Cognitive reserve and clinical progression in Alzheimer disease
A paradoxical relationship

Anna Catharina van Loenhoud, MSc, Wiesje Maria van der Flier, PhD, Alle Meije Wink, PhD, Ellen Dicks, MSc, Colin Groot, MSc, Jos Twisk, PhD, Frederik Barkhof, MD, PhD, Philip Scheltens, MD, PhD, and Rik Ossenkoppele, PhD, for the Alzheimer's Disease Neuroimaging Initiative


Abstract

Objective
To investigate the relationship between cognitive reserve (CR) and clinical progression across the Alzheimer disease (AD) spectrum.

Methods
We selected 839 β-amyloid (Aβ)–positive participants with normal cognition (NC, n = 175), mild cognitive impairment (MCI, n = 437), or AD dementia (n = 227) from the Alzheimer’s Disease Neuroimaging Initiative (ADNI). CR was quantified using standardized residuals (W scores) from a (covariate-adjusted) linear regression with global cognition (13-item Alzheimer’s Disease Assessment Scale–cognitive subscale) as an independent variable of interest, and either gray matter volumes or white matter hyperintensity volume as dependent variables. These W scores, reflecting whether an individual’s degree of cerebral damage is lower or higher than clinically expected, were tested as predictors of diagnostic conversion (i.e., NC to MCI/AD dementia, or MCI to AD dementia) and longitudinal changes in memory (ADNI-MEM) and executive functions (ADNI-EF).

Results
The median follow-up period was 24 months (interquartile range 6–42). Corrected for age, sex, APOE4 status, and baseline cerebral damage, higher gray matter volume-based W scores (i.e., greater CR) were associated with a lower diagnostic conversion risk (hazard ratio [HR] 0.22, p < 0.001) and slower decline in memory (β = 0.48, p < 0.001) and executive function (β = 0.67, p < 0.001). Stratified by disease stage, we found similar results for NC (diagnostic conversion: HR 0.30, p = 0.038; ADNI-MEM: β = 0.52, p = 0.028; ADNI-EF: β = 0.42, p = 0.077) and MCI (diagnostic conversion: HR 0.21, p < 0.001; ADNI-MEM: β = 0.43, p = 0.003; ADNI-EF: β = 0.59, p < 0.001), but opposite findings (i.e., more rapid decline) for AD dementia (ADNI-MEM: β = −0.91, p = 0.002; ADNI-EF: β = −0.77, p = 0.081).

Conclusions
Among Aβ-positive individuals, greater CR related to attenuated clinical progression in pre-dementia stages of AD, but accelerated cognitive decline after the onset of dementia.
A major limitation in the care and management of Alzheimer disease (AD) is the inability to provide patients with an accurate prognosis. Cognitive reserve (CR), the brain’s capability to preserve cognition despite underlying cerebral damage, may be a key determinant of clinical progression. Previous studies have demonstrated a paradox: while CR is associated with a delayed symptom onset, it is related to accelerated cognitive decline in advanced AD stages. Other studies, however, found no longitudinal CR associations.

These inconclusive results have several explanations. First, various studies used lifestyle proxies to operationalize CR (e.g., education, occupation, or physical activity), but there is no consensus about the validity of these indirect measures. Second, presence of AD was often established through clinical evaluation without consideration of AD biomarkers, likely resulting in samples confounded with non-AD participants. Third, few studies comprehensively investigated the relationship between CR and clinical progression across the entire AD spectrum, thus hampering a direct comparison between disease stages.

We optimized the examination of CR in relation to AD-related clinical progression by (1) using a (more direct) neuroimaging measure of CR, based on cognitive and structural brain measures, (2) selecting β-amyloid (Aβ)–positive individuals only, and (3) including participants across the entire AD spectrum (i.e., normal cognition [NC], mild cognitive impairment [MCI], and AD dementia). We hypothesized that greater CR related to attenuated diagnostic conversion risk and slower cognitive decline in early disease stages, but exacerbated clinical progression among clinically advanced participants.

Methods

Participants

All participants were recruited within the Alzheimer’s Disease Neuroimaging Initiative (ADNI) and enrolled at sites in the United States and Canada across 3 funding cycles (ADNI-1, ADNI-GO, and ADNI-2). They were aged between 55 and 91 years at baseline, English or Spanish speakers, non-depressed, and classified as NC, MCI, or AD dementia. Follow-up visits were scheduled every 6 months in the first 2 years, and annually thereafter. Cognitive status was evaluated with the Clinical Dementia Rating scale, Mini-Mental State Examination (MMSE), and Wechsler Memory Scale Logical Memory II. Cognitive assessments, physical examinations, and MRI scans were considered in determining diagnostic status. We selected all NC, MCI, and AD dementia participants with positive AD biomarkers (11C–Pittsburgh compound B or 18F-florbetapir PET if available, or CSF Aβ42), structural MRI, and neuropsychological data at baseline. We excluded individuals with MRI segmentation failures (n = 2), development of non-AD dementia or other neurologic disorders during follow-up (n = 5), and >6 months between baseline MRI and neuropsychological assessment (n = 2). Our final Aβ-positive sample (n = 839) included 175 NC, 437 MCI, and 227 AD dementia participants (figure 1). Follow-up data were available for 96% of the dataset. The remaining 4% was included to improve estimation of baseline effects in statistical models.

Standard protocol approvals, registrations, and patient consents

Written informed consent was obtained for all participants, and study procedures were approved by the institutional review board at each of the participating centers. ADNI is listed in the ClinicalTrials.gov registry (ADNI-1: NCT00106899; ADNI-GO: NCT01078636; ADNI-2: NCT0123197).

MRI acquisition and preprocessing

Structural MRI scans were acquired on 1.5T and 3T scanners from GE Healthcare (Cleveland, OH), Philips Medical Systems (Best, the Netherlands), and Siemens Medical Solutions (Malvern, PA) (adni.loni.usc.edu/methods/mri-analysis/mri-acquisition/). We downloaded T1 scans preprocessed with gradwarping, B1 correction, or N3 scaling for the present analyses. We further preprocessed these images using SPM12, before extracting gray matter (GM) volumes for all 839 participants. We computed intracranial volume (ICV) as the sum of GM, white matter, and CSF. White matter hyperintensity (WMH) volumes were segmented from proton density and T2 images (ADNI-1) or 2D–fluid-attenuated inversion

Glossary

Aβ = β-amyloid; AD = Alzheimer disease; ADAS-Cog = Alzheimer’s Disease Assessment Scale–cognitive subscale; ADNI = Alzheimer’s Disease Neuroimaging Initiative; ADNI-EF = standardized composite score based on a clock drawing task, animal and vegetable category fluency, the Trail-making task, Digit Span Backwards from the Wechsler Memory Scale–Revised, and the digit–symbol substitution task from the Wechsler Adult Intelligence Test–Revised; ADNI-MEM = standardized composite score based on the Alzheimer’s Disease Assessment Scale–cognitive subscale word list learning task, the Rey Auditory Verbal Learning Test, Logical Memory from the Wechsler Memory Scale; AIC = Akaike Information Criterion; ANOVA = analysis of variance; CI = confidence interval; CR = cognitive reserve; GM = gray matter; HR = hazard ratio; ICV = intracranial volume; IQR = interquartile range; MCI = mild cognitive impairment; MMSE = Mini-Mental State Examination; NC = normal cognition; WMH = white matter hyperintensity; VIF = variance inflation factor.
recovery (ADNI-GO/2) by the ADNI core laboratory using automated techniques and downloaded from the website for 796/839 participants. We log-transformed the WMH volumes to account for non-normality.

**Structural brain measures**

To quantify cerebral damage, we extracted GM volume on a whole-brain level and from bilateral temporoparietal and hippocampal masks (based on the Automated Anatomical Labeling atlas; figure e-1, doi.org/10.5061/dryad.cb0qk1g) and obtained global WMH volumes. As described below, these measures were included as dependent variables in models to create W scores, our operational measure of CR. Furthermore, GM and WMH volumes served as individual predictors of clinical progression in statistical models. For this purpose, we corrected for premorbid brain size by regressing out the effect of ICV, and used standardized GM or WMH volume residuals for further statistical analysis.

**Neuropsychological assessment**

Cognitive measures in our study were 13-item Alzheimer’s Disease Assessment Scale–cognitive subscale (ADAS-Cog), ADNI-MEM, and ADNI-EF. ADAS-Cog 13 is a global cognitive screening test that includes 13 subtests for memory, language, praxis, attention, visuoconstruction, and orientation. ADNI-MEM is a composite score based on the Rey Auditory Verbal Learning task, word list learning and recognition tasks from ADAS-Cog, recall from Logical Memory I of the Wechsler Memory Test–Revised, and 3-word recall item from the MMSE. ADNI-EF consists of Category Fluency (i.e., animals and vegetables), Trail-Making Test part A and B, Digit Span Backwards, Wechsler Adult Intelligence Scale–Revised Digit–Symbol Substitution, and 5 Clock Drawing items. In statistical analyses, we multiplied ADNI-MEM and ADNI-EF scores by 100 to facilitate the presentation of results.

**Calculation of W scores**

Our neuroimaging approach takes the general relationship between cognition and cerebral damage as a starting point, and uses this information to estimate an individual’s expected degree of cerebral damage based on cognitive performance. This expected value is then compared with an individual’s actual degree of cerebral damage, and the difference is referred to as a W score, our operational measure of CR. Cerebral damage was quantified using 4 measures: GM volume, either whole-brain (1) or within specific temporoparietal (2) or hippocampal masks (3), and global WMH volume (4). These measures were used as dependent variables in separate linear regressions that further included global cognitive performance (ADAS-Cog 13) as an independent variable of interest, and age, sex, ICV, and scanner field strength as covariates. Education was not added as a covariate, since this is one of the contributing factors of CR (i.e., our concept of interest). We obtained standardized residuals from these regressions and multiplied them by -1 to derive W scores (this multiplication was omitted for WMH volumes, as higher volumes already reflect greater cerebral damage). Specifically, we used the following formula: \[ W \text{ score} (\text{CR}) = -1 \times (\text{observed cerebral damage} - \text{predicted cerebral damage}/\text{SD}). \] Higher W scores reflect greater CR, because they indicate that...
a relatively high degree of cerebral damage is tolerated at a given level of cognitive function. We obtained 4 different W scores for further analysis (i.e., 3 GM-based W scores, 1 WMH-based W score).

Data availability
All imaging, demographics, and neuropsychological data used in this article are publicly available and were downloaded from the ADNI website (adni.loni.usc.edu). Upon request, we will provide a list of ADNI participant identifications for replication purposes.

Statistical analysis
Participant characteristics
Statistical analyses were performed in SPSS 22 (Chicago, IL). Analysis of variance (ANOVA), χ², and Kruskal-Wallis tests were used to compare disease stages by demographic variables, follow-up time, cognitive performance, ICV, and W scores. We used a multiple imputation procedure for the 0.3% of cognitive measures missing at baseline. This procedure was performed using the fully conditional specification method in SPSS, which is an iterative Markov Chain Monte Carlo approach suitable for arbitrary patterns of missing data. We imputed 15 data sets and included demographic, clinical, and neuropsychological variables and ICV as predictors in the model.

Relationships among W scores, structural brain measures, and education
First, we correlated W scores with the (ICV-adjusted) structural brain measures from which they were calculated (i.e., GM or WMH volume). We expected that these variables were interrelated because the higher the degree of cerebral damage, the greater the potential manifestation of CR. Moreover, as mathematically explained elsewhere, residuals from linear regressions are inherently correlated with their dependent variables. Such a correlation would suggest that the relationship between W scores and clinical progression cannot be correctly interpreted without taking into account the presence of cerebral damage, and thus we evaluated whether these structural measures could be included as covariates in subsequent statistical models, without violating the assumptions of multicollinearity. We considered a correlation r <0.8 between W scores and their related structural brain measure³⁹ and a variance inflation factor (VIF) <5 for these variables in combination with other covariates (age, sex, APOE4 status)⁴⁰ an acceptable level of multicollinearity. Furthermore, we determined the relationship of W scores and structural brain measures with education. As education is a well-known lifestyle proxy of CR, W scores should show a (positive) relationship with education, while this is not necessarily the case for the structural brain measures. We therefore grouped participants with lower (≤14 years, n = 264), intermediate (15–17 years, n = 286), and higher education (>18 years, n = 289) and performed ANOVAs to examine differences among these groups in W scores (ICV-adjusted) and GM and WMH volumes.

W scores and diagnostic conversion
We performed Cox regressions among predementia participants (NC, n = 175; MCI, n = 437) to study relationships between (continuous) W scores and diagnostic conversion (i.e., change in baseline diagnosis from NC to MCI/AD dementia, or from MCI to AD dementia). Diagnostic changes to a less advanced disease stage (e.g., MCI to NC, which sporadically occurred in the sample) was not considered a conversion. We ran 4 models for each W score, and included the related structural measure (i.e., ICV-adjusted whole-brain, temporoparietal, or hippocampal GM volume or WMH volume), age, sex, and APOE4 status as covariates. These analyses were also stratified for MCI and NC. We used 95% confidence intervals (CIs) to compare W score effects between disease stages; differences were considered significant when the effect size for one group fell outside of the CI of the other group.⁴¹

W scores and cognitive decline
To examine clinical progression in more psychometric detail, we performed linear mixed models with memory (ADNI-MEM) and executive functioning (ADNI-EF) as outcome variables and (continuous) W score, time (i.e., months), and W score × time as predictors. The interaction term was the effect of interest, as it reflected whether CR moderated the relationship between time and cognitive performance. Our models contained random intercepts and slopes. Note that, in contrast to the Cox regressions, the AD dementia group (n = 227) was now also included for analysis. Again, we ran separate models for each W score in the total group and stratified by disease stage, while correcting for structural measures (ICV-adjusted GM or WMH volume, and their interaction with time) as well as age, sex, and APOE4 status.

In secondary analyses, we repeated the main linear mixed models, but only included time points that were collected until the (first) date of diagnostic conversion. To illustrate the relevance of this analysis, consider a participant with MCI at baseline, a diagnostic conversion to AD dementia at 48 months, and a total follow-up of 108 months. In stratified analyses, this participant would be classified into the MCI group, although 56% of the data actually concern clinical progression during the AD dementia stage. Constraining the linear mixed models to exclusively consider data from the initial diagnosis leads to results that are completely non-overlapping with respect to disease stage (and thus more easily interpretable).

Comparison with other models of clinical progression
If our W scores were related to clinical progression, we would assume that this is an effect of CR. However, an alternative explanation would be that W scores are a derivative measure of cerebral damage, which presumably also relates to clinical progression. Although we partly accounted for this by including structural brain measures as covariates in the statistical models, we explicitly aimed to investigate whether W scores
added explanatory value to the prediction of clinical progression. We therefore compared our main models of clinical progression (which included both structural brain measures and W scores) to models in which W scores were not included. We used the Akaike Information Criterion (AIC) \(^{42}\) to evaluate which of the 2 models showed a better fit to the data. AIC assesses the model fit while also penalizing the number of predictors, thus favoring models with the optimal trade-off between parsimony and goodness-of-fit. A lower AIC value indicates a better model. We also assessed whether including W scores in models with a cognitive measure (ADAS-Cog 13) resulted in significant AIC improvements.

### Sensitivity analyses

We performed additional analyses to ensure that our main findings were robust to different methodologic decisions. We recalculated the W scores in 2 ways: (1) we allowed a nonlinear relationship between cognition and cerebral damage by adding a quadratic term (i.e., ADAS-Cog 13^2\(^{3}\)), as the CR theory describes a curve that reflects a relationship between cerebral damage and clinical symptoms that is initially flat and changes exponentially; (2) we created W scores based on ADNI-MEM and ADNI-EF as cognitive measures and instead used ADAS-Cog 13 as an outcome measure of cognitive decline. These recalculated W scores were replaced with the original W scores in the Cox regressions and linear mixed models, and we assessed whether this replacement led to relevant changes.

### Results

#### Participant characteristics

Table 1 provides an overview of our Aβ-positive sample (mean age 73.9 years; median education 16 years). Demographic characteristics were similar across disease stages, except that the proportion of APOE4 positivity was higher among more advanced disease stages (NC: 45.7%, MCI: 63.2%, AD dementia: 73.6%),\(^{43}\) and the NC group included more female participants (58.3%) relative to MCI and AD dementia groups (42.3% and 43.6%, respectively). NC participants also had a lower ICV (1,482.1 ± 145.1 cm\(^3\)) compared to MCI participants (1,519.3 ± 157.6 cm\(^3\)), and their GM-based hippocampal W scores (−0.27 ± 0.91) were lower than in MCI and AD dementia participants (0.07 ± 1.01, 0.6 ± 1.00, respectively). The other W scores were comparable across disease stages. On all cognitive measures, there were significant baseline differences between disease stages in the expected direction (i.e., AD dementia < MCI < NC). Similarly, CSF biomarkers were more abnormal with increasing clinical severity (i.e., lower CSF Aβ and higher CSF p-tau and t-tau). The median follow-up was 24 months (interquartile range [IQR] 6–42) in the total sample, with a shorter duration in the AD dementia group (median 12 months, IQR 6–18) compared to the other 2 groups (median 36 months, IQR 24–48). Diagnostic conversion to MCI occurred in 21% of the NC sample, and 32% of all predementia cases showed (additional) diagnostic conversion to AD dementia.

#### Relationships among W scores, structural brain measures, and education

As expected, each W score showed a significant correlation with the structural brain measure from which it was derived (table e-1, doi.org/10.5061/dryad.cb0qk1g). Assumptions of multicollinearity were not violated (i.e., r < 0.8 and VIF < 5 in all cases), thus we included the structural brain measures as covariates in the statistical models assessing the effect of W scores on clinical progression. ANOVA revealed significant differences between education groups for GM-based whole-brain (\(p = 0.043\)) and temporoparietal W scores (\(p = 0.047\)), such that highly educated participants showed higher W scores (i.e., greater CR) than those with lower education. The other 2 W scores and all structural brain measures did not show relationships with education (table 2). Since the association with education is an important validation step for the use of W scores as an operational measure of CR, we will only discuss results from the GM-based whole-brain and temporoparietal W scores here. However, all main statistical analyses were also performed in the GM-based hippocampal and WMH-based W scores (tables e-2–e-4, doi.org/10.5061/dryad.cb0qk1g). For simplicity, we will refer to “whole-brain/temporoparietal W scores” in the remaining Results sections, thus omitting the term “GM-based.”

#### W scores and diagnostic conversion

Cox regressions in predementia participants (\(n = 612\)) showed a negative association between whole-brain W scores and diagnostic conversion risk (222 conversions; whole brain: hazard ratio [HR] 0.22, \(p < 0.001\); temporoparietal: HR 0.25, \(p < 0.001\), indicating that greater CR related to a lower risk. Stratified by disease stage, this finding was present in both the NC group (38 conversions; whole brain: HR 0.30, \(p = 0.038\); temporoparietal: HR 0.30, \(p = 0.021\)) and the MCI group (184 conversions; whole brain: HR 0.21, \(p < 0.001\); temporoparietal: HR 0.25, \(p < 0.001\)). An overview of all W score models and the effect of each individual predictor on diagnostic conversion is available (table e-2; doi.org/10.5061/dryad.cb0qk1g).

#### W scores and cognitive decline

Linear mixed models in the total sample (\(n = 839\)) showed that higher whole-brain W scores were related to better baseline performance on ADNI-MEM (\(\beta = 121.19, p < 0.001\)) and ADNI-EF (\(\beta = 86.78, p < 0.001\)). Furthermore, these scores declined over time (ADNI-MEM: \(\beta = −1.35, p < 0.001\); ADNI-EF: \(\beta = −1.28, p < 0.001\)). Importantly, this relationship was moderated by CR, such that participants with higher W scores (i.e., greater CR) showed an attenuated decline in ADNI-MEM (\(\beta = 0.48, p < 0.001\)) and ADNI-EF (\(\beta = 0.67, p < 0.001\)). When stratifying by disease stage, these interaction effects for ADNI-MEM remained similar among NC (\(\beta = 0.52, p = 0.028\)) and MCI participants (\(\beta = 0.43, p = 0.003\)), but were inverted in the AD dementia group (\(\beta = −0.91, p = 0.002\)). A similar pattern was observed for ADNI-EF (NC: \(\beta = 0.42, p = 0.077\); MCI: \(\beta = 0.59, p < 0.001\); AD dementia: \(\beta = −0.77, p = 0.081\)), although these effects were slightly weaker and mostly nonsignificant. The results for AD dementia
participants were significantly different from the other 2 groups (figure 2). Results for temporoparietal W scores were highly comparable (table 3). An overview of all W score models and the effect of each individual predictor on cognitive decline is available (tables e-3 and e-4, doi.org/10.5061/dryad.cb0qk1g). Finally, repeating our main analyses using only within-disease stage time points yielded results in the same direction for both whole-brain and temporoparietal W scores. The effects for NC participants were no longer significant (table e-5, doi.org/10.5061/dryad.cb0qk1g).

**Table 1** Characteristics of the total sample and according to disease stages

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>NC</th>
<th>MCI</th>
<th>AD dementia</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>839</td>
<td>175</td>
<td>437</td>
<td>227</td>
<td>—</td>
</tr>
</tbody>
</table>

**Demographics**

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y</td>
<td>73.9 (7.2)</td>
<td>74.6 (5.9)</td>
<td>73.3 (7.1)</td>
<td>74.3 (8.1)</td>
<td>0.086</td>
</tr>
<tr>
<td>Sex, % female</td>
<td>46.0</td>
<td>58.3</td>
<td>42.3</td>
<td>43.6</td>
<td>0.001†</td>
</tr>
<tr>
<td>APOE4, % positive</td>
<td>62.3</td>
<td>45.7</td>
<td>63.2</td>
<td>73.6</td>
<td>&lt;0.001±k</td>
</tr>
</tbody>
</table>

**Education, y**

|                      | 16.0 (14.0–18.0) | 16.0 (14.0–18.0) | 16.0 (14.0–18.0) | 16.0 (13.5–18.5) | 0.051   |

**Cognition**

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MMSE</td>
<td>26.7 (2.8)</td>
<td>29.1 (1.1)</td>
<td>27.5 (1.8)</td>
<td>23.2 (2.0)</td>
<td>&lt;0.001±k</td>
</tr>
<tr>
<td>ADAS-Cog 13</td>
<td>19.4 (10.2)</td>
<td>9.5 (4.3)</td>
<td>17.5 (6.8)</td>
<td>30.7 (8.3)</td>
<td>&lt;0.001±k</td>
</tr>
<tr>
<td>ADNI-MEM</td>
<td>−0.0 (0.9)</td>
<td>1.0 (0.6)</td>
<td>0.1 (0.7)</td>
<td>−0.9 (0.6)</td>
<td>&lt;0.001±k</td>
</tr>
<tr>
<td>ADNI-EF</td>
<td>−0.1 (1.0)</td>
<td>0.6 (0.7)</td>
<td>0.1 (0.8)</td>
<td>−0.9 (0.8)</td>
<td>&lt;0.001±k</td>
</tr>
</tbody>
</table>

**Brain measures**

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CSF Aβ42, ng/Lb</td>
<td>140.7 (29.4)</td>
<td>153.9 (31.6)</td>
<td>141.1 (29.6)</td>
<td>129.6 (21.9)</td>
<td>&lt;0.001±k</td>
</tr>
<tr>
<td>CSF p-tau, ng/Lb</td>
<td>47.0 (25.9)</td>
<td>39.3 (22.7)</td>
<td>46.8 (23.7)</td>
<td>53.7 (30.5)</td>
<td>&lt;0.001±k</td>
</tr>
<tr>
<td>CSF t-tau, ng/Lb</td>
<td>107.3 (58.5)</td>
<td>76.9 (39.0)</td>
<td>107.7 (58.3)</td>
<td>131.2 (61.1)</td>
<td>&lt;0.001±k</td>
</tr>
<tr>
<td>ICV, cm³</td>
<td>1,508.6 (160.0)</td>
<td>1,482.1 (145.1)</td>
<td>1,519.3 (157.6)</td>
<td>1,504.8 (173.3)</td>
<td>0.034†</td>
</tr>
<tr>
<td>W score (whole-brain)c</td>
<td>0.0 (1.0)</td>
<td>0.0 (1.0)</td>
<td>−0.0 (1.0)</td>
<td>0.0 (1.0)</td>
<td>0.807</td>
</tr>
<tr>
<td>W score (temporoparietal)d</td>
<td>0.0 (1.0)</td>
<td>−0.0 (0.9)</td>
<td>−0.0 (1.0)</td>
<td>0.1 (1.0)</td>
<td>0.475</td>
</tr>
<tr>
<td>W score (hippocampal)e</td>
<td>0.0 (1.0)</td>
<td>−0.3 (0.9)</td>
<td>0.1 (1.0)</td>
<td>0.1 (1.0)</td>
<td>&lt;0.001†</td>
</tr>
<tr>
<td>W score (WMH)f</td>
<td>0.0 (1.0)</td>
<td>0.0 (1.0)</td>
<td>0.0 (1.0)</td>
<td>−0.0 (1.1)</td>
<td>0.893</td>
</tr>
</tbody>
</table>

**Clinical progression**

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Follow-up, moa</td>
<td>24.0 (6.0–42.0)</td>
<td>36.0 (24.0–48.0)</td>
<td>36.0 (24.0–48.0)</td>
<td>12.0 (6.0–18.0)</td>
<td>&lt;0.001±k</td>
</tr>
<tr>
<td>Conversion to MCI, n (%)g</td>
<td>37 (4.4)</td>
<td>37 (21.1)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Conversion to AD, n (%)h</td>
<td>195 (23.2)</td>
<td>11 (6.3)</td>
<td>184 (42.1)</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Abbreviations: AD = Alzheimer disease; ADAS-Cog = Alzheimer’s Disease Assessment Scale–cognitive subscale; ADNI = Alzheimer's Disease Neuroimaging Initiative; ADNI-EF = standardized composite score based on a clock drawing task, animal and vegetable category fluency, the Trail-making task, Digit Span Backwards from the Wechsler Memory Scale–Revised, and the digit–symbol substitution task from the Wechsler Adult Intelligence Test–Revised; ADNI-MEM = standardized composite score based on the Alzheimer’s Disease Assessment Scale–cognitive subscale word list learning task, the Rey Auditory Verbal Learning Test, Logical Memory from the Wechsler Memory Scale–Revised, and the Mini-Mental State Examination recall task; GM = gray matter; ICV = intracranial volume, in mL; IQR = interquartile range; MCI = mild cognitive impairment; MMSE = Mini-Mental State Examination; NC = normal cognition; WMH = white matter hyperintensities.

Data are presented as mean (SD) unless indicated otherwise.

* Data expressed as median (IQR).
† These data were measured in CSF (available for 789 [β-amyloid and p-tau] and 776 [t-tau]/839 cases).
§ W score based on GM volume in a whole-brain mask.
∥ W score based on GM volume in a temporoparietal mask.
¶ W score based on GM volume in a hippocampal mask.
* W score based on WMH volume (available for 796/839 cases).
* This number includes 10 NC participants who additionally converted to AD at a later time point.
† Significant difference between NC vs MCI.
‡ Significant difference between NC vs AD dementia.
§ Significant difference between MCI vs AD dementia.
50
Similar results as observed in the primary analyses, although the sensitivity analyses yielded largely similar results as observed in the primary analyses, although we no longer found an exacerbating effect worse performance). The sensitivity analyses yielded largely similar results as observed in the primary analyses, although findings among the AD dementia group were not fully replicated.

**Discussion**

In the present study we investigated how CR, defined as the degree to which an individual’s observed cerebral damage was higher or lower than predicted from cognitive performance, affected clinical progression in AD. In an Aβ-positive sample of participants in different stages across the entire disease spectrum (i.e., NC, MCI, and AD dementia), we found that CR influenced diagnostic conversion risk and the rate of cognitive decline. Importantly, these effects were not uniform across disease stages. While CR was associated with attenuated clinical progression in NC and MCI participants, it related to exacerbated cognitive decline among individuals with AD dementia. A possible explanation for these findings is that every individual—regardless of CR—reaches the end stage of AD at roughly the same time or level of cerebral damage, and thus a delayed symptom onset or slower initial decline is inherently counterbalanced by a faster progression towards this end point. Understanding these complex longitudinal associations of CR with clinical progression may improve the accuracy of individualized prognoses for patients with AD.

Although previous literature in the context of AD has not consistently shown a relationship between CR and clinical progression, there are several studies within early stages of AD that also reported CR-related attenuated decline or delay in symptom onset. These studies defined CR by educational level, composite scores, or a “residual approach”–based measure (comparable to our W score method). Likewise, multiple studies have previously demonstrated that after the (delayed) onset of clinical symptoms, patients with higher CR show a steeper functional decline. This was found for symptomatic individuals with higher education, occupation, premorbid...
Figure 2 Trajectories of memory and executive functions by level of cognitive reserve

The Y-axis represents estimated marginal means. Cognitive reserve groups were created (for visualization purposes) by calculating tertiles in each disease stage—(A) normal cognition, (B) mild cognitive impairment, and (C) Alzheimer disease (AD) dementia—based on mean gray matter (GM)-based W scores in the temporoparietal mask. The lines displayed correspond to APOE4-positive men with a (disease stage–specific) average age and intracranial volume–adjusted GM volume in the temporoparietal cortex. In the linear mixed models, W scores were included as a continuous variable, and memory and executive functions composite scores were multiplied by 100. Note that the maximum follow-up time for AD dementia participants is shorter (i.e., 24 months) compared to the other 2 groups (i.e., 120 months). ADNI = Alzheimer’s Disease Neuroimaging Initiative; ADNI-EF = standardized composite score based on a clock drawing task, animal and vegetable category fluency, the Trail-making task, Digit Span Backwards from the Wechsler Memory Scale–Revised, and the digit–symbol substitution task from the Wechsler Adult Intelligence Test–Revised; ADNI-MEM = standardized composite score based on the Alzheimer’s Disease Assessment Scale–cognitive subscale word list learning task, the Rey Auditory Verbal Learning Test, Logical Memory from the Wechsler Memory Scale–Revised, and the Mini-Mental State Examination recall task.
Table 3: Relationship of W scores with cognitive decline in the total sample and according to disease stage

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>β</td>
<td>CI</td>
<td>p Value</td>
<td>β</td>
<td>CI</td>
<td>p Value</td>
<td>β</td>
<td>CI</td>
</tr>
<tr>
<td><strong>ADNI-MEM</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>−1.37</td>
<td>−1.49 to −1.24</td>
<td>&lt;0.001(^b)</td>
<td>−0.65</td>
<td>−0.98 to −0.31</td>
<td>&lt;0.001(^b)</td>
<td>−1.55</td>
<td>−1.71 to −1.39</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W score (temporoparietal)(^a)</td>
<td>108.65</td>
<td>103.36 to 113.95</td>
<td>&lt;0.001(^b)</td>
<td>98.50</td>
<td>76.78 to 120.21</td>
<td>&lt;0.001(^b)</td>
<td>101.07</td>
<td>92.45 to 109.69</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time × W score (temporoparietal)(^a)</td>
<td>0.45</td>
<td>0.25 to 0.66</td>
<td>&lt;0.001(^b)</td>
<td>0.48</td>
<td>−0.00 to 0.97</td>
<td>0.052</td>
<td>0.43</td>
<td>0.15 to 0.71</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ADNI-EF</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>−1.31</td>
<td>−1.46 to −1.17</td>
<td>&lt;0.001(^b)</td>
<td>−0.54</td>
<td>−0.88 to −0.20</td>
<td>0.002(^b)</td>
<td>−1.37</td>
<td>−1.56 to −1.19</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W score (temporoparietal)(^a)</td>
<td>74.23</td>
<td>66.21 to 82.24</td>
<td>&lt;0.001(^b)</td>
<td>51.15</td>
<td>23.61 to 78.70</td>
<td>&lt;0.001(^b)</td>
<td>52.14</td>
<td>38.75 to 65.53</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time × W score (temporoparietal)(^a)</td>
<td>0.72</td>
<td>0.48 to 0.95</td>
<td>&lt;0.001(^b)</td>
<td>0.40</td>
<td>−0.10 to 0.90</td>
<td>0.115</td>
<td>0.70</td>
<td>0.39 to 1.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: AD = Alzheimer disease; ADNI = Alzheimer’s Disease Neuroimaging Initiative; ADNI-EF = standardized composite score based on a clock drawing task, animal and vegetable category fluency, the Trail-making task, Digit Span Backwards from the Wechsler Memory Scale–Revised, and the digit–symbol substitution task from the Wechsler Adult Intelligence Test–Revised; ADNI-MEM = standardized composite score based on the Alzheimer’s Disease Assessment Scale–cognitive subscale word list learning task, the Rey Auditory Verbal Learning Test, Logical Memory from the Wechsler Memory Scale–Revised, and the Mini-Mental State Examination recall task; CI = confidence interval; GM = gray matter; NC = normal cognition; MCI = mild cognitive impairment. Regression coefficients are unstandardized. Apart from the predictors listed above, these models included age, sex, APOE\(^4\) status, and GM volume (temporoparietal), as well as its interaction with time (see table e-2, doi.org/10.5061/dryad.cb0qk1g, for an overview of all predictors). We multiplied ADNI-MEM and ADNI-EF scores by 100 to facilitate the interpretation and presentation of the results.

\(^a\) W score based on GM volume in a temporoparietal mask.

\(^b\) Significant effect.
of dementia in all participants. Results showed that higher CR, as defined by years of education, related to a shorter time period between the onset of accelerated decline and the dementia diagnosis. These results were interpreted as follows: individuals with higher CR remained in phases of gradual cognitive decline for longer (i.e., presumably reflecting NC and MCI), but the transition between (late) MCI and dementia was relatively steep. Similar to our findings, the authors thus demonstrated that the inflection from attenuation to exacerbation took place around (or just prior to) the onset of dementia.

Strengths of this study include our neuroimaging method to capture CR. W scores are not based on lifestyle proxies, but derived from more direct measures that constitute the core of the concept (i.e., cognition and cerebral damage). Moreover, although CR is often conceptualized as a brain quality that an individual has gradually developed and carried since adulthood, our W scores only reflect an individual’s current degree of CR, based on how well cerebral damage is tolerated at the present level of cognitive function. This is important, as CR presumably changes over time either due to alterations in lifestyle (e.g., decreases/increases in physical activity) or as a direct consequence of disease progression. While W scores may capture these changes, they could be missed when less dynamic CR measures are used, such as education. That is, once education is completed, this lifestyle proxy remains the same, regardless of midlife and late life exposures. Another advantage of our study is the fact that AD was established through both clinical consensus and the evaluation of PET or CSF biomarkers.

Our study has several limitations. First, there was a relatively short follow-up duration for participants with AD dementia (median 12 months), possibly because repeated cognitive testing is more challenging at clinically advanced stages. As cognitive changes may be harder to capture over shorter follow-up durations, this could have caused a bias towards the detection of CR effects in participants who were less affected. Second, the sample sizes in our stratified analyses were unequal, resulting in lower thresholds for significance in MCI (n = 437) vs AD dementia (n = 227) or NC (n = 175). In our interpretation of the results, we therefore also considered effect sizes, which are not affected by sample size. Third, although we ensured that our entire sample was ApoE-positive, this was based on PET imaging for some individuals, and on CSF measures for others. While high concordance between these 2 approaches has been demonstrated, other data suggest that they cannot be used interchangeably. Fourth, as expected, our W scores were collinear with the structural brain measures from which they were derived. On a group level, individuals with higher W scores (i.e., greater CR) thus had more cerebral damage. To assess the unique association of CR with clinical progression, we therefore adjusted the statistical models for the level of cerebral damage at baseline. Although the assumptions for multicollinearity were not violated, interdependence of predictors generally affects the accuracy of the estimation of coefficients in multivariate models to at least some degree. However, even despite the presumed inflation of CIs and standard errors, the effects for W scores still reached significance, which is likely related to our large sample size (i.e., >800 participants).

A more general limitation inherent to the W score method is that its ability to capture CR is a direct consequence of the measures included in the underlying model. If the...
neuroimaging or cognitive measures do not adequately reflect AD-related changes, then the resulting W scores become less accurate. As an example, we used different GM and WMH volumes to quantify cerebral damage, but only 2 of them (i.e., whole brain and temporoparietal GM volume) yielded W scores that related to education. We argue that hippocampal GM-based and WMH-based W scores were less valid measures of CR, likely because the structural measures in these underlying models did not optimally reflect cerebral damage across the AD spectrum. In fact, adjusted for covariates, ADAS-Cog 13 showed no correlation with WMH volume \((r = 0.00, p = 0.983)\) in our total sample, and while such a correlation did exist with hippocampal GM volume \((r = -0.56, p < 0.001)\), it was considerably lower for participants with AD dementia \((r = -0.33, p < 0.001)\). This suggests that hippocampal GM volume reaches a plateau in advanced clinical stages and thus has a more restricted dynamic range than whole-brain or temporoparietal GM volume. On a related note, the ADAS-Cog 13 (which was used as a cognitive measure to calculate W scores) covers multiple cognitive domains, but executive functions are underrepresented.\(^{47}\) Participants with executive dysfunction may therefore have received ADAS-Cog 13 scores that underestimated their true level of cognitive impairment. This could have led to less accurate W scores, which could in turn explain why the results for ADNI-EF were generally weaker compared to ADNI-MEM. Although our neuroimaging measure thus contains some variance of no interest, the nature of our results and their compatibility with the theoretical model nevertheless suggest that the (GM-based whole-brain and temporoparietal) W scores truly reflected CR.

Gaining a better understanding of how trajectories of clinical progression differ as a function of CR has both scientific and clinical value. Our findings help to clarify inconsistencies in the existing literature, and may further contribute to the conceptualization of CR.\(^{1}\) These scientific steps are ultimately important for the development of nonpharmacologic strategies to delay the onset of clinical AD by enhancing CR. Increased awareness of associations between CR and clinical progression also has implications for clinical trials. Matching patients based on their level of CR may improve the ability to identify intervention effects on clinical progression. Without adequately controlling for initial differences between treatment groups in CR, both false-positive and -negative results can arise (i.e., an apparent effect of treatment that is truly related to CR, or a true effect that is masked due to CR). Moreover, CR could act as a moderator between the intervention and clinical outcome, such that individuals with low CR benefit more (or less) than those with high CR. Finally, taking CR into account in a clinical setting may facilitate prognostic accuracy for individual patients. As evidenced by the increasing interest in biomarker-based prognostic tools for clinicians,\(^{48,49}\) a neuroimaging approach such as the W score method—especially when CR is estimated from readily available markers (e.g., medial temporal lobe atrophy and MMSE scores)—could be instrumental in this process.

**Study funding**
Research of VUMC Alzheimer Center is part of the Neurodegeneration program of Amsterdam Neuroscience. The VUMC Alzheimer Center is supported by Alzheimer Nederland, Het Genootschap, and Stichting VUMC funds. W.M. van der Flier holds the Pasman chair. F. Barkhof is supported by the NIHR UCLH biomedical research center. This research was funded by DELA and the Internationale Stichting Alzheimer Onderzoek (ISAO) (to R.O.).

**Acknowledgment**
Data collection and sharing for this project was funded by the Alzheimer’s Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer’s Association; Alzheimer’s Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; FujiRebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lundbeck; Merck & Co., Inc.; Mesro Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health. The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer’s Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

**Disclosure**
A.C. van Loenhoud, W.M. van der Flier, A.M. Wink, E. Dicks, C. Groot, and J. Twisk report no disclosures relevant to the manuscript. F. Barkhof is supported by the NIHR UCLH biomedical research center. P. Scheltens and R. Ossenkoppele report no disclosures relevant to the manuscript. Go to Neurology.org/N for full disclosures.

**Publication history**
Received by Neurology October 19, 2018. Accepted in final form March 8, 2019.
References


Cognitive reserve and clinical progression in Alzheimer disease: A paradoxical relationship
Anna Catharina van Loenhoud, Wiesje Maria van der Flier, Alle Meije Wink, et al.
Neurology 2019;93:e334-e346 Published Online before print July 2, 2019
DOI 10.1212/WNL.0000000000007821

This information is current as of July 2, 2019

Updated Information & Services
including high resolution figures, can be found at:
http://n.neurology.org/content/93/4/e334.full

References
This article cites 48 articles, 6 of which you can access for free at:
http://n.neurology.org/content/93/4/e334.full#ref-list-1

Citations
This article has been cited by 1 HighWire-hosted articles:
http://n.neurology.org/content/93/4/e334.full##otherarticles

Subspecialty Collections
This article, along with others on similar topics, appears in the following collection(s):
Alzheimer's disease
http://n.neurology.org/cgi/collection/alzheimers_disease
Volumetric MRI
http://n.neurology.org/cgi/collection/volumetric_mri

Permissions & Licensing
Information about reproducing this article in parts (figures,tables) or in its entirety can be found online at:
http://www.neurology.org/about/about_the_journal#permissions

Reprints
Information about ordering reprints can be found online:
http://n.neurology.org/subscribers/advertise