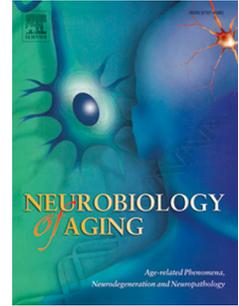# Accepted Manuscript

Cognition or genetics? Predicting Alzheimer's disease with practice effects, APOE genotype and brain metabolism

Javier Oltra-Cucarella, Miriam Sánchez-SanSegundo, Rosario Ferrer-Cascales

**Cognition or genetics? Predicting Alzheimer's disease with practice effects, APOE genotype and brain metabolism**

Javier Oltra-Cucarella, Miriam Sánchez-SanSegundo, Rosario Ferrer-Cascales, for the Alzheimer's Disease Neuroimaging Initiative*

Department of Health Psychology, University of Alicante (Spain). Campus de San Vicente del Raspeig s/n, 03690 San Vicente del Raspeig, Alicante, Spain

Text word count: 4,057

Abstract word count: 169

Running title: Practice effects predict AD in aMCI

Corresponding author: Miriam Sánchez-SanSegundo, Department of Health Psychology, University of Alicante (Spain). Campus de San Vicente del Raspeig s/n, 03690 San Vicente del Raspeig, Alicante, Spain, Phone number: +34 965.90.34.00 – Ext 2236,Email address: miriam.sanchez@ua.es

Alternate Corresponding author: Rosario Ferrer-Cascales, Department of Health Psychology, University of Alicante (Spain). Campus de San Vicente del Raspeig s/n, 03690 San Vicente del Raspeig, Alicante, Spain, Phone number: +34 965.90.94.20, Email address: rosario.ferrer@ua.es

**ABSTRACT**

As practice effects are common in neuropsychological assessment, this study analyzed their utility to identify individuals with amnestic Mild Cognitive Impairment (aMCI) at the greatest risk for Alzheimer's disease (AD-risk), and compared practice effects with APOE and brain metabolism biomarkers. We regressed Auditory Verbal Learning Test delayed recall (AVLT-DR) at six months on baseline AVLT-DR scores in 394 individuals with normal cognition (NC) from the ADNI database, and dichotomized 816 individuals with aMCI as showing (PE+) or not showing practice effects (PE-) when the discrepancy between observed and predicted scores was found in less than 10%, 7% and 5% of NC. Cox regressions analyzed the AD-risk at 6 years. More than 60% of aMCI were PE+. Controlling for age, sex, education, and baseline MMSE and AVLT-DR scores, the AD-risk was associated with PE- (HR=1.93), lower brain metabolism (HR=0.95) and APOE genotype (HR=1.92), with narrower risk estimates for PE-. The lack of practice effects during a six months period might be as precise as biomarkers for predicting the 6-year AD-risk.

## 1. INTRODUCTION

The annual risk estimates of progression from aMCI to AD (AD-risk) is around 8% in specialist clinical settings and 7% in community studies (Mitchell and Shiri-Feshki, 2009), which implies that a majority of individuals with aMCI will remain stable or even revert to normal. The clinical implication of this is that repeated cognitive testing is essential to identify those individuals with progressing cognitive decline, and hence at the greatest AD-risk (Albert et al., 2011).

Repeated cognitive testing can lead to incorrect conclusions if an individual's performance is compared against the same normative data on two occasions, because an increase in performance is expected for a number of cognitive tests due to the exposure to the same test in a previous occasion. This phenomenon, known as *practice effects* (Duff, 2012), has been documented in several populations including MCI (Calamia et al., 2012). Practice effects on memory tests have been reported in individuals with aMCI within the same session (Duff et al., 2012) and over periods of one week (Duff et al., 2017a), eighteen months (Campos-Magdaleno et al., 2017) and even 5 years (Gavett et al., 2016), and have proven useful to identify individuals with aMCI who will show larger cognitive decline after one year follow-up (Duff et al., 2011). However, some researchers have reported no practice effects in individuals with aMCI over different periods (Darby et al., 2002; Schrijnemaekers et al., 2006), so their utility remains controversial.

Practice effects have been associated with APOE ε4 genotype (Machulda et al., 2013; Zehnder et al., 2009) and with brain metabolism, which in turn have been associated with the AD-risk. It is known that APOE ε4 carriers, mostly those carrying two copies of

the allele, have an increased AD-risk compared to APOE ε4 non-carriers and carriers of

APOE ε2 and ε3 alleles (Elias-Sonnenschein et al., 2011; Qian et al., 2017; Yu et al.,

2014). In one study on practice effects (Machulda et al., 2013), APOE carriers failed to

sustain their initial practice effects over one year, with a level of performance similar

to baseline after approximately 6 years of follow-up. Regarding brain metabolism,

although data on the accuracy of FDG-PET are highly variable (Smailagic et al., 2015),

FDG-PET has been suggested as a more sensitive tool than cognitive scores for

predicting AD in aMCI (Herholz et al., 2011). FDG-PET has been associated with

practice effects on tests of visual and verbal memory, with more brain

hypometabolism being associated with worse cognitive performance and lower

practice effects (Duff et al., 2015, 2014). However, the associations between practice

effects and the AD-risk, and also the differential predictive value of AD for practice

effects, APOE genotype and brain metabolism was not analyzed in either of these

previous studies.

In the present study, we aimed to analyze whether practice effects in aMCI over two

successive assessments can help to identify individuals at the greatest AD-risk, and also

to compare the predictive value of practice effects with APOE genotype and brain

metabolism measured with FDG-PET. We expected that individuals with aMCI who did

not show practice effects would have the greatest AD-risk compared to those with

aMCI who did show practice effects (Duff et al., 2011; Hassenstab et al., 2015). Due to

the lack of previous comparisons among practice effects, APOE and FDG-PET, we could

not make a priori hypotheses about a superior predictive value for any of the variables

analyzed.

4

## 2. METHODS

### 2.1. Sample data

Data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu) were used in this study. The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of the ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD. For up-to-date information, see www.adni-info.org.

The Normal cognition group (NC) included 394 participants (48.2% females) aged 56 to 89 years with no depression or metabolic diseases, no cognitive complaints, a Clinical Dementia Rating scale (CDR) score = 0, Mini-Mental State Examination (MMSE) score equal or higher than 24, normal education-corrected Logical Memory (LM) delayed recall scores, and no significant impairments in activities of daily living. In the aMCI group (Petersen et al., 1999), 816 participants (40.8% females) aged 55 to 91 years with no metabolic diseases had subjective cognitive complaints, MMSE score ≥24, CDR score = 0.5 (mandatory memory box score ≥0.5), abnormal education-corrected LM delayed recall scores, general cognition and functional performance largely intact, and did not meet criteria for dementia. Five participants (0.6%) had mild depressive symptoms. All participants underwent physical and neurological examinations, screening laboratory tests, and provided blood samples for DNA and APOE testing. The ethical committee at each participating site approved the project, and all ADNI participants provided written consent before enrollment at each site.

*2.2. Procedure*

Data from 394 NC participants free of any type of dementia during a 6-year follow-up period (range: 6-72 months) were used to regress Auditory Verbal Learning Test (AVLT) delayed recall scores at six months on baseline RAVLT delayed recall scores. Then, baseline AVLT scores, the intercept, beta coefficient and standard error of the regression equation were used to predict six-months AVLT delayed recall scores in each aMCI participant. Predicted retest scores were subtracted from observed retest scores, and this discrepancy was divided by the standard error of a predicted score for a new case $[(T_{observed} - T_{predicted})/S_{n+1}]$ according to Crawford and Garthwaite (2007). The standardized discrepancy was compared against a t distribution with n-2 degrees of freedom (Crawford and Garthwaite, 2007), which is preferred over a normal distribution because it treats the sample used to build the regression equation as a sample and not as a population, and it has a lower rate of type I error compared to a z distribution (Crawford and Garthwaite, 2005). However, as the sample used to build the regression equation was large, p-values associated with discrepancies were similar to those obtained using a normal distribution (data not shown).

The p-value associated to the $t_{n-2}$ statistic can be interpreted as the percentage of individuals from the sample used to build the regression equation showing a similar or more extreme discrepancy (Crawford and Garthwaite, 2007), which could be interpreted as a percentile from a distribution of discrepancies. Based on statistical cut-off points used in the literature to define objective cognitive impairment, we used the bottom 10%, 7% and 5% of the NC group, which correspond to z-scores of approximately -1.28, -1.5 and -1.64 respectively for a one-sided test. Participants

6

showing a negative discrepancy at or below any of the cut-points were labeled as not showing practice effects (PE-), and as showing practice effects (PE+) if the discrepancy was above the cut-points or positive.

### 2.3. *FDG-PET measures*

For information about neuroimaging data acquisition see http://adni.loni.usc.edu/data-samples/pet/. Participants glucose levels were measured two hours after the last ingestion, and FDG-PET scans performed 30 minutes after intravenous administration of $[^{18}F]$-FDG if blood glucose level was <180 mg/dL (9.9 mmol/L). The variable FDG from the ADNIMERGE file was analyzed, which indicates the baseline average FDG uptake of angular, temporal and posterior cingulate gyri. Individuals with AD have lower values than MCI and NC, so higher values of FDG-PET indicate higher cerebral metabolism (Landau et al., 2011). To facilitate interpretation of the results from a regression model (see Section 2.5), FDG-PET values were multiplied by 100 for values to show the difference in the AD-risk for one unit increase in FDG-PET metabolism. Using this scale is also easier to interpret than using exponentiated values. The aMCI sample decreased to 627 participants due to missing FDG-PET values.

### 2.4. *Outcome*

We analyzed the difference in the AD-risk (McKhann et al., 2011, 1984) during a 6 year follow-up period.

## 2.5. Statistical analysis

Raw baseline and 6-months AVLT-DR scores were analyzed separately for NC and MCI groups with paired t-tests and Pearson's correlation coefficients. Demographic and continuous variables were compared between groups with ANOVAs, which provide eta-squared measures as effect size. Values of 0.01 are considered a small effect, values of 0.06 are considered a medium effect, and values of 0.14 are considered a large effect (Richardson, 2011). Sex ratio was compared with a chi-square test.

The AD-risk was compared with hazard ratios (HR) from a multivariable backward stepwise Cox proportional regression model. The first step included age, sex and education as demographic variables. The second step included PE+ and PE- groups, with baseline AVLT and MMSE scores as covariates. Including baseline AVLT scores allowed us controlling whether the AD-risk for PE- and PE+ groups was above and beyond baseline AVLT-DR scores, as in previous research (Duff et al., 2017b, 2015, 2011; Gavett et al., 2016; Hassenstab et al., 2015). The third step added APOE and FDG-PET biomarkers.

We performed a separate Cox regression model for each cut-off points used to define PE-. We did not to use a model with all the three cut-off points because of a high collinearity, with variance inflation factors being 5.20 to 10.92. Lastly, the risk of having at least one copy of the APOE ε4 allele was compared between PE+ and PE- groups using odds ratios (OR). All statistical analyses were performed using SPSS v.23, with alfa level set at 0.05.

2.6. *Analyses with missing data*

To check if our results would replicate in a sample without missing values, we used the

Missing Values Analysis (MVA) to highlight patterns of missing values and also to

replace them in the dataset (Tabachnick and Fidell, 2013). First, we analyzed whether

missingness in FDG-PET measures was related to each of the variables introduced in

the regressions analyses plus progression status using the Separate Variance t-test.

Whether the data were missed at random was analyzed with the Little's Missing

Completely at Random test (MCAR). We then imputed FDG-PET values using MVA

regression to estimating missing values. As estimates from regression can only be used

if estimated values fall within the range of values for complete cases (Tabachnick and

Fidell, 2013), we analyzed with an independent samples t-test the differences in the

distribution of scores between cases with complete values and cases with estimated

values. Lastly, we repeated the Cox proportional hazard regression for the 5% cut-off.

2.7. *Comparison of the risk of AD according to clinical profile*

Using the dataset with complete and estimated values, we categorized participants

into one of four groups according to the clinical profile: 1) participants showing no

practice effects, 2) participants having at least one APOE allele, 3) participants with

both conditions, and 4) participants showing practice effects and not having APOE

alleles. This latter group was used as the reference group for comparisons. The AD-risk

was compared among groups using a Cox regression with age, sex, education, and

MMSE and AVLT baseline scores.

## 3. RESULTS

NC and MCI groups differed on age, level of education, sex ratio, MMSE scores, baseline AVLT scores, 6-months AVLT scores and brain metabolism (Table 1). Differences were negligible for education, small for age and FDG-PET values, medium for baseline and 6-months AVLT scores, and large for MMSE scores. Compared to MCI, NC were slightly older, were more educated, had higher MMSE scores, higher baseline and 6-months AVLT-DR scores, more brain hypermetabolism, and more female participants. Participants with MCI had a statistically significantly higher probability of having at least one copy of the APOE ε4 allele. Individuals in the aMCI group were followed for an average period of 38.94 months (SD=21.59, range 6 to 72 months).

NC test-retest AVLT scores were similar (mean difference (MD) = 0.16, 95%CI = -0.09, 0.42, $t(393) = 1.28$, $p = .202$), with both scores being statistically significantly correlated ($r(394) = 0.44$, $p < .001$). Test scores predicted retest scores (Intercept=6.92, standard error = 2.18, $\beta$=0.45, $p < .001$). In the MCI group, retest scores were statistically significantly lower than test scores ($MD = 0.43$, 95%CI = 0.22, 0.64, $t(815) = 4.07$, $p < .001$), with both scores being statistically significantly correlated ($r(816) = 0.64$, $p < .001$). The percentage of participants obtaining a higher retest AVLT-DR score were 32.5% and 33.8% for NC and aMCI groups respectively ($\chi^2(1, N = 1,210) = 0.21$, $p = .644$).

### 3.1. aMCI groups based on practice effects

The number of PE- were smaller as the cut-off point used to define impairment was more restrictive, from 257 (31.5%) for a 10% cut-off point to 221 (27.1%) for a 7% cut-off point and 196 (24%) for a 5% cut-off point. Participants in the PE- group were more likely to have at least one copy of the APOE ε4 allele using a 10% cut-off (OR = 1.69, 95%CI = 1.26, 2.29, $p$ = .001), a 7% cut-off (OR = 1.86, 95%CI = 1.35, 2.55, $p$ < .001) and a 5% cut-off (OR = 1.95, 95%CI = 1.40, 2.72, $p$ < .001) than participants in the PE+ group. One hundred and seventy-five (21.4%) individuals in the aMCI sample progressed to AD. The percentage of progressors to AD in the PE-/PE+ groups was 35/15.2%, 37.6/15.5% and 39.3/15.8% for 10%, 7% and 5% cut-off points respectively.

### 3.2. Risk of progression to AD

Results from Cox regressions (table 2) indicated that only age was associated with the AD-risk among demographics, irrespective of the cut-off point. When cognitive variables were added to the model, lower baseline AVLT and MMSE scores, and PE- were associated with an increased AD-risk. When biomarkers were added to the model, more hypometabolism from FDG-PET and having at least one APOE e4 allele were associated with an increased AD-risk. The model including all the variables showed that having two APOE e4 alleles and not showing practice effects were the variables with the highest risk estimates, with PE- showing less heterogeneity and narrower confidence intervals and APOE becoming non-significant. As results were identical for the three cut-off points used to define the no practice effects groups,

table 2 shows results for the 5% cut-off point as commonly reported when using a z-score equal or lower than -1.645 (Duff, 2012) for a one-tailed test.

### 3.3. Analysis of missing FDG-PET values

Separate Variance t-test showed that FDG-PET missing values were associated with age ($p$ = .006), education ($p$ = .022), MMSE ($p$ < .001), baseline AVLT-DR scores ($p$ < .001), 6-months AVLT scores ($p$ < .001), follow-up ($p$ < .001), and progression to AD ($p$ < .001), but not with gender (p = .486) or APOE (p = .661). Participants with missing FDG-PET values were slightly older, were less educated, had a slightly lower MMSE and lower baseline and 6 months AVLT scores, and were followed for a shorter follow-up period. Little's MCAR test showed that values were not missing completely at random ($\chi^2$(6, N = 816) = 79518.03, $p$ < .001). Independent samples t-test showed that estimated values ($M$ = 123.89, $SD$ = 12.04) fell in the range of values for complete cases ($M$ = 124.43, $SD$ = 13.44, $t$(814) = -0.54, $p$ = .620). Cox proportional hazard regression replicated the results of analyses with values for complete cases. Older age (HR = 1.03, p = .016), lower MMSE (HR = 0.83, p < .001) and baseline AVLT scores (HR = 0.92, p < .001), having at least one ε4 allele, not showing practice effects and having a lower brain metabolism were associated with an increased AD-risk. However, although PE- again showed the most precise estimate of the AD-risk, the hazard ratio was higher for PE- (HR = 1.89, p < .001) than for APOE (HR for 1 ε4 allele = 1.64, p = .004; HR for 2 ε4 alleles = 1.65, p = .040) and FDG-PET (HR = 0.97, p < .001).

*3.4. Comparison of the AD-risk according to clinical profile*

Cox proportional hazard regression showed that the AD-risk was higher for those showing no practice effects only (HR = 2.49, p = .001), those having at least one APOE4 allele only (HR = 1.97, p = .002), and those having both conditions (HR = 3.41, p < .001) compared to those showing practice effects and with no APOE4 alleles. Figure 1 shows how the absolute AD-risk for participants with aMCI over a 6 year follow-up changes according to the available information. As shown, if an individual meets standard criteria for aMCI at baseline assessment, the expected AD-risk at six years without any additional information is 21.4%. If APOE testing is added to the medical record, the AD-risk increases for individuals with at least one APOE allele, with a risk estimate twice as high as that for APOE negative individuals. Adding data on practice effect obtained on a follow-up visit six months after baseline assessment again modifies the risk estimates. The AD-risk among APOE negative individuals is three times as high for individuals not showing practice effects, and is also higher than the risk estimate for aMCI at baseline. Among individuals with at least one APOE allele, those not showing practice effects have the greatest AD-risk, with a risk estimate twice as high as that for aMCI diagnosis at baseline. Interestingly, the AD-risk for APOE negative individuals not showing practice effects is higher than the AD-risk for APOE positive individuals showing practice effects.

13

## 4. DISCUSSION

This study aimed to analyze the value of practice effects on a list learning test between two assessments, conducted 6 months apart, for predicting progression from aMCI to AD during a 6 years follow-up period. We found that individuals with aMCI who did not show practice effects as expected based on a healthy comparison group had a significantly higher AD-risk, and also that the risk estimate for practice effects was higher than that for FDG-PET data and similar and more precise than that for APOE genotype.

Our results partially agree with previous research analyzing practice effects in aMCI. Contrary to what has been previously reported using verbal (Duff et al., 2008) and visual memory tests (Duff et al., 2007), we did not find significant improvements on AVLT delayed recall scores at retest in the aMCI sample. A possible explanation for this discrepancy could be related to the test-retest interval, which is inversely related to practice effects (Calamia et al., 2012). Whereas the studies by Duff et al. (2008, 2007) analyzed practice effects during 1-week and 2-week period, a 6-month period was used in this study. However, factors other than the length of follow-up must be considered, as practice effects have been reported for periods of 18 months after baseline assessment (Campos-Magdaleno et al., 2017). A possible explanation is that the large sample size in this study mitigated statistical characteristics of cognitive performance such as regression to the mean, which has also been associated with practice effects (Duff, 2012). The most important finding, however, is that more than 65% of individuals with aMCI showed practice effects when methods other than raw scores were used to define practice effects. Despite there being up to a third of

14

individuals with aMCI showing higher retest scores relative to baseline, 25% of the aMCI sample did not show significant practice effects. It is thus important to identify the frequency of test-retest changes in raw scores to more reliably identify practice effects.

This work used the regression-based reliable change index with that purpose, and used a $t_{n-2}$ distribution to identify the distribution of discrepancies between observed and predicted scores. This procedure has been suggested to more reliable identify true change compared to a z distribution, specially for small samples (Crawford and Garthwaite, 2007). This work is the first that analyzes practice effects for three different cut-off points along the distribution of discrepancies, and the one reporting that the estimate of the AD-risk in aMCI is similar for the lower 10%, 7% and 5% distribution of discrepancies between observed and predicted raw scores.

The main finding is that cognitive data, assessed through practice effects, were at least as useful for predicting AD over 6 years as genetic and biomarker data, results that were replicated when FDG-PET missing values were imputed using variables associated with missingness. Although practice effects were related to genetic data as previously reported (Duff et al., 2017b; Machulda et al., 2013), with the PE- group being more likely to have at least one copy of the APOE ε4 allele, the lack of practice effects outperformed FDG-PET in the identification of individuals at the greatest AD-risk, and showed a similar and more precise risk estimate than APOE genotype. This results are in line with those reported by Hassenstab et al. (2015), who found that APOE was not significant to predict worsening of clinical symptoms of cognitive impairment, and cannot support that FDG-PET are more sensitive than cognitive scores

15

for predicting AD in aMCI (Herholz et al., 2011). Duff et al. (2015) suggested that practice effects can be a proxy of certain biomarkers, and our data add that practice effects and biomarkers might be a useful combination to identify individuals at the greatest AD-risk during a 6-year follow-up. However, risk estimates for APOE and FDG-PET may be biased due to the association of these two variables in the ADNI database (Landau et al., 2013). Our results are also in line with the findings reported by Machulda et al. (2013), who found that APOE carriers' performance was similar to baseline after an average follow-up period of 6 years. Thus, future works will replicate whether practice effects are similar or even superior to genetic and biomarker data for predicting progression to AD in different follow-up periods.

Although promising, these results have some limitations. The ADNI project included only one cognitive test to define cognitive impairment in aMCI. Recent diagnostic approaches have reported both NC and MCI misdiagnoses when cognitive impairment is defined using several tests (Edmonds et al., 2015; Oltra-Cucarella et al., 2018). Second, to avoid circularity between test-retest regressions and risk analyses, we used a simple regression to predict retest AVLT-DR scores using baseline AVLT-DR scores without including other covariates. The use of a simple regression could have an impact on the discrepancy between observed and predicted scores, with multiple regressions providing a larger discrepancy (Duff et al., 2017a). However, the calculation of the standard error for a new case after conducting multiple regressions is computationally much more complicated than for simple regressions (Crawford et al., 2012), so using a simple regression may be more feasible to estimate the AD-risk in clinical practice where correlations between predictors are seldom available. The small and negligible differences between NC and MCI on age and education suggest that

16

these variables might have a little effect on the prediction of retest scores. Future research will clarify whether the use of multiple regressions provides additional information for the identification of individuals at an increased AD-risk over simple regressions.

Our results have important clinical implications that must be highlighted. The most recent criteria for MCI due to AD include biomarkers to define levels of certainty that MCI is a prodromal stage of AD (Albert et al., 2011). Our results show that practice effects can provide an estimate of the AD-risk over and above biomarker and genetic data, even if raw scores (particularly extreme low scores) increase at the second assessment in a proportion of individuals. Our findings could also be useful for interpreting the results of clinical trials (Brooks and Loewenstein, 2010), as it has been shown that it is important not only to identify changes in raw scores over a 6 months period but also to identify whether negative discrepancy between observed and expected scores are uncommon in healthy individuals who do not progress to AD. Showing a negative discrepancy at the bottom 10% of healthy individuals must warn about the increased AD-risk in individuals with aMCI. This would help to reduce the possibility that practice effects mask treatment effects (Goldberg et al., 2015), would help to identify eligible individuals for intervention trials, and also to interpret the presence of increased retest scores in those receiving cognitive or pharmacological interventions.

**Conclusions**

Rather than a source of error, the analysis of practice effects may be a valuable tool for the identification of individuals with aMCI at the greatest AD-risk. Practice effects on a verbal memory test may identify as accurately as genetic and biomarker data individuals with aMCI at the greatest AD-risk in a 6 year follow-up. Our results warrant further research with samples diagnosed with MCI using standard criteria, and for different test-retest periods and follow-up intervals.

18

**ACKNOWLEDGEMENTS**

**FIGURE CAPTIONS**

**Figure 1.** Estimates of the absolute AD-risk at six years according to clinical profile

NC: normal cognition. MCI: mild cognitive impairment. APOE+: individuals with one or more APOE allele. APOE-: individuals with no APOE allele. PE+: individuals showing practice effects on the 6-months Auditory Verbal Learning Test test-retest. PE-: individuals not showing practice effects on the 6-months Auditory Verbal Learning Test test-retest

**REFERENCES**

Albert, M.S., DeKosky, S.T., Dickson, D., Dubois, B., Feldman, H.H., Fox, N.C., Gamst, A.,

Holtzman, D.M., Jagust, W.J., Petersen, R.C., Snyder, P.J., Carrillo, M.C., Thies,

B., Phelps, C.H., 2011. The diagnosis of mild cognitive impairment due to

Alzheimer's disease: Recommendations from the National Institute on Aging-

Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's

disease. Alzheimers Dement. 7, 270–279.

https://doi.org/10.1016/j.jalz.2011.03.008

Brooks, L.G., Loewenstein, D.A., 2010. Assessing the progression of mild cognitive

impairment to Alzheimer's disease: current trends and future directions.

Alzheimers Res. Ther. 2. https://doi.org/10.1186/alzrt52

Calamia, M., Markon, K., Tranel, D., 2012. Scoring Higher the Second Time Around:

Meta-Analyses of Practice Effects in Neuropsychological Assessment. Clin.

Neuropsychol. 26, 543–570. https://doi.org/10.1080/13854046.2012.680913

Campos-Magdaleno, M., Facal, D., Lojo-Seoane, C., Pereiro, A.X., Juncos-Rabadán, O.,

2017. Longitudinal Assessment of Verbal Learning and Memory in Amnestic

Mild Cognitive Impairment: Practice Effects and Meaningful Changes. Front.

Psychol. 8. https://doi.org/10.3389/fpsyg.2017.01231

Crawford, J.R., Garthwaite, P.H., 2007. Using regression equations built from summary

data in the neuropsychological assessment of the individual case.

Neuropsychology 21, 611–620. https://doi.org/10.1037/0894-4105.21.5.611

Crawford, J.R., Garthwaite, P.H., 2005. Evaluation of Criteria for Classical Dissociations

in Single-Case Studies by Monte Carlo Simulation. Neuropsychology 19, 664–

678. https://doi.org/10.1037/0894-4105.19.5.664

Crawford, J.R., Garthwaite, P.H., Denham, A.K., Chelune, G.J., 2012. Using regression equations built from summary data in the psychological assessment of the individual case: Extension to multiple regression. Psychol. Assess. 24, 801–814. https://doi.org/10.1037/a0027699

Darby, D., Maruff, P., Collie, A., McStephen, M., 2002. Mild cognitive impairment can be detected by multiple assessments in a single day. Neurology 59, 1042–1046. https://doi.org/10.1212/WNL.59.7.1042

Duff, K., 2012. Evidence-Based Indicators of Neuropsychological Change in the Individual Patient: Relevant Concepts and Methods. Arch. Clin. Neuropsychol. 27, 248–261. https://doi.org/10.1093/arclin/acr120

Duff, K., Atkinson, T.J., Suhrie, K.R., Dalley, B.C.A., Schaefer, S.Y., Hammers, D.B., 2017a. Short-term practice effects in mild cognitive impairment: Evaluating different methods of change. J. Clin. Exp. Neuropsychol. 39, 396–407. https://doi.org/10.1080/13803395.2016.1230596

Duff, K., Beglinger, L., Schultz, S., Moser, D., Mccaffrey, R., Haase, R., Westervelt, H., Langbehn, D., Paulsen, J., 2007. Practice effects in the prediction of long-term cognitive outcome in three patient samples: A novel prognostic index. Arch. Clin. Neuropsychol. 22, 15–24. https://doi.org/10.1016/j.acn.2006.08.013

Duff, K., Beglinger, L.J., Van Der Heiden, S., Moser, D.J., Arndt, S., Schultz, S.K., Paulsen, J.S., 2008. Short-term practice effects in amnestic mild cognitive impairment: implications for diagnosis and treatment. Int. Psychogeriatr. 20. https://doi.org/10.1017/S1041610208007254

Duff, K., Chelune, G., Dennett, K., 2012. Within-Session Practice Effects in Patients

Referred for Suspected Dementia. Dement. Geriatr. Cogn. Disord. 33, 245–249.

https://doi.org/10.1159/000339268

Duff, K., Foster, N.L., Hoffman, J.M., 2014. Practice Effects and Amyloid Deposition:

Preliminary Data on a Method for Enriching Samples in Clinical Trials. Alzheimer

Dis. Assoc. Disord. 28, 247–252.

https://doi.org/10.1097/WAD.0000000000000021

Duff, K., Hammers, D.B., Dalley, B.C.A., Suhrie, K.R., Atkinson, T.J., Rasmussen, K.M.,

Horn, K.P., Beardmore, B.E., Burrell, L.D., Foster, N.L., Hoffman, J.M., 2017b.

Short-Term Practice Effects and Amyloid Deposition: Providing Information

Above and Beyond Baseline Cognition. J. Prev. Alzheimers Dis. 4, 87–92.

https://doi.org/10.14283/jpad.2017.9

Duff, K., Horn, K.P., Foster, N.L., Hoffman, J.M., 2015. Short-Term Practice Effects and

Brain Hypometabolism: Preliminary Data from an FDG PET Study. Arch. Clin.

Neuropsychol. 30, 264–270. https://doi.org/10.1093/arclin/acv018

Duff, K., Lyketsos, C.G., Beglinger, L.J., Chelune, G., Moser, D.J., Arndt, S., Schultz, S.K.,

Paulsen, J.S., Petersen, R.C., McCaffrey, R.J., 2011. Practice Effects Predict

Cognitive Outcome in Amnestic Mild Cognitive Impairment. Am. J. Geriatr.

Psychiatry 19, 932–939. https://doi.org/10.1097/JGP.0b013e318209dd3a

Edmonds, E.C., Delano-Wood, L., Clark, L.R., Jak, A.J., Nation, D.A., McDonald, C.R.,

Libon, D.J., Au, R., Galasko, D., Salmon, D.P., Bondi, M.W., 2015. Susceptibility

of the conventional criteria for mild cognitive impairment to false-positive

diagnostic errors. Alzheimers Dement. 11, 415–424.

https://doi.org/10.1016/j.jalz.2014.03.005

Elias-Sonnenschein, L.S., Viechtbauer, W., Ramakers, I.H.G.B., Verhey, F.R.J., Visser, P.J., 2011. Predictive value of APOE- 4 allele for progression from MCI to AD-type dementia: a meta-analysis. J. Neurol. Neurosurg. Psychiatry 82, 1149–1156. https://doi.org/10.1136/jnnp.2010.231555

Gavett, B.E., Gurnani, A.S., Saurman, J.L., Chapman, K.R., Steinberg, E.G., Martin, B., Chaisson, C.E., Mez, J., Tripodis, Y., Stern, R.A., 2016. Practice Effects on Story Memory and List Learning Tests in the Neuropsychological Assessment of Older Adults. PLOS ONE 11, e0164492. https://doi.org/10.1371/journal.pone.0164492

Goldberg, T.E., Harvey, P.D., Wesnes, K.A., Snyder, P.J., Schneider, L.S., 2015. Practice effects due to serial cognitive assessment: Implications for preclinical Alzheimer's disease randomized controlled trials. Alzheimers Dement. Diagn. Assess. Dis. Monit. 1, 103–111. https://doi.org/10.1016/j.dadm.2014.11.003

Hassenstab, J., Ruvolo, D., Jasielec, M., Xiong, C., Grant, E., Morris, J.C., 2015. Absence of practice effects in preclinical Alzheimer's disease. Neuropsychology 29, 940–948. https://doi.org/10.1037/neu0000208

Herholz, K., Westwood, S., Haense, C., Dunn, G., 2011. Evaluation of a Calibrated 18F-FDG PET Score as a Biomarker for Progression in Alzheimer Disease and Mild Cognitive Impairment. J. Nucl. Med. 52, 1218–1226. https://doi.org/10.2967/jnumed.111.090902

Landau, S.M., Harvey, D., Madison, C.M., Koeppe, R.A., Reiman, E.M., Foster, N.L., Weiner, M.W., Jagust, W.J., 2011. Associations between cognitive, functional, and FDG-PET measures of decline in AD and MCI. Neurobiol. Aging 32, 1207–1218. https://doi.org/10.1016/j.neurobiolaging.2009.07.002

Landau, S.M., Lu, M., Joshi, A.D., Pontecorvo, M., Mintun, M.A., Trojanowski, J.Q., Shaw, L.M., Jagust, W.J., for the Alzheimer's Disease Neuroimaging Initiative, 2013. Comparing positron emission tomography imaging and cerebrospinal fluid measurements of β-amyloid: CSF and Amyloid PET Imaging. Ann. Neurol. 74, 826–836. https://doi.org/10.1002/ana.23908

Machulda, M.M., Pankratz, V.S., Christianson, T.J., Ivnik, R.J., Mielke, M.M., Roberts, R.O., Knopman, D.S., Boeve, B.F., Petersen, R.C., 2013. 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. 27, 1247–1264. https://doi.org/10.1080/13854046.2013.836567

McKhann, G.M., Drachman, D., Folstein, M., Katzman, R., Price, D., Stadlan, E.M., 1984. 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 34, 939–939. https://doi.org/10.1212/WNL.34.7.939

McKhann, G.M., Knopman, D.S., Chertkow, H., Hyman, B.T., Jack, C.R., Kawas, C.H., Klunk, W.E., Koroshetz, W.J., Manly, J.J., Mayeux, R., Mohs, R.C., Morris, J.C., Rossor, M.N., Scheltens, P., Carrillo, M.C., Thies, B., Weintraub, S., Phelps, C.H., 2011. 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. 7, 263–9. https://doi.org/10.1016/j.jalz.2011.03.005

Mitchell, A., Shiri-Feshki, M., 2009. Rate of progression of mild cognitive impairment to

dementia - Meta-analysis of 41 robust inception cohort studies. Acta Psychiatr.

Scand. 119, 252–265. https://doi.org/10.1111/j.1600-0447.2008.01326.x

Oltra-Cucarella, J., Sánchez-SanSegundo, M., Lipnicki, D.M., Sachdev, P.S., Crawford,

J.D., Pérez-Vicente, J.A., Cabello, L., Ferrer-Cascales, R., 2018. Using the base

rate of low scores helps to identify progression from amnestic MCI to AD. J. Am.

Geriatr. Soc. https://doi.org/10.1111/jgs.15412

Petersen, R., Smith, G., Waring, S., Ivnik, R., Tangalos, T., Kokmen, E., 1999. Mild

cognitive impairment: Clinical characterization and outcome. Arch. Neurol. 56,

303–308. https://doi.org/10.1001/archneur.56.3.303

Qian, J., Wolters, F.J., Beiser, A., Haan, M., Ikram, M.A., Karlawish, J., Langbaum, J.B.,

Neuhaus, J.M., Reiman, E.M., Roberts, J.S., Seshadri, S., Tariot, P.N., Woods,

B.M., Betensky, R.A., Blacker, D., 2017. APOE-related risk of mild cognitive

impairment and dementia for prevention trials: An analysis of four cohorts.

PLOS Med. 14, e1002254. https://doi.org/10.1371/journal.pmed.1002254

Richardson, J.T.E., 2011. Eta squared and partial eta squared as measures of effect size

in educational research. Educ. Res. Rev. 6, 135–147.

https://doi.org/10.1016/j.edurev.2010.12.001

Schrijnemaekers, A.M.C., de Jager, C.A., Hogervorst, E., Budge, M.M., 2006. Cases with

Mild Cognitive Impairment and Alzheimer's Disease Fail to Benefit from

Repeated Exposure to Episodic Memory Tests as Compared with Controls. J.

Clin. Exp. Neuropsychol. 28, 438–455.

https://doi.org/10.1080/13803390590935462

Smailagic, N., Vacante, M., Hyde, C., Martin, S., Ukoumunne, O., Sachpekidis, C., 2015.
18F-FDG PET for the early diagnosis of Alzheimer's disease dementia and other
dementias in people with mild cognitive impairment (MCI). Cochrane Database
Syst. Rev. https://doi.org/10.1002/14651858.CD010632.pub2

Tabachnick, B.G., Fidell, L.S., 2013. Using Multivariate Statistics., Sixth Ed. ed. Pearson
Education Inc, New Jersey.

Yu, J.-T., Tan, L., Hardy, J., 2014. Apolipoprotein E in Alzheimer's Disease: An Update.
Annu. Rev. Neurosci. 37, 79–100. https://doi.org/10.1146/annurev-neuro-
071013-014300

Zehnder, A.E., Bläsi, S., Berres, M., Monsch, A.U., Stähelin, H.B., Spiegel, R., 2009.
Impact of APOE status on cognitive maintenance in healthy elderly persons. Int.
J. Geriatr. Psychiatry 24, 132–141. https://doi.org/10.1002/gps.2080

Table 1. Demographic, cognitive, genetic and biomarker data

| | NC (n = 394) | MCI (n = 816) | | |
| --- | --- | --- | --- | --- |
| | Mean (SD) | Mean (SD) | $t/\chi^2$ | $P$ |
| Age | 74.83 (5.73) | 73.06 (7.47) | 4.16 | < .001 |
| Years of education | 16.30 (2.73) | 15.92 (2.86) | 2.32 | .026 |
| Sex (M/F) | 204/190 | 483/333 | 5.95 | .015 |
| MMSE | 29.06 (1.14) | 27.56 (1.82) | 14.91 | < .001 |
| AVLT baseline | 12.82 (2.42) | 10.49 (3.48) | 11.91 | < .001 |
| AVLT 6-months | 12.66 (2.44) | 10.06 (3.65) | 12.82 | < .001 |
| Ethnicity | | | | |
| Not hispano/latino, n (%) | 378 (95.9) | 787 (96.4) | | |
| Hispano/latino, n (%) | 14 (3.6) | 25 (3.1) | | |
| Unknown, n (%) | 2 (0.5) | 4 (0.5) | 0.21 | .902 |
| Married, n (%) | 268 (68) | 632 (77.5) | 19.28 | .001 |
| APOE ε4 | | | | |
| 0, n (%) | 287 (72.8) | 397 (48.7) | | |
| 1, n (%) | 96 (24.4) | 327 (40.1) | | |
| 2, n (%) | 11 (2.8) | 92 (11.3) | 68.72 | < .001 |
| FDG-PET | 130.70 (11.47) | 124.43 (13.44) | 6.71 | < .001 |

NC: normal cognition. MCI: mild cognitive impairment. MMSE: Mini-Mental State Examination. AVLT: Auditory Verbal Learning Test. FDG-PET: fluorodeoxyglucose positron emission tomography. APOE: apolipoprotein E

Table 2. Cox proportional hazard ratios of risk for Alzheimer's disease

| | Model 1 | | Model 2 | | Model 3 | |
|---|---|---|---|---|---|---|
| | HR (95%CI) | *p* | HR (95%CI) | *p* | HR (95%CI) | *p* |
| Age | 1.06 (1.03, 1.09) | .000 | 1.04 (1.00, 1.07) | .025 | 1.04 (1.00, 1.07) | .022 |
| Education | 0.99 (0.92, 1.08) | .933 | | | | |
| Sex | 0.89 (0.57, 1.39) | .623 | | | | |
| MMSE | | | 0.79 (0.70, 0.90) | .000 | 0.85 (0.75, 0.97) | .017 |
| AVLT-DR | | | 0.89 (0.83, 0.95) | .000 | 0.92 (0.86, 0.97) | .005 |
| FDG-PET | | | | | 0.95 (0.94, 0.97) | .000 |
| 1 APOE ε4 allele | | | | | 1.51 (0.92, 2.47) | .105 |
| 2 APOE ε4 alleles | | | | | 1.92 (0.99, 3.71) | .052 |
| PE $t_{n-2}$ 10% | -- | | 1.73 (1.10, 2.72) | .017 | 1.58 (1.00, 2.48) | .047 |
| PE $t_{n-2}$ 7% | -- | | 1.94 (1.23, 3.06) | .004 | 1.79 (1.13, 2.81) | .012 |
| PE $t_{n-2}$ 5% | -- | | 2.10 (1.33, 3.33) | .002 | 1.93 (1.22, 3.05) | .005 |

MMSE: Mini-Mental State Examination. AVLT-DR: baseline Auditory Verbal Learning Test delayed recall scores. FDG-PET: fluorodeoxyglucose positron emission tomography (N = 815). APOE: apolipoprotein E-4. PE $t_{n-2}$: practice effects using a $t_{n-2}$ distribution as defined in the text. HR: hazard ratio. CI: confidence interval. Separate backward stepwise Cox regressions were performed for PE $t_{n-2}$ 10%, PE $t_{n-2}$ 7% and PE $t_{n-2}$ 5%
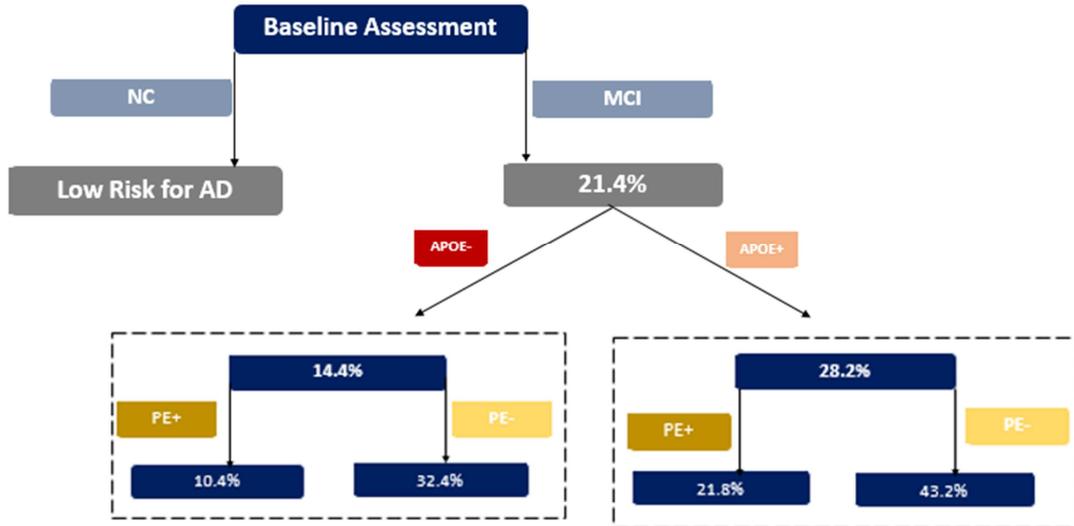
Highlights

1. Practice effects can identify individuals at the greatest risk for Alzheimer's disease

2. Practice effects might be as useful as APOE or FDG-PET to identify progressors to AD

3. These results are important for both clinical trials and clinical settings

4. Whether the same results are replicated in different samples warrants further

    research

The authors disclose:

1. No conflicts of interest

2. That sources of financial support have been disclosed in the manuscript: Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

3. That the data contained in the manuscript being submitted have not been previously published, have not been submitted elsewhere and will not be submitted elsewhere while under consideration at Neurobiology of Aging

4. Appropriate approval and procedures were used concerning human subjects both for the ADNI project and at each participating site