Cortical gyriﬁcation in relation to age and cognition in older adults

Sander Lamballais a, Elisabeth J. Vinke a,b, Meike W. Vernooij a,b, M. Arfan Ikram a, Ryan L. Muetzel a,c,*

a Department of Epidemiology, Erasmus MC University Medical Center Rotterdam, the Netherlands
b Department of Radiology and Nuclear Medicine, Erasmus MC University Medical Center Rotterdam, the Netherlands
c Department of Child and Adolescent Psychiatry, Erasmus MC University Medical Center Rotterdam, the Netherlands

A R T I C L E   I N F O

Keywords:
• Gyriﬁcation
• Aging
• Cognition
• Structural MRI

A B S T R A C T

Gyriﬁcation of the cerebral cortex changes with aging and relates to development of cognitive function during early life and midlife. Little is known about how gyriﬁcation relates to age and cognitive function later in life. We investigated this in 4397 individuals (mean age: 63.5 years, range: 45.7 to 97.9) from the Rotterdam Study, a population-based cohort. Global and local gyriﬁcation were assessed from T1-weighted images. A measure for global cognition, the g-factor, was calculated from ﬁve cognitive tests. Older age was associated with lower gyriﬁcation (mean difference per year = −0.0021; 95% conﬁdence interval = −0.0025; −0.0017). Non-linear terms did not improve the models. Age related to lower gyriﬁcation in the parietal, frontal, temporal and occipital regions, and higher gyriﬁcation in the medial prefrontal cortex. Higher levels of the g-factor were associated with higher global gyriﬁcation (mean difference per g-factor unit = 0.0044; 95% conﬁdence interval = 0.0015; 0.0073). Age and the g-factor did not interact in relation to gyriﬁcation (p > 0.05). The g-factor bilaterally associated with gyriﬁcation in three distinct clusters. The ﬁrst cluster encompassed the superior temporal gyrus, the insular cortex and the postcentral gyrus, the second cluster the lingual gyrus and the precuneus, and the third cluster the orbitofrontal cortex. These clusters largely remained statistically signiﬁcant after correction for cortical surface area. Overall, the results support the notion that gyriﬁcation varies with aging and cognition during and after midlife, and suggest that gyriﬁcation is a potential marker for age-related brain and cognitive decline beyond midlife.

1. Introduction

Gyriﬁcation is one of the most fundamental and distinguishing properties of the human cerebral cortex. The folding patterns of the cortex are highly heritable (Docherty et al., 2015), evolutionarily conserved, and similar amongst closely related animal species (Zilles et al., 2013). Abnormalities in gyriﬁcation, such as polymicrogyria and pachygryria, lead to altered brain function, which can manifest as impairments in speech and cognition. Similarly, both global and regional abnormalities in gyriﬁcation have been found in patients with autism (Duret et al., 2018; Blanken et al., 2015), schizophrenia (Matsuda and Ohi, 2018; Cao et al., 2017), and bipolar disorder (Cao et al., 2017). A deeper understanding of gyriﬁcation may therefore lead to better insight into a broad range of diseases.

Gyriﬁcation changes with age and in turn affects cognitive function (Cao et al., 2017; White et al., 2010; Gregory et al., 2016; Hogstrom et al., 2013). The global degree of gyriﬁcation is often expressed as the Gyrification Index (GI). The GI peaks during childhood, rapidly declines during adolescence and the decline slows down as adulthood progresses (Cao et al., 2017; White et al., 2010; Gregory et al., 2016). Regional patterns of gyriﬁcation can be quantiﬁed with the Local Gyrification Index (LGI) (Schaer et al., 2008). The regions surrounding the angular gyrus, i.e. the parietal cortex, seem most prone to age-related decline in the LGI (Hogstrom et al., 2013). The association between the LGI and cognition has been studied in both pediatric and adult cohorts, and it showed the strongest effect in the frontal and parietal regions as well as the temporoparietal junction (Gregory et al., 2016). These ﬁndings consolidate the relevance of gyriﬁcation in the normal development of the brain.

Several knowledge gaps still remain. Limited work exists on cortical gyriﬁcation during middle adulthood, i.e. 40–65 years of age, and late adulthood, i.e. beyond 65 years of age. Other aspects of the cerebral

* Corresponding author. Department of Child and Adolescent Psychiatry, Erasmus MC University Medical Center Rotterdam, PO Box 2040, 3000, CA, Rotterdam, the Netherlands.
E-mail address: r.muetzel@erasmusmc.nl (R.L. Muetzel).

https://doi.org/10.1016/j.neuroimage.2020.116637
Received 9 October 2019; Received in revised form 31 January 2020; Accepted 12 February 2020
Available online 17 February 2020
1053-8119/© 2020 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
cortex – such as cortical surface area – change significantly during middle and late adulthood (Vinke et al., 2018). Furthermore, atrophy of the cerebral cortex seems to accelerates towards the end of life (Battaglini et al., 2019), and the rates of atrophy differ between brain regions (Vinke et al., 2018). How gyriﬁcation changes during late life and how the changes are distributed across the brain remains to be elucidated. Similarly, cognitive function declines in aging, which in turn may affect if and how cognition and gyriﬁcation relate. Finally, most previous studies were performed in clinical samples or clinic-based settings, limiting the external validity of the ﬁndings. The use of population-based studies would allow for better generalization of the results.

The aim of the present study was to elucidate the associations of age and cognition with gyriﬁcation during middle and late adulthood. The study was performed using data from the Rotterdam Study cohort, a prospective population-based cohort study of individuals aged 40 years and higher. We hypothesized that age and the GI showed a non-linear association across middle and late adulthood, where the rate of loss of gyriﬁcation accelerates with age. Furthermore, based on previous volumetric work, we expected to ﬁnd that the shape of the association between age and the LGI differed across the brain, with regions near the angular gyrus showing the fastest decline towards the end of life. Finally, we hypothesized that cognition positively associated with the GI, and with the LGI in frontal and temporal regions.

2. Methods

2.1. Study population

The Rotterdam Study is a prospective cohort study based in the Ommoord district of Rotterdam, the Netherlands, that has been ongoing since 1989 (Ikram et al., 2017). The second recruitment wave started in 2000, and the third wave in 2006. All participants are re-invited for an interview and in-person examinations every 4–6 years. The study has included 14,926 participants 45 years of age and older. Neuroimaging was introduced in 2005 (Ikram et al., 2015). The current study population included individuals who were eligible to participate in a research center visit between 2006 and 2015 with cognitive testing and neuroimaging (n = 6647). Of these, 38 had no cognitive test battery data, 980 had incomplete data, 417 did not participate in the MRI study, in 462 the image surface tessellation in FreeSurfer failed, and 145 were excluded due to poor quality of the T1-weighted images. We further excluded participants with prevalent stroke (n = 126) or prevalent dementia (n = 82). The ﬁnal sample consisted of 4397 participants. A ﬂow chart of the study population is shown in Supplementary Fig. 1. The Rotterdam Study has been approved by the Medical Ethics Committee of the Erasmus MC (registration number MEC 02.1015). All participants provided written informed consent.

2.2. Assessment of cognitive function

All participants underwent a cognitive test battery (Hoogendam et al., 2014). The battery consisted of ﬁve tests, each assessing different cognitive domains. The ﬁrst test was the 15-word learning test (15WLT), to assess verbal learning and verbal memory (Bleecker et al., 1988). The 15WLT consisted of three trials where 15 words were presented visually, and after each trial participants had to name all words they could remember (i.e. immediate recall). At least 10 min after the third trial, participants were again asked to name all words that they could still remember (i.e. delayed recall). We used the number of words in the delayed condition as the measurement outcome. The second test was the Stroop task (Houx et al., 1993), a task that assesses selective attention and automaticity. Participants had to read aloud the names of colors (red, green, blue, yellow) as fast and flawless as possible. The words were printed on paper in mismatching colors (e.g. “blue” printed in the color red) to interfere with the naming process. The time to read all words was adjusted for the number of errors by calculating the time per word and adding one-and-a-half that time for each error. Thus, the Stroop task is inversely coded compared to the other tests, where a higher score relates worse performance. The total was then log transformed and used as the outcome measure. The third test was the letter-digit substitution test (LDST) (van der Elst et al., 2006), in which participants have to write down the corresponding digits next to letters according to a dictionary table. This assesses processing speed as well as executive function. Fourth, a word ﬂuency test (WFT) was administered to assess efﬁciency of searching long term memory (Welsh et al., 1994). Participants had to name as many animal species in a span of 1 min, with the total number of unique species as the outcome. Finally, we administered the Purdue pegboard test (PPB) (Tiffin and Asher, 1948), where participants had to place small metal pins into holes across three trials: left hand only, right hand only, and both hands. The sum of the number of pins over all trials was used as a measure for ﬁne motor dexterity and psychomotor ability. To summarize all tests into a single score for global cognition, known as the g-factor, we used principal component analysis and isolated the ﬁrst component (Deary, 2014). The g-factor explained 50.6% of the variance amongst the cognitive tests which is in agreement with previous literature (Deary, 2012).

2.3. Image acquisition

Neuroimaging was performed on a 1.5T magnetic resonance imaging (MRI) scanner with an eight-channel head coil (GE Sigma Excite, General Electric Healthcare, Milwaukee, USA). The imaging sequences have been described extensively elsewhere (Ikram et al., 2015). Axial T1-weighted images were collected using a 3D Spoiled Gradient Recalled sequence (TR = 13.8 ms, TE = 2.8 ms, T1 = 400 ms, ﬂip angle = 20°, bandwidth = 12.5 kHz, voxel size = 0.8 mm isotropic). The images were subsequently stored in an extensible neuroimaging archive toolkit (XNAT) database (Marcus et al., 2007).

2.4. Image processing

Images were processed using the FreeSurfer analysis suite (version 6.0) (Fischl, 2012). The standard reconstruction was conducted, where non-brain tissue was removed, voxel intensities were corrected for B1 ﬁeld inhomogeneities, voxels were segmented into white matter, gray matter and cerebrospinal ﬂuid, and surface-based models of gray and white matter were generated. The GI was calculated as the ratio between the outer contour of the cortex and the pial surface of the whole cerebrum. The LGI was estimated at each vertex along the cortical ribbon (Schaer et al., 2008, 2012), and each vertex was automatically assigned an anatomical label according to a predeﬁned atlas (Desikan et al., 2006). All measures were co-registered to a standard stereotaxic space and smoothed with a full-width half-max Gaussian kernel, 5 mm for the LGI given inherent smoothness and 10 mm for all other measures.

A multistep procedure was used to identify datasets of insufﬁcient quality for analysis. First, we used an automated tool to obtain a quality metric for each T1-weighted scan that assesses artifacts related to motion (White et al., 2018). Next, we visually inspected FreeSurfer reconstructions from 200 randomly selected scans. The visual ratings consisted of inspecting segmented brain images in the coronal, sagittal and axial directions, as well as 3D reconstructions of the pial surface. The segmentation was rated as a fail if FreeSurfer did not succeed to consistently trace the white and pial surfaces. Next, we established that the automated quality metric value predicted strongly whether a test passed or failed. We subsequently set a threshold above which all scans were of sufﬁcient quality, and all scans below the threshold were excluded.

2.5. Measurement of covariates

Hypertension was deﬁned as a resting blood pressure exceeding 140/90 mmHg or the use of blood pressure lowering medication. Blood pressure was measured twice with a sphygmomanometer after 5 min of...
rest, and the average of the two measurements was used. Use of blood pressure lowering medication was derived from information collected by a physician at the research center. Alcohol use was assessed during home interviews with questions based on beer, wine, liquor and other alcoholic beverages such as sherry and port. Based on these data, an established method was used to calculate alcohol in grams per day (Vliegenthart et al., 2002). BMI was calculating using the height and weight obtained during the research center visit. Smoking status was obtained during the home visit interview and classified as never smokers, past smokers or current smokers. Education level was assessed during the visit interview and classified into four categories according to the United Nations Educational, Scientific and Cultural Organization classification: primary (no or primary education), low (unfinished secondary and lower vocational), intermediate (secondary or intermediate vocational) or high education (higher vocational or university).

2.6. Statistical analyses

All statistical analyses were performed in R 3.4.3 (R Core Team, R, 2016). To assess the relation of age and cognition with the GI we used linear regression models. Surface-based LGI analyses were performed to study the spatial distributions of these associations along the cortex. This was done with vertex-wise analyses using the R package QDECR (https://github.com/slamballais/QDECR). Resulting p-value maps were corrected for multiple comparisons at the vertex level using Gaussian Monte Carlo Simulations (Hagler et al., 2006). Surface-based analyses on cortical thickness and similar measures may show non-Gaussian patterns of spatial correlations, which would increase the false positive rate higher than 0.05. We therefore set the cluster forming threshold to \( p < 0.001 \), as this has shown high correspondence with actual permutation testing across all surface measures (Greve and Fischl, 2018). We further applied Bonferroni correction to account for analyzing both hemispheres separately (i.e. \( p < 0.025 \) cluster-wise).

Age-related atrophy of the brain accelerates with age, which may also affect cortical gyriﬁcation. We therefore studied three types of associations between age and cortical gyriﬁcation: (1) a linear age term, (2) orthogonal linear and quadratic age terms, (3) a B-spline for age with two or three degrees of freedom. The spline knot for the two-fold spline was set at the median age, and the knots of the three-fold spline at the first and second tertiles. The shape of the relationship between age and gyriﬁcation was assessed in two steps by evaluating model fit. The linear and non-linear model ﬁts for the GI were compared by calculating the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) for each model. Next, we created a linear model for age and the LGI, and additionally a non-linear model depending on the AIC and BIC for the GI models.

Specific domains of cognition map to different functional regions of the cerebral cortex (Taylor et al., 2015). Therefore, in addition to the g-factor we also studied whether the scores from the individual cognitive tests associated with GI and LGI. Furthermore, to inspect whether the association between cognition and gyriﬁcation changes with age we created a separate model with an interaction term for age and the g-factor.

Models were adjusted for covariates to account for potential con founding. The age analyses were corrected for sex and for study cohort, i.e. the ﬁrst, second or third cohort of the Rotterdam Study. The cognition analyses were adjusted in three separate models, which allows for the impact of each set of new confounders on model estimates to be described. Model 1 was adjusted for sex, cohort, age at cognitive testing and age difference between cognitive testing and the MRI scan. The way that age entered the model as a covariate — linear, quadratic or with splines — was dependent on the results from the analyses on age and the GI. Model 2 was additionally adjusted for hypertension, alcohol intake, smoking status and BMI. Lastly, Model 3 was additionally adjusted for education level. The p-values for the associations between the potential confounders and the global gyriﬁcation index are shown in Table 1. To assess whether image quality could affect the results we ran sensitivity analyses with the image quality metric for each scan as a covariate.

Gyriﬁcation is calculated as the ratio of the pial surface and the outer surface of the brain (Schaer et al., 2012). However, the cortical surface area itself has also been shown to relate to cognitive function (Cox et al., 2018). Any association between the LGI and cognition may therefore be driven by cortical surface area, and potentially by cortical thickness as well. To further assess this, we performed a sensitivity analysis per cluster. In each model we deﬁned cognition as the outcome, and both the mean LGI and the mean cortical surface area or the mean cortical thickness of each cluster as the determinants. The models were further corrected for all covariates as used in Model 3. We then assessed whether the association between cognition and LGI remained statistically signiﬁcant, taking into account cortical surface area or thickness.

All covariates had less than 1% missing data except for alcohol use (4.8%). In order to maximize power, missing covariate data were imputed thirty times using multiple imputation by chained equations (van Buuren and Groothuis-Oudshorn, 2011). Imputed models were subsequently pooled per vertex according to Rubin’s rules (Rubin, 1987). We also performed a non-response analysis to examine whether the individuals who were not included into the ﬁnal sample were in any way different than those who were included (Table 1). This was done through logistic regression, where inclusion was entered as the outcome and all other variables were included as predictors.

Of note, in all models, we deﬁned age or cognition as the determinants (predictors) and gyriﬁcation as the outcome, as limitations in vertex-wise analyses generally only allow for the vertex measure to be modeled as the outcome. Thus, while cognition is generally considered a consequence of brain structure, due to limitations in the vertex-wise software it was deﬁned as a determinant of gyriﬁcation in the models. As a sensitivity analysis, we created models for each statistically significantly associated predictor.

Table 1  Baseline characteristics of the study population. The excluded sample (n = 2250) were all participants that were eligible for cognitive testing and the neuromaging study, but did not end up in the ﬁnal sample (see Supplementary Fig. 1).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Included sample</th>
<th>Excluded sample</th>
<th>p-value[^1]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at MRI (years)</td>
<td>63.5 ± 10.1</td>
<td>69.5 ± 11.3</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>G-factor</td>
<td>0.00 ± 1.00</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>Cohort (%)</td>
<td>RS-I</td>
<td>15.3</td>
<td>33.9</td>
</tr>
<tr>
<td></td>
<td>RS-II</td>
<td>25.3</td>
<td>26.7</td>
</tr>
<tr>
<td></td>
<td>RS-III</td>
<td>59.4</td>
<td>39.5</td>
</tr>
<tr>
<td></td>
<td>Sex, female (%)</td>
<td>55.3</td>
<td>57.8</td>
</tr>
<tr>
<td>Time between cognition and MRI (years)</td>
<td>0.3 ± 0.4</td>
<td>0.2 ± 0.3</td>
<td>.847</td>
</tr>
<tr>
<td>Hypertensive (%)[^2]</td>
<td>61.4</td>
<td>76.1</td>
<td>.053</td>
</tr>
<tr>
<td>Alcohol per day (grams)[^2]</td>
<td>9.2 ± 10.1</td>
<td>8.5 ± 9.6</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Body mass index (kg/m^2)[^2]</td>
<td>27.4 ± 4.1</td>
<td>28.1 ± 4.8</td>
<td>.003</td>
</tr>
<tr>
<td>Smoking status (%)[^2]</td>
<td>Never</td>
<td>30.9</td>
<td>29.9</td>
</tr>
<tr>
<td></td>
<td>Past</td>
<td>49.1</td>
<td>49.4</td>
</tr>
<tr>
<td></td>
<td>Current</td>
<td>20.0</td>
<td>20.7</td>
</tr>
<tr>
<td>Education level (%)[^2]</td>
<td>Primary</td>
<td>7.8</td>
<td>13.1</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>37.7</td>
<td>41.0</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>30.3</td>
<td>28.4</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>24.2</td>
<td>17.5</td>
</tr>
<tr>
<td>Mean GI</td>
<td>2.55 ± 0.08</td>
<td>2.52 ± 0.09</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

[^1] Missingness of data for all variables was below 1% except for alcohol consumption (4.8%).
[^2] Differences between inclusion and exclusion were tested through multiple logistic regression.
significant LGI cluster where the cluster-wise mean LGI was defined as the determinant and the g-factor as the outcome.

All reported results focus on the beta coefficients and the 95% confidence intervals (CIs) rather than p-values. Confidence intervals give insight into the range of values within which the true parameter will likely be, whereas p-values do not (Greenland et al., 2016). Any reported result that is stated as statistically significant will have a p-value below the threshold of 0.05.

3. Results

Baseline characteristics of the study population (n = 4397) are displayed in Table 1. The mean age of the participants was 63.5 years (SD: 10.1, range: 45.7 to 97.9) and 55.3% were female. We analyzed whether any differences were present between individuals included in the analysis and those who were eligible for MRI but did not end up in the final sample (Supplementary Fig. 1). Excluded participants tended to be older (mean = 6.2 years), more often from the first cohort of the Rotterdam Study (33.9% versus 15.3%), were less likely to drink alcohol (mean = −0.7 g per day), had a higher BMI (mean = 0.7 kg/m²), were more likely to have only primary education (13.1% versus 7.8%) and had a lower GI (mean = −0.03).

3.1. Age and global gyriﬁcation

A scatterplot of age and the GI is shown in Fig. 1A, and the results of the different models are shown in Table 2. In the linear model one year increase in age associated with a −0.0021 (95% CI: -0.0025; -0.0017) lower GI. For the 2nd polynomial model both the linear term (p < 0.001) and the quadratic term (p = 0.027) were also statistically significant. All
spline coefficients were statistically significant for the natural cubic splines with both two and three degrees of freedom. The AIC and BIC of all models were highly similar, suggesting that the linear model and the two-fold spline model with the LGI. Fig. 2 displays the vertex-wise associations between age and the LGI. In the linear model the LGI decreased with age in the parietal, temporal, occipital and frontal regions. The effect sizes were generally larger than those found when examining the association between age and the GI. A second cluster arose in the frontal pole and medial prefrontal cortex, where the LGI increased with age. The significant clusters were similar across hemispheres in both size and strength. The two-degree spline model differed from the linear model. The first spline fold, i.e. ages between 45.7 and 61.6 years, associated negatively with the LGI in the parietal, frontal, temporal and occipital regions. Unlike the linear model, no cluster was present near the frontal pole or the medial prefrontal pole. In the second spline fold, i.e. ages between 61.6 and 97.9 years, the negative associations were more restricted to the temporal and parietal regions, and the lateral part of the frontal cortex. In addition, a positive cluster was present in the medial prefrontal cortex and the frontal pole, stronger than in the linear model. The age-gyrification association shows a clear deviation in its shape in the medial prefrontal gyrus compared to other regions (Supplementary Fig. 2). The findings were robust upon further correction for the image quality metric (Supplementary Fig. 3).

3.4. Cognition and local gyri

Fig. 3 displays the vertex-wise associations of the g-factor with the LGI for the three adjustment models. Associations between g-factor and the LGI were mostly present in three clusters: (1) the superior temporal gyrus, the insular cortex and the postcentral gyrus, (2) the lingual gyrus, the precuneus and the pericalcarine cortex and (3) the orbitofrontal

### Table 2
The associations between age and GI.

<table>
<thead>
<tr>
<th>Model</th>
<th>Type</th>
<th>β</th>
<th>95% CI</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>Linear</td>
<td>−0.0021</td>
<td>−0.0025; −0.0017</td>
<td>−10252.60</td>
<td>−10214.27</td>
</tr>
<tr>
<td>Quadratic</td>
<td>1st polynomial</td>
<td>−0.0053</td>
<td>−0.0081; −0.0024</td>
<td>−10255.47</td>
<td>−10210.75</td>
</tr>
<tr>
<td></td>
<td>2nd polynomial</td>
<td>−0.000025</td>
<td>−0.000003; 0.000046</td>
<td>−10256.33</td>
<td>−10211.61</td>
</tr>
<tr>
<td>Spline (2)</td>
<td>1st spline</td>
<td>−0.1235</td>
<td>−0.2376; −0.1020</td>
<td>−10203.46</td>
<td>−10254.57</td>
</tr>
<tr>
<td></td>
<td>2nd spline</td>
<td>−0.0703</td>
<td>−0.0947; −0.0458</td>
<td>−10203.46</td>
<td>−10254.57</td>
</tr>
<tr>
<td></td>
<td>3rd spline</td>
<td>−0.0826</td>
<td>−0.0550; −0.1102</td>
<td>−10203.46</td>
<td>−10254.57</td>
</tr>
</tbody>
</table>

GI = Gyrification index.
CI = Confidence interval.
AIC = Akaike information criterion.
BIC = Bayesian information criterion.

* The models were adjusted for study cohort and sex.

* Adjusted for age at MRI scan (years), study cohort, sex and age difference between cognitive testing and MRI scan (years).
* Additional adjustment for hypertension (yes/no), alcohol intake (grams per day), BMI and smoking status (never/past/current).
* Additional adjustment for education level (primary/low/intermediate/high).
* The Stroop task is inversely coded compared to the other tests, where a higher score relates worse performance.

### Table 3
The associations between cognition and the GI.

<table>
<thead>
<tr>
<th>Domain</th>
<th>GI (Model 1)</th>
<th>GI (Model 2)</th>
<th>GI (Model 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>β</td>
<td>95% CI</td>
<td>β</td>
</tr>
</tbody>
</table>

#### WLT = 15 Word learning test.
#### LDST = Letter digit substitution test.
#### WFT = Word fluency test.
#### PPB = Purdue pegboard test.
#### CI = Confidence interval.

* Adjusted for age at MRI scan (years), study cohort, sex and age difference between cognitive testing and MRI scan (years).
* Additional adjustment for hypertension (yes/no), alcohol intake (grams per day), BMI and smoking status (never/past/current).
* Additionally adjusted for education level (primary/low/intermediate/high).
* The Stroop task is inversely coded compared to the other tests, where a higher score relates worse performance.
These clusters roughly presented bilaterally. Similar patterns were found for the associations between the individual cognitive tests and the LGI (Fig. 4). The LGI of the cuneate gyrus, insular cortex and superior temporal gyrus were all associated with the Stroop task, the LDST and the WFT, although more so on the right than the left hemisphere. Additionally, the WFT also associated with the LGI in several additional regions, namely the supramarginal gyrus on both hemispheres and the lateral orbitofrontal cortex, the angular gyrus and the superior parietal
4. Discussion

We show in a large population-based setting that global gyrification of the cerebral cortex decreases during middle and late adulthood. This decline in gyrification is mainly driven by regions close to the Sylvian fissure. A specific cluster within the medial prefrontal cortex showed more gyrification with increasing age, particularly during late adulthood. Furthermore, we also found that global cognition positively associates with gyrification, in particular in the temporal regions, the lingual gyrus and the cuneus.

The findings for age and gyrification are in line with previously reported findings in smaller or younger age samples. A study from 2017 attempted to map the life course trajectory of the GI in a cross-sectional sample of 881 participants (Cao et al., 2017). The authors found that the GI trajectory can be described as a negative logarithmic function, with the decline in gyrification slowing with age. However, their sample included only about 30 participants above the age of 60. Another study reported on the association of age and gyrification in 322 healthy adults of whom 116 were aged 60 or older (Hogstrom et al., 2013). They found that the LGI had non-linear associations with age in certain brain regions, especially the orbitofrontal and dorsomedial prefrontal cortex. In particular, LGI in these regions seemed to increase towards the end of life. Our study builds upon these findings, with a much larger number of participants beyond the age of 70 years, allowing us to more precisely study how gyrification changes during late adulthood. We found that the association is essentially linear, which matches a negative logarithmic life course pattern (Cao et al., 2017). We also established non-linear patterns in the regional surface-based patterns, albeit the increase in LGI was only seen in the medial prefrontal cortex and only towards the end of life. These findings further consolidate the global and local dependence of gyrification on age.

Gyrification associates with cognition during and after midlife, and this association does not change with age. Furthermore, in half of the significant clusters we found that cortical surface area is likely driving the associations. This is not surprising as loss of surface area leads to lower folding thus lower gyrification within that region. Still, three out of six LGI clusters remained associated with cognition after adjusting for surface area, suggesting that gyrification harbors independent information. Furthermore, the pattern of LGI clusters within our study was similar amongst the individual cognitive tests, suggesting that the LGI captures a more general aspect of cognitive function. Gyrification may therefore play a unique role in cognitive function, which could prove useful in the study of normal and abnormal cognitive aging. For example, other cortical characteristics such as thickness and surface area have distinct contributions to cognitive decline as seen in Alzheimer’s disease (Yang et al., 2019; Ossenkoppele et al., 2019; Dickerson et al., 2009), yet any such contributions from gyrification remain to be elucidated.

The temporal lobe has previously been linked to cognitive processes...
such as language (Price, 2012) and memory (Jeneson and Squire, 2012) as well as psychiatric disorders like adulthood autism spectrum disorders (Kohli et al., 2019) and schizophrenia (Palaniyappan and Liddle, 2012; Nesvag et al., 2014; Madre et al., 2019). Interestingly, these disorders have also been linked to abnormal gyri cation. Previous studies have found links between genes underlying cognition function and temporal lobe structure (Tan et al., 2019; Ersland et al., 2012), although the results are inconsistent (van der Essen, 1997), which states that axonal tension pulls the gyral walls inwards, thus folding the cortex. The axonal tension may depend on the image quality metric as a covariate.

Table 4
The associations of the mean LGI and mean surface area with g-factor per previously identified LGI cluster. Each model contained both LGI and surface area. The associations of the mean LGI and surface area were standardized in order to be able to compare the magnitude of their effects. All models were adjusted for age at MRI scan (years), study cohort, sex, age difference between cognitive testing and MRI scan (years), hypertension (yes/no), alcohol intake (grams per day), BMI (kg/m²), smoking status (never/past/current), and education level (primary/low/intermediate/high).

<table>
<thead>
<tr>
<th>Hemisphere</th>
<th>#</th>
<th>Location</th>
<th>Standardized mean LGI</th>
<th>Standardized mean surface area</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>β</td>
<td>95% CI</td>
</tr>
<tr>
<td>Left</td>
<td>1</td>
<td>Temporal</td>
<td>0.025</td>
<td>[0.006; 0.055]</td>
</tr>
<tr>
<td>Left</td>
<td>2</td>
<td>Cuneus</td>
<td>0.047</td>
<td>[0.016; 0.077]</td>
</tr>
<tr>
<td>Left</td>
<td>3</td>
<td>Orbitofrontal</td>
<td>0.025</td>
<td>[-0.001; 0.050]</td>
</tr>
<tr>
<td>Right</td>
<td>1</td>
<td>Temporal</td>
<td>0.045</td>
<td>[0.014; 0.076]</td>
</tr>
<tr>
<td>Right</td>
<td>2</td>
<td>Cuneus</td>
<td>0.028</td>
<td>[-0.005; 0.073]</td>
</tr>
<tr>
<td>Right</td>
<td>3</td>
<td>Orbitofrontal</td>
<td>0.030</td>
<td>[0.005; 0.054]</td>
</tr>
</tbody>
</table>

Table 5
The associations of the mean LGI and mean cortical thickness with g-factor per previously identified LGI cluster. Each model contained both LGI and cortical thickness. Both the LGI and cortical thickness were standardized in order to be able to compare the magnitude of their effects. All models were adjusted for age at MRI scan (years), study cohort, sex, age difference between cognitive testing and MRI scan (years), hypertension (yes/no), alcohol intake (grams per day), BMI (kg/m²), smoking status (never/past/current), and education level (primary/low/intermediate/high).

<table>
<thead>
<tr>
<th>Hemisphere</th>
<th>#</th>
<th>Location</th>
<th>Standardized mean LGI</th>
<th>Standardized mean thickness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>β</td>
<td>95% CI</td>
</tr>
<tr>
<td>Left</td>
<td>1</td>
<td>Temporal</td>
<td>0.063</td>
<td>[0.038; 0.088]</td>
</tr>
<tr>
<td>Left</td>
<td>2</td>
<td>Cuneus</td>
<td>0.069</td>
<td>[0.045; 0.084]</td>
</tr>
<tr>
<td>Left</td>
<td>3</td>
<td>Orbitofrontal</td>
<td>0.044</td>
<td>[0.020; 0.068]</td>
</tr>
<tr>
<td>Right</td>
<td>1</td>
<td>Temporal</td>
<td>0.075</td>
<td>[0.050; 0.100]</td>
</tr>
<tr>
<td>Right</td>
<td>2</td>
<td>Cuneus</td>
<td>0.058</td>
<td>[0.034; 0.083]</td>
</tr>
<tr>
<td>Right</td>
<td>3</td>
<td>Orbitofrontal</td>
<td>0.049</td>
<td>[0.026; 0.073]</td>
</tr>
</tbody>
</table>

Y gyrification may also associate with age due to more technical aspects of the data collection itself. Head motion may affect the relation between age and GI (Madan and Kersinger, 2017). The reasoning for this is that older participants tend to move more with their head while in the MRI. A previous study confirmed this and also found that head motion related to LGI, although the association was not very strong (Madan, 2018). We attempted to minimize the impact of head motion on the analyses by conservatively excluding all raw images with suboptimal quality and further by performing sensitivity analyses with the image quality metric as a covariate.

The study has several limitations. First of all, we relied on a cross-sectional study design to examine age effects on gyrification. Cross-sectional estimation of age-related changes may yield inaccurate estimates compared to longitudinal designs (Pfefferbaum and Sullivan, 2015). Second, changes in gyrification likely cause changes in cognition, but the models were specified with cognition as the determinant and gyri cation as the outcome due to limitations in the vertex-wise analysis modeling. Rerunning the models per cluster with proper specification did show that the LGI clusters indeed associated with cognition, suggesting that the models hold under proper specification. Third, the cognitive test battery that was used does not cover all aspects of cognitive function. Due to the emphasis on verbal tests we were not able to fully assess the scope of cognition and gyri cation. Fourth, in the case of cognition there may be reverse causality, as higher intelligence tends to lead to a healthier lifestyle and thus better brain health. We corrected for a number of variables related to lifestyle and their effect on the association was minimal. Despite this there could still be residual confounding by other variables that we did not account for. Fifth, while we excluded those with...
prevalent stroke and dementia, there could be other medical conditions and confounders that could bias the results. For example, traumatic brain injury and substance abuse disorders are known to accelerate brain and cognitive aging (Wood, 2017; Cole et al., 2015; Mende, 2019), and could subsequently affect the association of age and cognition with gyrification. Our study also had several strengths. First, this is the largest sample size to date in a study of gyrification and age or cognition, leading to sufficient statistical power to find associations, and unraveled new regional differences. Second, the individuals were sampled from a wide age-range, thus enabling making accurate inferences about gyrification even in the later phases of late adulthood. Third, the sample was drawn from a population-based cohort, thus the findings can be generalized beyond a clinical setting.

In conclusion, gyrification globally decreases linearly with age across the entirety of adulthood, and gyrification in the medial prefrontal cortex increases towards the end of life. Furthermore, gyrification increases with higher levels of cognitive performance in some clusters irrespective of surface area. These findings consolidate the importance of gyrification in normal brain function. Whether gyrification is a viable marker for abnormal brain aging and cognitive decline towards the end of life remains to be elucidated.

Disclosures

This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (project: ORACLE, grant agreement No: 678543; and project: EuroPOND, grant agreement No: 666992), and from the Sophia Foundation (project: 518-20). Supercomputing resources were supported by the NWO Physical Sciences Division (Exacte Wetenschappen) and SURFsara (Cartesius compute cluster, www.surfsara.nl). The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam.

Declaration of competing interest

The authors report no potential conflicts of interest.

CRediT authorship contribution statement

Sander Lamballais: Conceptualization, Software, Formal analysis, Writing - original draft, Visualization. Elisabeth J. Vinke: Conceptualization, Formal analysis, Writing - review & editing. Meike W. Verhoonii: Resources, Writing - review & editing. M. Arfan Ikram: Conceptualization, Resources, Funding acquisition, Writing - review & editing. Ryan L. Muetzel: Conceptualization, Writing - original draft, Supervision.

Acknowledgements

The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/neuroimage.2020.116637.

References


