Received: 13 March 2024 | Revised: 15 July 2024 | Accepted: 17 July 2024

DOI: 10.1002/dad2.70008

Alzheimer's & Dementia
Diagnosis, Assessment & Disease Monitoring

RESEARCH ARTICLE

Validation of regression-based change formulae for mild cognitive impairment and Alzheimer's disease

Kevin Duff PhD[1] | **Deborah Sevigny-Resetco**[2]

[1]Layton Aging & Alzheimer's Disease Center, Department of Neurology, Oregon Health & Science University, Portland, Oregon, USA

[2]Department of Psychiatry, Department of Behavioral Neuroscience, Oregon Health & Science University, Portland, Oregon, USA

Correspondence
Kevin Duff, Layton Aging & Alzheimer's Disease Center, Oregon Health & Science University, 3181 SW Sam Jackson Park Road (Mail Code: CR131), Portland, OR 97239, USA.
Email: duffk@ohsu.edu

Funding information
National Institute on Aging; National Institute on Alcohol Abuse and Alcoholism, Grant/Award Numbers: P30AG66518, T32-AA007468; National Institutes of Health; NIA/NIH, Grant/Award Number: U24 AG072122; ADRCs, Grant/Award Numbers: P30 AG062429, P30 AG066468, P30 AG062421, P30 AG066509, P30 AG066514, P30 AG066530, P30 AG066507, P30 AG066444, P30 AG066518, P30 AG066512, P30 AG066462, P30 AG072979, P30 AG072972, P30 AG072976, P30 AG072975, P30 AG072978, P30 AG072977, P30 AG066519, P30 AG062677, P30 AG079280, P30 AG062422, P30 AG066511, P30 AG072946, P30 AG062715, P30 AG072973, P30 AG066506, P30 AG066508, P30 AG066515, P30 AG072947, P30 AG072931, P30 AG066546, P20 AG068024, P20 AG068053, P20 AG068077, P20 AG068082, P30 AG072958, P30 AG072959

Abstract

INTRODUCTION: Identification of cognitive decline is critical in older adults at risk for dementia. In a 2020 study reported in *Archives of Clinical Neuropsychology*, Kiselica and colleagues developed standardized regression-based (SRB) change formulae for the Uniform Data Set 3.0 Neuropsychological Battery in cognitively unimpaired older adults. However, validation of their applicability in impaired individuals is needed.

METHODS: Using longitudinal data on 5974 participants (cognitively unimpaired, mild cognitive impairment, dementia) from the National Alzheimer's Coordinating Center, SRB change scores were calculated for each individual and compared across groups.

RESULTS: Across 6 to 24 months, minimal cognitive change was observed in cognitively unimpaired participants. Modest declines were seen in those with mild cognitive impairment and substantial declines in those with dementia. Change scores were negatively correlated with the Clinical Dementia Rating scale. In impaired individuals, SRB scores indicated more decline in those with positive amyloid scans.

DISCUSSION: Validation of SRB scores affords greater confidence in employing them in clinical and research settings.

KEYWORDS
change, dementia, longitudinal, mild cognitive impairment, neuropsychology, normative data

Highlights

- Validation of regression-based cognitive change scores in impaired samples.
- Clear differences on change scores across three groups (intact, MCI, dementia).
- Largely stable scores in intact participants, but notable decline in MCI and dementia.
- Moderate to strong relationship between change scores and the Clinical Dementia Rating scale sum of boxes.

1 | BACKGROUND

The identification of cognitive change is particularly relevant in older adults at risk for dementia. Clinically, cognitive decline can suggest the development of a cognitive disorder or progression of a disease state (eg, mild cognitive impairment [MCI] to dementia). In research settings, change can help evaluate the effects of interventions in this population, including identifying stability for disease-modifying agents. As

such, access to empirically validated methods of quantifying cognitive change is needed.

One method of quantifying cognitive change is the standardized regression-based (SRB) change formulae,[1] in which an individual's follow-up (ie, Time 2) cognitive test score is predicted from his/her initial (ie, Time 1) score and other relevant variables (eg, age, education, retest interval). Notable discrepancies between observed and predicted Time 2 scores can indicate decline (if sufficiently negative) or improvement (if sufficiently positive). Such SRB models have provided insights into diagnosis,[2] prognosis,[3–5] treatment response,[6] and brain pathology[7–10] in older adults with MCI and dementia.

Using a large sample of cognitively unimpaired older adults from the National Alzheimer's Coordinating Center (NACC) database, Kiselica et al.[11] developed SRB change formulae (hereafter referred to as SRBs) for the Uniform Data Set 3.0 Neuropsychological Battery to assess change across 6 to 24 months. Following the SRB model of McSweeny et al.,[1] these SRBs accounted for 14% to 62% of the variance of the Time 2 cognitive scores, using predictors of the respective Time 1 score, age, sex, education, race, and retest interval. These SRB change scores were preliminarily validated on a separate, randomly selected group of cognitively unimpaired older adults by (1) examining base rates of change using different cutoffs (eg, −1.645 vs −1.96), (2) examining correlations of change scores on similar tests (eg, animal vs vegetable fluency), and (3) examining correlations of change scores with a measure of cognitive and daily functioning. Although these analyses did demonstrate that the SRB change scores performed as expected in this cognitively unimpaired sample, further validation in a clinical sample is needed to give clinicians and researchers more confidence in the value and generalizability of these formulae.

The current project sought to further validate the SRBs developed by Kiselica et al.[11] by applying them to a more recent version of the NACC database and examining their results in those judged to be cognitively unimpaired or who had MCI or dementia. It was hypothesized that individuals classified as having dementia at Time 1 would have the largest amount of decline via SRBs, the cognitively unimpaired participants would have the least decline, and the MCI participants would fall between the other two groups. Secondary analyses comprised the following: (1) The base rates of improvement, stability, and decline were examined in these groups, with more decline expected in the MCI and dementia groups; (2) SRBs were correlated with an overall measure of cognition and daily functioning, the Clinical Demetia Rating (CDR) scale, where these two variables were expected to be negatively correlated; and (3) SRB z-scores were compared in those individuals who had amyloid positron emission tomography (PET) scan results, with amyloid positive individuals expected to have worse/lower SRBs than their amyloid negative counterparts. If further validated in clinical groups, then the application of SRBs might offer more sensitive metrics for tracking change in clinical settings and clinical trials.

RESEARCH IN CONTEXT

1. **Systematic review**: The authors reviewed the literature using traditional (eg, PubMed) sources. While regression-based cognitive change has not been widely studied, a relatively recent paper by Kiselica et al. (2020) was the impetus for this work. Relevant citations are appropriately cited.
2. **Interpretation**: Our findings demonstrate that the regression-based change scores developed by Kiselica et al. (2020) provide useful information about cognitive change in impaired samples. Such information allows clinicians and researchers to more confidently employ these methods.
3. **Future directions**: Although the manuscript validates the change scores of Kiselica et al. (2020), there is need for additional studies to expand this work. Examples include further understanding of (1) the role of shorter and longer retest intervals in determining cognitive change; (2) the impact of non-English test administration on cognitive change; and (3) the applicability of these change scores to younger, more ethnically diverse, and more etiologically diverse samples.

2 | METHODS

2.1 | Participants

The sample comprised 5974 older adults drawn from the NACC database, which was from the June 2023 data freeze. Of the 48,605 individuals in the database, 1190 below the age of 50 were excluded. Additionally, 4062 were excluded who did not complete the visit in English; 15,976 were excluded with a follow-up visit of less than 6 months or more than 2 years from their baseline visit; 19,840 were excluded because they did not complete the Uniform Data Set Version 3.0 Neuropsychological Battery at a baseline visit (Time 1) and the next follow-up visit (Time 2); 283 were excluded because they did not have a contemporaneous CDR score; and 1280 were excluded because they were in the original sample in Kiselica et al.[11] Of the 5974 individuals included in this sample, 769 had amyloid PET results for the secondary analyses.

2.2 | Measures

The cognitive tests in the NACC Uniform Data Set 3.0 Neuropsychological Battery,[12] for which Kiselica et al.[11] developed SRBs, include Montreal Cognitive Assessment (MoCA)[13]; Trail Making Test, Parts A and B[14]; Craft Story Immediate and Delayed Recall[15]; Benson Figure

Copy and Recall[16]; Category Fluency (animals and vegetables); Phonemic Fluency (sum of F-words and L-words); Number Span (Forward and Backward); and Multilingual Naming Test (MiNT).[17,18] These tests, which assess global cognition, attention and processing speed, visuospatial abilities, language, memory, and executive functioning, have been described in detail elsewhere.[12,19] For each test, higher scores indicate better cognition, with the exception of the Trail Making Tests, where lower scores indicate better cognition. Raw cognitive scores were used for analyses. Due to unknown factors, not all participants had scores for each test, but all available cognitive data were used.

Cognitive and functional impairment was assessed with the CDR,[20] which rates memory, orientation, judgment and problem-solving, community affairs, home and hobbies, and personal care, using input from the participant and a collateral source. The sum of the box scores was used for analyses, with higher scores indicating more cognitive and functional impairment.

Within the Uniform Data Set 3.0, clinicians indicated if participants had a result for an amyloid PET scan, and if so, whether it was positive or negative, according to local standards. No additional details about the scan acquisition, processing, or quantified values were available.

2.3 | Procedures

As a national repository of data, the NACC database has been approved by the appropriate institutional research ethics committees and certify that studies were performed in accordance with ethical standards as laid down in the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards.

Participants were categorized as cognitively unimpaired, MCI, or dementia based on the opinion of a single clinician or consensus group from NACC using participant/collateral report, CDR, and cognitive test scores. It is unclear the extent to which SRBs were employed in clinical consensus diagnosis at Alzheimer's Disease Research Centers, but it is not believed that they were routinely used.

2.4 | Data analyses

The SRBs of Kiselica et al.[11] were applied to the 5974 cases in the current dataset. These formulae predicted an individual's Time 2 raw score from the following predictors: respective Time 1 raw score, age (in years), sex (coded as male = 1 or female = 2), education (in years), race (coded as White = 1 or non-White = 2), and retest interval (in days). The predicted Time 2 score was subtracted from the observed Time 2 score, and the difference was divided by the standard error of the estimate (SEE) in Kiselica's original SRBs (ie, [Time $2_{observed}$ − Time $2_{predicted}$]/SEE). This process yielded an SRB z-score for each of the cognitive scores obtained for every individual, with negative values indicating that the individual showed more decline relative to the cognitively unimpaired participants in Kiselica's sample and positive values indicating more improvement relative to that sample. Following convention,[21] these z-scores were trichotomized as "decline" being <−1.645, "stable" being −1.644 to 1.644, or "improvement" being >1.645.

For the primary analysis, the three groups (cognitively unimpaired, MCI, dementia) were compared on the SRB z-scores for each of the 13 cognitive test scores with analysis of variance (ANOVA) tests. Post hoc analyses were made with Bonferroni corrections. Secondary analyses comprised the following: (1) The three groups were compared on the trichotomized SRB z-scores (decline, stable, improve) with chi-square tests; (2) SRB z-scores were correlated with the CDR sum of boxes with Spearman correlations, as CDR values tended to be skewed; and (3) SRB z-scores were compared between individuals with positive versus negative amyloid PET scan results with independent *t*-tests by group. In all analyses, demographic variables were not included as covariates, as these variables are included in the prediction of Time 2 scores in Kiselica et al.[11] An alpha level of 0.05 was used for analyses.

3 | RESULTS

3.1 | Preliminary comparisons between groups

Demographic information for each group is presented in Table 1. There were significant age differences by group, MCI > cognitively unimpaired > dementia ($F(2,5971) = 35.95$, $p < 0.001$). Education differences were also identified by group, cognitively unimpaired > MCI > dementia ($F(2,5934) = 51.12$, $p < 0.001$). Regarding sex, the cognitively unimpaired group had significantly more females than males, and the dementia and MCI groups had more males than females ($\chi^2(2) = 158.05$, $p < 0.001$). Each group had significantly more White-identifying participants than non-White participants ($\chi^2(2) = 115.94$, $p < 0.001$). The mean retest interval between Time 1 and Time 2 was significantly different ($F(2,5971) = 13.54$, $p < 0.001$), with the cognitively unimpaired group having the longest interval. Not surprisingly, the three groups differed on the CDR sum of boxes at Time 1 ($F(2,5971) = 3542.20$, $p < 0.001$) and Time 2 ($F(2,5971) = 3899.04$, $p < 0.001$) in the expected direction (dementia > MCI > cognitively unimpaired; Table 1). Similarly, Time 1 and Time 2 cognitive test scores (see Table 2) showed the expected significant differences between the groups (cognitively unimpaired > MCI > dementia). After the SRBs of Kiselica et al.[11] were applied to the entire sample, the Predicted Time 2 scores (see Table 2) also showed the expected significant group differences.

3.2 | Primary analyses comparing groups on SRB z-scores

As displayed in Table 3, the cognitively unimpaired group showed minimal change from Time 1 to Time 2, as demonstrated by their mean SRB z-scores near 0 (−0.14 to 0.12). Conversely, those in the MCI group

TABLE 1 Demographics.

	Development sample	Cognitively unimpaired	MCI	Dementia
N	627	2255	1855	1864
Age***	69.98 (7.69)	70.29 (7.44)[a]	71.69 (8.00)[b]	69.45 (9.16)[c]
Education (years)***	16.50 (2.43)	16.54 (2.44)[a]	16.23 (2.69)[b]	15.71 (2.78)[c]
Sex (% male)***	35.40%	35.60%	52.00%	52.60%
Race (% White)***	76.60%	78.90%	82.80%	91.20%
Retest interval (days)***	423.32 (84.18)	430.24 (87.31)[a]	421.49 (82.08)[b]	416.94 (80.95)[b]
CDR T1***	0.01 (0.09)	0.16 (0.39)[a]	1.40 (1.05)[b]	5.14 (3.28)[c]
CDR T2***	0.02 (0.09)	0.20 (0.49)[a]	1.89 (1.66)[b]	7.09 (4.24)[c]

Note: Each subscript letter denotes values significantly different from one another. Race and sex assessed with Pearson's χ^2; age, education, retest interval, CDR T1 and T2 assessed with ANOVAs.
Abbreviations: ANOVA, analysis of variance; CDR, clinical dementia rating scale; MCI, mild cognitive impairment; T1, Time 1; T2, Time 2.
***$p < 0.001$.

showed a moderate amount of decline, with mean SRB z-scores ranging from −1.07 to −0.15. Those classified as having dementia demonstrated the most decline across time, with mean SRB z-scores ranging from −3.35 to −0.64. These differences were statistically significant across groups (eg, all p-values < 0.001, η^2 ranged from 0.08 to 0.41).

3.2.1 | Secondary analyses of trichotomized SRB z-scores

When the three groups were compared on trichotomized SRB z-scores (decline: <1.645; stable: −1.645 to +1.645; improvement: >1.645) with chi-square tests, significant differences were observed for each test score (all p-values < 0.001, see Table 4). Using these cutoffs, normally distributed scores should yield 5% decline, 90% stable, and 5% improvement, which were the results for the cognitively unimpaired participants. However, those with MCI showed 1 to 6 times as many decliners as expected compared to Kiselica's cohort. Those with dementia showed 3 to 10 times as many decliners as expected.

3.2.2 | Secondary analyses of correlations of SRB z-scores with the CDR sum of boxes

The correlations between the SRB z-scores and the CDR sum of boxes for all participants are seen in Table 5. Consistent with expectations, all correlations are negative, indicating that lower/more negative SRB z-scores (ie, more cognitive decline) are associated with higher CDR sum of boxes scores (ie, worse cognitive and daily functioning). The largest correlations between SRB z-scores and CDR scores were for the MoCA, and the smallest correlations were for Number Span Forward. The average correlation between the SRB z-scores and the Time 2 CDR was slightly but significantly higher than the average correlation between the SRB z-scores and the Time 1 CDR ($r_{CDR1} = -0.42$, $r_{CDR2} = -0.46$, $z = -2.71$, $p = 0.003$).

3.2.3 | Secondary analyses of SRB z-scores by amyloid status

In the cognitively unimpaired group, 215 were reported to be amyloid negative and 56 amyloid positive. Consistent with the primary analysis, there were no significant differences between the amyloid positive and negative cognitively unimpaired individuals on any of the 13 SRB z-scores (all p-values > 0.05; see Table S1).

In the MCI group, there were 100 amyloid negative individuals and 166 amyloid positive. These two subgroups differed on five of the 13 SRB z-scores: MoCA ($p < 0.001$), Benson Figure Recall ($p < 0.001$), Category Fluency for vegetables ($p < 0.001$), Craft Story Immediate Recall ($p = 0.005$), and Craft Story Delayed Recall ($p < 0.001$). For each of these differences, those reported to be amyloid positive had significantly lower SRB z-scores (see Table S2).

In the dementia group, 50 amyloid negative individuals were compared to 182 amyloid positive individuals. These two subgroups differed on six of the 13 SRB z-scores: MoCA ($p = 0.015$), Benson Figure Copy ($p = 0.014$), Benson Figure Recall ($p < 0.001$), Trail Making Test, Part A ($p = 0.006$), Craft Story Immediate Recall ($p < 0.001$), and Craft Story Delayed Recall ($p < 0.001$). For each of these differences, those reported to be amyloid positive had significantly lower SRB z-scores (see Table S3).

4 | DISCUSSION

The current project sought to further validate the SRBs developed by Kiselica et al.[11] by applying them to a more recent version of the NACC database and examining their suitability with individuals who were judged to be cognitively unimpaired, as well as extending analyses to individuals that had MCI or dementia. Consistent with the primary hypothesis, individuals classified as having dementia at Time 1 had the largest amount of cognitive decline at Time 2, as seen through the SRBs of Kiselica et al.[11] The cognitively unimpaired participants

TABLE 2 Time 1, Time 2, and predicted Time 2 scores by group.

	Cognitively unimpaired			MCI			Dementia		
	Time			Time			Time		
	1	2	Predicted 2	1	2	Predicted 2	1	2	Predicted 2
MoCA	26.36 (2.53)	26.33 (2.70)	26.26 (1.72)	22.53 (3.50)	21.9 (4.20)	24.06 (2.17)	16.04 (5.98)	14.12 (6.66)	20.58 (3.41)
Number Span Forward	8.35 (2.31)	8.35 (2.34)	8.16 (1.53)	7.55 (2.28)	7.40 (2.27)	7.69 (1.52)	6.47 (2.34)	6.09 (2.44)	7.06 (1.51)
Number Span Backward	8.00 (2.14)	8.00 (2.16)	8.00 (1.33)	5.90 (2.04)	5.86 (2.10)	6.63 (1.27)	4.31 (2.20)	4.03 (2.31)	5.82 (1.29)
Trails A	31.52 (11.52)	32.04 (12.09)	30.77 (7.33)	40.63 (20.03)	43.75 (24.92)	35.89 (11.51)	65.68 (39.97)	72.27 (43.92)	48.76 (21.88)
Trails B	81.33 (37.94)	84.52 (43.89)	84.64 (26.72)	126.4 (71.03)	131.24 (73.65)	109.41 (42.89)	185.63 (91.56)	187.64 (92.77)	138.61 (51.74)
MiNT	30.2 (2.05)	30.26 (2.15)	30.44 (1.64)	28.58 (3.37)	28.30 (3.94)	29.24 (2.58)	24.00 (6.69)	24.00 (7.85)	26.00 (4.93)
Category Fluency—animals	21.65 (5.42)	21.23 (5.44)	21.44 (3.77)	16.92 (5.16)	16.11 (5.35)	18.39 (3.54)	11.55 (5.56)	10.34 (5.89)	15.15 (3.59)
Category Fluency—vegetables	15.06 (4.18)	14.77 (4.14)	15.06 (2.54)	11.13 (3.84)	10.54 (4.07)	12.77 (2.29)	7.05 (3.97)	6.24 (4.17)	10.73 (2.24)
Phonemic Fluency (F and L)	28.69 (7.90)	29.02 (8.06)	29.52 (6.21)	24.21 (8.75)	23.99 (8.65)	26.10 (6.89)	18.3 (9.20)	17.06 (9.33)	21.78 (7.07)
Craft Story Immediate Recall	22.22 (6.20)	22.4 (6.33)	22.54 (3.51)	15.42 (6.81)	15.17 (7.06)	18.81 (3.76)	9.18 (6.36)	8.28 (6.63)	15.53 (3.49)
Craft Story Delayed Recall	19.35 (6.28)	19.79 (6.40)	19.92 (3.79)	11.25 (7.23)	11.00 (7.71)	15.25 (4.23)	5.20 (6.19)	4.41 (6.17)	11.89 (3.63)
Benson Figure Copy	15.52 (1.27)	15.50 (1.33)	15.64 (0.48)	14.95 (1.97)	14.72 (2.14)	15.48 (0.64)	12.93 (4.33)	11.87 (5.01)	15.00 (1.23)
Benson Figure Recall	11.38 (2.75)	11.38 (2.89)	11.55 (1.64)	7.38 (4.00)	7.15 (4.35)	9.37 (2.28)	3.98 (4.25)	3.33 (4.14)	7.62 (2.47)

Note. In each cell, values are presented as mean (standard deviation).
Abbreviations: MCI, mild cognitive impairment; MiNT , Multilingual Naming Test; MoCA, Montreal Cognitive Assessment.

TABLE 3 Z-scores by group.

	Cognitively unimpaired	MCI	Dementia	Analysis (ANOVA)
MoCA	0.05 (1.04)	−1.07 (1.52)	−3.35 (2.26)	$F(2,3914) = 1380.12, \eta^2 = 0.41^{***}$
Number Span Forward	0.12 (0.98)	−0.15 (1.00)	−0.64 (1.07)	$F(2,4680) = 208.83, \eta^2 = 0.26^{***}$
Number Span Backward	−0.01 (0.95)	−0.44 (1.01)	−1.05 (1.02)	$F(2,4660) = 397.54, \eta^2 = 0.08^{***}$
Trails A	−0.12 (1.00)	−0.88 (2.13)	−2.91 (3.60)	$F(2,3934) = 450.12, \eta^2 = 0.19^{***}$
Trails B	0.01 (0.97)	−0.67 (1.66)	−1.75 (2.17)	$F(2,3418) = 297.60, \eta^2 = 0.15^{***}$
MiNT	−0.11 (1.18)	−0.78 (1.97)	−2.75 (3.66)	$F(2,3815) = 393.56, \eta^2 = 0.15^{***}$
Category Fluency—animals	−0.04 (1.02)	−0.58 (1.00)	−1.30 (1.03)	$F(2,4918) = 589.14, \eta^2 = 0.19^{***}$
Category Fluency—vegetables	−0.08 (1.01)	−0.69 (1.00)	−1.45 (0.96)	$F(2,4862) = 719.07, \eta^2 = 0.23^{***}$
Phonemic Fluency (F and L)	−0.08 (0.99)	−0.39 (1.06)	−0.99 (1.10)	$F(2,4599) = 273.36, \eta^2 = 0.17^{***}$
Craft story Immediate Recall	−0.01 (0.97)	−0.68 (1.06)	−1.44 (0.98)	$F(2,4598) = 724.73, \eta^2 = 0.15^{***}$
Craft Story Delayed Recall	−0.01 (0.93)	−0.77 (1.02)	−1.41 (0.81)	$F(2,4538) = 805.88, \eta^2 = 0.24^{***}$
Benson Figure Copy	−0.14 (1.02)	−0.61 (1.57)	−2.61 (3.49)	$F(2,3860) = 433.82, \eta^2 = 0.18^{***}$
Benson Figure Recall	−0.06 (0.98)	−0.94 (1.35)	−1.86 (1.19)	$F(2,3779) = 697.75, \eta^2 = 0.27^{***}$

Note: In each cell values are presented as mean (standard deviation).
Abbreviations: ANOVA, analysis of variance; MCI, mild cognitive impairment; MiNT, Multilingual Naming Test; MoCA, Montreal Cognitive Assessment.
$^{***}p < 0.001$.

TABLE 4 Chi-square analyses for groups using trichotomized SRB z-cores by group.

	Cognitively unimpaired	MCI	Dementia	Analysis (χ^2)
MoCA	6/90/4	32/66/2	75/25/0	$\chi^2(4) = 1335.28$, Cramer's V = 0.41***
Number Span Forward	3/91/6	5/91/4	16/83/1	$\chi^2(4) = 218.31$, Cramer's V = 0.15***
Number Span Backward	4/92/4	11/87/3	27/72/1	$\chi^2(4) = 381.33$, Cramer's V = 0.20***
Trails A	7/92/1	20/79/1	48/50/2	$\chi^2(4) = 633.74$, Cramer's V = 0.28***
Trails B	5/94/1	21/76/3	47/50/3	$\chi^2(4) = 523.82$, Cramer's V = 0.28***
MiNT	6/91/3	22/75/4	51/46/3	$\chi^2(4) = 658.60$, Cramer's V = 0.29***
Category Fluency—animals	5/90/5	13/85/2	41/59/1	$\chi^2(4) = 755.77$, Cramer's V = 0.28***
Category Fluency—vegetables	5/91/4	16/82/1	45/54/1	$\chi^2(4) = 838.35$, Cramer's V = 0.29***
Phonemic Fluency (F and L)	5/91/4	10/87/3	27/72/1	$\chi^2(4) = 357.35$, Cramer's V = 0.20***
Craft Story Immediate Recall	5/91/4	18/80/2	47/53/1	$\chi^2(4) = 795.13$, Cramer's V = 0.29***
Craft Story Delayed Recall	4/92/4	21/78/1	46/54/0	$\chi^2(4) = 779.78$, Cramer's V = 0.29***
Benson Figure Copy	8/92/1	17/82/1	43/56/1	$\chi^2(4) = 490.88$, Cramer's V = 0.25***
Benson Figure Recall	6/91/3	31/68/2	66/34/0	$\chi^2(4) = 994.12$, Cramer's V = 0.36***

Note: In each cell, the percentages of cases that decline/remained stable/improved are reported. Some cells may add up to more than 100% due to rounding.
Abbreviations: MiNT, Multilingual Naming Test; MoCA, Montreal Cognitive Assessment; SRB, standardized regression-based.
$^{***}p < 0.001$.

demonstrated minimal decline on cognitive tests between visits, with mean SRB z-scores being very close to 0. Those classified with MCI had mean SRB z-scores of approximately −0.5, or about half of a standard deviation unit of cognitive decline compared to Kiselica's original sample. Those classified as having dementia at Time 1 had SRB z-scores in the −1 to −2 range, which is 1 to 2 standard deviation units below Kiselica's sample. Furthermore, when the amount of cognitive change was trichotomized as decline/stable/improve, the percentage of cases falling into each of these three categories mirrored the primary findings: very few cases classified as having dementia and MCI improved over time, and most declined, which supports the notion that an absence of practice effects is an indicator of late life cognitive dysfunction.[22,23] Conversely, those classified as cognitively unimpaired at Time 1 tended to show the expected distribution of 5% declining, 90% remaining stable, and 5% improving by Time 2. These results provide additional

TABLE 5 CDR correlations by cognitive test.

	CDR Time 1	CDR Time 2
MoCA	−0.63***	−0.68***
Number Span Forward	−0.26***	−0.28***
Number Span Backward	−0.37***	−0.40***
Trails A	−0.45***	−0.51***
Trails B	−0.37***	−0.44***
MiNT	−0.36***	−0.40***
Category Fluency—animals	−0.43***	−0.47***
Category Fluency—vegetables	−0.47***	−0.51***
Phonemic Fluency (F and L)	−0.31***	−0.34***
Craft Story Immediate Recall	−0.47***	−0.53***
Craft Story Delayed Recall	−0.51***	−0.58***
Benson Figure Copy	−0.34***	−0.36***
Benson Figure Recall	−0.53***	−0.57***

Abbreviations: CDR, Clinical Demetia Rating; MiNT, Multilingual Naming Test; MoCA, Montreal Cognitive Assessment.
***$p < 0.001$.

validation of the neuropsychological change scores presented in Kiselica et al.[11] and support their application in clinical and research settings and clinical trials.

These results are consistent with other studies that examined SRBs in cognitively unimpaired and impaired samples. For example, Lin et al.[24] developed SRBs on a subset of the Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort. Reliable decline was much more prevalent in those with late MCI and Alzheimer's disease compared to those with early MCI, subjective cognitive impairment, or no cognitive impairment. In another ADNI sample, more decline across 6 months (as measured with SRBs) was associated with an increased risk of Alzheimer's disease after 6 years in individuals with amnestic MCI.[4] Therefore, these SRB models appear appropriate for capturing the expected decline in older adults with cognitive impairments.

Ceiling effects and regression to the mean likely contribute to differences in predicted Time 2 scores across the groups.[25] For example, with Time 1 scores that were closer to the maximum value, the cognitively unimpaired participants had less room to improve at Time 2 (ie, ceiling effects). Additionally, with more extreme scores at Time 1, the MCI and dementia participants were more likely to "regress" to mean values of the entire sample, leading to higher Predicted Time 2 scores. Since there are many (sometimes competing) factors involved in the assessment of cognitive change, the quantification of change with SRB methodology can clarify this complex phenomenon.

As additional validation of the SRB formulae of Kiselica et al.,[11] associations between change scores and an overall measure of cognition and daily functioning, the CDR sum of boxes, was examined. Consistent with expectations, strong correlations between these two sets of scores were seen, with worsening cognition (ie, more negative SRB z-scores) being associated with reports of worsening cognitive abilities and daily functioning (ie, increasing sum of boxes on the CDR). These associations were strongest on SRBs that evaluated global cognition (eg, MoCA, $r^2 = 0.63$ with CDR Time 1) and delayed recall (eg, Benson Figure Recall, $r^2 = 0.53$ with CDR Time 1), and less notable on SRBs assessing simple attention (eg, Number Span Forward, $r^2 = 0.26$ with CDR Time 1). These correlations were also stronger with the CDR at Time 2 compared to the CDR at Time 1 (eg, MoCA: $r^2 = 0.68$ vs 0.57, respectively). CDR sum of boxes has been linked to the risk of developing dementia,[26,27] so its association with poorer cognitive change over time is consistent with prior work.

When individuals who were reported to be biomarker positive via amyloid PET were compared to those reported to be amyloid negative, there were no significant differences on SRB z-scores in the cognitively unimpaired individuals, which suggests similar cognitive trajectories across their first follow-up in this longitudinal study. Future investigations might examine if these two subgroups remain comparable over longer periods of time (eg, 2 to 4 years) or if they diverge at some point. Conversely, those with amyloid results in the MCI and dementia groups showed more of this divergence. In those classified as MCI, individuals with amyloid positivity showed more decline on measures of global cognition, semantic fluency, and memory compared to their amyloid negative peers. Declines in these specific cognitive measures would be consistent with the progression to Alzheimer's disease.[28–30] Similarly, in the dementia group, the primarily amyloid positive participants demonstrated significantly worse SRB z-scores (ie, more decline) than individuals without excessive amyloid on PET imaging. In addition to the measures showing decline in MCI, Benson Figure Copy and Trail Making Test, Part A showed more decline with amyloid positivity in dementia. As some cognitive scores reach their floor (eg, minimum values on delayed recall trials) in the dementia group, they may lose sensitivity in tracking further decline and other measures with more range (eg, simple attention and construction) become more sensitive. However, given the small sample sizes and dichotomous amyloid status, these findings need to be replicated with larger cohorts and more detailed imaging results.

For those not familiar with the utility of SRBs, two case examples might be illustrative. During a medical appointment, a 77-year-old Hispanic female patient with 12 years of education presented with memory complaints. Her provider administered the MoCA, and the patient scored 30/30 at this baseline visit (Time 1). One year later (ie, 365 days, Time 2), she continued to complain of memory problems, so the MoCA was repeated, with a score of 26/30. Using Kiselica's coefficients, the patient's Time 1 raw score, her demographics, and time between assessments, her regression-based Time 2 score on the MoCA was predicted to be about 27 (ie, Predicted Time 2 = 13.11 + [−0.03*77] + [0.30*2] + [0.13*12] + [−0.85*2] + [−0.002*365] + [0.55*30] = 27.03). When this patient's observed and predicted Time 2 MoCA scores were compared and standardized, her "decline" of 4 raw points on the MoCA was still within the "stable" range (ie, SRB z-score = [26 to 27.03]/1.99 = −0.52), when considering her unique characteristics compared to the cognitively unimpaired NACC participants in Kiselica et al.[11] However, if this same patient's Time 1 score was in the dementia range (eg, 22) and her Time 2 score dropped by only 3 points (ie, 19), then her change on the MoCA would be more indicative of "decline" using the SRBs, with her follow-up score being predicted as about 23 (ie, Predicted Time 2 = 13.11 + [−0.03*77] + [0.30*2] + [0.13*12] + [−0.85*2] + [−0.002*365] + [0.55*22] = 22.63).

The comparison of her observed and predicted follow-up scores breaks the traditional threshold for decline of −1.645 (SRB z-score = [19 to 22.63]/1.99 = −1.82). Given the complexities of cognitive change and the many variables that can influence it, standard cutoffs to reflect change (eg, dropping by 2 points over a year) are likely to lead to both false positives and false negatives in these determinations. More sensitive and tailored approaches, like SRBs, can allow for more accurate determinations of change. To aid in the use of such change methods, Kiselica et al.[11] developed an SRB tool to perform these calculations, and this Excel spreadsheet is available in the online supplementary material of that paper.

As seen in the examples above, the SRBs are more computationally challenging than the more commonly used simple difference method of examining change. In the simple difference method, the difference between Time 1 and Time 2 scores is calculated (ie, Time 2 - Time 1). Although easier to generate, the results of the simple difference method lack clear meaning and scale. As in the examples above, it is unclear what a 3- to 4-point decline on the MoCA means (eg, is it normal or abnormal compared to peers?). Furthermore, without a common scale, a 3- to 4-point decline on the MoCA is unlikely to mean the same thing as a 3- to 4-point decline on the Trail Making Test or Craft Story Recall. Lastly, the simple difference method does not take into account factors that influence change, like the level of the Time 1 score, demographic variables, or the amount of time between assessments. The SRB method does provide meaning and scale, and it accounts for these factors, which yields a more sensitive accounting of change across time. For a review of these methods, see Duff.[25]

The current work is not without limitations. First, not all participants in the current sample were administered all cognitive tests. Since the NACC neuropsychological battery has evolved over time, some measures have been around for a long time (eg, Trail Making Test, Category Fluency) and others are newer (eg, MoCA, MiNT). Second, to match the inclusion criteria of Kiselica et al.,[11] participants were excluded from the current analyses if they were below the age of 50, did not complete the visit in English, had a follow-up visit of less than 6 months or more than 2 years from baseline visit, or did not have a contemporaneous CDR. Although these exclusions removed a relatively small number of participants, the generalizability of these findings to those individuals who would have been excluded is unclear. Recently published normative data on this cognitive battery in Spanish-speaking adults[31] may allow for SRBs for this important demographic group. Third, the current study utilized all-cause MCI and dementia, without consideration for subtypes or etiology. Finally, no biomarker data was used in confirming etiologies, although a report of amyloid status was used to compare cognitive trajectories in a small subset of the participants. As such, these results are not specific to amnestic MCI or Alzheimer's disease dementia, but they appear to apply to the range of neurodegenerative conditions that present at NACC sites. Despite these limitations, the SRB z-scores of Kiselica et al.[11] appear validated in impaired samples, which provides evidence for their use in clinical settings and research venues to enrich trials and/or more sensitively identify cognitive change.

ACKNOWLEDGMENTS

The project described was supported by research grants from the National Institute on Aging and the National Institute on Alcohol Abuse and Alcoholism: P30AG66518 and T32-AA007468. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Aging, National Institute on Alcohol Abuse and Alcoholism, or the National Institutes of Health. Additionally, the NACC database is funded by NIA/NIH Grant U24 AG072122. NACC data are contributed by the NIA-funded ADRCs: P30 AG062429 (PI James Brewer, MD, PhD), P30 AG066468 (PI Oscar Lopez, MD), P30 AG062421 (PI Bradley Hyman, MD, PhD), P30 AG066509 (PI Thomas Grabowski, MD), P30 AG066514 (PI Mary Sano, PhD), P30 AG066530 (PI Helena Chui, MD), P30 AG066507 (PI Marilyn Albert, PhD), P30 AG066444 (PI John Morris, MD), P30 AG066518 (PI Jeffrey Kaye, MD), P30 AG066512 (PI Thomas Wisniewski, MD), P30 AG066462 (PI Scott Small, MD), P30 AG072979 (PI David Wolk, MD), P30 AG072972 (PI Charles DeCarli, MD), P30 AG072976 (PI Andrew Saykin, PsyD), P30 AG072975 (PI David Bennett, MD), P30 AG072978 (PI Neil Kowall, MD), P30 AG072977 (PI Robert Vassar, PhD), P30 AG066519 (PI Frank LaFerla, PhD), P30 AG062677 (PI Ronald Petersen, MD, PhD), P30 AG079280 (PI Eric Reiman, MD), P30 AG062422 (PI Gil Rabinovici, MD), P30 AG066511 (PI Allan Levey, MD, PhD), P30 AG072946 (PI Linda Van Eldik, PhD), P30 AG062715 (PI Sanjay Asthana, MD, FRCP), P30 AG072973 (PI Russell Swerdlow, MD), P30 AG066506 (PI Todd Golde, MD, PhD), P30 AG066508 (PI Stephen Strittmatter, MD, PhD), P30 AG066515 (PI Victor Henderson, MD, MS), P30 AG072947 (PI Suzanne Craft, PhD), P30 AG072931 (PI Henry Paulson, MD, PhD), P30 AG066546 (PI Sudha Seshadri, MD), P20 AG068024 (PI Erik Roberson, MD, PhD), P20 AG068053 (PI Justin Miller, PhD), P20 AG068077 (PI Gary Rosenberg, MD), P20 AG068082 (PI Angela Jefferson, PhD), P30 AG072958 (PI Heather Whitson, MD), P30 AG072959 (PI James Leverenz, MD).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest. Author disclosures are available in the supporting information.

CONSENT STATEMENT

All human subjects provided informed consent or assent. If assent was taken, then informed consent was obtained from a legally authorized representative of the subject.

ORCID

Kevin Duff PhD https://orcid.org/0000-0002-9336-2400

REFERENCES

1. McSweeny AJ, Naugle RI, Chelune GJ, Luders H. T scores for change: an illustration of a regression approach to depicting change in clinical neuropsychology. *Clin Neuropsychol*. 1993;7:300-312.
2. Duff K, Beglinger LJ, Van Der Heiden S, et al. Short-term practice effects in amnestic mild cognitive impairment: implications for diagnosis and treatment. *Int Psychogeriatr*. 2008;20:986-999.

3. Duff K, Lyketsos CG, Beglinger LJ, et al. Practice effects predict cognitive outcome in amnestic mild cognitive impairment. *Am J Geriatr Psychiatry*. 2011;19:932-939.
4. Oltra-Cucarella J, Sanchez-SanSegundo M, Ferrer-Cascales R, Alzheimer Disease Neuroimaging I. Predicting Alzheimer's disease with practice effects, APOE genotype and brain metabolism. *Neurobiol Aging*. 2022;112:111-121.
5. Guevara JE, Kurniadi NE, Duff K. Assessing longitudinal cognitive change in mild cognitive impairment using estimated standardized regression-based formulas. *J Alzheimers Dis*. 2023;95:509-521.
6. Duff K, Beglinger LJ, Moser DJ, Schultz SK, Paulsen JS. Practice effects and outcome of cognitive training: preliminary evidence from a memory training course. *Am J Geriatr Psychiatry*. 2010;18:91.
7. Duff K, Foster NL, Hoffman JM. Practice effects and amyloid deposition: preliminary data on a method for enriching samples in clinical trials. *Alzheimer Dis Assoc Disord*. 2014;28:247-252.
8. Duff K, Hammers DB, Dalley BCA, et al. Short-term practice effects and amyloid deposition: providing information above and beyond baseline cognition. *J Prev Alzheimers Dis*. 2017;4:87-92.
9. Duff K, Anderson JS, Mallik AK, et al. Short-term repeat cognitive testing and its relationship to hippocampal volumes in older adults. *J Clin Neurosci*. 2018;57:121-125.
10. Duff K, Horn KP, Foster NL, Hoffman JM. Short-term practice effects and brain hypometabolism: preliminary data from an FDG PET Study. *Arch Clin Neuropsychol*. 2015;30:264-270.
11. Kiselica AM, Kaser AN, Webber TA, Small BJ, Benge JF. Development and preliminary validation of standardized regression-based change scores as measures of transitional cognitive decline. *Arch Clin Neuropsychol*. 2020.
12. Weintraub S, Besser L, Dodge HH, et al. Version 3 of the Alzheimer disease centers' neuropsychological test battery in the uniform data set (UDS). *Alzheimer Dis Assoc Disord*. 2018;32:10-17.
13. Nasreddine ZS, Phillips NA, Bedirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc*. 2005;53:695-699.
14. Partington J, Leiter R. Partington pathways test. *Psychol Service Cent J*. 1949;1:11-20.
15. Craft S, Newcomer J, Kanne S, et al. Memory improvement following induced hyperinsulinemia in Alzheimer's disease. *Neurobiol Aging*. 1996;17:123-130.
16. Possin KL, Laluz VR, Alcantar OZ, Miller BL, Kramer JH. Distinct neuroanatomical substrates and cognitive mechanisms of figure copy performance in Alzheimer's disease and behavioral variant frontotemporal dementia. *Neuropsychologia*. 2011;49:43-48.
17. Ivanova I, Salmon DP, Gollan TH. The multilingual naming test in Alzheimer's disease: clues to the origin of naming impairments. *J Int Neuropsychol Soc*. 2013;19:272-283.
18. Gollan TH, Weissberger GH, Runnqvist E, Montoya RI, Cera CM. Self-ratings of spoken language dominance: a multi-lingual naming test (MINT) and preliminary norms for young and aging spanish-english bilinguals. *Biling (Camb Engl)*. 2012;15:594-615.
19. Besser L, Kukull W, Knopman DS, et al, Neuropsychology Work Group D. Clinical Core leaders of the National Institute on Aging-funded USAsDC. Version 3 of the National Alzheimer's Coordinating Center's Uniform Data Set. *Alzheimer Dis Assoc Disord*. 2018;32:351-358.
20. Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology*. 1993;43:2412-2414.
21. Chelune GJ, Naugle RI, Luders H, Sedlak J, Awad IA. Individual change after epilepsy surgery: practice effects and base-rate information. *Neuropsychology*. 1993;7:41-52.
22. Hassenstab J, Ruvolo D, Jasielec M, Xiong C, Grant E, Morris JC. Absence of practice effects in preclinical Alzheimer's disease. *Neuropsychology*. 2015;29:940-948.
23. Duff K, Chelune G, Dennett K. Within-session practice effects in patients referred for suspected dementia. *Dement Geriatr Cogn Disord*. 2012;33:245-249.
24. Lin SS, Fletcher E, Gavett BE, Alzheimer's Disease Neuroimaging I. Reliable change in neuropsychological test scores is associated with brain atrophy in older adults. *J Neuropsychol*. 2021;15:274-299.
25. Duff K. Evidence-based indicators of neuropsychological change in the individual patient: relevant concepts and methods. *Arch Clin Neuropsychol*. 2012;27:248-261.
26. Grober E, Qi Q, Kuo L, Hassenstab J, Perrin RJ, Lipton RB. Stages of objective memory impairment predict Alzheimer's disease neuropathology: comparison with the clinical dementia rating scale-sum of boxes. *J Alzheimers Dis*. 2021;80:185-195.
27. Tzeng RC, Yang YW, Hsu KC, Chang HT, Chiu PY. Sum of boxes of the clinical dementia rating scale highly predicts conversion or reversion in predementia stages. *Front Aging Neurosci*. 2022;14:1021792.
28. Prado CE, Watt S, Treeby MS, Crowe SF. Performance on neuropsychological assessment and progression to dementia: a meta-analysis. *Psychol Aging*. 2019;34:954-977.
29. Olmos-Villasenor R, Sepulveda-Silva C, Julio-Ramos T, et al. Phonological and semantic fluency in Alzheimer's disease: a systematic review and meta-analysis. *J Alzheimers Dis*. 2023;95:1-12.
30. Carter SF, Caine D, Burns A, Herholz K, Lambon Ralph MA. Staging of the cognitive decline in Alzheimer's disease: insights from a detailed neuropsychological investigation of mild cognitive impairment and mild Alzheimer's disease. *Int J Geriatr Psychiatry*. 2012;27:423-432.
31. Marquine MJ, Parks A, Perales-Puchalt J, et al. Demographically-adjusted normative data among Latinos for the version 3 of the Alzheimer's disease centers' neuropsychological test battery in the uniform data set. *Alzheimers Dement*. 2023;19:4174-4186.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Duff K, Sevigny-Resetco D. Validation of regression-based change formulae for mild cognitive impairment and Alzheimer's disease. *Alzheimer's Dement*. 2024;16:e70008. https://doi.org/10.1002/dad2.70008