Smoking is associated with lower brain volume and cognitive differences: A large population analysis based on the UK Biobank

Progress in Neuropsychopharmacology & Biological Psychiatry 123 (2023) 110698

Zeqiang Linli a, b, c, Edmund T. Rolls d, e, Wei Zhao a, b, Jujiao Kang f, Jianfeng Feng e, f,*, Shuixia Guo a, b, *

MOE-LCSM, School of Mathematics and Statistics, Hunan Normal University, Changsha, PR China
Key Laboratory of Applied Statistics and Data Science, Hunan Normal University, College of Human Province, Changsha, PR China
School of Mathematics and Statistics, Guangdong University of Foreign Studies, Guangzhou, PR China
Oxford Centre for Computational Neuroscience, Oxford, UK
Department of Computer Science, University of Warwick, Coventry, UK
Centre for Computational Systems Biology, Fudan University, Shanghai, PR China

A B S T R A C T

The evidence about the association of smoking with both brain structure and cognitive functions remains inconsistent. Using structural magnetic resonance imaging from the UK Biobank (n = 33,293), we examined the associations between smoking status, dosage, and abstinence with total and 166 regional brain gray matter volumes (GMV). The relationships between the smoking parameters with cognitive function, and whether this relationship was mediated by brain structure, were then investigated. Smoking was associated with lower total and regional GMV, with the extent depending on the frequency of smoking and on whether smoking had ceased: active regular smokers had the lowest GMV (Cohen’s d = −0.362), and former light smokers had a slightly smaller GMV (Cohen’s d = −0.060). The smaller GMV in smokers was most evident in the thalamus. Higher lifetime exposure (i.e., pack-years) was associated with lower total GMV (β = −31.18, p = 8.35 × 10−56). In those who ceased smoking, the duration of abstinence was associated with a larger total GMV (β = 139.57, p = 2.36 × 10−56). It was further found that reduced cognitive function was associated with smoker parameters and that the associations were partially mediated by brain structure. This is the largest scale investigation we know of smoking and brain structure, and these results are likely to be robust. The findings are of associations between brain structure and smoking, and in the future, it will be important to assess whether brain structure influences smoking status, or whether smoking influences brain structure, or both.

1. Introduction

Over the decades, smoking has become one of the biggest threats to world health. There were about 1.3 billion smokers worldwide in 2020, accounting for about 16% of the world population, and about 8 million people die from smoking every year (WHO, 2021). Besides numerous negative health outcomes including circulatory and respiratory diseases (Jha et al., 2013), smoking might lead to multiple neurobiological and neurocognitive abnormalities, which may be through effects on brain structure (Debette et al., 2011; Durazzo et al., 2014; Myklebust et al., 2008). Hence, it is increasingly important to understand in a large study with many participants to produce robust findings, more precisely what the associations are between smoking, brain structure, and cognitive functions.

Structural magnetic resonance imaging (sMRI) is widely used to assess brain structural differences in vivo. Numerous studies with relatively small samples have reported widespread structural differences in smokers compared to nonsmokers, and high smoking may cumulatively be associated with more serious cognitive decline and brain alterations, including the prefrontal cortex (Brody et al., 2004; Chaarani et al., 2019; Ding et al., 2015; Fritz et al., 2014; Liao et al., 2012; Morales et al., 2012; Zhang et al., 2011; Zhong et al., 2016), anterior cingulate gyrus (Fritz et al., 2014; Liu et al., 2015; Morales et al., 2012; Pan et al., 2013), thalamus (Ding et al., 2015; Liao et al., 2012; Morales et al., 2012;
It is also noteworthy that most studies have focused on the differences between current smokers and non-smokers, or between those who have ever smoked and non-smokers. Indeed, smoking behaviors influenced by a variety of individual factors are so complex that it is simplistic to divide the population into smokers and non-smokers. For example, besides regular active smokers, and non-smokers who have never smoked, some people may smoke lightly in their lifetime, or some people may relapse to smoking again after quitting, etc. Smokers who lightly smoke comprise more than a quarter of the smoking population (Morrell and Cohen, 2000), representing an important target group for the cessation of smoking, taking into account potential differences between them and regular smokers in terms of smoking motivation and quitting-related cognition (Robertson et al., 2016). A study showed that different smoking habits and higher rates of cigarette smoking increased the risk of cardiovascular disease in women with atherosclerosis light smokers and almost risk free was found in those who had stopped for 15 years (Amiriet al., 2019). Nevertheless, few studies have compared the brain structure of smoker groups with different smoking habits to non-smokers. Analyzing smokers in a broader way based on measures of their smoking could provide a deeper understanding of the relationship between smoking and brain structure, which of course needs a large study population (Amiriet al., 2019; Oelsner et al., 2020; Schane et al., 2010).

Previous investigations have shown that smoking can be associated with lower cognitive function (Anstey et al., 2007; Mons et al., 2013), but we know of no previous investigation of whether the association between smoking and brain structure is related to altered cognitive function in smokers. In addition, though former smokers now outnumber current smokers in many countries (Oelsner et al., 2020), few studies have investigated brain structure and cognitive function in those who have stopped smoking, and this has important implications for public health and informing prevention strategies for smoking.

Based on prior literature, the current study makes new contributions in several ways: (1) the use of a more complete and large-scale sample (n=33,293) makes the results reliable and robust; (2) comprehensive consideration of the association of different smoking habits and brain structure provides a better understanding on their relationships; (3) exploration of the association of smoking and cognitive function and whether the relationships are mediated through brain structures in the same population is important for understanding of the brain mechanisms and cognitive function.

Therefore, the present study aimed to investigate the relationship between three smoking parameters (smoking status, the amount of smoking, and duration of smoking abstinence), and brain structure and cognitive function in middle-aged and older adults from the UK Biobank, one of the largest neuroimaging databases in the world. Specifically, the objectives of this study were (1) to investigate the different GMV between controls and 6 groups of smokers categorized by their smoking characteristics; (2) to explore the relationship between brain structure and the smoking amount and smoking abstinence; (3) to examine the relationship between smoking parameters and cognitive function, and whether this relationship is mediated by brain structure.

2. Materials and methods

2.1. Participants

The UK Biobank (http://www.ukbiobank.ac.uk) is a large prospective population-based cohort study that recruited approximately 500,000 community volunteers between 2006 and 2010 across the UK. Participants were recruited to collect a range of questionnaires about detailed phenotypic information including diet, lifestyle, anthropometric and cognitive function assessments, and biological samples, including blood and medical records obtained from the NHS registries. Since 2014, a subset of participants have been invited back to collect brain MRI scans, and questionnaires about diet, lifestyle, and cognitive function assessments, with 38,562 participants (aged from 44 to 81 at the time of their scans) available in the current study. Structural MRIs were collected across three imaging centers (62% of the samples were acquired in the Headache site, 25% of samples were acquired in the Newcastle site, and 13% of samples were acquired in the Reading site) that were equipped with identical scanners.

For the purposes of the study, as shown in Fig. S1, participants were excluded if they had (1) reported neuropsychological disorders at the time of assessment such as bipolar disorder, depression, and mania; (2) missing or unclear smoking data (for example, unclear or missing current and previous smoking habits); (3) missing key demographic covariates (age); (4) poor quality of MRI (i.e., image quality rating [IQR] was lower than 75%). Detailed information on the exclusion procedures is presented in the Supplementary Material. Following exclusions, there were 33,293 subjects with sMRI data included in the following analyses.

2.2. Neuroimaging data collection and preprocessing

The UK Biobank used a standard Siemens Skyra 32-channel 3T scanner (Siemens Medical Solutions, Germany) for all magnetic resonance brain imaging, with 1111 resolution and a field of view of 208 256 256 (http://biobank.ctsu.ox.ac.uk/crystal/field.cgi?id=2367).

All UK Biobank structural MRI data were preprocessed in the CAT12 toolbox with default settings, including: (1) the T1-weighted images were segmented into GM, white matter (WM), and non-brain voxels (cerebrospinal fluid, skull) using the “new-segment” routine; (2) population templates (GM, WM) were generated from each of the datasets separately using the DARTEL algorithm; (3) the gray-matter images were aligned to a nonlinear transformation field and normalized to the MNI space; (4) the normalized images were then smoothed with an 8 mm full width at half maximum Gaussian kernel with the resulting voxel size 1.5mm³. Spatially normalized, smoothed, and Jacobian-scaled gray matter images were obtained for each subject. The estimated total intracranial volume (ITV) was calculated as the summation of the gray matter, white matter, and cerebrospinal fluid volumes in the native space. This study focused on the total gray matter volume (GMV) and regional GMV for 166 regions of interest (ROIs) defined by the automated anatomical labeling atlas and AAL3 (anatomical regions are listed in Table S1) (Rolls et al., 2020).

2.3. Research variables

2.3.1. Smoking variables

In this study, the 33,293 participants were divided into 6 smoker groups and a control group according to the questionnaires about their smoking characteristics. Specifically, as shown in Table S2, 1254 subjects were classified as “Current smoker” who currently smoke on most or all days according to the data field 1239 “Current tobacco smoking” (https://biobank.ctsu.ox.ac.uk/crystal/field.cgi?id=1239; 240 subjects were classified as Relapsed smoker” who previously smoked on most or all days but with lightly smoking now according to both data
fields 1239 and 1249 “Past tobacco smoking” (https://biobank.ctsu.ox.ac.uk/crystal/field.cgi?id=1249); 6749 subjects were classified as “Ex-smoker” who previously smoked on most or all days and had quit smoking currently according to the data fields 1239 and 1249. Ac- cording to data fields 1239, 1249, and 2644 “Light smokers, at least 100+ smokes in lifetime” (https://biobank.ctsu.ox.ac.uk/crystal/field.cgi?id=2644), 418 subjects were classified as “Current-Light smoker” who have smoked lightly from the past until now, and smoked at least 100 times in total; 3751 subjects were classified as “Ex-Light smoker” who used to smoke lightly with smoking at least 100 times in total but quit smoking currently and 6214 subjects were classified as “V-Light smoker” those who had very light smoked and failed to meet the standards of light smokers (i.e., a total of at least 100 times in their lifetime). Finally, 14,667 participants who never smoked were classified as “Control” according to both data fields 1239 and 1249.

The other two key smoking parameters calculated were: ii) pack-years was calculated as cigarettes per day divided by 20 and then times the number of years smoked and was only available for current, relapsed, and ex-smokers; ii) quitting duration, which was calculated as age at the time of data collection minus the age when the participant stopped smoking on most days and this was only available for relapsed, and ex-smokers.

2.3.2. Cognitive measures

The UK Biobank contains a series of cognitive measures which were specifically designed or modified for use by the UK Biobank cognitive neuroscience expert working group. Despite the non-standard nature of these tests and the limited psychometric information, the UK Biobank cognitive data have been used in numerous scientific publications (Hagenaars et al., 2016; Kendall et al., 2017; Miller et al., 2016). The present study included 7 cognitive measures with continuous test scores, namely reaction time, fluid intelligence, numeric memory, pairs matching, symbol digit substitution, trail making, and paired-associate learning.

Reaction time assessment is based on 12 rounds of the card-game “Snap”. The participant is shown two cards at a time; if both cards are the same, they press a button. Box that is on the table in front of them as quickly as possible. The internal consistency reliability of these trials, measured by Cronbach’s α, was 0.85. The score of this cognitive variable for use is the mean duration to the first press of the snap-button summed over rounds in which both cards matched. It gives a measure of the raw processing and reaction speed of a participant that the larger the value, the slower the reaction speed.

Fluid intelligence assessment involved participants answering 13 multiple-choice questions which were designed to access verbal and numerical reasoning (Cronbach α reliability 0.62). Participants who didn’t answer all of the questions within the allotted 2-min limit are scored as zero for each of the unattempted questions. The score of this cognitive variable for use is a simple unweighted sum of the number of correct answers given to the 13 fluid intelligence questions. It reflects the verbal and numerical reasoning ability of a participant that the larger the value, the stronger the reasoning ability.

Numerical memory assesses numeric short-term memory. The partici- pant was shown a 2-digit number to remember. The number then dis- appeared and after a short while they were asked to enter the number on the screen. The number became one digit longer each time they remembered correctly (up to a maximum of 12 digits). The score of this cognitive function used in the present study was the longest number correctly recalled during the numeric memory test. It provides a measure of working memory such that the larger the value, the better the memory.

The pairs matching test was used to assess visual memory. Paired-associate cards were presented face down on the screen and the participant is asked to recall as many as possible in the fewest tries. Multiple rounds were conducted. The first round used 3 pairs of cards and the second 6 pairs of cards. The score of this cognitive variable for use was the number of incorrect matches in the round. It gives a measure of visual memory such that the larger the value, the worse the memory.

Symbol digit substitution was used to measure processing speed. The participant was presented with one grid linking symbols to single-digit integers and a second grid containing only the symbols. They were then asked to indicate the numbers attached to each of the symbols in the second grid using the first one as a key. The values of this cognitive test were the numbers of symbols correctly matched to digits by the participant that the larger the value, the better the cognition.

The trail-making test is a neuropsychological test of visual attention and task switching. Participants were asked to connect scattered circles containing a sequence of numbers (Trail A) and then to connect circles containing numbers or letters by alternating between them in ascending sequence (Trail B). We used the time taken to complete these tests for our analyses. It reflects the ability of visual search speed, scanning, speed of processing, and mental flexibility, as well as executive function- ing, and has been widely used in many studies that the larger the value, the longer the completing time.

Paired-associate learning is an as classic memory paradigm that is used to understand how people encode and retrieve newly formed associations between stimuli, which has most commonly been used to examine and understand the mechanisms of learning and forgetting of information. In the paired-associate learning task, the participants were shown 12 pairs of numbers (for 30 in total) and then, after an interval (in which they did a different test), presented with the first word of 10 of these pairs and asked to select the matching second word from a choice of 4 alternatives. We used the number of word pairs correctly associated for our study, which provides a measure of verbal declarative memory such that the larger the value, the better the memory.

It is worth emphasizing that the cognitive scores for use in the study are recommended by the UK Biobank and/or followed previous studies, and most of them have been proven to have substantial concurrent validity and test-retest reliability (Fawns & Ritchie, 2020).

2.3.3. Mental health

Given that anxiety, low well-being, and irritability or mania are common mental health symptoms for smokers (Moylan et al., 2013; Stickley et al., 2015) and based on the data category 136 “Mental Health”, the relevant measures (Anxiety, Happiness and subjective well-being, and Mania) were included in this study. Specifically, these vari- ables were calculated according to the data categories 139, 140, and 147 in the UK Biobank website, and were converted into binary variables based on their value distribution (e.g., high anxiety vs. low anxiety).

2.3.4. Other covariates

Variables, as follows known to be correlated with GMV and/or cognitive function, were included as covariates in all analyses: age (Luo et al., 2020), sex (Gennatas et al., 2017), handedness (Jang et al., 2017), ethnicity (Tang et al., 2010), body-mass index (BMI) (Hammer and Batty, 2019), alcohol drinking frequency (Piumatti et al., 2018; Zahr and Pfefferbaum, 2017), imagesites (Alfaro-Almagro et al., 2021), and TIV (Barnes et al., 2010).

Detailed information on these variables, including smoking vari- ables, cognitive measures, mental health symptoms, and demographic variables can be found in the Supplementary Material.

2.4. Statistical analysis

2.4.1. Association of smoking parameters and brain GMV

Two sample two-tailed t-tests were used to test whether smoking status was associated with GMV after removing the confounding effects of age, sex, handedness, ethnicity, BMI, alcohol status, imaging site, and TIV. These comparisons were divided into two categories: Smokers vs. Control and Between-smoker comparisons. The former comparison
category is to test whether smoking status was significantly associated with brain volume using the reposed GMV to compare each of the six smoking subgroups to controls, while the latter one is to examine the differences in GMV between the different groups of smokers. Effect sizes were calculated with Cohen’s d (Cohen, 2013).

Following a previous study (Karase et al., 2015), we used linear regressions to examine the relationship between the cumulative amount of smoking (i.e., pack-years) and brain GMV in ever-smokers (current and ex-smokers) with reference to controls (pack-years 0); and to explore the relationship between the quitting duration and brain GMV in Ex-smokers with reference to current smokers (duration 0), with adjustment for potential confounding effects including pack-years.

The above analyses were conducted to test the associations of smoking parameters with total and regional (166 cortical and subcortical regions) GMV. Separate comparisons or models were run for each brain region. The false-discovery rate (FDR) method described by Benjamini and Hochberg (Benjamini and Hochberg, 1995) was used to adjust for multiple comparisons when statistical tests were performed on each of the 166 brain regions. The Bonferroni correction procedure was used for comparisons oftional GMV.

2.4.2. Association of smoking parameters and cognitive function

We then modeled the associations between smoking parameters and cognitive functions by using linear regression, with the cognitive mean as the dependent variable and the smoking measures as independent variables, with adjustment for potential confounders of age, sex, handedness, ethnicity, BMI, alcohol drinking frequency, imaging site, and TIV. Separate Bonferroni corrections were conducted for each cognitive function.

2.4.3. Mediation analysis

To test the hypothesis of whether the relationships between the smoking parameters (independent: X) and cognition (dependent: Y) were mediated through brain structures (mediator: M), a mediation analysis with a standard 3-variable path model was performed using the R package mediation (http://CRAN-R-project.org/package/medication) (Baron and Kenny, 1986). Estimates were calculated for the total relationship of smoking on cognition (X –> Y), the relationship of smoking on brain GMV (X –> M), and the relationship of brain GMV on cognition adjusting for smoking (X –> Y). The significance of the mediation was estimated by the bias-corrected bootstrap approach (with 1000 random samplings). In this analysis, we focused on those cognitive tests that were significantly associated with smoking parameters (i.e., reaction time, symbol digit substitution test, and paired-associate learning). Confounding variables as in the association analysis were regressed out in the mediation model. The Bonferroni correction procedure was performed for mediation analysis of total GMV, while separate BH-FDR corrections were conducted for brain region statistical analyses (P_{FDR} < 0.05).

Detailed descriptions of the statistical analyses can be found in the Supplementary Material. Statistical analyses were performed using R, version 4.0.4 (https://www.r-project.org). Mapping results were visualized with Circos (version 0.69, http://circos.ca/), ggplot2 (version 3.3.5), and BrainNet Viewer (Xia et al., 2013).

3. Results

3.1. Participant characteristics

Of 33,293 participants, the mean age at enrollment was 63.73 (SD 7.53) years. 15,651 (47.0%) were male and 1038 (3.1%) were non-white people. The cohort included 1254 current smokers (3.77%), 240 relapsed smokers (0.72%), 6749 ex-smokers (20.27%), 418 curr-light smokers (1.26%), 3751 ex-light smokers (11.27%), 6214 v-light smokers (18.66%), and 14,667 controls (44.05%). Smokers showed the worst mental health condition; compared to controls, significantly larger percentage of smokers had relatively more anxiety (49.0% vs. 49.5%–63.6%), lower well-being (56.9% vs. 57.5%–73.3%), and higher manic status (23.8% vs. 25.8%–41.9%), with the magnitude related to the smoking frequency and smoking abstinence. Detailed participant characteristics are provided in Table 1.

3.2. Smoking is associated with lower GMV

3.2.1. Smoking status: smoker versus control

Except for the “V-Light smoker” group, smokers had significantly smaller total GMV than controls (Table 2, Fig. S2). Compared to the controls, the “Current smoker” group demonstrated the lowest total GMV (Bonferroni corrected p-value, i.e., P_{Bonferroni} = 1.18 × 10^{-30}, Cohen’s d = -0.362), followed by the “Relapsed smoker” (P_{Bonferroni} = 4.62 × 10^{-34}, Cohen’s d = -0.224), the “Ex-smoker” (P_{Bonferroni} = 6.62 × 10^{-36}, Cohen’s d = -0.151), the “Curr-Light smoker” (P_{Bonferroni} = 2.67 × 10^{-32}, Cohen’s d = -0.145), and the “Ex-Light smoker” groups (P_{Bonferroni} = 6.92 × 10^{-43}, Cohen’s d = -0.060).

Regionally, multiple brain regions showed significantly smaller volumes in smokers than controls (Fig. 1). Specifically, in the “Current smoker” group, there were extensive areas with smaller volume (155/166, FDR corrected p-value, i.e., P_{FDR} < 0.001, Cohen’s d range: [-0.426, 0.405]) among which the largest effect sizes were in the thalamus, fusiform gyrus (FFG), middle cingulate cortex (MCC), para-hippocampal gyrus (PHG), amygdala, lingual gyrus (LING), and prefrontal and temporal cortices (Table S3). In the “Relapsed smoker” group, 61 areas had smaller volumes mainly including multiple thalamic nuclei, and frontal and temporal cortices (P_{FDR} < 0.05, Cohen’s d range: [-0.248, 0.442], Table S4). In the “Ex-smoker” group, there were also widespread smaller GMVs but they were slightly less severe (140/166, P_{FDR} < 0.001, Cohen’s d range: [0.050, 0.140]) with the top effect sizes mainly in the putamen, amygdala, olfactory cortex, insula, MCC, and prefrontal and temporal cortices (Table S5). The “Curr-Light smoker” group showed smaller volume in 54 brain regions including the sub- stantia nigra (SN), thalamus, amygdala, hippocampus, PHG, and ventral tegmental area (VTA) (P_{FDR} < 0.05, Cohen’s d range: [-0.220, -0.109], Table S6). In the “Ex-Light smoker” group, 16 areas had slightly lower volume, mainly distributed in the anterior cingulate cortex (ACC), and frontal lobe (P_{FDR} < 0.001, Cohen’s d range: [-0.104, -0.072], Table S7). No significant regional differences were found in the “V-Light smoker” group compared to the control group after FDR correction.

3.2.2. Smoking status: between-smoker groups

We next examined the differences in GMV between the different groups of smokers (Table S8) with effects of the same possible confounding variables regressed out as above. The association of smoking abstinence with brain GMV was investigated by conducting two comparison (i.e., Ex-smoker vs. Current smoker and Ex-Light smoker vs. Curr-Light smoker). The effect of pack-years was additionally regressed out for the comparison of “Ex-smoker vs. Current smoker”. The results showed that the Ex-smoker group had a significantly larger total GMV than the Current smoker group (P_{Bonferroni} 1.30 \times 10^{-04}, Cohen’s d = 0.157). The association of smoking frequency with brain GMV was then investigated by conducting another two comparisons (i.e., Current smoker vs. Curr-Light smoker, Ex-smoker vs. Ex-Light smokers). The results were unsurprising in that Current smokers had smaller total GMV than Curr-Light smokers (P_{Bonferroni} 7.42 \times 10^{-14}, Cohen’s d = -0.211), and Ex-smokers had smaller total GMV than Ex-Light smokers (P_{Bonferroni} 2.97 \times 10^{-02}, Cohen’s d = 0.090). The results at the ROI level are shown in Fig. S3. Specifically, compared with the “Current smoker” group, the “Ex-smoker” group showed 55 regions with larger volume including multiple thalamic nuclei (P_{FDR} < 0.001, Cohen’s d range: [0.109, 0.252], Table S9), and 1 region (i.e., right locus coeruleus [LC]) with smaller volume (Cohen’s d = -0.167). Compared with the “Curr-Light smoker” group, the “Ex-Light smoker” group also showed larger volume in 4 regions (P_{FDR} <
Current - The meaning of NA is not clear.

Z. Linli et al.
Progress in Neuropsychopharmacology & Biological Psychiatry 123 (2023) 110698

Table 1
Demographic variables of different groups.

<table>
<thead>
<tr>
<th>Demographics</th>
<th>Overall (n = 33,293)</th>
<th>Current smoker (n = 12,554)</th>
<th>Relapsed smoker (n = 240)</th>
<th>Ex-smoker (n = 6749)</th>
<th>Curr-Light smoker (n = 418)</th>
<th>Ex-Light smoker (n = 3751)</th>
<th>V-Light smoker (n = 6214)</th>
<th>Control (n = 14,667)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (mean [SD])</td>
<td>63.73 (7.53)</td>
<td>61.84 (7.29)</td>
<td>62.39 (7.37)</td>
<td>65.64 (7.21)</td>
<td>61.72 (7.64)</td>
<td>64.36 (7.48)</td>
<td>63.25 (7.58)</td>
<td>63.12 (7.52)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Male (%)</td>
<td>15.65 (47.0)</td>
<td>64.4 (51.4)</td>
<td>130 (54.2)</td>
<td>3557 (52.7)</td>
<td>259 (62.0)</td>
<td>1918 (51.1)</td>
<td>2725 (43.9)</td>
<td>6418 (43.8)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Handedness: non-right (%)</td>
<td>3671 (11.0)</td>
<td>146 (11.6)</td>
<td>28 (11.7)</td>
<td>783 (11.6)</td>
<td>39 (9.3)</td>
<td>440 (11.7)</td>
<td>655 (10.5)</td>
<td>1580 (10.8)</td>
<td>0.208</td>
</tr>
<tr>
<td>Ethnic: non-white (%)</td>
<td>1038 (31)</td>
<td>46 (3.7)</td>
<td>11 (4.6)</td>
<td>135 (2.0)</td>
<td>16 (3.8)</td>
<td>107 (2.9)</td>
<td>166 (2.7)</td>
<td>557 (3.8)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Note: Group comparison p-values were calculated based on the available variable categories, that is, chi-square tests were used for categorical variables (with continuity correction) and analyses of variance were used for continuous variables.

The meaning of NA is not applicable. There are different available sample sizes for each mental health score or cognitive measure in analysis.

Table 2
Comparisons of total GMV between smokers and controls.

<table>
<thead>
<tr>
<th>Group</th>
<th>t-value</th>
<th>p-value</th>
<th>P(1-tail)</th>
<th>CohenD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current vs. Control</td>
<td>11.942</td>
<td>1.97</td>
<td>×</td>
<td>−0.362</td>
</tr>
<tr>
<td>Relapsed smoker vs. Control</td>
<td>3.549</td>
<td>4.62</td>
<td>×</td>
<td>−0.224</td>
</tr>
<tr>
<td>Ex-smoker vs. Control</td>
<td>10.803</td>
<td>6.62</td>
<td>×</td>
<td>−0.151</td>
</tr>
<tr>
<td>Curr-Light smoker vs. Control</td>
<td>2.849</td>
<td>4.45</td>
<td>×</td>
<td>−0.145</td>
</tr>
<tr>
<td>Ex-Light smoker vs. Control</td>
<td>3.282</td>
<td>1.15</td>
<td>×</td>
<td>−0.060</td>
</tr>
<tr>
<td>V-Light smokers vs. Control</td>
<td>−0.451</td>
<td>6.52</td>
<td>×</td>
<td>−0.007</td>
</tr>
</tbody>
</table>

Note: Comparisons of each smoker group against never-smoking control in total GMV regressed out age, sex, handedness, ethnicity, BMI, alcohol drinking frequency, sites, and TIV. The P(1-tail) were obtained by Bonferroni correcting.

smoker” group, the “Ex-smoker” group showed smaller volume in 5 regions (P<0.001, Cohen’s d range: [0.108, 0.085], Table S12), among which the top effect sizes were in the middle temporal gyrus, putamen, and cerebellum. Note that since both the pack-years and duration of quitting smoking are unavailable for the 2 Light smoker groups, the comparison involving the 2 Light smoker groups in this section should be considered exploratory.

3.2.3. Pack-years and quitting duration

As illustrated in Table S13, those who had ever smoked (current, relapsed, and ex-smokers) showed a negative association between pack-years and total GMV (β = 230.54, p < 0.001, Fig. 2A). There was a significant 0.05, Cohen’s d range: [0.170, 0.188], Table S10), mainly in the SN and cerebellum. Compared with the “Curr-Light smoker” group, the “Current smoker” group showed smaller volume in 71 regions (P<0.05, Cohen’s d range: [0.253, -0.127], Table S11), mainly in multiple thalamic nuclei, and larger volume in right LC (d=0.157), and bilateral VTA (d=0.128).
and 0.129 respectively). Compared with the "Ex-Light positive correlation of the duration of quitting smoking and total GMV ($\beta = 16.15, p = 5.62 \times 10^{-25}$); and this correlation was weakened (lower $\beta$ value) when additionally controlling for pack-years ($\beta = 139.57, p = 2.36 \times 10^{-18}$, Fig. 2B). Fig. 2E shows that the longer the smoking abstinence time, the smaller the difference of GMV from controls, and that this association was stronger for those who smoked for more pack-years. There were 154 ROIs negatively correlated and 3 ROIs positively correlated with pack-years ($P < 0.001$, Fig. 2C, Table S14). Many of these ROIs with stronger negative associations (bigger standardized coefficients) were in thalamic areas; while the ROIs with positive associations were in frontal areas.
3.3. Associations between smoking and cognition

Table S16 summarizes the 7 cognitive measures including the available sample sizes in the analysis. We only consider the samples without missing values to establish the model in the analysis.

Of the 7 cognitive functions, reaction time, symbol digit substitution scores, and paired association learning scores were significantly related to the smoking parameters (Table S17). Specifically, compared with controls, the “Current smoker” group had significantly longer reaction times ($\beta = 8.287, P_{\text{Bonferroni}} = 0.033$), and reaction time positively correlated with pack-years ($\beta = 0.291, P_{\text{Bonferroni}} = 0.026$), and negatively correlated with quitting duration ($\beta = -0.287, P_{\text{Bonferroni}} = 0.024$). As regards symbol digit substitution scores, the “Current smoker” and “Ex-smoker” groups showed significantly smaller scores than controls ($\beta = 1.089 - P_{\text{Bonferroni}} = 0.026$, respectively), and significant associations with the pack-years and quitting duration were found ($\beta = -0.035, P_{\text{Bonferroni}} = 2.08 \times 10^{-6}, \beta = 0.020, P_{\text{Bonferroni}} = 1.41 \times 10^{-5}$, respectively). Compared with controls, the “Current smoker” group had significantly lower paired-associate learning scores ($\beta = -0.380, P_{\text{Bonferroni}} = 2.14 \times 10^{-4}$) while the “Ex-Light smoker” group showed larger scores ($\beta = 0.306, P_{\text{Bonferroni}} = 9.08 \times 10^{-1}$), and paired-associate learning scores negatively correlated
significantly better cognitive functioning than the current smoker group in the fields of reaction time ($\beta = -9.896, P_{\text{Bonferroni}} < 0.001$), and paired-association learning ($\beta = 0.384, P_{\text{Bonferroni}} < 10^{-5}$). The current smoker group had lower paired-association learning scores than the Curr-Light smoker, while the Ex-smoker group had higher trail-making scores than the Ex-Light smoker. Detailed results are presented in Table S18.

### 3.4. Results of the mediation analysis

We performed mediation analysis for both the total and regional GMV which was significantly associated with both smoking parameters and cognitive function (Fig. S4). The results indicated that the total GMV significantly mediated the relationship between the smoking variables and symbol digit substitution scores (Fig. 3); that is total GMV partially mediated the association of smoking status (Proportion of mediation = 6.5%, $P_{\text{Bonferroni}} < 0.001$, Fig. 3A); pack-years (Proportion of mediation = 7.4%, $P_{\text{Bonferroni}} < 0.001$, Fig. 3B); and duration of quitting smoking (9.0%, $P_{\text{Bonferroni}} < 0.001$, Fig. 3C) with symbol digit substitution scores. Similarly, total GMV significantly mediated the smoking status (Proportion of mediation = 12.0%, $P_{\text{Bonferroni}} < 0.001$), pack-years (Proportion of mediation = 10.4%, $P_{\text{Bonferroni}} < 0.001$), and duration of quitting smoking (Proportion of mediation = 15.2%, $P_{\text{Bonferroni}} < 0.001$) and paired-associate learning scores. The results of the mediation analysis of total GMV performed for reaction time were non-significant. Performing mediation analysis for those ROIs significantly related to both smoking variables and measures of cognition, we found many ROIs with significant mediation effects ($P_{\text{Bonferroni}} < 0.05$) as shown in Fig. 3D-F (Table S19–21) for symbol and digit substitution, in Fig. S5A (Table S22–24) for reaction time, and in Fig. S5B (Table S25–27) for paired-associate learning, in which the thalamic nuclei were the most prominent mediators.

### 4. Discussion

The present study quantifies in detail the associations of smoking with brain GMV and cognition in a large neuroimaging dataset. We found that smokers with different smoking statuses showed different
Fig. 3. Mediation by total and regional GMV of the association between three smoking parameters ([A, D] Smoking status, [B, E] pack-years, and [C, F] quitting duration) and symbol digit substitution scores. (A)-(C) Mediation analysis on smoking parameters, total GMV, and symbol digit substitution score. Path A: the association between the smoking parameters and the mediator (total GMV); Path B: the association between the mediator and the outcome (symbol digit substitution scores) controlling for the smoking parameter; Path C shows the association (total effect) of smoking parameters and the symbol digit substitution scores when the total GMV was not taken into account. Path C indicates the direct association between the smoking parameters and the outcome (symbol digit substitution scores) controlling for the mediator, which shows a significant reduction in the regression coefficient when the association with the total GMV was taken into account (direct effect). Path AB shows that taking total GMV into account explains about 6-8% of the association between smoking exposure parameters and symbol digit substitution scores (mediation effect). (D)-(E) shows significant results of the mediation analysis on smoking parameters, regional GMV, and symbol digit substitution score. The color bar represents the percentage of the mediation effect that could be explained by the mediator (regional GMV). The percentage of the mediation effect was measured by the formula: 100*(total effect - direct effect)/total effect). The significance of the mediation was estimated by the bias-corrected bootstrap approach (with 1000 random samplings; P<0.05).

extents of smaller brain volume depending on smoking frequency and smoking abstinence. Higher lifetime exposure (i.e., pack-years) was associated with smaller GMV, while after quitting smoking, smokers with a significantly larger GMV had a longer quitting duration. Furthermore, smoking was associated with impaired cognitive functions measured by reaction time, the symbol digit substitution test, and paired associate learning, and these associations were partially mediated by brain structure. Findings from this study robustly develop an under-standing of the association of smoking with brain volume and cognition. A recent study based on the UK Biobank has found that smoking is associated with lower total GMV consistent with the current findings (Gray et al., 2020), but that study mainly focused on current smokers or ever-smokers (current plus former smokers) as with most previous studies (Elbehjani et al., 2019). The current study expands that recent study, considering a wider range of smoking statuses and reporting a clearer relationship between smoking and brain structure. While different smoking statuses all showed many discrepant brain regions, with many overlapping areas, it is noteworthy that the magnitude of the group differences (measured by Cohen’s d, as shown in Fig. 1) is different depending on the smoking status. Light smoking is associated with slightly smaller global and regional GMV, mainly distributed in cortical and subcortical structures including the thalamus, MCC, hip-pocampus, PHG, amygdala, SN, VTA, and red nucleus, all of which have lower volumes in regular smokers (except for the VTA). These regions include parts of the mesocorticolimbic system involved in a variety of functions including reward, and reinforcement learning (Berntson and Kringelbach, 2015; Grall-Bonnet and Sauvaget, 2014; Yager et al., 2015). The SN and VTA contain dopaminergic neurons, influenced by and providing important signals to other regions of the reward system (e.g., orbitofrontal cortex and amygdala) (Rolls, 2017, 2018). Smoking may be related to the differences in reward-related areas (Cheng et al., 2019). For those with higher smoking intensity and frequency, lower GMV was also found in the cerebral cortex and including the frontal and temporal lobes. These smaller regions (Table S3) have been reported in the previous literature but have not been found simultaneously (Brody et al., 2004; Hanlon et al., 2016). In this study, the combination of the statistical power of the large brain-imaging sample of the UK Biobank and the use of a highly robust metric (i.e., GMV) made it possible to
discover widespread associations between smoking and brain GMV (Madan and Kensinger, 2017). Lower brain volume in chronic and reg- ular smokers has also been associated with a higher risk for the neuro- cognitive disorder (Karas et al., 2003; Knight et al., 2016; Lee et al., 2013). Very interestingly, the current investigation revealed a high correlation between the lower GMV and the amount of smoking ("Cur- rent smoker" group, r = 0.94, p < 0.001, Fig. S7).

A smoking amount-GMV negative association between pack-years and GMV has been reported in previous studies consistent with the current study (Cox et al., 2019; Durazzo et al., 2017; Fritz et al., 2014; Peng et al., 2018), using data from the UK Biobank, Cox et al. (2019) looked at several cardiovascular risk factors and found that a greater number of cigarette pack-years was associated with smaller total GMV and reduced volume of the thalamus, basal ganglia, hippocampus, and several cortical regions. Also based on the UK Biobank, Gray et al. (2020) considered the association of brain structure with smoking duration and cigarettes per day instead of cigarette pack-years, only reporting that longer smoking duration is associated with smaller total GMV, as their investigation included numerous covariates. Cigarettes per day does not capture the potential association of smoking accumu- lation and brain structure while the duration of smoking does not take into account the level of actual smoke exposure. Pack-years used in this study may be a trade-off that synthesizes information from these two smoking characteristics (i.e., by using the product of cigarettes per day and smoking duration). All these smoking characteristics are helpful to draw a comprehensive conclusion for smoking-related studies. One strength of the present study is that we also showed those with only a little smaller GMVs had longer durations for quitting smoking in the same dataset. Similar evidence of a positive association between the duration of quitting and cortical thickness has previously been reported (Karama et al., 2015). Interestingly, regions with stronger smoking amount-GMV associations usually had a larger abstinence-GMV association.

Regionally, the thalamus was a prominent brain region in which low volume was associated with smoking, and which mediated the associa- tion between smoking and cognition. In smokers who still smoke now including the “Current smoker”, “Relapsed smoker”, and “Curt-Light smoker” groups, and the “Current smoker” and the “Curt-Light smoker” groups, the thalamus is the most significant brain region with the top effect size (i.e., Cohen’s d) and the magnitude is related to the extent of smoking. The thalamus is a brain region with the highest density of nicotinic acetylcholine receptors (nAChRs) (Mukherjee et al., 2018), and is involved in many cognitive functions including arousal, sustained attention, and behavioral inhibition (Huang et al., 2018). Reduced cholinergic function can impair cognition by reducing the firing rates in cortical attractor networks (Rolls and Deco, 2015). The present investigation describes an association between GMV and smoking, and does not reveal the direction of any effects (Parvaz et al., 2022). One hypothesis is that with low thalamic volume in some in- dividuals, there may be less excitation because of presumably fewer nAChRs, and these individuals may compensate for that by self- administering nicotine. Part of the reward value of nicotine is pro- duced by actions on nAChRs that increase dopamine release in for example the ventral striatum (Wills et al., 2022). However, the highest concentration of nicotine receptors is in the thalamus, and although chronic nicotine exposure leads to nAChR upregulation (Dubrott et al., 2015), the reduced thalamic volume in smokers reported here may mean that smokers might compensate for a lower thalamic volume and consequently fewer nAChRs by self-delivering nicotine. This may have beneficial effects on attention that could be influenced by the thalamic nAChRs (Sottille et al., 2017). An alternative hypothesis is that those who self-administer nicotine may reduce the gray matter volume of some brain regions, including the thalamus. In that situation, the constant bombardment of these nAChRs by long-term regular nicotine exposure may make it a prime target for potential morphometric anomalies. Differences in thalamic volume and functional connectivity have been related to whether smokers relapse after quitting (Wang et al., 2020),

but it is noted that the pharmacology of nicotine receptors and smoking is complex, with at least the rewarding aspects of nicotine thought to be related to increased dopamine release from dopaminergic neurons (Wills et al., 2022). In any case, we note that lower thalamic gray matter volume is not specific to smoking, and is found with other drugs of abuse including alcohol, cocaine, methamphetamine, opioids, cannabis, and synthetic cannabinoids (Huang et al., 2018).

It is noteworthy that given the number of participants in each group, we found highly statistically significant mean differences between the different smoking status groups in regions like the thalamus. Therefore, it is important to report the effect size (measured by Cohen’s d) in addition to the p-value as it quantifies the magnitude of a group-dif- ference, while a low p-value by itself only confirms its existence (Sulli- van and Feinn, 2012). According to Cohen’s (2013), d values of 0.2 represent small effects, values between 0.4 and 0.6 moderate effects, and d values of 0.8 or higher large effects. In this study, the magnitude of the group differences between smokers and controls was a marginal to small effect size, even for the difference for active regular smokers (i.e., Cur- rent smoker - Control), and is generally smaller than that in patients with neuropsychiatric diseases in which the effect sizes of group differences are small to moderate (Thompson et al., 2020), such as schizophrenia (van Ep et al., 2016), depression (Koolschijn et al., 2009), and bipolar disorder (Hibar et al., 2016). Therefore, the association of smoking and brain volume should not be overstated, but small effects in medicine cannot be important in terms of human health. Therefore, the findings of this paper may provide meaningful implications to under- stand neural mechanisms of smoking.

Of the 7 cognitive functions we examined in the current study, re- action time, symbol digit substitution scores, and paired association learning scores were significantly related to smoking. These tests of cognitive function reflect the speed with which tasks can be performed and also learning, and the ability to focus attention may be involved (Jaeger, 2018). As regards the remaining cognitive measures including fluid intelligence, numeric memory, pairs matching, and trail making, they are more a reflection of memory, reasoning ability, and executive functioning (Salthouse, 2011), and were not strongly associated with smoking. A possible implication is that the aspects of cognition most strongly associated with smoking are functions that probably reflect attention,alertness, and fast learning, with nicotine likely because of its cholinergic functions (Rolls et al., 2022; Rolls and Deco, 2015) to improve these aspects of cognition and performance.

There are several limitations to this work. First, since the results of this study are based on cross-sectional data, it is not possible to infer the causal relations of the associations between smoking and brain morphological variation identified here. Second, although “Light smokers” were recruited here, detailed smoking frequency or other smoking information is not available for this smoker group. Third, although the comprehensive analysis was based on the largest sample with neuroimaging data, there may be some potential sample bias in the study samples we used or even in the UK Biobank. Fry et al. (2017) have demonstrated that UK Biobank’s 500,000 participants are gener- ally healthier, leaner, and smoke less than their fellow countrymen and women, suffering less heart and kidney disease and cancer. Therefore, the UK Biobank is not representative of the whole UK population. In our study sample, since some subjects were excluded due to exclusion criteria, we inevitably lost useful information on smoking for this group of subjects. However, a valid assessment of smoking-brain structure relationships may be widely generalizable to enormous populations of people given the very large scale of the sample investigated. Fourth, some important covariates were not included in the current study that might be correlated with both brain structure and cognitive measures, such as education, physical exercise, and other drug addiction. Due to the absence of detailed years of education for participants in the UK Biobank, and the inadequate coverage of the relevant data field of other drug addiction and physical exercise on the study population, we were unable to further explore the possible confounding effect of these
covariables which may have an impact on the association between smoking with brain structure and cognitive measures. In addition, smoking and depression are often co-morbid, and both are associated with lower brain gray matter volume and impaired cognitive performance, but this was not considered in this study, although we have excluded those subjects diagnosed with depression.

5. Conclusions

In summary, based on the largest smoking-related dataset with sMRI data, we report associations between smoking, lower brain gray matter volume (GMV), and cognition. The magnitude of the association de-pends on the amount of smoking, and smaller differences in brain vol-ume are associated with the duration with which individuals could quit smoking. We also showed that smoking is associated with reduced cognitive ability (e.g., symbol-digit substitution scores and reaction time), and that this effect was partly mediated by the lower brain GMV. This study leads to new concepts including a possible role of thalamic nicotinic receptors in smoking, and provides a better understanding between the brain, smoking, and cognition that may be useful in the prevention of smoking and its treatment.

Authors contribution

Zeqiang Linli: Conceptualization, Formal analysis, Funding acquisition, Methodology, Validation, Visualization, Writing - original draft. Edmund T. Rolls: Conceptualization, Formal analysis, Validation, Writing - original draft, Writing - review & editing. Wei Zhao and Jujuang Kang: Data curation, Validation, Writing - original draft. Shuxia Guo and Jianfeng Feng: Conceptualization, Funding acquisition, Supervision, Writing - original draft, Writing - review & editing. All authors critically revised the manuscript and approved the final version.

Data and materials availability

All UK Biobank data used in this work were obtained under Data Access Application 19,542 and are available to eligible researchers through the UK Biobank (www.biobank.ac.uk).

Ethical statement

The UK Biobank received ethical approval from the research ethics committee (REC reference 11/NW/0382). Written informed consent was obtained from each subject.

Declaration of Competing Interest

The authors declare that they have no conflict of interest.

Data availability

The authors do not have permission to share data.

Acknowledgments

Shuxia Guo is supported by the National Natural Science Foundation of China (NSFC) (No. 12071124), Shanghai Municipal Science and Technology Major Project (No.2018SHZDZX01), Key Laboratory of Computational Neuroscience and Brain-Inspired Intelligence (LCNBI), and ZJ Lab. Jianfeng Feng is supported by the National Key R&D Pro-gram of China (nos. 2018YFC1312900 and 2019YFA0709502), the Shanghai Municipal Science and Technology Major Project (no. 2018SHZDZX01), the ZJ Lab, Shanghai Center for Brain Science and Brain-Inspired Technology and the 111 Project (no. B18015). Zeqiang Linli is supported by the Guangdong University of Foreign Studies (299-X5222211).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.pnpbp.2022.110968.

References


WHO. 2021. Tobacco. WHO.


