Predicting Cognitive Change in Older Adults: The Relative Contribution of Practice Effects

Kevin Duff1,*, Leigh J. Beglinger2, David J. Moser2, Jane S. Paulsen2, Susan K. Schultz2, Stephan Arndt2

1Department of Neurology, University of Utah, Utah, USA
2Department of Psychiatry, University of Iowa, Utah, USA

*Corresponding author at: Center for Alzheimer’s Care, Imaging, and Research, Department of Neurology, University of Utah, 650 Komas Drive, #106-A, Salt Lake City, UT 84108, USA. Tel.: +1-801-581-9983; fax: +1-801-581-2483.
E-mail address: kevin.duff@hsc.utah.edu (K. Duff).

Accepted 8 December 2009

Abstract

Assessing cognitive change in older adults is a common use of neuropsychological services, and neuropsychologists have utilized several strategies to determine if a change is “real,” “reliable,” and “meaningful.” Although standardized regression-based (SRB) prediction formulas may be useful in determining change, SRBs have not been widely applied to older adults. The current study sought to develop SRB formulas on a group of 127 community-dwelling older adults for several widely used neuropsychological measures. In addition to baseline test scores and demographic information, the current study also examined the role of short-term practice effects in predicting test scores after 1 year. Consistent with prior research on younger adults, baseline test performances were the strongest predictors of future test performances, accounting for 25%–58% of the variance. Short-term practice effects significantly added to the predictability of all nine of the cognitive tests examined (3%–22%). Future studies should continue extending SRB methodology for older adults, and the inclusion of practice effects appears to add to the prediction of future cognition.

Keywords: Predicting cognition; Practice effects

Introduction

Assessing cognitive change over time is a common task for neuropsychologists working in geriatric settings. Declines in the cognitive abilities of elderly patients could occur due to several reasons, including progressive neurological conditions (e.g., Alzheimer’s disease), complications of surgical procedures (e.g., delirium following joint replacement, coronary artery bypass), or exacerbations of chronic medical conditions (e.g., hypothyroidism). Additionally, stability or improvements in cognitive abilities due to some intervention (e.g., medication, cognitive rehabilitation) is an equally important use of geriatric neuropsychological testing.

Clinicians have a variety of choices when determining if a “real” and “meaningful” change in cognition has occurred in their patients. For example, there are several statistical formulas that utilize a test’s reliability, in combination with documented practice effects, to develop confidence intervals of “normal” change (e.g., Reliable Change Index, Jacobson & Truax, 1991; Zegers & Hafkenscheid, 1994; practice-adjusted Reliable Change Index, Chelune, Naugle, Luders, Sedlak, & Awad, 1993; for a review of these formulas, see Bruggemans, Van de Vijver, & Huysmans, 1997, or McCaffrey, Duff, & Westervelt, 2000). Alternatively, standardized regression-based (SRB) formulas are another option for assessing if a “real” and “meaningful” change has occurred. SRBs, introduced by McSweeney, Naugle, Chelune, and Luders (1993), use multiple regression algorithms to predict follow-up test performances using demographic variables and baseline test performances. SRBs have some distinct advantages over other methods. For example, SRBs can account for baseline testing performance, regression to the mean, and demographic factors, whereas Reliable Change Indexes cannot. SRBs, however, also pose some unique challenges (e.g., relatively large samples are needed to develop the formulas). In general, SRBs have demonstrated greater sensitivity than...
the other methods of assessing change (Barr, 2002; Temkin, Heaton, Grant, & Dikmen, 1999), although this finding has not been entirely consistent (Heaton et al., 2001).

SRB methodology has been applied to a number of patient samples within neuropsychology, including epilepsy (Hermann et al., 1996; McSweeny et al., 1993; Sawrie, Chelune, Naugle, & Luders, 1996), cardiac conditions (Bruggemans et al., 1997), and concussion (Barr & McCrea, 2001). This same methodology has also been applied to samples of neurologically healthy adults (Temkin et al., 1999) to develop “normal” change algorithms. Across multiple studies using SRBs, baseline cognitive performance on a test has consistently been found to be the best predictor of follow-up test performance. Demographic variables (e.g., age, education) tend to slightly improve prediction accuracy. Unfortunately, SRB methodology has not been widely applied to geriatric settings. Most of the patients and controls in the aforementioned studies tended to be younger (e.g., <50 years old), which might limit the generalizability of findings to older adults.

Since this methodology was first introduced for geriatric patients (Sawrie, Marson, Boothe, & Harrell, 1999), only a handful of studies have specifically developed SRBs for use with this age range. Although each of these studies contributes to a scant literature on SRBs in older adults, each study also has notable limitations for their broader applicability in geriatric neuropsychology. For example, one study (Tombaugh, 2005) only used brief cognitive screening measures (e.g., Mini-Mental Status Examination) and other studies (Duff et al., 2004, 2005; Duff, Schoenberg, 2008; Raymond, Hinton-Bayre, Radel, Ray, & Marsh, 2006) used circumscribed batteries (i.e., MicroCog, Repeatable Battery for the Assessment of Neuropsychological Status). Two studies (Frerichs & Tuokko, 2005; Tombaugh, 2005) used very long retest intervals (i.e., 5 years), whereas another (Raymond et al., 2006) used very short intervals (i.e., 2 and 12 weeks). All of the studies focused on “normal” change by utilizing community-dwelling samples of older adults without cognitive impairments. Lastly, none of these studies examined the possible influence of short-term practice effects on SRBs. Recently, short-term practice effects have demonstrated some prognostic value (Duff et al., 2007), and these practice effects might add to the accuracy of SRBs (e.g., contributing above and beyond the baseline test scores).

The current study sought to extend the literature on SRBs in older adults by developing change formulas on a battery of commonly used neuropsychological measures. A clinically relevant retest interval of 1 year was employed. Both cognitively intact and mildly impaired participants were used to capture a wider range of cognitive change across time. Finally, short-term practice effects were considered as an additional predictor variable in the SRB models. It was hypothesized that baseline test scores would best predict 1-year test scores, and that short-term practice effects would significantly contribute to these prediction models.

**Materials and Methods**

**Participants**

One hundred and twenty-seven community-dwelling older adults participated in the current study, and these participants have been described previously (Duff, Beglinger, et al., 2008). Briefly, these individuals were recruited from senior centers and independent living facilities to prospectively study practice effects in older adults. Their mean age was 78.7 (7.8) years and their mean education was 15.5 (2.5) years. Most were women (81.1%) and all were Caucasian. Premorbid intellect at baseline was average (Wide Range Achievement Test-3 [WRAT-3] Reading: $M = 108.4$ [6.0]). To be classified as amnestic MCI, all participants had to complain of memory problems (i.e., self-reported as yes/no during an interview). These participants had objective memory deficits (i.e., age-corrected scores at or below the 7th percentile on two delayed recall measures [described below] relative to a premorbid intellectual estimate [WRAT-3 Reading]). The 7th percentile is 1.5 $SD$ below the mean, which is a typical demarcation point for cognitive deficits in MCI. Cognition was, otherwise, generally intact (i.e., non-memory age-corrected scores above the 7th percentile) and no functional impairments (e.g., assistance needed with managing money, taking medications, driving) could be reported. To be classified as “cognitively intact,” all objective memory and non-memory performances were at least above the 7th percentile. All data were reviewed by two neuropsychologists (KD and LJB). Fifty-three individuals (42% of sample) were classified with amnestic MCI according to existing criteria (Petersen et al., 1999), and the remainder were classified as cognitively normal. No one was classified as demented (i.e., both impaired memory and other cognitive domains). All classifications were made following the 1-week visit, so examiners were “blinded” to classification at the baseline and 1-week visits. However, only baseline cognitive performances were used in these classifications.

**Procedures**

All participants provided informed consent prior to participation, and all procedures were approved by the local Institutional Review Board. During a baseline visit, all participants completed the following measures: Brief Visuospatial Memory
Test-Revised (BVMT-R), Hopkins Verbal Learning Test-Revised (HVLT-R), Controlled Oral Word Association Test (COWAT), animal fluency, Trail Making Test Parts A and B (TMTA and TMTB), Symbol Digit Modalities Test (SDMT), WRAT-3 Reading subtest, and the 30-item Geriatric Depression Scale (GDS). After 1 week, the battery was repeated, with the exception of the WRAT-3 Reading subtest. After 1 year, the battery was again repeated, with the exception of the WRAT-3 Reading subtest. Alternate forms were purposefully not utilized on re-evaluation, as the study sought to maximize practice effects.

Data Analyses

WRAT-3 Reading scores were age-corrected standard scores using normative data from the test manual. All other values are raw scores.

A validity check on the classification of MCI or normal cognition was performed with two MANCOVAs on baseline cognitive scores. In the first MANCOVA, the two groups were compared on immediate and delayed recall measures from the battery (i.e., BVMT-R and HVLT-R). In the second MANCOVA, the two groups were compared on non-memory measures (i.e., COWAT, animals, TMTA, TMTB, and SDMT). If these individuals represented amnestic MCI and cognitively intact peers, then they should be different on the first MANCOVA (i.e., memory) and comparable on the second MANCOVA (i.e., non-memory). Since individuals with MCI can decline to dementia on follow-up, we also calculated two additional MANCOVAs on the 1-year scores, again comparing these two groups on memory and non-memory measures. In all these analyses, age and baseline WRAT-3 Reading scores were used as covariates, as the groups differed on both of these variables ($p < .05$). The groups were comparable on education, gender, and baseline GDS scores.

Practice effects scores were generated for nine cognitive variables from the repeated battery: BVMT-R Total learning across three trials, BVMT-R Delayed Recall, HVLT-R Total learning across three trials, HVLT-R Delayed Recall, COWAT total words across three 60 s trials, animal fluency total words across one 60 s trial, TMTA seconds to completion, TMTB seconds to completion, and SDMT total correct in 90 s. To generate practice effects scores, the baseline score was subtracted from the 1-week score. For example, the BVMT-R Delayed Recall practice effects score was BVMT-R Delayed Recall at 1-week (BVMT-R Delayed Recall at Baseline). No change from baseline to 1 week would be reflected in a practice effects score of approximately zero. A practice effects score with a positive sign (e.g., +21.5) would reflect an increase in scores from baseline to 1 week. A practice effects score with a negative sign (e.g., −12.3) would reflect a decrease in scores from baseline to 1 week. For most of the measures, a positively signed practice effect is expected (as this demonstrates improvement from baseline to 1 week); however, negatively signed practice effects would be expected on TMTA and TMTB (as this also demonstrates improvement across time). The magnitude of change from one assessment to another was measured with effect sizes (e.g., 1-week score − baseline score/SD of the difference score one week − baseline).

To assess if practice effects contributed to the prediction of 1-year follow-up test scores above and beyond baseline test scores, nine separate stepwise multiple regressions were calculated, one for each criterion variables (i.e., 1-year scores). The predictor variables in each model were age, education, gender, WRAT-3 Reading score, MCI status, baseline score, and practice effects score. For example, demographic variables, WRAT-3 Reading, MCI status, baseline HVLT-R Total, and HVLT-R Total practice effects score were regressed on 1-year HVLT-R Total. Age was coded as years old at baseline. Gender was coded as men = 0 and women = 1. Education was coded as years. MCI status was coded as MCI = 1, and cognitively normal = 0. The inclusion of these demographic variables has been empirically supported in prior work (Duff et al., 2004; McCaffrey et al., 2000; Rapport, Brines, Axelrod, & Theisen, 1997). It was decided to include both cognitively normal and impaired participants into the same regression models to increase the range of test scores, but also include MCI status as a predictor variable to see if practice effects differentially predicted follow-up scores in these two groups. The data were screened for univariate and multivariate outliers with box plots, standardized scores (not to exceed ±3.0), and Mahalanobis distances ($p < .001$). Linearity and multicollinearity were assessed with scatterplots and variance inflation factors (not to exceed 2.5). Normal probability plots were also examined for distribution of error.

Results

After controlling for age and baseline WRAT-3 Reading scores, the two subgroups of participants were comparable on all non-memory tests at the baseline assessment, multivariate $F(5,115) = 1.00, p = .42$. Consistent with existing criteria (Petersen et al., 1999), the amnestic MCI subgroup performed significantly below their healthy peers on all baseline tests of immediate and delayed memory, multivariate $F(4, 120) = 38.48, p < .001$. The results of these two MANCOVAs support the classification of participants as amnestic MCI or intact at baseline. This same pattern was observed with 1-year scores, with the two subgroups being comparable on the non-memory MANCOVA, multivariate $F(5, 116) = 0.63, p = .67$, but significantly
Table 1. Repeated measures on the entire sample

<table>
<thead>
<tr>
<th>Cognitive scores</th>
<th>Baseline</th>
<th>One week</th>
<th>Practice effects</th>
<th>One year</th>
<th>ES&lt;sub&gt;oneweek−baseline&lt;/sub&gt;</th>
<th>ES&lt;sub&gt;oneyear−baseline&lt;/sub&gt;</th>
<th>ES&lt;sub&gt;oneyear−oneweek&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>BVMT-R Total</td>
<td>14.8 (6.9)</td>
<td>23.3 (8.6)</td>
<td>8.5 (5.4)</td>
<td>15.9 (7.9)</td>
<td>1.57</td>
<td>0.21</td>
<td>−1.30</td>
</tr>
<tr>
<td>BVMT-R Delayed Recall</td>
<td>5.7 (3.4)</td>
<td>8.1 (3.1)</td>
<td>2.4 (2.2)</td>
<td>6.4 (3.3)</td>
<td>1.09</td>
<td>0.29</td>
<td>−0.71</td>
</tr>
<tr>
<td>HVLT-R Total</td>
<td>23.6 (5.5)</td>
<td>27.9 (5.8)</td>
<td>4.3 (4.0)</td>
<td>24.5 (6.1)</td>
<td>1.08</td>
<td>0.18</td>
<td>−0.63</td>
</tr>
<tr>
<td>HVLT-R Delayed Recall</td>
<td>6.9 (3.4)</td>
<td>9.0 (2.7)</td>
<td>2.1 (2.4)</td>
<td>7.5 (3.6)</td>
<td>0.88</td>
<td>0.21</td>
<td>−0.58</td>
</tr>
<tr>
<td>COWAT</td>
<td>38.9 (10.8)</td>
<td>39.8 (12.0)</td>
<td>1.0 (7.3)</td>
<td>39.7 (12.4)</td>
<td>0.12</td>
<td>0.09</td>
<td>−0.01</td>
</tr>
<tr>
<td>Animals</td>
<td>17.5 (5.3)</td>
<td>17.9 (5.3)</td>
<td>0.5 (4.0)</td>
<td>18.0 (6.0)</td>
<td>0.10</td>
<td>0.09</td>
<td>0.02</td>
</tr>
<tr>
<td>TMTA</td>
<td>43.6 (15.6)</td>
<td>39.6 (12.8)</td>
<td>−4.0 (9.9)</td>
<td>44.0 (19.2)</td>
<td>−0.40</td>
<td>−0.03</td>
<td>0.30</td>
</tr>
<tr>
<td>TMTB</td>
<td>113.3 (52.9)</td>
<td>100.7 (51.6)</td>
<td>−12.6 (38.1)</td>
<td>113.0 (67.1)</td>
<td>−0.33</td>
<td>−0.01</td>
<td>0.27</td>
</tr>
<tr>
<td>SDMT</td>
<td>40.6 (12.4)</td>
<td>42.3 (10.3)</td>
<td>1.7 (10.1)</td>
<td>40.1 (11.7)</td>
<td>0.32</td>
<td>−0.06</td>
<td>−0.25</td>
</tr>
<tr>
<td>GDS</td>
<td>4.0 (3.3)</td>
<td>4.0 (3.9)</td>
<td>n/a</td>
<td>4.7 (4.5)</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>

Notes: BVMT-R = Brief Visuospatial Memory Test-Revised; HVLT-R = Hopkins Verbal Learning Test-Revised; COWAT = Controlled Oral Word Association Test; TMTA = Trail Making Test Part A; TMTB = Trail Making Test Part B; SDMT = Symbol Digit Modalities Test; GDS = Geriatric Depression Scale; ES = effect size. Values represent means and standard deviations (in parentheses). All scores are raw scores. Practice effects scores are 1 week−baseline.

different on the memory MANCOVA, multivariate $F(4, 118) = 12.34, p < .001$. These results further support the classification of the participants, but also indicate that MCI subgroup did not significantly decline on non-memory measures across 1 year relative to the intact subgroup. Baseline, 1-week, and 1-year scores from the repeated measures are presented in Table 1. Practice effects scores are also presented in Table 1. Finally, effect sizes for changes in cognitive variables are also reported in Table 1.

All nine of the 1-year cognitive scores were significantly predicted by their respective models, and the stepwise regression models are presented in Table 2. Baseline scores significantly contributed to the prediction of One Year scores in all nine models, accounting for 25%–58% of the variance (i.e., top row of adjusted $R^2$ values in Table 2). Practice effects scores also significantly contributed to all nine of the models, adding an additional 3%–22% of the variance (i.e., difference between top and second rows of adjusted $R^2$ values in Table 2). For each cognitive score, the final model’s $F$, adjusted $R^2$, standard error of the estimate, constant, and unstandardized β weights for relevant variables are listed.

Discussion

In this study, a series of regression models were performed on a sample of nondemented, community-dwelling older adults to predict 1-year follow-up performances on a battery of commonly used neuropsychological measures. Consistent with the existing literature (Duff et al., 2004, 2005; Duff, Schoenberg, 2008; Hermann et al., 1996; McSweeny et al., 1993; Sawrie et al., 1996; Temkin et al., 1999), the best predictor of follow-up performance (i.e., 1-year scores) was initial performance (i.e., baseline scores) on that same measure. Across cognitive measures, baseline scores shared between 25% and 58% of the variance with 1-year scores (Duff et al., 2007; Howieson et al., 2008). In the present study, 1-week practice effects have routinely been viewed as error variance in retesting paradigms, recent research suggests that these improvements in test scores might have diagnostic and prognostic utility in neuropsychologically impaired older samples. Diagnostically, several researchers have observed that individuals with MCI tend to benefit less from practice than healthy peers (Cooper, Lacritz, Weiner, Rosenberg, & Cullum, 2004; Darby, Maruff, Collie, & McStephen, 2002; Duff, Beglinger, et al., 2008; Yan & Dick, 2006; Zehnder, Blasi, Berres, Spiegel, & Monsch, 2007). Prognostically, an absence of practice effects has been linked to eventual decline in MCI (Duff et al., 2007; Howieson et al., 2008). In the present study, 1-week practice effects on all nine of the cognitive variables examined significantly improved predictions of test scores at 1 year.

To our knowledge, this is the first study to examine the possible influence of short-term practice effects on SRBs. Although practice effects have routinely been viewed as error variance in retesting paradigms, recent research suggests that these improvements in test scores might have diagnostic and prognostic utility in neuropsychologically impaired older samples. Diagnostically, several researchers have observed that individuals with MCI tend to benefit less from practice than healthy peers (Cooper, Lacritz, Weinreb, Rosenberg, & Cullum, 2004; Darby, Maruff, Collie, & McStephen, 2002; Duff, Beglinger, et al., 2008; Yan & Dick, 2006; Zehnder, Blasi, Berres, Spiegel, & Monsch, 2007). Prognostically, an absence of practice effects has been linked to eventual decline in MCI (Duff et al., 2007; Howieson et al., 2008). In the present study, 1-week practice effects on all nine of the cognitive variables examined significantly improved predictions of test scores at 1 year.

The short-term practice effects used in the current study might allow clinicians and researchers to identify these “at-risk” individuals even sooner than reported previously (e.g., Howieson et al. examined practice effects across a 1-year interval). Practice effects also might have implications for interpreting the results of longitudinal studies (Salthouse & Tucker-Drob, 2008). Admittedly, the relative contribution of practice effects above and beyond baseline scores was small in these analyses (e.g., 3%–22% of shared variance, see the second row of $R^2$ values for each cognitive score in Table 2). The practice effects retest interval in the current study was 1 week, and future studies might investigate if shorter or longer retest intervals can lead to practice effects with greater contributions to prediction accuracy. For example, Attix and colleagues (2009) used change across 1 year to better determine cognitive trajectories across longer periods of time.
The combined effects of baseline scores and practice effects in predicting future cognition deserve some additional comment. As can be seen in the last three columns of Table 1, the magnitude of practice effects varies by test and retest intervals. Although others have commented on this fact (McCaffrey et al., 2000; Salthouse & Tucker-Drob, 2008), the largest differences occurred on measures of learning and memory. For example, the effect sizes for the BVMT-R and HVLT-R between baseline and 1 week averaged 1.15, whereas the effect sizes across this same interval for non-memory tests averaged 0.25. Similarly, the longer the retest interval, the smaller the practice effect (e.g., average effect size: Baseline and 1 week = 0.65, 1 week and 1 year = −0.45, baseline and 1 year = 0.13). This decreasing magnitude across time is probably also related to the number of assessments, as practice effects tend to decrease by the third assessment point for some tests (Beglinger et al., 2005).

The combined effects of baseline scores and practice effects in predicting future cognition deserve some additional comment. It should not be too surprising that baseline test scores predict future test scores, as the majority of cognitive abilities do not normally change a lot over the course of 1 or 2 years. For example, the average correlation between baseline and 1-year scores in the current study was 0.64, and the majority of 1-year scores are within a couple of points of their respective baseline
scores (Table 1). In this way, the baseline score provides a fair amount of information about an expected follow-up score. Practice effects, however, appear to provide some indication of expected change from that baseline level. Short-term improvements in test scores might suggest the presence of additional cognitive reserve or plasticity. An absence of practice effects or short-term declines (i.e., negative practice effects) might suggest neuropsychological dysfunction. Although these hypotheses require further study, practice effects seem to be another clinically relevant variable.

Prior SRB studies have observed that demographic variables (e.g., age, education, and gender) provide a small but statistically significant role in predicting the follow-up cognitive scores. For example, in their report on SRBs for the 12 subtests of the Repeatable Battery for the Assessment of Neuropsychological Status, Duff and colleagues (2005) found that age contributed to eight SRBs, education contributed to five, gender contributed to two, and race contributed to one. The results of the present study were quite different. As can be seen in Table 2, demographic variables only contributed to three of the nine models (i.e., gender contributed to TMTB, age contributed to HVLT-R Delayed Recall, and COWAT). It is possible that some restriction in the range of demographic variables in the current study lead to their exclusion from SRB models. For example, since all participants in the current study were Caucasian, race would not contribute to any of the models. However, other demographic variables in the current sample seemed to have sufficient variation (e.g., age: 65–96 years, education: 8–24 years, and 20.4% men). It is also possible that the variance captured by demographic variables in prior studies is now accounted for in the practice effects score.

Another aspect of the current study that warrants comment is the composition of the sample used to develop the SRBs. Prior studies have tended to use relatively homogeneous samples to generate SRBs. For example, in their original study on SRBs, McSweeny and colleagues (1993) used only seizure patients to develop change formulas. Conversely, Temkin and colleagues (1999) used only neurological healthy individuals to predict follow-up cognitive scores. The current study used both healthy elders and those classified with amnestic MCI. In some ways, these two subsamples do reflect a single group: Non-demented, community-dwelling elders. However, almost by definition, one group suffers from at least “mild” memory problems, whereas the other group does not. It was our intent to combine both the subsamples to increase the variability of cognitive scores, which increases the potential of developing SRBs that would be applicable across a broad segment of the older adult population. In a related vein, Heaton and colleagues (2001) have observed that SRBs and other change formula developed on healthy samples might be less applicable in clinical samples. In their work, the authors developed change formulas on healthy adults, but then examined their validity in patients with schizophrenia (who were presumed to be relatively stable). Fewer than expected numbers of these patients with schizophrenia were identified as “not changing” across time. Heaton and colleagues suggested that samples used to develop SRBs might include individuals who are neurological stable, but not necessarily cognitively normal, so that a wide range of baseline and follow-up scores are represented. Unwittingly, we might have achieved this directive, as our combined sample of amnestic MCI and healthy elders contained a broader range of cognitive functioning that might yield more accurate prediction formulas in clinical samples. Nonetheless, we also included MCI status as another variable in the regression models, which allows us to see if memory impairment might differentially affect retesting performances. In the current analyses, MCI status only contributed to one of the regression models, BVMT-R Total Recall. In this model, the negative β weight seems to indicate that being identified as amnestic MCI lowers the expected follow-up score on this measure of visual learning. The general lack of further cognitive decline in our MCI sample across 1 year (as indicated by the two MANCOVAs on 1-year scores) may have also “restricted the range” of this variable.

Despite the potential benefits of using SRBs (e.g., increased precision in the assessment of change), it should be noted that we are not advocating for a strictly psychometric assessment of any patient (i.e., based solely on test data). Complementing historical information, behavioral observations, and laboratory results are also vital pieces of clinical information. We are, however, attempting to provide the necessary psychometric information to assist in the clinical decision-making process. For those interested in utilizing this information, a copy of the computer program used to calculate the predicted scores, difference scores, and test of the differences’ significance can be obtained from the first author. It should also be noted that some findings suggest that SRBs are no better than other indexes of reliable change in patient samples (Heaton et al., 2001).

Some limitations with the present study are acknowledged. As with most regression-based prediction formulas (Tabachnick & Fidell, 1996), less accurate estimates are possible for individuals whose cognitive functioning falls at the extremes (e.g., <2nd percentile or >98th percentile) at baseline. In these cases, the prediction equations are more susceptible to regression to the mean and other fluctuations. However, the present study utilized a sample with both intact and impaired participants, which might lessen the chance of these statistical fluctuations. Caution should be exercised when using these formulas outside the demographic and situational parameters of the sample (e.g., <64 or >96 years old; relatively brief or extended retest intervals; non-Caucasians). Since all subjects in this study were evaluated at baseline, 1 week, and 1 year, it is unknown how accurate these SRBs would be for a patient who did not have the 1-week assessment. Finally, the stability of the regression equations needs to be validated in an independent sample, as the reliability of practice effects, especially in impaired samples, is not known.
In conclusion, the present SRB algorithms have the potential to provide more accurate assessments of cognitive change in older adults by considering the influence of initial performance, practice effects, and other demographic factors. These equations were developed on measures widely used in neuropsychological practice. These formulas also, for the first time, specifically utilize the short-term practice effects, which appear to lead to more accurate predictions of follow-up cognitive scores. Although validation of the effectiveness of these formulas in clinical samples is needed, they have the potential to contribute to the clinical decision-making process.

Funding

The project described was supported a research grant (R03 AG025850-01 and K23 AG028417-01A2) from the National Institute on Aging. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Aging or the National Institutes of Health.

Conflict of Interest

None declared.

References


