# Supplementary Material

**An Operational Definition of 'Abnormal Cognition' to Optimize the Prediction of Progression to Dementia: What Are Optimal Cut-Off Points for Univariate and Multivariate Normative Comparisons?**

## SUPPLEMENTARY MATERIAL 1

**Supplementary Table 1.** Characteristics of the 10 neuropsychological test variables in ANDI that were used for the current analyses.

| | N in ANDI norms | % Male | Age range | Demographic variables |
|---|---|---|---|---|
| MMSE | 16128 | 55 | 18-97 | A+S+E |
| RAVLT – total score of 5 trials | 5017 | 50 | 18-97 | A+S+E |
| RAVLT – delayed recall score | 4540 | 49 | 18-97 | A+S+E |
| Animal fluency (1 minute) | 5783 | 40 | 18-96 | A+E |
| COWAT letter fluency | 2894 | 48 | 18-97 | A+S+E |
| TMTa | 3216 | 47 | 18-97 | A+S+E |
| TMTb | 3320 | 46 | 18-97 | A+S+E |
| Stroop card I – color | 1783 | 41 | 18–91 | A+S+E |
| Stroop card II – word | 2147 | 43 | 18-91 | A+S+E |
| Stroop card III – color-word interference | 2132 | 43 | 18-91 | A+S+E |

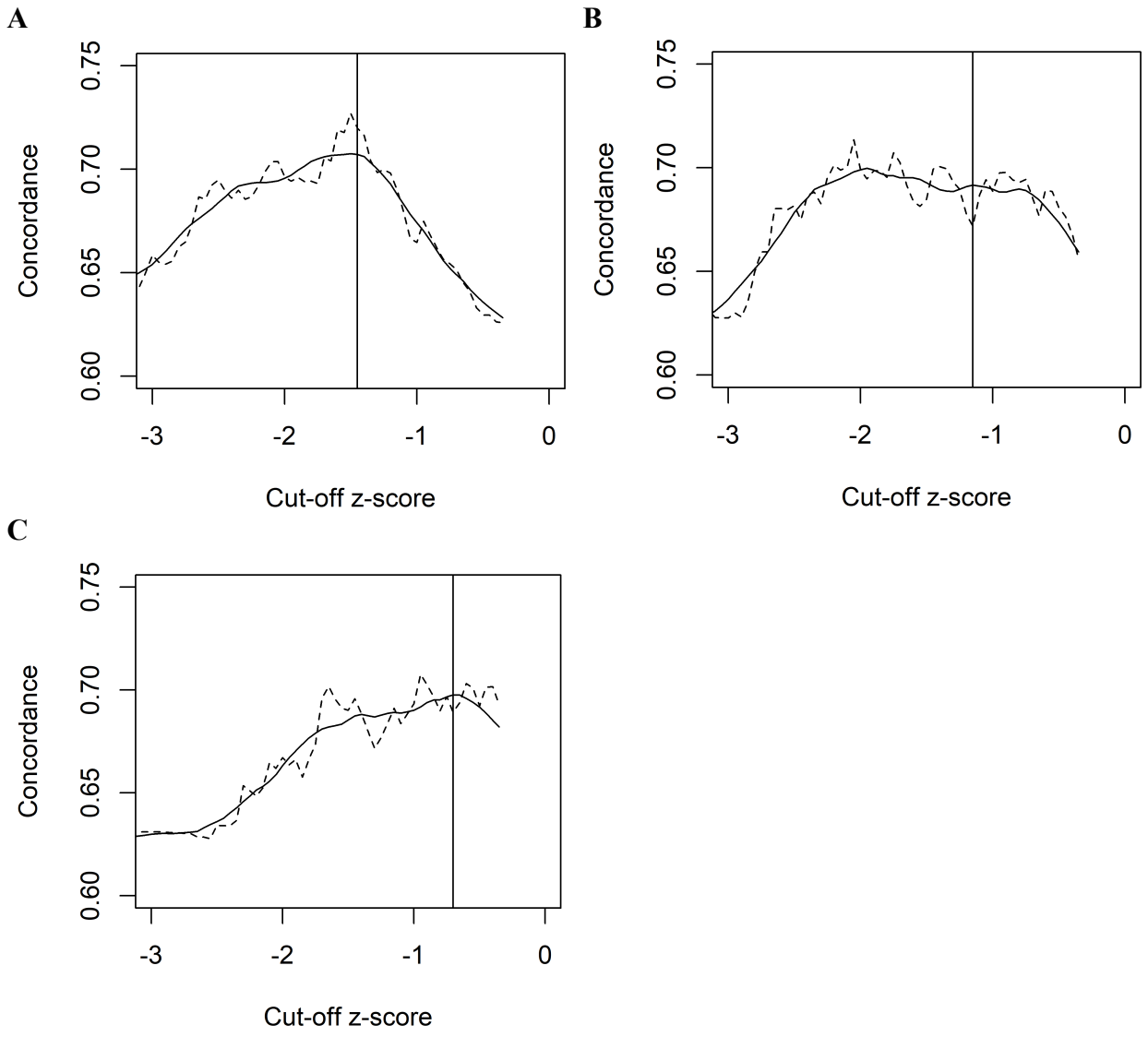
MMSE, Mini-Mental State Examination; RAVLT, Rey Auditory Verbal Learning Test; COWAT, Controlled Oral Word Association Test; TMT, Trail Making Test; Stroop, Stroop Color-Word Test. A, age; S, sex; E, level of education were necessary for demographic corrections (de Vent NR, Agelink van Rentergem JA, Schmand BA, Murre JM; ANDI Consortium, Huizenga HM (2016) Advanced Neuropsychological Diagnostics Infrastructure (ANDI): a normative database created from control datasets. *Front Psychol* **7**, 1601).

**SUPPLEMENTARY MATERIAL 2**

## Calibration of number and magnitude of deviations

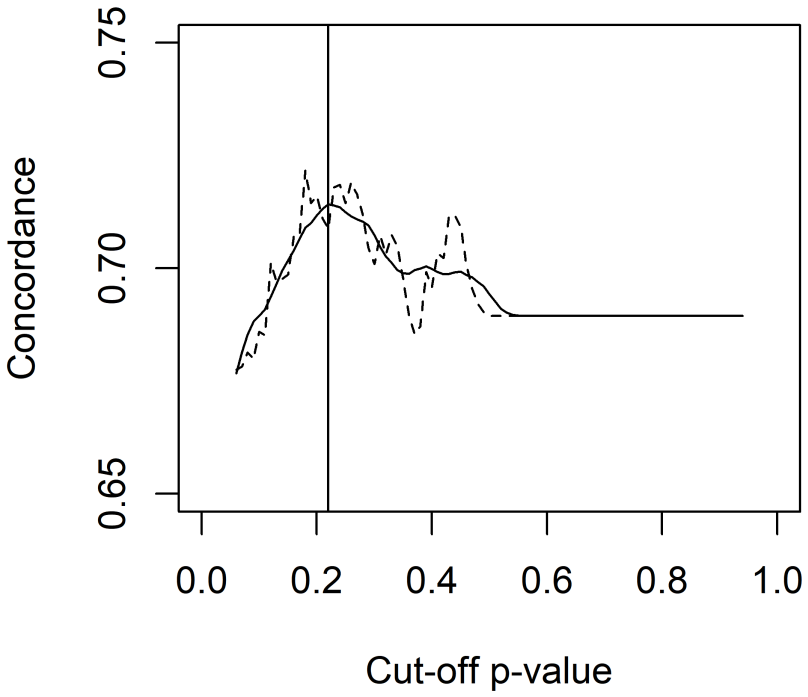
For one deviation, visual inspection of Supplementary Figure 1A suggests that the concordance increases until it reaches a maximum level around $z$ = -1.5, after which it plateaus until about $z$ = -2.5. In this plateau, the cut-off scores all produce similar results in terms of predicting progression to dementia. A formal analysis (see Supplementary Material 3) was performed to find the border of the plateau which confirms that $z$ = -1.45 is indeed the point where the test statistic starts to plateau. For two deviations (Supplementary Figure 1B), a similar situation occurs: the test statistic increases until around $z$ = -1 and then has a plateau to $z$ = -2.2; a formal analysis indicates the plateau begins at $z$=-1.15. For three deviations (Supplementary Figure 1C), it peaks at $z$ = -0.7 and plateaus until about -2.0; a formal analysis confirms the plateau begins at $z$ = -0.70.

**Supplementary Figure 1.** Concordance (y-axis) as a function of magnitude of cut-off z-scores (x-axis) for prediction of development of dementia given (A) one, (B) two, and (C) three deviations. Both the actual concordance scores (dashed line) and a smoothed function (running mean ± 6; solid line) are plotted. The vertical line indicates the cut-off point found by the formal analysis of the start of the plateau.



For the MNC status we calibrated how large the profile abnormality had to be to best predict progression to dementia. Supplementary Figure 2 shows how well the MNC-status predicts progression to dementia as a function of the cut-off *p*-values. It shows a clear peak indicating the optimal cut-off value. A *p*-value of 0.22 (one-sided; as we are trying to diagnose abnormality) as to be selected as the best value to predict progression to dementia.

**Supplementary Figure 2.** Cut-off *p*-values (x-axis) for the MNC-status method. The y-axis shows the concordance. The dotted line represents the raw curve and the continuous line is the smoothed function (running mean ± 6). The vertical line indicates the point where the concordance is highest.



**SUPPLEMENTARY MATERIAL 3**

**Finding start of plateau**

In order to find the start of the plateau for each of the criteria (1, 2, or 3 deviations) depicted in Supplementary Figure 3, we calculated the second derivative. If the second derivative is maximally negative, the concordance has reached its plateau (when reading from right to left). The first derivative was calculated according to:

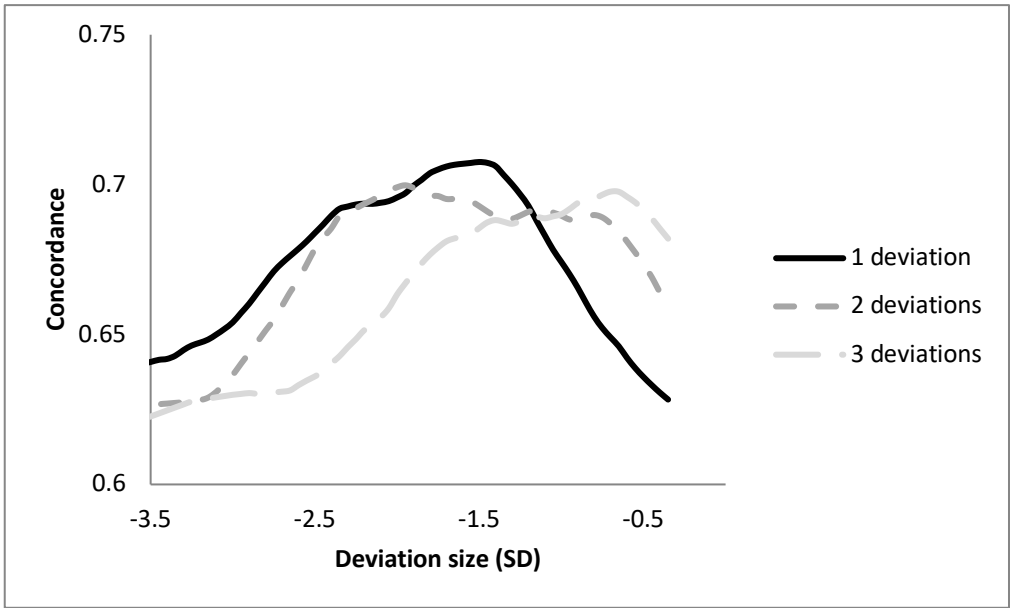
$$\frac{dy_i}{dx} = \frac{y_{i+2} - y_{i-2}}{x_{i+2} - x_{i-2}}$$

where $y_i$ is the $i^{\text{th}}$ concordance and $x_i$ is $i^{th}$ deviation size. The first derivative denotes the slope (positive means increasing concordance, negative means decreasing predictive ability) and is plotted in Figure 8. Then the second derivative was calculated according to:

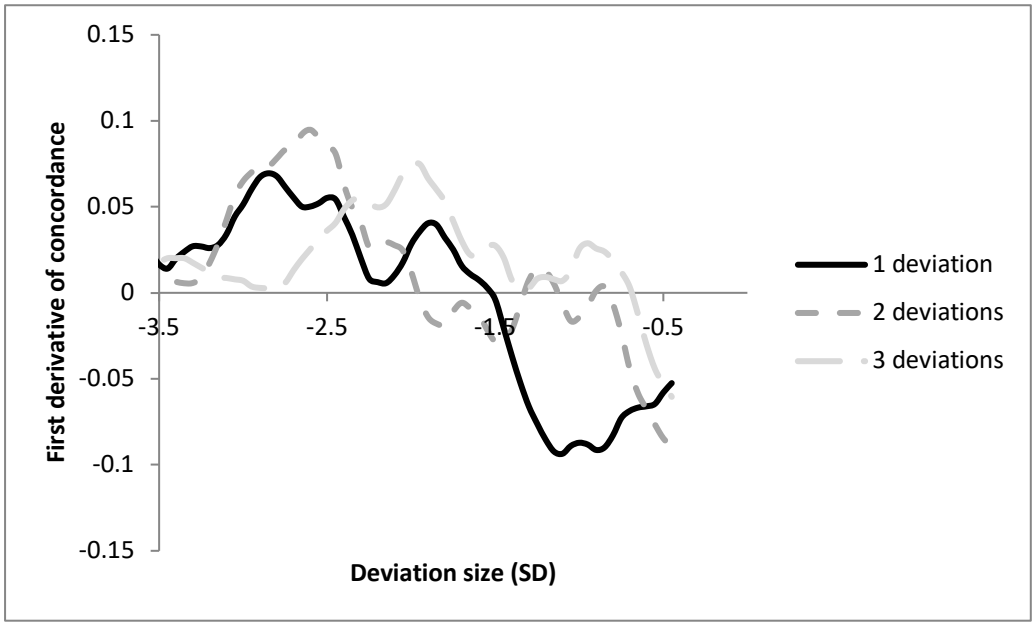
$$\frac{d^2y_i}{dx^2} = \frac{dy_{i+2}/dx_{i+2} - dy_{i-2}/dx_{i-2}}{x_{i+2} - x_{i-2}}$$

This denotes the change in slope of the predictor value. This is plotted in Supplementary Figure 5.

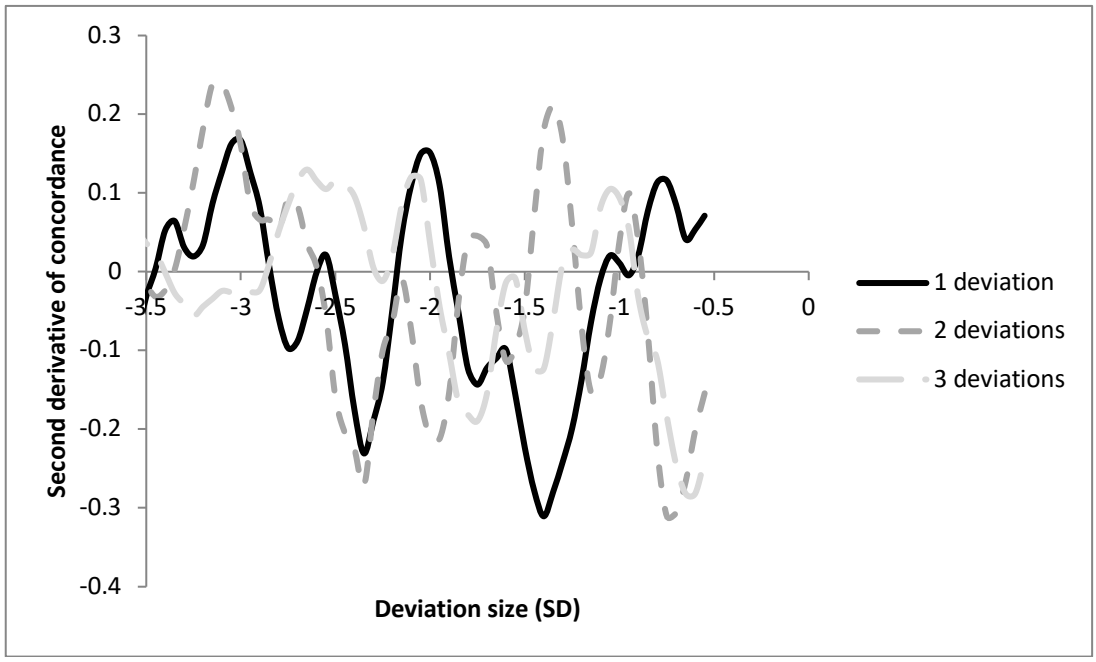
**Supplementary Figure 3.** Concordance as a function of deviation size for one, two and three deviations.




**Supplementary Figure 4.** First derivative of concordance as a function of deviation size for one, two and three deviations



**Supplementary Figure 5.** Second derivative of concordance as a function of deviation size for one, two and three deviations. We used this to calibrate the optimal cut-off points, that is for one deviation it is -1.45, for two -1.15 and for three deviations -0.70.



**SUPPLEMENTARY MATERIAL 4**

**Diagnostic overlap between methods**

**Supplementary Table 2.** Diagnostic overlap expressed by Cohen's k for each combination of cut off scores.

| *Combination* | *Cohen's k* | *95% CI* | *p* |
|---|---|---|---|
| **1 deviation (z < -1.45) and 2 deviations (z < -1.15)** | 0.69 | 0.63 - 0.77 | <0.001 |
| **1 deviation (z < -1.45) and 3 deviations (z < -0.70)** | 0.66 | 0.59 - 0.74 | <0.001 |
| **2 deviations (z <-1.15) and 3 deviations (z < -0.70)** | 0.68 | 0.61 - 0.75 | <0.001 |
| **MNC status (p < 0.22) and 1 deviation (z < - 1.45)** | 0.61 | 0.54 - 0.68 | <0.001 |
| **MNC status (p < 0.22) and 2 deviations (z < -1.15)** | 0.62 | 0.54 - 0.69 | <0.001 |
| **MNC status (p < 0.22) and 3 deviations (z < -0.70)** | 0.48 | 0.40 - 0.56 | <0.001 |