Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring 4 (2016) 149-158

Cognitive & Behavioral Assessment

Modeling practice effects in healthy middle-aged participants of the Alzheimer and Families parent cohort

Gonzalo Sánchez-Benavides[a,*], Juan D. Gispert[a,b], Karine Fauria[a], José Luis Molinuevo[a], Nina Gramunt[a]

[a]Barcelonaβeta Brain Research Center, Pasqual Maragall Foundation, Barcelona, Spain
[b]Universitat Pompeu Fabra, Barcelona, Spain

Abstract

Introduction: Repetitive administration of neuropsychological tests can lead to performance improvement merely due to previous exposure. The magnitude of such practice effects (PEs) may be used as a marker of subtle cognitive impairment because they are diminished in healthy individuals subsequently developing Alzheimer's disease (AD).

Methods: To explore the relationship between sociodemographic factors, AD family history (FH), and *APOE* ε4 status, and the magnitude of PE, four subtests of the Wechsler Adult Intelligence Scale-IV were administered twice to 400 middle-aged healthy individuals, most of them first-degree descendants of AD patients.

Results: PEs were observed in all measures. Sociodemographic variables did not show a uniform effect on PE. Baseline score was the strongest predictor of change, being inversely related to PE magnitude. Significant effects of the interaction term *APOE* ε4*Age in processing speed and working memory were observed.

Discussion: PEs exert a relevant effect in cognitive outcomes at retest and, accordingly, they must be taken into consideration in clinical trials. The magnitude of PE in processing speed and working memory could be of special interest for the development of cognitive markers of preclinical AD.

Keywords: Alzheimer's disease; Preclinical; Cognition; Practice effects; *APOE*; Family history

1. Introduction

Learning from previous experience is at the core of cognitive ability in humans because it provides clear advantages for adaptation [1]. When individuals are repeatedly exposed to a problem or a task, they are expected to improve their performance since they may have developed strategies and memories that help them solve it better and/or in a more efficient way. Tasks offered to an examinee during a neuropsychological assessment are not free of these learning effects, which may influence the interpretation of cognitive change. Even when subject's ability, mood and motivation, and exploration conditions remain stable, prior experience with the tasks could still lead to improvements in performance. These improvements that are merely due to previous experience are referred to as practice effects (PEs) [2]. Although PE were classically viewed as a psychometric confound that should be minimized or adjusted for, it has been more recently suggested that they could represent a useful cognitive variable. Reduced or absence of PE at short intervals have been shown to enable the distinction of individuals with and without cognitive impairment [3–5] and to predict their long-term cognitive outcomes, as shown by Duff and colleagues with an interval of one-week between assessments [6–8]. Recent studies exploring PE at longer-term intervals (e.g., annual assessments with several

*Corresponding author. Tel.: +34-93-316-0990; Fax: +34-93-316-0996.
E-mail address: gsanchezb@fpmaragall.org

http://dx.doi.org/10.1016/j.dadm.2016.07.001

follow-ups) have found similar results: PE are attenuated in asymptomatic subjects that either progressed to mild cognitive impairment (MCI) [9] or to symptomatic Alzheimer's disease (AD) [10]. As a whole, these reports suggest that reduced PE may serve as a valuable indicator of preclinical AD Stage III, since, in addition to positive AD biomarkers, subtle cognitive changes would be present before meeting criteria for a clinical diagnosis (i.e., MCI) [11]. Thus, the study of PE is of special interest because they may be indicative of subtle cognitive changes in persons performing within "psychometrically normal" ranges (i.e., subjects whose baseline and follow-up scores are *per se* not suggestive of cognitive impairment).

Similar to most cognitive variables, the magnitude of PE seems to be influenced by sociodemographic factors, such as age and education. In a thorough meta-analysis, Calamia et al. reported a consistent negative effect of age in PE [12], although some studies have not observed such relationship [13–15]. Although less studied, the level of formal education has also shown disagreeing results. Although some studies found a positive influence [16], other failed to find such relationship [13,17]. Another key variable that is frequently taken into account in PE studies is the length of the time interval between assessments. As previously mentioned, a wide range of intervals have been studied, encompassing administrations within the same day (e.g., [3]), one-week retests (e.g., [18]), and sessions spared by a year or more (e.g., [9]). The general evidence suggests that shorter intervals are related to higher gains at retest, being this improvement virtually zero after 5 years (see [12]), although some reports have found evidence of PE after 7 or more years [19].

Other variables, such as gene pool, can likely influence the magnitude of PE. Presence of the *APOE* ε4 allele of the *Apolipoprotein E* (*APOE*) gene has been related to a small but consistent decrease in cognitive performance in healthy adults [20]. The *APOE* gene genotype is also known to influence the risk of developing late-onset AD, with subjects carrying one or two *APOE* ε4 alleles presenting a 3-fold and 12-fold increased risk, respectively [21]. However, the effect of *APOE* ε4 status only accounts for less than one third of the estimated disease heritability [22] and other genetic and nongenetic factors, such as environmental exposures, lifestyle or nutrition, also modulate the risk of suffering AD. The concept of family history of AD (FH) captures both genetic and nongenetic factors in measuring AD risk (reviewed in [23]). Some studies suggest that FH and *APOE* are independent and additive risk factors for developing the disease [24,25] and that both can be useful as markers to stratify healthy subjects in different risk level groups [26].

Few studies have addressed the impact of carrying an *APOE* ε4 allele and/or having FH of AD in the magnitude of PE. Zehnder in 2009 and Donix in 2012 reported a negative association between the *APOE* ε4 allele and PE in memory tasks [27,28], but, more recently, Jonaitis et al. did not find any relationship in a larger sample with an extended cognitive test battery [26]. By contrast, in this latter performed in the context of the Wisconsin Registry for Alzheimer's Prevention (WRAP) study, Jonaitis et al. did find a slight attenuation of PE related to the number of previous visits in positive FH healthy middle-aged subjects.

In this scenario, in which the effect of subject characteristics on PE remain unclear, and taking into account the possible utility of such cognitive outcome as a marker of preclinical AD, further knowledge on the possible moderator effect of individual variables in PE is necessary. In this study, we aimed to provide further data on the topic by studying the impact of age, sex, education, risk of AD related to FH of the disease, and *APOE* ε4 status, on PE when re-testing 1 to 3 months from baseline.

2. Methods

This study was carried out as part of a wider research platform: the Alzheimer and Families (ALFA; Clinicaltrials.gov Identifier: NCT01835717) parent cohort. Details of the study along with an extended description of inclusion and exclusion criteria are described elsewhere [29]. ALFA participants are cognitively healthy men and women aged between 45 and 74 years, most of them first-degree descendants of AD patients. Participants included in the parental ALFA cohort during the first four months of recruitment (April–July 2013) were consecutively offered the possibility of attending a second visit 6 weeks (±2) after and were included in the present study.

The study was approved by the Ethics Committee of the "Parc de Salut Mar" (Barcelona, Spain) and conducted in accordance to the directives of the Spanish Law 14/2007, of 3rd of July, on Biomedical Research. All participants signed an informed consent form and had a close relative, who also granted their consent, volunteering to participate in the functional assessment of the participant.

2.1. Participants and procedure

In the context of a validation study of a memory task performed within the ALFA parent cohort, 400 individuals aged between 45 and 65 years from this cohort were administered twice four subtest of the Wechsler Adult Intelligence Scale-IV in two visits (Visit 1, V1; Visit 2, V2) separated by a time interval of 6 weeks (±2 weeks). These tests were administered in the time between immediate and delayed recall at both visits. To diminish possible rehearsal in the intervisit interval, participants were not told that they would repeat in V2 exactly the same tests as in V1. Mood state was recorded by means the Goldberg Anxiety and Depression Scale (GADS) at both visits.

Information about vascular risk factors was also collected. The REGICOR cardiovascular risk function, an adaptation of the Framingham function validated in a Spanish sample [30] that estimate participants' risk of suffering coronary disease events at 10 years was calculated. In

addition, the probability of dementia in 20 years using the CAIDE (Cardiovascular Risk Factors, Aging, and Dementia) score [31] was estimated. We calculated both model I and model II CAIDE scores, which differ in the inclusion (model II) or not (model I) of the *APOE* status. The genotype of nine individuals could not be determined.

FH of AD was considered positive when at least one of the participant's parents developed AD before 75 years (FH+). If the participant does not have any parent with AD or they developed AD after 74 years, they were classified as negative (FH−). This 75-year cutoff was selected based on the evidence from recent studies on age-dependent genetic risk that propose that genetic load exerts a higher influence before this age (see [32,33]).

2.2. Cognitive measures

Four WAIS-IV subtests were administered using the Spanish version of the scale [34]. The administered subtests were as follows: The digit span subtest (score range, 0–48) measures short-term and working memory and is composed of three parts: forward (score range, 0–16), backward (score range, 0–16), and sequencing (score range, 0–16). The coding subtest (score range, 0–135) measures, among others, processing speed, short-term visual memory, visual perception, coordination, and attention. The matrix reasoning subtest (score range, 0–26) measures fluid intelligence, visual organization, ability to identify part-whole relationships, and simultaneous processing. Finally, the visual puzzles subtest (score range, 0–26) assesses visuospatial and fluid intelligence, spatial manipulation, and the anticipation of relationships between parts. All tests were administered by trained neuropsychologists following the instructions of the published manual. More information about the tasks can be consulted in the Supplementary Materials section of the current manuscript and in the technical manual [34].

2.3. Statistical analyses

Initial descriptive analyses were performed. Raw scores with wider ranges instead of scaled scores were used because they are more sensitive to change. Pearson correlations between V1 and V2 scores and Cohen's *d* effect size indices were computed. Simple discrepancy scores between visits were calculated by subtracting V1 score from V2 score. Higher discrepancy scores (i.e., positive) result from improvement in V2, whereas lower ones (i.e., negative) emerge from decline in performance at V2. V2/V1 ratios were also computed. Paired *t* tests were used to assess whether change between visits was significant.

As an initial approach to analyze the effect of individuals risk for AD in discrepancy scores group means were compared depending on *APOE* ε4 status (group 0, no *APOE* ε4 alleles and group 1, at least 1 *APOE* ε4 allele) and FH (0 for absent and 1 for at least having one parent that developed AD before 75 years). Because discrepancy scores were approximately normal (absolute values of skewness and kurtosis <1, except for Coding, which showed a slight leptokurtic distribution, kurtosis = 1.37), parametric tests were used. *APOE* ε4 and FH groups were compared by means *t* tests for independent samples. Subjects were also grouped depending on the number of *APOE* ε4 alleles (0, 1, 2), and discrepancy scores were compared among groups using a one-way ANOVA and post hoc tests (Tukey). The associations between raw discrepancy scores and age, education, intervisit interval, and baseline score were initially analyzed using simple regression analyses. Univariate general linear models were constructed to further explore the influence of variables related to AD risk, namely *APOE* ε4 status and FH of AD in the discrepancy of scores between visits. In these models, *APOE* ε4 status and FH were entered as random factors and intervisit interval, age, education, sex and baseline score vascular risk scores and discrepancy in GADS as covariates. Simple discrepancy scores in digit span total, coding, matrix reasoning, and visual puzzles were considered as dependent variables. Main effects and interactions of interest (i.e., *APOE* ε4*FH, *APOE* ε4*Age, and FH*Age) were studied. Subsequent models keeping covariates, main factors of interest (*APOE* ε4 status and FH), and significant interactions under $P = .01$ were finally fitted. Sequential type I sum of squares were used to conform the principle of marginality, in which higher order terms are entered after all corresponding lower order terms. Models were constructed with covariates first, *APOE* ε4 status and FH (in this order) in second term and, finally, two-way interaction terms. In final models, significant main effects were only interpreted when no significant interaction ($P > .05$) was present. As previously stated, with the exception of the selection step of relevant interactions ($P > .01$), significance threshold was set at $P = .05$. R statistical software (v.3.2.0) was used to conduct the analyses and to plot graphs.

3. Results

The study sample was composed of slightly more women (60%) than men and included a noticeable percentage of *APOE* ε4 carriers (36.3%). The percentage of carriers was higher in the FH+ group (39.6% vs. 31.7%), but this difference did not reach statistical significance (χ^2 [1,N = 388] = 2.25, $P = .13$). The mean years of formal education were 14. Retest interval ranged from 15 to 78 days with a mean interval of 46 days (SD = 10.35). Only 2.5% of the individuals were retested in <4 weeks and 14.8% in >8 weeks. A more detailed description of the sample, including scores in cognitive screening tests and cardiovascular risk scores, is found in Table 1. Table 2 shows these demographic and scale characteristics of the study participants according to their *APOE* and family history status.

Table 3 shows the raw scores at both visits and the discrepancy indices. As a whole, the group showed statistically significant improvements at V2 in all WAIS-IV subtests. The mean increase was of 6.5% (mean of the

Table 1
Sociodemographic, basic cognition, vascular-related risk scores, and genetic description of the study sample

	Mean	SD
Age (y)	53.43	5.39
Education (y)	13.96	3.41
MMSE	29.06	1.06
MIS	7.79	0.52
Semantic fluency (animals)	23.28	5.22
Interval between visits (d)	46.00	10.35
REGICOR (n = 359)*	3.68	1.97
CAIDE model I (n = 360)†	2.33	3.08
CAIDE model II (n = 353)‡	2.34	3.66
	Count/total	Percentage
Females	240/400	60.0
At least one *APOE* ε4 allele	142/391	36.3
Presence of family history of AD	357/400	89.3
AD onset <75 (FH+)	231/354	65.3
AD onset ≥75	123/354	34.7
No history of AD	43/400	10.7
No history of AD or AD onset ≥75 (FH−)	166/397	41.8

Abbreviations: MMSE, mini mental state examination; MIS, memory impairment screen; REGICOR, *Registre Gironí del Cor* function; CAIDE, Cardiovascular Risk Factors, Aging, and Dementia function; SD, standard deviation.

*Risk of suffering coronary disease events at 10 years in percentage.

†Model I, probability in percentage, of dementia in 20 years without *APOE* status.

‡Model II, probability in percentage of dementia in 20 years with *APOE* status.

percentage of improvement derived from the change ratios in each of the four main variables). Presence of anxiety and depression symptoms in the sample, as assessed by the GADS, showed minimum, nonsignificant, mean change between visits (V2-V1 GADS anxiety, M = −0.10, SD = 1.45, t(399) = 1.45, P = .149; V2-V1 GADS depression, M = −0.03, SD = 0.68, t(399) = 0.880, P = .380).

3.1. Means discrepancy and AD risk profile

No mean differences in the discrepancy indices in cognitive scores (V2–V1) were observed between groups when individuals were classified based on their risk for AD profile, neither between *APOE* ε4 carriers and noncarriers nor between individuals with and without FH. When coding participants' *APOE* status based on the number of *APOE* ε4 alleles (0, 1 or 2), differences remained nonsignificant.

3.2. Sociodemographic, baseline performance, and intervisit interval effects on PE

Univariate simple linear regression analyses were applied to explore the relationship between discrepancy scores and age, education, intervisit interval, and baseline performance. Fig. 1 shows the scatterplots and the results of simple univariate regression analyses between discrepancy scores and age, education, intervisit interval, and baseline score. In these models, age did not predict any discrepancy score, whereas education only significantly predicted the discrepancy in visual puzzles scores (the more the education the higher the PE). Intervisit interval predicted discrepancy in coding and matrix reasoning (in both cases, shorter intervals were associated to higher PE). Baseline scores were the strongest predictors of discrepancy scores, showing an inverse relationship in all the cases.

3.3. Effect of AD risk variables on PE

All *APOE* ε4 and FH main effects resulted nonsignificant. However, the interaction term *APOE* ε4*Age was found to be statistically significant for the coding discrepancy score (F(1,379) = 5.20, b = −0.304, P = .023, η_p^2 = 0.014), and quasi-significant for the digit span one (F(1,379) = 3.64, b = −0.124, P = .057, η_p^2 = 0.010). In

Table 2
Sample characteristics according to *APOE* and family history status

	APOE ε4		Family history of AD onset <75	
	Carriers	Noncarriers	Positive	Negative
N	142	249	231	166
Age, y, mean (SD)	52.84 (5.26)	53.84 (5.43)	52.59 (5.20)	54.62 (5.49)*
Education, y, mean (SD)	14.01 (3.37)	13.95 (3.42)	14.06 (3.30)	13.81 (3.57)
% females	54.93	61.85	56.71	64.46
MMSE, mean (SD)	29.11 (1.07)	29.04 (1.04)	29.12 (0.98)	28.96 (1.15)
MIS, mean (SD)	7.80 (0.48)	7.79 (0.51)	7.80 (0.51)	7.78 (0.50)
Semantic fluency (animals), mean (SD)	24.15 (5.09)	22.80 (5.22)*	23.10 (5.03)	23.52 (5.49)
Interval between visits (d), mean (SD)	45.76 (10.42)	46.31 (10.46)	46.74 (10.74)	45.10 (9.78)
REGICOR (n = 359)† %, mean (SD)	3.75 (1.95)	3.64 (2.01)	3.52 (1.89)	3.88 (2.07)
CAIDE model I (n = 360)‡ %, mean (SD)	2.44 (3.35)	2.30 (2.96)	2.16 (2.72)	2.59 (3.55)
CAIDE model II (n = 353)§ %, mean (SD)	3.42 (4.92)	1.72 (2.49)*	2.13 (2.90)	2.66 (4.56)

Abbreviations: MMSE, mini mental state examination; MIS, memory impairment screen; REGICOR, Registre Gironí del Cor function; CAIDE, cardiovascular risk factors, aging, and dementia function; SD, standard deviation.

*P < .05 in independent samples t test.

†Risk of suffering coronary disease events at 10 years.

‡Model I, probability of dementia in 20 years without *APOE* status.

§Model II, probability of dementia in 20 years with *APOE* status.

Table 3
V1 and V2 scores, discrepancy indices, correlations, and effect sizes

	V1	V2	V2-V1	Ratio V2/V1	*r*	*d*
Digit span total	25.73 (5.35)	26.60 (5.35)*	0.87 (3.49)	1.046 (0.153)	0.788*	0.16
Digit span forward	8.79 (2.18)	9.01 (2.27)*	0.22 (1.95)	1.052 (0.248)	0.615*	0.10
Digit span backward	8.30 (2.16)	8.59 (2.17)*	0.29 (1.84)	1.064 (0.238)	0.639*	0.13
Digit span sequencing	8.65 (2.11)	9.01 (2.27)*	0.36 (1.95)	1.081 (0.358)	0.589*	0.16
Coding	69.86 (13.71)	74.27 (14.36)*	4.46 (6.81)	1.070 (0.112)	0.883*	0.31
Matrix reasoning	17.51 (4.18)	18.11 (4.14)*	0.60 (3.52)	1.066 (0.262)	0.643*	0.14
Visual puzzles	14.64 (4.23)	15.25 (4.21)*	0.61 (3.30)	1.078 (0.255)	0.694*	0.15

NOTE. Means and (standard deviations) are shown. *$P < .05$. Paired *t* tests were applied to test the difference in means between V1 and V2. *r*, Pearson correlation. *d*, Cohen's effect size. For digit span, variables direct scores are displayed.

addition, the interaction term FH*Age significantly predicted matrix reasoning change ($F(1,379) = 5.28$, $b = 0.134$, $P = .022$, $\eta_p^2 = 0.014$). Simple slopes stratifying by *APOE* ε4 status or FH as appropriate were calculated and can be seen in Fig. 2. *APOE* ε4 carriers displayed a decline in PE with advancing age, whereas noncarriers did not show such pattern. With regard to FH, individuals categorized as negative showed an attenuation of PE with advancing age, whereas participants that had at least one of their parents that had suffered AD before the age of 75 years did not have such attenuation. To further explore the possible causes of such unexpected pattern, means in baseline score, interval between visits, age, and education were compared between FH+ and FH− groups. Significant mean differences were found for age, being the FH+ group younger than the FH− (FH+ [$M = 52.6$, $SD = 5.2$], FH− [$M = 54.6$, $SD = 5.5$], $t(395) = 3.75$, $P < .001$), and mean baseline scores, being higher in the FH+ group (FH+ [$M = 18.0$, $SD = 3.8$], FH− [$M = 16.9$, $SD = 4.6$], $t(395) = -2.7$, $P = .007$). As previously stated, no statistically significant differences between groups in the number of *APOE* ε4 carriers was found. Next, we also analyzed whether the same results were obtained when FH was introduced before *APOE* ε4 status in the model. This did not produce relevant changes in the output.

Finally, four additional models were constructed adding interaction terms between scores at V1 and age, education, *APOE* ε4 status, and FH. The V1*Age term on coding discrepancy scores was the only significant interaction ($F(1,376) = 6.75$, $b = -0.012$, $P = .01$, $\eta_p^2 = .018$). Vascular risk scores (REGICOR and CAIDE I and II) did not significantly predict discrepancy scores in any model, but we found significant moderate negative Pearson correlations between them and performance in Coding at V1 (REGICOR r = −0.25, CAIDE model I r = −0.33, CAIDE model II r = −0.35) and low ones (r value around −0.10) with digit span, matrix reasoning, and visual puzzles.

4. Discussion

This study aimed at exploring the effects of sociodemographic variables, baseline performance, *APOE* ε4 status, and FH of AD in the magnitude of PE in nonmemory measures in a sample of cognitively normal middle-aged adults. Globally, a mean improvement of 6.5% in performance was observed at V2. We calculated discrepancy scores (V2–V1) and studied the effect of subjects' characteristics. Age did not predict PE in any variable, whereas education predicted PE in visuospatial reasoning. Baseline score was in all cases the strongest predictor of change showing an inverse relationship to the magnitude of PE. AD risk factors, namely *APOE* ε4 and FH, did not show significant main effects in PE. However, we observed that *APOE* ε4 carriers displayed less improvement at V2 in processing speed and working memory when associated with increasing age. In addition, FH− subjects showed attenuated PE also associated with increasing age.

We have observed consistent PE for most variables. All of them yielded significance in the paired *t* test for mean differences. The higher effect size between visits was found for coding ($d = .31$) and the lower for digit span forward ($d = .10$). These findings are in agreement with the suggestion that timed tests that require an infrequently practiced response, such as substitution tasks, tend to be more susceptible to PE [35]. Although the comparison with previous reports is not straightforward, because studies differ in subjects' characteristics, test-retest intervals, and types of used tasks, the magnitude of the PE found in this study mirrors prior evidence, either using the same WAIS subtests [34–36] or other comparable neuropsychological measures [15,17].

Sociodemographic variables were overall poor predictors of gains or losses at retest. Age did not significantly predict any discrepancy score in univariate regressions, whereas education showed an effect only in visual puzzles discrepancy score. This globally negative findings match those previous studies that reported no relevant influences of age and education on PE [13–15,37]. However, it is worth mention that our results could be substantially driven by the narrow age range and the homogeneous middle-upper education of the studied sample.

Baseline performance exerts a strong effect in discrepancy outcomes. Previous literature suggests that when it comes to PE, the "richer get richer" rule applies [37], under the assumption that high performers would take more advantage from previous exposition to the task. In contrast, we found that participants with lower baseline scores tend to

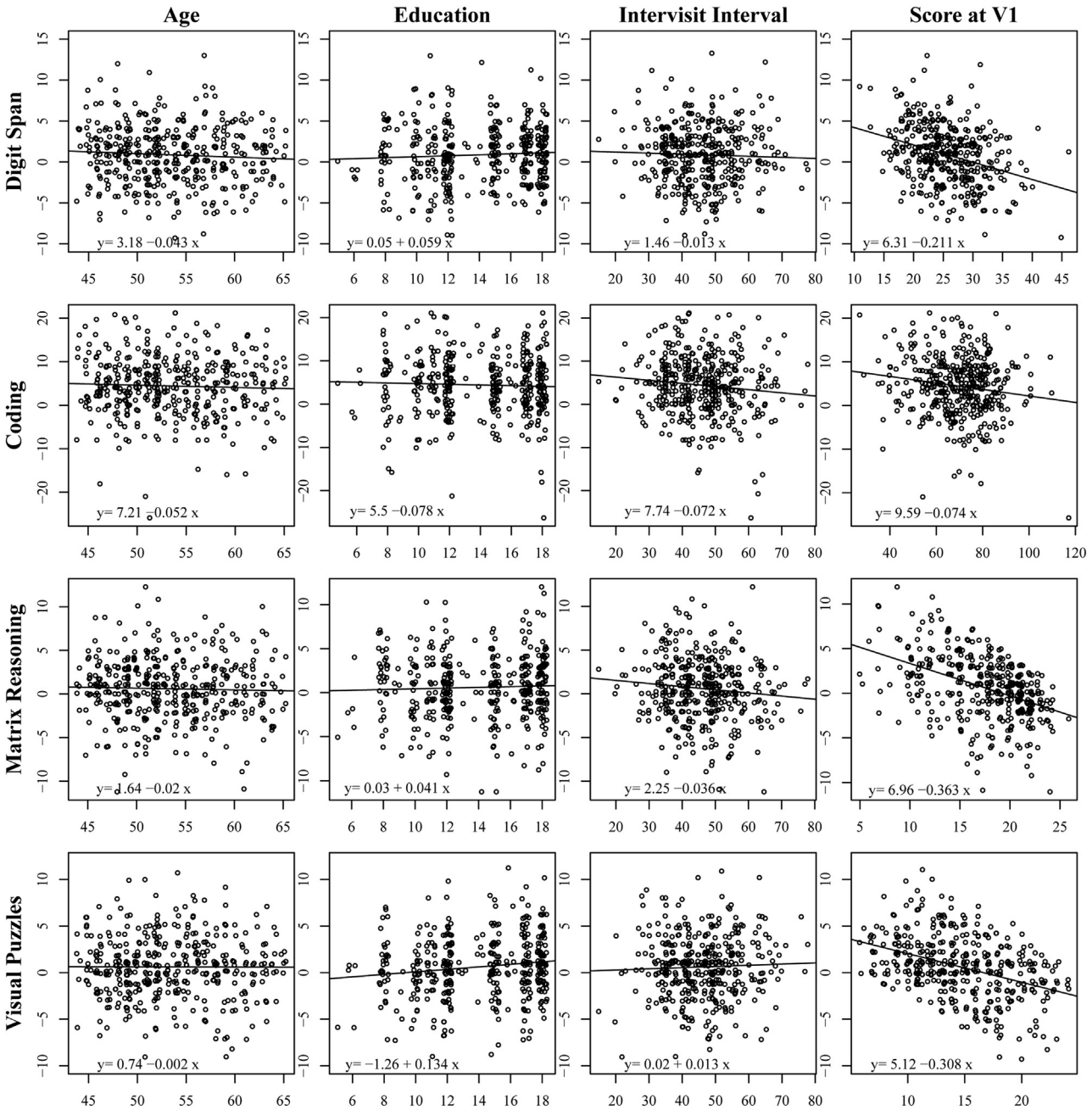

Fig. 1. Scatterplots and results of simple univariate regression between discrepancy scores and age, education, intervisit interval, and baseline scores. Significant coefficients: Education for visual puzzles discrepancy score [$b = 0.134$, $t(398) = 2.79$, $P = .006$, $R^2 = 0.019$]. Intervisit interval for coding discrepancy score [$b = -0.072$, $t(398) = -2.21$, $P = .028$, $R^2 = 0.012$] and matrix reasoning [$b = -0.036$, $t(398) = -2.11$, $P = .035$, $R^2 = 0.011$]. Baseline for discrepancy scores in digit span [$b = -0.211$, $t(398) = -6.84$, $P < .001$, $R^2 = 0.105$]; coding [$b = -0.074$, $t(398) = -3.01$, $P < .001$, $R^2 = 0.022$]; matrix reasoning [$b = -0.363$, $t(398) = -9.54$, $P < .001$, $R^2 = 0.186$]; and visual puzzles [$b = -0.308$, $t(398) = -8.57$, $P < .001$, $R^2 = 0.156$].

improve more than high performers at V2, in agreement with some other reports [38]. One plausible explanation for this is related to the "regression to the mean" phenomenon, in which extreme values at baseline tend to lie closer to the mean at retest. In addition, the special features of the studied measures can also contribute to this effect. WAIS scales are performance-based tests designed to obtain the maximum performance of a given subject in a single application. Therefore, such maximum-performance approach would minimize the capability of improvement at retest. In any case, our results confirmed that PEs exert a nondismissible influence in performance that should be accounted for and consciously addressed for a proper interpretation of cognitive outcomes. This issue is fundamental for clinical trials in which, in addition to the use of a control group, additional recommendations for controlling for PE have been

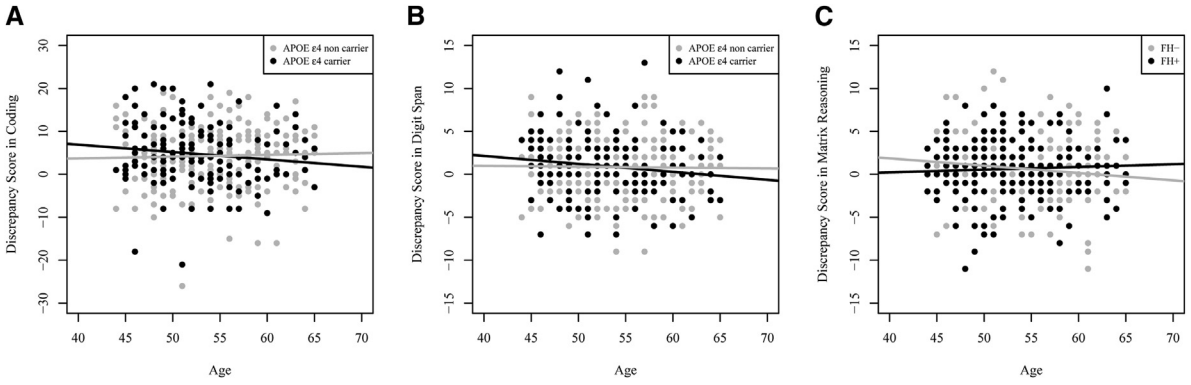

Fig. 2. Simple slopes showing the relationship between PE and age stratifying by *APOE* ε4 [(A) for the discrepancy score [V2 minus V1] in coding and (B) for digit span] or by FH [(C) for the discrepancy matrix reasoning].

specifically proposed for their application in preclinical AD studies. These include multiple prebaseline testing, the use of reliable change indices and alternate forms, and practice-insensitive tests [39].

4.1. Risk factors for AD

No differences in PE were observed between AD risk groups, neither when grouping by the presence or the absence of the *APOE* ε4 allele, nor by having or not a relevant FH antecedent as defined in our study. Our results partially concur with Jonaitis et al. findings in the context of the WRAP study [26]. They did not observe significant *APOE* ε4 effects on PE but found an interaction effect between FH and the number of previous assessments in the factors speed and flexibility and working memory. Interestingly, we observed significant effects on tasks tapping these same domains for the interaction term between *APOE* ε4 and age in PE (i.e., coding and digit span), with carriers showing less improvement at V2 as they age. Despite obvious differences between both studies and results, the convergence of domains in which diminished PE related to both risk factors gives additional credit to the importance of studying PE in tasks tapping processing speed and working memory for the detection of initial signs of subtle cognitive decline.

In addition, our results also partially reproduce a counterintuitive pattern observed by Jonaitis et al. They found a significant effect of the interaction between FH and age consisting in a worse age-related change in speed and flexibility in subjects without antecedents of AD [26]. We found a similar pattern for the PE in matrix reasoning, in which FH− subjects displayed attenuated PE as they age. Jonaitis et al. discussed their results in the context of controlling the effect of age as a linear covariate and suggested that this effect might sometimes show a nonlinear (quadratic) behavior. Should the effects of aging be nonlinear, accounting for this factor with a linear term would not completely remove this component, and as consequence, this may cause some unexpected behavior in the dependent variable. Another possible explanation of such unexpected result could be a high degree of collinearity between the FH and *APOE* covariates. In this event, if the two variables are strongly correlated, it might be impossible to dissociate their distinct impact on the dependent variable. To evaluate these two alternatives, we further explored the reliability of the interaction with two additional statistical models. When age was entered as a quadratic factor, the direction of the interaction was reversed (FH+ showed less PE when age). On the other hand, the interaction remained unchanged when we removed *APOE* status as a factor, therefore discarding collinearity as a potential explanation. These results suggest that the interaction effect between age and FH on cognition may arise from a nonlinear effect of aging. Conversely, no changes in the direction of the effects of the interaction term *APOE* ε4*Age were observed. Taken together, these results highlight the importance of appropriately covariating the effect of aging on the measurements of cognitive outcomes. Previous evidence from neuroimaging studies has shown that healthy aging is a nonlinear process and therefore is better modeled using a quadratic approach (e.g., [40–43]). In addition, cognitive aging has also shown to follow a quadratic pattern (e.g., [44]).

Our definition of FH that is dependent on parental age of onset of AD could have also had an impact in our results. Previous studies have limited the age of onset of the AD antecedent to <85 [45] or <80 [46] because "dementia occurring at a very old age is less likely to have a strong genetic component" [45]. In our study, a cutoff of 75 years was selected based on the evidence from recent studies on age-dependent genetic risk that propose that genetic load exerts a higher influence before this age (see [32,33]). In addition, our recruitment strategy for the wider parent cohort (mainly first-degree descendants of AD patients) may have also introduced a potential bias in the present study.

Our results support the idea that working memory may be a key domain in the detection of subtle cognitive decline in preclinical AD. Recent studies have found a greater decline only in the attention/working memory domain in subjects without cognitive impairment that had evidence of neuropathologic AD change [47]. In addition, working

memory showed the highest heritability among siblings with parental history of AD, with *APOE* ε4 status explaining not much of this heritability, and it is suggested to be a cognitive target for future genetic studies [48]. Because PEs are also suggested to be useful as a subtle cognitive marker sensitive to preclinical AD, a more thorough evaluation of PE in working memory could be of high interest. Within this preclinical AD phase, characterized by pathophysiological changes occurring decades before the emergence of clinical symptoms, a late stage that encompasses "subtle cognitive impairment" has been proposed (stage III) [11]. Although the concept of "subtle cognitive impairment" is still not fully operationalized, some works that have recently addressed the topic, have consistently proposed the use of intra-subject measures, either by studying cognitive decline during time or the magnitude of PE [10,26]. In addition, the latter has been demonstrated to be an inexpensive, useful tool to identify subjects for the enrichment of clinical trial samples aiming at recruiting amyloid-positive participants [49].

The main limitation of the present study is its transversal nature. The longitudinal follow-up of participants, currently underway, will allow us to assess for the ability of PE to predict cognitive decline. Furthermore, these studies will help ascertain whether PE in the cognitive domains under study here may be used as cognitive markers for preclinical AD.

In addition, in this study, we measured PE with four WAIS-IV subtests, whereas episodic memory, which is the domain that showed the most consistent PE, is not included in the current analyses. A preliminary study in a much smaller sample from the ALFA cohort in which we analyzed PE on the memory binding test showed a 10% to 26% of improvement in V2 [50]. Nevertheless, the study of practice effects (PE) in nonmemory domains is still highly relevant in subjects at risk of AD. Machulda et al. [9] found that subjects that remained cognitively stable had PE in the memory, language, attention, and visual reasoning domains at retest. In contrast, subjects who declined (either to incident MCI or dementia in the following 6 years) showed diminished PE in the language, attention, and visual reasoning domains but not in memory. Similarly, Wilson et al. [51] pointed out that a change in semantic and working memory preceded the decline in other domains.

The narrow age range of participants may also be seen *a priori* as a limitation. Nonetheless, we found relevant AD risk and age interaction effects in PE, and these results point out to a possible strong interaction between *APOE* ε4 and FH status and age that will be confirmed in longitudinal studies. Finally, as noted by Jonaitis et al. [26], although studies on healthy middle-aged adults do not have the development of MCI or dementia as endpoints, they are useful to seek for antecedent cognitive phenotypes stratifying the samples by known markers of dementia risk, such as *APOE* ε4 or FH status.

In summary, our main result is that PEs are partially affected by individual characteristics and baseline performance in healthy middle-aged subjects. AD risk factors, namely *APOE* ε4 and FH, interactions with age show an effect in the magnitude of PE in some tasks. In line with previous reports, our results suggest that PE in the processing speed and working memory domains could be of special interest for the development of cognitive markers of preclinical AD. Our data add evidence on the fact that PEs exert a relevant effect in cognitive outcomes at retest, and accordingly, they must be taken into consideration in clinical trials for preclinical AD (for example using multiple prebaseline testing or alternate forms of the same test).

Acknowledgments

The research leading to these results has received funding from "la Caixa" Foundation. Additional funding was obtained from Fondo de Investigación Sanitaria (FIS), Instituto de Salud Carlos III (ISC-III) under grant PI12/00326. J.D.G. holds a 'Ramón y Cajal' fellowship (RYC-2013-13054). We thank the raters and nurse for data collection. We thank Carolina Minguillón for text editing and continuous support and Andreia Carvalho for her logistic support.

Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.dadm.2016.07.001.

RESEARCH IN CONTEXT

1. Systematic review: Previous reports on the relationship of individuals' characteristics on practice effects (PE) are cited throughout the manuscript.
2. Interpretation: The relationship between sociodemographic factors, Alzheimer's disease (AD) family history and *APOE* ε4 status, and the magnitude of PE in 400 participants (45–65 years), most of them first-degree descendants of AD patients, was assessed for four subtests of the WAIS-IV. PEs were observed in all measures. Sociodemographic variables did not show a uniform effect. Baseline score was the strongest predictor of change, being inversely related to PE magnitude. Significant effects of the interaction term *APOE* ε4-age in processing speed and working memory were observed.
3. Future directions: PEs must be considered in clinical trials in preclinical AD (for example with multiple prebaseline testing or by suing alternate forms). PE in processing speed and working memory could be of special interest for the development of cognitive markers of preclinical AD.

References

[1] Koziol LF, Budding DE. Subcortical Structures and Cognition. New York: Springer; 2009.

[2] McCaffrey R, Duff K, Westervelt HJ. Practitioner's Guide to Evaluating Change with Neuropsychological Assessment Instruments. New York: Springer Science & Business Media; 2000.

[3] Darby D, Maruff P, Collie A, Mcstephen M. Mild cognitive impairment can be detected by multiple assessments in a single day. Neurology 2002;59:1042–6.

[4] Cooper DB, Epker M, Lacritz L, Weine M, Rosenberg RN, Honig L, et al. Effects of practice on category fluency in Alzheimer's disease. Clin Neuropsychol 2001;15:125–8.

[5] Cooper DB, Lacritz LH, Weiner MF, Rosenberg RN, Cullum CM. Category fluency in mild cognitive impairment: reduced effect of practice in test-retest conditions. Alzheimer Dis Assoc Disord 2004; 18:120–2.

[6] Duff K, Beglinger LJ, Schultz SK, Moser DJ, McCaffrey RJ, Haase RF, et al. Practice effects in the prediction of long-term cognitive outcome in three patient samples: A novel prognostic index. Arch Clin Neuropsychol 2007;22:15–24.

[7] Duff K, Beglinger LJ, Van Der Heiden S, Moser DJ, Arndt S, Schultz SK, et al. Short-term practice effects in amnestic mild cognitive impairment: implications for diagnosis and treatment. Int Psychogeriatr 2008;20:986–99.

[8] Duff K, Lyketsos CG, Beglinger LJ, Chelune G, Moser DJ, Arndt S, et al. Practice effects predict cognitive outcome in amnestic mild cognitive impairment. Am J Geriatr Psychiatry 2011;19:932–9.

[9] Machulda MM, Pankratz VS, Christianson TJ, Ivnik RJ, Mielke MM, Roberts RO, et al. Practice effects and longitudinal cognitive change in normal aging vs. incident mild cognitive impairment and dementia in the Mayo Clinic Study of Aging. Clin Neuropsychol 2013;27:1247–64.

[10] Hassenstab J, Ruvolo D, Jasielec M, Xiong C, Grant E, Morris C. Absence of Practice Effects in Preclinical Alzheimer's Disease. Neuropsychology 2015;29:940–8.

[11] 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.

[12] Calamia M, Markon K, Tranel D. Scoring higher the second time around: meta-analyses of practice effects in neuropsychological assessment. Clin Neuropsychol 2012;26:543–70.

[13] Duff K, Callister C, Dennett K, Tometich D. Practice effects: a unique cognitive variable. Clin Neuropsychol 2012;26:1117–27.

[14] Gross AL, Benitez A, Shih R, Bangen KJ, Glymour MMM, Sachs B, et al. Predictors of Retest Effects in a Longitudinal Study of Cognitive Aging in a Diverse Community-Based Sample. J Int Neuropsychol Soc 2015;21:506–18.

[15] Bartels C, Wegrzyn M, Wiedl A, Ackermann V, Ehrenreich H. Practice effects in healthy adults: a longitudinal study on frequent repetitive cognitive testing. BMC Neurosci 2010;11:118.

[16] Stuss DT, Stethem LL, Poirier CA. Comparison of three tests of attention and rapid information processing across six age groups. Clin Neuropsychol 1987;1:139–52.

[17] Van der Elst W, Van Boxtel MPJ, Van Breukelen GJP, Jolles J. Detecting the significance of changes in performance on the Stroop Color-Word Test, Rey's Verbal Learning Test, and the Letter Digit Substitution Test: the regression-based change approach. J Int Neuropsychol Soc 2008;14:71–80.

[18] Duff K. Evidence-based indicators of neuropsychological change in the individual patient: relevant concepts and methods. Arch Clin Neuropsychol 2012;27:248–61.

[19] Salthouse TA, Schroeder DH, Ferrer E. Estimating Retest Effects in Longitudinal Assessments of Cognitive Functioning in Adults Between 18 and 60 Years of Age. Dev Psychol 2004;40:813–22.

[20] Wisdom NM, Callahan JL, Hawkins KA. The effects of apolipoprotein E on non-impaired cognitive functioning: A meta-analysis. Neurobiol Aging 2011;32:63–74.

[21] Verghese PB, Castellano JM, Holtzman DM. Apolipoprotein E in Alzheimer's disease and other neurological disorders. Lancet Neurol 2011;10:241–52.

[22] Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet 2013; 45:1452–8.

[23] Donix M, Small GW, Bookheimer SY. Family history and APOE-4 genetic risk in Alzheimer's disease. Neuropsychol Rev 2012; 22:298–309.

[24] Hayden KM, Zandi PP, West NA, Tschanz JT, Maria C, Corcoran C, et al. Effects of Family History and APOE ε4 Status on Cognitive Decline in the Absence of AD: The Cache County Study. Arch Neurol 2009;66:1378–83.

[25] Loewenstein DA, Acevedo A, Ownby R, Agron J, Barker WW, Isaacson R, et al. Using different memory cutoffs to assess mild cognitive impairment. Am J Geriatr Psychiatry 2006;14:911–9.

[26] Jonaitis EM, Koscik RL, La Rue A, Johnson SC, Hermann BP, Sager MA. Aging, Practice Effects, and Genetic Risk in the Wisconsin Registry for Alzheimer's Prevention. Clin Neuropsychol 2015; 29:426–41.

[27] Zehnder AE, Bläsi S, Berres M, Monsch AU, Stähelin HB, Spiegel R. Impact of APOE status on cognitive maintenance in healthy elderly persons. Int J Geriatr Psychiatry 2009;24:132–41.

[28] Donix M, Ercoli LM, Siddarth P, Brown JA, Martin-Harris L, Burggren AC, et al. Influence of Alzheimer disease family history and genetic risk on cognitive performance in healthy middle-aged and older people. Am J Geriatr Psychiatry 2012;20:565–73.

[29] Molinuevo JL, Gramunt N, Gispert JD, Fauria K, Esteller M, Minguillon C, et al. The ALFA project: A research platform to identify early pathophysiological features of Alzheimer's disease. Alzheimer's & Dementia: Translational Research & Clinical Interventions 2016; 2:82–92.

[30] Marrugat J, Vila J, Baena-Díez JM, Grau M, Sala J, Ramos R, et al. Relative Validity of the 10-Year Cardiovascular Risk Estimate in a Population Cohort of the REGICOR Study. Rev Esp Cardiol 2011; 64:385–94.

[31] Kivipelto M, Ngandu T, Laatikainen T, Winblad B, Soininen H, Tuomilehto J. Risk score for the prediction of dementia risk in 20 years among middle aged people: a longitudinal, population-based study. Lancet Neurol 2006;5:735–41.

[32] Van Cauwenberghe C, Van Broeckhoven C, Sleegers K. The genetic landscape of Alzheimer disease: clinical implications and perspectives. Genet Med 2016;18:421–30.

[33] Sleegers K, Bettens K, De Roeck A, Van Cauwenberghe C, Cuyvers E, Verheijen J, et al. A 22-single nucleotide polymorphism Alzheimer risk score correlates with family history, onset age, and cerebrospinal fluid Aβ42. Alzheimers Dement 2015;11:1452–60.

[34] Wechsler D. Escala de Inteligencia Wechsler para adultos IV (Spanish version). Madrid: Pearson; 2012.

[35] Estevis E, Basso MR, Combs D. Effects of practice on the Wechsler Adult Intelligence Scale-IV across 3- and 6-month intervals. Clin Neuropsychol 2012;26:239–54.

[36] Basso MR, Carona FD, Lowery N, Axelrod BN. Practice effects on the WAIS-III across 3- and 6-month intervals. Clin Neuropsychol 2002; 16:57–63.

[37] Rapport LJ, Brines DB, Theisen ME, Axelrod BN. Full scale IQ as mediator of practice effects: The rich get richer. Clin Neuropsychol 1997;11:375–80.

[38] Duff K, Beglinger LJ, Schoenberg MR, Patton DE, Mold J, Scott JG, et al. Test-retest stability and practice effects of the RBANS in a community dwelling elderly sample. J Clin Exp Neuropsychol 2005; 27:565–75.

[39] 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. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring 2015;1:103–11.

[40] Kennedy KM, Erickson KI, Rodrigue KM, Voss MW, Colcombe SJ, Kramer AF, et al. Age-related differences in regional brain volumes: A comparison of optimized voxel-based morphometry to manual volumetry. Neurobiol Aging 2009;30:1657–76.

[41] Jernigan TL, Gamst AC. Changes in volume with age - Consistency and interpretation of observed effects. Neurobiol Aging 2005;26:1271–4.

[42] Allen JS, Bruss J, Brown CK, Damasio H. Normal neuroanatomical variation due to age: The major lobes and a parcellation of the temporal region. Neurobiol Aging 2005;26:1245–60.

[43] Good CD, Johnsrude IS, Ashburner J, Henson RN, Friston KJ, Frackowiak RS. A voxel-based morphometric study of ageing in 465 normal adult human brains. Neuroimage 2001;14:21–36.

[44] Hedden T, Gabrieli JDE. Insights into the ageing mind: a view from cognitive neuroscience. Nat Rev Neurosci 2004;5:87–96.

[45] Debette S, Wolf PA, Beiser A, Au R, Himali JJ, Pikula A, et al. Association of parental dementia with cognitive and brain MRI measures in middle-aged adults. Neurology 2009;73:2071–8.

[46] Honea RA, Swerdlow RH, Vidoni ED, Goodwin J, Burns JM. Reduced gray matter volume in normal adults with a maternal family history of Alzheimer disease. Neurology 2010;74:113–20.

[47] Monsell SE, Mock C, Hassenstab J, Roe CM, Cairns NJ, Morris JC, et al. Neuropsychological changes in asymptomatic persons with Alzheimer disease neuropathology. Neurology 2014;83:434–40.

[48] Darst BF, Koscik RL, Hermann BP, La Rue A, Sager MA, Johnson SC, et al. Heritability of cognitive traits among siblings with a parental history of Alzheimer's disease. J Alzheimers Dis 2015;45:1149–55.

[49] Duff K, Foster NL, Hoffman JM. Practice effects and amyloid deposition: preliminary data on a method for enriching samples in clinical trials. Alzheimer Dis Assoc Disord 2014;28:247–52.

[50] Gramunt N, Sánchez-Benavides G, Buschke H, Lipton RB, Masramon X, Gispert JD, et al. Psychometric Properties of the Memory Binding Test: Test-Retest Reliability and Convergent Validity. J Alzheimers Dis 2016;50:999–1010.

[51] Wilson RS, Leurgans SE, Boyle PA, Bennett DA. Cognitive decline in prodromal Alzheimer disease and mild cognitive impairment. Arch Neurol 2011;68:351–6.