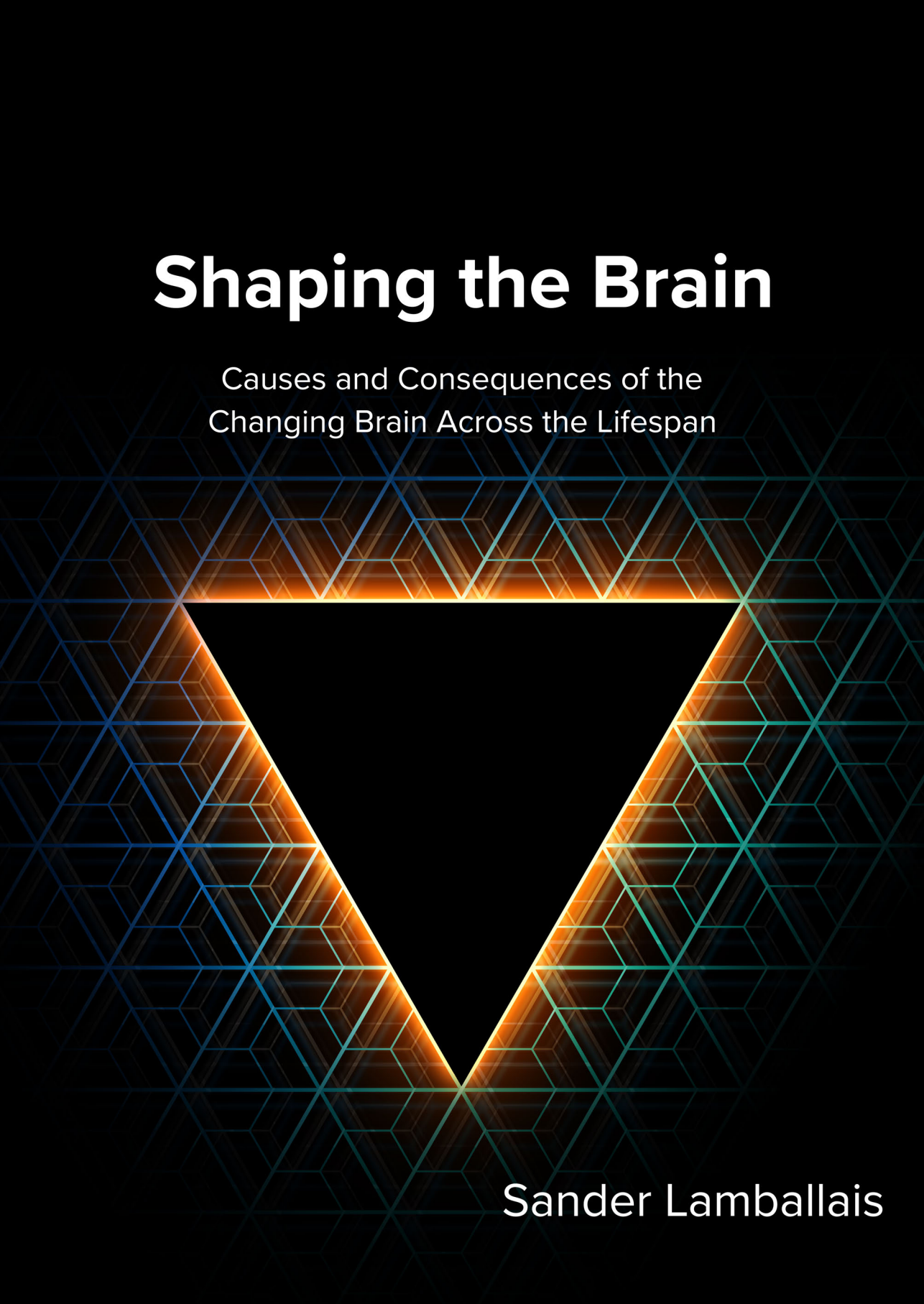
Shaping the Brain

Causes and Consequences of the Changing Brain Across the Lifespan



Sander Lamballais

SHAPING THE BRAIN

Causes and consequences of the changing brain across the lifespan

Sander Lamballais

Shaping the brain: causes and consequences of the changing brain across the lifespan

Author: S. Lamballais

ISBN: 978-94-6419-410-4

Cover design: S. Lamballais

Layout: S. Lamballais

Printing: Gildeprint BV, Enschede

Shaping the Brain: Causes and Consequences of the Changing Brain Across the Lifespan

Het vormen van de hersenen: de oorzaken en consequenties van het veranderende brein gedurende het hele leven

Proefschrift

ter verkrijging van de graad van doctor aan de

Erasmus Universiteit Rotterdam

op gezag van de

rector magnificus

Prof. dr. A.L. Bredenoord

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

dinsdag 25 januari 2022 om 15.30 uur

door

Sander Lamballais Tessensohn

geboren te Krimpen aan den IJssel.

Erasmus University Rotterdam

Promotiecommissie:

Promotoren:	Prof. dr. M.A. Ikram
	Prof. dr. H.W. Tiemeier
Overige leden:	Prof. dr. N.E.M. van Haren
	Prof. dr. M. Ewers
	Prof. dr. H.E. Hulshoff Pol
Copromotoren:	Dr. T.J.H. White
	Dr. H. Adams

Paranimfen:	Drs. R.F. Hussainali
	Mr. C.L. de Mol

This work is dedicated to Anna, the love of my life,

and to you, the reader who never stopped inspiring me

TABLE OF CONTENT

MANUSCRIPTS THAT FORM THE BASIS OF THIS THESIS

Chapter 2.1: **Lamballais, S.***, Adank, M. C.*, Hussainali, R. F., Schalekamp-Timmermans, S., Vernooij, M. W., Luik, A. I., Steegers, E. A. P. & Ikram, M. A. (2021) Design and overview of the ORigins of Alzheimer's disease aCross the Life coursE (ORACLE) Study. *Eur J Epi* **36**, 117-127, doi: 10.1007/s10654-020-00696-3.

Chapter 2.2: **Lamballais, S.** & Muetzel, R. L. (2021) QDECR: a flexible, extensible vertex-wise analysis framework in R. *Front Neuroinform,* doi: 10.3389/fninf.2021.561689.

Chapter 3.1: **Lamballais, S**., Jansen, P. R., Labreqcue, J. A., Ikram, M. A. & White, T. (2021) Genetic scores for adult subcortical volumes associate with subcortical volumes during infancy and childhood. *Hum Brain Mapp* **42**, 1583-1593, doi: 10.1002/hbm.25292.

Chapter 3.2: **Lamballais, S.**, Muetzel, R. L., Ikram, M. A., Tiemeier, H., Vernooij, M. W., White, T. & Adams, H. H. H. (2020) Genetic burden for late-life neurodegenerative disease and its association with early-life lipids, brain, behavior, and cognition. *Front Psychiatry* **11**, 33, doi: 10.3389/fpsyt.2020.00033.

Chapter 4.1: **Lamballais, S.***, Sajjad, A.*, Leening, M. J. G., Gaillard, R., Franco, O. H., Mattace-Raso, F. U. S., Jaddoe, V. W. V., Roza, S. J., Tiemeier, H. & Ikram M. A. (2018) Association of blood pressure and arterial stiffness with cognition in 2 population-based child and adult cohorts. *J Am Heart Assoc* **7**, e009847, doi: 10.1161/JAHA.118.009847.

Chapter 4.2: Hussainali, R. F.*, **Lamballais, S.***, Schalekamp-Timmermans, S., Roeters van Lennep, J. E., Steegers, E. A. P. & Ikram, M. A. (*manuscript in preparation*) The association of vascular risk factors during early adulthood with subsequent cognition and brain structure.

Chapter 5.1: **Lamballais, S.**, Delaney, S., Muetzel, R. L., White, T., Ikram, M. A. & Tiemeier, H. (*manuscript in preparation*) Childhood adversity and accelerated aging of the brain.

Chapter 5.2: **Lamballais, S.**, Vinke, E. J., Vernooij, M. W., Ikram, M. A. & Muetzel, R. L. (2020) Cortical gyrification in relation to age and cognition in older adults. *Neuroimage* **212**, 116637, doi: 10.1016/j.neuroimage.2020.116637.

Chapter 5.3: **Lamballais, S.**, Zijlmans, J. L., Vernooij, M. W., Ikram, M. K., Luik, A. I. & Ikram, M. A. (2020) The risk of dementia in relation to cognitive and brain reserve. *J Alzheimers Dis* **77**, 607-618, doi: 10.3233/JAD-200264.

GENERAL INTRODUCTION

1

> *"To form an understanding like ours, nature needed more time than for fabricating animal minds. We had to pass through infancy to reach the age of reason. We had to experience the trials and tribulations of anility to derive from them the advantages that characterize man."*
>
> – Julien Offray De La Mettrie (L'Homme Plante, 1748)

In 1748, Julien Offray de La Mettrie posited that the superior mind of human beings over other animals arises from their altricial nature[1]. Where precocial animals are born with the capacities to be mobile and explore their environment, altricial animals are born relatively unmatured, lacking mobility and the capacity to explore (**Figure 1**). However unmatured the human brain may be at birth, it continues to develop throughout childhood and far into adulthood, to ultimately become the most intelligent amongst all animals.

The human brain develops along a general trajectory across the lifespan[2]. The gray matter of the brain – the tissue rich in cellular bodies – rapidly grows in utero and starts to decrease in volume from middle childhood onwards. The white matter of the brain, which primarily contains axons that connect brain regions, continues to increase in volume well into

Figure 1 | Example of precocial (left) versus altricial (right) species of birds (top) and mammals (bottom).

adulthood. As individuals age into their 40's and 50's, their brains start to accumulate neuropathological changes, such as lesions in the white matter and focal, small intracerebral hemorrhages[3]. As the brain ages further, it starts to show more signs of decline, including a reduction in volume and the accumulation of neuropathological burden. In turn, these changes can throughout adulthood lead to cognitive and functional decline and may result in the onset of late-life dementia, with Alzheimer's disease accounting for most dementia cases.

Individuals show large variations in how their brains change over a lifetime. Certain individuals suffer from vast losses in brain volume or decline of function before the age of 50 years, whereas others surpass the age of 100 years while accumulating minimal neuropathological burden[4,5]. Understanding what causes these differences in trajectories could have significant consequences for society. First, it would become possible to predict whose brain will age without significant problems, and whose brain will be burdened by pathology and functional decline. Second, several of these causes can be modified[6], such as vascular risk factors[7]. Combined, this would enable identification and treatment of individuals who are at a high risk of suffering from structural and functional decline of the brain later in life. However, to realize this, we will need a much better grasp of how the brain develops and declines over the lifespan.

Our understanding of the brain has developed rapidly over the last decades. This has primarily been driven by a number of advances. First, technological advances have led to new generations of in vivo brain imaging, i.e., mapping of the brain in living organisms. For example, magnetic resonance imaging enables the mapping of structural and functional aspects of the brain with submillimeter resolution[8,9]. Second, these imaging techniques have become much cheaper to perform as well as more readily available. This has led to the third advance, the advent of large-scale neuroimaging studies and publicly available neuroimaging datasets. Certain studies published in the last few years have included over 30,000 participants with neuroimaging data[10-12], a number that was unimaginable just 20 years ago. Finally, the number of studies with multiple measurements over time has rapidly increased in recent years[13-15]. This has led to direct observations of how specific aspects of the brain change over time, as well as exploration of the factors that predict those changes.

The etiology of lifelong brain change – both development and deterioration – remains elusive. We aimed to uncover new aspects of brain changes by utilizing data from population-based studies from early, mid and late-life.

SCOPE OF THIS THESIS

The general aim of this thesis is to identify and understand different mechanisms that shape the brain across the lifespan. The thesis will primarily focus on genetic mechanisms (**chapter 3**), vascular mechanisms (**chapter 4**) and concepts related to reserve (**chapter 5**).

Genetic mechanisms

Brain structure is highly heritable and multiple genes have been identified that shape specific regions of the brain[10-12]. Similarly, late-life neurodegenerative diseases like Alzheimer's disease are also moderately heritable. In the case of Alzheimer's disease, several specific genes have been identified[16], such as APOE, that significantly affect the risk to develop the disease.

These studies primarily focused on adult samples. However, the brain undergoes its most rapid changes during early life. The genes related to brain structure and late-life neurodegenerative disease may influence brain structure and function during early childhood or even during prenatal development. In **chapter 3.1**, we used the results from a genome-wide association study on subcortical volumes in adults to calculate genetic scores for subcortical volumes in infants and young children. In **chapter 3.2**, we calculated genetic scores for late-life neurodegenerative diseases and examined how these related to early-life measures of brain structure, cognition performance, behavioral measures, and serum lipid levels.

Vascular mechanisms

Vascular pathology directly affects brain health. One of the strongest risk factors for developing dementia is the occurrence of stroke[17], a disruption of the blood supply to part of the brain leading to rapid cell death. Furthermore, high levels of blood pressure during midlife associate with lower levels of cognitive function later in life, independent of the late-life blood pressure levels[18,19]. Such studies have primarily focused on mid and late-life populations. However, vascular health is strongly influenced by genetics, and certain vascular measures like blood pressure have stable trajectories across life. Thus, vascular risk factors may exert their effect earlier in life. In **chapter 4.1**, we studied whether blood pressure and arterial stiffness associated with cognition during early childhood, and contrasted this to the associations found during mid and late-life. In **chapter 4.2**, we examined how vascular risk factors across adulthood relate to cognition and brain structure nearly 15 years later.

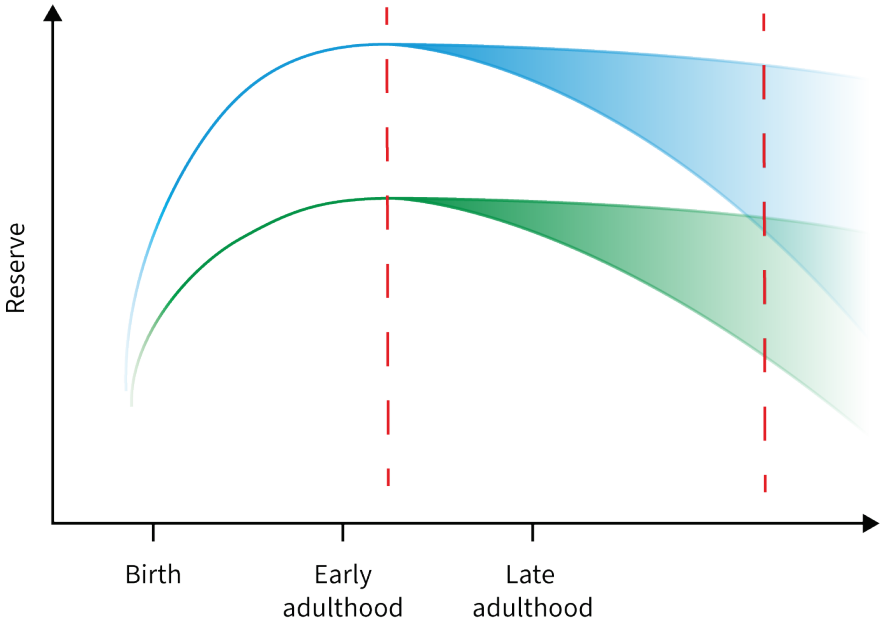
Reserve

Individuals differ in their susceptibility to develop dementia during late-life. These differences likely have their origin decades before the onset of actual symptoms. For example, early life educational attainment seems to modify the risk to develop dementia. A mechanism that

explains the effect of education on the onset of dementia is reserve[20], which is the capacity to buffer the effect of age-related decline and neuropathological burden. A distinction can be made between cognitive and brain reserve. Cognitive reserve describes the use of cognitive mechanisms to buffer functional decline. Brain reserve is usually defined as the sheer neural capital that a person has, e.g., the volume of the brain, the number of neurons and other structural features. Both brain and cognitive reserve presumably change across the lifespan (**Figure 2**). For example, cognitive reserve is hypothesized to build up with higher levels of educational attainment, occupational complexity, and late-life engagement in cognitive and social activities[21].

We aimed to understand different aspects of brain change in relation to reserve and related concepts. In **chapter 5.1**, we examined how adverse life experiences during early childhood affected subsequent brain aging during adolescence. **In chapter 5.2**, we studied how the gyrification of the cerebral cortex changes across mid and late-life, and how it related to levels of cognitive functions. Finally, in **chapter 5.3** we address the question of whether early-life or late-life reserve levels are more relevant for understanding who will develop dementia, and how cognitive and brain reserve interact in affecting the incidence of dementia.

Figure 2 | Potential trajectories for low (green) and high (blue) levels reserve.



SETTING

The research questions within this thesis lie along the lifespan. Questions related to early-life (**chapters 3.1, 3.2, 4.1, 5.1**) were primarily studied in the Generation R Study[22], a prospective population-based birth cohort study based in Rotterdam, the Netherlands. The primary aim of the study was to gain insight into the prenatal and early-life determinants of normal and abnormal health during childhood and adolescence. It thus started with the inclusion of pregnant women and consisted of extensive measures of both parents and their child – before and after birth. The Generation R Study has generated a wealth of data that have led to over a thousand research manuscripts that address different aspects of normal and abnormal health during early life. In this thesis, we focused on measures from prenatal development, early childhood, and middle childhood.

Questions related to late-life (**chapters 4.1, 5.2, 5.3**) were explored in the Rotterdam Study[23], another prospective population-based cohort study based in Rotterdam, the Netherlands. Where Generation R focuses on the beginnings of life and its consequences for health, the Rotterdam Study focuses on determinants of health during the end of life. The participants were aged 40 years and older at baseline and participated in a myriad of measurements derived from nearly all modern fields of medicine. The participants are followed using medical records and repeat visits to the research center to understand how their health relates to subsequent health outcomes.

Generation R and the Rotterdam Study cover the beginning and the end of the lifespan, respectively. However, the age gap between the studies – roughly from 13 to 40 years – represents a significant portion of the lifespan and of brain development and deterioration. In **chapter 2.1**, we present the design and overview of the ORACLE Study, a study nested in the Generation R Study that focuses on the parents of the Generation R children. At inclusion in the Generation R Study, the age range of the parents spanned from the late teens to the late 40's, essentially filling the gap between the Generation R children and the Rotterdam Study. We introduced a separate research visit for these parents nearly 15 years after inclusion, to study the longitudinal determinants of brain health and function during adulthood (**chapter 4.2**).

Epidemiology has defined three primary sources of bias, or systematic error, in study design: confounding bias, selection bias and information bias[24]. For each of these biases, steps can be taken during design of the study or during analysis that minimize these sources of bias; this has been done in the Generation R Study and in the Rotterdam Study. For example, confounding bias arises when factors that influence both the determinant and outcome of a

model are not accounted for. Both Generation R and the Rotterdam Study collected extensive data on factors that may be confounders for specific research questions. By incorporating these into the analysis models, the effect of confounding can be reduced. However, as the field of neuroimaging did not arise from epidemiology, the most used analysis tools also do not have features to account for different sources of bias. To bridge these fields, we developed a neuroimaging analysis tool that incorporates such epidemiological principles into the analysis of neuroimaging data (**chapter 2.2**).

Finally, in **chapter 6** I will reflect the work done in this thesis, in particularly the implications of the findings and the methodological considerations of the conducted studies. Finally, I will discuss the future perspectives and directions that are relevant for understanding brain health from a population-based perspective.

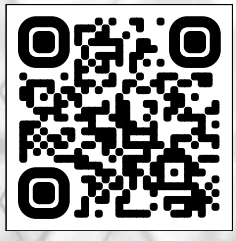
METHODOLOGICAL ADVANCES

2.1

Design and overview of the ORigins of Alzheimer's disease aCross the Life coursE (ORACLE) Study

Sander Lamballais*, Maria C. Adank*, Rowina F. Hussainali, Sarah Schalekamp-Timmermans, Meike W. Vernooij, Annemarie I. Luik, Eric A. P. Steegers, M. Arfan Ikram

Adapted from:



ABSTRACT

Brain development and deterioration across the lifespan are integral to the etiology of late-life neurodegenerative disease. Factors that influence the health of the adult brain remain to be elucidated and include risk factors, protective factors, and factors related to cognitive and brain reserve. To address this knowledge gap, we designed a life-course study on brain health, which received funding through the EU ERC Programme under the name Origins of Alzheimer's Disease Across the Life course (ORACLE) Study. The ORACLE Study is embedded within Generation R, a prospective population-based cohort study of children and their parents, and links this with the Rotterdam Study, a population-based study in middle-aged and elderly persons. The studies are based in Rotterdam, the Netherlands. Generation R focuses on child health from fetal life until adolescence with repeated in-person examinations but has also included data collection on the children's parents. The ORACLE Study aims to extend the parental data collection in nearly 2000 parents with extensive measures on brain health, including neuroimaging, cognitive testing, and motor testing. Additionally, questionnaires on migraine, depressive symptoms, sleep, and neurological family history were completed. These data allow for the investigation of longitudinal influences on adult brain health as well as intergenerational designs involving children and parents. As a secondary focus, the sampling is enriched by mothers (n = 356) that suffered from hypertensive disorders during pregnancy to study brain health in this high-risk population. This article provides an overview of the rationale and the design of the ORACLE Study.

INTRODUCTION

The number of Alzheimer's disease (AD) cases is projected to double or even triple by 2050[25,26], which emphasizes the urgency to disentangle the etiology of AD and to develop effective preventative strategies. Although AD has an onset late in life, the risk to develop AD is influenced by both early-life and adulthood factors[27-29], including cognitive and academic performance[30-33], cardiovascular health[34-36], lifestyle factors[6,37-43] and life events[44-47]. These factors likely affect the susceptibility to develop AD through mechanisms such as cognitive and brain reserve[20,48]. These mechanisms have been hypothesized to reduce or buffer the effect of brain pathology and aging.

The influence that the risk factors have on the incidence of AD likely depends on the life phase[49]. For example, hypertension during midlife has more influence on the risk of AD than hypertension later in life[50]. However, it is unclear whether hypertension during earlier phases of life also affect the incidence of dementia, and to what extent. Similarly, most studies on risk factors and compensatory mechanisms have primarily focused on midlife and beyond. It remains to be elucidated whether the risk factors already exert their effect on the etiology of AD during earlier phases of adulthood, and through which mechanisms.

Through the EU ERC Programme, funding was secured for a program entitled The Origins of Alzheimer's Disease Across the Life course (ORACLE) Study, which aims to further elucidate the age at which AD risk factors start affecting brain health and to further understand the underlying mechanisms. The ORACLE Study is embedded within the Generation R Study[22], a prospective birth cohort established in 2002 that focuses on health development from fetal life until early adulthood. The parents of the children had a mean age of 30.9 years (standard deviation: 5.7) at study intake and participated in extensive measures of their health. The ORACLE Study started in 2017, as a dedicated research visit for the parents to conduct extensive measures on brain health, including neuroimaging and cognitive testing.

The ORACLE Study has several aims. The first aim is to elucidate the associations of established and promising AD risk factors that were collected during previous Generation R visits with cognitive and brain measures from the ORACLE research visit. Such factors will include vascular risk factors like blood pressure and lipid profiles, mental health metrics such as depressive-like symptoms, and lifestyle factors like daily exercise. A second aim is to consider how brain health develops over a lifetime. The ORACLE Study bridges the age gap between the children of the Generation R Study and participants from the Rotterdam Study[23], a prospective cohort study in individuals aged 45 years and older. By combining these three study populations, brain health can be studied across the life course. Furthermore,

intergenerational effects can be examined by combining neuroimaging data from children and parents of the Generation R Study. A third aim is to study hypertensive disorders of pregnancy (HDP) and their impact on subsequent brain health and cognitive performance, as Generation R has extensive prenatal data. HDP affect 2-10% of all pregnancies and have been implicated as a potential risk factor for dementia[51]. However, little is known on how it influences subsequent brain health. The ORACLE Study will provide a unique opportunity to investigate the role of AD risk factors and HDP in brain health across adulthood.

In this manuscript, we give an overview of the measures that are collected from the parents as part of the ORACLE Study.

STUDY OVERVIEW

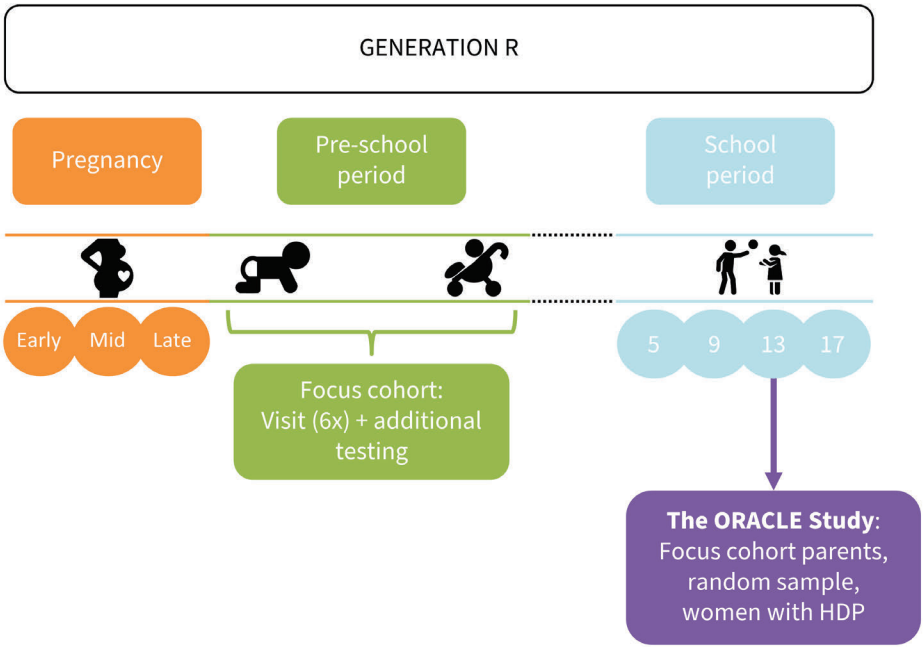
Generation R Study

The ORACLE Study is embedded in the Generation R Study, a population-based prospective cohort study from fetal life onwards based in Rotterdam, the Netherlands[22,52]. The Generation R Study was designed to identify early environmental and genetic causes of normal and abnormal growth, development, and health from fetal life until young adulthood[22]. All pregnant women living in Rotterdam with an expected delivery date between April 2002 and January 2006 were invited to participate. A total of 9,778 pregnant mothers and 6,347 partners were recruited into the Generation R cohort, which led to 9,749 live born children. A subset of 1,232 children and their parents – the "Focus subcohort" – have partaken in additional detailed measurements of both fetal and postnatal growth and development. A schematic overview of the Generation R Study is shown in **Figure 1**. All measures that have been performed in the parents of the Generation R Study have been described elsewhere[22,52]. The Generation R Study and the ORACLE Study have been approved by Medical Ethical Committee of the Erasmus Medical Center Rotterdam, the Netherlands. All participants have to provide written informed consent before participating in the study.

The ORACLE Study

The ORACLE Study is designed to test how factors during early adulthood affect brain health and structure at later ages. The parents within the Generation R Study provide a unique opportunity to tackle such questions. The parents had a mean age of 30.9 years (standard deviation: 5.7) at study baseline and data were collected for multiple factors that play a role in later adulthood brain health and the etiology of AD, like cardiovascular functioning, lifestyle, and life events. To explore whether and how these factors measured during early adulthood affect consequent brain health in middle adulthood, we introduced the ORACLE Study. It

Figure 1 | Flow of Generation R and the ORACLE Study. Parents were invited three times during pregnancy and returned to the research center with their children 5, 9 and 13 years after pregnancy. Additional detailed measurements of fetal and postnatal growth and development have been conducted in a subgroup of children (n = 1232, known as the 'Focus cohort') and their parents at 32 weeks gestational age and the postnatal ages of 1.5, 6, 14, 24, 36 and 48 months. The ORACLE Study started as part of the 13-year research visit.



consists of a parental research visit, with the goal to conduct cognitive testing, extensive neuroimaging, and an assortment of other physiological and functional measures.

The ORACLE Study started in May 2017 and is still ongoing. The aim is to recruit 2,000 parents whose children had also participated in the most recent wave of the Generation R Study (2016 – 2019). We have invited all parents from the Focus subcohort. The sample is further supplemented by randomly selecting parents from the whole Generation R cohort until 2,000 individuals have participated.

The ORACLE Study aims to form a bridge between two population-based cohorts: The Generation R Study[22] and the Rotterdam Study[23]. The Generation R Study focuses on early life (childhood, adolescence), the ORACLE Study includes individuals during early and mid adulthood, and the Rotterdam Study covers mid adulthood until the end of life. By combining these three studies, brain health can roughly be studied from a life-course perspective. To

harmonize the studies, the ORACLE Study has adapted the cognitive test battery from the Rotterdam Study as well as a similar set of brain magnetic resonance imaging (MRI) sequences.

Hypertensive disorders of pregnancy

A secondary aim is to assess whether HDP affects the structure and function of the post-pregnancy brain. As such, we aim to invite all mothers who had experienced HDP during their index pregnancy. HDP was determined for every pregnancy during the initial phase of the Generation R Study. Obstetric records were obtained from the midwife and hospital registries[22,53]. HDP was defined as pre-eclampsia and gestational hypertension. We used the criteria according to the International Society for the Study of Hypertension in Pregnancy of 2001[54]. Therefore, gestational hypertension was defined as development of a systolic blood pressure ≥ 140 mmHg or a diastolic blood pressure ≥ 90 mmHg without proteinuria after 20 weeks of gestation in previous normotensive women. Preeclampsia was defined as a new onset of hypertension with a SBP ≥ 140 mmHg or a DBP ≥ 90 mmHg and proteinuria (≥300 mg/day) at or after 20 weeks of gestational age. A total of 356 women with HDP were eligible for inclusion, and they were invited irrespective of being part of the Focus subcohort.

MEASURES

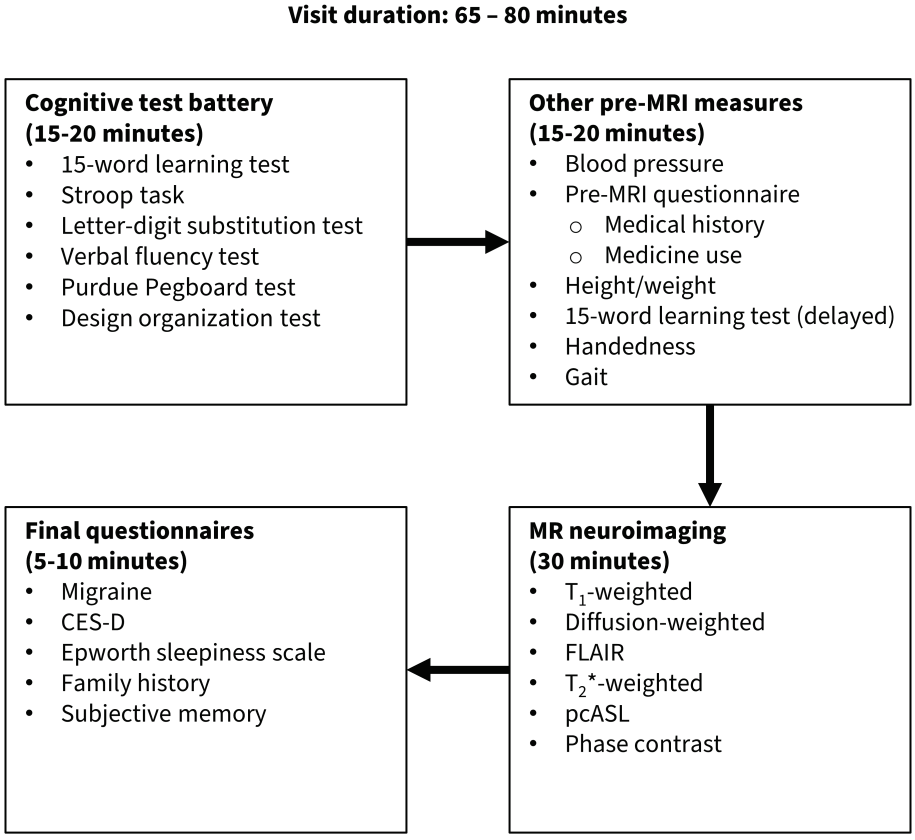
An overview of the ORACLE research visit and all measures is given in **Figure 2**. The visit starts with a cognitive test battery consisting of six tests (see Section 3.2): [1] the 15-word learning test, [2] the Stroop task, [3] a letter-digit substitution test, [4] a verbal fluency test, [5] the Purdue pegboard test and [6] the design organization test. The cognitive battery is followed by an assessment of gait (Section 3.3), a blood pressure measurement (Section 3.4) and questionnaires (Section 3.5). The participants are then scanned in an MRI scanner, and the session lasts for 30 minutes (Section 3.1). The total visit duration is approximately 65 to 80 minutes.

Neuroimaging

The aims of the neuroimaging are to:

- map and quantify the structure of the gray and white matter of the brain;
- map and quantify markers of cerebrovascular disease, i.e., white matter hyperintensities, brain infarcts (lacunar and cortical) and microbleeds;
- map and quantify the cerebral blood flow.

Figure 2 | Schematic overview of the ORACLE Study research visit. CES-D = Center for Epidemiological Studies – Depression scale; MRI = magnetic resonance imaging; FLAIR = fluid-attenuated inversion recovery; pcASL = pseudo-continuous arterial spin labeling.



We chose sequences and acquisition parameters that were comparable to the acquisition protocols of the Generation R Study and the Rotterdam Study while constraining the scanning time to 30 minutes. The finalized scan protocol is shown in **Table 1**.

Participants are excluded if they have any contraindications for the MRI, like metal implants or claustrophobia.

Scanner and equipment

As of 2013, the Generation R Study has a dedicated MRI scanner in the Erasmus MC-Sophia hospital[55], the same hospital that houses the Generation R research center. We are performing the ORACLE Study on this MRI scanner as well, to make the images more comparable to the

Table 1 | Parameters for the MRI sequences within the ORACLE Study.

Sequence	Comment	Mode	Time (m:s)	TR/TE (ms)	TI (ms)	Flip angle (°)	FOV (cm^2)	Matrix	Number of slices	Slice thickness (mm)	BW (kHz)
IR-FSPGR	ARC acceleration = 2	3D	4:48	8.8/3.4		10	22	220x220	196	1	25
DWI	35 directions; b = 1000 mm^2/s, b_0 = 3; PED P/A	3D	5:59	8200/84			24	120x120	65	2	250
DWI	b_0 = 3; PED A/P	3D	0:39	8200/84			24	120x120	65	2	250
FLAIR		2D	4:42	8761/120	2281		25	320x224	64	2.5	50
T2*	GRE	3D	5:54	min/23.9		12	25	320x224	160	1	41.67
pcASL	3 PLDs: 1,000, 1,570 and 2,460 ms	3D	3:16	5591/10.7			24		36	4	62.50
PC	Carotid and basilar flow; VENC = 80 cm/s	3D	1:53	min/3.8		10	23	256x160	20	2	41.67

A/P = anterior-posterior; ARC = autocalibrating reconstruction for Cartesian imaging; BW = bandwidth; DWI = diffusion weighted imaging; FLAIR = fluid-attenuated inversion recovery; FOV = field of view; GRE = gradient echo; IR-FSPGR = inversion recovery fast spoiled gradient recalled; P/A = posterior-anterior; PC = phase contrast; pcASL = pseudo-continuous arterial spin labeling; PED = phase-encoding direction; PLD = postlabeling delay; TE = echo time; TI = inversion time; TR = repetition time; VENC = velocity encoding.

images of the Generation R Study children. The scanner is a 3T GE Discovery MR750w MRI System (General Electric, Milwaukee, WI, USA) with the GE DV24 software package. The software package has intentionally not been updated since 2014, to ensure that the images remain relatively unchanged over the years. Images are obtained using an eight-channel head coil.

Several steps are undertaken to ensure comfort of the participants and to reduce participant motion. To reduce noise levels the participants are given earplugs and additionally headphones if this fits into the head coil. To ensure immobility of the head we use bilateral soft cushioning. A stiff cushion is also placed under the legs of the participants to reduce discomfort during scanning. Participants with back problems are offered additional pillows and other support. Finally, all participants are shown the same nature documentary during the scanning unless they prefer not to.

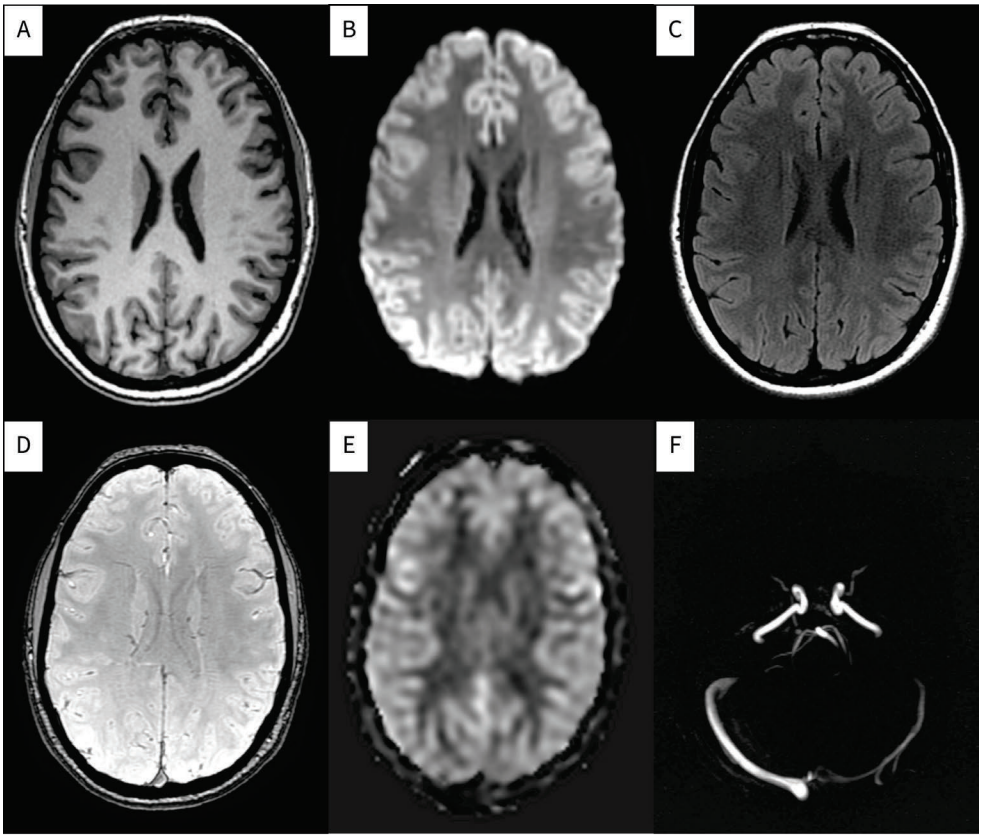
Image acquisition

The session starts with a three-plane localizer for positioning, and an ASSET scan to enable parallel imaging. T_1-weighted images to assess the structure of the brain are obtained using a 3D axial inversion recovery fast spoiled gradient recalled sequence (1 x 1 x 1 mm^3). The sequence is further accelerated by a factor of 2 using autocalibrating reconstruction for Cartesian imaging (ARC). White matter microstructural integrity is assessed with an axial spin echo sequence with an echo planar imaging (EPI) readout (2 x 2 x 2 mm^3). The gradient is set at b = 1000 m/s^2 in 35 directions with a posterior-anterior phase encoding direction. In addition, 3 sets of images with a gradient of b = 0 m/s^2 are collected. To be able to perform susceptibility distortion correction we also collect 3 sets of b = 0 m/s^2 images with an anterior-posterior phase encoding direction.

White matter lesions and infarcts are visualized using a 2D axial fluid-attenuated inversion recovery (FLAIR) sequence (0.8 x 1.1 x 2.5 mm^3). To visualize microbleeds we use a T_2*-weighted sequence (0.8 x 1.1 x 1.0 mm^3). Both are based on the sequence parameters of the Rotterdam Study to promote cross-study comparisons[13]. We further assess local blood perfusion with a pseudo-continuous arterial spin labeling (pcASL) sequence with three postlabeling delays (1,000, 1,570 and 2,460 ms). The total cerebral blood flow is quantified using an ungated 3D phase contrast sequence (velocity encoding = 80 cm/s) and based on the blood flow through the carotids and the basilar arteries.

Examples of each sequence have been compiled into **Figure 3**. Once the scanning session is completed, the images are send and stored in an XNAT storage instance[56].

Figure 3 | Example images from all sequences in the MRI protocol. (A) T_1-weighted, (B) diffusion-weighted, (C) FLAIR, (D) T_2*-weighted, (E) pcASL, (F) phase contrast.



Quality assessment

An initial assessment of general image quality is made by the radiographer during the scanning for the T_1-weighted sequence. The quality of the T_1-weighted scans is classified as poor, questionable, good, or excellent. In most cases, poor quality scans seem to be due to excessive movement. If the scans are rated as poor or questionable, the radiographer instructs the participant to try to lay as still as possible, and that the T_1-weighted sequence will be repeated.

Image quality of all sequences is further assessed during and after image processing with manual and automated methods. T_1-weighted segmentations are visually inspected by trained raters and each scan receives a fail/pass rating on a global level. This is done by rating a subset of slices in all three orientations (i.e., axial, coronal and sagittal) as well as the 3D reconstructions of the white and pial surfaces.

T_1-weighted images are also processed through a validated and automated quality assessment pipeline[57]. Images with low ratings from this pipeline are reinspected by the raters and excluded if justified. Furthermore, the quality rating from the automated pipeline will be used as a covariate in MRI-related projects, to see whether the analyses are confounded by subtle differences in image and processing quality.

Incidental findings

A by-product of population-based neuroimaging is the detection of incidental findings (IFs), i.e., abnormalities that are unrelated to the aims of the study but could bear clinical relevance. A meta-analysis study estimated the prevalence of such findings for brain MRI to be 2.7%, with neoplastic findings in about 0.7% of individuals[58]. However, it remains unclear what the best course of action is for incidental findings. For example, a follow-up study found that small meningiomas, the most common neoplastic finding, tend to remain stable and thus clinical intervention is not needed[59]. The UK Biobank also found that most clinical referrals of incidental findings did not benefit the participants while being accompanied by side effects such as emotional distress[60]. Incidental findings are thus to be expected but should be approached conservatively, to minimize burden on the participants.

Like the Generation R child neuroimaging waves, we opted for a two-layered approach in incidental finding detection and management[55]. The first layer takes place during the scanning session itself. The radiographer scrolls through the T_1-weighted images once, to detect any gross abnormalities. If any are detected, a certified neuroradiologist is contacted (author: MWV). The second layer is post-hoc inspection of the T_1-weighted images, the DWI images, the FLAIR images and the T2*-weighted images by trained personnel. A certified neuroradiologist subsequently checks all findings (author: MWV). When a finding is deemed clinically relevant, it is discussed in a broader consensus meeting where the decision is made to refer the patient for further diagnostic testing and/or follow-up or clinical intervention.

Cognitive testing

All participants are tested individually by trained examiners. The participants are told that they will take part in several tests and are asked to try their best. If participants ask about the purpose of the test, the experimenter indicates that they cannot disclose this during the test battery. If participants indicate that they want to restart or that they feel like they are doing poorly, the experimenter encourages them to continue and finish the test. Spare glasses are available for visually impaired participants who did not bring their own. Audio is recorded for the entire cognitive test battery if the participant consented. If needed, a stopwatch is used to measure time.

15-word learning test

The 15-word learning test is a neuropsychological test to assess the ability of verbal learning, retrieval, and recognition of verbal memory[61]. The test consists of three subtasks[62]. The first subtask consists of three trials where the same 15 unrelated words were used as stimuli. During each trial, the words are shown one by one on a computer screen, with a presentation time of 2 seconds per word, in the same order. After every trial, the participant is asked to name as many words as they can remember (immediate recall), and the trial is ended once the participant cannot recall any more words.

At least 20 minutes after the third trial, the participant is asked to name as many presented words as they can remember (delayed recall). Once the participant named all words that they can recall, they are presented with 15 previously shown words and 15 new words on the computer screen, one by one. The participants are asked to answer 'Yes' or 'No' to whether the item belonged in the list of immediate recall (recognition). The number of correctly recalled words in each trial are scored as the main outcome measures[61].

Stroop task

The Stroop task is a neuropsychological test used to assess the ability to inhibit cognitive interference. Cognitive interference occurs when the processing of a specific stimulus feature impedes the simultaneous processing of a second stimulus attribute, known as the Stroop Effect[63,64]. The Stroop task that is being used consists of three subtasks, and 40 stimuli for each subtask are distributed evenly in a 4 by 10 matrix[65]. The first subtask shows words of different colors (red, yellow, blue, or green) in black writing. The second subtask shows rectangles solidly colored in either red, yellow, blue, or green. The third and last task shows the words of different colors (red, yellow, blue, or green) in dissimilar ink color[65]. Every task must be read aloud as quickly as possible without mistakes. There is no time limit to complete each subtask. The time in seconds needed for a subtask is given as dependent measures (reading subtask, color naming subtask, and color-word interference subtask respectively)[61].

Letter-digit substitution test

The Letter-digit substitution test is a neuropsychological test to assess the ability of processing speed and executive function[61]. In the test, a key in which the numbers 1 to 9 are paired with a letter is given at the top of the worksheet[66]. Beneath the key, letters are given in a random order and the participant needs to pair the letters with the number according to the key. The first ten items are used for practice to ensure that the participant understands the test instructions. Following the practice round, the participant needs to pair as many numbers

as possible with the letters within 60 seconds in the given order. The number of correct items is used as the outcome of the test, with the maximum being 125 points[61].

Verbal fluency test

The verbal fluency test is a neuropsychological test to assess the ability of the efficiency of searching in long-term memory[61]. An animal-based task is used, where participants are asked to generate as many animal names as possible within 60 seconds[67]. The number of correct animals is used as the outcome of the test[61]. Furthermore, the examiners write out the answers in full for post-hoc construction of semantic networks across all participants[68,69]. This can be done to determine the degree to which participants cluster within and switch between semantic categories of animals[69].

Purdue pegboard test

The Purdue pegboard test is a neuropsychological test to assess dexterity and fine motor skill[61]. The Purdue pegboard is a rectangular board consisting of two columns of 25 holes[70,71]. Above these columns are reservoirs for metal pins. The participant needs to move the pins one at a time into the holes of the column on the same side of the board as the hand being used. They start at the top of the column and place as many pins as possible within 30 seconds. The task consists of three subtasks[70,71]. The first task is with the dominant hand, followed by the non-dominant hand, and then using the left and right hand simultaneously. The numbers of moved pins for each subtask are used as the outcome measure[61].

Design organization test

The design organization test is a neuropsychological test to assess visuospatial ability[61,72]. The participant is presented with nine designs consisting of squares. Six different squares are used: 1 black square, 1 white square, and 4 squares that are half-black and half-white divided along the diagonal in different orientations. Numbers from 1 to 6 are assigned to each type of square, with a key at the top of the sheet. Participants must convert the pattern designs to corresponding number designs. They must fill in as many numbers as possible within 120 seconds. The first nine patterns consist of five designs with 2 by 2 squares, followed by four designs with 3 by 3 squares, and thus a maximum of 56 points can be scored. The number of correct items completed is used as the outcome measure[61].

Gait assessment

The walking pattern, or gait, is a complex sequence of movements integrating sensory information and motor commands[73-75]. Gait is considered an accurate reflection of general health and is influenced by many organ systems such as the central and peripheral nervous

system, cardiovascular system, and musculoskeletal system[76-78]. Gait is assessed with an electronic walkway using pressure sensors (GAITRite; Sparta, NJ: 4.88-m active area; 120-Hz sampling rate) and is considered accurate to determine gait parameters. Participants perform a standardized gait protocol consisting of four different walking conditions: normal walk, turning, tandem walk, and dual task. In the normal walk, participants walk six times over the walkway at their own pace. In turning, participants walk at their usual pace, turn halfway, and return to the starting position. In the tandem walk, participants walk tandem (heel-to-toe) on the walkway. During the dual task, participants must count down from a 100 by seven while walking over the walkway. Raters inspect all recordings and identify individual footsteps. The GAITRite software returns a broad range of parameters that are commonly summarized into seven independent gait domains: rhythm, phases, variability, pace, tandem, turning, and base of support.

Until February 2018, the gait data were collected with the wireless GAITRite CIRFace system. We experienced intermittent technical issues primarily related to loss of wireless signal. Since March 2018, the gait data has been collected with the wired GAITRite RE system. Both systems have the same settings and characteristics, and so the data from both systems will be used for analysis.

Blood pressure and anthropometry

Systolic and diastolic blood pressure are measured with the validated automatic sphygmomanometer Omron 907 (OMRON, Matsusaka Co., Ltd., Japan)[79]. All participants are seated in upright position with back support. Blood pressure is measured two times over a 60-second interval, and the mean blood pressures are used for further analysis[80]. Furthermore, body weight (kilograms) and body height (centimeters) are measured after participants take off their footwear.

Questionnaires

Migraine

Migraine is assessed with a validated screening questionnaire[81]. This questionnaire includes five questions asking whether the participant had (i) severe headaches in the past 12 months, (ii) what the headache severity was, (iii) whether the participant had suffered from headaches which were preceded by visual disturbances, (iv) whether the participant had been diagnosed with migraine by a physician, and (v) whether the participant had ever used anti-migraine medication.

Depressive symptoms

Depressive symptoms are measured using the validated Dutch version of the Center for Epidemiologic Studies-Depression Scale (CES-D)[82]. The CES-D comprises 20 items, each with a possible score of 0-3, and the score ranges from 0-60.

Sleep propensity

Sleep propensity, one's readiness to transition from an awake state to sleep, is measured with a Dutch version of The Epworth Sleepiness Scale[83]. The scale comprises eight items, each with a possible score of 0-3, and the score ranges from 0 to 24. Higher scores indicate higher levels of sleep propensity.

Handedness

Handedness is assessed using a modified version of the Edinburgh Handedness Inventory[84]. This inventory contains questions on which hand someone prefers for a range of activities such as writing, holding a fork, and striking a match. An option was added for both hands, to capture ambidexterity. Furthermore, an item on eyedness ("which eye would you look with through a telescope") and footedness ("which leg would you use to kick a ball") were included as well. From these items, a laterality index ranging from -1 (preference for left) to 1 (preference for right) is calculated.

Subjective memory complaints

Subjective memory complaints are assessed with four self-reported question. The first question is "do you have more difficulty remembering things?" If participants answer "Yes", three follow-up questions are asked: the year in which these problems seemed to start, whether the problems started suddenly (no/yes), and whether the problems have changed over time (no/yes).

Other self-reported data

Additional information is obtained from participants through a semi-structured interview and a separate questionnaire. Participants are asked about their sleep during the last night (shorter, longer or the same as usual), and consumption of caffeine, nicotine, alcohol, and drugs in the past 24 hours. Furthermore, participants are asked about their medical status (presence of diseases) as well as medication use. We specifically ask about neurological and psychiatric diseases and disorders. Women are asked about their menstrual cycle, i.e., whether they menstruate and when they last menstruated. Finally, participants are asked

whether their biological mother or biological father have been diagnosed with dementia, Parkinson's disease, schizophrenia, or bipolar disease.

CURRENT STATE OF RECRUITMENT

Below we present the progress of the ORACLE Study from May 2017 up to September 2019. In total, 1579 individuals have been reached for an invitation to take part in the ORACLE Study, and 1307 (83%) have accepted that invitation. Of the respondents, 169 were HDP women, and 135 (80%) participated. Most participants were female (n = 842, 64.4%). The mean age during the visit was 46.4 years for the women (standard deviation: 4.4, range: 33.0 – 60.9) and 49.0 for the men (standard deviation: 5.0, range: 36.5 – 72.0). The mean follow-up time since the intake of the Generation R Study was 14.9 years (standard deviation: 0.8 years). Most participants (80.6%) reported to be of Dutch ancestry, 15.0% as non-Western and 4.4% as other Western. Finally, at study intake 58.8% of the participants had a university degree, 37.8% only finished secondary (vocational) education, and 3.4% only finished primary school or had no degree.

An overview of the participation rates for each measurement is given in **Table 2**. Out of 1307 participants, 1280 (97.9%) completed at least one MRI sequence. The remaining participants did not participate in neuroimaging due to claustrophobia (n = 11), contraindications (n = 9), technical scan issues (n = 5) and participants having to leave early (n = 2). Of the scanned individuals, 1180 (92.2%) had complete data on all sequences (99.8% T_1-weighted, 98.8% diffusion-weighted imaging, 99.1% FLAIR, 97.4% T_2*-weighted, 95.2% pcASL and 96.4% phase contrast). Early termination of a scan session was most commonly due to anxious feelings. Furthermore, the pcASL and the phase contrast were sometimes skipped due to time constraints if the T_1-weighted sequence had to be rescanned for quality purposes. Finally, the diffusion-weighted sequence was not performed in the first 11 participants due to technical issues.

Participation rates for the other tests were generally close to 100% (**Table 2**). Several measures have lower participation rates as they were introduced at a later time point during data collection, i.e., the Purdue pegboard test, the design organization test, the blood pressure measurements, and some of the questionnaires. In addition, the gait assessment experienced technical issues during the initial phase of the data collection, which resulted in a relatively low participation rate (78.0%). Other non-participation can be explained by participant refusal or inability to participate (e.g., muscle problems for the Purdue pegboard test), technical problems with the equipment (e.g., blood pressure measurements), or skipping of questions in questionnaires.

Table 2 | Participation rates for each measurement of the ORACLE Study up to September 2019.

Measurement	Eligible (n)	Available (n)	Participation (%)
MRI			
T_1-weighted	1280	1277	99.8
Diffusion-weighted imaging	1269	1254	98.8
FLAIR	1280	1269	99.1
T_2^*-weighted	1280	1247	97.4
pcASL	1280	1219	95.2
Phase contrast	1280	1234	96.4
Cognitive testing			
15-word learning test	1307	1300	99.5
Stroop task	1307	1300	99.5
Letter-digit substitution test	1307	1300	99.5
Verbal fluency test	1307	1300	99.5
Purdue pegboard test	840[a]	820	97.6
Design organization test	839[a]	830	98.9
Other measures			
Gait assessment	1307	1016	78.0
Blood pressure	905[a]	889	98.2
Anthropometry (weight, height)	1307	1302	99.6
Migraine questionnaire	905[a]	872	96.4
CES-D	905[a]	885	97.8
ESS	905[a]	884	97.7
Handedness	1307	1302	99.6
Subjective memory complaints	905[a]	879	97.1

[a] = These measures were introduced later in the study.

DISCUSSION

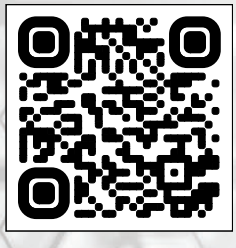
Given the high response rates and the high quality of the collected data, the ORACLE Study will likely reach its goal of 2000 participants with a few more months of data collection. Once the data collection is completed, the ORACLE Study will be open for collaborative projects. All requests for collaboration can be directed to study PI Professor M. Arfan Ikram (m.a.ikram@erasmusmc.nl).

2.2

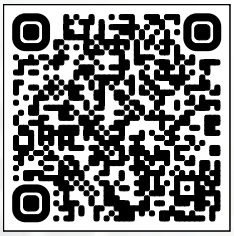
QDECR: a flexible, extensible vertex-wise analysis framework in R

Sander Lamballais, Ryan L. Muetzel

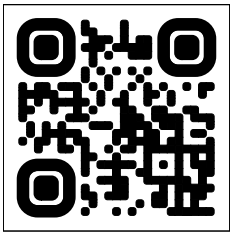
Adapted from:



Supplement:



Website:



ABSTRACT

The cerebral cortex is fundamental to the functioning of the mind and body. *In vivo* cortical morphology can be studied through magnetic resonance imaging in several ways, including reconstructing surface-based models of the cortex. However, existing software for surface-based statistical analyses cannot accommodate "big data" or commonly used statistical methods such as the imputation of missing data, extensive bias correction, and non-linear modeling. To address these shortcomings, we developed the QDECR package, a flexible and extensible R package for group-level statistical analysis of cortical morphology. QDECR was written with large population-based epidemiological studies in mind and was designed to fully utilize the extensive modeling options in R. QDECR currently supports vertex-wise linear regression. Design matrix generation can be done through simple, familiar R formula specification, and includes user-friendly extensions for R options such as polynomials, splines, interactions, and other terms. QDECR can handle unimputed and imputed datasets with thousands of participants. QDECR has a modular design, and new statistical models can be implemented which utilize several aspects from other generic modules which comprise QDECR. In summary, QDECR provides a framework for vertex-wise surface-based analyses that enables flexible statistical modeling and features commonly used in population-based and clinical studies, which have until now been largely absent from neuroimaging research.

INTRODUCTION

The cerebral cortex is integral to human psychological and physical functioning, and has been studied for centuries[85]. Modern neuroimaging techniques have enabled non-invasive assessment of the cortex, which has led to a myriad of studies elucidating the antecedents and consequences of typical and atypical cortical features. One common method for obtaining *in vivo* brain scans is with magnetic resonance imaging (MRI). The images begin as a grid of 3-dimensional grayscale pixels (voxels) which are commonly further processed by classifying (segmenting) the voxels into gray matter, white matter, and cerebrospinal fluid components[86]. Other tools subsequently trace and isolate the cortex to create surface-based representations of the brain as a series of interconnected points (vertices) forming a 2D mesh[87,88]. Specific characteristics of the cortex like its thickness or curvature can be derived from these surface representations at unique locations across the cortical mantle. These maps tend to consist of hundreds of thousands of vertices, allowing the study of the cerebral cortex at a fine-grain resolution. With advent of population neuroscience[89] and the introduction of several large-scale open-access neuroimaging initiatives[90,91], it is imperative that the neuroimaging community has a repertoire of tools available which are able to accommodate these massive, high dimensional data sets.

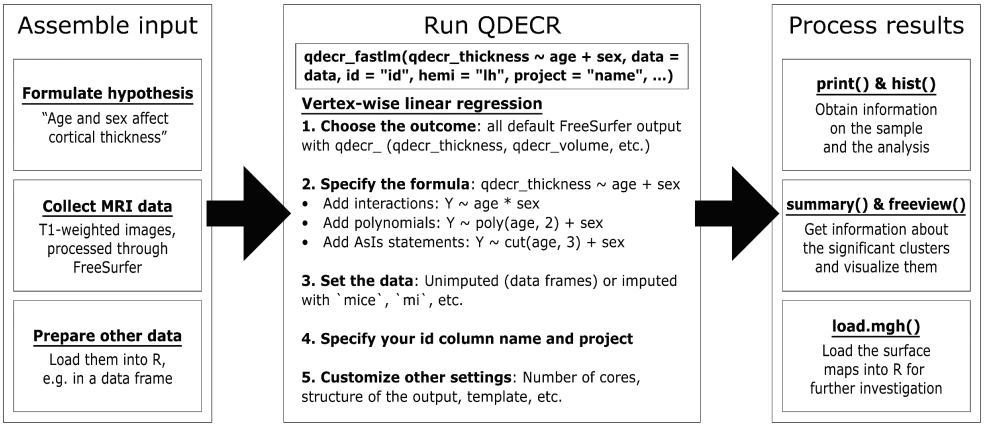
Tools for vertex-wise analyses of brain imaging data have been around since the creation of surface-based cortical models. One widely used tool which implements the linear model is Qdec (www.surfer.nmr.mgh.harvard.edu), which stands for Query, Design, Estimate, Contrast. Qdec is bundled with FreeSurfer[88], an open source software suite designed to generate surface-based maps of the brain from structural MRI data. Qdec facilitates whole-brain, vertex-wise analyses from a graphical user interface. Though brilliantly user-friendly, the interface has limitations including model specification (e.g., restrictions on the number of continuous and categorical variables that can be used) and handling of missing data. Qdec is the front-end interface which was built on top of the mri_glmfit program; a tool which was written in C++, works from the command line, and can handle larger datasets and more complicated design matrices. Another tool, SurfStat[92], was developed in Matlab and has a number of user-friendly features that mri_glmfit does not have, including formula-based creation of design matrices. SurfStat is still widely used (e.g., [93]), but has not been updated since 2008 and requires a Matlab license.

The field of neuroimaging is rapidly developing, particularly with studies generally growing in sample size due to the advent of open databases, consortia collaborations, and population neuroscience initiatives[89]. These studies, increasingly more epidemiologic in nature, require analytical tools that can handle statistical and epidemiological characteristics like big

datasets, correction for confounding and selection bias, allowing imputed data to account for missingness, and creating more flexibility in statistical model specification[94]. The previously described vertex-wise analysis tools were designed with vertex-wise analyses as their core purpose, but may lack features that are now crucial to begin integrating into neuroimaging as common practice to ensure proper analysis and interpretation of the data. Furthermore, expansion of features of those tools is not always straightforward as they were not designed in a modular fashion. Lastly, each tool is designed within a software framework which was not originally designed with statistical computing in mind.

We designed the QDECR package, a flexible and extensible R package for vertex-wise analyses. R is a programming language designed around statistical computing[95] and has become increasingly popular in academia and neuroimaging[96]. More importantly, R has a standardized syntax for statistical modeling, arguably has the most extensive statistical functionality of all existing programming languages, and its codebase is improving and expanding every day through a large user base and open source framework. We designed QDECR to fully utilize R's existing statistical infrastructure. We also designed QDECR to use the same user-friendly syntax as all other R modeling functions, so new users are immediately familiar with the QDECR syntax. Finally, we designed it to be an extensible and modular framework, where advanced users can implement their own type of statistical analyses on a vertex-wise level while still using core features of the framework (e.g., reading data, generating figures). In this manuscript we will describe the structure and features of the QDECR package.

Figure 1 | General workflow of the QDECR package, from data collection to post-processing.



MATERIALS AND METHODS

General workflow

The general workflow of the QDECR package is shown in **Figure 1**.

Input: FreeSurfer-processed images

Two sources of data need to be available to conduct analyses. The first stream of data consists of the T_1-weighted MR images, which have been fully processed through the FreeSurfer analysis suite (i.e., `recon-all`). Of note, after running the full primary `recon-all` reconstruction, users also need to run the `qcache` processing with `recon-all`. Activating the `-qcache` flag will co-register a given dataset into a standard coordinate system and spatially smooth the surface-based maps with a set of full width half max (FWHM) values. It outputs these maps in .mgh file format (Massachusetts General Hospital) in the "surf" subdirectory of the FreeSurfer output. As part of the FreeSurfer installation, the `SUBJECTS_DIR` environmental variable is set to indicate the directory where all the subject data are stored; QDECR will recognize this environmental variable, and users can set (or override) it via an optional argument when calling QDECR.

Input: Phenotype / covariate data

The second stream of data involves the other information relevant to the research question, specifically the phenotypic information of interest and covariates (e.g., age and sex). These data should be loaded into R with the user's method of choice, and ideally stored as a standard (imputed) data frame object. Furthermore, the phenotype data must include the identifier which was used to store the MRI data to link the two data types during analysis (i.e., the identifier variable the MRI data are stored on).

Analysis with QDECR

The next step is to run one of the analysis functions from the QDECR package, for example `qdecr_fastlm` for linear regression. At minimum, the following input arguments need to be specified:

- `formula`: a formula object, specifying the linear model to be used;
- `data`: a data frame containing the non-vertex (e.g., phenotype/covariate) data related to the research question;
- `id`: the name of the column in the data frame that identifies each subject;
- `hemi`: a hemisphere (`lh` for left, `rh` for right);
- `project`: a project name used for labeling output files.

As an example, to run a vertex-wise analysis to study the effect of age and sex on cortical thickness of the left hemisphere, the R code would be:

```
qdecr_fastlm(qdecr_thickness ~ age + sex, data = pheno, id = "id", hemi
= "lh",  project = "test")
```

where `pheno` is an R data frame object containing at least the columns `id`, `age`, and `sex`. All rows in `pheno` must correspond to an existing MRI session in `SUBJECTS_DIR`. During the analysis, information on the input data as well as the progress of the analysis will be printed on the console. The analysis will generate a number of files on disk (**Table 1**). The output of the analysis can be stored directly into an R variable, or it can be loaded back in at a later point in time.

In the current version of QDECR, results which are corrected for multiple comparisons (i.e., tests across all vertices) are by default also saved. This is done automatically in `qdecr_fastlm` using pre-cached smoothed Gaussian Monte Carlo – known as MCZ – simulations on a cluster level[97]. The cluster-forming threshold can be changed by specifying the `mcz_thr` argument (default = `0.001` based on previous work showing correspondence with full permutation tests[98]). The cluster-wise p-values are further corrected for performing additional analyses (e.g., in both hemispheres), which can be set with the `cwp_thr` argument (default = `0.025`, which is 0.05 Bonferroni corrected for running both left and right hemispheres).

The `qdecr_fastlm` function has many, additional arguments that users can specify. Information on the function and its arguments can be obtained by calling `?qdecr_fastlm`. Several arguments may be of particular interest to users. First, users can a path to `SUBJECTS_DIR` into `qdecr_fastlm` directly with the `dir_subj` argument, and to `FREESURFER_HOME` with the `freesurfer_home` argument. Second, the `target` argument allows for specification of the target template to use. Users can input templates that are available in `SUBJECTS_DIR`, but note that `-qcache` must be run with whichever target template has been specified. By default, the `fsaverage` template is used. Third, users can differentiate which level of smoothing (i.e., FWHM) should be used by using the `fwhm` argument, which is set to `10` by default.

Inspection of QDECR output

The output of the analysis can be explored with an array of functions within the QDECR package. Most of these functions were built on top of commonly used R functions. For example, the `print` and `summary` functions – which are familiar to most R users – can be used to extract information about the analysis and the significant brain areas (clusters) identified,

Table 1 | Overview of the output files. The `project` .rds file will have the name of the outputted project name. The `stack*` names will be replaced with the stack number, e.g., `stack1`. The `th*` names will be replaced with the threshold of the cluster-wise threshold. All `.cache.` files are output from `mri_surfcluster`.

Name	Per stack	Description
`project`.rds	no	A file that stores the qdecr output object and can be reloaded with `qdecr_load`.
finalMask.mgh	no	The final mask that was used for the analyses.
fwhm.dat	no	A file containing the estimated smoothness.
significant_clusters.txt	no	Contains all significant clusters (the output of `summary(vw, annot = TRUE)`).
stack_names.txt	no	Contains the link for variable name-stack number.
stack*.coef.mgh	yes	Contains the vertex-wise regression coefficients from the linear regression.
stack*.se.mgh	yes	Vertex-wise standard errors from the linear regression.
stack*.t.mgh	yes	Vertex-wise t-values from the linear regression.
stack*.p.mgh	yes	Vertex-wise p-values from the linear regression.
stack*.cache.th*.abs.sig.cluster.mgh	yes	Vertex-wise log10-transformed p-values of the cluster-wise significance.
stack*.cache.th*.abs.sig.cluster.summary	yes	Text file with summary information about clusters from the `mri_surfcluster` call.
stack*.cache.th*.abs.sig.masked.mgh	yes	Vertex-wise values after setting the non-cluster vertices to zero.
stack*.cache.th*.abs.sig.ocn.annot	yes	Vertex-wise annotations for the clusters to which each vertex belongs.
stack*.cache.th*.abs.sig.ocn.mgh	yes	Vertex-wise values for the clusters to which each vertex belongs.
stack*.cache.th*.abs.sig.voxel.mgh	yes	Vertex-wise value for the corrected voxel-wise significance.

respectively. Furthermore, QDECR provides functions to visualize the data (e.g., `hist` and `freeview`).

Internal structure

The field of neuroimaging has grown extensively over the last few decades, and a wealth of analysis methods have been developed. This can be daunting for new users, who would benefit from user-friendly and restricted analysis software. However, such restrictions may deter advanced users who require flexibility in applying their methods. QDECR was designed with both audiences in mind: straightforward and intuitive to use for beginners, yet flexible and extensible for advanced users. To achieve this, QDECR was designed to contain six modules:

1. **Input checking**. All input arguments undergo integrity checks. For example, provided paths are checked if they already exist in the system, and datasets are checked for the presence of the variables in the model specification, i.e., the formula.
2. **Model preparation**. The first steps of the statistical modeling are done here. The user-specified model is created and all steps that can be done before the vertex-wise calculations are processed. For example, in linear regression the portion of the design matrix that is common to all vertices is generated here.
3. **Loading the vertex-wise data**. In this step, the dataset and the provided paths are used to load in the vertex-wise data into a file-backed matrix.
4. **Vertex-wise analysis**. This module builds upon step 2 and runs the statistical model for every vertex. The output of each model is stored in dedicated file-backed matrices.
5. **Multiple testing correction**. Once all analyses are done, multiple testing correction can be applied across all vertices.
6. **Output generation**. An R object is compiled to contain all the information on the QDECR call, and output files are generated on disk to store the results more permanently.

These modules are implemented into the `qdecr` function. At its core, `qdecr` can handle any statistical model that is entered as long as a model preparation module (module #2) and a corresponding vertex-wise analysis module (module #4) exist. Functions like `qdecr_fastlm` are wrappers that automatically use the appropriate modules in `qdecr` to perform vertex-wise linear regression. Thus, users who only want to perform analyses do not have to think about any of the modules nor the underlying `qdecr` function, while advanced users can use the framework to implement new types of models more easily.

Formula objects

An important part of regression modeling is the creation of a design matrix. R uses formula objects in building design matrices. Formulas usually have three components:

1. the left-hand side (the outcome or dependent variable);
2. the right-hand side (the determinants or independent variables);
3. a tilde to separate the sides.

For example, in the formula `qdecr_thickness ~ age + sex`, the qdecr_thickness is defined as the outcome variable, and age and sex are denoted as the determinant variables. Using R formula objects for design matrix creation has several strengths:

- Categorical variables (like `sex`) are automatically recoded. By default, the levels will be dummy-coded according to the default behavior of linear regression in R, but other contrasts are available.
- Interaction terms can be introduced using the `*` (main effects plus interaction) or `:` (interaction only) symbols, for example `qdecr_thickness ~ age * sex`.
- Variables can be customized within the formula, for example by adding polynomial terms (e.g., `poly(age, 3)`), adding splines (e.g., `splines::ns(age, 3)`), standardizing a variable (e.g., `scale(age)`) and recoding of variables (e.g., `cut(age, 3)`).

New features for formulas can be seamlessly introduced, such has been done with the `Formula` package[99].

Thus, R formula objects – when used properly – allow for intuitive and extremely powerful behavior related to the creation of a design matrix. QDECR builds upon these principles, and in general most functions that manipulate formula objects will automatically work in QDECR as well, offering users continuity in syntax they already know from R.

Note that QDECR can handle all vertex-wise measures that FreeSurfer outputs by default, and the names are simply the FreeSurfer-assigned names preceded by "qdecr_" (**Table 2**), such as `qdecr_thickness` and `qdecr_area_pial`. The only modification is the “w-g.pct” file, which is written as `qdecr_w_g.pct` as the hyphen (or minus sign) has a specific meaning in R formula objects.

In certain cases, users may choose to create custom surface maps (e.g., functional activation maps). QDECR can be used to analyze those maps, by specifying the `custom_measure` argument of `qdecr_fastlm`. Users should supply the name of the vertex measure preceded by "qdecr_" (e.g., `qdecr_radialdistance`). Furthermore, the surface files must be placed in the “surf” directory of the FreeSurfer output of each participant. Finally, the surface files must follow the same naming convention as the other surface maps (e.g., "lh.radialdistance.fwhm10.fsaverage.mgh").

Table 2. Overview of the surface-based measures.

Surface measure in FreeSurfer	Name in QDECR	Description
area	qdecr_area	Surface area of the white matter surface
area.pial	qdecr_area.pial	Surface area of the pial matter surface
curv	qdecr_curv	Smoothed mean curvature
jacobian_white	qdecr_jacobian_white	The Jacobian of the spherical transformation
sulc	qdecr_sulc	Average convexity compared to the average surface
thickness	qdecr_thickness	Cortical thickness; the distance between the white and pial surfaces
volume	qdecr_volume	Cortical volume
w-g.pct	qdecr_w_g.pct	Gray to white signal intensity ratio
white.H	qdecr_white.H	Mean curvature of the white surface
white.K	qdecr_white.K	Gaussian curvature of the white surface

Statistical modeling of linear regression

The base model implemented in QDECR is a vertex-wise linear regression model with the vertex-wise metric, e.g., cortical thickness, as the outcome. At each vertex, a least squares regression would be performed:

$$\beta = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathrm{y},$$

where **X** is an *N* (subjects) x *p* (variables) matrix that is the design matrix, y is an *N* x 1 vector with the values at a given vertex, and β is a *p* x 1 vector of regression coefficients.

Running a linear regression for each separate vertex using the default `lm` function from the `stats` package would take a significant amount of time in R as R is an interpreted programming language. In interpreted languages the interpreting and execution of a line of code requires some operation time. Given the thousands of vertices that maps exist of, and given that linear regressions take milliseconds to perform, the compute time can become hours to days. The regression coefficients for all *m* vertices can be determined in a single step with the formula:

$$\mathbf{B} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{Y},$$

Where **Y** is an *n* x *m* (number of vertices) matrix with all vertex-wise values, and **B** is a *p* x *m* matrix of regression coefficients. Note that the vertex measures are the outcome, and thus the design matrix for all vertices is identical. To decrease run time QDECR therefore only builds the design matrix once.

The matrices may become exceedingly large given thousands of vertices, thousands of participants, and tens of imputed datasets. This would then exceed the RAM size of the RAM size of consumer grade computers. To avoid this, QDECR internally splits **Y** into 'chunks', or smaller partitions, so that the analyses can be run in smaller sets limiting the amount of required RAM at a given moment. By default, `qdecr_fastlm` creates chunks of 1000 vertices, but the size can be fine-tuned to a given setup (e.g., number of subjects, RAM availability, imputed datasets, etc.) with the `chunk_size` argument.

Handling imputed data

Missing covariate or phenotypic data in datasets can pose problems for statistical analyses. Previous vertex-wise tools require complete data, and thus any subjects with missing covariate or phenotypic data would have to be excluded for analysis, which could lead to loss of power and an increase in bias[100]. Rather than using only complete observations, methods have been developed to impute the missing data, typically under the assumption that the missingness is random and that the missingness can be predicted from other available data. To account for uncertainty in the imputation process, the imputation is repeated to generate several imputed datasets. For users who decide imputations are useful and feasible for their set of analyses, QDECR was designed to handle such imputed datasets from the most commonly used R imputation packages. Internally, QDECR uses a function called `imp2list` that converts any prespecified data object to a list of data frames. Consequently, `qdecr_fastlm` accepts the following object types for its `data` argument: Data frames, matrices, and lists of data frames, as well as imputed objects from the following R packages: mice[101], mi[102], amelia[103], and missForest[104]. Furthermore, users can implement methods for new classes by converting their object to a list of datasets. In regression analyses, the estimates across the imputed datasets are pooled using Rubin's rules[105].

Proof of principle for vertex-wise linear regression

To illustrate the QDECR package we performed vertex-wise analyses in 1000 randomly selected participants from the UK Biobank[91]. The participants had a mean age of 63.9 years (standard deviation: 7.7, range: 47.1 – 80.0) and 52.8% was female. The T_1-weighted images were processed with FreeSurfer version 6.0[88]. Additionally, in order to facilitate reproducible benchmarking and testing, a set of 10,000 simulated surface-based cortical thickness maps have also been made publicly available alongside a full installation of QDECR at Code Ocean[106], and can be freely explored and tested via the web interface (https://codeocean.com/capsule/6804031/tree/v2). A full tutorial on how to use QDECR can be found in the Supplemental Materials (Section 1) and via the GitHub repository (https://github.com/slamballais/QDECR).

RESULTS

Age Associations: Example in the UK Biobank

Within the UK Biobank sample, we aimed to study the association between age and vertex-wise cortical thickness, adjusted for sex. This can be achieved by running the following code in R:

```
vw <- qdecr_fastlm(qdecr_thickness ~ age + sex, data = pheno, id = "id",
hemi = "lh", project = "test_project")
```

The formula (i.e., `qdecr_thickness ~ age + sex`) captures cortical thickness as the dependent variable, and age and sex as the independent variables. The variable `pheno` contains the information on age and sex per participant, and the identifier is `id`. By specifying `hemi = "lh"`, we specify that the left hemisphere should be analyzed. Finally, the project name `test_project` is used, which will be incorporated in the names of the files that will be written to disk. The output is stored in the variable `vw`.

Once the analysis is done, a summary of the analysis can be viewed with `print(vw)` (**Figure 2A**). The significant clusters can be tabulated with `summary(vw)` (**Figure 2B**), which shows that a number of clusters have a significant association with age. Further inspection of the vertex-wise data can be done with `hist(vw)` (**Figure 3A**), which generates a histogram of the vertex-wise mean cortical thickness. The results can be visualized with the FreeSurfer FreeView tool by typing `freeview(vw)` (**Figure 3B**). Within the FreeView visualization it is clear that cortical thickness generally decreases with age, particularly in the temporal lobe and the precentral gyrus.

Age Associations: Comparison to mri_glmfit

To demonstrate the accuracy and consistency of the tool, we compared the results from the previous paragraph to those of mri_glmfit using the same data from the UKBB. A linear model for the association of age and sex with cortical thickness was generated in mri_glmfit with a DOSS (Different Offset, Same Slope) design. A full description of the input and code is given in the Supplementary Materials (Section 2). The vertex-wise regression coefficients – or betas – for age were examined. The mean absolute difference in the regression coefficients for age across the tested vertices was $7.9 \cdot 10^{-7}$, which arose from differences in rounding. The correlation between the betas was near perfect (Pearson's $r = 1.00$, Spearman's $\rho = 0.9999954$).

Figure 2 | Examination of the QDECR analysis on the association of age and cortical thickness in a subset of the UK Biobank. (A) The `print` function returns the key information of the analyzed project, including the input arguments, the included formula, the size of the dataset and the number of included vertices. (B) An example of output from the `summary` function. Each row represents a statistically significant cluster.

A

```
> vw
[input] Hemisphere:                          [lh]
[input] Project name:                        [test_project]
[input] Project final name:                  [lh.test_project.thickness]
[input] Number of cores:                     [1]
[input] Target:                              [fsaverage]
[input] dir_out_tree:                        [TRUE]
[input] clobber:                             [FALSE]
[input] fwhm:                                [10]
[paths] Subjects dir:                        [/data/subj]
[paths] Freesurfer home dir:                 [/opt/freesurfer]
[paths] Temp dir:                            [/data/]
[paths] Output dir:                          [/data/]
[paths] Path to default mask:                [/opt/R/library/3.4/QDECR/extdata/lh.fsaverage.cortex.mask.mgh]
[data ] N subjects:                          [1000]
[data ] N data points:                       [1000]
[data ] N datasets:                          [1]
[data ] Vertices loaded:                     [163842]
[model] Model:                               [RcppEigen::fastLm]
[model] Vertex data:                         [thickness]
[model] Formula:                             [qdecr_thickness ~ age + sex]
[mask ] Mask origin:                         [/opt/R/library/3.4/QDECR/extdata/lh.fsaverage.cortex.mask.mgh]
[mask ] Masked vertices:                     [149955]
[stack] Stack 1:                             [(Intercept)]
[stack] Stack 2:                             [age]
[stack] Stack 3:                             [sexmale]
[post ] Final mask path:                     [/data/lh.test_project.thickness/finalMask.mgh]
[post ] Final N vertices:                    [143392]
[post ] Estimated fwhm:                      [14]
[post ] Mean thickness per vertex:           [2.46066378601519]
[post ] SD thickness per vertex:             [0.264084935927758]
[post ] Mean thickness per subject:          [2.46066378601519]
[post ] SD thickness per subject:            [0.429482000166482]
```

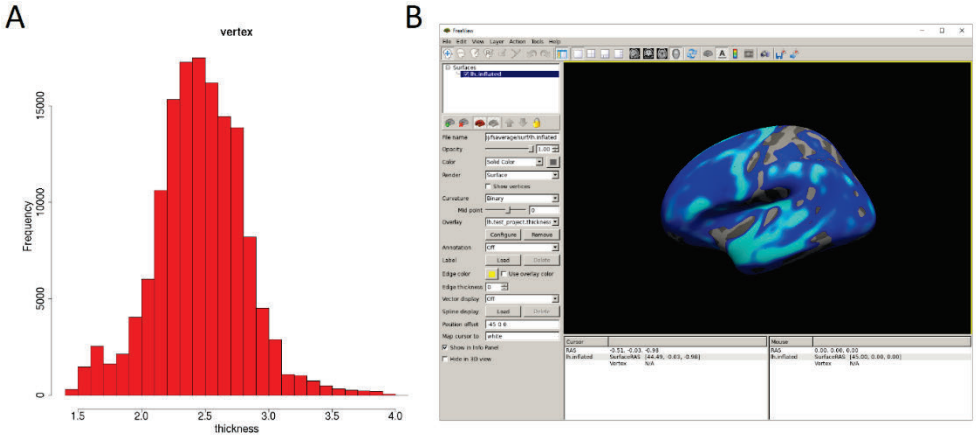
B

```
> summary(vw)
      variable cluster n_vertices mean_thickness mean_coefficient     mean_se
1  (Intercept)       1     143355       2.463277      2.783936728 0.068587854
2          age       1     105068       2.460255     -0.005839711 0.001052069
3      sexmale       1       7527       2.470069     -0.076713039 0.017572838
4      sexmale       2       5622       2.463866     -0.083187846 0.016560827
5      sexmale       3       3278       2.390507     -0.070495092 0.016238139
6      sexmale       4       2434       2.431146     -0.099360468 0.021204048
7      sexmale       5       1798       2.516486     -0.085744572 0.022282601
8      sexmale       6        931       2.484556     -0.061612288 0.016227576
9      sexmale       7       1313       2.512338     -0.088589393 0.019355318
10     sexmale       8        725       2.362900     -0.064267313 0.016279876
11     sexmale       9       1012       2.475424     -0.070191394 0.017521944
12     sexmale      10        707       2.581274     -0.061434269 0.015175473
13     sexmale      11        922       2.376967      0.077048640 0.016970790
14     sexmale      12        844       2.439652     -0.071649750 0.016546104
15     sexmale      13        579       2.623549     -0.067766057 0.015788556
16     sexmale      14        804       2.495248      0.054577059 0.012038274
17     sexmale      15        553       2.353505     -0.066493435 0.016591723
18     sexmale      16        658       2.441398     -0.074464551 0.018972812
19     sexmale      17        285       2.460376      0.043094365 0.011006375
```

Performance benchmark

To demonstrate how QDECR performs in terms of compute time in comparison to other tools, we used simulated data to benchmark the performance of `qdecr_fastlm`. We studied the influence of sex and age on cortical thickness of the left hemisphere in 100, 500, 1000, 5,000 and 10,000 participants. Further, we tested the impact of multiple imputation by generating 1, 33 or 100 imputed datasets. Lastly, we studied the impact of parallel processing by using 1 or 4 CPU cores. **Figure 4A** contains the results of the benchmark. The time it took to perform

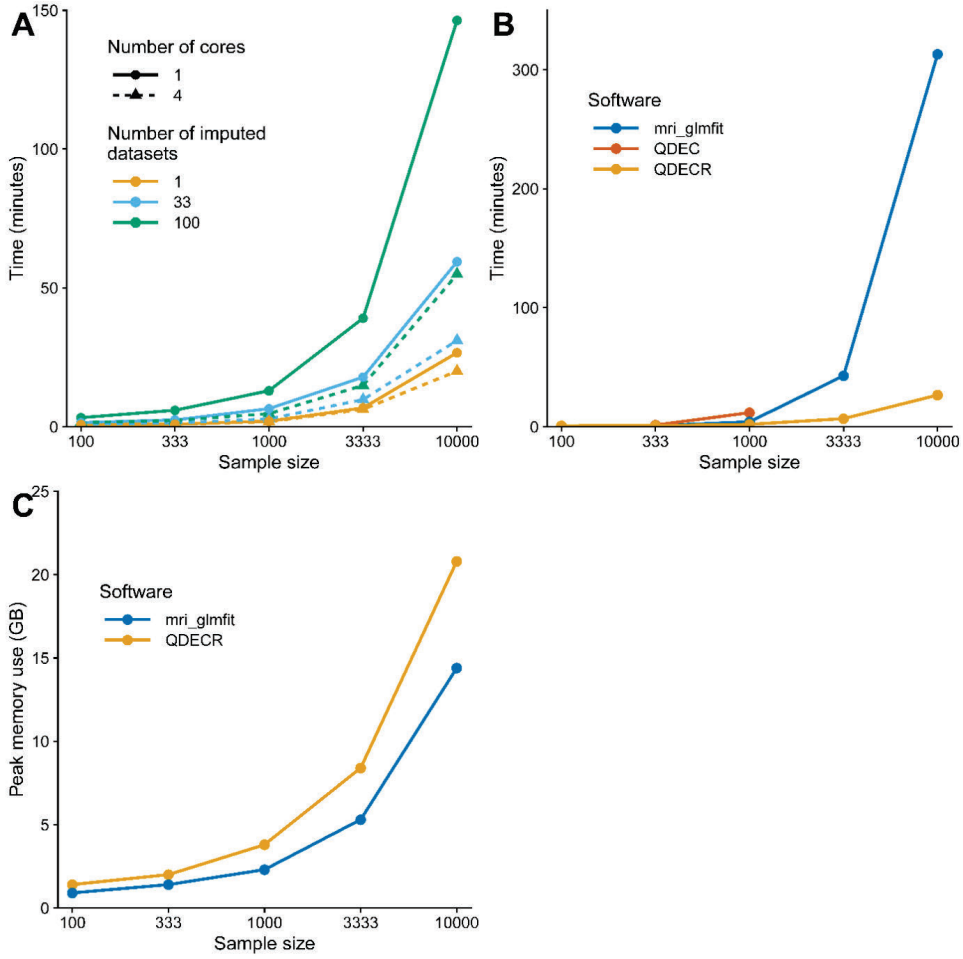
Figure 3 | Plots of the QDECR analysis results. (A) A histogram of the vertex-wise mean cortical thickness for the study sample, generated with `hist`. (B) FreeView can be called with the `freeview` function. This panel displays the vertex-wise map for the effect of age on cortical thickness.



an analysis on unimputed data ranged from 0.8 minutes for a sample of 100 datasets to 26.7 minutes for a sample of 10,000 datasets. When the dataset with 10,000 was further imputed 100 times for missing covariates, the time increased to 146.4 minutes (i.e., 448% longer than unimputed). The performance can be boosted by recruiting more CPU cores for the analysis. For example, using 4 cores compared to 1 core on 10,000 datasets with 100 imputed datasets dropped the analysis time from 146.4 to 54.9 minutes (i.e., 62.5% reduction).

We further compared the QDECR performance with Qdec and mri_glmfit (**Figure 4B**). Both Qdec and mri_glmfit were faster than QDECR on samples of 100 individuals, and slower for samples with 333 participants and more. Furthermore, Qdec was not able to finish the analyses on 3,333 and 10,000 participants due to errors that arose when merging the underlying MRI data into a single set. While mri_glmfit did succeed in running the analyses, compared to QDECR it was much slower for the set of 3,333 individuals (42.8 minutes vs 6.8 minutes with QDECR) and 10,000 individuals (313.1 minutes vs 26.7 minutes with QDECR). Finally, we compared the peak memory use of QDECR with that of mri_glmfit (**Figure 4C**). Overall, QDECR had a higher peak memory use than mri_glmfit. Where QDECR reached a peak memory use of 8.4 GB for 3,333 individuals and 20.8 GB for 10,000 individuals, mri_glmfit used 5.3 GB and 14.4 GB, respectively.

Figure 4 | Computation time benchmark of the QDECR package. (A) Displays the computation time of an analysis with QDECR when varying sample size, number of cores (threads) used and number of imputed datasets that were included. (B) Compares the computation time for mri_glmfit, QDEC and QDECR when using a single core (thread) and unimputed data. For QDEC we used the time that the analysis took, without loading or assembling of the data or the multiple testing correction. For mri_glmfit we measured the time it took to run three commands: [1] mri_preproc to assemble the FreeSurfer output, [2] mri_glmfit to run the analysis and [3] mri_glmfit-sim to perform the multiple testing correction. **Note: QDEC returned errors when attempting to run the analyses on 3,333 and 10,000 participants and is thus not represented for those sample sizes in panel B. (C) Compares the peak memory use in gigabytes (GB) for mri_glmfit and QDECR when using a single core (thread) and unimputed data.



DISCUSSION

QDECR provides a framework to perform vertex-wise analyses in R. It has the same base functionality as other vertex-wise tools and adds several functionalities. We have shown that QDECR runs faster on large datasets than other tools and can additionally handle imputed datasets to minimize bias or loss of power due to exclusion of participants. Moreover, we aimed to maximize user friendliness for individuals familiar with R through the implementation of formula objects to handle design matrix specification and through writing functions with similar arguments as other base functions. Finally, QDECR sets the stage for further development of statistical applications to study the cerebral cortex in population neuroscience settings.

QDECR has several limitations. The primary focus of the package has been to implement vertex-wise analyses in R. In contrast, glmfit has a myriad of options for the MRI data available. It works with both voxel-wise and vertex-wise data with all volume files that are recognized by the FreeSurfer `mri_convert` function (.mgh, .nii, etc.). It also has several options related to the analysis that are not available in QDECR yet, such as different methods for multiple testing correction (e.g., permutation testing) and weighted least squares. However, QDECR is still in development, and these options will likely be available in the future. Another limitation for part of the potential users is that QDECR is only available in R. Qdec will therefore remain more feasible for users with little programming experience, and MATLAB and Python users would have to learn basic R skills to use it. Furthermore, developing new modules requires mastery of R. Still, we opted for R as it provides an ideal environment to further develop the statistical options for vertex-wise analyses. Furthermore, R is gaining popularity in medical research, especially with the advent of Bioconductor[107] and more recently Neuroconductor[96].

While QDECR presents a substantial contribution to vertex-wise analyses, several areas of opportunity for expansion and improvement exist. First, at the moment only the general linear model is implemented. We envision logistic regression, linear mixed models, and structural equation models to be the next key targets for future implementation. Next, though QDECR relies on multiple testing correction that is native to the FreeSurfer library, new modeling techniques may require new methods for adjusting analyses for multiple comparisons. Thus, another target for development is implementing permutation testing and other state-of-the-art methods in the field of neuroimaging. Finally, QDECR in its first implementation can handle the mgh file format and assumes a FreeSurfer image reconstruction. In the future, new methods should be implemented to accommodate other file format types (e.g., Minc/Civet[108]), and allow for 3D voxel data in addition to surface data (e.g., Nifti format data).

CONCLUSION

QDECR extends the capabilities of existing whole-brain vertex-wise statistical software for neuroimaging data analysis, allowing for larger (population-based) datasets, incorporation of novel epidemiological and statistical concepts, and elegant expansion within the widely used and open-source R statistical framework.

Software availability statement

The code is made available at https://github.com/slamballais/QDECR. Supporting information and further tutorials can be found at https://www.qdecr.com. QDECR uses the GNU General Public License version 3.0 (GPL-3).

Data availability statement

The data used to generate the figures will not be made available. However, the code and simulated data are available in a dedicated Code Ocean code capsule[106].

Authors' contributions

R.L.M. conceived the package. Both authors designed the package. S.L. adapted all the functions and code and optimized the performance. Both authors wrote the manuscript and approved the final manuscript.

Funding

The work was supported by the European Union's Horizon 2020 Research and Innovation programme (project: ORACLE, grant agreement No: 678543), the Sophia Foundation (grant S18-20 [R.L.M.]) and the Erasmus University Fellowship [R.L.M.].

Acknowledgements

This research has been conducted using the UK Biobank Resource under application number 23509. We would like to thank Doug Greve for giving us extensive comments and feedback on the design of the package. We would also like to thank Heath Pardoe, who wrote the code for the `load.mgh` and `save.mgh` functions that are included in the package.

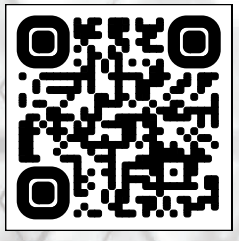
GENETIC MECHANISMS

3.1

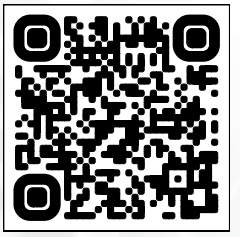
Genetic scores for adult subcortical volumes associate with subcortical volumes during infancy and childhood

Sander Lamballais, Philip R. Jansen, Jeremy A. Labrecque, M. Arfan Ikram, Tonya White

Adapted from:



Supplement:



ABSTRACT

Individual differences in subcortical brain volumes are highly heritable. Previous studies have identified genetic variants that underlie variation in subcortical volumes in adults. We tested whether those previously identified variants also affect subcortical regions during infancy and early childhood. The study was performed within the Generation R Study, a prospective birth cohort. We calculated polygenic scores based on reported GWAS for volumes of the accumbens, amygdala, brainstem, caudate nucleus, globus pallidus, putamen, and thalamus. Participants underwent cranial ultrasound around 7 weeks of age (range: 3-20), and we obtained metrics for the gangliothalamic ovoid, a predecessor of the basal ganglia. Furthermore, the children participated in a magnetic resonance imaging (MRI) study around the age of 10 years (range: 9-12). A total of 340 children had complete data at both examinations. Polygenic scores primarily associated with their corresponding volumes at 10 years of age. The scores also moderately related to the diameter of the gangliothalamic ovoid on cranial ultrasound. Mediation analysis showed that the genetic influence on subcortical volumes at 10 years was only mediated for 16.5-17.6% of the total effect through the gangliothalamic ovoid diameter at 7 weeks of age. Combined, these findings suggest that previously identified genetic variants in adults are relevant for subcortical volumes during early life, and that they affect both prenatal and postnatal development of the subcortical regions.

INTRODUCTION

Subcortical brain structures are a collection of diverse regions with a myriad of functions. Subtle differences in these regions have been implicated in cognitive function, emotion regulation and certain psychiatric disorders like schizophrenia and autism spectrum disorder[109-112]. Like the rest of the brain[2,113-115], the subcortex develops over a lifetime, with emphasis on expansive growth during the prenatal period, reorganization during early life and atrophy during late life. The development of subcortical regions is partly shaped by stochastic events[116] and by environmental factors such as early-life adverse events[117-119]. Subcortical development is also strongly influenced by genetics, as the heritability has been estimated to be around 60-80% for subcortical volumes[120,121]. Further elucidating the role of genetics in the developmental trajectories of subcortical regions may in turn provide insight into the functional consequences of the subcortex.

The strong heritability of subcortical volumes likely reflects the combined effect of weak signals across the genome rather than a single gene. This was confirmed through genome-wide association studies (GWAS)[122], which aimed to establish associations between a given phenotype and a wide range of markers of genomic variation such as single nucleotide polymorphisms (SNPs). In 2015, Hibar and colleagues performed GWAS on the volumes of several subcortical regions, and identified a total of seven SNPs that specifically related to the volumes of either the putamen, the caudate nucleus or the hippocampus[122]. More recently Satizabal and colleagues performed a GWAS in over 40,000 adults and extended the findings to 25 genetic loci, potentially implicating 62 different genes[11]. These loci provide an opportunity to further understand the genetic influences underlying subcortical development.

The GWAS informs about the genetic underpinnings of subcortical volumes, but its study population consisted primarily of adults. It is not clear whether these findings would generalize for studies on subcortical volumes during childhood or even infancy. The GWAS likely captured signal that related to all aspects of changes in subcortical volumes, i.e., early-life development, the height of the peak in subcortical volume and the degenerative processes leading to a decline in volume with age. Furthermore, children may reach the same peak subcortical volume during adulthood, but may do so at different rates, which would also further diminish the value of the identified genetic markers in a pediatric population.

We therefore aimed to study two related questions. First, we aimed to study whether the genetic markers that were identified during adulthood relate to subcortical volumes earlier in life, i.e., infancy and early childhood. Second, if the genetic markers do associate with early-life subcortical volumes, it remains unclear during which life phase the genetic effects

primarily take place. Subcortical volumes during infancy likely correlate strongly to subcortical volumes during early childhood. If the genetic effects mainly occur during fetal development, then the associations would still carry over to early childhood. We would want to study whether the genetic markers still associate with subcortical volumes during early childhood when considering subcortical volumes during infancy.

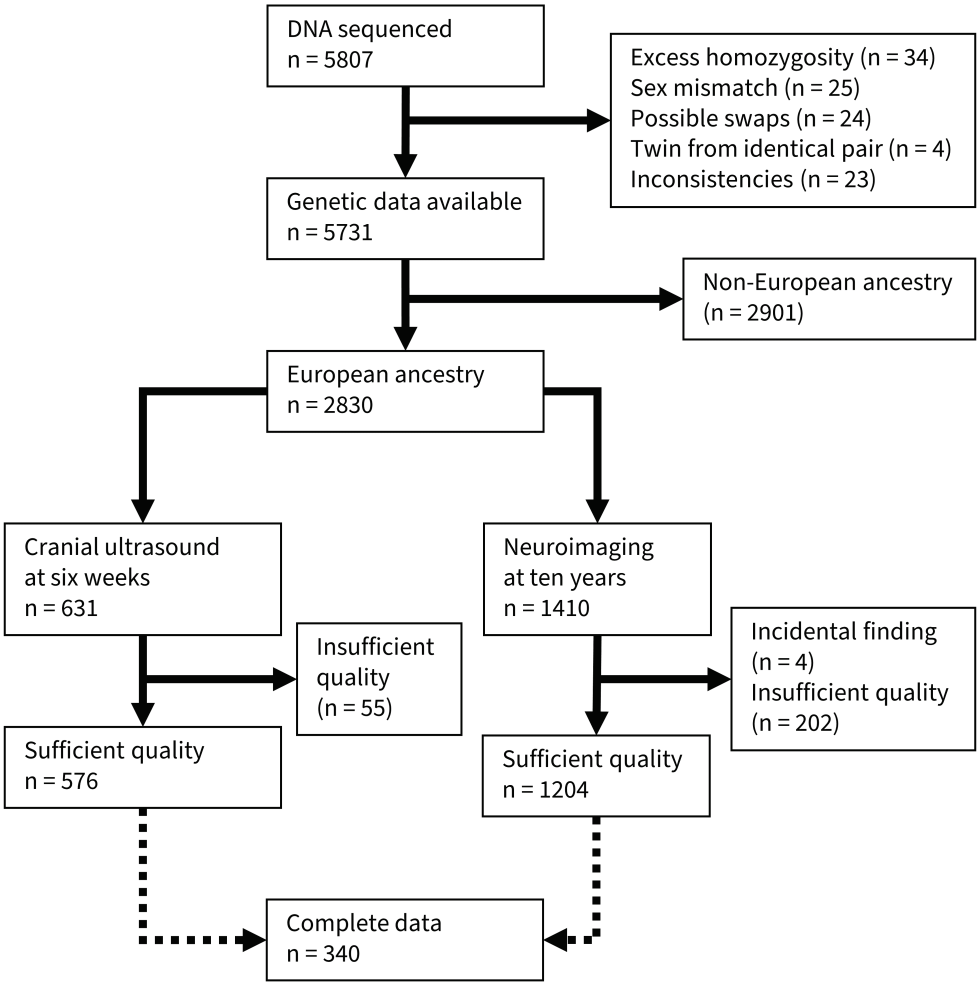
Within this context, we aimed to study during which life phases the adult-derived genomic loci are relevant. We performed the current study in the Generation R Study, a prospective birth cohort based in Rotterdam, the Netherlands[22]. A subset of the children of Generation R was invited for a cranial ultrasound around seven weeks of age. From these images we quantified the gangliothalamic ovoid (GTO) diameter (GTOD), the diameter of a subcortical structure that develops into the basal ganglia. At around ten years of age, the children were additionally invited for an MRI scan of the brain. Next, we utilized a technique known as polygenic scores (PGS), a score calculated based on GWAS results that quantifies a person's genetic predisposition for a trait, in this case the subcortical volumes. Within Generation R we obtained genotypic information, and for each child we calculated PGS for each subcortical region. We then assessed whether the PGS associated with the ultrasound metrics during infancy and the MRI scan during childhood. Finally, to see whether the PGS associations with subcortical volumes during childhood were mediated by the PGS associations with infant subcortical volumes we formulated a causal mediation model. Importantly, the GTOD measurement is coarser than the MRI measurements given that the GTOD is a diameter while the MRI measurements are volumetric, like the original GWAS. While this will reduce the accuracy to estimate the mediation, we do not expect systematic noise in the GTOD measurement that could bias the estimation. Furthermore, the analysis should provide some insight into whether the infant subcortical measures mediate the association between the PGS and the childhood subcortical volumes.

METHODS

Study population

The study was conducted within the Generation R cohort, a prospective birth cohort based in Rotterdam, the Netherlands[22]. At approximately seven weeks of age, a subset of the children partook in cranial ultrasound measurements. At approximately ten years of age, most of these children also visited the research center for MRI scans of the brain[55]. A flowchart of the sample is shown in **Figure 1**. Out of 2,830 children of European ancestry and with genetic data available, 576 had data on the seven-week ultrasound and 1,204 had data on the ten-year MRI. In total, 340 children had data on both measures.

Figure 1 | Flowchart of the study population.



Ethics statement

The study was approved by the Medical Ethical Committee of the Erasmus MC University Medical Center in Rotterdam, and written informed consent was obtained from all primary caregivers of the participants.

Genotyping and polygenic risk scores

All genotyping and quality control procedures have been described elsewhere[123,124]. Samples were collected from cord blood at birth (Illumina 610K Quad Chip) or from venipuncture at a visit to the research center at the age of around six years (Illumina 660K Quad Chip). The

Illumina 610K and 660K samples were merged based on their overlapping SNPs. Quality control was performed using PLINK (version 1.9)[125]. SNPs were removed if the minor allele frequency was below 1%, the Hardy-Weinberg disequilibrium p-value below > 1×10^{-6} or a SNP call rate of less than 98%. Individual data were removed in cases of genetic and phenotypic sex mismatch, excess rates of homozygosity of the genotypes (> 4 standard deviations) and genotype quality (> 5% missing). Pairs of individuals with identical genetic information but who were not identical twins were removed, as they were likely samples that were processed twice. For identical twins, the twin with the lower call rate was removed.

The current study was based on a GWAS on subcortical volumes[11]. The GWAS included 40,000 participants of predominantly European ancestry, and they did not overlap with the present study sample. They performed a GWAS for each subcortical volume. We utilized the summary statistics to calculate a PGS per subcortical region. PGS can be calculated by summing the effect size of each SNP for a given person[126]. However, SNPs have proximal dependence and tend to co-occur the closer they are to each other, a principle known as linkage disequilibrium (LD). A common method to adjust for this is a sliding window approach, where the SNP with the lowest p-value within that range is kept and all other SNPs are pruned. We used a more refined method named LDpred[126], which utilizes LD information from an external dataset to improve prediction accuracy. Furthermore, different scores can be calculated depending on the fraction *P* SNPs that are assumed to be causal. We used different values for the *P* parameter: 1.0, 0.5, 0.1, 0.05, 0.01 and 0.005. The 1000 Genomes project was used as the external LD dataset[127]. To account for the European ancestry in the original GWAS we calculated the genomic components of our study population with the multi-dimensional scaling function of PLINK[124,125], and we subsequently selected participants of European ancestry.

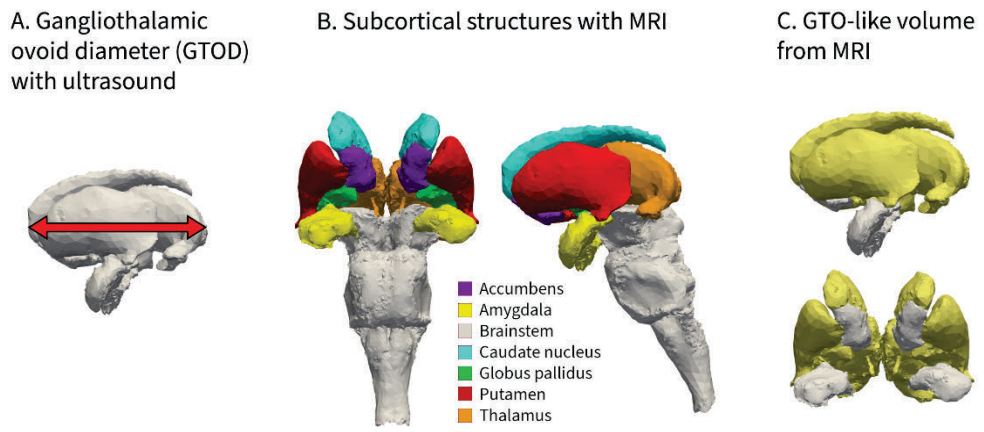
Ultrasound image acquisition and processing

Three-dimensional cranial ultrasound was performed at approximately seven weeks of post-natal age using a multifrequency electronic transducer (3.7-9.3 MHz) with a scan angle of 146° (Voluson 730 Expert, GE Healthcare, Waukesha, WI, USA). The procedure has been described elsewhere[128,129]. In brief, the probe was placed on the anterior fontanel. The volume box was positioned at the level of the foramen of Monro, and a pyramid-shaped volume of brain tissue was imaged. Images were subsequently used for measuring the GTOD (**Figure 2A**), which was done by two trained raters. The reliability of the measurements was high (Cronbach's $\alpha = 0.83$), and thus the average of the diameters of the two raters were used for analysis.

MRI Image acquisition and processing

The MRI imaging and image processing at the 10-year visit has been described elsewhere[55]. In brief, structural brain images were obtained on a 3T GE Discovery MR750w MRI System (General Electric, Milwaukee, WI, USA) using an 8-channel receive-only head coil. T_1-weighted images were obtained using an inversion recovery-prepared fast spoiled gradient recalled (IF-SPGR) sequence (T_R = 8.77 ms, T_E = 3.4 ms, T_I = 600 ms, flip angle = 10°, field of view = 220 x 220 mm, acquisition matrix = 220 x 220, slice thickness = 1 mm, number of slices = 230, bandwidth = 25 kHz). The images were processed through the FreeSurfer analysis suite, version 6.0[130]. After removal of non-brain tissue and normalizing voxel intensities for B_1 homogeneities, the images were segmented into prespecified cortical and subcortical regions, and their volumes were determined. We extracted summary information on volume of the nucleus accumbens, the amygdala, the brainstem, the caudate nucleus, the globus pallidus, the putamen, and the thalamus (**Figure 2B**). We averaged the volume of the left and right hemispheric subcortical structures as this was done in the original GWAS on subcortical volume as well and we did not expect a lateralized effect.

Figure 2 | Visual representation of the subcortical structures under study. Panel A shows how the gangliothalamic ovoid diameter was determined from the ultrasound at seven weeks. Panel B displays the subcortical regions that were obtained from the MR images at the ten-year visit. Panel C shows how we approximated a GTO-like volume using the MR image data, i.e., by summing the volumes for the caudate nucleus, the globus pallidus, the putamen, and the thalamus. The images were based on data from the MIDA model[145].



Statistical analysis

The study focused on four sets of analyses. Analyses were performed using R, version 3.5.1[95].

Subcortical PGS and childhood MRI

In the first set of analyses, we used linear regression models to explore whether PGS for subcortical regions associated with subcortical volumes obtained with MRI at around 10 years of age. We first assessed whether PGS and volumes for the same region were associated (within-region associations). Next, we considered whether PGS and volumes of differing regions were associated (between-region associations). We created a separate linear regression for each combination of PGS and subcortical volumes across all values for parameter P, yielding 7 x 7 x 6 = 294 models. In each model, the PGS and the volumes were standardized to increase comparability of the findings. Furthermore, to control for potential confounding, we adjusted for age during the MRI (in years), intracranial volume (in ml), sex and the first ten genomic components. To adjust for multiple testing, we applied a Bonferroni correction to the 7 x 7 different models within a given parameter P. The PGS based on different values for parameter P tended to correlate very strongly (median Pearson's r = 0.91), thus due to the conservative nature of the Bonferroni correction, we did not additionally correct for each parameter of P. The alpha level was thus set at 0.05 / 49 = 0.00102.

GTO PGS and infant US

For the second set of analyses, we used linear regression models to see whether the subcortical PGS associated with GTOD as obtained from the ultrasound during infancy. As the GTO is a combination of different subcortical regions, we created a PGS for the GTO by summing the PGS of several regions together. Given that these subcortical regions differ in size, we calculated a relative GTO PGS according to the following formula:

$$PGS_{(GTO)} = 0.22 \cdot PGS_{(caudate)} + 0.10 \cdot PGS_{(pallidus)} + 0.28 \cdot PGS_{(putamen)} + 0.40 \cdot PGS_{(thalamus)}$$

The weights were obtained by considering the mean relative size of the subcortical regions as obtained from the MRI around the age of 10 years, as we did not have information on the relative size of these regions from the ultrasound during infancy. Like the first set of analyses, the variables were standardized, and the models were corrected for gestational age at birth (in weeks), age at ultrasound (in weeks), head circumference (in cm), sex and the first ten genomic components. Given that only one analysis was performed per value for parameter P, we did not apply correction for multiple testing and the alpha level was set at 0.05.

GTO PGS and childhood MRI

As a third set of analyses, we wanted to assess whether the GTO PGS associated with the equivalent of the GTO at age 10 years. We therefore calculated a GTO-like volume, where the volumes of the caudate nucleus, the pallidi, the putamen and the thalami were summed together (**Figure 2C**). Then, we used linear regression to assess the association between the GTO PGS and this GTO-like volume. The models were specified similarly as the first set of analyses. Given that only one analysis was performed per value for parameter P, we did not apply correction for multiple testing and the alpha level was set at 0.05.

Mediation analyses

For the fourth set of analyses, we explored whether the associations between the PGS and the subcortical volumes at 10 years of age were mediated by the GTOD during infancy. The genetic loci captured by the PGS may drive both prenatal and postnatal effects on neurodevelopment. Any relevance to prenatal development should be expressed by stronger associations between a GTO PGS and the GTOD from the postnatal ultrasound at 7 weeks than the association between a GTO PGS and the GTO-like volume from the MRI at 10 years. More critically, if the PGS primarily affects postnatal development then the PGS should associate with GTO-like volume from MRI at 10 years when also considering the GTOD from the postnatal ultrasound at 7 weeks. We therefore created a causal mediation model with the GTO PGS as the determinant, the GTOD from the 7-weeks postnatal ultrasound as the mediator and the GTO-like volume at age ten as the outcome[131]. Causal mediation allows for estimation of the relative contributions of the natural direct effect of determinant on the outcome and the natural indirect effect that runs from the determinant to the outcome through the mediator. The confidence intervals were estimated through nonparametric bootstrapping. Causal mediation assumes that all confounding factors for the direct effect and the indirect effect are accounted for. We therefore corrected for sex, age, intracranial volume and the first ten genomic components when estimating the exposure-outcome association. For the mediator-outcome association we additionally corrected for maternal education at birth of the child (low, medium, or high) and birth weight of the child in grams. The analyses were performed using the R `mediation` package[132].

Samples used in the analyses

The main measurements – i.e., cord blood at birth or venipuncture at the research center, postnatal ultrasound, and the MRI visit – were performed at very different phases of the children's lives. In addition, the postnatal ultrasound was only administered in a subsample of the whole Generation R Study. As such, the number of children with complete data was

relatively small compared to the number of children that had genetic data and one of the two imaging measures (**Figure 1**). The sample size therefore differs per analysis, with 1204 children for the PGS-MRI associations, 576 children for the PGS-GTOD associations, and 340 children for the PGS-MRI associations mediated by the GTOD.

Sensitivity analyses

Previous literature suggests that subcortical development differs between men and women, with local differences in rate and variability independent from total brain size[133-136]. We reanalyzed the PGS and MRI subcortical volume associations by adding an interaction term for the different PGS and sex. Furthermore, we reanalyzed all associations using only the set of 340 children with complete data. Finally, to make the GTO-like volume more comparable to the GTOD in the mediation analyses, we modeled the GTO-like volume to as a spherical volume to derive a diameter. The estimated diameter was defined as $2\sqrt[3]{\frac{3v}{4\pi}}$, where v was the GTO-like volume.

To assure that the FreeSurfer subcortical segmentations were of good quality, each subcortical structure segmentation was visually inspected in 929 of the 1201 children [137]. Most regions had >98% children with sufficient quality, except for the caudate (94.4%), the putamen (88.3%) and the pallidum (90.9%). We therefore performed stringent sensitivity analyses where children were only included if their subcortical segmentation had been assessed, and if they passed quality assessment for all regions, i.e., 761 out of 929 children. In the 340 children with both ultrasound and MRI data, 272 had their segmentation quality rated, and of those 219 had sufficient segmentation quality to be included in the sensitivity analysis.

RESULTS

Characteristics of the study population

The population characteristics are displayed in **Table 1**. In the whole study population, the cranial ultrasound took place at a mean age of 6.8 (SD = 1.8) weeks, with a mean head circumference of 38.6 (SD = 1.5) cm and a mean GTOD of 4.3 (SD = 0.2) cm. Participants on average were 10.2 (SD = 0.6) years of age during the MRI scan, with a mean intracranial volume of 1540 (SD = 138) ml. PGS at different parameter P values for any given region were highly correlated (median Pearson's r = 0.91) (**Supplementary Figure 1**), also in those with only complete data (**Supplementary Figure 2**).

Participants with complete data did not differ from participants with data on only MRI characteristics. However, compared to participants with only ultrasound they were slightly

Table 1 | Characteristics of the study population. The values represent the means and standard deviations unless stated otherwise.

	Partial data set (n = 1430)	Complete data set (n = 340)
General characteristics		
Gender, boy (n, %)	730 (51.0%)	169 (49.7%)
Gestational age at birth (weeks)	40.1 (1.6)	40.2 (1.5)
Birth weight (g)[a]	3533 (524)	3530 (510)
Maternal age at birth (years)[a]	32.1 (4.0)	32.2 (3.9)
Maternal education at birth (n, %)[a]		
Low	118 (8.4%)	29 (8.6%)
Middle	321 (22.8%)	69 (20.5%)
High	966 (68.8%)	238 (70.8%)
Ultrasound (n)	**566**	**340**
Age (weeks)	6.8 (1.8)	6.5 (1.6)
Head circumference (cm)	38.6 (1.5)	38.5 (1.5)
GTO diameter (cm)	4.3 (0.2)	4.3 (0.2)
Magnetic resonance imaging (n)	**1201**	**340**
Age (years)	10.2 (0.6)	10.2 (0.6)
Intracranial volume (ml)	1540 (138)	1531 (138)
Accumbens nucleus volume (ml)	1.4 (0.2)	1.4 (0.2)
Amygdala volume (ml)	3.6 (0.4)	3.6 (0.4)
Brainstem volume (ml)	18.9 (1.8)	18.9 (1.7)
Caudate nucleus volume (ml)	8.3 (1.0)	8.2 (0.9)
Globus pallidus volume (ml)	3.9 (0.4)	3.9 (0.4)
Putamen volume (ml)	10.8 (1.1)	10.8 (1.1)
Thalamus volume (ml)	15.2 (1.3)	15.1 (1.3)

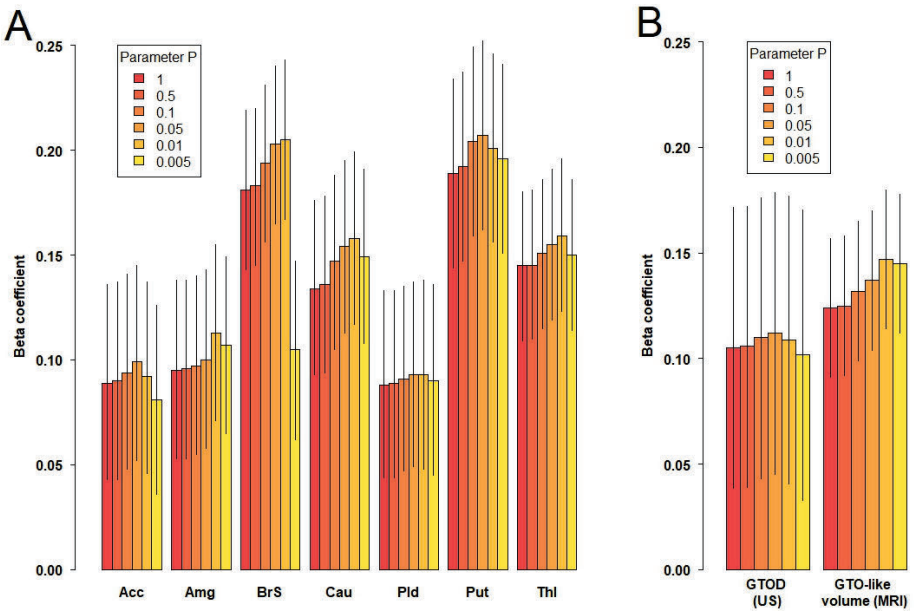
[a] = These variables are not included in the analyses but are used to show the differences between the partial and complete sets.

younger (mean difference = 0.3 weeks, $p < 0.001$) and had a smaller head circumference (mean difference = 0.1 cm, $p = 0.016$).

Subcortical volume PGS and childhood MRI

The volumes of all regions were correlated, with most Pearson's r values ranging between 0.4 and 0.6 (**Supplementary Figure 3**). The region-specific associations between the PGS and mean volumes are shown in **Figure 3**. All regional volumes were significantly associated with their PGS, with the strongest associations in the putamen ($\beta = 0.207$, 95% CI = [0.162; 0.252], $\Delta R^2 = 4.7\%$) and the brainstem ($\beta = 0.205$, 95% CI = [0.167; 0.243], $\Delta R^2 = 4.5\%$). Most importantly, the associations for all regions were generally constant across different values for parameter

Figure 3 | The beta coefficients for the association between subcortical PGS and the brain. The colored bars represent the different values for the *P* parameter for LDpred. (A) shows the results for PGS and their associated regions during early childhood. (B) shows the associations for the GTO PGS and the corresponding measures from ultrasound and MRI. Acc = accumbens; Amg = amygdala; BrS = brainstem; Cau = caudate nucleus; MRI = magnetic resonance imaging; Pld = pallidus; Put = putamen; Thl = thalamus; US = ultrasound.



P. The interaction term of PGS with sex was not significant in any of the associations. The correlations were similar in the children with only complete data (**Supplementary Figure 4**).

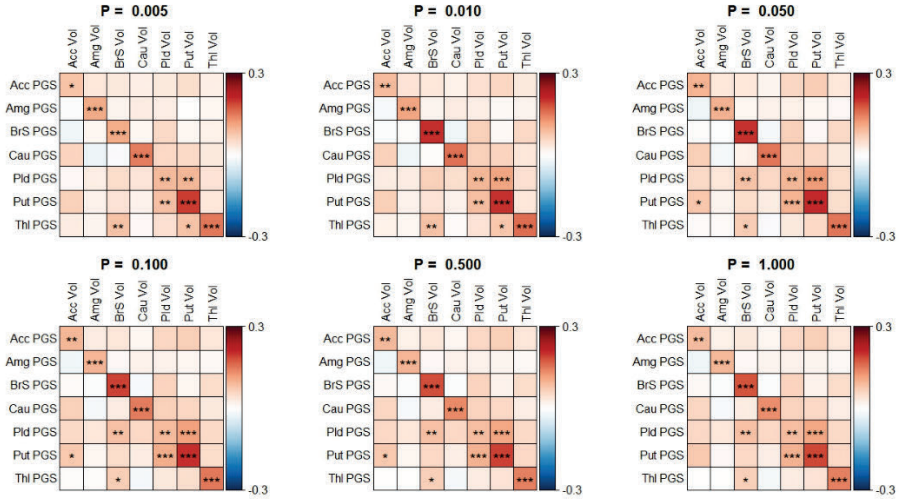
Region-nonspecific associations are shown in **Figure 4** and **Supplementary Table 1** for PGS based on all parameters P. After Bonferroni correction only 3 to 4 out of 42 region-nonspecific associations reached statistical significance across parameter *P* values, compared to 7 out of 7 for the region-specific associations. Two of the region-nonspecific associations that were statistically significant for all parameter *P* values were located within the globus pallidus and the putamen (both $p < .001$), which are neighboring regions. Sensitivity analyses showed that the pattern of associations generally remained similar after excluding participants with lower region-specific segmentation quality and those without segmentation quality ratings (**Supplementary Figure 5**). However, the region-specific association between the PGS and volume of the nucleus accumbens did not reach statistical significance anymore in all but the parameter P = 0.05 level. In children with only complete data only the region-specific

associations were statistically significant, and the accumbens and pallidus associations did not reach statistical significance at any level of parameter P (**Supplementary Figure 6**). Similar patterns emerged in children with only complete data after excluding those with poor image segmentation quality and those without segmentation quality ratings (**Supplementary Figure 7**).

GTO PGS as a predictor for infant US and childhood MRI

We constructed a GTO PGS by combining the scores for the caudate nucleus, the putamen, the globus pallidus and the thalamus. We further combined the volumes of these regions as obtained from the MR images at ten years to approximate a GTO-like volume. As **Figure 3B** and **Supplementary Table 2** show, the GTO PGS at all levels of parameter P seemed to associate with similar effect size to the GTOD and the GTO-like volume from childhood MRI. For example, at parameter P = 1.00 one standard deviation increase in the PGS led to a 0.105 (95% CI = [0.039; 0.171]) increase in standardized GTOD and 0.124 (95% CI = [0.091; 0.157]) increase in standardized GTO-like volume.

Figure 4 | Heatmaps for associations between the subcortical PGS and the MRI-based subcortical volumes. Each heatmap corresponds to a different value for parameter P. The rows represent the PGS. Columns represent the standardized volumes for the subcortical regions as obtained from the MRI segmentations. The associations were Bonferroni corrected. Acc = accumbens; Amg = amygdala; BrS = brainstem; Cau = caudate nucleus; Pld = pallidus; Put = putamen; Thl = thalamus; Vol = volume. * p-value < 0.05; ** p-value < 0.01; *** p-value < 0.001.



The GTOD from ultrasound also related to the GTO-like volume from childhood MRI, with each standard deviation increase in GTOD associating with a 0.27 standardized increase in GTO-like volume (95% CI = [0.17; 0.36], ΔR^2 = 3.9%).

Mediation analyses

To assess the relative contribution of the GTO PGS to GTOD and the GTO-like volume we constructed a causal mediation model. The results for the mediation analyses at all parameter P values are shown in **Table 2**. We found that the effect of the combined PGS on the estimated GTO volume at ten years of age was partly mediated by the GTOD at seven weeks of age at parameter P = 0.05 and higher. At those thresholds, the natural direct effect ranged between 82.4% and 83.5%, and the natural indirect effect between 16.5% and 17.6%. Substituting the GTO-like volume for its estimated diameter did not attenuate or change the results, with the proportion of the mediated effect increasing by less than 0.5% for all parameter P values (**Supplementary Table 3**).

Table 2 | Results for the mediation analyses. The percentages represent the proportions of the natural direct effect and the natural indirect, i.e., mediated, effect.

Parameter *P*	Direct effect (95% CI)	Indirect effect (95% CI)	p-value
0.005	87.4% (67.4 – 100.0)	12.6% (0.0 – 32.6)	.134
0.010	85.9% (68.4 – 100.0)	14.1% (0.0 – 31.6)	.058
0.050	83.5% (60.7 – 97.8)	16.5% (2.2 – 39.3)	.018
0.100	83.0% (61.4 – 97.7)	17.0% (2.3 – 38.6)	.028
0.500	82.5% (59.8 – 98.0)	17.5% (2.0 – 40.2)	.036
1.000	82.4% (58.7 – 97.5)	17.6% (2.5 – 41.3)	.028

DISCUSSION

The genetic loci related to subcortical volumes in an adult sample also associate with subcortical volume during infancy and early childhood. Furthermore, the associations during early childhood are partly independent from those during infancy, suggesting that the previously identified loci affect both prenatal and postnatal development.

The region-specific PGS primarily associated with the region-specific volumes from MRI and had limited associations with other regions. The previously identified genetic loci therefore generalize to earlier life phases, even infancy, and could aid in predictions based on subcortical volumes. The coefficients obtained from the analyses may be considered modest in size. In general, a standard deviation change in the PGS only led to about 0.1 standard

deviations increase in the associated subcortical volume. Interestingly, this effect was similar to the effect size of sex and the effect size of a year increase in age in the early childhood MRI. In addition, PGS tend to have effect sizes around this magnitude[138]. This suggests that the PGS did capture a relevant effect in prediction of subcortical volume in childhood.

The mediation analysis suggested that the genetic effects on subcortical volumes during infancy only explain a relatively small part of the volumes during early childhood. A theoretical consideration is that the analyses were adjusted for global scale, i.e., head circumference during infancy and intracranial volume during early childhood, which was also done in the original GWAS[11]. The analyses therefore focus on the relative size of subcortical regions. Combined, this implies that the genetic scores capture a change in the relative subcortical volumes that occurs during early childhood, that is at least partly independent of the volumes during infancy. Little is known, however, about growth trajectories of subcortical volume relative to intracranial volume during early life. A study of 48 postmortem fetal brains found that the deep nuclei – which roughly correspond to the subcortical regions of interest – did not increase in size relative to the supratentorial volume between gestational weeks 20 to 31[139]. By contrast, a study by the ENIGMA consortium showed that the subcortical volumes corrected for intracranial volume rapidly change particularly during early childhood[140]. Further work is needed to study whether and how trajectories of prenatal subcortical volumes relate to postnatal trajectories, and how the underlying genetic mechanisms differ.

Another explanation for the relatively weak mediating effect relates to individual differences in brain development. Subcortical volumes during early childhood may already relate very closely to the peak volumes as seen during adulthood, given that the brain reaches its maximum size at approximately 10 to 12 years of age[141,142]. However, individuals reach these end states of development at different rates[133,143]. Individuals who reach the same peak volume may do so at very different ages due to differences in developmental rates during in utero and early childhood. Studying regional brain volumes during infancy may therefore be confounded by differences in developmental trajectories that are perhaps driven by different genetic mechanisms, thus leading to an underestimation of the mediating effect. These differences in developmental trajectories may be amplified in subcortical regions like the thalamus, which develop throughout adolescence and reach their peak volume during adulthood[144]. Further work on infant brain development is needed to address these questions.

The study had several limitations. First, the original GWAS were based on MRI data whereas the imaging during infancy was done using cranial ultrasound. These methods have vastly different levels of precision, with the ultrasound potentially suffering from more measurement error compared to the MRI. In addition, the ultrasound data focused on the GTO, which does not have a proper equivalent structure in the MRI analyses and the GWAS. These differences in

measurement error and region of interest have likely led to an underestimation of the mediating effect of subcortical volumes during infancy. Second, the exclusion of those genetically classified as non-European leads to a severe reduction in sample size and further limits the generalizability of the findings. Third, the sample with complete data was only about a quarter of the sample with useable data. We tried to account for this limitation by also showing the analyses for both datasets, and we further showed that the characteristics of the samples and the effect sizes of each model are comparable. Fourth, any residual confounding in the mediation analysis may lead to overestimation of the mediating effect[131]. We corrected for educational attainment of the mother and birth weight of the child in the association between infant and early childhood GTO volume, but other confounders could still be unaccounted for. Fifth, the GWAS on subcortical volumes focused on SNPs, which cover only part of the genetic variation between individuals. If all genetic variation underlying subcortical volumes could be captures in a score, then the identified associations may have been different. However, given the low prevalence of other sources of genetic variation – like indels and copy-number variants – it would require massive sample sizes to identify those. Our study also benefitted from a number of strengths. As the sample size was reasonably large, the statistical power was high enough to detect relatively small effects. For example, for the associations between the PGS and early childhood MRI volumes, the a priori power to detect 2% explained variance on top of the base model was 98.1% for unadjusted analyses and 78.8% for the Bonferroni adjusted analyses. Finally, the Generation R study provided a unique sample with prospectively collected neuroimaging during infancy and childhood.

In conclusion, we established that genetic variants that were found to be related to subcortical volume during adulthood also relate to subcortical volume during infancy and early childhood. We further showed that these effects are partly independent, suggesting that the genetic loci may modulate specific stages of postnatal development.

Data availability statement

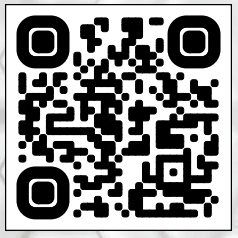
The data used in this study will not be made publicly available due to legal and informed consent restrictions. Reasonable requests for access to the data can be directed to the Director of the Generation R Study, Vincent Jaddoe (generationr@erasmusmc.nl), in accordance with the local, national and European Union regulations.

3.2

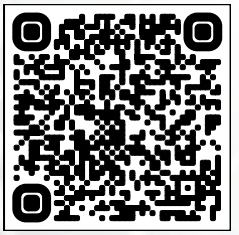
Genetic burden for late-life neurodegenerative disease and its association with early-life lipids, brain, behavior, and cognition

Sander Lamballais, Ryan L. Muetzel, M. Arfan Ikram, Henning Tiemeier, Meike W. Vernooij, Tonya White, Hieab H. H. Adams

Adapted from:



Supplement:



ABSTRACT

Background

Genetics play a significant role in the etiology of late-life neurodegenerative diseases like Alzheimer's disease, Parkinson's disease, and frontotemporal dementia. Part of the individual differences in risk for these diseases can be traced back decades before the onset of disease symptoms. Previous studies have shown evidence for plausible links of apolipoprotein E (APOE), the most important genetic marker for Alzheimer's disease, with early-life cognition and neuroimaging markers. We aimed to assess whether genome-wide genetic burden for the aforementioned neurodegenerative diseases plays a role in early-life processes.

Methods

We studied children from the Generation R Study, a prospective birth cohort. APOE genotypes and polygenic genetic burdens for Alzheimer's disease, Parkinson's disease, and frontotemporal dementia were obtained through genome-wide genotyping. Non-verbal intelligence was assessed through cognitive tests at the research center around the age of 6 years, and educational attainment through a national school performance test around the age of 11 years. The Child Behavior Checklist was administered around the age of 10 years, and data from the anxious/depressed, withdrawn/depressed, and the internalizing behavior problems scales were used. Children participated in a neuroimaging study when they were 10 years old, in which structural brain metrics were obtained. Lipid serum profiles, which may be influenced by APOE genotype, were assessed from venal blood obtained around the age of 6 years. The sample size per analysis varied between 1,641 and 3,650 children due to completeness of data.

Result

We did not find evidence that APOE genotype or the polygenic scores impact on childhood nonverbal intelligence, educational attainment, internalizing behavior, and global brain structural measures including total brain volume and whole brain fractional anisotropy (all $p > 0.05$). Carriership of the APOE ε2 allele was associated with lower and APOE ε4 with higher low-density lipoprotein cholesterol concentrations when compared to APOE ε3/ε3 carriers.

Conclusion

We found no evidence that genetic burden for late-life neurodegenerative diseases associates with early-life cognition, internalizing behavior, or global brain structure.

INTRODUCTION

Genetic factors play a significant role in the etiology of late-life neurodegenerative diseases like Alzheimer's disease (AD)[16,146], Parkinson's disease (PD)[147] and frontotemporal dementia (FTD)[148]. With the exception of rare Mendelian forms of diseases, cases arise due to multifactorial processes where many genetic variants confer risk of neurodegeneration, in combination with non-genetic factors. The clinical onset of the aforementioned diseases tends to be preceded by years of deterioration of cognition and brain structure[149-151] as well as an increased incidence of depressive and psychiatric symptoms[152-154]. For AD, these differences may even extend decades before the onset of disease[37,155,156], which could partly be explained by individual differences in the genetic burden for AD. As the genome is stable throughout life, the genes implicated in late-life neurodegenerative disease may already lead to subtle differences during childhood.

The apolipoprotein E epsilon 4 allele (APOE ε4) is the strongest common genetic variant for AD[157-159]. The APOE gene plays a role in lipoprotein metabolism, and has been shown to affect lipid serum profiles during adulthood[160-164], and potentially during childhood[165,166]. As APOE also increases the risk for AD, its role in early-life cognition and brain markers has also been studied. The studies on APOE ε4 and cognition during adolescence and early adulthood have reported mixed results, with some reporting lower cognitive function, some higher, and most reporting no difference[167]. Additionally, a number of studies showed that APOE ε4 may relate to lower brain volumes during infancy and childhood, particularly in regions affected in AD such as the hippocampus[168-173]. Overall, APOE ε4 may associate with early-life processes, but this needs to be elucidated further.

With the advent of genome-wide association studies (GWAS) there has been an increase in the number of genes identified for neurodegenerative disease. GWAS has led to the discovery of at least 30 new genetic loci for AD[16,146], at least genetic 24 loci for PD[147] and at least three loci for FTD[148]. The disease burden per locus can be combined into a single score, known as polygenic risk scores (PGRS), to assess the genetic burden a person has for that disease[174]. The genetic burden for AD, PD and FTD may relate to early-life processes, which can be studied using PGRS. However, few studies exist that assesses the effect of such PGRS on early-life markers.

To obtain a more comprehensive overview of the relevance in early-life of genes related to late-life neurodegenerative disease we performed a comprehensive study within the Generation R birth cohort. We assessed the APOE genotype and created PGRS for AD, PD and FTD. Given the existing literature we hypothesized that these genetic predispositions to late-

life neurodegenerative disorders associate with early-life non-verbal intelligence quotient (IQ), educational attainment, internalizing behavior, and neuroimaging markers, and that APOE and the AD PGRS associate with lipid profiles.

METHODS

Participants

The data was obtained from the Generation R cohort, a prospective birth cohort based in Rotterdam, The Netherlands[22]. Pregnant women in Rotterdam were at their first prenatal visit approached to participate. A total of 9,901 children were born as part of the Generation R cohort and were invited to participate in questionnaires and research center visits beginning in 2002 to the present day.

DNA was sequenced from blood obtained from the umbilical cord or with blood samples collected around 6 years of age, and genetic data was available for 5725 children. In the case of sibling pairs (n = 235 pairs) we included the oldest sibling. This led to a sample of 5490 children. The focus of the current study was on cognitive function, brain structure and blood lipid profiles. Non-verbal IQ was measured at approximately 6 years (n = 3650) and educational attainment at 11 years of age (n = 1641). The Childhood Behavior Checklist (CBCL) was administered around the age of 10 years with data for the anxious/depressed scale (n = 1867), the withdrawn/depressed scale (n = 1862) and the internalizing problems scale (n = 1859) used for this study. Magnetic resonance imaging (MRI) of the brain was done when the children were approximately 10 years of age, collecting both T_1-weighted (n = 1962) and diffusion-weighted images (n = 1832). Blood lipid profiles were determined with blood samples obtained around the age of 6 years (n = 2749). A flow chart of the study population is shown in **Supplementary Figure 1**.

Ethics Statement

The study was conducted in accordance with the guidelines as proposed in the World Medical Association Declaration of Helsinki and was approved by the Medical Ethics Committee of the Erasmus MC (registration number MEC 02.1015). Written informed consent was obtained from primary caregivers on behalf of the child.

Genotyping, APOE ε4 and PGRS

DNA sample collection, genotype calling procedures and subsequent quality control have been described elsewhere[123,124]. In brief, samples were either collected from cord blood at birth (Illumina 610K Quad Chip) or from venipuncture at a visit to the research center when children were between the age of 5 till 8 years (Illumina 660K Quad Chip). Single nucleotide

polymorphisms were filtered for minor allele frequency < 0.01, Hardy-Weinberg disequilibrium p < .00001 and missing rate > 0.05. To be able to account for population stratification, we calculated the first ten genomic components using the multi-dimensional scaling function of PLINK[124,125].

APOE carriership status was assessed from the genotyped data and based on the nucleotide combinations of two single nucleotide polymorphisms: rs429348 and rs7412. A thymine at both locations is classified as APOE ε2, one thymine and one cytosine as APOE ε1 or APOE ε3, and both cytosines as APOE ε4. As APOE ε1 and APOE ε3 cannot be distinguished we classified both as APOE ε3. We considered APOE ε3/ε3 to be the reference category as this is the most prevalent genotype.

PGRS for AD, PD and FTD were calculated using PRSice-2[175] . The scores were based on summary statistics from the largest GWA studies for each respective neurodegenerative disease[147,148,176]. PGRS are generally calculated for different thresholds of statistical significance in the summary statistics. As we did not have an a priori hypothesis on the optimal threshold, we calculated PGRS based on single nucleotide polymorphisms below the following p-value thresholds: 0.000001, 0.000005, 0.00001, 0.000005, 0.00001, 0.00005, 0.0001, 0.0005, 0.01, 0.05, 0.1, 0.5 and 1.0. Strand flips were corrected, and we used clumping to build the score using independent loci.

Non-verbal IQ and educational attainment

Two measures for cognitive function were available. The first was an assessment of non-verbal IQ at the research visit around the age of 6 years. Participants completed two subtests of the Snijders-Oomen Non-verbal Intelligence Test-Revised (SON-R 2½-7)[177]: "Mosaics", a spatial visualization task, and "Categories", an abstract reasoning task. The raw scores were converted to IQ scores using age and sex-specific norms. As both tasks specifically assess non-verbal cognition, we considered these scores as non-verbal IQ scores. The correlation between IQ derived from the whole test battery and IQ derived from just the “Mosaics” and “Categories” tests has been shown to be high (r = 0.86)[178].

The measure of cognitive function was the educational attainment score obtained at the age of 11 years. The 'Centraal Instituut voor Toetsontwikkeling' (CITO) test is administered in the majority primary schools in The Netherlands and is completed during the final year of primary school. The CITO test generally consists of two main skill domains: language and mathematics. The raw test scores for both domains were obtained for most Generation R children that took the CITO test during the years 2014 to 2017 and that were still part of Generation R at the time. As the test difficulty tends to vary slightly each year we summed the raw domain scores to a total score for each child, standardized the scores for all children within a given year and finally

combined the stratified distributions into one distribution. This method yielded standardized scores that were comparable across testing years.

Child Behavior Checklist

Behavioral problems were assessed using the CBCL for ages 6 to 18[179]. The CBCL is a validated and reliable 113-item inventory that uses caregiver-reported information to assess behavioral problems in children. The procedure and specific characteristics for Generation R have been described elsewhere[55]. For this study we considered mother-reported data on the anxious/depressed, the withdrawn/depressed and the internalizing problems scales.

Image acquisition and processing

Image acquisition has been described elsewhere[55]. In brief, structural brain MR images were obtained on a single 3T GE Discovery MR750w MRI System (General Electric, Milwaukee, WI, USA) utilizing an 8-channel receive-only head coil. T_1-weighted images were collected using a 3D inversion recovery-prepared fast spoiled gradient recalled sequence (T_R = 8.77 ms, T_E = 3.4 ms, T_I = 600 ms, flip angle = 10°, Field of view = 220 x 220 mm, Acquisition matrix = 220 x 220, slice thickness = 1 mm, number of slices = 230, bandwidth = 25 kHz). Diffusion-weighted images consisted of 3 b_0 volumes and 35 diffusion directions using an echo planar imaging sequence (T_R = 12,500 ms, T_E = 72 ms, Field of view = 240 x 240 mm, Acquisition matrix = 120 x 120, slice thickness = 2 mm, number of slices = 65, b = 900 s/mm^2).

T_1-weighted images were processed through the FreeSurfer analysis suite, version 6.0.0[88]. The procedure has been described elsewhere[180]. Briefly, non-brain tissue was removed, voxel intensities were normalized for B1 inhomogeneity, whole-brain tissue segmentation was performed, and a surface-based model of the cortex was reconstructed. For each participant we obtained metrics for total brain volume, cortical gray matter volume, cerebrospinal fluid volume and mean cortical thickness. For analyses of APOE status and the AD PGRS we additionally focused on volumes of the hippocampus, the entorhinal cortex, the middle temporal gyrus and the parahippocampal gyrus. For the PD PGRS we also considered volumes of the nucleus accumbens, the caudate nucleus, the globus pallidi and the putamen. Finally, for the FTD PGRS we also looked at the frontal and the temporal lobes, and in particular the volume, the mean thickness, and the surface area. For all lateralized structures we took the mean of both sides.

DTI images were processed through the FMRIB Software Library (FSL), version 5.0.9[181]. The full procedure is described elsewhere[180]. Briefly, non-brain tissue was removed and images were corrected for eddy-current artifacts and translations/rotations resulting from head motion. Diffusion tensors were fitted at each voxel using the RESTORE method from the Camino diffusion MRI toolkit[182]. We further performed probabilistic white matter fiber tractography in

native space for each participant using the FSL plugin AutoPtx to identify connectivity distributions of a number of well-known fiber bundles[183]. Average fractional anisotropy and mean diffusivity values were then computed for each white matter tract. Global measures for fractional anisotropy and mean diffusivity were obtained by performing factor analyses on the tract-specific values[184].

Lipid profiles

Lipid profiles of the children were assessed from venous blood acquired during the research visits around the age of 6 years after a thirty minute fast. Serum total cholesterol, high-density lipoprotein cholesterol (HDL-c) and triglyceride concentrations were derived with the Roche Cobas 8000 analyzer (Roche Diagnostics GmbH, Penzburg, Germany), and low-density lipoprotein cholesterol (LDL-c) was estimated using the Friedewald equation[185]. We considered these lipids in relation to APOE status and the AD PGRS as the APOE gene plays a significant role in lipid metabolism[186], whereas we did not have such a prior expectation for PD and FTD.

Statistical Analysis

Statistical analyses were performed with the R statistical package, version 3.5.2[95]. We used multiple linear regression for all outcomes, correcting for age at outcome measurement, sex of the child, maternal education (low, intermediate, or high) and the first ten genomic components. The latter was done to take into account the underlying genetic structure of the population. The serum lipid models were additionally adjusted for body mass index (BMI) at the time of the venous puncture. The volumetric neuroimaging models, i.e., cortical volume, CSF volume and the disease-specific regional brain volumes, were additionally adjusted for total brain volume. Furthermore, we applied square-root transformations to the CBCL scales to better satisfy the linearity assumption of linear regression.

Polygenic burden may only affect those whose burden is above a certain threshold, thus leading to non-linearity of an association. We assessed this through two approaches: [1] dichotomization of the top PGRS decile versus the rest of the population, [2] fitting restrictive cubic splines on the PGRSs to assess any non-linearity in the association.

Use of PGRS in the Generation R Study requires a critical consideration of ethnicity. The GWAS from which we used the summary statistics were based on populations of European ancestry. Findings from GWAS and by extension PGRS are specific to the ethnicity of the original study population. The Generation R Study is based in the city of Rotterdam, where about half of all individuals are of non-European ancestry. We focused our main findings on the complete population, but we additionally stratified our analyses for European ancestry to check for any effects related specifically to ethnicity. We additionally performed sensitivity analyses where

we did not correct for the first ten genomic components, to see whether improper correction for population stratification is relevant for studies on APOE and studies on AD, PD, and FTD PGRS[187].

Table 1 | Characteristics of the study population.

Characteristics	All (N = 5490)	European ancestry (N = 2651)	Non-European ancestry (N = 2839)
APOE genotype (%)			
ε2/ε2	0.5	0.5	0.6
ε2/ε3	11.5	10.4	12.6
ε2/ε4	2.3	2.5	2.1
ε3/ε3	64.9	60.6	69.0
ε3/ε4	19.1	21.6	16.8
ε4/ε4	1.6	2.2	1.1
Visit around 6 years			
non-verbal IQ (mean, SD)	101 (15)	105 (14)	97 (15)
Total cholesterol (mean, SD) (mmol/L)	4.2 (0.6)	4.2 (0.6)	4.3 (0.7)
HDL-c (mean, SD) (mmol/L)	1.3 (0.3)	1.3 (0.3)	1.4 (0.3)
LDL-c (mean, SD) (mmol/L)	2.4 (0.6)	2.3 (0.6)	2.4 (0.6)
Triglycerides[a] (geometric mean, SD) (mmol/L)	1.0 (0.5)	1.1 (0.5)	1.5 (0.5)
Visits at 10 and 11 years			
CITO score, standardized (mean, SD)	0.0 (1.0)	0.2 (0.9)	-0.3 (1.1)
Total brain volume (mean, SD) (cm^3)	1200 (118)	1225 (113)	1168 (116)
Cortical volume (mean, SD) (cm^3)	574 (59)	588 (56)	556 (59)
Cerebrospinal fluid volume (mean, SD) (cm^3)	0.9 (0.2)	0.9 (0.2)	0.9 (0.2)
Mean cortical thickness (mean, SD) (mm)	2.67 (0.08)	2.68 (0.08)	2.67 (0.08)
Global FA, standardized (mean, SD)	0.00 (1.00)	0.11 (0.97)	-0.15 (1.02)
Global MD, standardized (mean, SD)	0.00 (1.00)	-0.03 (0.98)	0.04 (1.03)
Anxious/depressed scale (mean, SD)	2.2 (2.7)	2.2 (2.7)	2.2 (2.6)
Withdrawn/depressed scale (mean, SD)	1.1 (1.6)	1.1 (1.5)	1.1 (1.8)
Internalizing problems scale (mean, SD)	4.7 (5.0)	4.5 (4.8)	5.1 (5.4)

APOE = apolipoprotein epsilon; FA = fractional anisotropy; HDL-c = high density lipoprotein cholesterol; IQ = intelligence quotient; LDL-c = low density lipoprotein cholesterol; MD = mean diffusivity.

[a] Triglyceride serum values were log-transformed.

Multiple testing correction was considered on three levels: [1] the PGRS for AD, PD or FTD, [2] the PGRS thresholds, and [3] the outcome measures. We did not expect dependence amongst the PGRS of the neurodegenerative diseases. Therefore, we applied a Bonferroni correction across AD, PD and FTD. As the PGRS thresholds were strongly intercorrelated as well as some of the outcome measures, we applied a False Discovery Rate (FDR) correction within a given disease. The p-values reported below are those after the FDR correction.

Data availability statement

The datasets for this manuscript are not automatically publicly available due to legal and informed consent restrictions. Reasonable requests to access the datasets should be directed to the Director of the Generation R Study, Vincent Jaddoe (v.jaddoe@erasmusmc.nl), in accordance with the local, national and European Union regulations.

RESULTS

Population characteristics

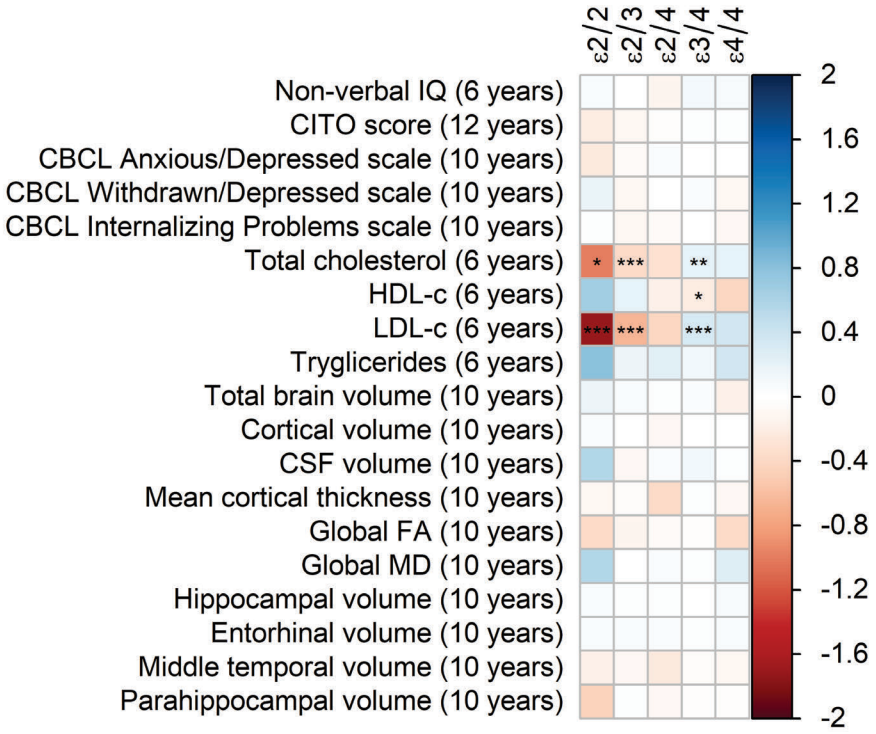
Table 1 shows the characteristics of the total study population and stratified by European ancestry and non-European ancestry. Overall, the most common APOE genotypes were ε3/ε3 (64.9%), ε2/ε3 (11.5%) and ε3/ε4 (19.1%), whereas the other genotypes were much less common, i.e., ε2/ε2 (0.5%), ε2/ε4 (2.3%) and ε4/ε4 (1.6%). These numbers were similar for those with European and non-European ancestry.

APOE and PGRS for AD

Figure 1-**2** display the results of the associations of all relevant outcomes with APOE genotype and AD PGRS, respectively. Neither APOE genotype nor any AD PGRS associated with non-verbal IQ during the six-year visit or the CITO score at 11 years (all $p_{corrected} > 0.05$). The APOE genotype and AD PGRS also did not relate to global brain metrics such as total brain volume and CSF volume, nor with the connectivity metrics global fractional anisotropy and mean diffusivity (all $p_{corrected} > 0.05$). APOE genotype and the AD PGRS also did not associate with region-specific metrics for the hippocampus, the entorhinal cortex, the medial temporal gyrus and the parahippocampal region. Finally, APOE genotype and AD PGRS did not show any statistically significant associations with the CBCL scales anxious/depressed or withdrawn/depressed, nor with the internalizing problems scale (all $p_{corrected} > 0.05$).

The APOE genotype associated with serum lipid profiles. Compared to the APOE ε3/ε3 genotype, those with APOE ε2/ε3 had lower total cholesterol concentrations ($\beta = -0.32$, SE = 0.06, $p_{corrected} < 0.001$), lower LDL-c concentrations ($\beta = -0.57$, SE = 0.06, $p_{corrected} < 0.001$) and

Figure 1 | Heatmap showing the regression coefficients between APOE genotype and all phenotypes. The E3/3 genotype is used as a reference for the other genotypes. All coefficients are standardized. The reported p-values were corrected for multiple testing. IQ = Intelligence Quotient; CBCL = Child Behavior Checklist; HDL = high-density lipoprotein; LDL = low-density lipoprotein; CSF = cerebrospinal fluid; FA = fractional anisotropy; MD = mean diffusivity; * = 0.05; ** = 0.01; *** = 0.001.

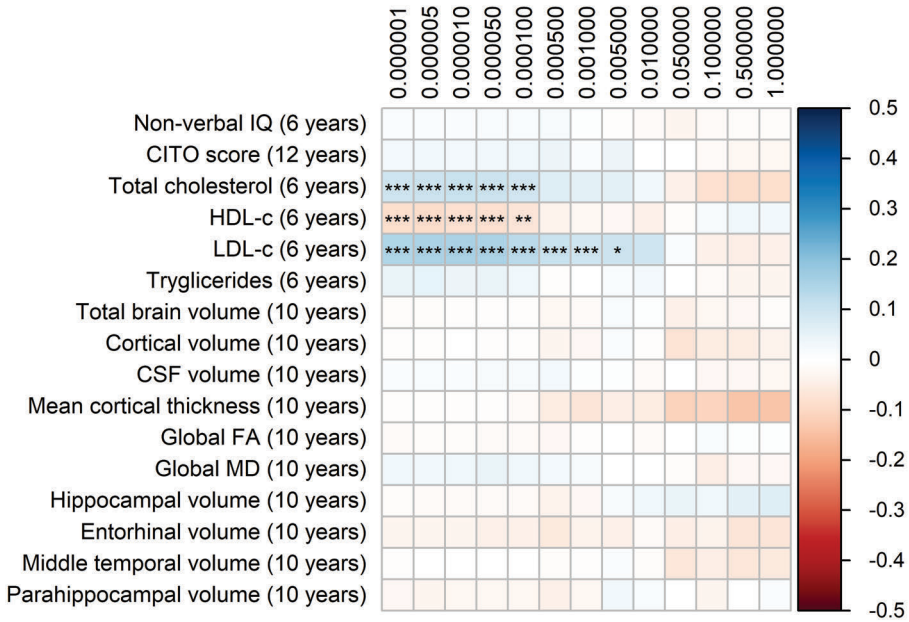


higher HDL-c concentrations (β = 0.22, SE = 0.06, $p_{corrected}$ = 0.01). The APOE ε2/ε2 followed the exact same pattern but with even larger differences.

Compared to the APOE ε3/ε3 genotype those with ε3/ε4 had higher total cholesterol concentrations (β = 0.19, SE = 0.05, $p_{corrected}$ = 0.003), higher LDL-c concentrations (β = 0.16, SE = 0.05, $p_{corrected}$ < 0.001) and lower HDL-c concentrations (β = 0.26, SE = 0.05, $p_{corrected}$ = 0.02). These differences were similar and larger when comparing the APOE ε3/ε3 genotype with the APOE ε4/ε4 genotype.

Triglycerides were higher in all genotypes compared to APOE ε3/ε3, although none of these were statistically significance (all $p_{corrected}$ > 0.05).

Figure 2 | Heatmap showing the regression coefficients between the AD PGRS and all phenotypes. Each score is based on a different threshold for inclusion of SNPs into the score. All coefficients are standardized. The reported p-values were corrected for multiple testing. IQ = Intelligence Quotient; CBCL = Child Behavior Checklist; HDL = high-density lipoprotein; LDL = low-density lipoprotein; CSF = cerebrospinal fluid; FA = fractional anisotropy; MD = mean diffusivity; * = 0.05; ** = 0.01; *** = 0.001.

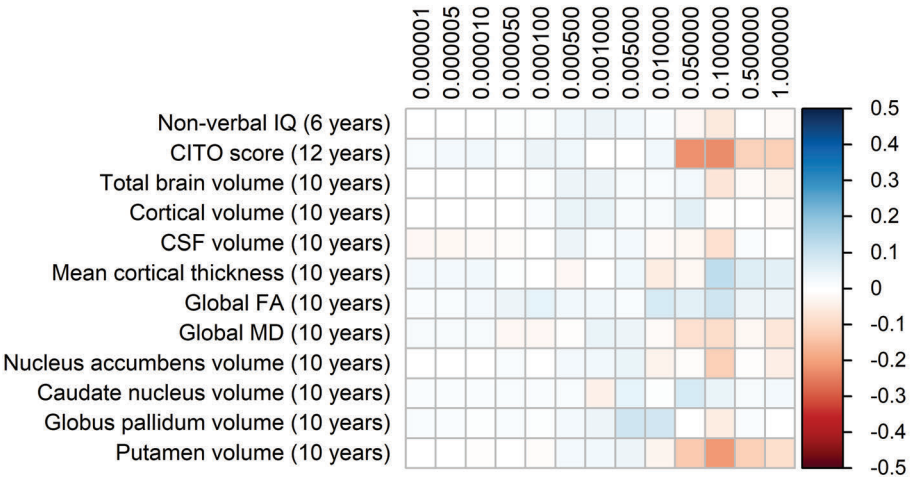


The AD PGRS also associated with serum lipid profiles, but only at stricter PGRS thresholds, i.e., PGRS thresholds below 0.001. However, these associations disappeared upon including the APOE genotype as a covariate (all $p_{corrected} > 0.05$). Furthermore, the results did not differ when using the top-decile PGRS dichotomization rather than the continuous PGRS, or when modeling cubic splines.

PGRS for PD and FTD

The results for the PD and FTD PGRS are shown in **Figures 3-4**, respectively. We found no support for associations of scores at any threshold with non-verbal IQ, educational attainment, internalizing behavior scales or neuroimaging markers. Furthermore, we did not find evidence for associations of the PD scores with the volumes of the nucleus accumbens (ß

Figure 3 | Heatmap showing the regression coefficients between the PD PGRS and all phenotypes. Each score is based on a different threshold for inclusion of SNPs into the score. All coefficients are standardized. The reported p-values were corrected for multiple testing. IQ = Intelligence Quotient; CBCL = Child Behavior Checklist; HDL = high-density lipoprotein; LDL = low-density lipoprotein; CSF = cerebrospinal fluid; FA = fractional anisotropy; MD = mean diffusivity; * = 0.05; ** = 0.01; *** = 0.001.

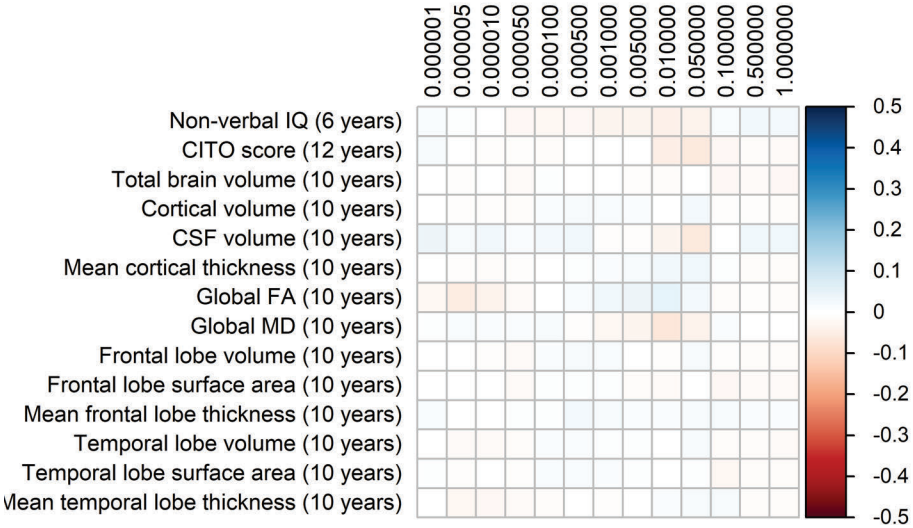


for PGRS at 0.05 threshold = -0.08, $SE_{0.05}$ = 0.12, $p_{corrected}$ = 1.00), the caudate nucleus ($ß_{0.05}$ = 0.03, $SE_{0.05}$ = 0.12, $p_{corrected}$ = 1.00), the globus pallidus ($ß_{0.05}$ = -0.07, $SE_{0.05}$ = 0.13, $p_{corrected}$ = 1.00) or the putamen ($ß_{0.05}$ = -0.13, $SE_{0.05}$ = 0.12, $p_{corrected}$ = 1.00). Similarly, we did not observe any associations of the FTD scores with the volumes of the frontal ($ß_{0.05}$ = -0.00, $SE_{0.05}$ = 0.03, $p_{corrected}$ = 1.00) or temporal lobes ($ß_{0.05}$ = 0.01, $SE_{0.05}$ = 0.03, $p_{corrected}$ = 1.00).

Population structure

All analyses were performed in all available participants and were controlled for the first 10 genomic components. We further stratified the analyses for European versus non-European ancestry (**Figures 5A & 5D**, **Supplementary Figure 2**), and the effect estimates were generally similar. We additionally reran the analyses without correcting for the genomic components, and this led to stark changes in the results (**Figures 5B & 5E**, **Supplementary Figure 3**). The higher the PRS threshold, the more statistically significant findings were present in the analyses not corrected for genomic components compared to when we did correct for genomic components. We further split the uncorrected analyses for European versus non-European ancestry, to see whether one of these groups was driving the sudden change in findings (**Figures 5C & 5F**, **Supplementary Figure 4**). Within the uncorrected analyses for

Figure 4 | Heatmap showing the regression coefficients between the FTD PGRS and all phenotypes. Each score is based on a different threshold for inclusion of SNPs into the score. All coefficients are standardized. The reported p-values were corrected for multiple testing. IQ = Intelligence Quotient; CBCL = Child Behavior Checklist; HDL = high-density lipoprotein; LDL = low-density lipoprotein; CSF = cerebrospinal fluid; FA = fractional anisotropy; MD = mean diffusivity; * = 0.05; ** = 0.01; *** = 0.001.

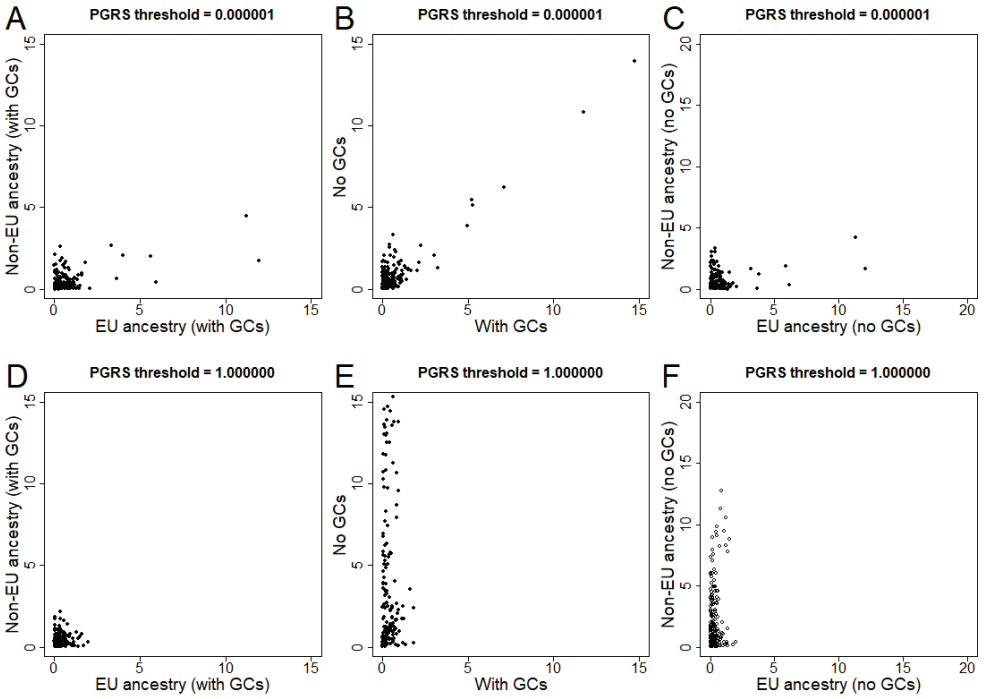


Individuals of non-European ancestry we find an inflation of the number of statistically significant findings, whereas this was not the case for individuals of European ancestry.

DISCUSSION

None of the measures for genetic burden for AD, PD or FTD were associated with childhood non-verbal IQ, educational attainment, internalizing behavior, global brain structure, or disease-specific regional brain structures. Although genetic burden for late-life neurodegenerative disease has been linked to brain structure and cognitive function during late-life, we find no evidence that these affect the same processes during early life. Furthermore, we provided clear evidence that the APOE genotype affects lipid profiles during childhood. Finally, we showed that improper control of the ethnic structure of the population through genomic components can lead to false positive associations when considering PGRS for AD, PD and FTD.

Figure 5 | Scatterplots of log-transformed p-values when comparing different analyses. The results are shown for PGRS thresholds 0.000001 (top row) and 1.000000 (bottom row). Comparisons are shown for European versus non-European (A, D), correcting for genomic components or not (B, E), and European versus non-European when not correcting for genomic components (C, F). EU = European; GC = Genomic component.



We found no support for the link between AD genetic burden and global, hippocampal and temporal regions although previous studies have provided evidence for such links during infancy[172,173], childhood[168,171,188,189] and early adulthood[169,190-197]. The support for such associations does seem stronger in studies during early adulthood than during childhood. This suggests that the genetic burden for AD becomes more relevant with age, and that there are cumulative processes at play which may only become apparent after early-life. For example, the pathological burden in APOE ε4 carriers may be increased due to accumulation of lipoprotein[198,199], reduced neuronal reparative capacity[200,201] and altered responses to neuroinflammatory processes[202].

Interpreting the role of APOE in childhood brain development is further complicated by the inconsistency of findings. Shaw and colleagues found in children and adolescents that APOE ε4 carriers had thinner entorhinal cortices than non-carriers[168]. Chang and colleagues also studied children and adolescents, but they found that ε4 carriers had larger hippocampi than

non-carriers[171]. Additionally, they report that ε4 carriers compared to non-carriers had larger volumes for the cuneus, the temporal pole, the lateral occipital pole, and the medial orbitofrontal cortex. The two studies that report that APOE affects brain structure in infants show that ε4 carriers have smaller hippocampi than non-carriers[172,173]. Furthermore, they report that APOE affects regions that are very different from those reported by Chang and colleagues[171]. More recently, Axelrud and colleagues reported that the AD PGRS relates to hippocampal volume in Brazilian children aged 6 to 14 years old[188]. However, the source GWAS on which the PGRS in the latter study was based was performed in a population of European ancestry[176]. As we have shown, using the AD PGRS in populations of non-European ancestry leads to false positive findings. Indeed, Axelrud and colleagues could not replicate their findings in a separate Canadian population of 1,024 adolescents. In summary, previous findings have been inconsistent, which suggests that AD genetic burden may only affect early-life brain structure under specific circumstances or that the effect is unlikely to be clinically relevant.

We did not find evidence in our study to suggest that AD genetic burden affects cognitive functioning during childhood. Previous studies on this topic report mixed results, but several larger studies also did not find evidence for such a link. Taylor and colleagues studied cognition in the ALSPAC study[165]. The only pattern observed was that APOE ε4 carriers performed better on cognitive tests than those with a APOE ε3/ε3 genotype, although not statistically significant. In our study we found no evidence to support this. More recently, Weissberger and colleagues meta-analyzed data from 9234 individuals aged 2 to 40 years old and found no association of APOE ε4 carriership with intelligence, attention, executive function, language, memory, processing speed and visuospatial abilities[203]. In our study we confirmed this finding at two timepoints in childhood (around 6 years of age and around 11 years of age), and we also extend the findings to APOE ε2 genotypes and to broader AD genetic burden. Taken together, the literature and this study suggest that AD genetic burden does not affect cognition during early life.

The role of the APOE gene in serum lipid profiles during early life has been studied before. In 2011, a study by Taylor and colleagues assessed the relation between APOE status and serum lipid profiles in 2875 children aged 8 to 11 years from the ALSPAC cohort, a prospective birth cohort study[165]. They showed that carriership of APOE ε2 was associated with reduced cholesterol and increased triglyceride levels compared to APOE ε3/ε3, whereas APOE ε4 carriers had both elevated cholesterol and triglyceride levels. In 1997, Kallio and colleagues showed that cord blood from 42 APOE ε4 carriers contained higher concentrations of cholesterol than 13 carriers of APOE ε2[166]. In addition, LDL levels rose steeper during the first year of life in the APOE ε4 carriers than in the APOE ε2 carriers. In our study we had similar

findings for ε2/ε3 and ε3/ε4 but not for ε2/ε2 and ε4/ε4 genotypes, likely because those genotypes were uncommon within the current study population. Our findings further consolidate the causal role of APOE genotype in serum lipid levels even during early life.

We found no evidence that PD and FTD genetic burden influences early-life processes. However, the etiology and pathogenesis of PD and FTD are poorly understood, and less is known on the preclinical disease stage compared to AD. It is therefore not clear how genetic burden for PD or FTD would influence early-life processes. As both syndromes can occur through dominant autosomal inheritance, it should be possible to investigate families of PD or FTD patients to identify such processes. However, we were unable to identify any such study in the literature. Another route would be to look at healthy carriers of known genetic risk variants for either PD or FTD to identify affected processes. For example, the G2019S mutation in the LRRK2 gene, the gene most widely associated with Parkinson's disease, has been studied in healthy controls. Different studies found this gene to be associated with lower executive functioning[204], changes in gait[205], olfactory dysfunction[206]. However, all these studies were small and exploratory. To the best of our knowledge, no studies focusing on FTD candidate genes in healthy controls are available. Further work is needed to elucidate whether PD and FTD genetic burden play a role in other domains during early-life, for example brain function rather than brain structure.

The etiology of AD, PD and FTD extend beyond lipid profiles, the brain, behavior, and cognition, thus raising the question which other processes could be relevant during childhood. For example, cerebrospinal fluid markers levels such as Tau and phosphorylated Tau are affected by APOE ε4 carriership in demented individuals[207-210]. In addition, ApoE protein levels in cerebrospinal fluid, but not blood serum, depend on the APOE genotype[211]. Another avenue for further research is inflammatory markers such as C-reactive protein, interleukin-6 and α1-antichymotrypsin, which have shown predictive value for the onset of all-cause dementia[212]. Further assessment of endophenotypes closely related to specific gene function may provide more stable findings related to early life.

Our findings may have been limited by several aspects of the study design. First, we relied on cross-sectional data. Brain growth follows non-linear trajectories, reaching a peak at around the age when the children in our study underwent neuroimaging[142]. The genetic burdens for neurodegenerative disease may affect the trajectories of brain development, which would only be detectable through longitudinal studies. Alternatively, the genetic burden for late-life neurodegenerative disease may not express until later in childhood or adolescence, and the study population may simply be too young for the research questions at hand. Second, the number of individuals with ε2/ε2 or ε4/ε4 genotypes was relatively low, thus we were likely underpowered to establish any small effects for those genotypes. Third, we administered a

limited number of cognitive tests around the age of 6, limiting our investigation to non-verbal IQ. AD is generally characterized by a loss of memory function, for which we did not have an adequate test in children.

Our study also had clear strengths. The size of our study population ensured sufficient power to detect relatively small effects related to the common APOE genotypes and the AD, PD and FTD PGRS. Furthermore, we provide an unambiguous case for proper control of population stratification, which was only possible due to the large proportion of participants of non-European ancestry. Finally, the Generation R Study is a representative sample from the general population, which vastly improves the generalizability to a community-dwelling population of European descent.

In conclusion, we found no evidence to support the role of genetic burden for late-life neurodegenerative disease in early-life cognitive performance, internalizing behavior, and brain metrics. APOE genotype was related to blood lipid profiles. Genetic burden for AD, PD and FTD did not relate to cognition or brain structure. These findings suggest that the etiology of late-life neurodegenerative disease becomes only relevant later in life.

VASCULAR MECHANISMS

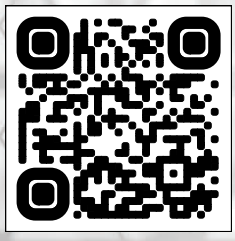
4

4.1

Association of blood pressure and arterial stiffness with cognition in 2 population-based child and adult cohorts

Sander Lamballais*, Ayesha Sajjad*, Maarten J. G. Leening, Romy Gaillard, Oscar H. Franco, Francesco U. S. Mattace-Raso, Vincent W. V. Jaddoe, Sabine J. Roza, Henning Tiemeier. M. Arfan Ikram

Adapted from:



ABSTRACT

Background

High blood pressure levels and higher arterial stiffness have been shown to be associated with lower cognition during adulthood, possibly by accumulative changes over time. However, vascular factors may already affect the brain during early life.

Methods and Results

We examined the relation between cognition and vascular factors within 5853 children from the Generation R Study (mean age 6.2 years) and 5187 adults from the Rotterdam Study (mean age 61.8 years). Diastolic and systolic blood pressure and arterial stiffness were assessed, the latter by measuring pulse-wave velocity and pulse pressure. For cognition, the Generation R Study relied on nonverbal intelligence, whereas the Rotterdam Study relied on a cognitive test battery to calculate the g-factor, a measure of global cognition. In the Generation R Study, standardized diastolic blood pressure showed a significant association with standardized nonverbal intelligence (β = -0.030, 95% confidence interval = [-0.054; -0.005]) after full adjustment. This association held up after excluding the top diastolic blood pressure decile (β = -0.042 [-0.075; -0.009]), suggesting that the relation holds in normotensives. Within the Rotterdam Study, standardized cognition associated linearly with standardized systolic blood pressure (β = -0.036 [-0.060; -0.012]), standardized pulse-wave velocity (β = -0.064 [-0.095; -0.033]), and standardized pulse pressure (β = -0.044 [-0.069; -0.020], and nonlinearly with standardized diastolic blood pressure (quadratic term β = -0.032 [-0.049; -0.015]) after full adjustment.

Conclusions

Blood pressure and cognition may already be related in the general population during early childhood, albeit differently than during adulthood.

INTRODUCTION

Dementia is a disease posing a huge burden on societies worldwide, and the number of cases is predicted to double by 2040[213]. It is a multifactorial disease, with the role of cardiovascular risk factors increasingly recognized[214]. Blood pressure and arterial stiffness have been of particular interest since these are easily measured and amenable to standard and inexpensive treatments. Interestingly, the effect of high blood pressure may accumulate over time, as mid-life[214-216],persistent hypertension into late life[18] and longer exposure to hypertension[19] have been shown to associate with dementia. Furthermore, mid-life blood pressure and arterial stiffness are also inversely associated with cognition among healthy individuals during later life[217,218], implying that vascular factors and cognitive functioning relate on a clinical as well as a preclinical level.

The previously described studies have focused on mid- and late-life populations but the relevance of the associations during early life remains to be elucidated. Given that blood pressure and arterial stiffness in individuals follow stable trajectories[219-222], partly determined through genetic predisposition[223,224], it is conceivable that the earliest adverse associations between the vascular factors and cognition may be discernable at young age already[225]. Previous studies investigating the link between blood pressure and cognition during childhood and adolescence have primarily focused on hypertensive versus normotensive populations[226-228]. We therefore hypothesize that the associations between vascular factors and cognitive functioning may already be present during childhood as well as within the normal ranges of blood pressure. Finally, given the cumulative effects of vascular risk factors during life, we also hypothesize that the magnitude of the association increases with age.

Hence, we aimed to evaluate whether higher levels of systolic (SBP) and diastolic blood pressure (DBP) and arterial stiffness associated with worse cognition during childhood and mid- to late adulthood. Arterial stiffness was measured directly via carotid-femoral pulse wave velocity (PWV) and indirectly using pulse pressure (PP)[229]. We studied the early life relation in the pediatric Generation R birth cohort and used data from the Rotterdam Study cohort with individuals aged 45 years and over to establish a benchmark for comparison.

METHODS

The data, analytic methods, and study materials will not be made available readily to other researchers for purposes of reproducing the results or replicating the procedure due to legal and informed consent restrictions. Specific requests for consideration can be made to the

respective studies. The first and the corresponding authors had full access to all datasets within this study.

Study population

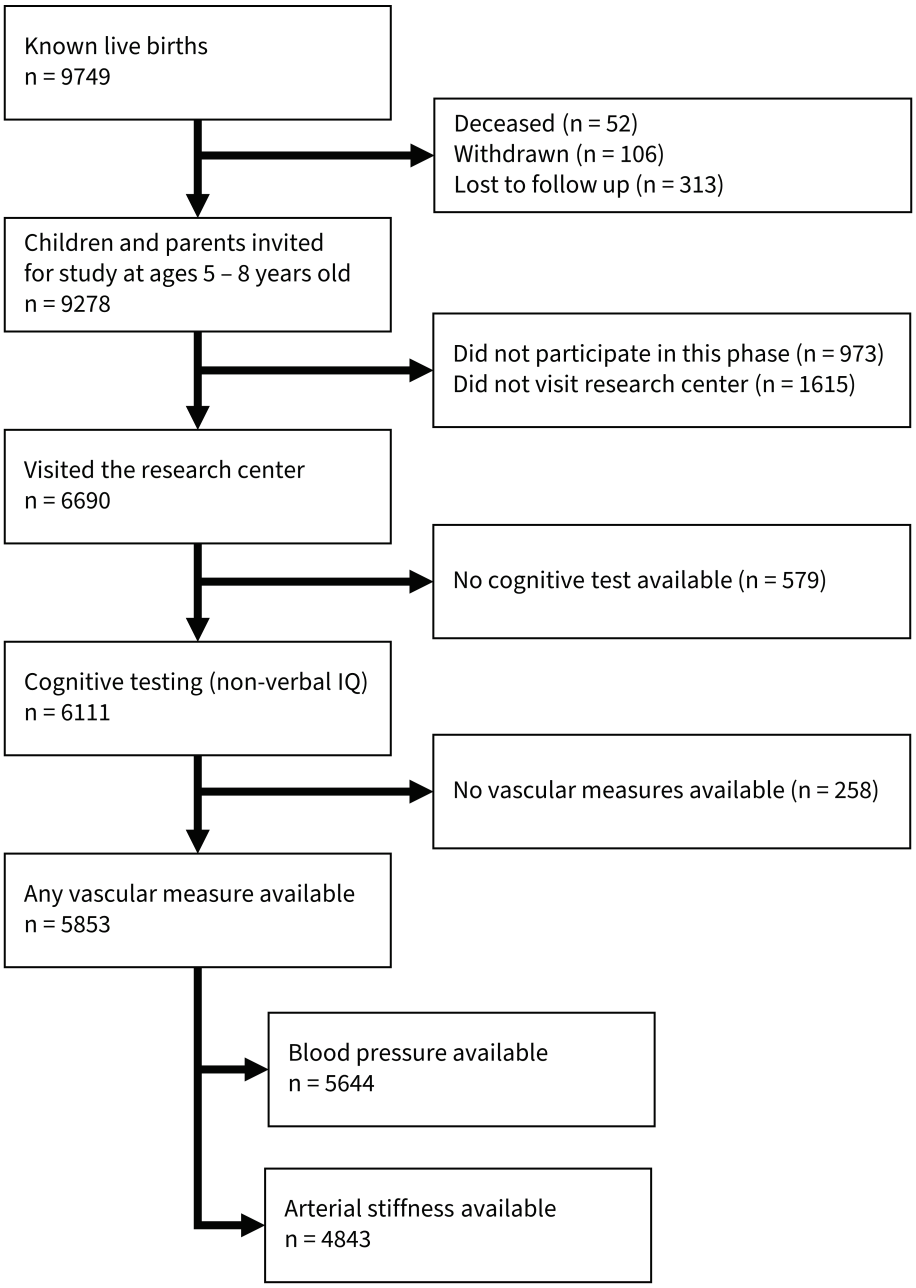
The Generation R Study is a population-based birth cohort in Rotterdam, the Netherlands[22]. In short, 9,745 children were born between April 2002 and January 2006 from mothers who were enrolled during pregnancy or immediately after birth of the child. Of those, 6,690 children visited the research center at the age of five to eight years for follow-up data collection. For this study, we selected 5,853 (mean age is 6.2 ± 0.5 years) children with available data on the intelligence quotient (IQ) measure and at least one of the vascular measures, including carotid-femoral PWV, SBP, and DBP (**Figure 1**). The study was conducted in accordance with the guidelines as proposed in the World Medical Association Declaration of Helsinki and was approved by the Medical Ethical Committee of the Erasmus MC University Medical Center in Rotterdam. Written informed consent was obtained from all primary caregivers of the participants.

The Rotterdam Study is a prospective population-based cohort that started in the Ommoord District of The Netherlands[230]. The first three cohorts – RS-I, RS-II and RS-III – started in 1990, 2000 and 2006 and included 7,983, 3,011 and 3,932 participants, respectively. SBP, DBP and PWV were measured during the third visit of RS-I (RS-I-3), the first visit of RS-II (RS-II-1) and the first visit of RS-III (RS-III-1). Cognitive testing was introduced in 2002, and therefore cognition in the first two cohorts was assessed in a later research phase than the vascular measures, namely in the fourth visit of RS-I (RS-I-4, mean time lag = 4.5 years) and the second visit of RS-II (RS-II-2, mean time lag = 4.1 years). After exclusion of 527 participants with a history of dementia at the time of the vascular measures the final population consisted of 5,187 individuals (**Figure 2**). The Rotterdam Study has been approved by the Medical Ethics Committee of the Erasmus MC University Medical Center in Rotterdam and by the Dutch Ministry of Health, Welfare and Sport. All participants provided written informed consent to participate in the study and to have their information obtained from treating physicians.

Measurement of blood pressure and carotid-femoral pulse wave velocity

In Generation R we measured blood pressure four times at the right brachial artery, in supine position, with 1-minute intervals using the validated automatic sphygmanometer Datascope Accutor Plus (Paramus, NJ)[231]. SBP and DBP were determined by excluding the first measurement and averaging the other measurements. In the Rotterdam Study, blood pressure was measured twice before measurement of PWV. Blood pressure was measured twice with a sphygmomanometer after 5 minutes of rest, and the mean was taken as the participant's reading.

Figure 1 | Flow chart of inclusion for the Generation R Study.



In both studies we assessed carotid-femoral PWV, the reference method to assess aortic stiffness[232], using an automatic device (Complior; Artech Medical, Pantin, France) with participants in supine position. Piezoelectric sensors placed on the skin close to the carotid (proximal) and femoral (distal) artery. PWV was defined as the ratio between the distance travelled by the pulse wave and the time delay between the carotid and femoral pressure waveforms, as expressed in meters per second[233]. To cover a complete respiratory cycle, the mean of at least ten consecutive pressure waveforms was used in the analyses. PWV can be measured reliably with good reproducibility in pediatric populations[234]. Finally, we calculated the pulse pressure by subtracting DBP from SBP.

Cognitive Function

The ethnic profiles of the studies differ significantly. The Rotterdam Study consists of roughly 96% ethnically Dutch participants whereas the Generation R Study only does for roughly 57%. For Generation R we therefore focused on two subtests of the Snijders-Oomen Non-verbal Intelligence Test–Revised (SON-R 2½-7)[177]: 'Mosaics', which tapped into spatial visualization abilities, and 'Categories', which assessed abstract reasoning abilities. The raw scores were converted to non-verbal IQ using age and sex-specific norms. These scores correlated well with IQ scores derived from the total test (r = 0.86) [178] and with the distribution of IQ in the general population[177].

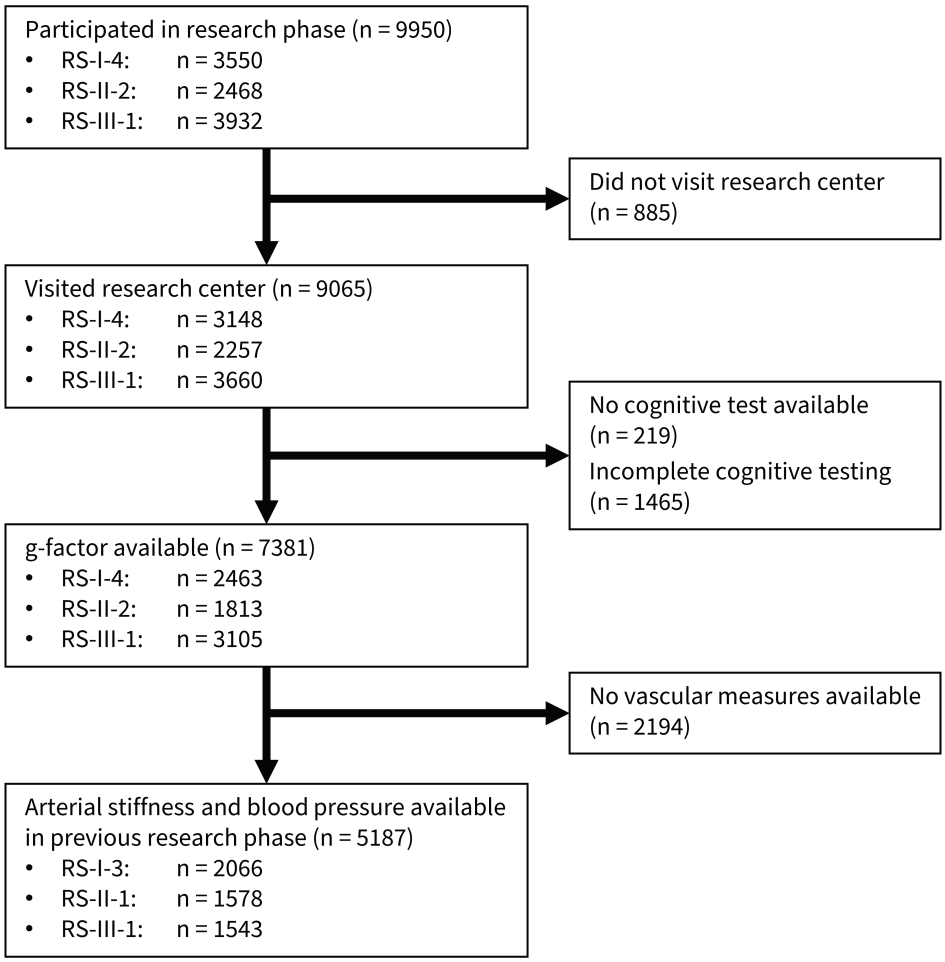
In the Rotterdam Study we focused on a much broader range of cognitive domains to gain a comprehensive understanding of cognitive function in non-demented elderly[61]. We were interested in the general underlying structure of cognition[235], generally known as the g-factor, a stable concept related to intelligence. For the Rotterdam Study we calculated the g-factor by applying principal component analysis to scores from five cognitive tests: Color-word interference Stroop task, letter digit substitution test, verbal fluency test, delayed recall score of the 15-word learning test, and the Purdue pegboard test[61]. The g-factor was defined as the first principal component as returned by the analysis and explained 54.3% of the variance, which is similar to other studies in the literature[236].

As the two studies do not have overlapping scales and as the Rotterdam Study only has one non-verbal non-motor cognitive test, we decided to focus the analyses on general cognition i.e., non-verbal IQ in Generation R and the g-factor in the Rotterdam Study, rather than the subscales.

Measurement of Covariates

For Generation R we included information obtained from midwives and hospital registries on child sex, birth weight (in grams), gestational age at birth and complications during delivery.

Figure 2 | Flow chart of inclusion for the Rotterdam Study.



Body mass index (BMI) of the child was based on height and weight as measured during the visit. Child ethnicity was based on parental countries of birth. Maternal age, maternal smoking during pregnancy and maternal education were assessed by questionnaires. Diet quality was determined by a food frequency questionnaire sent to the children at the age of eight (median age = 8.1 years). The algorithm to score adherence to Dutch dietary guideline has been previously described[237] and is based on sufficient intake of vegetables, fruit, whole grains, fish, legumes, nuts, dairy, oils and soft fats, low intake of sugar-containing beverages, and high-fat and processed meat. The average hours of physical activity per week was based on a parent-

reported questionnaire describing time spent on walking, cycling, physical education, swimming, playing outside and sports participation[238].

In the Rotterdam Study the covariates were measured during the same examination phases as the vascular measures. Smoking status and education were obtained during home interviews. BMI was based on height and weight during the research center visit. Diabetes mellitus was defined as having a fasting glucose level of ≥7.0 mmol/L, or ≥11.1 mmol/L if only non-fasten serum samples were available, or using blood glucose–lowering medication. Fasting glucose levels were available for over 97% of the study population. Data on indication for use of blood pressure-lowering medication were based on information collected by a physician at the research center. Adherence to Dutch dietary guidelines was determined via food frequency questionnaires with a similar algorithm as in Generation R[239]. The food frequency questionnaire was not administered for RS-I-3, so the data from RS-I-1 was used, which was collected four to ten years earlier. Alcohol use was assessed during home interviews with questions based on beer, wine, liquor and moderately strong alcohol types like sherry and port. The algorithm to calculate alcohol in grams per day is provided elsewhere[240]. For RS-I-3 and RS-II-1 physical activity was assessed using a validated adapted version of Zutphen Physical Activity Questionnaire[241] and expressed in MET hours per week[242]. For RS-III-1 physical activity was assessed using the LASA Physical Activity Questionnaire, and expressed in MET hours per week[243]. Due to the difference between the questionnaires we standardized the MET hours per cohort. Finally, the time interval between the two visits was included in the models as it represents the aging between the two visits, and cognition generally declines with age.

Statistical analysis

Initially, both ordinary linear regression and iteratively reweighted least squares were used to analyze the data, the latter using Huber-White standard errors. As the models did not noticeably differ in their estimates and standard errors, we decided to report the results from the iteratively reweighted least squares models. To increase the comparability between studies we standardized the determinants and outcomes. All levels of associations are presented with their 95% confidence intervals (CIs).

Generation R used four models:

- Model 1: Adjusted for sex of the child and age of the child during the visit.
- Model 2: Model 1 further adjusted for birth weight, BMI of the child during visit, ethnicity of the child, gestational age at birth, diet quality score and physical activity.
- Model 3: Model 2 further adjusted for prenatal or perinatal maternal variables, i.e., education level at birth of the child, age at birth of the child, parity and smoking during pregnancy.

- Model 4: Model 3 further adjusted for maternal diabetes mellitus and hypertension during the pregnancy.

Additionally, pre-eclampsia has been consistently associated with elevated blood pressure in the offspring[244], thus we ran a sensitivity analysis excluding children with mothers experiencing pre-eclampsia during pregnancy in order to ensure that the association was not accounted for by this population.

The Rotterdam Study used two models:

- Model 1: Adjusted for age, sex, cohort, and time interval between the measures.
- Model 2: Model 1 further adjusted for education level, BMI, smoking status, diabetes mellitus status, use of blood pressure lowering medication, diet quality score, alcohol intake and physical activity standardized per cohort.

For both studies we additionally adjusted the models with PWV as the determinant for heart rate and mean arterial pressure (DBP + 1/3 × (SBP - DBP)).

Several studies have found non-linear associations between blood pressure and measures of cognition[245]. We therefore performed sensitivity analyses with quadratic terms for DBP and SBP in both the Generation R and the Rotterdam Study cohorts. In addition, stratification for antihypertensive drug use has shown that the association between blood pressure and cognition is diminished in those that use antihypertensive drugs[246], so we stratified for it in our final sensitivity analysis.

All adjustment variables had less than 5% missing values except for maternal smoking, diabetes mellitus and hypertension during pregnancy, with 13.4%, 14.6% and 14.7% missingness, respectively. This pattern is due to the questions being part of a prenatal questionnaire which was not filled out by all participants who were included postpartum. Missing values were imputed 100 times using chained equations and the model fits for each imputed dataset were subsequently pooled[105]. Statistical analyses were performed in R 3.3.3[247]. The package mice 2.30[101] was used for multiple imputation, and MASS 7.3-45[248], sandwich 2.3-4[249] and lmtest 0.9-35[250] to create the iteratively reweighted least squares models.

RESULTS

Generation R

Characteristics of the Generation R Study population stratified by ethnicity are provided in **Table 1**. The mean carotid-femoral PWV was 5.5 (0.9) m/s, and the mean SBP, DBP, and PP were 103 (8), 61 (7) and 42 (7) mmHg, respectively. The average non-verbal IQ score was 101 (15) points. Excluded participants, i.e., that took part in the research phase but did not have data on cognition, had younger mothers at birth (29.4 versus 30.6 years, $p < 0.05$) who were less likely to have obtained higher education (40.4% versus 46.9%, $p < 0.05$) and were less likely to be of Dutch ethnicity (53.8% versus 57.4%, $p < 0.05$).

Table 1 | Characteristics of the Generation R Study participants.

Characteristics (N = 5853)	**Dutch (n = 3350)**	**Other Western (n = 510)**	**Non-Western (n = 1993)**
IQ score	104 ± 15	101 ± 15	96 ± 15
Systolic blood pressure (mmHg)	102 ± 8	103 ± 9	104 ± 9
Diastolic blood pressure (mmHg)	60 ± 7	61 ± 7	62 ± 7
Pulse pressure (mmHg)	42 ± 6	42 ± 7	42 ± 7
Pulse wave velocity (m/s)	5.5 ± 0.9	5.5 ± 0.9	5.6 ± 0.9
Male (%)	49.4	45.0	51.8
Age (y)	6.1 ± 0.4	6.1 ± 0.5	6.3 ± 0.6
Birth weight (g)	3451 ± 584	3406 ± 562	3316 ± 556
Body mass index (kg/m^2)	15.9 ± 1.5	16.2 ± 1.8	16.7 ± 2.2
Gestational age at birth (w)	39.8 ± 1.9	39.8 ± 2.0	39.6 ± 1.8
Diet Quality Score	4.5 ± 1.2	4.5 ± 1.2	4.4 ± 1.3
Physical activity (hours / week)	2.2 ± 1.2	2.2 ± 1.3	1.9 ± 1.2
Characteristics of mothers at partum			
Education (%)			
Low	2.8	7.4	22.1
Intermediate	39.6	34.7	55.2
High	57.6	57.9	22.7
Age (y)	31.4 ± 4.7	31.0 ± 4.9	29.1 ± 5.6
Nulliparous (%)	61.1	56.0	47.6
Pregnancy smoking (%)			
Never	74.8	75.1	74.9
Until aware	10.0	10.9	6.2
Continued	15.2	14.0	18.9
Diabetes mellitus (%)	0.2	0.0	0.9
Hypertension (%)	1.0	1.1	1.9

Table 2 | Associations between standardized blood pressure, standardized pulse pressure, standardized pulse wave velocity and standardized child non-verbal IQ scores within the Generation R Study. The table represent the results from the iteratively reweighted least squares models. All beta values represent the change in the outcome when increasing the value of a determinant by one standard deviation.

non-verbal IQ	Model[a]	β	95% confidence interval	
			lower	upper
SBP	1	-0.059	-0.084	-0.033
	2	-0.027	-0.053	-0.002
	3	-0.018	-0.043	0.008
	4	-0.018	-0.043	0.007
DBP	1	-0.068	-0.094	-0.043
	2	-0.040	-0.065	-0.015
	3	-0.030	-0.055	-0.006
	4	-0.030	-0.054	-0.005
PP	1	-0.001	-0.028	0.026
	2	0.009	-0.016	0.035
	3	0.011	-0.015	0.036
	4	0.010	-0.015	0.036
PWV	1	-0.036	-0.061	-0.010
	2	-0.017	-0.043	0.008
	3	-0.015	-0.039	0.009
	4	-0.015	-0.039	0.009

DBP = diastolic blood pressure; SBP = systolic blood pressure; PP = pulse pressure; PWV = pulse wave velocity.

[a] Model 1 consisted of age (years) and sex. The models for PWV were additionally corrected for pulse rate before measurement (beats/min) and mean arterial pressure (mmHg).

Model 2 consisted of all variables from model 1, birth weight (grams), BMI (kg/m^2), ethnicity of the child, gestational age at birth, diet quality score and physical activity (hours per week).

Model 3 consisted of all variables from model 2, education of the mother at birth of the child, age of the mother at birth of the child, and parity.

Model 4 consisted of all variables from model 3, smoking status, diabetes mellitus during pregnancy and hypertension during pregnancy.

Table 2 shows the associations between standardized PWV, SBP, DBP, and PP with standardized non-verbal IQ in 6-year-old children. When corrected for sex and age of the child, all three vascular measures were negatively associated with the non-verbal IQ score. However, only DBP remained statistically significantly associated after adjusting for all child and maternal covariates (β = -0.030, 95% CI = [-0.054; -0.005]). This relationship was not modified by gender and ethnicity. After exclusion of 114 children (2.2%) whose mothers experienced pre-eclampsia during pregnancy, the association remained statistically significant and the effect size was not affected (β = -0.032, 95% CI = [-0.059; -0.005]). To ensure that the association

held up in a normotensive population we excluded the top decile for DBP, and the effect size seemed to be unaffected (β = -0.42, 95% CI = [-0.075; -0.009]).

Rotterdam Study

Characteristics of the Rotterdam Study participants stratified by cohort are provided in **Table 3**. The carotid-femoral PWV had a mean of 11.6 (3.0) m/s, and SBP, DBP, and PP averaged at 144 (22), 84 (10) and 60 (16) mmHg, respectively. For participants from RS-I and RS-II the median interval between the measurements of exposures and the outcome was 4.4 years (IQR = 4.1-4.7 years). Excluded participants, i.e., those with incomplete data on cognition and the vascular measures, were more likely to use blood pressure lowering medication (35.4% versus 29.3%, $p < 0.05$), meet the criteria for diabetes mellitus (14.1% versus 9.8%, $p < 0.05$) and were more likely to have obtained lower levels of education (42.5% versus 33.0%, $p < 0.05$).

Table 3 | Characteristics of the Rotterdam Study participants.

Characteristics (N = 5187)	RS-I cohort (n = 2066)	RS-II cohort (n = 1578)	RS-III cohort (n = 1543)
Cognitive function g-factor	-0.52 ± 0.93	-0.08 ± 0.88	0.47 ± 0.86
Systolic blood pressure (mmHg)	149 ± 20	150 ± 20	130 ± 18
Diastolic blood pressure (mmHg)	84 ± 10	86 ± 10	80 ± 10
Pulse Pressure (mmHg)	64 ± 16	64 ± 16	50 ± 11
Pulse wave velocity (m/s)	13.0 ± 2.8	12.2 ± 2.8	9.1 ± 1.6
Age (y)	63.6 ± 5.7	63.1 ± 6.6	58.0 ± 7.3
Sex, Male (%)	42.0	44.0	42.7
Time interval between measurements (y)	4.5 ± 0.4	4.1 ± 0.5	0.0 ± 0.0
Education (%)			
Low	43.3	30.1	26.0
Medium	44.7	51.6	46.2
High	12.0	18.3	27.8
Body mass index (kg/m^2)	26.8 ± 3.8	27.1 ± 3.9	27.4 ± 4.3
Smoking (%)			
Never	33.0	30.1	33.3
Past	51.7	51.4	46.3
Current	15.3	18.5	20.4
Diabetes Mellitus (%)	10.0	10.8	8.5
Blood pressure lowering medication (%)	36.5	24.4	27.7
Diet quality score	6.9 ± 1.8	6.2 ± 1.9	6.9 ± 1.9
Alcohol (grams / day)	11.2 ± 15.4	11.0 ± 14.3	8.8 ± 9.5

Table 4 shows the associations between standardized carotid-femoral PWV, PP, and blood pressure with the standardized g-factor. For DBP no significant associations were shown in model 1 and 2 (β = -0.006, 95% CI = [-0.028; 0.017]). In contrast, carotid-femoral PWV (β = -0.064, 95% CI = [-0.095; -0.033]), SBP (β = -0.036, 95% CI = [-0.060; -0.012]) and PP (β = -0.044, 95% CI = [-0.069; -0.020]) showed statistically significant negative associations with the g-factor. These relationships were not modified by gender. For comparison, a year increase in age led to a 0.062 standard deviation decrease of the g-factor (95% CI = [-0.065; -0.058]).

Non-linear associations

Within Generation R the quadratic terms for neither DBP nor SBP reached statistical significance (p > 0.05). Within the Rotterdam Study the quadratic term for DBP (β = -0.032, CI 95% = [-0.048; -0.015]) but not SBP (β = -0.013, CI 95% = [-0.029; 0.002]) reached statistical significance. This suggests the presence of a non-linear relationship between DBP and the g-factor where more extreme values of DBP, i.e., both at the lower and higher ends, were associated with a quadratic decrease in the g-factor.

Table 4 | Associations between standardized blood pressure, standardized pulse pressure standardized pulse wave velocity and the standardized g-factor within the Rotterdam Study. The table represent the results from the iteratively reweighted least squares models. All beta values represent the change in the outcome when increasing the value of a determinant by one standard deviation.

g-factor	Model[a]	β	95% confidence interval	
			lower	upper
SBP	1	-0.059	-0.084	-0.033
	2	-0.036	-0.060	-0.012
DBP	1	-0.021	-0.045	0.001
	2	-0.006	-0.028	0.017
PP	1	-0.065	-0.091	-0.039
	2	-0.044	-0.069	-0.020
PWV	1	-0.080	-0.112	-0.047
	2	-0.064	-0.095	-0.033

DBP = diastolic blood pressure; SBP = systolic blood pressure; PP = pulse pressure; PWV = carotid-femoral pulse wave velocity.

[a] Model 1 consisted of age (years), sex, cohort and time difference between exposure and outcome measurements (years). The models for PWV were additionally corrected for pulse rate before measurement (beats/min) and mean arterial pressure (mmHg).

Model 2 consisted of all variables from model 1, education, BMI (kg/m^2), smoking status, diabetes mellitus, blood pressure lowering medication usage, diet quality score, alcohol intake (grams / day) and physical activity standardized per cohort (MET hours / week).

Antihypertensive drug use

Within those that did not use antihypertensive drugs we found similar associations as described above, i.e., linear associations for PWV (β = -0.082, 95% CI = [-0.119; -0.046]) and SBP (β = -0.031, 95% CI = [-0.060; -0.002]) with cognition, and a statistically significant quadratic term for DB P (β = -0.022, 95% CI = [-0.038; -0.005]). In those using antihypertensive drug none of the linear terms reached statistical significance. However, the coefficient for the quadratic DBP term remained statistically significant (β = -0.055, 95% CI = [-0.078; -0.031]).

DISCUSSION

In both the pediatric and elderly cohorts vascular measures associated with cognition. In particular, DBP but not SBP, PP or PWV was negatively associated with non-verbal IQ amongst children of 6 years old. In middle-aged and elderly, both arterial stiffness and SBP were negatively associated with the g-factor, while DBP showed an inverted non-linear relation. Thus, the exact nature of the vascular-cognitive relation might depend on the life phase.

Cognition in elderly has been previously linked to blood pressure[217,245]. However, most studies dichotomized blood pressure measures into presence or absence of hypertension, and the cut-offs used varied amongst the studies. In addition, most of those studies also did not adjust for antihypertensive drug use, which has been suggested to modify the cognitive-vascular relationship[246]. In the current study, we showed that the associations were statistically significant for continuous blood pressure measures and when adjusting for use of antihypertensive drugs.

The findings within the Generation R cohort suggest that the vascular-cognitive relation may extend to earlier life phases as well. The association between DBP and non-verbal IQ in the pediatric cohort showed a similar effect size as the SBP and arterial stiffness associations with the g-factor within the Rotterdam Study. While blood pressure and cognition have been studied in pediatric populations, those studies have generally focused on hypertensive cases[226-228,251]. Lande and colleagues[226] showed among 5077 children aged 6 to 16 years that elevated blood pressure levels seems to relate to a digit span test, although this effect disappeared after multiple testing correction. In addition, Adams and colleagues[227] showed in 201 children that those with sustained primary hypertension were more likely to suffer from learning disabilities. While the effect sizes cannot be directly compared due to differing determinants and outcomes the current study does show that the association holds for normotensive populations. It also bolsters the idea that the association between the vascular and cognitive performance may have its roots in childhood.

Different mechanisms have been proposed to link hypertension and arterial stiffness to cognition. Cerebrovascular disease burden is a likely mechanism during adulthood. For example, brain plaque and tangle burden seemingly mediate the effect of diastolic blood pressure on cognition[252]. Additionally, the relation between arterial stiffness and memory may be mediated by cerebrovascular resistance and white matter hyperintensities[253]. However, such pathways may depend on aging[254] and could thus not be relevant for explaining the findings in our pediatric sample.

A more likely mechanism is cerebrovascular reactivity, which encompasses the vasodilatory and vasoconstrictive ability of cerebral vessels. Indeed, Settakis and colleagues showed that hypertensive adolescents (aged 14-18 years) had reduced cerebrovascular reactivity compared to normotensives after a 30 second breath hold[255]. The most profound difference was found for the diastolic blood flow velocity. Another study by Wong and colleagues in hypertensive and normotensive children (aged 7-20 years) showed that diastolic blood pressure related more strongly to cerebrovascular reactivity than systolic blood pressure after a CO_2 challenge test[256]. Hypertensive status has also been linked to reduced cerebrovascular reactivity in regions related to the default mode network[257], which in turn plays a role in cognitive functioning[258]. The association between DBP and non-verbal IQ in our pediatric sample may thus be mediated by cerebrovascular reactivity.

Lower DBP seemed to be associated with lower levels of cognitive function in the adults but not the children. Several mechanisms can underlie this difference. First, blood pressure in children may be more tightly controlled, especially at lower values, than in adults[259,260]. As such the lower end of the blood pressure distribution in children does not reach levels at which they become detrimental to the brain. Second, brains of children may better withstand low blood pressure, for instance because of better compensatory mechanisms of the small peripheral vessels. Third, perhaps prolonged exposure to hypotension is necessary for it to impair cognitive function. Further studies are needed to confirm and explain this finding.

The association between arterial stiffness and the g-factor in elderly could be related to the SBP findings. Interestingly, the relation between SBP and arterial stiffness may in fact be bidirectional[261,262]. The former increases pulsatile aortic wall stress, leading to stretching and thus stiffening of elastic lamellae of the large arteries. Conversely, arterial stiffness has been shown to predict hypertension in mid to late adulthood[263-265].

The findings show statistical significance, but the question remains whether they warrant clinical implications since one standard deviation change in DBP within Generation R pediatric sample roughly translated to half of a non-verbal IQ point change. It remains to be seen whether treatment of high blood pressure in childhood would have beneficial effects on

cognition later in life. Our study does underscore that the detrimental effects of blood pressure, however small those may be, might have an origin already in early life. The findings therefore give etiological insights into the interplay between the vascular system and cognition during the life course. Effect sizes found in the Rotterdam Study adult sample can be interpreted more clearly, with one standard deviation increase in SBP and PWV having about the same effect on the g-factor as aging half a year and one year, respectively. Such findings do warrant further investigation of any causal benefit of blood pressure lowering medication and lifestyle changes on cognition.

Several limitations should be taken into account. First, the comparability of results between the cohorts is hampered due to the difference in measures of cognition. The children were tested for non-verbal IQ due to the diverse ethnic background of the children, while the g-factor did include a verbal component. In addition, the g-factor may capture other aspects of cognition that were not assessed in the children. Second, both cohorts were studied cross-sectionally, which increased the comparability between the cohorts but did not allow a clearer, developmental narrative, and also prevented any causal interpretations. In particular, the hypothesized relationship could actually be reversed, with higher levels of cognition being associated with healthier diets and lower levels of sedentary behavior[266,267]. Fourth, the determinants and outcome were not measured in same visit for the majority of the Rotterdam Study population. This may have led to survivor bias due to selective attrition between the visits, and potentially residual confounding due to time-dependent covariates. Finally, our cohorts do not cover the age ranges of 6 to 45 years, and the exact life course relation between the vascular and cognitive systems remains to be elucidated. However, both cohorts were recruited in the same study area and rely on similar methodologies, strengthening the comparability of the findings.

We found that blood pressure and arterial stiffness showed significant associations with cognition during mid- and late adulthood. In addition, we showed that DBP during childhood also associates with non-verbal cognition, even after excluding the top DBP decile. Thus, we have provided evidence that the associations between cognition and vascular factors hold for the general population and across the age spectrum, warranting further investigation in the exact mechanisms that govern the associations over the whole life course.

4.2

The association of vascular risk factors during early adulthood with subsequent cognition and brain structure

Rowina F. Hussainali*, Sander Lamballais*, Sarah Schalekamp-Timmermans, Jeanine E. Roeters van Lennep, Eric A. P. Steegers, M. Arfan Ikram

ABSTRACT

Background

The risk to develop Alzheimer's disease during late-life is influenced by vascular risk factors during midlife and late-life. These factors partly exert their effect by increasing the burden of neuropathology and affecting cognitive function. We aimed to assess whether vascular risk factors during early adulthood also affect subsequent brain structure and cognition.

Methods

We studied 1629 participants from the ORACLE Study, a sub-study within the population-based Generation R Study cohort. At baseline (2002-2006) (mean age = 32.3, standard deviation = 4.9), vascular risk factors were measured, including blood pressure, body mass index, and smoking for all participants, and HDL-c, LDL-c, triglycerides, and remnant cholesterol for female participants. During follow-up (2017-2020), the participants underwent a cognitive test battery as well as an MRI scan of the brain. Brain-related measures included total brain volume, ventricular volume, and log-transformed white matter lesion (WML) volume. All measures were standardized.

Findings

Vascular risk factors during early adulthood were not associated with subsequent cognition, total brain volume or ventricular volume. However, they did associate with a larger WML after multiple testing corrections (β_{SBP} = 0.07, 95% confidence interval = [0.02, 0.11]); β_{DBP} = 0.07 [0.02, 0.11]); β_{BMI} = 0.06 [0.02, 0.11]; $\beta_{smoking}$ = 0.13 [0.03, 0.23]). The interactions of age with systolic and diastolic blood pressure reached statistical significance. When stratified by age, the association of systolic and diastolic blood pressure with WML volume was present in all groups including participants above 30 years.

Interpretation

Vascular risk factors during early adulthood seem to relate to higher levels of neuropathological burden around 15 years later, which in turn is a risk factor for the development of Alzheimer's disease. If the findings are replicated, blood pressure monitoring and interventions may be needed from age 30 onwards to minimize subsequent neuropathological burden.

INTRODUCTION

The most common form of dementia is Alzheimer's disease (AD), which is a late-life neurodegenerative disease with a large health burden. The risk to develop AD is largely influenced vascular risk factors (VRFs) like hypertension, cholesterol levels, obesity, and smoking during midlife and late-life[38,49,50,268,269]. The VRFs likely influence the incidence of dementia by causing lesions in the brain[270,271], which consequently lead to decline of cognitive function[217,272,273], both of which are important markers for AD[274]. VRFs are modifiable, and a recent meta-analysis of observational studies found that the use of antihypertensive drugs reduced the risk of dementia in patients with hypertension[7]. There is thus an urgent need to better understand how VRF and brain health are related.

Although the influence of midlife VRF on the incidence of dementia is widely established, less is known about the relevance of VRF during earlier phases of adulthood. A high VRF burden during early adulthood could lead to the accumulation of physiological damage to the brain in subsequent years and increase the predisposition to further structural lesions during midlife. One study of 463 participants found that VRF burden from the age of 36 years onwards associated with higher neuropathological burden and smaller brain volumes at 70 years of age[275]. Similarly, a study in 3000 participants aged 18 to 30 years showed that poorer levels of VRFs associated with poorer overall cognitive functioning 25 years later[276,277]. However, VRF may already exert their effects on a shorter time scale. Furthermore, both studies considered global metrics for VRFs, even though the individual VRFs may have differing effects with age.

In the current study, we aimed to determine the associations of specific VRFs during early adulthood associated with subsequent cognition and brain structure roughly 15 years later. We explored the different vascular risk factors individually and investigated at which age the effects of vascular risk factors exposure become apparent.

METHODS

Participants

This study was performed as part of the Generation R Study, a population-based prospective birth cohort study based in Rotterdam, the Netherlands[22]. Women were recruited during pregnancy and their children were subsequently followed up throughout childhood. The women participated in a wide range of measurements, including on VRFs, at different time points. When the children were around the age of 14 years, a subset of parents was invited to participate in a parent-focused research visit: the origins of Alzheimer's disease across the life-course (ORACLE) Study[278].

The present study focuses on data from two visits per parent: The research visit that took place between 2002 and 2006 (first visit) and the ORACLE Study which took place from 2017 to 2020 (the second visit). In total, 2084 participants visited. For the analyses, we excluded participants with incomplete data on the VRFs during the first visit (n = 443). We included participants with either complete cognitive or neuroimaging data for the second visit. In the cognition analyses we excluded participants who were colorblind (n = 4), were physically not capable to complete some of the tests (n = 12), lacked motivation during the tests (n = 3), did not follow any of the instructions (n = 26) or had missing or incomplete data on any of the tests (n = 62). For neuroimaging, we excluded participants who were not scanned (n = 79), had incidental findings that interfered with the analyses (n = 9), or whose scans were of too low quality (n = 43). Information on cognitive testing was available and complete in 1534 participants, and neuroimaging data in 1510 participants, yielding 1629 unique participants. A detailed flow chart of the study population is shown in **Figure 1**.

The study was designed in accordance with the guidelines set by the World Medical Association Declaration of Helsinki. The study has been approved by the Medical Ethics Committee of the Erasmus Medical Center, University Medical Center, Rotterdam, the Netherlands. Written informed consent was obtained for all participants.

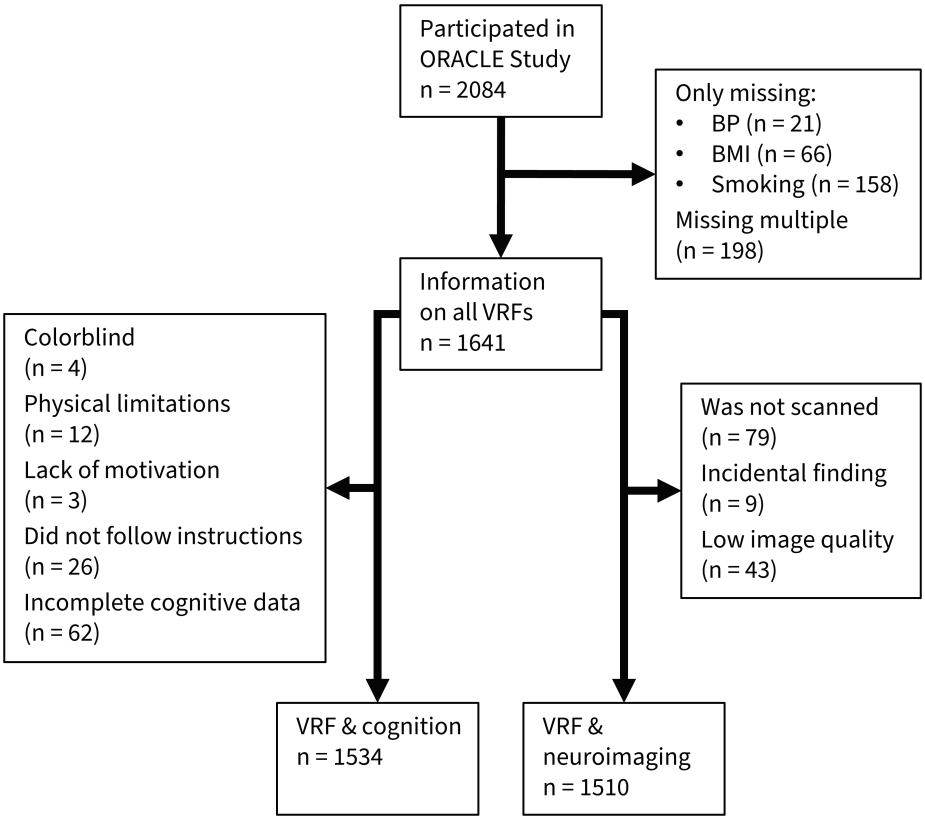
Vascular risk factors

During the first visit, measurements for a number of VRFs were collected. For the present study we focused on SBP, DBP, body mass index (BMI), and smoking. For the women, information on serum lipid profiles was also available, i.e., high density lipoprotein cholesterol (HDL-c), low density lipoprotein cholesterol (LDL-c), triglycerides (TG) and remnant cholesterol.

Blood pressure

During the first visit, blood pressure was measured using the validated automatic sphygmomanometer Datascope Accutorr Plus (Paramus, NJ, USA). During the second visit, the validated Omron 907 automated digital oscillometric sphygmomanometer (OMRON, Matsusaka Co., Ltd., Japan) was used. A cuff was placed around the right upper arm. In case of an upper arm exceeding 32 centimeters, a larger cuff (32-44 cm or 42-52 cm) was used. At both visits, the mean value of two blood pressure readings over a five-minute interval was documented for each participant. Blood pressure was measured by trained research assistants.

Figure 1 | Flow chart of the study population.



BMI and smoking

During both visits, height (cm) and weight (kg) were measured without shoes. From this information, BMI (kg/m^2) was calculated. Information on smoking at baseline was obtained through self-report questionnaires. The participants were asked whether they had smoked cigarettes in the last months, and if so, how many. For the purpose of our analyses, we focused on whether the participants had smoked or not.

Lipid profiles

Non-fasting cholesterol and TG levels were determined from blood samples collected during the first visit for the women. Lipid profiles are not available for the men. The procedure has been previously described[52]. In brief, blood was collected through venous puncture, were processed within 4 hours after collection and subsequently stored at -80 °C. Total cholesterol, HDL-c and TG levels were determined with the Vital Scientific (Merck) Selectra E Chemistry

Analyzer (Vital Scientific N. V., Dieren, The Netherlands). LDL-c was estimated using the Friedewald equation[185] and remnant cholesterol was calculated by subtracting HDL-c and LDL-c from the total cholesterol levels.

Cognitive Assessment

All participants underwent a cognitive test battery during the second visit. A detailed description of the procedure can be found elsewhere[278]. Briefly, cognitive assessment consisted of six tests, assessing various domains of cognition. At onset of the second research visit study, the assessment consisted of four tests. The first test was the 15-word learning test (15WLT), to assess verbal learning and verbal memory. The second test was the Stroop task, a task that assesses selective attention and automaticity. The third test was the letter-digit substitution test, which assesses processing speed as well as executive function. The fourth test was a verbal fluency test, which assesses efficiency of searching the long-term memory. After several hundred participants, a fifth and sixth test were added: the Purdue pegboard test (PPT) to measure dexterity and fine motor skills, and the design organization test (DOT) to assess visuospatial ability.

To summarize all tests into a single score for global cognition, known as the g-factor[279], we used principal component analysis and isolated the first component. The g-factor explained 64.1% of the variance amongst the cognitive tests which is in agreement with previous literature[280]. Primary analyses for cognition focused on the g-factor, and secondary analyses involved separate models for each cognitive test score.

Image Acquisition and Image Processing

During the second visit, structural magnetic resonance imaging (MRI) was obtained on a 3T GE Discovery MR750w MRI System (General Electric, Milwaukee, WI, USA) with an 8-channel receive-only head coil. We collected T_1-weighted with an inversion recovery-prepared fast spoiled gradient recalled sequence (T_R = 8.77 ms, T_E = 3.4 ms, T_I = 600 ms, flip angle = 10°, Field of view = 220 x 220 mm, acquisition matrix = 220 x 220, slice thickness = 2 mm, number of slices = 230).

We processed the T_1-weighted images through the FreeSurfer analysis suite, version 6.0. Non-brain tissue was removed and the voxel intensities were normalized for B1 inhomogeneity. Next, the images were segmented into gray matter, white matter, and cerebrospinal fluid. Ventricular volume was calculated by taking the sum of the lateral ventricles, the third ventricle and the fourth ventricle. The brain measures of interest were the total brain volume, gray matter volume, white matter volume, total ventricular volume, and total white matter lesion (WML) volume. The fluid-attenuated inversion recovery (FLAIR) sequence was manually inspected on the presence of cortical and lacunar infarcts. Cortical infarcts were defined as

infarcts involving the cortical grey matter, regardless the size. Lacunar infarcts were classified as such if an infarct was present within the white matter, cerebellum, basal ganglia, and thalamus, if larger than 3mm and smaller than 15 mm. To confirm an infarct, it had to be visible on both the FLAIR and the T1 sequence. To assess cerebral microbleeds (CMB), we manually inspected the 3D T2*-weighted gradient-recalled echo sequence. CMBs present themselves as small hypointense foci with a maximum size up to 10 mm.

Other measurements

Information on ethnicity and education level was obtained through self-reported questionnaires administered at the first visit. Ethnicity was recoded to three levels: Dutch, other Western and non-Western. Education level was also recoded into three levels: low (no or primary education), intermediate (lower/intermediate vocational education or secondary education) and high (higher vocational education or university). Information on blood pressure lowering medication was obtained through questionnaires around the time of the first visit.

Statistical Analysis

Statistical analyses were performed in R, version 3.6.2[95]. The primary goal was to examine the associations of VRFs at the first visit with cognition and brain structure at the second visit. To achieve this, regression models were constructed for every combination of VRFs and continuous outcome measures. As for models with a binary outcome, logistic regression models were used. To control for confounding the models were adjusted for age at the first visit, the age difference between the first and the second visit, ethnicity, and education level. Analyses for total brain volume, gray matter volume, white matter volume, ventricular volume, WML volume, cortical and lacunar infarcts, and CMBs were further adjusted for intracranial volume. All measures were standardized. Multiple testing was corrected for across all associations using the Benjamini-Hochberg procedure[281].

The associations may depend on age, with the associations becoming more apparent towards later adulthood. As such, we aimed to test whether the associations changed with age. To do so, we added interaction terms for age and VRFs for models that were statistically significant in the previous analyses. We further applied the Johnson-Neyman procedure to visualize how age affected the associations of interest[282].

Missing covariate data was imputed 50 times with the package mice using chained equations (50 iterations) and the estimates from the models will be pooled according to Rubin's rules[101,105,283].

RESULTS

Population characteristics

The population characteristics for the first and second visit are shown in **Table 1**. In total, 1629 subjects were included in the study. The mean age of the participants was 32.3 (standard deviation = 4.9) years at the first visit and 47.3 (4.9) years at the second visit.

Table 1 | Characteristics of the study population.

Characteristics	All (N = 1629) M ± SD, N (%)	Women (n = 1074) M ± SD, N (%)	Men (n = 555) M ± SD, N (%)
Age at first visit (years)	32.3 ± 4.9	31.4 ± 4.5	34.1 ± 5.1
Age at second visit (years)	47.3 ± 4.9	46.4 ± 4.5	49.1 ± 5.2
Ethnicity			
Dutch	1216 (74.7)	742 (69.2)	474 (85.4)
Other Western	326 (20.0)	264 (24.6)	62 (11.2)
Non-Western	86 (5.3)	67 (6.2)	19 (3.4)
Education level			
Low	212 (13.1)	142 (13.3)	70 (12.8)
Intermediate	459 (28.4)	314 (29.4)	145 (26.5)
High	944 (58.5)	611 (57.3)	333 (60.8)
Measurements at first visit			
SBP (mmHg)	121.6 ± 13.9	117.2 ± 12.2	130.1 ± 13.0
DBP (mmHg)	70.4 ± 10.1	68.8 ± 9.6	73.3 ± 10.2
BMI (kg/m^2)	24.6 ± 3.9	24.4 ± 4.2	25.1 ± 3.2
Smoking			
No	1206 (74.0)	863 (80.4)	343 (61.8)
Yes	423 (26.0)	211 (19.6)	212 (38.2)
HDL-c (mmol/L)		1.8 ± 0.3	
LDL-c (mmol/L)		2.4 ± 0.7	
Triglycerides (mmol/L)		0.6 ± 0.2	
Remnant cholesterol (mmol/L)		1.3 ± 0.5	

At the first visit, the participants had an average SBP of 121.6 (13.9) mmHg, DBP of 70.4 (10.1) mmHg, and a BMI of 24.6 (3.9) kg/m^2. As for the women, they had an average HDL-c of 1.8 (0.3) mmol/L, LDL-c of 2.4 (0.7) mmol/L, TG of 1.3 (0.5) mmol/L, and remnant cholesterol of 0.6 (0.2) mmol/L. The g-factor had a mean value of 0 (1) as it is a standardized metric. The average total brain volume was 992 (104) cm^3, a mean ventricular volume of 4217 (1986) mm^3 and a mean WML volume of 883 (635) mm^3.

Vascular risk factors and cognition

Results from the associations models of standardized VRFs from the first visit with cognition from the second visit are shown in **Table 2**. Although all VRFs were negatively associated with the g-factor, none of these associations reached statistical significance. When considering the individual cognitive tests, higher levels of BMI were statistically significantly associated with poorer performance on the Stroop task (β = 0.08, 95% CI = [0.03; 0.12]). Higher levels of SBP (β = -0.13, 95% CI = [-0.19; -0.06]) and DBP (β = -0.11, 95% CI = [-0.17; -0.05]) were statistically significantly associated with a lower score on the PPB test.

Vascular risk factors and brain structure

The results for brain structure are shown in **Table 3**. None of the VRFs were statistically significantly associated with subsequent total brain volume or total ventricular volume (all p > 0.05). Higher levels of SBP and DBP as well as smoking were related to a larger WML volume (β_{SBP} = 0.08, 95% CI = [0.03; 0.12]; βD_{BP} = 0.06, 95% CI = [0.02; 0.10]; $\beta_{smoking}$ = 0.15, 95% CI = [0.06; 0.24]). None of the VRFs were statistically significantly associated with CMBs, cortical infarcts, or lacunar infarcts.

Age-dependent associations

We explored how the identified associations changed with age by adding an interaction term for age and the VRFs to the models. We found that age statistically significantly interacted with SBP, DBP, and BMI in their association with WML volume. These associations were visualized using the Johnson-Neyman procedure and are displayed in **Figure 2**. Through the Johnson-Neyman procedure, we found that SBP, DBP, and BMI showed statistically significant associations with WML volume from around the age of 30 onwards. Furthermore, the strength of the associations increased with age.

Table 2 | Standardized associations between the VRFs and the cognitive measures. Bolded associations were statistically significant after multiple testing corrections.

Per SD difference	g-factor β (95% CI)	15WLT β (95% CI)	Stroop 3 β (95% CI)	LDST β (95% CI)	WFT β (95% CI)	DOT β (95% CI)	PPB β (95% CI)
SBP	0.01	-0.01	0.00	0.02	-0.01	0.00	-0.13
	(-0.04; 0.06)	(-0.06; 0.05)	(-0.04; 0.05)	(-0.03; 0.07)	(-0.06; 0.05)	(-0.06; 0.05)	(-0.19; -0.06)
DBP	0.02	0.00	0.03	0.01	0.02	-0.02	-0.11
	(-0.03; 0.07)	(-0.05; 0.05)	(-0.01; 0.08)	(-0.04; 0.05)	(-0.03; 0.06)	(-0.07; 0.03)	(-0.17; -0.05)
BMI	0.00	-0.03	0.08	-0.03	0.03	-0.01	-0.08
	(-0.04; 0.05)	(-0.08; 0.02)	(0.03; 0.12)	(-0.07; 0.02)	(-0.02; 0.08)	(-0.06; 0.04)	(-0.14; -0.02)
Smoking	0.01	-0.06	0.00	-0.04	0.05	-0.08	-0.12
	(-0.10; 0.12)	(-0.17; 0.05)	(-0.10; 0.10)	(-0.14; 0.06)	(-0.06; 0.15)	(-0.20; 0.04)	(-0.25; 0.01)
Mothers only							
HDL-c	0.01	0.02	-0.03	0.02	-0.01	0.05	-0.01
	(-0.04; 0.06)	(-0.04; 0.07)	(-0.07; 0.02)	(-0.03; 0.07)	(-0.06; 0.05)	(-0.01; 0.11)	(-0.07; 0.06)
LDL-c	-0.03	0.04	0.03	-0.02	-0.02	0.04	-0.06
	(-0.08; 0.03)	(-0.01; 0.09)	(-0.01; 0.08)	(-0.07; 0.02)	(-0.07; 0.03)	(-0.01; 0.10)	(-0.13; 0.00)
Triglycerides	-0.03	-0.03	0.05	-0.03	-0.01	-0.01	-0.01
	(-0.08; 0.03)	(-0.08; 0.03)	(0.00; 0.10)	(-0.08; 0.02)	(-0.06; 0.04)	(-0.07; 0.05)	(-0.07; 0.06)
Remnant cholesterol	-0.02	-0.03	0.05	-0.03	-0.01	-0.01	-0.01
	(-0.08; 0.03)	(-0.08; 0.03)	(0.00; 0.10)	(-0.08; 0.02)	(-0.06; 0.04)	(-0.07; 0.05)	(-0.07; 0.06)

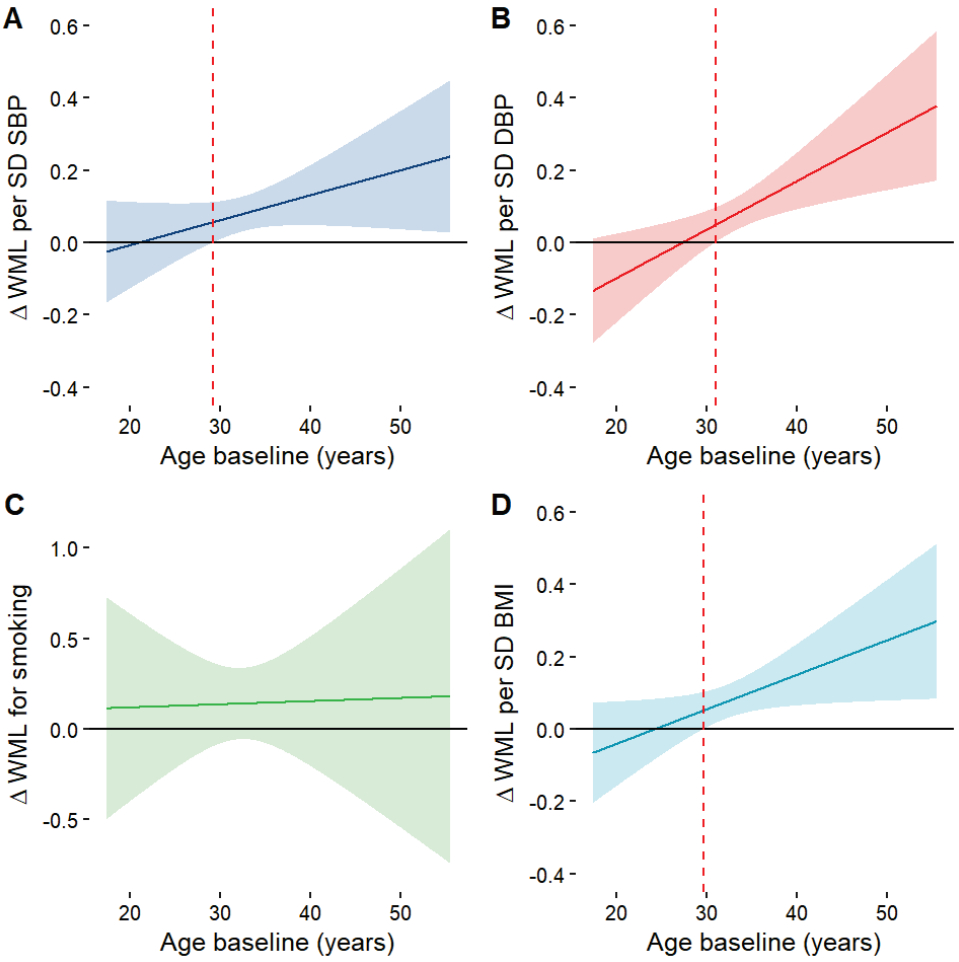
15WLT = 15-word learning test; BMI = body mass index; CI = confidence interval; DBP = diastolic blood pressure; HDL-c = high-density lipoprotein cholesterol; LDL-c = low-density lipoprotein cholesterol; LDS = letter digit substitution test; SBP = systolic blood pressure; SD = standard deviation; WFT = word fluency test.

Table 3 | Standardized associations between the VRFs and brain metrics. Bolded associations were statistically significant after multiple testing corrections.

Per SD difference	Total brain volume β (95% CI)	Ventricular volume β (95% CI)	White matter lesion volume β (95% CI)	Any cerebral microbleeds β (95% CI)	Any cortical infarcts β (95% CI)	Any lacunar infarcts β (95% CI)
SBP	0.00	0.04	0.08	0.01	0.00	0.01
	(-0.03; 0.02)	(-0.01; 0.08)	(0.03; 0.12)	(-0.01; 0.03)	(0.00; 0.01)	(0.00; 0.02)
DBP	-0.01	0.05	0.06	0.01	0.00	0.00
	(-0.03; 0.01)	(0.01; 0.09)	(0.02; 0.10)	(-0.01; 0.03)	(0.00; 0.01)	(-0.01; 0.01)
BMI	0.01	0.03	0.06	0.00	0.00	0.01
	(-0.01; 0.04)	(-0.01; 0.07)	(0.02; 0.10)	(-0.02; 0.02)	(0.00; 0.01)	(0.00; 0.02)
Smoking	-0.07	0.12	0.15	0.00	-0.01	0.01
	(-0.12; -0.02)	(0.02; 0.22)	(0.06; 0.24)	(-0.04; 0.04)	(-0.02; 0.00)	(-0.01; 0.03)
Mothers only						
HDL-c	-0.01	0.00	-0.02	0.00	0.00	0.00
	(-0.03; 0.02)	(-0.05; 0.05)	(-0.06; 0.03)	(-0.02; 0.01)	(0.00; 0.00)	(-0.01; 0.00)
LDL-c	-0.01	0.00	0.01	0.00	0.00	0.00
	(-0.04; 0.01)	(-0.04; 0.05)	(-0.03; 0.06)	(-0.01; 0.02)	(0.00; 0.00)	(-0.01; 0.01)
Triglycerides	-0.01	0.02	0.01	0.00	0.00	0.00
	(-0.03; 0.01)	(-0.03; 0.07)	(-0.04; 0.05)	(-0.02; 0.02)	(-0.01; 0.00)	(-0.01; 0.01)
Remnant cholesterol	-0.01	0.02	0.01	0.00	0.00	0.00
	(-0.03; 0.01)	(-0.03; 0.07)	(-0.04; 0.05)	(-0.02; 0.02)	(-0.01; 0.00)	(-0.01; 0.01)

CI = confidence interval; BMI = body mass index; DBP = diastolic blood pressure; HDL-c = high-density lipoprotein cholesterol; LDL-c = low-density lipoprotein cholesterol; SBP = systolic blood pressure; SD = standard deviation.

Figure 2 | Age-dependence of the coefficients of the association between VRFs and WML volume. The plots were obtained using the Johnson-Neyman procedure. The x-axis denotes age at baseline, the y-axis denotes the coefficient of the association for SBP (A), DBP (B), smoking (C), and BMI (D). The dashed red line indicates the age from which the association is statistically significant. BMI = body mass index; DBP = diastolic blood pressure; SBP = systolic blood pressure; SD = standard deviation.



DISCUSSION

In this study we found that VRFs during early adulthood are associated with brain health 15 years later. In particular, we showed that having higher blood pressure levels and higher BMI

at 30 years and older is statistically significantly associated with more WMLs after 15 years. The findings provide a clear age cut-off at which blood pressure and BMI start to play a formative role in brain health. Finally, VRFs between the age of 25 and 50 years also seemed to be associated with performance on specific cognitive tests after 15 years, albeit we found no evidence that this association changes with age nor did we find such an association with global cognitive functioning.

The WML-related findings build upon the existing literature. Blood pressure has previously been associated with WML load and WML progression in individuals aged 50 years and older[270,284-289]. However, studies on VRFs and WML in individuals before the age of 50 years are sparse. Work from the Insight 46 study showed that VRF burden at the age of 36 years, but not blood pressure by itself, associated with WMLs around the of 70 years. Our findings make a case that VRFs from the age of 30 years already relate to WML 15 years later. Furthermore, we also show that blood pressure and BMI are the likely VRFs that carry this association.

The relevance of blood pressure in the formation of WMLs has been established[290]. Hypertension leads to thickening of the vessel walls and narrowing of the vascular lumen in the arteries that run through the white matter[291]. Consequently, this could lead to local ischemic events and thus lead to an infarction. Indeed, formation of WMLs is in part preceded by the onset of small subcortical infarcts[292]. Furthermore, lacunar infarcts are more likely to occur within or on the edge of existing WMLs[293]. Notably, the regions close to WMLs have also show structural and vascular changes[294], suggesting that WMLs are markers for more widespread regional pathological effects.

The associations may already exist to some extent before the age of 30 years[295]. However, the vast majority of our study population was 30 years or older, and thus we may have been underpowered to detect such an association. Furthermore, WMLs tend to be limited in volume across early and middle adulthood[3,58,296], so a longer follow-up time than 15 years is needed to test whether VRFs at a younger age relates to WML volume load and progression. Furthermore, the findings raise the question of whether intervening on BMI and high blood pressure can reduce the WML volume or whether they have a common cause. Longitudinal studies suggest that antihypertensive drugs reduce the progression of WML volume in elderly hypertensive patients[297-299]. Further work is needed to explore the implications of our findings, and the extent to which they can be intervened upon.

Much less is known about the association between blood pressure and fine motor dexterity as assessed by the PPB test. WMLs have previously been shown to affect dexterity[300-302], and so WMLs may mediate the association between blood pressure and the performance on the PPB test. An alternative explanation is that there may be residual confounding by physical exercise.

The right types of physical exercise reduce blood pressure over time[303,304], and physical exercise can also improve hand strength and hand dexterity[305,306]. However, no data was available on the levels of physical activity in the participants of our study. As such, further work is needed to confirm the association between blood pressure and the PPB, and to elucidate the underlying mechanism.

Several limitations should be taken into account when interpreting the findings. First, most of the participants had a Dutch ancestry, which limits the extent to which the findings can be generalized to other ethnicities. Second, an interval of 15 years may not have been optimal for the detection of the associations. For example, VRFs during early adulthood may be predictive of the development of microbleeds and infarcts at an older age, which can only be detected with a longer follow-up time. Third, the associations focused on lipid levels were only tested in women, which limited the sample size as well as the generalizability of the findings to men. Finally, no information was available on cognitive performance and neuroimaging measure at the time of the first visit, and thus we cannot account for any associations that were present at the baseline visit. This study also had several strengths. It was part of Generation R, which has been set up to collect a wide range of information on the participants. As such, we were able to properly account for any confounding variables that we expect may play a role. Furthermore, the data were collected prospectively, and thus the VRF measures are unlikely to suffer from information bias. Finally, our study was set up to examine a wide range of VRFs, cognitive measures and neuroimaging markers in a systematic manner.

In conclusion, blood pressure and BMI during early adulthood associate with higher WML load during midlife. Further work is needed to replicate and potentially extend these findings, and to elucidate their clinical implications.

RESERVE MECHANISMS

5.1

Childhood adversity and accelerated aging of the brain

Sander Lamballais, Scott Delaney, Ryan L. Muetzel, Tonya White, M. Arfan Ikram, Henning Tiemeier

ABSTRACT

Background

Adverse childhood experiences (ACEs), in particular of a threatening nature, accelerate biological aging and pubertal development during childhood. This may arise from an evolutionary drive to maximize survival and can manifest as earlier menarche or worse cardiovascular risk profile. However, within this framework it is unclear whether aging is accelerated or decelerated in the brain as assessed with neuroimaging.

Methods

The study was performed in 3,247 children from the Generation R Study, a prospective birth cohort based in the Netherlands. Threatening ACEs were quantified through parent-reported exposure to physical violence. Economic disadvantage was studied as a non-threatening ACE. Brain aging was determined using magnetic resonance imaging data from the children when they were aged between 8 and 13 years. We trained a brain age model in one half of the sample using gradient boosting and tested the association between the ACEs and brain aging in the other half. This splitting process was repeated 10,000 times.

Results

The predicted brain age correlated with the chronological age (Pearson's $r = 0.75$). Brain aging was mainly predicted by regions that are known to develop during puberty. Threatening ACEs were always positively associated with brain aging (mean $\beta = 0.065$), but the findings only reached statistical significance ($p < 0.05$) in 15.8% of the splits. Economic disadvantage did not seem to associate with accelerated or decelerated brain aging (statistical significance in 1.4% of the splits).

Conclusions

Threatening ACEs accelerate rather than decelerate brain aging during puberty, although further conclusive work is needed.

INTRODUCTION

Adverse childhood experiences (ACEs) have profound consequences for the physiological and psychological development of children and adolescents[307,308]. Examples of ACEs include physical and sexual abuse, physical and psychological neglect, and community crime and violence. Half of the adult population have experienced at least one ACE[309-311], though the numbers vary by country, study context and type of ACE. Furthermore, ACEs have different consequences for the victim, and different frameworks have been proposed to classify them. McLaughlin & Sheridan proposed to classify ACEs along two axes: threat and deprivation[312]. The treat axis describes the degree of threat that the event has to one's physical being, while deprivation denotes the absence of cognitive and social constructs. Threat and deprivation have unique contributions to the psychological and physical consequences of ACEs[313-316]. Socioeconomic status (SES) is tangential to threat and deprivation and may compound the negative psychological and physiological effects of ACEs.

ACEs have also been shown to accelerate aging. Generally, a distinction can be made between chronological aging – the number of calendar years – and biological aging. The latter describes the concept that body function changes with age, and someone's biological age is the age at which their body functions. Proxies for biological aging include telomere length, epigenetic expression and pubertal development assessments like Tanner stages[317]. All these markers associate with accelerated aging in children and adolescents exposed to threatening ACEs[316,318]. This is not the case for depriving ACEs[316,318]. Evolutionary, threat could accelerate aging because it signals an unsafe environment for the organism to survive in[319,320]. By accelerating maturation of the reproductive system[321], the organism achieves independence and can reproduce at an earlier age, thus maximizing survival.

The brain has also been proposed as a mode for understanding biological age of the brain[322,323]. Statistical models can be trained to predict someone's age from magnetic resonance imaging (MRI) brain data[322,323]. The difference between the chronological age and the brain age is denoted as the Brain Age Gap Estimate (BrainAGE). Given that other markers of biological aging associate with threatening ACEs, BrainAGE should as well. To our knowledge this has not been studied. Furthermore, from an evolutionary perspective it is unclear whether threat should accelerate or decelerate brain aging. Threat diverts energy towards systems that are important for survival, so if the brain is one of those systems, then it would lead to accelerated brain aging. However, the brain could also be one of the systems from which energy is diverted, which should slow down maturation of neural tissue. Finally, SES and depriving ACEs may also affect the degree at which the brain matures.

The aim of our study was to investigate whether ACEs and SES decelerate or accelerate brain aging. To do so, we utilized data from the Generation R Study, a large, prospective birth cohort. Generation R provides the perfect context to answer the questions at hand as information is available on ACEs, and neuroimaging is available for children between the ages of 8 and 15 years. We looked at economic disadvantage as a marker for both depriving ACEs and SES influences. Finally, to further validate the findings, we show that our model for brain aging relates to pubertal development in an external cohort: the Adolescent Brain Cognitive Development (ABCD) study.

METHODS

Study population

The study was performed as part of the Generation R Study, a prospective birth cohort based in Rotterdam, the Netherlands[22]. The aim of the cohort was to study the determinants of normal and abnormal fetal and childhood development. Pregnant women were recruited between 2002 and 2006, and 9,901 babies were born as part of the study. The study consisted of questionnaires and physical measures during and after pregnancy. We focused on children who participated in at least one of the neuroimaging research visits at the ages of 10 or 13 years (n = 4,479)[55]. Participants were further excluded if they had dental braces or if the quality of the images was too low (n = 735). For children who participated in both visits we opted to only use the data from the 13-year visit. We further excluded children who had no information on ACEs (n = 460) and for twin pairs we randomly excluded one of the two twins per pair (n = 37). The final sample consisted of 3,247 children with neuroimaging data and information on threat (n = 2,546) or economic disadvantage (n = 2,737). The full flow chart of the study population is shown in **Supplementary Figure 1**.

The study was conducted in accordance with the guidelines as set by the World Medical Association Declaration of Helsinki and was approved by the Medical Ethics Committee of the Erasmus MC (registration number MEC 02.1015). Written consent was obtained from the primary caregivers on behalf of the child, and additionally from the child if the measurements were performed at or after the age of 12 years to comply with national laws.

Physical violence

We focused on physical violence as an example of a high threat ACE. During the 10-year research phase the primary caregivers, which was in most cases the mother, were given questionnaires that contained two items from the Alabama Parenting Questionnaire (APQ)[324]. The two items were “you spank your child with your hand when he/she has done something wrong” and “you slap your child when he/she has done something wrong”. Participants had

to rate on a 5-point Likert scale – ranging from "never" to "always" – how often this typically occurred in the household.

A second source of information on physical violence was the standardized life events interview during the 10-year research center visit. The primary caregiver of the child was interviewed and ask questions about notable life events that may have occurred. The caregiver had to indicate to the best of their knowledge at which age this happened and whether the event had a lasting impact on the child. For the analyses within this paper, we focused on the dichotomous item "has anyone ever used physical violence against your child?".

From these items we created a dichotomized variable that coded the presence of slapping or spanking of the child as obtained from the APQ, or an encounter with physical violence that still influenced the child as obtained from the life events interview.

Economic disadvantage

We used economic disadvantage as a marker for a mix between SES and depriving ACEs. Economic disadvantage cannot be measured directly, and so three measures were combined as a proxy: [1] financial income, [2] financial difficulties, and [3] adjustments to the financial circumstances of the household[325]. Financial income at birth was obtained through questionnaires, and the data was dichotomized at the amount of social security around that time (€1,200.-). Financial difficulties were assessed through the question: "In the past year, have you experienced any difficulty in paying for food, rent, electricity bill and suchlike?", with "no", "some difficulty", and "great difficulty" as the response options. Adjustments to financial circumstances were addressed in 13 questions asking whether participants possessed certain items (e.g., a refrigerator or a phone) or performed certain actions (e.g., buying new clothes or going out), and whether this was due to financial reasons or other reasons. Economic disadvantage was coded as a dichotomous item, where individuals with an income below €1,200 and any experience of financial difficulties or adjustments due to financial reasons were coded as experiencing economic disadvantage.

Image acquisition & processing

Neuroimaging data from the 10-year and the 13-year research visit were used for this study[55]. Both visits were performed on the same 3T GE Discovery MR750w MRI System (General Electric, Milwaukee, WI, USA) with an 8-channel receive-only head coil. T_1-weighted images were collected with an inversion recovery-prepared fast spoiled gradient recalled sequence (T_R = 8.77 ms, T_E = 3.4 ms, T_I = 600 ms, flip angle = 10°, field of view = 220 x 220 mm, acquisition matrix = 220 x 220, slice thickness = 1 mm, number of slices = 230, bandwidth = 25 kHz).

The images were processed using the FreeSurfer analysis suite, version 6.0.0[88]. The procedure for image processing has been described previously[326]. The images were stripped of non-brain tissue, voxel intensities were normalized for B1 field inhomogeneity, and the brain tissue was segmented into gray matter, white matter, and cerebrospinal fluid. The segmented maps were further tessellated to reconstruct surface-based models of the cerebral cortex. To obtain information on the cortical thickness, surface area and volumes of cortical regions, the surface maps were registered to the Desikan-Killiany atlas[327]. Finally, volumes of the subcortical regions and the cerebellum were obtained.

BrainAGE calculation

There is no golden standard procedure to predict brain age from MR images[323,328]. We opted to use a gradient boosting model[329-331]. Boosting is a method that reduces bias in the prediction and has been successfully applied to brain aging and other machine learning questions. The idea behind boosting is to create a weak classifier that does not very accurately predict the outcome from the data, reweighing poorly classified data points to have more weight, and then repeating this process a number of iterations. In gradient boosting, the data points are replaced by the residuals of the prediction model, thus each sample is reweighted to the extent that the predicted value differs from the observed value. The classifiers from each round are then combined into a single model.

The model was trained to predict age based on the volumes of the subcortical regions and the cerebellum, and on the thickness, surface area and volume of the cerebral regions. The trained model could be applied in a new dataset to predict the age based on the same brain characteristics. The difference between the predicted brain age and the chronological age was used to define BrainAGE[322].

Validation with pubertal development in the ABCD study

The predictions made by BrainAGE models are assumed to be a proxy for the biological age of the brain. To check whether our model indeed predicted brain aging we compared the predicted ages with another marker of biological aging: pubertal development. This was done in the Adolescent Brain Cognitive Development (ABCD) study, a 21-site cohort study in the United States of America. The total sample consists of 11,875 children between the ages of 9 and 11 years.

To assess pubertal development, both the child and one of their parents separately completed the Pubertal Development Scale[332]. The scale contains five questions that each contain drawings of four different stages of pubertal development of the genitalia and other regions of the body. The person who fills in the questionnaire has to indicate for each question which stage the child is in, and the final score is calculated by taking the average of all answers. The

self-reported scale is not as accurate as physician observation for determining the exact pubertal stage, but it has been shown to distinguish between prepubertal and pubertal children in large epidemiological studies[333]. We only used the parent reported data as it has been shown to be more reliable in ABCD[334], and had less missing data.

The neuroimaging acquisition and processing of the ABCD-study have been described elsewhere[90,335]. In brief, T_1-weighted images were collected using a 3D inversion prepared RF-spoiled gradient echo scan. The images were corrected for gradient nonlinearity distortions and further processed with FreeSurfer version 5.3[88]. The other steps of the image processing match those described for Generation R.

We excluded children without T_1-weighted images or insufficient image quality (n = 485). We further excluded children without parent-reported data on the pubertal development scale (n = 776). Finally, we selected one of each multiple birth sets, i.e., twins and triplets (excluded n = 660). The final sample consisted of 8,912 children.

Covariates

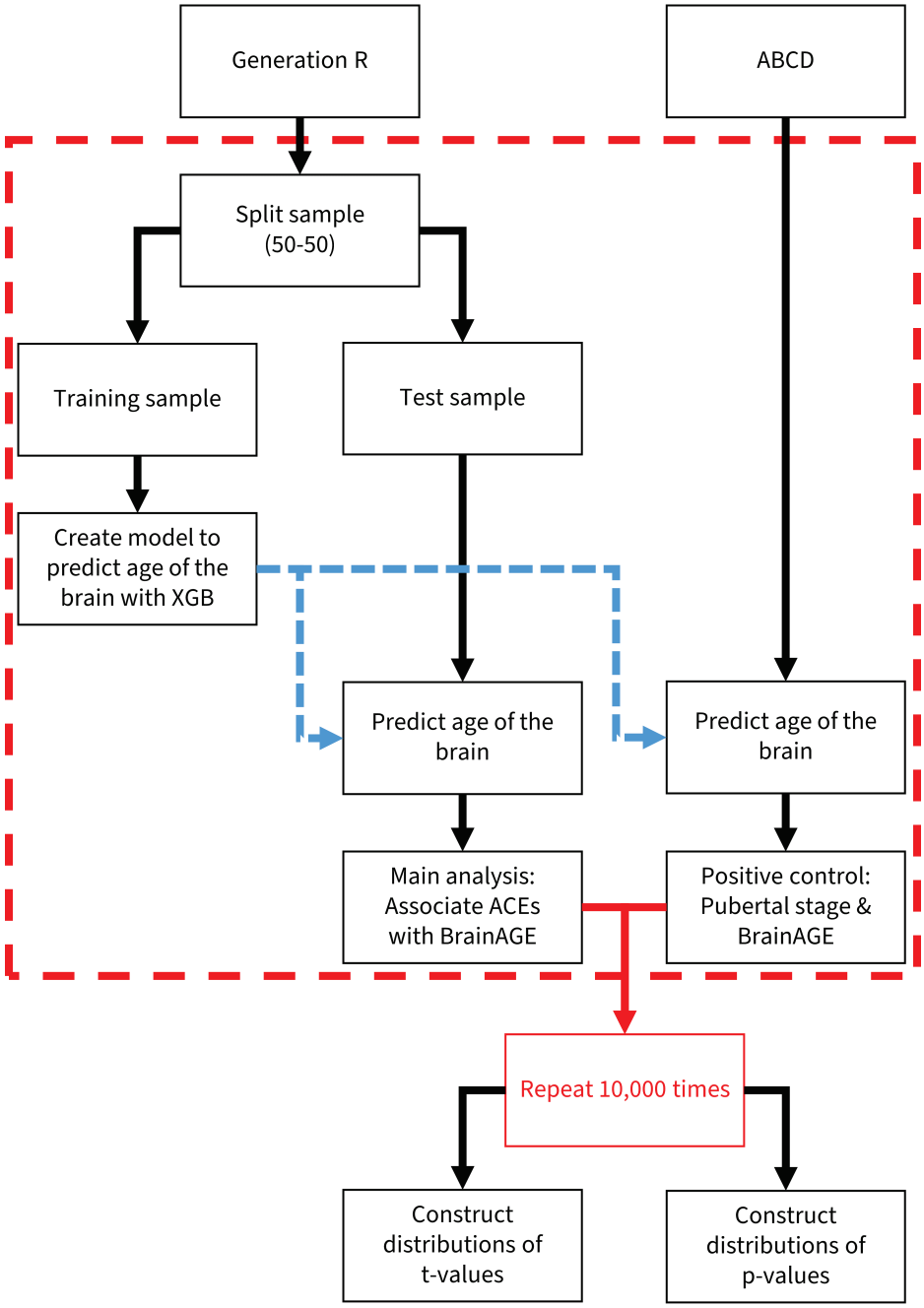
Date of birth and sex of the children were obtained from the medical records. Information on maternal education at birth of the child was obtained through questionnaires and was divided into three levels: low (no or only primary education), intermediate (high school or vocational education) and high (higher vocational education to university). Ethnicity of the child was based on the ethnicity of the parents, which was obtained through parent-report questionnaires.

Statistical analysis

An overview of the study design is given in **Figure 1**. We first split the Generation R sample into two halves: a training set and a test set. The training set was used to train a brain age model. To do so, we constructed a gradient boosting model with the *xgboost* R package[336]. The model used a linear booster and was trained for 1000 rounds. Lambda, alpha and maximum depth values were determined through five-fold cross validation. The other parameters of *xgboost* were set to their default values. The trained gradient boosting model was then used to predict the age of the brain of each individual in the test set. BrainAGE was defined as the difference between the observed and the predicted brain age. The model was evaluated in the test set by calculating the Pearson's r correlation coefficient, the mean absolute error (MAE) and the root square mean error (RSME).

A linear regression model was constructed in the test set to study whether threat and economic disadvantage associated with BrainAGE. The BrainAGE variable was standardized.

Figure 1 | Schematic representation of the study design.



Potential confounding was controlled for by adjusting for sex of the child, chronological age of the child during the MRI scanning, ethnicity of the child (Dutch / other Western / non-Western) and education level of the mother at birth of the child. We included chronological age as a covariate as it has been shown to improve the validity of models that study BrainAGE[337].

The results of the analyses depend on how the training and test set are generated. We therefore repeated the splitting of the sample, the generation of the BrainAGE model and the linear regression analyses a total of 10,000 times. Distributions of the t-values and the p-values of threat and economic disadvantage were then further examined.

To infer which brain variables were important for the prediction of brain aging, we extracted the relative importance of all variables for all 10,000 gradient boosting models. Finally, for each iteration we extracted the RMSE, the MAE and the Pearson's r of the chronological versus predicted age for the test set, as well as the p-values and regression coefficients for both threat and economic disadvantage.

To further inspect the validity of the brainAGE setup, we included a positive control condition, i.e., a condition where statistical significance is expected. We thus aimed to see whether BrainAGE was related to another measure of biological aging. This was done in the ABCD study, by examining the association between parent-reported pubertal development and brainAGE.

Information on maternal education was missing in 1.65% of the study population. These missing data were accounted for by multiple imputation with chained equations[101], with which 20 datasets were generated. The regression models were constructed for every dataset, and the coefficients were subsequently pooled according to Rubin's rules[105]. All statistical analyses were performed in R 3.6.1[95].

RESULTS

Study population characteristics

The characteristics of the 3,247 Generation R children are shown in **Table 1**. The scans were obtained at a mean age of 11.0 years (standard deviation (SD) = 1.7). Of the participants 50.6% were girls, 64.6% had a Dutch ethnicity and 55.3% had mothers with a high level of educational attainment. The excluded children were more likely to have a non-Western ethnicity (43.7% versus 26.5%, $p < .001$) and a low level of maternal education (27.7% versus 15.1%, $p < .001$) than the included children.

Table 1 | Characteristics of the Generation R Study population.

	Included sample N = 3,284	Excluded sample N = 1,195	Non-response analysis p-value
Sex, girl	50.6	50.6	0.707
Ethnicity			< 0.001
Dutch	64.6	48.6	
Other Western	8.9	7.7	
Non-Western	26.5	43.7	
Maternal education at birth of child, %			< 0.001
Low	15.1	27.7	
Intermediate	29.6	34.0	
High	55.3	38.3	
Age at MRI (years), mean (SD)	11.0 (1.7)	10.9 (1.7)	0.376

Overall, 30.0% of the children have experienced a threatening ACE during their life. Of those, 3.0% had someone use physical violence against them where the child was still affected by the event. In addition, for 27.2% of the caregivers indicated that they pinched or slapped their child when they did something wrong. During pregnancy, 10.3% of individuals had an income of lower than €1,200.- and reported financial difficulties or adjustments made due to financial reasons.

BrainAGE within Generation R

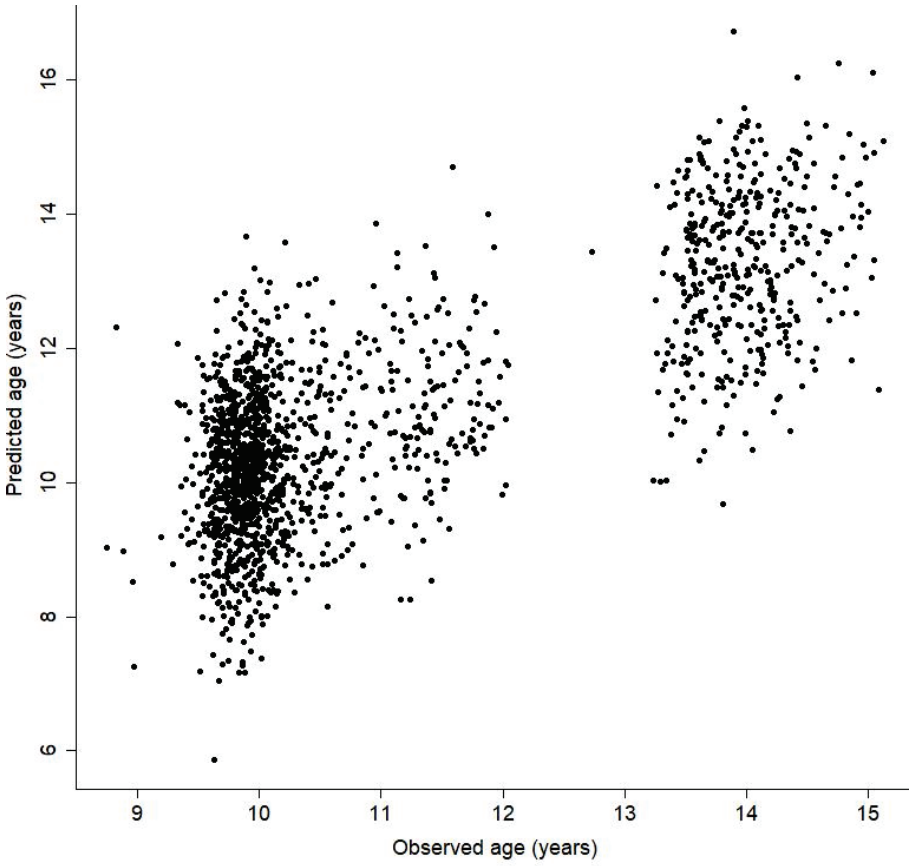
The full study design is shown in **Figure 1**. An example of a scatterplot of the observed and predicted age within the test set is shown in **Figure 2**. Across the 10,000 splits the average Pearson's r was 0.75 (SD: 0.007), the average MAE was 0.91 (SD: 0.01) and the average RMSE was 1.14 (SD: 0.02) (**Supplementary Figure 2**).

The 20 most important regions in the gradient boosting model are described in **Table 2**. The top regions were mostly located in the prefrontal cortex (inferior frontal gyrus, middle frontal gyrus, superior frontal gyrus, precentral and postcentral gyrus) and subcortically (accumbens, globus pallidus and hippocampus). Only the thickness of the right precentral gyrus was present in the top 20 in all 10,000 models. More importantly, the brain aging prediction was not driven by a single region, but rather was based on a myriad of regions across the brain.

Association between ACEs and BrainAGE

The results for the ACEs and BrainAGE associations across all 10,000 splits are displayed in **Figure 3**. Overall, threat generally had a positive association with BrainAGE with a mean coefficient of 0.065 (SD: 0.039), with 15.8% of the runs having a p-value below 0.05. On the

Figure 2 | Example of BrainAGE within the Generation R Study. The dataset is derived from two separate research visits, which also explains the gap in the data. The scatterplot is based on one of the 10,000 splits that were performed.

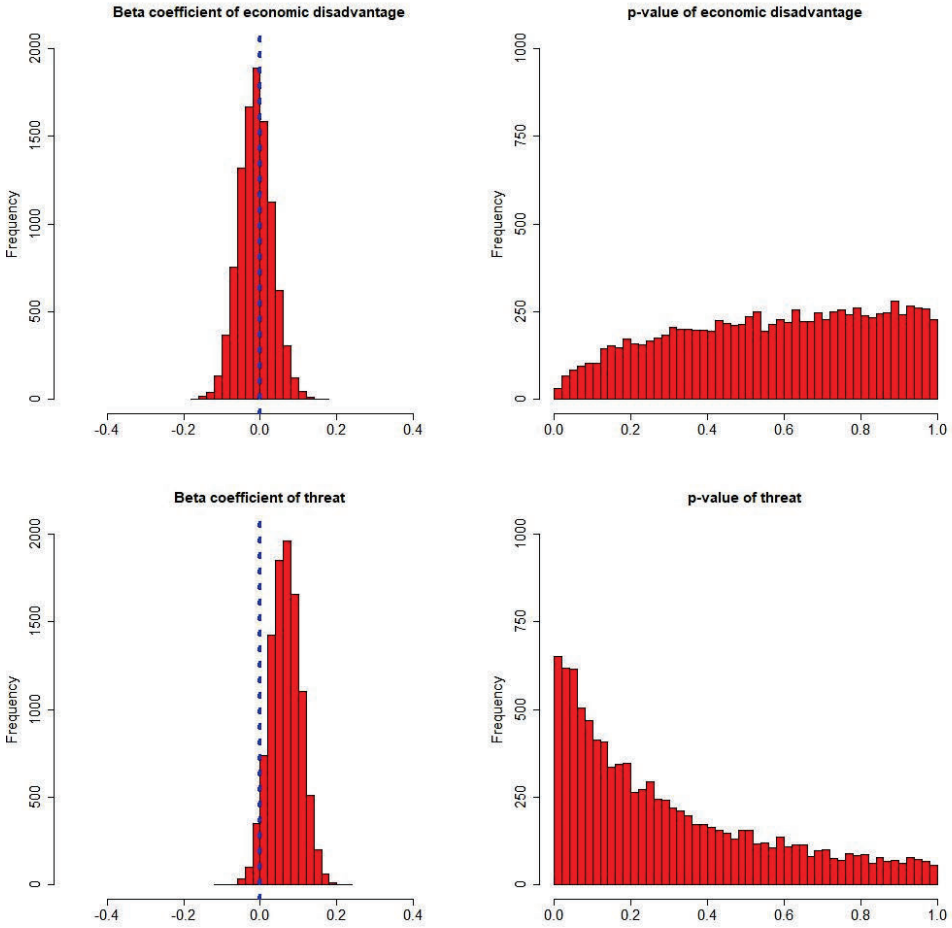


other hand, the average coefficient for economic disadvantage was -0.012 (SD: 0.043), which only reached statistical significance in 1.4% of the 10,000 runs.

Pubertal development and BrainAGE in the ABCD study

The scans from the ABCD study were obtained around a mean age of 9.9 years (SD: 0.6) and the sample consisted for 47.5% of girls. When the Generation R trained BrainAGE models were applied to the ABCD cohort, puberty versus prepuberty was associated with a higher brainAGE.

Figure 3 | Histograms of the regression coefficients and p-values for the association between ACEs and BrainAGE across the 10,000 splits.



DISCUSSION

The findings support the notion that threatening ACEs associate with accelerated biological aging of the brain. Furthermore, no evidence was found to support an association between economic deprivation and accelerated or decelerated brain aging.

The findings of this study build upon the existing literature[313-316]. Previous studies were mostly performed in populations with an enrichment of ACEs, with high rates of abuse, experienced or witnessed structural violence, and psychological and physical deprivation. Our study sample was drawn from the general population, and the experienced threatening ACEs were

Table 2 | Top 20 variables that were important for the brain aging prediction across the 10,000 splits.

	Region	Side	Metric	Importance (mean ± SD)	Present in top 20 (%)
1	Precentral gyrus	Right	Thickness	0.17 ± 0.02	100.00
2	Lateral orbitofrontal gyrus	Left	Area	-0.15 ± 0.03	98.56
3	Rostral middle frontal gyrus	Left	Thickness	-0.15 ± 0.02	99.59
4	Lateral orbitofrontal gyrus	Left	Thickness	-0.14 ± 0.02	99.42
5	Postcentral gyrus	Left	Volume	-0.13 ± 0.02	96.30
6	Accumbens area	Right	Volume	-0.13 ± 0.02	96.56
7	Pars orbitalis	Right	Area	0.12 ± 0.03	89.34
8	Pars opercularis	Right	Area	0.12 ± 0.02	88.72
9	Superior frontal gyrus	Right	Volume	0.12 ± 0.02	85.00
10	Caudal anterior cingulate	Left	Area	0.11 ± 0.03	68.12
11	Postcentral gyrus	Left	Thickness	0.11 ± 0.02	75.15
12	Precentral gyrus	Left	Area	0.10 ± 0.02	61.61
13	Caudal anterior cingulate	Left	Volume	-0.10 ± 0.03	58.08
14	Globus Pallidus	Right	Volume	0.10 ± 0.02	59.38
15	Pars orbitalis	Right	Volume	-0.10 ± 0.02	54.88
16	Superior parietal gyrus	Right	Thickness	-0.10 ± 0.02	45.57
17	Superior frontal gyrus	Left	Volume	0.09 ± 0.02	45.35
18	Lingual gyrus	Right	Thickness	-0.09 ± 0.02	51.63
19	Superior frontal gyrus	Right	Area	-0.09 ± 0.02	39.80
20	Hippocampus	Right	Volume	0.09 ± 0.02	51.63

SD = Standard deviation

generally of a lower severity than in the previous studies. Thus, even exposure to low levels of threatening ACEs seems to lead to accelerated aging of the body and of the brain.

The brain aging models mostly relied on regions that have previously been identified to change during childhood and adolescence[338]. A total of 11 out of the 20 most important metrics were derived from regions within the frontal cortex, and another 2 out of 20 measures were from the postcentral gyrus, a region directly bordering the frontal cortex. The relative importance of these metrics matches the previous literature on childhood brain development[338-340], and further validates the model that was used. Furthermore, the findings suggests that threatening ACEs influence the biological aging as a whole rather than just regions that are related to emotional processing and emotional expression, as has been suggested previously[341].

While we found evidence to support an association between threatening ACEs and accelerated brain aging, it is less clear what such an acceleration entails. At least two different mechanisms could have led to these findings. The first mechanism assumes that children follow a fixed trajectory of brain development. If this is the case, then brain aging would be accelerated by simply moving more quickly along this trajectory. A second mechanism would be that certain phases of brain development are simply shortened or even skipped. During early childhood, the microstructure of the white matter greatly changes due to pruning of connections and refinement of neural networks. Threatening ACEs could simply trigger the onset of the next phase in development before such refinement has finalized. In other words, if the threatening ACEs had not occurred in an individual's life, then their brains may simply spend more time refining all the networks and connections before jumping into a next phase.

The hypothesis that threat triggers the next phase of development is supported by the finding that threatening ACEs also accelerate the onset of pubertal development[316,318]. The transition to puberty has been association with sudden changes to the macro and microstructure of the brain[342]. Thus, pubertal development may be the mediating mechanism underlying threatening ACEs and brain maturation. Further work is needed to take into account all three factors – threatening ACEs, pubertal development, and brain aging – into a s single study.

Our study design suffered from several limitations. First, the assessment of threatening ACEs was very limited in scope. No questions on physical and sexual abuse were asked. These questions were intentionally not included, as the Generation R Study is a cohort focused on all aspects of health in the general population. The questions were deemed too extreme for the context in which they were assessed. Furthermore, the questions would have been parent-reported rather than child-reported, which could lead to underestimating of the prevalence of abuse and consequently information bias. A second limitation of the study was the narrow age range on which the brain aging models were build. With most children being between 9 and 10 years old during the neuroimaging visit, the model had little variance in the age distribution to be trained with. Still, we showed that the predicted brain age was related to another marker of biological aging in an external cohort, thus validating the brain age metric in this study. A third limitation was the lack of an external cohort to train the brain age model in. The model was built and tested in the Generation R Study, while ideally the two steps are done in separate cohorts. The 10,000 splits approach was introduced to assess the volatility of the findings, and consequently the findings are probabilistic in nature, i.e., statistical significance in a given fraction of the splits. Building and testing the model in separate cohorts would have allowed us to perform a single test to assess the relation between ACEs and brain aging. The study also has several strengths. First and foremost, the large sample size allowed for proper specification and estimation of a brain aging model. Second, all samples were obtained on the

same scanner, thus avoiding any scanner-related variability which could introduce bias in the brain aging model. Third, the brain aging model was validated in an external cohort with a different marker of biological aging. Finally, the measures for economic disadvantage were obtained prospectively, which reduced any risk of recall bias.

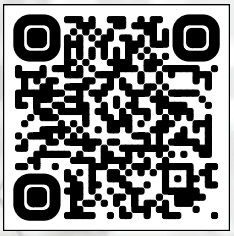
In conclusion, threatening ACEs accelerate aging of the adolescent brain, while we found no support for the idea that depriving ACEs and SES affect brain aging. It remains to be elucidated what underlying mechanism links threat to accelerated aging, and whether this mechanism can be intervened on to avoid early maturation of the brain.

5.2

Cortical gyrification in relation to age and cognition in older adults

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

Adapted from:



Supplement:

ABSTRACT

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 T_1-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.

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[343], evolutionarily conserved, and similar amongst closely related animal species[344]. Abnormalities in gyrification, such as polymicrogyria and pachygyria, 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[345,346], schizophrenia[347,348], and bipolar disorder[348]. 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[348-351]. 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[348-350]. Regional patterns of gyrification can be quantified with the Local Gyrification Index (LGI)[352]. The regions surrounding the angular gyrus, i.e. the parietal cortex, seem most prone to age-related decline in the LGI[351]. 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[350]. 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 to 65 years of age, and late adulthood, i.e., beyond 65 years of age. Other aspects of the cerebral cortex – such as cortical surface area – change significantly during middle and late adulthood[3]. Furthermore, atrophy of the cerebral cortex seems to accelerates towards the end of life[353], and the rates of atrophy differ between brain regions[3]. 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.

METHODS

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[354]. 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 to 6 years. The study has included 14,926 participants 45 years of age and older. Neuroimaging was introduced in 2005[13]. 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 = 6,647). 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 4,397 participants. A flow chart of the study population is shown in **Supplementary Figure 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.

Assessment of cognitive function

All participants underwent a cognitive test battery[61]. 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[355]. 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 minutes 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[356], 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)[66], 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[357]. Participants had to name as many animal species in a span of one minute, with the total number of unique species as the outcome. Finally, we administered the Purdue pegboard test (PPB)[71], 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[358]. The g-factor explained 50.6% of the variance amongst the cognitive tests which is in agreement with previous literature[280].

Image acquisition

Neuroimaging was performed on a 1.5T magnetic resonance imaging (MRI) scanner with an eight-channel head coil (GE Signa Excite, General Electric Healthcare, Milwaukee, USA). The imaging sequences have been described extensively elsewhere[13]. Axial T_1-weighted images were collected using a 3D Spoiled Gradient Recalled sequence (T_R = 13.8 ms, T_E = 2.8 ms, T_I = 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[56].

Image processing

Images were processed using the FreeSurfer analysis suite (version 6.0)[88]. The standard reconstruction was conducted, where non-brain tissue was removed, voxel intensities were corrected for B_1 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[352,359], and each vertex was automatically assigned an anatomical label according to a predefined atlas[327]. All measures were co-registered to a standard stereotaxis 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 T_1-weighted scan that assesses artifacts related to motion[57]. 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 "fail" if FreeSurfer failed 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.

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 minutes 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[240]. BMI was calculating using the height and weight obtained during the research center visit. Smoking status was obtained during home interviews and individuals were classified as never smokers, past smokers, or current smokers. Education level was assessed during the home 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).

Statistical analyses

All statistical analyses were performed in R 3.4.3[95]. 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[97]. 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[98]. 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[98]. 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[360]. 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 confounding. 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 **Supplementary 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[359]. However, the cortical surface area itself has also been shown to relate to cognitive function[361]. 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%). To maximize power, missing covariate data were imputed thirty times using multiple imputation by chained

equations[101]. Imputed models were subsequently pooled per vertex according to Rubin's rules[105]. 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.

Table 1 | **Baseline characteristics of the study population**. The excluded sample (n = 2,250) were all participants that were eligible for cognitive testing and the neuroimaging study but did not end up in the final sample (see **Supplementary Figure 1**).

Characteristics	**Included sample N = 4,397**	**Excluded sample N = 2,250**	**p-value[b]**
Age at MRI (years)	63.5 ± 10.1	69.5 ± 1.13	< .001
G-factor	0.00 ± 1.00		
Cohort (%)			< .001
RS-I	15.3	33.9	
RS-II	25.3	26.7	
RS-III	59.4	39.5	
Sex, female (%)	55.3	57.8	.278
Time between cognition and MRI (years)	0.3 ± 0.4	0.2 ± 0.3	.847
Hypertensive (%)[a]	61.4	76.1	.053
Alcohol per day (grams)[a]	9.2 ± 10.1	8.5 ± 9.6	<.001
Body mass index (kg/m^2)[a]	27.4 ± 4.1	28.1 ± 4.8	.003
Smoking status (%)[a]			.406
Never	30.9	29.9	
Past	49.1	49.4	
Current	20.0	20.7	
Education level (%)[a]			.002
Primary	7.8	13.1	
Low	37.7	41.0	
Intermediate	30.3	28.4	
High	24.2	17.5	
Mean GI	2.55 ± 0.08	2.52 ± 0.09	< .001

GI = gyrification index; MRI = magnetic resonance imaging.

[a] = Missingness of data for all variables was below 1% except for alcohol consumption (4.8%).

[b] = Differences between inclusion and exclusion were tested through multiple logistic regression.

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 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[362]. Any reported result that is stated as statistically significant will have a p-value below the threshold of 0.05.

RESULTS

Baseline characteristics of the study population (n = 4,397) 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 Figure 1**). Excluded participants tended to be older (mean = 6.2 years), were more often from the first cohort of the Rotterdam Study (33.9% versus 15.3%), were less likely to drink alcohol (mean = -0.7 grams per day), had a higher BMI (mean = 0.7 kg/m^2), were more likely to have only primary education (13.1% versus 7.8%) and had a lower GI (mean = -0.03).

Age and global gyrification

A scatterplot of age and the GI is shown in **Figure 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 fit sufficiently describes the association of age during mid and late adulthood with the GI.

Cognition and global gyrification

A scatterplot of the g-factor and the GI is shown in **Figure 1B**, and regression coefficients for the g-factor and all separate cognitive tests are shown in **Table 3** for all three adjustment models. Higher levels of the g-factor were associated with a higher GI, with similar results

Figure 1 | Scatterplot of age (A) and cognition (B) with GI. For age, four models were plotted (linear, quadratic, and the two and three spline models). The lines are not fully visible due to the extensive overlap. The plot for cognition only shows the linear model.

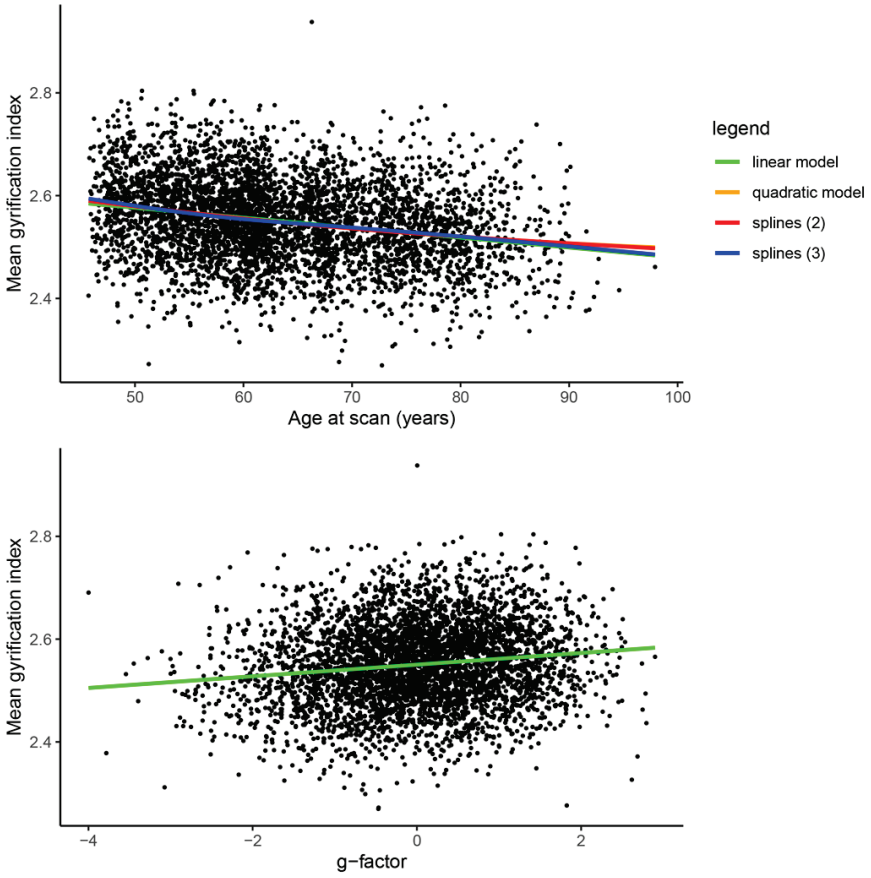


Table 2 | The associations between age and GI.

Model[a]	Type	β	95% CI	AIC	BIC
Linear	Linear	-0.0021	-0.0025; -0.0017	-10252.60	-10214.27
Quadratic	1st polynomial	-0.0053	-0.0081; -0.0024	-10255.47	-10210.75
	2nd polynomial	-2.5e-5	-3.0e-5; -4.6e-5		
Spline (2)	1st spline	-0.1235	-0.2376; -0.1020	-10256.33	-10211.61
	2nd spline	-0.0703	-0.0947; -0.0458		
Spline (3)	1st spline	-0.0552	-0.0384; -0.0718	-10203.46	-10254.57
	2nd spline	-0.1179	-0.0915; -0.1442		
	3rd spline	-0.0826	-0.0550; -0.1102		

AIC = Akaike information criterion; BIC = Bayesian information criterion; CI = confidence interval; GI = gyrification index.

[a] The models were adjusted for study cohort and sex.

Table 3 | The associations between cognition and the GI.

Domain	Model 1[a]		Model 2[b]		Model 3[c]	
	β	95% CI	β	95% CI	β	95% CI
g-factor	0.0045	0.0018; 0.0073	0.0047	0.0019; 0.0075	0.0044	0.0015; 0.0073
15WLT	-0.0001	-0.0009; 0.0007	-0.0001	-0.0010; 0.0007	-0.0003	-0.0011; 0.0006
Stroop[d]	-0.0090	-0.0171; -0.0010	-0.0090	-0.0172; -0.0010	-0.0081	-0.0164; -0.0002
LDST	0.0005	0.0001; 0.0009	0.0005	0.0002; 0.0009	0.0005	0.0001; 0.0009
WFT	0.0009	0.0005; 0.0013	0.0009	0.0005; 0.0013	0.0009	0.0005; 0.0013
PPB	0.0001	-0.0004; 0.0007	0.0002	-0.0004; 0.0007	0.0002	-0.0004; 0.0007

CI = confidence interval; LDST = letter digit substitution test; PPB = Purdue pegboard test; WFT = word fluency test; WLT = 15-word learning test.

[a] = Adjusted for age at MRI scan (years), study cohort, sex and age difference between cognitive testing and MRI scan (years).

[b] = Additionally adjusted for hypertension (yes/no), alcohol intake (grams per day), BMI and smoking status (never/past/current).

[c] = Additionally adjusted for education level (primary/low/intermediate/high).

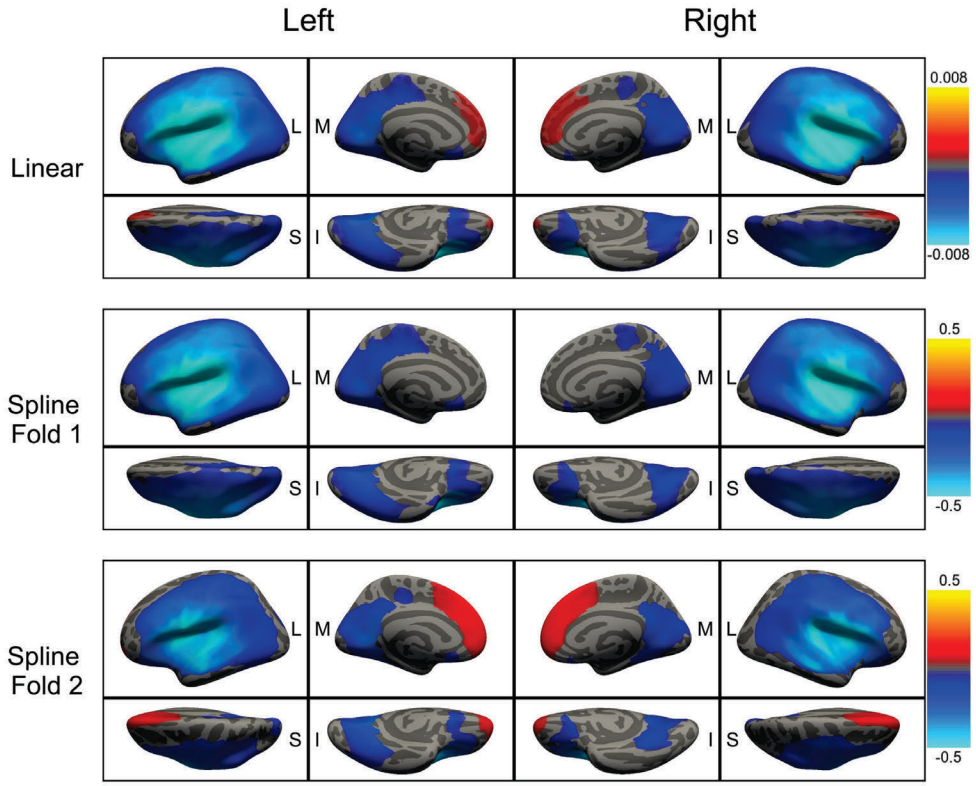
[d] = The Stroop task is inversely coded compared to the other tests, where a higher score relates worse performance.

across Model 1 (β = 0.0045, 95% CI = 0.0018; 0.0073) to Model 3 (β = 0.0044, 95% CI = 0.0015; 0.0073). We examined the individual cognitive tests to see which cognitive tests drove most of the association. Three cognitive tests yielded statistically significant results, namely the LDST (β = 0.0005, 95% CI = 0.0001; 0.0009), the WFT (β = 0.0009, 95% CI = 0.0005; 0.0013) and the Stroop task (β = -0.0081, 95% CI = -0.0164; -0.0002). Of these, the Stroop task had the strongest association with gyrification. Of note is that the association with the Stroop task was negative due to the lower scores on the Stroop task reflecting higher cognitive performance. Finally, the interaction term between age and cognition did not reach statistical significance in any of the models (all $p_{unadjusted} > 0.05$), thus the magnitude of the association between cognition and the GI was stable during and after midlife.

Age and local gyrification

To determine the precise spatial extent of associations between age and gyrification, we performed surface-based vertex-wise analyses. Due to the similar fits between the models of age and the GI we opted to further investigate the linear model and the two-fold spline model with the LGI. **Figure 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

Figure 2 | Vertex-wise associations between age – the linear model and the two-fold spline model – and the local gyrification index (LGI). The color scale represents the regression coefficients. The models were adjusted for study cohort and sex. L = Lateral; M = Medial; S = Superior; I = Inferior.

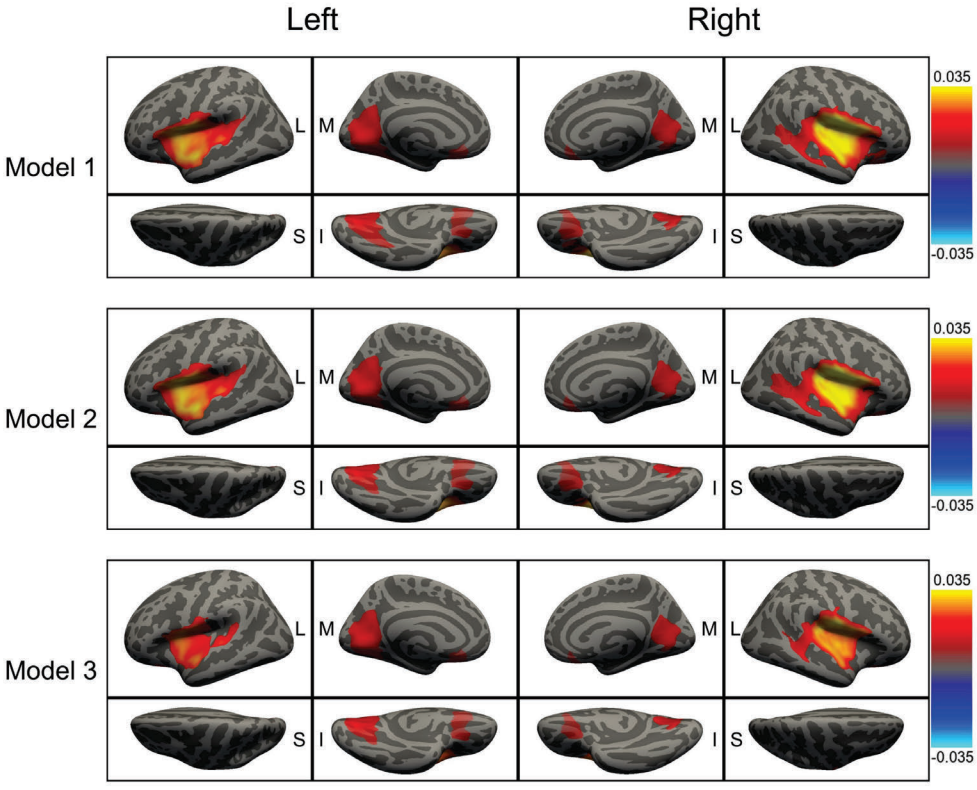


hemispheres in both size and strength. The two-degree spline model differed from the linear model. The first spline fold, i.e., ages between 45.7 and 61.6 years, associated negatively with the LGI in the parietal, frontal, temporal, and occipital regions. Unlike the linear model, no cluster was present 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 Figure 2**). The findings were robust upon further correction for the image quality metric (**Supplementary Figure 3**).

Cognition and local gyrification

Figure 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 gyrus. These clusters roughly presented bilaterally. Similar patterns were found for the associations between the individual cognitive tests and the LGI (**Figure 4**). The LGI of the cuneate gyrus, insular cortex and superior temporal gyrus were all associated with the Stroop task, the LDST

Figure 3 | Vertex-wise associations between the g-factor and the LGI. The color scale represents the regression coefficients. Model 1 was adjusted for age (linear term), study cohort, sex and the time difference between the cognitive test battery and the MRI visit. Model 2 was additionally adjusted for hypertension status, alcohol intake, BMI, and smoking. Model 3 was additionally adjusted for education level. L = Lateral; M = Medial; S = Superior; I = Inferior.



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 gyrus on the right hemisphere. Further correction for the image quality metric did not affect these findings (**Supplementary Figure 4**). Finally, as cognition was originally specified as the determinant, we constructed cluster-wise models with the cluster-wise mean gyrification as the determinant and cognition as the outcome. For the identified LGI clusters the associations with cognition remained unattenuated (all clusters $p < .00001$).

Figure 4 | Associations between the individual cognitive tests and the LGI. The color scale represents the regression coefficients. The images show the results for Adjustment Model 3. The Stroop task is inversely coded compared to the other tests, where a higher score relates worse performance. No statistically significant clusters were identified for the 15-word learning test or the Purdue pegboard test, thus these are not displayed. LDST = Letter digit substitution test; WFT = Word fluency test. L = Lateral; M = Medial; S = Superior; I = Inferior.

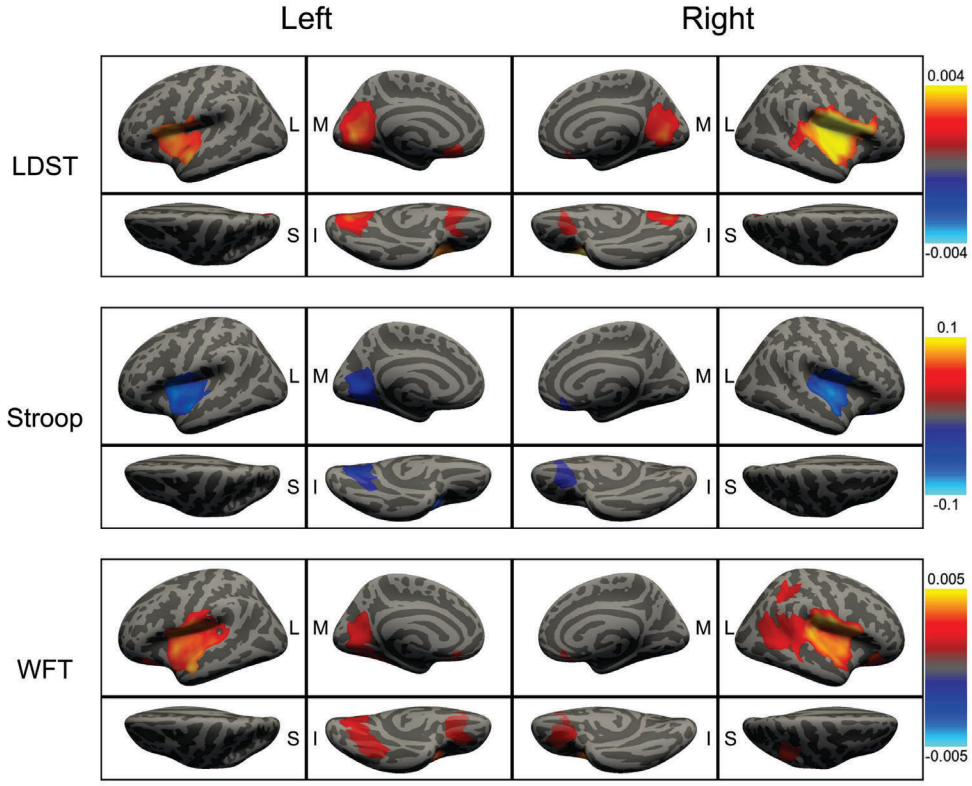


Table 4 | The associations of the mean LGI and mean surface area with g-factor per previously identified LGI cluster. Each model contained both LGI and surface area. Both the LGI and surface area were standardized to be able to compare the magnitude of their effects. All models were adjusted for age 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/m^2), smoking status (never/past/current), and education level (primary/low/intermediate/high).

Hemisphere	#	Location	Standardized mean LGI			standardized mean surface area		
			β	95% CI	p	β	95% CI	p
Left	1	Temporal	0.025	[-0.006; 0.055]	.115	0.056	[0.024; 0.088]	.001
Left	2	Cuneus	0.047	[0.016; 0.077]	.002	0.025	[-0.005; 0.056]	.106
Left	3	Orbitofrontal	0.025	[-0.001; 0.050]	.056	0.067	[0.038; 0.096]	< .001
Right	1	Temporal	0.045	[0.014; 0.076]	.004	0.041	[0.009; 0.073]	.012
Right	2	Cuneus	0.028	[-0.005; 0.073]	.094	0.029	[-0.004; 0.063]	.085
Right	3	Orbitofrontal	0.030	[0.005; 0.054]	.019	0.061	[0.034; 0.089]	< .001

Both vertex-wise cortical surface area and thickness associate with age (**Supplementary Figure 5**) and the g-factor (**Supplementary Figure 6**). To assess whether the associations between cognition and the LGI were driven by cortical surface area or thickness, we created a new model for each significant LGI cluster with the g-factor as the outcome and both the mean LGI and the mean cortical thickness or surface area as determinants. Associations remaining after concurrent adjustment for cortical surface area or thickness suggest an independence between LGI the other measures. The results are shown in **Table 4** for surface area and in **Table 5** for cortical thickness. After adjustment for surface area, the LGI remained associated with the g-factor in the left hemisphere in the cluster near the cuneus ($p_{unadjusted}$ = .002) but not the clusters in the orbitofrontal cortex ($p_{unadjusted}$ = .056) or the temporal cortex ($p_{unadjusted}$ = .115). In the right hemisphere the association between the LGI and cognition was unaffected by surface area in both the orbitofrontal ($p_{unadjusted}$ = .019) and the temporal clusters ($p_{unadjusted}$ = .004), but not in the cluster in the cuneus ($p_{unadjusted}$ = .094). In all clusters, the LGI was unaffected by additional corrections for cortical thickness.

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[348]. 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[351]. 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[348]. We also established non-linear patterns in the regional surface-based patterns,

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 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/m^2), smoking status (never/past/current), and education level (primary/low/intermediate/high).

Hemisphere	#	Location	Standardized mean LGI			standardized mean thickness		
			β	95% CI	p	β	95% CI	p
Left	1	Temporal	0.063	[0.038; 0.088]	< .001	0.062	[0.039; 0.085]	< .001
Left	2	Cuneus	0.069	[0.045; 0.094]	< .001	0.045	[0.022; 0.068]	< .001
Left	3	Orbitofrontal	0.044	[0.020; 0.068]	< .001	-0.025	[-0.049; -0.000]	.047
Right	1	Temporal	0.075	[0.050; 0.100]	< .001	0.061	[0.037; 0.085]	< .001
Right	2	Cuneus	0.058	[0.034; 0.083]	< .001	0.050	[0.027; 0.073]	< .001
Right	3	Orbitofrontal	0.049	[0.026; 0.073]	< .001	-0.003	[-0.027; 0.021]	.790

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[363-365], yet any such contributions from gyrification remain to be elucidated.

The temporal lobe has previously been linked to cognitive processes such as language[366] and memory[367] as well as psychiatric disorders like adulthood autism spectrum disorders[368] and schizophrenia[369-371]. Interestingly, these disorders have also been linked to abnormal gyrification[345-348]. Genetic mechanisms underlying cognitive processes and neuropsychiatric disorders may also affect cortical morphology in the temporal region, and in particular gyrification. Previous studies have found links between genes underlying cognition function and temporal lobe structure[372,373], although the results are inconsistent[374]. Thus, further work is needed to elucidate the presence of a genetic pleiotropic link between gyrification and function of the temporal lobe.

Several mechanisms could explain how gyrification changes with age. One explanation is that during brain development the cortical surface buckles due to differential rates of growth of cortical layers[375], and the opposite may occur during adulthood. The rate of atrophy is higher in gray than white matter during early and mid adulthood[376,377]. Gray matter atrophy is mostly through the reduction of surface area of the cortex, which leads to more shallow sulci and consequently a lower GI. The rate of atrophy of white matter starts to exceed the rate for gray matter during late adulthood[3], which could in turn lead to an increase in gyrification with age. Interestingly, a previous study found that after the age of 60 years the cingulate cortex thickens and that the rate of thinning of the medial prefrontal cortex declines[378], which could explain the increased gyrification of the medial prefrontal cortex in our study. Indeed, we see a similar thinning of the cingulate cortex. However, our findings also suggest that gyrification overall

keeps decreasing during older adulthood, thus other mechanisms than the different gray and white matter atrophy rates are also likely involved.

Another plausible explanation is the "axon tension" theory[379], which states that axonal tension pulls the gyral walls inwards, thus folding the cortex. The axonal tension may depend on the health status of the axon, and damage to axons could lead to reduced tension and consequently decreased gyrification. White matter microstructure decreases with age[380,381] and white matter lesions accumulate from mid adulthood onwards[3], and could explain the decrease in gyrification. However, further experimental work has discredited axonal tension as a cause of cortical folding. For example, if axonal tension causes gyrification then cutting the gyrus transaxially should unfold the gyrus, and experiments have shown that this is not the case[382]. Thus, further work is needed to elucidate the causes of cortical (un)folding during adulthood.

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[383]. 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[384]. 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, 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[385]. Second, changes in gyrification likely cause changes in cognition, but the models were specified with cognition as the determinant and gyrification 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 gyrification. 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[386-388], 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 unravel 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.

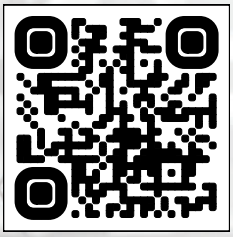
In conclusion, gyrification globally decreases linearly with age across the entirety of adulthood, and gyrification in the medial prefrontal cortex increases towards the end of life. Furthermore, gyrification increases with higher levels of cognitive performance in some clusters irrespective of surface 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.

5.3

The risk of dementia in relation to cognitive and brain reserve

Sander Lamballais, Jendé L. Zijlmans, Meike W. Vernooij, M. Kamran Ikram, Annemarie I. Luik, M. Arfan Ikram

Adapted from:



ABSTRACT

Background

Individual differences in the risk to develop dementia remain poorly understood. These differences may partly be explained through reserve, which is the ability to buffer cognitive decline due to neuropathology and age.

Objective

To determine how much early and late-life cognitive reserve (CR) and brain reserve (BR) contribute to the risk of dementia.

Methods

4,112 dementia-free participants (mean age = 66.3 years) from the Rotterdam Study were followed up for on average 6.0 years. Early-life CR and BR were defined as attained education and intracranial volume, respectively. Late-life CR was derived through variance decomposition based on cognition. Late-life BR was set as the total non-lesioned brain volume divided by intracranial volume.

Results

Higher early-life CR (hazard ratio = 0.48, 95% CI = [0.21; 1.06]) but not early-life BR associated with a lower risk of incident dementia. Higher late-life CR (hazard ratio = 0.57, 95% CI = [0.48; 0.68]) and late-life BR (hazard ratio = 0.54, 95% CI = [0.43; 0.68]) also showed lower levels of dementia. Combining all proxies into one model attenuated the association between early-life CR and dementia (hazard ratio = 0.56, 95% CI = [0.25; 1.25]) whereas the other associations were unaffected. These findings were stable upon stratification for sex, age, and APOE ε4. Finally, high levels of late-life CR and BR provided additive protection against dementia.

Conclusion

The findings illustrate the importance of late-life over early-life reserve in understanding the risk of dementia and show the need to study CR and BR conjointly.

INTRODUCTION

Susceptibility to develop dementia varies greatly across individuals and is thought to be influenced by both early and late-life factors[6,30-32,39,389-393]. These factors shape the risk of dementia through a mechanism called reserve. Cognitive reserve (CR) buffers the effects of dementia-related pathology through cognitive flexibility or recruitment of alternative neural networks while brain reserve (BR) buffers pathology through the structure of the brain[20,48]. CR and BR are generally estimated through proxies. Educational attainment is the most commonly used proxy for the maximum attained CR throughout early-life[31,32]. BR is commonly estimated through intracranial volume (ICV) as it a strong proxy for maximal brain volume early in life[20,394]. However, reserve changes as people grow older, and these fluctuations cannot be captured with educational attainment and ICV. Recently, a method was proposed to quantify CR in late-life from cross-sectional data[395,396]. It uses structural equation modeling based on demographics, accumulated white matter lesions as a measure of neuropathology and cognitive functioning. BR in late-life can be estimated through brain volume[20,397].

Against this background, several knowledge gaps need to be addressed. First, both early-life and late-life proxies of reserve have been associated with onset of dementia. However, given that reserve builds up and diminishes with age, it remains unclear whether early-life or late-life reserve proxies are more relevant to study for the risk of developing dementia. Second, even though CR and BR are conceptually related, they have generally been studied in isolation. It is unclear whether they independently protect against dementia or if one effect is confounded by the other. Finally, most studies on CR and BR focused on patients with memory complaints, mild cognitive impairment or dementia at baseline[398]. The extent to which these associations hold for dementia-free individuals in the general population is unclear.

The present study aimed to elucidate the relative importance of early-life and late-life CR and BR in the risk of developing dementia. The study was performed as part of the Rotterdam Study, a population-based prospective cohort.

MATERIALS AND METHODS

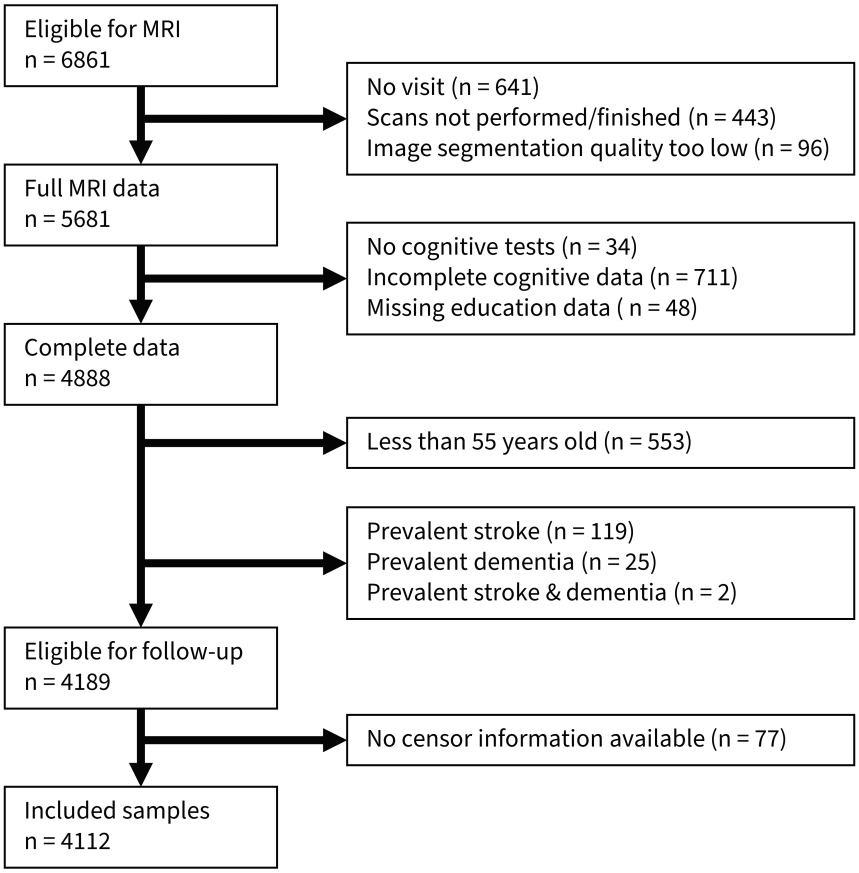
Study population

The Rotterdam Study is a prospective population-based cohort in the Ommoord district of Rotterdam, the Netherlands[354]. The study started in 1989 with the RS-I cohort and encompassed 7,983 participants aged 55 years and older. In 2000, the study was expanded with the RS-II cohort, which consisted of an additional 3,011 individuals aged 55 years and older. In 2006, the RS-III cohort started and included 3,932 participants aged 45 years and

older. At study entry and subsequent follow-up visits, the participants partook in a home visit as well as one or several visits to a dedicated research center. MR neuroimaging was included into the core protocol from 2005 onwards. For the participants included in the current analysis, we used their first MRI exam between 2005 and 2014 as the baseline[13]. Cognitive tests closest to the MRI with respect to date of assessment were used.

A flow chart of the inclusion process is presented in **Figure 1**. A total of 4,888 participants had complete data on both the MRI and the cognitive test battery. We excluded all participants below the age of 55 years old (n = 553) as none of them developed dementia during follow-up.

Figure 1 | Flowchart of the selection process.



Furthermore, we excluded all individuals who were diagnosed with stroke (n = 119), dementia (n = 25) or both (n = 2) before follow-up started. Finally, participants were excluded if no information on follow-up was available (n = 77). The final sample used in the analyses included 4,112 participants.

Standard protocol approvals, registrations, and patient consents

The Rotterdam Study has been approved by the Medical Ethics Committee of the Erasmus MC and by the Dutch Ministry of Health, Welfare and Sport (Population Screening Act WBO, license number 1071272-159521-PG). All participants provided written informed consent to participate in the study and to have their information obtained from their treating physicians.

Operationalization of reserve

The primary aim of this study was to investigate the role of early-life and late-life reserve in the incidence of dementia. As participants were included during middle and late adulthood, proxies were used for early-life CR and BR. Educational attainment was used as a proxy for early-life CR and ICV as a proxy for early-life BR. Late-life CR was estimated using a structural equation model similar to the model used in the study by Petkus and colleagues[396]. For late-life BR measurements were chosen that closely related to early-life BR, i.e., ICV. Hence, late-life BR was defined as the proportion of MRI-based normal appearing brain tissue. Non-lesioned brain volume is conceptually close to ICV as a proxy for early-life BR, as both describe the total neural capacity. Late-life BR was calculated according to the equation: [total brain volume – white matter lesion volume] / ICV. The white matter lesion volume was removed from the total brain volume as BR denotes the neural capacity to buffer neuropathology, and white matter lesions are a type of neuropathology and should thus not be counted towards BR. The metric was divided by ICV for two reasons. First, this proportional metric has been shown to have a stronger association with dementia than brain volume by itself[399]. Second, the strong correlation between brain volume and ICV (Pearson's r = 0.90) introduced unwanted multicollinearity in the models that contained both ICV and late-life BR.

Educational attainment and cognitive function

Education level was assessed at baseline and classified into four categories according to the UNESCO classification. The levels used were primary education ("primary"), lower/intermediate general education or lower vocational education ("low"), intermediate vocational education or higher education ("intermediate"), and higher vocational education or university ("high"). Primary education was used as the reference class. This was because previous studies have indicated that educational attainment particularly reduces the risk of incident dementia at lower levels of educational attainment[400].

To assess cognitive function all participants underwent a cognitive test battery. The protocol has been described elsewhere[61]. In brief, the battery consisted of a 15-word verbal learning test[355], the Stroop test[356], the letter-digit substitution test[66], a verbal fluency task[357], and the Purdue pegboard Test[71].

Late-life CR

Late-life CR was estimated using a structural equation model (**Figure 2**), which was based on a previous study [396]. In brief, CR can be described as the level of cognitive function when considering the degree of neuropathology and age-related decline. The structural equation

model therefore defined late-life CR as a latent variable that is calculated by controlling the cognitive scores for demographic factors and brain-related factors including neuropathology. This approach is known as variance decomposition and has been validated as an appropriate method to assess late-life CR[395,396,401].

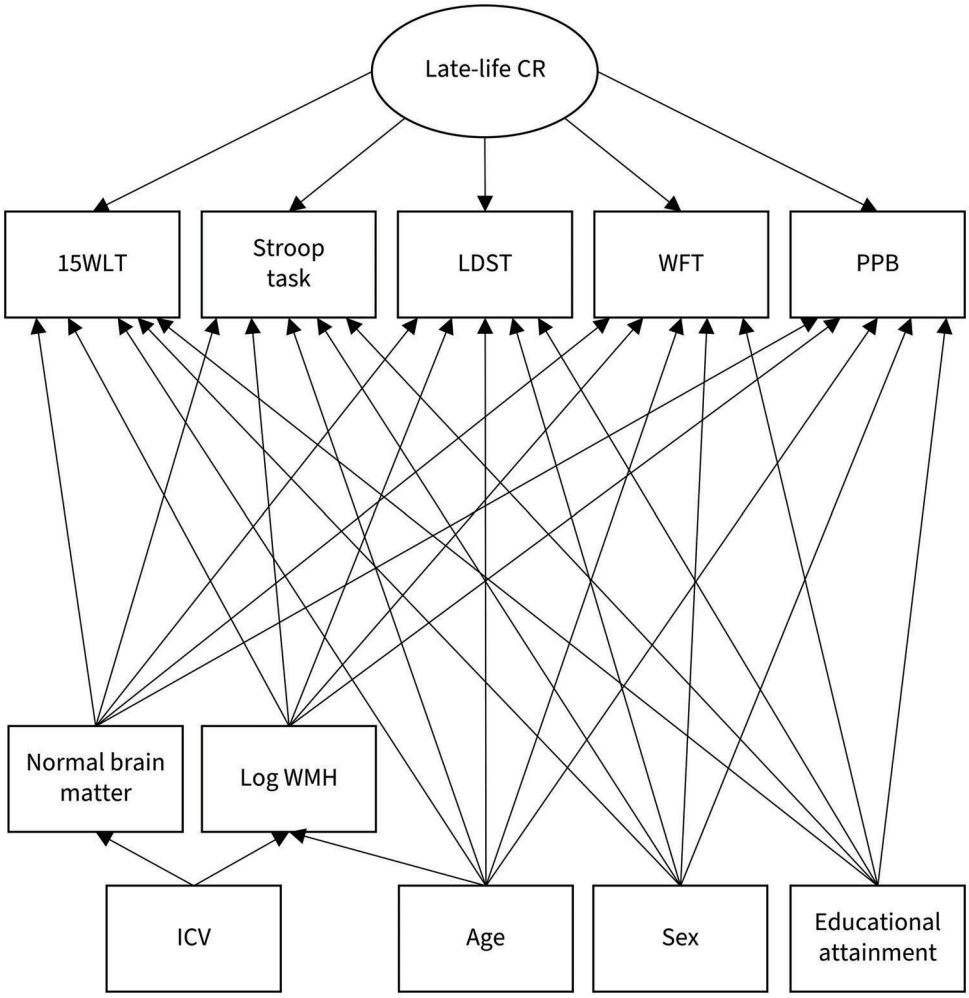
The full model was designed as follows. Scores were obtained for the five cognitive tests and for each we regressed out the effects of age, sex, educational attainment, the natural log of the white matter lesion volume and the volume of normal appearing brain matter. Normal appearing brain matter and the log of the white matter lesions were adjusted for intracranial volume, age, sex, and educational attainment. Late-life CR was defined as a latent variable on which all cognitive test scores loaded. Thus, CR was the difference between the actual cognitive functioning and the predicted level of cognition based on demographic factors and brain structure. The distribution of the Stroop test times had a strong positive skew, which led to violation of the normality assumption of the residuals in the structural equation model. This was amended by log transforming the Stroop test time distribution. The model fit was evaluated using the comparative fit index (CFI) > 0.95, the Tucker Lewis Index (TLI) > 0.95 and the root-mean-squared error of approximation (RMSEA) < 0.06, in accordance with commonly used criteria[402].

Assessment of intracranial volume and brain volume

Neuroimaging was performed on a 1.5 Tesla MRI scanner with an eight-channel head coil (GE Signa Excite, General Electric Healthcare, Milwaukee, USA). The imaging sequence and processing details have been described extensively elsewhere[13]. In brief, images from three sequences were utilized, i.e., a T1-weighted sequence, a proton density-weighted sequence, and a fluid-attenuated inversion recovery (FLAIR) sequence.

The images were segmented into grey matter, cerebrospinal fluid and white matter using an automated processing algorithm based on a k-nearest-neighbor classifier, which has been described previously[403]. In brief, the classifier automatically registers six manually segmented

Figure 2 | Structural equation to estimate late-life cognitive reserve. Comparative Fit Index = 0.991; Tucker-Lewis Index = 0.965; root mean square error of approximation = 0.050 [95% CI = 0.043; 0.056].



atlases non-rigidly to the images. By considering the *k* nearest voxels, voxel-wise tissue class probabilities can be calculated, which are then used to classify the voxel to a tissue type. All segmentations were visually inspected, and corrections were applied where necessary. Total brain volume was estimated by summing total gray and white matter volumes. ICV was calculated as the sum of the total brain volume and the cerebrospinal fluid volume.

White matter lesions were quantified through an automated method[404]. In brief, white matter lesions are typically detected on FLAIR images as hyperintense regions in the white matter.

The automated method utilized the gray matter classification to determine the optimal intensity threshold for the white matter lesions on the FLAIR images, thus enabling automated segmentation[404].

Assessment of incident dementia

Participants were screened for dementia at study entry and subsequent center visits with the Mini-Mental State Examination[405] and the Geriatric Mental Schedule organic level[406]. Those with a Mini-Mental State Examination score below 26 or Geriatric Mental Schedule score above 0 underwent a physician interview and additional testing using the Cambridge examination for mental disorders in the elderly (CAMDEX)[407]. In addition, the entire cohort was continuously under surveillance for dementia through electronic linkage of the study database with medical records from general practitioners and the regional institute for outpatient mental health care. Study physicians biannually evaluate all records and combine information from medical records with in-person screening to draw up individual case reports. In these reports, the physicians covered all gathered relevant information to establish the presence, probability, and subtype of dementia. A consensus panel led by a consultant neurologist established the final diagnosis according to standard criteria for dementia using the revised third edition of the Diagnostic and Statistical Manual of Mental Disorders version III (DSM-III-R)[408].

Stroke

Participants were excluded if they had prevalent stroke, i.e., if they were diagnosed with stroke before the start of follow-up (see "Study Population"). The stroke ascertainment has been described previously[409]. In brief, during the baseline interview participants were asked whether they had suffered a stroke, as diagnosed by a physician. These answers were then verified through medical records. After the baseline interview, participants were continuously monitored for the occurrence of events through automatic linkage with the files of the general practitioners. If an event occurred, a consensus panel would consider the event based on information from the general practitioner and hospital discharge records if available. The event was defined as a probable stroke if typical clinical symptoms were present but no clinical information on neuroimaging was available. MRI scans obtained as part of the Rotterdam Study were not used in the consensus panel to maintain consistency of ascertainment among those that did and did not participate in the MRI scanning. For the current paper, all participants with definite or probable stroke before the start of follow-up were excluded.

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 minutes of rest, and the mean was taken as the participant's

reading. Use of blood pressure lowering medication was derived from information collected at the home interview. Alcohol use was assessed during the home interview, and assessed consumption of beer, wine, liquor, and other alcohol types. To define the total alcohol intake, the number of each beverage consumed was converted to alcohol intake in grams per day and then summed up. The algorithm to calculate alcohol in grams per day is described elsewhere[240]. Smoking status was obtained during home interviews and defined as never, former, and current. Body mass index (BMI) was calculating using the height and weight obtained with calibrated scales during the research center visit. The number of APOE ε4 alleles was determined by DNA sequencing procedures which have been described elsewhere[410].

Statistical analysis

The primary aim of the study was to examine the associations of early-life and late-life CR and BR with the risk of developing dementia. To do so, Cox regression models were constructed. The beginning of follow-up time was defined as the moment of completion of the cognitive measures or the MRI visit, whichever occurred last. The censor date was based on the date of dementia diagnosis, date of loss to follow-up, date of death, or January 1st 2016, whichever came first. All analyses were performed with age as the time scale. All reserve proxies but educational attainment were standardized.

To adjust for potential confounding, the models included the following covariates: Cohort (RS-I, RS-II or RS-III), sex, age difference between cognitive testing and MRI scan (years), hypertension (yes/no), alcohol intake (g/day), smoking status (never/past/current), BMI (kg/m^2) and APOE ε4 allele count. Given the strong correlation between ICV and body height, body height was also considered as a confounder. However, adjusting for body height may lead to overadjustment and thus an underestimation of the association between ICV and the incidence of dementia[394]. Furthermore, the results were unattenuated when body height was included as a covariate (data not shown). Therefore, only the findings from the models that were not adjusted for body height were reported.

To study the relative contribution of all factors in the incidence of dementia, three different types of Cox regression models were constructed. In Model 1, to examine the effect of each proxy in isolation, each reserve proxy was examined in a separate Cox regression model:

- **Early-life CR**: $\ln\left(\frac{h_1(t)}{h_0(t)}\right) = \beta_{edu} \cdot edu + \beta_{covariates} \cdot covariates$
- **Early-life BR**: $\ln\left(\frac{h_1(t)}{h_0(t)}\right) = \beta_{ICV} \cdot ICV + \beta_{covariates} \cdot covariates$
- **Late-life CR**: $\ln\left(\frac{h_1(t)}{h_0(t)}\right) = \beta_{llCR} \cdot llCR + \beta_{covariates} \cdot covariates$
- **Late-life BR**: $\ln\left(\frac{h_1(t)}{h_0(t)}\right) = \beta_{llBR} \cdot llBR + \beta_{covariates} \cdot covariates$

where $\frac{h_1(t)}{h_0(t)}$ is the hazard ratio, ln is the natural log, the βs are the regression coefficients, llCR is the late-life CR and llBR is the late-life BR. In Model 2, the CR proxies were combined into one model and the BR proxies into another to consider the relative importance of early-life and late-life proxies:

- **CR**: $\ln\left(\frac{h_1(t)}{h_0(t)}\right) = \beta_{edu} \cdot edu + \beta_{llCR} \cdot llCR + \beta_{covariates} \cdot covariates$
- **BR**: $\ln\left(\frac{h_1(t)}{h_0(t)}\right) = \beta_{ICV} \cdot ICV + \beta_{llBR} \cdot llBR + \beta_{covariates} \cdot covariates$

In Model 3, all four reserve proxies were combined into a single model to see whether CR and BR exert independent effects on the risk of dementia:

- $\ln\left(\frac{h_1(t)}{h_0(t)}\right) = \beta_{edu} \cdot edu + \beta_{ICV} \cdot ICV + \beta_{llCR} \cdot llCR + \beta_{llBR} \cdot llBR + \beta_{covariates} \cdot covariates$

For each model, the hazard ratios, 95% confidence intervals (95% CI) and p-values are reported.

To see how CR and BR interact in their protective effects, two new models were created: one containing early-life CR and BR, and one containing late-life CR and BR. In both models, both variables were dichotomized along their mean values and created four groups of individuals based on their levels of CR and BR. We consequently looked at whether these groups differed in their risk of dementia.

To see how stable the findings were, the analyses were stratified by for sex, age at baseline (below or equal and above 77 years of age) and APOE ε4 carriership to see how stable the associations were. Missing data were imputed twenty times using chained equations and the estimates from the models were pooled subsequently[101,105,411]. All statistical analyses were performed in R 3.4.2 using the 'survival' package[95,412]. The structural equation model was built with the 'lavaan' package[413].

Data availability statement

The datasets for this manuscript are not automatically publicly available due to legal and informed consent restrictions. Reasonable requests to access the datasets should be directed to the management team of the Rotterdam Study (secretariat.epi@erasmusmc.nl), which has a protocol for approving data requests.

RESULTS

Table 1 shows the baseline characteristics of the 4,112 participants. The mean age at baseline was 66.3 years (standard deviation = 8.7), and 2,278 (55.4%) were women. The median

Table 1 | Baseline characteristics of the study population (N = 4,112).

Characteristics	M ± SD or N (%)
RS cohort	
RS-I	806 (19.6%)
RS-II	1,134 (27.6%)
RS-III	2,172 (52.8%)
Sex, female	2,278 (55.4%)
Age at start of follow-up in years	66.3 ± 8.7
Years difference between cognitive testing and MRI scan (median; IQR)	0.13 ± 0.26
Hypertension, yes[a]	2,696 (65.7%)
Alcohol in g/day[a]	9.7 ± 10.6
Smoking status[a]	
Never	1,277 (31.2%)
Past	2,096 (51.1%)
Current	722 (17.6%)
Body mass index in kg/m2[a]	27.4 ± 4.1
APOE ε4 number of alleles[a]	
0	2,850 (71.3%)
1	1,053 (26.4%)
2	92 (2.3%)
Education level	
Primary	333 (8.1%)
Low	1,617 (39.3%)
Medium	1,215 (29.5%)
High	947 (23.0%)
ICV in cm^3	1,139.4 ± 115.9
Late-life BR ((BV – WMHV) / ICV)	0.82 ± 0.04

BV = brain volume; ICV = intracranial volume; IQR = interquartile range; RS = Rotterdam Study; WMHV = white matter hyperintensity volume.

[a] Missingness was present in the variables hypertension (0.2%), alcohol intake (6.3%), smoking status (0.4%), body mass index (0.2%) and APOE ε4 allele count (2.8%).

duration between the cognitive testing and the MRI scanning was 0.13 years (interquartile range: 0.26; from 0.08 to 0.34). Most participants achieved low to intermediate education (primary = 8.1%, low = 39.3%, intermediate = 29.5%, high = 23.1%). The mean ICV was 1,139.4 cm^3 (standard deviation = 115.9 cm^3). Both late-life CR and BR were standardized and had a mean of 0 (standard deviation = 1). The structural equation for late-life CR had a satisfactory fit (CFI = 0.991, TLI = 0.965, RMSEA = 0.050 [95% CI = 0.043; 0.056]).

Incidence of dementia and separate models of the reserve proxies

During 24,631 person-years of follow-up (mean follow-up time = 6.0 years, SD = 2.8) a total of 110 participants developed dementia. As expected, the incidence rate increased with age (0.09% at 70 years, 0.80% at 80 years and 3.91% at 90 years).

Table 2 shows the results of the Cox regressions for the risk of dementia associated with each reserve proxy. In Model 1, compared to those with only a primary education, the risk of dementia was lower in those with low education (hazard ratio (HR) = 0.68 [95% CI = 0.38; 1.24], p = .20), intermediate education (HR = 0.59 [0.31; 1.12], p = .11) and high education (HR = 0.48 [0.21; 1.06], p = .07), although none of these associations reached statistical significance. Similarly, higher ICV associated with a reduced risk of dementia (HR = 0.94 [0.74; 1.19], p = .62), but this also did not reach statistical significance. Both late-life CR (HR = 0.57 [0.48; 0.68], p < .0001) and late-life BR (HR = 0.54 [0.43; 0.68], p < .0001) associated with a lower risk of dementia.

Incidence of dementia and combined models of the reserve proxies

In Model 2, the CR proxies were combined into one model and the BR proxies into another model. The effect for educational attainment was attenuated ($HR_{primary\text{-}vs\text{-}low}$ = 0.77 [0.42; 1.39], p = 0.38; $HR_{primary\text{-}vs\text{-}intermediate}$ = 0.73 [0.38; 1.40], p = 0.34; $HR_{primary\text{-}vs\text{-}high}$ = 0.56 [0.25; 1.25], p = 0.15), whereas late-life CR remained strongly associated with a lower risk of dementia (HR = 0.58 [0.48; 0.68], p < .0001). The combined model for the BR proxies did not change the results compared to the separate models. In Model 3, all proxies were combined into a single model. The results of Model 2 and Model 3 were very similar, with both late-life CR (HR = 0.62 [0.52; 0.74], p < .0001) and late-life BR (HR = 0.59 [0.46; 0.75], p < .0001) associating with a lower risk of dementia.

Interaction between CR and BR

Participants were divided into four groups: high CR with high BR, high CR with low BR, low CR with high BR and low CR with low BR. This was done for both early-life and late-life reserve proxies separately. The survival curves for each early-life group are shown in **Figure 3A**, and

Table 2 | Hazard ratios of incident dementia related to early and late-life proxies of reserve. A total of 110 events occurred during a mean follow-up time of 6.0 years (SD: 2.8). ICV, late-life CR and late-life BR were standardized. Included covariates were cohort, sex, age difference between cognitive testing and MRI scan (years), hypertension (yes/no), alcohol intake (g/day), smoking status (never/past/current), BMI (kg/m2) and APOE ε4 allele count.

Life phase	Domain	Measure	Model 1[a] HR (95% CI)	Model 2[b] HR (95% CI)	Model 3[c] HR (95% CI)
Early-life	CR	Education			
		Primary	(reference)	(reference)	(reference)
		Low	0.68 (0.38; 1.24)	0.77 (0.42; 1.39)	0.82 (0.45; 1.49)
		Intermediate	0.59 (0.31; 1.12)	0.73 (0.38; 1.40)	0.80 (0.41; 1.56)
		High	0.48 (0.21; 1.06)	0.56 (0.25; 1.25)	0.56 (0.24; 1.28)
Early-life	BR	ICV, per SD	0.94 (0.74; 1.19)	0.85 (0.67; 1.07)	0.89 (0.70; 1.12)
Late-life	CR	Predicted CR, per SD	0.57 (0.48; 0.68)	0.58 (0.48; 0.68)	0.62 (0.52; 0.74)
Late-life	BR	Proportion of healthy brain tissue, per SD	0.54 (0.43; 0.68)	0.52 (0.41; 0.66)	0.59 (0.46; 0.75)

HR = hazard ratio; CR = cognitive reserve; BR = brain reserve; ICV = intracranial volume.

[a] = Separate models for each of the four proxies.

[b] = Separate models for the CR proxies and the BR proxies.

[c] = Combination of all four proxies into a single model.

for each late-life group in **Figure 3B**. No statistically significant differences were found between the early-life groups (all p > .05). In contrast, differences between the groups were present in the late-life proxies. Compared to the high CR high BR group, the high CR low BR group did not differ in incidence of dementia (HR = 1.14 [0.58; 2.25], p = .70), whereas both the low CR with high BR group (HR = 2.24 [1.21; 4.13], p = .01) and the low CR with low CR group (HR = 3.93 [2.24; 6.87], p < .0001) had higher rates of dementia.

Sensitivity analyses

Within model 3, the effect of stratification on age, sex and APOE ε4 carriership was further assessed. All HRs for these sensitivity analyses are displayed in **Figure 4**. The proxies for reserve did not show statistically significant differences in their associations with incident dementia for those aged below 77 years old versus those 77 years and older (all p > .05). Moreover, no differences were found when comparing the associations in men and women (all p > .05). For APOE ε4 carriership we found that carriers showed a weaker association between the late-life CR and incidence of dementia (HR = 0.86 [0.63; 1.16], p = .32) than non-carriers (HR = 0.52 [0.41; 0.66], p < .0001), although in the same direction.

Figure 3 | **Dementia survival curves for the low and high CR and BR strata.** Panel A displays the curves for early-life reserve, and Panel B for late-life reserve. The variables were dichotomized along the mean. CR = Cognitive reserve; BR = Brain reserve.

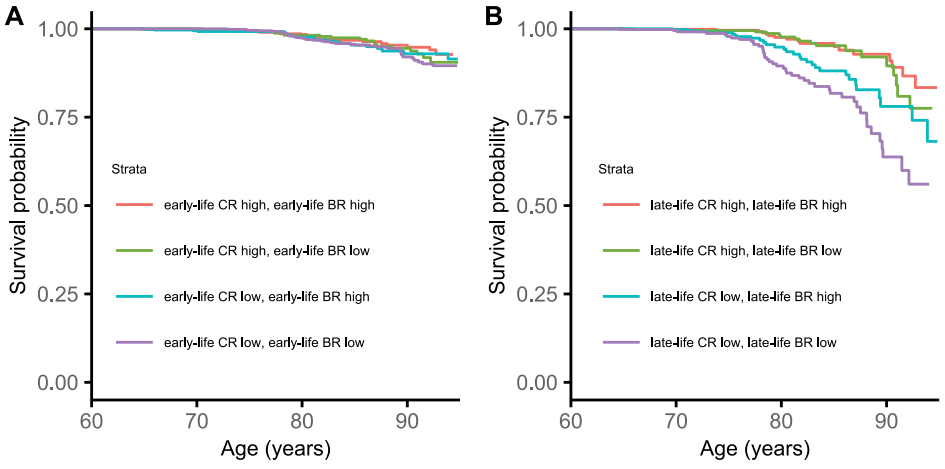
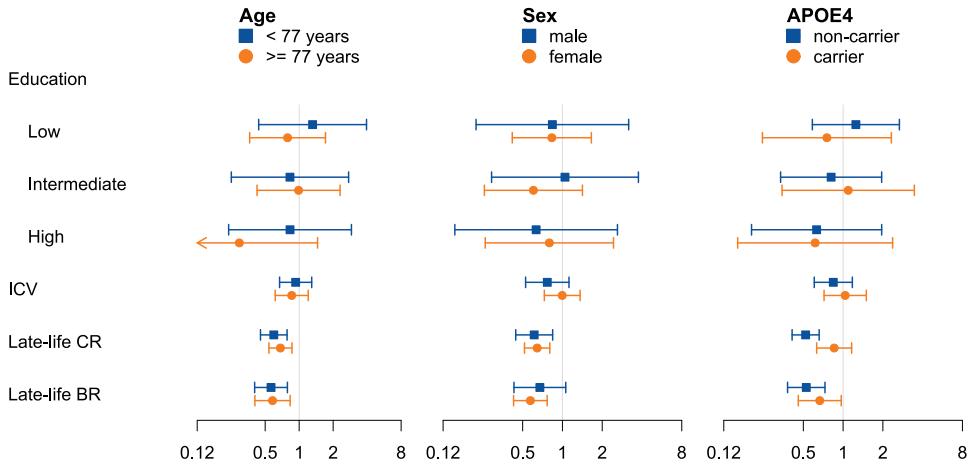


Figure 4 | Hazard ratios for incident dementia and their 95% CIs, stratified for age, sex and APOE ε4 carriership.



DISCUSSION

Late-life constructs of reserve – both CR and BR – related more strongly to incidence of dementia than early-life constructs. Furthermore, while early-life CR reduced the risk of dementia, this effect was attenuated when considering late-life CR simultaneously. Finally, late-life CR and late-life BR protected both independently and interactively against the incidence of dementia.

Previous literature on reserve has mostly focused on proxies that are relatively constant throughout life, such as educational attainment and ICV[20,394]. Educational attainment has been extensively studied in both patients and dementia-free populations, and it is an integral part of most theories related to resilience towards dementia. For example, higher levels of education could lead to healthier lifestyle choices and therefore better cardiovascular and brain health[414]. Still, in a systematic review of 88 population-based studies only 51 studies found a statistically significant association between education and the prevalence or incidence of dementia[400]. Our findings also showed a protective effect of education on dementia, but the association did not reach statistical significance. The practical utility of educational attainment as a proxy for CR should therefore be examined more critically, especially considering the existence of latent variable methods to estimate CR from a wider range of variables.

The comparison between reserve proxies from different life phases has been studied previously[31,32]. One study analyzed data from 7,574 individuals on primary school grades, educational attainment, and occupational complexity[31]. The analyses showed that high levels of these factors independently lowered the risk of dementia to a comparable degree. These findings were extended by a similar study in 602 participants of the Kungsholmen Project, a prospective longitudinal cohort[32]. Late-life CR was defined as a latent variable based on physical, social, and mental activity during late life. They found that early-life, mid-life, and late-life CR related to the incidence of dementia. More importantly, when all CR measures were combined into a single model the early-life and mid-life CR effects were attenuated. The present study extended these findings by using a more direct measure of late-life CR, i.e., structural equation modeling that incorporates sociodemographic and brain factors.

A recent meta-analysis showed that a larger ICV tends to correlate with better cognitive function after adjusting for pathology[394]. Several other studies have suggested that individuals with a larger ICV show slower rates of cognitive decline[415,416]. Still, ICV does not seem to differ between dementia patients and healthy controls[417-420]. The present study also did not yield support for a link between ICV and dementia. Rather, we found that late-life BR – which was defined as the amount of healthy brain tissue – protects against incident dementia with a comparable effect as late-life CR. Moreover, the effects for CR and BR seemed to be additive, where higher levels of reserve in either decreased the risk for dementia. These findings emphasize the need to reconsider the role of ICV as the most appropriate proxy for BR[394], especially when more dynamic measures likely have greater clinical relevance.

The present study had several limitations. First, education and ICV are commonly used proxies for early-life CR and BR, but they arguably do not reflect early-life reserve accurately. For example, educational attainment is strongly influenced by sociodemographic and cultural factors. During the 1950's and 1960's women did not necessarily have equal access to education as men did[421], regardless of their cognitive and academic functioning, which may compromise the value of education as a CR proxy. A second limitation of the study is the limited number of cases within the study. While the Rotterdam Study included 1,741 incident dementia cases by the end of 2016, most of these occurred before brain MRI scanning was introduced into the Rotterdam Study. A third limitation is that no information was available on earlier phases of cognitive decline, such as mild cognitive impairment. Reserve likely plays a role during those phases too, and the effects of BR and CR may differ from those observed in this study. Finally, the ascertainment of dementia relied on the DSM-III-R criteria rather than newer guidelines such as those proposed by the NIA-AA[422]. The reason for relying on the DSM-III-R criteria is because they have been in place since the beginning of the Rotterdam Study in 1989. By maintaining the same standards and using the same screening tools over the last 30

years, the Rotterdam Study data have been instrumental in understanding how the incidence of dementia develops over time. However, this does mean that the findings should be interpreted in the context of the DSM-III-R criteria.

The study also had several strengths. The dementia ascertainment procedures were extensive, with continuous screening of medical records and follow-up based on research center results. In addition, loss to follow-up was low in participants who partook in both cognitive testing and the brain MRI scan. The number of missed dementia cases was therefore likely low. Finally, we were able to correct for a wide range of factors that may have confounded the associations.

CONCLUSION

Late-life reserve proxies are relatively more important than early-life reserve proxies in the incidence of dementia. Additionally, late-life CR and late-life BR have independent and additive protective effects against incidence of dementia. This study emphasizes the need to consider CR and BR simultaneously to further elucidate the etiology of dementia.

GENERAL DISCUSSION

6

> *"In the company of blockheads, the mind gets rusty for lack of use. In tennis, one's return is rotten when the serve is bad. I would prefer an intelligent man to have had no education at all, provided he is still young enough, than to have been badly educated. A man with a poorly trained mind is like an actor spoiled in the provinces."*
>
> – Julien Offray De La Mettrie (L'Homme Machine, 1748)

The human brain develops and deteriorates across a lifespan. In this thesis, I aimed to further elucidate the role of genetic, vascular, and reserve-related mechanisms in shaping the brain. For the genetic mechanisms, we focused on polygenic scores for late-life neurodegenerative disease and subcortical volumes and their relevance for the brain during early life. For the vascular mechanisms, we first explored the association of blood pressure and arterial stiffness with cognition during early and late life. Next, we examined how vascular risk factors during early and mid adulthood relate to cognition and brain structure over 15 years later. For reserve-related mechanisms, we first considered how early life adverse events shape the early-life brain. Next, we characterized the age-related differences in cortical gyrification across adulthood. Finally, we showed how early-life and late-life brain and cognitive reserve (CR) and brain reserve (BR) relate to the incidence of late-onset dementia. These studies primarily relied on data from the Generation R Study for early-life data, and the Rotterdam Study for later-life data. We also established the ORACLE Study, a neuroimaging visit focused on the parents of the Generation R children, to explore early and mid adulthood risk factors and their role in subsequent brain health. Here, I present a discussion of the findings of this thesis.

FINDINGS IN PERSPECTIVE

On genetic mechanisms

The etiology of late-life neurodegenerative disease may have its roots early in life. As the genetic code is largely stable across the lifespan, it has been speculated that the genetic burden for late-life neurodegenerative diseases affects early-life brain structure and function. Our findings suggest that this is unlikely, at least for the measures that we examined. Polygenic scores for Alzheimer's disease, Parkinson's disease and frontotemporal dementia did not associate with cognitive function, global brain structure or brain regions that have been implicated in those diseases. As expected, APOE did influence serum lipid levels during early childhood. Abnormal levels of lipid levels are seen as a vascular risk factor, that may in turn affect brain structure and function later in life. Thus, while genetic burden for late-life neurodegenerative diseases do not seem to affect brain function and global and regional brain structure during early-life, they may indirectly affect the brain later in life.

We further explored the role of genetics in early-life by calculating polygenic scores for subcortical volume in Generation R. We found that the scores related to subcortical volumes during infancy – as measured with cranial ultrasound – and during early childhood – as measured with MRI. Through mediation analysis, we were able to show that the effect of the genetic scores on subcortical volumes during childhood was only partly mediated by the subcortical volumes during infancy. This suggests that the scores capture genetic mechanisms across different developmental phases, primarily birth. It also suggests that reserve – at least in subcortical regions – is partly shaped by genetics during early life.

On vascular mechanisms

Previous literature has shown that midlife vascular characteristics like blood pressure and arterial stiffness associate with subsequent cognition levels. Our findings from **chapter 4.1** provide further support for this association, as higher late-life systolic blood pressure (SBP) and arterial stiffness were associated with poorer cognitive performance. Furthermore, diastolic blood pressure (DBP) showed a non-linear association, with both low and high DBP were associated with lower cognition. Given that the etiology of late-life neurodegenerative disease has its roots in early life, we considered whether the association between vascular risk factors and cognition is also present during early-life. Our findings suggest that higher DBP does relate to lower levels of cognition during early childhood, but with a weaker effect than what we observed in the late-life population.

To explore the longitudinal nature of the association between vascular risk factors and brain structure and function, we used data from the ORACLE Study. The ORACLE Study enabled the study of vascular factors roughly 15 years before assessment of cognition and brain structure. Data were available from a wider range of vascular risk factors: SBP, DBP, body mass index, smoking low-density lipoproteins, high-density lipoproteins (HDL), non-HDL levels, and triglycerides. Interestingly, in this sample blood pressure did not seem to associate with subsequent cognition. This may be because the sample size was too small to detect the effect. Indeed, we found much larger effect sizes than with the cross-sectional approach in **chapter 4.1**, but the standard errors were also proportionally larger. Additionally, we found that blood pressure did associate with subsequent performance on the Purdue pegboard test, and with white matter lesion volume. Further stratification by age showed that blood pressure from the age of 30 years onwards was associated with subsequent white matter hyperintensities. This may have been due to the limited follow-up; associations could appear at younger ages if the follow-up time was longer. Overall, these findings do suggest that vascular risk factors associate with brain structure and cognition across the lifespan.

On mechanisms related to reserve

Levels of CR and BR build up during early life. This is likely due to both genetic and environmental factors. Adverse childhood experiences have previously been linked to earlier onset of puberty, particularly if they were of a threatening nature[318]. Our findings suggest that threatening adverse childhood experiences also accelerate aging of the adolescent brain. This may affect reserve levels in different ways. First, the build-up of reserve slows down significantly during early adulthood. Early maturation of the adolescent brain may induce such a state of slower reserve build-up at an earlier age and thus there would be less time to accumulate reserve. This would lead to an earlier and lower peak in reserve levels compared to children whose brain matures later. Second, adverse childhood experiences may lead to negative consequences in other aspects of life, such as academic performance and cognitive and psychosocial development. This may lead to a reduced amount of reserve to be build, which would appear as a seemingly earlier maturation of the brain. This mechanism would also affect later factors that may increase reserve, such as educational attainment and occupational complexity. It is important to elucidate which mechanisms are at play, as the consequences of the adverse childhood experiences on reserve may propagate beyond adolescence.

We further studied how early-life and late-life reserve relate to the incidence of dementia. In the Rotterdam Study, we defined early-life CR as educational attainment and early-life BR as intracranial volume. Late-life CR was determined through variance decomposition of cognition based on demographics, white matter lesion volume, and late-life CR. Finally, late-life BR was set as the non-lesioned brain volume divided by the intracranial volume. Importantly, we found that late-life reserve levels are more relevant than early-life reserve levels. As the early-life reserve proxies are strongly predictive of late-life proxies, they may affect the risk of dementia partly through their influence on the late-life reserve levels, although further work is needed to demonstrate this. In addition, we considered the relative importance of CR and BR. We found that late-life CR and BR both had a comparable influence on the risk of dementia. Finally, the late-life reserve levels seemed to interact with each other, and thus having low levels of both will lead to a more than additively higher risk to develop dementia. The findings related to brain reserve were somewhat surprising, as the measure for brain reserve was relatively simple, i.e., gross brain volume. Brain reserve can be better approximated with more refined techniques like positron emission tomography or more specific measures such as regional/hippocampal volumes[423-425]. More refined measures of brain reserve may show even stronger associations with incidence of dementia, and could further change our understanding of CR and BR.

To better understand refined features of the brain, we studied the population characteristics of cerebral gyrification using QDECR. The focus was primarily on older age, as few studies included individuals over the age of 60 years. We showed that global gyrification has a linear negative association with age rather than a non-linear association from middle adulthood onwards. Furthermore, we mapped regional patterns of age-related changes in gyrification. This revealed one area, the medial prefrontal cortex, that showed more gyrification with age during late adulthood. Furthermore, we found that gyrification in primarily temporal regions were related to cognition. Finally, we showed that these associations were independent of other well-known characteristics of the cerebral cortex, i.e., cortical thickness and surface area. These findings point to cortical gyrification as a possible biomarker for age-related and cognition-related states of the brain.

METHODOLOGICAL CONSIDERATIONS

On the definition of reserve

The idea of a brain-related reserve stems from a study published by Katzman and colleagues in 1988[426]. Through post-mortem examination, they identified elderly individuals with minimal cognitive decline yet with extensive pathological features of Alzheimer's disease. This gave rise to the idea of a reserve, a buffer capacity to mitigate the negative consequences of neuropathology. In the same year, a book chapter was published that speculated on the idea of CR. The chapter stated that "psychosocial factors act primarily to reduce the margin of intellectual reserve to a level where a more modest level of brain pathology results in a diagnosable dementia."[427] Over 30 years have passed, and numerous studies have been conducted with the aim to elucidate what CR is. However, no study has yielded a concrete framework that directly defines or measures CR.

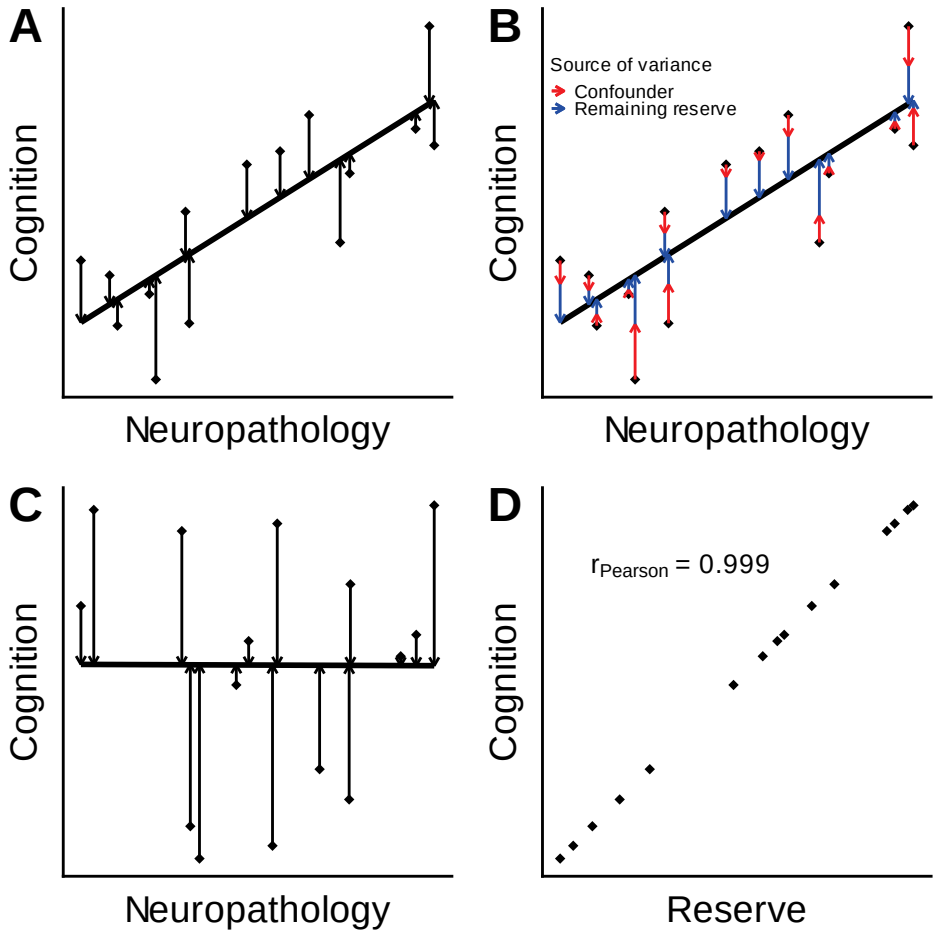
CR cannot be measured directly, and it thus has two definitions: a theoretical/conceptual definition and an operational definition. According to the white paper on reserve[20], CR is conceptually defined as "the adaptability of cognitive processes that helps to explain differential susceptibility of cognitive abilities or day-to-day function to brain aging, pathology or insult." The operational definition of CR states how it can be measured, for which two established methods exist. The first is to use proxy measures, such as educational attainment or occupational complexity, to see how they modulate the association between pathological brain features and the symptomatology of dementia. However, this method has severe drawbacks. The primary drawback is that it defines CR as the extent to which the CR proxy moderates the association between neuropathology and some factor. Consequently, CR cannot be quantified, and this method only allows researchers to see whether CR has an effect

or not. The second method uses structural equations to model CR as a latent construct[395,396], which generally entails that cognitive performance is modeled as a consequence of demographic factors, neuropathology, and the unobserved CR. This method boils down to the idea that, given fixed levels of neuropathology, CR is defined as the remaining unexplained effect on the outcome. It is significantly more powerful than the first method, as it assigns a quantity to CR.

Here, I will argue that the current conceptual framework of CR is flawed and ipso facto hinders any progress in understanding it. My arguments stem primarily from two key observations: misspecification of definitions and residuality. Even though CR has a theoretical and an operational definition, they are inseparable. The theoretical definition as stated in the Whitepaper consists of two parts: "the adaptability of cognitive processes" and all the words that follow. The second part is essentially identical to the operational definition. It simply states that the first part – i.e., the adaptability – must explain the operational definition, i.e., why differences exist between people in their functioning when considering neuropathology. Thus, the theoretical definition encapsulates the operational definition. What remains is "the adaptability of cognitive processes", which should explain the individual differences in susceptibility, i.e., the operational definition. The "adaptability of cognitive processes" is a valid definition, but it is not part of the operational definition, as the operational definition only addresses the susceptibility that is caused by the lack of adaptability. Furthermore, the existing methods to operationalize CR actually quantify the susceptibility, not the adaptiveness of cognitive processes per se. Combined, this implies that the theoretical and operational definitions are redundant – the theoretical definition encapsulates the operational one – and insufficient – the operational definition does not address the actual "adaptability of cognitive processes".

A second problem is the residual nature of the operational definition. In the structural equation method, CR is the residual variance that remains after removing the effects of demographics and neuropathology (**Figure 1A**). The residual variance can vary for a multitude of reasons. Any confounding factor that is not adjusted for will contribute to the residual variance and will unintentionally be labeled as CR (**Figure 1B**). For example, if the neuropathological trait is linked to sex and sex were not adjusted for, then the sex-related residual variance will incorrectly be interpreted as CR. Furthermore, the residual model implicitly assumes that the included neuropathology measures associate with the cognitive measures. If they do not associate (**Figure 1C**), then most variance will be unexplained (i.e., residual), and the CR measure will correlate highly with the cognitive measures (**Figure 1D**). Paradoxically, cognitive measures that do not associate with common neuropathology

Figure 1 | (A) Plot of neuropathology versus cognition. The residuals (arrows) can be labelled as CR. (B) Further partitioning of the residuals. (C) Residuals when there is no association between the neuropathology measure and cognition. (D) Correlation between CR from panel C and cognition.



measures but associate with dementia symptoms will seem to be the strongest proxies for CR according to the residual method, despite not relating to CR at all.

As a consequence of the arguments above, little research can be conducted to further elucidate the mechanisms of CR. If CR was to be elucidated through its theoretical definition, then the focus would be on exploring the adaptiveness of cognitive processes. However, the current operationalizations do not address adaptiveness. If instead CR was to be elucidated from its operational definition, then it would be important to identify a way to quantify CR

directly. However, as stated, no direct measures exist. Further elucidation is not truly possible when CR is defined as a residual, which is the case for as long as the operational definitions force a dependence between neuropathology and CR. Without acknowledging these limitations, the field will not progress in grasping what CR actually entails.

To improve upon the current definitions, we must consider how they arose. The residual approach to CR considers the adaptability of cognitive processes by taking out the contribution of neuropathology (and other factors) from cognitive function. This begs the question: Is it possible – if only in theory – to quantify adaptability of cognitive processes? The answer is yes. Keep in mind that cognitive processes arise from the brain and depend heavily on the brain's structural and functional organization. Cognition may vary depending on the structural and functional connections in the brain on the level of individual neurons, neural populations, cortical columns, and brain regions. Thus, it should in theory be possible to derive the level of cognition with full knowledge of the state of a brain. Similarly, adaptability of cognitive processes depends on reorganization of the brain on structural and functional levels. The degree to which a person's brain can reorganize in response to aging and neuropathology is the essence of cognitive reserve. If we can quantify this compensatory reorganization ability, then we would have a direct measure of cognitive reserve without relying on measures of cognition or neuropathology. This idea has been proposed several times and has been considered in great detail elsewhere[428]. While it is currently not clear how such a reorganization ability can be measured, further work on this topic can revolutionize the concept of CR.

On the surface of the brain

The field of neuroimaging is rapidly developing. Cohorts such as ABCD Study and the UK Biobank provide high quality brain images of thousands of individuals to thousands of research groups who are capable of analyzing such datasets. More than ever, the field needs to discuss the ramifications of study design in neuroimaging, and the factors that need to be considered. Here, I will discuss two topics: parcellation of the brain and epidemiological bias.

Parcellation of the brain

Realistically, no factor equally affects every cell in the brain equally, nor do all cells in the brain affect a common outcome. Effects are usually refined to specific regions or cell-types. For example, Alzheimer's disease has been extensively linked to the hippocampus[429], a brain region that plays a central role in memory[430]. Precise spatial localization of effects is thus important.

A major consideration to achieve proper spatial mapping is the subdivision of brain measurements into clusters or brain regions. The shape of the cerebral cortex can be captured through tessellation of the surface, resulting in a vertex-wise, surface-based map[431]. Using brain atlases, the maps can be further parcellated or divided into specific regions. The first brain atlas was created by Korbinian Brodmann, and was based on histological studies with a focus on cytoarchitectural differences of brain regions[432]. In 2006, a structural MRI-based atlas was published that is now known as the Desikan-Killiany atlas[327], which is still widely used to study specific brain regions in MRI-based studies. The Desikan-Killiany atlas was also used in this thesis in chapter 3.2 to study regions related to late-life neurodegenerative diseases, and in chapter 5.1 to segment the cerebral cortex for the study of brain age. Over the years, more refined and more detailed atlases. This was done by improving the quality of the dataset by using larger samples and multi-modal data[433], or by focusing on specific types of study such as neonates[434].

Unfortunately, the use of atlases has clear limitations. First, no clear guidelines or criteria exist regarding when to use which atlases. Second, the different atlases may have large effects on the results of analyses. For example, the team of FreeSurfer updated their hippocampus atlas between version 5.3 and 6.0[88]. The latter used an atlas that was based on high-resolution images, i.e., 0.1 mm^3 isometric voxels, that yielded significantly more accurate segmentation of the hippocampus and its surrounding region[435]. Thus, the choice of atlas may heavily influence the accuracy of the analyses, and thus the precision and validity of the conclusions drawn from a study. Third, atlases assume that the division of brain regions is identical between the sample on which the atlas is based and the sample that is being studied. In fact, the boundaries of regions from functional atlases have been shown to change depending on the underlying task from which the functional data are derived[436]. This implies that relying on individual atlases may not be sufficient. It may be beneficial to incorporation of multiple atlases as has been done for the subcortex in MaGeTbrain[437], or to even tailor atlases to the individual scans. Combined, these limitations suggest that analyses may yield differing results depending on the atlas used. It is unclear how the results of the atlas-based analyses in this thesis would change had we used other atlases than the Desikan-Killiany atlas.

An alternative approach is vertex-wise analysis, where all vertices are studied individually instead. While this still requires the mapping of all scans to a common template, it does not use atlases to create measures for specific brain regions. Given that vertex-wise analysis makes fewer assumptions about the distribution of brain regions, it is especially useful for studies where the relevant brain regions are unknown or poorly understood. However, vertex-wise analyses also have much lower statistical power as they tend to involve hundreds of thousands of vertices and thus high false positive rates. Furthermore, the analysis of vertices

ignores the underlying cytoarchitectural properties of that region, which can be very relevant for understanding the research question at hand. Vertex-wise analyses are not performed because we care about individual vertices, but to give a rough indication from which region effects may arise. Ironically, the findings from vertex-wise analyses can only be interpreted by linking the statistically significant vertices to a region or biological construct, which is done through atlases.

In this thesis we relied on atlases for segmentation of subcortical regions (chapter 3.1), regions related to late-life neurodegenerative disease (chapter 3.2) and whole-brain segmentation (chapter 5.1). It is unclear how the results of these chapters would have changed had we used different atlases. Moreover, brain-related prediction models – like the brain age model in chapter 5.1 – may improve further if atlas-based data and vertex-wise data are combined, as both approaches capture different types of biologically relevant information. However, to what extent the results would have changed given other approaches will remain unknown. Thus, further work is needed to elucidate the importance of atlas choice in neuroimaging studies, and consensus is needed on how atlases are to be selected and applied.

Bias

Historically, neuroimaging studies have primarily been conducted as part of neurobiological or neuropsychological research. In recent years, neuroimaging data from a number of large population-based cohorts have become publicly available, such as the UK Biobank. In certain subfields of neuroimaging this will lead to a paradigm shift, with the focus moving from small, clinical samples to large, population-based analyses. It is therefore important to reflect how principles that are commonly considered in Epidemiology are relevant for neuroimaging studies. Here, I will focus on the three canonical sources of systematic error or bias: confounding bias, selection bias and information bias.

Confounders are a set of factors that are causes of both the determinant and the outcome of a model. By not adjusting appropriately for confounders, the association will be biased and consequently lead to incorrect inferences on the association between the determinant and the outcome. In the field of neuroimaging, most studies consider four groups of confounders[438]: biological parameters (e.g. sex and head size), age at time of scanning, acquisition-related parameters (e.g. study site) and quality-related parameters (e.g. head motion).

However, the brain is shaped by numerous factors, which could all act as confounders in specific research questions. Outside of the four aforementioned confounding groups, little attention has been given to other potential confounders such as socioeconomic status, physical activity, and mental health. Mental health has its origins in the brain, but has been

shown to shape the brain as well[180]. Crucially, confounding and causal directions may vary from region to region and from vertex to vertex. For example, suffering from amblyopia – the lack of input from one eye to the brain – if untreated will lead to structural and functional changes that are specific to the occipital lobe. Another example would be brain networks, e.g., those related to reward, influencing lifestyle choices such as smoking, which in turn could be risk factors that affect the whole brain. It may be necessary to have separate confounding adjustment for different regions of the brain, or to develop other frameworks that can account for such mechanisms. The field would greatly benefit from formalization of confounder adjustment, in particularly a deeper understanding of which factors to consider for which brain imaging modalities, and how to adjust for them.

Selection bias arises when the participants included into the study differ from the population that the researchers attempted to recruit from, and when those differences affect the research question. A classic example of selection bias is the healthy worker effect. In 1965, when studying the mortality rate of amongst gas workers, Doll and colleagues discovered that gas workers as a whole had lower mortality rates than the general population[439]. This was surprising, as the idea already prevailed that coal-related exposures affect health. The problem, as it turns out, was that bias was present as the gas workers had physically demanding jobs. As such, those working were also on average more fit than the general population, which contains a heterogeneity of healthy and unhealthy people. Selection bias arose from the fact that the group of gas workers had been selected with criteria that were not accounted for. In the end, Doll and colleagues did adequately identify the problem, and showed that exposure differences within the group of gas workers did – as expected – increase mortality due to lung cancer and chronic bronchitis.

Selection bias is likely present in neuroimaging studies. Certain causes of selection bias are unique to the use of the MRI scanner. For example, individuals with contraindications such as metal implants or dental braces may not or cannot be scanned. Metal implants may be present due to cardiovascular problems, which could produce biases in studies exploring the role of vascular factors in shaping the brain. Dental braces may produce biases as access to proper dental care and the required funds for dental braces may result in fewer children from poorer families having braces. Indeed, children in Generation R whose parents had a lower-than-average income at study baseline were also less likely to have dental braces during early adolescence (data not shown). Participants may also decline to participate in neuroimaging studies due to fears related to being scanned. Claustrophobia can impair participants to lay still or even enter the MRI scanner, and certain individuals refuse to participate to MRI studies due to misconceptions of the risks that MRI scanning holds. Overall, these factors combined may only affect a fraction of the total sample – about 5-10% – but the presence of these factors

is consistent across studies. Furthermore, these factors may be amplified in specific scenarios, as the most recent wave of neuroimaging in the children of Generation R saw exclusion of around 30-40% of the sample due to dental braces (data not shown). The effects of selection bias in neuroimaging should thus be seriously considered and evaluated.

The relevance of selection bias likely varies by study. For example, the Rotterdam Study and Generation R capture a large and representative portion of the source (urban) populations at baseline. Both studies also collect information about participants that may relate to attrition, such as demographics and questions on personality and psychopathology. While both studies suffer from attrition, the available information can be used to infer how severe the bias is. Furthermore, these data can be used to attempt to correct for a part of the bias, for example through inverse probability weighting[440]. This will be much more difficult in the UK Biobank, where the response rate for participation was only 5.5% and potentially less representative of the general population[441]. Even so, all these population-based studies likely suffer from the neuroimaging-related causes of selection bias that I described in the previous paragraph. One study attempted to assess the influence of sample composition by weighing neuroimaging data from 1,162 children to approximate the general population of the United States[442]. The unweighted results suggested later maturation of cortical and subcortical regions than the weighted results and other studies in the literature. While this study focuses on generalizability, i.e., external validity, it illustrates how selection bias, i.e., internal validity, may also be compromised if we make invalid assumptions about how our study populations relate to the populations we intend to study. Despite all this, the role of selection bias has not been systematically studied or explored in neuroimaging or neuroepidemiology, and thus the consequences remain largely to be elucidated.

Information bias in neuroimaging has been studied more extensively. It has commonly been studied in the context of multisite studies, where differences in acquisition parameters may lead to slight differences in the images and thus in the processed data[443,444]. Differences in acquisition were limited in Generation R and the Rotterdam Study, as both studies rely on a single MRI scanner, and the acquisition parameters and software versions have not or barely been updated since the start of the study. Another source of information bias arises from the quality of the images given identical scanning parameters[445,446]. This commonly occurs due to susceptibility artifacts and motion artifacts. Susceptibility artifacts arise from signal deviation, e.g., distortions or changes. They arise due to moving fluids (e.g., blood) and air (e.g., air in the nasal cavity), but may also appear when ferromagnetic objects are still present (e.g., bonded dental retainers or dental implants). Motion artifacts can arise due to head motion, which may be more common in individuals with anxiety for the MRI or in those with the disease of interest, such as in attention-deficit/hyperactivity disorder[447,448] or Parkinson's disease. These artifacts

may affect the image quality, and in turn affect the results from processing said images. In fact, head motion has been shown to lead to an underestimation of the thickness of the cerebral cortex[57]. While methods have been introduced to account for head and body motion during scanning acquisition, they have not been implemented in the Generation R Study or the Rotterdam Study. An alternative solution would be to remove images of poor quality, as was done in this thesis, but this may lead to further selection bias if the causes of the artifacts related to the research question at hand.

FUTURE PERSPECTIVES

Diminishing returns

We do not fully understand human neurodevelopment and have not mapped all the factors that affect it. However, is it truly necessarily to elucidate every aspect of neurodevelopment before we claim that we understand it fully? Our understanding of brain development is lightyears ahead of De La Mettrie's statement from 1748 (see **chapter 1**)[1]. Still, his views on the shaping of the adult brain are not far from ours. A crucial point is outlined by Isaac Asimov in his essay "The Relatively of Wrong"[449]. In it, he argues that our view of the Earth has changed over time from being perceived as flat to being assumed a sphere. The Earth turned out to be an "oblate spheroid", given that it bulged around the equator, and yet further observations have found that the bulge is asymmetrical. While each incremental model represented reality more accurately, their practical implications were also increasingly more limited. To assume that the Earth is flat is still more wrong than to assume that the Earth is a sphere.

The field of Neuroepidemiology – and perhaps population-based research as a whole – is approaching a similar limit. Over time, bigger questions will have found an answer, and remaining research questions will address more niche topics whose potential impact will be increasingly smaller. For example, this can be seen when comparing the 2017 and 2020 reports of the Lancet commission on dementia prevention, intervention, and care[49,450]. Compared to the 2017 report, the 2020 report has added three new modifiable risk factors to their list of dementia risk factors: heavy alcohol use (45-65 years), traumatic brain injury (45-65 years) and air pollution (> 65 years). These factors cover a combined population attributable fraction of 6.5%, or about 1 in 15. While it cannot be denied that 1 in 15 is significant, the contribution of the three risk factors should be considered in context. First, the effect sizes of heavy alcohol use and air pollution are the smallest of all risk factors, and thus require significantly larger studies to detect accurately. Second, the gains were primarily made in the category of midlife, i.e., between 45 and 65 years. It is extremely costly to start a study with a large sample of middle-aged individuals that have to be followed up for decades in order to see if they develop

dementia. As time goes on, we will see fewer new additions to the potential late-life risk factors, and an increasing focus on earlier stages of life, as we have done with the ORACLE Study (**chapters 2.1 and 4.2**).

To prevent our scientific returns to diminish, I argue that several solutions need to be considered. First, data collection within new and established cohorts should largely be standardized to be able to combine results from those cohorts during later stages. This would include the design of the studies, the collection, storage and processing of data, and the retention of and communication with participants. This may need to be supported through centralized committees and teams of experts who can aid in setting up new or maintaining existing cohort studies. Of note is that studies should still be free to choose the tools of their liking, but the choices must be justified and balanced against the need for unification of the data. Second, as longitudinal data collection is costly, existing ongoing cohort studies should be incentivized to improve their data quality procedures and to obtain retrospective and prospective informed consent from participants to make the data of the cohort publicly available. Third, previously collected and analyzed data may still have value through combining it with other data or by utilizing novel analytical methods. Thus, funding organizations should make more funding available for methodological endeavors that aim to maximize the knowledge utilization of existing and future data. Finally, irrespective of whether and how the other solutions are implemented, the returns that we will see from future studies will on average decrease, up to the point where we will have to consider whether they are worth our investments. At its core, population-based research currently is about observation rather than intervention. When the factors that affect brain aging are generally understood, the need to implement interventions will increase while the need for further observational findings will decrease. As neuroepidemiologists, we must consider the future of our field, what will be the main advances that we expect to make in the coming decades, and how those could potentially translate into meaningful interventions in the general population. While such an approach does not prevent diminishing returns in Neuroepidemiology, it would ensure that our gains are still maximally utilized for the benefit of society.

Polygenic profiling

Polygenic risk scores allow utilization and application of the results of genome-wide association studies to clinical and other settings. By combining the effects of tens to millions of genetic variants into one number, polygenic scores capture an overarching signal into a number that is easy to interpret. Several challenges remain in the application of polygenic risk scores. Here, I will discuss problems and avenues related to accounting for etiological pathways, transancestral effects and causal variants.

Pathways

A prominent limitation of polygenic scores is that they combine variants that all may exert their effect through differing mechanisms, and thus it is unclear which mechanisms are at play. While this may not make much of a difference for prediction studies, it impairs interpretability of the findings in etiological studies. Several methods have been proposed to account for this limitation by creating subscores for specific underlying pathways, for example by hypothesis-driven grouping of selected variants by pathway[451,452] or by data-driven decomposition of phenotypic and genotypic data into components[453]. These methods are limited in that they only utilize independent top hits rather than all or large sets of genetic variants. Furthermore, both methods rely on external data to define the pathways, and thus they are limited by how representative those data are. Another limitation to these methods is that they cannot account for the complexity of pathways that exist. We may want to account for multiple polygenic scores simultaneously, but these could all be related through shared effects of genetic variants, i.e., pleiotropy, or overlapping pathways. One method that comes close to account for this is metaGRS[454], which relies on calculating polygenic scores for multiple endophenotypes and subsequently combining them into one score through regularization. Further developments in these kinds of methods will elevate the polygenic risk score from merely a predictive score to a multifaceted tool that can be used to elucidate the etiology of complex diseases.

Transancestral portability

A second limitation of polygenic scores is that current genome-wide studies generally rely on European samples and may thus not be applicable or relevant for non-European individuals[455-457], also known as poor generalizability or – more topically – poor transancestral portability. This effect could arise due to shared causal effects that are masked by differences in genetic architecture, or by true differences in the genetic variants that affect the phenotype of interest. The most straightforward way to address improve transancestral portability is to perform genome-wide studies in populations with heterogenous ethnicities. Such an approach can be used to elucidate transancestral genetic effects and thus improve the generation of polygenic scores[458]. The number of such studies outside of neuroepidemiology has steadily increased over the last years[457,459,460]. Within neuroepidemiology, there are a number of consortia that aim to address the transancestral nature of brain-related disease, such as the MEGASTROKE consortium for stroke[461], and the UNITED consortium for neurodegeneration[462]. Still, these consortia will likely not provide a transancestral alternative to other brain-related phenotypes, e.g., brain development or brain structure. Another approach to improve transancestral portability is by combining the results of genome-wide studies that were

performed in non-overlapping ethnic groups. More research groups in non-European countries are amassing large, genetic databases, such as the BioBank Japan Project that focuses on 47 common diseases in over 200,000 participants[463]. Results from their genome-wide studies can be combined with European-centered studies through methods that estimate which effects are common to both populations, and which effects are unique. This has been done through adjustment for the ancestral linkage disequilibrium structure[458], and by accounting for ethnical differences in cell-type specific regulatory elements[464]. These methods are still in their infancy, but they show the potential of methodology in addressing and amending the European-focused nature of polygenic score research.

Causal variants

A third limitation for the application of polygenic scores is the handling of causal variants. Genome-wide studies tend to identify hundreds or thousands of genetic variants that associate with the phenotype, many of which are clustered in distinct regions. The focus is typically put on lead variants, i.e., variants within a region that have the most statistically significant association or that stand out through some other metric. However, these lead variants are generally not the variants that cause the phenotypic association[465]. For example, the true causal variant may not have been assessed, and thus the identified lead variant is a mere proxy. Another consideration is that genome-wide studies tend to focus on single nucleotide polymorphisms, which in certain cases may be a marker for accumulated complex structural variants such as copy number variants[466]. It may occur that a single nucleotide polymorphism associates with a phenotype simply because a subset of individuals has a correlated structural variant in close proximity. The misidentification of lead variants as causal variants may impair the applicability of polygenic scores. Lead variants may associate with causal variants in one population but not the other. In other words, polygenic scores will appear high if a person has all the risk variants, but in reality that person may not have any of the true causal risk variants. This problem could be addressed by improving efforts for genetic fine-mapping, i.e., using external information such as linkage disequilibrium or functional annotation to understand which variants likely lead to a change in gene expression. Another potential avenue is the fact that the degree to which genetic variants affect gene expression can vary by cell type[467], for example due to cell-type differences in expression of regulatory elements. Methods to account better model causally informed polygenic scores are still in their infancy, and further work would greatly improve the utility of polygenic scores in clinical settings.

Methods and models

Ever increasing corpora of neuroimaging data have become available through population-based cohorts and other studies. While such data immediately lead to novel insights, these tend to be "low-hanging fruit", i.e., projects that are straightforward to conduct. What remains are the "high-hanging fruits", i.e., projects that require more complex approaches in both methods and models. Large leaps in our understanding of the brain can be made through the application of novel methods or the novel application of established methods. A commonly lauded group of methods is deep learning[468,469]. These methods are generally used to find patterns in structured and unstructured data without making explicit assumptions about the patterns. Deep learning methods are particularly useful for spatially organized data, such as the 3-dimensional voxel-wise brain, the surface-based vertex-wise cortex and the linear structure of the genome. Other high-hanging fruit projects will be reached through methods that can consider whole datasets or multiple datasets at once. Examples include canonical correlation analysis (CCA) for correlating components in two or more datasets[470,471] and multi-omics factor analysis for omics data[472].

While the implementation of such methods may seem straightforward – we simply apply the math to a new dataset – this is not necessarily the case. It may take years before the limitations of methods are noted. This is clearly illustrated by an example CCA. In 2017, CCA was used in patients with depression to explore the association of functional MRI data and depressive symptoms[473]. From the CCA results, the researchers concluded that there were four neurophysiological subtypes amongst the patients, each with a distinct pattern of functional connectivity in the brain. However, subsequent attempts by other groups to replicate these findings were without avail. As it turns out, CCA requires specific adjustments for inflated false positive results[474,475]. This illustrates that application of novel methods can lead to groundbreaking insights, but that proper application of such methods may take years to refine.

Further advances in modeling will also prove to be essential. The BrainAGE model provides a new perspective to consider how the brain is shaped[323,476], yet it is based on a simple idea. The concept arose from a desire to better understand biological aging of the brain, as it seemed to play a role in a number of late-life neurodegenerative diseases. Still, progress was limited as the biological age of the brain was generally assessed with global measures, such as total brain volume or brain-wide atrophy rates. BrainAGE was innovative as it used high dimensional data, which was accomplished by using relevance vector regression, a type of Bayesian regression that had previously not been used in this context. BrainAGE is an example of how novel applications of methods allow us to generate conceptual models to address unsolved

problems in the field. Stated in another way, innovative models may only arise if we as a field are aware of the problems that remain to be elucidated.

CONCLUDING REMARKS

In this chapter, I gave a brief overview of the chapters and their implications, the methodological considerations, and speculations on future directions of research and academia. Overall, this thesis provides an array of insights into the causes and consequences of brain development and deterioration. With this, we have taken another step towards an understanding of the brain and the factors that shape it.

SUMMARY / SAMENVATTING

7

SUMMARY

Our brains change over a lifetime. During early-life the brain develops, and as life continues we can see that the various tissues within the brain start to decay and accumulate damage. This, in turn, can have implications for our daily functioning and the development of diseases. Interestingly, although humans all tend to follow the same overall trajectory, there are numerous individual deviations from this trajectory when considering the whole lifespan. Brain health is widely studied, but plenty of questions still remain, which I aimed to address in this thesis. My general aim was to understand the causes and consequences of brain health across the lifespan. In particular, we focused on early-life, i.e. childhood, and late-life, i.e. older adulthood. To do so, we used data from the Generation R Study – a birth cohort from fetal life until adolescence – and the Rotterdam Study – a cohort from 45 years until the end of life.

The scope of my research aim required additional data collection and the implementation of new tools. In **Chapter 2.1**, we introduced the ORACLE Study. The study included over 2,000 parents from Generation R who had participated in broad phenotyping of their health through physical measures and questionnaires, and who visited our research center for further cognitive and neuroimaging assessment roughly 15 years later. The ORACLE Study bridges the gap between the children from the Generation R Study, and the individuals from the Rotterdam Study aged 45 years and older. The data present a unique opportunity to explore how health during early and middle adulthood relates to and affects subsequent brain health. In **Chapter 2.2**, we introduced the QDECR tool, which is an analysis tool for imaging data of the brain. The amount of neuroimaging data within population-based cohorts has increased drastically over the last decade, but few tools provided functionality for analyses that were common within population-based research. QDECR was designed to address these problems; for example, it allows researchers to apply common methods for confounding and selection bias correction. Furthermore, it provides a framework that can be expanded with new statistical models.

Next, we explored how genetic, vascular, and other mechanisms affect brain health and cognition.

To study the genetic mechanisms, we utilized polygenic scores for brain health, and related them to early-life brain structure and other relevant outcomes in Generation R. Polygenic scores are scores that represent one's genetic predisposition to develop a specific trait. They are typically calculated based on previously performed genome-wide association studies, which explore how the genome relates to a trait. In **Chapter 3.1**, we calculated polygenic scores for different subcortical regions of the adult brain and showed that these scores related

to subcortical brain volumes during both infancy and the school age. Furthermore, the volumes during infancy only weakly mediated the effect of the polygenic scores on the volumes during school age, suggesting that the polygenic scores represent multiple genetic mechanisms across different developmental stages. We also explored whether genetic burden for late-life neurodegenerative diseases related to school age brain structure, cognitive function, and behavior (**Chapter 3.2**). Overall, we found no clear evidence that this was the case for polygenic scores for Alzheimer's disease, Parkinson's disease, and frontotemporal dementia. Contrary to earlier findings, polygenic scores for Alzheimer's disease did not relate to volumetric measures of the frontal cortex, the hippocampus, and other implicated brain regions. We did show a strong association between alleles of the Apolipoprotein E gene – one of the best known causal genes for Alzheimer's disease – with blood serum lipid levels, demonstrating that the genes can have an effect during early life, just not necessarily in areas where their influence is exerted during late-life.

For vascular mechanisms, we primarily focused on vascular risk factors and their association with brain structure and cognition. Vascular risk factors are a set of factors that infer a higher risk to develop cerebrovascular disease, and include high blood pressure, arterial stiffness, obesity, and smoking. These, in turn, have been shown to relate to a higher incidence of late-life neurodegenerative disease. In **Chapter 4.1**, we studied the association of blood pressure and arterial stiffness with cognitive functioning in both the Generation R Study and the Rotterdam Study. In the children, only higher levels of diastolic blood pressure were associated with lower non-verbal intelligence. In the adults, both blood pressure and arterial stiffness related to global cognitive function. In particular, we found that diastolic blood pressure had a U-shaped association with cognition, in that both lower and higher levels related to poorer cognitive function. In **Chapter 4.2**, we utilized data from the ORACLE Study to explore whether vascular risk factors during early adulthood, i.e. 18 to 35 years of age, relate to brain structure and cognition roughly 15 years later. We found that this was not the case, with some exceptions. For example, higher systolic and diastolic blood pressure during early adulthood seemed to relate to more white matter lesions during the subsequent visit. If these findings are replicated, this would provide an earlier mechanism relating vascular risk factors and neuropathology than previously shown.

Finally, we explored a number of other mechanisms, collectively referred to as "reserve mechanisms". Reserve is the idea that humans have a capacity to deal with the consequences of neuropathology and aging. Brain reserve usually refers to the mass of the brain, the number of neurons, and other neurobiological quantities. Cognitive reserve refers to our flexibility, either in cognition through efficient cognitive strategies or in biology through recruitment of alternative neural networks to counteract other losses. In **Chapter 5.1**, we explored the effect

of adverse childhood events on the maturation of the brain during adolescence. Adverse childhood events have been shown to accelerate the onset of puberty, particularly when of a threatening rather than depriving nature. We studied whether this is also true for aging of the brain. For each child in Generation R, we used machine learning to estimate the age of their brains. Then, we showed that weak markers for threatening but not depriving childhood events seemed to relate to a lower age of the brain. The findings point to a mechanism that may affect cognitive and brain reserve later in life. In **Chapter 5.2**, we examined the gyrification or folding of the cerebral cortex from middle adulthood towards the end of life. We showed that gyrification reduces linearly with age, although at varying rates in different regions, and gyrification even seems to increase in other regions. Furthermore, we showed that the degree of gyrification is related – in certain regions – to specific cognitive tasks. In **Chapter 5.3**, we then explored how markers for early-life and late-life cognitive and brain reserve relate to the incidence of dementia in the Rotterdam Study. We showed that late-life markers of both cognitive and brain reserve were much more strongly associated with the incidence of dementia in the next few years than early-life markers. Furthermore, late-life cognitive and brain reserve seemed to contribute to an equal extent, and they seemed to interact in that having both low cognitive and brain reserve related to a much higher chance of being diagnosed with dementia in subsequent years.

Finally, the findings are discussed in **Chapter 6**. In particularly, the chapter provides an exposition of the limitations and strengths of the studies. I also outlined my reasoning that the current definition of cognitive reserve is flawed, and that we need to find a way to more clearly define what it entails biologically. The chapter also covers the relevance of the way we study the brain, as well as the role of confounding, selection bias, and information bias in neuroimaging research. Finally, I covered my views on the future directions of the field as a whole, particularly focusing on the direction of Neuroepidemiology, the potential and obstacles for genetic profiling of disease, and the relevance of innovations in methods and models for our understanding of the causes and consequences of the changing brain.

SAMENVATTING

Onze hersenen veranderen door het leven heen. Tijdens het vroege leven ontwikkelen de hersenen zich, en door het leven heen zien we dat verschillende delen van de hersenen beginnen te vervallen en schade op te bouwen. Dit heeft weer implicaties voor ons dagelijkse functioneren en het ontstaan van ziektes. Hoewel alle mensen ongeveer hetzelfde traject van hersenontwikkeling volgen, wijken we allemaal op onze eigen manier af van dit traject als we naar het hele leven kijken. Hoewel de gezondheid van de hersenen veel is bestudeerd, zijn er nog veel openstaande vragen, en een deel daarvan probeer ik te beantwoorden in deze dissertatie. Mijn overkoepelende doel was om de oorzaken en gevolgen van de gezondheid van de hersenen te vinden door het leven heen. Hierbij hebben we ons gericht op het vroege leven, c.q. kindertijd, en het late leven, c.q. de latere volwassenheid. Om dit te doen, hebben we gebruik gemaakt van data van de Generation R Study – een geboortecohort van het foetale leven tot aan de adolescentie – en de Rotterdam Study – een cohort van 45 jaar tot het einde van het leven.

Het doel van mijn onderzoek vereiste aanvullende dataverzameling en het implementeren van nieuwe methodes. In **Hoofdstuk 2.1** introduceerden we de ORACLE Study. Deze studie omvat 2.000 ouders uit Generation R die eerder deel hebben genomen aan uitgebreide fenotypering van hun gezondheid door fysieke metingen en vragenlijsten, en die ons onderzoekscentrum na 15 jaar opnieuw bezochten voor verdere cognitieve en hersenmetingen. De ORACLE Study overbrugt het gat in leeftijd tussen de kinderen van Generation R en de deelnemers van 45 jaar en ouder uit de Rotterdam Study. De data representeren een unieke kans om te onderzoeken hoe de gezondheid tijdens de vroege en latere volwassenheid samenhangt met de hersengezondheid later in het leven. In **Hoofdstuk 2.2** introduceren we QDECR, een analysetool voor hersenscans. In het afgelopen decennium is er binnen bevolkingsonderzoek een ongekende toename geweest van het aantal gemaakte hersenscans, maar er zijn weinig tools beschikbaar die hersenscans analyseren volgens epidemiologische principes. QDECR was ontwikkeld om dit mogelijk te maken voor onderzoekers, zodat men bijvoorbeeld gemakkelijk kan corrigeren voor bias zoals confounding en selectiebias. Daarnaast is QDECR een framework wat uitgebreid kan worden met nieuwe statistische modellen.

Vervolgens onderzochten we hoe genetische, vasculaire en andere mechanismen invloed hebben op hersengezondheid en cognitie.

We bestudeerden genetische mechanismes met behulp van polygene scores voor hersengezondheid, en we keken of deze samenhingen met de hersenstructuur en andere processen tijdens het vroege leven. Polygene scores zijn scores die iemands genetische aanleg

voor een bepaalde uitkomst kwantificeren. Ze worden meestal berekend aan de hand van genoombrede associatiestudies, die weer gebruikt worden om een beeld te krijgen van hoe het genoom samenhangt met die uitkomst. In **Hoofdstuk 3.1** berekenden we binnen Generation R polygene scores voor subcorticale gebieden van het volwassen brein, en toonden aan dat deze scores samenhingen met subcorticale volumes in zowel babies als schoolkinderen. Daarnaast vonden we dat de volumes in babies hoogstens deels de volumes in de schoolkinderleefitjd konden verklaren, wat suggereert dat de polygene scores genetische mechanismen vangen uit verschillende ontwikkelingsfases. In **Hoofdstuk 3.2** onderzochten we of polygene scores voor neurodegeneratieve ziektes op latere leeftijd samenhingen met de hersenstructuur, cognitie en gedrag in schoolkinderen. Hier vonden we geen bewijs voor met scores voor de ziektes van Alzheimer en Parkinson, noch voor frontotemporale dementie. Dit was tegenstrijdig met eerdere bevindingen in de literatuur. Wel vonden we een associatie tussen het Apolipoproteïne E gen – de meest bekende oorzaak van de ziekte van Alzheimer – met lipides in het bloed, wat aantoont dat deze genen wel degelijk een effect hebben tijdens het vroege leven, maar niet per sé in dezelfde gebieden als in het latere leven.

Qua vasculaire mechanismes hebben we ons primair gericht op vasculaire risicofactoren en hun associatie met hersenstructuur en cognitie. Vasculaire risicofactoren zijn een aantal factoren die een hoger risico geven op cerebrovasculaire ziektes, waaronder een hoge bloeddruk, vaatwandstijfheid, obesitas en roken. Deze factoren hangen ook samen met een hogere incidentie van neurodegeneratieve ziektes op latere leeftijd. In **Hoofdstuk 4.1** bestudeerden we of bloeddruk en vaatwandstijfheid associeerden met cognitie in zowel Generation R als de Rotterdam Study. In de kinderen vonden we dat hogere diastole bloeddruk samenhing met lagere non-verbale intelligentie. In de Rotterdam Study vonden we dat zowel bloeddruk als vaatwandstijfheid samenhingen met intelligentie. Bovendien vonden we dat diastole bloeddruk een U-vormige associatie had met cognitie, waarbij zowel lage als hoge diastole bloeddruk associeerde met een slechter cognitief functioneren. In **Hoofdstuk 4.2** gebruikten we data van de ORACLE Study om te onderzoeken of vasculaire risicofactoren tijdens de vroege volwassenheid (18-35 jaar) associeerden met hersenstructuur en cognitie ongeveer 15 jaar later. We vonden dat dit niet het geval was, met een aantal uitzonderingen. Zo vonden we bijvoorbeeld dat hogere systole en diastole bloeddruk samenhingen met meer wittestoflaesies op latere leeftijd. Als deze bevindingen gerepliceerd worden, dan wijst dit op een link tussen vasculaire risicofactoren en hersenschade op een jongere leeftijd dan voorheen werd gedacht.

Als laatste keken we naar andere mechanismen, samengevat als “reserve mechanismen”. Reserve is het concept dat mensen een capaciteit hebben om om te gaan met de gevolgen van

hersenschade en veroudering. Breinreserve slaat op de massa van de hersenen, het aantal neuronen, en andere biologische kwantiteiten. Cognitieve reserve slaat op flexibiliteit, enerzijds in cognitie door handige cognitieve strategieën en anderzijds in biologie door het rekruteren van alternatieve, compensatoire hersennetwerken. In **Hoofdstuk 5.1** keken we naar het effect van ingrijpende jeugdervaringen op ontwikkeling van het adolescente brein. Eerder is aangetoond dat ingrijpende jeugdervaringen de puberteit vervroegen, vooral als het om ervaringen van dreigende aard gaat in plaats van depriverend. Wij onderzochten of dit ook gold voor de hersenen. Voor elk kind van Generation R bepaalden we via machine learning de leeftijd van de hersenen. Vervolgens toonden we aan dat zwakke markers voor jeugdervaringen van dreigende maar niet depriverende aard inderdaad samenhingen met een oudere leeftijd van de hersenen. Deze bevindingen wijzen op een mechanisme wat mogelijk cognitieve en breinreserve op latere leeftijd beïnvloedt. In **Hoofdstuk 5.2** brachten we de gyrificatie – de mate van golving in de cortex – in kaart voor het latere leven. We vonden dat gyrificatie lineair met leeftijd afneemt, hoewel het tempo verschilt per hersengebied, en er ook gebieden zijn waar het toeneemt. Verder vonden we dat gyrificatie in bepaalde hersengebieden samenhing met specifieke cognitietaken. In **Hoofdstuk 5.3** keken we in de Rotterdam Study hoe markers voor cognitieve en breinreserve in het vroege en late leven samenhingen met de incidentie van dementie. We vonden dat zowel cognitieve als breinreserve uit het late leven sterker samenhingen met de incidentie van dementie dan die uit het vroege leven. De associaties waren voor cognitieve en breinreserve uit het late leven even sterk. Tevens vonden we dat ze interacteerden, dus dat de kans op een dementiediagnose aanzienlijk toeneemt als beide reserves laag zijn.

Alle bevindingen worden in **Hoofdstuk 6** verder besproken. Deze geeft een overzicht van de limitaties en sterke kanten van de studies. Verder geef ik een uiteenzetting waarom het concept van reserve onvolmaakt is, en dat we moeten zoeken naar een beschrijving die gegrond is in de biologie. Het hoofdstuk dekt verder de relevantie van de manier waarop we het brein bestuderen, en de rol van confounding, selectiebias en informatiebias in hersenwetenschappelijk onderzoek. Als laatste dekt het hoofdstuk mijn kijk op de toekomstige richting van het veld als geheel, met name de richting van de neuroepidemiologie, de potentie en obstakels voor genetische profilering van ziekte, en de relevantie van innovaties in methodes en modellen voor ons begrip van de oorzaken en gevolgen van het veranderde brein.

BIBLIOGRAPHY

1 De La Mettrie, J. O. (1994) *Man a Machine; and, Man a Plant*. (Hackett Publishing Company).

2 Stiles, J. & Jernigan, T. L. (2010) The basics of brain development. *Neuropsychol Rev* **20**, 327-348, doi:10.1007/s11065-010-9148-4.

3 Vinke, E. J. *et al.* (2018) Trajectories of imaging markers in brain aging: the Rotterdam Study. *Neurobiol Aging* **71**, 32-40, doi:10.1016/j.neurobiolaging.2018.07.001.

4 Den Dunnen, W. F. *et al.* (2008) No disease in the brain of a 115-year-old woman. *Neurobiol Aging* **29**, 1127-1132, doi:10.1016/j.neurobiolaging.2008.04.010.

5 Takao, M., Hirose, N., Arai, Y., Mihara, B. & Mimura, M. (2016) Neuropathology of supercentenarians - four autopsy case studies. *Acta Neuropathol Commun* **4**, 97, doi:10.1186/s40478-016-0368-6.

6 Peters, R. *et al.* (2019) Combining modifiable risk factors and risk of dementia: a systematic review and meta-analysis. *BMJ Open* **9**, e022846, doi:10.1136/bmjopen-2018-022846.

7 Ding, J. *et al.* (2020) Antihypertensive medications and risk for incident dementia and Alzheimer's disease: a meta-analysis of individual participant data from prospective cohort studies. *Lancet Neurol* **19**, 61-70, doi:10.1016/S1474-4422(19)30393-X.

8 Lusebrink, F., Sciarra, A., Mattern, H., Yakupov, R. & Speck, O. (2017) T1-weighted in vivo human whole brain MRI dataset with an ultrahigh isotropic resolution of 250 mum. *Sci Data* **4**, 170032, doi:10.1038/sdata.2017.32.

9 Barisano, G. *et al.* (2019) Clinical 7 T MRI: Are we there yet? A review about magnetic resonance imaging at ultra-high field. *Br J Radiol* **92**, 20180492, doi:10.1259/bjr.20180492.

10 Grasby, K. L. *et al.* (2020) The genetic architecture of the human cerebral cortex. *Science* **367**, doi:10.1126/science.aay6690.

11 Satizabal, C. L. *et al.* (2019) Genetic architecture of subcortical brain structures in 38,851 individuals. *Nature Genetics*, doi:10.1038/s41588-019-0511-y.

12 Hibar, D. P. *et al.* (2017) Novel genetic loci associated with hippocampal volume. *Nat Commun* **8**, 13624, doi:10.1038/ncomms13624.

13 Ikram, M. A. *et al.* (2015) The Rotterdam Scan Study: design update 2016 and main findings. *Eur J Epidemiol* **30**, 1299-1315, doi:10.1007/s10654-015-0105-7.

14 Miller, K. L. *et al.* (2016) Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nat Neurosci* **19**, 1523-1536, doi:10.1038/nn.4393.

15 German National Cohort, C. (2014) The German National Cohort: aims, study design and organization. *Eur J Epidemiol* **29**, 371-382, doi:10.1007/s10654-014-9890-7.

16 Jansen, I. E. *et al.* (2019) Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat Genet* **51**, 404-413, doi:10.1038/s41588-018-0311-9.

17 Kuzma, E. *et al.* (2018) Stroke and dementia risk: A systematic review and meta-analysis. *Alzheimers Dement* **14**, 1416-1426, doi:10.1016/j.jalz.2018.06.3061.

18 McGrath, E. R. *et al.* (2017) Blood pressure from mid- to late life and risk of incident dementia. *Neurology* **89**, 2447-2454, doi:10.1212/WNL.0000000000004741.

19 Abell, J. G. *et al.* (2018) Association between systolic blood pressure and dementia in the Whitehall II cohort study: role of age, duration, and threshold used to define hypertension. *Eur Heart J*, doi:10.1093/eurheartj/ehy288.

20 Stern, Y. *et al.* (2018) Whitepaper: Defining and investigating cognitive reserve, brain reserve, and brain maintenance. *Alzheimers Dement*, doi:10.1016/j.jalz.2018.07.219.

21 Perneczky, R. *et al.* (2019) Translational research on reserve against neurodegenerative disease: consensus report of the International Conference on Cognitive Reserve in the Dementias and the Alzheimer's Association Reserve, Resilience and Protective Factors Professional Interest Area working groups. *BMC Med* **17**, 47, doi:10.1186/s12916-019-1283-z.

22 Kooijman, M. N. *et al.* (2016) The Generation R Study: design and cohort update 2017. *Eur J Epidemiol* **31**, 1243-1264, doi:10.1007/s10654-016-0224-9.

23 Ikram, M. A. *et al.* (2020) Objectives, design and main findings until 2020 from the Rotterdam Study. *Eur J Epidemiol*, doi:10.1007/s10654-020-00640-5.

24 Rothman, K. J., Greenland, S. & Lash, T. L. (2008) *Modern Epidemiology*. 3rd edition edn, (Lippincott Williams & Wilkins).

25 Prince, M. J. *et al.* (2015) The burden of disease in older people and implications for health policy and practice. *Lancet* **385**, 549-562, doi:10.1016/S0140-6736(14)61347-7.

26 G. B. D. Neurological Disorders Collaborator Group. (2017) Global, regional, and national burden of neurological disorders during 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet Neurol* **16**, 877-897, doi:10.1016/S1474-4422(17)30299-5.

27 Seifan, A., Schelke, M., Obeng-Aduasare, Y. & Isaacson, R. (2015) Early Life Epidemiology of Alzheimer's Disease--A Critical Review. *Neuroepidemiology* **45**, 237-254, doi:10.1159/000439568.

28 Ritchie, K., Ritchie, C. W., Yaffe, K., Skoog, I. & Scarmeas, N. (2015) Is late-onset Alzheimer's disease really a disease of midlife? *Alzheimers Dement (N Y)* **1**, 122-130, doi:10.1016/j.trci.2015.06.004.

29 Braak, H. & Braak, E. (1997) Frequency of stages of Alzheimer-related lesions in different age categories. *Neurobiol Aging* **18**, 351-357, doi:10.1016/s0197-4580(97)00056-0.

30 Dekhtyar, S., Wang, H. X., Fratiglioni, L. & Herlitz, A. (2016) Childhood school performance, education and occupational complexity: a life-course study of dementia in the Kungsholmen Project. *Int J Epidemiol* **45**, 1207-1215, doi:10.1093/ije/dyw008.

31 Dekhtyar, S. *et al.* (2015) A Life-Course Study of Cognitive Reserve in Dementia--From Childhood to Old Age. *Am J Geriatr Psychiatry* **23**, 885-896, doi:10.1016/j.jagp.2015.02.002.

32 Wang, H. X., MacDonald, S. W., Dekhtyar, S. & Fratiglioni, L. (2017) Association of lifelong exposure to cognitive reserve-enhancing factors with dementia risk: A community-based cohort study. *PLoS Med* **14**, e1002251, doi:10.1371/journal.pmed.1002251.

33 Larsson, S. C. *et al.* (2017) Modifiable pathways in Alzheimer's disease: Mendelian randomisation analysis. *BMJ* **359**, j5375, doi:10.1136/bmj.j5375.

34 Horder, H. *et al.* (2018) Midlife cardiovascular fitness and dementia: A 44-year longitudinal population study in women. *Neurology* **90**, e1298-e1305, doi:10.1212/WNL.0000000000005290.

35 Walker, K. A. *et al.* (2019) Association of Midlife to Late-Life Blood Pressure Patterns With Incident Dementia. *JAMA* **322**, 535-545, doi:10.1001/jama.2019.10575.

36 Gregson, J. *et al.* (2019) Blood pressure and risk of dementia and its subtypes: a historical cohort study with long-term follow-up in 2.6 million people. *Eur J Neurol*, doi:10.1111/ene.14030.

37 Albanese, E. *et al.* (2017) Body mass index in midlife and dementia: Systematic review and meta-regression analysis of 589,649 men and women followed in longitudinal studies. *Alzheimers Dement (Amst)* **8**, 165-178, doi:10.1016/j.dadm.2017.05.007.

38 Kivimaki, M. *et al.* (2018) Body mass index and risk of dementia: Analysis of individual-level data from 1.3 million individuals. *Alzheimers Dement* **14**, 601-609, doi:10.1016/j.jalz.2017.09.016.

39 Najar, J. *et al.* (2019) Cognitive and physical activity and dementia: A 44-year longitudinal population study of women. *Neurology* **92**, e1322-e1330, doi:10.1212/WNL.0000000000007021.

40 Palta, P. *et al.* (2019) Leisure-time physical activity sustained since midlife and preservation of cognitive function: The Atherosclerosis Risk in Communities Study. *Alzheimers Dement* **15**, 273-281, doi:10.1016/j.jalz.2018.08.008.

41 Andel, R. *et al.* (2008) Physical exercise at midlife and risk of dementia three decades later: a population-based study of Swedish twins. *J Gerontol A Biol Sci Med Sci* **63**, 62-66, doi:10.1093/gerona/63.1.62.

42 Kivimaki, M. *et al.* (2019) Physical inactivity, cardiometabolic disease, and risk of dementia: an individual-participant meta-analysis. *BMJ* **365**, l1495, doi:10.1136/bmj.l1495.

43 Russ, T. C., Lee, I. M., Sesso, H. D., Muniz-Terrera, G. & Batty, G. D. (2019) Five-decade trajectories in body mass index in relation to dementia death: follow-up of 33,083 male Harvard University alumni. *Int J Obes (Lond)* **43**, 1822-1829, doi:10.1038/s41366-018-0274-z.

44 Sindi, S. *et al.* (2017) Midlife Work-Related Stress Increases Dementia Risk in Later Life: The CAIDE 30-Year Study. *J Gerontol B Psychol Sci Soc Sci* **72**, 1044-1053, doi:10.1093/geronb/gbw043.

45 Donley, G. A. R., Lonnroos, E., Tuomainen, T. P. & Kauhanen, J. (2018) Association of childhood stress with late-life dementia and Alzheimer's disease: the KIHD study. *Eur J Public Health* **28**, 1069-1073, doi:10.1093/eurpub/cky134.

46 Sindi, S. *et al.* (2017) Midlife Work-Related Stress is Associated with Late-Life Gray Matter Volume Atrophy. *J Alzheimers Dis Rep* **1**, 219-227, doi:10.3233/ADR-170035.

47 Gilsanz, P. *et al.* (2019) Stressors in Midlife and Risk of Dementia: The Role of Race and Education. *Alzheimer Dis Assoc Disord* **33**, 200-205, doi:10.1097/WAD.0000000000000313.

48 Cabeza, R. *et al.* (2018) Maintenance, reserve and compensation: the cognitive neuroscience of healthy ageing. *Nat Rev Neurosci*, doi:10.1038/s41583-018-0068-2.

49 Livingston, G. *et al.* (2020) Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *Lancet* **396**, 413-446, doi:10.1016/S0140-6736(20)30367-6.

50 Ou, Y. N. *et al.* (2020) Blood Pressure and Risks of Cognitive Impairment and Dementia: A Systematic Review and Meta-Analysis of 209 Prospective Studies. *Hypertension* **76**, 217-225, doi:10.1161/HYPERTENSIONAHA.120.14993.

51 Basit, S., Wohlfahrt, J. & Boyd, H. A. (2018) Pre-eclampsia and risk of dementia later in life: nationwide cohort study. *BMJ* **363**, k4109, doi:10.1136/bmj.k4109.

52 Kruithof, C. J. *et al.* (2014) The Generation R Study: Biobank update 2015. *European Journal of Epidemiology* **29**, 911-927, doi:10.1007/s10654-014-9980-6.

53 Coolman, M. *et al.* (2010) Medical record validation of maternally reported history of preeclampsia. *J Clin Epidemiol* **63**, 932-937, doi:10.1016/j.jclinepi.2009.10.010.

54 Brown, M. A., Lindheimer, M. D., de Swiet, M., Van Assche, A. & Moutquin, J. M. (2001) The classification and diagnosis of the hypertensive disorders of pregnancy: statement from the International Society for the Study of Hypertension in Pregnancy (ISSHP). *Hypertens Pregnancy* **20**, IX-XIV, doi:10.1081/PRG-100104165.

55 White, T. *et al.* (2018) Paediatric population neuroimaging and the Generation R Study: the second wave. *Eur J Epidemiol* **33**, 99-125, doi:10.1007/s10654-017-0319-y.

56 Marcus, D. S., Olsen, T. R., Ramaratnam, M. & Buckner, R. L. (2007) The Extensible Neuroimaging Archive Toolkit: an informatics platform for managing, exploring, and sharing neuroimaging data. *Neuroinformatics* **5**, 11-34, doi:10.1385/ni:5:1:11.

57 White, T. *et al.* (2018) Automated quality assessment of structural magnetic resonance images in children: Comparison with visual inspection and surface-based reconstruction. *Hum Brain Mapp* **39**, 1218-1231, doi:10.1002/hbm.23911.

58 Morris, Z. *et al.* (2009) Incidental findings on brain magnetic resonance imaging: systematic review and meta-analysis. *BMJ* **339**, b3016.

59 Bos, D. *et al.* (2016) Prevalence, Clinical Management, and Natural Course of Incidental Findings on Brain MR Images: The Population-based Rotterdam Scan Study. *Radiology* **281**, 507-515, doi:10.1148/radiol.2016160218.

60 Gibson, L. M. *et al.* (2017) Impact of detecting potentially serious incidental findings during multi-modal imaging. *Wellcome Open Res* **2**, 114, doi:10.12688/wellcomeopenres.13181.3.

61 Hoogendam, Y. Y., Hofman, A., van der Geest, J. N., van der Lugt, A. & Ikram, M. A. (2014) Patterns of cognitive function in aging: the Rotterdam Study. *Eur J Epidemiol* **29**, 133-140, doi:10.1007/s10654-014-9885-4.

62 Van der Elst, W., van Boxtel, M. P., van Breukelen, G. J. & Jolles, J. (2005) Rey's verbal learning test: normative data for 1855 healthy participants aged 24-81 years and the influence of age, sex, education, and mode of presentation. *J Int Neuropsychol Soc* **11**, 290-302, doi:10.1017/S1355617705050344.

63 Scarpina, F. & Tagini, S. (2017) The Stroop Color and Word Test. *Front Psychol* **8**, 557, doi:10.3389/fpsyg.2017.00557.

64 Stroop, J. R. (1935) Studies of interference in serial verbal reactions. *Journal of Experimental Psychology* **18**, 643-662, doi:10.1037/h0054651.

65 Van der Elst, W., Van Boxtel, M. P., Van Breukelen, G. J. & Jolles, J. (2006) The Stroop color-word test: influence of age, sex, and education; and normative data for a large sample across the adult age range. *Assessment* **13**, 62-79, doi:10.1177/1073191105283427.

66 Van der Elst, W., Van Boxtel, M. P., Van Breukelen, G. J. & Jolles, J. (2006) The Letter Digit Substitution Test: normative data for 1,858 healthy participants aged 24-81 from the Maastricht Aging Study (MAAS): influence of age, education, and sex. *J Clin Exp Neuropsychol* **28**, 998-1009, doi:10.1080/13803390591004428.

67 Troyer, A. K. (2000) Normative Data for Clustering and Switching on Verbal Fluency Tasks. *Journal of Clinical and Experimental Neuropsychology* **22**, 370-378, doi:10.1076/1380-3395(200006)22:3;1-v;ft370.

68 Troyer, A. K., Moscovitch, M. & Winocur, G. (1997) Clustering and switching as two components of verbal fluency: evidence from younger and older healthy adults. *Neuropsychology* **11**, 138-146.

69 Goni, J. *et al.* (2011) The semantic organization of the animal category: evidence from semantic verbal fluency and network theory. *Cogn Process* **12**, 183-196, doi:10.1007/s10339-010-0372-x.

70 Podell, K. (2011) in *Encyclopedia of Clinical Neuropsychology* (eds Jeffrey S. Kreutzer, John DeLuca, & Bruce Caplan) 2086-2088 (Springer New York).

71 Tiffin, J. & Asher, E. J. (1948) The Purdue pegboard; norms and studies of reliability and validity. *J Appl Psychol* **32**, 234-247.

72 Killgore, W. D. S., Glahn, D. C. & Casasanto, D. J. (2005) Development and Validation of the Design Organization Test (DOT): A Rapid Screening Instrument for Assessing Visuospatial Ability. *Journal of Clinical and Experimental Neuropsychology* **27**, 449-459, doi:10.1080/13803390490520436.

73 Callisaya, M. L. *et al.* (2009) A population-based study of sensorimotor factors affecting gait in older people. *Age Ageing* **38**, 290-295, doi:10.1093/ageing/afp017.

74 Pearson, K. G. (2004) Generating the walking gait: role of sensory feedback. *Prog Brain Res* **143**, 123-129, doi:10.1016/S0079-6123(03)43012-4.

75 Sahyoun, C., Floyer-Lea, A., Johansen-Berg, H. & Matthews, P. M. (2004) Towards an understanding of gait control: brain activation during the anticipation, preparation and execution of foot movements. *Neuroimage* **21**, 568-575, doi:10.1016/j.neuroimage.2003.09.065.

76 Abellan van Kan, G. *et al.* (2009) Gait speed at usual pace as a predictor of adverse outcomes in community-dwelling older people an International Academy on Nutrition and Aging (IANA) Task Force. *J Nutr Health Aging* **13**, 881-889.

77 Verlinden, V. J., Van der Geest, J. N., Hofman, A. & Ikram, M. A. (2014) Cognition and gait show a distinct pattern of association in the general population. *Alzheimers Dement* **10**, 328-335, doi:10.1016/j.jalz.2013.03.009.

78 Rosano, C. *et al.* (2011) High blood pressure accelerates gait slowing in well-functioning older adults over 18-years of follow-up. *J Am Geriatr Soc* **59**, 390-397, doi:10.1111/j.1532-5415.2010.03282.x.

79 Khawaja, R. A., Qureshi, R., Mansure, A. H. & Yahya, M. E. (2010) Validation of Datascope Accutorr Plus using British Hypertension Society (BHS) and Association for the Advancement of Medical Instrumentation (AAMI) protocol guidelines. *J Saudi Heart Assoc* **22**, 1-5, doi:10.1016/j.jsha.2010.03.001.

80 Gaillard, R. *et al.* (2014) Risk factors and consequences of maternal anaemia and elevated haemoglobin levels during pregnancy: a population-based prospective cohort study. *Paediatr Perinat Epidemiol* **28**, 213-226, doi:10.1111/ppe.12112.

81 Launer, L. J., Terwindt, G. M. & Ferrari, M. D. (1999) The prevalence and characteristics of migraine in a population-based cohort: the GEM study. *Neurology* **53**, 537-542, doi:10.1212/wnl.53.3.537.

82 Radloff, L. S. (1977) The CES-D scale: a self-report depression scale for research in the general population. *Applied Psychological Measurement* **1**, 385-401.

83 Johns, M. W. (1991) A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep* **14**, 540-545, doi:10.1093/sleep/14.6.540.

84 Oldfield, R. C. (1971) The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* **9**, 97-113.

85 Fan, X. & Markram, H. (2019) A Brief History of Simulation Neuroscience. *Front Neuroinform* **13**, 32, doi:10.3389/fninf.2019.00032.

86 Wright, I. C. *et al.* (1995) A voxel-based method for the statistical analysis of gray and white matter density applied to schizophrenia. *Neuroimage* **2**, 244-252, doi:10.1006/nimg.1995.1032.

87 Dale, A. M. & Sereno, M. I. (1993) Improved Localizadon of Cortical Activity by Combining EEG and MEG with MRI Cortical Surface Reconstruction: A Linear Approach. *J Cogn Neurosci* **5**, 162-176, doi:10.1162/jocn.1993.5.2.162.

88 Fischl, B. (2012) FreeSurfer. *Neuroimage* **62**, 774-781, doi:10.1016/j.neuroimage.2012.01.021.

89 Paus, T. (2010) Population neuroscience: why and how. *Hum Brain Mapp* **31**, 891-903, doi:10.1002/hbm.21069.

90 Casey, B. J. *et al.* (2018) The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. *Dev Cogn Neurosci* **32**, 43-54, doi:10.1016/j.dcn.2018.03.001.

91 Alfaro-Almagro, F. *et al.* (2018) Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank. *Neuroimage* **166**, 400-424, doi:10.1016/j.neuroimage.2017.10.034.

92 Worsley, K. J. *et al.* (2009) SurfStat: a Matlab toolbox for the statistical analysis of univariate and multivariate surface and volumetric data using linear mixed effects models and random fied theory. *Neuroimage* **47**, S102.

93 Albaugh, M. D. *et al.* (2019) Amygdalar reactivity is associated with prefrontal cortical thickness in a large population-based sample of adolescents. *PLoS One* **14**, e0216152, doi:10.1371/journal.pone.0216152.

94 Smith, S. M. & Nichols, T. E. (2018) Statistical Challenges in "Big Data" Human Neuroimaging. *Neuron* **97**, 263-268, doi:10.1016/j.neuron.2017.12.018.

95 R Core Team. (2016) *R: A language and environment for statistical computing*. (R Foundation for Statistical Computing).

96 Muschelli, J. *et al.* (2019) Neuroconductor: an R platform for medical imaging analysis. *Biostatistics* **20**, 218-239, doi:10.1093/biostatistics/kxx068.

97 Hagler, D. J., Jr., Saygin, A. P. & Sereno, M. I. (2006) Smoothing and cluster thresholding for cortical surface-based group analysis of fMRI data. *Neuroimage* **33**, 1093-1103, doi:10.1016/j.neuroimage.2006.07.036.

98 Greve, D. N. & Fischl, B. (2018) False positive rates in surface-based anatomical analysis. *Neuroimage* **171**, 6-14, doi:10.1016/j.neuroimage.2017.12.072.

99 Zeileis, A. & Croissant, Y. (2010) Extended Model Formulas in R: Multiple Parts and Multiple Responses. *J Stat Softw* **34**, 1-13.

100 Rubin, D. B. (1976) Inference and Missing Data. *Biometrika* **63**, 581-590, doi:DOI 10.1093/biomet/63.3.581.

101 Van Buuren, S. & Groothuis-Oudshoorn, K. (2011) mice: Multivariate Imputation by Chained Equations in R. *J Stat Softw* **45**, 1-67.

102 Su, Y. S., Gelman, A., Hill, J. & Yajima, M. (2011) Multiple Imputation with Diagnostics (mi) in R: Opening Windows into the Black Box. *J Stat Softw* **45**, 1-31.

103 Honaker, J., King, G. & Blackwell, M. (2011) Amelia II: A Program for Missing Data. *J Stat Softw* **45**, 1-47.

104 Stekhoven, D. J. & Buhlmann, P. (2012) MissForest--non-parametric missing value imputation for mixed-type data. *Bioinformatics* **28**, 112-118, doi:10.1093/bioinformatics/btr597.

105 Rubin, D. B. (1987) *Multiple imputation for nonresponse in surveys*. (John Wiley & Sons).

106 Lamballais, S. & Muetzel, R. L. (2020) QDECR: a flexible, extensible vertex-wise analysis framework in R. *Code Ocean*, doi:10.24433/CO.2177760.v1.

107 Huber, W. *et al.* (2015) Orchestrating high-throughput genomic analysis with Bioconductor. *Nat Methods* **12**, 115-121, doi:10.1038/nmeth.3252.

108 Ad-Dab'bagh, Y. *et al.* (2006) *The CIVET image-processing environment: a fully automated comprehensive pipeline for anatomical neuroimaging research*.

109 Janacsek, K. *et al.* (2020) Sequence learning in the human brain: A functional neuroanatomical meta-analysis of serial reaction time studies. *Neuroimage* **207**, 116387, doi:10.1016/j.neuroimage.2019.116387.

110 Van Erp, T. G. *et al.* (2016) Subcortical brain volume abnormalities in 2028 individuals with schizophrenia and 2540 healthy controls via the ENIGMA consortium. *Mol Psychiatry* **21**, 547-553, doi:10.1038/mp.2015.63.

111 Tian, Q. *et al.* (2017) The brain map of gait variability in aging, cognitive impairment and dementia-A systematic review. *Neurosci Biobehav Rev* **74**, 149-162, doi:10.1016/j.neubiorev.2017.01.020.

112 Van Rooij, D. *et al.* (2018) Cortical and subcortical brain morphometry differences between patients with autism spectrum disorder and healthy individuals across the lifespan: results from the ENIGMA ASD working group. *Am J Psychiatry* **175**, 359-369, doi:10.1176/appi.ajp.2017.17010100.

113 Somerville, L. H. (2016) Searching for signatures of brain maturity: what are we searching for? *Neuron* **92**, 1164-1167, doi:10.1016/j.neuron.2016.10.059.

114 Walhovd, K. B. *et al.* (2011) Consistent neuroanatomical age-related volume differences across multiple samples. *Neurobiol Aging* **32**, 916-932, doi:10.1016/j.neurobiolaging.2009.05.013.

115 Gennatas, E. D. *et al.* (2017) Age-related effects and sex differences in gray matter density, volume, mass, and cortical thickness from childhood to young adulthood. *J Neurosci* **37**, 5065-5073, doi:10.1523/JNEUROSCI.3550-16.2017.

116 White, T. (2019) Brain Development and Stochastic Processes During Prenatal and Early Life: You Can't Lose It if You've Never Had It; But It's Better to Have It and Lose It, Than Never to Have Had It at All. *J Am Acad Child Adolesc Psychiatry* **58**, 1042-1050, doi:10.1016/j.jaac.2019.02.010.

117 Calem, M., Bromis, K., McGuire, P., Morgan, C. & Kempton, M. J. (2017) Meta-analysis of associations between childhood adversity and hippocampus and amygdala volume in non-clinical and general population samples. *Neuroimage Clin* **14**, 471-479, doi:10.1016/j.nicl.2017.02.016.

118 Cassiers, L. L. M. *et al.* (2018) Structural and functional brain abnormalities associated with exposure to different childhood trauma subtypes: A systematic review of neuroimaging findings. *Front Psychiatry* **9**, 329, doi:10.3389/fpsyt.2018.00329.

119 Frodl, T. *et al.* (2017) Childhood adversity impacts on brain subcortical structures relevant to depression. *J Psychiatr Res* **86**, 58-65, doi:10.1016/j.jpsychires.2016.11.010.

120 Batouli, S. A., Trollor, J. N., Wen, W. & Sachdev, P. S. (2014) The heritability of volumes of brain structures and its relationship to age: a review of twin and family studies. *Ageing Res Rev* **13**, 1-9, doi:10.1016/j.arr.2013.10.003.

121 Jansen, A. G., Mous, S. E., White, T., Posthuma, D. & Polderman, T. J. (2015) What twin studies tell us about the heritability of brain development, morphology, and function: a review. *Neuropsychol Rev* **25**, 27-46, doi:10.1007/s11065-015-9278-9.

122 Hibar, D. P. *et al.* (2015) Common genetic variants influence human subcortical brain structures. *Nature* **520**, 224-229, doi:10.1038/nature14101.

123 Jaddoe, V. W. *et al.* (2007) The Generation R Study Biobank: a resource for epidemiological studies in children and their parents. *Eur J Epidemiol* **22**, 917-923, doi:10.1007/s10654-007-9209-z.

124 Medina-Gomez, C. *et al.* (2015) Challenges in conducting genome-wide association studies in highly admixed multi-ethnic populations: the Generation R Study. *Eur J Epidemiol* **30**, 317-330, doi:10.1007/s10654-015-9998-4.

125 Purcell, S. *et al.* (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-575, doi:10.1086/519795.

126 Vilhjalmsson, B. J. *et al.* (2015) Modeling linkage disequilibrium increases accuracy of polygenic risk scores. *Am J Hum Genet* **97**, 576-592, doi:10.1016/j.ajhg.2015.09.001.

127 The 1000 Genomes Project Consortium. (2015) A global reference for human genetic variation. *Nature* **526**, 68-74, doi:10.1038/nature15393.

128 Roza, S. J. *et al.* (2008) Foetal growth determines cerebral ventricular volume in infants The Generation R Study. *Neuroimage* **39**, 1491-1498, doi:10.1016/j.neuroimage.2007.11.004.

129 Herba, C. M. *et al.* (2010) Infant brain development and vulnerability to later internalizing difficulties: the Generation R study. *J Am Acad Child Adolesc Psychiatry* **49**, 1053-1063, doi:10.1016/j.jaac.2010.07.003.

130 Fischl, B. *et al.* (2004) Automatically parcellating the human cerebral cortex. *Cereb Cortex* **14**, 11-22.

131 VanderWeele, T. J. (2016) Mediation Analysis: A Practitioner's Guide. *Annu Rev Publ Health* **37**, 17-32, doi:10.1146/annurev-publhealth-032315-021402.

132 Tingley, D., Yamamoto, T., Hirose, K., Keele, L. & Imai, K. (2014) mediation: R Package for Causal Mediation Analysis. *J Stat Softw* **59**.

133 Raznahan, A. *et al.* (2014) Longitudinal four-dimensional mapping of subcortical anatomy in human development. *Proc Natl Acad Sci U S A* **111**, 1592-1597, doi:10.1073/pnas.1316911111.

134 Kiraly, A. *et al.* (2016) Male brain ages faster: the age and gender dependence of subcortical volumes. *Brain Imaging Behav* **10**, 901-910, doi:10.1007/s11682-015-9468-3.

135 Herting, M. M. *et al.* (2018) Development of subcortical volumes across adolescence in males and females: A multisample study of longitudinal changes. *Neuroimage* **172**, 194-205, doi:10.1016/j.neuroimage.2018.01.020.

136 Wierenga, L. M. *et al.* (2018) A key characteristic of sex differences in the developing brain: greater variability in brain structure of boys than girls. *Cereb Cortex* **28**, 2741-2751, doi:10.1093/cercor/bhx154.

137 Weeland, C. J. *et al.* (2020) Brain Morphology Associated With Obsessive-Compulsive Symptoms In 2,551 Children From the General Population. *J Am Acad Child Adolesc Psychiatry*, doi:10.1016/j.jaac.2020.03.012.

138 Jansen, P. R. *et al.* (2019) Polygenic Scores for Neuropsychiatric Traits and White Matter Microstructure in the Pediatric Population. *Biol Psychiatry Cogn Neurosci Neuroimaging* **4**, 243-250, doi:10.1016/j.bpsc.2018.07.010.

139 Scott, J. A. *et al.* (2011) Growth trajectories of the human fetal brain tissues estimated from 3D reconstructed in utero MRI. *Int J Dev Neurosci* **29**, 529-536, doi:10.1016/j.ijdevneu.2011.04.001.

140 Dima, D. *et al.* (2020) Subcortical Volume Trajectories across the Lifespan: Data from 18,605 healthy individuals aged 3-90 years. *bioRxiv*, 2020.2005.2005.079475, doi:10.1101/2020.05.05.079475.

141 Giedd, J. N. (2004) Structural magnetic resonance imaging of the adolescent brain. *Ann N Y Acad Sci* **1021**, 77-85, doi:10.1196/annals.1308.009.

142 Shaw, P. *et al.* (2008) Neurodevelopmental trajectories of the human cerebral cortex. *J Neurosci* **28**, 3586-3594, doi:10.1523/JNEUROSCI.5309-07.2008.

143 Karolis, V. R. *et al.* (2017) Volumetric grey matter alterations in adolescents and adults born very preterm suggest accelerated brain maturation. *Neuroimage* **163**, 379-389, doi:10.1016/j.neuroimage.2017.09.039.

144 Wang, Y., Xu, Q., Luo, J., Hu, M. & Zuo, C. (2019) Effects of Age and Sex on Subcortical Volumes. *Front Aging Neurosci* **11**, 259, doi:10.3389/fnagi.2019.00259.

145 Iacono, M. I. *et al.* (2015) MIDA: A Multimodal Imaging-Based Detailed Anatomical Model of the Human Head and Neck. *PLoS One* **10**, e0124126, doi:10.1371/journal.pone.0124126.

146 Kunkle, B. W. *et al.* (2019) Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Abeta, tau, immunity and lipid processing. *Nat Genet* **51**, 414-430, doi:10.1038/s41588-019-0358-2.

147 Nalls, M. A. *et al.* (2014) Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. *Nat Genet* **46**, 989-993, doi:10.1038/ng.3043.

148 Ferrari, R. *et al.* (2014) Frontotemporal dementia and its subtypes: a genome-wide association study. *Lancet Neurol* **13**, 686-699, doi:10.1016/S1474-4422(14)70065-1.

149 Jack, C. R., Jr. *et al.* (2013) Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. *Lancet Neurol* **12**, 207-216, doi:10.1016/S1474-4422(12)70291-0.

150 Postuma, R. B. & Berg, D. (2019) Prodromal Parkinson's Disease: The Decade Past, the Decade to Come. *Mov Disord* **34**, 665-675, doi:10.1002/mds.27670.

151 Staffaroni, A. M. *et al.* (2019) Individualized atrophy scores predict dementia onset in familial frontotemporal lobar degeneration. *Alzheimers Dement*, doi:10.1016/j.jalz.2019.04.007.

152 Cherbuin, N., Kim, S. & Anstey, K. J. (2015) Dementia risk estimates associated with measures of depression: a systematic review and meta-analysis. *BMJ Open* **5**, e008853, doi:10.1136/bmjopen-2015-008853.

153 Barnes, D. E. *et al.* (2012) Midlife vs late-life depressive symptoms and risk of dementia: differential effects for Alzheimer disease and vascular dementia. *Arch Gen Psychiatry* **69**, 493-498, doi:10.1001/archgenpsychiatry.2011.1481.

154 Almeida, O. P., Hankey, G. J., Yeap, B. B., Golledge, J. & Flicker, L. (2017) Depression as a modifiable factor to decrease the risk of dementia. *Transl Psychiatry* **7**, e1117, doi:10.1038/tp.2017.90.

155 Li, G. *et al.* (2017) Cognitive Trajectory Changes Over 20 Years Before Dementia Diagnosis: A Large Cohort Study. *J Am Geriatr Soc* **65**, 2627-2633, doi:10.1111/jgs.15077.

156 Elias, M. F. *et al.* (2000) The preclinical phase of alzheimer disease: A 22-year prospective study of the Framingham Cohort. *Arch Neurol* **57**, 808-813, doi:10.1001/archneur.57.6.808.

157 Stocker, H., Mollers, T., Perna, L. & Brenner, H. (2018) The genetic risk of Alzheimer's disease beyond APOE epsilon4: systematic review of Alzheimer's genetic risk scores. *Transl Psychiatry* **8**, 166, doi:10.1038/s41398-018-0221-8.

158 Strittmatter, W. J. *et al.* (1993) Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. *Proc Natl Acad Sci U S A* **90**, 1977-1981.

159 Corneveaux, J. J. *et al.* (2010) Association of CR1, CLU and PICALM with Alzheimer's disease in a cohort of clinically characterized and neuropathologically verified individuals. *Hum Mol Genet* **19**, 3295-3301, doi:10.1093/hmg/ddq221.

160 El-Lebedy, D., Raslan, H. M. & Mohammed, A. M. (2016) Apolipoprotein E gene polymorphism and risk of type 2 diabetes and cardiovascular disease. *Cardiovasc Diabetol* **15**, 12, doi:10.1186/s12933-016-0329-1.

161 Kritharides, L., Nordestgaard, B. G., Tybjaerg-Hansen, A., Kamstrup, P. R. & Afzal, S. (2017) Effect of APOE epsilon Genotype on Lipoprotein(a) and the Associated Risk of Myocardial Infarction and Aortic Valve Stenosis. *J Clin Endocrinol Metab* **102**, 3390-3399, doi:10.1210/jc.2017-01049.

162 Alvim, R. O. *et al.* (2010) APOE polymorphism is associated with lipid profile, but not with arterial stiffness in the general population. *Lipids Health Dis* **9**, 128, doi:10.1186/1476-511X-9-128.

163 Lin, S. K. *et al.* (2004) Association of apolipoprotein E genotypes with serum lipid profiles in a healthy population of Taiwan. *Ann Clin Lab Sci* **34**, 443-448.

164 Shatwan, I. M. *et al.* (2018) Association of apolipoprotein E gene polymorphisms with blood lipids and their interaction with dietary factors. *Lipids Health Dis* **17**, 98, doi:10.1186/s12944-018-0744-2.

165 Taylor, A. E. *et al.* (2011) IQ, educational attainment, memory and plasma lipids: associations with apolipoprotein E genotype in 5995 children. *Biol Psychiatry* **70**, 152-158, doi:10.1016/j.biopsych.2010.10.033.

166 Kallio, M. J. *et al.* (1997) Apoprotein E phenotype determines serum cholesterol in infants during both high-cholesterol breast feeding and low-cholesterol formula feeding. *J Lipid Res* **38**, 759-764.

167 O'Donoghue, M. C., Murphy, S. E., Zamboni, G., Nobre, A. C. & Mackay, C. E. (2018) APOE genotype and cognition in healthy individuals at risk of Alzheimer's disease: A review. *Cortex* **104**, 103-123, doi:10.1016/j.cortex.2018.03.025.

168 Shaw, P. *et al.* (2007) Cortical morphology in children and adolescents with different apolipoprotein E gene polymorphisms: an observational study. *Lancet Neurol* **6**, 494-500, doi:10.1016/S1474-4422(07)70106-0.

169 O'Dwyer, L. *et al.* (2012) Reduced hippocampal volume in healthy young ApoE4 carriers: an MRI study. *PLoS One* **7**, e48895, doi:10.1371/journal.pone.0048895.

170 Filippini, N. *et al.* (2009) Distinct patterns of brain activity in young carriers of the APOE-epsilon4 allele. *Proc Natl Acad Sci U S A* **106**, 7209-7214, doi:10.1073/pnas.0811879106.

171 Chang, L. *et al.* (2016) Gray matter maturation and cognition in children with different APOE epsilon genotypes. *Neurology* **87**, 585-594, doi:10.1212/WNL.0000000000002939.

172 Knickmeyer, R. C. *et al.* (2014) Common variants in psychiatric risk genes predict brain structure at birth. *Cereb Cortex* **24**, 1230-1246, doi:10.1093/cercor/bhs401.

173 Dean, D. C., 3rd *et al.* (2014) Brain differences in infants at differential genetic risk for late-onset Alzheimer disease: a cross-sectional imaging study. *JAMA Neurol* **71**, 11-22, doi:10.1001/jamaneurol.2013.4544.

174 Khera, A. V. *et al.* (2018) Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet* **50**, 1219-1224, doi:10.1038/s41588-018-0183-z.

175 Euesden, J., Lewis, C. M. & O'Reilly, P. F. (2015) PRSice: Polygenic Risk Score software. *Bioinformatics* **31**, 1466-1468, doi:10.1093/bioinformatics/btu848.

176 Lambert, J. C. *et al.* (2013) Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet* **45**, 1452-1458, doi:10.1038/ng.2802.

177 Tellegen, P. J., Winkel, M., Wijnberg-Williams, B. & Laros, J. A. (2005) *SON-R 2,5-7: Snijder-Oomen niet-verbale intelligentietest*. (Hogrefe Uitgevers).

178 Langeslag, S. J. *et al.* (2013) Functional connectivity between parietal and frontal brain regions and intelligence in young children: the Generation R study. *Hum Brain Mapp* **34**, 3299-3307, doi:10.1002/hbm.22143.

179 Achenbach, T. M. & Rescorla, L. A. (2000) *Manual for the ASEBA preschool forms and profiles*. Vol. 30 (Burlington, VT: University of Vermont, Research center for children, youth …).

180 Muetzel, R. L. *et al.* (2018) Tracking Brain Development and Dimensional Psychiatric Symptoms in Children: A Longitudinal Population-Based Neuroimaging Study. *Am J Psychiatry* **175**, 54-62, doi:10.1176/appi.ajp.2017.16070813.

181 Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W. & Smith, S. M. (2012) Fsl. *Neuroimage* **62**, 782-790, doi:10.1016/j.neuroimage.2011.09.015.

182 Chang, L. C., Jones, D. K. & Pierpaoli, C. (2005) RESTORE: robust estimation of tensors by outlier rejection. *Magn Reson Med* **53**, 1088-1095, doi:10.1002/mrm.20426.

183 De Groot, M. *et al.* (2013) Improving alignment in Tract-based spatial statistics: evaluation and optimization of image registration. *Neuroimage* **76**, 400-411, doi:10.1016/j.neuroimage.2013.03.015.

184 Muetzel, R. L. *et al.* (2015) White matter integrity and cognitive performance in school-age children: A population-based neuroimaging study. *Neuroimage* **119**, 119-128, doi:10.1016/j.neuroimage.2015.06.014.

185 Friedewald, W. T., Levy, R. I. & Fredrickson, D. S. (1972) Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clin Chem* **18**, 499-502.

186 Huang, Y. & Mahley, R. W. (2014) Apolipoprotein E: structure and function in lipid metabolism, neurobiology, and Alzheimer's diseases. *Neurobiol Dis* **72 Pt A**, 3-12, doi:10.1016/j.nbd.2014.08.025.

187 Hellwege, J. N. *et al.* (2017) Population Stratification in Genetic Association Studies. *Curr Protoc Hum Genet* **95**, 1 22 21-21 22 23, doi:10.1002/cphg.48.

188 Axelrud, L. K. *et al.* (2018) Polygenic Risk Score for Alzheimer's Disease: Implications for Memory Performance and Hippocampal Volumes in Early Life. *Am J Psychiatry* **175**, 555-563, doi:10.1176/appi.ajp.2017.17050529.

189 Quiroz, Y. T. *et al.* (2015) Brain Imaging and Blood Biomarker Abnormalities in Children With Autosomal Dominant Alzheimer Disease: A Cross-Sectional Study. *JAMA Neurol* **72**, 912-919, doi:10.1001/jamaneurol.2015.1099.

190 Foley, S. F. *et al.* (2017) Multimodal Brain Imaging Reveals Structural Differences in Alzheimer's Disease Polygenic Risk Carriers: A Study in Healthy Young Adults. *Biol Psychiatry* **81**, 154-161, doi:10.1016/j.biopsych.2016.02.033.

191 DiBattista, A. M., Stevens, B. W., Rebeck, G. W. & Green, A. E. (2014) Two Alzheimer's disease risk genes increase entorhinal cortex volume in young adults. *Front Hum Neurosci* **8**, 779, doi:10.3389/fnhum.2014.00779.

192 Alexopoulos, P. *et al.* (2011) Hippocampal volume differences between healthy young apolipoprotein E epsilon2 and epsilon4 carriers. *J Alzheimers Dis* **26**, 207-210, doi:10.3233/JAD-2011-110356.

193 Nao, J. *et al.* (2017) Adverse Effects of the Apolipoprotein E epsilon4 Allele on Episodic Memory, Task Switching and Gray Matter Volume in Healthy Young Adults. *Front Hum Neurosci* **11**, 346, doi:10.3389/fnhum.2017.00346.

194 Konishi, K. *et al.* (2016) APOE2 Is Associated with Spatial Navigational Strategies and Increased Gray Matter in the Hippocampus. *Front Hum Neurosci* **10**, 349, doi:10.3389/fnhum.2016.00349.

195 Alexander, G. E. *et al.* (2012) Gray matter network associated with risk for Alzheimer's disease in young to middle-aged adults. *Neurobiol Aging* **33**, 2723-2732, doi:10.1016/j.neurobiolaging.2012.01.014.

196 Mormino, E. C. *et al.* (2016) Polygenic risk of Alzheimer disease is associated with early- and late-life processes. *Neurology* **87**, 481-488, doi:10.1212/WNL.0000000000002922.

197 Heise, V., Filippini, N., Ebmeier, K. P. & Mackay, C. E. (2011) The APOE varepsilon4 allele modulates brain white matter integrity in healthy adults. *Mol Psychiatry* **16**, 908-916, doi:10.1038/mp.2010.90.

198 Hanson, A. J. *et al.* (2013) Effect of apolipoprotein E genotype and diet on apolipoprotein E lipidation and amyloid peptides: randomized clinical trial. *JAMA Neurol* **70**, 972-980, doi:10.1001/jamaneurol.2013.396.

199 Heinsinger, N. M., Gachechiladze, M. A. & Rebeck, G. W. (2016) Apolipoprotein E Genotype Affects Size of ApoE Complexes in Cerebrospinal Fluid. *J Neuropathol Exp Neurol* **75**, 918-924, doi:10.1093/jnen/nlw067.

200 Ignatius, M. J. *et al.* (1986) Expression of apolipoprotein E during nerve degeneration and regeneration. *Proc Natl Acad Sci U S A* **83**, 1125-1129.

201 Arendt, T. *et al.* (1997) Plastic neuronal remodeling is impaired in patients with Alzheimer's disease carrying apolipoprotein epsilon 4 allele. *J Neurosci* **17**, 516-529.

202 Tzioras, M., Davies, C., Newman, A., Jackson, R. & Spires-Jones, T. (2018) Invited Review: APOE at the interface of inflammation, neurodegeneration and pathological protein spread in Alzheimer's disease. *Neuropathol Appl Neurobiol*, doi:10.1111/nan.12529.

203 Weissberger, G. H., Nation, D. A., Nguyen, C. P., Bondi, M. W. & Han, S. D. (2018) Meta-analysis of cognitive ability differences by apolipoprotein e genotype in young humans. *Neurosci Biobehav Rev* **94**, 49-58, doi:10.1016/j.neubiorev.2018.08.009.

204 Thaler, A. *et al.* (2012) Lower cognitive performance in healthy G2019S LRRK2 mutation carriers. *Neurology* **79**, 1027-1032, doi:10.1212/WNL.0b013e3182684646.

205 Mirelman, A. *et al.* (2011) Gait alterations in healthy carriers of the LRRK2 G2019S mutation. *Ann Neurol* **69**, 193-197, doi:10.1002/ana.22165.

206 Saunders-Pullman, R. *et al.* (2011) Olfactory dysfunction in LRRK2 G2019S mutation carriers. *Neurology* **77**, 319-324, doi:10.1212/WNL.0b013e318227041c.

207 Tapiola, T. *et al.* (1998) CSF tau is related to apolipoprotein E genotype in early Alzheimer's disease. *Neurology* **50**, 169-174.

208 Van Harten, A. C. *et al.* (2017) CSF ApoE predicts clinical progression in nondemented APOEepsilon4 carriers. *Neurobiol Aging* **57**, 186-194, doi:10.1016/j.neurobiolaging.2017.04.002.

209 Koch, G. *et al.* (2017) CSF tau is associated with impaired cortical plasticity, cognitive decline and astrocyte survival only in APOE4-positive Alzheimer's disease. *Sci Rep* **7**, 13728, doi:10.1038/s41598-017-14204-3.

210 Hohman, T. J. *et al.* (2018) Sex-Specific Association of Apolipoprotein E With Cerebrospinal Fluid Levels of Tau. *JAMA Neurol* **75**, 989-998, doi:10.1001/jamaneurol.2018.0821.

211 Cruchaga, C. *et al.* (2012) Cerebrospinal fluid APOE levels: an endophenotype for genetic studies for Alzheimer's disease. *Hum Mol Genet* **21**, 4558-4571, doi:10.1093/hmg/dds296.

212 Darweesh, S. K. L. *et al.* (2018) Inflammatory markers and the risk of dementia and Alzheimer's disease: A meta-analysis. *Alzheimers Dement* **14**, 1450-1459, doi:10.1016/j.jalz.2018.02.014.

213 Ferri, C. P. *et al.* (2005) Global prevalence of dementia: a Delphi consensus study. *Lancet* **366**, 2112-2117, doi:10.1016/S0140-6736(05)67889-0.

214 Skoog, I. *et al.* (1996) 15-year longitudinal study of blood pressure and dementia. *Lancet* **347**, 1141-1145.

215 Whitmer, R. A., Sidney, S., Selby, J., Johnston, S. C. & Yaffe, K. (2005) Midlife cardiovascular risk factors and risk of dementia in late life. *Neurology* **64**, 277-281, doi:10.1212/01.WNL.0000149519.47454.F2.

216 Joas, E. *et al.* (2012) Blood pressure trajectories from midlife to late life in relation to dementia in women followed for 37 years. *Hypertension* **59**, 796-801, doi:10.1161/HYPERTENSIONAHA.111.182204.

217 Novak, V. & Hajjar, I. (2010) The relationship between blood pressure and cognitive function. *Nat Rev Cardiol* **7**, 686-698, doi:10.1038/nrcardio.2010.161.

218 Pase, M. P., Herbert, A., Grima, N. A., Pipingas, A. & O'Rourke, M. F. (2012) Arterial stiffness as a cause of cognitive decline and dementia: a systematic review and meta-analysis. *Intern Med J* **42**, 808-815, doi:10.1111/j.1445-5994.2011.02645.x.

219 Allen, N. B. *et al.* (2014) Blood pressure trajectories in early adulthood and subclinical atherosclerosis in middle age. *JAMA* **311**, 490-497, doi:10.1001/jama.2013.285122.

220 Shear, C. L., Burke, G. L., Freedman, D. S. & Berenson, G. S. (1986) Value of childhood blood pressure measurements and family history in predicting future blood pressure status: results from 8 years of follow-up in the Bogalusa Heart Study. *Pediatrics* **77**, 862-869.

221 Klumbiene, J., Sileikiene, L., Milasauskiene, Z., Zaborskis, A. & Shatchkute, A. (2000) The relationship of childhood to adult blood pressure: longitudinal study of juvenile hypertension in Lithuania. *J Hypertens* **18**, 531-538.

222 Ferreira, I., van de Laar, R. J., Prins, M. H., Twisk, J. W. & Stehouwer, C. D. (2012) Carotid stiffness in young adults: a life-course analysis of its early determinants: the Amsterdam Growth and Health Longitudinal Study. *Hypertension* **59**, 54-61, doi:10.1161/HYPERTENSIONAHA.110.156109.

223 Parmar, P. G. *et al.* (2016) International Genome-Wide Association Study Consortium Identifies Novel Loci Associated With Blood Pressure in Children and Adolescents. *Circ Cardiovasc Genet* **9**, 266-278, doi:10.1161/CIRCGENETICS.115.001190.

224 Justice, A. E. *et al.* (2016) Genome-wide association of trajectories of systolic blood pressure change. *BMC Proc* **10**, 321-327, doi:10.1186/s12919-016-0050-9.

225 Ditto, B., Seguin, J. R. & Tremblay, R. E. (2006) Neuropsychological characteristics of adolescent boys differing in risk for high blood pressure. *Ann Behav Med* **31**, 231-237, doi:10.1207/s15324796abm3103_4.

226 Lande, M. B., Kaczorowski, J. M., Auinger, P., Schwartz, G. J. & Weitzman, M. (2003) Elevated blood pressure and decreased cognitive function among school-age children and adolescents in the United States. *J Pediatr* **143**, 720-724, doi:10.1067/S0022-3476(03)00412-8.

227 Adams, H. R., Szilagyi, P. G., Gebhardt, L. & Lande, M. B. (2010) Learning and attention problems among children with pediatric primary hypertension. *Pediatrics* **126**, e1425-1429, doi:10.1542/peds.2010-1899.

228 Lande, M. B. *et al.* (2017) Neurocognitive Function in Children with Primary Hypertension. *J Pediatr* **180**, 148-155 e141, doi:10.1016/j.jpeds.2016.08.076.

229 Laurent, S. *et al.* (2006) Expert consensus document on arterial stiffness: methodological issues and clinical applications. *Eur Heart J* **27**, 2588-2605, doi:10.1093/eurheartj/ehl254.

230 Hofman, A. *et al.* (2015) The Rotterdam Study: 2016 objectives and design update. *Eur J Epidemiol* **30**, 661-708, doi:10.1007/s10654-015-0082-x.

231 Wong, S. N., Tz Sung, R. Y. & Leung, L. C. (2006) Validation of three oscillometric blood pressure devices against auscultatory mercury sphygmomanometer in children. *Blood Press Monit* **11**, 281-291, doi:10.1097/01.mbp.0000209082.09623.b4.

232 Vlachopoulos, C., Aznaouridis, K. & Stefanadis, C. (2010) Prediction of cardiovascular events and all-cause mortality with arterial stiffness: a systematic review and meta-analysis. *J Am Coll Cardiol* **55**, 1318-1327, doi:10.1016/j.jacc.2009.10.061.

233 Asmar, R. *et al.* (1995) Assessment of arterial distensibility by automatic pulse wave velocity measurement. Validation and clinical application studies. *Hypertension* **26**, 485-490.

234 Donald, A. E. *et al.* (2010) Determinants of vascular phenotype in a large childhood population: the Avon Longitudinal Study of Parents and Children (ALSPAC). *Eur Heart J* **31**, 1502-1510, doi:10.1093/eurheartj/ehq062.

235 Johnson, W., te Nijenhuis, J. & Bouchard, T. J. (2008) Still just 1 g: Consistent results from five test batteries. *Intelligence* **36**, 81-95, doi:10.1016/j.intell.2007.06.001.

236 Plomin, R. (2001) The genetics of g in human and mouse. *Nat Rev Neurosci* **2**, 136-141, doi:10.1038/35053584.

237 Van der Velde, L. A. *et al.* (2018) Diet quality in childhood: the Generation R Study. *Eur J Nutr*, doi:10.1007/s00394-018-1651-z.

238 Wijtzes, A. I. *et al.* (2014) Sedentary behaviors, physical activity behaviors, and body fat in 6-year-old children: the generation R study. *Int J Behav Nutr Phys Act* **11**, 96, doi:10.1186/PREACCEPT-1946502959127020.

239 Voortman, T. *et al.* (2017) Adherence to the 2015 Dutch dietary guidelines and risk of non-communicable diseases and mortality in the Rotterdam Study. *Eur J Epidemiol* **32**, 993-1005, doi:10.1007/s10654-017-0295-2.

240 Vliegenthart, R. *et al.* (2002) Alcohol consumption and risk of peripheral arterial disease: the Rotterdam study. *Am J Epidemiol* **155**, 332-338.

241 Caspersen, C. J., Bloemberg, B. P., Saris, W. H., Merritt, R. K. & Kromhout, D. (1991) The prevalence of selected physical activities and their relation with coronary heart disease risk factors in elderly men: the Zutphen Study, 1985. *Am J Epidemiol* **133**, 1078-1092.

242 Koolhaas, C. M. *et al.* (2016) Physical Activity Types and Coronary Heart Disease Risk in Middle-Aged and Elderly Persons: The Rotterdam Study. *Am J Epidemiol* **183**, 729-738, doi:10.1093/aje/kwv244.

243 Stel, V. S. *et al.* (2004) Comparison of the LASA Physical Activity Questionnaire with a 7-day diary and pedometer. *J Clin Epidemiol* **57**, 252-258, doi:10.1016/j.jclinepi.2003.07.008.

244 Davis, E. F. *et al.* (2012) Cardiovascular risk factors in children and young adults born to preeclamptic pregnancies: a systematic review. *Pediatrics* **129**, e1552-1561, doi:10.1542/peds.2011-3093.

245 Reitz, C. & Luchsinger, J. A. (2007) Relation of Blood Pressure to Cognitive Impairment and Dementia. *Curr Hypertens Rev* **3**, 166-176, doi:10.2174/157340207781386747.

246 Farmer, M. E. *et al.* (1990) Longitudinally measured blood pressure, antihypertensive medication use, and cognitive performance: the Framingham Study. *J Clin Epidemiol* **43**, 475-480.

247 Team, R. D. C. (2016) *R: A language and environment for statistical computing*. (R Foundation for Statistical Computing).

248 Venables, W. N. & Ripley, B. D. (2002) *Modern Applied Statistics with S*. Fourth Edition edn, (Springer).

249 Zeileis, A. (2004) Econometric computing with HC and HAC covariance matrix estimators. *J Stat Softw* **11**, 1-17.

250 Zeileis, A. & Hothorn, T. (2002) Diagnostic Checking in Regression Relationships. *R News* **2**, 7-10.

251 Lande, M. B., Kupferman, J. C. & Adams, H. R. (2012) Neurocognitive alterations in hypertensive children and adolescents. *J Clin Hypertens (Greenwich)* **14**, 353-359, doi:10.1111/j.1751-7176.2012.00661.x.

252 Roussotte, F. F. *et al.* (2018) In Vivo Brain Plaque and Tangle Burden Mediates the Association Between Diastolic Blood Pressure and Cognitive Functioning in Nondemented Adults. *Am J Geriatr Psychiatry* **26**, 13-22, doi:10.1016/j.jagp.2017.09.001.

253 Cooper, L. L. *et al.* (2016) Cerebrovascular Damage Mediates Relations Between Aortic Stiffness and Memory. *Hypertension* **67**, 176-182, doi:10.1161/HYPERTENSIONAHA.115.06398.

254 Elias, M. F. *et al.* (2009) Arterial pulse wave velocity and cognition with advancing age. *Hypertension* **53**, 668-673, doi:10.1161/HYPERTENSIONAHA.108.126342.

255 Settakis, G. *et al.* (2003) Cerebrovascular reactivity in hypertensive and healthy adolescents: TCD with vasodilatory challenge. *J Neuroimaging* **13**, 106-112.

256 Wong, L. J. *et al.* (2011) Hypertension impairs vascular reactivity in the pediatric brain. *Stroke* **42**, 1834-1838, doi:10.1161/STROKEAHA.110.607606.

257 Haight, T. J. *et al.* (2015) Vascular risk factors, cerebrovascular reactivity, and the default-mode brain network. *Neuroimage* **115**, 7-16, doi:10.1016/j.neuroimage.2015.04.039.

258 Mak, L. E. *et al.* (2017) The Default Mode Network in Healthy Individuals: A Systematic Review and Meta-Analysis. *Brain Connect* **7**, 25-33, doi:10.1089/brain.2016.0438.

259 Parati, G., Di Rienzo, M., Coruzzi, P. & Castiglioni, P. (2009) Chronic hypotension and modulation of autonomic cardiovascular regulation. *Hypertens Res* **32**, 931-933, doi:10.1038/hr.2009.150.

260 Van den Berg, M. E. *et al.* (2018) Normal Values of Corrected Heart-Rate Variability in 10-Second Electrocardiograms for All Ages. *Front Physiol* **9**, 424, doi:10.3389/fphys.2018.00424.

261 Franklin, S. S. (2005) Arterial stiffness and hypertension: a two-way street? *Hypertension* **45**, 349-351, doi:10.1161/01.HYP.0000157819.31611.87.

262 Mitchell, G. F. (2014) Arterial stiffness and hypertension: chicken or egg? *Hypertension* **64**, 210-214, doi:10.1161/HYPERTENSIONAHA.114.03449.

263 Liao, D. *et al.* (1999) Arterial stiffness and the development of hypertension. The ARIC study. *Hypertension* **34**, 201-206.

264 Dernellis, J. & Panaretou, M. (2005) Aortic stiffness is an independent predictor of progression to hypertension in nonhypertensive subjects. *Hypertension* **45**, 426-431, doi:10.1161/01.HYP.0000157818.58878.93.

265 Kaess, B. M. *et al.* (2012) Aortic stiffness, blood pressure progression, and incident hypertension. *JAMA* **308**, 875-881, doi:10.1001/2012.jama.10503.

266 Buckley, J., Cohen, J. D., Kramer, A. F., McAuley, E. & Mullen, S. P. (2014) Cognitive control in the self-regulation of physical activity and sedentary behavior. *Front Hum Neurosci* **8**, 747, doi:10.3389/fnhum.2014.00747.

267 Allan, J. L., McMinn, D. & Daly, M. (2016) A Bidirectional Relationship between Executive Function and Health Behavior: Evidence, Implications, and Future Directions. *Front Neurosci* **10**, 386, doi:10.3389/fnins.2016.00386.

268 Anstey, K. J., Ashby-Mitchell, K. & Peters, R. (2017) Updating the Evidence on the Association between Serum Cholesterol and Risk of Late-Life Dementia: Review and Meta-Analysis. *J Alzheimers Dis* **56**, 215-228, doi:10.3233/JAD-160826.

269 Zhong, G., Wang, Y., Zhang, Y., Guo, J. J. & Zhao, Y. (2015) Smoking is associated with an increased risk of dementia: a meta-analysis of prospective cohort studies with investigation of potential effect modifiers. *PLoS One* **10**, e0118333, doi:10.1371/journal.pone.0118333.

270 Debette, S. *et al.* (2011) Midlife vascular risk factor exposure accelerates structural brain aging and cognitive decline. *Neurology* **77**, 461-468, doi:10.1212/WNL.0b013e318227b227.

271 Pase, M. P. *et al.* (2018) Vascular risk at younger ages most strongly associates with current and future brain volume. *Neurology* **91**, e1479-e1486, doi:10.1212/WNL.0000000000006360.

272 Sabia, S., Marmot, M., Dufouil, C. & Singh-Manoux, A. (2008) Smoking history and cognitive function in middle age from the Whitehall II study. *Arch Intern Med* **168**, 1165-1173, doi:10.1001/archinte.168.11.1165.

273 Dahl, A. K. *et al.* (2013) Body mass index across midlife and cognitive change in late life. *Int J Obes (Lond)* **37**, 296-302, doi:10.1038/ijo.2012.37.

274 Bos, D. *et al.* (2018) Cerebral small vessel disease and the risk of dementia: A systematic review and meta-analysis of population-based evidence. *Alzheimers Dement* **14**, 1482-1492, doi:10.1016/j.jalz.2018.04.007.

275 Lane, C. A. *et al.* (2019) Associations Between Vascular Risk Across Adulthood and Brain Pathology in Late Life: Evidence From a British Birth Cohort. *JAMA Neurol*, 1-9, doi:10.1001/jamaneurol.2019.3774.

276 Reis, J. P. *et al.* (2013) Cardiovascular health through young adulthood and cognitive functioning in midlife. *Ann Neurol* **73**, 170-179, doi:10.1002/ana.23836.

277 Yaffe, K. *et al.* (2014) Early adult to midlife cardiovascular risk factors and cognitive function. *Circulation* **129**, 1560-1567, doi:10.1161/CIRCULATIONAHA.113.004798.

278 Lamballais, S. *et al.* (2020) Design and overview of the Origins of Alzheimer's Disease Across the Life course (ORACLE) study. *Eur J Epidemiol*, doi:10.1007/s10654-020-00696-3.

279 Deary, I. J. (2014) The Stability of Intelligence From Childhood to Old Age. *Curr Dir Psychol Sci* **23**, 239-245, doi:10.1177/0963721414536905.

280 Deary, I. J. (2012) Intelligence. *Annu Rev Psychol* **63**, 453-482, doi:10.1146/annurev-psych-120710-100353.

281 Benjamini, Y. & Hochberg, Y. (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B (Methodological)* **57**, 289-300.

282 Johnson, P. O. & Fay, L. C. (1950) The Johnson-Neyman technique, its theory and application. *Psychometrika* **15**, 349-367, doi:10.1007/BF02288864.

283 Bodner, T. E. (2008) What improves with increased missing data imputations? *Structural Equation Modeling***.15**, pp, doi:10.1080/10705510802339072.

284 Van Dijk, E. J. *et al.* (2004) The association between blood pressure, hypertension, and cerebral white matter lesions: cardiovascular determinants of dementia study. *Hypertension* **44**, 625-630, doi:10.1161/01.HYP.0000145857.98904.20.

285 Verhaaren, B. F. *et al.* (2013) High blood pressure and cerebral white matter lesion progression in the general population. *Hypertension* **61**, 1354-1359, doi:10.1161/HYPERTENSIONAHA.111.00430.

286 Scharf, E. L. *et al.* (2019) Cardiometabolic Health and Longitudinal Progression of White Matter Hyperintensity: The Mayo Clinic Study of Aging. *Stroke* **50**, 3037-3044, doi:10.1161/STROKEAHA.119.025822.

287 Hajjar, I. *et al.* (2011) Hypertension, white matter hyperintensities, and concurrent impairments in mobility, cognition, and mood: the Cardiovascular Health Study. *Circulation* **123**, 858-865, doi:10.1161/CIRCULATIONAHA.110.978114.

288 Allan, C. L. *et al.* (2015) Lifetime hypertension as a predictor of brain structure in older adults: cohort study with a 28-year follow-up. *Br J Psychiatry* **206**, 308-315, doi:10.1192/bjp.bp.114.153536.

289 Ma, Y. *et al.* (2020) Blood Pressure Variability and Cerebral Small Vessel Disease: A Systematic Review and Meta-Analysis of Population-Based Cohorts. *Stroke* **51**, 82-89, doi:10.1161/STROKEAHA.119.026739.

290 Alateeq, K., Walsh, E. I. & Cherbuin, N. (2021) Higher Blood Pressure is Associated with Greater White Matter Lesions and Brain Atrophy: A Systematic Review with Meta-Analysis. *J Clin Med* **10**, doi:10.3390/jcm10040637.

291 Laurent, S. & Boutouyrie, P. (2015) The structural factor of hypertension: large and small artery alterations. *Circ Res* **116**, 1007-1021, doi:10.1161/CIRCRESAHA.116.303596.

292 Loos, C. M., van Oostenbrugge, R. J. & Staals, J. (2012) The appearance of a new white matter lesion adjacent to the old infarct in first-ever lacunar stroke patients: a two-year follow-up study with MRI. *Cerebrovasc Dis* **34**, 443-445, doi:10.1159/000344003.

293 Duering, M. *et al.* (2013) Incident lacunes preferentially localize to the edge of white matter hyperintensities: insights into the pathophysiology of cerebral small vessel disease. *Brain* **136**, 2717-2726, doi:10.1093/brain/awt184.

294 Maillard, P. *et al.* (2011) White matter hyperintensity penumbra. *Stroke* **42**, 1917-1922, doi:10.1161/STROKEAHA.110.609768.

295 Lamballais, S. *et al.* (2018) Association of Blood Pressure and Arterial Stiffness With Cognition in 2 Population-Based Child and Adult Cohorts. *J Am Heart Assoc* **7**, e009847, doi:10.1161/JAHA.118.009847.

296 Nyquist, P. A. *et al.* (2015) Age differences in periventricular and deep white matter lesions. *Neurobiol Aging* **36**, 1653-1658, doi:10.1016/j.neurobiolaging.2015.01.005.

297 Dufouil, C. *et al.* (2001) Longitudinal study of blood pressure and white matter hyperintensities: the EVA MRI Cohort. *Neurology* **56**, 921-926, doi:10.1212/wnl.56.7.921.

298 Sprint Mind Investigators for the SPRINT Research Group *et al.* (2019) Association of Intensive vs Standard Blood Pressure Control With Cerebral White Matter Lesions. *JAMA* **322**, 524-534, doi:10.1001/jama.2019.10551.

299 Dufouil, C. *et al.* (2005) Effects of blood pressure lowering on cerebral white matter hyperintensities in patients with stroke: the PROGRESS (Perindopril Protection Against Recurrent Stroke Study) Magnetic Resonance Imaging Substudy. *Circulation* **112**, 1644-1650, doi:10.1161/CIRCULATIONAHA.104.501163.

300 Sachdev, P. S., Wen, W., Christensen, H. & Jorm, A. F. (2005) White matter hyperintensities are related to physical disability and poor motor function. *J Neurol Neurosurg Psychiatry* **76**, 362-367, doi:10.1136/jnnp.2004.042945.

301 Riaz, M. *et al.* (2021) What does hand motor function tell us about our aging brain in association with WMH? *Aging Clin Exp Res* **33**, 1577-1584, doi:10.1007/s40520-020-01683-0.

302 Koppelmans, V., Hirsiger, S., Merillat, S., Jancke, L. & Seidler, R. D. (2015) Cerebellar gray and white matter volume and their relation with age and manual motor performance in healthy older adults. *Hum Brain Mapp* **36**, 2352-2363, doi:10.1002/hbm.22775.

303 Cornelissen, V. A. & Smart, N. A. (2013) Exercise training for blood pressure: a systematic review and meta-analysis. *J Am Heart Assoc* **2**, e004473, doi:10.1161/JAHA.112.004473.

304 Inder, J. D. *et al.* (2016) Isometric exercise training for blood pressure management: a systematic review and meta-analysis to optimize benefit. *Hypertens Res* **39**, 88-94, doi:10.1038/hr.2015.111.

305 Labott, B. K., Bucht, H., Morat, M., Morat, T. & Donath, L. (2019) Effects of Exercise Training on Handgrip Strength in Older Adults: A Meta-Analytical Review. *Gerontology* **65**, 686-698, doi:10.1159/000501203.

306 Zaccagni, L. *et al.* (2020) Handgrip Strength in Young Adults: Association with Anthropometric Variables and Laterality. *Int J Environ Res Public Health* **17**, doi:10.3390/ijerph17124273.

307 Hughes, K. *et al.* (2017) The effect of multiple adverse childhood experiences on health: a systematic review and meta-analysis. *Lancet Public Health* **2**, e356-e366, doi:10.1016/S2468-2667(17)30118-4.

308 Bellis, M. A. *et al.* (2019) Life course health consequences and associated annual costs of adverse childhood experiences across Europe and North America: a systematic review and meta-analysis. *Lancet Public Health* **4**, e517-e528, doi:10.1016/S2468-2667(19)30145-8.

309 Carlson, J. S. *et al.* (2019) Prevalence of adverse childhood experiences in school-aged youth: A systematic review (1990-2015). *International Journal of School & Educational Psychology. 2019, pp. No Pagination Specified. Jan Int J Sch Educ Psychol*, doi:10.1080/21683603.2018.1548397.

310 Crouch, E., Probst, J. C., Radcliff, E., Bennett, K. J. & McKinney, S. H. (2019) Prevalence of adverse childhood experiences (ACEs) among US children. *Child Abuse Negl* **92**, 209-218, doi:10.1016/j.chiabu.2019.04.010.

311 Merrick, M. T., Ford, D. C., Ports, K. A. & Guinn, A. S. (2018) Prevalence of Adverse Childhood Experiences From the 2011-2014 Behavioral Risk Factor Surveillance System in 23 States. *JAMA Pediatr* **172**, 1038-1044, doi:10.1001/jamapediatrics.2018.2537.

312 Sheridan, M. A. & McLaughlin, K. A. (2014) Dimensions of early experience and neural development: deprivation and threat. *Trends Cogn Sci* **18**, 580-585, doi:10.1016/j.tics.2014.09.001.

313 Miller, A. B. *et al.* (2018) Dimensions of deprivation and threat, psychopathology, and potential mediators: A multi-year longitudinal analysis. *J Abnorm Psychol* **127**, 160-170, doi:10.1037/abn0000331.

314 Machlin, L., Miller, A. B., Snyder, J., McLaughlin, K. A. & Sheridan, M. A. (2019) Differential Associations of Deprivation and Threat With Cognitive Control and Fear Conditioning in Early Childhood. *Front Behav Neurosci* **13**, 80, doi:10.3389/fnbeh.2019.00080.

315 McLaughlin, K. A., Sheridan, M. A. & Lambert, H. K. (2014) Childhood adversity and neural development: deprivation and threat as distinct dimensions of early experience. *Neurosci Biobehav Rev* **47**, 578-591, doi:10.1016/j.neubiorev.2014.10.012.

316 Sumner, J. A., Colich, N. L., Uddin, M., Armstrong, D. & McLaughlin, K. A. (2019) Early Experiences of Threat, but Not Deprivation, Are Associated With Accelerated Biological Aging in Children and Adolescents. *Biol Psychiatry* **85**, 268-278, doi:10.1016/j.biopsych.2018.09.008.

317 Jylhava, J., Pedersen, N. L. & Hagg, S. (2017) Biological Age Predictors. *EBioMedicine* **21**, 29-36, doi:10.1016/j.ebiom.2017.03.046.

318 Colich, N. L., Rosen, M. L., Williams, E. S. & McLaughlin, K. A. (2020) Biological aging in childhood and adolescence following experiences of threat and deprivation: A systematic review and meta-analysis. *Psychol Bull* **146**, 721-764, doi:10.1037/bul0000270.

319 Belsky, J., Steinberg, L. & Draper, P. (1991) Childhood experience, interpersonal development, and reproductive strategy: and evolutionary theory of socialization. *Child Dev* **62**, 647-670, doi:10.1111/j.1467-8624.1991.tb01558.x.

320 Ellis, B. J. (2004) Timing of pubertal maturation in girls: an integrated life history approach. *Psychol Bull* **130**, 920-958, doi:10.1037/0033-2909.130.6.920.

321 Belsky, J., Ruttle, P. L., Boyce, W. T., Armstrong, J. M. & Essex, M. J. (2015) Early adversity, elevated stress physiology, accelerated sexual maturation, and poor health in females. *Dev Psychol* **51**, 816-822, doi:10.1037/dev0000017.

322 Franke, K., Ziegler, G., Kloppel, S., Gaser, C. & Alzheimer's Disease Neuroimaging, I. (2010) Estimating the age of healthy subjects from T1-weighted MRI scans using kernel methods: exploring the influence of various parameters. *Neuroimage* **50**, 883-892, doi:10.1016/j.neuroimage.2010.01.005.

323 Franke, K. & Gaser, C. (2019) Ten Years of BrainAGE as a Neuroimaging Biomarker of Brain Aging: What Insights Have We Gained? *Front Neurol* **10**, 789, doi:10.3389/fneur.2019.00789.

324 Shelton, K. K., Frick, P. J. & Wootton, J. (1996) Assessment of parenting practices in families of elementary school-age children. *J Clin Child Psychol* **25**, 317-329, doi:DOI 10.1207/s15374424jccp2503_8.

325 Rijlaarsdam, J. *et al.* (2013) Economic disadvantage and young children's emotional and behavioral problems: mechanisms of risk. *J Abnorm Child Psychol* **41**, 125-137, doi:10.1007/s10802-012-9655-2.

326 Muetzel, R. L. *et al.* (2019) Frequent Bullying Involvement and Brain Morphology in Children. *Front Psychiatry* **10**, 696, doi:10.3389/fpsyt.2019.00696.

327 Desikan, R. S. *et al.* (2006) An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* **31**, 968-980, doi:10.1016/j.neuroimage.2006.01.021.

328 Cole, J. H., Marioni, R. E., Harris, S. E. & Deary, I. J. (2019) Brain age and other bodily 'ages': implications for neuropsychiatry. *Mol Psychiatry* **24**, 266-281, doi:10.1038/s41380-018-0098-1.

329 Hogestol, E. A. *et al.* (2019) Cross-Sectional and Longitudinal MRI Brain Scans Reveal Accelerated Brain Aging in Multiple Sclerosis. *Front Neurol* **10**, 450, doi:10.3389/fneur.2019.00450.

330 Richard, G. *et al.* (2018) Assessing distinct patterns of cognitive aging using tissue-specific brain age prediction based on diffusion tensor imaging and brain morphometry. *PeerJ* **6**, e5908, doi:10.7717/peerj.5908.

331 Kaufmann, T. *et al.* (2019) Common brain disorders are associated with heritable patterns of apparent aging of the brain. *Nat Neurosci* **22**, 1617-1623, doi:10.1038/s41593-019-0471-7.

332 Petersen, A. C., Crockett, L., Richards, M. & Boxer, A. (1988) A self-report measure of pubertal status: Reliability, validity, and initial norms. *J Youth Adolesc* **17**, 117-133, doi:10.1007/BF01537962.

333 Rasmussen, A. R. *et al.* (2015) Validity of self-assessment of pubertal maturation. *Pediatrics* **135**, 86-93, doi:10.1542/peds.2014-0793.

334 Thijssen, S., Collins, P. F. & Luciana, M. (2019) Pubertal development mediates the association between family environment and brain structure and function in childhood. *Dev Psychopathol*, 1-16, doi:10.1017/S0954579419000580.

335 Hagler, D. J., Jr. *et al.* (2019) Image processing and analysis methods for the Adolescent Brain Cognitive Development Study. *Neuroimage* **202**, 116091, doi:10.1016/j.neuroimage.2019.116091.

336 Tianqi Chen *et al.* (2019) *xgboost: Extreme Gradient Boosting*.

337 Le, T. T. *et al.* (2018) A Nonlinear Simulation Framework Supports Adjusting for Age When Analyzing BrainAGE. *Front Aging Neurosci* **10**, 317, doi:10.3389/fnagi.2018.00317.

338 Giedd, J. N. *et al.* (1999) Brain development during childhood and adolescence: a longitudinal MRI study. *Nat Neurosci* **2**, 861-863, doi:10.1038/13158.

339 Shaw, P. *et al.* (2006) Intellectual ability and cortical development in children and adolescents. *Nature* **440**, 676-679, doi:10.1038/nature04513.

340 Walhovd, K. B., Fjell, A. M., Giedd, J., Dale, A. M. & Brown, T. T. (2017) Through Thick and Thin: a Need to Reconcile Contradictory Results on Trajectories in Human Cortical Development. *Cereb Cortex* **27**, 1472-1481, doi:10.1093/cercor/bhv301.

341 Hart, H. & Rubia, K. (2012) Neuroimaging of child abuse: a critical review. *Front Hum Neurosci* **6**, 52, doi:10.3389/fnhum.2012.00052.

342 Herting, M. M. & Sowell, E. R. (2017) Puberty and structural brain development in humans. *Front Neuroendocrinol* **44**, 122-137, doi:10.1016/j.yfrne.2016.12.003.

343 Docherty, A. R. *et al.* (2015) Does degree of gyrification underlie the phenotypic and genetic associations between cortical surface area and cognitive ability? *Neuroimage* **106**, 154-160, doi:10.1016/j.neuroimage.2014.11.040.

344 Zilles, K., Palomero-Gallagher, N. & Amunts, K. (2013) Development of cortical folding during evolution and ontogeny. *Trends Neurosci* **36**, 275-284, doi:10.1016/j.tins.2013.01.006.

345 Duret, P. *et al.* (2018) Gyrification changes are related to cognitive strengths in autism. *Neuroimage Clin* **20**, 415-423, doi:10.1016/j.nicl.2018.04.036.

346 Blanken, L. M. *et al.* (2015) Cortical morphology in 6- to 10-year old children with autistic traits: a population-based neuroimaging study. *Am J Psychiatry* **172**, 479-486, doi:10.1176/appi.ajp.2014.14040482.

347 Matsuda, Y. & Ohi, K. (2018) Cortical gyrification in schizophrenia: current perspectives. *Neuropsychiatr Dis Treat* **14**, 1861-1869, doi:10.2147/NDT.S145273.

348 Cao, B. *et al.* (2017) Lifespan Gyrification Trajectories of Human Brain in Healthy Individuals and Patients with Major Psychiatric Disorders. *Sci Rep* **7**, 511, doi:10.1038/s41598-017-00582-1.

349 White, T., Su, S., Schmidt, M., Kao, C. Y. & Sapiro, G. (2010) The development of gyrification in childhood and adolescence. *Brain Cogn* **72**, 36-45, doi:10.1016/j.bandc.2009.10.009.

350 Gregory, M. D. *et al.* (2016) Regional Variations in Brain Gyrification Are Associated with General Cognitive Ability in Humans. *Curr Biol* **26**, 1301-1305, doi:10.1016/j.cub.2016.03.021.

351 Hogstrom, L. J., Westlye, L. T., Walhovd, K. B. & Fjell, A. M. (2013) The structure of the cerebral cortex across adult life: age-related patterns of surface area, thickness, and gyrification. *Cereb Cortex* **23**, 2521-2530, doi:10.1093/cercor/bhs231.

352 Schaer, M. *et al.* (2008) A surface-based approach to quantify local cortical gyrification. *IEEE Trans Med Imaging* **27**, 161-170, doi:10.1109/TMI.2007.903576.

353 Battaglini, M. *et al.* (2019) Lifespan normative data on rates of brain volume changes. *Neurobiol Aging* **81**, 30-37, doi:10.1016/j.neurobiolaging.2019.05.010.

354 Ikram, M. A. *et al.* (2017) The Rotterdam Study: 2018 update on objectives, design and main results. *Eur J Epidemiol* **32**, 807-850, doi:10.1007/s10654-017-0321-4.

355 Bleecker, M. L., Bolla-Wilson, K., Agnew, J. & Meyers, D. A. (1988) Age-related sex differences in verbal memory. *J Clin Psychol* **44**, 403-411.

356 Houx, P. J., Jolles, J. & Vreeling, F. W. (1993) Stroop interference: aging effects assessed with the Stroop Color-Word Test. *Exp Aging Res* **19**, 209-224, doi:10.1080/03610739308253934.

357 Welsh, K. A. *et al.* (1994) The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part V. A normative study of the neuropsychological battery. *Neurology* **44**, 609-614.

358 Deary, I. J. (2014) The stability of intelligence from childhood to old age. *Current Directions in Psychological Science***.23**, pp, doi:10.1177/0963721414536905.

359 Schaer, M. *et al.* (2012) How to measure cortical folding from MR images: a step-by-step tutorial to compute local gyrification index. *J Vis Exp*, e3417, doi:10.3791/3417.

360 Taylor, P., Hobbs, J. N., Burroni, J. & Siegelmann, H. T. (2015) The global landscape of cognition: hierarchical aggregation as an organizational principle of human cortical networks and functions. *Sci Rep* **5**, 18112, doi:10.1038/srep18112.

361 Cox, S. R. *et al.* (2018) Brain cortical characteristics of lifetime cognitive ageing. *Brain Struct Funct* **223**, 509-518, doi:10.1007/s00429-017-1505-0.

362 Greenland, S. *et al.* (2016) Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *Eur J Epidemiol* **31**, 337-350, doi:10.1007/s10654-016-0149-3.

363 Yang, H. *et al.* (2019) Study of brain morphology change in Alzheimer's disease and amnestic mild cognitive impairment compared with normal controls. *Gen Psychiatr* **32**, e100005, doi:10.1136/gpsych-2018-100005.

364 Ossenkoppele, R. *et al.* (2019) Associations between tau, Abeta, and cortical thickness with cognition in Alzheimer disease. *Neurology* **92**, e601-e612, doi:10.1212/WNL.0000000000006875.

365 Dickerson, B. C. *et al.* (2009) Differential effects of aging and Alzheimer's disease on medial temporal lobe cortical thickness and surface area. *Neurobiol Aging* **30**, 432-440, doi:10.1016/j.neurobiolaging.2007.07.022.

366 Price, C. J. (2012) A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage* **62**, 816-847, doi:10.1016/j.neuroimage.2012.04.062.

367 Jeneson, A. & Squire, L. R. (2012) Working memory, long-term memory, and medial temporal lobe function. *Learn Mem* **19**, 15-25, doi:10.1101/lm.024018.111.

368 Kohli, J. S., Kinnear, M. K., Martindale, I. A., Carper, R. A. & Muller, R. A. (2019) Regionally decreased gyrification in middle-aged adults with autism spectrum disorders. *Neurology* **93**, e1900-e1905, doi:10.1212/WNL.0000000000008478.

369 Palaniyappan, L. & Liddle, P. F. (2012) Aberrant cortical gyrification in schizophrenia: a surface-based morphometry study. *J Psychiatry Neurosci* **37**, 399-406, doi:10.1503/jpn.110119.

370 Nesvag, R. *et al.* (2014) Reduced brain cortical folding in schizophrenia revealed in two independent samples. *Schizophr Res* **152**, 333-338, doi:10.1016/j.schres.2013.11.032.

371 Madre, M. *et al.* (2019) Structural abnormality in schizophrenia versus bipolar disorder: A whole brain cortical thickness, surface area, volume and gyrification analyses. *Neuroimage Clin* **25**, 102131, doi:10.1016/j.nicl.2019.102131.

372 Tan, P. K., Ananyev, E. & Hsieh, P. J. (2019) Distinct Genetic Signatures of Cortical and Subcortical Regions Associated with Human Memory. *eNeuro* **6**, doi:10.1523/ENEURO.0283-19.2019.

373 Ersland, K. M. *et al.* (2012) Gene-based analysis of regionally enriched cortical genes in GWAS data sets of cognitive traits and psychiatric disorders. *PLoS One* **7**, e31687, doi:10.1371/journal.pone.0031687.

374 Van der Lee, S. J. *et al.* (2019) A genome-wide association study identifies genetic loci associated with specific lobar brain volumes. *Commun Biol* **2**, 285, doi:10.1038/s42003-019-0537-9.

375 Richman, D. P., Stewart, R. M., Hutchinson, J. W. & Caviness, V. S., Jr. (1975) Mechanical model of brain convolutional development. *Science* **189**, 18-21.

376 Taki, Y. *et al.* (2011) Correlations among brain gray matter volumes, age, gender, and hemisphere in healthy individuals. *PLoS One* **6**, e22734, doi:10.1371/journal.pone.0022734.

377 Schippling, S. *et al.* (2017) Global and regional annual brain volume loss rates in physiological aging. *J Neurol* **264**, 520-528, doi:10.1007/s00415-016-8374-y.

378 Fjell, A. M. *et al.* (2014) Accelerating cortical thinning: unique to dementia or universal in aging? *Cereb Cortex* **24**, 919-934, doi:10.1093/cercor/bhs379.

379 Van Essen, D. C. (1997) A tension-based theory of morphogenesis and compact wiring in the central nervous system. *Nature* **385**, 313-318, doi:10.1038/385313a0.

380 Inano, S., Takao, H., Hayashi, N., Abe, O. & Ohtomo, K. (2011) Effects of age and gender on white matter integrity. *AJNR Am J Neuroradiol* **32**, 2103-2109, doi:10.3174/ajnr.A2785.

381 Burzynska, A. Z. *et al.* (2017) White Matter Integrity Declined Over 6-Months, but Dance Intervention Improved Integrity of the Fornix of Older Adults. *Front Aging Neurosci* **9**, 59, doi:10.3389/fnagi.2017.00059.

382 Xu, G. *et al.* (2010) Axons pull on the brain, but tension does not drive cortical folding. *J Biomech Eng* **132**, 071013, doi:10.1115/1.4001683.

383 Madan, C. R. & Kensinger, E. A. (2017) Test-retest reliability of brain morphology estimates. *Brain Inform* **4**, 107-121, doi:10.1007/s40708-016-0060-4.

384 Madan, C. R. (2018) Age differences in head motion and estimates of cortical morphology. *PeerJ* **6**, e5176, doi:10.7717/peerj.5176.

385 Pfefferbaum, A. & Sullivan, E. V. (2015) Cross-sectional versus longitudinal estimates of age-related changes in the adult brain: overlaps and discrepancies. *Neurobiol Aging* **36**, 2563-2567, doi:10.1016/j.neurobiolaging.2015.05.005.

386 Wood, R. L. (2017) Accelerated cognitive aging following severe traumatic brain injury: A review. *Brain Inj* **31**, 1270-1278, doi:10.1080/02699052.2017.1332387.

387 Cole, J. H., Leech, R., Sharp, D. J. & Alzheimer's Disease Neuroimaging, I. (2015) Prediction of brain age suggests accelerated atrophy after traumatic brain injury. *Ann Neurol* **77**, 571-581, doi:10.1002/ana.24367.

388 Mende, M. A. (2019) Alcohol in the Aging Brain - The Interplay Between Alcohol Consumption, Cognitive Decline and the Cardiovascular System. *Front Neurosci* **13**, 713, doi:10.3389/fnins.2019.00713.

389 Vemuri, P. *et al.* (2014) Association of lifetime intellectual enrichment with cognitive decline in the older population. *JAMA Neurol* **71**, 1017-1024, doi:10.1001/jamaneurol.2014.963.

390 Borenstein, A. R., Copenhaver, C. I. & Mortimer, J. A. (2006) Early-life risk factors for Alzheimer disease. *Alzheimer Dis Assoc Disord* **20**, 63-72, doi:10.1097/01.wad.0000201854.62116.d7.

391 Whalley, L. J. *et al.* (2000) Childhood mental ability and dementia. *Neurology* **55**, 1455-1459.

392 Stern, Y. *et al.* (1994) Influence of education and occupation on the incidence of Alzheimer's disease. *JAMA* **271**, 1004-1010.

393 Meng, X. & D'Arcy, C. (2012) Education and dementia in the context of the cognitive reserve hypothesis: a systematic review with meta-analyses and qualitative analyses. *PLoS One* **7**, e38268, doi:10.1371/journal.pone.0038268.

394 Van Loenhoud, A. C., Groot, C., Vogel, J. W., Van der Flier, W. M. & Ossenkoppele, R. (2018) Is intracranial volume a suitable proxy for brain reserve? *Alzheimers Res Ther* **10**, 91, doi:10.1186/s13195-018-0408-5.

395 Reed, B. R. *et al.* (2010) Measuring cognitive reserve based on the decomposition of episodic memory variance. *Brain* **133**, 2196-2209, doi:10.1093/brain/awq154.

396 Petkus, A. J. *et al.* (2019) General and domain-specific cognitive reserve, mild cognitive impairment, and dementia risk in older women. *Alzheimers Dement (N Y)* **5**, 118-128, doi:10.1016/j.trci.2019.02.003.

397 Stern, Y. (2017) An approach to studying the neural correlates of reserve. *Brain Imaging Behav* **11**, 410-416, doi:10.1007/s11682-016-9566-x.

398 Lawrence, E. *et al.* (2017) A Systematic Review of Longitudinal Studies Which Measure Alzheimer's Disease Biomarkers. *J Alzheimers Dis* **59**, 1359-1379, doi:10.3233/JAD-170261.

399 Bigler, E. D. & Tate, D. F. (2001) Brain volume, intracranial volume, and dementia. *Invest Radiol* **36**, 539-546.

400 Sharp, E. S. & Gatz, M. (2011) Relationship between education and dementia: an updated systematic review. *Alzheimer Dis Assoc Disord* **25**, 289-304, doi:10.1097/WAD.0b013e318211c83c.

401 Zahodne, L. B. *et al.* (2013) Quantifying cognitive reserve in older adults by decomposing episodic memory variance: replication and extension. *J Int Neuropsychol Soc* **19**, 854-862, doi:10.1017/S1355617713000738.

402 Hu, L. T. & Bentler, P. M. (1999) Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives. *Struct Equ Modeling* **6**, 1-55, doi:10.1080/10705519909540118.

403 Vrooman, H. A. *et al.* (2007) Multi-spectral brain tissue segmentation using automatically trained k-Nearest-Neighbor classification. *Neuroimage* **37**, 71-81, doi:10.1016/j.neuroimage.2007.05.018.

404 De Boer, R. *et al.* (2009) White matter lesion extension to automatic brain tissue segmentation on MRI. *Neuroimage* **45**, 1151-1161, doi:10.1016/j.neuroimage.2009.01.011.

405 Folstein, M. F., Folstein, S. E. & McHugh, P. R. (1975) "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* **12**, 189-198, doi:10.1016/0022-3956(75)90026-6.

406 Copeland, J. R. *et al.* (1976) A semi-structured clinical interview for the assessment of diagnosis and mental state in the elderly: the Geriatric Mental State Schedule. I. Development and reliability. *Psychol Med* **6**, 439-449.

407 Roth, M. *et al.* (1986) CAMDEX. A standardised instrument for the diagnosis of mental disorder in the elderly with special reference to the early detection of dementia. *Br J Psychiatry* **149**, 698-709.

408 Association, A. P. (1987) *Diagnostic and statistical manual of mental disorders*. Third edition revised edn.

409 Hollander, M. *et al.* (2002) Carotid plaques increase the risk of stroke and subtypes of cerebral infarction in asymptomatic elderly: the Rotterdam study. *Circulation* **105**, 2872-2877, doi:10.1161/01.cir.0000018650.58984.75.

410 Van der Lee, S. J. *et al.* (2018) The effect of APOE and other common genetic variants on the onset of Alzheimer's disease and dementia: a community-based cohort study. *Lancet Neurol* **17**, 434-444, doi:10.1016/S1474-4422(18)30053-X.

411 Bodner, T. E. (2008) What improves with increased missing data imputations? *Structural Equation Modeling* **15**, 651-675, doi:10.1080/10705510802339072.

412 Therneau, T. & Grambsch, P. (2000) *Modeling survival data: Extending the Cox model*. (Springer).

413 Rosseel, Y. (2012) lavaan: An R Package for Structural Equation Modeling. *J Stat Softw* **48**, 1-36.

414 Clare, L. *et al.* (2017) Potentially modifiable lifestyle factors, cognitive reserve, and cognitive function in later life: A cross-sectional study. *PLoS Med* **14**, e1002259, doi:10.1371/journal.pmed.1002259.

415 Perneczky, R. *et al.* (2010) Head circumference, atrophy, and cognition: implications for brain reserve in Alzheimer disease. *Neurology* **75**, 137-142, doi:10.1212/WNL.0b013e3181e7ca97.

416 An, H. *et al.* (2016) Large intracranial volume accelerates conversion to dementia in males and APOE4 non-carriers with mild cognitive impairment. *Int Psychogeriatr* **28**, 769-778, doi:10.1017/S104161021500229X.

417 Wolf, H., Julin, P., Gertz, H. J., Winblad, B. & Wahlund, L. O. (2004) Intracranial volume in mild cognitive impairment, Alzheimer's disease and vascular dementia: evidence for brain reserve? *Int J Geriatr Psychiatry* **19**, 995-1007, doi:10.1002/gps.1205.

418 Tate, D. F. *et al.* (2011) Intracranial volume and dementia: some evidence in support of the cerebral reserve hypothesis. *Brain Res* **1385**, 151-162, doi:10.1016/j.brainres.2010.12.038.

419 Jenkins, R., Fox, N. C., Rossor, A. M., Harvey, R. J. & Rossor, M. N. (2000) Intracranial volume and Alzheimer disease: evidence against the cerebral reserve hypothesis. *Arch Neurol* **57**, 220-224.

420 Edland, S. D. *et al.* (2002) Total intracranial volume: normative values and lack of association with Alzheimer's disease. *Neurology* **59**, 272-274.

421 Van Hek, M., Kraaykamp, G. & Wolbers, M. H. J. (2016) Comparing the gender gap in educational attainment: the impact of emancipatory contexts in 33 cohorts across 33 countries. *Educ Res Eval* **22**, 260-282, doi:10.1080/13803611.2016.1256222.

422 McKhann, G. M. *et al.* (2011) The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* **7**, 263-269, doi:10.1016/j.jalz.2011.03.005.

423 Pettigrew, C. *et al.* (2017) Cognitive reserve and cortical thickness in preclinical Alzheimer's disease. *Brain Imaging Behav* **11**, 357-367, doi:10.1007/s11682-016-9581-y.

424 Soldan, A. *et al.* (2015) Relationship of medial temporal lobe atrophy, APOE genotype, and cognitive reserve in preclinical Alzheimer's disease. *Hum Brain Mapp* **36**, 2826-2841, doi:10.1002/hbm.22810.

425 Soldan, A. *et al.* (2017) Cognitive reserve and long-term change in cognition in aging and preclinical Alzheimer's disease. *Neurobiol Aging* **60**, 164-172, doi:10.1016/j.neurobiolaging.2017.09.002.

426 Katzman, R. *et al.* (1988) Clinical, pathological, and neurochemical changes in dementia: a subgroup with preserved mental status and numerous neocortical plaques. *Ann Neurol* **23**, 138-144, doi:10.1002/ana.410230206.

427 Mortimer, J. A. (1988) in *Etiology of dementia of Alzheimer's type* 39-52 (John Wiley and Sons).

428 Medaglia, J. D., Pasqualetti, F., Hamilton, R. H., Thompson-Schill, S. L. & Bassett, D. S. (2017) Brain and cognitive reserve: Translation via network control theory. *Neurosci Biobehav Rev* **75**, 53-64, doi:10.1016/j.neubiorev.2017.01.016.

429 Ferreira, D., Nordberg, A. & Westman, E. (2020) Biological subtypes of Alzheimer disease: A systematic review and meta-analysis. *Neurology* **94**, 436-448, doi:10.1212/WNL.0000000000009058.

430 Jeffery, K. J. (2018) The Hippocampus: From Memory, to Map, to Memory Map. *Trends Neurosci* **41**, 64-66, doi:10.1016/j.tins.2017.12.004.

431 Dale, A. M., Fischl, B. & Sereno, M. I. (1999) Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage* **9**, 179-194, doi:10.1006/nimg.1998.0395.

432 Zilles, K. (2018) Brodmann: a pioneer of human brain mapping-his impact on concepts of cortical organization. *Brain* **141**, 3262-3278, doi:10.1093/brain/awy273.

433 Glasser, M. F. *et al.* (2016) A multi-modal parcellation of human cerebral cortex. *Nature* **536**, 171-178, doi:10.1038/nature18933.

434 Alexander, B. *et al.* (2019) Desikan-Killiany-Tourville Atlas Compatible Version of M-CRIB Neonatal Parcellated Whole Brain Atlas: The M-CRIB 2.0. *Front Neurosci* **13**, 34, doi:10.3389/fnins.2019.00034.

435 Iglesias, J. E. *et al.* (2015) A computational atlas of the hippocampal formation using ex vivo, ultra-high resolution MRI: Application to adaptive segmentation of in vivo MRI. *Neuroimage* **115**, 117-137, doi:10.1016/j.neuroimage.2015.04.042.

436 Salehi, M. *et al.* (2020) There is no single functional atlas even for a single individual: Functional parcel definitions change with task. *Neuroimage* **208**, 116366, doi:10.1016/j.neuroimage.2019.116366.

437 Chakravarty, M. M. *et al.* (2013) Performing label-fusion-based segmentation using multiple automatically generated templates. *Hum Brain Mapp* **34**, 2635-2654.

438 Alfaro-Almagro, F. *et al.* (2020) Confound modelling in UK Biobank brain imaging. *NeuroImage*, 117002, doi:10.1016/j.neuroimage.2020.117002.

439 Doll, R. *et al.* (1965) Mortality of Gasworkers with Special Reference to Cancers of the Lung and Bladder, Chronic Bronchitis, and Pneumoconiosis. *Br J Ind Med* **22**, 1-12, doi:10.1136/oem.22.1.1.

440 Nohr, E. A. & Liew, Z. (2018) How to investigate and adjust for selection bias in cohort studies. *Acta Obstetricia et Gynecologica Scandinavica* **97**, 407-416, doi:10.1111/aogs.13319.

441 Swanson, J. M. (2012) The UK Biobank and selection bias. *Lancet* **380**, 110, doi:10.1016/S0140-6736(12)61179-9.

442 LeWinn, K. Z., Sheridan, M. A., Keyes, K. M., Hamilton, A. & McLaughlin, K. A. (2017) Sample composition alters associations between age and brain structure. *Nat Commun* **8**, 874, doi:10.1038/s41467-017-00908-7.

443 Fortin, J. P. *et al.* (2018) Harmonization of cortical thickness measurements across scanners and sites. *Neuroimage* **167**, 104-120, doi:10.1016/j.neuroimage.2017.11.024.

444 Pomponio, R. *et al.* (2020) Harmonization of large MRI datasets for the analysis of brain imaging patterns throughout the lifespan. *Neuroimage* **208**, 116450, doi:10.1016/j.neuroimage.2019.116450.

445 Makowski, C., Lepage, M. & Evans, A. C. (2019) Head motion: the dirty little secret of neuroimaging in psychiatry. *J Psychiatry Neurosci* **44**, 62-68, doi:10.1503/jpn.180022.

446 Bolton, T. A. W. *et al.* (2020) Agito ergo sum: Correlates of spatio-temporal motion characteristics during fMRI. *Neuroimage* **209**, 116433, doi:10.1016/j.neuroimage.2019.116433.

447 Couvy-Duchesne, B. *et al.* (2016) Head Motion and Inattention/Hyperactivity Share Common Genetic Influences: Implications for fMRI Studies of ADHD. *PloS one* **11**, e0146271-e0146271, doi:10.1371/journal.pone.0146271.

448 Engelhardt, L. E. *et al.* (2017) Children's head motion during fMRI tasks is heritable and stable over time. *Developmental cognitive neuroscience* **25**, 58-68, doi:10.1016/j.dcn.2017.01.011.

449 Asimov, I. (1989) The relatively of wrong. *The Skeptical Inquirer* **14**, 35-44.

450 Livingston, G. *et al.* (2017) Dementia prevention, intervention, and care. *Lancet* **390**, 2673-2734, doi:10.1016/S0140-6736(17)31363-6.

451 Ahmad, S. *et al.* (2018) Disentangling the biological pathways involved in early features of Alzheimer's disease in the Rotterdam Study. *Alzheimers Dement* **14**, 848-857, doi:10.1016/j.jalz.2018.01.005.

452 Vinueza-Veloz, M. F. *et al.* (2020) Genetic risk for Alzheimer disease in children: Evidence from early-life IQ and brain white-matter microstructure. *Genes Brain Behav* **19**, e12656, doi:10.1111/gbb.12656.

453 Chasman, D. I., Giulianini, F., Demler, O. V. & Udler, M. S. (2020) Pleiotropy-Based Decomposition of Genetic Risk Scores: Association and Interaction Analysis for Type 2 Diabetes and CAD. *Am J Hum Genet* **106**, 646-658, doi:10.1016/j.ajhg.2020.03.011.

454 Abraham, G. *et al.* (2019) Genomic risk score offers predictive performance comparable to clinical risk factors for ischaemic stroke. *Nat Commun* **10**, 5819, doi:10.1038/s41467-019-13848-1.

455 Martin, A. R. *et al.* (2019) Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat Genet* **51**, 584-591, doi:10.1038/s41588-019-0379-x.

456 Dikilitas, O. *et al.* (2020) Predictive Utility of Polygenic Risk Scores for Coronary Heart Disease in Three Major Racial and Ethnic Groups. *Am J Hum Genet* **106**, 707-716, doi:10.1016/j.ajhg.2020.04.002.

457 Duncan, L. *et al.* (2019) Analysis of polygenic risk score usage and performance in diverse human populations. *Nat Commun* **10**, 3328, doi:10.1038/s41467-019-11112-0.

458 Shi, H. *et al.* (2020) Localizing Components of Shared Transethnic Genetic Architecture of Complex Traits from GWAS Summary Data. *Am J Hum Genet* **106**, 805-817, doi:10.1016/j.ajhg.2020.04.012.

459 Walters, R. K. *et al.* (2018) Transancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. *Nat Neurosci* **21**, 1656-1669, doi:10.1038/s41593-018-0275-1.

460 Kanai, M. *et al.* (2018) Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases. *Nat Genet* **50**, 390-400, doi:10.1038/s41588-018-0047-6.

461 Malik, R. *et al.* (2018) Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat Genet* **50**, 524-537, doi:10.1038/s41588-018-0058-3.

462 Adams, H. H. H., Evans, T. E. & Terzikhan, N. (2019) The Uncovering Neurodegenerative Insights Through Ethnic Diversity consortium. *Lancet Neurol* **18**, 915, doi:10.1016/S1474-4422(19)30324-2.

463 Nagai, A. *et al.* (2017) Overview of the BioBank Japan Project: Study design and profile. *J Epidemiol* **27**, S2-S8, doi:10.1016/j.je.2016.12.005.

464 Amariuta, T. *et al.* (2020) Improving the trans-ancestry portability of polygenic risk scores by prioritizing variants in predicted cell-type-specific regulatory elements. *Nat Genet* **52**, 1346-1354, doi:10.1038/s41588-020-00740-8.

465 Schaid, D. J., Chen, W. & Larson, N. B. (2018) From genome-wide associations to candidate causal variants by statistical fine-mapping. *Nat Rev Genet* **19**, 491-504, doi:10.1038/s41576-018-0016-z.

466 Collins, R. L. *et al.* (2020) A structural variation reference for medical and population genetics. *Nature* **581**, 444-451, doi:10.1038/s41586-020-2287-8.

467 Van Arensbergen, J. *et al.* (2019) High-throughput identification of human SNPs affecting regulatory element activity. *Nat Genet* **51**, 1160-1169, doi:10.1038/s41588-019-0455-2.

468 Vieira, S., Pinaya, W. H. & Mechelli, A. (2017) Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: Methods and applications. *Neurosci Biobehav Rev* **74**, 58-75, doi:10.1016/j.neubiorev.2017.01.002.

469 Eraslan, G., Avsec, Z., Gagneur, J. & Theis, F. J. (2019) Deep learning: new computational modelling techniques for genomics. *Nat Rev Genet* **20**, 389-403, doi:10.1038/s41576-019-0122-6.

470 Wang, H. T. *et al.* (2020) Finding the needle in a high-dimensional haystack: Canonical correlation analysis for neuroscientists. *Neuroimage* **216**, 116745, doi:10.1016/j.neuroimage.2020.116745.

471 Claes, P. *et al.* (2018) Genome-wide mapping of global-to-local genetic effects on human facial shape. *Nat Genet* **50**, 414-423, doi:10.1038/s41588-018-0057-4.

472 Argelaguet, R. *et al.* (2020) MOFA+: a statistical framework for comprehensive integration of multi-modal single-cell data. *Genome Biol* **21**, 111, doi:10.1186/s13059-020-02015-1.

473 Drysdale, A. T. *et al.* (2017) Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nat Med* **23**, 28-38, doi:10.1038/nm.4246.

474 Dinga, R. *et al.* (2019) Evaluating the evidence for biotypes of depression: Methodological replication and extension of. *Neuroimage Clin* **22**, 101796, doi:10.1016/j.nicl.2019.101796.

475 Winkler, A. M., Renaud, O., Smith, S. M. & Nichols, T. E. (2020) Permutation inference for canonical correlation analysis. *Neuroimage* **220**, 117065, doi:10.1016/j.neuroimage.2020.117065.

476 Smith, S. M. *et al.* (2020) Brain aging comprises many modes of structural and functional change with distinct genetic and biophysical associations. *Elife* **9**, doi:10.7554/eLife.52677.

APPENDICES

A. ACKNOWLEDGEMENTS

The general design of the Generation R Study is made possible by financial support from the Erasmus MC, University Medical Center, Rotterdam, the Netherlands Organization for Health Research and Development (ZonMw), and the Ministry of Health, Welfare and Sport. The ORACLE Study was primarily funded by the European Union's Horizon 2020 research and innovation programme (678543, ORACLE), with additional funding from the Preeclampsia Foundation (2016 Vision Grant) and the Stichting Coolsingel (Project number 471). We gratefully acknowledge the contribution of the children and parents, general practitioners, hospitals, midwives and pharmacies in Rotterdam.

The Rotterdam Study is supported by the Erasmus MC University Medical Center and Erasmus University Rotterdam; The Netherlands Organisation for Scientific Research (NWO); The Netherlands Organisation for Health Research and Development (ZonMw); the Research Institute for Diseases in the Elderly (RIDE); The Netherlands Genomics Initiative (NGI); the Ministry of Education, Culture and Science; the Ministry of Health, Welfare and Sports; the European Commission (DG XII); and the Municipality of Rotterdam. We gratefully acknowledge the contribution of the inhabitants, general practitioners and pharmacists of the Ommoord district.

For chapter 2.2, we would like to acknowledge the UK Biobank (application number 23509). In particular, we thank the participants for their time and dedication.

Part of the data used in the preparation of chapter 5.1 was obtained from the Adolescent Brain Cognitive Development (ABCD) Study (https://abcdstudy.org), held in the NIMH Data Archive (NDA). The ABCD Study is supported by the National Institutes of Health and additional federal partners under award numbers U01DA041022, U01DA041028, U01DA041048, U01DA041089, U01DA041106, U01DA041117, U01DA041120, U01DA041134, U01DA041148, U01DA041156, U01DA041174, U24DA041123, U24DA041147, U01DA041093, and U01DA041025. A full list of supporters is available at https://abcdstudy.org/federal-partners.html. Chapter 5.1 reflects the views of the authors and may not reflect the opinions or views of the NIH or ABCD consortium investigators.

The PhD trajectory was funded through the European Union's Horizon 2020 research and innovation programme (678543, ORACLE).

B. LIST OF PUBLICATIONS

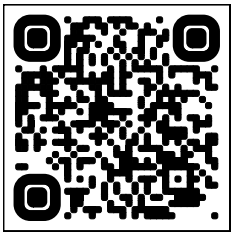
1. **Lamballais, S**. (2013) Optogenetics and its applications in psychology: manipulating the brain using light. *J Eur Psychol Stud* **4**, 87-100, doi: 10.5334/jeps.bc.
2. Van den Berg, W. E., **Lamballais, S.** & Kushner, S. A. (2015) Sex-specific mechanism of social hierarchy in mice. *Neuropsychopharmacology* **40**, 1364-1372, doi: 10.1038/npp.2014.319.
3. **Lamballais, S.***, Sajjad, A.*, Leening, M. J. G., Gaillard, R., Franco, O. H., Mattace-Raso, F. U. S., Jaddoe, V. W. V., Roza, S. J., Tiemeier, H. & Ikram, M. A. (2018) Association of blood pressure and arterial stiffness with cognition in 2 population-based child and adult cohorts. *J Am Heart Assoc* **7**, e009847, doi: 10.1161/JAHA.118.009847.
4. Chen, Z., Schoufour, J. D., Rivadeneira, F., **Lamballais, S.**, Ikram, M. A., Franco, O. H. & Voortman, T. (2019) Plant-based diet and adiposity over time in a middle-aged and elderly population: the Rotterdam Study. *Epidemiology* **30**, 303-310, doi: 10.1097/EDE.0000000000000961.
5. Muetzel, R. L., Mulder, R. H., **Lamballais, S.**, Cortes Hidalgo, A. P., Jansen, P., Güroğlu, B., Vernooij, M. W., Hillegers, M., White, T., El Marroun, H. & Tiemeier, H. (2019) Frequent bullying involvement and brain morphology in children. *Front Psychiatry* **10**, 696, doi: 10.3389/fpsyt.2019.00696.
6. Chen, Z., Franco, O. H., **Lamballais, S.**, Ikram, M. A., Schoufour, J. D., Muka, T. & Voortman, T. (2020) Associations of specific dietary protein with longitudinal insulin resistance, prediabetes and type 2 diabetes: the Rotterdam Study. *Clin Nutr* **39**, 242-249, doi: 10.1016/j.clnu.2019.01.021.
7. Koenraads, S. P. C., van der Schroeff, M. P., van Ingen, G., **Lamballais, S.**, Tiemeier, H., Chang, S. E., Baatenburg, R. J., White, T., Franken, M. C. & Muetzel, R. L. (2020) Structural brain differences in pre-adolescents who persist in and recover from stuttering. *Neuroimage Clin* **27**, 102334, doi: 10.1016/j.nicl.2020.102334.
8. **Lamballais, S.**, Zijlmans, J. L., Vernooij, M. W., Ikram, M. K., Luik, A. I. & Ikram, M. A. (2020) The risk of dementia in relation to cognitive and brain reserve. *J Alzheimers Dis* **77**, 607-618, doi: 10.3233/JAD-200264.
9. **Lamballais, S.**, Muetzel, R. L., Ikram, M. A., Tiemeier, H., Vernooij, M. W., White, T. & Adams, H. H. H. (2020) Genetic burden for late-life neurodegenerative disease and its association with early-life lipids, brain, behavior and cognition. *Front Psychiatry* **11**, 33, doi: 10.3389/fpsyt.2020.00033.

10. **Lamballais, S.**, Vinke, E. J., Vernooij, M. W., Ikram, M. A. & Muetzel, R.L. (2020) Cortical gyrification in relation to age and cognition in older adults. *Neuroimage* **212**, 116637, doi: 10.1016/j.neuroimage.2020.116637.
11. Amiri, M., **Lamballais, S.**, Geenjaar, E., Blanken, L. M. E., Tiemeier, H. & White, T. (2020) Environment-wide association study (E"WAS) of prenatal and perinatal factors associated with autistic traits: a population-based study. *Autism Res* **13**, 1582-1600, doi: 10.1002/aur.2372.
12. Van den Dries, M. A., **Lamballais, S.**, El Marroun, H., Pronk, A., Spaan, S., Ferguson, K. K., Longnecker, M. P., Tiemeier, H. & Guxens, M. (2020) Prenatal exposure to organophosphate pesticides and brain morphology and white matter microstructure in preadolescents. *Environ Res* **191**, 110047, doi: 10.1016/j.envres.2020.110047.
13. **Lamballais, S.***, Adank, M. C.*, Hussainali, R. F., Schalekamp-Timmermans, S., Vernooij, M. W., Luik, A. I., Steegers, E. A. P. & Ikram, M. A. (2021) Design and overview of the Origins of Alzheimer's Disease Across the Life course (ORACLE) Study. *Eur J Epidemiol* **36**, 117-127, doi: 10.1007/s10654-020-00696-3.
14. Steegers, C., Blok, E., **Lamballais, S.**, Jaddoe, V. W. V., Bernardoni, F., Vernooij, M. W., van der Ende, J., Hillegers, M., Ehrlich, S., Jansen, P., Micali, N., Dieleman, G. & White, T. (2021) The association between body mass index and brain morphology in children: a population-based study. *Brain Struct Funct* **226**, 787-800, doi: 10.1007/s00429-020-02209-0.
15. **Lamballais, S.**, Jansen, P. R., Labrecque, J. A., Ikram, M. A. & White, T. (2021) Genetic scores for adult subcortical volumes associate with subcortical volumes during infancy and childhood. *Hum Brain Map* **42**, 1583-1593, doi: 10.1002/hbm.25292.
16. López-Vicente, M., **Lamballais, S.**, Louwen, S., Hillegers, M., Tiemeier, H., Muetzel, R. L. & White, T. (2021) White matter microstructure correlates of age, sex and motor ability in a population-based sample of 3031 school-age children. *Neuroimage* **227**, 117643, doi: 10.1016/j.neuroimage.2020.117643.
17. Zijlmans, J. L., **Lamballais, S.**, Lahousse, L., Vernooij, M. W., Ikram, M. K. & Ikram, M. A. (2021) The interaction of cognitive and brain reserve with frailty in association with mortality: an observational cohort study. *Lancet Healthy Longev* **2**, E-194-E201, doi: 10.1016/S2666-7568(21)00028-3.
18. **Lamballais, S.** & Muetzel, R. L. (2021) QDECR: a flexible, extensible vertex-wise analysis framework in R. *Front Neuroinform* **15**, 561689, doi: 10.3389/fninf.2021.561689.
19. Van Arendonk, J., Yilmaz, P., Steketee, R., Zijlmans, J. L., **Lamballais, S.**, Niessen, W. J., Neitzel, J., Ikram, M. A. & Vernooij, M. W. (2021) Resistance to developing brain pathology due to vascular risk factors: the role of educational attainment. *Neurobiol Aging* **106**, 197-206, doi: 10.1016/j.neurobiolaging.2021.06.006.

20. Vilor-Tejedor, N., Garrido-Martin, D., Rodriguez-Fernandez, B., **Lamballais, S.**, Guigó, R. & Gispert, J. D. (2021) Multivariate Analysis and Modelling of multiple Brain endOphenotypes: Let's MAMBO! *Comput Struct Biotechnol J* **19**, 5800-5810, doi: 10.1016/j.csbj.2021.10.019.
21. Zijlmans, J. L., **Lamballais, S.**, Vernooij, M. W., Ikram, M. A. & Luik, A. I. (*In press*) Sociodemographic, lifestyle, physical and psychosocial determinants of cognitive reserve. *J Alzheimers Dis.*
22. Knol, M. J., Pawlak, M. A., **Lamballais, S.**, Terzikhan, N., Hofer, E., Xiong, Z., Klaver, C. C. W., Pirpamer, L., Vernooij, M. W., Ikram, M. A., Schmidt, R., Kayser, M., Evans, T. E. & Adams, H. H. H. (*In press*) Genetic architecture of orbital telorism. *Hum Mol Genetics*, doi: 10.1093/hmg/ddab334.

PubMed:

Web of Science (login required):



C. PHD PORTFOLIO

Name PhD candidate:	Sander Lamballais
Erasmus MC department:	Epidemiology
PhD period:	2016 –2021
Promotors:	M.A. Ikram
	H.W. Tiemeier
Copromotors:	T.J.H. White
	H. Adams

Activity	Year	ECTS*
PhD training		
Master of Science in Health Sciences, specialization Genetic Epidemiology		
Principles of Research in Medicine	2014	0.7
Genome Wide Association Analysis	2014	1.4
Principles of Genetic Epidemiology	2014	0.7
Genomics in Molecular Medicine	2014	1.4
Advances in Genomics Research	2014	0.4
Study Design	2014	4.3
Biostatistical Methods I: Basic Principles	2014	5.7
Biostatistical Methods II: Classical Regression Models	2014	4.3
Genetic-Epidemiologic Research Methods	2015	5.1
SNPs and Human Diseases	2015	1.4
Linux for Scientists	2015	0.6
Repeated Measurements in Clinical Studies	2015	1.4
Survival Analysis for Clinicians	2015	1.9
Advances in Genome-Wide Association Studies	2015	1.4
Family-Based Genetic Analysis	2015	1.4
A First Encounter With Next-Generation Sequencing Data	2015	1.4
Planning and Evaluation of Screening	2015	1.4
Psychology in Medicine	2015	1.4
English Language	2014	1.4
Introduction to Medical Writing	2014	1.1
Courses for the Quantitative Researcher	2014	1.4
Other Courses		
Causal Inference	2016	0.7
Causal Mediation Analysis	2016	0.7
History of Epidemiological Ideas	2017	0.7

Scientific Integrity	2018	0.3
FSL Course	2019	1.8
MR Safety Training, Level 1 & 2	2016	1.0
Adult Basic Life Support	2018	0.3
Pediatric Basic Life Support	2018	0.1
University Teaching Qualification	2020-2021	3.0
Conferences		
VASCOG 2016, Amsterdam, Netherlands	2016	0.5
ISMRM-ESMRMB 2018, Paris, France (poster)	2018	1.0
OHBM 2018, Rome, Italy (3 posters + oral)	2019	2.5
ResDem 2019, München, Germany	2019	0.5
OHBM 2020, Virtual (2 posters)	2020	1.5
OHBM 2021, Virtual (2 posters)	2021	1.5
Seminars		
Generation R Research Meetings	2016-2020	2.0
Departmental Seminars (Epidemiology)	2016-2020	2.0
Journal Club (Epidemiology)	2016-2020	2.0
2. Teaching activities		
Teaching		
Lecture in Minor “De Gekte Voorbij”	2017	
Teaching Assistant “Cognomics Brain Imaging Genetics Workshop”	2018	
Lecture in Minor “Genetica in de Maatschappij”	2020	
Lecture in Minor “Genetica in de Maatschappij”	2021	
3. Other activities		
Rater for incidental findings (1500+ scans)	2016-2020	4.0
Coordinator incidental findings, neuroimaging Generation R (children)	2016-2017	2.0
Coordinator incidental findings, neuroimaging Generation R (parents)	2017-2020	2.5
Lead PhD Neuroimaging Generation R (children)	2016-2017	
Co-coordinator ORACLE Study	2017-2020	
Assisting in Peer Review (6 manuscripts)	2018-2021	1.8
Peer review (6 manuscripts)	2019-2021	3.6

* 1 ECTS (European Credit Transfer System) equals a workload of 28 hours.

D. WORDS OF GRATITUDE / DANKWOORD

This thesis represents one of the most significant journeys that I have taken in my life. This journey would not have been possible without the help of many, both directly and indirectly. In the strongest terms, I want to thank the participants of Generation R Study, the Rotterdam Study, the ABCD Study and the UK Biobank for their contributions to science. Their time and effort have been invaluable to advances in epidemiological research. I am grateful for being given the opportunity to use the data from these studies for my thesis. I would also like to extent my gratitude to those who made and make these studies possible. I want to personally thank the data management teams of Generation R (**Claudia**, **Eline**, and **Marjolein**) and the Rotterdam Study (**Frank** and **Jolanda**), the secretariat team of Generation R (particularly **Patricia**) and the Neuroepidemiology group (**Gabriëlle** and **Erica**), the Generation R research center team (particularly **Karin**, **Lucienne,** and **Rukiye**), the people who have helped me with the MRI research (**Brendan**, **Floor**, **Juan**, **Maaike**, **Piotr**, **Sylvia,** and **Teun**), and those from the Generation R Management team that helped me along the way (particularly **Janine** and **Vincent**). A special shoutout goes to the students of the Radiology MRI team, some of whom I spent days on end with and who have made my journey much more interesting: **Chiara**, **Eline B**, **Eline van C**, **Eline W**, **Esmee**, **Femke**, **Freya**, **Jessica**, **Laura**, **Marco**, **Merel**, **Michiel**, **Nadine**, **Özlem**, and **Sarina**. Without all of your involvement, this journey would not have been possible. And, of course, a personal thanks to the (ex-)coordinator of the Generation R research center: **Ronald**. Ronald, dank je voor alle hulp en gezelligheid door de jaren heen, jouw nuchterheid en humor kwamen altijd op het goede moment!

My gratitude also extends to my supervisors: **Arfan Ikram**, **Henning Tiemeier**, **Tonya White**, and **Hieab Adams**.

Arfan, I vividly recall our first meeting, my job interview. Rather than topics typical to a job interview, we primarily discussed random scientific topics and our views on them. This set the tone for the rest of my PhD. You have been a marvelous supervisor and collaborator. You always provided me with the room to express and develop my thoughts and ideas, while always making sure I did not stray too far from where I should be heading. Even when we had our disagreements, I always felt that we were on the same page. I think this is because I am also an R-fan. Thank you for all your help and guidance, both professionally and personally.

Henning, the first time we met was over 10 years ago, and while I could list any of numerous anecdotes, I'd like to do this like our meetings: short and to the point, with some banter at the end. Across the last decade, you have been an academic role model for me throughout different phases of my own academic development. Thank you for all the conversations and

discussions; your genuine enthusiasm for science has always been a massive source of inspiration for me. I hope to have many more of those discussions, although we seem to have futureproofed that part.

Tonya, at the end of the day, you are the main reason why I ended up writing this thesis. At the end of my MSc in Health Sciences I was at a crossroads of what I should be pursuing: my postponed dreams of molecular neuroscience, my newfangled passion for epidemiological research, or – realistically – a job outside of academy. Not only did you convince me to stay, you made it possible to do so. Even more so, you were the person to notify me of the vacancy for a PhD on a project regarding reserve. But obviously, you played a much bigger role in my academic career. Our scientific discussions – be it on the autism spectrum or neuroimaging or anything else – were always inspiring, often lasting hours longer than planned, but always filled with new and creative angles and ideas. Your creativity and genuine optimism have greatly shaped my views on science. Thank you for everything.

Hieab, I do not even know where to start. Or when. Or, well, how. It has been an absolute pleasure to work with you. You are an endless stream of ideas and inspiration, you have a deep grasp of all the topics that we ever discuss, your sky does not have a limit, yet still you have time to share memes, come up with hilarious puns ("zinc finger licking good", come on), and have "major" updates? You are a true force of nature. Thanks for everything, both from during my PhD and after.

I would also like to thank the members of the assessment committee: **Neeltje van Haren**, **Michael Ewers**, and **Hilleke Hulshoff Pol**. I would like to thank you all for your participating in this journey. You have all been an inspiring along the way, so it was an honor to have you in my committee. In particular, I want to separately thank **Neeltje** for her role earlier in my PhD:

Beste Neeltje, ik weet nog goed toen ik aan het einde van mijn PhD was en nog weinig zekerheid had over wat ik erna kon gaan doen. Ik klopte bij jou aan, om te kijken of je mensen kende die me verder konden helpen. In plaats daarvan nam je direct meer dan een half uur van je tijd om mijn situatie samen door te lopen, te kijken wat mijn beste opties waren, en om me gerust te stellen dat het wel in orde zou komen. Dit gesprek illustreerde wat ook uit andere situaties bleek, namelijk je passie voor de academie en vooral de mensen erin. Bedankt voor al je hulp, en voor het deelnemen aan de laatste fase van mijn PhD traject.

Over the years, I have met many wonderful people at Generation R that I would like to thank: **Alex**, **Andrea**, **Annemarie**, **Bing**, **Clair**, **Cees**, **Charlotte**, **Chen**, **Desi**, **Desiree**, **Elisabet**, **Elize**, **Fer**, **Gosia**, **Hanan**, **Hannah**, **Ivonne**, **Jeremy**, **Koen**, **Laura**, **Lisa**, **Lisanne**, **Lizzie**, **Lorenza**, **Louk**, **Mannan**, **Marjolein D**, **Michiel**, **Monica**, **Natalie**, **Nikita**, **Nienke**, **Pauline**, **Philip**, **Raisa**, **Rosa**, **Runyu**, **Ryan**, **Sara**, **Scott**, **Simone**, **Sunayna**, **Suzanne**, **Tessa**, and **Yllza**. A special

mention goes out to **Bing**, **Elisabet**, **Isabel**, **Lorenza**, **Louk**, and **Rowina**. A big part of my PhD was the ORACLE Study, AKA the "Generation R parent MRI", so I would like to thank the students (**Demi**, **Evelien**, **Floor**, **Mariska**, **Mellan**, and **Isabel**) as well as the PhDs (**Annemarijne**, **Isabel** again, **Jende**, and **Rowina** again).

Out of all these names, there are several people I would like to thank personally.

Ryan, Ryan, Ryan. I think it's fair to grant you the title of bromotor, because you've been a driving force behind a lot of the work during my PhD. From the MRI protocol to the confounding paper, from random brainstorm sessions about everything neuroepi to the borrels in Na. It has been an interesting collaboration, to say the least, and it's fun to think about how far we've come from where we were 5 years ago. Thanks for everything; I'm looking forward to see what the next few years will bring for you :)

Philip, prs jansen, dit dankwoord zou natuurlijk niet compleet zijn zonder jou. En zeg nou zelf, wat topt op een vrijdagavond Mario uitspelen op kantoor? Veel dingen, maar toch. We hebben veel leuke problemen mogen tackelen, van polygenic scores tot hoe je een piano door een raam kunt takelen, en het blijft één van mijn trotste momenten dat ik de Philip Jansen pubquiz heb gewonnen. Dank je dat je er altijd was tijdens mijn PhD om me te helpen als ik dat nodig had, en bovenal bedankt voor alle gezelligheid.

Rowina en **Louk**. In plaats van 2 paragrafen per persoon krijgen jullie 1 paragraaf per 2 personen. Come on, hoe kan ik jullie apart bespreken? Kort gezegd wil ik jullie bedanken voor al jullie support en liefde, voor alle slechte grappen (de lat lag soms laag) en worthy suggestions. Real Talk is pas een paar jaar oud, maar het voelt aan als een eeuwigheid. Of het nou ging over werk of over andere zaken in het leven, jullie waren er altijd, for better or for worse. Love you guys, truly! Louk, voldoet dit aan dQw4w9WgXcQ?

Of course, the remaining members of the "meme group" (still not sure who came up with that name). **Isabel**, we kunnen een traantje of twee (drie?) laten, maar ik blijf het grappig vinden dat ik je op rare uren kan appen over een random stuk R code en dat je vervolgens alles neerlegt om erover te discussiëren. Dank je voor alle hulp en gezelligheid! **Elisabet**, korte versie: bedankt. Lange versie: bedankt! **Lorenza**, I'm trying really hard to think of a good joke, but you'd come up with a better one anyway. Thanks for all the fun conversations and monologues. **Bing**, it was always very nice talking to you and discussing various topics in life, your spirit always brought mine up. Thanks for the gezelligheid!

Yllza, thanks for all the double espressos ;) But in all seriousness, thanks for all the fun discussions and conversations that we had. I vividly recall our meetings starting off with you offering me Clipper tea, which is actually still the tea brand that I buy now and is sitting in front

of me as I type this. But even when we moved to Doppio, it was always a pleasure to discuss everything and anything with you. I'm happy to see where you ended up, and I'm sure we'll have a coffee soon!

Annemarijne, het heeft even geduurd, maar we zijn er toch doorheen gekomen. Het was een leuke samenwerking tijdens de ouder MRI, en het was altijd leuk om te zien hoe we elkaar aanvulden op verschillende vlaken. Bedankt voor alle gezelligheid :)

Desiree, jou kan ik natuurlijk ook niet vergeten! Ik mis onze vele besprekingen (en roddelsessies) over de MRI en het leven, maar ik ben in ieder geval blij om te zien waar we allebei zijn beland. Dank je voor alle gezellige dim sum sessies, moeten we vaker doen!

Mariska, ook jou wilde ik toch even apart bedanken. Vooral ook omdat we elkaar tegen blijven komen, zoals in de Albert Heijn :) Het was altijd een genot om met je samen te werken, en het was heel erg leuk om je te zien groeien over de jaren heen. Heel erg veel succes met je carrière als psycholoog!

Part of my PhD was as a member of the Neuroepidemiology group of the Rotterdam Study, where I also met a lot of inspiring people. I would like to thank **Alis**, **Amber**, **Arno**, **Daniel**, **Eline**, **Frank W**, **Gena**, **Hazel**, **Hoyan**, **Isabelle**, **Janine**, **Joyce**, **Julia**, **Kamran**, **Kimberly**, **Lana**, **Lisanne**, **Lotte**, **Meike**, **Noor**, **Paloma**, **Pauline**, **Pinar**, **Silvan**, **Thom**, and **Zhangling**.

Dear **Zhangling**, I hope you (and Yicheng!) are doing well. It has been a pleasure working with you on various projects, but most of all it was always great to be around you. Your kindness and cheerfulness were always a pleasure. Thank you for everything :)

Pinarrrrrrrrrr, het wordt maar weer eens tijd om de volgende (game)sessie te plannen! Jij natuurlijk ook bedankt voor al je gezelligheid en voor alle (klaag)gesprekken die we hebben gehad; duidelijk een gemis in mijn dag tegenwoordig. Hopelijk gaat alles goed met je, en we spreken elkaar snel weer!

Lana, ik blijf het bizar toevallig vinden dat we elkaar weer eens tegen zijn gekomen binnen het onderzoek. Ik kan alleen maar zeggen dat het voor mij een positieve verrassing was. Ik vond het altijd heel erg gezellig om even bij te praten over alles en niets, dank je dat je er altijd was als ik je nodig had! Het enige jammere is dat we geen paper meer hebben geschreven, maar dat is mijn schuld :P Ik spreek je hopelijk snel weer!

Of course, the journey of a PhD is shaped by more than the PhD trajectory itself. I want to acknowledge all the other people in my life that had a significant influence on where I stand now.

First, my previous mentors.

Beste **Katinka**, ik weet niet of ik dit ooit duidelijk heb uitgesproken, maar zonder jou had ik hier misschien nooit gestaan. Ik weet nog dat het een willekeurig gesprek was bij T-12, toen je aangaf dat ik als BSc student best kon helpen bij onderzoek. Vanaf daar hebben we nog jaren samengewerkt, een ervaring die mij en mijn denken heeft gevormd tot wat het is. Dank je voor al je begeleiding, door de jaren heen!

Dear **Steven**, I want to thank you for having me as part of your group. Not only did you give me ample opportunity to develop myself and hone my skills, I also vividly remember numerous discussions that we had. I don't know if you recall this, but there was this one time I had spent days on a formula that explained the behavior we saw in the tube test. When I presented it to you, you immediately understood what I meant and rewrote it to be much more informative in a matter of minutes. Moments like those inspired me to always strive for the best, and to step past my own boundaries. Thank you for all your help and inspiration.

Next, I want to specifically acknowledge my friends from the NIHES master, the "gen-epics": **Arda**, **Elena**, **Eliana**, **Fjorda**, and **Kate**. Of course, I need to give some extra credit where credit is due ;) **Arda**, dude, thanks for always being there for me. It was always good fun during NIHES, and every time we met afterwards was very valuable to me, like meeting a long lost brother. I hope you're doing great, and stay gezellig ;)! **Eliana**, kiddo, from Guitar Hero to watching fireworks in the Na building while the code was giving errors… It felt like randomly falling face-first while walking. It's been quite a journey. Thanks for all your help along the way, and for all the fun times! **Kate**, it was always an inspiration to watch you during your PhD and seeing your dedication to the craft. I miss hanging out in your office, because every time I came by there were new stories and books to be discussed. I hope you're doing well, and I'm sure you'll do great wherever you go :)

Finally, I want to thank those that stood closest to me. **Ma** en **pa**, bedankt voor al jullie steun door de jaren heen. Jullie hebben er alles aan gedaan om te zorgen dat ik precies kon doen wat ik wilde in het leven. Hier heb ik – gelukkig – gretig gebruik van gemaakt. Bedankt voor alle inspiratie in het leven! **Tim**, het is een eer om zo'n rolmodel als broer te hebben. Tenminste iemand die snapt dat random NES en SNES feitjes het belangrijkste zijn in het leven. En natuurlijk ook **Smita**, bedankt voor alle steun!

En als allerlaatste natuurlijk **Anna**, mijn eeuwige steun en toeverlaat. Bedankt voor alle R tips. Nu ik hier toch sta, wil ik je bedanken voor al je steun door de jaren heen, als collega en vooral als partner. Je bent het zonnetje in mijn leven, en ik kijk uit naar waar het leven ons zal brengen. Gelukkig mogen we deze reis voortzetten in een tweede boekje ;)

E. ABOUT THE AUTHOR

Sander Lamballais was born in Krimpen a/d IJssel in 1991. After completing VWO at the Marnix Gymnasium (Rotterdam), he started his BSc degree in Psychology at the Erasmus University Rotterdam in 2009. During these years, he gained research experience under the supervision of Katinka Dijkstra (Psychology, EUR) and Ype Elgersma (Neuroscience, EMC), where he worked on false memory and tuberous sclerosis, respectively.

In 2012 and after obtaining his BSc degree he joined the Neuroscience MSc program at the Erasmus Medical Center. Under the supervision of Steven Kushner, he studied social dominance in mice using the tube test, showing that the paradigm does not primarily assess social dominance. He learned how to program in MATLAB to create a computational model that provided an alternative mechanism to explain the observed behavior. In 2014, he joined the MSc program in Health Sciences, with a specialization in Genetic Epidemiology. Under the supervision of dr. Tonya White, he explored whether a polygenic score for Alzheimer's Disease associated with hippocampal volume in the children of Generation R.

After completing his MSc degree, he worked with Tonya White to explore non-genetic determinants of autism symptoms within Generation R. In 2016, he started a PhD under the supervision of Arfan Ikram, Henning Tiemeier, and Tonya White, and in 2018 Hieab Adams joined the team as well. The aim of the PhD was to further elucidate the determinants and consequences of cognitive reserve across the lifespan. He was heavily involved in the neuroimaging of both children and parents within the Generation R Study. In 2020, Sander started a part-time and subsequently full-time position at the department of Clinical Genetics, in the group of Hieab Adams. Here, he has focused on the development of methods to innovate the use of genetic and genomic data. In his free time he likes to game and to work on his book "R for nonprogrammers".

Website:

Stack Overflow: