

Short-Term Practice Effects and Brain Hypometabolism: Preliminary Data from an FDG PET Study

Kevin Duff^{1,*}, Kevin P. Horn², Norman L. Foster¹, John M. Hoffman²

¹Center for Alzheimer's Care, Imaging, and Research, Department of Neurology, University of Utah, Salt Lake City, USA

²Center for Quantitative Cancer Imaging, Huntsman Cancer Institute, University of Utah, Salt Lake City, USA

*Corresponding author at: Center for Alzheimer's Care, Imaging and Research, University of Utah, Department of Neurology, 650 Komas Drive #106-A, Salt Lake City, UT 84108, USA. Tel: +1-801-581-2483; fax: +1-801-581-2483
E-mail address: kevin.duff@hsc.utah.edu (K. Duff).

Accepted 17 March 2015

Abstract

Practice effects are improvements in cognitive test scores due to repeated exposure to the same tests. Typically viewed as error, short-term practice effects have been shown to provide valuable clinical information about diagnosis, prognosis, and treatment outcomes in older patients with mild cognitive impairments. This study examined short-term practice effects across one week and brain hypometabolism on fluoro-2-deoxyglucose (FDG) positron emission tomography (PET) in 25 older adults (15 intact, 10 Mild Cognitive Impairment). Averaged cerebral brain metabolism on FDG PET was correlated with multiple cognitive scores at baseline in those with Mild Cognitive Impairment, and short-term practice effects accounted for additional variance in these same subjects. The relationship between brain metabolism and cognition (either at baseline or practice effects) was minimal in the intact individuals. Although needing replication in larger samples, short-term practice effects on tests of executive functioning and memory may provide valuable information about biomarkers of Alzheimer's disease.

Keywords: Practice effects/reliable change; Neuroimaging (functional); Elderly/Geriatrics/Aging

Introduction

Practice effects are improvements in cognitive test performance due to repeated exposure to the same testing materials (McCaffrey & Westervelt, 1995). It is suspected that practice effects invoke multiple cognitive abilities, including learning and memory, intellect, and novelty (Duff, Callister, Dennett, & Tometich, 2012; Rapport, Brines, Axelrod, & Theisen, 1997; Suchy, Kraybill, & Franchow, 2011). In a recent meta-analysis of these improvements due to repeat testing, Calamia, Markon, and Tranel (2012) provided evidence of how complicated practice effects can be. For example, they reported that practice effects tend to be more salient on certain types of tests (e.g., memory, processing speed) than others (e.g., visuospatial abilities, verbal knowledge), with considerable variability within cognitive domains. They also reported that age, retest interval, and use of alternate forms could influence the amount and direction of practice effects. Despite this, short-term practice effects are quite robust in healthy older adults (Beglinger et al., 2005; Calamia et al., 2012), but they tend to be smaller or absent in those with mild cognitive impairment (MCI) and dementia (Cooper et al., 2001; Cooper, Lacritz, Weiner, Rosenberg, & Cullum, 2004; Duff et al., 2008; Duff, Chelune, & Dennett, 2012). Although practice effects are typically viewed as error variance in repeated assessments, short-term practice effects may hold unique information about cognition. For example, using practice effects across brief periods (e.g., hours to days), relationships have been found with diagnosis (Darby, Maruff, Collie, & McStephen, 2002; Duff et al., 2008; Fernandez-Ballesteros, Zamarron, & Tarraga, 2005), prognosis (Duff et al., 2007, 2011), and treatment response (Duff, Beglinger, Moser, Schultz, & Paulsen, 2010; Fernandez-Ballesteros et al., 2012) in older cohorts of cognitively intact and impaired individuals. In these studies, individuals with lower-than-expected practice effects are more likely to be diagnosed with a cognitive disorder, more likely to cognitively worsen over time, and less likely to respond to a cognitive intervention.

Little published research has been done, however, to compare practice effects to biomarkers of Alzheimer's disease (AD), which would be a natural extension of this line of investigation. In one study, Machulda and colleagues (2014) reported that APOE-4 status was related to diminishing practice effects across 15 months in a large cohort of cognitively normal individuals. In another study, poor practice effects across 1 week on a visual memory test was associated with greater amyloid deposition in the brain (Duff, Foster, & Hoffman, 2014). Finally, individuals with mild cognitive impairments who were no longer benefitting from practice after 2 weeks showed reduced global cerebral blood flow to more challenging cognitive tasks studied with O¹⁵-water positron emission tomography (PET) (Boles Ponto, Magnotta, Moser, Duff, & Schultz, 2006). Clearly, more work is needed in this area, as practice effects can vary depending on retest interval, cognitive domain assessed, individual tests administered, and patient diagnosis (Calamia et al., 2012). However, if this prior work can be validated, then practice effects might serve as a proxy measure of certain biomarkers in clinical trials of AD, as well as in trials of MCI, a group at greater risk for progressing to AD (Manly et al., 2008; Tabert et al., 2006).

Given the relative paucity of literature on the relationship between short-term practice effects on cognitive testing and biomarkers of AD pathology, the purpose of this study was to examine this relationship using ¹⁸F-fluoro-2-deoxyglucose (FDG) PET in a sample of non-demented older adults. It was hypothesized that brain hypometabolism on FDG PET would be correlated with smaller short-term practice effects on cognitive testing.

Methods

Participants

Twenty-five older adults (mean age = 74.6 [6.8] years, mean education = 16.1 [3.2] years, 18 females/7 males, and all Caucasian) were enrolled in this study. These individuals were all recruited from senior centers and independent living facilities to participate in studies on memory and aging. All reported to be functionally independent in activities of daily living, and this was corroborated by a knowledgeable informant. Based on objective cognitive testing (described below), a neuropsychologist (KD) classified the majority of these individuals as cognitively intact ($n = 15$), with the remainder being characterized as MCI ($n = 10$) according to existing criteria (Petersen et al., 1999). The MCI individuals exhibited at least an amnesic profile, with memory functioning falling 1.5 *SDs* below expectations. Exclusion criteria for this study included: history of neurological disease known to affect cognition (e.g., stroke, head injury with loss of consciousness of >30 min, seizure disorder, demyelinating disorder, etc.); dementia based on DSM-IV criteria; current or past major psychiatric illness (e.g., schizophrenia, bipolar affective disorder); 30-item Geriatric Depression Score >15; history of substance abuse; current use of cholinesterase inhibitors, other cognitive enhancers, antipsychotics, or anticonvulsant medications; history of radiation therapy to the brain; history of significant major medical illnesses, such as cancer or AIDS; uncontrolled diabetes or blood glucose >180 mg/dl on the day of the FDG PET scan, claustrophobia, and currently pregnant or lactating.

Procedures

The local institutional review board approved all procedures and all participants provided informed consent before data collection commenced. As part of a larger study, all participants completed a neuropsychological battery of commonly used tests designed to characterize cognitive status across several domains, including learning and memory (Hopkins Verbal Learning Test—Revised [HVLTR] and Brief Visuospatial Memory Test—Revised [BVMT-R]), attention and processing speed (Trail Making Test Part A [TMTA] and Symbol Digit Modalities Test [SDMT]), and executive functioning (Trail Making Test Part B [TMTB] and Controlled Oral Word Association Test [COWAT]). An estimate of premorbid intellect (Reading subtest of the Wide Range Achievement Test—IV) and screening test of global cognition (Mini-Mental Status Examination) were also included. For each of these cognitive scores, higher values indicate better performance, with the exception of TMTA and TMTB, which were the opposite. All neuropsychological tests in Table 1 were repeated after ~1 week to quantify practice effects. Alternate forms were purposely not used in this study to maximize practice effects. A 1-week retest interval was chosen based on prior studies showing that practice effects across this period provided useful information about the diagnosis, prognosis, and treatment recommendations in individuals with MCI (Duff et al., 2007, 2008, 2010, 2011).

Following completion of the cognitive battery, participants underwent FDG PET imaging. FDG was produced under PET cGMP standards and the studies were conducted under an approved IND. Patients fasted for 4 h and blood glucose was measured and confirmed to be <180 mg/dl (9.9 mmol/l) prior to the administration of FDG. Patients were injected with 370 mBq (10 mCi) of FDG and allowed to rest comfortably in a quiet room with eyes open and ears un-occluded during uptake. Imaging was then

Table 1. Cognitive test scores, practice effects z -scores, and correlations with FDG PET composite

Cognitive measure	Baseline	1 week	Practice effects z -score	r (d)	pr (d)
BVMT-R Total Recall					
Intact	18.7 (6.4)	26.7 (6.2)	−0.2 (0.7)	.49 (1.1)	−0.21 (0.4)
MCI	12.1 (7.3)	21.3 (7.6)	−0.1 (0.9)	−.64* (1.7)	−0.09 (0.2)
BVMT-R Delayed Recall					
Intact	7.7 (2.7)	9.9 (1.7)	0.1 (0.6)	.20 (0.4)	0.10 (0.2)
MCI	5.3 (3.3)	7.8 (3.1)	−0.4 (1.0)	−.29 (0.6)	−0.38 (0.8)
HVLT-R Total Recall					
Intact	29.2 (4.3)	31.5 (3.5)	−0.2 (0.5)	.24 (0.5)	0.24 (0.5)
MCI	23.8 (5.4)	27.0 (6.3)	−0.3 (0.7)	−.70* (2.0)	−0.53 (1.2)
HVLT-R Delayed Recall					
Intact	9.4 (2.4)	10.6 (1.9)	0.0 (0.5)	.48 (1.1)	0.16 (0.3)
MCI	5.5 (3.6)	9.1 (2.5)	0.2 (1.0)	−.40 (0.9)	−0.51 (1.2)
TMTA (time to complete)					
Intact	37.4 (12.5)	43.9 (42.6)	1.0 (4.5)	−.01 (0.0)	0.31 (0.6)
MCI	50.1 (23.8)	46.8 (22.1)	0.5 (2.8)	.12 (0.2)	−0.30 (0.6)
TMTB (time to complete)					
Intact	84.8 (26.1)	77.1 (27.5)	−0.1 (0.5)	−.44 (1.0)	0.19 (0.4)
MCI	112.1 (56.8)	111.4 (74.1)	0.6 (1.4)	.79* (2.6)	0.63 (1.6)
SDMT (correct in 90'')					
Intact	44.4 (6.6)	48.0 (9.1)	0.2 (1.0)	−.13 (0.3)	0.12 (0.2)
MCI	35.8 (9.1)	37.7 (9.8)	−0.5 (0.9)	−.56 (1.3)	−0.52 (1.2)
COWAT (correct across three 60'' trials)					
Intact	38.7 (9.5)	41.1 (11.8)	0.1 (1.2)	.08 (0.2)	0.08 (0.2)
MCI	34.7 (8.1)	35.4 (12.5)	−0.3 (1.3)	−.64 (1.7)	−0.02 (0.0)

Notes: Baseline and 1 week scores are raw scores (means and SD s), where higher scores indicate better cognitive performance for all tests except TMTA and TMTB, which is opposite. Practice effects z -scores are based on standardized regression-based change formulae (see Methods for additional details), which quantifies how much change an individual shows across 1 week on a particular test compared with 167 non-demented older adults who were also administered this battery at baseline and 1 week. For all the practice effects, z -scores, except TMTA and TMTB, higher values indicate larger practice effects compared with the peer group. For TMTA and TMTB practice effects z -scores, the opposite was true (i.e., lower values indicate larger practice effects).

r = Pearson correlation between baseline cognitive scores and FDG PET composite score. pr = partial Pearson correlation between practice effects z -score and FDG PET composite score, controlling for baseline cognitive score. d = Cohen's d effect size. MCI = mild cognitive impairment, BVMT-R = Brief Visuospatial Memory Test—Revised, HVLT-R = Hopkins Verbal Learning Test—Revised, TMTA = Trail Making Test Part A, TMTB = Trail Making Test Part B, SDMT = Symbol Digit Modalities Test, COWAT = Controlled Oral Word Association Test.

* $p < .05$.

performed 30 min after the injection of FDG. Emission imaging time was 30 min. Two different PET tomographs were used for the imaging, a GE Advance PET scanner or a GE Discovery DST PET/CT scanner. The two PET tomographs have essentially equivalent performance characteristics as determined by phantom brain studies. The images were acquired in 3D mode with measured attenuation correction utilizing Germanium rod sources (GE Advance) or with CT attenuation correction (GE DST PET/CT). The image reconstruction used for both PET tomographs was iterative, 30–32 subsets, loop filter of 2.0, and diameter of 25.6 cm. FDG PET data were analyzed using CortexID, a software application available from GE Healthcare. Cortex ID is an FDA approved brain image analysis software designed to quantify FDG uptake measured with PET and PET-CT brain scans. The software aids clinicians in the assessment of FDG PET scans by providing automated analysis through quantification of the comparison of local peak FDG uptake activity values, at standardized bilateral anatomical locations, compared with the corresponding reference normal peak activity in age-stratified normal control subjects. The resulting quantification is presented through 3D Stereotactic Surface Projection maps of the brain as initially described by Minoshima, Frey, Koeppe, Foster, and Kuhl (1995). CortexID allows for user generated information on the relative changes in metabolic activity normalized to the pons presented as regional and global z -scores between a subject's images and age-stratified controls. A global z -score composite of FDG uptake from CortexID ("averaged cerebral") was used in these analyses for multiple reasons. First, using a single composite score reduced the number of statistical comparisons in this small study and lowers the risk of Type I error. Second, the pattern of FDG PET hypometabolism seen in AD tends to be quite widely distributed (e.g., posterior cingulate, parietal, temporal, and frontal regions) (Chen et al., 2011; Herholz et al., 2002; Hoffman et al., 2000). Third, we view practice effects as a more global measure of cortical integrity and cognitive plasticity that does not neatly localize (Duff, Callister, et al., 2012). In the global z -score composite, higher values indicate lower brain metabolism.

Statistical Analyses

Raw scores on all cognitive tests were used in these analyses. To generate a practice effects score on each measure, a standardized regression-based prediction formula was developed using an independent sample of 167 non-demented older adults (mean age = 78.6 years, mean education = 15.4 years, 81% female) who were also administered this battery at baseline and 1 week (Duff, 2014). This methodology, which allowed us to predict one week scores from observed baseline scores, is frequently used in the assessment of neuropsychological change (Duff, 2012). Briefly, stepwise linear regression models predicted 1 week scores from baseline scores, age, years of education, and gender. The resulting formulae were applied to the current sample to yield predicted 1 week scores. The predicted 1 week scores were subtracted from the observed 1 week scores and divided by the standard error of the estimate of the regression equations to yield z -scores that reflects how much the predicted score deviated from the observed score for each participant. In this calculation, a z -score close to 0 indicates that the current participant showed a “normal” amount of practice effects compared with peers, whereas a z -score of -0.5 indicates that the current participant showed smaller practice effects (i.e., about a half of an SD less) compared with peers. Additional details of this practice effects score, including its calculation, can be obtained from the first author or in Duff (2014).

In the primary analyses, Pearson correlations were calculated between the FDG PET composite and the baseline scores on each of the cognitive tests. Partial correlations were next calculated between the FDG PET composite and the practice effects z -scores, controlling for the baseline cognitive scores. These analyses were computed separately for the two groups (intact and MCI). An α value of 0.05 was used for these comparisons. However, due to the small sample sizes of the two groups, effect sizes (e.g., Cohen’s d) were calculated for trends heading towards statistical significance.

Results

Participants in the two groups were comparable in age ($t[23] = 0.47, p = .64$, intact mean = 75.1 [5.9] years, MCI mean = 73.8 [8.2] years) and education ($t[23] = 0.40, p = .69$, intact mean = 15.9 [3.4] years, MCI mean = 16.4 [2.9] years), although not in gender distribution ($\chi^2[1] = 4.0, p = .04$, intact = 13 females/2 males, MCI = 5 females/5 males). Premorbid intellect was comparable between the two groups (WRAT-IV Reading, $t[23] = 1.12, p = .25$, intact mean = 109.3 [7.1] standard score points, MCI mean = 114.0 [12.9] standard score points), as was performance on a cognitive screening measure (MMSE, $t[23] = 0.58, p = .57$, intact mean = 28.3 [1.7], MCI mean = 27.9 [1.4]). Consistent with group classification, delayed recall was significantly lower in the MCI participants at the baseline visit (HVLTR-Delayed Recall: $t[23] = 4.4, p < .001$, intact mean = 106.8 [11.9] standard score points, MCI mean = 78.1 [20.9] standard score points; BVMT-R delayed recall: $t[23] = 2.2, p = .04$, intact mean = 98.5 [16.3] standard score points, MCI mean = 83.2 [19.4] standard score points). Across the eight repeated cognitive scores, the cognitive intact subjects tended to have slightly higher practice effects z -scores than those classified as MCI, although none were statistically significant (see Table 1). Finally, the global composite z -score for FDG PET, normalized to the pons, was comparable between the two groups ($t[23] = 1.17, p = .25$, intact mean = 0.21 [0.55], MCI mean = 0.58 [1.05]).

The correlations between FDG PET metabolism and baseline cognitive test scores for the intact and MCI groups are presented in Table 1. In the individuals classified as MCI, brain metabolism was significantly correlated the following cognitive scores at baseline: BVMT-R Total Recall ($r = -.64, p = .04$), HVLTR Total Recall ($r = -.70, p = .03$), and TMTB ($r = .79, p = .01$). For each of these correlations, better baseline cognitive performance was associated with more hypometabolism on FDG PET. In the intact subjects, no correlations were statistically significant at $p < .05$.

The partial correlations between FDG PET metabolism and the practice effects z -scores, controlling for baseline cognitive scores, for the two groups are also presented in Table 1. Although none of these partial correlations were statistically significant in either group, there were noteworthy trends in the MCI sample. For example, in the MCI subjects, practice effects on TMTB accounted for a sizeable amount of variance in brain metabolism, above and beyond the baseline TMTB score ($pr = 0.63, p = .09$). Similarly, after accounting for the baseline cognitive score, practice effects across 1 week added to the variance of FDG PET metabolism for HVLTR Total Recall ($pr = -0.53, p = .18$), HVLTR Delayed Recall ($pr = -0.51, p = .19$), BVMT-R Delayed Recall ($pr = -0.38, p = .31$), and SDMT ($pr = -0.52, p = .19$). Admittedly, these partial correlations are not statistically significant, but they all trend in the expected direction showing that individuals with more hypometabolism showed poorer (i.e., lower) practice effects across 1 week on these tests. The partial correlations in the intact subjects were smaller (e.g., TMTB: $pr = 0.19, p = .51$; HVLTR Total Recall: $pr = 0.24, p = .40$).

Discussion

Given the limited literature on the relationship between short-term practice effects on cognitive testing and biomarkers of AD pathology, the current study sought to examine the association between these improvements in test scores across 1 week and brain

metabolism using FDG PET in older adults. We first examined the relationship between baseline cognitive scores and brain metabolism. In the MCI sample, multiple baseline cognitive test scores were related to brain metabolism, with better cognition being associated with higher metabolism. However, there was considerable variability in these relationships, with 1%–62% of the variance being shared. We next examined the additional contribution of practice effects on these same cognitive tests. Again in the MCI sample, practice effects across 1 week on TMTB explained an additional 40% of the variance in FDG PET findings, above and beyond baseline TMTB scores. That is, whereas baseline TMTB scores provided valuable information about brain metabolism, practice effects on TMTB augmented our knowledge about brain metabolism in individuals classified as MCI. Practice effects on HVLTR, BVMT-R, and SDMT also seemed to add to our understanding of FDG PET findings in those with MCI. The relationship between brain metabolism and cognition (either at baseline or practice effects) was minimal in the cognitively intact individuals. Although preliminary, these results suggest that practice effects may be useful as potential proxy measures of certain biomarkers in clinical trials of MCI.

As noted earlier, few studies have examined how practice effects link to biomarkers in AD, but the current results are largely consistent with those studies. In the only other PET study that we could identify, [Boles Ponto and colleagues \(2006\)](#) found reduced global cerebral blood flow in older individuals who were showing diminished practice effects on the BVMT-R in an O^{15} -water PET study. Using short-term practice effects on the BVMT-R, [Duff and colleagues \(2014\)](#) reported that brain amyloid deposition was five times higher in individuals with low practice effects compared with those with high practice effects. Finally, cognitively intact APOE-4 carriers were less likely than non-carriers to maintain practice effects on a memory composite across time ([Machulda et al., 2014](#)). All of these studies, including the current one, suggest that less-than-expected practice effects provide useful information about biomarkers of AD.

Interestingly, only a few cognitive variables in the current battery were significantly associated with brain hypometabolism on FDG PET in this cohort. For example, in the MCI subjects, cerebral metabolism was correlated with baseline scores on the BVMT-R, HVLTR, and TMTB. Practice effects on these same measures and the SDMT seemed to account for additional variance in the FDG PET composite, above and beyond the baseline test scores in the subjects classified as MCI. These findings are in agreement with the existing literature on practice effects. For example, in a meta-analysis of practice effects, [Calamia and colleagues \(2012\)](#) noted that practice effects on test of memory and processing speed tended to be the largest. These results are also consistent with findings in studies on baseline cognitive testing. For example, in the multi-site, international Alzheimer's Disease Neuroimaging Initiative trial, [Ewers and colleagues \(2013\)](#) observed that a combination of FDG PET and TMTB results best predicted progression from cognitively intact to MCI. In the same cohort, specific AD-related parietotemporal hypometabolic patterns were associated with executive dysfunction (including the TMTB) in MCI and AD patients, and other patterns were associated with memory difficulties (including tests similar to the HVLTR) in MCI patients ([Habeck et al., 2012](#)). The FDG PET findings of patients with schizophrenia best predicted TMTB performance, especially in the bilateral frontal gyri ([Horacek et al., 2006](#)). Overall, these studies suggest that tests like TMTB and HVLTR tend to be associated with brain metabolism measured with FDG PET in various neuropsychiatric conditions.

A few notes about practice effects also seem necessary. First, our calculation of short-term practice effects was not based on a simple difference score (e.g., time 2 – time 1). Instead, we utilized regression-based models to predict time 2 scores from time 1 scores using a large independent cohort ([Duff, 2014](#)). The predicted time 2 score was compared with the observed time 2 score to yield a practice effects z -score that shows how far an individual's practice effect deviated from their peers. Practice effects z -scores close to 0 indicate little deviation compared with peers, whereas z -scores of ± 1.0 indicate considerably more deviation. These regression-based models and the practice effects z -scores that they yield tend to be more sensitive indicators of change than the simple difference score ([Barr, 2002](#); [Temkin, Heaton, Grant, & Dikmen, 1999](#)). This also was observed in our data. For example, in Table 1, the simple difference of the SDMT in intact subjects was +3.5 (i.e., 1 week – baseline = 3.5 more items correct after 1 week), whereas the practice effects z -score was 0.2 (i.e., 2/10 an SD better across 1 week compared with 167 peers). Conversely, on this same test in MCI subjects, the simple difference of +2 was “normatively” poorer (e.g., about a half an SD poorer than peers). Therefore, the regression-based calculation, which also controls for baseline level of test performance and demographic variables, may be more sensitive than the simple difference method. Second, using this regression-based model of practice effects, the current sample largely performed around expectations, with mean practice effects z -scores being no more than one half an SD from 0 for nearly all tests. The lone exception was TMTA, on which the intact sample fell below expectations compared with the peer group (mean practice effects z -score = +1.0, which is reversed because higher scores indicate poorer performance on this test). It is not clear why the current sample showed smaller-than-expected practice effects on this one test, but it does raise some concerns about the generalizability of this sample. Finally, we view practice effects as a complex cognitive phenomenon. They can vary as a function of retest interval, cognitive domain, characteristics of the specific tests, and patient diagnosis ([Calamia et al., 2012](#)). Moreover, we view them as a global measure of cortical integrity and cognitive plasticity ([Duff, Callister, et al., 2012](#)).

The potential utility of these findings deserves some mention. We realize that our use of a short retest interval is an atypical clinical scenario. However, prior work has indicated that this brief retest interval can provide some clarity about diagnosis, longer-term prognosis, and even treatment recommendations (Duff et al., 2007, 2008, 2010, 2011). Furthermore, there are clinical scenarios where brief repeat testing may be used. For example, clinicians may seek to track short-term improvements within a rehabilitation setting, examine benefits of fast-acting medications, or assess rapidly evolving neurological conditions (e.g., delirium). The regression-based practice effects z -scores formulae have recently been published (Duff, 2014), which makes them available for clinical use. We do not suggest that knowing someone's practice effects across 1 week will negate the need for an FDG PET if the latter is clinically indicated. Our findings and those of others (Boles Ponto et al., 2006; Duff et al., 2014; Machulda et al., 2014) only found associations between practice effects and biomarkers of AD and MCI. This should not be interpreted to mean that practice effects are ready to supplant any current and well-established biomarkers. But this is clearly an avenue for future investigation that may gain clinical relevance if these initial findings are supported.

Despite the interesting findings, some limitations of this study should be acknowledged. First, these results should be viewed as preliminary due to the small sample size. Larger studies with a wider range of cognitive functioning (including practice effects) would better test this hypothesis. Secondly, the sample was relatively homogeneous (e.g., all Caucasian, highly educated, mostly female, healthy enough to complete a PET scan), and the ability to generalize these findings to a more diverse group is unknown. Third, our sample was quite mild in their cognitive dysfunction, with most subjects being cognitively intact and a minority of them being classified as MCI. As noted earlier, it is unclear if these findings would replicate in more impaired samples. Studies suggest that multidomain MCI (e.g., impairments in memory and language) are at greater risk for progression to AD (Manly et al., 2008; Tabert et al., 2006). However, since clinical trials appear to be moving towards earlier points in the disease spectrum, there may be less relevance of similar findings in more advanced cases. Fourth, a global composite of the FDG PET was used in the current study for a number of reasons. However, this limits our ability to examine regional differences in hypometabolism and how they relate to short-term practice effects. Future studies should include larger numbers of subjects to allow for comparisons across different brain regions (e.g., frontal association cortex, temporal association cortex, anterior and posterior cingulate). Fifth, structural imaging was not part of this research protocol. However, the co-registration of FDG PET and structural MRI may better characterize disease states and predict progression (Musiek et al., 2012). Similarly, APOE status was not determined in this cohort, but prior work has found associations between this genetic marker of AD and practice effects (Machulda et al., 2014). Finally, it should be noted that FDG PET is a neuroimaging technique used to inform clinical practice about brain metabolism. The results of this neuroimaging procedure are not, in isolation, diagnostic, and they require correlation with clinical observation and evaluation. Regardless of these limitations, the current study found notable relationships between brain hypometabolism assessed with FDG PET and baseline cognitive test scores, and 1-week practice effects added to our understanding of brain metabolism in individuals with MCI. This growing body of findings supports the future examination of practice effects as a potential proxy measure in clinical trials in AD and MCI.

Funding

Funding for this project was provided by an anonymous foundation, NIH NIA K23 AG028417, the Center for Quantitative Cancer Imaging at the Huntsman Cancer Institute, and the University of Utah Center for Alzheimer's Care, Imaging and Research.

References

- Barr, W. B. (2002). Neuropsychological testing for assessment of treatment effects: Methodologic issues. *CNS Spectrums*, 7 (4), 300–302, 304–306.
- Beglinger, L. J., Tangphao-Daniels, O., Kareken, D. A., Zhang, L., Mohs, R., & Siemers, E. R. (2005). Neuropsychological test performance in healthy elderly volunteers before and after donepezil administration: A randomized, controlled study. *Journal of Clinical Psychopharmacology*, 25 (2), 159–165.
- Boles Ponto, L. L., Magnotta, V. A., Moser, D. J., Duff, K. M., & Schultz, S. K. (2006). Global cerebral blood flow in relation to cognitive performance and reserve in subjects with mild memory deficits. *Molecular Imaging and Biology*, 8 (6), 363–372.
- Calamia, M., Markon, K., & Tranel, D. (2012). Scoring higher the second time around: Meta-analyses of practice effects in neuropsychological assessment. *The Clinical Neuropsychologist*, 26 (4), 543–570.
- Chen, K., Ayutyanont, N., Langbaum, J. B., Fleisher, A. S., Reschke, C., Lee, W., et al. (2011). Characterizing Alzheimer's disease using a hypometabolic convergence index. *Neuroimage*, 56 (1), 52–60.
- Cooper, D. B., Epker, M., Lacritz, L., Weiner, M., Rosenberg, R. N., Honig, L., et al. (2001). Effects of practice on category fluency in Alzheimer's disease. *The Clinical Neuropsychologist*, 15 (1), 125–128.
- Cooper, D. B., Lacritz, L. H., Weiner, M. F., Rosenberg, R. N., & Cullum, C. M. (2004). Category fluency in mild cognitive impairment: Reduced effect of practice in test-retest conditions. *Alzheimer Disease and Associated Disorders*, 18 (3), 120–122.
- Darby, D., Maruff, P., Collie, A., & McStephen, M. (2002). Mild cognitive impairment can be detected by multiple assessments in a single day. *Neurology*, 59 (7), 1042–1046.

- Duff, K. (2012). Evidence-based indicators of neuropsychological change in the individual patient: Relevant concepts and methods. *Archives of Clinical Neuropsychology*, 27 (3), 248–261.
- Duff, K. (2014). One-week practice effects in older adults: Tools for assessing cognitive change. *The Clinical Neuropsychologist*, 28 (5), 714–725.
- Duff, K., Beglinger, L., Schultz, S., Moser, D., McCaffrey, R., Haase, R., et al. (2007). Practice effects in the prediction of long-term cognitive outcome in three patient samples: A novel prognostic index. *Archives of Clinical Neuropsychology*, 22 (1), 15–24.
- Duff, K., Beglinger, L. J., Moser, D. J., Schultz, S. K., & Paulsen, J. S. (2010). Practice effects and outcome of cognitive training: Preliminary evidence from a memory training course. *The American Journal of Geriatric Psychiatry*, 18 (1), 91.
- Duff, K., Beglinger, L. J., Van Der Heiden, S., Moser, D. J., Arndt, S., Schultz, S. K., et al. (2008). Short-term practice effects in amnesic mild cognitive impairment: Implications for diagnosis and treatment. *International Psychogeriatrics*, 20 (5), 986–999.
- Duff, K., Callister, C., Dennett, K., & Tometch, D. (2012). Practice effects: A unique cognitive variable. *The Clinical Neuropsychologist*, 26 (7), 1117–1127.
- Duff, K., Chelune, G., & Dennett, K. (2012). Within-session practice effects in patients referred for suspected dementia. *Dementia and Geriatric Cognitive Disorders*, 33 (4), 245–249.
- 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 Disease and Associated Disorders*, 28 (3), 247–252.
- Duff, K., Lyketsos, C. G., Beglinger, L. J., Chelune, G., Moser, D. J., Arndt, S., et al. (2011). Practice effects predict cognitive outcome in amnesic mild cognitive impairment. *The American Journal of Geriatric Psychiatry*, 19 (11), 932–939.
- Ewers, M., Brendel, M., Rizk-Jackson, A., Rominger, A., Bartenstein, P., Schuff, N., et al. (2013). Reduced FDG-PET brain metabolism and executive function predict clinical progression in elderly healthy subjects. *NeuroImage: Clinical*, 4, 45–52.
- Fernandez-Ballesteros, R., Botella, J., Zamarron, M. D., Molina, M. A., Cabras, E., Schettini, R., et al. (2012). Cognitive plasticity in normal and pathological aging. *Clinical Interventions in Aging*, 7, 15–25.
- Fernandez-Ballesteros, R., Zamarron, M. D., & Tarraga, L. (2005). Learning potential: A new method for assessing cognitive impairment. *International Psychogeriatrics*, 17 (1), 119–128.
- Habeck, C., Risacher, S., Lee, G. J., Glymour, M. M., Mormino, E., Mukherjee, S., et al. (2012). Relationship between baseline brain metabolism measured using [(1)(8)F]FDG PET and memory and executive function in prodromal and early Alzheimer's disease. *Brain Imaging and Behavior*, 6 (4), 568–583.
- Herholz, K., Salmon, E., Perani, D., Baron, J. C., Holthoff, V., Frolich, L., et al. (2002). Discrimination between Alzheimer dementia and controls by automated analysis of multicenter FDG PET. *Neuroimage*, 17 (1), 302–316.
- Hoffman, J. M., Welsh-Bohmer, K. A., Hanson, M., Crain, B., Hulette, C., Earl, N., et al. (2000). FDG PET imaging in patients with pathologically verified dementia. *Journal of Nuclear Medicine*, 41 (11), 1920–1928.
- Horacek, J., Dockery, C., Kopecek, M., Spaniel, F., Novak, T., Tislerova, B., et al. (2006). Regional brain metabolism as the predictor of performance on the Trail Making Test in schizophrenia. A 18FDG PET covariation study. *Neuroendocrinology Letters*, 27 (5), 587–594.
- Machulda, M. M., Pankratz, V. S., Christianson, T. J., Ivnik, R. J., Mielke, M. M., Roberts, R. O., et al. (2014). Practice effects and longitudinal cognitive change in normal aging vs. incident mild cognitive impairment and dementia in the Mayo Clinic Study of Aging. *The Clinical Neuropsychologist*, 27 (8), 1247–1264.
- Manly, J. J., Tang, M. X., Schupf, N., Stern, Y., Vonsattel, J. P., & Mayeux, R. (2008). Frequency and course of mild cognitive impairment in a multiethnic community. *Annals of Neurology*, 63 (4), 494–506.
- McCaffrey, R. J., & Westervelt, H. J. (1995). Issues associated with repeated neuropsychological assessments. *Neuropsychology Review*, 5 (3), 203–221.
- Minoshima, S., Frey, K. A., Koeppe, R. A., Foster, N. L., & Kuhl, D. E. (1995). A diagnostic approach in Alzheimer's disease using three-dimensional stereotactic surface projections of fluorine-18-FDG PET. *Journal of Nuclear Medicine*, 36 (7), 1238–1248.
- Musiek, E. S., Saboury, B., Mishra, S., Chen, Y., Reddin, J. S., Newberg, A. B., et al. (2012). Feasibility of estimation of brain volume and 2-deoxy-2-(18)F-fluoro-D-glucose metabolism using a novel automated image analysis method: Application in Alzheimer's disease. *Hellenic Journal of Nuclear Medicine*, 15 (3), 190–196.
- Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., & Kokmen, E. (1999). Mild cognitive impairment: Clinical characterization and outcome. *Archives of Neurology*, 56 (3), 303–308.
- Rapport, L. J., Brines, D. B., Axelrod, B. N., & Theisen, M. E. (1997). Full scale IQ as a mediator of practice effects: The rich get richer. *The Clinical Neuropsychologist*, 11 (4), 375–380.
- Suchy, Y., Kraybill, M. L., & Franchow, E. (2011). Practice effect and beyond: Reaction to novelty as an independent predictor of cognitive decline among older adults. *Journal of the International Neuropsychological Society*, 17 (1), 101–111.
- Tabert, M. H., Manly, J. J., Liu, X., Pelton, G. H., Rosenblum, S., Jacobs, M., et al. (2006). Neuropsychological prediction of conversion to Alzheimer disease in patients with mild cognitive impairment. *Archives of General Psychiatry*, 63 (8), 916–924.
- Temkin, N. R., Heaton, R. K., Grant, I., & Dikmen, S. S. (1999). Detecting significant change in neuropsychological test performance: a comparison of four models. *Journal of the International Neuropsychological Society*, 5 (4), 357–369.