Predicting cognitive resilience from midlife lifestyle and multi-modal MRI: A 30-year prospective cohort study

Anya Topiwala1,*, Sana Suri1,2, Charlotte Allan1, Vyara Valkanova1, Nicola Filippini1, Claire E. Sexton1, Verena Heise1, Enikő Zsoldos1, Abda Mahmood1, Archana Singh-Manoux1,2, Clare E. Mackay1, Mika Kivimäki3, Klaus P. Ebmeier1

1 Department of Psychiatry, University of Oxford, Oxford, United Kingdom, 2 Wellcome Centre for Integrative Neuroimaging, Oxford, United Kingdom, 3 Centre for Research in Epidemiology and Population Health, INSERM, Villejuif, France, 4 Department of Epidemiology and Public Health, University College London, London, United Kingdom

Abstract

Background

There is significant heterogeneity in the clinical expression of structural brain abnormalities, including Alzheimer’s disease biomarkers. Some individuals preserve their memory despite the presence of risk factors or pathological brain changes, indicating resilience. We aimed to test whether resilient individuals could be distinguished from those who develop cognitive impairment, using sociodemographic variables and neuroimaging.

Methods

We included 550 older adults participating in the Whitehall II study with longitudinal data, cognitive test results, and multi-modal MRI. Hippocampal atrophy was defined as Scheltens Scores > 0. Resilient individuals (n = 184) were defined by high cognitive performance despite hippocampal atrophy (HA). Non-resilient participants (n = 133) were defined by low cognitive performance (≤ 1.5 standard deviations (S.D.) below the group mean) in the presence of HA. Dynamic and static exposures were evaluated for their ability to predict later resilience status using multivariable logistic regression. In a brain-wide analysis we tested for group differences in the integrity of white matter (structural connectivity) and resting-state networks (functional connectivity).

Findings

Younger age (OR: 0.87, 95% CI: 0.83 to 0.92, p < 0.001), higher premorbid FSIQ (OR: 1.06, 95% CI: 1.03 to 1.10, p < 0.0001) and social class (OR 1 vs. 3: 4.99, 95% CI: 1.30 to 19.16, p = 0.02, OR 2 vs. 3: 8.43, 95% CI: 1.80 to 39.45, p = 0.007) were independently associated with resilience. Resilient individuals could be differentiated from non-resilient participants by higher fractional anisotropy (FA), and less association between anterior and posterior...
resting state networks. Higher FA had a significantly more positive effect on cognitive performance in participants with HA, compared to those without.

**Conclusions**

Resilient individuals could be distinguished from those who developed impairments on the basis of sociodemographic characteristics, brain structural and functional connectivity, but not midlife lifestyles. There was a synergistic deleterious effect of hippocampal atrophy and poor white matter integrity on cognitive performance. Exploiting and supporting neural correlates of resilience could offer a fresh approach to postpone or avoid the appearance of clinical symptoms.

**Introduction**

The term resilience is frequently used in a psychological context, referring to the ability to cope in the face of stressful life events. There is recent evidence that the concept of resilience is also meaningful and relevant for cognitive outcomes [1, 2]. Some individuals appear to maintain high memory function despite the presence of risk factors for impairment [3] or pathological brain changes, indicating resilience. For example, approximately one fifth of individuals positive for Alzheimer biomarkers have normal cognitive function, regardless of whether these are neuropathological findings on autopsy [4], amyloid on positron emission tomography (PET) imaging [5] or abnormal levels of cerebrospinal fluid (CSF) tau and amyloid [2, 6]. Similarly, the extent of functional deficits following stroke and traumatic brain injury can differ markedly [7, 8].

Two related concepts, brain reserve and cognitive reserve, have been cited as mechanisms to explain resilience, although the terms are often used interchangeably in the literature [9]. Reserve is proposed to moderate between brain measures and cognition. Brain reserve refers to differences in the anatomical properties of the brain, such as its size, which confer a passive buffer against the effects of damage [10, 11]. Cognitive reserve theory was initially born out of epidemiological studies not examining the brain. Lower dementia risk was reported in those with certain socio-demographic characteristics, such as high education or intelligence, now often used as proxies for reserve [12, 13]. Cognitive reserve is proposed to reflect an active ability to optimize performance through the differential recruitment of brain networks, altered brain metabolism [14, 15], or alternative cognitive strategies. Functional brain imaging has revealed posterior-to-anterior or bilateral shifts in activation, and network de-differentiation in older adults, who successfully maintain their cognitive function [16, 17]. Impaired connectivity may predict cognitive impairment in Alzheimer’s disease [18]. Whether such strategies can compensate for structural brain adversity has not been investigated. Mounting evidence suggests that plasticity is possible even in the adult brain [19], making the distinction between passive brain reserve and active cognitive reserve models somewhat blurred and artificial [20]. Here we use the term resilience, which describes the observed phenomenon without mechanistic connotation.

Despite the vital implications for clinical care and research agenda, predictors and mechanisms underlying resilience have not been established [21]. Insights into cognitive resilience are likely to inform the interpretation and clinical significance of neuroimaging findings. Additionally, enhancing and supporting potential neural correlates of resilience may offer a fresh therapeutic approach, for example in Alzheimer’s disease, where current research is focusing with limited success on drugs to reduce the primary pathology [21–23]. There have
been recent attempts to identify socio-demographic features of individuals resilient to high
genetic Alzheimer’s disease risk [3, 24], beta amyloid and tau [25], and grey matter reductions
[26, 27]. However, the importance of midlife lifestyle and clinical factors, and structural and
functional brain connectivity has not yet been explored.

In our study, we defined and characterized resilience on the basis of cognitive performance
in the presence of an established biomarker for Alzheimer’s disease (hippocampal atrophy, HA).
Amongst those with hippocampal atrophy, we sought to identify socio-demographic,
clinical and brain corollaries of good (“resilient”) versus poor (“non-resilient”) outcomes. Our
first hypothesis, was that high social class and high full-scale IQ (FSIQ), often cited as markers
of cognitive reserve [10], would be associated with higher resilience [12, 28]. Alternative cogni-
tive strategies or differential network recruitment may facilitate maintained cognition. Hence
our second hypothesis was that resilient individuals would have superior structural brain con-
nectivity and distinct patterns of functional connectivity.

Materials and methods

Subjects and socio-demographic and lifestyle data

Five hundred and fifty participants were randomly selected from the Whitehall II cohort for
the imaging sub-study (2012–2016) [29]. Comparison to the larger sample is made in S1 Table.
Apart from self-exclusion due to inability to attend the examination in Oxford and take part in
an MRI study, post-hoc exclusions were due to incomplete or poor-quality images, gross struc-
tural abnormality, or missing cognitive or confounder data (Fig 1, S2 Table). Socio-demo-
graphic, health and lifestyle variables were measured prior to MRI over a follow-up period of
(Phase 7), 2007–9 (Phase 9), and 2011 (Phase 11) (see S2 Table). Age, sex, active bilingualism,
smoking, alcohol consumption, physical activity, depressive symptoms (General Health Ques-
tionnaire) and the Framingham Stroke Risk Score (FRS), were assessed by self-report ques-
tionnaire or clinical examination. Social class was determined according to occupation in
1991–3. Social networks were measured using questions derived by Berkman and Syme [30]
and summarized on a scale. Current cognitive function and subjective memory complaints
were assessed prior to the MRI scan with Montreal Cognitive Assessment (MoCA), Trail Mak-
ing Test (TMT A and B), Rey-Osterrieth Complex Figure (RCF) copying, RCF immediate and
delayed recall, Hopkins Verbal Learning Test (HVLT-R) total immediate (HVLT TR) and
delayed (HVLT DR) recall, Digit Span (DSF/DSB/DSS) and Digit Coding (all from the Wechs-
ler Adult Intelligent Scale-IV), lexical and semantic fluency, Boston Naming Test (BNT) and
Test of Premorbid Function (TOPF). Short-term memory (recall of a 20-word list) was tested
TOPF with sex and education adjustment [31]. History of Major Depressive Disorder was
assessed prior to the MRI scan using the Structured Clinical Interview for DSM-IV.

Participants gave written informed consent to be approached for the study and gave sepa-
rate written informed consent for every procedure included in the study, following a protocol
and consent form approved by the respective Ethics of Research Committees. Capacity for
consent was assessed by research workers trained in this procedure (EZ, AM). No formal psy-
chological tests or assessments were used to determine whether participants were able to pro-
vide written informed consent. The ethics committee approved this procedure.

Defining resilient and non-resilient participants

The resilient group of participants was defined by preserved cognitive function in the presence
of HA (<1.5 S.D. below mean on all cognitive tests). HA was selected on the basis of its known
Fig 1. Selection of sample and definition of resilient and non-resilient groups. Hippocampal atrophy was defined using the Scheltens' scale (>0) [32]. High cognitive performance was defined as <1.5 S.D below the mean on any cognitive test [33]. Low cognitive performance was defined as ≥1.5 S.D. below the mean on at least one cognitive test. Abbreviations—N—number, MRI—magnetic resonance imaging, DTI—diffusion tensor imaging.

https://doi.org/10.1371/journal.pone.0211273.g001
association with cognitive impairment in the literature. It has been identified as one of the five most robust biomarkers for Alzheimer’s disease and is a supportive sign in the revised NINCDS-ADRDA diagnostic criteria [5, 34]. We also empirically tested its relation to cross-sectional and longitudinal cognition in our sample (see S3 & S4 Tables, S1 Fig). HA was identified using the well-validated and widely used semi-quantitative Scheltens visual rating scale as any atrophy score \( > 0 \) [32, 35, 36] (S2 Fig). Although we have hippocampal volume measures, we elected to use a visual rating scale to identify hippocampal atrophy for a number of reasons: Visual ratings are highly sensitive and specific [32], in some cases more sensitive than volumetry [37]. They are clinically widely applicable, unlike volumetry, which necessitates specialist software and operator correction of segmentation. While volumetry assesses absolute, or at best relative volume (in relation to intracranial or grey matter volume), visual rating scales indicate anatomical abnormalities, such as visible and enlarged choroid fissures, enlarged lateral ventricles and hippocampal thickness. In other words, visual ratings account for grey matter, and inverse proportional changes in cerebrospinal fluid. Therefore, high scores may better reflect a pathological change, rather than smaller premorbid size. In contrast, volumetry (usually done after automatic segmentation of grey matter) only accounts for grey matter making such a distinction impossible. Inter- and intra-rater reliability is high (ICC = 0.8–0.9 and 0.7–0.9 respectively from our data) [28]. Reproducibility between quantitative approaches is not straightforward. Differences in operator judgement, scanner sequence, magnet strength and automated software can significantly impact results.

The (non-resilient) group of participants was defined by impaired cognitive performance in the presence of HA (defined as \( \geq 1.5 \) S.D. below mean [33] in the absence of validated normative data for all tests in this age group) on at least one test. Whilst links between hippocampal volume and verbal memory are well described [38, 39], as also demonstrated in the predictive power of hippocampal atrophy for the trajectory of verbal memory in the Whitehall II study (see S4 Table and S1 Fig) we chose additionally to verbal memory to consider other cognitive domains following empirical analyses demonstrating relationships between hippocampal volume and performance on multiple tests (see S3 and S5 Tables, S1 Text). In the absence of collateral and concurrent medical information about health and functional ability, we were unable to make a diagnosis of Mild Cognitive Impairment or dementia. However, the absence of a reported clinical diagnosis, ranges of impairment, participants’ ability to give informed consent and to make their way to Oxford from across the country independently, suggest that there was nobody with a probable diagnosis of dementia.

**MRI analysis**

All MRI scans were acquired between 2012 and 2015 at the University of Oxford Wellcome Centre for Integrative Neuroimaging (WIN), using a 3 Tesla Siemens Verio scanner. T1-weighted MPRAGE, fluid-attenuated inversion recovery (FLAIR), 60-direction diffusion tensor (DTI) and multi-band EPI 3T MRI sequences were used. T1-weighted images were processed using FSL tools (www.fmrib.ox.ac.uk/fsl) [40] and ‘fsl_anat (Beta version)’ (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/fsl_anat).

**Hippocampal metrics.** Two hippocampal metrics were examined. First, the presence of hippocampal atrophy, used to identify resilient and non-resilient subjects, was assessed using the Scheltens visual rating scale. Scores \( > 0 \) were used as an indication of atrophy [41]. Three clinicians, blind to measured volumes and participant characteristics independently rated each subject. Discrepancies were resolved following a consensus meeting. Second, hippocampal volumes were automatically extracted using FIRST [42], corrected for intracranial volume (ICV, by dividing by total intracranial volume’ 100, as is standard in the literature) and averaged.
across sides. Volume measures were only used for comparison (see S4 Fig), and when exploring accessory hypotheses of associations between cognitive performance measures and hippocampal volume.

**Voxel-based morphometry analysis.** To discount the possibility that the non-resilient subjects had a more advanced disease process (indicated by grey matter differences) than the resilient subjects, group differences in brain-wide grey matter density were examined. Voxel-based morphometry (VBM) is an objective method to compare grey matter density between groups in each voxel (smallest distinguishable image volume) of the structural image.

To avoid overfitting, confounders were included in each model if they were: 1) an established risk factor for the outcome measure, 2) not thought to lie on the causal chain between exposure and outcome. Premorbid FSIQ and social class were included to isolate brain differences independent of cognitive reserve. Adjustment was made in the VBM analysis for: age, sex, FRS, alcohol consumption, social class and FSIQ.

**Diffusion tensor imaging analysis.** Diffusion tensor images indicate the directional preference of water diffusion in neural tissue and allow inferences about the structural integrity of white matter tracts. Images were corrected for head movement and eddy currents and brain masks generated using BET. Fractional anisotropy (FA), mean (MD), axial (AD) and radial diffusivity (RD) maps were generated using DTIFit (http://fsl.fmrib.ox.ac.uk/fsl/fdt). Tract-Based Spatial Statistics (TBSS) were used [43] to perform voxelwise statistical analysis. Pre-processing prepared images for registration to standard space. Mean and skeletonized FA, MD, RD and AD images were created and thresholded. Lastly each FA, MD, RD and AD image was projected onto the relevant skeleton. To detect group differences between resilient and non-resilient subjects, a generalised linear model (GLM) was applied using permutation-based non-parametric testing (randomise) [44], correcting for multiple comparisons across space (threshold-free cluster enhancement, TFCE). Additionally for subsequent analysis, masks of the corpus callosum body were created using the ICBM-DTI-81 white-matter labels atlas [45] and used to extract mean FA indices. Adjustment for confounding included: age, sex, FRS, FSIQ, social class, alcohol consumption and depressive symptoms.

**Resting state functional connectivity analysis.** We used a hypothesis-free method to identify group differences in the synchronicity of responses across the brain between functionally distinct resting state network nodes [46]. Participants were scanned on a 3T Siemens Magnetom Verio (Erlangen, Germany) scanner with a 32-channel head coil, at the FMRIB Center, Oxford. T1-weighted structural MRI (multi-echo MPRAGE sequence with motion correction) and multiband echo-planar imaging rs-fMRI scans (voxel $\frac{1}{2}\text{mm}$ isotropic, TR $\frac{1}{4}$ 1.3 s, acquisition time $\frac{1}{4}$ 10 min 10 s, multi-slice acceleration factor $\frac{1}{4}$ 6, number of volumes $\frac{1}{4}$ 460) were acquired. Rs-fMRI data were pre-processed (motion correction, brain extraction, high-pass temporal filtering at 100s, field-map correction) using FSL tools. MELODIC pre-processing includes motion correction, brain extraction, high-pass temporal filtering (cut-off 150 seconds) and field map correction [40]. To reduce noisy components, a data-cleaning approach was used. Following single-subject independent components analysis (ICA), FMRIB’s ICA-based X-noiseifier (FIX) was used to automatically classify and regress out artefactual components [47]. FIX was “trained” on hand-classified ICA components on a matched training set (Whitehall II MB6). Data were registered affine to structural images using FLIRT [48]. FNIRT was used to register images into standard (MNI) space [49]. High dimensionality (dimensions = 100) ICA was performed on the pre-processed images to produce a study-specific template of spatial maps [50]. This template was used to extract time series using the dual regression approach [51]. Time courses were fed into FSLNets (v0.6) to perform network modeling [46]. Nodes were classified as ‘good’ (n = 58) or ‘bad’ (n = 42) (white matter, physiological noise, MRI or movement artifacts) [52]. The netmat was created and partial
correlations were calculated [46]. Matlab was used to reorder the nodes after a hierarchical clustering of the group-average correlation netmat (S3 Fig). A nodes x edges matrix was created. Group differences were examined with randomise, controlling for multiple comparisons (family-wise error), age, sex, FSIQ and social class.

**Other statistical analyses**

All analyses outlined below were done with R version 3.4.0 [53]. Descriptive data were summarized for all subjects, and separately by group, according to variable type and distribution (Table 1).

Logistic regression was used to identify significant predictors of resilience status (binary outcome: ‘resilient’ versus ‘non-resilient’, see above). Social class and FSIQ were added to the model as predictors. Education was not included due to concerns about multicollinearity with the concomitant inclusion of premorbid FSIQ (which was more closely associated with resilience (see Table 1), was education adjusted and suffered fewer ceiling effects). Longitudinal

Table 1. Sample characteristics.

<table>
<thead>
<tr>
<th></th>
<th>Mean (S.D.)¹, median (IQR)² or N (%)³</th>
<th>Comparison of resilient and non-resilient groups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Whole sample (N = 511)</td>
<td>Resilient (N = 184)</td>
</tr>
<tr>
<td>Age at scan, years</td>
<td>69.5 (5.3)¹</td>
<td>69.1 (4.9)¹</td>
</tr>
<tr>
<td></td>
<td>3.7 (2.5 to 4.9), p&lt;0.0001⁴</td>
<td></td>
</tr>
<tr>
<td>Female, N (%)</td>
<td>101 (19.8)²</td>
<td>30 (16.3)³</td>
</tr>
<tr>
<td></td>
<td>0.015, -0.09 to 0.1, p = 0.9⁶</td>
<td></td>
</tr>
<tr>
<td>Education, years</td>
<td>14 (5)²</td>
<td>16 (6)²</td>
</tr>
<tr>
<td></td>
<td>9931.5, p = 0.004³</td>
<td></td>
</tr>
<tr>
<td>FSIQ⁸</td>
<td>118.0 (10.3)²</td>
<td>122.3 (7.9)²</td>
</tr>
<tr>
<td></td>
<td>-5.7 (-7.9 to -3.6), p&lt;0.0001⁴</td>
<td></td>
</tr>
<tr>
<td>Social class, N (%)</td>
<td>4</td>
<td>0 (0)</td>
</tr>
<tr>
<td></td>
<td>14.8, p = 0.01²</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>4 (0.8)</td>
</tr>
<tr>
<td></td>
<td>15 (11.3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>392 (76.7)</td>
</tr>
<tr>
<td></td>
<td>102 (76.6)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>80 (15.7)³</td>
</tr>
<tr>
<td></td>
<td>16 (12.0)³</td>
<td></td>
</tr>
<tr>
<td>Systolic blood pressure, mmHg</td>
<td>120.6 (13.6)¹</td>
<td>120.6 (13.9)¹</td>
</tr>
<tr>
<td></td>
<td>1.4 (-1.8 to 4.5), p = 0.4⁴</td>
<td></td>
</tr>
<tr>
<td>Diastolic blood pressure, mmHg</td>
<td>75.6 (9.6)¹</td>
<td>75.5 (9.8)¹</td>
</tr>
<tr>
<td></td>
<td>0.8 (-1.3 to 3.0), p = 0.4⁴</td>
<td></td>
</tr>
<tr>
<td>History of Major Depressive disorder, N (%)¹⁰</td>
<td>90 (19.6)²</td>
<td>54 (29.3)³</td>
</tr>
<tr>
<td></td>
<td>3.8, (0.002 to 0.2), p = 0.05⁶</td>
<td></td>
</tr>
<tr>
<td>General Health Questionnaire, anxiety score</td>
<td>3.9 (2.6)¹</td>
<td>3.6 (2.3)¹</td>
</tr>
<tr>
<td></td>
<td>0.1 (-0.5 to 0.7), p = 0.7³</td>
<td></td>
</tr>
<tr>
<td>Alcohol, weekly units</td>
<td>8 (12)²</td>
<td>10 (13)²</td>
</tr>
<tr>
<td></td>
<td>11489, p = 0.4³</td>
<td></td>
</tr>
<tr>
<td>Moderate exercise, weekly N</td>
<td>2.3 (0.9)¹</td>
<td>2.1 (0.8)¹</td>
</tr>
<tr>
<td></td>
<td>0.01 (-0.2 to 0.2), p = 0.9⁴</td>
<td></td>
</tr>
<tr>
<td>Club attendance, weekly N¹¹</td>
<td>2.9 (1.2)¹</td>
<td>2.8 (1.2)¹</td>
</tr>
<tr>
<td></td>
<td>0.06 (-0.2 to 0.4), p = 0.7⁴</td>
<td></td>
</tr>
</tbody>
</table>

¹ Mean (standard deviation).
² Median (interquartile range).
³ Number (percentage).
⁴ Mean difference (95% confidence interval), p value.
⁵ W statistic, p value.
⁶ Odds ratio (95% confidence intervals), p value.
⁷ Pearson statistic, p value.
⁸ Estimated from Test of Premorbid Function with education and sex adjustment.
⁹ Social class based on occupation, Phase 3: 4 = skilled manual, 3 = skilled non-manual, 2 = managerial, 1 = professional.
¹⁰ Structured Clinical Interview for DSM IV (SCID) at time of MRI scan.
¹¹ At time of MRI scan.

Baseline (unless otherwise stated) characteristics of whole sample with DTI and adequate confounder information, and resilient and non-resilient subgroups.

**Abbreviations**: N–number, S.D.–standard deviation, IQR–interquartile range, FSIQ–full-scale intelligence quotient, DTI–diffusion tensor imaging.

https://doi.org/10.1371/journal.pone.0211273.t001
data on putative confounding factors were initially included in the model based on knowledge of the literature: age, sex, exercise (mean weekly hours of moderate activity over phases), Framingham Risk Score (mean 10 year stroke risk (%) over phases), alcohol (mean units per week over phases), smoker (proportion of phases), depressive symptoms (mean GHQ score over phases), and social networks (mean score over phases). Non-significant predictors (p>0.05) were excluded from the final model presented.

In order to test whether MRI-based predictors of resilience identified from preceding analyses were associated specifically with resilience rather than just better verbal memory, we used multiple linear regression on the whole sample with DTI (with and without HA) (n = 511). To reduce multiple testing, we used HVLT DR as measure for verbal memory. This was on the basis of a strong established association with the hippocampus in the literature [54] and in our empirical analysis (S3 Table). We hypothesized that a true MRI predictor of resilience would in the presence of hippocampal atrophy differ in its ability to predict memory function (dependent variable) from its predictive power in participants without hippocampal atrophy. We tested this adding the interaction term: predictor \* hippocampal volume to the independent variables. In the interests of a parsimonious analysis, we chose a single MRI measure to reflect 'white matter integrity', i.e. FA of the body of the corpus callosum. First, significant group differences existed in this region as shown by the TBSS analysis (Fig 2), and second, as the corpus callosum is a large tract separate from grey matter, it increases confidence that the FA measurement reflects only white matter. Models were adjusted for age, sex, Framingham Risk Score, alcohol consumption, depressive symptoms, premorbid FSIQ and social class. All hypothesis tests were two-sided and statistical significance was deemed as p<0.05.

**Results**

**Identification of resilient subjects**

The MRI Substudy sample was comparable to the whole Whitehall II Phase 11 cohort from which it had been selected randomly (S1 Table). A relative excess of women took part in the imaging study. Overall mean years of education and depressive symptoms were higher, BP measures lower in the subsample. Availability of lifetime risk data across previous phases of the sub-sample was satisfactory (S2 Table). A summary of participant characteristics is given in Table 1. The distribution of Scheltens hippocampal atrophy scores is illustrated in S2 Fig. Of 527 participants in the MRI Study, 317 showed Scheltens scores > 0 suggesting a degree of hippocampal atrophy. Of these, we identified 184 individuals with the resilience phenotype and 133 who were non-resilient participants. Mean intracranial volume-adjusted volumes were associated with atrophy scores, but this association was by no means perfect (S4 Fig). Seventy-two of those we defined as non-resilient reported subjective complaint about their cognition at the time of scanning, as opposed to twenty-eight in the resilient group. Empirical justification for our resilience definition is included in S1 Text: There was no significant group difference in hippocampal volume (as expected based on the selection methods), nor large disparities in grey matter density (S5 Fig). Therefore, it is unlikely that our resilience phenotype was an artefact of a less advanced disease process with less wide-spread volumetric changes. A small group (n = 42) of subjects had impairment on cognitive testing in the absence of small hippocampi, implying extra-hippocampal changes, thus not fitting our definitions of resilient or non-resilient participants, and therefore outside of our formulated research question.

**Sociodemographic correlates of resilience**

Higher premorbid FSIQ (OR per 1-point increase: 1.06, 95% CI: 1.03 to 1.10, p<0.0001) and social class (OR for 1 vs. 3: 4.99, 95% CI: 1.30 to 19.16, p = 0.02, OR for 2 vs. 3: 8.43, 95% CI:
1.80 to 39.45, \( p = 0.007 \) independently predicted resilience (after controlling for age and sex; Table 2). These odds were not confounded by exercise habits, social networks, vascular risk (Framingham Risk Score), depressive symptoms, bilingualism, alcohol consumption or smoking, measured repeatedly over the 30 years, which also did not independently predict resilience.

**Structural connectivity in resilient and non-resilient participants**

In the brain-wide TBSS analyses, resilient individuals showed increased fractional anisotropy (FA), and reduced mean diffusivity (MD) and radial diffusivity (RD) in multiple tracts compared with non-resilient individuals, but no differences in axial diffusivity (AD; Fig 2). Furthermore, in the separate regression analyses of the entire MRI sample irrespective of
hippocampal atrophy (n = 511), there was a significant interaction ($\beta$ = -14.9.08, 95% CI: -23.21 to -6.64, $P = 0.0004$) between corpus callosum FA and hippocampal volume in predicting verbal memory (Fig 3). Strikingly, in subjects with large hippocampi (i.e. >2.25% ICV), corpus callosum FA did not predict cognitive performance. Amongst those with smaller hippocampi, FA became increasingly important (S1 Fig). In contrast, interaction terms for hippocampal volume and social class or FSIQ were not significant.

**Functional connectivity in resilient and non-resilient participants**

The functional connectivity between nodes #52 and #35 was weaker in the resilient group compared to the non-resilient group (Fig 4). One node (#52) of the pair was in frontal pole of the central executive network (CEN), and the other (#35) in the inferior parietal lobe (IPL) of the default mode network [55]. The population average connectivity matrix (S3 Fig) indicates that for the entire sample, the functional association between nodes #52 and #35 was weakly negative. However, there is not necessarily a relationship between the strength (or even sign) of the population average connection between two nodes and the extent to which inter-subject variability in the connection strength correlates with a given behavioural measure [46].

**Discussion**

We combined multimodal neuroimaging and functional cognitive data to define a phenotype of resilience. Sociodemographic factors, but not lifestyle, were associated with resilience. Resilient subjects’ brains were characterized by higher structural integrity of white matter tracts and weaker inter-network functional connections.

**Comparison with other studies**

As predicted, high premorbid FSIQ and social class (in addition to age) were independently associated with resilience, in the presence of HA [28]. However, it appears these effects can be explained by shared variance with cognitive test performance (general intelligence or ‘G’

---

**Table 2. Sociodemographic correlates of resilience.**

<table>
<thead>
<tr>
<th>Predictor</th>
<th>N</th>
<th>Unadjusted odds ratio (95% CI)</th>
<th>Adjusted$^3$ odds ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td>0.88 (0.84 to 0.92), $p&lt;0.001$</td>
<td>0.87 (0.83 to 0.92), $p&lt;0.001$</td>
<td></td>
</tr>
<tr>
<td>Male Sex</td>
<td>266</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female Sex</td>
<td>51</td>
<td>1.04 (0.57 to 1.91), $p = 0.90$</td>
<td>2.22 (0.98 to 5.01), $p = 0.06$</td>
</tr>
<tr>
<td>FSIQ$^1$</td>
<td>1.07 (1.04 to 1.10), $p&lt;0.001$</td>
<td>1.06 (1.03 to 1.10), $p&lt;0.001$</td>
<td></td>
</tr>
<tr>
<td>Social class$^2$</td>
<td>3</td>
<td>9.38 (2.70 to 32.59), $p&lt;0.001$</td>
<td>8.43 (1.80 to 39.45), $p = 0.007$</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5.15 (1.66 to 16.00), $p = 0.005$</td>
<td>4.99 (1.30 to 19.16), $p = 0.02$</td>
</tr>
</tbody>
</table>

1 Premorbid FSIQ (education and sex adjusted) according to Test of Premorbid Function performed at the time of scan.
2 Social class based on occupation at Phase 3: 1 = professional, 2 = managerial, 3 = skilled non-manual. No subjects included in this subset analysis were in social class 4 (manual).
3 For all other predictors in model lists in column one. Results from logistic regression analyses, predicting odds of resilience (visually rated hippocampal atrophy and absent cognitive impairment, n = 184) vs. non-resilient participants (visually rated hippocampal atrophy and cognitive impairment on $\geq$1 test, n = 133).

Abbreviations: CI–confidence interval, $p$–p value, FSIQ–full-scale intelligence quotient.

https://doi.org/10.1371/journal.pone.0211273.t002
factor), rather than a synergistic compensatory effect. In the Whitehall II sample at least, such factors, often referred to as indices of cognitive reserve, reflect higher baseline memory performance rather than greater compensatory capacity [10]. Our finding is in agreement with previous Whitehall II findings in dementia [56]. It also raises concerns that links between cognitive
A

\[ p = 0.0418 \]

B

![Box plots comparing groups R and NR](image)

R

NR
reserve indices and **cognitive impairment** could result from detection bias, if only cross-sectional cognitive performance is examined.

White matter microstructure appears to partially mitigate the effect of HA on verbal memory in our sample, suggesting a potential ‘synergistic’ deleterious effect of hippocampal and white matter damage on cognitive performance. It is not clear whether resilient participants in our study have avoided age-related decline in white matter (reduced “brain battering”) and have greater capacity to tolerate damage (brain reserve), or whether differences are due to plasticity. Compensatory changes, including increased number of axons or increased myelination are possible, as have been reported in response to brain lesions [57].

Our functional connectivity results appear to challenge the general notion that “stronger is better”. Most resting-state fMRI studies to date have only described intra-network connections, often studying large brain networks as a whole. Our approach allowed us to examine both intra- and inter-network connectivity, which has been of greater interest in recent studies. While we observed no group differences in intra-network connections, participants in the resilient group appeared to have weaker functional association (in the form of an inverse partial correlation) between the CEN and DMN compared with non-resilient subjects. Functional brain networks in the elderly typically become less distinct due to an increase in inter-network strength of connections along with a relative decrease in intra-network connections in older compared to younger adults [58]. Here, a relatively weaker inter-network connection between the CEN and default mode nodes was associated with the resilience phenotype, suggesting that this group may maintain more distinct network structure in older age.

The CEN has been linked to demanding externally orientated tasks [59] and the DMN, mediates internally directed self-referential processes [60]. There are several reports that anti-correlation between the CEN and default mode networks is important for brain function [61, 62] and correlates with higher cognitive performance [63]. However, others have described a u-shaped relationship with aging, with compensatory increases in segregation of networks linked to reductions in structural integrity [64]. Our results appear to be more consistent with the latter, however without longitudinal imaging, it is impossible to determine whether the functional differences observed here are a compensation for pathological changes within the brain (HA) or represent lifelong differences between resilient and non-resilient groups. We therefore interpret the functional connectivity results with caution.

**Strengths and limitations**

We consider three advances of our work over existing studies. First, the depth and duration of the phenotyping allowed a thorough long-term examination of lifestyle contributions to resilience, largely ignored by previous cross-sectional studies [25, 27]. Second, the breadth of the cognitive tests and measurement of longitudinal decline have not been considered in published work [25]. Third, use of multi-modal MRI and voxel-based brain analyses allowed us to examine the relationship between resilience, specifically defined, and structural and functional brain connectivity.

In spite of these strengths, some limitations need to be highlighted. The brain correlates of resilience we identified were in the context of analyses examining potential moderators of the...
effect of hippocampal atrophy on cognition. Other brain/cognition relationships may have entirely different moderating variables.

We used cross-sectional hippocampal atrophy ratings, so cannot exclude the possibility, of unlikely face validity, that these reflect lifelong differences rather than age-related or even pathological change. Whilst longitudinal imaging could be helpful, scan intervals of more than a couple of years add further confounders, because of developments in the physical scanner and imaging sequences. Links between hippocampal volume and current and future cognition are well established [38], and replicated in our sample. Whilst HA is a diagnostic marker for Alzheimer’s disease [34], we do not wish to claim subjects in our sample had specific markers of Alzheimer’s pathology in the absence of knowledge of tau or amyloid status. HA has also been reported in depression, neurotoxic injuries, other causes of dementia, and in moderate drinkers [41]. On the other hand, hippocampal sparing has been described in atypical Alzheimer’s disease [65] so we may have systematically excluded some participants with early Alzheimer’s pathology but large hippocampi. However, as our primary interest was in exploring tolerance to hippocampal atrophy, irrespective of its etiology, we do not think this detracts from the findings. We used a clinical visual rating scale to group subjects on the basis of HA, as there are no established ‘normal’ vs. ‘abnormal’ cut offs for volumetric data. The Scheltens scale has been widely used and is well validated, is applicable to clinical settings, and aligns to some extent with volumetric measures in our sample [32, 35, 36].

Standard deviation cut-offs rather than normative values were used to categorize subjects on cross-sectional cognitive performance. Our cut-offs do not necessarily generalize to other samples or indicate pathological deficits. Many of the tests we used do not have robust normative data. Even for the better investigated MoCA, there is no fine-grained breakdown for effects of age or education, and little data on older subjects relevant to our sample. There is continuing debate about what constitutes the ‘best’ normative comparison group for neurocognitive tests, and studies vary on the stringency with which they exclude cognitively impaired subjects. As with any observational study, we cannot deny the possibility that some individuals we defined as ‘resilient’, will develop cognitive impairment after our study period (right censoring).

We are uncertain about the generalizability of our findings, as the Whitehall II cohort is not representative of the wider population. By example, it has a low proportion of women, reflective of the 1980s UK civil service. Lastly, we cannot exclude a selection bias. Those most cognitively impaired are not unlikely to have been included in the study, which required travel to Oxford and a neuropsychological assessment. Although this may have attenuated some of the associations, it is an unlikely source of major bias.

**Implications of findings**

This study provides one avenue of investigation into how brain findings may translate into future clinical symptoms. We envisage several potential clinical and research implications of our study. First, clinical interpretation of neuroimaging findings must take account of sociodemographic factors. Second, future research in cognitive impairment and dementia should not ignore multiple synergistic pathological contributions. Finally, a fresh approach to delaying symptoms in those with pathological changes may be exploiting and supporting neural correlates of resilience, such as white matter microstructure. To date, therapeutic research has been focusing on agents to reduce a single primary pathology, with limited success [22]. Exercise [66], vascular risk [67] and brain stimulation with electrical or magnetic currents have been associated with white matter integrity [68], and even neuroplasticity. Fruitful leads may therefore include tight control of cardiovascular risk factors with antihypertensives and lipid-
lowering drugs, physical training programs or even brain stimulation. The approach may be particularly suitable in individuals whose scans have revealed early indicators of structural damage but exhibit minimum functional deterioration. Even small increases in resilience may postpone or avoid onset of symptoms in a clinically significant fashion.

Supporting information

S1 Table. Comparison of the imaging sample with the whole Whitehall II cohort at Phase 11. The MRI sample consisted of all those in the Oxford imaging sub-study with at least a structural MRI image (but not necessarily diffusion tensor imaging) and cross-sectional socio-demographic and lifestyle data.

S2 Table. Completeness of lifestyle and health data at each study phase.

S3 Table. Cross-sectional cognitive performance and hippocampal volume. P-values result from hypothesis tests (likelihood tests) comparing regression models with and without hippocampal volume. N = 511. Models were adjusted for: age, sex, FSIQ and social class.

S4 Table. Longitudinal verbal memory decline and hippocampal atrophy status. Estimates are from mixed effects models (binomial regression) fitted longitudinal memory test scores over time (study Phase 11 to 12). Models were adjusted for: age, sex, FSIQ, social class, time from study baseline and hippocampal atrophy status. Estimates represent the odds of memory recall. The ‘Time x hippocampal atrophy’ interaction term tests whether there was a significant difference in memory decline over time between subjects with and without hippocampal atrophy.

S5 Table. Cross-sectional cognitive test data for resilient and non-resilient groups.

S1 Fig. Modelling longitudinal change in memory test scores according to hippocampal size for a typical subject. Predictions are made on the basis of mixed effects models (using N = 511 subjects) with cognitive testing (memory recall/20 words) performed before MRI at study Phase 11 and after the MRI in Phase 12, for a mean age of 75 years (at scan) and premorbid IQ of 118, according to cross-sectional hippocampal atrophy. Hippocampal atrophy is defined on the basis of the Scheltens scale (score >0).

S2 Fig. Distribution of Scheltens ratings for hippocampal atrophy on whole imaged sample. Counts indicate number of participants from the whole imaged sample with hippocampal atrophy (left (L) and right (R) sides) as rated by the semi-quantitative Scheltens scale. N = 511.

S3 Fig. Functional connectome consisting of hierarchical ordering of the 58 functionally distinct brain regions as network nodes, and the connectivities between these as network edges. The strengths of the estimated network edges are elements in a N_{nodes} x N_{nodes} network matrix. Boxes above the diagonal signify full correlations (thought to represent direct and indirect connections) and those below the diagonal represent partial correlations (used to infer only direct connections between nodes). Based on data from 317 subjects.
S4 Fig. Relationship between Scheltens ratings of hippocampal atrophy and volumetric hippocampal measures. Violin plots (outline includes all data points; red diamond indicates mean) display Scheltens visual rating scores against automated hippocampal (as a % of total intracranial volume) for left and right sides. Red diamonds indicate mean hippocampal volumes. N = 511. (PDF)

S5 Fig. Grey matter in resilient and non-resilient subjects. Voxel-based morphometry analysis results. Orange/red blobs represent voxels where there is significantly higher grey matter density in resilient (N = 184) compared to cognitively impaired (N = 133) groups. Colour bar indicates 1-p values. Analysis was adjusted for: age, sex, alcohol consumption, Framingham Risk Score, FSIQ, social class and TFCE-corrected. (PDF)

S1 Text. Empirical justification for choice of resilience metrics. (PDF)

Acknowledgments
We would like to thank Matteo Bastiani for his comments on diffusion tensor imaging.

Author Contributions
Conceptualization: Anya Topiwala, Archana Singh-Manoux, Clare E. Mackay, Mika Kivimäki, Klaus P. Ebmeier.
Data curation: Enikő Zsoldos, Archana Singh-Manoux, Mika Kivimäki, Klaus P. Ebmeier.
Formal analysis: Anya Topiwala, Nicola Filippini, Claire E. Sexton.
Funding acquisition: Archana Singh-Manoux, Mika Kivimäki, Klaus P. Ebmeier.
Investigation: Anya Topiwala.
Methodology: Anya Topiwala, Klaus P. Ebmeier.
Project administration: Enikő Zsoldos, Abda Mahmood.
Supervision: Archana Singh-Manoux, Clare E. Mackay, Mika Kivimäki, Klaus P. Ebmeier.
Visualization: Anya Topiwala.
Writing – original draft: Anya Topiwala.
Writing – review & editing: Sana Suri, Charlotte Allan, Vyara Valkanova, Nicola Filippini, Claire E. Sexton, Verena Heise, Enikő Zsoldos, Abda Mahmood, Archana Singh-Manoux, Clare E. Mackay, Mika Kivimäki, Klaus P. Ebmeier.

References


