Supplemental Table A: Neuropsychological tests and their specific components used for DELTA score calculation

|  |  |  |  |
| --- | --- | --- | --- |
| **Cognitive Domain** | **Test Name** | **Test Component(s)** | **Score Range** |
| Memory | Rey Auditory Verbal Learning Test (AVLT) | Delayed Recall | 0-15 |
| % Retention  (Delayed Recall / Trial 6) | 0-100+% |
| WMS-R Logical Memory (LM) | Delayed Recall | 0-25 |
| % Retention (Delayed Recall / Immediate Recall) | 0-100+% |
| Language | Animal Fluency | Total Correct | 0-(no max) |
| 30-Item Boston Naming Test (BNT30) | Total Correct (spontaneous + semantic cue) | 0-30 |
| Executive Function | Clock Drawing | Total Score | 0-5 |
| Trail Making Test | Trails B/A Proportion (Trails B Completion Time / Trails A Completion Time) | >1.0 |

Supplemental Table B: Change in DELTA group status based on BL DELTA group for the Robust Cognitively Normal subsample. Values represent the percentage of participants with follow-up DELTA scores within each DELTA group. DELTA subgroups were determined based on the distribution of DELTA scores in the RCN sample.

Abbrev: DELTA – Discrepancy-based Evidence for Loss of Thinking Abilities, RCN – Robust Cognitively Normal

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **RCN Baseline DELTA Group (Score Range)** | **Follow-Up DELTA Group (Score Range)** | **Year 1** | **Year 2** | **Year 3** |
| *“Level of Evidence”* | *“Level of Evidence”* | *(N=187)* | *(N=176)* | *(N=95)* |
| No Evidence (0)  N=189 | No Evidence (0) | 86.6 | 83.5 | 88.4 |
| Low (1-3) | 11.8 | 16.5 | 11.6 |
| Moderate (4-6) | 1.6 | 0.0 | 0.0 |
| Strong (7-9) | 0.0 | 0.0 | 0.0 |
| Very Strong (10-15) | 0.0 | 0.0 | 0.0 |
|  |  | *(N=63)* | *(N=57)* | *(N=35)* |
| Low (1-3)  N=63 | No Evidence (0) | 68.3 | 56.1 | 54.3 |
| Low (1-3) | 28.6 | 33.3 | 42.9 |
| Moderate (4-6) | 3.2 | 8.8 | 2.9 |
| Strong (7-9) | 0.0 | 1.8 | 0.0 |
| Very Strong (10-15) | 0.0 | 0.0 | 0.0 |

**Appropriate Use of the Current DELTA Score (Supplemental)**

The DELTA score presented in this study can only be used with white/Caucasian individuals due to the highly white/Caucasian ADNI cohort. The neuropsychological test battery must mirror the tests administered in the ADNI cohort (Table 1). While regression-based normative methods theoretically allow for application beyond age, education, and word-reading ability ranges included in the normative sample, we caution against this (Brooks et al., 2011). The most appropriate cases will be over age 65 (IQR of RCN sample = 70.8-78.2) and have at least some college education (IQR of RCN sample = 15-19 years of education).

**Novel Aspects of the DELTA Approach (Supplemental)**

We did not set out to create a new definition for “MCI” or any other clinical syndrome. Rather, the goal of the DELTA approach was advancing the clinical utility of psychometrically sound, algorithmic methods that most closely mirrors well-validated methods in the MCI literature (Bondi et al., 2014; Jak et al., 2016). The DELTA score provides a finer-grained, continuous metric for characterizing the strength of evidence for cognitive decline. Relative to other approaches, there are several novel components of the DELTA score that we believe incrementally advances neuropsychological test score interpretation (Iverson & Brooks, 2011):

* We incorporated a person-specific, performance-based metric (word-reading ability) into regression equations for predicting “premorbid” test scores, which added value over traditional normative reference factors like age, gender, and years of education on some tests.
* We attempted to reduce confounding effects of overlapping cognitive functions on specific test scores. For example, rather than reliance on delayed memory recall score(s) only, we required individuals demonstrate evidence of forgetfulness such that they lost at least 50% of initially-learned details. We also parsed out the “executive” set-shifting component of Trails B by dividing the Trails B time by the individual’s Trails A time.
* We expanded upon methods defining cognitive status using a single cut-point, like z<-1.0, by assigning different levels of evidence for cognitive decline (i.e., higher or lower DELTA scores) that account for z-Discrep scores along multiple percentile cutoffs (16th, 7th, and 2nd).
* We further individualize the DELTA score by adjusting the required z-Discrep scores as a function of predicted test scores (“Low,” “Average,” or “High”). While the 16th, 7th, and 2nd percentile cutoffs are applied uniformly in the DELTA scoring criteria, the z-Discrep values that correspond to these percentiles vary by predicted raw score. This should raise the confidence that a given discrepancy between one’s observed and predicted score represents true change and is not an artifact of having above or below average estimated premorbid intellect (Horton, 1999; Tremont, Hoffman, Scott, & Adams, 1998).
* We developed and provide an automated, clinic-ready scoring program that not only calculates an individual’s DELTA score but also includes data informing risk for future functional decline and for positive AD biomarkers.