most strongly associated with cognitive and functional declines at later stages [22].

Given the amount of recent accumulated knowledge on normal and abnormal function of biomarker progressions, it is not surprising that computer-processable disease models are taking the lead in drug and biomarker discovery efforts [23]. As an illustration [24], proposed two computer-processable cause-and-effect models are based on the Biological Expression Language (http://www.openbel.org/), which support the automatic reasoning of interlinked molecules, and normal and abnormal biological processes. They argued that computer-processable disease models should be based on cause-and-effect regulatory effects that link upstream causal entities to downstream bioclinical effects. In agreement with that group, we believe that computer-processable disease model approaches would be enhanced with the addition of quantitative, real-life, complex activities of daily living, a computerized cognitive performance data set, such as our complex instrumental activities of daily living (iADL) marker with day-out task (DOT) and dual-task walk (NAV) profiles.

The aim of this study was to examine the relative ability of individual biomarker progressions in relation to our complex iADL marker of longitudinal cognitive and functional declines. We used 5-year longitudinal data at each stage of AD to assess which progressions are associated with such declines. To conduct a fair comparison, analogous to a recent study [21], we standardized all biomarkers and presented

clinical values corresponding to each standard deviation. We hypothesized that the fine-grained staging potential of our complex iADL marker could improve clinical trial designs by predicting (1) conversion from cognitively normal (CN-stable) to MCI (CN-converters) and (2) conversion from MCI (MCI-stable) to AD (MCI-converters), allowing the recruitment of high-risk populations with higher accuracy.

2. Methods

2.1. Data source

Data used in the preparation of this article were obtained from two independent data sets: the Greek Association for Alzheimer's Disease and Related Disorders outpatient memory clinic, belonging to the Third Neurological Clinic of the Aristotle University of Thessaloniki and Virtual Reality Medical Center, San Diego. The study was approved by the Institutional Ethics Review Board at each participating institution, and written consent was obtained from all participants, in accordance with the Declaration of Helsinki.

We downloaded data from the initiative's database on March 5, 2015, and included the following for the current analysis: CSF, FDG-PET, and amyloid PET biomarker and MRI scans at baseline and follow-up that met global quality control criteria. Finally, we performed a binary classification of cognitively normal subjects into those who have a stable diagnosis of cognitively normal (CN-stable) and those who convert to MCI (CN-converters). The same procedure was used for MCI subjects into those who have a stable diagnosis of MCI (MCI-stable) and those who convert to AD (MCI-converters). Stable subjects were defined as those with a cognitively normal or MCI diagnosis who remained the same at the end of the 12-, 24-, 36-, 48-, or 60-month follow-up.

2.2. Participants

A total of 350 people were enrolled from which 65 were excluded due to incident vascular events, and 75 were not able to complete the full duration of the study and were considered dropouts. In the end, 215 subjects with valid data for our variables of interest, from which 71 with normal cognition, 61 with MCI, and 83 with AD using baseline diagnosis, were used in this study. Subjects with normal cognition did not meet criteria for dementia or MCI [25,26], had a mini-mental state examination (MMSE) score between 24 and 30, and a global clinical dementia rating (CDR) [27] score of 0. MCI subjects had a CDR of 0.5, MMSE score between 24 and 30, evidence of objective memory loss or a memory complaint (as measured by education adjusted scores on the Wechsler Memory Scale Logical Memory II), absence of significant other cognitive domains impairment, essentially preserved activities of daily living, and absence of dementia. Mildly demented AD participants had MMSE scores between 20 and 26, global CDR scores of 0.5 or 1.0, and met the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association criteria for probable AD [28]. To minimize the influence of vascular pathology, we used the Hachinski ischemic index and excluded participants with significant vascular disease burden at study baseline. Baseline and slopes of progression of the following biomarkers were examined: MRI total brain volumes, hippocampal volumes, ventricular volumes, white matter hyperintensity volumes, CSF t-tau protein and amyloid beta (Aβ)42 levels, cortical thickness of selected regions (precuneus and medial temporal cortical thickness—the latter being the summary variable obtained by adding averaged means for left and right entorhinal, perirhinal, and posterior parahippocampal cortical region thicknesses [29]), and FDG-PET. The CSF t-tau and phosphorylated tau data were log transformed to improve normality. Group biomarker characteristics are summarized in Table 1.

2.3. Neuropsychological examination

All the subjects were assessed with a standardized neuropsychological test battery. MMSE was used to assess global cognitive functioning. Trajectories of memory declines were examined with Alzheimer's Disease Neuroimaging Initiative-memory (ADNI-Mem), which was computed by the use of different word lists in the Rey auditory verbal learning test, the Alzheimer's Disease Assessment Scale-cognitive subscale (ADAS-Cog), and by Logical Memory I data, similar to Crane et al. 2012 [30]. Alzheimer's Disease Neuroimaging Initiative-executive functioning (ADNI-Exe) was used to measure executive decline trajectories and included category fluency (animals), category fluency (vegetables), trails A and B, digit span backward, Wechsler Adult Intelligence Scale-Revised digit symbol substitution, and five clock drawing items (circle, symbol, numbers, hands, and time), similar to Crane et al. 2008 [31]. Considerations for compiling ADNI-Mem and ADNI-Exe included the following: (1) coverage of the domains of interest (memory, executive functions, attention, and visuospatial abilities); (2) ability to measure change over a 2–5 year period; (3) compatibility with previous ADNI biomarker progression studies; (4) being efficient and practical, with low demands for use in our multi-site setting; and (5) avoid ceiling or floor effects. The ADNI-Mem and ADNI-Exe scores are psychometrically optimized, robust, previously validated composite scores of memory and executive function, respectively, with high external validity [32].

2.4. DOT and NAV marker tasks

The complex iADL marker of this study was a complex activity of daily living, which previous research showed as a valid and reliable early indicator of cognitive decline in elderly persons [20,33]. In summary, complex iADL is a set of naturalistic tasks that required coordination of information by eliciting medium-to-high cognitive control,

Table 1
Clinical characteristics at baseline of the subjects (means with SDs)

Characteristics	Normal at baseline		MCI at baseline		AD at baseline	
	Number of assessments available, mean (range)	Baseline values, mean (SD)	Number of assessments available, mean (range)	Baseline values, mean (SD)	Number of assessments available, mean (range)	Baseline values, mean (SD)
n	71		61		83	
Age	N/A	72.0 (9.3)	N/A	72.2 (8.4)	N/A	74.7 (9.8)
Years of education	N/A	14.1 (4.9)	N/A	13.6 (5.4)	N/A	14.4 (4.6)
Female (%)	N/A	53.0	N/A	55.0	N/A	56.0
APOE ε4 (ε4 allele present) (%)	N/A	27.5	N/A	56.2	N/A	71.3
CSF t-tau (pg/mL)	4.2 (1–6)	68.2 (32.2)	4.3 (1–6)	103.1 (50.5)	4.2 (1–6)	123.2 (50.1)
CSF Aβ42 (pg/mL)	4.2 (1–6)	211.1 (52.5)	4.3 (1–6)	164 (56.9)	4.2 (1–6)	141.8 (40.3)
FDG-PET	4.2 (1–6)	1.3 (0.1)	4.2 (1–6)	1.2 (0.1)	4.3 (1–6)	1.1 (0.1)
Brain volume (cm^3)						
WMH	3.4 (1–6)	8.1E24 (3E23)	3.6 (1–6)	8.3E24 (3E23)	4.3 (1–6)	3.3E23 (3E23)
Hippocampal	4.2 (1–6)	3.5 (0.5)	4.3 (1–6)	2.8 (0.5)	4.3 (1–6)	2.6 (0.5)
Ventricular	4.2 (1–6)	17.6 (9.2)	4.3 (1–6)	19.1 (9.8)	4.3 (1–6)	23.1 (10.9)
Total brain	4.2 (1–6)	1072.4 (110.6)	4.3 (1–6)	1053.3 (117.8)	4.3 (1–6)	998.3 (120.0)
WMH/ICV	3.8 (1–6)	6E25% (1.6E24%)	3.6 (1–6)	6E25% (1.7E24%)	4.3 (1–6)	7E25% (1.6E24%)
Hippocampal/ICV	4.2 (1–6)	0.2% (0.03%)	4.3 (1–6)	0.2% (0.03%)	4.3 (1–6)	0.2% (0.03%)
Ventricular/ICV	4.2 (1–6)	1.2% (0.5%)	4.3 (1–6)	1.4% (0.6%)	4.3 (1–6)	1.6% (0.8%)
Total brain/ICV	4.2 (1–6)	69.2% (4.1%)	4.3 (1–6)	66.9% (4.3%)	4.3 (1–6)	66.6% (4.2%)
Thickness (mm)						
Precuneus thickness	4.2 (1–6)	2.1 (0.2)	4.3 (1–6)	2.0 (0.2)	4.3 (1–6)	2.0 (0.3)
Medial temporal thickness*	4.2 (1–6)	5.9 (0.5)	4.3 (1–6)	5.5 (0.7)	4.3 (1–6)	4.9 (0.7)

Abbreviations: SD, standard deviations; MCI, mild cognitive impairment; AD, Alzheimer's disease; N/A, not applicable; APOE, apolipoprotein E; CSF, cerebrospinal fluid; t-tau, total tau; Aβ, amyloid beta; FDG-PET, fluorodeoxyglucose-positron emission tomography; WMH, white matter hyperintensity; ICV, intracranial volume.

*Summary variable by adding averaged means for left and right entorhinal, perirhinal, and posterior parahippocampal cortical region thickness.

such as inhibition of external stimuli or processing speed (e.g., reaction time at interactive events), which is believed to be affected by aging [20] (Appendix).

2.5. FDG-PET metrics and analysis

18F-FDG PET imaging was performed by the Virtual Reality Medical Center, at two different sites in San Diego, CA, USA. At both sites, the FDG-PET images were acquired using a 24 rings General Electric 3D PET/CT device (Discovery ST PET with Light Speed CT), isotropic resolution of 5.99 mm; 15.7-cm axial field of view (FOV); 70-cm transaxial FOV. Following a previously published procedure [12], each FDG-PET image underwent a stringent quality control procedure to assess image quality. We used the FORE-Iterative algorithm to reconstruct images using 48 subsets with five iterations and xy-z filter (cutoff of 4 mm), yielding a 128 × 128 matrix with a pixel size of 1.95 mm (Appendix).

2.6. MRI imaging and analysis

Details of the MRI methodology have previously been described [20]. Cross-sectional regional measures of brain volume for the hippocampus, entorhinal cortex, middle temporal gyrus, fusiform, ventricles, and whole brain, as well as total intracranial volume, were collected on a 1.5-T scanner using a standardized back-to-back 3D magnetization prepared rapid gradient echo (MP-RAGE) protocol: sagittal plane, TR/TE/TI, 2400/3/1000 ms, flip angle 8°, 24-cm FOV, 192 × 192 in-plane matrix, 1.2-mm slice thickness. All regional volumes were normalized by dividing by total intracranial volume for each subject and calculated at baseline using FreeSurfer version 4.3 (http://surfer.nmr.mgh.harvard.edu/).

2.7. CSF biomarkers

Aβ42 and t-tau protein concentrations were our CSF biomarkers. CSF was collected in polypropylene tubes and obtained by lumbar puncture performed with a 20- or 24-gauge spinal needle between L4 and L5 or L3 and L4. The samples were maintained at +4°C and afterward, centrifuged at 2000 × *g* for 5 minutes, then aliquoted and stored at −80°C. Finally, a commercially available enzyme-linked immunosorbent assay (Innogenetics, Ghent, Belgium) was used to determine Aβ42 and t-tau protein concentrations.

2.8. Statistical analysis

Longitudinal trajectories of biomarkers and association between biomarker progressions and outcome were calculated using SPSS 23.0 for Windows (IBM Corporation, New York, USA). Analysis followed a recently

published methodology [21] that estimates the individual-specific slopes of each biomarker's longitudinal trajectory using mixed-effects models. Similar to that work, we considered changes in diagnosis when subjects remained in a new diagnostic category for at least two follow-up assessments. Therefore, we used estimates and observed baseline progressions as predictors of cognitive decline at either ADNI-Mem or ADNI-Exe. When applying our mixed-effects models, we controlled for age, sex, apolipoprotein E (*APOE*) ε4 allele status (ε4 present or absent), years of education, and practice effects (Appendix).

3. Results

When generally examining each biomarker progression across all diagnosis groups, we found that explanatory abilities of both biomarker and our complex iADL marker values increased with progression from normal through MCI to AD. In addition, for both ADNI-Mem and ADNI-Exe, differences in explanatory capacities between biomarker progressions were larger in the MCI and AD groups than those observed among cognitive normal subjects. When examining each biomarker explanatory ability with regard to ADNI-Mem changes, the highest portion of variability across diagnosis groups was explained by progressions of the complex iADL marker, followed by the FDG-PET scores and ventricular volume loss, respectively. However, with regard to ADNI-Exe changes, the order was complex iADL marker, ventricular volume loss, and FDG-PET scores that had the highest explanatory ability. The ability of t-tau to predict memory or executive declines was either null or added further variability to the model. Finally, progression of the CSF Aβ42 level was associated more strongly with memory decline during the MCI stage and executive decline during the AD stage.

In the following sections, we summarize our results per cognitive domain (ADNI-Mem and ADNI-Exe) with tables that show the proportion of variability in ADNI-Mem and ADNI-Exe declines over time, explained by each biomarker progression (Tables 2 and 3). Tables should be read as follows: Among MCI subjects, for example, one standard deviation larger expansion in novel computerized marker scores is associated with a 1.40 further decline in memory scores each year (slope effect: −0.18), and marker progression explained 76% of variability in cognitive decline; whereas for executive scores, the novel computerized marker scores explained 86.3% of variability in executive decline. When biomarker progression values, such as the ones mentioned previously, are a positive percentage, then the corresponding predictor explains the variation in outcome progression, whereas when the inclusion of the predictor adds more estimation error instead of improving model fitting, we noted N/A.

3.1. Associations between biomarker, computerized screening marker progressions and memory decline

When examining the memory domain for normal subjects, the complex iADL marker progressions explained variability at almost 50%, whereas FDG-PET progression, total brain, precuneus, and medial temporal lobe thickness progression explained variability in memory declines, but only to a limited extent: 2%, 6%, 2.9%, and 4.67%, respectively. Among MCI subjects, besides the complex iADL marker scores, other biomarkers that explained the most variability were progression of ventricular volumes (43.8%), followed by shrinkage of medial temporal cortical thickness (32.7%), whole-brain thickness (26%), and hippocampal atrophy (23.4%). Among AD subjects, much higher proportions of variability were explained by biomarker progressions, especially novel complex iADL marker, FDG-PET progression, and total and ventricular brain volume atrophy.

3.2. Associations between biomarker, computerized screening marker progressions and executive function decline

When considering the executive domain for normal subjects, the complex iADL marker progressions again explained more variability at 64%, whereas ventricular volume (23.1%) and total brain thickness (13.1%) showed the highest association with executive function declines among normal subjects. CSF Aβ42 progression, FDG-PET progression, medial temporal lobe, and precuneus thickness progression explained the variability in executive functions to some extent (3.40%–11.70%). Among MCI subjects, as with memory declines, the complex iADL marker had the stronger associations with executive declines at 86.3%. Ventricular volume progression and FDG-PET progression followed in explaining the variability of executive declines (47.5% and 37.5%, respectively). CSF Aβ42 explained slightly more variability in executive declines than in memory declines with 17.6% compared with 12.6% in memory declines. Among AD subjects, the novel marker and ventricular volume progression explained the highest variability of executive function declines (95.7% and 83.3% respectively), followed by FDG-PET scores progression (44.1%). As expected, neither hippocampal volume, medial temporal lobe, and precuneus thickness progression explained the variability of ADNI-Exe decline as much among AD subjects, although they did for ADNI-Mem.

3.3. Prediction of clinical outcomes

According to our findings, the complex iADL marker explained by far the most variability among both memory and executive declines. Because the data collected with our "complex iADL marker" were raw discriminative values and not binary scores, we calculated both the balanced accuracy (BAC) and the area under the curve (AUC). BAC was

Table 2
Proportion of decline in memory function (ADNI-Mem) explained by each biomarker progression

	Normal group		Among MCI*		Among AD	
Biomarker	% Variability explained by biomarkers	Standardized effect size	% Variability explained by biomarkers	Standardized effect size	% Variability explained by biomarkers	Standardized effect size
Novel computerized marker (1 SD = 1.4)	49.00	−0.18	76.00	−0.21	82.00	−0.29
t-tau progression (1 SD = 0.21)	N/A	—	N/A	—	N/A	—
Aβ42 progression (1 SD = 0.15)	1.00	0.08	12.60	−0.41	10.00	−0.10
FDG-PET progression (1 SD = 0.19)	2.00	0.10	17.80	0.13	84.10	0.22
Log_WMH/ICV progression (1 SD = 0.04)	N/A	—	N/A	—	6.10	0.10
HPCV/ICV progression (1 SD = 0.08)	N/A	—	23.40	0.15	35.00	0.19
Ventricles/ICV progression (1 SD = 0.13)	N/A	—	43.80	−0.18	73.80	−0.21
wbrain/ICV progression (1 SD = 0.11)	6.00	0.07	26.00	0.10	30.40	0.19
pthickness progression (1 SD = 0.10)	2.94	0.05	6.58	0.09	8.31	0.11
mtthickness progression (1 SD = 0.13)	4.67	0.12	32.70	0.16	42.80	0.21

Abbreviations: ADNI-Mem, Alzheimer's Disease Neuroimaging Initiative-memory; MCI, mild cognitive impairment; AD, Alzheimer's disease; SD, standard deviation; t-tau, total tau; Aβ, amyloid beta; FDG-PET, fluorodeoxyglucose-positron emission tomography; WMH, white matter hyperintensity; ICV, intracranial volume; HPCV, hippocampal volume; wbrain, total brain volume; pthickness, precuneus thickness; mtthickness, medial temporal cortical thickness; *APOE*, apolipoprotein E.

NOTE. Brain volumes were divided by ICV. Controlling for age at baseline, sex, education, *APOE* ε4 allele (at least one vs. none), and practice effects.

NOTE. N/A: Variability increased instead of decreased or had no changes, after inclusion of the predictors in the model. For instance, including these variables, goodness of fit of the model compared with the null model did not improve because they did not explain the variability of cognitive outcomes or caused more estimation errors instead of explaining the variability.

*To capture changes in diagnosis from MCI to AD during the follow-up, an indicator variable (before AD coded as 0, after AD coded as 1) was included as a control variable to factor in the shift in slopes in cognitive decline among MCI.

defined as the average of the sensitivity and the specificity, obtained by thresholding the prediction values at zero, whereas AUC was calculated using the trapezoid method. Table 4 in the following lists the BAC, sensitivity, and specificity, and AUC over different follow-up durations. The AUC was comparable with previous studies using the "complex iADL marker" for shorter follow-up durations [20]. Table 5 lists the hazard ratio and statistical significance of each variable in the Cox proportional hazards model. Complex iADL marker scores procession was a significant

Table 3
Proportion of decline in executive function (ADNI-Exe) explained by each biomarker progression

	Normal group		Among MCI*		Among AD	
Biomarker	% Variability explained by biomarkers	Standardized effect size	% Variability explained by biomarkers	Standardized effect size	% Variability explained by biomarkers	Standardized effect size
Novel computerized marker (1 SD = 1.4)	64.00	−0.23	86.30	−0.35	95.70	−0.42
t-tau progression (1 SD = 0.21)	N/A	—	N/A	—	N/A	—
Aβ42 progression (1 SD = 0.15)	3.40	−0.10	17.60	−0.31	22.40	−0.30
FDG-PET progression (1 SD = 0.19)	6.00	0.10	37.50	0.14	44.10	0.22
Log_WMH/ICV progression (1 SD = 0.04)	N/A	—	N/A	—	N/A	—
HPCV/ICV progression (1 SD = 0.08)	N/A	—	14.20	0.05	N/A	—
Ventricles/ICV progression (1 SD = 0.13)	23.10	−0.19	47.50	−0.24	83.30	−0.32
wbrain/ICV progression (1 SD = 0.11)	13.10	0.04	36.20	0.11	20.80	0.29
pthickness progression (1 SD = 0.10)	5.40	0.05	6.55	0.12	4.13	0.16
mtthickness progression (1 SD = 0.13)	11.70	0.10	22.00	0.16	23.90	0.22

Abbreviations: ADNI-Exe, Alzheimer's Disease Neuroimaging Initiative-executive functioning; MCI, mild cognitive impairment; AD, Alzheimer's disease; SD, standard deviation; t-tau, total tau; Aβ, amyloid beta; FDG-PET, fluorodeoxyglucose-positron emission tomography; WMH, white matter hyperintensity; ICV, intracranial volume; HPCV, hippocampal volume; wbrain, total brain volume; pthickness, precuneus thickness; mtthickness, medial temporal cortical thickness; *APOE*, apolipoprotein E.

NOTE. Brain volumes were divided by ICV. Controlling for age at baseline, sex, education, *APOE* ε4 allele (at least one vs. none), and practice effects.

NOTE. N/A: Variability increased instead of decreased or had no changes, after inclusion of the predictors in the model. For instance, including these variables, goodness of fit of the model compared with the null model did not improve because they did not explain the variability of cognitive outcomes or caused more estimation errors instead of explaining the variability.

*To capture changes in diagnosis from MCI to AD during the follow-up, an indicator variable (before AD coded as 0, after AD coded as 1) was included as a control variable to factor in the shift in slopes in cognitive decline among MCI.

Table 4
Classification results for discriminating MCI-stable versus MCI-converters and CN-stable versus CN-converters using novel computerized marker progression at different assessments

Results	Balanced accuracy (%)	Sensitivity (%)	Specificity (%)	AUC	n-c/n-s
MCI-converters versus MCI-stable					
12 mo	87	84	87	0.87	14/47
24 mo	88	92	88	0.91	19/42
36 mo	91	93	89	0.91	22/39
48 mo	94	94	95	0.94	27/34
60 mo	93	94	93	0.94	37/24
CN-converters versus CN-stable					
12 mo	89	83	91	0.85	1/70
24 mo	91	87	93	0.88	6/65
36 mo	92	85	95	0.89	9/62
48 mo	93	88	94	0.88	10/61
60 mo	96	100	94	0.95	12/59

Abbreviations: MCI, mild cognitive impairment; CN, cognitively normal; AUC, area under receiver operating characteristic curve; n-c, number of converters; n-s, number of stable subjects.

hazard for conversion from both MCI to AD, and cognitively normal to MCI.

4. Discussion

This study was designed to investigate whether the definition of the preclinical AD phenotype can be improved with the addition of a "complex everyday executive function disruptions" measure through a novel complex iADL marker. When examined longitudinally together with preclinical AD biomarker progressions, our novel marker progressions explained more variability of declines in memory and executive functions in normal, MCI, and AD subjects.

When interpreting these results, it is important to stress that our analysis was based on biomarker progressions of subjects at risk, who subsequently phenoconvert to clinical AD, and not predefined cutoff values. Comparatively exploring the two major cognitive domains investigated, namely memory and executive functions, we produced comparable results to a previous study regarding the individual biomarker progressions' capabilities to explain cognitive declines [21]. We found robust associations of both ADNI-Mem and ADNI-Exe domain declines, for all individual biomarker progressions of the MCI and AD groups. Our complex iADL marker correlated with FDG-PET score changes, ventricular volume increases, whole-brain volume declines, and medial temporal cortical thinning progressions, explaining cognitive progressions in concert with previous research [20,37].

When examining individual biomarker progressions for the normal subjects in our study, we found that the complex iADL marker, CSF Aβ42, and t-tau biomarker progression values, as well as precuneus thickness and medial temporal cortical thinning progressions, explained more variability in ADNI-Exe than ADNI-Mem trajectory declines. Given the small number of data points in this study, e.g., progression from CN-stable to MCI and MCI to AD and the fact that we controlled for vascular brain disease, this finding shows that certain functional or structural biomarkers change relatively late in a long disease process. This trend was more apparent in memory declines than executive functional declines. The interpretation of this finding includes the possibility that executive dysfunction in complex activities is more common among people with AD risk alleles during the presymptomatic stage, and recruitment into trials should take such phenotypes into account. Other studies have also observed differences in memory and executive functioning during the presymptomatic stage among people with AD risk alleles [30,38]. Such difference might be linked to different genetic architecture and different susceptibility to medications designed to modify the underlying biology [39]. As our knowledge and understanding of these phenomena grow, we could limit enrollment of subjects at risk using biomarkers in combination

Table 5
Hazard ratios with 95% confidence intervals for conversion from MCI to AD, and cognitively normal to MCI, obtained by fitting uncorrected and corrected Cox proportional hazards model

End Points	Hazard ratio (CI)	*P* value	Corrected hazard ratio (CI)	Corrected *P* value
MCI to Alzheimer's disease progression				
Computerized marker	1.45 (1.20–1.93)	1.23×10^{-5}*	1.37 (1.14–1.45)	.002*
Age	1.01 (0.98–1.05)	.82	0.99 (0.98–1.04)	.53
Education	0.99 (0.93–1.07)	.61	0.99 (0.91–1.06)	.55
APOE ε4 carrier	1.55 (1.03–2.39)	.053	1.39 (0.88–2.19)	.15
Male	0.73 (0.34–1.15)	.23	0.81 (0.43–1.23)	.49
Cognitively normal to MCI progression				
Computerized marker	1.76 (1.32–2.34)	1.55×10^{-5}*	1.69 (1.26–2.35)	.012*
Age	1.01 (0.94–1.15)	.85	1.00 (0.94–1.19)	.83
Education	1.02 (0.92–1.19)	.71	1.01 (0.87–1.15)	.82
APOE ε4 carrier	3.05 (1.19–7.89)	.018*	2.10 (1.03–6.20)	.12
Male	1.90 (0.85–4.99)	.23	1.44 (0.56–4.22)	.52

Abbreviations: MCI, mild cognitive impairment; AD, Alzheimer's disease; CI, confidence interval; *APOE*, apolipoprotein E.
NOTE. *$P < .05$.

Table 6
Alzheimer's disease biomarker costs, adopted from [8].

Biomarker method	Patient discomfort	Risk	Est. cost per 1000 subjects, $*	Additional considerations
Cerebrospinal fluid	Significant	Moderate to high	350,000–1,000,000	Risks include significant headache (in 40%), back or leg pain (in 11%), and rare meningitis, epidural abscess, or subdural hematoma. Requisite: skill of staff performing procedure.
Neuroimaging	Mild to moderate	Low		
sMRI			400,000–800,000	Claustrophobia, need for lying still for long periods of time, expensive facility and imaging equipment, specialized staff, significant time for post hoc analysis, and variability between facilities.
fMRI			600,000–900,000	
PET			1,000,000–2,000,000	
SPECT			1,000,000–2,000,000	
MRS			700,000–1,000,000	
Blood based	Minimal	Low	40,000–100,000	Possible bruising at the site of venipuncture and vasovagal reaction.
Computerized novel screening marker	Minimal	Low	20,000–80,000†	To date, there is no single, universally accepted computerized screening system that satisfies all needs in the detection of cognitive impairment.

Abbreviations: Est., estimated; sMRI, structural magnetic resonance imaging; fMRI, functional magnetic resonance imaging; PET, positron emission tomography; SPECT, single-photon emission computed tomography; MRS, magnetic resonance spectroscopy.

*Cost calculations based on available online information regarding estimated individual testing charges. These are procedural charges only and do not include the costs of assays performed using cerebrospinal fluid or blood-based analyses or the personnel charges for time spent in association with imaging or fluid-based bioinformatic analyses.

†Cost estimations per Annual Wellness Visit based on Alzheimer's Association recommendations for operationalizing the detection of cognitive impairment in a primary care setting [41].

with the computerized screening marker progression, which gives early indication of this specific subgroup [40]. Most importantly, with the increasing pressure to develop accurate biomarkers of preclinical AD, costs and benefits associated with the various biosignatures need to be considered (Table 6). Our findings provide further evidence that a novel computerized screening marker can be used in combination with or independent of CSF Aβ42 and hippocampal measurements to best identify patients at risk, optimizing the costs/benefits ratio in clinical trials [8,42,43].

This study has several strengths. To the best of our knowledge, this is the first study that provided individual biomarker and a complex iADL marker progressions predicting conversion from CN-stable to MCI and MCI to AD. Previous attempts have focused either on overt disease stages or a subset of biomarkers [4,38,44,45]. The follow-up period in our study was longer than in most previous studies, and potential differences in findings need to be interpreted taking this into account. Finally, we used a systematic approach to model building, and all biomarker/clinical data were collected and processed with uniform standard criteria to address the challenges posed by the multiplicity of potential biomarkers of interest.

Some limitations should be taken into account. First, although we tried to minimize the influence of vascular pathology, we cannot rule out the possible role of subclinical vascular pathology. Second, our sample size was too small for a genome-wide search that could better explain the differences observed. Previous studies estimated required sample sizes per arm for longitudinal trials among subjects with normal cognition when examining biomarker baseline values [46,47]. Based on those studies, enriching with FDG-PET baseline values gives the smallest sample size of 1039 when the outcome is a CDR-sum of boxes. Sample size for our study was computed a priori based on similar studies in the literature [26] and in accordance to Kelley and Rausch recommendations [48] to obtain sufficiently narrow confidence intervals for the model parameters of interest.

Ultimately, as research is linking amyloid and tau pathology as consequences of the AD, instead of the driving mechanism, the focus moves toward trials in presymptomatic populations. Complex iADL impairments can be another observable consequence, and it is particularly important to define useful markers predictive of further cognitive declines. In that context, our results provide support for two major conclusions. First, progressions from a complex iADL marker scores can be used to define at-risk presymptomatic populations as an inexpensive low-risk solution. Second, once a clear definition of the preclinical phenotype is provided, the complex iADL marker progressions can further refine the enrollment of people with MCI progressing to AD, ensuring that those people will benefit early from future interventions.

Acknowledgments

This research was funded by the European Union's Seventh Framework Programme (FP7/2017-2013), the Novel Serious Game Diagnostic Marker initiative project team at the Virtual Reality Medical Center, USA, the Greek

Association for Alzheimer's Disease and Related Disorders outpatient memory clinic financed by the Greek Ministry of Finance and the European IMI project PharmaCog (Prediction of Cognitive Properties of New Drug Candidates for Neurodegenerative Diseases in Early Clinical Development). PharmaCog started in 2010 and is a partnership of 32 academic and industry actors from seven countries. It is one of the most ambitious European projects for tackling bottlenecks in Alzheimer's disease research and drug discovery. Finally, the authors acknowledge all the executive, clinical, and administrative support personnel from Virtual Reality Medical Center and the Greek Association for Alzheimer's Disease and Related Disorders Memory Clinics for their support in quality assurance and data acquisition.

RESEARCH IN CONTEXT

- Systematic review: We reviewed available English language literature in PubMed up to June 2015 using the term "predictive biomarker" to find studies that examined predictors of cognitive declines at each stage of Alzheimer's disease (AD).
- Interpretation: The performance of our computerized screening marker progressions is comparable with that of more established and widely accepted biomarkers, such as fluorodeoxyglucose-positron emission tomography score, precuneus, and medial temporal cortical thickness progression and precuneus values. Persons at risk for AD could benefit through the use of multiple, diverse assessment tools, such as the computerized screening marker capable of reliably identifying cognitive changes at the earliest stages.
- Future directions: Initial strategic steps for integrating computerized screening markers into future development of diagnostic and therapy trial technologies are (1) establish a transsectoral multidisciplinary global network of collaborating investigators as an international shared resource and (2) build the technological platform for managing such a resource.

References

[1] Toyn J. What lessons can be learned from failed Alzheimer's disease trials? Expert Rev Clin Pharmacol 2015;8:1–3.
[2] Weiner MW, Veitch DP, Aisen PS, Beckett LA, Cairns NJ, Green RC, et al. The Alzheimer's Disease Neuroimaging Initiative: A review of papers published since its inception. Alzheimers Dement 2012; 8:S1–68.
[3] Insel PS, Mattsson N, Donohue MC, Mackin RS, Aisen PS, Jack CR, et al. The transitional association between β-amyloid pathology and regional brain atrophy. Alzheimers Dement 2015; 11:1171–9.
[4] Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, et al. Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011;7:280–92.
[5] Sattlecker M, Kiddle SJ, Newhouse S, Proitsi P, Nelson S, Williams S, et al. Alzheimer's disease biomarker discovery using SOMAscan multiplexed protein technology. Alzheimers Dement 2014;10:724–34.
[6] Prestia A, Caroli A, Wade SK, van der Flier WM, Ossenkoppele R, Van Berckel B, et al. Prediction of AD dementia by biomarkers following the NIA-AA and IWG diagnostic criteria in MCI patients from three European memory clinics. Alzheimers Dement 2015;11:1191–201.
[7] Selkoe DJ. Resolving controversies on the path to Alzheimer's therapeutics. Nat Med 2011;17:1060–5.
[8] Fiandaca MS, Mapstone ME, Cheema AK, Federoff HJ. The critical need for defining preclinical biomarkers in Alzheimer's disease. Alzheimers Dement 2014;10:S196–212.
[9] Noel-Storr AH, Flicker L, Ritchie CW, Nguyen GH, Gupta T, Wood P, et al. Systematic review of the body of evidence for the use of biomarkers in the diagnosis of dementia. Alzheimers Dement 2013; 9:e96–105.
[10] Hill DL, Schwarz AJ, Isaac M, Pani L, Vamvakas S, Hemmings R, et al. Coalition against major diseases/European Medicines Agency biomarker qualification of hippocampal volume for enrichment of clinical trials in predementia stages of Alzheimer's disease. Alzheimers Dement 2014;10:421–4293.
[11] Brickman AM, Schupf N, Manly JJ, Stern Y, Luchsinger JA, Provenzano FA, et al. APOE ε4 and risk for Alzheimer's disease: Do regionally distributed white matter hyperintensities play a role? Alzheimers Dement 2014;10:619–29.
[12] Prestia A, Caroli A, Herholz K, Reiman E, Chen K, Jagust WJ, et al. Diagnostic accuracy of markers for prodromal Alzheimer's disease in independent clinical series. Alzheimers Dement 2013;9:677–86.
[13] Edmonds EC, Delano-Wood L, Clark LR, Jak AJ, Nation DA, McDonald CR, et al. Susceptibility of the conventional criteria for mild cognitive impairment to false-positive diagnostic errors. Alzheimers Dement 2015;11:415–24.
[14] Lehmann S, Schraen S, Quadrio I, Paquet C, Bombois S, Delaby C, et al. Impact of harmonization of collection tubes on Alzheimer's disease diagnosis. Alzheimers Dement 2014;10:S390–3942.
[15] Boccardi M, Bocchetta M, Ganzola R, Robitaille N, Redolfi A, Duchesne S, et al. Operationalizing protocol differences for EADC-ADNI manual hippocampal segmentation. Alzheimers Dement 2015;11:184–94.
[16] Frisoni GB, Jack CR. HarP: the EADC-ADNI harmonized protocol for manual hippocampal segmentation. A standard of reference from a global working group. Alzheimers Dement 2015;11:107–10.
[17] Young AL, Oxtoby NP, Daga P, Cash DM, Fox NC, Ourselin S, et al. A data-driven model of biomarker changes in sporadic Alzheimer's disease. Brain 2014;137:2564–77.
[18] Rembach A, Faux NG, Watt AD, Pertile KK, Rumble RL, Trounson BO, et al. Changes in plasma amyloid beta in a longitudinal study of aging and Alzheimer's disease. Alzheimers Dement 2014; 10:53–61.
[19] Lim YY, Maruff P, Pietrzak RH, Ellis KA, Darby D, Ames D, et al. Aβ and cognitive change: Examining the preclinical and prodromal stages of Alzheimer's disease. Alzheimers Dement 2014;10:743–7511.
[20] Tarnanas I, Tsolaki M, Nef T, Muri R, Mosimann UP. Can a novel computerized cognitive screening test provide additional information for early detection of Alzheimer disease? Alzheimers Dement 2014; 10:790–8.
[21] Dodge HH, Zhu J, Harvey D, Saito N, Silbert LC, Kaye JA, et al. Biomarker progressions explain higher variability in stage-specific

cognitive decline than baseline values in Alzheimer disease. Alzheimers Dement 2014;10:690–703.
[22] Drachman DA. The amyloid hypothesis, time to move on: Amyloid is the downstream result, not cause, of Alzheimer's disease. Alzheimers Dement 2014;10:372–80.
[23] Donohue MC, Jacqmin-Gadda H, Le Goff M, Thomas RG, Raman R, Gamst AC, et al. Estimating long-term multivariate progression from short-term data. Alzheimers Dement 2014;10:S400–10.
[24] Kodamullil AT, Younesi E, Naz M, Bagewadi S, Hofmann-apitius M. Computable cause-and-effect models of healthy and Alzheimer's disease states and their mechanistic differential analysis. Alzheimers Dement 2015;2:1–11.
[25] Geuze E, Vermetten E, Bremner JD. MR-based in vivo hippocampal volumetrics: 1. Review of methodologies currently employed. Mol Psychiatry 2005;10:147–59.
[26] Caroli A, Prestia A, Chen K, Ayutyanont N, Landau SM, Madison CM, et al. Summary metrics to assess Alzheimer disease-related hypometabolic pattern with 18F-FDG PET: Head-to-head comparison. J Nucl Med 2012;53:592–600.
[27] McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR, 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.
[28] McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. 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;34:939–44.
[29] Dickerson BC, Wolk DA. Biomarker-based prediction of progression in MCI: Comparison of AD signature and hippocampal volume with spinal fluid amyloid-β and tau. Front Aging Neurosci 2013;5:55.
[30] Crane PK, Carle A, Gibbons LE, Insel P, Mackin RS, Gross A, et al. Development and assessment of a composite score for memory in the Alzheimer's Disease Neuroimaging Initiative (ADNI). Brain Imaging Behav 2012;6:502–16.
[31] Crane PK, Narasimhalu K, Gibbons LE, Pedraza O, Mehta KM, Tang Y, et al. Composite scores for executive function items: Demographic heterogeneity and relationships with quantitative magnetic resonance imaging. J Int Neuropsychol Soc 2008;14:746–59.
[32] Gibbons LE, Carle AC, Mackin RS, Harvey D, Mukherjee S, Insel P, et al. A composite score for executive functioning, validated in Alzheimer's Disease Neuroimaging Initiative (ADNI) participants with baseline mild cognitive impairment. Brain Imaging Behav 2012; 6:517–27.
[33] Tarnanas I, Papagiannopoulos S, Kazis D, Wiederhold M, Widerhold B, Tsolaki M. Reliability of a novel serious game using dual-task gait profiles to early characterize aMCI. Front Aging Neurosci 2015;7:50.
[34] Tarnanas I, Schlee W, Tsolaki M, Müri R, Mosimann U, Nef T. Ecological validity of virtual reality daily living activities screening for early dementia: Longitudinal study. JMIR Serious Games 2013; 1:e1.
[35] Landau SM, Harvey D, Madison CM, Koeppe RA, Reiman EM, Foster NL, et al. Associations between cognitive, functional, and FDG-PET measures of decline in AD and MCI. Neurobiol Aging 2011;32:1207–18.
[36] Della Rosa PA, Cerami C, Gallivanone F, Prestia A, Caroli A, Castiglioni I, et al. A standardized [18F]-FDG-PET template for spatial normalization in statistical parametric mapping of dementia. Neuroinformatics 2014;12:575–93.
[37] Kvartsberg H, Duits FH, Ingelsson M, Andreasen N, Öhrfelt A, Andersson K, et al. Cerebrospinal fluid levels of the synaptic protein neurogranin correlates with cognitive decline in prodromal Alzheimer's disease. Alzheimers Dement 2015;11:1180–90.
[38] Silbert LC, Dodge HH, Perkins LG, Sherbakov L, Lahna D, Erten-Lyons D, et al. Trajectory of white matter hyperintensity burden preceding mild cognitive impairment. Neurology 2012;79:741–7.
[39] Herrup K, Carrillo MC, Schenk D, Cacace A, Desanti S, Fremeau R, et al. Beyond amyloid: Getting real about nonamyloid targets in Alzheimer's disease. Alzheimers Dement 2013;9:452–4581.
[40] Malhotra A, Younesi E, Bagewadi S, Hofmann-Apitius M. Linking hypothetical knowledge patterns to disease molecular signatures for biomarker discovery in Alzheimer's disease. Genome Med 2014;6:97.
[41] Cordell CB, Borson S, Boustani M, Chodosh J, Reuben D, Verghese J, et al. Alzheimer's Association recommendations for operationalizing the detection of cognitive impairment during the Medicare Annual Wellness Visit in a primary care setting. Alzheimers Dement 2013; 9:141–50.
[42] Vellas B, Carrillo MC, Sampaio C, Brashear HR, Siemers E, Hampel H, et al. Designing drug trials for Alzheimer's disease: What we have learned from the release of the phase III antibody trials: A report from the EU/US/CTAD Task Force. Alzheimers Dement 2013;9:438–44.
[43] Van Harten AC, Visser PJ, Pijnenburg YAL, Teunissen CE, Blankenstein MA, Scheltens P, et al. Cerebrospinal fluid Aβ42 is the best predictor of clinical progression in patients with subjective complaints. Alzheimers Dement 2013;9:481–7.
[44] Ye BS, Seo SW, Yang JJ, Kim HJ, Kim YJ, Yoon CW, et al. Comparison of cortical thickness in patients with early-stage versus late-stage amnestic mild cognitive impairment. Eur J Neurol 2014;21:86–92.
[45] Lista S, Garaci FG, Ewers M, Teipel S, Zetterberg H, Blennow K, et al. CSF Aβ1-42 combined with neuroimaging biomarkers in the early detection, diagnosis and prediction of Alzheimer's disease. Alzheimers Dement 2014;10:381–92.
[46] Grill JD, Di L, Lu PH, Lee C, Ringman J, Apostolova LG, et al. Estimating sample sizes for predementia Alzheimer's trials based on the Alzheimer's Disease Neuroimaging Initiative. Neurobiol Aging 2013;34:62–72.
[47] Leoutsakos JM, Bartlett AL, Forrester SN, Lyketsos CG. Simulating effects of biomarker enrichment on Alzheimer's disease prevention trials: Conceptual framework and example. Alzheimers Dement 2014; 10:152–61.
[48] Kelley K, Rausch JR. Sample size planning for longitudinal models: Accuracy in parameter estimation for polynomial change parameters. Psychol Methods 2011;16:391–405.

Appendix

DOT and NAV marker tasks

The complex iADL marker used in this study consisted of two modules simulating complex activities of daily living (CADL): the three-dimensional (3D) immersive reality day-out task (DOT) and the 3D immersive reality spatial navigation only task (NAV). The DOT was a complex task breakdown followed by a rehearsal exercise of a virtual apartment building fire evacuation drill. The drill included six different scenarios of increasing difficulty, where participants navigated the virtual environment using a first-person perspective and simple hand pointing gestures for forward, backward, and left, and right lateral movement, respectively. They could also use natural finger pointing and grabbing gestures to select, pick, drop, and move objects inside the virtual environment and had to complete each within 8 minutes. All participant movement within the virtual building was recorded at 10 Hz and represented as a series of x, y, z coordinates, with actions annotated and time stamped.

The DOT naturalistic actions script was based on an ordered list of right and wrong actions that was prepared by an occupational psychologist and was used to examine executive function and prospective memory as well as planning and reasoning in a complex emergency routine. The fire evacuation drill setting had six different simulated fire situations (from easy to more difficult) taking place at a virtual apartment block with three levels and five apartments per level. The task put a medium-to-high load on the cognitive control processes with which older adults prioritize, organize, initiate, and complete a number of subroutines (e.g., pick-up the phone and call the fire department, sound the fire-alarm) to evacuate safely and in the fastest possible way from an apartment level (e.g., second floor) to the ground area (e.g., determine and gather information on the size of the fire, avoid smoke, avoid wrong actions such as using the elevator). In this sense, DOT is a CADL, which previous research showed is a valid and reliable indicator of cognitive decline in elderly persons [20].

The NAV task took place at the same virtual apartment block but with the player challenged in different aspects of executive function, such as volition, self-awareness, planning, inhibition of dominant response, external distraction during response control, and dual-task coordination [33]. The goal at difficulty level 1 was to navigate from point A to point B, after the route was demonstrated by a first-person perspective camera walk through without iteration. The NAV task took place with six levels of difficulty, with the addition of one more point of destination per difficulty level—for example, level 3 has three points to reach, level 4 has four, and so forth. Each level had a starting position (start) and an end position (goal) and multiple ways to arrive from start to goal. Participants were asked to make their way from start to goal in the shortest time possible.

Both tasks took place at the same virtual reality enriched environment and placed a medium-to-high demand on higher-order cognitive control processes. One of the cognitive control processes with the highest demand was to follow a mental strategy to reach a goal in parallel with behavioral performance monitoring while inhibiting environmental stressors—for example, virtual distractors requiring the player to count stimuli while walking, a process that typically involves high cognitive control. Scoring was computed for both DOT and NAV and has been described before [34]. It basically consists of a quantitative ratio of efficacy that is computed by an algorithm that follows four activity parameters in real-time: (1) omission of one of the activities, (2) repetition of the same activity, (3) incorrect order in performing the activities, and (4) number of attempts before completing a given activity. In addition to the quantitative score mentioned previously, we also get a graphical representation, which illustrates the length of time participants spent at each specific set of x, y, z coordinates in the 3D space. Fig. 1 shows a graphical example of mean group completion performance at DOT.

The order of participating in either the DOT or NAV tasks was random, and both started after each participant had 5 minutes to read written instructions detailing the task, virtual building layout, and task rules. Then, participants practiced the virtual environment using gestures to move around the building and completed three practice runs involving object collection, button pressing, unlocking the stairwell door with a key code, and folder sorting. This also allowed participants to familiarize themselves with the building. None of the practice runs were used in the main task. The practice session took in total approximately 20 minutes.

Participants in this study played all difficulty levels of DOT and NAV in a baseline session (visit 1) and then at the end of the 12-, 24-, 36-, 48-, or 60-month follow-up. To standardize performance during different difficulty levels of the DOT and NAV tasks, we used an algorithm that detects subtle changes in intra-individual variability, which we validated previously [33]. This algorithm reflects a transient, within-person change in behavioral performance, and more particularly, latency (variability across response time performance scores) and accuracy based (variability across accuracy scores—correct vs. wrong responses) and has been reportedly associated with early functional and cognitive decline [20,33,34].

Fluorodeoxyglucose-positron emission tomography metrics and analysis

Alzheimer's disease (AD)–related hypometabolism was computed with the pre-defined regions of interest (MetaROIs) average [35] using the SPM8 15O-H2O PET template [36]. For each fluorodeoxyglucose-positron emission tomography

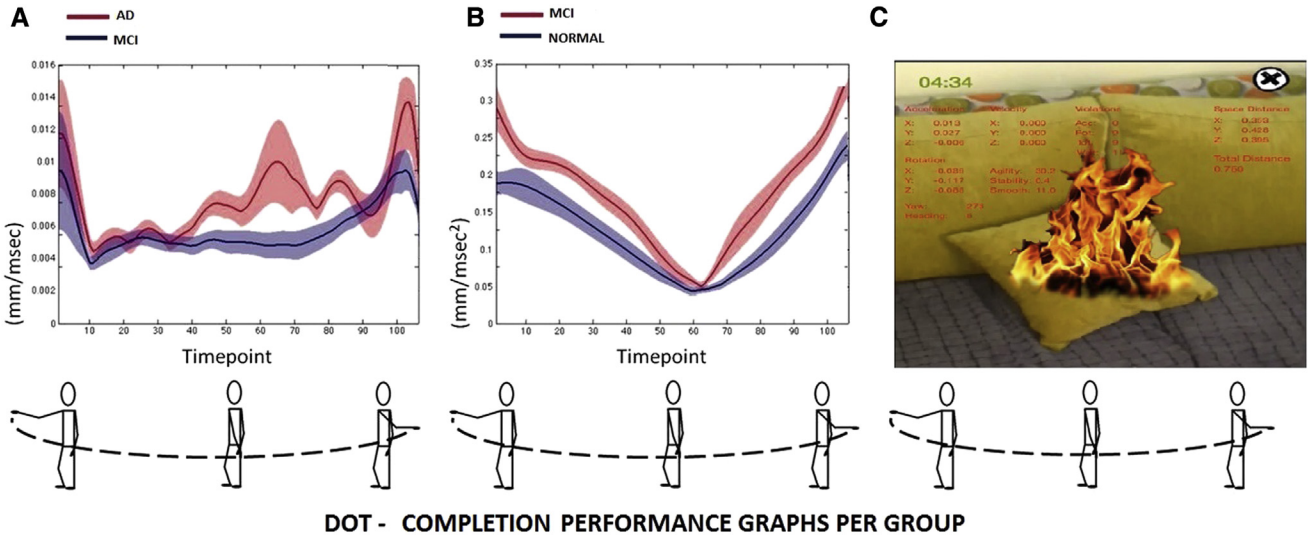

Fig. 1. The graphical representation which illustrates the mean group completion performance profiles of the complex iADL (DOT) from a tablet PC while the users were navigating in 3D space. (A) This is the mean completion time values for the MCI and AD groups while the user was interacting with the different tasks during the complex iADL. (B) This is the mean completion time values for the normal and MCI groups while the user was interacting with the different tasks during the complex iADL, and (C) this is an actual screenshot of DOT, where the user is required to perform fire safety skills and emergency evacuation in the presence of a room fire. A time countdown at the upper left corner is providing gamification and extra pressure for the completion of the task. Abbreviations: iADL, instrumental activities of daily living; DOT, day-out task; 3D, three dimensional; MCI, mild cognitive impairment; AD, Alzheimer's disease.

(FDG-PET) image, we computed metaROI average as the average of the mean counts in five metaROI volumes on spatially and intensity-normalized PET images. All volume metrics were based on voxel-by-voxel analysis based on a meta-analysis of studies [26] carrying out direct whole-brain contrasts of FDG-PET data (summary variable by averaged mean for left and right temporal, right, and left angular and posterior cingulate cortices) and reporting Z-scores or T-values in voxels, showing significantly different mean FDG uptake between patients (AD or mild cognitive impairment [MCI]) and controls at each time point.

Statistical analysis

To examine the amount, magnitude, and timing of practice effects, we used a piecewise approach to observe changes in cognitive slopes. For example, because the normative magnitude and precise timing of when practice effects subside are not established, we allowed slope changes between baseline and 6-month assessment, between 6-month assessment and 12-month follow-up assessment, and so on. This analysis indicated that the Alzheimer's Disease Neuroimaging Initiative-memory function model fitness is best if we assume that the practice effect peaked at the 6-month assessment, and for executive function (ADNI-Exe) at the 12-month assessment, similar to the recent work [21]. We included this change point in the model to control for practice effects. In addition, the variability in cognitive decline (i.e., individual differences in slopes explained by the subject-specific baseline biomarker progressions) was compared with variability in our novel computerized screening market progressions. All biomarker variables were standardized so that estimated effects on cognitive decline could be meaningfully compared across different biomarkers and our novel computerized screening marker. Standardization was done using baseline means and standard deviations, and each model included time in months (0, 6, 12, 18, 24, and 36 months from baseline). Each brain volume was divided by intracranial volume (ICV), and white matter hyperintensity/ICV was first normalized due to common skewed distribution with a log transformation.

Finally, we applied the same mixed-effects models to each stage separately for normal cognition, MCI, and AD. We examined the predictive effect of progressions on cognitive declines within each group, with intercept and time treated as random effects in all models. Changes in diagnoses in the MCI group from MCI to AD were indicated using a variable. All models used restricted maximum likelihood for estimation and assumed an unstructured within-subject error covariance structure. To examine the overall fit of the models, we used (1) formal fit criteria and (2) residual plots inspected visually. All results were considered significant at $P < .05$.