The adolescent brain and depression: A neuroimaging approach to understanding biological and psychosocial risk factors

Niamh MacSweeney

THE UNIVERSITY of EDINBURGH

Doctor of Philosophy

The University of Edinburgh

2022

Thesis Abstract

Adolescence is a period of significant neurodevelopment and increased vulnerability to the onset of depression. However, the neural underpinnings of depression during adolescence and the associated risk factors are not well understood. The aim of this PhD research was to fill this knowledge gap by examining biological and psychosocial factors associated with the emergence of depression during adolescence.

Using a large, population-based sample, the Adolescent Brain Cognitive Development (ABCD) Study, my doctoral work found that depression in early adolescence is associated with similar neuroimaging findings (cortical and white matter microstructural features) to those seen in adult depression samples. Further, the work in this thesis demonstrated that earlier pubertal timing is associated with an increased risk for later depression in adolescence. While earlier pubertal timing was also related to structural brain features, brain structure was not found to mediate the observed association between early pubertal timing and later depressive symptoms. This finding highlights the important role that other aspects of a young person's biology, psychology and social world may play, and should be explored in future work.

This thesis also investigated how dynamic functional brain networks relate to irritability in adolescent depression using a co-produced youth-researcher design. In this pilot study, I first worked with young people to develop a novel fMRI irritability task that reflected the social nature of irritability in adolescence. Using a local sample of youth with depressive symptoms, I found that dynamic functional brain networks differed between the irritability task and a standard resting state scan, which provides preliminary evidence for validation of this novel task. Finally, my work demonstrated that properties of dynamic brain networks related to emotion regulation and cognitive control were associated with youth depressive symptoms and irritable mood.

Taken together, the findings of this thesis suggest that neuroanatomical differences may be present early in the disease course of depression and that biological factors, such as early pubertal development, relate to depression risk. Moreover, this work provides preliminary

evidence to suggest that alterations in dynamic brain network properties are associated with depressive symptoms and irritability in adolescence. Further, this doctoral research highlights the importance of co-produced study designs in developmental cognitive neuroscience. This work makes an important contribution to our understanding of the factors associated with the emergence of depression during adolescence, which lays a strong foundation upon which to base future longitudinal research.

Lay Summary

Adolescence is a time of immense change for our biology, psychology, and social world. These changes bring with them both opportunities and risks. On one hand, they allow young people to move away from childhood and carve their own identities as independent adults. On the other hand, adolescence is also a time when mental health problems, like depression, are most likely to emerge. However, how the brain is associated with the onset of depression and its associated risk factors during adolescence is not well understood. If we can understand how aspects of our biology and behaviour are associated with the emergence of depression during adolescence, we might identify better targets and timings for the treatment and prevention of depression.

In this thesis, I used brain imaging (MRI) data from volunteers in a large population study of adolescents (The Adolescent Brain Cognitive Development Study). My first aim was to examine how brain structure is associated with depression in early adolescence. I found that differences in brain structure, especially in regions involved in emotion regulation and cognitive control, were associated with higher levels of depressive symptoms in youth aged 9-11 years. Overall, the brain structural alterations that were related to adolescent depression were like those observed in adults with depression. However, there were also some differences in brain structure specific to adolescent depression. This suggests that brain structural alterations may be present early in the disease course of depression and that some of these differences may be specific to adolescent-onset depression.

The second aim of this thesis was to examine whether earlier pubertal timing is associated with an increased risk for depression in adolescence, and how brain structure might affect this relationship. Pubertal timing refers to an individual's pubertal development relative to their same-age, same-sex peers. Previous research has found that earlier pubertal timing is associated with an increased risk for depression in both males and females but the role of brain structure in this association had remained unclear. I replicated previous research using a large, demographically diverse sample and found that individuals, aged 10-11 years, who began puberty before their peers were more likely to report higher levels of depression two

years later, when they were aged 12-13 years. I also explored whether specific aspects of brain structure played a role in this association, but I did not find that this was the case. This highlights the need to explore the role that other biological (e.g., genetics, brain function), psychological (e.g., self-esteem), and social factors (e.g., peer and family relations) may play in the association between earlier pubertal timing and increased depression risk in adolescence.

The final aim of this thesis was to explore how brain function was associated with irritability in adolescence, and how this related to depressive symptoms. Irritability is a core symptom of adolescent depression and an early indicator of emotion regulation difficulties. However, existing research on irritability typically overlooks the social nature of adolescence. Therefore, I worked with young people to design an irritability task that aimed to reflect the experience of irritability as a young person today. I then recruited an independent sample of young people who underwent a functional MRI scan while performing our novel irritability task. The task involved reading a series of irritating scenarios and imagining being in those situations as vividly as possible. To validate the task, I first investigated whether patterns of brain activity differed between the irritability task and a scan when the brain is at rest (i.e., the participant looks at a cross on a screen for the duration of the scan). I found that the patterns of brain activity differed across the two conditions, and a brain network involved in cognitive control and goal-oriented behaviour was more likely to be occupied during the irritability task. This suggests that our novel task may induce a state of mind related to emotion regulation. I also found that certain patterns of brain activity were associated with depressive symptoms and irritable mood, which may provide insight into how alterations in brain activity could contribute to the emergence of depression in adolescence.

Ultimately, the work in this thesis has advanced our understanding of the features of brain structure and function that may be associated with depression in adolescence, and how other aspects of social behaviour (e.g., irritable mood) relate to mental health difficulties. To further develop our understanding, we need to examine these associations over time to distinguish the factors that shape positive developmental patterns (e.g., mental wellbeing) and those that increase risk for maladaptive trajectories, such as the onset of depression. The

overarching aim of this research is to identify youth that diverge from a positive developmental trajectory at the earliest stage possible so that we can divert them away from ill-health towards wellbeing.

Declaration of Originality

I declare that this thesis is my own composition and that it has not been submitted for any other degree or professional qualification at this university or any other institution. Parts of the work comprising this thesis have been previously published. The included publications are my own work, expect where indicated otherwise.

The work presented in Chapter 3 has been published in *EClinicalMedicine*. Author contributions are as follows: XS, **NM**, LR and HCW conceived and designed the research; XS led the formal analysis, and **NM** and XS were responsible for result interpretation and visualisation; XS and **NM** drafted the manuscript; all authors reviewed the article.

The work presented in Chapter 4 has received an "in principle acceptance" in *Developmental Cognitive Neuroscience* as a Stage 1 Registered Report. The Stage 2 manuscript has been submitted for review. Author contributions are as follows: NM and HCW conceived and designed the research; **NM** and JA conducted the statistical analyses; **NM** drafted the article; **NM**, JA, AES, XS, HC, BC, RMR, ASFK, LR, and HCW contributed to the methodology; all authors reviewed the article.

The work presented in Chapter 5 has been submitted for publication and is available as a preprint on PsyArXiv. Author contributions are as follows: **NM**, PL, SZ, SC, LR and HCW conceived and designed the research; **NM** and PL conducted the formal analysis (literature review); **NM** drafted the article; all authors reviewed the article.

Signed:

Date: 2nd December 2022

Covid-19 Note

The work contained within this thesis was undertaken between April 2020 and November 2022. When I began my doctoral studies in January 2019, my original PhD project was entirely data collection-based and involved recruiting young people with depressive symptoms from community and clinical settings in Scotland. Therefore, the first 15 months of my PhD were focused on developing a study protocol and preparing an NHS ethics application (IRAS ID: 274402). We obtained a favourable ethics opinion from the West of Scotland Research Ethics Committee (REC) on February 4th, 2020. Recruitment for the study began thereafter but was halted on March 16th, 2020, due to a university-wide decision to suspend all empirical studies that involved in-person data collection, which remained in place until November 2020. Due to the many unknowns surrounding the development of the Covid-19 pandemic, my supervisors and I decided to switch the focus of the PhD to working with pre-existing data. Thankfully, resources like the Adolescent Brain Cognitive Development (ABCD) Study were suited to addressing the original aims of my PhD. Following the lifting of data-collection restrictions in November 2020, I began recruitment and data collection for the pilot neuroimaging study that comprises Chapters 5-7 of this thesis.

Acknowledgements

Firstly, I would like to express my sincere thanks to my primary supervisor, Heather Whalley, for her guidance and support throughout my PhD. Her kind and calm disposition helped see me through difficult points during my studies. I always enjoyed the chats about Highland adventures that punctuated our supervision meetings — the fact that this culminated in an impromptu supervision at the top of Liathach, Torridon, seems like a very fitting conclusion to this journey! I would also like to thank my secondary supervisors, Stephen Lawrie and Alex Kwong for their assistance and advice during my PhD. Thanks also to Stella Chan.

Many thanks to all members, past and present, of the Division of Psychiatry at Kennedy Tower. I would especially like to thank Judith Allardyce for being so generous with her time and expertise — I could not have completed this work without her guidance and support. Thank you also to Amelia Edmondson-Stait and Hannah Casey for their patience and help with my (many) coding woes. I would also like to thank Aleks Stolicyn, Liana Romaniuk, and Xueyi Shen for their expertise and assistance with the statistical analysis in this thesis. Many thanks to Nikolaj Krær Høier and Anders Jespersen for their friendship, Scandi vibes, and ample supply of caffeine, baked goods, and plants (Anders!) on the 7th floor. Thanks especially to Niko for making every day in the office together a joy during the final year of my PhD — your friendship made all the difference, tusind tak min ven!

A special thanks to Laura Klinkhamer, firstly for her friendship, and also for being the best Open Research partner-in-crime! Founding the Edinburgh branch of ReproducibiliTea together has been one of the most rewarding experiences of my academic career and I am so proud of the Open Research community that we helped create. I am fortunate to have met many wonderful and inspiring people through ReproducibiliTea and the Edinburgh Open Research Initiative — thank you for making adopting open research practices a little less daunting!

Thank you to the trustees and volunteers at Mental Health Research UK, without whose support this doctoral work would not have been possible. Thanks also to the staff at the

Clinical Research Imaging Centre for their assistance with our imaging data collection. I would also like to thank the participants of our local study and the ABCD Study, as well as the ABCD Study co-ordinators. I would like to extend a special thanks to our youth researchers Perrine Louvet and Simal Zafar, and to all the other young people that made the research in Chapters 5-7 of this thesis possible. I learned a great deal from you all.

Attending the Modelling Developmental Change ABCD workshop in July 2021 was the highlight of my PhD and introduced me to a wonderful community of bright, curious, and generous minds. A special thanks to Kate Mills and Michelle Byrne for organising, and Landry Goodgame-Huffman for just being awesome!

Thank you to my science teachers, Denis Maguire and Dan McKenna, for their support during my first encounter with the world of research (even if I didn't know it at the time!) at the BT Young Scientist Exhibition 2011. And of course, thanks to Clare Kelly at Trinity College Dublin who played a pivotal role in shaping my very early research career and who inspired me to pursue a PhD.

Thanks also to my friends and family in Scotland and Ireland, and especially to Kadi Vaher whose company during the long library days of our theses write ups made the process a great deal easier! Thank you too to Taylor, especially the release of Midnights, for soundtracking this thesis.

Of course, tusen tusen takk to Sunniva Bøstrand for accompanying me on this PhD journey — I am so proud of us for doing this hard thing together. Here's to manifesting many more future adventures, min skatt.

I am so very lucky to be surrounded by fierce and fabulous women whose "oomph and sparkle" has been an unwavering source of support, encouragement, and inspiration over the course of my PhD. No words can really do justice how grateful I am for your friendship so I'll end with a heartfelt thanks to Zandie, Anna, Julie, Alice, Aoife, Sarah, Sunniva, Naomi, and Line.

Thank you to my dear sister, Aoife, for always being there for me and making me laugh when I needed it most. Thanks to my dogs, Lilly and Dubh, for being there from school exams through to my PhD thesis. My biggest thanks go to my parents, Erina and Terry, for always encouraging me to ask questions and for providing me with every opportunity I could have imagined. Your curiosity and sense of adventure in life is my greatest inspiration and I am so proud to be your daughter.

Lastly, I would like to express my gratitude to the mountains and seas of Ireland and Scotland. There was never a PhD stress that couldn't be remedied by a jaunt up a hill or a dip in the sea!

Publications

Publications included in this thesis:

(*joint-first)

Shen, X.*, **MacSweeney, N.*,** Chan, S. W., Barbu, M. C., Adams, M. J., Romaniuk, L., McIntosh, A.M. & Whalley, H. C. (2021). Brain structural associations with depression in a large early adolescent sample (the ABCD Study®). *EClinicalMedicine.* *Open Access Version.*

MacSweeney, N., Allardyce, J., Edmondson-Stait, A., Shen, X., Casey, H., Chan, S. W. Y., Cullen, B., Reynolds, R., Kwong, A. S. F., Lawrie, S., Romaniuk, L., Whalley, H. C. (in-principle acceptance). The role of brain structure in the association between pubertal timing and depression risk in early adolescence: A registered report. *Developmental Cognitive Neuroscience* (Stage 2 submitted).

MacSweeney, N., Louvet, P., Zafar, S., Chan, S. W. Y. Kwong, A. S. F., Lawrie, S., Romaniuk, L., Whalley, H.C., Keeping up with the kids: Towards a more ecologically valid study of irritability in youth depression and its underlying neural circuitry (under review). Preprint Version.

Other publications completed during my doctoral studies
(*joint-first)

Saragosa-Harris, N.*, M., Chaku, N.*, **MacSweeney, N.*,** Guazzelli Williamson, V.*, Scheuplein, M., Feola, B., … Mills, K. L. (2022). A practical guide for researchers and reviewers using the ABCD Study and other large longitudinal datasets. *Developmental Cognitive Neuroscience.* *Open Access Version.*

O'Rourke, S., Whalley, H., Janes, S., **MacSweeney, N**., Skrenes, A., Crowson, S., MacLean, L., & Schwannauer, M. (2020). The development of cognitive and emotional maturity in adolescents and its relevance in judicial contexts. Scottish Sentencing Council UK. *Commissioned report.* *Online Version.*

Achievements and Awards

Winner, Good Research Citizenship Award at the Good Research Practice Awards, University of Edinburgh for work with Edinburgh ReproducibiliTea as co-founder and organiser (2022)

Highly Commended, British Neuroscience Association Student Credibility Prizes for open research work (2022)

Awardee (Co-I), Student Experience Grant (£5,000), University of Edinburgh, to organise the Edinburgh Open Research Initiative and ReproducibiliTea Conference (2022)

Awardee, Guarantors of Brain Travel Grant (£600) to support invited research visit to Prof. Christian K. Tamnes' Research Group, University of Oslo (2021)

Winner, Delegates' Choice Award, MQ Mental Health Science Summit (£200) for poster entitled "Understanding irritability in adolescent depression: Development of a novel fMRI task using a co-produced youth-research design" (2021)

Representative, University of Edinburgh, LERU Doctoral Summer School, Trinity College Dublin. Theme: Re-evaluating the Role of the Expert **(**2021)

Awardee (PI), Royal Society STEM Partnership Grant (£3,000) for project entitled "Does our biology influence our mood?" in collaboration with Musselburgh Grammar School (2019)

Finalist, University of Edinburgh 3 Minute Thesis Competition, The Whirlwind of Adolescent Depression, YouTube Video (2019)

Scholar, Mental Health Research UK Children and Young People PhD Scholarship (£90k)

Table of Contents

1 General Introduction 1

1.1 Outline 1

1.2 The dynamic world of adolescence 1

1.3 Depression in adolescence 2

1.4 Risk factors for depression — a brief overview 5

1.5 The developing brain 9

1.6 Puberty: The biological catalyst of adolescence 18

1.7 The adolescent brain and depression: previous research and the current thesis 22

1.8 Thesis aims 25

2 ABCD Methods 27

2.1 Chapter introduction 27

2.2 The ABCD Study 27

2.3 ABCD measures 29

2.4 Overview of key statistical methods 37

2.5 Summary 39

3 Brain Structural Associations with Depression in Adolescence 40

3.1 Chapter introduction 40

3.2 Abstract 43

3.3 Introduction 44

3.4 Methods 46

3.5 Results 51

3.6 Discussion 58

3.7 Chapter conclusion 65

4 The Role of Brain Structure in the Association Between Pubertal Timing and Depression Risk 66

4.1 Chapter introduction 66

4.2 Abstract 69

4.3 Introduction 71

4.4 Materials and methods 79

4.5 Data analysis plan 86

4.6 Pilot analyses 92

4.7 Pilot results 95

4.8 Main results 99

4.9 Discussion 111

4.10 Data access 121

4.11 Acknowledgements 121

4.12 CRediT statement 122

5 Irritability in Adolescent Depression — A Narrative Literature Review 124

5.1 Chapter introduction 124

5.2 Abstract 127

5.3 Literature review 128

5.4 CRediT statement 137

5.5 Chapter conclusion 138

6 Irritability in Adolescent Depression — Pilot Study Methods 139

6.1 Chapter introduction 139

6.2 Data source 139

6.3 Funding and ethics 141

6.4 Study procedure 141

6.5 Study materials 142

6.6 Imaging data pre-processing 146

6.7 Leading Eigenvector Dynamics Analysis (LEiDA) 154

6.8 Chapter conclusion 159

7 Exploring dynamic functional brain networks in adolescent depression using a co-produced novel irritability task 160

7.1 Chapter introduction 160

7.2 Introduction 161

7.3 Methods 164

7.4 Results 171

7.5 Discussion 186

7.6 Chapter conclusion 192

8 General Discussion 193

8.1 Chapter Introduction 193

8.2 Summary of findings 193

8.3 Limitations, methodological considerations, and future directions 197

8.4 Concluding remarks 207

9 Appendices 209

10 References 307

Key Terms

AAL	Automated Anatomical Labelling
ABCD STUDY	Adolescent Brain Cognitive Development Study
ACS	American Community Survey
ADHD	Attention deficit hyperactivity disorder
ARI	Affective Reactivity Index
ASD	Autism spectrum disorder
AT	AtlasTrack
BMI	Body mass index
BOLD	Blood-oxygen-level-dependent
BRAINAGE	Brain age gap estimate
CBCL	Child Behaviour Checklist
DAIC	ABCD Data Analytics and Informatics Centre
DAWBA	Development and Well-Being Assessment
DEAP	Data Exploration and Analysis Portal
DFC	Dynamic functional connectivity
DK	Desikan-Killiany Atlas
DLPFC	Dorso-lateral prefrontal cortex
DMDD	Disruptive mood dysregulation disorder
DMN	Default mode network
DPL	Dynamic BOLD phase-locking matrix
DS	Depressive symptoms
DTI	Diffusion tensor imaging
ECC	Eddy current correction
ENIGMA	Enhancing NeuroImaging Genetics through Meta-Analysis
EPI	Echo planar imaging
FA	Fractional anisotropy
FC	Functional connectivity
FD	Framewise displacement

FMRI	Functional magnetic resonance imaging
FOV	Field of view
FPN	Fronto-parietal network
GLM	Generalised linear model
HALFPIPE	*H*armonised *Anal*ysis of *F*unctional MRI *pipe*line
ICA-AROMA	Independent Components Analysis-based Automatic Removal of Motion Artefacts
ICV	Intracranial volume
IRB	Institutional Review Board
IRR	Incidence rate ratio
K-SADS	Kiddie Schedule for Affective Disorders and Schizophrenia
LEIDA	Leading Eigenvector Dynamics Analysis
LME	Linear mixed effect model
MD	Mean diffusivity
MDD	Major depressive disorder
MFQ	Mood and Feelings Questionnaire
MNI	Montreal Neurological Institute
MPFC	Medial prefrontal cortex
NIMH	National Institute for Mental Health
ODD	Oppositional defiant disorder
PDS	Pubertal Development Scale
PHQ-9	Patient Health Questionnaire, depression module
PL	Phase-locking
PT	Pubertal timing
QA	Quality assessment
QC	Quality control
ROI	Region of interest
RS-FMRI	Resting state functional magnetic resonance imaging
RSN	Resting state network
SEM	Structural equation modelling
SN	Salience network

TE	Echo time
TR	Repetition time
TS	Tanner Stage Line Drawings
WBV	Whole brain volume
YPAG	Young Person Advisory Group

List of Figures

Figure 1.1 – The spectrum of depression. 2
Figure 1.2 – Core clinical symptoms of depression (according to the DSM-5). 3
Figure 1.3 — Illustration of cortical thickness measurement. 10
Figure 1.4 — Neurodevelopmental milestones across the lifespan. 12
Figure 1.5 — Illustration of isotopic (A) and anisotropic (B) water diffusion. 14
Figure 1.6 — The physical and hormonal changes that characterise puberty. 20
Figure 2.1 — Map of ABCD Study sites. 29
Figure 2.2 — ABCD Neuroimaging Protocol. 31
Figure 2.3 — Illustration of the Desikan-Killiany (DK) Atlas brain parcellation 33
Figure 3.1 — Associations between Major depressive disorder (MDD), depressive symptoms (DS) and general measures of cortical and white-matter structures. 52
Figure 3.2 — P-value plots for associations between depressive symptoms (DS) and measures for regional brain regions. 54
Figure 3.3 — P-value plots for associations between Major depressive disorder (MDD) and measures for single brain regions. 55
Figure 3.4 — Associations between socio-environmental factors and absolute discrepancies of caregiver and child reports on depressive symptoms (DS). 57
Figure 4.1 — Pubertal Development Scale quality checking decision tree at Year 1 82
Figure 4.2 — Effect of pubertal timing on depressive symptoms without considering mediation 87
Figure 4.3— Effect of pubertal timing on depressive symptoms including mediation of brain structure. 88
Figure 4.4 — Pilot results: Cortical and subcortical regions of interest for the pubertal timing ~ brain structure association. 95
Figure 4.5 — Pilot results: Cortical and subcortical regions of interest for the depressive symptoms ~ brain structure association. 97
Figure 4.6 —Main results: Frequencies (N) for parent summary scores from the Pubertal Development Scale (PDS). 100
Figure 4.7— Main results: Incidence Rate Ratios (IRRs) for the association between pubertal timing and youth depressive symptoms. 102
Figure 4.8 — Exploratory results: Significant cortical associations with earlier pubertal timing in female youth. 106

Figure 4.9 — Exploratory results: Mediation paths and statistics for main effect of pubertal timing and depressive symptoms, mediated through accumbens area volume. 108
Figure 4.10 — Exploratory results: Mediation paths and statistics for main effect of parental depression and youth depressive symptoms, mediated through pubertal timing. 109
Figure 6.1 — HALFpipe Workflow. 148
Figure 6.2 — An example of the HALFpipe interactive web-based quality assessment (QA) tool for an individual participant. 153
Figure 6.3 — Illustrative description of Leading Eigenvector Dynamics Analysis (LEiDA). 157
Figure 7.1 — Significance of between-condition differences in phase-locking (PL) state probability and duration as a function of k. 173
Figure 7.2 — Boxplots showing the differences in phase-locking (PL) state probability between the rest and irritability conditions for all clustering configurations k = 2 to 20. 174
Figure 7.3 — Boxplots showing the differences in phase- locking (PL) state dwell time (duration) between the rest and irritability conditions for all clustering configurations k = 2 to 20. 175
Figure 7.4 — Phase-locking (PL) brain states for a clustering configuration of k = 11. 177
Figure 7.5 — Repertoire of recurrent phase-locking (PL) states obtained with a clutering configuration of k = 11. 178
Figure 7.6 — Overlap of the 10 non-global phase-locking (PL) states with seven canonical resting-state networks (RSNs) as per Yeo et al. (2011). 179
Figure 7.7 — Phase-locking (PL) state 2 for K = 11, which is characterised by regions of the default mode network (DMN). 181
Figure 7.8 — Phase-locking (PL) state 10 for K = 11, which is characterised by regions of the fronto-parietal network (FPN). 182
Figure 7.9 —Dynamic brain states associated with depressive symptoms and irritable mood. 185

List of Tables

Table 3.1 — Sample sizes and demographic features for MDD and depressive symptoms (DS). 51
Table 4.1 — Hypotheses tested in this registered report. 77
Table 4.2 — Name and description of study variables used in this registered report. 80
Table 4.3 — Model specifications for pubertal timing and depressive symptoms association. 87
Table 4.4 — Pilot analyses: Model specifications for brain structural measures. 93
Table 4.5 — Pilot results: Pubertal timing-brain structure models and associated statistics with significant ROI associations. 96
Table 4.6 — Pilot results: Depressive symptoms ~ brain structure models and associated statistics with significant ROI associations. 98
Table 4.7 — Main results: Descriptive statistics for sample. 99
Table 6.1 — The 18 most highly rated scenarios that were used as stimuli in our irritability task. ... 145
Table 7.1 —Irritability pilot study: Sample descriptive statistics. 171

Introduction to Thesis

Where can it be found again,

An elsewhere world, beyond

Maps and atlases,

Where all is woven into

And of itself, like a nest

Of crosshatched grass blades?

- Seamus Heaney (*Human Chain*, 2010)

The burden of depression falls heavily on youth. Adolescence is a peak time for the emergence of depressive disorders and the rates of adolescent depression are rising. However, *why* depression is more likely to emerge during this period and *who* it is most likely to affect, is not well understood. Adolescence is also a time of immense biological, psychological, and social change — it is an "elsewhere world" that has only recently been appreciated as a unique developmental period that bridges childhood and adulthood. The dynamic nature of our biology, psychology, and social world during this phase of life render it both a challenging and opportunistic area of research. On one hand, unravelling the complex interplay of change "where all is woven into and of itself/ like a nest of crosshatched grass blades" is no easy task. The changes during adolescence occur on multiple levels (e.g., neuronal, hormonal, behavioural) and in a variety of settings (home, school, peer, and parent relationships). Conversely, the dynamic nature of this period also creates myriad opportunities to develop interventions that target multiple levels of change so that we can create an optimal world in which adolescents can thrive.

The advent of neuroimaging techniques has provided unparalleled insight into the structural and functional development of the adolescent brain — the cornerstone of the adolescent (and human) experience. With 100 billion neurons and 100 trillion synapses, the human brain is arguably the most complex biological system in existence. Understanding how the structure of the brain — neurons, synapses, and associated cells and molecules — translates to brain function and in turn, behaviour, and how this is influenced by biological and environmental factors, is the overarching mission of the field of cognitive neuroscience.

To date, the field has developed "maps and atlases" of the human brain as a first step towards bettering our understanding of brain structure and function. However, to understand how developmental outcomes, especially maladaptive ones like depression, emerge from these maps and atlases we must move beyond them. To do so, a holistic framework that studies brain development within the transactional interplay of concurrent biological, psychological, and social change is needed. Considering this, and the heightened vulnerability to depression during adolescence, the overarching aim of this thesis was to contribute to this collective aim

by investigating biological and psychosocial factors associated with depression during adolescence.

Specifically, the first two aims of this thesis were to use a large developmental cohort study to examine how brain structure is associated with the emergence of depression during adolescence, and how this relates to other biological factors such as pubertal development. Using these data, I first assessed how cortical measures and white matter microstructure relate to depression in early adolescence using reports from both caregivers and youth themselves. I then examined how earlier pubertal timing is associated with later depression risk, and whether certain aspects of brain morphometry mediate this relationship. The third and fourth aims of this thesis relate to examining dynamic functional brain networks associated with depression during adolescence, and how this is associated with psychological factors like irritability. In this pilot study, I first validated a novel fMRI task targeting irritability that was co-produced with young people. I then examined how dynamic brain states relate to behavioural measures, such as depressive symptoms and irritable mood.

In Chapter 1, I introduce adolescence and the biological, psychological, and social changes that characterise this developmental period. I then discuss how these changes are interwoven and contribute to the increased risk for the onset of depression during adolescence. Further, I discuss the motivation for studying brain structure, function, and pubertal development as ways to better our understanding of adolescent depression and identify tractable targets for intervention. I also provide an overview of the constructs of interest in the current thesis. In Chapter 2, I describe the cohort study used in Chapters 3 & 4 of this thesis: The Adolescent Brain Cognitive Development (ABCD) Study®. In Chapter 3, I investigate brain structures associated with depression in early adolescence. In Chapter 4, I examine how earlier pubertal timing relates to later depression during adolescence, and whether brain structure mediates this association. In Chapter 5, I provide a narrative literature review of the neural correlates of irritability in adolescence and highlight the need for neuroimaging study designs that reflect the social nature of adolescence. I then discuss the motivation for adopting a co-produced youth-research design for our pilot fMRI study. In Chapter 6, I describe the characteristics of our pilot study and the methodology employed. In Chapter 7, I use a data-

driven approach to explore the dynamic functional brain networks associated with our novel fMRI task, and whether characteristics of these brain states relate to behavioural measures. Finally, in Chapter 8, I discuss the main findings of this thesis considering its limitations and highlight directions for future work.

1 General Introduction

1.1 Outline

A substantive body of research has shown that a series of developmental processes contribute to the heightened vulnerability to the onset and maintenance of depression during adolescence. Specifically, these include genetic risk, neuromaturation, hormonal changes, and social development, which interact with environmental factors to confer varying degrees of risk. In this chapter, I first describe the period of adolescence before outlining the epidemiology of depression during this developmental phase and its associated risk factors. I then discuss typical structural and functional brain development, and the neuroimaging paradigms used to examine these domains. I also provide an overview of pubertal development and how this relates to neuromaturation. Finally, I discuss the evidence that links deviations from normative development in these domains to depression during adolescence, which provides the rationale for the current thesis. I end this chapter by outlining the main aims of this doctoral work.

1.2 The dynamic world of adolescence

Adolescence, a life phase spanning the ages 10-24, is the developmental period that bridges childhood and adulthood (Sawyer et al., 2018). It is characterised by immense biological growth and significant social role transitions that allow youth to move away from the security of childhood and begin to forge their own identities as adults. In many ways, this state of flux mirrors the foundational growth, learning, and neuromaturation that occurs in the first few years of life (Shonkoff et al., 2012). The importance of early life experiences in shaping later developmental outcomes is widely recognised and has shaped global policy and practices (Black & Hurley, 2014). However, it is only in recent years that adolescence has been recognised as a second “sensitive period” (Dahl et al., 2018). Like early childhood, the dynamic nature of adolescence means that a young person’s life can quickly pivot in both positive and negative directions. Thus, this critical period of development is a time when interventions and policy changes could have a potent effect and allow young people to put their best foot

forward as they enter the world of adulthood. A significant barrier to young people having an active and meaningful role in society during adolescence is mental health disorders, such as depression.

1.3 Depression in adolescence

Globally, depression is a leading cause of illness and disability and is associated with significant personal, societal and economic costs (Global Burden of Diseases, 2018). Depression in adolescence is of particular concern due to its recurrent disease course and association with an increased risk for comorbid physical and mental health conditions, as well as concurrent and later psychosocial difficulties (Malhi & Mann, 2018). Depression can be defined as a variety of mood related symptoms and behaviours that exist along a spectrum (see Figure 1.1; Thapar et al., 2022). At one end of this continuum, we have symptoms that are a normative response to life events — for example, feeling sad or having concentration difficulties is a normal reaction to relationship difficulties or academic stress. These reactions are often adaptive and can even enable effective coping. However, as we move further along this spectrum, we encounter mood and behavioural states that fall outside normative fluctuations, a transition to which is associated with functional impairment and degree of coping (Foulkes, 2022; Thapar et al., 2022).

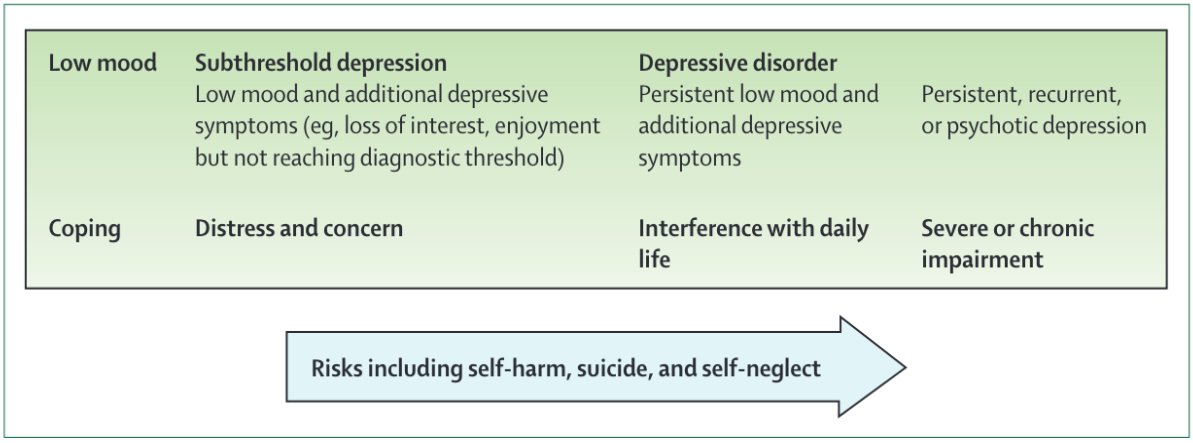


Figure 1.1 – The spectrum of depression. Figure taken from Thapar et al., 2022 (Copyright: Elsevier Ltd).

Depressive symptoms that do not meet full criteria for major depressive disorder (MDD) are referred to as sub-threshold depression/depressive symptoms, which can negatively impact

quality of life and are a risk factor for a later depressive disorder (Bertha & Balázs, 2013). Depressive disorders fall at the other end of the spectrum and are characterised by mood and behaviours that are longer lasting and significantly impair daily functioning. Specifically, accordingly to the Diagnostic and Statistical Manual of Mental Disorders (Fifth Edition) (DSM-5), a diagnosis of MDD is defined by depressed mood (or irritable mood in adolescent MDD) that is present nearly every day for most of the day, *or* a marked diminished interest or enjoyment (termed anhedonia) in all, or almost all activities, every day for most of the day, for a period of at least 2 weeks. A range of other symptoms accompany these core symptoms (see Figure 1.2). To meet diagnostic criteria for MDD, five of these symptoms, including at least one cardinal symptom, must be present and interfere with daily life functioning (American Psychiatric Association, 2013b). While these diagnostic criteria can be helpful in both clinical and research settings, it is evident that depression is a highly heterogenous condition (Fried, 2015; Fried & Nesse, 2015).

Core clinical symptoms of major depressive disorder

- Low mood, or irritable mood in adolescents
- Diminished interest or pleasure in all, or almost all, activities
- Significant weight loss or gain, or decrease or increase of appetite
- Insomnia or hypersomnia
- Psychomotor agitation or retardation
- Loss of energy or fatigue
- Excessive feelings of guilt or worthlessness
- Diminished ability to think or concentrate, or indecisiveness
- Recurrent thought of death, suicidal ideas, or suicide attempt

Figure 1.2 – Core clinical symptoms of depression (according to the DSM-5).

There is a growing appreciation that depression is not a single disorder which has resulted in the emergence of research on depression subtypes defined by features such as: primary symptoms (e.g., low mood alongside sleep difficulties), age of onset (e.g., adolescence versus adulthood), nature of onset (e.g., after a stressful life event), whether it is a single episode or recurrent/chronic, and treatment response (Harald & Gordon, 2012). While promising,

subtyping depression has not yet revealed differential causes and treatment responses. This could be due to several factors including overlapping subtypes — adolescent onset depression is more likely to be recurrent, for example — as well as the current tools used to measure depression. Existing depression measures have come under increasing scrutiny with some researchers, such as Eiko Fried and colleagues, arguing for a radical overhaul of the theoretical and methodological foundations of depression measurement so that they reflect the developments made in our understanding of depression over recent decades (e.g., the importance of depression subtypes) (Fried et al., 2022). Although it is important to acknowledge the heterogeneity of depression, and the challenges that are therefore inherent to researching depression, the focus of depression within this thesis is not on subtyping depression nor examining depressive disorder specifically. Instead, I adopt a dimensional approach to studying depression that emphasises symptom severity ranging from low-mild, moderate through to severe, a decision that was shaped by the sample characteristics of both the cohort study and locally collected sample used in this thesis.

The data used in this thesis comprises youth in the early to mid-stages of adolescence (ages 9 – 13 years; the ABCD Study sample) as well as the later stages of this developmental period (ages 16-20 years; our locally collected sample). A recent large-scale meta-analysis of epidemiological studies found that the peak age of onset for depressive disorders is 20.5 years (Solmi et al., 2022). Given that ABCD is a population-based rather than a clinical sample, the proportion of young people with a diagnosis of MDD is likely to be relatively low. Our local study was also a community-based sample that aimed to recruit youth with a range of depressive symptoms. Adopting a dimensional approach to the study of adolescent depression allows us to capture greater variation in depressive symptoms — an important consideration for this age range given that it is a period when the early signs of emotional distress are likely to emerge. This can then lay a strong foundation to chart the trajectories of depressive symptoms: whether they remit or instead recur and transition from sub-threshold depression to a depressive disorder. Importantly, the dynamic nature of adolescence, both in terms of biological and social factors, means that it is an ideal window of opportunity in which to intervene before depression becomes chronic and young people become embedded in a life trajectory with an increased risk for poorer developmental outcomes.

1.4 Risk factors for depression — a brief overview

Although the aetiology of depression remains unclear, there is a relatively good understanding of the multiple factors that affect an individual's risk for depression. Crucially, there is no single factor that determines whether a person will develop depression. Rather, it is likely that myriad risk factors, each with their own probabilistic risk effect, interact with each other over time to determine the degree of depression risk (Boyce et al., 2020). Such risk factors operate at the individual-level, the family- and peer-level, and at the population level.

1.4.1 Individual-level factors

Evidence from twin-studies suggests that around 40% of the variance in depression risk is accounted for by genetic factors (McIntosh et al., 2019). Recent genome wide association studies in adults have identified over a hundred genetic variants, each with a small effect size, that contribute to risk for depression. These genes were associated with synaptic structure and neurotransmission, especially in prefrontal brain regions (Howard et al., 2019). Genetic risk is also important in adolescent depression, especially in terms of symptom severity and rate of change (Jami et al., 2022; Kwong et al., 2021). However, a complex gene-environment interaction is also at play here because genetic vulnerability to depression is correlated with exposure to environmental stressors (Rutter, 2010). This means that individuals with a higher genetic risk are more likely to be exposed to social stressors, which thus creates additional depression risk. Other aspects of our biology, such as pubertal timing, can place an individual at an increased risk for depression. Specifically, youth that begin puberty ahead of their peers are more likely to experience depression in adolescence, compared to those that begin puberty around the same time as their peers (Ullsperger & Nikolas, 2017). This risk effect is likely due to a combination of biological and psycho-social factors and their interaction with each other (Pfeifer & Allen, 2021).

There are many other individual-level risk factors associated with depression some of which include thinking styles and behavioural traits such as neuroticism (Hakulinen et al., 2015), low positive emotionality (Khazanov & Ruscio, 2016), and rumination (Cano-López et al., 2022).

Unsurprisingly, these modifiable factors have been the focus of psychological interventions for depression. Moreover, a number of comorbid mental and physical health difficulties may increase risk for depression such as: a history of anxiety (Rice et al., 2017) and irritability (Vidal-Ribas & Stringaris, 2021) in childhood, a diagnosis of a neurodevelopmental disorder (Hollocks et al., 2019; Meinzer et al., 2014) (e.g., attention deficit hyperactivity disorder (ADHD) and autism), a history of a chronic physical illness involving the central nervous system (Pinquart & Shen, 2011) (e.g., migraine or epilepsy), obesity (Rao et al., 2020) and sleep disruption (Marino et al., 2021). Like genetic factors, the variance in depression risk explained by each of these risk factors is small, and some of these associations (e.g., sleep disturbance and obesity) may be bidirectional or explained by confounding factors (Rao et al., 2020).

1.4.2 Family-level and peer-level risk factors

Having a family history of depression is one of the most common risk factors for depression. Indeed, approximately 40% of individuals who have a parent with depression will develop depression themselves, and this risk is greatest for those with a history of multi-generational, chronic, and early-onset depression (Maciejewski et al., 2018). The transmission of depression across generations is likely due to a combination of genetic and environmental mechanisms, such as offspring being exposed to current parental depression (McAdams et al., 2015). Importantly, there are a number of resilience-promoting factors that can mitigate this risk many of which pivot around strong social connectedness — high quality relationships with other family members and friends, and participation in school and sporting activities (Collishaw et al., 2016; Stein et al., 2014). On the other hand, negative social experiences have been associated with later adolescent depression. These social stressors often relate to early life adversities (LeMoult et al., 2020; Norman et al., 2012), such as neglect, abuse, stressful life events (e.g., death of loved one or experiencing a serious illness), bullying (Moore et al., 2017), and social isolation (Achterbergh et al., 2020). Although there has been much research recently on the association between adolescent social media use and depression, findings are currently inconclusive suggesting that although there may be some benefits (e.g., increased perceived social support), they are accompanied by risks, such as damaging social comparison, cyberbullying, and addiction (Ivie et al., 2020).

1.4.3 Population-level risk factors

A common thread linking the social stressors mentioned above is poverty and social deprivation, which are undeniably the most widely studied community-level stressors associated with depression (Stirling et al., 2015). Individuals who grow up in poverty or in neighbourhoods with a high crime rate, are homeless, or are a refugee or displaced due to war, are at a much higher risk for developing psychiatric disorders, including depression (Kessler et al., 2010). A recent longitudinal randomised control trial by Sheridan et al. (2022) that used data from the Bucharest Early Intervention Project, demonstrated the causal impact of early deprivation (in this case, institutional care in early childhood) on cortical brain development across middle childhood and adolescence, particularly in prefrontal regions and in white matter tracts connecting prefrontal and parietal regions (Sheridan et al., 2022). These findings provide a possible neurobiological explanation for the enduring impact of exposure to adversity early in life on multiple developmental outcomes, such as increased risk for psychopathology.

Although a different line of research, it is important to note the almost global experience of the Covid-19 pandemic in 2020-21, during which the prevalence of depression and anxiety in young people doubled (especially in older adolescents and females), compared with pre-pandemic estimates (Racine et al., 2021). These stressors were found to disproportionally affect minority ethnic/racial and gender/sexuality groups, which in part may be related to racism and peer victimisation, and the increased social deprivation and stress that often accompanies such prejudices (Amos et al., 2020). Moving forward, it will be interesting to examine the impact of a stressful life event like a global pandemic on the developmental trajectories of children and young people today. Although there are several challenges in addressing such a research question (e.g., varying Covid-19 restrictions within countries, disruption of data collection, and the differential impact of Covid restrictions on family life), cohort studies like ABCD continued to collect data during and after the pandemic, which will allow the prospective longitudinal investigation of a natural experiment like Covid-19.

1.4.4 Linking risk factors to biology

Together, these separate findings highlight that many risk factors operate cumulatively to exert deleterious effects on adolescent mental health outcomes. However, despite this evidence, most research has tended to examine these risk factors independently. It has been suggested that there is great promise in a multi-level integrative approach that combines epidemiological and aetiological research to identify modifiable risk factors that can prevent the onset of youth mental health problems. These can then be used as treatment targets for youth already experiencing difficulties as well as in prevention and intervention efforts (Allen & Dahl, 2015; Pfeifer & Allen, 2021). To this end, research should aim to move beyond describing these risk factors and instead attempt to unravel the developmental and neurobiological mechanisms that may transmit the effects of these myriad risk factors to depression during adolescence.

One next step for research is to thus investigate how features of the developing adolescent brain relate to the emergence of depression and understand how risk factors such as early pubertal timing and irritable mood may contribute to depressive problems during this period. Research from large-scale neuroimaging studies have demonstrated robust brain structural alterations in adult depression. However, the temporal origins of these morphometrical differences earlier in development remain unclear. Adolescence has a biological beginning with the onset of puberty, an event that is infused with significance for mental health risk. The increased vulnerability to internalising difficulties from puberty onwards, especially for females, paired with the substantive body of evidence linking earlier pubertal timing and increased risk for depression, highlight the potential prominent role of puberty in advancing our understanding of the aetiology of adolescent depression. Nonetheless, the role of brain structure in understanding the association between earlier pubertal timing and increased risk for depression is not well understood. In addition to biological factors, irritable mood is a hallmark of adolescent depression and an early sign of emotion regulation difficulties. However, existing research on the neural correlates of irritability typically overlook the social nature of adolescence. Considering this, the focus of this thesis will be on examining how the adolescent brain, and its association with pubertal timing and irritability, relate to depression risk during this key developmental period.

1.5 The developing brain

Animal studies provided the first evidence that early in development, the brain undergoes a period of synaptic proliferation, whereby there is a deluge of synapse formation such that the synaptic density (number of synapses per unit volume of brain tissue) is markedly higher than the adult brain (Lund et al., 1977; Rakic et al., 1986). We see a similar pattern in humans — brain size increases four-fold between birth and preschool age, and is approximately 90% of adult brain volume by the age of six years (Stiles & Jernigan, 2010). The brain then undergoes a protracted period of synaptic pruning during childhood and adolescence (Huttenlocher & Dabholkar, 1997), which relates to changes in grey matter (brain tissue containing the neuronal cell bodies), white matter (brain tissue comprising myelinated nerve fibres), as well as functional reorganisation (Stiles & Jernigan, 2010). Importantly, these changes relate to developmental milestones in behaviour, such as the development of motor skills (Hadders-Algra, 2018) and higher-order cognitive functions, like theory of mind (Richardson et al., 2018).

The emergence of neuroimaging methods, such as magnetic resonance imaging (MRI), has equipped researchers with an invaluable tool to advance our understanding of structural and functional brain development. MRI allows the *in vivo* quantification of myriad brain properties in a non-invasive manner (Lerch et al., 2017). Here, I provide an overview of neuroimaging methods used to quantify brain structure and function and discuss the associated literature on typical structural and functional brain development during adolescence. This will help situate our understanding of how deviations from typical development can relate to depression during adolescence.

1.5.1 Grey matter development during adolescence

Structural MRI can be used to measure aspects of brain morphometry, including cortical and subcortical volume, cortical thickness, surface area, and sulcal depth. Methods used to measure cortical volume vary across software tools and are usually voxel- (e.g., FSL; Jenkinson et al., 2012) or surface-based (e.g., FreeSurfer; Dale et al., 1999; Fischl et al., 2002) approaches. The former involves counting the number of voxels in the brain (or in a particular

brain structure). On the other hand, the surface-based approach, which is used in the ABCD Study, calculates the volume inside the pial surface (the border between cortical grey matter and cerebrospinal fluid) minus the volume inside the white surface (the boundary between grey matter and white matter) minus the tissue volume (voxel count) of subcortical areas. Cortical thickness and surface area are the substrates of cortical volume. Cortical thickness is measured by calculating the vertex-wise closest distance between the white and pial surface (See Figure 1.3; Fischl & Dale, 2000). The measurement of cortical surface area is based either on the white surface or pial surface, which is mapped onto a template brain (Chen et al., 2012). The amount of expansion or contraction needed for an individual's brain to map successfully onto the template brain is used as a measure of the vertex-wise cortical surface area. Finally, sulcal depth, an indicator of the shape of the cortical surface, is defined as the distance between the central cortical surface and its convex hull relative to a mid-surface that crosses the cortical surface (Fischl et al., 1999; Yun et al., 2013).

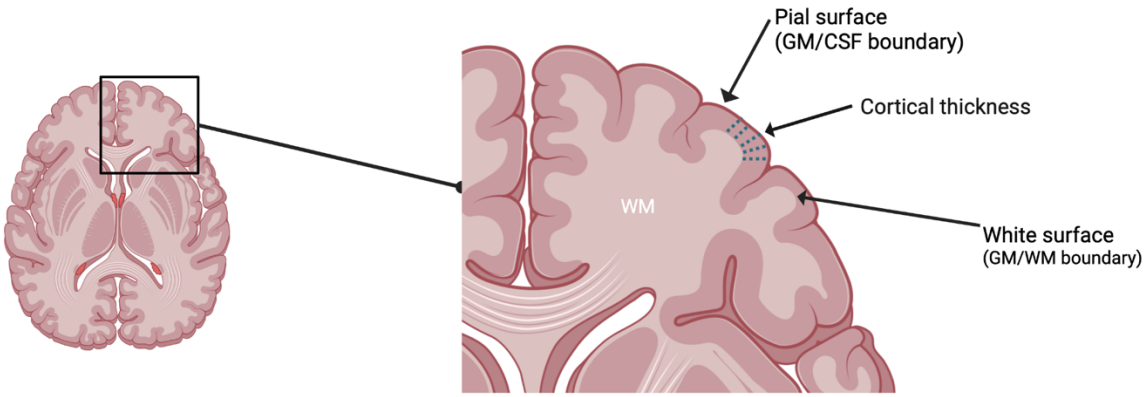


Figure 1.3 — Illustration of cortical thickness measurement. Cortical thickness is the vertex-wise closest distance between the white and pial surface of the brain. GM = grey matter; WM = white matter; CSF = cerebrospinal fluid. Figure created with BioRender.com.

Longitudinal developmental neuroimaging studies have made a remarkable contribution to our understanding of grey matter development across childhood and adolescence. Together, this work has shown that grey matter changes are non-linear and vary across brain regions (for recent review see Norbom et al., 2021). Early longitudinal work from the National Institute of Mental Health in the US reported that cortical volume development in the first two decades of life followed an inverted U-shaped trajectory, with frontal and parietal regions

reaching peak volume in early adolescence (around age 12 years) while the volume of temporal regions peaked later in mid-adolescence (Giedd et al., 1999; Gogtay et al., 2004; Lenroot et al., 2007; Raznahan et al., 2011). However, more recent work in independent longitudinal samples (aged 8 to 30 years (N = 391, scans = 852) and 7 to 23 years (N = 135, scans = 202) found that cortical volume decreases monotonously in a non-linear manner across childhood and adolescence. These results suggest that cortical volumetric reductions begin much earlier than previously reported (Mills et al., 2016; Wierenga et al., 2014).

As mentioned, cortical thickness and surface area give rise to cortical volume. However, these morphometric features show a degree of genetic distinction (Winkler et al., 2010) and follow different developmental trajectories (Lyall et al., 2015; Tamnes et al., 2017; Wierenga et al., 2014). Like changes in the field's understanding of cortical volume, cortical thickness is also considered to peak much earlier in development (i.e., early childhood) (Lyall et al., 2015) than previously thought (Raznahan et al., 2011), and then follow a monotonic decreasing trajectory throughout mid-childhood, adolescence, and beyond (Frangou et al., 2022; Tamnes et al., 2017; Vidal-Pineiro et al., 2020; Wierenga et al., 2014). Further, cortical thinning demonstrates spatiotemporal variation across brain regions with association cortices demonstrating a more protracted period of maturation compared to sensory regions (Norbom et al., 2021; Tamnes et al., 2017).

Cortical surface area also exhibits this pattern of development whereby sensory areas expand greatly in the first two years of life (Li et al., 2013), which is followed by the continued expansion of higher-order cortical regions that peak in late-childhood/early adolescence before stabilising by mid-adolescence, and slightly decreasing thereafter (Ducharme et al., 2016; Tamnes et al., 2017; Wierenga et al., 2014). Several neurobiological processes have been put forward to explain the developmental patterns in cortical thinning and surface area seen in MRI studies, including reorganisation of dendritic arbour and increased intracortical myelination (Natu et al., 2019; Patel et al., 2019). Importantly, the observed morphometric development is likely the product of several overlapping neurobiological mechanisms (Norbom et al., 2021). Although the direct study of these cellular processes is difficult due to the small number of post-mortem histological studies in this age range (Hagler et al., 2022),

emerging brain transcriptomic work (Patel et al., 2020) and longitudinal research focused on individual differences in neuromaturation (albeit requiring samples with thousands of individuals) (Marek et al., 2022) may help elucidate the neurobiological processes underpinning the observed MRI findings (Norbom et al., 2021).

Recent collaborative work on brain charts for the human lifespan (Bethlehem et al., 2022) has offered insight into the neurodevelopmental milestones across the lifespan (see Figure 1.4), which will undoubtedly be a valuable resource for future research.

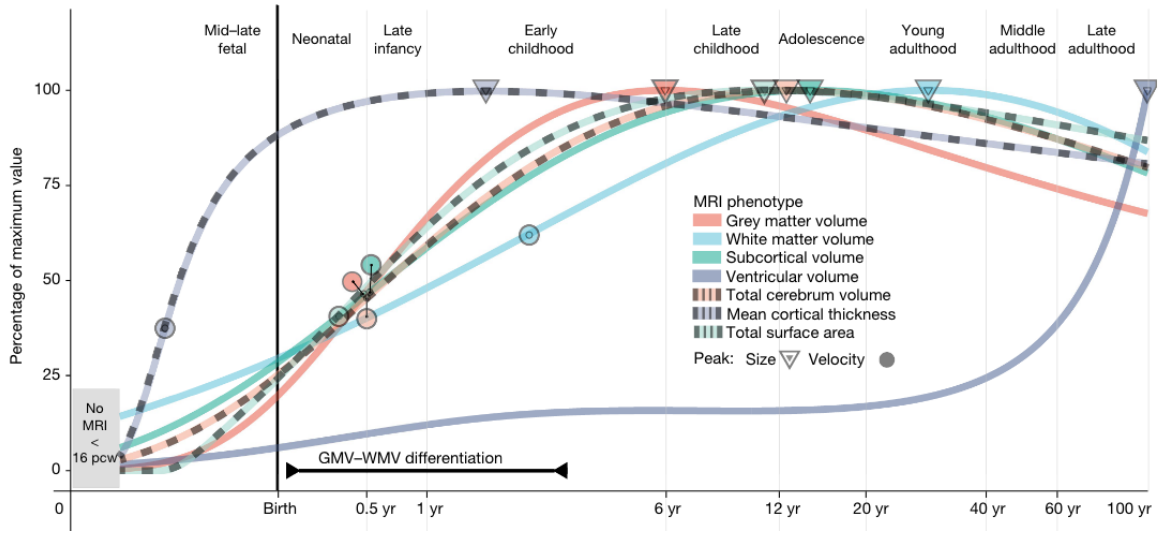


Figure 1.4 — Neurodevelopmental milestones across the lifespan. This figure is a graphical summary of the findings from Bethlehem et al. (2022). The normative trajectories of the median (50th centile) for each global MRI phenotype as a function of age (log-scaled) are shown. Circles depict the peak rate of growth milestones for each phenotype. Triangles depict the peak volume of each phenotype. Figure taken from Bethlehem et al. (2022) and figure caption adapted from the original paper.

The work by Bethlehem and colleagues (2022) suggests that subcortical brain volume demonstrates an overall non-linear increase across childhood and adolescence but has a later volumetric peak in mid-adolescence compared to cortical volume (mid-childhood). However, there are regional and sex associated differences in the developmental patterns of subcortical brain areas. For example, findings from an international collaborative project (Herting et al., 2018) comprising three independent longitudinal datasets, suggest that the volume of subcortical regions related to sensory, motor, and cognitive function (e.g., caudate, thalamus,

putamen, and nucleus accumbens) decreases subtly across this developmental period, while amygdala volume demonstrates a modest increase.

Regarding sex differences, males were found to demonstrate a steeper increase in amygdala volume compared to females, while females showed a decrease in nucleus accumbens and putamen volume but males showed no or little change across this age range (Herting et al., 2018). Although these results are mostly consistent with previous work (Lenroot et al., 2007; Raznahan et al., 2014; Wierenga et al., 2018), some studies have found conflicting results, such as minimal change or decrease in amygdala volume across adolescence (Dennison et al., 2013; Wierenga et al., 2018), and an increase in putamen volume (Wierenga et al., 2018). Further, even within the collaborative study by Herting and colleagues (2018), inconsistencies were found across samples in terms of the pattern of change (i.e., whether a linear, quadratic, or cubic model best fitted the development trajectories of these brain regions), which may be due to differences in population characteristics, sampling strategies and scanning protocols. Together, these findings underscore the need for a “team-science” approach to developmental science (e.g., harmonised study protocols, where possible, and analysis pipelines) (Zanolie et al., 2022) as well as a shift in perspective to focus on the pattern (i.e., stability/rate and direction) of change rather than trying to fit a specific model term (Herting et al., 2018).

1.5.2 White matter development during adolescence

Diffusion tensor imaging (DTI) measures the diffusivity of water molecules within brain tissue, providing insight into the white matter microstructure and structural connectivity of the brain. Unlike grey matter, which has predominantly isotropic water diffusion (i.e., water diffusion occurs equally in all directions (see Figure 1.5a), white matter tracts have anisotropic diffusion, whereby water diffusion occurs along the direction of the fibre (see Figure 1.5b). From this measurement of the restricted diffusion of water molecules, certain scalars can be derived such as mean diffusivity (MD) and fractional anisotropy (FA). MD refers to the average amount of diffusion along the three main diffusion axes ($\frac{\lambda_1 + \lambda_2 + \lambda_3}{3}$). FA is a measure of the degree of diffusion in the principal direction compared to the two orthogonal directions and is therefore a scalar value between 0 (i.e., equal diffusion in all directions) and 1 (i.e., diffusion occurs in one direction only).

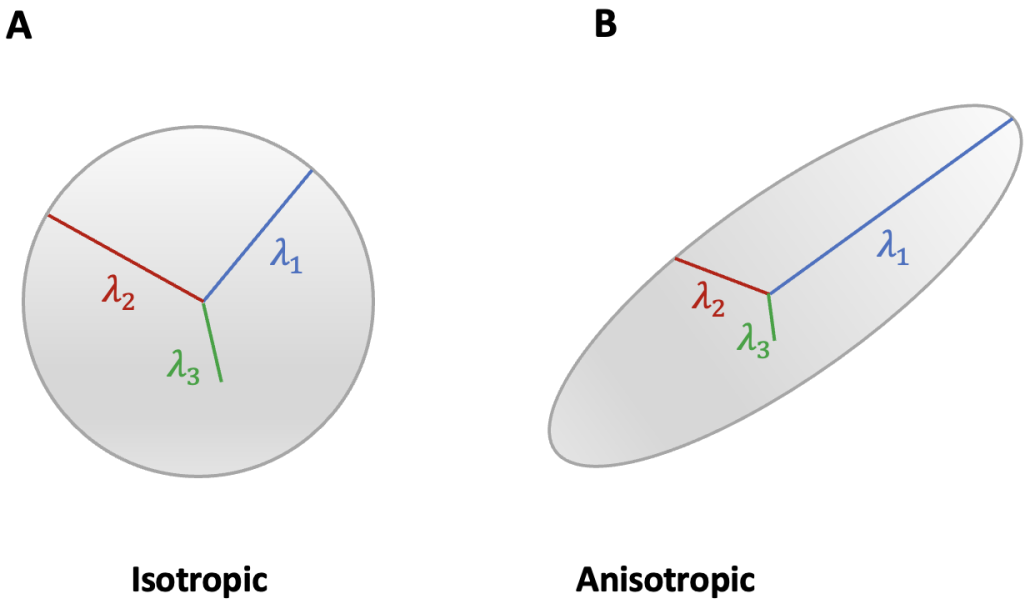


Figure 1.5 — Illustration of isotopic (A) and anisotropic (B) water diffusion.

Up until relatively recently, higher FA values and lower MD values were thought to reflect greater white matter density, with some known exceptions at crossing white matter fibres (e.g., at the junction of the corpus callosum and corona radiata) (Tuch et al., 2003), where this interpretation was recognised as problematic. Importantly, the biophysical underpinnings of

these white diffusion measures are influenced by factors such as fibre diameter, fibre density, myelination and membrane permeability (Beaulieu, 2002). This biophysical complexity has gained more attention in recent years and highlighted the possible oversimplified interpretation of what we can infer about the "integrity" of white matter microstructure from these DTI measures (Tamnes, Roalf, et al., 2018), with some researchers arguing against the use of this term entirely (Jones, et al., 2013). While more advanced diffusion MRI methods, such as neurite orientation dispersion and density imaging (NODDI; Zhang et al., 2012), were beyond the scope of the current thesis, these techniques can provide more fine-grained detail about white matter microstructural properties and will be important for research going forward.

Although the precise neurobiological mechanisms that give rise to DTI metrics are not fully understood, a substantive body of research has shown widespread increases in white matter volume during childhood and adolescence, suggesting increased myelination and axonal packing (for a recent review see Lebel & Deoni, 2018). This white matter microstructural maturation is thought to serve the development and refinement of higher-order cognitive processes (e.g., inhibitory control) (Simmonds et al., 2014). However, there is a paucity of longitudinal research and findings from cross-sectional studies have been inconsistent (Goddings et al., 2021). Nonetheless, there is evidence to support the hypothesis that FA increases and MD decreases across childhood and adolescence, before plateauing by young adulthood. Like grey matter development, these microstructural changes are thought to be non-linear (Lebel et al., 2019; Lebel & Deoni, 2018; Pohl et al., 2016).

1.5.3 Functional brain development during adolescence

Functional MRI (fMRI) quantifies hemodynamic changes in the brain via a blood-oxygen-level-dependent (BOLD) signal, which is used as a proxy measure of brain activity with temporal and spatial resolution in seconds and millimetres, respectively. fMRI is used to examine fluctuations in brain activity during specific tasks (e.g., working memory or reward processing tasks) or while the brain is at rest, which is known as resting-state fMRI. Within the domain of resting-state fMRI, which is the perspective taken to study brain function in the current thesis, functional connectivity (FC) is the most common property of the brain analysed.

Resting-state FC refers to patterns of co-activation in different brain regions at rest and is a measure of the correlation of activity in distinct brain areas over time (Friston, 2011).

From this work, parcellations of the brain as a dynamic system of functionally distinct but complementary networks have emerged (e.g., Yeo et al., 2011). Due to their role in typical and atypical development, some of the most widely studied brain functional networks include: 1) the default mode network (DMN; which includes the ventromedial prefrontal cortex, posterior cingulate cortex, and precuneus) and plays a key role in self-directed thought, including introspection and autobiographical memory (Andrews-Hanna et al., 2014); 2) the fronto-parietal network (FPN; which consists of the dorsolateral prefrontal cortex and posterior parietal cortex) and supports goal-directed behaviour, such as cognitive control and decision making (Zanto & Gazzaley, 2013); and 3) the salience network (SN; comprising the insula and dorsal anterior cingulate cortex) and is involved in attending to salient stimuli in one's environment and supporting the response to such stimuli by relaying information between the DMN and FPN (Corbetta et al., 2008).

Historically, brain function has been studied from a static perspective, in which an average FC measure is calculated across the entire time-series of the resting-state scan. While this approach has greatly developed the field's understanding of the functional connectome (i.e., the brain's collective set of functional connections) (Biswal et al., 2010; Yeo et al., 2011), static FC methods do not capture the instantaneous waxing and waning of brain network activity over time (Iraji et al., 2021). The past decade has thus seen an emphasis on developing methods that capture the inherent dynamic nature of functional brain networks, termed dynamic FC (dFC; Cabral, Kringelbach, et al., 2017; Calhoun et al., 2014). Although many different methods comprise this line of research, such as the sliding window approach (Allen et al., 2014; Handwerker et al., 2012), co-activation pattern analysis (Karahanoğlu & Van De Ville, 2015; Liu et al., 2013; Tagliazucchi et al., 2012), and phase-coherence pattern analysis (Cabral, Vidaurre, et al., 2017; Glerean et al., 2012; Hellyer et al., 2015), these studies have collectively shown that brain activity involves time-varying, reoccurring, configurations of the coupling and uncoupling of brain regions. These spatiotemporal patterns have revealed important information that can assist our understanding of the processes underlying typical

and atypical behaviour (Cabral, Vidaurre, et al., 2017; Iraji et al., 2021; Sakoğlu et al., 2010; Zalesky et al., 2014).

Indeed, studying the brain through a network or systems lens is important from a developmental perspective given that disruptions to widespread brain connections are strongly associated with psychopathology (Vanes & Nosarti, 2022; Vértes & Bullmore, 2015). Although there is a paucity of longitudinal research on typical functional development (Ernst et al., 2015), relative to how this is disrupted in atypical development (e.g., mental health disorders), the extant evidence suggests that brain function becomes more integrated and efficient across development (Bassett & Sporns, 2017; Ernst et al., 2015; Kundu et al., 2018). There seems to be a general shift from "local" (i.e., anatomically proximal) connections, which dominate during childhood and early adolescence, to a more "distributed" functional architecture from young adulthood onwards, whereby distal connections strengthen (Edde et al., 2021; Fair et al., 2009). This refinement of the brain's functional architecture is thought to support the development of cognitive processes, such as emotion regulation and inhibitory control (Ernst et al., 2015). These long-range connections comprise many of the known resting state networks (e.g., DMN, FPN, SN) and research suggests that regions within these networks become more connected across development. For example, a large cross-sectional study by Truelove-Hill et al. (2020) found an increase in the connectivity within several networks across adolescence such as the DMN, FPN, and SN (Truelove-Hill et al., 2020). Longitudinal work also provides evidence that connectivity between functionally-related brain regions strengthens during development, such as the subcortico-subcortical connections (e.g., between the hippocampus, amygdala, nucleus accumbens and putamen) and cortico-cortical connections (e.g., ventral anterior cingulate, dorsal anterior cingulate, frontal medial, and subcallosal) (van Duijvenvoorde et al., 2019).

While the existing research broadly supports the idea that within-network connections strengthen across development, our understanding of between-network connectivity is less clear. Some research suggests that functional networks become more segregated over time, such that networks involved in affective-motivational processes (e.g., fronto-limbic connections) function in an increasingly independent manner (Fareri et al., 2015; van

Duijvenvoorde et al., 2016, 2019). However, work by Marek et al. (2015) suggests there is an increase in between-network connectivity across adolescence (as well as a concurrent decrease in within-network connectivity) (Marek et al., 2015).

Taken together, these findings highlight the complex nature of structural and functional brain development during adolescence and underscore the need for further work in this area. The multi-modal nature of the methods employed in the current thesis make a direct and timely contribution to this body of research. Before discussing how brain structure and function relate to depression during adolescence, our attention now turns to the key biological event, or rather the *series* of events, that propel a young person away from childhood into the flux of adolescence — puberty.

1.6 Puberty: The biological catalyst of adolescence

Although puberty can be regarded as a biological event, it is infused with personal and social significance. The surge in hormones that characterise the beginning of adolescence play a central role in a series broader biological, psychological, and social changes that prepare a young person for reproductive maturity (Crone & Dahl, 2012). Alongside the physical changes typically associated with puberty (see Figure 1.6a), this period of development is also characterised by changes to motivation and desires, changes in sleep patterns and circadian rhythm, as well as myriad other social, behavioural and emotional changes, such as influential peer relationships and romantic relationships (Andrews et al., 2021; Crone & Dahl, 2012).

From a biological perspective, as illustrated in Figure 1.6b, pubertal development consists of two phases: adrenarche and gonadarche, which are triggered by the hypothalamic-pituitary-adrenal (HPA) and hypothalamic-pituitary-gonadal (HPG) axes, respectively (for reviews see Abreu & Kaiser, 2016 and Crone & Dahl, 2012). Adrenarche represents the earliest signs of pubertal development, usually occurring between the ages 6-10 years (earlier for females) and is characterised by the increased secretion of the androgen, dehydroepiandrosterone (DHEA) and its sulphate (DHEAS), from the zona reticularis of the adrenal gland (Biro et al., 2014). DHEA levels continue to increase until the early 20s and are responsible for the

development of a number of secondary sex characteristics including, pubic hair growth and changes in body odour and skin features (e.g., acne) (Havelock et al., 2004).

The second phase of pubertal development, gonadarche, also occurs earlier in females, typically between the ages of 9-14 years, while the onset for males is usually between 10-15 years of age (McAnarney, 1992). Gonadarche involves the production of sex steroid hormones (gonads) such as oestrogen and testosterone, via the pulsatile release of gonadotrophin-releasing hormone (GnRH) from the hypothalamus (during sleep), which then stimulates the release of follicle stimulating and lutenising hormones (FSH and LH) from the pituitary gland. The HPG axis is first active in the prenatal and early postnatal life but is then made dormant by inhibitory inputs from the hypothalamus. Although the precise mechanisms that re-awaken the HPG axis are not fully understood, it is thought that it arises through interactions with neural systems implicated in metabolism, energy storage, and sleep regulation. Important agents identified include the hormone leptin and kisspeptins, a family of neuropeptides (Abreu & Kaiser, 2016). Testosterone and oestrogen enable reproductive maturity and are responsible for the development of additional sex characteristics, such as testicular development and voice deepening in males, and breast development and menstruation in females. A third neuroendocrine axis that is part of pubertal maturation is the release of growth hormone (GH) from the pituitary gland, which is responsible for the rapid physical growth that characterises adolescence.

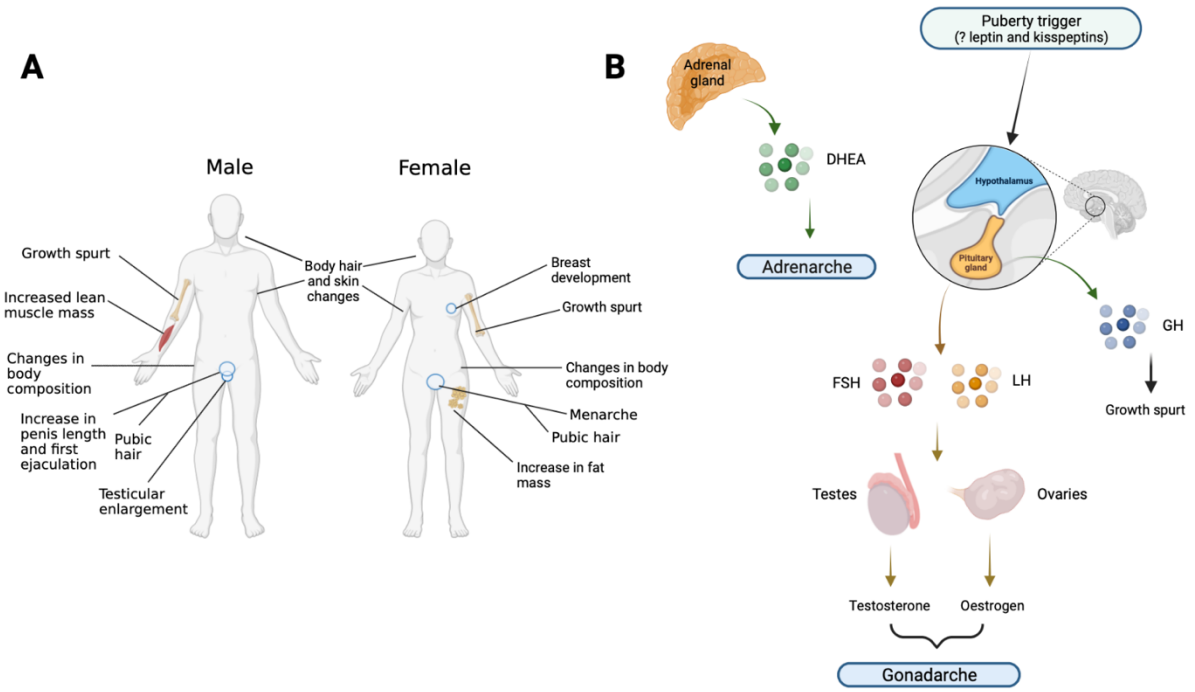


Figure 1.6 — The physical and hormonal changes that characterise puberty. A) Physical changes associated with puberty for males and females; B) Hormonal changes that occur during puberty. DHEA = Dehydroepiandrosterone; FSH = Follicle stimulating hormone; LH = Lutenising hormone; GH = Growth hormone.

The gold standard measurement of pubertal development is via physical examination by a clinician, as self-reported pubertal development measures can be subject to biases (Shirtcliff et al., 2009). In research settings, especially in large scale studies like ABCD, a physical examination may not always be possible. Therefore, pubertal development is frequently assessed via self- (or parent-) report, such as the Pubertal Development Scale (PDS; Petersen et al., 1988) or the Tanner Scale (Marshall & Tanner, 1969, 1970), and also via hormonal measures (Goddings et al., 2019). It is important to note that questionnaire-based assessments of puberty (e.g., PDS) can conflate three distinct (but related) neuroendocrine processes, a nuance that should be acknowledged when attempting to unravel how pubertal development relates to other aspects of adolescence, such as brain development and mental health risk.

Although all individuals progress through the same stages of puberty, there are marked individual differences in the timing and pattern (i.e., tempo) of pubertal maturation. There are a number of factors that are associated with pubertal onset including genetics, nutritional status, adoption, and emotional well-being (for a recent review see Mancini et al., 2022).

Indeed, findings from genetic studies suggest that 50-80% of variance in pubertal timing may have genetic underpinnings (Abreu & Kaiser, 2016). Further, the decrease in the average age of menarche in the US and some parts of Europe between the mid-19th and mid-20th century is thought to be linked to general improvements in health, nutrition, and living conditions (Wyshak & Frisch, 1982). However, the heterogenous research methods employed in these earlier studies limits the generalisability of such findings to other populations, including present day youth. This underscores the need for further longitudinal research using more recent data and ideally involving studies with harmonised indices of pubertal development across diverse samples. Crucially, it is the variation in pubertal timing, rather than pubertal stage, that is associated with psychopathology during adolescence (Graber, 2013; Ullsperger & Nikolas, 2017). Pubertal timing can be defined as pubertal development relative to same-age, same-sex peers, such that an individual can be categorised as developing ahead (early), in-line (on-time) or after (late) their peers. Given that this thesis is focused on investigating how features of the adolescent brain relate to depression during this period, the role of pubertal maturation in this association will be examined through the lens of pubertal timing.

Beyond age-related changes, research has shown that pubertal development impacts neuromaturation (Vijayakumar et al., 2018). Regarding brain structural development, earlier pubertal timing (measured via physical and hormonal measures) has been associated with reduced cortical thickness and cortical volume in the prefrontal cortex, the cingulate cortex, and the temporal lobe (Koolschijn et al., 2014; Pfefferbaum et al., 2016). These brain regions are involved in cognitive processes known to undergo significant development during adolescence such as cognitive control, decision making, and emotion regulation. While advanced pubertal maturation has been associated with subcortical changes, such as increased amygdala and hippocampal volume and decreased volume of striatal regions, there is a dearth of research examining pubertal timing and brain structural development specifically (Goddings et al., 2019). Further, findings on the association between white matter microstructure and pubertal timing have been inconsistent, although there is some degree of support across studies for a positive association between pubertal timing and FA (see Vijayakumar et al. (2018) for a review).

Research on how pubertal development relates to functional brain development is an emerging line of research. Given that examining functional connectivity and pubertal timing in the context of adolescent depression was beyond the scope of the current thesis, the associated literature is not reviewed extensively here. In brief, of the handful of fMRI studies that do exist, most have focused on reward processing and socio-emotional development related regions (e.g., striatal regions and the amygdala) given the significant maturation of these domains during adolescence (Andrews et al., 2021). Findings from these studies have been mixed with some reporting increased activation in striatal regions and decreased activation of the amygdala in more pubertally advanced youth, while others have found the opposite pattern of results, or no puberty-related activation changes (for recent review see Galván, 2021).

In conclusion, it is evident from the existing literature that pubertal and brain development interact in a complex manner during adolescence, and much more longitudinal work is needed to better understand this association. A major challenge in this area of research is the multitude of methods used to assess pubertal development, which may in themselves contribute to the heterogenous findings reported. Importantly, multi-verse analysis approaches have gained popularity in recent years (Barendse et al., 2021) and will be key to producing robust and reproducible findings in the future.

1.7 The adolescent brain and depression: previous research and the current thesis

The research discussed to date undoubtedly shows that adolescence is a time of immense biological change. It is therefore unsurprisingly that deviations from typical biological development have been the focus of research on the vulnerability to mental health disorders, such as depression, during adolescence. However, the research undertaken to date has been limited in several ways.

Firstly, it is only in recent years that we have had sufficiently powered neuroimaging studies to begin to understand how features of the adolescent brain relate to the emergence of

depression during this period. Large population-based studies (N = >1,000), such as work from the ENIGMA consortium (Enhancing NeuroImaging Genetics through Meta-Analysis) have demonstrated widespread brain structural alterations in adult MDD. For instance, lower hippocampal volume, decreased cortical thickness in frontal areas (Schmaal et al., 2017), and differences in white matter microstructure in fronto-limbic and fronto-thalamic tracts (van Velzen et al., 2020) have been reported in MDD cases compared to controls. However, since these studies were conducted in adults, it is not possible to ascertain whether the observed brain structural deviations are a cause or consequence of depression. Although there have been some recent efforts to examine the brain structural associations with depression in adolescence, which report lower total surface area and regional reductions in frontal regions in depressed cases, findings have been inconsistent, and this work has largely comprised older adolescent cases (≥ 16 years) (Arnone et al., 2012; Kempton et al., 2011; Lai, 2013; Reynolds et al., 2014; Serafini et al., 2014; Shad et al., 2012). As a result, the earlier origins of depression-related brain structural alterations in adolescence remain understudied, likely owing to a lack of large neuroimaging studies comprising adolescent samples. Longitudinal work (involving at least three distinct timepoints of data) is needed to chart brain development alongside depression trajectories across adolescence so that we can identify multi-factorial profiles that confer risk for or promote resilience to depression over time. However, until such data becomes available through studies like ABCD, a first step in achieving this goal is to examine the origins of depression in early adolescence, which will provide a strong foundation upon which to base future longitudinal work.

A second key limitation of existing research within the field to date is the examination of risk factors independently. While this approach may be due to the targeted focus of independent studies in the past, as opposed to the broad scope of cohort studies like ABCD, it is important that our study designs reflect the dynamic and interactive world of adolescence. Although a large body of research has demonstrated that earlier pubertal timing is associated with a heightened vulnerability to depression during adolescence (Conley et al., 2012; Ge & Natsuaki, 2009; Hamilton et al., 2014; Pfeifer & Allen, 2021; Ullsperger & Nikolas, 2017), the neurobiological mechanisms underpinning this relationship are not well understood. A number of theoretical models have been proposed to explain this association such as the

"maturation disparity hypothesis", which posits that youth that begin puberty ahead of their peers experience psychological distress due to a mismatch between their accelerated physical development and asynchronous development of frontal and limbic brain regions (Brooks-Gunn et al., 1985; Ge & Natsuaki, 2009; Ullsperger & Nikolas, 2017). However, neuroimaging and pubertal development data from a large sample of adolescents (>500 youths) has only recently been made available through ABCD. Although other longitudinal adolescent cohort studies like IMAGEN do exist, baseline data collection started in mid-adolescence when the youth were aged 14 years old. Thus, ABCD has the crucial advantage of following youth from the early (or even pre-pubertal) stages of adolescence right through the young adulthood. In time, this will allow us to unravel the complex interplay between pubertal development, neuromaturation, and the social environment of youth, and how this relates to the onset of psychopathology during adolescence.

Although "Big Data" research is well-positioned (and powered) to detect subtle effects such as individual differences in development and the associated contributing factors, the breadth of studies like ABCD can come at the cost of phenotypic depth. As such, cohort studies may be better conceptualised as tools for generating hypotheses that can then be formally tested in smaller studies that target a specific mechanism (Saragosa-Harris et al., 2022), as the measures available may not match the specific research question at hand. Therefore, the bespoke nature of small-scale studies, and the accompanying creativity, helps keep our research questions timely and our methods appropriate.

The opportunity to co-produce research with young people is an ideal example of a strength of small-scale studies. Co-production helps us ensure that our research questions take us towards the world of adolescence rather than away from it (MacSweeney et al., 2019; Whitmore & Mills, 2021). For example, although irritability is regarded as a cardinal symptom of adolescent depression, the social and interactive context in which irritability occurs during this developmental period has typically been overlooked in study designs to date. Both irritability and depression have been associated with disruptions to the integration of large-scale functional brain networks involved in emotion processing and cognitive control, such as the DMN, FPN, and SN. However, existing fMRI paradigms targeting irritability (e.g.,

frustrative non-reward paradigms) do not tend to capture the rich social tapestry of the world in which this irritable mood occurs. This highlights the need for novel study designs that better reflect the social nature of adolescence.

1.8 Thesis aims

By leveraging the unique strengths offered by large and small-scale neuroimaging studies, the current thesis makes a direct contribution to existing knowledge gaps in the field by addressing the following four aims, of which the first two involve data from the ABCD Study while the latter two use data from a locally collected pilot study:

1. The first aim of this thesis was to examine the temporal origins of brain structural associations with depression in early adolescence using baseline data from the ABCD Study when youth are aged 9-11 years (Chapter 3).

2. Building on work described in Chapter 3, the second aim of this thesis was to investigate whether brain structure mediated the association between earlier pubertal timing and later depression using three waves of follow-up data from the ABCD Study when the youth are aged 9-13 years (Chapter 4).

3. The third aim of this thesis was to review current findings on the neural circuitry of irritability in adolescent depression, highlight directions for future research, and emphasise the importance of co-produced research with young people to improve the ecological validity of research within the field (Chapter 5).

4. Directly addressing the avenues for further research identified in Chapter 5, the final aim of this thesis was to develop a novel fMRI paradigm targeting irritability in a sample of local youth with depressive symptoms, using a co-produced youth-researcher design. Adopting a dynamic functional connectivity approach, I then tested the validity of this task as way of inducing irritable mood, and how features of dynamic

functional brain networks relate to depressive symptoms and irritable mood in a sample of adolescents aged 16-18 years (Chapter 7).

2 ABCD Methods

2.1 Chapter introduction

In this chapter, I first provide an overview of the Adolescent Brain Cognitive Development Study® data, as used in Chapters 3 & 4 of this thesis. I then describe the neuroimaging measures and associated quality control protocols, as well as the depression and puberty measures in this cohort study. Finally, I outline the rationale for the statistical methods used in this work.

Elements of this chapter were adapted from a paper published in *Developmental Cognitive Neuroscience* in June 2022 where I was joint-first author. This paper arose from discussions during the 2021 Modelling Developmental Change in the ABCD Study Workshop (https://abcdworkshop.github.io).

2.2 The ABCD Study

The Adolescent Brain Cognitive Development (ABCD) Study® is the largest longitudinal developmental neuroimaging study to date with ~11,800 9-10-year-olds recruited at baseline between 2016 and 2018. The baseline cohort are being followed up for ten years with data collected annually (non-imaging measures) and biannually (imaging measures and bioassays), as well as mid-year phone interviews. The original objective of the ABCD Study was to investigate risk and resilience factors related to the development of substance use disorders (e.g., cannabis). However, the scope of the project has since evolved and is now focused on examining the biopsychosocial correlates of mental and physical health in the second decade of life (Barch et al., 2021). Undoubtedly, the ABCD Study is a valuable resource to the field of developmental cognitive neuroscience with immense potential to provide evidence-based policy recommendations to better the lives of young people and their families (Feldstein Ewing et al., 2018).

The ABCD Study provides an annual data release which includes tabulated behavioural, questionnaire and imaging data along with detailed release notes. The projects included in this thesis used data from release 2.0.1 (Chapter 3; 08/07/2019) and release 4.0 (Chapter 4; 27/10/2021). Access to the ABCD Study data was granted by the National Institute for Mental Health (NIMH) under Data User Certificate ID: 10607.

2.2.1 Recruitment

Participants (and their parents/guardians) were recruited from 21 nationally distributed sites across the United States (see Figure 2.1) with the aim of creating a population-level, socio-demographically-diverse sample (Garavan et al., 2018a). The primary recruitment strategy for ABCD was school-based, whereby recruitment materials were given to all children within the target age range to take home to their caregivers. Interested families contacted the study site directly, underwent a brief phone screening and if eligible, their baseline assessment was scheduled. Participants were also recruited through summer-camps and youth groups to avoid a recruitment lag during the summer months. Further, researchers at each site were encouraged to use their local knowledge to engage with under-recruited youth (e.g., minority or low-income families).

ABCD Study exclusion criteria included non-English proficiency in the young person, general MRI contraindications, a history of a major neurological disorder, traumatic brain injury, extreme premature birth (<28 weeks gestational age), a diagnosis of schizophrenia, intellectual disability, moderate to severe autism spectrum disorder, or substance abuse disorder (Karcher et al., 2018).

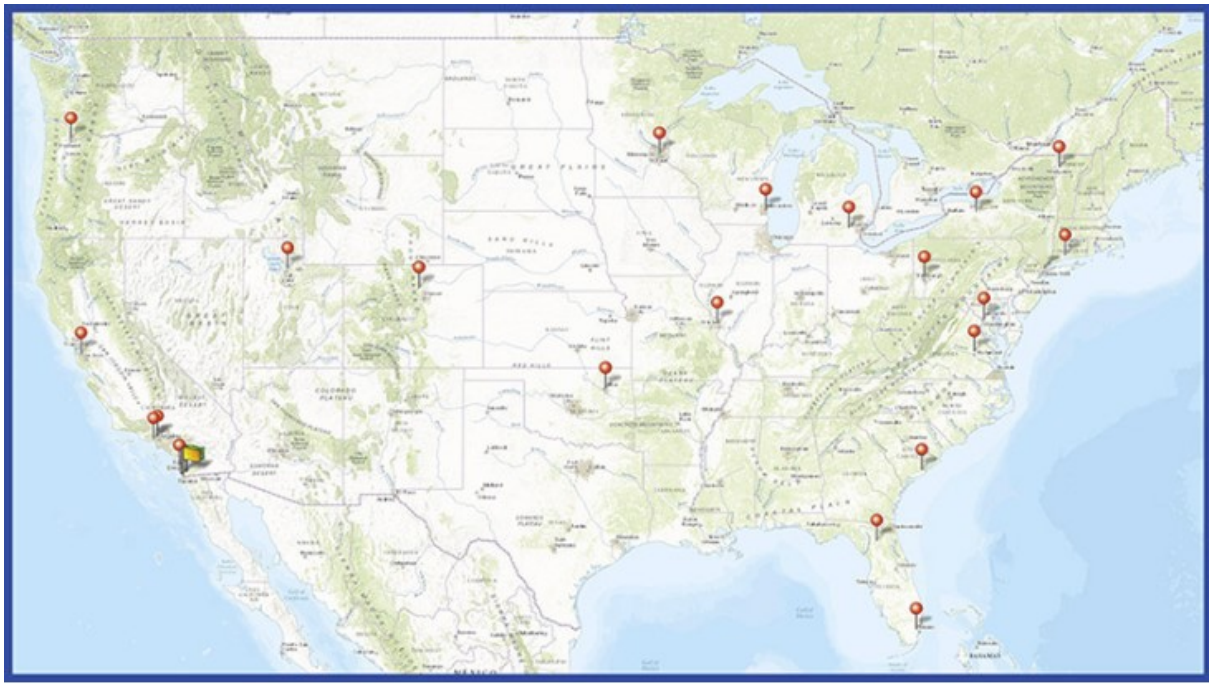


Figure 2.1 — Map of ABCD Study sites. Source: www.abcdstudy.org.

2.2.2 Funding and ethics

The ABCD Study is supported by the National Institutes of Health and additional federal partners under award numbers U01DA041048, U01DA050989, U01DA051016, U01DA041022, U01DA051018, U01DA051037, U01DA050987, U01DA041174, U01DA041106, U01DA041117, U01DA041028, U01DA041134, U01DA050988, U01DA051039, U01DA041156, U01DA041025, U01DA041120, U01DA051038, U01DA041148, U01DA041093, U01DA041089, U24DA041123, U24DA041147. Most ABCD research sites obtained ethical approval via a central Institutional Review Board (IRB) at the University of California, San Diego, with some sites obtaining local IRB approval (Auchter et al., 2018).

2.3 ABCD measures

The ABCD Study includes a wide range of data on youth and their families' mental and physical health, environmental context, behaviour, genes, and neurocognitive development. The main measures used in the current thesis are briefly outlined below. Full descriptions of these measures can be found in Chapters 3 & 4.

2.3.1 Neuroimaging measures

ABCD adopted an optimised MRI acquisition protocol to measure brain structure and function, which builds upon efforts made by other Big Data studies such as the Human Connectome Project (HCP; https://www.humanconnectome.org) and the Paediatric, Imaging, Neurocognition, and Genetics (PING) Study (Jernigan et al., 2016). This protocol is compatible with all three 3 tesla (T) scanner types used across sites: Siemens Prisma, General Electric 750, and Phillips. The imaging protocol includes 3D T1 and T2 weighted (T1w and T2w) and diffusion weighted (DTI) images for measures of brain structure, and resting state and task-based functional MRI for measures of brain function (Casey et al., 2018a). This thesis used the T1w and DTI data, which are detailed below.

2.3.1.1 Scanning protocol

A T1w sequence acquired 176 contiguous 1.0mm slices (matrix = 256 x 256, FoV = 256mm, flip angle = 8°) using RF-spoiled gradient echo scanning for cortical and subcortical segmentation. A high angular resolution diffusion imaging (HARDI) scan (81 1.7mm slices, matrix = 140 x 140, FoV = 240mm, flip angle = 77/78/90° (scanner dependent)) using multiband echo planar imaging (EPI), with multiple b-values, and fast integrated B_0 distortion correction (reversed polarity gradient method (RPG) was collected for white matter tract segmentation and diffusion measurement (Holland et al., 2010). A standard adult-size head coil (32 or 64 channel depending on scanner) was used over a nonstandard customised coil. There is empirical evidence to support its use for this age group and it avoided a number of analysis and practical challenges, such as the issue of a customed head coil being confounded with age (Burgund et al., 2002; Kang et al., 2003). Further, at ages 9-10 years when baseline data was collected, the brain is 90-95% of the adult brain size (Casey et al., 2018a).

The ABCD Neuroimaging Protocol is outlined in Figure 2.2. Participants completed the scanning session in one or two visits depending on scanner and participant availability. The scanning protocol was piloted in a subsample of participants across sites and data showed that scanning in a single or double session did not significantly affect task performance or tiredness levels (Casey et al., 2018a) (see Figure 2.2). To minimise motion, the young person's head was stabilised with foam padding. Further, real-time motion correction and motion

monitoring was used on T1w acquisitions (White et al., 2010), which has been shown to significantly reduce motion-related image degradation (Tisdall et al., 2016).

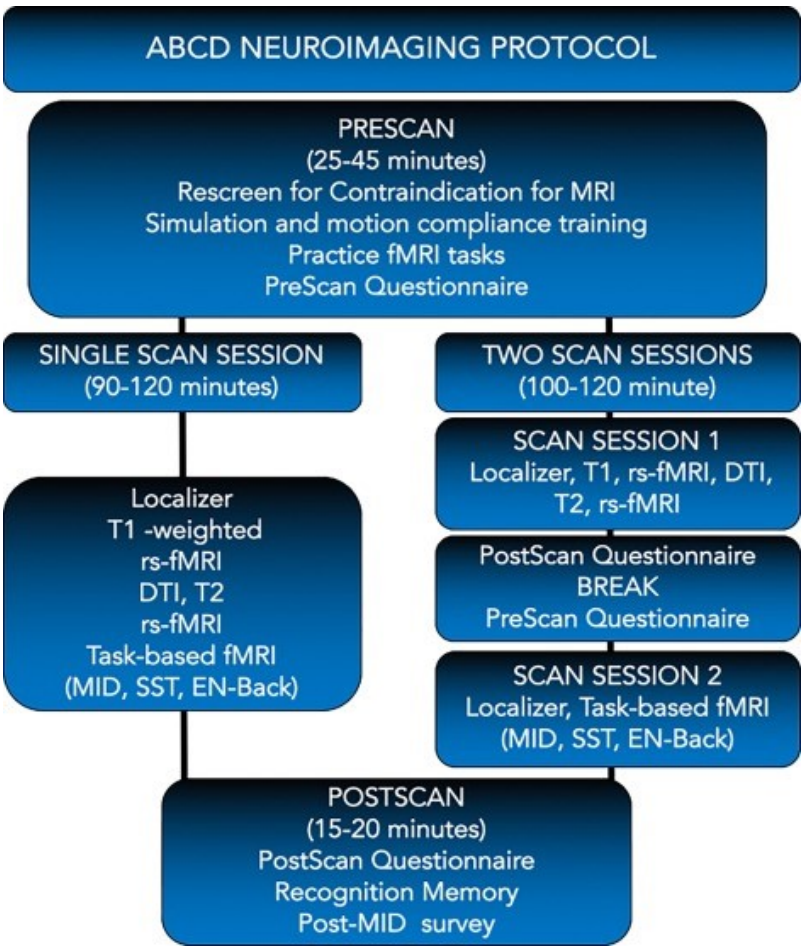


Figure 2.2 — ABCD Neuroimaging Protocol. Source: Casey et al. (2018). Copyright: Developmental Cognitive Neuroscience. MID = Monetary incentive delay task; SST = Stop signal task; EN-back = N-back task.

2.3.1.2 Imaging pre-processing and quality assessment

The minimally processed T1w and DTI data provided by ABCD in releases 2.0.1 and 4.0 were used for Chapters 3 & 4, respectively. This modality specific pre-processing (outlined below) was carried out by the ABCD Data Analytics and Informatics Centre (DAIC) including converting raw DICOM files to compressed files, distortion and motion correction, alignment to standard space, initial quality control, and post-processing quality control (QC).

Before undergoing pre-processing, all image series were visually inspected by at least two trained ABCD technicians to check for indicators of poor image quality (e.g., excessive head

motion, severe ghosting, blurring, or ringing). As these artefacts make accurate brain segmentation impossible, all images that failed the initial QC were removed from further processing and analysis by the ABCD team.

2.3.1.3 T1w pre-processing

First, T1w images were corrected for gradient non-linearity distortions according to scanner-specific non-linear transformation protocols. To account for intensity non-uniformity typical in MRI, where there are inconsistent intensity variations across brain tissue depending on the distance from the head coil (i.e., receive coil bias), an in-house bias field correction method was applied. Next, the images were registered and resampled to standard space using an in-house reference brain. This reference brain has 1.0mm isotropic voxels and was generated by averaging T1w images from 500 adults after they had been nonlinearly registered to a template brain image using discrete cosine transformations (Friston et al., 1995).

FreeSurfer v5.3 was used for cortical surface reconstruction and subcortical segmentation (Dale et al., 1999; Fischl et al., 2002, 2004; Fischl & Dale, 2000; Ségonne et al., 2004). Morphometric measures used included cortical volume, thickness, area, and sulcal depth, and volumetric measures only for subcortical brain regions. FreeSurfer has been validated for use in youth (Ghosh et al., 2010) and has already been used successfully in other large-scale paediatric neuroimaging studies (e.g., Paediatric Imaging, Neurocognition, and Genetics (PING) Study; Jernigan et al., (2016)). The FreeSurfer pipeline labels subcortical structures using an automated volumetric segmentation procedure according to Talairach atlas (Fischl et al., 2002). In ABCD, this framework yielded 30 labelled subcortical measures. Cortical grey matter and underlying white matter structures are labelled according to surface-based nonlinear registration to the Desikan-Killiany (DK) atlas based on gyral structures (Desikan et al., 2006). In the DK atlas, a gyrus is defined as running between the bottom of two adjacent sulci, which parcellates the brain into 34 regions per hemisphere (see Figure 2.3). The Destrieux atlas is also available in FreeSurfer and divides the brain into gyral and sulcal regions based on the curvature value of the surface (Destrieux et al., 2010). This method provides 74 sulco-gyral structures per hemisphere. Although both atlases are well-validated and widely used (Hagler et al., 2019), the DK atlas was chosen over the Destrieux for its fewer number of

brain regions to limit the number of models tested in the whole brain analyses undertaken in Chapters 3 & 4. The DK atlas was also chosen for consistency with other large-scale neuroimaging studies such as the UK Biobank and the Lothian Birth Cohort.

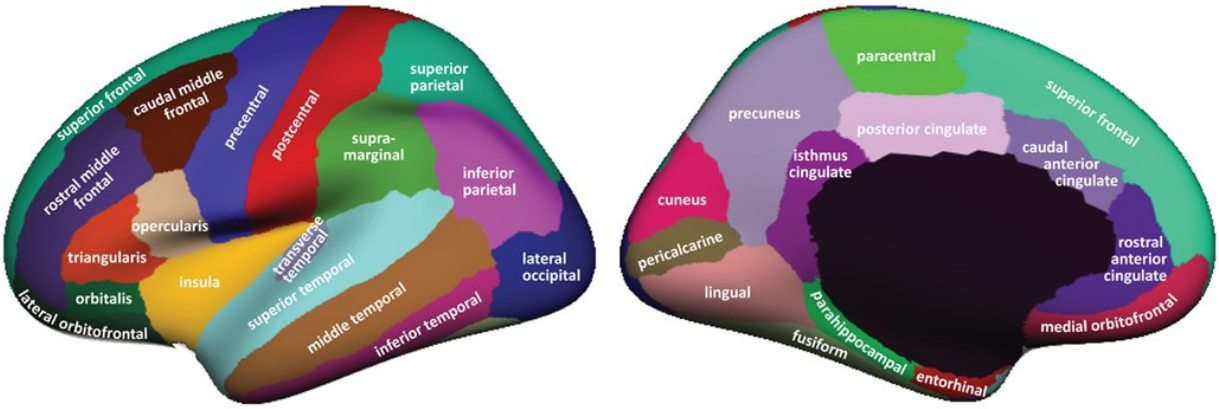


Figure 2.3 — Illustration of the Desikan-Killiany (DK) Atlas brain parcellation ; (34 regions per hemisphere). Source: FreeSurfer.com.

2.3.1.4 T1w quality control

Given the sheer volume of scans collected in ABCD, an exhaustive manual review of every scan was not practical. Instead, ABCD used an AI-guided automated quality assessment protocol whereby specific QC metrics (e.g., bad registration and brain cut-off) were used to identify scans to be sent for manual review. The manually reviewed subsample also included a random sample as well as participants with data that had been flagged as statistical outliers. This accounts for ~7% of ABCD participants with imaging data. Reviewers assigned a binary (0 = reject |1 = accept) QC score as well as rating the severity of the problem for five types of artefact: motion, intensity inhomogeneity, white matter underestimation, pial overestimation, and magnetic susceptibility artefact. Ratings included absent, mild, moderate, or severe, and were labelled 0 to 3, respectively. In the most recent ABCD data release 4.0., the DAIC have created a variable that indicates whether the T1w imaging data has been recommended for inclusion (0 = exclude, 1 = include). This variable encompasses all QC criteria (e.g., initial QC, FreeSurfer QC) which were handled individually in previous releases.

2.3.1.5 DTI pre-processing

In DTI, eddy currents are created due to a changing magnetic field in the conductor. This can lead to distortion of the images and therefore, eddy current correction (ECC) needs to be applied. EEC uses a model-based approach where the pattern of distortion is predicted across the entire set of diffusion weighted images, based on diffusion gradient orientations and amplitudes (Zhuang et al., 2006). In ABCD, to correct for head motion the post-ECC diffusion images were adjusted for head rotation (Leemans & Jones, 2009). Further, average frame-wise displacement values were also calculated which were used as a covariate in my statistical analyses to account for residual effects of head motion (Yendiki et al., 2014). Spatial and intensity distortions induced by B_0 field inhomogeneity in EPI images were minimised using a validated and accurate method that relies on reversing phase-encoding polarities (Holland et al., 2010). The diffusion weighted images were then resampled with 1.7mm isotropic resolution (equal to the DTI acquisition resolution) with a fixed rotation and translation relative to the corresponding T1w image. This resulted in a standard orientation for the diffusion weighted images which produces more consistent diffusion orientations across participants. This improved the registration accuracy of the DTI images to the T1w images (Hagler et al., 2019).

AtlasTrack (AT), a probabilistic atlas-based method, was used for the automated segmentation of major white matter tracts. The AT fibre atlas contains probabilities and orientation information for specific long-range projection fibres. Additional tracts not included in the original AT atlas were added by the ABCD team. T1w images for each participant were registered to the AT atlas using discrete cosine transformations, and the diffusion weighted orientations for each participant were then compared to the pre-defined AT orientations and tract location probabilities. Following this atlas registration, several standard microstructural measures relating to the white matter tissue properties using DTI were calculated, including fractional anisotropy (FA) and mean diffusivity (MD) which I used in later statistical analyses. ABCD provides two types of DTI model fits, inner shell (includes 6 directions at b = 500 s/mm^2 and 15 directions at b = 1000 s/mm^2 and full/multi shell (includes all gradient directions and strengths (directions = 6, 15, 60; b = 500, 1000/2000, 6000 s/mm^2). Although novel multi shell diffusion techniques, such as Restriction Spectrum Imaging (RSI)

allow the estimation of both restricted and hindered diffusion within individual voxels (White et al., 2013), I used the inner shell fit in my analyses to keep my methods consistent with existing large-scale studies (e.g., ENIGMA, UKB).

2.3.1.6 DTI quality control

Like the QC protocol applied to the T1w data, the DTI data underwent automated AI-guided QC which identified a subsample of scans for subsequent manual review. However, this sampling approach was only applied from Release 3.0 onwards. Therefore, in Chapter 3, which used data from release 2.0.1, the scans that underwent manual QC comprised the subset of participants available in release 1.1. Trained reviewers assigned a binary (0 = reject |1 = accept) QC score and rated the severity of the problem for five artefact types: B_0 warping, motion, full head coverage, registration with T1w image, and accuracy of fibre tract segmentation. Ratings included absent, mild, moderate, or severe, and were labelled 0 to 3, respectively. Like the T1w data from release 4.0, the modality specific inclusion QC variable was used for the DTI data in Chapter 4. This is a binary variable (0 = exclude | 1 = include) that covers the individual DTI QC criteria handled individually in earlier releases.

Therefore, for both T1w and DTI data, the QC approach taken in this thesis varies slightly between Chapters 3 & 4 and is detailed within each chapter. The quality assessment protocol adopted by ABCD is described in full in their annual release notes (https://nda.nih.gov/study.html?id=1299) and in the ABCD imaging protocol paper by Hagler and colleagues (Hagler et al., 2019).

2.3.2 Depression measures

In brief, the primary measures used to quantify youth depression were the Kiddie-Schedule for Affective Disorders and Schizophrenia (K-SADS; Kaufman et al., 1997) and the Child Behaviour Checklist (CCBL; Achenbach, 2011) for Chapters 3 and 4, respectively. Additional details of the depression measures used can be found in the individual thesis chapters.

In Chapter 3, a computerised version of the K-SADS was used in the ABCD Study to assess lifetime (past and/or current) MDD and depressive symptoms in youth at baseline (Kaufman et al., 1997). The K-SADS was completed by both youth and their caregiver separately and was self-administered. The computerised version of the K-SADS has been shown to have good to high reliability (AUC = 0.89 – 1.00) compared to the clinician administered version (Townsend et al., 2020). To quantify MDD, we used the MDD diagnosis binary measure created by the ABCD team. We also created an additional measure of depressive symptom (DS) severity based on the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) criteria, which categorised DS as “severe”, “moderate”, “mild” and “none”. This DS measure was validated using the CBCL.

In November 2021, the ABCD team reported in the 4.0 release notes that an error had been discovered with the algorithm used by the ABCD team to generate an MDD diagnosis from the K-SADS. The algorithm used did not include impairment in the diagnostic criteria which will likely have led to the overestimation of MDD diagnoses in all ABCD data releases to date. The paper that comprises Chapter 3 was published before this error was reported, and thus the methods pertaining to this chapter describe the original analysis undertaken for the paper. At time of writing, this data error had not been fixed and thus, we were unable to re-run our analyses to examine the extent to which it may have affected our findings. This data error is discussed further in the conclusion of Chapter 3 following the main body of the paper.

In Chapter 4, the CBCL “withdrawn-depressed” syndrome (raw scores) parent report was used to examine current youth depressive symptoms. The CBCL is one of the most validated and widely used measures to assess internalising and externalising difficulties in young people (Achenbach, 2011; Achenbach & Rescorla, 2004). Further, a lifetime measure of depressive

symptoms was not appropriate for the analyses in this chapter given the proposed directionality of effects (i.e., earlier pubertal timing was hypothesised to be associated with later depressive symptoms).

In both Chapters 3 & 4, the Adult Self Report (ASR) scale in the Achenbach System of Empirically Based Assessment (Barch et al., 2018a) was used to account for potential biases introduced by the current mood of caregivers on the reporting of their child's psychopathology (Maoz et al., 2014).

2.3.3 Pubertal development measure

The assessment of pubertal development in this thesis is described in full within Chapter 4. In brief, the Pubertal Development Scale (PDS) was used to measure perceived pubertal development (Petersen et al., 1988). This is a five-item questionnaire that assesses the development of secondary sex characteristics, where each is rated on a 4-point scale (1 = no development; 2 = development has barely begun; 3 = development is definitely underway; and 4 = development is complete). Higher scores thus reflect more advanced pubertal development.

As previous research has shown that youth tend to over report pubertal development at younger ages (Schlossberger et al., 1992), the PDS parent report was used over the youth self-report. Moreover, there was a significant degree of missing data (~50%) in the PDS youth self-report data, which was another deciding factor in my measure choice. As the research questions and hypotheses in Chapter 4 pertained to pubertal timing specifically, the PDS total score was regressed on age for males and females separately, and the standardised residual obtained was used as a continuous measure of pubertal timing (Dorn et al., 2006; Hamilton et al., 2014).

2.4 Overview of key statistical methods

Detailed statistical methods sections can be found within Chapters 3 & 4. Here, I provide a brief rationale for the methods chosen.

A combination of generalised linear-mixed models (GLM) and linear mixed effects models (LME) were used to examine the association between adolescent depression and the biological and psychosocial factors of interest in this thesis. GLMs are an extension of the general linear model but provide a more flexible analytic framework that can characterise non-normal dependent variables. Given that the current dependent variables of interest were often binary (e.g., MDD case/control) or discrete counts (e.g., number of depressive symptoms), GLMs were used to avoid violating statistical assumptions (e.g., normally distributed residual values) (Gardner et al., 1995).

In Chapter 3, I used GLMs to model unilateral brain regions and LMEs to model bilateral brain regions, where hemisphere was treated as a within-participant fixed effect and participant ID was modelled as a random effect. As the analyses in this chapter pertained to an unrelated sample of participants, it was not necessary to account for family ID as a random factor. Further, site ID was modelled as a fixed effect.

However, in Chapter 4, I decided to adapt my analytic approach so that it better reflected the related structure of ABCD, maximising the data available for the mediation analyses using follow up data. Thus, GLMs were used with family ID and scanner ID (or site ID for non-imaging models) as random factors. Given that some ABCD sites have multiple scanners, scanner ID is the recommended variable to use to account for inter-site differences and is better modelled as a random factor. Using a related sample in Chapter 4 resulted in some adjustments to my analysis methods compared to Chapter 3. Namely, the mean of bilateral brain structures was used due to the additional complexity of the mixed effects analysis structure and non-convergence of the models using a complex random factor structure (i.e., inclusion of participant ID, family ID and scanner ID resulted in model convergence issues). Further, given that temporally separated variables were available in release 4.0, it was possible to conduct mediation analyses in Chapter 4 using a structural equation modelling (SEM) framework (Maxwell et al 2011). SEM has grown in popularity as a way to examine how a predictor variable X relates to some outcome variable Y via one (or many) intervening pathways, and how well different models fit the observed data (Hayes, 2009).

2.5 Summary

In this chapter, I outlined the neuroimaging measures and quality control procedures, and the depression and puberty measures in the ABCD cohort. I also provided rationale for the statistical methods employed in Chapters 3 & 4 of this thesis. In the next chapter, I introduce my first study in which I investigated brain structural associations with adolescent depression using baseline data from the ABCD Study.

3 Brain Structural Associations with Depression in Adolescence

3.1 Chapter introduction

Findings from large-scale neuroimaging studies in adults suggest that depression is associated with alterations in cortical measures and white matter microstructure. However, most research to date has been conducted in adults. Further, the handful of studies examining depression-related imaging features in adolescents have reported highly heterogeneous findings and have mostly involved older adolescent samples. Therefore, the temporal origins of cortical and white matter microstructural changes associated with the emergence of depression in adolescence remains largely unknown. Exploring how depression-related imaging features in early adolescence relate to findings in adults could provide important insight into the aetiology of depression and inform timings of potential interventions.

In this chapter, we therefore undertook whole brain exploratory analyses to examine associations between brain structure (cortical metrics and white matter microstructure) and youth depression ratings from both parent- and child-report. Here, we used baseline data from the ABCD Study (release 2.0.1) when youth were aged 9-11 years (note: the ABCD baseline age range is often reported as being 9-10 years, a discrepancy that may be due to whether the age of youth at the extreme of this range (e.g., 10.99 years) is rounded up or not). In Chapter 4, the baseline age range is reported as 9-10 years as per the initial ABCD protocol papers (Casey et al., 2018a; Hagler et al., 2019).

This study was published in *EClinicalMedicine* in November 2021. I co-led this study with Dr Xueyi Shen and we share first authorship on the associated paper. The manuscript of this study is included in this chapter in a Word document format. Author contributions are included within the manuscript. Minor edits have been made to the main manuscript (e.g., updated numbering for the main tables and figures) to keep the formatting consistent across

this thesis. The Supplementary Information for this paper/chapter, including the Supplementary Data, and a PDF of the published paper, are included in Appendix 1.

Brain structural associations with depression in a large early adolescent sample (the ABCD Study®)

Xueyi Shen[1*+], Niamh MacSweeney[1+], Stella W. Y. Chan[2], Miruna C Barbu[1], Mark J Adams[1], Stephen M Lawrie[1], Liana Romaniuk[1], Andrew M McIntosh[1], Heather C Whalley[1]

[1] Division of Psychiatry, University of Edinburgh, Edinburgh, United Kingdom

[2] Department of Clinical Psychology, University of Edinburgh, Edinburgh, United Kingdom

[+]Co-first authors

*Corresponding author:

Dr Xueyi Shen

E-Mail: xueyi.shen@ed.ac.uk

Address:

University of Edinburgh, Royal Edinburgh Hospital, Morningside Park

Edinburgh EH10 5HF

United Kingdom

Keywords: Big data, Adolescent depression, Adolescent Brain Cognitive Development Study, Brain structure

3.2 Abstract

Background Depression is the leading cause of disability worldwide with >50% of cases emerging before the age of 25 years. Large-scale neuroimaging studies in depression implicate robust structural brain differences in the disorder. However, most studies have been conducted in adults and therefore, the temporal origins of depression-related imaging features remain largely unknown. This has important implications for understanding aetiology and informing timings of potential intervention.

Methods Here, we examine associations between brain structure (cortical metrics and white matter microstructural integrity) and depression ratings (from caregiver and child), in a large sample of early adolescents (9 to 11 years old) from the US-based, Adolescent Brain Cognitive Development (ABCD) Study®. Data was collected from 2016 to 2018.

Findings We report significantly decreased global cortical and white matter metrics, and regionally in frontal, limbic and temporal areas in adolescent depression (Cohen's d = -0·018 to -0·041, β = -0·019 to -0·057). Further, we report consistently stronger imaging associations for caregiver-reported compared to child-reported depression ratings. Divergences between reports (caregiver vs child) were found to significantly relate to negative socio-environmental factors (e.g., family conflict, absolute β = 0·048 to 0·169).

Interpretation Depression ratings in early adolescence were associated with similar imaging findings to those seen in adult depression samples, suggesting neuroanatomical abnormalities may be present early in the disease course, arguing for the importance of early intervention. Associations between socio-environmental factors and reporter discrepancy warrant further consideration, both in the wider context of the assessment of adolescent psychopathology, and in relation to their role in aetiology.

Funding Wellcome Trust (References: 104036/Z/14/Z and 220857/Z/20/Z) and the Medical Research Council (MRC, Reference: MC_PC_17209).

3.3 Introduction

Major depressive disorder (MDD) is a chief cause of disability (James et al., 2018) with a heritability of approximately 37% (Sullivan et al., 2000). This burden falls heavily on adolescents, as over 50% of depression cases emerge before the age of 25 (World Health Organisation, 2012). Adolescent depression is notably associated with a more severe illness course and can lead to the propagation of difficulties across the lifespan (Thapar et al., 2012). MDD is associated with disruptions in brain structure (Schmaal et al., 2017, 2020; Shen et al., 2017, 2019a). However, due to a lack of large-scale neuroimaging samples for adolescents, the origin and development of depression-related imaging features remains largely unknown.

Large population-based neuroimaging studies in adults have allowed unparalleled insight into the neurobiological underpinnings of depression (Shen et al., 2017, 2019a). For example, recent evidence from the ENIGMA (Enhancing NeuroImaging Genetics through Meta-Analysis) consortium demonstrated widespread structural abnormalities in MDD from large adult samples, including reduced hippocampal volume, decreased frontal cortical thickness (Schmaal et al., 2017; N=10,105) and altered fronto-limbic and fronto-thalamic tract microstructure (van Velzen et al., 2020; N=2,907). Since these highly powered studies have largely been conducted in adults, they preclude investigation of the neurobiology underlying the emergence and development of depression earlier in life. Given adolescence is the period of greatest risk for the development of depression (Thapar et al., 2012), as well as a time of immense neurodevelopmental change (Mills et al., 2016; Tamnes et al., 2017), it is a key period in which to investigate evidence for the emergence of these imaging features.

Findings from earlier studies on brain structural alterations in adolescent MDD have been highly heterogeneous (Arnone et al., 2012; Kempton et al., 2011; Lai, 2013; Reynolds et al., 2014; Serafini et al., 2014; Shad et al., 2012). A recent meta-analysis of imaging studies of MDD from ENIGMA, which included a relatively large adolescent population (N=507, age range 12-21 years), indicated lower global surface area and regional reductions in frontal areas in this younger sample of depressed cases (Schmaal et al., 2017). However, this subsample comprised primarily of participants from older adolescence to young adulthood, where 90% of the sample were aged ≥16 years, meaning earlier origins of depression related

brain imaging features remain underexplored. There have also been recent efforts to investigate whether white matter integrity disruptions seen in adult cases are present in adolescent depression. Although some studies report reduced white matter microstructural integrity in adolescents with depression, findings have lacked consistency in terms of regions (Bessette et al., 2014; Jones et al., 2019; LeWinn et al., 2014) and effects sizes (Aghajani et al., 2014; Henderson et al., 2013), likely due to small sample sizes (Bessette et al., 2014; Jones et al., 2019; LeWinn et al., 2014). Moreover, symptom heterogeneity may contribute to these disparate findings as reduced white matter integrity has been found to relate to depression subtypes (Cullen et al., 2019) as well as subthreshold depression (Vulser et al., 2018). It therefore remains unclear whether reduced white matter integrity is a hallmark of early depression pathophysiology during adolescence.

There is a significant degree of individual difference in symptom presentation and impairment in adolescents experiencing depressive symptoms, especially regarding the social context in which these difficulties manifest (e.g., home, school). Discordance between child and parent reports of psychopathology has been well documented (Achenbach, 2006; De Los Reyes, 2011) with some research suggesting that parents may under-report youth depressive symptoms compared to youth self-report (Eg et al., 2018). However, given inconsistent findings in the literature (De Los Reyes et al., 2013), a multiple-informant approach, which usually includes the young person and their parents, remains “best-practice” (De Los Reyes et al., 2015; Rausch et al., 2017). Notably, the associated implications of reporter discrepancy in youth psychopathology (Achenbach, 2006) has been understudied in the context of underlying neurobiological associations. We extend this existing work by looking at both parent and child reported symptoms and how these differentially associate with imaging features.

The current study therefore examines early associations between multi-modal structural imaging features and the emergence of MDD and depressive symptoms (DS) from the population-based, demographically diverse, Adolescent Brain Cognitive Development (ABCD) Study, using both caregiver and child reported symptoms (Casey et al., 2018a). The ABCD Study is a population-based longitudinal project that encompasses magnetic resonance

imaging (MRI) data and lifetime assessments of psychiatric disorders in 9-11-year-old US children (N = 8634, mean age = 9·91 years).

3.4 Methods

3.4.1 Participants

Data from the curated annual release 2.0.1 of the Adolescent Brain Cognitive Development (ABCD) Study were used. Participants were recruited from 21 sites across the United States (Garavan et al., 2018a). A total of N = 10,198 children (9-11 years) participated in the baseline assessment, which took place between September 1st 2016 and August 31st 2018. The unrelated participants with quality-controlled brain imaging measures (cortical measures or white matter measures) were included in the analysis (N = 8631, mean age = 9·91, standard deviation = 0·62, 52·3% were male). The study was approved by the National Institute of Mental Health Data Archive, United States (NIMH). Written consent was obtained from all participants. Data was accessed through the NDA data base (https://nda.nih.gov/general-query.html?q=query=featured-datasets:Adolescent%20Brain%20Cognitive%20Development%20Study%20(ABCD); Federal-Wide Assurance: FWA00018101). Further details can be found in Table 3.1. Demographic information for those with missing data can be found in Table S1.

3.4.2 Derived brain structural measures

Brain imaging data were acquired and processed by the ABCD team. A 3-T Siemens Prisma, General Electric 750 or Phillips scanner was used for data acquisition. A unified protocol for the scanning was used to harmonise between sites and scanners. Protocols used for data acquisition and processing are described elsewhere (Casey et al., 2018a). Standard pre-processing and quality check (QC) procedures were conducted according to the ABCD protocol. Participants with excessive head motion or poor data quality were excluded from the curated data release.

Two types of brain structural measures were used in the present study: grey matter cortical and white matter microstructural measures.

Cortical measures were generated using FreeSurfer 5.3 (https://surfer.nmr.mgh.harvard.edu/fswiki/FreeSurferWiki). Four types of cortical measures were used: surface area, thickness, volume and sulcal depth. First, global measures were generated for each cortical measure over the whole brain (see Figure S1). The Desikan-Killiany atlas was then used for parcellation of 34 bilateral regional cortical structures.

White matter microstructural measures included fractional anisotropy (FA) and mean diffusivity (MD). Global measures of FA and MD were generated over the whole brain. The AtlasTract was used to map boundaries of the 14 bilateral and 3 unilateral major tracts (Hagler et al., 2009). FA/MD values were then derived for each of the region.

Data with poor-quality raw T1/DTI scans and low post-processing QC scores were removed. As there were outlying values for white matter microstructural measures, we removed those with global FA and MD values 5 standard deviations from mean (Shen et al., 2017). Further details can be found in Supplementary Information.

3.4.3 Measures for Major Depressive Disorder and depressive symptoms in adolescents

Life-time Major Depressive Disorder (MDD) and depressive symptoms (DS) for children were assessed using a computerised version of the Kiddie Schedule for Affective Disorders and Schizophrenia (K-SADS; Kaufman et al., 1997) The scale included 28 binary items on current and past DS that reached clinical significance (Table S2). Questions were completed by parents and children separately and self-administered. A previous study of the computerised version for scales showed good to high reliability, with AUC = 0.89-1.00 comparing against clinician administered, computerised K-SADS diagnoses version (Townsend et al., 2020). Lifetime measures of MDD and DS were generated by combining reports on current and past symptoms (a positive answer for either current or past were grouped as positive for lifetime depression, and negative answers on both were grouped as negative, see Supplementary

Methods). Questionnaires were completed by both children and by a caregiver independently. A diagnosis of MDD was generated by the ABCD team for both child and caregiver reported symptoms separately (Barch et al., 2018a). We additionally created a measure of DS based on Diagnostic and Statistical Manual of Mental Disorders (DSM-V) criteria for the severity scale of depression (American Psychiatric Association, 2013b). Levels of DS included: 'severe', 'moderate', 'mild' and 'none of the above' (encoded as 3-0, respectively, see Table S2-3, Figure S4 and Supplementary Methods). DS assessed using the Child Behaviour Checklist (CBCL, a Likert-scale measure) based on reports by caregivers were also used to validate these measures (Brasil & Bordin, 2010) (see Supplementary Information). MDD and DS for unrelated children were included in the analysis. Sample sizes for these variables can be found in Table 3.1.

In addition to the DS reported by caregivers and children separately, we looked at the average reports and discrepancies of DS. Average DS was obtained by calculating the mean of DS reported by each caregiver-child pair. Discrepancy was generated by obtaining the absolute values of subtracting caregiver and child reports.

We also sought to control for potential biases introduced by the current mood of caregivers which could confound associations between their rating of depression in the adolescents with the brain structural measures. We used a subscale of DSM-V-oriented items for depressive problems from the Adult Self-Report (ASR) in the Achenbach System of Empirically Based Assessment (Barch et al., 2018a) (Supplementary Methods).

3.4.4 Measures of socio-environmental factors

As exploratory analyses to further understand potential socio-environmental factors relating to differences in caregiver and child report, we also examined absolute discrepancies in these reports with variables from the ABCD sample including cultural, social, family and school environment of children reported by both caregivers and children themselves, see Supplementary Methods and Table S4.

3.4.5 Statistical models

Statistical analyses were performed in Scientific Linux 2.6.32, using R 3.6.1.

Firstly, associations between MDD diagnosis (binary) and number of depressive items reported DS (continuous) in adolescents and brain structural measures were tested using a General Linear Model (GLM, 'glm' function) or Linear Mixed-effect model (LME, 'lme' function) (Pinheiro et al., 2007) in R. For unilateral brain measures, a GLM was used. For bilateral brain measures, an LME model was used with hemisphere set as a repeated measure (Shen et al., 2017). Covariates included age, age^2, sex, ethnicity, study site, recent social deprivation, and additional imaging covariates: head motion (data field: 'fsqc_qu_motion') and hemisphere for the LME models (see Table S5). To further test if current caregivers' mood confounded associations between depression in adolescents and brain structural measures, we added ASR scores of caregivers as a covariate and compared the results with the main model.

The analyses of associations with brain structural measures followed a hierarchical order from global measures at the whole-brain level to individual structures. For cortical measures, this included whole brain cortical volume, mean thickness, total surface area, mean sulcal depth, followed by individual brain regions. For white matter microstructural measures (FA/MD), the global 'g' measures were first tested, followed by individual tracts. The p values were corrected using family-wise error correction with the FDR (false discovery rate) method (Benjamini & Yekutieli, 2001), using the 'p.adjust' function in R. This was applied for each brain measure category and each reporter separately.

In addition to the main models, we conducted analyses on the mean and discrepancy of DS reported by caregivers and children. Average reports and discrepancies for DS was generated for each child-caregiver pair. Results for the associations between the average severity and general/regional brain measures are shown in the Supplementary Information.

We conducted sensitivity analyses to test potential confounding effects of MRI sites, scanner manufacturers and anti-depressant use in the adolescents (methods and results reported in the Supplementary Information).

To examine origins of the discrepancy of reports, we tested which socio-environmental factors were associated with the discrepancy. R function 'kappa2' from package 'irr' was used to estimate Cohen's Kappa for testing agreement between caregiver and child reports (https://www.rdocumentation.org/packages/irr/versions/0.84.1). GLM models were used and covariates kept consistent with models above with the exception of removing imaging covariates (methods reported in Supplementary Information). As these measures were more likely to be independent tests rather than correlated (e.g. brain structural measures), we applied Bonferroni-correction (Abdi, 2007).

The current study adheres to the STROBE reporting guidelines.

3.4.6 Role of funding sources

Our funding sources (Wellcome Trust and Mental Health Research UK) were not involved in the study preparation/design, analysis/interpretation of data, nor in the writing and submission of this report.

3.5 Results

3.5.1 Major Depressive Disorder (MDD), depressive symptoms (DS) and brain measures

Description of individuals meeting criteria for MDD and reporting DS at mild level and above by both parent (MDD: *N*=194 (2.82%), DS: *N*=654 (7.57%) and child (MDD: *N*=180 (2.60%), DS: *N*=687 (7.98%)) are reported in Table 3.1. For caregiver report, youth who had DS of mild, moderate and severe type were 0.69% (*N*=60), 4.47% (*N*=386), and 2.41% (*N*=208), respectively. For child self-report, individuals reporting mild, moderate and severe DS were 0.69% (*N*=59), 4.91% *(N*=423), and 2.38% (*N*=205), respectively.

			N	Age		Sex (% of Male)
				Mean	SD	
Total sample			8634	9.91	0.62	52.3%
MDD	**Reported by caregivers**	**Case**	194	10.02	0.6	53.6%
		Control	6683	9.89	0.62	51.6%
	Reported by children	**Case**	180	9.95	0.63	58.9%
		Control	6744	9.9	0.62	51.4%
DS	**Reported by caregivers**	**Severe**	60	10.1	0.56	51.7%
		Moderate	386	9.91	0.61	54.1%
		Mild	208	9.96	0.62	58.2%
		None of the above	7980	9.91	0.62	52.1%
	Reported by children	**Severe**	59	10	0.61	61.0%
		Moderate	423	9.91	0.62	53.9%
		Mild	205	9.86	0.6	52.7%
		None of the above	7926	9.91	0.62	52.1%

Table 3.1 — Sample sizes and demographic features for MDD and depressive symptoms (DS).

3.5.1.1 Global imaging metrics

Caregiver report: Global results are shown in Figure 3.1 and Supplementary Data 1. MDD diagnosis for the child as reported by caregivers was associated with significantly lower total cortical volume (Cohen's d=-0·022, p=0·013) and global FA (Cohen's d=-0·027, p=7·96×10^{-4}).

Increasing summary measures of DS reported by caregivers were associated with decreased total cortical volume (β=-0·037, p=5·18×10^{-5}), total surface area (β=-0·034, p=1·41×10^{-4}) and global FA (β=-0·023, p=4·06×10^{-3}).

Child report: Associations between imaging measures and the child reported measures are also shown in Figure 3.1. In general, the associations were weaker than the above reports by caregivers, with fewer significant associations (see Figure 3.1 and Supplementary Data 1). MDD diagnosis based on child report was significantly associated with increased cortical sulcal depth (Cohen's d=0·020, p=0·040) and lower global FA (Cohen's d=-0·018, p=0·022), and DS were associated with lower total cortical volume (β=-0·027, p=2·93×10^{-3}), smaller total surface area (β=-0·020, p=0·022) and greater sulcal depth (β=0·023, p=0·023).

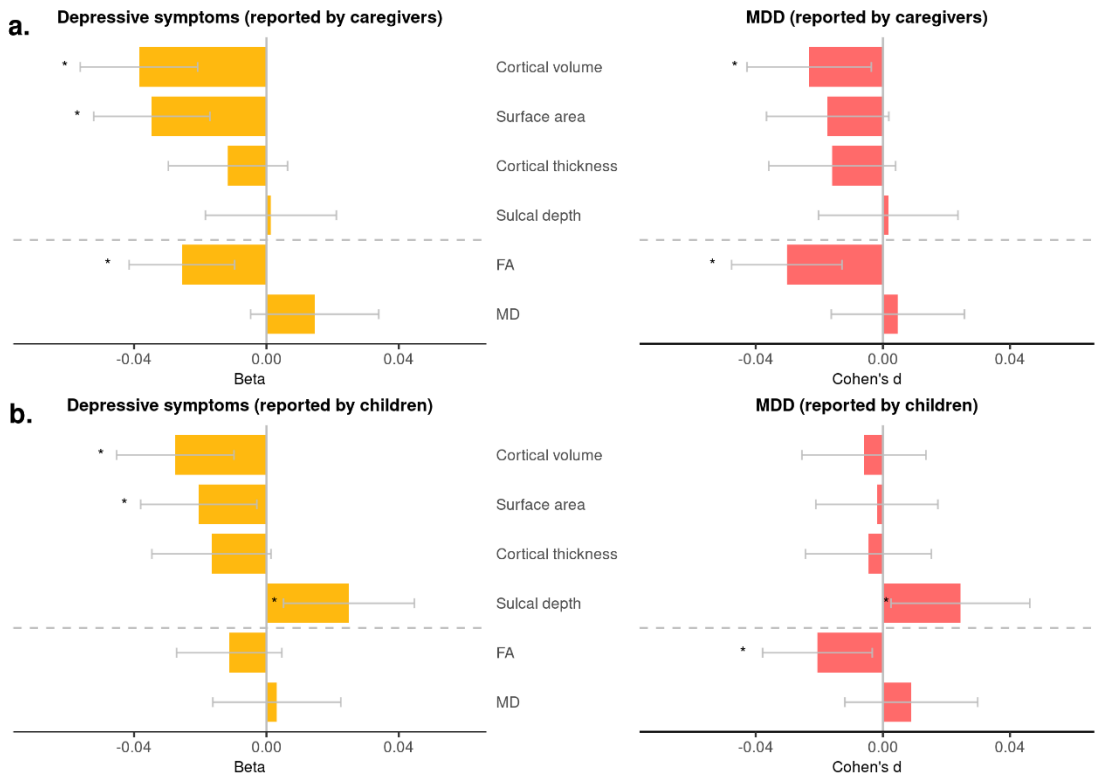


Figure 3.1 — Associations between Major depressive disorder (MDD), depressive symptoms (DS) and general measures of cortical and white-matter structures. X-axes represent standardised effect sizes, and y-axes represent each general measure of brain structure. Error bars represent the 95% confidence interval. Panel (a) shows the results for MDD/depressive symptoms reported by caregivers on children, and panel (b) shows the results for MDD/depressive symptoms reported by children themselves.

3.5.1.2 Regional brain metrics

Caregiver report: Regional results are shown in Figure 3.2 & Figure 3.3 and Supplementary Data 2-5. For cortical measures, MDD diagnosis-based reports by caregivers was associated with reduced volumes of the caudal middle frontal lobe, entorhinal cortex, superior frontal lobe, superior temporal lobe, and temporal pole (Cohen's d range: -0·019 to -0·029, p_{FDR} range: 0·043 to 0·012). Volumes in caudal middle frontal lobe and superior frontal lobe were also associated with DS, along with volumes in other regions that include inferior parietal lobe, middle temporal lobe and precentral gyrus (β range: -0·019 to -0·024, p_{FDR} range: 0·041 to 0·012). Smaller surface area of similar regions was associated with higher DS, which include caudal middle frontal lobe, inferior parietal lobe, middle temporal lobe and superior frontal lobe (β range: -0·020 to -0·024, p_{FDR} range: 0·049 to 0·013). Sulcal depth of rostral anterior cingulate was also associated with higher DS (β=0·029, p_{FDR}=0·004).

For white matter microstructural measures, MDD diagnosis by caregivers was associated with lower FA in uncinate fasciculus, inferior longitudinal fasciculus, inferior-fronto-occipital fasciculus, superior longitudinal fasciculus, temporal superior longitudinal fasciculus, parietal superior longitudinal fasciculus, superior cortico-striate tract and inferior frontal superior frontal cortex (Cohen's d range: -0·016 to -0·036, p_{FDR} range: 0·049 to $3{\cdot}74\times10^{-4}$). Increased DS were associated with reductions of white matter microstructural integrity in the inferior longitudinal fasciculus, superior longitudinal fasciculus, parietal superior longitudinal fasciculus, superior cortico-striate tract (β range: -0·021 to -0·026, p_{FDR} range: 0·045 to 0·021).

Child report: The only significant association between reports by children and individual brain regions was for surface area of transverse temporal and MDD diagnosis (Cohen's d=0·027, p_{FDR}=0·019). No significant association was found for white matter microstructural measures (p_{FDR}>0·11).

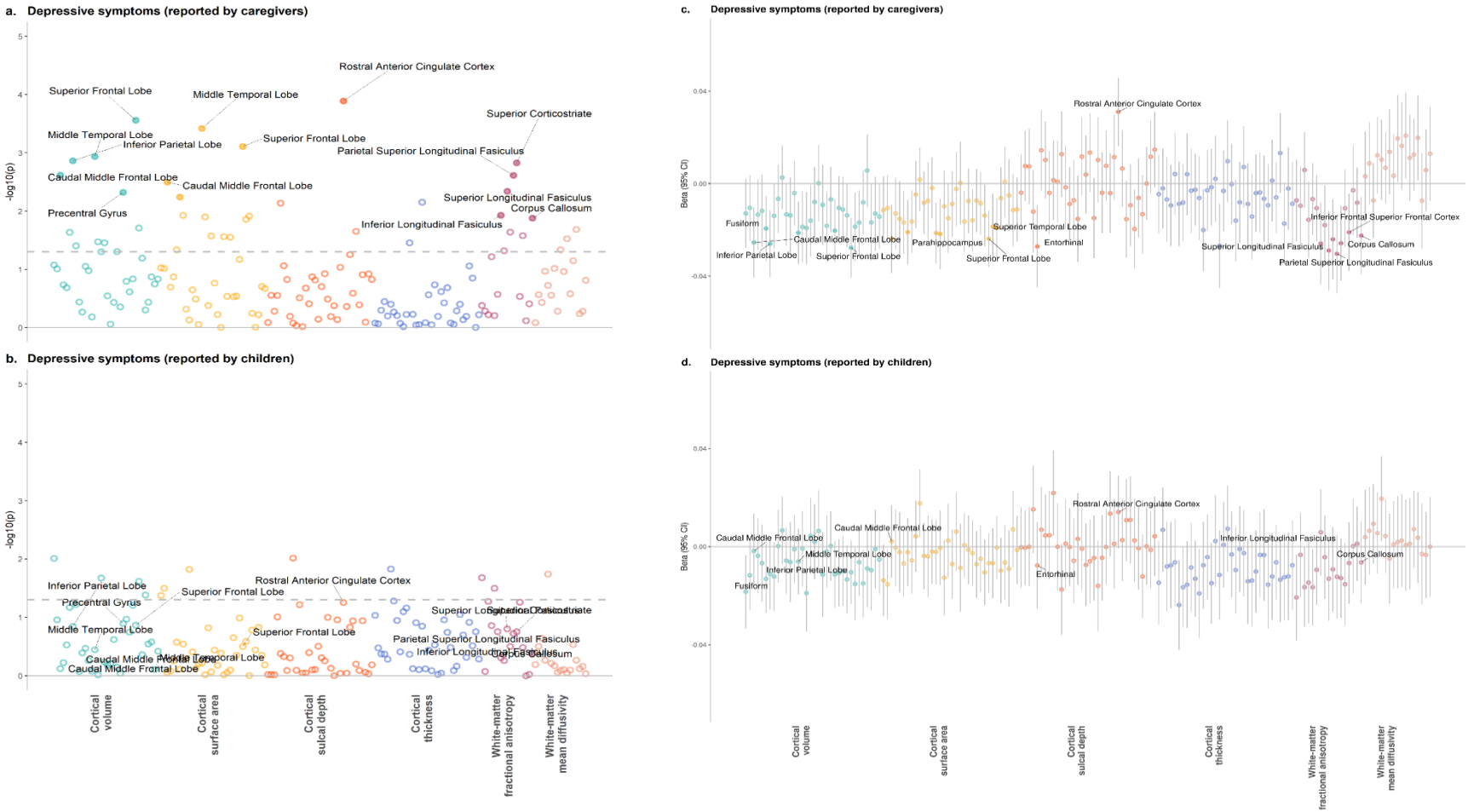


Figure 3.2 — P-value plots for associations between depressive symptoms (DS) and measures for regional brain regions. X axes represent individual brain structural measures, and y axes represent -log10 transformed p-values. Panels (a) and (b) present the p-value statistics for DS reported by caregivers on children and for symptoms reported by children themselves, respective. Panels (c) and (d) show the standardised regression coefficients and 95% confidence intervals for DS reported by caregivers on children and for symptoms reported by children themselves, respectively. Solid dots represent variables significantly associated with DS after FDR-correction. For clarity, threshold for nominal significance before FDR-correction is shown as the grey dashed line in panels (a) and (b).

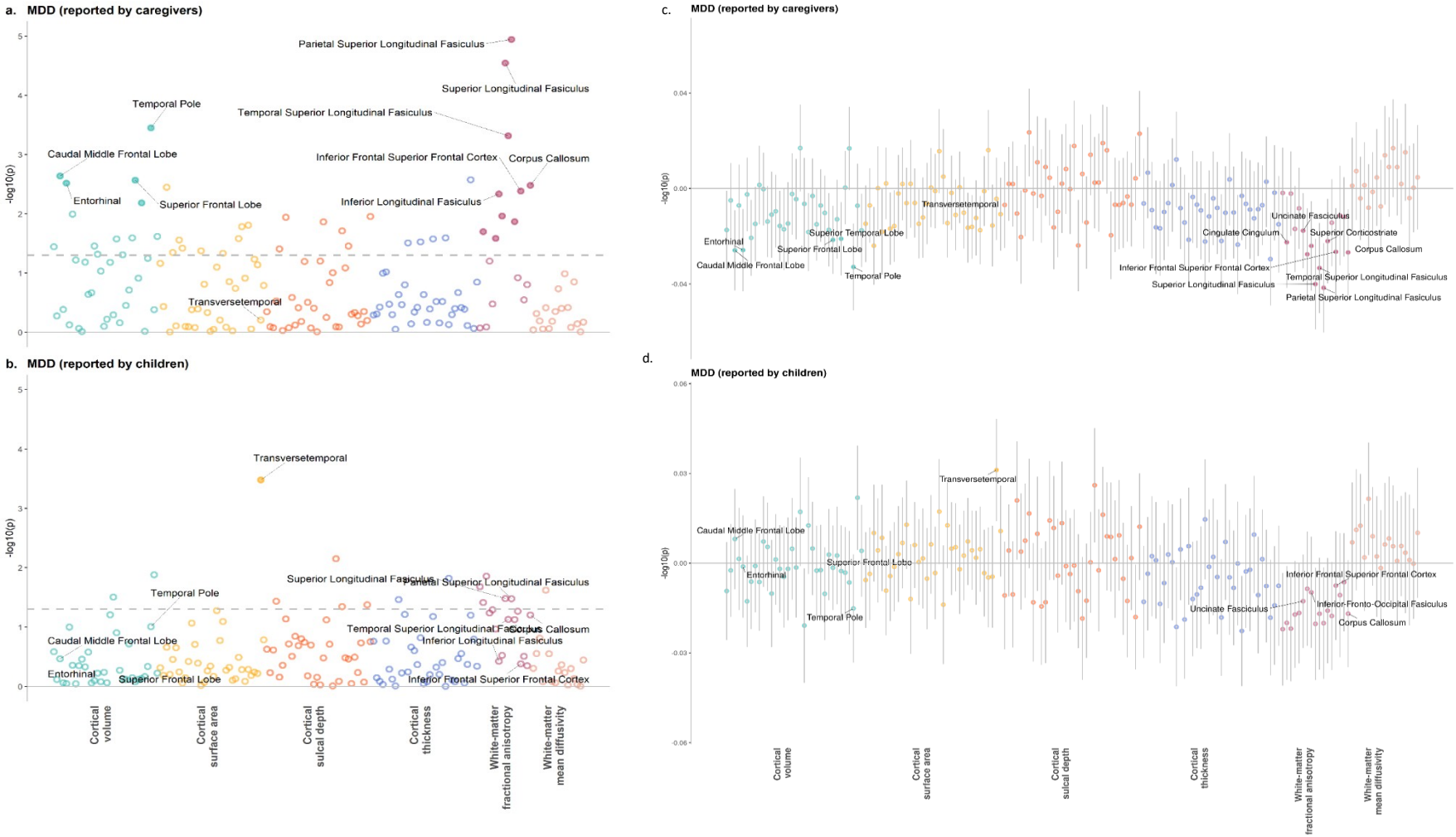


Figure 3.3 — P-value plots for associations between Major depressive disorder (MDD) and measures for single brain regions. X axes represent individual brain structural measures, and y axes represent -log10 transformed p-values. Panels (a) and (b) present the p-value statistics for MDD reported by caregivers on children and for symptoms reported by children themselves, respective. Panels (c) and (d) show the standardised regression coefficients and 95% confidence intervals for MDD reported by caregivers on children and for MDD reported by children themselves, respectively. Solid dots represent variables significantly associated with MDD after FDR-correction. For clarity, threshold for nominal significance before FDR-correction is shown as the grey dashed line in panels (a) and (b).

3.5.1 The association between adolescent MDD, DS and brain measures controlling for mood of caregivers

We conducted an additional sensitivity analysis to test if associations between caregiver reports of MDD/DS in the children remained significant after controlling for measures of current depression in the caregivers. All associations remained significant for global brain measures (Figure S5). Results for single brain regions can be found in Figure S6-7. Overall results with and without controlling for ratings on depressive scale in caregivers showed high correlation (across all association tests between individual brain measures and MDD/depression symptoms reported by children and caregivers, r = 0.996 for standardised effect sizes, r = 0.984 for p-values, see Figure S8). For those associations that were significant without controlling for mood of caregivers, all remained in the same direction and 90.3% remained significant after FDR-correction.

Additional sensitivity analyses indicated results remained significant after controlling for medication (in child) and were consistent across sites/scanner (see Supplementary Information, Figures S9-16). Findings also remained robust (in terms of comparison to the other caregiver report results) after controlling for the magnitude of reporter discrepancy (see Figures S17-19).

3.5.2 Discrepancy between caregiver and child report of depressive symptoms and associations with socio-environmental factors

A significant but low agreement of DS was observed between child and caregiver reports of depression in the child (unweighted Cohen's Kappa=0.06, $p=5{\cdot}29\times10^{-12}$). See Figure S4. Among the caregiver-child pairs, 92 pairs showed large discrepancy (one reported severe DS and the other none), 705 showed moderate discrepancy (discrepancy = two levels) and 7802 pairs showed low or no discrepancy (discrepancy <= one level, contains 7408 pairs that both reported DS lower than mild). Additional analyses that examined reporter discrepancy across different K-SADS items indicated a heterogenous pattern of reporter discrepancy across the entire range of diagnostic items. See Figure S20.

Greater discrepancy in caregiver and child report of DS was positively associated with sleep disturbance and family conflict reported by caregivers on children, and family conflict and school disengagement reported by children (β range: 0·060 to 0·169, $p_{Bonferroni}$ range: $2{\cdot}31\times10^{-5}$ to $5{\cdot}26\times10^{-18}$). Greater agreement, reflected by negative associations, was found with increased neighbourhood safety and prosocial behaviour of children reported by caregivers, as well as acceptance by caregiver/parent and secondary caregiver, school environment, school involvement and caregiver/parent monitoring reported by children (β range: -0·048 to -0·096, $p_{Bonferroni}$ range: $3{\cdot}86\times10^{-7}$ to $9{\cdot}15\times10^{-56}$). See Figure 3.4.

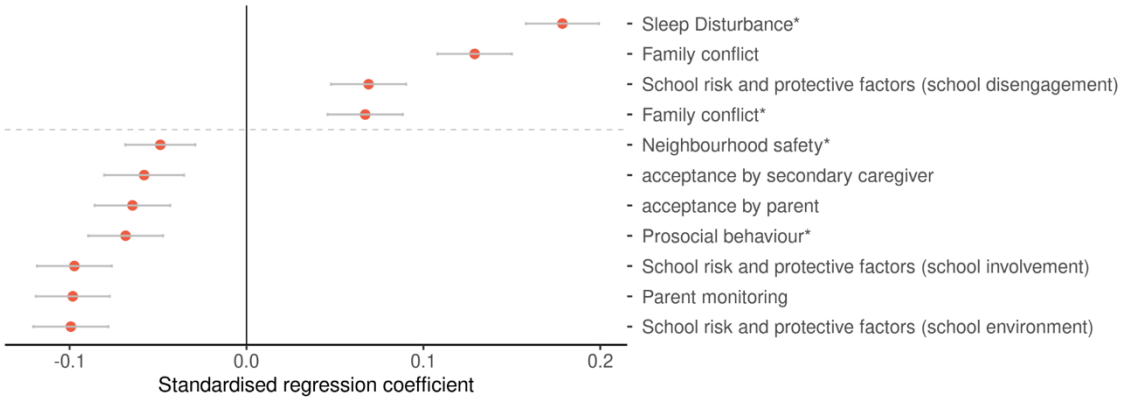


Figure 3.4 — Associations between socio-environmental factors and absolute discrepancies of caregiver and child reports on depressive symptoms (DS). Variables marked with an asterisk are caregiver reports and the rest are child reports. The x-axis shows the standardised regression coefficients. The y-axis shows the variables significantly associated with absolute discrepancies of caregiver and child reports ($p_{bonferroni}<0.05$). Error bars represent 95% confidence intervals. A positive regression coefficient represents a positive relationship between the given trait and the absolute discrepancy reported by caregivers and children, and a negative regression coefficient represent a relationship between the trait and less discordance between caregiver and child reports.

3.6 Discussion

The current study leverages the largest available sample to report brain structural differences and their association with caregiver and child reports of depression in early adolescence. We demonstrated that MDD and DS (as reported by caregivers) were associated with similar imaging findings as seen in adult samples, including reduced global cortical volume and global fractional anisotropy (FA) (MDD: Cohen's d range: -0·022 to -0·027, DS: β range: -0·029 to -0·057). Our findings also suggest that surface area differences, which have been less consistently reported in adult studies, may be a feature of depressive symptoms in adolescents. Reports of depression in children given by caregivers consistently demonstrated stronger associations with cortical structure and white matter microstructure compared to child report. Finally, reporter discrepancy was positively associated with family conflict and school disengagement (β range: 0·060 to 0·169). Higher levels of prosocial behaviour in both school and family environments were linked to lower reporter discordance (β range: -0·048 to -0·096). Sensitivity analyses demonstrated that our results were not related to potential confounders such as anti-depressant medication use (child and caregiver) and scanner and site differences.

Along with the global cortical and white matter differences described above, we also report regionally reduced white matter microstructural integrity in fronto-limbic circuits such as the superior longitudinal fasciculus and cortico-striate tract. This is consistent with previous large scale, population-based studies in adult MDD suggesting that aberrant patterns of white-matter microstructure are present at the early stages of the disease (Shen et al., 2017). Reduced microstructural integrity in these association fibres has been previously found to be related to compromised cognitive control, which may underpin clinical features of depression (Cox et al., 2016; Holleran et al., 2020; Shen et al., 2017).

The present investigation also found reduced cortical surface area in adolescents with depression, globally and regionally, including a diffuse pattern of localised surface area deficits (MDD: Cohen's d=-0·021, DS: β=-0·066), but not in cortical thickness. Notably, unlike the overall pattern of results, surface area reductions are less commonly reported in adult depression compared to cortical thickness reductions (Schmaal et al., 2017). This therefore

implies that surface area reductions may be specifically related to the onset and risk factors of depression in early life stages. Similar findings were also shown in another cohort study looking at older adolescents by Schmaal et al. (2017). While not as commonly reported in adult populations as cortical thickness, surface area reductions have shown to be genetically correlated with MDD (Grasby et al., 2020) and are associated with early risk factors of MDD, such as early life trauma and low birth weight (Opel et al., 2019; Skranes et al., 2013). Regions that demonstrate cortical surface area abnormalities, such as the precentral gyrus, inferior parietal gyrus, and the superior frontal gyrus may be more vulnerable to the effects of delayed maturation in adolescent depression due to reduced synaptic pruning and dendritic growth over this period (Amlien et al., 2016; Wierenga et al., 2014). Longitudinal studies are needed to understand the origin and development of depression-related, surface area brain features during adolescence and across the life course.

Similarities were observed in the association between depression and brain measures across caregiver and child reports. For example, DS from both reports was significantly associated with global cortical volume (caregiver: β= -0·037; child: β= -0·027) and surface area (caregiver: β= -0·034; child: β= -0·020), while MDD status was associated with decreased whole brain FA (caregiver: β= -0·030; child: Cohen's d= -0·021). The effect sizes were similar to those found in adults (Shen et al., 2019a). However, the current study also revealed that reports of depression by caregivers on adolescents demonstrated stronger and more numerous associations with brain structural measures than by adolescent self-report. These were not biased by medication or by current mood of the caregivers themselves. A difficulty in the diagnosis of depression in adolescents is the integration of reports from both caregivers and adolescents (Achenbach, 2006; De Los Reyes, 2011). Although we found agreement between caregiver and child reports across individual depressive items, there were indications of important differences. In line with previous work (Blakemore, 2008; De Los Reyes, 2011; Lewis et al., 2012), internalising and somatic type symptoms (e.g., self-esteem, guilt) were more commonly reported by child than caregiver, while decreased concentration and functional impairments were reported more by caregiver than child. Given the stronger neuroimaging associations found for caregiver report of depression, we consider that cognitive and

functional impairments may be more strongly connected with these early neurobiological changes — these hypotheses should be tested in future work.

Depressive symptoms reported by caregivers and children showed significant but low correlation, and caregiver's report showed greater associations with brain structural measures. Our findings demonstrate that divergences in origins of reporting relate to environmental and societal factors such as family conflict and social cohesion (Kelly et al., 2016). These findings reveal the importance of a supportive environment in defining caregiver-child reporter differences; factors such as child-perceived parental support and acceptance also imply secure attachment styles (Chorot et al., 2017). It is possible that contextual associations between environmental factors and reporter discrepancy may be associated with developmental processes specific to adolescents. Therefore, whilst we cannot completely exclude possible contributions of broader socio-environmental factors, we consider it unlikely that current socio-economic status was driving our main neuroimaging findings, as we have controlled for these in our main analysis. Future longitudinal work should examine the neurobiological consequences of these external societal factors to better understand their role in the origins of the disorder, as well as the potential for environmental intervention.

Although this study benefits from the large imaging sample size, there are limitations. The ABCD cohort is currently a cross-sectional sample. Longitudinal research is needed to facilitate investigating causal effects in these relationships and to inform case-control differences in developmental courses. Further, MDD diagnosis was unavailable in the current data release (2.0.1) for ~20% of participants due to the inclusion of subclinical participants that did not reach criteria for case or control categorisation (as conducted by ABCD study team). Additional analysis however suggests minimal bias between individuals with and without this missing data (see Table S1 in the Supplementary Information for further detail). Missing data will remain a challenge for community-based population cohorts like ABCD and its treatment will warrant important consideration going forward. While the current study uses both a binary and continuous measure of depression, depressive symptoms are notably highly heterogeneous (Fried & Nesse, 2015). This heterogeneity can have pronounced research and

clinical implications; for example, individual symptoms may differentially impact impairment of psychosocial functioning (Fried & Nesse, 2015) and distinctive patterns for longitudinal trajectories of individual symptoms may have heterogeneous underlying neurobiological mechanisms. Future work should examine depressive symptom heterogeneity in the context of brain structure especially during adolescence when subclinical symptoms may manifest, and uncertainties exist around subsequent formal diagnoses.

Although we appreciate the above determinants of heterogeneity for depression, it is important to focus on the neurobiological associations directly linked with the overall diagnosis and severity as a first step, given that disease prediction using neuroimaging phenotypes are predominantly trained and investigated in adult samples (Jahanshad et al., 2013; Thompson, 2019; Thompson et al., 2014). The present findings on the early origins of depression showed distinctive patterns compared to results from adults, which provides strong rationale for separating investigations on diagnosis for adolescent depression, as well as its prediction and treatment. Further, the current findings were generally robust against influence from comorbidity. However, some associations, for example, those found in general cortical grey matter measures, attenuated after controlling for comorbidity. Reasons for this may include shared genetic and environmental risk factors between major psychiatric disorders (Sullivan & Geschwind, 2019). Future studies using genetic and epigenetic data may be able to interrogate cross-disorder associations more directly.

Small effect sizes found for the associations in the present study are likely to be contributed by the heterogeneity of disease manifestations and presentation of subtypes. Small effect sizes are a challenge in large neuroimaging research due to the small amount of variance explained by each individual variable (Dick et al., 2021; Milham et al., 2017). However, big data research also allows for the identification of subtle effects, and neurobiological associations with depression are indeed consistently small in large-sample studies (Schmaal et al., 2017, 2020; Shen et al., 2017, 2019a; van Velzen et al., 2020). These subtle effects may not be statistically detectable in small-scale studies, which also have the caveat of potentially inflated effect sizes due to sample selection bias (Milham et al., 2017). However, the advent of machine learning techniques that examine multiple neuroimaging variables simultaneously

in large multi-site studies holds promise of a move towards the identification of clinically relevant neuroimaging disease markers (Nunes et al., 2020). A further limitation is the young age of the adolescents in the current sample (aged 9-11). It is likely that the cascade of neurobiological changes associated with the onset of puberty may have a further significant impact on the neural circuits implicated in depression (Dahl et al., 2018; Pfeifer & Allen, 2021). Future research is needed to explore any interaction effects between pubertal development, brain measures, and depression in adolescents.

Our findings demonstrate similarities between adult and adolescent imaging features of depression which collectively suggest that cortical and white matter microstructural abnormalities are present early in the disease course of depression and that some of these may extend throughout the lifespan (Schmaal et al., 2017). We demonstrate that these depression-related imaging features are not related to medication in this early adolescent sample. Our results also show evidence of decreased surface area, which may imply an adolescent-specific vulnerability. Investigating the origins of these differences may further the understanding of the aetiology of depression over this highly sensitive neurodevelopmental period and thus, help identify at-risk youth. Future longitudinal studies may further inform causal relationships between depression during adolescence and brain structural development.

Contributors

XS, NM, LR and HCW were responsible for the project conceptualisation, methodology and validation. XS and MJA carried out the data curation. XS and HCW were responsible for the decision to submit. XS was responsible for formal analysis. XS and NM were responsible for writing the original draft and visualisation. XS, NM, SYC, MCB, SML, LR, HCW reviewed versions of the manuscript. XS, NM, LR and HCW were responsible for manuscript editing and review. AMM, SML and HCW were responsible for supervision, project administration, resources, and funding acquisition. XS, NM, MCB, MJA, LR, AMM and HCW had full access to the raw data used in this study.

Declaration of interests

AMM receives two separate Wellcome Trust awards (104036/Z/14/Z and 220857/Z/20/Z), a Sackler trust grant, illumina and Janssen speaker fees. HCW is a co-recipient of a Wellcome Trust award (104036/Z/14/Z). NM is the recipient of a Mental Health Research UK PhD studentship. There are no conflicts of interest declared by other authors.

Data sharing statement

Data used in the preparation of this article were obtained from the Adolescent Brain Cognitive Development (ABCD) Study® (https://abcdstudy.org), held in the NIMH Data Archive (NDA). A full list of supporters is available at https://abcdstudy.org/federal-partners.html. Qualified researchers can request access to ABCD shared data through the NDA portal. Scripts for the analyses in this project can be found on this Github repository: https://github.com/xshen796/ABCD_MDD_brain.

Acknowledgements

The authors wish to thank all the families who participate in the ABCD Study and all of the ABCD Study staff who make this work possible. We would also like to thank our funders, Wellcome Trust and Mental Health Research UK.

For the purposes of open access, the authors have applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Data used in the preparation of this article were 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 U01DA041048, U01DA050989, U01DA051016, U01DA041022, U01DA051018, U01DA051037, U01DA050987, U01DA041174, U01DA041106, U01DA041117, U01DA041028, U01DA041134, U01DA050988, U01DA051039, U01DA041156, U01DA041025, U01DA041120, U01DA051038, U01DA041148, U01DA041093, U01DA041089, U24DA041123, U24DA041147. A full list of supporters is available at https://abcdstudy.org/federal-partners.html. A listing of participating sites and a complete listing of the study investigators can be found at https://abcdstudy.org/consortium_members/. ABCD consortium investigators designed and implemented the study and/or provided data but did not necessarily participate in the analysis or writing of this report. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH or ABCD consortium investigators. The ABCD data repository grows and changes over time. The ABCD data used in this report came from NDA Study ID: 1260; DOI: 10.15154/1522853.

Funding

Wellcome Trust (References: 104036/Z/14/Z and 220857/Z/20/Z) and the Medical Research Council (MRC, Reference: MC_PC_17209).

3.7 Chapter conclusion

Before exploring how other biological and social factors relate to the adolescent brain and depression, we needed to trace the roots of depression in the early adolescent brain. The work outlined in this chapter directly addressed this aim and laid a strong foundation upon which to base the work undertaken in Chapter 4. Following the publication of our paper, there are a few important considerations to highlight:

Firstly, we interpret lower FA values as representing reduced white matter microstructural integrity. As discussed in Chapter 1, this interpretation is increasingly being regarded as overly simplistic due to the complex biophysical mechanisms that give rise to DTI measures. Thus, we have avoided using the term "integrity" when discussing white matter microstructural measures in the remainder of this thesis.

Secondly, in November 2021, the ABCD team reported in the 4.0 release notes that an error had been discovered with the algorithm used to calculate the K-SADS MDD diagnosis. Specifically, the algorithm used did not include impairment in the diagnostic criteria which will likely have led to the overestimation of MDD diagnoses in all ABCD data releases to date. The paper that comprises Chapter 3 had already been published when this error was reported. Given that the noted MDD issue has yet to be fixed at the time of writing (and is not expected to be until release 5.0, due in early to mid 2023), we were not able to re-run our analyses to specifically test the extent to which this data error may have influenced our findings. We have made the code for all analyses in this chapter publicly available and encourage other research teams to test the replicability of our findings once this data error has been fixed. However, our study used both a binary (MDD) and continuous (depressive symptom severity) measure of depression, the latter of which was not affected by the algorithm error. Importantly, we found similar brain structural associations across both depression indices and therefore, we feel that this data error is unlikely to have impacted the main findings of the study.

4 The Role of Brain Structure in the Association Between Pubertal Timing and Depression Risk

4.1 Chapter introduction

Our findings from Chapter 3 suggest that brain structural alterations exist early in the disease course of depression. Although adolescence is characterised by significant brain morphological changes, beyond age related changes, pubertal maturation is also associated with brain structural changes. Importantly, youth that begin puberty ahead of their peers (i.e., have earlier pubertal timing) are at an increased risk for depression during adolescence. However, the role that brain structure may play in the relationship between pubertal timing and depression risk is not well understood.

In this chapter, I used data from the ABCD Study to undertake pre-registered analyses to first test whether earlier pubertal timing measured when youth were aged 10-11 years is associated with increased depressive symptoms two years later, when youth were aged 12-13 years. I then tested whether certain cortical, subcortical and white-matter microstructural measures, that were identified via pilot analyses, mediated this association. This study was conducted as a registered report, whereby the Stage 1 manuscript, comprising the Introduction, Methods, Data Analysis Plan, and Pilot Analyses, was submitted for review to *Developmental Cognitive Neuroscience* in December 2021. Following peer review and revision, our study was awarded an "in-principle acceptance" in August 2022, after which I was able to begin the main analysis and interpret our findings (Stage 2). I hope that the registered report format of this work, and the extensive detail given on the study design, materials and analyses will increase the replicability and reproducibility of this research.

Our registered report is included in full in this thesis chapter and was submitted for Stage 2 review at *Developmental Cognitive Neuroscience* in December 2022. Supplementary

Information and Supplementary Data for this registered report/chapter is included in Appendix 2.

The role of brain structure in the association between pubertal timing and depression risk in an early adolescent sample (the ABCD Study®): A registered report

Authors: Niamh MacSweeney[1*], Judith Allardyce[1], Amelia Edmondson-Stait[1], Xueyi Shen[1], Hannah Casey[1], Stella W. Y. Chan[2], Breda Cullen[3], Rebecca M. Reynolds[4], Sophia Frangou[5,6], Alex S. F. Kwong[1,7,8], Stephen M. Lawrie[1], Liana Romaniuk[1], Heather C. Whalley[1]

[1] Division of Psychiatry, University of Edinburgh, Edinburgh, United Kingdom

[2] School of Psychology and Clinical Language Sciences, University of Reading

[3] Institute of Health and Wellbeing, University of Glasgow, Glasgow, United Kingdom

[4] University of Edinburgh/British Heart Foundation Centre for Cardiovascular Science, Queen's Medical Research Institute, Edinburgh, UK

[5] Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, USA

[6] Department of Psychiatry, Djavad Mowafaghian Centre for Brain Health, University of British Columbia, Vancouver, BC, Canada

[7] MRC Integrative Epidemiology Unit at the University of Bristol, United Kingdom

[8] Population Health Sciences, Bristol Medical School, University of Bristol, UK

*Corresponding author:

Niamh MacSweeney

E-Mail: niamh.macsweeney@ed.ac.uk

Address:

University of Edinburgh, Royal Edinburgh Hospital, Morningside Park

Edinburgh EH10 5HF, United Kingdom

4.2 Abstract

Background

Earlier pubertal timing (PT) is associated with higher rates of depressive disorders in adolescence. Neuroimaging studies report brain structural associations with both pubertal timing and depression. However, whether brain structure mediates the relationship between PT and depression remains unclear.

Methods

The current registered report examined associations between PT (indexed via perceived pubertal development), brain structure (cortical and subcortical metrics, and white matter microstructure) and depressive symptoms (DS) in a large sample (N = ~5,000) of young adolescents (aged 9-13 years) from the Adolescent Brain Cognitive Development (ABCD) Study. We used three waves of follow-up data when the youth were aged 10-11 years, 11-12 years, and 12-13 years, respectively. We used generalised linear-mixed models (H1) and structural equation modelling (H2 & H3) to test our hypotheses.

Hypotheses

We hypothesised that earlier PT at Year 1 would be associated with increased DS at Year 3 (H1), and that this relationship would be mediated by global (H2a-b) and regional (H3a-g) brain structural measures at Year 2. Global measures included reduced cortical volume, thickness, surface area and sulcal depth. Regional measures included reduced cortical thickness and volume in temporal and fronto-parietal areas, increased cortical volume in the ventral diencephalon, increased sulcal depth in the pars orbitalis, and reduced fractional anisotropy in the cortico-striatal tract and corpus callosum. These regions of interest were informed by our pilot analyses using baseline ABCD data when the youth were aged 9-10 years.

Results

Earlier pubertal timing was associated with increased depressive symptoms two years later. The magnitude of effect was stronger in female youth and the association remained

significant when controlling for parental depression, family income, and BMI in females but not in male youth. On the other hand, our hypothesised brain structural measures did not mediate the association between earlier pubertal timing and later depressive symptoms.

Conclusion

The present results demonstrate that youth, particularly females, who begin puberty ahead of their peers are at an increased risk for adolescent-onset depression. Future work should explore additional biological and socio-environmental factors that may affect this association so that we can identify targets for intervention to help these at-risk youth.

4.3 Introduction

Adolescence is a period of increased vulnerability to mental health conditions, particularly internalising difficulties such as depression (Malhi & Mann, 2018; Thapar et al., 2012). Earlier-onset of depression is associated with a more severe illness course (Thapar et al., 2012) and with a range of psychosocial and physical difficulties which perpetuate across the lifespan (Clayborne et al., 2019; Fergusson & Woodward, 2002; Naicker et al., 2013). Given the emergence of depression during the adolescent period, the role that pubertal development may play in this heightened vulnerability has garnered increasing attention (Graber, 2013; Hamlat et al., 2019; Pfeifer & Allen, 2021; Ullsperger & Nikolas, 2017). Earlier pubertal timing has been associated with increased risk for depression in both males and females (Graber, 2013; Hamlat et al., 2020; Mendle et al., 2010; Ullsperger & Nikolas, 2017). Further, genetic studies have found that earlier age of menarche is implicated in depression (Howard et al., 2019). Adolescence is also a time of immense neurobiological change (Mills et al., 2016; Tamnes et al., 2017; Vijayakumar et al., 2016) and brain structural differences have been found in both adults (Schmaal et al., 2017, 2020; Shen et al., 2017, 2019a) and adolescents (Schmaal et al., 2017; Shen et al., 2021) with depression. However, the role that neural mechanisms may play in the relationship between pubertal timing and depression risk is not well understood. Here, we therefore examine whether brain structure mediates the association between pubertal timing and depressive symptoms in a large sample of adolescents (aged 9-13 years) from the Adolescent Brain Cognitive Development (ABCD) Study®.

4.3.1 Defining and measuring pubertal timing

Pubertal timing measures pubertal development relative to same-age, same-sex peers, such that an individual can be categorised as developing ahead (early), in-line (on-time) or after (late) their peers. Measures of pubertal timing are most often derived from methods used to assess pubertal status, such as the Pubertal Development Scale (PDS; Petersen et al., 1988) and Tanner Stage Line Drawings (TS; Marshall & Tanner, 1969, 1970). However, other measures used include age of menarche and sex hormone measures (Goddings et al., 2019; Ullsperger & Nikolas, 2017). Pubertal maturation as assessed by the PDS and TS focuses on

the development of secondary sex characteristics (e.g., testicular, breast, and pubic hair development), which stem directly from changes in sex hormones. These measures are completed by a clinician (TS), or via self- (or parent-) report (PDS/TS). Most often, a pubertal timing score is derived by regressing a pubertal status score on chronological age to calculate a sex-specific residual for each person (Barendse et al., 2021; Dorn & Biro, 2011; Mendle et al., 2010; Ullsperger & Nikolas, 2017). The residual score represents how much an individual's pubertal development deviates from what is expected for their age with positive and negative scores indicating earlier and later timing, respectively. It is worth noting that pubertal development consists of two phases: adrenarche, usually occurring between the ages 6-9 years (Biro et al., 2014), and gonadarche, which typically takes place between the ages 9-14 years for females and 10-15 years for males.

4.3.2 Pubertal timing and psychopathology

Historically, research on pubertal timing effects on psychopathology has highlighted that youth, particularly females (Graber, 2013; Hamlat et al., 2019; Hamilton et al., 2014), who undergo puberty earlier than their peers are at an increased risk for psychopathology (Conley et al., 2012; Ge & Natsuaki, 2009; Hamilton et al., 2014). However, a recent meta-analysis suggests that earlier pubertal timing is detrimental to both sexes and that later pubertal timing is not significantly associated with psychopathology (Ullsperger & Nikolas, 2017). Although a number of conceptual models (Brooks-Gunn et al., 1985, 1994; Petersen et al., 1988) have been proposed to explain the association between earlier pubertal timing and increased risk for psychopathology, the "maturation disparity hypothesis" (Brooks-Gunn et al., 1985; Ge et al., 2001; Ge & Natsuaki, 2009), has received the most empirical support (Graber, 2013; Ullsperger & Nikolas, 2017). The maturation disparity hypothesis posits that early developing youth experience psychological distress due to an incongruity between their accelerated physical development and asynchronous maturation of cognitive and emotional brain regions.

Importantly, the psychological and social changes that occur during adolescence such as heightened self-awareness and social sensitivity (Blakemore & Mills, 2014; Pfeifer & Peake, 2012), increased risk-taking behaviour and impulsivity (Bjork & Pardini, 2015; Defoe et al.,

2015; Romer, 2010) as well as greater peer influence on behaviour (Albert et al., 2013; Blakemore, 2018; Knoll et al., 2015) are likely underpinned by the distinct developmental trajectories of temporal and limbic areas (involved in emotion and reward processing) and prefrontal regions (involved in cognitive control) (Albert et al., 2013; Casey et al., 2008; Mills et al., 2014; Steinberg, 2008). It has been postulated that earlier developing youth therefore experience a greater discordance in the mismatch between the earlier developing affective regions and the more protracted development of cognitive regions (Ge & Natsuaki, 2009; Ullsperger & Nikolas, 2017), which may place them at an increased risk for mental health difficulties. Given that the onset of puberty is about 18 months earlier for females than males, this maturation disparity hypothesis may also explain the preponderance of depression (2:1) in females compared to males from adolescence onwards (Conley et al., 2012; Hankin, 2006, 2015; Hankin & Abramson, 1999). Although the maturation disparity hypothesis best accounts for the extant findings, a more nuanced model that considers the role of biological and psychosocial factors as potential mediators or moderators in the association between earlier pubertal timing and increased risk for psychopathology is needed.

4.3.3 Pubertal timing and brain structure

Research on typical neurodevelopment demonstrates a reduction in grey matter volume and cortical thickness during adolescence, while cortical surface area increases throughout childhood before plateauing by mid-adolescence, and slightly decreasing thereafter (Bethlehem et al., 2022; Ducharme et al., 2016; Mills et al., 2016; Vijayakumar et al., 2016; Wierenga et al., 2014). These patterns of human brain development, were recently evidenced in a collaborative paper involving >100 studies and >123,000 MRI scans (Bethlehem et al., 2022), which is the largest aggregated sample to date. However, research has also shown that pubertal development impacts neurodevelopment beyond age-related changes (Vijayakumar et al., 2018). For example, a number of studies demonstrate extensive negative associations between pubertal timing (indexed via physical and hormonal measures) and cortical volume and thickness, mainly in regions implicated in cognitive control, decision making, and emotion regulation, such as the prefrontal cortex, anterior cingulate cortex, and the temporal lobe (Koolschijn et al., 2014; Pfefferbaum et al., 2016). Notably, few studies have examined surface area changes during puberty as surface area is often obscured when investigating volumetric

estimates — a product of cortical surface area and thickness (Vijayakumar et al., 2016). Given that surface area and cortical thickness reflect distinct neurobiological processes (Wierenga et al., 2014) and are genetically independent of each other (Winkler et al., 2010), examining pubertal timing in relation to surface area maturation may reveal novel associations.

Regarding associations between subcortical measures and pubertal development, research has focused on the amygdala, hippocampus, and striatal regions given their role in emotion regulation, reward processing, and decision making (Bhanji & Delgado, 2014; Dalgleish, 2004). A number of cross-sectional and some longitudinal studies have reported that more advanced pubertal maturation is associated with an increase in volume of the amygdala and hippocampus and a decrease in volume in striatal areas (Blanton et al., 2012; Hu et al., 2013; Goddings et al., 2014, 2019). Although these findings provide insight into the role of puberty in subcortical brain development, there is a dearth of research that specifically examines pubertal timing (i.e., pubertal development relative to same-age, same-sex peers) (Goddings et al., 2019) and its association with brain morphological changes (Koolschijn et al., 2014; Neufang et al., 2009; Peper et al., 2009). Further, longitudinal data has shown that there are unique but co-existing age effects that complicate examining the relationship between puberty and structural brain development (Goddings et al., 2019). For example, a recent longitudinal study demonstrated a positive linear association between perceived pubertal maturation (indexed via TS) and the hippocampus, amygdala, caudate and pallidum. However, these associations did not remain significant when age was controlled for (Vijayakumar et al., 2021). Further inconsistencies have emerged in the literature when utilising different measures of pubertal development (Koolschijn et al., 2014; Vijayakumar et al., 2018), and also in studies using large age ranges (Satterthwaite et al., 2014; Urošević et al., 2014). There is also some research suggesting that pituitary gland volume mediates the association between earlier pubertal timing and increased risk for depression in adolescence but more research is needed on this topic (Whittle et al., 2012). Additionally, the current literature does not allow for the identification of clear sex differences in the association between cortical and subcortical structure and pubertal timing, likely owing to the paucity of longitudinal, large-scale research in this area (Herting & Sowell, 2017; Vijayakumar et al., 2018).

There is less research examining the association between pubertal development and white matter microstructure (Goddings et al., 2019) and findings are mixed (Vijayakumar et al., 2018). There is some degree of support for a positive association between pubertal timing and fractional anisotropy (FA; Herting et al., 2012; Peper et al., 2015). However, findings have been inconsistent when considering the relation between pubertal hormones and FA, and between all indices of pubertal development (physical maturation and hormonal measures) and mean diffusivity (MD; Herting et al., 2012; Peper et al., 2009). These discrepancies may be attributed to the various diffusion tensor imaging (DTI) techniques employed and the relatively small sample sizes. Future large-scale neuroimaging research that leverages harmonised protocols and considers the unique and contemporaneous effect of age is needed to disentangle the associations between pubertal timing and white matter microstructure.

Large, population-based research projects, such as the ABCD Study®, directly address limitations of earlier research (e.g., small sample sizes, inconsistent protocols) and will allow us to investigate how brain changes across adolescence are related to developmental outcomes, especially the emergence of mental health difficulties (Casey et al., 2018a). The ABCD Study includes magnetic resonance imaging (MRI) data, assessments of psychiatric disorders, and hormonal and physical puberty measures in 9-10-year-old US children at baseline (N= ~11,800). Our previous work with the ABCD Study® explored the temporal origins of depression-related imaging features and demonstrated that depression ratings in early adolescence were associated with similar cortical and white matter microstructural differences to those seen in adult samples (Shen et al., 2021). These findings suggest that neuroanatomical abnormalities may be present early in the disease course. However, the cascade of neurobiological changes associated with the onset of puberty may have an important role in risk for depression and may allow further mechanistic insight into the origins of these depression-related imaging features (Dahl et al., 2018; Pfeifer & Allen, 2021).

4.3.4 Proposed study

While research has shown that earlier pubertal timing is associated with an increased risk for depression, the underlying neurobiological mechanisms remain unclear. The goal of the present study was to investigate whether brain structure (cortical and subcortical metrics, and white matter microstructural measures) mediates the association between earlier pubertal timing (indexed via perceived physical development) and increased depressive symptoms in a young adolescent sample. We first tested if earlier pubertal timing when youth were aged 10-11 years (Year 1) was associated with higher depressive symptoms two years later when they were aged 12-13 years (Year 3). We then examined whether specific brain structural metrics at Year 2 (identified via our pilot analyses, described in detail in *Pilot Analyses*), mediated the association between earlier pubertal timing and later depressive symptom severity. Given the differences in the average age of puberty-onset for males and females, we ran our models separately for males and females.

Specifically, our key hypotheses were that earlier pubertal timing at Year 1 would be associated with increased depressive symptoms at Year 3 (**H1**), and that this relationship would be mediated by global **(H2a-b)** and regional measures **(H3a-g)** outlined in Table 4.1. These regions of interest were consistent with existing literature on puberty- and depression-related imaging features in adolescence (Schmaal et al., 2017; Shen et al., 2021; Vijayakumar et al., 2018). We did not make formal hypotheses about sex differences in the current study due to inconsistent findings in the literature.

Research Q	Hypothesis	Analysis Test	Effect of Interest	Threshold for evidence
Is earlier pubertal timing associated with later depression?	**H1:** Earlier pubertal timing will be associated with later higher depressive symptoms	Generalised linear mixed effect model	Beta value and p value	ß ≥0.01 and p ≤0.05
Does brain structure mediate the association between earlier pubertal timing and later depression?	Informed by our pilot analyses, the association between earlier pubertal timing and increased depressive symptoms will be mediated by: *Global measures* **H2a**: Reduced global cortical volume, surface area, thickness and sulcal depth **H2b:** Reduced global FA *Regional measures* **H3a:** Reduced cortical thickness in temporal regions, namely, the middle temporal gyrus and insula **H3b:** Reduced cortical thickness in frontal regions namely, the lateral orbito-frontal cortex and middle frontal gyri **H3c:** Reduced cortical volume in temporal regions, namely, middle temporal gyrus and bank of the superior temporal sulcus **H3d:** Reduced cortical volume in fronto-parietal regions, namely, the middle frontal and postcentral gyri **H3e** Reduced FA in the cortico-striatal tract and corpus collosum **H3f:** Increased sulcal depth in the pars orbitalis **H3g:** Increased volume in the ventral diencephalon	Multi-level structural equation model	Indirect effect in mediation model	ß ≥0.01 and p ≤0.05

Table 4.1 — Hypotheses tested in this registered report.

The results of this multi-modal study will inform our understanding of how pubertal timing and brain structure may be associated with depression during adolescence. Undertaking this project as a registered report with shared analytic code applied to an openly available dataset will further increase the replicability and reproducibility of this work.

4.4 Materials and methods

4.4.1 Participants

The data used in the current study were drawn from the ABCD curated annual data release 4.0. and used Year 1, Year 2, and Year 3 follow up data. ABCD participants were recruited from 21 sites across the United States (Garavan et al., 2018a). Approximately N = ~11,800 9–10-year-olds participated in the baseline assessment. We included individuals with quality-controlled pubertal development measures (physical) at Year 1 and quality-controlled brain imaging measures (cortical and subcortical size/metrics, and white matter measures) at Year 2. Missing depression outcome and covariate data were handled using appropriate methods (see *Missing Data* section). This resulted in a final sample of approximately N = ~5,000 individuals, which represents about 50% of the full sample. This smaller sample size can be attributed to the partial follow-up data available at the time of the 4.0 data release (Fall 2021).

To inform hypotheses for the current registered report, specifically the brain regions of interest (ROIs), we conducted pilot analyses using data from the baseline timepoint, when youth were aged 9-10 years (N = 9,339, males = 4802, females = 4537) from data release 4.0. Participants were included in the pilot analyses if they had quality controlled pubertal development, depression, and brain imaging measures. Given that the main analyses did not use any baseline data, we did not anticipate that this decision would significantly impact our findings. Findings from the pilot analyses are reported in the *Pilot Analyses* section.

4.4.2 Measures

All variables (excluding imaging variables) as per the NDA ABCD data dictionary/Data Exploration and Analysis Portal (DEAP) portal field names can be found in Table 4.2 below.

Field name(s)	Description
pds_1_p; pds_2_p; pds_3_p; pds_f4_p; pds_f5b_p	Caregiver: PDS female items, which were summed to generate PDS total score.
pds_1_p; pds_2_p; pds_3_p; pds_m4_p; pds_m5_p.	Caregiver: PDS male items, which were summed to generate PDS total score

cbcl_scr_syn_withdep_r	CBCL withdrawn-depressed syndrome subscale raw score
age_years	Age of child in years
anthro_bmi_calc	Body mass index (BMI) (DEAP field name, non-NDA)
race.6level	6-level derived race variable (white, black, Asian, AIAN/NHPI, other, mixed)
demo_comb_income_p	Household income
asr_scr_depress_r	Parental mood
acs_raked_propensity_score	Imputed raked propensity weight. The raked propensity weight merges the ACS and ABCD data (with missing data imputed), estimates the propensity model, computes and scales/trims the propensity weights and finally rakes the scaled weights to final ACS control totals by age, sex and race/ethnicity.
site_id_l	ABCD study site
mri_info_deviceserialnumber	Scanner ID
dmri_dti_meanmotion	DTI average framewise displacement in mm
rel_family_id	Family ID

Table 4.2 — Name and description of study variables used in this registered report.

Field names are the column names in the original curated ABCD data release or in the DEAP portal.

4.4.2.1 Independent variable — pubertal timing measure

Protocols previously outlined by Cheng et al., (2021) and Herting et al., (2021) were consulted in the preparation of the pubertal timing measures.

The Pubertal Development Scale (PDS) was used to examine the perceived development of secondary sex characteristics such as growth spurts, body hair growth, skin changes, breast development and menarche in girls, and voice changes and growth of facial hair in boys. In line with existing work on puberty measures in the ABCD Study, and given previous research

showing that youth tend to over-report their perceived physical development at younger ages (Schlossberger et al., 1992), caregiver PDS report was utilised instead of child self-report in the current analysis. The PDS includes five-items, and each characteristic is rated on a 4-point scale (1 = no development; 2 = development has barely begun; 3 = development is definitely underway; and 4 = development is complete; except menstruation, which is coded 1 = has not begun, 4 = has begun). Thus, higher scores reflect more advanced pubertal maturation. We did not examine age of menarche in the current analysis due to the small number of females (3% and 13%) in the ABCD sample that have reached this developmental milestone at baseline and Year 1 data collection, respectively (Herting et al., 2021).

In line with existing research on pubertal timing, the PDS total score was regressed on age for girls and boys separately and the standardised residual obtained constituted the continuous measure of pubertal timing (Dorn et al., 2006; Hamilton et al., 2014). Only participants with complete data for 5/5 PDS items were included in the analysis.

Changes to the sample size at each stage of the quality control process can be found in Figure 4.1.

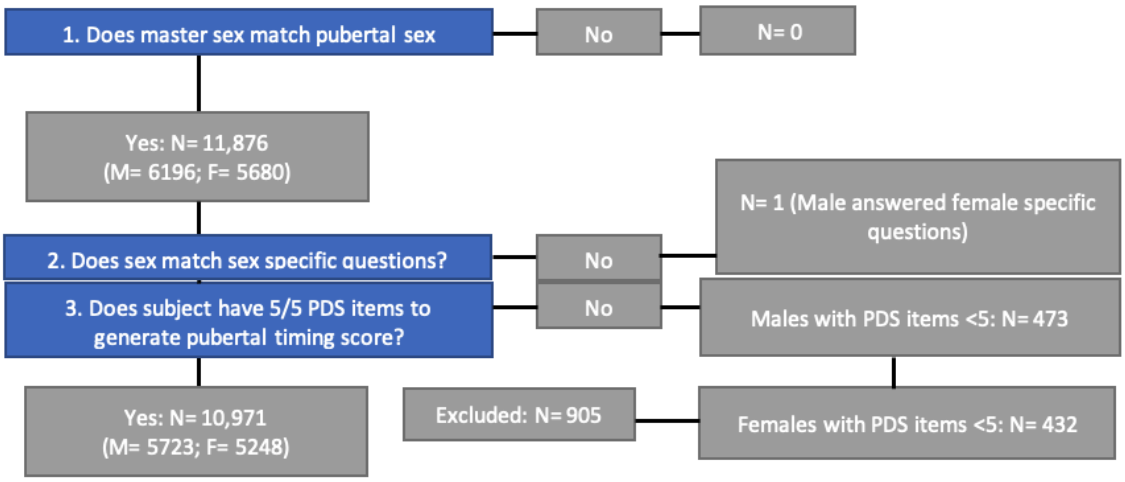


Figure 4.1 — Pubertal Development Scale quality checking decision tree at Year 1 (curated annual release 4.0). Notes: Master sex = Q: Sex of subject, Answer: Male, Female, Other, Not Reported

4.4.2.2 Dependent variable — depressive symptoms

Depressive symptoms (DS) for children were assessed using the Child Behaviour Check List (CBCL) parent report. The CBCL is one of the most widely used measures to examine internalising and externalising difficulties in youth (Achenbach & Rescorla, 2004). It comprises raw scores as well as standardised (T-scores) based on national norms in young people aged 6-18 years. We quantified current depressive symptoms using the CBCL "withdrawn-depressed" syndrome subscale raw scores, which examine depressive symptoms within the past two weeks. Raw scores were chosen over T-scores for our analyses because they reflect all the variation in symptoms that occur in the sample. Due to the substantial percentage of individuals in a normative sample who obtain low scores on the CBCL syndrome subscales, the T-score assignments begin at 50 which means that all individuals in the lowest 50% are assigned a T-score of 50.

4.4.2.3 Hypothesised mediator — brain structural measures

Brain imaging data were acquired and processed by the ABCD team. A 3T Siemens Prisma, General Electric 750 or Phillips scanner was used for data acquisition. A unified protocol for the scanning was used to harmonise between sites and scanners (Casey et al., 2018a).

Protocols used for data acquisition and processing are described elsewhere (Casey et al., 2018a; Hagler et al., 2019). In brief, T1-weighted data was acquired by magnetisation-prepared rapid acquisition gradient echo scans with a resolution of 1×1×1 mm^3, which was used for generating cortical and subcortical structural measures, and diffusion-weighted data was obtained by high angular resolution diffusion tensor imaging (DTI) scans, used for generating white matter microstructural measures.

Imaging data was quality controlled according to recommended QC criteria outlined by ABCD in the 4.0 release notes: "MRI Quality Control (QC) and Recommended Image Inclusion Criteria". ABCD have created a data structure *abcd_imgincl01* that provides modality-specific summary imaging inclusion flags that indicate whether an individual meets all QC criteria for the modality, and are scored as 1 = include, 0 = exclude. These summary variables account for factors such as imaging QC and post-processing (see public release notes for full description). Public release notes are available here: https://nda.nih.gov/ftpDownload?id=JiZtiu3lTujb6V2rEo64cnE&ownerId=kpLEjup90qoGybQOCIG1FPQ&ownerName=collection). We included individuals that met all the recommended inclusion criteria (i.e., score = 1) on the "imgincl_t1w_include" variable for the T1w data and the "imgincl_dmri_include" variable for the DTI data. To account for additional motion artefact in the DTI data, included a measure of mean framewise displacement (variable name: dmri_dti_meanmotion) in our DTI models.

Three types of brain structural measures were used in the present study: grey matter cortical and subcortical metrics, and white matter microstructural measures. The derivation of brain structural measures followed a hierarchical order from global measures at the whole-brain level to individual structures.

Cortical measures were generated using Freesurfer 5.3 (https://surfer.nmr.mgh.harvard.edu/fswiki/FreeSurferWiki). Four types of cortical measures were used: surface area, thickness, volume and sulcal depth. First, global measures were generated for each cortical measure over the whole brain. The Desikan-Killiany atlas was used for parcellation of 34 bilateral cortical structures and 16 bilateral subcortical structures. For

bilateral brain structures, we generated an average measure across the left and right hemisphere to use in our analyses.

White matter microstructural measures included fractional anisotropy (FA) and mean diffusivity (MD). Global measures of FA and MD were generated over the whole brain. The AtlasTrack was used to map boundaries of the 14 bilateral and 3 unilateral major tracts (Hagler et al., 2009). FA/MD values were then derived for each tract.

4.4.3 Covariates

Research examining how puberty is related to developmental processes and outcomes indicates that several factors may shape these associations . Examining variability in these constructs is crucial to understanding their unique contributions to developmental and psychological outcomes (Saragosa-Harris et al., 2022). For example, differences in race/ethnicity and body mass index (BMI) have been associated with early pubertal timing (e.g., earlier age of menarche and onset of breast development) (Biro et al., 2013; Chumlea et al., 2003) and an increased risk for depression (Anderson & Mayes, 2010; Quek et al., 2017). Further, youth raised in families with low socioeconomic status, especially those with significant financial hardship, are at an increased risk for psychopathology (Bradley & Corwyn, 2002; Herting et al., 2021; Peverill et al., 2021) (as evidenced in a meta-analysis of US population-based studies by Peverill et al. 2021). Research has also demonstrated that parental mood can influence the reporting of their child's psychopathology (Maoz et al., 2014).

Although the imaging QC protocol outlined by ABCD excluded participants with excessive head motion across all imaging modalities (i.e., structural and DTI), motion-related confounds have been found to systematically impact structural connectivity measures derived from DTI data (Baum et al., 2018). Children of the same age may exhibit developmental differences in cranial or brain size, which need be considered to determine whether regional differences are independent of global effects (Mills et al., 2016; O'Brien et al., 2011). While there is currently no consensus on whether to use intracranial volume (ICV) or whole brain volume (WBV), some research suggests that WBV may be a more viable measure to use as it has been found to be

more stable across developmental samples than ICV (Mills et al., 2016). The importance of controlling for site effects to account for inter-site variability has been well documented (Feaster et al., 2011). Although ABCD data collection took place at 21 sites, 30 different MRI scanners were used during data collection as some sites had more than one scanner. Therefore, the potential for variability across scanners is also important to consider (Saragosa-Harris et al., 2022). As ABCD has been oversampled for twins and siblings, it is important to account for relatedness between individuals when using the related sample (Gelman & Hill, 2006; Iacono et al., 2018).

Taking these findings together, when modelling the relationships between pubertal timing, brain structure and depressive symptoms, we included the following variables as *fixed effects* in our models: age, race/ethnicity, BMI, household income; parental current mood, and a DTI average framewise displacement measure (for DTI models only). We included family ID and site (for non-imaging models) and scanner ID (for all other models) as *random effects.*

4.4.4 Consideration of outcome neutral conditions

Effect sizes (β values) and FDR (false discovery rate) corrected p values (where appropriate) were the main parameters of interest in the main analyses of the current study. The minimum effect size of interest, $\beta \geq 0.01$ and a $p \leq 0.05$, was considered statistically significant. This was informed by our previous work using the ABCD sample, which examined baseline cross-sectional brain structural associations with depression ratings in adolescence (Shen et al., 2021).

4.5 Data analysis plan

4.5.1 Main analyses

All analyses were conducted in R Version 4.1. and Mplus Version 8.8. Scripts for all analyses can be found at https://github.com/niamhmacsweeney/ABCD_puberty_depression.

The analytic approach comprised two steps: **H1**) examine the associations between pubertal timing and depressive symptoms in adolescents (see Figure 4.2) and **H2&3**) determine whether brain structural measures (identified via pilot analyses, see Pilot analyses section) mediates this association (see Figure 4.3). All hypotheses are outlined in Table 4.1.

We note that regardless of the effect sizes for our first association test (i.e., the total effect), we still conducted the mediation analysis in an attempt to accurately quantify any indirect effects (Agler & De Boeck, 2017). Further, although the models in our pilot analyses were run separately for males and females, and thus generated some varying ROIs, our main analyses used *all* the ROIs identified from both male and female models due to non-hypothesised sex differences in the current study.

When testing our hypotheses, we first ran our base model and if results met our specified threshold for evidence (see Table 4.1), we ran our full model structure. This allowed us to explore whether our main associations were attenuated by the presence of additional covariates.

Hypothesis 1 (H1): Pubertal timing -> depressive symptoms

Independent variable: Pubertal timing at Year 1 (youth aged 10-11 years). This was indexed by the PDS, which is a continuous measure.

Dependent variable: Youth current depressive symptoms at Year 3 (aged 12-13 years), as reported by caregiver. Depressive symptoms were indexed using the CBCL “withdrawn-depressed” syndrome subscale, which is in count data format.

Generalised linear mixed effects models (GLM) were conducted to test the associations, using the 'lmerTest' function in R (Kuznetsova et al., 2017). Models to test H1 are listed in Table 4.3.

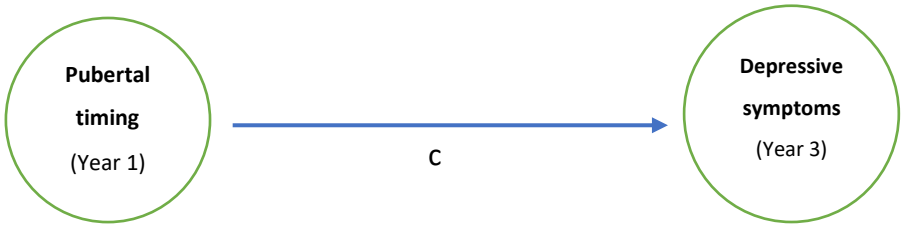


Figure 4.2 — Effect of pubertal timing on depressive symptoms without considering mediation (c: total effect).

Sex specific models	**Covariates (Base model)**	**Covariates (Full model)**
Depressive symptoms ~ pubertal timing	Fixed: age, race/ethnicity Random: family, site	Fixed: age, race/ethnicity, BMI, site, household income, parental current mood Random: family, site

Table 4.3 — Model specifications for pubertal timing and depressive symptoms association.

Models were conducted for males and females separately.

Total number of tests across all levels of model adjustment: **2** (females) and **2** (males).

Hypothesis 2 (H2): Pubertal timing -> brain structural measures -> depressive symptoms

We tested whether brain structural ROIs measured at Year 2 partially and significantly mediated associations between pubertal timing at Year 1 and depressive symptoms at Year 3 (see Figure 4.3). ROIs were determined based on our pilot analyses (see Pilot analyses section) and are listed in Table 4.5 & Table 4.6. We note that due to the observation of bidirectional effects with regards to the inferior longitudinal fasciculus for pubertal timing–(increased FA) and depressive symptoms–(increased MD) brain structural associations in the pilot analyses, we excluded it as an ROI from the mediation analysis as we did not expect it to have a mediating effect. Pilot analyses results are reported in full in the Supplementary Data.

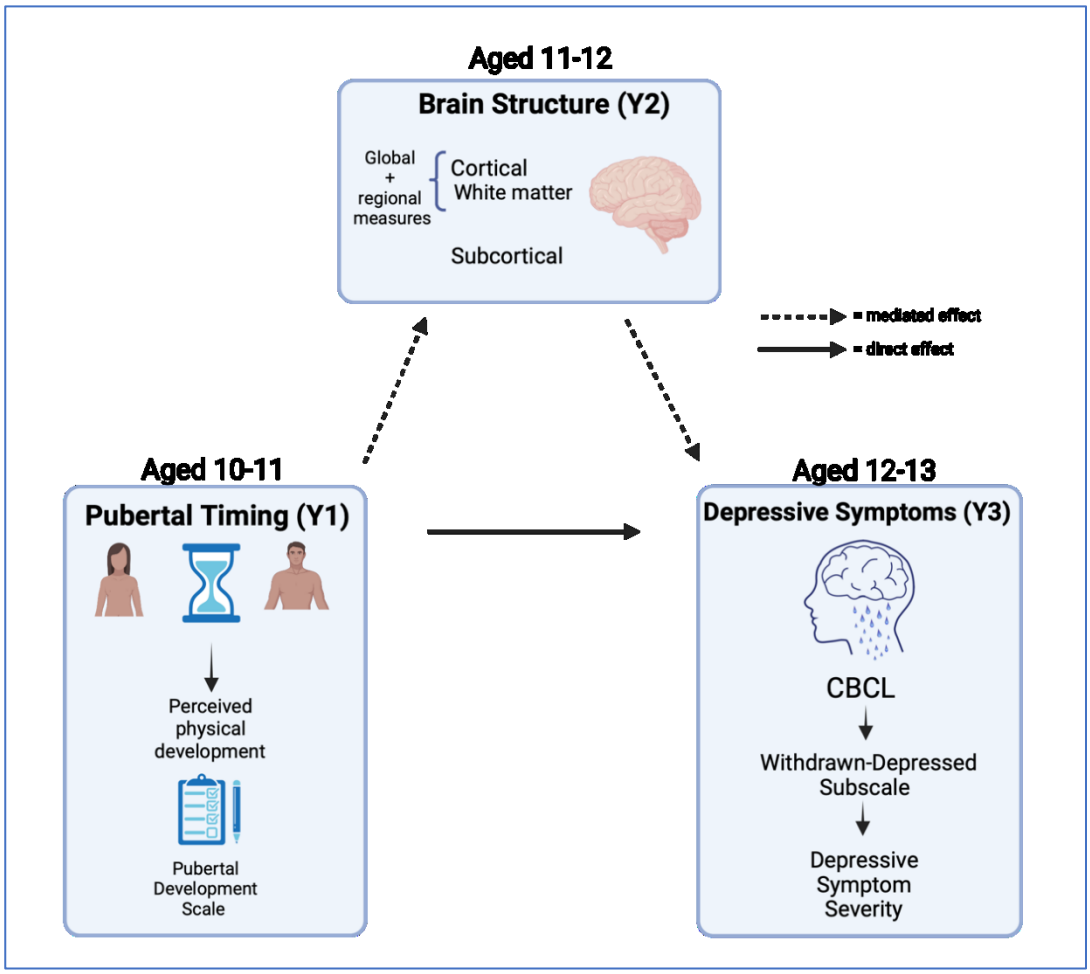


Figure 4.3— Effect of pubertal timing on depressive symptoms including mediation of brain structure.

As detailed in the Pilot analyses section (see Table 4.5 & Table 4.6), several brain structural measures were found to be significantly associated with pubertal timing and depressive symptoms, and thus informed the following hypotheses:

The association between earlier pubertal timing and increased depressive symptoms would be mediated by:

4.5.2 Global measures

- **H2a**: Reduced global cortical volume, surface area, thickness and sulcal depth.

- **H2b:** Reduced global FA

4.5.3 Regional measures

- **H3a:** Reduced cortical thickness in temporal regions, namely, the middle temporal gyrus and insula.

- **H3b:** Reduced cortical thickness in frontal regions namely, the lateral orbito-frontal cortex and middle frontal gyrus.

- **H3c:** Reduced cortical volume in temporal regions, namely, middle temporal gyrus and bank of the superior temporal sulcus.

- **H3d:** Reduced cortical volume in fronto-parietal regions, namely, the middle frontal and postcentral gyri.

- **H3e** Reduced FA in the cortico-striatal tract and corpus collosum.

- **H3f:** Increased sulcal depth in the pars orbitalis

- **H3g:** Increased cortical volume in the ventral diencephalon.

We ran mediation analysis using multi-level structural equation modelling (MLSEM) with Mplus software (Preacher et al., 2010) and via the "MplusAutomation" package in R (Hallquist & Wiley, 2018). MLSEM enables the stratification of within individual, between family and site/scanner variance, therefore allowing us to capture these random effects.

This model characterised associations between pubertal timing, brain structural ROIs and depressive symptoms. We undertook single and multiple mediator models, depending on the brain ROI. Multiple mediator models allowed us to examine the proportion of variance in the pubertal timing-depression associations uniquely explained by all brain structural ROIs and allowed for comparisons between different ROIs. For this analysis, we simultaneously entered individual brain structural ROIs as covarying mediators. A combined cluster variable was used to model random effects in Mplus due to model parameter requirements.

The primary outcomes of interest were the direct effect between the pubertal timing measure and depressive symptoms, and the indirect paths between these two variables that are mediated by brain structural ROIs. Statistical significance of this indirect effect was used to indicate that a significant mediation of the total effect is present. An effect was considered statistically significant when $p \leq 0.05$ and there was a minimum effect size (β value ≥ 0.01). As outlined in the Consideration of outcome neutral conditions section, this minimum effect size was based on our previous work where effect sizes for brain structural associations with depression ratings in the ABCD sample were found to be in the region of 0.01-0.03 (Shen et al., 2021). Bootstrapping with 1000 repetitions was used to calculate robust standard errors.

The Base model included age, race/ethnicity, and DTI motion as fixed effects, and family ID and scanner ID as random effects. The full model included the same random effects as the base model but with the additional fixed effects: WBV, BMI, household income, and parental current mood.

4.5.4 Sensitivity analysis

In addition to the main models, we conducted sensitivity analysis to examine the association between earlier pubertal timing and the potential change (or rather worsening) of depressive

symptoms between timepoints (i.e., Year 1 and Year 3), by including Year 1 depressive symptoms as an additional covariate in our full model as a sensitivity analysis. Further, to examine potential demographic and socio-economic bias in the selection of ABCD participants, we included a population-weighting variable in our full model that calibrates ABCD weighted distributions to nationally representative controls from the American Community Survey (ACS).

4.5.5 *Missing data*

Missing outcome and covariate data were handled using appropriate methods (Matta et al., 2018). Multiple imputation by chained equations (MICE) was used to treat missing data for H1 using the "mice" package in R (Buuren & Groothuis-Oudshoorn, 2011). For H2 and H3, full information maximum likelihood (FIML) estimation in Mplus was used to handle missing data in our mediation analyses. As sensitivity analyses, we compared our imputed analysis approach to complete case analysis for H1.

4.5.6 Exploratory analyses

To identify any additional relevant brain structural measures that may not have been identified in the pilot analyses due to the use of baseline data only, we also undertook exploratory whole brain analysis to examine whether any other brain structural measures (at Year 2) mediated the association between pubertal timing (at Year 1) and depressive symptoms (at Year 3). Multiple comparison correction (FDR method) was applied using the "p.adjust" function in R and applied to each brain measure category separately. These analyses were considered post-hoc and thus reported as exploratory findings.

4.5.7 Project timeline

Our Stage One registered report obtained an in-principle acceptance in August 2022 and we submitted our Stage 2 manuscript for review in December 2022.

4.6 Pilot analyses

4.6.1 Statistical model specifications

Given inconsistent findings in terms of the brain regions associated with pubertal development and depressive symptoms in adolescence, we conducted pilot analysis on baseline data from the ABCD Study to identify ROIs for our second and third hypotheses. We note that although our main analyses used imaging data from Year 2 and depressive symptoms from the Year 3 follow up, the pilot analyses used baseline data *only* for all measures (N = 9,339, males = 4802, females = 4537, mean age = 9.91 years, SD = 0.62). This was to avoid handling any of the follow-up data given that the study is a registered report. Due to the non-longitudinal nature of the pilot analyses, we used complete case analysis.

The pilot analyses consisted of identifying ROIs that were significantly associated with *baseline* measures of 1) pubertal timing and/or 2) depressive symptoms. An association was considered significant if the nominal (un-corrected) p-value ≤0.0001 for the pubertal timing-brain models and a p-value ≤0.005 for the depression-brain models. Nominal thresholds were chosen since these were pilot analyses conducted to inform the main analyses. We note that due to the more numerous and stronger associations observed for the pubertal timing-brain models, a more conservative p-value threshold was chosen to limit the number of ROIs carried through to the main analyses.

Associations between pubertal timing (indexed via PDS) and brain structural measures (cortical, subcortical, and white matter microstructure) were examined using GLMs via the 'lmerTest' (Kuznetsova et al., 2017). Covariates included: age, race/ethnicity, WBV, DTI motion, (for DTI models) as fixed effects and individual ID and scanner ID as random effects.

Associations between depressive symptoms and brain structural measures were examined using the same model specifications as described above.

The analyses of associations with brain structural measures followed a hierarchical order from global measures at the whole-brain level to individual structures. For cortical measures, this included whole brain cortical volume, mean thickness, total surface area, mean sulcal depth,

followed by individual brain regions. For subcortical measures, this included volumes for subcortical regions as specified in the FreeSurfer segmentation. For white matter microstructural measures (FA/MD), the global 'g' measures were first tested, followed by individual tracts.

These brain measures were examined for 1) pubertal timing and 2) depressive symptoms. Models were run separately for males and females. See Table 4.4 for model specifications.

Type	**Measures**	**Number of variables**	**Covariates (Base model)**
Global brain measures	Mean whole-brain cortical thickness	1 unilateral	Fixed: Age + race/ethnicity. Random: Family ID +scanner ID
	Total whole-brain surface area	1 unilateral	
	Mean whole-brain sulcal depth	1 unilateral	
	Total whole-brain volume	1 unilateral	
	Global total white matter fractional anisotropy	1 unilateral	
	Global total white matter mean diffusivity	1 unilateral	
Regional brain measures	Cortical thickness	34 bilateral	Fixed: Age + race/ethnicity + WBV. Random: Family ID +scanner ID
	Cortical surface area	34 bilateral	
	Cortical sulcal depth	34 bilateral	
	Cortical volume	34 bilateral	
	Subcortical regions	16 bilateral	
	White matter fractional anisotropy	14 bilateral, 3 unilateral	Fixed: Age + race/ethnicity + DTI motion. Random: Family ID +scanner ID
	White matter mean diffusivity	14 bilateral, 3 unilateral	

Table 4.4 — Pilot analyses: Model specifications for brain structural measures.

To examine pubertal timing ~ brain structure associations, we ran **6** models for global brain measures, and **184** models for regional brain measures for males and females separately.

To test associations between depressive symptoms and brain structure, we ran **6** models for global brain measures and **184** models for regional brain measures for both males and females separately.

4.7 Pilot results

The association between pubertal timing and brain structural measures at baseline

Global brain metrics

Earlier pubertal timing was associated with lower global cortical thickness and sulcal depth in females (β range: -0·054 to -0.066; $p_{uncorrected}$: = 5.55×10^{-6} to 5.12×10^{-5}). No significant relationships were found in males.

Regional brain metrics

For females, earlier pubertal timing was associated with decreased cortical thickness and volume in temporal and frontal regions, and increased cortical volume in the ventral diencephalon (β range: -0.0425 to 0.0931; $p_{uncorrected}$ range: 6.73×10^{-7} to 0.0001) (see Table 4.5 and Figure 4.4).

For males, earlier pubertal timing was also associated with decreased cortical thickness in the lateral orbitofrontal cortex (β: -0.0525; $p_{uncorrected}$: 0.0001) (see Table 4.5).

Regions meeting criteria for ROIs (p≤0.0001) are reported in Table 4.5 and illustrated in Figure 4.4. See Supplementary Data 1&2 for full details of all models for males and females, respectively.

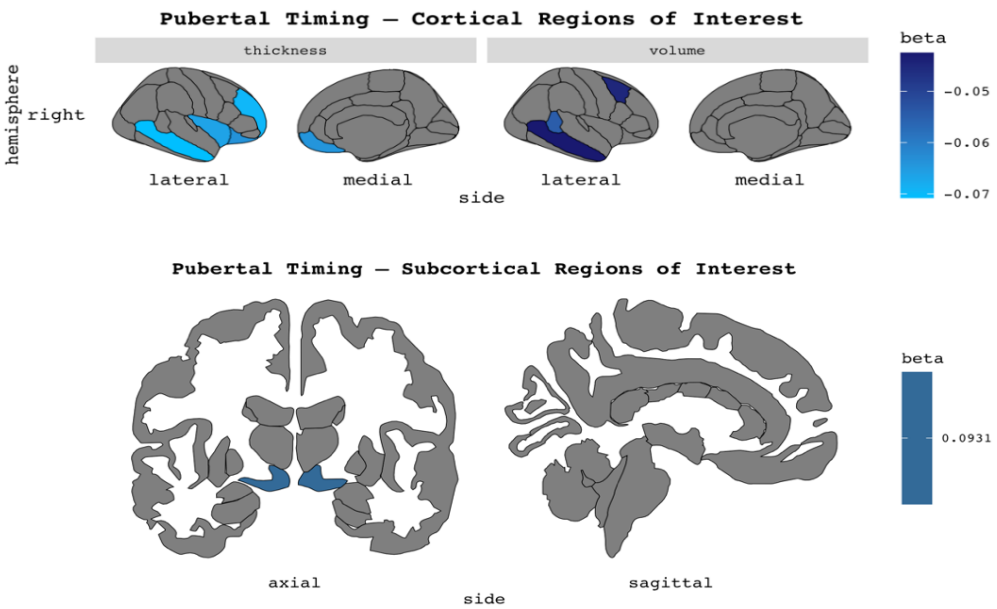


Figure 4.4 — Pilot results: Cortical and subcortical regions of interest for the pubertal timing ~ brain structure association.

Pubertal timing	Brain Structure	beta	std	t.value	$p_{uncorrected}$
Females					
PDS	Global cortical thickness	-0.0659	0.0145	-4.5476	$5.55x10^{-6}$
	Global sulcal depth	-0.0540	0.0133	-4.0536	$5.12x10^{-5}$
	Insula (thickness)	-0.0663	0.0148	-4.4816	$7.58x10^{-6}$
	Lateral orbitofrontal cortex (thickness)	-0.0635	0.0150	-4.2293	$2.39x10^{-5}$
	Middle temporal gyrus (thickness)	-0.0708	0.0142	-4.9754	$6.73x10^{-7}$
	Medial orbitofrontal cortex (thickness)	-0.0640	0.0148	-4.3262	$1.55x10^{-5}$
	Rostral middle frontal gyrus (thickness)	-0.0696	0.0149	-4.6687	$3.11x10^{-6}$
	Bank of the superior temporal sulcus (volume)	-0.0545	0.0129	-4.2148	$2.54x10^{-5}$
	Caudal middle frontal gyrus (volume)	-0.0449	0.0120	-3.7279	0.0001
	middle temporal gyrus (volume)	-0.0425	0.0105	-4.0528	$5.14x10^{-5}$
	Ventral diencephalon (volume)	0.0931	0.0101	9.2557	$3.12x10^{-20}$
Males					
PDS	Lateral orbitofrontal cortex (thickness)	-0.0525	0.0141	-3.7237	0.0001

Table 4.5 — Pilot results: Pubertal timing-brain structure models and associated statistics with significant ROI associations.

The association between depressive symptoms and brain structural measures at baseline

Global brain metrics

Increased depressive symptoms were associated with reduced global cortical volume and surface area for males and females and reduced global FA for females only (β range: -0.0636 to -0.0290; $p_{uncorrected}$ range: 5.67×10^{-8} to 0.0037). No significant relationships were observed in males.

Regional brain metrics

For females, increased depressive symptoms were associated with lower FA in the corpus callosum and parietal superior cortico-striate tract, and increased sulcal depth in the pars orbitalis (β range: -0.0280 to 0.0375; $p_{uncorrected}$ range: 0.0051 to 0.0057).

For males, increased depressive symptoms were associated with decreased cortical surface area and volume in the postcentral gyrus, and reduced volume in the middle temporal gyrus (β range: -0.0256 to -0.0285; $p_{uncorrected}$ range: 0.0014 to 0.0042).

Regions that met criteria to be classified as ROIs (uncorrected p-value ≤0.005) are shown in Table 4.6. Cortical ROIs are also illustrated in Figure 4.5. Please see Supplementary Data 3&4 for a complete list of all depression-brain structure models and associated statistics.

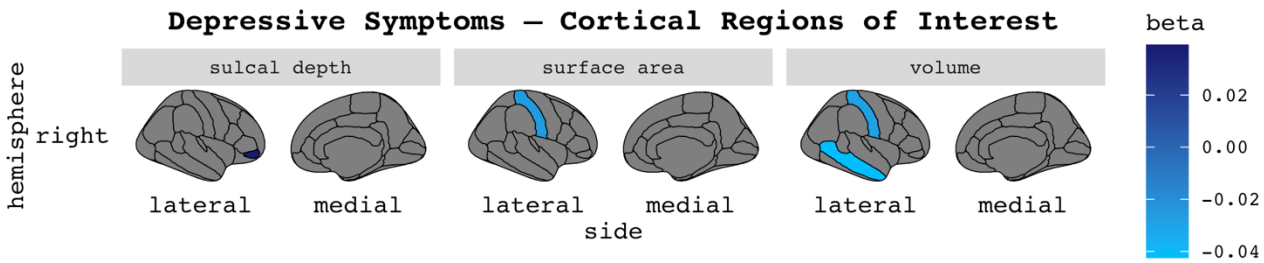


Figure 4.5 — Pilot results: Cortical and subcortical regions of interest for the depressive symptoms ~ brain structure association. Note: DTI ROIs (corpus callosum and parietal superior corticostriate tract) are not illustrated in Figure 5.5 but listed in Table 4.6.

Depressive Symptoms	Brain Structure	beta	std	t.value	$p_{uncorrected}$
Females					
CBCL Withdrawn Depressed (raw score)	Global surface area	-0.0396	0.0122	-3.2330	0.0012
	Global cortical volume	-0.0393	0.0122	-3.2238	0.0013
	Global FA	-0.0290	0.0100	-2.9071	0.0037
	Pars orbitalis (Sulcal depth)	0.0393	0.0133	2.9420	0.0033
	Corpus Callosum (FA)	-0.0280	0.0101	-2.7656	0.0057
	Parietal superior corticostriate (FA)	-0.0327	0.0117	-2.7990	0.0051
Males					
CBCL Withdrawn Depressed (raw score)	Global cortical surface area	-0.0607	0.0118	-5.1263	3.06×10^{-7}
	Global cortical volume	-0.0636	0.0117	-5.4364	5.67×10^{-8}
	Postcentral gyrus (surface area)	-0.0256	0.0090	-2.8623	0.0042
	Middle temporal gyrus (volume)	-0.0283	0.0089	-3.1887	0.0014
	Postcentral gyrus (volume)	-0.0285	0.0093	-3.0636	0.0022

Table 4.6 — Pilot results: Depressive symptoms ~ brain structure models and associated statistics with significant ROI associations.

4.8 Main results

Sample characteristics are presented in Table 4.7

Characteristic	F, N = 2,726[1]	M, N = 3,001[1]	p-value[2]
Age (Y1)	10.98 (0.63)	11.00 (0.64)	0.082
Age (Y2)	11.96 (0.64)	11.98 (0.65)	0.2
Missing (N)	34	35	
Age (Y3)	12.87 (0.64)	12.91 (0.65)	0.021
PDS total score	10.37 (3.09)	7.72 (2.12)	<0.001
Youth depressive symptoms	1.49 (2.19)	1.30 (1.90)	0.012
Missing (N)	46	46	
Race/ethnicity			0.072
White	1,876 / 2,696 (69.58%)	2,175 / 2,975 (73.11%)	
Black	296 / 2,696 (10.98%)	278 / 2,975 (9.34%)	
Asian	70 / 2,696 (2.60%)	73 / 2,975 (2.45%)	
AIAN/NHPI	23 / 2,696 (0.85%)	17 / 2,975 (0.57%)	
Other	105 / 2,696 (3.89%)	111 / 2,975 (3.73%)	
Mixed	326 / 2,696 (12.09%)	321 / 2,975 (10.79%)	
Missing (N)	30	26	
BMI (Y1)	19.43 (4.24)	19.43 (4.10)	0.8
Missing (N)	51	43	
Household income			0.6
<$5000	47 / 2,558 (1.84%)	66 / 2,813 (2.35%)	
$5,000-$11,999	63 / 2,558 (2.46%)	69 / 2,813 (2.45%)	
$12,000-$15,999	55 / 2,558 (2.15%)	43 / 2,813 (1.53%)	
$16,000-$24,999	89 / 2,558 (3.48%)	111 / 2,813 (3.95%)	
$25,000-$34,999	142 / 2,558 (5.55%)	133 / 2,813 (4.73%)	
$35,000-$49,999	212 / 2,558 (8.29%)	233 / 2,813 (8.28%)	
$50,000-$74,999	364 / 2,558 (14.23%)	388 / 2,813 (13.79%)	
$75,000-$99,999	404 / 2,558 (15.79%)	438 / 2,813 (15.57%)	
$100,000-$199,999	848 / 2,558 (33.15%)	964 / 2,813 (34.27%)	
>$200,000	334 / 2,558 (13.06%)	368 / 2,813 (13.08%)	
Missing (N)	168	188	
DTI mean FD	1.20 (0.43)	1.24 (0.53)	0.14
Missing (N)	566	476	
Parent depressive symptoms	3.96 (3.59)	3.95 (3.54)	>0.9
Missing (N)	35	38	

[1] Mean (SD); n / N (%)
[2] Wilcoxon rank sum test; Pearson's Chi-squared test

Table 4.7 — Main results: Descriptive statistics for sample. Y1 = year 1; Y2 = year 2; Y3 = year 3. Youth depressive symptoms = CBCL withdrawn depressed total raw score; AIAN/NHPI = American Indian/Alaska Native/Native Hawaiian and other Pacific Islander; Household income = yearly gross household income; DTI mean FD = mean framewise displacement from year 2 DTI data; Parent depressive symptoms = Depressive Problems ASR DSM-5-Oriented Scale.

Frequencies for perceived pubertal development and youth depressive symptoms, based on parent report, are shown in Figure 4.6

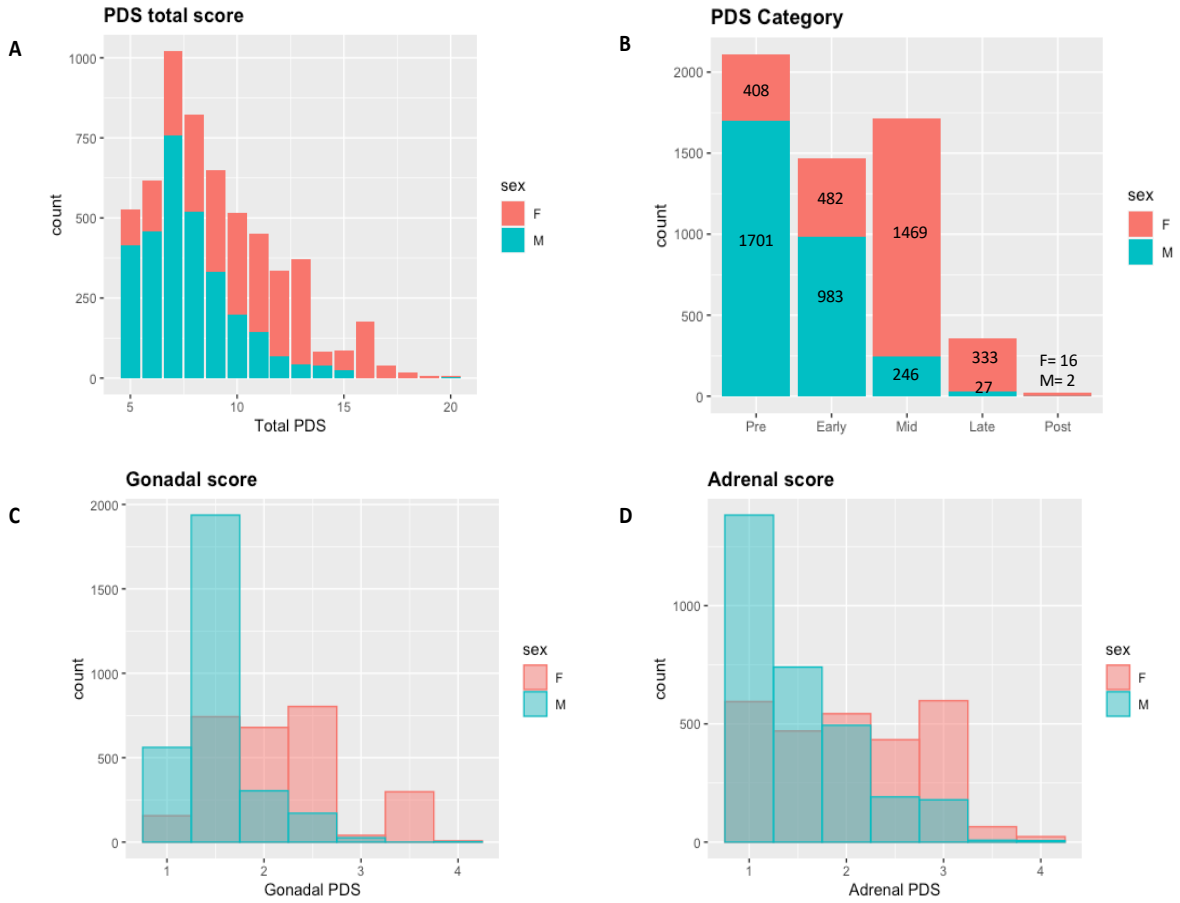


Figure 4.6 —Main results: Frequencies (N) for parent summary scores from the Pubertal Development Scale (PDS). (A) Total pubertal development score; (B) PDS Category score counts ranging from pre- to post- pubertal. Note that PDS category score (variables: "pds_p_ss_female_category" and "pds_p_ss_male_category" data were not available for 60 participants in ABCD release 4.0 so N = 5667 (full sample N = 5727) for Figure 4.6b. (C) Gonadal score averaging gonadal PDS items and ranging from 1 = not begun to 4 = complete; (D) Adrenal score averaging adrenal PDS items ranging from 1 = not begun to 4 = complete.

4.8.1 A note on the interpretation of effect sizes

Given that our outcome (youth depressive symptoms) is a count variable (with a Poisson distribution), it was not standardised in our main analyses. However, due to the large variance observed across our variables (especially the imaging variables), all other numeric variables (except the population weight propensity score) were converted to z-scores before running our models to ensure consistency across our hypotheses. Therefore, unless otherwise stated, the estimated ß values should be interpreted as follows: for one unit change in the predictor variable, the difference in the logs of the expected counts is predicted to change by the respective ß value, given that the other predictor variables in the model are held constant. To further aid the interpretation of our main results, we report incidence rate ratios (IRR) alongside the ß values and associated p-values. The incidence rate ratio is the exponential of the reported ß coefficient and can be interpreted as a relative risk.

4.8.2 Hypothesis 1

Our first hypothesis tested whether earlier pubertal timing at year 1 (youth aged 10-11 years) is associated with higher depressive symptoms at year 3 (youth aged 12-13). Our analyses demonstrate support for this hypothesis such that both females and males who started puberty earlier than their peers were more likely to report higher depressive symptoms two years later. Basic model: Females: ß = 0.27; [IRR = 1.31]; $p < 2 \times 10^{-16}$; males: ß = 0.08 [IRR = 1.09]; $p = 0.005$. In our fully adjusted model, which included additional covariates (BMI, household income, and parental depressive symptoms), our main effect size attenuated slightly for females (ß = 0.20 [IRR = 1.22], $p < 2 \times 10^{-16}$). However, the observed effect size for males (ß = 0.04; [IRR = 1.045], $p = 0.15$) no longer met the threshold for evidence (as per the definition outlined in Table 4.1) when these additional factors were accounted for in our model. The results (reported using IRRs) for Hypothesis 1 for females and males (base and fully adjusted models) are illustrated in Figure 4.7. All statistics (e.g., ß values, IRRs, standard errors, and *p* values) for Hypothesis 1 are reported in Tables S1 & S2 in the Supplementary information.

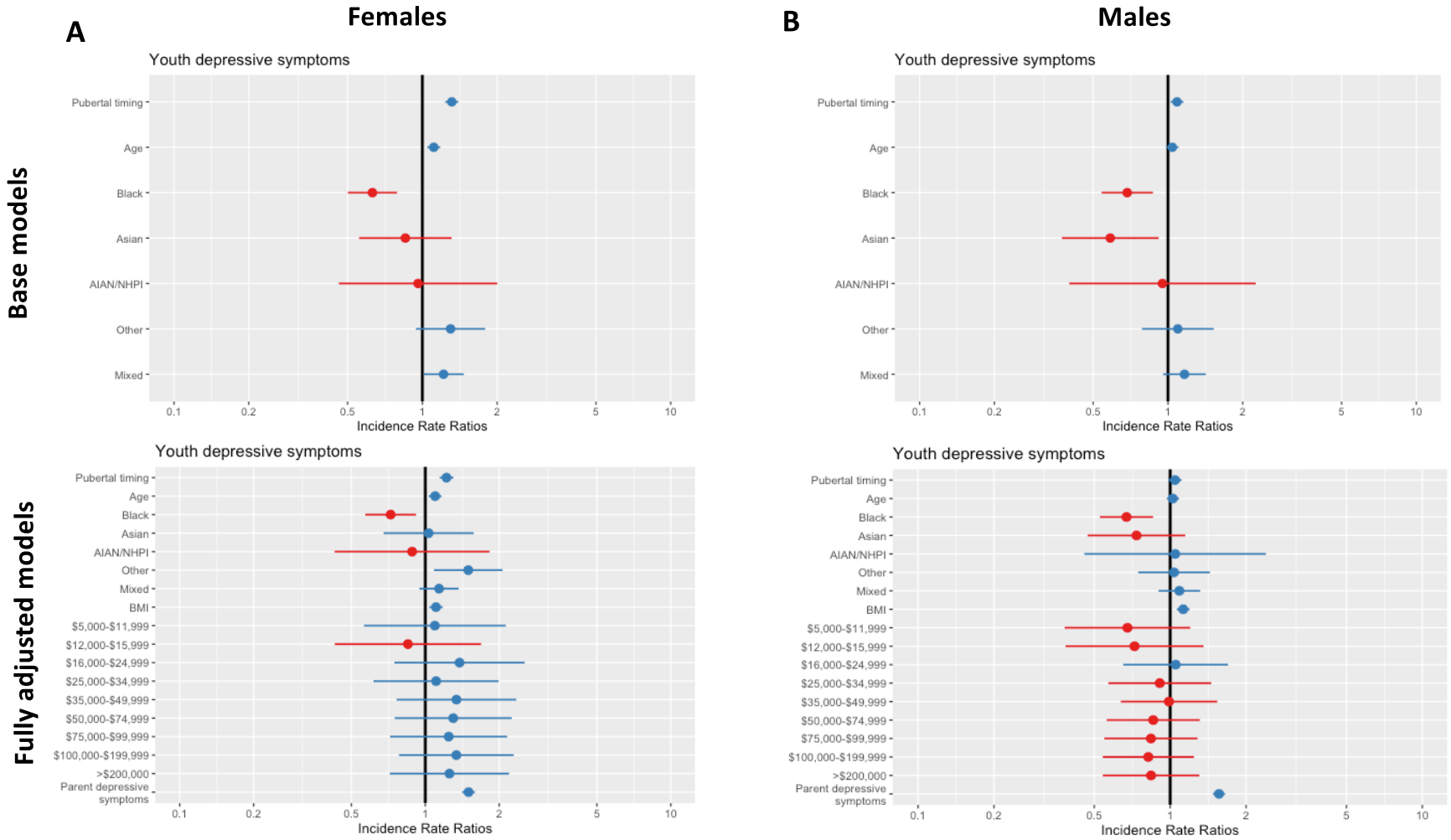


Figure 4.7— Main results: Incidence Rate Ratios (IRRs) for the association between pubertal timing and youth depressive symptoms. Results for females are shown in (A) and males are shown in (B). Base models are shown in top panel and fully adjusted models are presented in the bottom panel. The neutral line or vertical intercept line is shown in bold and indicates no effect. Blue IRRs indicate a greater depression risk while red IRRs represent a decreased depression risk. Error bars represent 95% confidence intervals.

4.8.3 Exploratory analyses related to Hypothesis 1

To explore whether there were specific aspects of pubertal development that were driving the association between earlier pubertal timing and increased depression risk, we examined adrenal and gonadal PDS items separately and generated an average score for each. As previously described by Shirtcliff et al., 2009, and adopted by others using the ABCD puberty data (e.g., Herting et al., 2021), a gonadal PDS score was generated for females by averaging growth spurt, breast development and menarche PDS items (variables: pds_1_p, pds_f4_p, pds_f5b_p), and for males by averaging growth spurt, deepening of voice, and facial hair PDS items (variables: pds_1_p, pds_m4_p, pds_m5_p). Adrenal PDS scores were calculated for both sexes by averaging pubic, body hair and skin changes PDS items (pds_2_p, pds_3_p). A pubertal timing score for adrenal and gonadal measures was then obtained by regressing the adrenal and gonadal PDS scores on age, and using the residual obtained as the timing measure. This was done for males and females separately.

Using the same model set up as H1, we ran two independent models with adrenarcheal timing (AT) and gonadarcheal timing (GT) scores as predictors. For females, both AT and GT were significantly associated with later youth depression (Base model: AT: ß = 0.23 [IRR = 1.26]; $p < 0.001$; GT: ß = 0.24 [IRR = 1.28]; $p < 0.001$), and these effects remained significant in the fully adjusted model (AT: ß = 0.17 [IRR = 1.18]; $p < 0.001$; GT: ß = 0.17 [IRR = 1.19]; $p < 0.001$). For males, only AT was significantly associated with later youth depression (Base model: AT: ß = 0.10 [IRR = 1.11]; $p = 0.001$; GT: ß = 0.05 [IRR = 1.04]; $p = 0.154$). Neither association was significant in the fully adjusted model for males. Results are reported in full in Tables S3-S6 in the Supplementary Information.

As further post-hoc analysis, we included both GT and AT scores in the same model (base model specification) to investigate whether one aspect of pubertal development was significantly associated with later youth depression, above and beyond the other. First, Spearman's rank correlation was computed to assess the relationship between AT and GT measures (Females: Spearman's $\rho = 0.61$; Males: Spearman's $\rho = 0.44$). Model specifications: $youth\ depression \sim gonadal\ timing + adrenal\ timing + age + race + 1|\ site\ ID + 1|\ family\ ID$.

Our results suggest that, in females, both GT and AT contribute significantly to the association between earlier pubertal timing and youth depression (GT, controlling for AT, ß = 0.18 [IRR = 1.19]; $p < 0.001$; AT: ß = 0.12 [IRR = 1.13]; p = 0.001). We note here that the effect size for GT is slightly larger than AT. For males, our results suggest that the association between earlier pubertal timing and later youth depression is being driven by AT rather than GT (GT controlling for AT, ß = -0.006 [IRR = 0.994]; $p = 0.85$; AT: ß = 0.10 [IRR = 1.11]; $p = 0.003$). Results are reported in full in the Supplementary Information, Tables S7 and S8.

4.8.4 Hypothesis 2 — global brain measures

Our second hypothesis tested whether global brain structural measures at year 2 (specifically, lower global volume, cortical thickness, surface area, and sulcal depth (H2a) and FA (H2b)) mediated the association between earlier pubertal timing at year 1 and higher depressive symptoms at year 3. In both females and males, we did not find support for these hypotheses in the current analyses due to the absence of an indirect effect (Females: H2a: ß = -0.001 [IRR = 0.99], $p = 0.89$; H2b: ß = 0.001 [IRR = 1.00], $p = 0.57$; Males: H2a: ß = 0.003 [IRR = 1.00], $p = 0.43$; H2b: ß = -0.001 [IRR = 0.99], $p = 0.39$) from our predictor (pubertal timing) to our outcome (youth depressive symptoms) through our hypothesised mediator (brain structure).

4.8.5 Hypothesis 3 — regional brain measures

Our third hypothesis investigated whether regional brain structural measures at year 2 (identified via our pilot analyses and listed in Table 4.1) mediated the association between earlier pubertal timing at year 1 and later youth depressive symptoms at year 3. Overall, for both females and males, our results did not find support for our hypotheses due to an absence of an indirect effect (Females: ß range: -0.008 to 0.004 [IRR range = 0.99 to 1.00], p range: 0.05 to 0.89; Males: ß range: -0.004 to 0.003 [IRR range = 0.99 to 1.00], p range: 0.27 to 0.85). Of note, H3a, that reduced cortical thickness mediates the association between earlier pubertal timing and increased youth depression, partially met our threshold for evidence (ß ≥ 0.01, $p \leq 0.05$) such that the p value was marginally within the criteria, but the ß value was not (ß = -0.008, $p = 0.05$). Given the proximity of this p value to our predefined threshold for

evidence (outlined in Table 4.1) and the very small observed effect size, we did not consider this support for our H3a hypothesis.

All model statistics, including the direct and total effects, for single and multiple mediator models for Hypotheses 2 and 3 are reported in full in Supplementary Data 5 & 6 (females) and 7 & 8 (males). We have also included the Mplus output files for all models in the Supplementary Information for this registered report.

4.8.6 Exploratory analyses related to Hypothesis 2 & 3

We undertook exploratory analyses to identify any additional brain structural measures that may mediate the association between earlier pubertal timing and later increased depressive symptoms. Further, we also wanted to identify brain structural features related to pubertal timing and depressive symptoms in this large sample of early adolescents. We re-ran the base models specified in the pilot analyses (with the removal of WBV a covariate due to its potential effect on regional brain estimates (Mills et al., 2016)) using pubertal timing data from year 1, imaging data from year 2, and depressive symptom data from year 3. For this exploratory analysis, we used the base model set up (as per the approach taken in our pilot analyses), so that the models were consistent across the pilot and main analyses.

4.8.6.1 Whole brain exploratory analyses: Brain structural associations with pubertal timing

Standardised beta values for the pubertal timing — brain structure associations that remained significant after correction for multiple comparison ($p_{FDR} \leq 0.001$) are reported here in the main text. Relevant statistics (ß values, standard errors, uncorrected and corrected p-values) for all pubertal timing – brain structure models are reported in Supplementary Data 9 (females) & 10 (males).

For females, we found that earlier pubertal timing was associated with reduced global cortical thickness (ß = -0.10; $p_{FDR} = 4.4 \times 10^{-5}$) and global cortical volume (ß = -0.09; $p_{FDR} = 1.3 \times 10^{-5}$). Regionally, earlier pubertal timing was associated with reduced cortical thickness and volume

in temporal, frontal and parietal regions (ß range: -0.12 to -0.08; p_{FDR} range: 2.2×10^{-6} to 0.0008). See Figure 4.8.

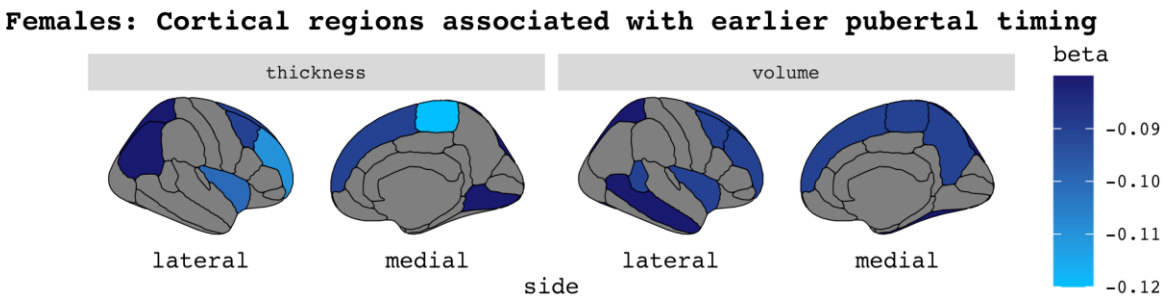


Figure 4.8 — Exploratory results: Significant cortical associations with earlier pubertal timing in female youth. $p_{FDR} \leq 0.001$.

For males, earlier pubertal timing was not significantly associated with global brain measures. The only regional brain measure that demonstrated a significant association with earlier pubertal timing was the increased volume of the ventral diencephalon (ß = -0.07; p_{FDR} = 0.001).

4.8.6.2 Brain structural associations with depressive symptoms

Beta values for the brain structure – depression symptoms associations that remained significant after correction for multiple comparison ($p_{FDR} \leq 0.05$) are reported in the main text. We note that a less conservative threshold for multiple comparison correction was used here compared to the pubertal timing — brain structure models due to the more numerous and stronger associations found for the latter association. Relevant statistics for all brain structure — depression models can be found in Supplementary Data 11 (females) & 12 (males).

For females, no significant associations were found between global brain measures and depressive symptoms. However, our results demonstrated several significant regional associations with depression, namely, reduced volume in the accumbens area (ß = -0.105 [IRR = 0.90], p_{FDR} = 0.024, increased sulcal depth in the bank of the superior temporal sulcus (ß =

0.133 [IRR = 1.14], p_{FDR} = 0.003) and precuneus (ß = 0.12 [IRR = 1.12], p_{FDR} = 0.020), as well as increased MD in the inferior fronto-occipital fasciculus (ß = 0.11 [IRR = 1.12], p_{FDR} = 0.050).

For males, we did not find any significant associations between global brain measures and depression symptoms. Regarding regional associations, depressive symptoms were associated with reduced volume in the accumbens area (ß = -0.10 [IRR = 0.90], p_{FDR} = 0.012), pallidum (ß = -0.08 [IRR = 0.92], p_{FDR} = 0.052), and thalamus (ß = -0.08 [IRR = 0.92], p_{FDR} = 0.056), as well as reduced surface area in the medial orbitofrontal gyrus (ß = -0.11 [IRR = 0.89], p_{FDR} = 0.029).

4.8.6.3 Exploratory mediation analysis

Unlike the pilot analyses, whereby we included any ROIs that were associated with increased depressive symptoms and/or earlier pubertal timing, we adopted a more streamlined approach in our exploratory analyses given that we had full access to the data and any findings would be reported as post-hoc. Therefore, we included any ROI that demonstrated a significant association (after correction for multiple comparisons, specified above) with *both* earlier pubertal timing at year 1 and increased depressive symptoms at year 3. The only brain measure that met this criterion was lower volume of the accumbens area in females. Model 1: Accumbens area ~ pubertal timing (Y ~ X): ß = -0.086, p_{FDR} = 0.001); Model 2: Depressive symptoms ~ accumbens area (Y ~ X): ß = -0.105, p_{FDR} = 0.02). In males, lower accumbens area volume was associated with increased depressive symptoms (ß = -0.104, p_{FDR}= 0.013) but not with pubertal timing (ß = -0.02, p_{FDR}= 0.36). However, for completeness, we tested whether the accumbens area mediated the association between earlier pubertal timing and later depression in both females and males.

For both females and males, we did not find any evidence of a mediating effect of accumbens area volume on the association between earlier pubertal timing and increased depressive symptoms (Females indirect effect: ß = 0.005, p = 0.14; Figure 4.9a; Males indirect effect: ß = 0.004, p = 0.09; Figure 4.9b).

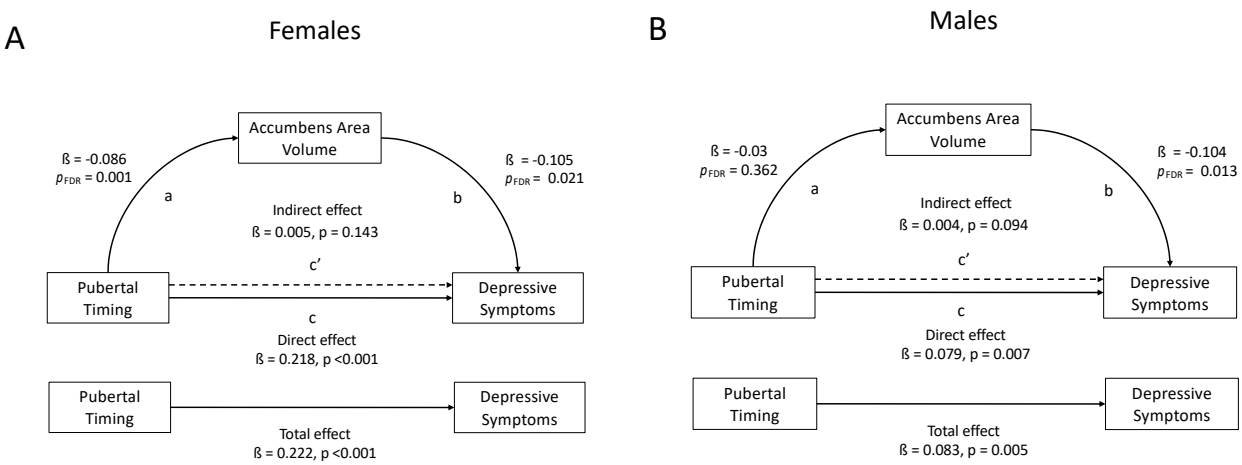


Figure 4.9 — Exploratory results: Mediation paths and statistics for main effect of pubertal timing and depressive symptoms, mediated through accumbens area volume. Results for females are shown in (A) and males are shown in (B).

4.8.6.4 Additional mediation

Given that the current findings did not provide support for a mediating effect of brain structure on the association between earlier pubertal timing and later youth depression, we undertook post-hoc analyses to investigate whether other factors may have a mediating effect. While a host of factors could have been tested, given that exploratory analysis was not the primary focus of this registered report, we decided to only examine variables already included in the study design. Given that pubertal timing and parental depression demonstrated the strongest associations with our outcome of interest, we examined whether rather than being a predictor, pubertal timing mediated the association between early risk factors for depression, namely, parental depression and youth depression. Thus, we investigated whether pubertal timing at year 1 (youth aged 10-11 years) mediated the association between parental depression at baseline (youth aged 9-10 years) and youth depression at year 3 (youth aged 12-13 years).

As illustrated in Figure 4.10, we found that earlier pubertal timing mediated (indirect effect: ß = 0.014, p = 0.002) the positive association between parental depression and later youth depression in females only.

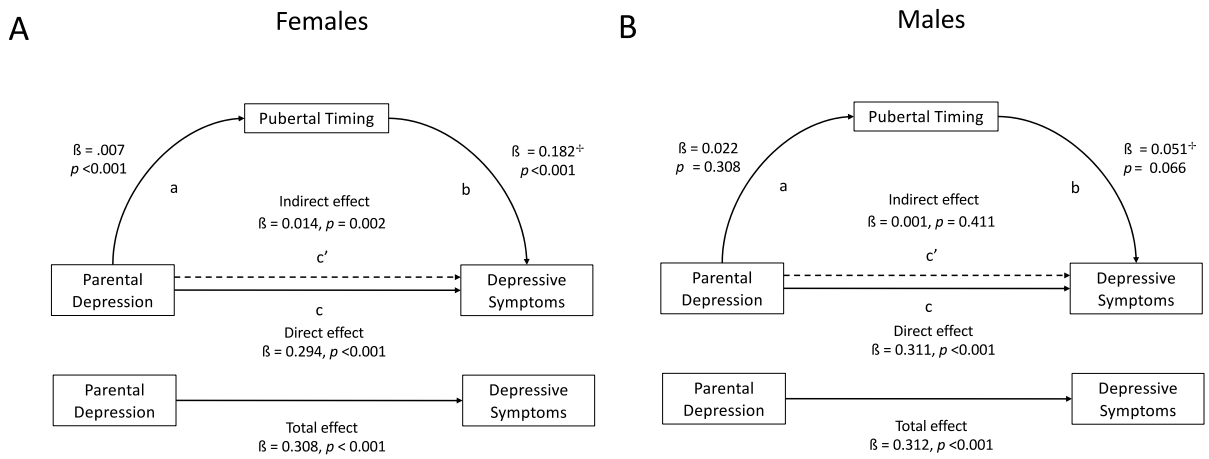


Figure 4.10 — Exploratory results: Mediation paths and statistics for main effect of parental depression and youth depressive symptoms, mediated through pubertal timing. Results for females are shown in (A) and males are shown in (B). + = effect size controlling for parental depression.

4.8.7 Missing data

For H1, we imputed missing outcome and covariate data using multiple imputation by chained equations. Compared to complete case analysis, similar effect sizes were observed when missing data was imputed (see Tables S9 & S10 in the Supplementary Information). Details of our imputation methods can be found in the Supplementary Information.

4.8.8 Sensitivity analysis

We examined the association between earlier pubertal timing and the potential change (or rather worsening) of depressive symptoms between timepoints (i.e., Year 1 and Year 3) by including Year 1 youth depressive symptoms as an additional covariate in our base and fully adjusted models. For females, we found that earlier pubertal timing was significantly associated with the worsening of depressive symptoms over time (Base model: ß = 0.17 [IRR = 1.17]; p <0.001; fully adjusted model: ß = 0.16 [IRR = 1.73]; p <0.001). For males, earlier pubertal timing was not significantly associated with the worsening of depressive symptoms over time (Base model: ß = 0.05 [IRR = 1.05]; p =0.06; fully adjusted model: ß = 0.05 [IRR = 1.05]; p = 0.06). Base and fully adjusted models for females and males are reported in full in Tables S11 and S12 in the Supplementary information.

We also found that when a population stratification variable was added to our full model as a regression weight, our main effects were similar to the main analysis. However, the addition of the population weight in our models did inflate the effect sizes of some factor levels in our race/ethnicity variable, which may be due to a mismatch between the representation of this ethnicity in the US population versus the ABCD sample. See Supplementary Information, Tables S13 (females) and S14 (males).

4.9 Discussion

In the present study, we investigated whether earlier pubertal timing is associated with an increased risk for later depressive symptoms in adolescence, and whether certain *a priori* brain structural measures mediated this association, in a large, demographically diverse sample of youth. We found that earlier pubertal timing when youth were aged 10-11 years was significantly associated with increased depressive symptoms two years later, when youth were aged 12-13 years. Thus, the first hypothesis of this registered report was supported. Although this association was found in both female and male youth, the observed association was stronger for female adolescents compared to males. Further, in females but not males, this association remained significant when controlling for other factors associated with depression risk, such as family income, parental depression, and BMI. On the other hand, we did not find evidence that the hypothesised brain structural measures (Hypotheses 2 & 3) mediated the association between earlier pubertal timing and later depressive symptoms.

Our exploratory analyses demonstrated brain structural associations with earlier pubertal timing and youth depressive symptoms, although the effects were more numerous and widespread between pubertal timing and brain structure. Of note, we also did not find any evidence of brain structural mediation when examining regions beyond those specified *a priori*. However, when we broadened the lens taken to investigate depression risk within this sample, our post-hoc analyses revealed that earlier pubertal timing mediated the association between parental depression and increased youth depressive symptoms.

The current results advance our understanding of how pubertal timing relates to brain structural maturation and depression risk beyond age-related changes using one of the largest available samples to date. Taken together, the findings of this registered report suggest that while a robust association exists between earlier pubertal timing and increased depressive symptoms, particularly for females, brain structure does not mediate this association. Our results highlight the need to consider additional biological factors (e.g., genetics), other brain metrics (e.g., brain function, brain age gap estimates (AGE)) and socio-environmental risk factors, when examining the association between earlier pubertal timing

and increased depression risk, and the differential impact they may have across sexes. Given the significant burden of depression in adolescence and beyond, further research is urgently needed in this area so that we can better understand how to support young people as they navigate this formative developmental transition.

4.9.1 Earlier pubertal timing is associated with later youth depressive symptoms

A substantive body of evidence has demonstrated that youth that begin puberty ahead of their peers are at an increased risk of psychopathology, including depression. Using one of the largest sample sizes to date (N = ~ 5,300), our findings extend prior work and emphasise the detrimental effects of accelerated pubertal maturation on youth mental health outcomes. On a high-level, our findings lend some support to the maturation disparity hypothesis that has been proposed to explain the effects of early pubertal timing on psychopathology risk (Brooks-Gunn et al., 1985; Ge et al., 2001; Ge & Natsuaki, 2009). This conceptual model does not explicitly predict sex differences and posits that both early maturing males and females are at an increased risk for mental health difficulties during adolescence due to an incongruency in their physical, cognitive, social, and emotional development. Findings from a recent meta-analysis by Ullsperger and Nikolas (2017) provide the strongest empirical support for this conceptual framework. This pattern of results has also been replicated in recent studies, whereby early pubertal timing was found to predict the onset and recurrence of depression in male and female adolescents (Hamlat et al., 2020; McNeilly et al., 2022).

Although we found significant associations between earlier pubertal timing and increased depressive symptoms in both sexes, the magnitude of effect was consistently greater in females. Importantly, at the time pubertal development was assessed in our study (youth aged 10-11 years), most males were pre-pubertal or in the early stages of puberty, while the females exhibited a much broader spread of pubertal maturation. Thus, the greater magnitude of effect for pubertal timing and later psychopathology in females that we observed could reflect a temporal effect, such that the distress associated with the experience of maturing ahead of your peers does not manifest straight away. This reasoning may explain why the youth that are most pubertally mature in our sample (i.e., the females) exhibit a

stronger association with depressive symptoms. Further, our sensitivity analyses revealed that when earlier depressive symptoms were included in our model, earlier pubertal timing was associated with the worsening of symptoms (between ages 10-11 years and 12-13 years) in females but not males, which may reflect the significantly higher incidence of depressive symptoms in females compared to males in the current sample.

Another possible explanation for the difference in magnitude of effect between sexes, is that there are distinct biological processes underpinning the composite early pubertal timing effect on psychopathology. Rather than using a global measure of pubertal timing, our exploratory analyses examined how specific aspects of pubertal timing, namely, adrenarcheal and gonadarcheal timing were associated with depressive symptoms. Our results demonstrated that both aspects of pubertal timing were associated with depressive symptoms in females, which is consistent with prior findings (Barendse et al., 2021). However, when looking at males, a significant association was found between depressive symptoms and adrenarcheal timing but *not* gonadarcheal timing. This suggest that in the current sample of early adolescent males, the observed association between early pubertal timing and later depressive symptoms was being driven by adrenarcheal aspects of pubertal maturation, which occur earlier than gonadarcheal changes.

Taken together, these findings indicate that gonadarcheal processes may strengthen the association between earlier pubertal timing and depressive symptoms in adolescence. However, additional follow up data in males is needed to properly test this hypothesis. Of note, while the males and females in the ABCD sample are the same chronological age, there are marked differences in their pubertal progression. Thus, if we examined the association between pubertal timing and depressive symptoms in males when they are at the same "pubertal age" as females (i.e., when they are one or two years older), the same gonadarcheal/adrenarcheal effect may be observed. As multiple timepoints of ABCD data become available across adolescence, longitudinal modelling (e.g., latent growth curve analysis) should be used to test sex differences in pubertal timing, as well as tempo, and their relation to depression risk.

Crucially, the measures of adrenarcheal and gonadarcheal timing used in our analysis were proxy measures for underlying hormonal changes, where there is a known temporal delay between fluctuations in hormone changes and the associated physical changes (Bordini & Rosenfield, 2011). Recent research by Barendse & Byrne et al., 2021, using a multiverse analysis approach (in a female sample only), did not find an association between hormone-based pubertal timing measures and increased risk for internalising psychopathology. The authors conclude that hormone-based measures of pubertal timing may not contribute to the aspects of pubertal timing that are associated with adolescent mental health difficulties. While examining hormonal assays of pubertal maturation was beyond the scope of this registered report, as previously mentioned, myriad biological and socio-environmental factors are likely to moderate or mediate the observed effects. An important next step for the field will thus be to broaden the scope of existing multiverse approaches so that they explore direct and indirect effects associated with multi-modal measures of pubertal timing and psychopathology across sexes, and crucially, how they develop over time.

It must also be noted that the work of others, including meta-analytic findings from Ullsperger and Nikolas (2017), has shown that sex does not moderate the association between pubertal timing and depression risk in adolescence (Hamlat et al., 2020; McNeilly et al., 2022). For example, using baseline data from ABCD, McNeilly and Saragosa-Harris et al. (2022) have recently shown that the effect sizes in the association between earlier pubertal timing and internalising difficulties are similar in male and female youth aged 9-10 years. However, a direct comparison of results is difficult given differences in how pubertal timing is defined and measured across studies. Like the current study, some studies (e.g., Barendse et al., 2021), quantify pubertal timing using the residual obtained from regressing a pubertal development score on age. Then, depending on the model outcome (e.g., depression), age is often added as an additional covariate to examine the association between the predictor (e.g., pubertal timing) and the outcome *beyond* age related changes. However, other studies (e.g., McNeilly et al., 2022) have quantified pubertal timing via pubertal development summary scores (e.g., higher PDS scores = earlier pubertal timing) and use age as a covariate in their models. While such approaches may be appropriate when the age range of the sample is narrow (Vijayakumar et al., 2018), the significant heterogeneity in how pubertal timing and age-

related effects are assessed, may explain the inconsistencies (e.g., sex differences) reported in the literature. Fortunately, the ongoing emphasis on protocol papers that outline considerations for researchers studying pubertal development, such as that by Cheng and colleagues (2021), will greatly aid the harmonisation of analysis pipelines and make result comparison easier.

It may also be that the inconsistent findings in the current literature reflect a general direct effect between earlier pubertal timing and depressive symptoms that is comprised of sex-specific vulnerabilities that vary in magnitude across adolescence. For example, previous research has shown that ethnicity, life stress, and cognitive processes (e.g., rumination) moderate the risk of earlier pubertal timing for later psychopathology according to sex (Alloy et al., 2016; Hamilton et al., 2014). Indeed, the findings of the current study underscore the importance of considering this nuance. While our base models were significant across males and females, important sex-differences emerged when additional socio-environmental risk factors (e.g., BMI) were controlled for in our models. This provides further evidence to suggest that while early maturing youth are at an increased risk for depression in adolescence, there may be sex-specific biological and social/environmental mechanisms that influence this risk. Future research that adopts a biopsychosocial conceptual framework is needed to refine existing theories so that they better reflect the complex interplay of risk (and resilience promoting) factors that underpin the association between pubertal timing and psychopathology in adolescence (Ullsperger & Nikolas, 2017).

4.9.2 Brain structure does not mediate the association between earlier pubertal timing and later depressive symptoms

In this study, we did not find that cortical, subcortical, or white matter microstructural measures mediated the association between earlier pubertal timing and increased depressive symptoms in adolescents. This absence of a mediating effect was observed in both the brain structural measures originally hypothesised, and tested in our confirmatory analyses, as well as those examined in our exploratory analyses. Together, these results suggest that although pubertal timing is associated with alterations in brain morphology above and beyond age-related changes, these brain structural features do not appear to mediate the increased risk

for later depressive symptoms in early developing youth. Future work should explore whether other neuroimaging features, such as brain function, which has been largely understudied in the field (Pfeifer & Allen, 2021), mediates the association between earlier pubertal timing and depression risk.

Importantly, our exploratory work extends existing research by examining brain structural associations with pubertal timing in the largest available sample to date (N = ~5,000). Crucially, our analyses pertain to brain structural associations with pubertal timing specifically which is distinct from examining pubertal development controlling for age, although as discussed, this distinction is often overlooked in the extant literature. In line with existing research, we found that earlier pubertal timing was associated with lower global cortical volume and thickness. Moreover, our results demonstrated regional reductions in cortical volume and thickness in both frontal (e.g., middle frontal gyri), temporal (e.g., the insula, bank of the superior temporal sulcus), and parietal regions (e.g., the precuneus, inferior and superior parietal gyri, paracentral gyrus), which is consistent with prior findings that have used both physical and hormonal pubertal measures (Goddings et al., 2019; Vijayakumar et al., 2018). Further, we also found that a decrease in the volume of the nucleus accumbens was related to earlier pubertal timing which replicates earlier work (Goddings et al., 2014). Although a positive association between pubertal timing and FA has been reported previously (Herting et al., 2012; Peper et al., 2015), we did not find that white matter microstructure was associated with pubertal timing in the current sample.

It is important to note however, that the above findings were for female adolescents only and the only significant association between earlier pubertal timing in males was with increased volume of the ventral diencephalon. One possible interpretation of this pattern of results is that there is a temporal delay between hormonal changes and downstream effects on brain structure, and due to differences in the age of puberty onset between sexes, we do not yet see brain structural effects in males. Our pilot analyses provide some preliminary evidence for this interpretation such that we observed increased volume of the ventral diencephalon in females at ages 9-10 years but not in males at the same age, and we see the opposite pattern of results a year later in our main analyses. Given that the ventral diencephalon

houses the hypothalamus, the key endocrine structure responsible for the surge in gonadal hormones during puberty, the temporal differences we see in volumetric increases across males and females may reflect these distinct pubertal maturation timelines. Future longitudinal work is needed to further understand how pubertal timing affects brain structural development over time, which will help elucidate whether the sex differences observed in the current study attenuate as the pubertal stages of females and males align.

Our earlier work (Shen et al., 2021), and that of others (Schmaal et al., 2017), has demonstrated that global and regional alterations in cortical and white matter microstructural measures are associated with depression in adolescence. The results of the current study are somewhat aligned with these earlier findings such that we see differences in some temporal, parietal and frontal regions, as well as in fronto-occipital white matter tracts, with a consistent directionality of effects. However, these results were not consistent across females and males, and the depression-related imaging features were less widespread compared to existing findings, highlighting the need for further work on this topic. This discrepancy in results could be due to several factors including the significantly smaller sample of the present results as well as differences in the depression outcome measures and statistical analysis methods used. Nonetheless, we also report alterations in subcortical areas such as reduced volume of the nucleus accumbens (in both sexes), pallidum, and thalamus (males only). While hippocampal volume reductions is the most consistently reported depression-related subcortical region (Schmaal et al., 2016), research on typical subcortical development suggests a subtle decrease in the volume of the nucleus accumbens, pallidum, and thalamus across adolescence (Herting et al., 2018). Thus, depression-related subcortical features found in the present sample could reflect an accelerated neurodevelopment, which is thought to be related to early life stress and/or the emergence of depressive symptoms in adolescence (Callaghan & Tottenham, 2016; Ho & King, 2021). Indeed, this interpretation could also be extended to the associations demonstrated between earlier pubertal timing and brain structure.

4.9.3 Looking back to move forward: Expanding our conceptual model linking pubertal timing, neurodevelopment, and mental health outcomes

While the focus of this registered report was to test the mediating role of brain structure in the association between early pubertal timing and increased depression risk, it is also important to consider the factors associated with accelerated pubertal development to begin with. The findings of our exploratory work, where earlier pubertal timing was found to mediate the association between parental depression and youth depression, underscore the need for further work in this area. In fact, recent work by Colich and McLaughlin (2022) proposes that earlier pubertal development may be a mechanism that relates early-life adversity with the emergence of internalising difficulties in adolescence (Colich & McLaughlin, 2022). Thus, the conceptual framework of the current study should be expanded in future work to consider socio-environmental factors (e.g., early life adversity) that predict earlier pubertal timing. Given that knowledge gaps still exist in our understanding of normative brain development during adolescence (particularly in terms of brain function), global neuroimaging metrics, such as "BrainAGE" may better capture deviations (e.g., acceleration) from typical neuromaturation, and on what scale this occurs (globally, or in particular brain networks) (Colich & McLaughlin, 2022; Popescu et al., 2021). Emerging research has already begun to explore such questions by examining brain maturation and puberty in early adolescence using deep learning brain age prediction models (Holm et al., *preprint*). Further, applying such methods in large longitudinal datasets like ABCD, will help the field better distinguish how early life experiences relate to pubertal timing, how this in turn shapes neurodevelopment, and the ways in which this may contribute to vulnerability to mental health difficulties during adolescence.

4.9.4 Limitations and future directions

The current study is not without limitations. Firstly, the puberty and depression measures used were both parent-report. We prioritised parent-report of youth pubertal development (over youth self-report) because youth have been found to over-report their pubertal development in the early stages of adolescence (Schlossberger et al., 1992). Further, there

was a large number of "I don't know" responses in the early waves of ABCD puberty data collection (Cheng et al., 2021). However, adolescent report may better capture the more intimate body changes associated with puberty, especially in the later pubertal stages (Dorn et al., 1990). Importantly, self- (and parent-) report measures of pubertal development assess the outcome of prolonged systemic hormonal effects and thus are limited in their ability to make inferences about the biological mechanisms relating pubertal timing to neurodevelopment (Goddings et al., 2019). Compared to physician assessments of pubertal development and picture-based measures like the Tanner Stages (Marshall & Tanner, 1969, 1970), there are number of measurement error considerations related to the PDS. For example, PDS responses mix rate of change and stage, such that someone experiencing rapid pubertal changes (i.e., tempo) might be more likely to select the response "definitely underway" compared to someone with a more protracted pubertal development. Similarly, the yearly interval between assessments in ABCD may mean that aspects of pubertal development may be described as "complete" even though further changes could occur later. As longitudinal data become available in ABCD, such considerations warrant attention so that we can map and interpret patterns of pubertal maturation as accurately as possible, while acknowledging the limitations of the measures available.

Akin to the limitations associated with parent-report of pubertal development, examining adolescents' self-report of depressive symptoms, and how this relates to parent report, is an important consideration for future work. Discrepancy between child and parent reports of psychopathology has been well documented (Achenbach, 2006; De Los Reyes, 2011) and suggests that parents may under-report youth depressive symptoms compared to youth self-report (Eg et al., 2018). Socio-environmental factors, such as family conflict and social cohesion have been found to be associated with greater and less reporter discordance, respectively (Kelly et al., 2016). Given the substantive body of research relating early life stress to both early pubertal timing and youth psychopathology (Colich & McLaughlin, 2022), multi-informant approaches should be incorporated into future analysis designs where possible.

Although the temporal distance between pubertal timing, brain structure, and youth depression measures was a strength of the analyses undertaken in the current study, due to

the varying availability of follow-up data, we did not examine any changes in our variables of interest between timepoints. Moreover, a brain structure previously found to mediate the association between earlier pubertal timing and depression in adolescence, namely, the pituitary gland (Whittle et al., 2012), was not available in the brain parcellations provided in the ABCD curated data release and was thus not tested as a potential mediator. We recommend using the raw ABCD imaging data to test the pituitary gland specifically as a mediator in further research. Further, subsequent work should also reflect the temporality inherent to development by examining domains such as pubertal tempo (the rate at which pubertal development occurs), and how this relates to both brain structural and functional changes across adolescence, as well as depressive symptom trajectories. Charting individual differences in development has gained increasing attention in recent times but longitudinal studies with multiple timepoints are necessary to generate developmental pathways (Bethlehem et al., 2022; Mills et al., 2021). For example, do differences in pubertal timing represent a stable risk factor that predicts the emergence of depression or do other factors (e.g., early life stress, loneliness) exert varying degrees of influence during adolescence.

4.9.5 Conclusion

The current study makes a significant contribution to our understanding of how pubertal development relates to psychopathology by directly testing an outstanding question in the field, as identified in a recent review paper Pfeifer & Allen (2021). That is, we examined a specific feature of brain development, brain structure, and tested whether it mediated the association between pubertal timing and depressive symptoms. However, central to the new conceptual model proposed by Pfeifer & Allen (2021) is the consideration of neural, social, and pubertal processes simultaneously and how they co-evolve and interact over time. Thus, the design of future studies should try to reflect this complex interplay of factors as much as possible. Longitudinal cohort studies like ABCD will be key to answering such research questions due to the large sample size and multi-modal nature of the measures collected. Further, ABCD could be used to replicate and extend the important multiverse findings from studies with smaller samples (e.g., Barendse & Byrne, 2021). Importantly, adopting a “team science” approach will be crucial to the success of this effort so that we can advance our understanding of the aspects of pubertal development that drive mental health

vulnerabilities during adolescence in a reproducible and collaborative manner (Zanolie et al., 2022).

4.10 Data access

N.M and X.S have had access to the ABCD annual curated data release 2.0.1 through their project entitled, "Brain structural associations with depression in a large early adolescent sample (the ABCD Study) (Shen et al., 2021). N.M and X.S also have access to the curated data releases 3.0 and 4.0. X.S looked at the baseline depression measures outlined in the current project. N.M has accessed the **baseline** puberty, depression, imaging, and socio-environmental variables outlined in the current project for the purposes of data quality control and pre-processing. N.M looked at year 1, year 2 and year 3 follow up data to determine sample sizes for the main analyses. Prior to the main analyses, the only statistical models run by N.M were those outlined in the pilot analyses. All scripts (R and Mplus) used in this registered report are available on the GitHub repository for this project: https://github.com/niamhmacsweeney/ABCD_puberty_depression .

N.M and all co-authors self-certify that they did not observe any of the statistical models outlined in the confirmatory analysis until after the in-principle acceptance was issued.

4.11 Acknowledgements

NM is the recipient of a Mental Health Research UK PhD studentship. HCW is supported by a JMAS SIM fellowship from the Royal College of Physicians of Edinburgh and by an ESAT College Fellowship from the University of Edinburgh. RMR is supported by the British Heart Foundation (Grant number: RE/18/5/34216).

Data used in the preparation of this article were obtained from the Adolescent Brain Cognitive Development^{SM} (ABCD) Study (https://abcdstudy.org), held in the NIMH Data Archive (NDA). This is a multisite, longitudinal study designed to recruit more than 10,000 children age 9-10 and follow them over 10 years into early adulthood. The ABCD Study® is supported by the

National Institutes of Health and additional federal partners under award numbers U01DA041048, U01DA050989, U01DA051016, U01DA041022, U01DA051018, U01DA051037, U01DA050987, U01DA041174, U01DA041106, U01DA041117, U01DA041028, U01DA041134, U01DA050988, U01DA051039, U01DA041156, U01DA041025, U01DA041120, U01DA051038, U01DA041148, U01DA041093, U01DA041089, U24DA041123, U24DA041147. A full list of supporters is available at https://abcdstudy.org/federal-partners.html. A listing of participating sites and a complete listing of the study investigators can be found at https://abcdstudy.org/consortium_members/. ABCD consortium investigators designed and implemented the study and/or provided data but did not necessarily participate in the analysis or writing of this report. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH or ABCD consortium investigators. The ABCD data repository grows and changes over time. The NDA DOI for this study is 10.15154/1523126. DOIs can be found at https://nda.nih.gov/general-query.html?q=query=studies%20~and~%20dataRepositories=Adolescent%20Brain%20Cognitive%20Development%20~and~%20orderBy=id%20~and~%20orderDirection=Ascending.

N.M attended the ABCD Modelling Developmental Change Workshop (grant number: R25MH125545) in July 2021 and would like to extend a special thank you to the organisers (Dr Kate Mills and Dr Michelle Byrne) and participants for their support in the development of this project.

4.12 CRediT statement

Niamh MacSweeney: Conceptualisation; data curation; formal analysis; investigation; methodology; validation; visualisation; writing – original draft; writing – review & editing; project administration.

Judith Allardyce: Formal analysis; methodology; writing – review & editing

Amelia Edmonston-Stait: methodology, writing – review & editing.

Xueyi Shen: Data curation; methodology; writing – review & editing

Hannah Casey: Methodology, writing – review & editing.

Stella W. Y. Chan: Writing – review & editing; funding acquisition

Breda Cullen: Methodology; writing – review & editing

Rebecca M. Reynolds: Methodology; writing – review & editing

Sophia Frangou: Writing – review & editing

Alex S. F. Kwong: Methodology; writing – review & editing; supervision

Stephen M. Lawrie: Methodology; writing – review & editing, supervision; funding acquisition

Liana Romaniuk: Methodology; writing – review & editing; supervision

Heather C. Whalley: Conceptualisation; methodology; writing – review & editing; supervision; funding acquisition

5 Irritability in Adolescent Depression — A Narrative Literature Review

5.1 Chapter introduction

The work contained within this thesis has thus far focused on brain structural associations with depressive symptoms in adolescence, and how other biological factors, such as early pubertal timing relate to depression risk. However, as previously discussed, an individual's vulnerability to depression arises, and unfolds, within a complex interplay of biological, psychological, and social processes. In addition to brain structural alterations, research has shown that differences in brain function are implicated in depression. In this chapter, I focus on a specific aspect of behaviour, irritability, which is an additional cardinal symptom of adolescent depression. High levels of irritable mood in childhood/adolescence predict later depression in adolescence and young adulthood. It has therefore been suggested that irritability could be an early indication of emotion regulation difficulties and thus, a promising intervention target. However, the neural underpinnings of irritability in adolescent depression remain underexplored and existing research methods typically overlook the social context in which irritability occurs during adolescence.

While many advantages accompany population-based cohort studies like ABCD (e.g., large sample size/increased statistical power), the breadth of measures available can come at the cost of phenotypic depth. Additionally, the measures available within a pre-existing dataset may not be suited to answering your research question of interest. For example, the measures available within ABCD did not allow me to address the limitations within the extant literature on the neural basis of irritability in adolescent depression. This highlights the continued value of small-scale studies, whereby researchers can curate a study that serves their specific research question, and importantly, employ methodologies that may not be as feasible within large-scale studies, such as co-production.

Thus, as a prelude to Chapters 6 & 7, this chapter contains a narrative review that brings together current research on irritability in adolescent depression and the associated neurobiology, and highlights directions for further research. Specifically, I draw attention to how existing studies on the neural basis of irritability typically use paradigms that overlook the social context in which irritability occurs. Given that adolescence is a time during which behaviour is significantly impacted by one's surroundings, I propose an innovative research design centred on co-produced research with young people. I argue that this creative approach will ensure that our research questions and methodologies accurately reflect the lives of young people, which will improve construct and ecological validity within the field. Given the significant mental health challenges faced by young people today, we urgently need to develop novel approaches to better understand adolescent depression and identify tractable targets for intervention.

The article contained within this chapter has been submitted for publication and is available as a pre-print on PsyArXiv. In line with the views expressed in our paper, we are delighted to have two youth researchers (P.L and S.Z) as co-authors on this manuscript who were central to the development of this project. Author contributions are included within the manuscript.

Keeping up with the kids: Towards a more ecologically valid study of irritability in youth depression and its underlying neural circuitry

Authors: Niamh MacSweeney[1*], Perrine Louvet[1†], Simal Zafar[2†], Stella W. Y. Chan[3], Alex S. F. Kwong[1,4,5], Stephen M. Lawrie[1], Liana Romaniuk[1], Heather C. Whalley[1].

[1] Division of Psychiatry, Centre for Clinical Brain Sciences, University of Edinburgh, UK

[2] School of Life Sciences, University of Glasgow, UK

[3] School of Psychology and Clinical Language Sciences, University of Reading, UK

[4]MRC Integrative Epidemiology Unit at the University of Bristol, UK

[5] Population Health Sciences, Bristol Medical School, University of Bristol, UK

[†]Youth-researchers

* Corresponding author: Niamh MacSweeney (niamh.macsweeney@ed.ac.uk)

Address: University of Edinburgh, Royal Edinburgh Hospital, Morningside Place, Edinburgh, EH10 5HF, UK

All authors have no conflicts of interest to declare.

5.2 Abstract

Irritability is a core symptom of adolescent depression, characterised by an increased proneness to anger or frustration. Irritability in youth is associated with future mental health problems and impaired social functioning, suggesting that it may be an early indicator of emotion regulation difficulties. Adolescence is a period during which behaviour is significantly impacted by one's environment. However, existing research on the neural basis of irritability typically use experimental paradigms that overlook the social context in which irritability occurs. Here, we bring together current findings on irritability in adolescent depression and the associated neurobiology and highlight directions for future research. Specifically, we emphasise the importance of co-produced research with young people as a means to improve the ecological validity of research within the field. Ensuring that our research design and methodology accurately reflect to lives of young people today lays a strong foundation upon which to better understand adolescent depression and identify tractable targets for intervention.

5.3 Literature review

Adolescence, a life phase spanning the ages 10-24 years, is a time of increased risk for the emergence of depressive disorders, which have a peak onset age of 19.5 years (Sawyer et al., 2018; Solmi et al., 2021). Importantly, rates of adolescent depression are rising. In the US, rates increased from 8.1% in 2009 to 15.8% in 2019 (Daly, 2022). Compared to adult-onset depression, adolescent-onset depression is associated with a more recurrent illness course and a host of physical and psycho-social difficulties with longer term consequences (Malhi & Mann, 2018; Thapar et al., 2012). It is therefore unsurprising that it is a leading cause of illness and disability for this age group (James et al., 2018). Taken together, these findings suggest that there are increasing unmet needs of adolescents with mental health difficulties (Wilson & Dumornay, 2022).

Unlike major depressive disorder (MDD) in adults, where low mood and anhedonia are primary diagnostic symptoms, irritability is considered an additional cardinal symptom specific to MDD in adolescence (American Psychiatric Association, 2013b). Irritability can be defined as low frustration tolerance and an overreaction to blocked goal attainment relative to same-age peers (Avenevoli et al., 2015; Stringaris et al., 2013). While this can represent a normative behaviour in adolescence, it becomes a pathological feature when associated with persistent functional impairment. Several behavioural studies demonstrate that high irritability in childhood and youth predicts later depression, suicidality, and impaired social functioning in adulthood (Leibenluft & Stoddard, 2013; Stringaris et al., 2013).This suggests that irritability may be an early indicator of emotion regulation difficulties and an actionable target for the prevention of downstream mental illness.

Existing definitions of irritability (i.e., proneness to anger/frustration) have shaped the primary experimental paradigms used in neuroimaging research on irritability, typically using frustrative non-reward and threat response tasks. Although these methodological approaches likely induce an irritated mood, the question remains as to whether they sufficiently tap into the broader social context in which the irritable mood occurs. This question is particularly pertinent when studying irritability in adolescence, a period during which mood and behaviour are heavily impacted by one's social environment (Blakemore &

Mills, 2014; Sawyer et al., 2018). Here, we therefore seek to draw together current findings on adolescent irritability and its underlying neurobiology, not as a formal literature review (available elsewhere, e.g., Lee et al., 2022) but as a means to discuss opportunities for future directions in this field. Specifically, we highlight the value of co-produced research with young people as a way to ensure that our research design and methodology accurately reflects the lives of adolescents (MacSweeney et al., 2019; Whitmore & Mills, 2021). Improving the ecological validity of our research will help maximise the chances of identifying tractable targets for intervention that are appropriate for today's youth.

5.3.1 Defining irritability in the context of depression

According to current psychiatric nosology, irritability can be categorised as being chronic or episodic (Leibenluft et al., 2006). Chronic irritability in adolescence represents a young person's baseline mood. Chronic irritability is considered a defining feature of disruptive mood dysregulation disorder (DMDD), whereby clinically significant irritability must have been present for at least 12 months (American Psychiatric Association, 2013b). Conversely, episodic irritability, which refers to changes from baseline mood, is more often seen in mood disorders such as depression and bipolar disorder (American Psychiatric Association, 2013b). The Diagnostic and Statistical Manual of Mental Disorders (Fifth Edition) (DSM-5) further distinguishes irritability into phasic and tonic. The former referring to behavioural outbursts of extreme anger from a high baseline, while the latter relates to angry mood lasting several days, months or weeks. Although typically described in chronic irritability, they may also be present within episodic irritability for the duration of the irritable mood (Vidal-Ribas & Stringaris, 2021). In terms of links between irritability and depression (see Vidal-Ribas & Stringaris, 2021 for further detail), the model with the most support is that of "shared risk factors". That is, shared risk factors including genetic risk, family history of depression, temperament characteristics, and negative parenting styles, influence both outcomes (Vidal-Ribas & Stringaris, 2021).

5.3.2 The neurobiology of irritability

While theoretical models of irritability can help contextualise factors associated with the emergence and development of irritability, identifying tractable targets for intervention requires an understanding of the neural mechanisms underpinning this salient and transdiagnostic marker of mental illness. The growing emphasis on adopting a translational neuroscience perspective is reflected in the significant increase in the number of studies published on the neural basis of irritability over the past decade (Lee et al., 2022; Nielsen et al., 2021). These studies have largely focused on exploring how irritability in childhood and adolescence, typically in clinical samples (e.g., youth with DMDD, attention deficit hyperactivity disorder (ADHD), internalising difficulties), relates to changes in blood-oxygen-level-dependent (BOLD) signal, measured via functional magnetic resonance imaging (fMRI).

5.3.3 Task fMRI studies

Task-based fMRI studies make up much of this relatively nascent field of research, which pivots upon three neurocognitive domains: threat processing/emotional reactivity, reward processing, and cognitive control, the latter of which is a much smaller body of work. Threat processing and reward processing constitute the two brain/behaviour pathways proposed by Brotman et al. (2017) in their translational neuroscience model of irritability (Brotman et al., 2017). Evidence for the threat processing pathway emerges from research which suggests that increased irritability is associated with an aberrant neural response in the amygdala, thalamus and insula, when youth are presented with emotionally threatening stimuli (e.g., angry or fearful faces) (Kryza-Lacombe et al., 2020; Tseng et al., 2016; Wiggins et al., 2016). Moreover, higher levels of irritability were found to be associated with more pronounced fluctuations in neural activation across task conditions (e.g., from congruent to incongruent trials), which may represent the additional effort required by youth with high irritability levels to process and respond to emotional stimuli in their environment.

Research supporting the reward processing pathway of irritability centres on frustrative non-reward tasks, whereby a frustrated psychological state is induced when the participant fails to receive a reward they have been conditioned to expect. Thus, the neural mechanisms of irritability are examined by inducing a frustrated state in real time and studying the associated

neural correlates. Using this rigged reward paradigm, research has shown that youth with high irritability exhibit aberrant neural responses in fronto-striatal regions, such as the prefrontal cortex (PFC), cingulate gyri, and caudate, compared to typically developing youth (Deveney et al., 2013; Perlman et al., 2015; Tseng et al., 2019). Further, a recent study by Scheinost and colleagues used connectome-based predictive modelling and found that during frustration trials, functional connectivity within motor-sensory, subcortical and salience networks, and between these networks and fronto-parietal networks, was associated with increased levels of irritability (Scheinost et al., 2021). The handful of studies that have examined the neural basis of irritability via cognitive control fMRI tasks, such as inhibitory control paradigms (e.g., stop signal task, Flanker task), suggest that youth with high levels of irritability exhibit inhibitory control deficits, which are reflected in aberrant patterns of neural activation in superior frontal gyri, temporal gyri, and the anterior cingulate cortex (Chaarani et al., 2020; Liuzzi et al., 2020).

Taken together, current paradigms used to examine irritability may be more likely to elicit fear or stress responses than more genuinely irritable ones, which highlights the need for tasks that aim to induce irritable mood specifically. Improving the construct validity of existing paradigms will have downstream effects on the ecological validity of the field by ensuring that our research methods capture the experience of youth irritability in the present day as accurately as possible.

5.3.4 Resting state fMRI studies

There are a limited number of resting state studies that have examined the neural correlates of irritability. In these studies, a questionnaire-based measure of irritability (e.g., the Affective Reactivity Index (ARI); Stringaris et al., 2012) is collected outside the scanner and then these behavioural measures are related to resting state imaging features. Most of the existing resting state studies focus on chronic irritability within the context of aggressive behaviour/temper outbursts in childhood-onset disorders such as ADHD, oppositional defiant disorder (ODD), and autism spectrum disorder (ASD) (Bennett et al., 2017; Gaffrey et al., 2021; Roy et al., 2018; Weathersby et al., 2019), as well as some examining irritability in DMDD and bipolar disorder (Stoddard et al., 2015). Similar to the task-fMRI irritability literature, these

studies suggest that the neural correlates of irritability comprise a diverse set of functional networks such as the default mode network (DMN), fronto-parietal network (FPN), executive control, sensory-motor, and visual networks (Nielsen et al., 2021). These networks support and coordinate cognitive processes associated with irritable mood, including self-referential behaviour (DMN), reward processing and emotion regulation (FPN, executive control network), and motor response (sensory motor network).

5.3.5 Next steps for fMRI-irritability research

A recent systematic review and meta-analysis by Lee et al. (2022) sought to determine whether there are convergent neural responses associated with irritability across the domains of threat, reward processing, and cognitive control (using task-fMRI studies only) (Lee et al., 2022). Interestingly, they found no evidence for convergence across these neurocognitive domains and posit that this may be due to the marked heterogeneity in clinical characteristics, task design, irritability measures, analysis methods, small sample sizes, and a lack of longitudinal research. These limitations also extend to the resting state irritability research (Nielsen et al., 2021). Taking these findings together, it remains unclear whether brain mechanisms underpinning irritability vary across disorders (Eshel & Leibenluft, 2020). Some evidence suggests that individual differences in dispositional (i.e., chronic) irritability may be more underpinned by amygdala-DMN connectivity than state (i.e., episodic) irritability, due to more consistent findings in this neural circuity in resting state (Fulwiler et al., 2012; Gaffrey et al., 2021) compared to task-based studies (Kryza-Lacombe et al., 2020; Stoddard et al., 2017). However, more research, ideally combining resting state and task-based paradigms in larger samples with harmonised protocols is needed.

The heterogeneity present across multiple domains (e.g., in samples and methodologies) sheds light on several important considerations, especially research involving developmental samples. Probing individual differences in irritability will allow us to better understand its bounds as a normative behaviour across development, and what might reflect concerning irritable mood. How this relates to other cognitive processes, such as emotion regulation, and the associated underlying neural circuitry, will pave the way forward for targeted intervention strategies. Further, there is an overall paucity of research on age-related (and sex-related)

changes associated with irritability — longitudinal and sufficiently powered cross-sectional studies that examine age interaction effects are also needed. It may be that the neural underpinnings of irritability vary across development and are related to other typical, as well as divergent, neurodevelopmental changes. Emerging research on brain growth charts for the human lifespan will be helpful in this effort (Bethlehem et al., 2022). Moreover, few studies have examined irritability in later adolescence (only 4/30 studies in the Lee et al. review had a mean age >15 years). Given the varying age of onset for mental health difficulties during adolescence (Solmi et al., 2021) and the distinct neuromaturation that characterises this life phase (Bethlehem et al., 2022), future studies should be designed in a developmentally sensitive way. Some large, longitudinal youth cohort studies such as IMAGEN and ABCD, include variables related to irritable mood alongside neuroimaging data, and have already contributed to our understanding of the neurobiology of youth irritability and psychopathology (e.g., Chaarani et al., 2020, using IMAGEN data).

While the large sample sizes of such cohort studies are well powered to detect more subtle effects (e.g., individual differences in irritability, underlying neural circuitry, and potential contributing factors), the breadth of measures included in such studies comes at the cost of phenotypic depth. For example, ABCD and IMAGEN do not include an irritability-specific questionnaire like the ARI. Instead, a measure of irritability is derived from individual items in broad mental health measures (e.g., DAWBA, CBCL, K-SADS). Thus, rather than a “panacea” to the many unknowns in developmental cognitive neuroscience, cohort studies may be better conceptualised as hypothesis generating tools that can inform directions for future studies (Saragosa-Harris et al., 2022). To develop a finer-grained characterisation of irritability, and the functional significance of altered neural circuitry — especially how it relates to psychopathology — we need construct-specific and ecologically valid experimental designs. Ideally, these designs would involve harmonised protocols across studies to minimise sources of error as much as possible. Initiatives like the ENIGMA Irritability Working Group are leading by example in this way.

In sum, the surge of studies on the neurobiology of irritability over the past decade has allowed us to outline the brain networks involved in this transdiagnostic symptom from which

myriad directions for future research have emerged. Before embarking upon these new avenues of research, we should reflect on *how* we plan to move forward to ensure that our journey takes us towards the world of young people rather than away from it.

5.3.6 The value of co-produced research

Co-produced research, whereby the target population of the study (e.g., adolescents) are involved in as many steps of the research project as possible, has gained increasing traction in recent times (https://www.ukri.org/about-us/policies-standards-and-data/good-research-resource-hub/research-co-production/). Initiatives like Young Persons' Advisory Groups (YPAGs) allow young people to be involved in research in an active, meaningful, and mutually beneficial way. As co-researchers, young people and researchers can work together to ensure that the research questions, methods, and dissemination of research findings are relevant to the lives of young people today (MacSweeney et al., 2019). Although co-produced research involves a considerable (and front-loaded) time investment, researchers should approach it like other best practices in research, such as open science (Whitmore & Mills, 2021). Transparent and rigorous research that is attuned to the issues and experiences of today's young people will be key tools in our effort to answer complex questions in developmental science (e.g., why is adolescence a period of significant vulnerability to the onset of mental health difficulties?). Thankfully, resources are now available to help researchers undertake effective and meaningful co-created research (Whitmore & Mills, 2021).

5.3.7 Towards a more ecologically valid study of irritability

Although there have been important commentaries on the neuroscience of irritability (Eshel & Leibenluft, 2020), the social context in which irritable mood occurs has been largely overlooked. Irritability occurs in social and interactive contexts between young people. However, existing fMRI paradigms like frustrative non-reward tasks and emotional faces tasks, may not appropriately capture the rich social tapestry of adolescence. To enhance both construct and ecological validity, future research on irritability should incorporate social context into the study design. For example, Lee et al., (2022) propose a frustrative social non-reward task that targets behaviours like social rejection. This work would complement

existing research on social exclusion in adolescence, which has used socially relevant tasks like Cyberball (Sebastian et al., 2010; Williams et al., 2000). Given that avoidance of social rejection drives adolescent decision-making and behaviour (Tomova et al., 2021), this research could provide novel insight into how the nuances of the adolescent social world relate to the emergence and development of irritability and related mental health difficulties. Importantly, this effort to align our research methods with the social world of adolescence could be strengthened even further by undertaking research that is co-produced with young people.

By asking young people questions like, "*What situations do you find irritating in your daily life*?", we could design studies that better reflect the experience of irritability as a young person. In turn, this could help disentangle the current heterogenous findings in irritability research. As mentioned by Lee et al. (2022), these insights could be incorporated into task-based fMRI, but there is also opportunity for "hybrid" resting-state paradigms. Recent calls for a "third-wave" of fMRI research propose the use of integrated fMRI paradigms, whereby task-like manipulations are paired with "traditional" resting state approaches (Finn, 2021). Naturalistic stimuli (e.g., movie watching) are some examples of integrated fMRI paradigms (Sonkusare et al., 2019), which allow researchers to regain some degree of experimental control, while acknowledging the dynamic patterns of brain function that arise from self-generated activity. Further, pairing these integrated paradigms with analyses capable of capturing fine grained temporal details, such as dynamic functional connectivity analysis, warrants consideration going forward. It has been argued that progress in our understanding of the human brain and behaviour is likely to emerge from these "third-wave" paradigms (Finn, 2021). However, this progress will be hampered if the paradigms are not ecologically valid. Co-produced fMRI paradigms will ensure that the construct of interest is studied in a way that reflects real-world experience. For example, when studying youth irritability, we could ask young people to come up with irritating scenarios based on their own experiences. These scenarios would then form the stimuli for an integrated fMRI paradigm, asking young people to read each irritating scenario and imagine the experience as vividly as possible while in the scanner. This protocol would be suited to a range of samples (e.g., healthy volunteers,

young people with mental health difficulties) but could also be adapted to suit different sample characteristics and research questions.

Importantly, novel paradigms like this would need to be validated against traditional task-based irritability paradigms (e.g., frustrative non-reward and threat response tasks) as well as behavioural measures of irritability (e.g., ARI). Given the lack of convergence in the neural correlates of irritability across neurocognitive domains (Lee et al., 2022), a novel, co-produced integrated fMRI task with improved ecological validity, holds great promise as way to better our understanding of youth irritability, identify tractable intervention targets, and move young people away from illness towards wellbeing.

5.4 CRediT statement

Niamh MacSweeney – Conceptualisation, Methodology, Formal analysis (literature review), Investigation, Writing – Original Draft, Writing – Review & Editing, Project administration.

Perrine Louvet — Conceptualisation, Formal analysis (literature review), Writing – Review & Editing.

Simal Zafar — Conceptualisation, Writing – Review & Editing.

Stella Chan — Conceptualisation, Writing – Review & Editing, Funding Acquisition.

Alex Kwong — Writing – Review & Editing.

Stephen Lawrie — Writing – Review & Editing.

Liana Romaniuk — Conceptualisation, Writing – Review & Editing, Supervision, Funding Acquisition.

Heather Whalley — Conceptualisation, Writing – Review & Editing, Supervision, Funding Acquisition.

5.5 Chapter conclusion

The literature reviewed and sentiments expressed in Chapter 5 directly inform the work in Chapters 6 & 7 of this thesis. In Chapter 6, I give an overview of the methods employed in our pilot study that adopted a co-produced youth-researcher design to develop a novel integrated fMRI paradigm that reflected the social nature of adolescence. Although our novel paradigm can be described as a "hybrid/integrated" paradigm because it comprises a "standard" resting state scan with a task-like manipulation, it is referred to as the "irritability task" hereafter in this thesis. I report the findings from this co-produced pilot study in Chapter 7.

6 Irritability in Adolescent Depression — Pilot Study Methods

6.1 Chapter introduction

This chapter introduces the data and methods used for Chapter 7 of this thesis. Firstly, I outline the source of the data. I then describe the various data types used including the fMRI task development, depression, and irritability measures. I also outline the scanning protocol, quality control, and pre-processing methods implemented. Finally, I give an overview of the statistical methods used to analyse this data, namely, Leading Eigenvector Dynamics Analysis (LEiDA; Cabral et al., 2017).

Many thanks to Liana Romaniuk, Laura Klinkhamer, and Hannah Casey for assisting with the quality assessment of the fMRI data.

Supplementary information for this chapter can be found in Appendix 3.

6.2 Data source

The data used for this chapter came from a pilot study entitled, *Development of novel neuroimaging markers for the detection of adolescent depression*. I led the recruitment and data collection for this study between November 2020 and May 2021, which was supported by Kimberley Atkinson through the NRS Mental Health Network. Data collection took place at the Clinical Research Imaging Centre (CRIC), University of Edinburgh. I also led the development of the co-produced irritability task in Summer 2020, which is described in Section 6.5.1 and in the Supplementary Information for this chapter.

6.2.1 Recruitment

Participants were recruited through schools and universities, third sector organisations (e.g., charities and youth groups), social media and via snowball sampling in the greater Edinburgh

area. Identified organisations were contacted by the study team and asked if they would distribute information about the study to 16–20-year-olds in the organisation. If an organisation agreed to participate, standard recruitment materials were provided and circulated by the organisation via mailing lists, social media, posters, and word of mouth. Interested individuals contacted the study team directly and were given a study information sheet. They then completed a brief online screening questionnaire, and if eligible, their study appointment was scheduled. All participants provided informed consent before taking part in the study. Participants were also given a picture of their brain for taking part in the study and travel costs were covered to increase the accessibility of the study. Recruitment materials can be found in the Supplementary Information.

6.2.2 Participants

In total, 30 participants aged 16-20 years (M_{age} = 18.86 years; SD = 0.83; 77.4% female) with self-reported depression took part in this study. Eligible participants were aged 16-20 years, free from MRI contraindications (e.g., dental braces, pacemakers), fluent English speakers, did not report a past or current clinical diagnosis of autism spectrum disorder, a neurological or genetic disorder, or known intellectual disability, and had a Mood and Feelings Questionnaire (MFQ; short version) total score ≥ 8. A total MFQ score ≥ 12 indicates the presence of depression in the respondent (Burleson Daviss et al., 2006; Costello & Angold, 1988). As our aim was to recruit a sample with a range of depressive symptoms, we selected a lower MFQ cut-off score ≥8 (range = 8 to 23, mean = 16.52, SD = 4.20) to ensure the presence of mild and moderate symptoms in our sample. The MFQ was used for screening purposes only — the main depression outcome measure (PHQ-9) for this study is described later in this chapter.

There was a data transfer error for one participant's imaging data, which excluded their data from further analysis and resulted in a final sample of N = 29.

6.3 Funding and ethics

This study was funded by a Wellcome Trust Institutional Strategic Support Fund grant (PI: H Whalley). The study protocol was approved by the Edinburgh Medical School Research Ethics Committee (Reference: 19-HV-061) in November 2020. Our ethics approval letter can be found in the Supplementary Information.

6.4 Study procedure

The study appointment lasted approximately 2.5 hours and comprised a 50-minute MRI scan, which was followed by a battery of questionaries (hosted on Online Surveys). Participants were offered a break after the scan and given water and a light snack.

6.4.1 Scanning protocol

A 3T Siemens (model: Magnetom Skyra Fit) MRI scanner with a 32-channel head coil was used for data acquisition at the CRIC, University of Edinburgh. Inflatable pads were used to immobilise the participant's head. The scanning procedure consisted of a T1-weighted sequence that yielded 192 contiguous 1.0mm slices (matrix = 256 x 256; FoV = 256mm; flip angle = 7°). This was followed by a functional imaging protocol using an axial gradient echoplanar imaging pulse sequence (EPI) [TR = 1400ms; TE = 30ms; matrix = 70 x 70; FoV = 210mm; flip angle = 68°, spatial resolution = 3mm isotropic). Sixty contiguous 3mm slices were collected during each TR using 2x GRAPPA acceleration. The functional imaging protocol consisted of two resting state scans, each lasting 6 minutes. The first scan was a standard resting state scan whereby the participant was asked to focus on a white cross on a dark screen. The second resting state scan was our novel irritability task whereby participants were asked to read a series of 18 irritating scenarios presented on the screen one at a time and to imagine being in each scenario as vividly as possible. Each scenario was presented for a period of 20 seconds. The development of this task is outlined in Section 6.5.1.Tasks were presented using a screen in the bore of the magnet and run using Presentation® software. After the resting state scans, participants completed a reward task and a self-referential recall task, but these data are not included in this thesis.

6.5 Study materials

6.5.1 Irritability task

As discussed in Chapter 5, existing research on the neural basis of irritability use experimental paradigms that often overlook the social context in which irritability occurs. Therefore, we aimed to improve the ecological validity of youth irritability neuroimaging research by developing a novel fMRI task that better reflected the experience of irritability as a young person today. To achieve this aim, we adopted a co-produced youth-researcher design whereby young people were involved in multiple stages of the task development, as outlined below.

In Stage 1 of the task development, a series of scenarios were generated by an independent group of youth (N = 25) at science communication and outreach events in the Edinburgh area. Specifically, young people were asked to complete the following sentence based on their own life experience and thoughts: *"I find it irritating when…."*. These scenarios were collected via a laptop/pen and paper. Young people were verbally asked by the research team whether they were aged 16-18 years before they completed the sentence prompt on scenarios that they found irritating. Some young people listed more than one scenario. We did not collect any identifying information from the young people, and we recorded all scenarios collected regardless of whether multiple scenarios were provided by the same young person.

In Stage 2 of the task development, the research team reviewed these scenarios, removed any duplicates, and corrected the scenarios for spelling and grammar. Given the prompt (i.e., *"I find it irritating when…"*) that we used in Stage 1, all the scenarios were similar in sentence length. However, on a few occasions, the verbatim phrasing provided by the young person could have been more succinct or an alternative sentence structure worked better. For example, "I find it irritating when I do housework" was changed to "*I find housework irritating".* We thus made minor edits so that the sentence length was as consistent and coherent as possible across the scenarios. We also provided some additional contextual detail in some scenarios so that they were easier to understand/imagine.

This resulted in a final sample of 51 irritating scenarios which were compiled into an online survey and rated by an independent group of young people (N = 61, aged 16-18 years). The mean irritability rating for each of these scenarios can be found in Table S1 in the Supplementary Information for this chapter.

In the final stage of task development, we asked an independent group of young people (N = 61; aged 16-18, also separate from study sample), to rate these scenarios on a 5-point scale where 1 = “not at all irritating” and 5 = “very irritating”. Scenarios were compiled into an online survey and distributed to young people via social media and local youth groups. We worked with a youth researcher (Simal Zafar) as part of a Nuffield Summer Research Placement, who led the initial distribution of this survey. The survey remained live for a period of 5 weeks between August and September 2020. The 18 most-highly rated scenarios were chosen as stimuli for the irritability task (see Table 6.1). The task was programmed in NBS Presentation (version 19) by Liana Romaniuk.

No.	Scenario	Rating (out of 5)
1	You're having a conversation with someone, and are trying to offer helpful advice, but they rudely keep talking over you and don't listen to what you're saying.	4.53
2	You're talking with your parents, and it’s obvious that they're not taking you seriously and are just patronising you.	4.23
3	You're doing something fun online, and the WiFi keeps disconnecting.	3.67
4	You're trying to get somewhere important, but you're stuck behind someone walking slowly.	3.36
5	You and your friends are having a private conversation. Your parents listen in but only hear part of it, and then get angry with you because they didn't hear and understand the whole thing.	3.89
6	You're in your room busy looking at your phone, and then someone just barges in and starts touching your things.	3.59
7	You're in one of your favourite shops, and you see someone being really rude to someone who works there, who is obviously doing their best.	4.46
8	Someone you thought you could trust tells you a blatant lie.	4.33
9	You've been planning to sort out your room today. Just when you're about to start, your parent tells you to sort your room out.	4.28
10	Your parent asks you to do the washing up for the third time, even though you've already said you'll do it.	3.93

11	You're choosing the subjects you'd like to study, and your parent just tell you what you "should" be studying, without listening.	3.56
12	You're not in a good mood, but you're forced to go to a family event and make awkward conversation.	3.11
13	A teacher expects you to complete a certain task, but you don't know how to do it.	3.98
14	Everyone expects you to know what you want to do in life, even though you are still a teenager.	3.66
15	You've been preparing for a test that means a lot to you, but on the day, you get stressed out and it doesn't go as well as you'd hoped.	4.06
16	There's something you wanted to do today, but you just can't get motivated for some reason.	3.77
17	You're eating out somewhere nice, but one of your friend's friends is eating with their mouth open.	3.68
18	It's been a long day, but you're stuck waiting somewhere. It's too hot, and now you feel hungry.	3.12

Table 6.1 — The 18 most highly rated scenarios that were used as stimuli in our irritability task.

Note: Scenarios are listed in random order here, and this order was used in the task across all participants.

6.5.2 Depressive symptom measure

Depressive symptom severity was measured using the depression module of the Patient Health Questionnaire (PHQ-9; Kroenke et al., 2001). The PHQ-9 is a self-administered 9-item questionnaire based on the Diagnostic and Statistical Manual (4th Edition; DSM-IV). It assesses the present degree of depression in individuals by asking them to rate each symptom criteria on a scale of "0" (not at all) to "3" (nearly every day). The PHQ-9 has been validated across a range of clinical settings and samples as a screener for Major Depressive Disorder (MDD) as well as being a measure of depressive symptom severity in both adults (Levis et al., 2019) and adolescents (Richardson et al., 2010).

6.5.3 Irritability measure

Youth irritability was measured using the Affective Reactivity Index (ARI) self-report questionnaire (Stringaris, Goodman, et al., 2012). The ARI is a six-item scale that examines the frequency, duration, and threshold for irritable mood over the past six months. The ARI also contains a seventh, impairment-of-functioning item. Each item is rated on a 3-point Likert scale ranging from "0" (not true) to "3" (certainly true). Items include "easily annoyed by others", "get angry frequently", and "stay angry for a long time". The ARI has demonstrated good internal consistency (Cronbach's α values ≥.80) and construct validity in both clinical and community samples (Stringaris, Goodman, et al., 2012; Tseng et al., 2017). The ARI is the most commonly used dimensional measure of irritability and has been used to quantify youth irritability in a number of neuroimaging studies (Liuzzi et al., 2020; Scheinost et al., 2021; Stoddard et al., 2017; Tseng et al., 2019). Of note, the ARI is thought to assess trait-like irritability rather than state-like irritable mood, the former of which is regarded as less context dependent.

6.6 Imaging data pre-processing

6.6.1 Introduction to HALFpipe

The *H*armonised *A*na*l*ysis of *F*unctional MRI *pipe*line (HALFpipe) version 1.2.1 was used for the pre-processing, quality assessment, and single-subject feature extraction of the imaging data used in Chapter 7 of this thesis. HALFpipe is a recent open-source and user-friendly tool developed by the Enhancing NeuroImaging Genetics through Meta-Analysis (ENIGMA) consortium (Waller et al., 2022). Given the reproducibility crisis in the field of neuroimaging (Gorgolewski et al., 2016; Poldrack et al., 2017) and many others (Baker, 2016), there has been increasing demand for the adoption of standardised pre-processing pipelines. Analysis pipelines are particularly key to fMRI research as myriad computational operations must be applied to fMRI data to generate interpretable results. Thus, there are many software tools available to carry out the necessary algorithmic processing and statistical modelling of fMRI data, such as SPM, AFNI, and FSL. Within these tools, there are many researcher "degrees of freedom" that can be applied at each step of the analysis pipeline, termed *analytic flexibility* (Poldrack et al., 2017). Although the rapid development of these computationally advanced

software tools has propelled the field of neuroimaging forwards, the abundance of choice and analytic flexibility has generated inconsistent findings, which can have significant effects on scientific conclusions (Botvinik-Nezer et al., 2020).

In an effort to improve reproducibility within neuroimaging, a series of default parameters (e.g., spatial smoothing thresholds, Gaussian filtering) have been established based on empirically-derived best practices (Grüning et al., 2018). This has been employed in well-established pipelines such as fMRIPrep, whereby a series of software tools are used together for different elements of the analysis workflow (Esteban et al., 2020). fMRIPrep has become the front runner in the field due to its adoption of open research best practices, namely, open-source availability, pleasant user experience, as well as its "glassbox" principle of transparency, so that the user gains an understanding of what is going on at each step of the pipeline. However, fMRIPrep is mostly limited to the pre-processing of fMRI data and therefore analytic flexibility remains for parameter selection of postprocessing analyses, such as feature extraction and model specification.

HALFpipe builds upon fMRIPrep by offering a standardised workflow from raw fMRI data through to feature extraction and group-level statistics. HALFpipe also includes additional features such as conversion of raw fMRI data to BIDS format, spatial smoothing, temporal filtering, advanced confound regression, and feature extraction. The HALFpipe workflow is shown in Figure 6.1. Like fMRIPrep, the HALFpipe software is containerised (e.g., Docker), which means that it comes packaged with all the software tools required for it to run, such as fMRIPrep, FSL, ANTs, FreeSurfer, and AFNI. This means that all users of a specific HALFpipe release will be using identical versions of the software tools housed within the container, which will aid reproducibility across different research and computing environments.

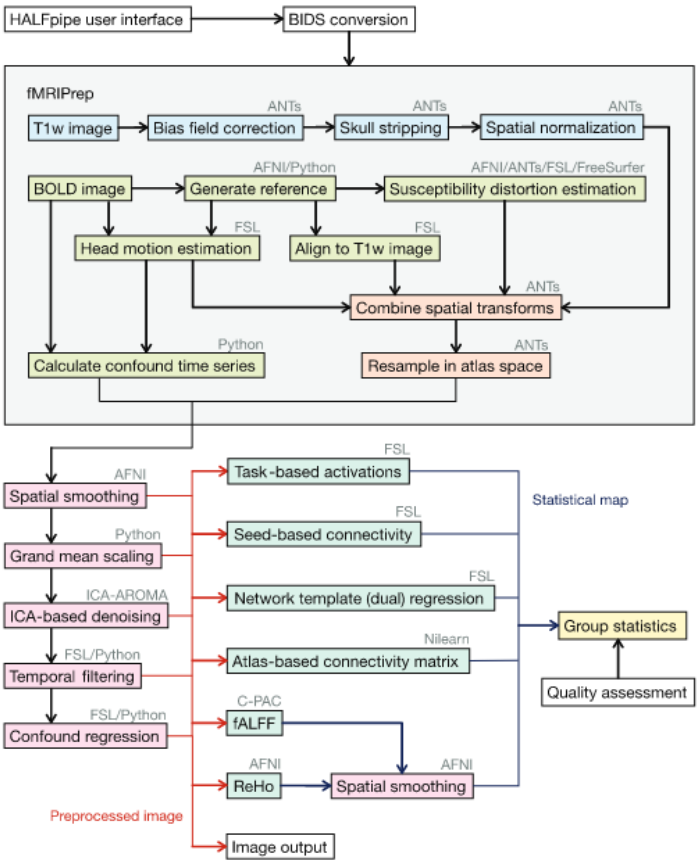


Figure 6.1 — HALFpipe Workflow. HALFpipe is configured in a user interface where the user is asked a series of questions about their data and the processing steps to perform. Data are then converted to BIDS format to allow standardised processing (white). After minimal preprocessing of the structural (blue) and functional (green and orange) data with fMRIPrep, additional preprocessing steps can be selected (pink). Using the preprocessed data, statistical maps can be calculated during feature extraction (turquoise). Finally, group statistics can be performed (yellow). Figure taken from Waller et al. 2022. The caption included here has been adapted slightly from the original paper.

6.6.2 Pre-processing specifications

Within HALFpipe, the main pre-processing of the imaging data was done with fMRIPrep. To note, six volumes were collected before the start of each scan to achieve T1 signal saturation, so it was not required to discard any additional volumes. First, the anatomical T1w images were skull-stripped via the *antsBrainExtraction.sh* function (from ANTs) using OASIS (Open Access Series of Imaging Studies) as a target template. The structural images were then normalised to standard space via the *antsRegistration* template (from ANTs), which HALFpipe defines as MNI152NLin2009cAsym (the most current and detailed template available) (Horn,

2016). HALFpipe then carries out additional pre-processing steps that denoises, filters and harmonises the functional data.

Firstly, the functional data was denoised by Independent Component Analysis-based Automatic Removal of Motion Artefacts (ICA-AROMA), which employs an independent component analysis algorithm that classifies components as signal or noise, with high accuracy and robustness (Pruim et al., 2015). ICA-AROMA requires a reference template that is defined in standard space template MNI152NLin6Asym, which is different to that used by fMRIPrep. HALFpipe implements a different approach to fMRIPrep to deal with the two different standard space templates, by using a pre-existing warp between the two standard templates (Horn, 2016), rather than estimating a second normalisation to the other template. Compared to fMRIPrep, this halves the processing time and avoids the need to potentially manually check both spatial registrations. ICA-AROMA was then performed on the resulting fMRI images in MNI152NLin6Asym space using the *ica_aroma_wf* workflow from fMRIPrep.

Next, grand mean scaling was applied where an image mean (the within-scan mean across all voxels and timepoints) is set to a predefined value of 10,000, which is calculated from the masked functional image. As the grand mean relates to scanner parameters such as amplifier power rather than neural mechanisms, grand mean scaling makes analysis results more interpretable and comparable across participants, sessions, and sites (Gavrilescu et al., 2002).

The noise components that were previously estimated using ICA-AROMA were then removed from the smoothed and grand-mean-scaled data using FSL's function *fsl_regfilt*. This method minimises removing shared variance between signal and noise components by calculating an ordinal least squares regression for each voxel, where signal and noise components are both included as regressors in the design matrix. As a result, the regression weights represent the unique variance of the noise components rather than the shared variance with the signal components. Next, a noise time series is calculated by multiplying the noise components by their regressor weights and adding them together. A denoised time series is then yielded by subtracting the noise time series from the voxel time series. Finally, a Gaussian-weighted

temporal filter set at 128s FWHM (HALFpipe's default) was applied to remove low-frequency drift via FSL *Feat.*

6.6.3 Quality assessment

Although there have been many recent efforts to automate the quality assessment (QA) of fMRI data via machine learning methods (Esteban et al., 2017) or by predefined image quality thresholds (Alfaro-Almagro et al., 2018), these methods have not yet reached a stage where they can replace the eyes of trained researchers. Given the relatively modest sample size of our dataset, manually quality assessing the data was not a large undertaking. This process was greatly aided by the interactive web app provided by HALFpipe via a single HTML file. Within this web app, reports are provided for each participant based on several pre-processing steps such as T1 skull stripping and normalisation, BOLD temporal signal-to-noise ratio (tSNR), motion confounds, ICA-based artefact removal, and spatial normalisation (see Figure 6.2). The images can be rated as "good", "uncertain" or "bad" according to detailed explanations outlined in the ENIGMA HALFpipe quality assessment manual at https://github.com/HALFpipe/HALFpipe#quality-checks.

Figure 6.2 shows an example of the quality control checks that were undertaken for each participant:

A. T1w skull stripping: We see a bias-field corrected anatomical image with an overlay of the brain mask, indicated by the red line. We checked to see if any brain regions were missing from the mask and to make sure that parts of the skull were not included in the mask.

B. T1w spatial normalisation: We see an anatomical image that has been resampled in standard space with an overlay of a brain atlas in standard space. We checked to see that the atlas regions closely matched the resampled image.

C. Echo planar imaging (EPI) tSNR: We see the temporal signal-to-noise ratio of the functional image post fMRIPrep pre-processing. We checked to determine whether the signal was distributed evenly across the brain, and whether there was any localised drop-out, motion striping artefacts, distortion, or asymmetry.

D. EPI confounds: Here, we see a carpet plot produced by fMRIPrep. This is a two-dimensional plot of the timeseries within a scan, with time shown on the x-axis and voxels on the y-axis. Voxels are categorised into grey matter (blue), subcortical grey matter (orange), cerebellum (green), and white matter (WM) and cerebrospinal fluid (CSF) (red). Above the carpet plot, we are given time courses of magnitude of framewise displacement, global signal, CSF global signal, WM global signal, and DVARS (the temporal change in root-mean square intensity). We looked for abrupt changes in the intensity of the heatmap based on the motion and signal parameters mentioned above, which may represent motion spikes. We also checked for extended signal changes which could have indicated artefacts caused by scanner hardware defects.

E. EPI ICA-based artefact removal: This report shows the time course of the average signal extracted from each ICA-component, which has been classified as either signal (green) or noise (red). A spatial map, timeseries and power spectrum has been created for each component. We checked that the components classified as noise did not correspond to known brain networks (e.g., DMN, salience network) or temporal patterns representing signal.

F. EPI spatial normalisation: We are shown the functional image post fMRIPrep pre-processing overlaid with a brain atlas in standard space. We checked that the atlas regions closely matched the functional image.

I undertook the initial QA and Laura Klinkhamer independently reviewed all the images. Further, a random sample of each rating band (good, uncertain, bad) were reviewed by Liana

Romaniuk and Hannah Casey. All reviewers had experience with structural and functional neuroimaging quality control. Given the overall good quality of the data (there were no bad ratings, and only a handful of participants were rated as "uncertain" on the EPI confounds checks), there was 100% consensus on the QA by all four raters. We used an average mean framewise displacement value of <.25mm as the motion cut-off for inclusion, which is a moderate to conservative threshold and in line with protocols employed by large scale neuroimaging studies (e.g., ABCD) (Hagler et al., 2019). Following QA, no participants were excluded from further analyses. QA statistics for participants are included in Tables S3 (irritability condition) & S4 (rest condition) in the Supplementary Information.

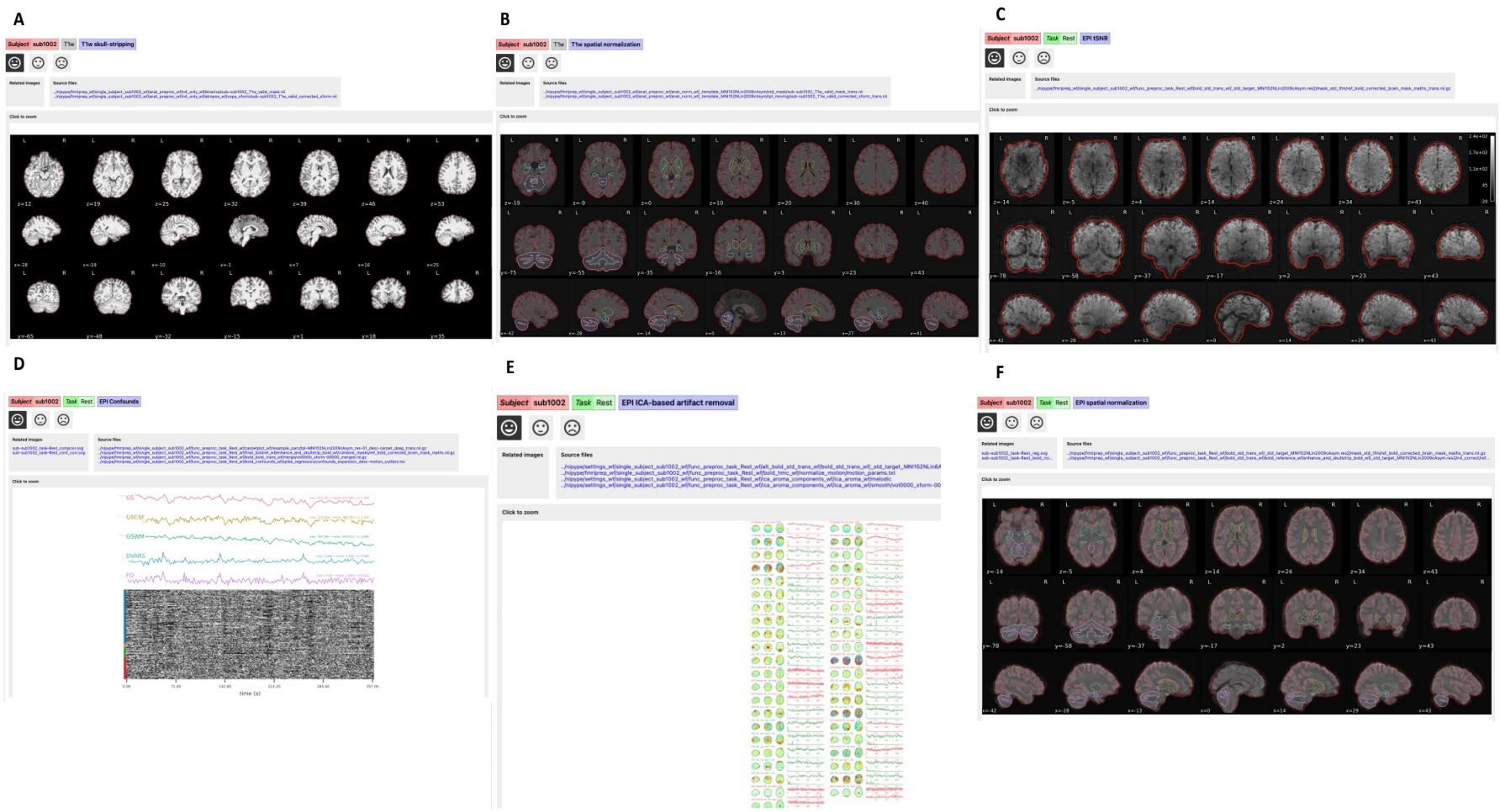


Figure 6.2 — An example of the HALFpipe interactive web-based quality assessment (QA) tool for an individual participant. This participant has a rating of "good" across all quality control check categories. A: T1w skull stripping; B: T1 spatial normalisation; C: Echo planar imaging (EPI) temporal signal-to-noise ratio (tSNR); D: EPI confounds; E: EPI Independent component analysis (IPA)-based artefact removal; F: EPI spatial normalisation.

6.6.4 Feature extraction

Although HALFpipe can extract several features that are commonly used in both task and resting state fMRI analysis, our analysis focused on the atlas-based connectivity matrix feature as this was the required input for the LEiDA approach, described below. Of note, spatial smoothing is not applied in HALFpipe for this feature extraction. Spatial smoothing is not recommended for network analysis with fMRI data, as it has been shown to influence the identity of hubs in functional brain networks (Alakörkkö et al., 2017).

Average time series were extracted from each region based on the Automated Anatomical Labelling Atlas 2 (AAL2; Rolls et al., 2015; Tzourio-Mazoyer et al., 2002). AAL atlases are widely used in resting state neuroimaging research and parcellate the human brain according to a spatially normalised high resolution single subject T1w image provided by the Montreal Neurological Institute (MNI). AAL atlases use a sulci-based approach to parcellate the brain whereby the main sulci are first delineated and then used as landmarks to define and label anatomical regions in 3D space. Using specially designed software, these anatomical regions were manually traced every 2mm on each axial slice from the MNI single subject. This resulted in a 3D reconstruction of the human brain which was parcellated into 120 (for the AAL2) anatomical regions.

6.7 Leading Eigenvector Dynamics Analysis (LEiDA)

A newly developed data-driven approach led by Dr Joana Cabral, Leading Eigenvector Dynamics Analysis (LEiDA; Cabral, Vidaurre, et al., 2017) was used to examine time-varying functional connectivity (FC). Central to this method is an understanding that the brain is a dynamic complex system. Research suggests that different groups of brain areas exhibit patterns of spontaneous correlated activity, which give rise to intrinsic functional brain networks (Cabral, Kringelbach, et al., 2017; Yeo et al., 2011). Thus, brain activity (i.e., BOLD time-series) can be expressed as a repertoire of coupled dynamical units or phase-locking mechanisms (phase-locking is a kind of synchrony in systems organised in time and space). On a high-level, LEiDA detects recurrent patterns of phase coherence, termed phase-locking (PL) states, and quantifies properties of these states, such as how often they are likely to be

occupied (probability of occurrence), and for how long (duration/dwell time). These PL states have been shown to map onto known resting state networks and relate to cognitive and emotional processing (Alonso Martínez et al., 2020; Cabral, Vidaurre, et al., 2017; Figueroa et al., 2019; Lord et al., 2019).

Below, the theory underpinning the LEiDA approach is described and illustrated in Figure 6.3. To allow other researchers to apply LEiDA to their own datasets, Cabral and colleagues have provided open-source MATLAB code and software tutorials, which were used in the current thesis (LEiDA GitHub repository: https://github.com/PSYMARKER/leida-matlab). As reflected in the LEiDA description below, a 100-region (versus 120) brain parcellation was used in our final analyses after regions with NaN values in the AAL2 connectivity matrix derived from our imaging data were removed. This is further described in Chapter 7, alongside the results from this analysis.

6.7.1 Dynamic BOLD Phasing Locking (PL) Analysis

For each condition (rest and irritability), a NxT BOLD dataset was derived where N = 100 is the number of brain regions and T= 257 is the number of repetition times (TR).

The first step in the LEiDA workflow involves band-pass filtering the BOLD signal between 0.01 and 0.1 Hz before computing an analytic BOLD signal phase, $\theta(n,t)$, for each brain region at each TR using the Hilbert Transform. The Hilbert transform expresses a given signal x in polar coordinates, where A represents the time-varying amplitude and θ is the time-varying phase or phase angle:

$$x(t) = A(t) * \cos(\theta(t))$$

As illustrated in Figure 6.3a, the BOLD signal phase of a region n over time is shown as $e^{-i\theta(t)}$ where $\sin(\theta(t))$ is the imaginary part of the analytic phase, and $\cos(\theta(t))$ is the real part (black dotted lines). The $\cos(\theta(t))$ of the phase angle still captures the fluctuations of the original BOLD signal (green) but now has a constant amplitude between -1 and 1. The red

arrows represent the Hilbert transformed phases at each TR, which can be projected onto a complex plane that is defined by the real and imaginary axes at $t = 0$.

These BOLD signal phases are then used to generate a whole-brain pattern of BOLD phase coherence at each single time point t by computing a dynamic BOLD PL matrix $dPL(n, p, t)$ which estimates the phase alignment between each pair of brain regions n and p at each time t using the following equation:

$$dPL(n, p, t) = \cos(\theta(n, t) = \ \theta(p, t))$$

At a given TR, if two brain areas, n and p have BOLD signals that are completely temporally aligned (i.e., the phase difference = 0°), they will have a PL value = 1, and can be expressed as follows:

$$dPL(n, p, t) = \cos(0°) = \ 1$$

Conversely, if the BOLD signals of n and p have a phase difference of 180° (in complex plane), the PL value will = -1, and can be written as:

$$dPL(n, p, t) = \cos(0°) = \ -1$$

Therefore, the dynamic PL matrix for each participant will be a 3D tensor of size $NxNxT$, where $N = 100$ is the number of brain regions and $T = 257$ is the total number of time points.

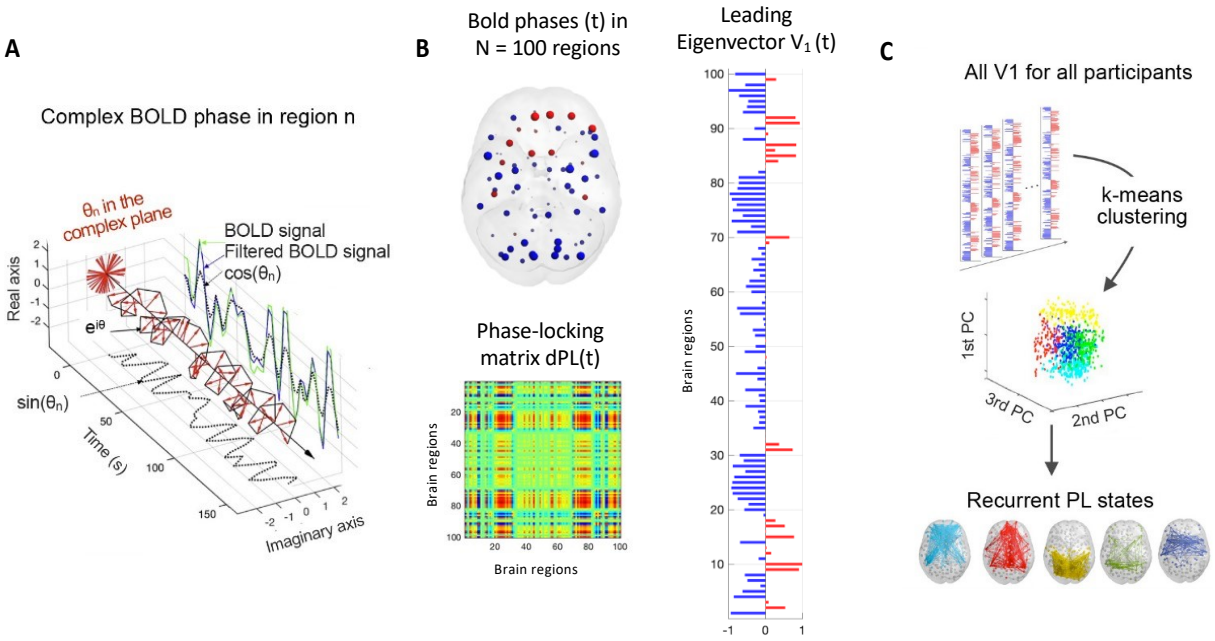


Figure 6.3 — Illustrative description of Leading Eigenvector Dynamics Analysis (LEiDA). LEiDA detects recurrent BOLD phase-locking (PL) patterns. ***A)*** *For a given region (n), the BOLD signal (green) is band-pass filtered between 0.01 and 0.1 Hz (blue) and then Hilbert transformed into an analytic signal, whose phase is represented over time by* $e^{-i\theta}$ *(black line) and at each TR (red arrows).* ***B)*** *At a single time point, BOLD phases in all N =100 regions can be represented in cortical space (top). The dPL(t) matrix captures the phase alignment between each pair of regions (bottom). The leading eigenvector of the dPL(t) matrix,* $V_1(t)$*, is the vector that best captures the main orientation of all phases, where each element in* $V_1(t)$ *relates to the projection of the phase of each region into* $V_1(t)$ *(right). Blue = regions aligned with* $V_1(t)$*; red = regions not aligned with* $V_1(t)$*.* ***C)*** *All the leading eigenvectors are concatenated across participants and inputted into a k-means clustering algorithm, which subsets the data points into a pre-defined number of clusters k. Each cluster centroid represents a recurrent PL state. dPL = dynamic phase locking. This figure was adapted from Alonso Martínez et al. (2020). Copyright: Frontiers in Neural Circuits.*

6.7.2 Calculating the leading eigenvector of the phase-locking matrix

The second step involves calculating the leading eigenvector for each dPL matrix at each time t. Therefore, for each $dPL(t)$, the leading eigenvector $V_1\ (t)$ is a N x 1 vector that represents the main *orientation*, or global mode, of BOLD phases across all brain regions. Each element in $V_1\ (t)$ captures the projection of the BOLD phase of each brain area onto the leading eigenvector. When all elements in $V_1\ (t)$ have the same sign, it means that all BOLD phases are aligned with the orientation of $V_1\ (t)$, which indicates the global mode guiding all BOLD signals. On the other hand, if the leading eigenvector $V_1\ (t)$ possesses elements with different signs (positive or negative), the BOLD signals are not aligned with the leading eigenvector, and this categorises brain regions into two groups (blue or red) depending on their phase

relationship (see Figure 6.3b) (Newman, 2006). The magnitude of each element in $V_1\,(t)$, indicates the strength of the brain areas' group membership of the group in which it has been placed (Newman, 2006). Given that V and -V represent the same eigenvector (i.e., they span the same one-dimensional subspace), a convention is used assuming that most of the elements have negative values (Alonso Martínez et al., 2020; Lord et al., 2019). LEiDA significantly reduces the dimensionality of the data by only considering the eigenvector associated with the leading eigenvalue instead of considering all elements of the N x $N\ dPL$ matrix (Cabral, Vidaurre, et al., 2017).

6.7.3 Detecting recurrent BOLD PL states

The next step involves detecting recurrent BOLD PL states using k-means clustering which partitions the set of leading eigenvectors into a predefined number of clusters k (see Figure 6.3c). Given that the optimal number of functional networks in the brain remains under debate, LEiDA allows the user to run the k-means clustering algorithm with a range of k values (e.g., 2 to 20), which divides the set of leading eigenvectors into $k = 2, 3, 4, \ldots, 20$. Prior research that has used the LEiDA approach has chosen a range of k values that reflect the range of functional networks often reported in resting state literature (Yeo et al., 2011). For each value of k, the clustering algorithm produces k cluster centroids with a shape of N x 1 vectors V_c, which denotes the average vector of each cluster. These centroids represent the recurrent BOLD PL states, which can be represented as a network in cortical space. The value of $V_c(n)$ is used to scale the colour of each brain region and links can then be plotted between brain areas that diverge from the global mode.

The clustering algorithm also appoints a single PL state to each TR across the time series by identifying the closest centroid V_c at each TR. From this, the 1) probability of occurrence: the proportion of timepoints assigned to a PL state during the scan and 2) duration/dwell time: the average number of consecutive time points assigned to a PL state across the scan, can be calculated. From these statistics, between-group differences can be calculated with permutations, bootstrapping, and multiple comparison corrections.

6.8 Chapter conclusion

This chapter outlined the measures and analysis methods used in Chapter 7, which is the final results chapter of this thesis. In Chapter 7, I report and discuss the results from our pilot study which explored dynamic functional brain networks in adolescent depression, and their relation to irritable mood, using the LEiDA approach.

7 Exploring dynamic functional brain networks in adolescent depression using a co-produced novel irritability task

7.1 Chapter introduction

The work outlined in this chapter directly addresses the research gaps identified and discussed in Chapter 5. Specifically, I pilot a novel fMRI task targeting irritability in a non-clinical sample of youth with depressive symptoms. As our goal was to design a task that captures the social context in which irritability occurs, the task was co-produced with young people. As a first step in validating this task, I adopted a data-driven analysis approach to examine whether dynamic properties (e.g., probability of occurrence and dwell time) of functional brain networks differed between the irritability task and a standard resting state scan. I also explored whether the properties of these dynamic functional brain networks related to depressive symptoms and irritable mood in this sample of adolescents.

Supplementary Information for this chapter can be found in Appendix 3, alongside the related material for this study that is outlined in Chapter 6.

7.2 Introduction

Adolescence, a life phase spanning the ages 10-24 years, is a period of increased vulnerability to the onset of mental health difficulties, including depression (Sawyer et al., 2018; Solmi et al., 2021). Adolescent-onset depression is associated with a more recurrent and severe illness course, which compounds the burden of depression across the lifespan and can lead to a host of psychosocial and physical difficulties (Malhi & Mann, 2018; Thapar et al., 2022). Unlike major depressive disorder (MDD) in adults, irritable mood is an additional cardinal symptom of adolescent MDD, alongside low mood, and anhedonia (American Psychiatric Association, 2013b). High levels of irritability in youth predict later depression, suicidality, and poor social functioning (Leibenluft & Stoddard, 2013; Stringaris et al., 2013; Stringaris, Zavos, et al., 2012), suggesting that irritability may be an early indicator of emotion regulation difficulties. Previous research suggests that both MDD and irritability are underpinned by disruptions to the coordination of large-scale functional brain networks involved in emotional processing and regulation (Kaiser et al., 2015; Nielsen et al., 2021). Importantly, adolescence is a period during which mood and behaviour are significantly impacted by one's social environment (Blakemore & Mills, 2014). However, existing fMRI paradigms targeting irritability typically overlook the social context in which irritability occurs. Here, we pilot a novel resting state paradigm targeting irritability using a co-produced youth-researcher design to explore dynamic functional brain networks associated with depressive symptoms and irritable mood in a non-clinical sample of youth.

Resting-state functional magnetic resonance imaging (rs-fMRI) studies have linked depression to aberrant functional connectivity (FC) patterns in several resting state networks (RSNs) involved in emotion regulation, processing, and attention. These RSNs include the default mode network (DMN), which comprises the medial prefrontal cortex (mPFC), lateral frontal cortex, temporal-parietal areas, the striatum, and the hippocampus. The DMN plays a key role in self-directed thought, including introspection and autobiographical memory (Andrews-Hanna et al., 2014). The fronto-parietal network (FPN), including the dorso-lateral PFC (dlPFC) and posterior parietal regions, supports goal-directed behaviour, such as cognitive control and decision making (Zanto & Gazzaley, 2013). Finally, the salience network (SN), comprising

the insula and mid-cingulate regions, supports attending to salient stimuli in one's environment (Corbetta et al., 2008). Specifically, meta-analytic research has shown that depression is related to hypoconnectivity within the FPN, and between the FPN and SN. Depression has also been characterised by hyperconnectivity within the DMN, between the DMN and FPN, and SN (Kaiser et al., 2015). Further, neural-behavioural research on RSNs in both adults and adolescents has shown that these dysfunctional FC patterns are associated with depression-related behaviours such as rumination and cognitive biases towards negative information (Kaiser et al., 2018, 2019; Marchitelli et al., 2022). Although research on RSNs associated with irritability is sparse, the existing studies suggest that the same RSNs are involved (Nielsen et al., 2021). Taken together, the heightened vulnerability to depression during adolescence may reflect maladaptive alterations in these functional networks.

Most research to date has studied depression-related FC from a static perspective, whereby an average FC measure is computed across the entire time-series. Importantly, this approach does not capture the spontaneous waxing and waning of brain network activity over time (Iraji et al., 2021). Over the past decade, there has been a growing emphasis on developing methods that capture the dynamic nature inherent to functional networks, an approach termed dynamic FC (dFC) (Calhoun et al., 2014; Sakoğlu et al., 2010). This work has shown that brain activity comprises time-varying, reoccurring patterns of the coupling and uncoupling of brain regions. Further, these recurrent spatiotemporal configurations (dynamic FC states) and their properties (e.g., probability of occurrence, dwell time, and transition profiles) contain important information that can assist our understanding of the processes underlying cognition and behaviour (Cabral, Vidaurre, et al., 2017; Iraji et al., 2021; Sakoğlu et al., 2010; Zalesky et al., 2014). However much more research is needed in this area. For example, depression has been associated with both increased (Kaiser et al., 2016; Long et al., 2020) and reduced (Marchitelli et al., 2022) dynamic variability in the limbic network, FPN and DMN. Although investigating the characteristics of dFC within a psychiatric context is still a nascent field of research (Iraji et al., 2021), emerging findings suggest that it could be used to successfully predict treatment outcomes to electroconvulsive therapy in MDD patients (Sendi et al., 2021).

Over the past two decades, several methods have been put forward to examine the dynamics of functional brain networks and how they relate to cognition and behaviour. A number of studies have proposed that the properties of dFC, such as the duration of time spent in certain brain networks, could serve as a biological marker for psychiatric disorders such as depression (Alonso Martínez et al., 2020; Figueroa et al., 2019) and schizophrenia (Farinha et al., 2022; Rabany et al., 2019). While the sliding window approach has been the most common method used to evaluate dFC (Allen et al., 2014; Handwerker et al., 2012), other techniques with higher temporal resolution have emerged in recent years, such as co-activation pattern analysis (Karahanoğlu & Van De Ville, 2015; Liu et al., 2013; Tagliazucchi et al., 2012) and phase-coherence pattern analysis (Cabral, Vidaurre, et al., 2017; Glerean et al., 2012; Hellyer et al., 2015). By definition, co-activation pattern analysis methods, such as point process analysis (Tagliazucchi et al., 2012) and innovation-driven co-activation patterns (iCAPs) analysis (Karahanoğlu & Van De Ville, 2015), are only sensitive to simultaneous blood-oxygen-level-dependent (BOLD) signals that exceed a certain threshold, which in turn are used to identify dynamic brain states. On the other hand, phase-coherence techniques are able to capture temporally delayed associations between brain regions, which may better reflect the ultra-slow, oscillating nature of resting state brain networks (Cabral, Kringelbach, et al., 2017; Deco & Kringelbach, 2016; Gutierrez-Barragan et al., 2019; Roberts et al., 2019).

Leading Eigenvector Dynamics Analysis (LEiDA) is a recently developed phase-coherence approach that relies of the detection of instantaneous (i.e., at each single timepoint) recurrent phase-locking (PL) patterns (Cabral, Vidaurre, et al., 2017). LEiDA focuses on the relative phase of the BOLD signal by determining how each BOLD phase projects onto the leading eigenvector of all BOLD phases at each timepoint across the scan. This approach represents a significant step forward in dFC research as it reduces the dimensionality of the data (from a NxN matrix to a 1xN vector), which allows for better convergence of the clustering algorithm. Importantly, LEiDA has been shown to detect PL states that overlap with known RSNs (Lord et al., 2019; Vohryzek et al., 2020). The dynamical properties of these PL states, such as occupancy probability, state dwell time/duration, and transition probabilities, have been associated with cognitive performance (Cabral, Vidaurre, et al., 2017), MDD history (Figueroa et al., 2019), non-clinical depressive symptom severity (Alonso Martínez et al.,

2020), as well as the effects of psilocybin (Lord et al., 2019). Taken together, these findings suggest that LEiDA is an ideal tool with which to examine RSNs related to behaviour and the associated dynamic functional properties of these PL states.

In this pilot study, we apply LEiDA, for the first time, to an adolescent dataset to investigate whether FC states differ between a resting state scan and a novel irritability task. We then examine if the dynamical properties of these PL states are associated with depressive symptom severity and irritability, as rated through behavioural questionnaires. Our study employs a co-produced youth-researcher design whereby young people (aged 16-20 years) were involved in the design and development of our irritability task. By adopting this approach, our aim was to design an irritability task that reflected what experiences young people find irritating in the present day.

7.3 Methods

7.3.1 Participants

As previously outlined in Chapter 6, N = 30 participants aged 16-20 years (M_{age} = 18.86 years; SD = 0.83; 77.4% female) with self-reported depressive symptoms took part in our pilot study. Eligible participants were aged 16-20 years, free from MRI contraindications, fluent English speakers, did not report a past or current clinical diagnosis of autism spectrum disorder, a neurological or genetic disorder, or known intellectual disability, and had a Mood and Feelings Questionnaire (MFQ) total score ≥ 8 (Burleson Daviss et al., 2006; Costello & Angold, 1988). N = 25 participants in the sample were medication naïve, while 5 individuals were either currently (N = 3) or had previously (N = 2) taken psychotropic medication for depression and/or anxiety.

Participants were recruited through schools and universities, third sector organisations (with the assistance of the NRS Mental Health Research Network), social media and via snowball sampling. The study protocol was approved by the Edinburgh Medical School Research Ethics Committee. All participants provided written informed consent prior to taking part in the study. Data collection took place between December 2020 and June 2021.

Data for one participant were unusable due to a data transfer error, which resulted in a final N = 29.

7.3.2 Mood-related measures

Current depressive symptom severity was measured using the depression module of the Patient Health Questionnaire (PHQ-9) (Kroenke et al., 2001). The PHQ-9 is self-administered and has been validated across a range of clinical settings and samples, including adolescents, as a screener for MDD (Arroll et al., 2010; Beard et al., 2016; Levis et al., 2019; Volker et al., 2016). Irritable mood was assessed using the Affective Reactivity Index (ARI) self-report questionnaire (Stringaris, Goodman, et al., 2012). The ARI assesses trait-like irritability and has shown good internal consistency (Cronbach's α values ≥.80) and construct validity in both clinical and community samples (Mulraney et al., 2014; Stringaris, Goodman, et al., 2012).

7.3.3 Irritability fMRI task

To design a task that reflected the experience of irritability as a young person, we adopted a co-produced youth-researcher design whereby young people were involved in multiple stages of the task development. In the first phase, we asked young people (N = 25; aged 16-18 years, independent from the scanned sample) at local educational outreach events to briefly describe irritating scenarios based on their own personal experience and/or feelings. Specifically, they were asked to complete the following sentence: *"I find it irritating when…"*. This resulted in 51 unique irritating scenarios (listed in Table S1 in the Supplementary Information). We then asked another independent group of young people (N = 61; aged 16-18 years), to rate these scenarios on a 5-point scale where 1 = "not at all irritating" and 5 = "very irritating" via an online survey, which was distributed to young people via social media and local youth groups. Importantly, the two groups involved in task development were independent of the study sample. The survey remained live for 5 weeks between August and September 2020. The 18 most-highly rated scenarios were chosen as stimuli for the irritability task (see Table 6.1 in Chapter 6), which was programmed in NBS Presentation® (version 19). Further details on the task development can be found in the Supplementary Information.

7.3.4 Scanning protocol

A 3T Siemens (Magnetom Skyra Fit) MRI scanner with a 32-channel head coil was used to obtain the brain images. The scanning procedure consisted of a T1-weighted sequence that yielded 192 contiguous 1.0mm slices (matrix = 256 x 256; FoV = 256mm; flip angle = 7°). This was followed by a functional imaging protocol using an axial gradient echoplanar imaging pulse sequence (EPI) [TR = 1400ms; TE = 30ms; matrix = 70 x 70; FoV = 210mm; flip angle = 68°, spatial resolution = 3mm isotropic). Sixty contiguous 3mm slices were collected during each TR using 2x GRAPPA acceleration. Two resting state scans were collected, each lasting six minutes. The first was a standard resting state scan whereby the participant was asked to focus on a white cross on a dark screen. This was followed by the irritability task scan whereby participants were asked to read a series of 18 irritating scenarios presented on the screen one at a time and to imagine being in each scenario as vividly as possible. Each scenario was presented for a period of 20 seconds. Visual stimuli for the scans were presented using a screen in the bore of the magnet and run using Presentation® software. After the resting state scans, participants completed two additional fMRI tasks: a self-referential recall memory task and a value choice reward task, not presented here. To aid noise reduction and reduce head mobility, participants were provided with ear plugs and foam padding to support the head. The total scanning time was approximately 50 minutes. Six pulse sequences took place before each scan acquisition, so removing dummy scans during image pre-processing was not required.

7.3.5 Image pre-processing and analysis

The ENIGMA *H*armonised *A*nalysis of *F*unctional MRI *pipe*line (HALFpipe) version 1.2.1 was used for the pre-processing, quality assessment, and single-subject feature extraction of the imaging data (Waller et al., 2022). HALFpipe is a semi-automated pipeline based on fMRIPrep (Esteban et al., 2020). The imaging pre-processing and quality control methods for this study are described in detail in Section 6.6.1 in Chapter 6. In the interest of chapter completeness, an overview is provided below.

To account for low frequency noise, images were bandpass filtered at 128s (using a Gaussian FWHM filter) and grand mean scaling was applied with a mean of 10,000. Motion artefacts

were regressed out using ICA-AROMA. All images were quality controlled (QC) by four independent raters. An average mean framewise displacement (FD) value of <.25mm was used as the threshold for inclusion. Following QC, no participants were excluded from further analyses. Detailed QC statistics can be found in Tables S2 & S3 in the Supplementary Information. Average timeseries were then extracted from each brain region for each participant based on the Automated Anatomical Labelling atlas 120 (AAL120) (Rolls et al., 2015; Tzourio-Mazoyer et al., 2002). These connectivity matrices were used as the input for LEiDA. Across all participants, any brain region with NaN values in the AAL120 connectivity matrix was removed (see Table S4, Supplementary Information, for the brain regions removed). This resulted in a 100-region parcellation of the brain.

Additional bandpass filtering between 0.01 and 0.1 HZ was applied to the BOLD signals for each of the 100 brain areas to remove high frequency components related to respiratory and cardiac signals. This focused on the most-meaningful frequency range of BOLD signal fluctuations (Biswal et al., 1995; Cabral, Vidaurre, et al., 2017).

7.3.6 Dynamic BOLD phase-locking analysis

For each experimental condition (resting state and irritability task), a NxT BOLD matrix was derived where N = 100 is the number of brain regions and T = 257 is the number of volumes. The BOLD signal phase $\theta(n,t)$ for each brain region at each TR was then computed using the Hilbert transform. The Hilbert transform expresses a given signal x in polar coordinates, where A represents the time-varying amplitude and θ is the time-varying phase or phase angle:

$$x(t) = A(t) * \cos(\theta(t))$$

These BOLD signal phases are then used to generate a whole-brain pattern of BOLD phase coherence at each single time point t by computing a dynamic BOLD PL matrix $dPL(n,p,t)$ which estimates the phase alignment between each pair of brain regions n and p at each time t using the following equation:

$$dPL(n,p,t) = \cos(\theta(n,t) = \theta(p,t))$$

At a given TR, if two brain areas, n and p have BOLD signals that are completely temporally aligned (i.e., the phase difference = 0°), they will have a PL value = 1. On the other hand, if the phase difference between brain regions n and p is 180°, the PL value will = -1. This produces a dPL 3D matrix for each participant of size $NxNxT$, where $N = 100$ is the number of brain regions and $T = 257$ is the total number of volumes.

7.3.7 Computing the leading eigenvector of the phase-locking matrix

A leading eigenvector for each dPL matrix at each time t is then calculated, where the leading eigenvector $V_1\,(t)$ is a N x 1 vector that represents the main orientation of BOLD phases across all brain regions. When all elements in $V_1\,(t)$ have the same sign, it means that all BOLD phases are aligned with the orientation of $V_1\,(t)$, which indicates the global mode guiding all BOLD signals. On the other hand, if the leading eigenvector $V_1\,(t)$ possesses elements with different signs (positive or negative), the BOLD signals are not aligned with the leading eigenvector, and this categorises brain regions into two groups (blue or red) depending on their phase relationship (Newman, 2006). The magnitude of each element in $V_1\,(t)$, indicates the strength of the brain areas' group membership of the group in which it has been placed (Newman, 2006). Given that V and -V represent the same eigenvector (i.e., they span the same one-dimensional subspace), a convention is used assuming that most of the elements have negative values (Alonso Martínez et al., 2020; Lord et al., 2019). LEiDA significantly reduces the dimensionality of the data by only considering the eigenvector associated with the leading eigenvalue instead of considering all elements of the N x N dPL matrix (Cabral, Vidaurre, et al., 2017; Lord et al., 2019; Vohryzek et al., 2020).

7.3.8 Detecting recurrent BOLD PL states

The main aim of this pilot study was to investigate whether the BOLD PL states differed between a standard resting state scan and our novel irritability task. Thus, to detect recurrent BOLD PL states we used LEiDA to apply k-means clustering to all leading eigenvectors $V_1\,(t)$ across all participants. This resulted in 257 x 29 x 2 = 14,906 leading eigenvectors (number of volumes x sample size x scan conditions). The clustering produces k clusters where each k represents a recurrent PL state, where higher values of k demonstrate more fine-grained

network configurations. Given that the optimal number of functional networks in the brain remains under debate, LEiDA allows the user to run the k-means clustering algorithm with a range of k values (e.g., $k = 2, 3, 4, \ldots, 20$). Prior research that has used the LEiDA approach has chosen a range of k values that reflect the series of functional networks often reported in resting state literature (Alonso Martínez et al., 2020; Figueroa et al., 2019; Lord et al., 2019). Here, we did not aim to determine the optimal number of PL states and instead, we were interested to see if there was a PL state that was consistently significantly different across the resting state and irritability task conditions. Thus, first we chose a wide range of k = 2 to 20 and explored how the PL states differed across scan conditions. We then chose the k number that demonstrated the highest proportion of significant differences between conditions. It is important to note that the k clustering assigns a single PL state to each timepoint, which is independent of the other values of k and results in independent clustering models (Figueroa et al., 2019).

7.3.9 Between condition differences

To investigate how the range of k PL states differed between the resting state and irritability conditions, for each participant we computed: 1) the fractional occupancy/probability: the proportion of timepoints assigned to a PL state during the scan, and 2) the dwell time/ duration: the average number of consecutive time points assigned to a PL state across the scan. We then compared these values across the two scan conditions using non-parametric permutation-based paired sample t-tests (10,000 permutations with 500 bootstrap samples per permutation). There are k number of hypotheses tested for the k PL states produced by k-means clustering. Thus, to correct for multiple comparisons, we set a significance threshold of $0.05/\ \Sigma(\mathrm{k})$.

7.3.10 Association with behavioural measures

To examine whether the properties of the PL states derived by LEiDA were associated with depressive symptoms (PHQ-9 total score) and irritable mood (ARI total score), we used the Spearman rank correlation method to investigate whether PL state probability or duration, for the rest and irritability condition separately, correlated with PHQ-9 and ARI total scores.

We also examined whether these PL states properties were correlated with age and motion. We corrected for multiple comparison using the Bonferroni correction method.

7.3.11 Code availability statement

The LEiDA analysis was conducted in MATLAB version R-2022a. We used the publicly available LEiDA code provided by Cabral and colleagues on the LEIDA MATLAB Toolbox GitHub: https://github.com/PSYCHOMARK/leida-matlab#leida-analysis-k. All code for the current project is available on the GitHub repository for this project: https://github.com/niamhmacsweeney/LEiDA_irritability_study

7.4 Results

7.4.1 Sample characteristics

N = 29 youth were included in the final analyses. Our sample revealed a wide range of depressive symptoms as indexed by the PHQ-9 total score (range = 3 to 24). Descriptive statistics of the sample are reported in Table 7.1. Although our sample contained more females than males, t-tests did not demonstrate any significant mean differences between groups on age ($t(9) = 0.03$, $p = 0.073$); PHQ-9 total score ($t(23) = -2.00$, $p = 0.057$); ARI total score ($t(16) = -0.61$, $p = 0.550$); and on motion measures ($t(11) = 1.08$, $p = 0.303$).

Characteristic	**Male**, N = 7[1]	**Female**, N = 22[1]
Age	18.90 (0.92) (17.73 - 20.04)	18.88 (0.87) (17.16 - 20.02)
Depressive symptoms (PHQ-9 total score)	8.86 (3.02) (5 - 13)	12.55 (6.77) (3 - 24)
Irritability (ARI total score)	2.71 (1.80) (0 - 5)	3.27 (2.86) (0 - 10)
Medication naïve	6 / 7 (86%)	19 / 22 (86%)
Average motion (mean framewise displacement)	0.13 (0.03) (0.10 - 0.19)	0.12 (0.03) (0.07 - 0.22)

[1] Mean (SD) (Range); n / N (%)

Table 7.1 —Irritability pilot study: Sample descriptive statistics.Average motion (mean framewise displacement) is the average motion across the irritability and rest conditions.

7.4.2 Detection of the most different brain states between rest and irritability conditions

As described in the Methods section, our aim was not to determine the optimal number of PL states. Instead, our goal was to validate the irritability task by examining whether the probability and duration of the PL states differed between the irritability task and the standard resting state scan. To do this, we chose the clustering configuration that demonstrated the highest proportion of PL states whose probability and duration were significantly different across the resting state and irritability conditions. In Figure 7.1, we illustrate, for each clustering configuration of the PL state samples into *k* PL state categories, the *p*-values obtained from the between-condition comparisons (using paired samples t-tests) in terms of the probability and duration of each PL state. We corrected for multiple comparisons by dividing the chosen *p*-value threshold of 0.05 by the sum of all tests performed ($0.05/\Sigma(\mathrm{k})$; blue line in Figure 7.1). Although this is a conservative approach, we found that all clustering configurations of $k \geq 8$, demonstrated at least one PL state whose probability of occurrence differed significantly between the irritability and rest conditions, after correction for multiple comparisons at this threshold. Of note in Figure 7.1, differences that withstood Bonferroni correction ($0.05/\mathrm{k}$) are indicated in green, while nominally significant associations are shown in red.

Across the range of k values, and for both probability of occurrence and dwell time, the clustering configuration that demonstrated the highest and most consistent *proportion* of significant differences between conditions was k = 11, as indicated by the grey box in Figure 7.1. Between-condition comparisons for all partition models (k = 2 to 20) are also shown in Figure 7.2 (probability of occurrence) and Figure 7.3 (dwell time). Like the blue line in Figure 7.1, the blue coloured boxplots represent the comparisons that were significant after correction for multiple comparisons at the $0.05/\Sigma(\mathrm{k})$ threshold, while the green coloured boxplots indicate comparisons that withstood Bonferroni correction. As illustrated in Figure 7.2 and Figure 7.3, the second (PL state 2) and tenth (PL state 10) columns demonstrated the most significant differences between conditions (i.e., they have the most blue/green coloured boxplots), the proportion of which was greatest for the clustering configuration k = 11.

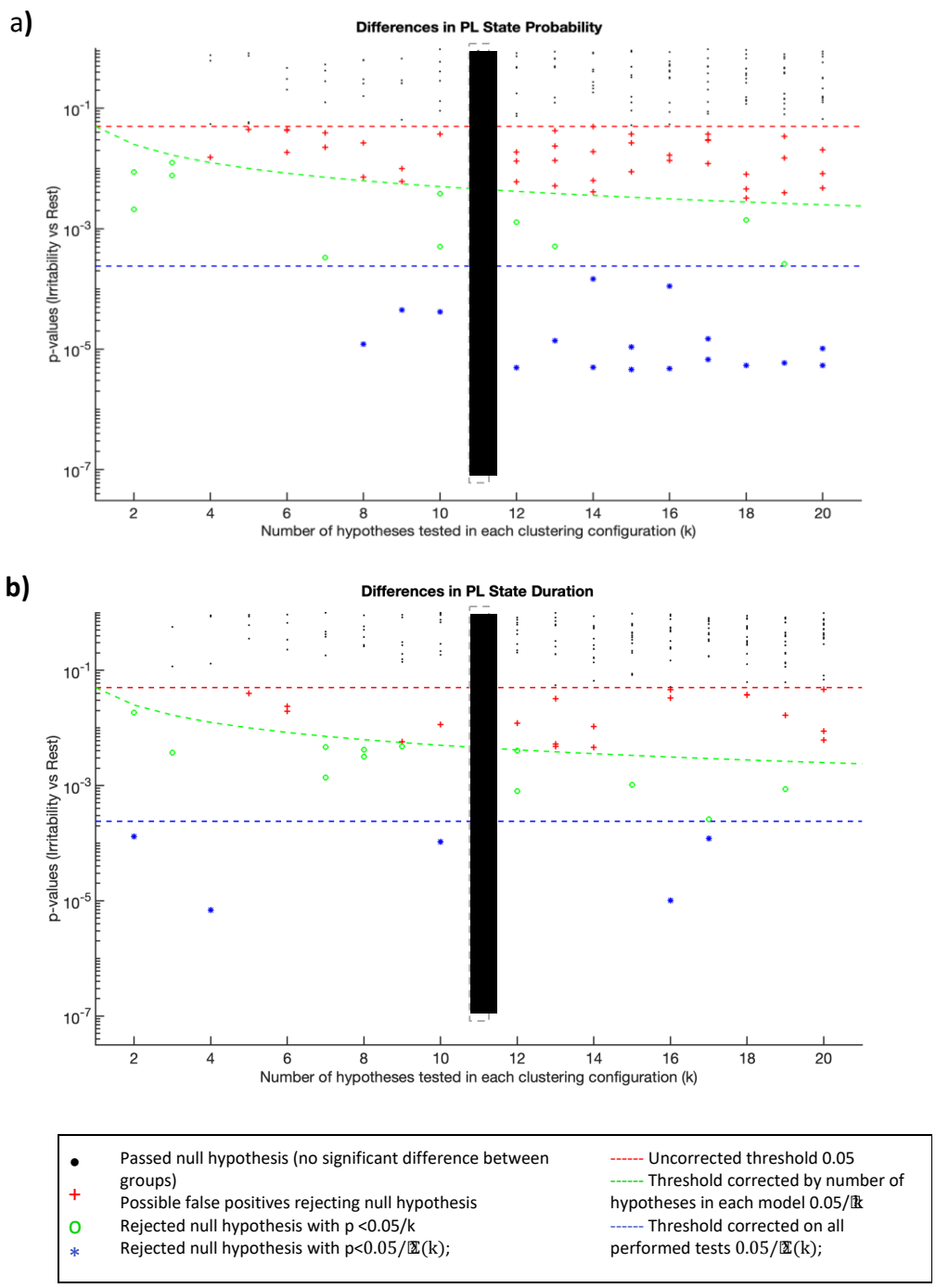
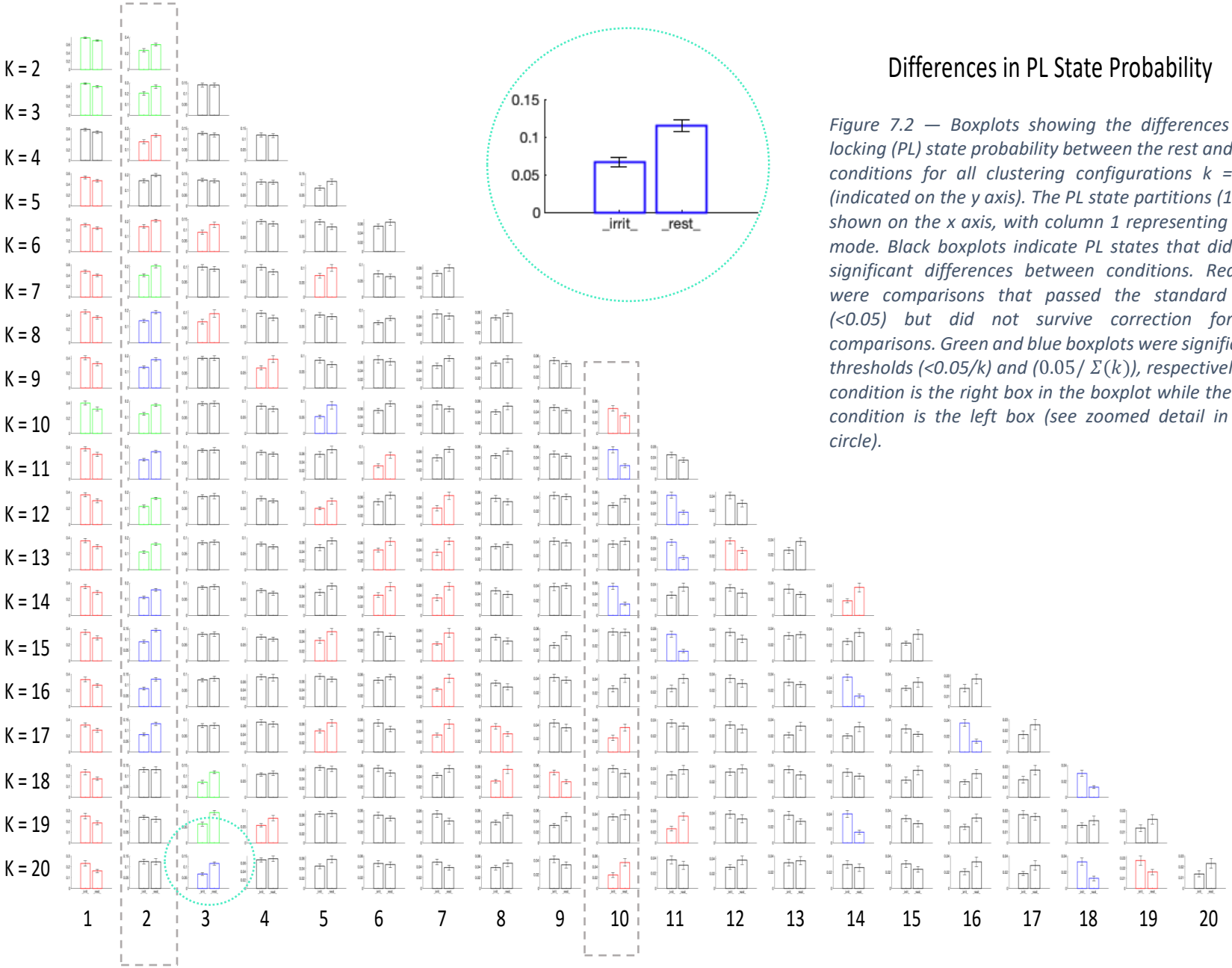
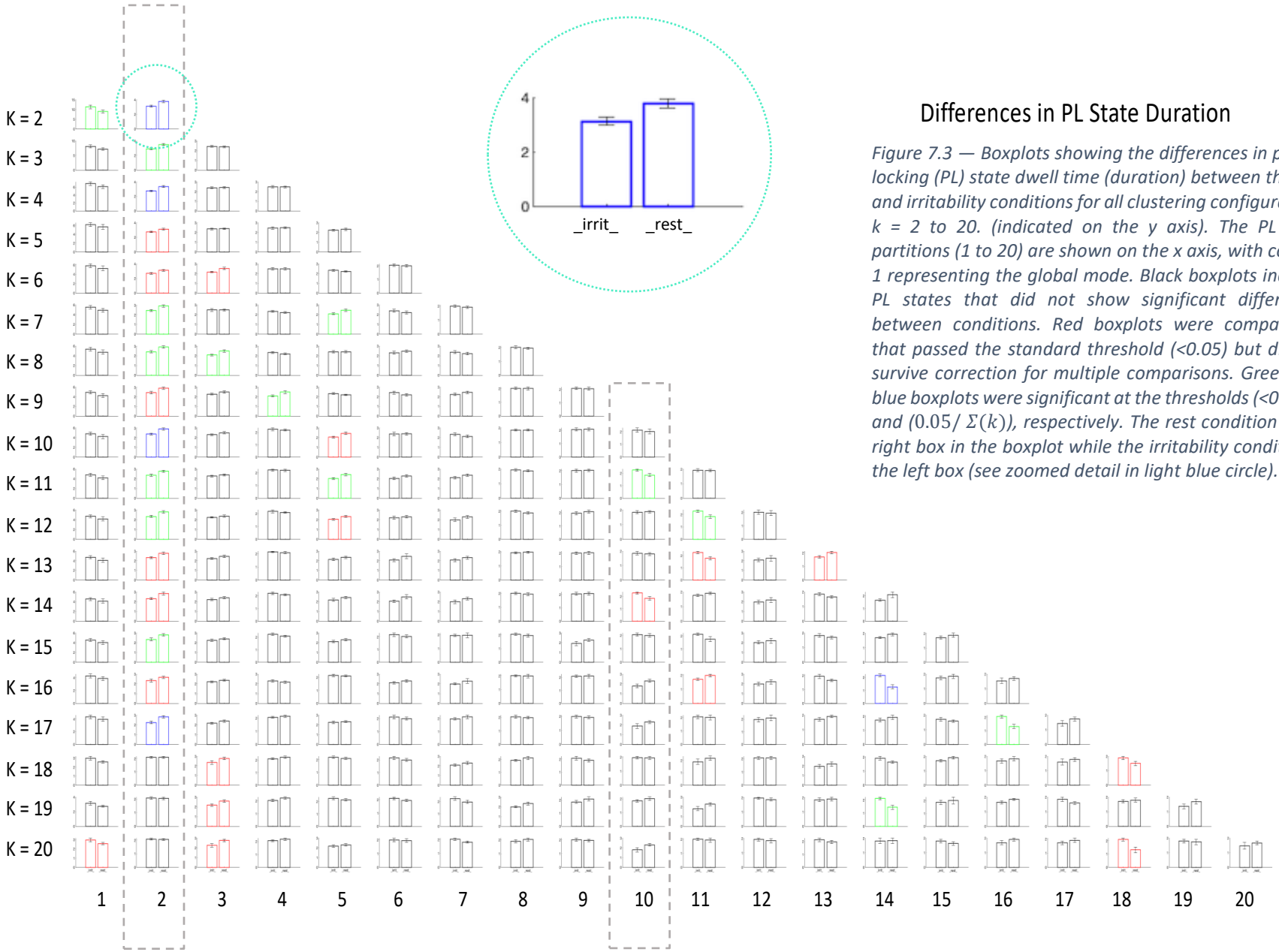


Figure 7.1 — Significance of between-condition differences in phase-locking (PL) state probability and duration as a function of k. For each clustering configuration of the sample into k = 2 to 20 states, we plot p-values associated with the between-condition comparison between the resting state and irritability task in both PL state probability (a) and duration (b). We find that, even though most PL states do not show significant differences between groups (black dots falling above the 0.05 threshold, red dashed line), for all k > 2, there is at least one PL state that falls below (or very close to) the corrected threshold by the number of clusters (<0.05/k, green dashed line) and/or the corrected threshold for all performed tests ($0.05/\Sigma(k)$, blue dashed line). The partition model that demonstrated the highest and most consistent proportion of significant differences between conditions across the range of k values, k = 11, is highlighted by the grey boxes.



Differences in PL State Probability

Figure 7.2 — Boxplots showing the differences in phase-locking (PL) state probability between the rest and irritability conditions for all clustering configurations k = 2 to 20. (indicated on the y axis). The PL state partitions (1 to 20) are shown on the x axis, with column 1 representing the global mode. Black boxplots indicate PL states that did not show significant differences between conditions. Red boxplots were comparisons that passed the standard threshold (<0.05) but did not survive correction for multiple comparisons. Green and blue boxplots were significant at the thresholds (<0.05/k) and ($0.05/\Sigma(k)$), respectively. The rest condition is the right box in the boxplot while the irritability condition is the left box (see zoomed detail in light blue circle).



Differences in PL State Duration

Figure 7.3 — Boxplots showing the differences in phase-locking (PL) state dwell time (duration) between the rest and irritability conditions for all clustering configurations k = 2 to 20. (indicated on the y axis). The PL state partitions (1 to 20) are shown on the x axis, with column 1 representing the global mode. Black boxplots indicate PL states that did not show significant differences between conditions. Red boxplots were comparisons that passed the standard threshold (<0.05) but did not survive correction for multiple comparisons. Green and blue boxplots were significant at the thresholds (<0.05/k) and $(0.05/\Sigma(k))$*, respectively. The rest condition is the right box in the boxplot while the irritability condition is the left box (see zoomed detail in light blue circle).*

7.4.3 Relevant PL states

In Figure 7.4, we demonstrate the complete repertoire of PL states that are produced by LEiDA when k = 11 is chosen. This shows distinct network configurations that occur, dissolve, and reoccur in all participants across the scan. Notably, the global state (PL state 1) was occupied most often and demonstrated the most variance across participants. In this PL state, all BOLD signals are aligned which represents a slowly changing global mode of BOLD activity. These networks are represented in matrix format in Figure 7.5 and shown alongside the links in the cortex that LEiDA used to derive the brain functional networks. Moreover, the networks returned by LEiDA overlap with previously described RSNs by Yeo et al. (2011), as illustrated in Figure 7.6.

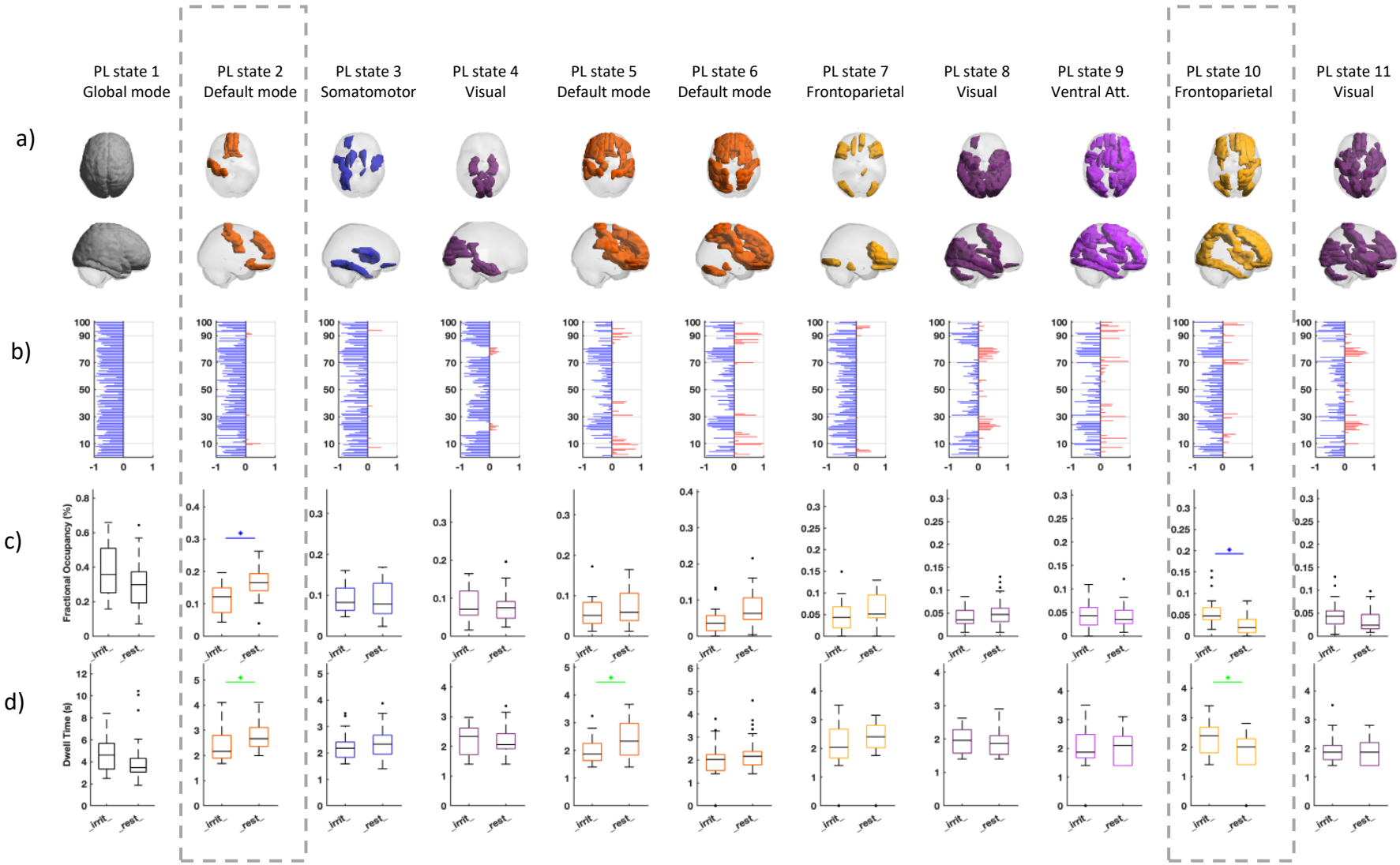


Figure 7.4 — Phase-locking (PL) brain states for a clustering configuration of k = 11. Panel a) shows each PL state represented in Yeo et al. cortical space and panel b) shows the PL states represented as BOLD phase projections into V_1 (Global mode or FC state 1) for the 100 brain regions included in the parcellation. Panels c) and d) show between-condition ("_irrit_" = irritability; "_rest_" = rest) comparisions of fractional occupancy (probability) and dwell time, which are coloured according to the Yeo et al. cortical space atlas. The PL states that demonstrated significant between-condition differences are highlighted in the grey boxes.

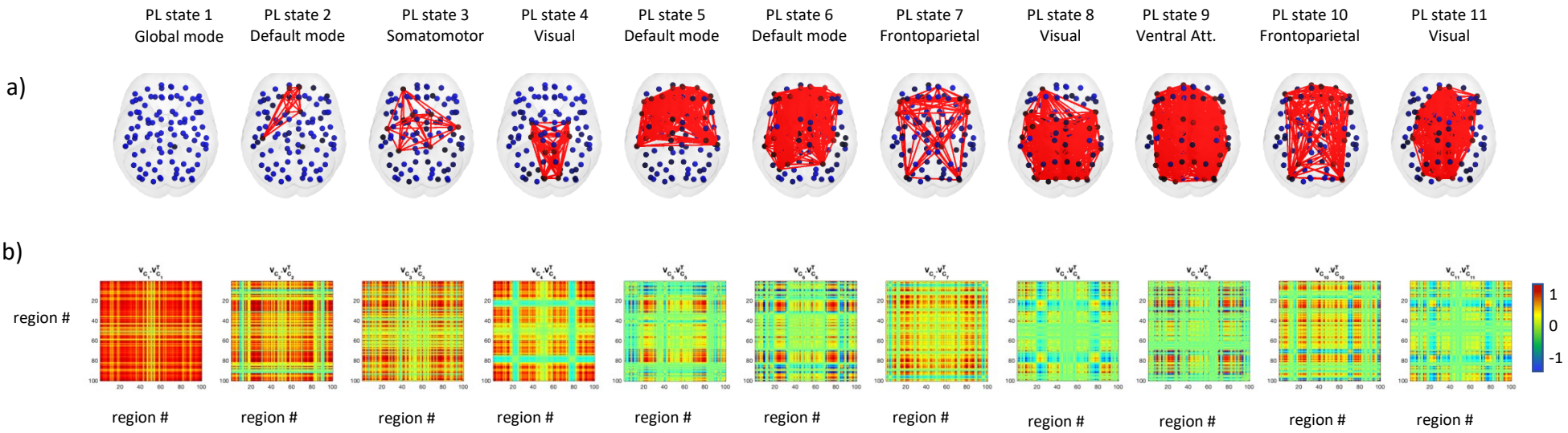


Figure 7.5 — Repertoire of recurrent phase-locking (PL) states obtained with a clutering configuration of k = 11. a) PL states represented in Yeo et al. (2011) cortical space. Functionally connected regions (represented as spheres) are coloured in blue and red links are plotted between AAL areas with >400 MNI voxels contributing to each resting state network. b) PL states are also represented as the outer product of V_c (where c is a PL state from 1 to 11 for k =11), which is a 100 x 100 matrix representing the number of brain regions, where positive (red) values indicate the product of Vc elements with the same sign, which can be positive or negative.

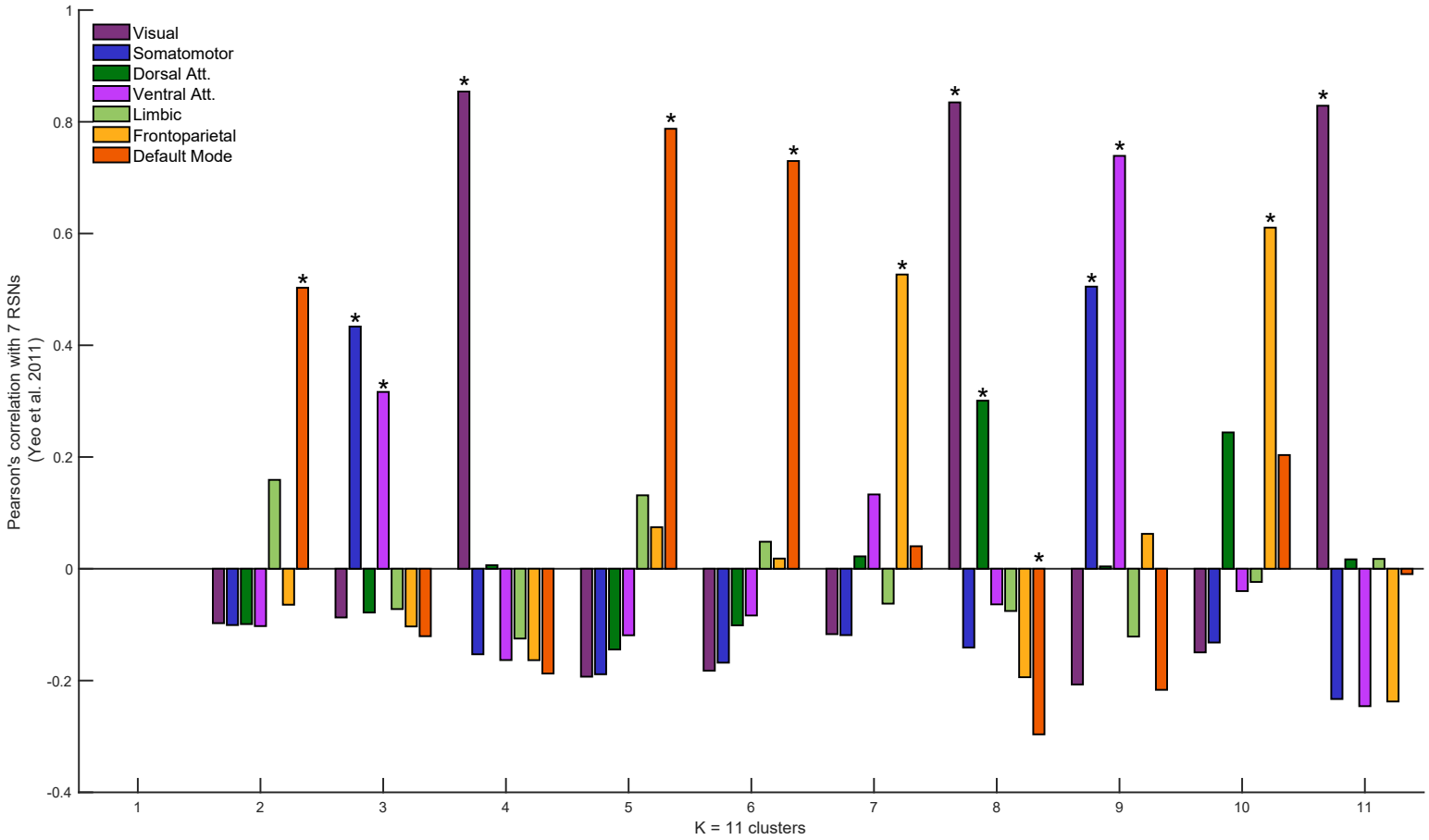


Figure 7.6 — Overlap of the 10 non-global phase-locking (PL) states with seven canonical resting-state networks (RSNs) as per Yeo et al. (2011). We compared the 10 non-global PL states obtained with LEiDA (for k = 11) with the RSNs by calculating the Pearson's correlations between 1 x 100 centroid vectors V_c (shown in Figure 7.5) and the 7 Yeo et al. (2011) RSNs. Asterisks denote significant correlations with p-value <0.001.

Our results demonstrated a PL state (PL state 2) that appeared consistently more and lasted longer during the resting state scan condition compared to irritability condition. This PL state comprised a brain network including medial frontal (right and left frontal superior gyri) and orbitofrontal regions (right and left olfactory cortex, right rectus, left orbitofrontal cortex) as well as postcentral areas (left postcentral gyrus) (Figure 7.7a). When this data-driven PL state was mapped onto the Yeo et al. (2011) RSN parcellation, it was derived as part of the default mode network (DMN) indicated in the colour orange in Figure 7.7b. As shown in Figure 7.7c and Figure 7.7d, respectively, PL state 2 occurred more frequently (16% compared to 12.4%, $p < 0.001$, ($0.05/\Sigma(k)$ correction threshold) and lasted longer when occupied (2.6 seconds compared to 2.2 seconds, $p < 0.001$, Bonferroni correction threshold).

The results returned by LEiDA revealed a second PL state (PL state 10) that, in contrast to PL state 2, appeared consistently more and was occupied for a longer duration during the irritability condition compared to the resting state scan condition. This PL state comprised a network made up mostly of medial and inferior frontal areas (right and left superior frontal gyri, left and right orbital frontal gyri), parietal regions (left and right parietal gyri, the left insula, the right postcentral gyrus), and occipital regions (left and right inferior occipital gyri, the right fusiform gyrus) (Figure 7.8a). Mapping this state onto the Yeo et al., RSN parcellation derived it as the fronto-parietal network (FPN), indicated in the mustard yellow colour in Figure 7.8b. As illustrated in Figure 7.8c and Figure 7.8d, respectively, PL state 10 occurred more frequently (4.9% compared to 2.2%, $p < 0.001$, ($0.05/\Sigma(k)$ correction threshold) and for a longer duration when occupied (2.4 seconds compared to 2 seconds, $p < 0.001$, Bonferroni correction threshold) during the irritability task compared to the resting state condition.

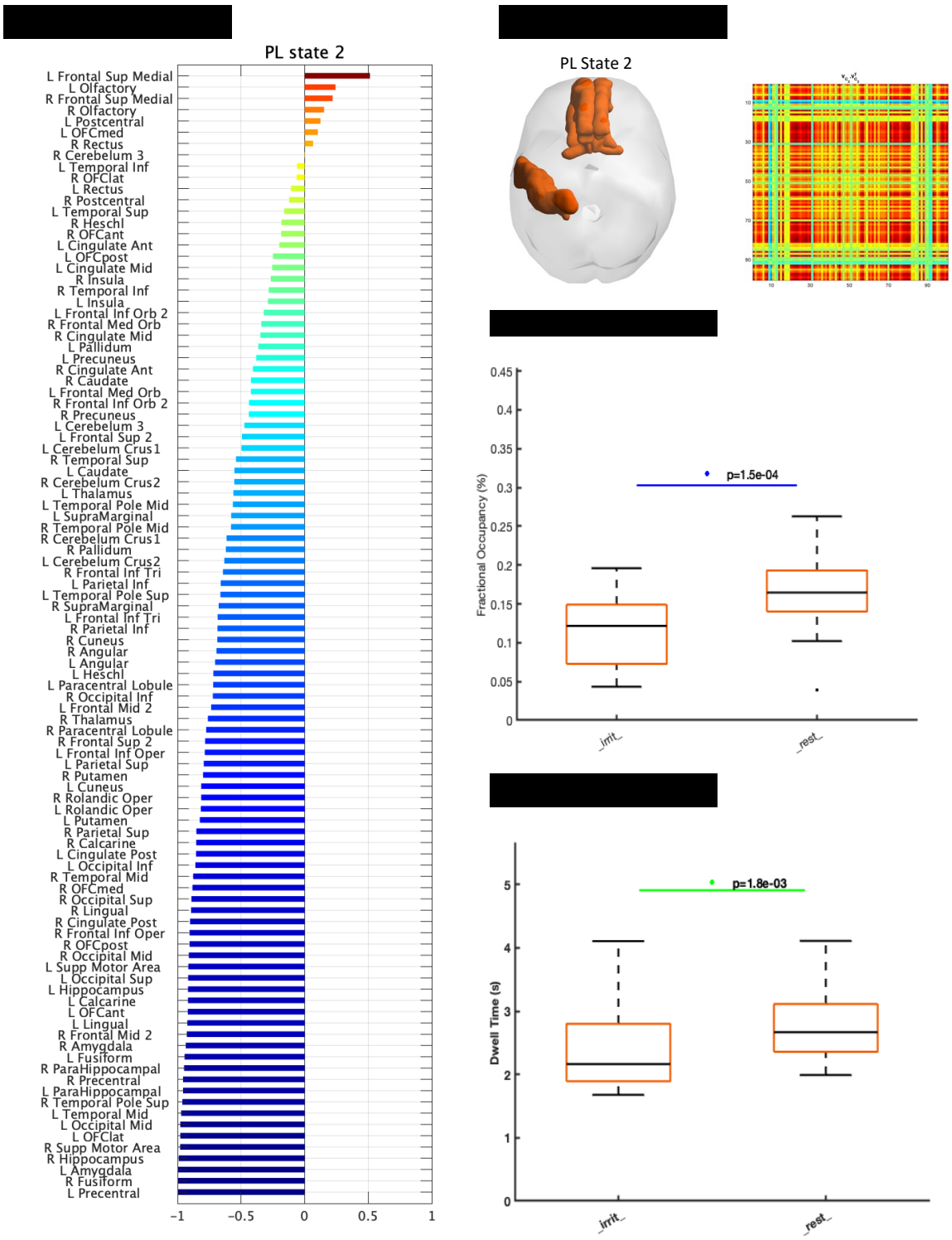


Figure 7.7 — Phase-locking (PL) state 2 for K = 11, which is characterised by regions of the default mode network (DMN). Panel a) shows the projections of the phase of each brain region onto the leading eigenvector (V_1) arranged in order of BOLD signal phase alignment. Brain areas whose phase diverges from V_1 comprise PL state 2 and are the first brain areas listed. Panel b) represents PL state 2 in cortical space and matrix format, while between-condition comparisons for fractional occupancy (probability) and dwell time (duration) are shown in panels c) and d), respectively. Irritability condition = "_irrit_" and rest condition = "_rest_".

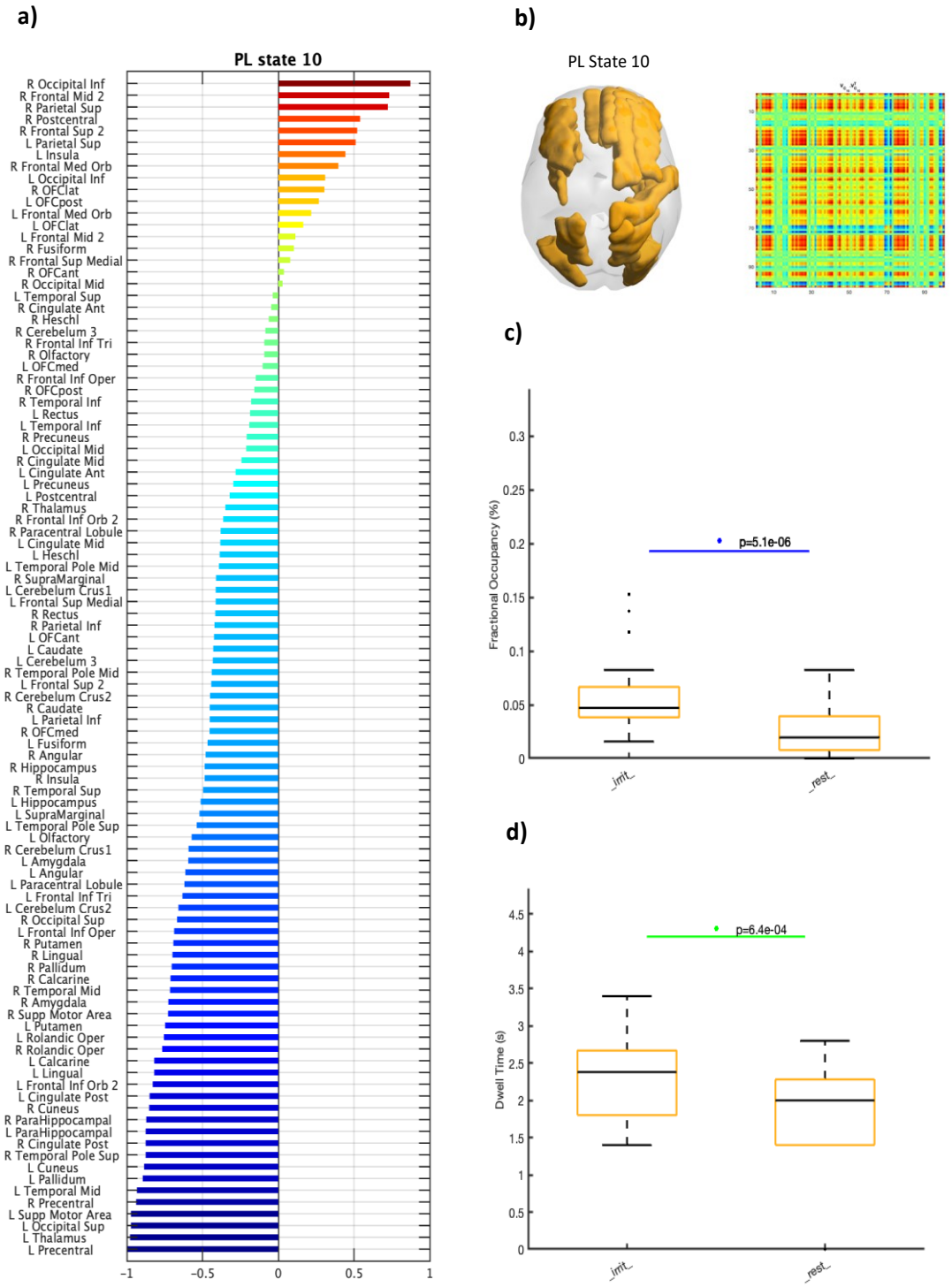


Figure 7.8 — Phase-locking (PL) state 10 for K = 11, which is characterised by regions of the fronto-parietal network (FPN). Panel a) shows the projections of the phase of each brain region onto the leading eigenvector (V_1) arranged in order of BOLD signal phase alignment. Brain areas whose phase diverges from V_1 comprise PL state 10 and are the first brain areas listed. Panel b) represents PL state 10 in cortical space and matrix format, while between-condition comparisons for fractional occupancy (probability) and dwell time (duration) are shown in panels c) and d), respectively. Irritability condition = "_irrit_" and rest condition = "_rest_".

7.4.4 Clinical relevance of PL states

To investigate whether the properties of the PL states derived by LEiDA were associated with depressive symptoms (PHQ-9 total score) and irritable mood (ARI total score), we tested whether PL state probability or duration correlated with PHQ-9 and ARI total scores, using the Spearman rank correlation method. This analysis was undertaken separately for the irritability condition and resting state condition.

7.4.4.1 Irritability condition — correlation between functional brain state properties and clinical measures

During the irritability task, for k = 11, we found that the duration of PL state 11 was negatively correlated with depressive symptoms (r(27) = -.49, p = 0.007). As illustrated in Figure 7.9a, this PL state is characterised by a network comprising frontal, occipital and limbic regions such as the frontal superior gyrus, calcarine, amygdala, hippocampus, cingulate, and insula. Notably, this correlation between PL state 11 and depressive symptoms did not remain significant after correction for multiple comparisons, using Bonferroni correction. For this clustering configuration (k = 11), no other significant correlations (at $p \leq 0.01$) were found between the brain state properties and depressive symptoms or irritable mood.

Moving beyond the partition model used in our earlier analyses (i.e., k =11), we explored whether the probability or duration of the PL states correlated with depressive symptoms or irritable mood across the remaining clustering configurations, i.e., k = 2 to 20 (excluding k = 11). Similar to our findings for k = 11, we found that for k = 20, the duration of a fronto-occipital-limbic network (PL 7) was negatively correlated with depressive symptoms (r(27) = -.64, p = 0.0001), with statistical significance surviving correction for multiple comparisons (see Figure 7.9b). Further, we found that irritable mood (ARI total score) was also negatively correlated with the probability of occurrence of PL states that comprised fronto-occipital-limbic networks, namely, PL 14 (r(27) = -.49, p = 0.007) and PL 15 (r(27) = -.48, p = 0.007) for k =17 (Figure 7.9c) and k =15 (Figure 7.9d), respectively. On the other hand, the duration of PL 15 for k = 18 (Figure 7.9e), comprising parts of the DMN, was positively associated with irritable mood (r(27) = .49, p = 0.006). These correlations did not remain significant after correction for multiple comparisons.

7.4.4.2 Rest condition — correlation between functional brain state properties and clinical measures

For the rest condition, the only significant correlation between the PL state properties and clinical measures was found between the probability of PL 18 for *k* = 20 (Figure 7.9f), which was negatively associated with depressive symptoms (r(27) = -.50, p = 0.005). This PL state comprised a network of fronto-parietal regions. Notably, this PL state was also found to be significantly different between the rest and irritability conditions. For example, during the irritability task, this brain state occurred more frequently compared to the rest condition (2.6% vs 0.5%). The statistical significance of this correlation did not withstand correction for multiple comparisons.

7.4.4.3 Irritability and rest conditions — correlations with age and motion

We also explored whether age and motion (indexed via an average mean framewise displacement measure across both conditions) were correlated with the probability or duration of PL states across all values of *k* in both conditions separately. In the irritability condition, age was positively associated with the duration of PL state 8 for k =17 (r(27) = 0.48, p = 0.008) and PL state 6 for k = 19 (r(27) = .53, p = 0.003), which constitute parietal and occipital-temporal networks, respectively. These correlations did not remain significant after correction for multiple comparisons was applied. Across both conditions, no other significant correlations (at p ≤ 0.01) were found between the PL state properties and age and motion.

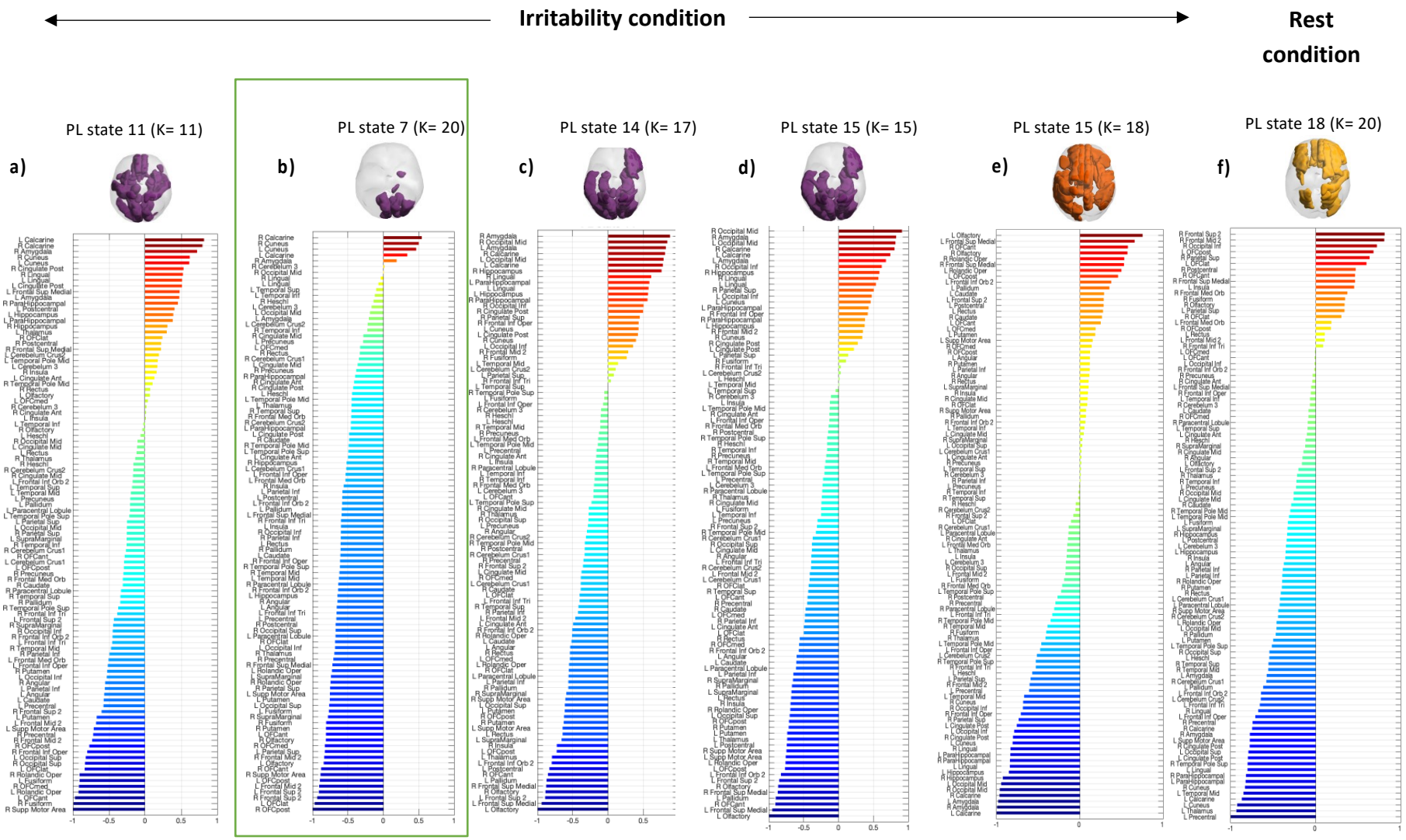


*Figure 7.9 —Dynamic brain states associated with depressive symptoms and irritable mood. Phase locking (PL) brain states that that were significantly associated with depressive symptoms (a-b) or irritable mood (c-e) in the **irritability task condition**, and irritable mood in the **rest condition** (f). The negative correlation between PL state 7 and depressive symptoms remained significant after correction for multiple comparisons, highlighted in the green box.*

7.5 Discussion

This study investigated differences in dynamic FC patterns during a standard resting-state scan compared to a novel irritability task in a sample of youth with depressive symptoms. Using the LEiDA approach, the expression of two PL states was significantly different between the rest and irritability conditions in terms of probability of occurrence and duration. Specifically, during the rest condition, there was increased occupancy and duration of a PL state defined by dominance of the DMN, and a decreased occupancy and duration of a PL state characterised by the FPN. The opposite pattern of results was found during the irritability condition — namely an increased occupancy and duration of a PL state overlapping with the FPN, and a decreased occupancy and duration of a PL state characterised by the DMN. These two states were found to be significantly different across a range of clustering configurations. Within each condition, we also examined whether the occupancy and duration of these dynamic functional brain states were correlated with depressive symptoms or irritable mood. Within the irritability task, the occupancy and duration of PL states characterised by a fronto-limbic-occipital network were negatively correlated with depressive symptoms and irritable mood across a range of partition models. This differs from the rest condition whereby a negative association was found between the occupancy of a PL state characterised by the FPN and youth depressive symptoms.

Taken together, our results suggest that our novel irritability task induces greater occupancy and duration of a dynamic brain state characterised by the FPN. The distinct PL patterns observed between the brain at rest and during the irritability task provide initial evidence for the validation of this task that was designed with young people with the aim of creating a paradigm that reflected the social nature of irritability in adolescence. Further, we found preliminary evidence to suggest that during the irritability task, individuals with higher depressive symptoms and irritability spend less time in brain states characterised by a fronto-limbic-occipital network. Given that these brain regions underpin psychological processes such as emotion regulation, our results suggest that youth with greater levels of depressive symptoms and/or irritability may have a reduced ability to regulate their emotions when presented with an irritating situation. Importantly, the observed associations are correlations

and therefore should not be interpreted as causal. Moreover, many of these correlations did not withstand correction for multiple comparisons. This highlights the need for further investigation, using a larger sample, into the behaviour underpinning the increased occupancy of these functional brain networks during our novel irritability task.

7.5.1 PL states differ during the irritability task compared to the resting-state condition

As this was a pilot study, our first aim was to validate our irritability task by investigating whether the time-varying properties of BOLD PL patterns during the irritability condition differed significantly from those observed during a standard resting-state scan. Our results show that parts of the DMN were more likely to be occupied, and stayed in for a longer time, during the resting state condition compared to the irritability task. These findings are consistent with the resting state literature which has consistently shown the dominance of the DMN when the brain is at rest and engaged in self-referential mental processes (Gusnard et al., 2001; Raichle et al., 2001). Although the DMN is comprised of a series of brain regions including the posterior cingulate, precuneus, medial prefrontal (mPFC) and inferior parietal cortices (Broyd et al., 2009), it is important to note that the PL state derived by LEiDA was characterised by only a subset of these regions. These brain regions were the medial superior frontal gyri, the left medial orbital gyrus, the olfactory cortex, the left postcentral gyrus, and the right gyrus rectus, which have been associated with higher order cognitive processes to perform functions such as planning, reasoning, and social interactions (Jobson et al., 2021; Xu et al., 2019).

An advantage of the data-driven nature of LEiDA is that it allows a finer grained parcellation of the dynamic nature of known RSNs and how they vary across contexts. Here, we show that (mainly) medial frontal regions of the DMN were occupied less frequently during the irritability task compared to the resting-state condition. Interestingly, there were other parcellations of the DMN (e.g., PL 5 and 6) whose frequency of occurrence did not differ significantly between conditions. While it is difficult to draw conclusions about the functional relevance of this finding in a pilot study such as ours, our results show that the dynamic

properties of certain sub-networks within the brain's functional architecture can vary substantially depending on the context.

Further proof of concept for our novel irritability task comes from our finding that a PL state characterised by the FPN was occupied more frequently, and for a longer duration, during the irritability condition compared to the resting-state condition. The FPN network is regarded as a distinct control network responsible for the rapid and flexible coordination of goal-oriented behaviour (Marek & Dosenbach, 2018). Moreover, research suggests that the FPN is a functional hub in charge of the integration of brain-wide behaviour to meet the demands of the environment or task at hand (Cole et al., 2013). During our novel irritability task, youth read a series of irritating scenarios (devised by an independent sample of youth) and were asked to imagine being in that situation as vividly as possible. The increased occupancy and duration of a brain state implicated in cognitive control during the irritability condition compared to the rest condition suggests that the task is impacting behaviour in potentially a few ways. On one hand, the higher levels of occupancy of the FPN during the irritability task could represent participants processing the new information presented in the task (i.e., the irritating scenario) and then co-ordinating their behaviour (i.e., imagining being in that scenario). However, if this state just arose by attending to the new information of the irritating scenario, we would have expected to see differences in the dorsal attention network between conditions.

Another alternative explanation of this pattern of dynamic functional brain networks is that the brain state induced by our task is associated with irritability. As discussed further in Section 7.5.2, within the irritability task, we observed more numerous associations with depressive symptom severity and irritable mood, compared to the resting state condition. Given that our task asks individuals to imagine being in the irritating scenario presented as vividly as possible, the observed pattern of results could reflect varying degrees of adaptive emotion regulation. For example, the negative association observed between the occupancy probability/duration of a limbic-frontal-occipital brain network and depressive symptom severity/irritable mood suggests that individuals with less severe depressive symptoms and/or irritable mood spend more time in brain networks involved in emotion regulation,

while the opposite pattern of results is seen for young people with more severe depressive symptoms and greater irritable mood.

Indeed, several studies — albeit from a static FC approach — have used frustrative non-reward paradigms to induce a frustrated/irritable mood in clinical- and community-based youth samples and have found changes in the neural activation of limbic regions, such as the anterior cingulate cortex, amygdala, and striatum (Tseng et al., 2019; Pawliczek et al., 2013; Deveney et al., 2013). While we do not see differences in the dynamic properties of the limbic network (as defined by the Yeo et al. (2011) RSN parcellation) between conditions (i.e., irritability vs rest), we do observe increased occupancy of a brain state (PL state 10) comprising regions involved in socio-emotional processing (e.g., the insula) during the irritability condition (Uddin et al., 2017). Given that our task was specifically designed to capture the social nature of irritability in adolescence, this may explain why there is not a greater degree of overlap with the aforementioned findings from frustrative non-reward studies, which often do not capture the social context in which irritability occurs (Lee et al., 2022). Further, another novelty of our study is that it examines the temporal dynamics of brain functional networks, which have not been investigated in an irritability context before. To better understand the neural basis of youth irritability, future research should examine different (social and non-social) irritability paradigms simultaneously using both static and dynamic FC approaches in a larger sample. Nonetheless, our study provides initial validation of an irritability task that captures the social nature of irritability during adolescence using a dynamic FC approach. Importantly, the current study highlights the value and feasibility of co-produced research as a way to design and develop fMRI paradigms that reflect the real-world experience of youth.

7.5.2 Differences in dynamic functional network properties correlate with depressive symptoms and irritable mood

We observed stronger and more numerous correlations between depressive symptoms and irritable mood and the probability and duration of brain networks in the irritability task compared to the resting-state condition. This provides further preliminary evidence that our irritability task may access a phenotypically relevant state. Looking at the functional networks

related to depressive symptoms, we find that the duration of a PL state characterised by limbic (e.g., the amygdala, hippocampus, cingulate, and insula), occipital and frontal regions is negatively associated with depressive symptoms and irritable mood during the irritability task. As our study is the first dynamic functional connectivity study focused on irritability, it is difficult to situate these findings within the broader literature, especially because exploring static FC approaches or traditional frustrative non reward paradigms alongside our novel task was beyond the scope of our pilot study. However, previous dynamic FC studies in adult depression have shown increased variability in fronto-limbic networks and the DMN (Kaiser et al., 2016; Long et al., 2020). Here, the shorter duration that youth with higher depressive symptoms spend in fronto-limbic networks could represent a similar increased variability, which is in line with the aforementioned studies. On the other hand, we found a positive correlation between duration of the DMN occupancy and depressive symptoms, suggesting reduced temporal variability. However, this finding is consistent with another recent study focused on adolescent-onset depression (Marchitelli et al., 2022), which found reduced temporal variability in the DMN but also in the limbic network and FPN. Together, these mixed findings underscore the need for further research in this area.

Though it is difficult to contextualise our irritability task findings with regards to the broader fMRI-depression literature, a handful of standard resting state studies using the LEiDA approach have emerged in recent years and provide some interesting considerations (Alonso Martínez et al., 2020; Figueroa et al., 2019). Alonso-Martínez and colleagues also used a non-clinical depression sample but grouped individuals into high and low depression during a resting state scan. They found that participants in the high-depression group spent more time in a brain state that connected the DMN (particularly the precuneus) and FP network, and less time in the visual and dorsal attention network. The authors suggests that, in line with other studies, the pattern of shifting from the DMN to other networks could reflect maladaptive thinking styles, such as heighted rumination (Hamilton et al., 2015; Marusak et al., 2017). The other depression relevant LEiDA study by Figueroa et al. focused on remitted MDD and found that individuals recovering from depression were less able to access a widespread brain network linking frontal, DMN, striatal and attention areas, and when this state was occupied, it lasted for a shorter duration. This finding could reflect a reduced ability

of individuals with remitted-MDD to engage cognitive control processes, especially those related to emotion regulation. The current findings demonstrate a similar pattern of results — during the rest condition, a brain state comprising a fronto-limbic-parietal network was negatively associated with depression symptoms, which may suggest impaired cognitive control. Importantly however, it is not appropriate to compare our behavioural findings directly to the related LEiDA studies because of the different analysis designs employed. Due to the pilot nature of our study, our non-clinical sample was not large enough to examine between-group differences in dynamic functional brain states between depressed and non-depressed youth. Further, the associations observed between functional brain states and clinical measures were correlational in nature, many of which did not survive correction for multiple comparison, and therefore any conclusions presented here are tentative. Notably, motion was not found to be associated with dynamic brain state properties in the current study which is important given the known concerns surrounding motion-related noise artefact in dFC research (Chen et al., 2017).

7.5.3 Strengths and limitations

While the novelty and co-produced nature of this study is a clear strength, it is not without its limitations. Firstly, the small sample size (N = 29) significantly limits our ability to draw firm conclusions about how dynamic functional brain networks relate to depressive symptoms and irritable mood in adolescence. However, our proof-of-concept study has laid a strong foundation upon which to build future research. We have shown that it is feasible to work with young people to design and develop an fMRI paradigm, an experience that has been rewarding and enriching for all members of our study team. Secondly, although certain brain states (e.g., limbic-fronto-occipital networks) present during our irritability task were correlated with irritable mood, the measure we used to assess irritability captured trait irritability (Stringaris, Goodman, et al., 2012). Our study would have been strengthened if we had also asked participants to rate how irritating they found the scenarios presented in the task, and then investigated how these ratings of state-irritability related to dynamic brain state properties. Additionally, the self-directed nature of the irritability task, with no direct behavioural read-out, means that we cannot be sure that the participants were doing what they were asked to do, nor how they were doing it. For example, were the participants really

imagining being in the irritating scenarios presented during the task; did the extent to which they engaged with the task vary over the task epoch and was it dependent on the nature of the irritating scenario. This caveat could have added variability to our data, especially in youth with depressive symptoms, given that concentration and motivation difficulties are symptoms of depression (American Psychiatric Association, 2013b). It is also important to note that the data for this study was collected during the Covid 19 pandemic and thus the mental health measures collected could have been inflated by the ongoing social and personal stress associated with this period (Kwong et al., 2021).

7.6 Chapter conclusion

This study highlights the promise of whole brain dynamic FC analyses as a way to examine the neural correlates of irritability in adolescent depression using a novel irritability task co-produced with young people. The distinct functional brain state patterns observed between the irritability task and resting state condition provide preliminary evidence that validate our task as a method to examine the neural basis of irritability in a way that reflects the social context in which irritability occurs. Further, we find that higher irritability and depressive symptoms were associated with a shorter duration of time spent in networks comprising frontal, limbic, parietal, and occipital regions. This could reflect increased dynamic variability and may underpin maladaptive cognitive processes, such as impaired emotion regulation and rumination, that characterise depression. Using a novel LEiDA approach together with our own co-produced task, our work lays a strong foundation upon which to explore aberrations in dynamic functional brain networks associated with core symptoms of high-burden diseases such as depression, which will help pave a path towards novel intervention targets.

8 General Discussion

8.1 Chapter Introduction

The overall aim of this thesis was to understand how brain structure and function, pubertal maturation, and irritability, all of which undergo immense change during adolescence, are associated with depression during adolescence. A central feature of this thesis was the use of multi-modal neuroimaging methods (structural and functional MRI), robust methodological approaches (e.g., multi-informant reports, registered report publishing format, co-produced research design), and whole brain analyses — which help characterise the biological and psychosocial factors associated with the increased vulnerability to depression in adolescence.

In this chapter, I first summarise the findings of this thesis before discussing these results in the context of the broader literature (a more specific discussion of the findings can be found in the individual results chapters). More specifically, I give an overview of the brain features associated with adolescent depression across the three studies in this thesis. I then situate these findings within conceptual frameworks that have been proposed to explain this pattern of results. Following this, I discuss the limitations of this thesis, including methodological concerns, and highlight directions for future research.

8.2 Summary of findings

As discussed in Chapter 1, adolescence is a peak time for the emergence of depressive disorders, and adolescent-onset depression is associated with a more recurrent and chronic illness course. Existing evidence points to both biological and socio-environmental risk in the aetiology of depression but the mechanistic pathways leading to depression remain poorly understood. If we can better understand how the immense biological, psychological, and social changes (and their associated mechanisms) that characterise adolescence contribute to this heightened vulnerability to depression, we will increase our chances at identifying

modifiable targets for intervention that can move young people away from mental illness towards wellbeing.

This thesis directly addressed several key questions that will help fill the identified knowledge gaps, the first of which was to examine how brain structure (cortical grey matter and white matter microstructure) is associated with the emergence of depression in early adolescence (youth aged 9-10 years), as discussed in Chapter 3. Using the largest available sample at the time, we found that youth depression (characterised as a MDD diagnosis and depressive symptom severity) was associated with similar alterations in brain structure to those seen in adult samples. For example, we found that youth depression was associated with reductions in global cortical volume and fractional anisotropy, as well alterations in individual white matter tracts such as the superior longitudinal fasciculus and cortico-striate tract. Alterations in these white matter microstructural measures have previously been associated with cognitive control difficulties and thus may underlie the clinical manifestations of depression. However, using this baseline ABCD data, we also saw that global and regional surface area reductions were associated with youth depression. This depression-related imaging feature is not commonly seen in adult depression but has been reported in other adolescent depression neuroimaging studies. Thus, surface area reductions could reflect a specific adolescent-onset vulnerability to depression and may be associated with known risk factors to depression early in development, such as early life trauma (Opel et al., 2019).

Our use of a multi-informant approach to assess youth depression revealed a number of important and unexpected associations that we explored further in post-hoc analyses. Namely, parent report of youth depression demonstrated stronger and numerous brain structural associations compared to youth report, which were not influenced by parental current mood or youth medication use. We found that youth were more likely to report somatic or internally focused symptoms (e.g., self-esteem, guilt), while adults more commonly reported youth concentration difficulties and functional impairments, which is consistent with previous literature (Lewis et al., 2012). Further, greater reporter discrepancy between parent and youth was found to be associated with factors such as family conflict and youth sleep disturbance, while a lower discordance in reports was related to better social

cohesion. Taken together, these findings suggest that while brain structural differences may exist early in the disease course of depression, associations between socio-environmental factors and reporter discrepancy warrant further consideration in the assessment of youth mental health (in both research and clinical settings), as well as their role in the aetiology of depression.

The work undertaken in Chapter 4 directly builds upon the findings in Chapter 3 by exploring how other key biological events of adolescence, namely, the onset of puberty, is associated with depression risk, and the role that brain structure may play in this association. This project was undertaken as a registered report, whereby our hypotheses and analysis plan underwent peer review prior to data analysis.

By leveraging the ongoing release of longitudinal data from ABCD, the results of this registered report demonstrated that youth with accelerated pubertal development (relative to same-age, same-sex peers) at ages 10-11 years were at an increased risk for depression two years later, when they were aged 12-13 years. There are marked differences in the puberty maturational timelines of males and females, whereby on average, females begin puberty about 18 months ahead of males. Given that the sample in this study were in the early to mid-stages of adolescence, and thus females were more likely to be further along in their pubertal maturation, the analyses were stratified by sex. Although earlier pubertal timing was associated with depressive symptoms in both males and females when controlling for age and race/ethnicity, the magnitude of effect was larger in female youth. Further, when we accounted for additional depression risk factors (e.g., family income, parental depression, and youth BMI), the association between earlier pubertal timing and depression was no longer significant in males. Controlling for age and race/ethnicity, our exploratory analyses demonstrated that while adrenarcheal and gonadarcheal timing equally contributed to the observed effect between earlier pubertal timing and depression in females, the effect was driven by the former aspect of pubertal development in males. This suggests that biological and socio-environmental factors may differentially influence depression risk in males and females that begin puberty ahead of their same-sex, same-age peers. Future work should consider objective (e.g., hormonal assays) and questionnaire-based assessments of pubertal

development alongside qualitative methods so that we can better understand the experience of going through puberty in the present day, and how this may differ across settings (e.g., school, home, amongst peers) and between males and females.

Unlike the first hypothesis of our registered report, we did *not* find that our *a priori* brain structural measures, which primarily consisted of reduced cortical thickness and volume in frontal and temporal regions, as well as lower global cortical and white matter microstructural metrics, mediated the association between earlier pubertal timing and later youth depression. Further, when we undertook exploratory whole brain analyses to examine additional brain regions not specified in our confirmatory analyses, we did not find any robust evidence of mediation in the current sample. However, our results do demonstrate widespread associations between earlier pubertal timing and reductions in cortical thickness and volume in frontal, parietal, and temporal regions, which may represent accelerated neurodevelopment. This work makes a significant contribution to the existing literature given that we specifically examined pubertal timing associations with brain development by using a residual timing score rather than only controlling for age, as discussed further in Chapter 4.

Taken together, our findings replicate prior work and demonstrate that accelerated pubertal development is a risk factor for adolescent-onset depression. Although we did not find that brain structure mediated this association, it is somewhat unsurprising that a unimodal measure of brain development was not found to significantly influence the association between earlier pubertal timing and psychology. Given the complex interaction between pubertal maturation, brain development, socio-environmental factors and mental health outcomes, future work that adopts a longitudinal network-based approach will likely yield the most fruitful, *and* clinically meaningful findings so that we can understand how the complex web of interactions that constitute development give rise to positive and negative outcomes.

In Chapter 5, the focus of this thesis shifts from utilising “Big Data” as a way to study adolescent brain development and depression to a small, locally collected sample. Although the sample size of this study was 0.24% the size of the ABCD sample, our local neuroimaging study had the key advantage of including a bespoke fMRI task that was co-produced with

young people to answer a specific research question. Co-producing developmental science can help improve the ecological validity of our research design and methods and ensure that they are relevant to the lives of young people today (Whitmore & Mills, 2021). Thus, the primary aim of this pilot study was to validate this novel irritability task by comparing functional brain network properties during this condition to a standard resting state scan. Given that the brain is a dynamic system whereby brain activity is known to fluctuate over time, we adopted a dynamic FC approach. We used the novel data-driven LEiDA method to explore how functional brain network properties, namely fractional occupancy and dwell time, varied across the irritability and rest conditions, and how these properties related to depressive symptoms and irritable mood.

In Chapter 7, we validate our novel irritability task by showing that time-varying properties of functional brain networks (as derived by LEiDA) differed across conditions. More specifically, we found that parts of the DMN were more likely to be occupied, and occupied for a greater length of time, during the rest condition compared to the irritability condition. On the other hand, a brain state characterised by the FPN was occupied more frequently, and for a longer duration, during the irritability condition compared to the rest condition. The DMN is thought to be more active at rest while the FPN is involved in the rapid and flexible coordination of goal-oriented behaviour. Thus, the current findings provide preliminary evidence that our co-produced irritability task induces a brain state different to that at rest, and one that may relate to irritability or emotion regulation. Further, we found that the occupancy and duration of brain states characterised by a fronto-limbic-occipital network were negatively associated with depressive symptoms and irritable mood in this sample. Although the generalisability of the findings from this pilot study are limited due to the small sample size, they highlight the value of co-produced research with young people and lay a strong foundation for future work of this kind in developmental science.

8.3 Limitations, methodological considerations, and future directions

More specific discussions of the limitations of the research undertaken in this thesis are contained within individual chapters. Here, I discuss broader limitations of developmental

cognitive neuroscience research as it stands today, which are relevant to the work in this thesis as well. I also highlight considerations for future research.

8.3.1 Depression-related neuroimaging features and current conceptual frameworks

Considering the existing literature on typical neurodevelopment, the results of the current thesis suggest that brain structural associations with depression may represent an accelerated neurodevelopment. For example, the global and/or regional volumetric and surface area reductions that were found to be associated with youth depression in Chapter 3 may indicate precocious development given that these cortical morphometric features have been found to decrease across childhood and into adolescence (Norbom et al., 2021; Tamnes et al., 2017). The depression related imaging features observed in Chapter 4 show a consistent pattern of results, particularly in the smaller volumes of some subcortical nuclei (e.g., nucleus accumbens) which have been found to exhibit subtle volumetric decreases across this period of development (Herting et al., 2018). Our findings on the brain structural features (e.g., reduced global and regional cortical volume and thickness) associated with earlier pubertal timing also suggest accelerated neuromaturation. It has been suggested that the advanced development of the brain (and other biological processes, e.g., puberty) may serve an adaptive function under adverse conditions (Callaghan & Tottenham, 2016; Colich et al., 2020).

Indeed, evidence for this "Stress Acceleration Hypothesis" as proposed by Callaghan and Tottenham (2016) has been found in cross-sectional research on the impact of early life adversity (e.g., trauma, socio-economic disadvantage) on neuromaturation. Although this conceptual framework was originally applied to emotional brain circuits (e.g., fronto-limbic), it is thought to apply more broadly to an overall accelerated biological ageing following adversity (Colich et al., 2020; Gur et al., 2019; Rakesh, Cropley, et al., 2021). This theoretical framework is complementary to more domain-specific developmental models, such as the maturation disparity hypothesis (Brooks-Gunn et al., 1985; Ge & Natsuaki, 2009). This accelerated biological ageing may manifest as early pubertal timing, which in turn creates a

mismatch between biological and cognitive/social development, and thus confers an increased vulnerability to poorer developmental outcomes like psychopathology.

Importantly however, longitudinal research is needed to truly assess developmental patterns and the factors that shape different trajectories. Unfortunately, large-scale longitudinal research on adolescent brain development with multi-modal data is scarce at present but will soon be available with the release of follow-data from cohort studies like ABCD. Therefore, any conclusions made about the biological mechanisms influencing developmental outcomes using cross-sectional data are tentative at best. Moreover, inconsistencies have emerged across different developmental domains. For instance, if accelerated neurodevelopment does occur due to early life stress, it is unclear why youth exposed to adversity do not exhibit more advanced cognitive function compared to their peers. In fact, they seem to be at a significant disadvantage as exposure to adversity early in life is associated with lasting differences in cognitive and socio-emotional development (Humphreys et al., 2015; Nelson et al., 2007).

The importance of considering the limitations of cross-sectional research to study development is underscored further by divergences between cross-sectional and longitudinal findings. For example, a recent review by Rakesh & Whittle (2021) found that while socio-economic disadvantage seemed to be associated with accelerated structural neuromaturation in cross-sectional studies, longitudinal research suggests the contrary. A number of longitudinal studies have found that youth who have experienced early life stress (e.g., poverty, social deprivation) are more likely to exhibit a pattern of brain development that suggests a maturational lag (i.e., a rate of change with a flatter slope) relative to young people who have not experienced environmental stress. This delayed maturation can be seen at a global level (e.g., cortical volume) as well as regionally, such as in the volume/thickness of the temporal, frontal, and parietal lobes, the hippocampus, and amygdala (Barch et al., 2020; Hair et al., 2015; Whittle et al., 2017). Moreover, these findings may help explain why youth from more deprived backgrounds tend to exhibit poorer cognitive function compared to their more privileged peers (Noble et al., 2007).

Synthesising the literature on brain functional development is more challenging relative to brain structural findings due to the wide array of tasks used to probe the same cognitive domain (e.g., N-back vs. delayed match-to-sample tasks are both used to assess working memory) and the small samples used in fMRI studies (Rakesh & Whittle, 2021). However, resting state findings are beginning to emerge from large studies like ABCD and suggest that early life deprivation is associated with widespread alterations in the functional connectivity of sensorimotor and executive control networks (Herzberg & Gunnar, 2020; Rakesh, Seguin, et al., 2021) . While this may represent atypical development, longitudinal work is first needed to chart typical functional brain development (both static and dynamic FC approaches) before we can establish deviations from the expected trajectory, and how they relate to developmental outcomes.

Collectively, the findings from this thesis and the extant body of evidence highlight that the relationship between the functional and structural features of the adolescent brain and the emergence of depression during adolescence is one of great complexity. Moving forward, our research methods must adopt an approach that reflects this nuance, a consideration that is gaining increased attention. A recent review by Ferschmann and colleagues (2022) calls for the field to adopt an "ecological neuroscience" perspective in the study of brain development that reflects the transactional interplay with a person's physical and social environment, and how this affects neurocognitive development (Ferschmann et al., 2022). If the overarching aim of our research as developmental cognitive neuroscientists is to understand the factors that influence brain development and how they in turn predict life trajectories, we need to prioritise refining our current methods. Although the associations between brain features and developmental outcomes are statistically significant, they explain a tiny proportion of the overall variance, which is problematic is we want to identify methods that facilitate causal inference (Marek et al., 2022).

8.3.2 Small effect sizes

Small effect sizes are consistently found in large-scale neuroimaging studies — the ABCD Study is no different, as reflected by the effect sizes reported in Chapters 3 & 4. There has been much debate about how meaningful these neuroimaging relationships are and what

they really add to our ability to predict developmental outcomes when compared to clinical and psychosocial data (Dick et al., 2021). The small amount of variance explained is a problem not just limited to neuroimaging research, genome wide association studies face the same difficulties in predicting complex traits, like psychiatric disorders (Timpson et al., 2018). On one hand, large sample sizes like that of ABCD are sufficiently powered to detect small, subtle effects, and some have argued that small effects can be meaningful as they accrue over time at the population level (Funder & Ozer, 2019). Further, due to the relatively small sample sizes in neuroimaging research up until very recently, the larger effect sizes previously reported are likely to have been greatly inflated. Given the complex web of interactions that shape development, as previously discussed, it is perhaps unsurprising that an individual measure (e.g., global cortical volume or the volume of the nucleus accumbens) only explains a small amount of variance in our outcome of interest (e.g., depression). Further, the vast majority of research within the field to date has studied neuromaturation and developmental outcomes from a group-level perspective, which is also likely to have contributed to the small effect sizes reported. If we are ever to use an "ecological neuroscience" approach that combines neuroimaging with other biological and psycho-social factors to reliably predict developmental outcomes, we need to be able to do so at the level of the *individual.*

8.3.3 Individual variability in brain development

Understanding individual variability in how the human brain changes across adolescence remains one of the most understudied areas in neuroscience today. Although large longitudinal MRI studies in developmental samples have only recently emerged, the focus of research in this area has been on group-level trajectories. As previously discussed, interpretations that have arisen from this work include that early life adversity is associated with a maturational lag in brain development, which in turn confers an increased vulnerability to psychopathology in adolescence. Even in cross-sectional work, like that presented in the current thesis, the reduced cortical thickness found to be associated with adolescent depression and earlier pubertal timing, has been interpreted as accelerated brain maturation, for example. While the observed patterns may indeed characterise specific developmental outcomes, the known heterogeneity in brain development across individuals means that it is impossible to draw any firm developmental conclusions (Becht & Mills, 2020). For example,

the significant variability in hippocampal volume in typically developing children and young adults makes it impossible to firmly conclude that depressed individuals with lower hippocampal volumes exhibit accelerated maturation (Tamnes, Bos, et al., 2018).

Recent work that has moved beyond group-level estimates suggests that individual variability in brain morphometric changes over time varies across structures, by sex, and is more likely to occur at the transitions into and out of adolescence (Mills et al., 2021; Tamnes, Bos, et al., 2018). Similarly, longitudinal investigations of resting state and task-based functional connectivity MRI have shown that there are considerable inter-individual differences in connectivity estimates across childhood and adolescence (Peters et al., 2016; Telzer et al., 2018; van Duijvenvoorde et al., 2019). Recognising the individual variation inherent to human development must be a, if not *the,* central consideration in our effort to identify how deviations from typical development relate to the emergence of psychopathology during adolescence. Further, knowing when these periods of increased variability occur will help us identify developmental windows when interventions may be their most effective or determine which cognitive processes are likely to be the most responsive (Becht & Mills, 2020). Given that at least three timepoints of data per individual are needed to model different growth trajectories, which are not yet available in the ABCD Study, exploring individual variation in brain development and how this relates to depression risk in adolescence, was not possible in the current thesis. However, the data needed to model individual developmental trajectories will soon be available (expected end of 2023), which makes it a very exciting time to be a researcher in developmental cognitive neuroscience!

While the examples discussed to date have centred on individual variation in brain development, this heterogeneity extends to myriad aspects of our biology. Studying individual differences in biological ageing have become increasingly popular in recent years as large amounts of biological and phenotypic data (e.g., Ns >100,000) have been made available to researchers through cohort studies like the UK Biobank. Normative modelling, an emerging machine learning method, can be used to chart population level patterns in brain development, for example, from which the extent that an individual's development deviates can then be calculated. A notable application of normative modelling was the publication of

"brain charts for the human lifespan" earlier this year (Bethlehem et al., 2022). Further, other researchers have harnessed the notion of individual variability to develop personalised body and brain "biological clocks" (Ferrucci et al., 2020; Khan et al., 2017) . This systems approach to the study of human development has provided insight into multi-organ aging, and how different biological systems interact. For example, it was found that heart and pulmonary age were most strongly associated with increased brain age, such that a one year increase in cardiovascular age, explained a 27 day increase in brain age (Tian et al., preprint). However, this work has only been conducted in adults to date and there are important developmental considerations when applying such methods in younger cohorts, such as the faster rate of neurodevelopment earlier in life relative to middle or later life (Vidal-Pineiro et al., 2021).

Nonetheless, as longitudinal data becomes available, the work presented in this thesis could be expanded to examine individual differences in biological ageing across multiple domains (e.g., brain age (indexed via functional and structural measures), pubertal tempo (physical and hormonal measures), and cognitive processes) and how this relates to developmental outcomes. For example, it remains unknown whether individual variation in accelerated biological maturation or degeneration is associated with depression risk or relapse during adolescence and into young adulthood. Understanding the factors associated with the continuity (or discontinuity) between different biological ageing processes may provide better insight into the biological mechanisms underpinning risk for psychopathology in adolescence. Crucially, this could help us detect which individuals are most at risk for depression and provide interventions informed by their biological aging profile and the socio-environmental context in which this is occurring. Intervening on youth depression before it becomes chronic or severe would allow young people to participate in society actively and more meaningfully during adolescence and beyond.

8.3.4 Weird results from W.E.I.R.D samples?

To date, most developmental cognitive neuroscience research (including the work in thesis) is based on a small slice of society from western, educated, industrialised, rich, and democratic (WEIRD) cultures (Henrich et al., 2010). Therefore, our understanding of the factors and experiences that shape brain development may only be relevant to a limited

proportion of the population. Whether our findings apply to other cultures or populations remains unknown at present and work in other settings is urgently needed. This is particularly pertinent in developmental research given that 90% of adolescents in the world live in low- and middle-income countries (LMICs; United Nations, 2017). Further, in the context of depression research, there may be cultural differences (e.g., language used to describe depressive symptoms, perceived stigma) between developed and developing countries, which will need to be taken into consideration when expanding research in these areas. Thankfully, several funding bodies (e.g., UK Medical Research Council, Wellcome Trust) have made health research in LMICs a strategic priority. For example, Generation Malawi is a new longitudinal cohort study of the mental and physical health of families in Malawi, whose study design was modelled on the Generation Scotland cohort in the UK.

As we broaden the scope of developmental cognitive neuroscience research globally, it is also important to reflect on the representativeness of the research samples that exist in our own communities. Within WEIRD cultures, the individuals that take part in research studies are not representative of the population at large. This phenomenon is known as selection bias and is a widely acknowledged problem in population-based cohort studies. For example, compared to the general population, a research participant in UK Biobank is more likely to live in a less socially deprived area, be more highly educated, of white ethnicity, female, and have fewer self-reported health conditions. (Fry et al., 2017). This selection bias also applies to developmental cohorts (Keiding & Louis, 2016). Aware of the aforementioned caveats of cohort study recruitment, the ABCD Study aimed to recruit a sample that reflected US population demographics by using a school-based probability sampling recruitment strategy (Garavan et al., 2018a). While an admirable effort was made by ABCD to recruit a representative sample, the final baseline sample, while demographically diverse, should not be considered representative of all 9–10-year-olds in the US. For example, 23% and 11% of ABCD participants came from households with a master's and doctoral degree, respectively, while the US population average for households with a master's degree is 12% and 5% for a doctoral degree (Garavan et al., 2018a). Further, household education, along with race and distance from study site were found to be the strongest predictors of sample attrition in

ABCD, which should be a key consideration for all researchers working with this dataset going forward (Feldstein Ewing et al., 2022).

8.3.5 Measurement considerations and applications

Considerations about the measures and methodologies used in this thesis have been discussed within individual chapters, including reporter discrepancy between parent and child report, limitations of existing measures of pubertal development, and static versus dynamic FC approaches. Here, I briefly extend the discussion on a consideration that applies to all aspects of developmental research, namely youth engagement and co-production, which will be essential in subsequent research.

As discussed in Chapter 5, co-producing research with young people will ensure that our research questions and methods are relevant to the lives of adolescents today. Given the significant emphasis on the environmental pathways that influence neurodevelopment, it is no longer sufficient to simply collect data from participants — we must also consider how the community might benefit from our research findings. The ABCD Study has working groups (e.g., the ABCD Outreach and Dissemination Working Group) whose key mission is to help participating families learn about the findings emerging from the study and how this might benefit communities. The voices and concerns of young people and their families should inform how we conduct and disseminate our research. For example, it will be important to consider how current affairs and societal issues (e.g., the rise of gun violence and restricted access to abortion services in the US, Black Lives Matter protests, the escalating climate crisis and associated climate anxiety, as well as the rise of political autocracy across the world) affect the developmental outcomes of young people today. Thankfully, the ABCD Study team are continuously adding new measures to their study protocol and data that capture the subtleties of the environment in which young people are growing up will be available to researchers. Moving forward, community advisory boards (e.g., Young Person Advisory Groups) can provide a framework for developmental neuroscientists to establish and sustain authentic partnerships between the community and academic institutions to ensure that we continue to collect data that are relevant to our study of development.

Adopting a co-produced partnership approach to research is more important than ever given the increasing levels of distrust in science, which has likely been driven by the spread of misinformation online (Philipp-Muller et al., 2022). However, the technological advances of today can also significantly aid our study of development. For example, intensive longitudinal research methods, such as ecological momentary assessment (EMA), whereby data are collected from individuals at multiple times throughout the day, are increasingly being used to study how subtle variations in behaviour relate to health outcomes. Smartphone app-based data collection has the potential to dramatically advance our understanding of the neuroscience of mental health by enabling the collection of a vast quantity of data (e.g., via EMA measures or wearables like Fitbits) from much larger and diverse samples (Gillan & Rutledge, 2021). This data could then be used to complement more traditional lab-based data collection (e.g., MRI). Broadening the scope of the methods used within this thesis, if we wanted to advance our understanding of why youth that begin puberty ahead of their peers are at an increased risk for depression, we could adopt an EMA approach. For instance, alongside the collection of MRI and biospecimen data (e.g., pubertal hormones), we could gather information on the social experience of going through puberty on a daily/weekly/monthly basis (e.g., body perception, self-esteem, loneliness, attitudes towards peers, family, sport and social activities). This could advance our understanding of factors that distinguish young people who navigate adolescence but remain well from those who develop mental health difficulties, and help inform targeted intervention strategies.

8.3.6 Shifting the focus away from risk factors

The work contained within this thesis, as well as the literature reviewed, has overwhelmingly focused on risk factors for youth depression. While this approach reflects the aims and objectives of this doctoral thesis, future work should consider the mechanisms associated with resilience promoting factors. While the burden of depression falls heavily on youth, most young people transition through adolescence without developing a mental health disorder. Given that it may be easier to facilitate the introduction of resilience promoting factors into the lives of young people (e.g., social cohesion, physical exercise, access to green space, adaptive coping strategies) than mitigating known risk factors for depression (e.g., social deprivation, family history, trauma), it is important that we understand the processes of

"what went right" rather than being overly focused on "what went wrong". By embracing the considerations previously discussed (e.g., longitudinal research focused on individual variability with an ecological neuroscience perspective as well as co-produced research), we can strengthen our chances of developing interventions that significantly change the worrying state of youth mental health today.

8.4 Concluding remarks

Where can it be found again,
An elsewhere world, beyond

Maps and atlases,
Where all is woven into

And of itself, like a nest
Of crosshatched grass blades?

-Seamus Heaney (*Human Chain*, 2010)

It will be no easy task, but the "elsewhere world" of adolescence and the factors that shape developmental outcomes can be "found" and understood. This is a challenge that we must tackle head on if we are going to identify modifiable factors that impact neurodevelopment so that we can develop interventions that translate to sustained improvements in youth mental health. But importantly, we only stand a chance at unravelling this complex "nest" of interactions if we adopt an ecological neuroscience approach by connecting the impact of physical, social, and environmental influences on neurodevelopment. This approach must also be transparent and reproducible by adopting open science practices and ecologically valid by co-producing this work with young people — the work contained within this thesis has strived to embrace these frameworks as much as possible. However, given the

complexities of development, “team-science” and collaboration will be crucial to reaching our collective goal of creating a world that gives young people every chance to flourish in their development as they fly the nest of adolescence and enter the strange and wonderful world of adulthood.

9 Appendices

Appendices 1-3 for this thesis are included in this section and are also available in an online repository here.

Appendix 1: Supplementary Information for Chapter 3

Brain structural associations with depression in a large early adolescent sample (the ABCD cohort)

Shen & MacSweeney *et al.*

Supplementary methods and results

Scanning protocols

Protocols used for data acquisition and processing were described elsewhere(Garavan et al., 2018b; Lopez-Leon et al., 2008). In brief, T1-weighted data was acquired by magnetisation-prepared rapid acquisition gradient echo scans with a resolution of 1×1×1 mm^3, which was used for generating cortical structural measures, and diffusion-weighted data was obtained by high angular resolution diffusion imaging scans, used for generating white matter microstructural measures.

Unrelated participants

Unrelated participants were selected by keeping the first entry of each individual family ID (rel_group_id: https://nda.nih.gov/general-query.html?q=query=data-element%20~and~%20searchTerm=name:%20%22rel_group_id%22).

QC criteria for brain structural measures

For cortical measures, QC measures on raw imaging was first applied to remove poor-quality raw T1 scans (data field: 'iqc_t1_ok_ser'). Participants that had low post-processing quality check scores (<1) for Freesurfer outcome identified (data field: 'fsqc_qc') were then removed.

For white matter microstructural measures, participants with poor-quality raw diffusion imaging scans (data field: 'iqc_dmri_ok_ser'), poor-quality raw T1 scans (data field: 'iqc_t1_ok_ser') and poor quality of FreeSurfer parcellation for T1 data (data field: 'fsqc_qc', potentially indicating low quality for T1 scans) were removed from further analyses. We also removed subjects that had low score for post-processing (data field: 'dmri_dti_postqc_qc'). Finally, as there were extreme values that caused the distributions of global white matter measures heavily skewed, we removed participants showing global FA and MD values 5 standard deviations from mean (see Figure S2 for sample sizes and Figure S3 for distribution before and after removing outliers).

We used the data field 'dmri_dti_postqc_qc' for screening data after preprocessing. The variable is a binary variable, covering 3733 people. Among them, 224 did not pass QC and 3509 passed. To test the impact of post-processing QC, first we showed the standardised values of global fractional anisotropy (FA) and mean diffusivity (MD) (after removing outlying values of -/+ 5 std away from mean) in Figure S3. Secondly,

we conducted analysis of this subsample (N=3733) with post-processing QC data available and for those who passed QC in this subsample (N=3509). The main model of testing associations between caregiver/child report of MDD/depressive symptoms (DS) and regional brain measures was used. As shown in Figure S21, overall correlation of the standard regression coefficient/Cohen's d for all the tests conducted were highly correlated (r=0.918). The uncorrected p-values were also highly correlated (r=0.789).

Deriving life-time MDD definition from current and past MDD status

The curated data from ABCD (Adolescent Brain Cognitive Development) study contains three MDD (Major Depressive Disorder) definitions: MDD current (field names: ksads_1_840_t and ksads_1_840_p for reports by children and caregivers, respectively), MDD past (field names: ksads_1_842_t and ksads_1_842_p for reports by youths and caregivers respectively) and MDD current in partial remission (field names: ksads_1_841_t and ksads_1_841_p). When defining the lifetime MDD definitions, participants who were ***either*** a case for MDD current or past definition would be identified as cases, and those who were controls for ***both*** current and past MDD definitions were identified as controls. Those who were cases for MDD current in partial remission were also identified as cases. Life-time definitions were derived for reports by caregivers and youths separately.

Deriving total scores of DS

DS were generated based on DSM-V (Diagnostic and Statistical Manual of Mental Disorders, 5th edition)(American Psychiatric Association, 2013a) criteria for the severity scale of. For each item, a binary outcome indicates whether the single symptom met clinical significance (1=Yes and 0=No). For each item, a life-time score was generated using the same method for generating lifetime MDD definition. The life-time score was then used for generating DS.

In total there were 28 items used, which covered 15 individual DS (see Table S2). Among the 15 individual DS, depressed mood, anhedonia and fatigue were core symptoms and the rest of 12 DS were secondary symptoms. Life-time scores of these individual DS were used to generate a total measure of DS that includes four severity levels: severe, moderate, mild and none of the above. A detailed description of the total measure of DS can be found in Table S3.

Agreement between caregiver and child report

We tested the agreement between caregiver and child report. Proportion of agreement for MDD diagnosis between caregiver and child report was estimated using the 'agree' function in R package 'irr' (version 0.84.1, https://cran.r-project.org/web/packages/irr/irr.pdf). Tolerance of disagreement was set as 0. Cohen's Kappa of DS between caregiver and child report was conducted using the 'kappa2' function from the 'irr' R package (version 3.6.2). For both analyses, only participants with non-empty values for both caregiver and child reports were included ($N_{MDD\ definition}$=8635, $N_{DS\ measure}$=8599).

Proportion of agreement between caregiver and child report of MDD diagnosis was 95.9% (Table 2). Among all participants, 8273 were identified as controls by both child and caregiver report (95.8% of the total sample), 182 were cases based on caregiver report but not child report (2.11%), 168 were cases according to child report but not caregiver (1.95%), and finally 12 were cases according to both reports (0.14%). Agreement between caregiver and child reported DS was reported in the main text.

Average and discrepancy of DS reported by caregivers and children and its association with brain structural measures

Average reports for the severity of depression was generated for each child-caregiver pair. Results for the associations between the average severity and general/regional brain measures are shown in Figures S17-19.

Discrepancy was generated by obtaining the absolute values of subtracting caregiver and child reports of DS (Figure S4). Associations between discrepancy of DS and general brain structural measures were tested (Figure S22). For regional measures, the associations with discrepancy of DS were tested on those brain measures that associated with caregiver report of depression (Figure S23).

Validating DS measured by KSADS and CBCL

In the present cohort, two types of mental health scales were used. Although we reported results based on KSADS, we tested DS assessed by CBCL (The Child Behaviour Checklist) as an additional analysis and compared the results produced using KSADS and CBCL. DS assessed using CBCL was based the DSM-5-oriented items.

KSADS-derived scores of DS reported by caregivers showed high correlations with CBCL's depression score based on the DSM-5 (Diagnostic and Statistical Manual of Mental Disorders, 5th Edition) scale (r=0.346). The KSADS DS reported by youths showed weak correlation with CBCL depression score (r=0.076). See Figures S24.

CBCL DMS-5-oriented score of depression reported by caregivers showed consistent results with DS assessed by KSADS reported by caregivers (Figures S25-26). Associations were found in general cortical surface area, volume and white matter FA (β ranged from -0.025 to -0.043, p ranged from 1.86×10^{-3} to 1.53×10^{-6}, Figure S25). Overall effect sizes for regional brain measures showed high correlation with results for KSADS-measured DS (r = 0.713, see Figure S26).

Depressive problems of caregivers

A subscale of DSM-5-oriented items for depressive problems from the Adult Self-Report (ASR) in the Achenbach System of Empirically Based Assessment was used for the one caregiver who accompanied the child to the study. We used the field 'asr_scr_depress_r' as a score for severity of depression in caregivers (N=8633). A total of 14 items were used for this scale(Barch et al., 2018b). There were five participants who did not answer any of the questions, and thus they were removed from analysis. All other participants completed the entire questionnaire. This scale of DS in caregivers was tested to indicate recent DS, and a high proportion of reporters were mothers (85.66% of all caregiver reporters, N_{Mother}=7412).

We tested the agreement between the ASR scale and self-reported history of depression using glm model, setting the binary variable of self-reported history as independent variable, and the ASR scale as dependent variable. Two measures for maternal risk showed the greatest agreement (OR=0.762, $p<2\times10^{-16}$), followed by measures for caregivers (OR=0.556, $p<2\times10^{-16}$). Measures for paternal risk showed the poorest agreement (OR=0.536, $p<4.49\times10^{-7}$).

Recent social deprivation

A measure of recent social deprivation was included in the analysis as a covariate. The measure was derived from items the parent demographics survey (https://nda.nih.gov/data_structure.html?short_name=abcd_lpds01), which was answered by parents. Participants were asked to give a 'Yes' (=1) or 'No' (=0) answer to seven questions: 1. Needed food but couldn't afford to buy it or couldn't afford to go out to get it? 2. Were without telephone service because you could not afford it? 3. Didn't pay the full amount of the rent or mortgage because you could not afford it? 4. Were evicted from your home for not paying the rent or mortgage? 5. Had services turned off by the gas or electric company, or the oil company wouldn't deliver oil because payments were not made? 6. Had someone who needed to see a doctor or

go to the hospital but didn't go because you could not afford it? And 7. Had someone who needed a dentist but couldn't go because you could not afford it?

The sum score of all the 'Yes' answers was used as the measure for social deprivation. The measure has a mean value of 0.936, with its minimum and maximum scores = 0 and 14, respectively. We included this variable as the main proxy that may confound MDD-related deficits in brain development instead of controlling more extensively for socio-environmental protective/risk factors, as the latter would have restricted our sample with complete data to less than half of the current sample (N=4,036).

Sensitivity analyses: site differences

Data was collected in 22 sites across the United States. In order to test whether site difference makes significant impact on data, we used leave-one-out method to test the robustness of results. This was conducted on associations between MDD/DS and general brain measures.

For each association (e.g. cortical volume ~ MDD reported by caregivers), the analysis was conducted 22 times. Each time, data from one site was taken out, and therefore analysis was performed on the remaining 21 sites. Effect sizes of leave-one-out analysis were then compared against the one found using the entire sample. Results are reported in Figures S9-12.

Impact of differences in scanning site were assessed using leave-one-out analysis for the significant associations found between MDD diagnosis/DS and general brain measures in the total sample, described above were assessed (Figures S9-12).

Caregiver report: All associations between MDD diagnosis/DS reported by caregivers and brain measures remained significant in all iterations of the leave-one-out analysis (p ranged from 0.037 to 3.44×10^{-9}). Therefore, the results regarding caregiver report are not likely to be driven by a single site.

Child report: Associations between DS reported by child and brain measures were also significant in all iterations (p ranged from 0.034 to 1.72×10^{-4}). For associations between MDD diagnosis reported by child and brain measures, 14 out of 22 iterations were significant (p ranged from 0.049 to 0.021) and 8 iterations were not significant (p ranged from 0.091 to 0.053) for sulcal depth, and 19 iterations were significant (p ranged from 0.047 to 0.007) and 3 were not significant (p ranged from 0.082 to 0.069) for FA. In summary, results for DS reported by child are not likely to be driven by sites, whereas results for MDD diagnosis are more heterogeneous across sites.

Sensitivity analysis: scanner differences

As the three type of scanners from the major manufacturers can show significant impacts on the estimation of intracranial volume(Casey et al., 2018b), we added scanner manufacturer as an additional covariate in a secondary sensitivity model. All significant associations found between general brain measures and MDD/DS remained significant (see Figure S13). Effect sizes for the associations between individual brain structural measures and MDD/DS showed high correlation with results of the main model (r = 0.999 for standardised effect sizes, r = 0.996 for p-values, see Figure S14).

Sensitivity analysis: impact of medication

The use of antidepressants was small in this sample (N=136). However, we also investigated the potential impact of these on our main finding by including the use of antidepressants as an additional covariate in a sensitivity analysis.

The ABCD Parent Medications Survey Inventory Modified from PhenX (short name: medsy01, URL: https://nda.nih.gov/data_structure.html?short_name=medsy01) was used to derive a variable for medication usage. The survey was used for caregivers to report medication intake by child in the past two weeks before the assessment took place. In order to estimate the effect of medication on brain imaging measures, we extracted reported medication intake that match the assessment date with the imaging assessment.

We used two types of data columns for screening medication usage. The first one is the general screening question, 'Did your child take any medications in the past two weeks and if so did you bring them with you?' (field name: brought_medications, URL: https://nda.nih.gov/general-query.html?q=query=data-element%20~and~%20searchTerm=name:%20%22brought_medications%22).

Available answers were: 0 = Yes (medications taken in the past two weeks); 1 = Yes (medication brought with the caregiver); 2 = Refused; and 3 = Took No Medications.

The second type of screening questions were reported medication names. Data columns with names of prescribed medications were used (field names: med1_rxnorm_p, med2_rxnorm_p, …, med15_rxnorm_p). A list of drug names were generated using the British National Formulary-70 (BNF 70) (https://www.bnf.org/products/books/) under the category of depression(Shen et al.,

2019b). We used these drug names as key words to find matched medications as antidepressants.

Combining the two types of questions, we generated a categorical variable to indicate medication usage:

0 = no medication brought (brought_medications==3) & no prescribed medication name reported (N=5320),

1 = medication was used for the past two weeks & no prescribed medication name reported (N=2665),

2 = medication was used for the past two weeks & at least one medication reported & none was included in the list of antidepressants (N=1942), and finally,

3 = medication was used for the past two weeks & at least one antidepressant was reported (N=136).

In the sensitivity analysis, this variable of medication usage was included as an additional covariate. Results for the association between general brain measures and MDD/DS can be found in Figure S15. Comparison of effect sizes and p values between the main model and the model controlling for medication usage, associated with individual brain measures can be found in Figure S16.

All associations found between general brain measures and MDD/DS remained significant (Figure S15). The effect sizes and p-values for the associations between measures of individual brain structures and MDD/DS were highly correlated with the results of the main model (r=0.989 for standardised effect sizes, r=0.965 for p-values, see Figure S16).

Sensitivity analysis: covarying comorbidity

We further looked at comorbid major psychiatric conditions (Bipolar I, Bipolar II, ADHD, Psychosis and Conduct disorder) reported by caregivers. Within the 194 cases, 132 reported one or more of these conditions in the present/past (68.04%) and 62 reported none (31.96%). There were 1612 MDD controls reported any of the above conditions (24.12%) and 5072 reported none of these conditions, including MDD (75.88%).

Further, a supplementary analysis was conducted to see if covarying for comorbidity may change the main findings regarding the associations between brain structural measures and parent report of depression. The additional covariate comorbidity was generated by summarising the comorbid major psychiatric conditions (1=with any comorbid condition and 0=none of the comorbid conditions reported).

Results can be found in Figures S27-28.

Measures for socio-environmental factors

Measures from ABCD sum scores culture & environment – caregiver (abcd_sscep01), ABCD sum scores culture & environment – youth (abcd_sscey01) and ABCD sum scores physical health – caregiver (abcd_ssphp01) were used. The scores for culture and environment reported by both caregivers and children were used because the questionnaires contain different items.

Item screening was conducted by removing items that have less than half of the total sample sizes (Discrimination Measure: dim_y_ss_mean and Mind Diet score: cna_p_ss_sum were thus removed from analysis). As we used the entire sample, therefore questionnaires that focus on one specific culture were not included in the analysis (ABCD Parent Mexican American Cultural Values Scale). For each participant, those that answered less than half of the questionnaire items in a single scale were set as NA for the given summary score.

A complete list of variables included in the analysis can be found in Table S4.

Table S1. Sample sizes and demographic features for those with and without MDD.

			N	Age		Sex (% of Male)
				Mean	SD	
MDD	**Reported by caregivers**	**Case + control**	6878	9.9	0.62	51.7%
		Missing data	1757	9.97	0.62	54.9%
	Reported by children	**Case + control**	6924	9.9	0.62	51.6%
		Missing data	1710	9.97	0.62	55.4%

Table S2. Items in the scale to assess Major Depressive Disorder (MDD) and depressive symptoms using the Kiddie Schedule for Affective Disorders and Schizophrenia (KSADS). When depressive symptoms were calculated, these variables were counted as one rather than two variables: insomnia and hypersomnia, decreased and increased appetite, weight loss and weight gain, and finally, psychomotor agitation and retardation.

Field name				DSM-V symptoms	Sub-items
Caregiver		Youth			
Past	Current	Past	Current		
ksads_1_1_p	ksads_1_2_p	ksads_1_1_t	ksads_1_2_t	Depressed mood (core symptom)	--
ksads_1_5_p	ksads_1_6_p	ksads_1_5_t	ksads_1_6_t	Anhedonia (core symptom)	--
ksads_1_159_p	ksads_1_160_p	ksads_1_159_t	ksads_1_160_t	Fatigue (core symptom)	--
ksads_1_161_p	ksads_1_162_p	ksads_1_161_t	ksads_1_162_t	Concentration disturbance	--
ksads_1_181_p	ksads_1_182_p	ksads_1_181_t	ksads_1_182_t	Decreased self-esteem	--
ksads_1_177_p	ksads_1_178_p	ksads_1_177_t	ksads_1_178_t	Guilt	--
ksads_1_179_p	ksads_1_180_p	ksads_1_179_t	ksads_1_180_t	Hopeless	--
ksads_1_183_p	ksads_1_184_p	ksads_1_183_t	ksads_1_184_t	Impairment in functioning due to depression	--
ksads_1_163_p	ksads_1_164_p	ksads_1_163_t	ksads_1_164_t	Indecision	--
ksads_1_3_p	ksads_1_4_p	ksads_1_3_t	ksads_1_4_t	Irritability	--
ksads_1_157_p	ksads_1_158_p	ksads_1_157_t	ksads_1_158_t	Disturbed sleep	Hypersomnia
ksads_22_141_p	ksads_1_156_p	ksads_22_141_t	ksads_1_156_t		Insomnia when depressed
ksads_1_171_p	ksads_1_172_p	ksads_1_171_t	ksads_1_172_t	Changed weight	Weight gain
ksads_1_167_p	ksads_1_168_p	ksads_1_167_t	ksads_1_168_t		Weight loss
ksads_1_165_p	ksads_1_166_p	ksads_1_165_t	ksads_1_166_t	Changed appetite	Decreased appetite
ksads_1_169_p	ksads_1_170_p	ksads_1_169_t	ksads_1_170_t		Increased appetite
ksads_1_173_p	ksads_1_174_p	ksads_1_173_t	ksads_1_174_t	Psychomotor symptoms	Psychomotor agitation in depressive disorder
ksads_1_175_p	ksads_1_176_p	ksads_1_175_t	ksads_1_176_t		Psychomotor retardation

ksads_23_953_p	ksads_23_964_p	ksads_23_953_t	ksads_23_964_t	Self-harm and suicidal thoughts/attempts	Aborted attempt
ksads_23_952_p	ksads_23_963_p	ksads_23_952_t	ksads_23_963_t		Interrupted attempt
ksads_23_951_p	ksads_23_962_p	ksads_23_951_t	ksads_23_962_t		Preparatory actions toward imminent suicidal behavior
ksads_23_945_p	ksads_23_956_p	ksads_23_945_t	ksads_23_956_t		Self-Injurious behavior without suicidal intent
ksads_23_949_p	ksads_23_960_p	ksads_23_949_t	ksads_23_960_t		Suicidal ideation active intent
ksads_23_948_p	ksads_23_959_p	ksads_23_948_t	ksads_23_959_t		Suicidal ideation active method
ksads_23_947_p	ksads_23_958_p	ksads_23_947_t	ksads_23_958_t		Suicidal ideation active non-specific
ksads_23_950_p	ksads_23_961_p	ksads_23_950_t	ksads_23_961_t		Suicidal ideation active plan
ksads_23_946_p	ksads_23_957_p	ksads_23_946_t	ksads_23_957_t		Suicidal ideation passive
ksads_23_954_p	ksads_23_965_p	ksads_23_954_t	ksads_23_965_t		Suicide attempt

Table S3. Measure of DS derived from individual items of DSM-V MDD symptomology and diagnosis of suicidality.

Severity level	**Criteria: (condition (a) OR condition (b)) AND condition (c)**		
	Condition (a)	**Condition (b)**	**Condition (c)**
Severe	Core symptoms = 3 + Secondary symptoms > 3	Any suicidal attempt	Identified as case
Moderate	Core symptoms >1 + Secondary symptoms = 2 to 3	Total symptoms = 7 to 8	Not identified as Severe
Mild	Core symptoms >1 + Secondary symptoms = 1 to 2	Total symptoms = 5 to 6	Not identified as Moderate/Severe
None of the above		None of the above	

Table S4. Measures of cultural and social environment, family environment, physical health and sociodemographic status. Field names are the column names used in the original ABCD curated data. For those measures that used multiple data fields, notes are provided for the methods of creating them.

Field name(s)	Description	Category	Note
meim_p_ss_total	Caregiver: Multi-group ethnic identity (total scale)	Cultural and Social Environment	
via_p_ss_hc	Caregiver: Vancouver index of acculturation (heritage culture)	Cultural and Social Environment	
via_p_ss_amer	Caregiver: Vancouver index of acculturation (American 'mainstream' culture)	Cultural and Social Environment	
nsc_p_ss_mean_3_items	Caregiver: Neighbourhood safety	Cultural and Social Environment	
srpf_y_ss_ses	Child: School risk and protective factors (school environment)	Cultural and Social Environment	
srpf_y_ss_iiss	Child: School risk and protective factors (school involvement)	Cultural and Social Environment	
srpf_y_ss_dfs	Child: School risk and protective factors (school disengagement)	Cultural and Social Environment	
fes_p_ss_fc_pr	Caregiver: Family conflict	Family Environment	
psb_p_ss_mean	Caregiver: Prosocial behaviour	Family Environment	
pmq_y_ss_mean	Child: Parent monitoring	Family Environment	

fes_y_ss_fc_pr	Child: Family conflict	Family Environment	
psb_y_ss_mean	Child: Prosocial behaviour	Family Environment	
crpbi_y_ss_parent	Child: acceptance by parent	Family Environment	
crpbi_y_ss_caregiver	Child: acceptance by secondary caregiver	Family Environment	
sds_p_ss_total	Caregiver: Sleep Disturbance Scale (total scale)	Physical Health	
pds_p_ss_female_category, pds_p_ss_male_category	Caregiver: Pubertal development scale	Physical Health	Female and male pubertal development scores were originally separated.
demo_comb_income_v2	Household income	Sociodemographic status	
parent1_edu, parent2_edu	Highest education of parents	Sociodemographic status	A higher education between two caregivers was extracted as the highest education of the household/between caregivers

Table S5. Statistical models used for association tests between MDD/DS of depression and brain structural measures.

Type	Measures	Number of variables	Covariates (MDD/Depressive symptoms)
General brain measures	Mean whole-brain cortical thickness	1 unilateral	age+age2+sex+motion(fsqc_qu_motion)+site+race/ethnicity+recent social deprivation
	Total whole-brain surface area	1 unilateral	
	Mean whole-brain sulcal depth	1 unilateral	
	Total whole-brain volume	1 unilateral	
	Global total white matter fractional anisotropy	1 unilateral	
	Global total white matter mean diffusivity	1 unilateral	
Regional brain measures	Cortical thickness	34 bilateral	age+age2+sex+motion(fsqc_qu_motion)+site+race/ethnicity+recent social deprivation+intracranial volume
	Cortical surface area	34 bilateral	
	Cortical sulcal depth	34 bilateral	
	Cortical volume	34 bilateral	
	White matter fractional anisotropy	14 bilateral	age+age2+sex+motion(fsqc_qu_motion)+site+race/ethnicity+recent social deprivation
	White matter mean diffusivity	14 bilateral	

Figure S1. Density plot of global measures of cortical surface area, thickness, volume and sulcal depth. The x-axis represents the standardised scores. The y-axis represents distribution density. For illustration purpose, the plots were generated using the 'ggplot' function in R package 'ggplot2', with a smoothing adjustment of 2 and an illustration alpha (for transparency) of 0.1.

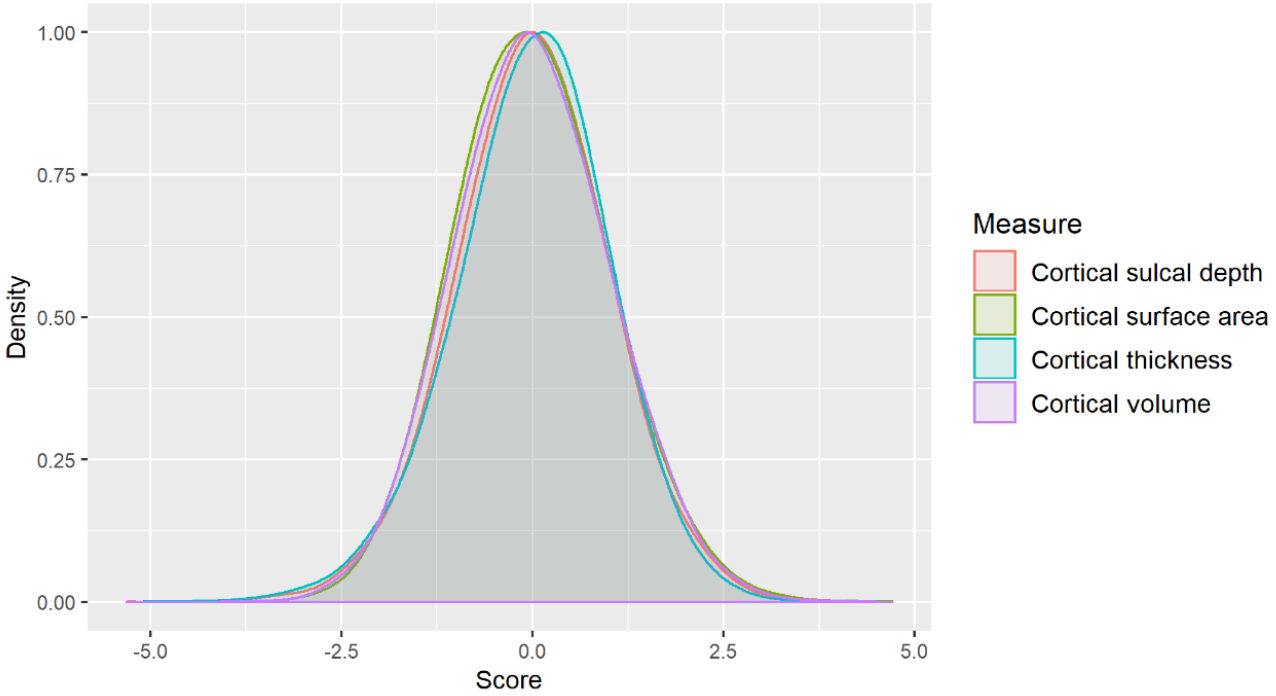


Figure S2. Sample sizes after each quality check (QC) step. For replication purpose, here in this chart we presented the field names used in the ABCD cohort. A general description of the fields used can be found in the Supplementary Methods. Detailed descriptions of each field can be found in the ABCD data dictionary (MRI raw data QC: https://nda.nih.gov/data_structure.html?short_name=mriqcrp102 and https://nda.nih.gov/data_structure.html?short_name=mriqcrp202; Freesurfer QC – cortical measures: https://nda.nih.gov/data_structure.html?short_name=freesqc01; dMRI post processing: https://nda.nih.gov/data_structure.html?short_name=dmriqc01) and in the quality check documentation (http://dx.doi.org/10.15154/1503209).

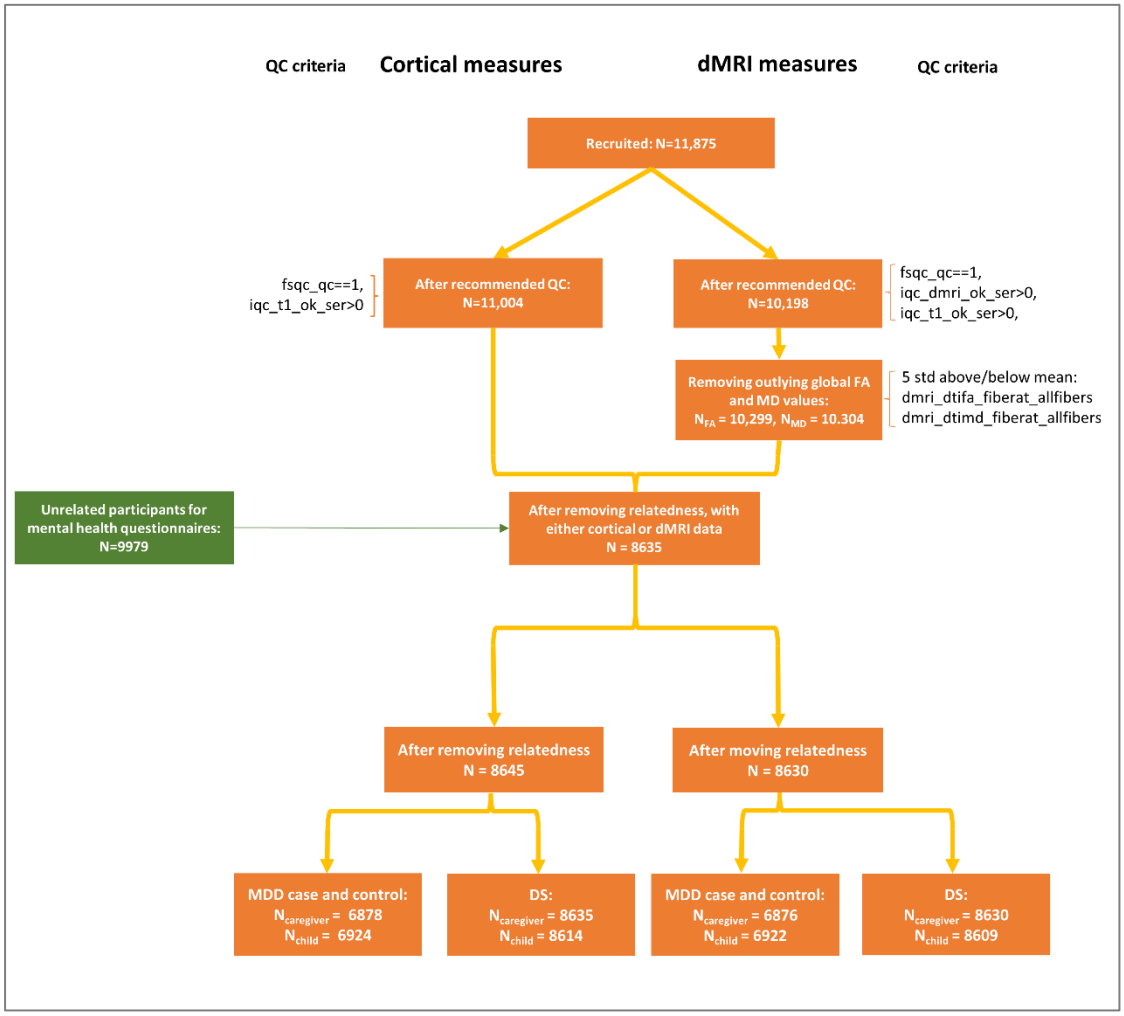


Figure S3. (a) Density plot of global measures of white matter fractional anisotropy (FA) and mean diffusivity (MD). The x-axis represents the standardised scores. The y-axis represents distribution density. In the left panel, data after removing those with poor data quality was used. As there were extreme values causing the distributions heavily skewed, we then removed participants with global values 5 standard deviations away from mean, which results in the distribution maps in the right panel. For illustration purpose, the plots were generated using the 'ggplot' function in R package 'ggplot2', with a smoothing adjustment of 2 and an illustration alpha (for transparency) of 0.1. (b) Comparison between the standardised values for general white matter microstructural measures that were kept or removed due to post-processing quality check (QC). Red dots represent values outside of 95% confidence interval.

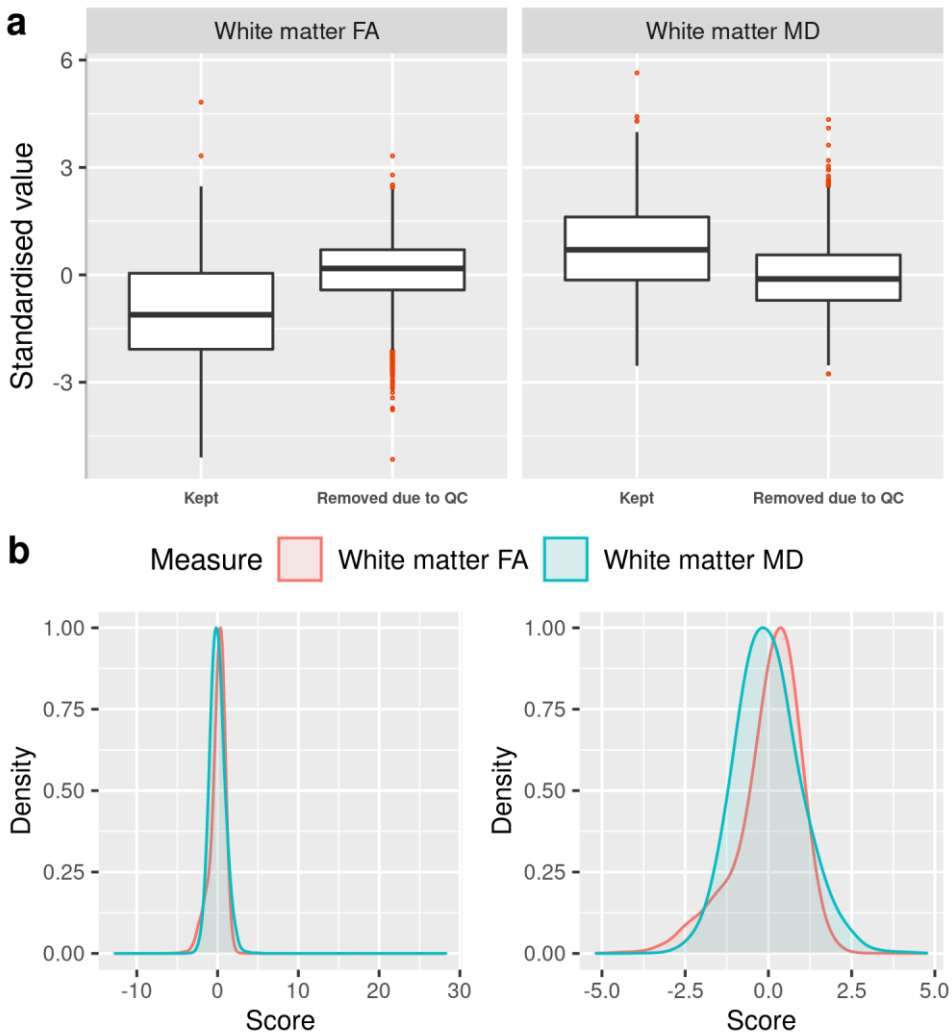


Figure S4. Histogram of DS reported by caregivers and children, absolute discrepancy and discrepancy between caregiver and child reports. The x-axis represents DS in panel a. In panels b and c the x-axes represent the absolute discrepancy and discrepancy between caregiver and child (caregiver - child) reports respectively. The y-axes represent distribution density.

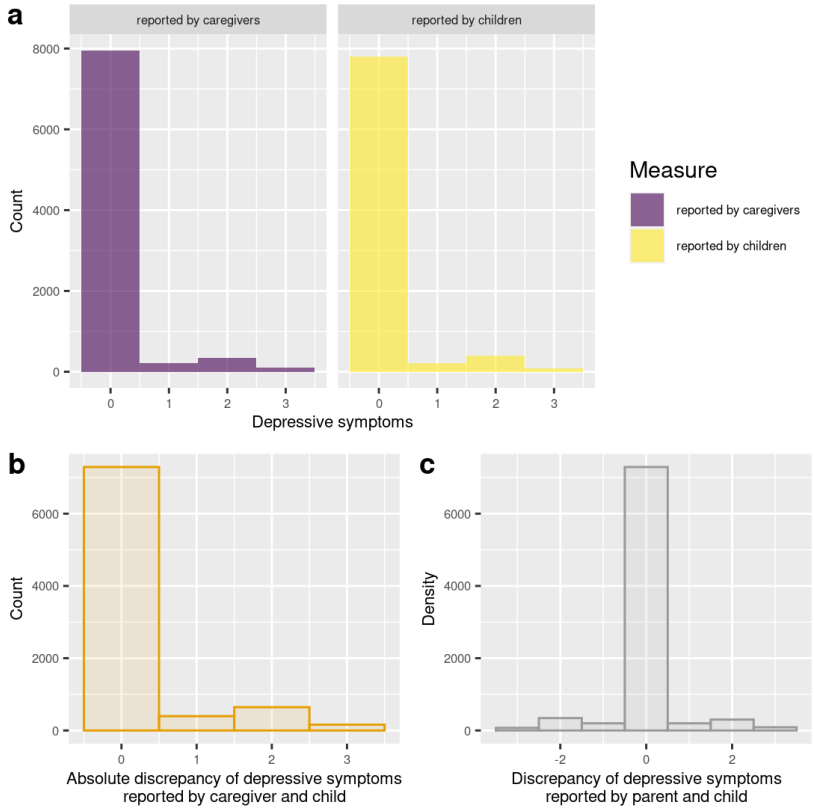


Figure S5. Associations between MDD, depressive symptoms and general measures of brain structures, controlling for ASR scale for severity of depression in caregivers. X-axes represent standardised effect sizes with error bars showing 95% confidence interval, and y-axes represent each general measure of brain structure. Panel a shows the results for MDD/depressive symptoms reported by caregivers on children, and panel b shows the results for symptoms reported by children themselves.

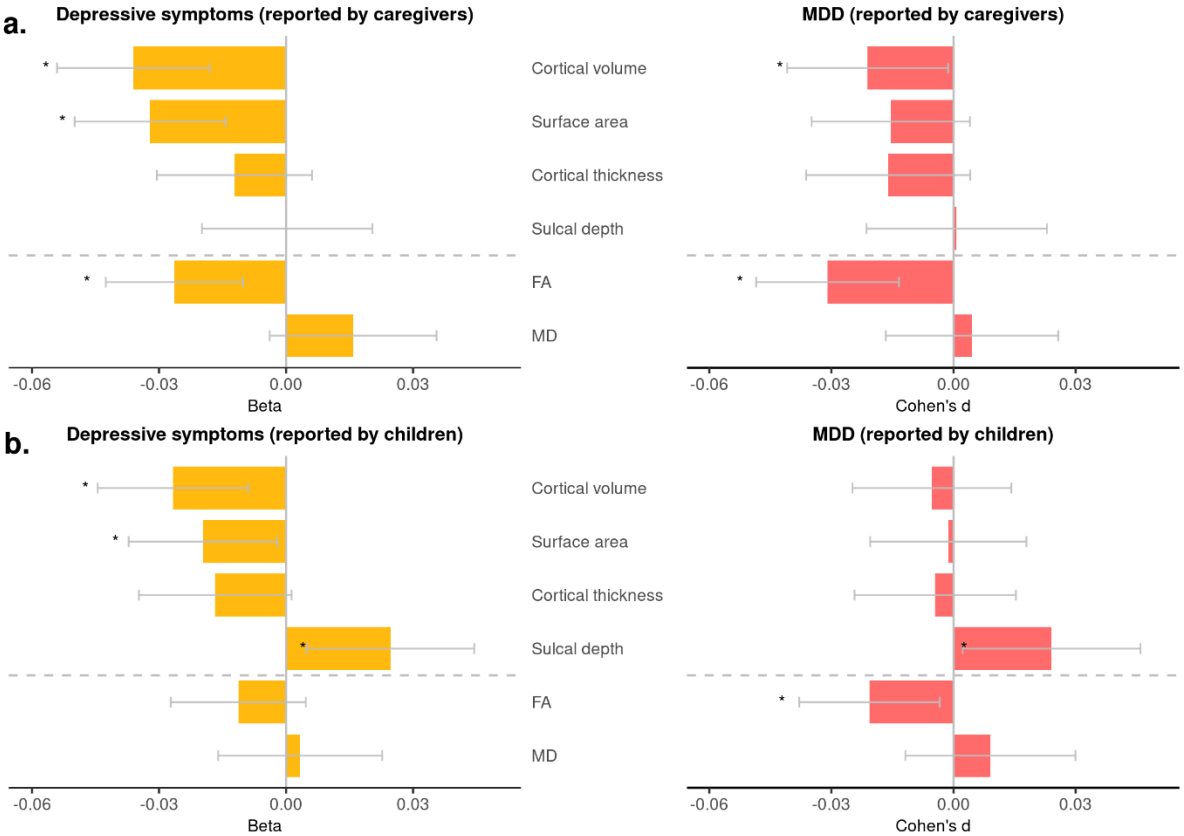


Figure S6. P-value plots for associations between depressive symptoms and measures for single brain regions, controlling for ASR scale for severity of depression in caregivers. X axes represent measures for brain structural measures, and y axes represent -log10 transformed p-values. Panel (a) shows the results for depressive symptoms reported by caregivers on children, and panel (b) shows the results for symptoms reported children themselves. Solid dots represent variables associated with depressive symptoms after FDR-correction. For clarity, threshold for significance after FDR-correction is shown as the pink dashed lines.

Figure S7. P-value plots for associations between MDD diagnosis and measures for single brain regions, controlling for ASR scale for severity of depression in caregivers. X axes represent measures for brain structural measures, and y axes represent -log10 transformed p-values. Panel (a) shows the results for depressive symptoms reported by caregivers on children, and panel (b) shows the results for symptoms reported by children themselves. Solid dots represent variables associated with depressive symptoms after FDR-correction. For clarity, threshold for significance after FDR-correction is shown as the pink dashed lines.

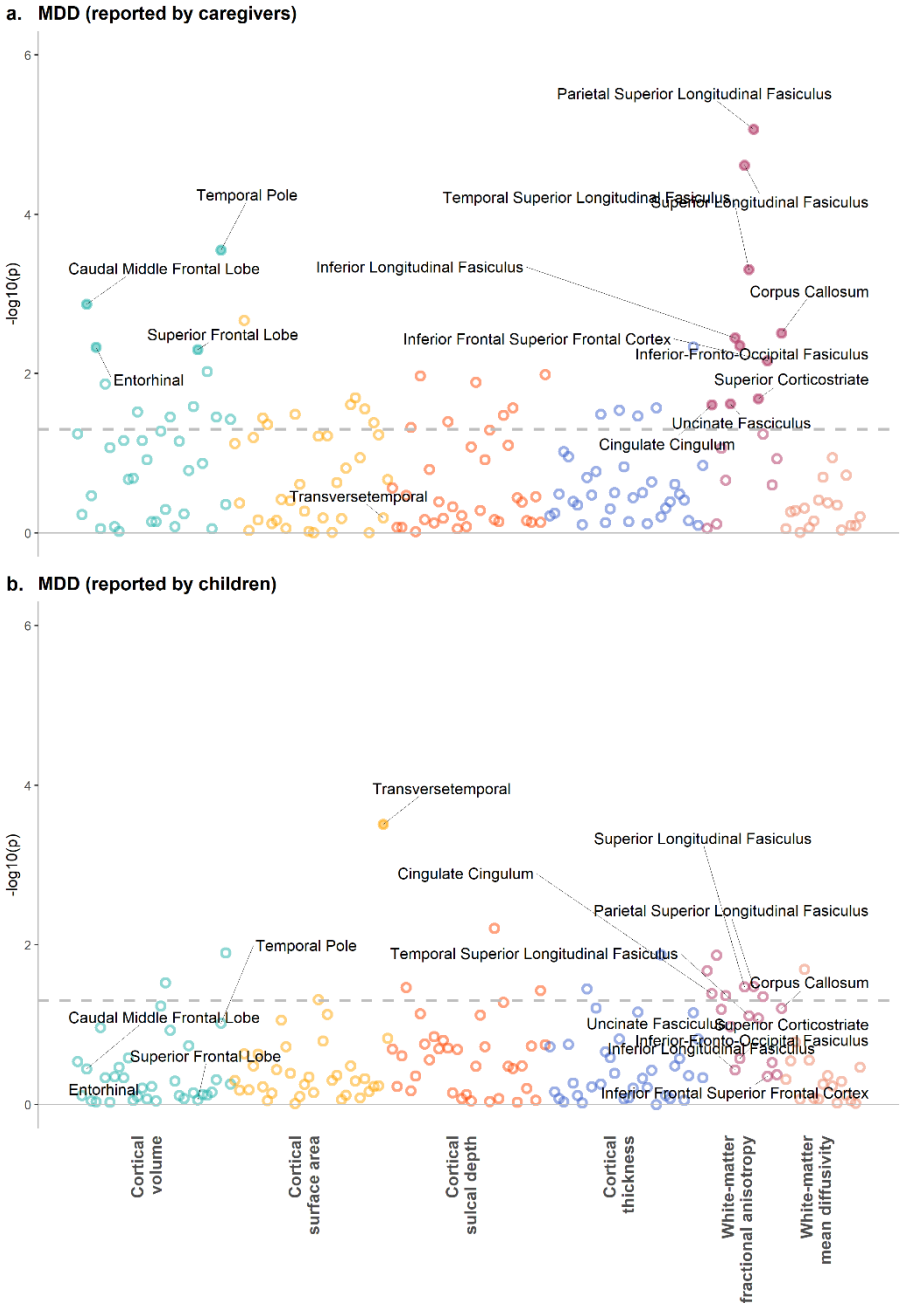


Figure S8. Correlations of effect sizes and p-values between the main analysis and the analysis controlling for ASR scale for severity of depression in caregivers (Supplementary Methods). X-axes represent statistics of the main model and the Y-axes represent statistics of controlling for severity of depression in caregivers. The left panel shows the correlation of standardised effect sizes (regression coefficient/Cohen's d depending on which independent variable was used – MDD/depressive symptoms), and the right panel shows the correlation of p-values. In the right panel, the grey dashed line shows the threshold of nominal significance ($p < 0.05$).

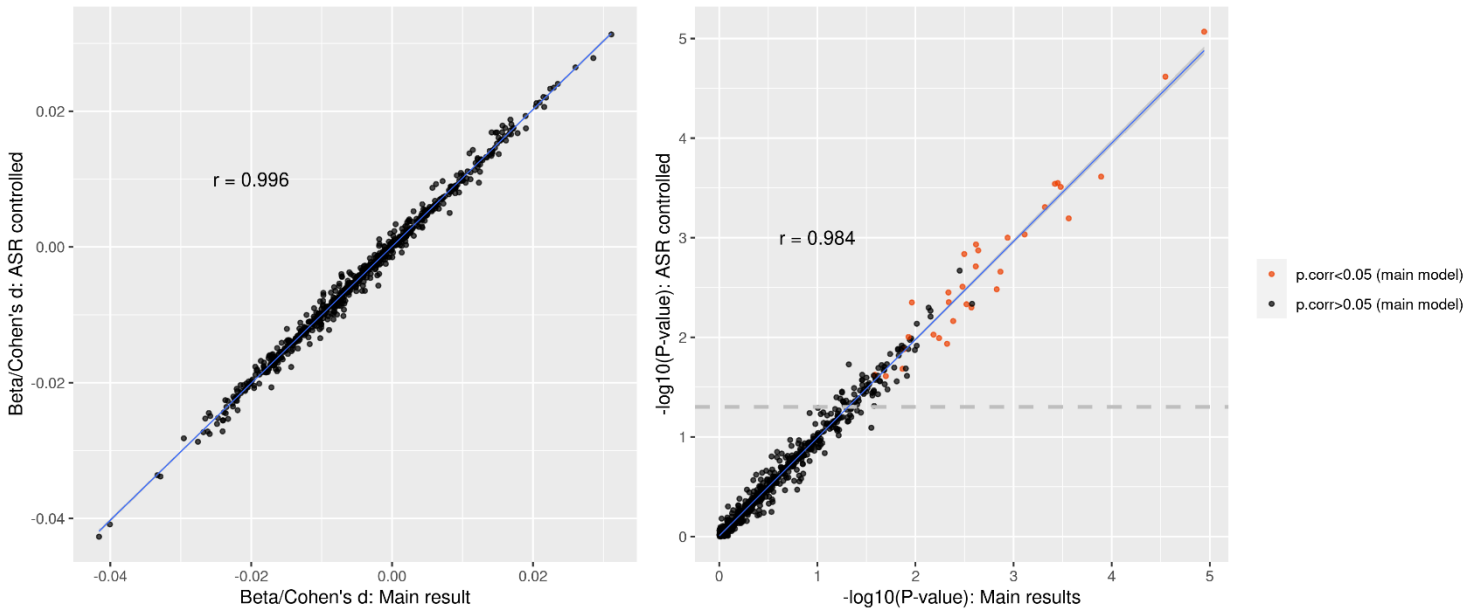


Figure S9. Leave-one-out analysis testing the association between general brain measures and depressive symptoms reported by caregivers. The x-axes represent standardised effect sizes. The y-axes represent the analysis leaving the given site out. The green dots represent the effect sizes by using the whole sample. The error bars represent 95% confidence intervals.

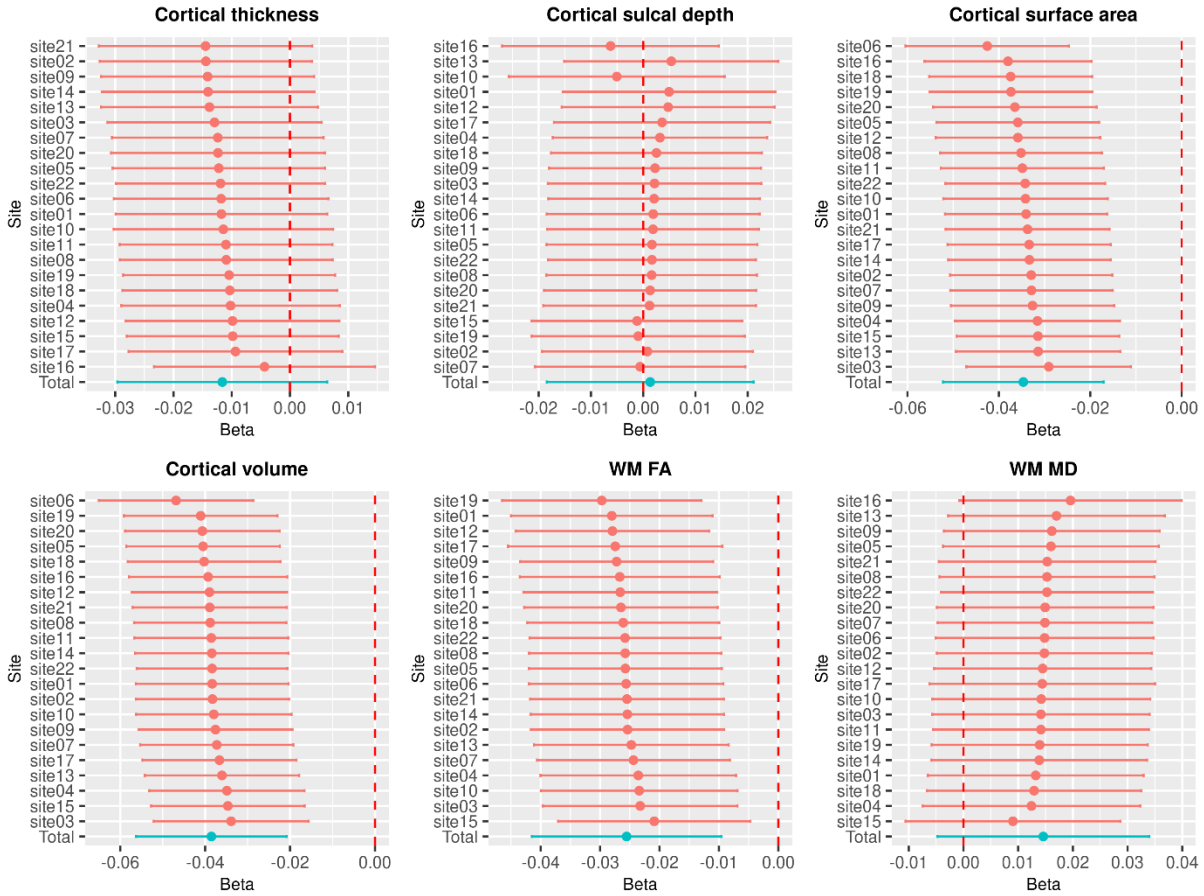


Figure S10. Leave-one-out analysis testing the association between general brain measures and depressive symptoms reported by children. The x-axes represent standardised effect sizes. The y-axes represent the analysis leaving the given site out. The green dots represent the effect sizes by using the whole sample. The error bars represent 95% confidence intervals.

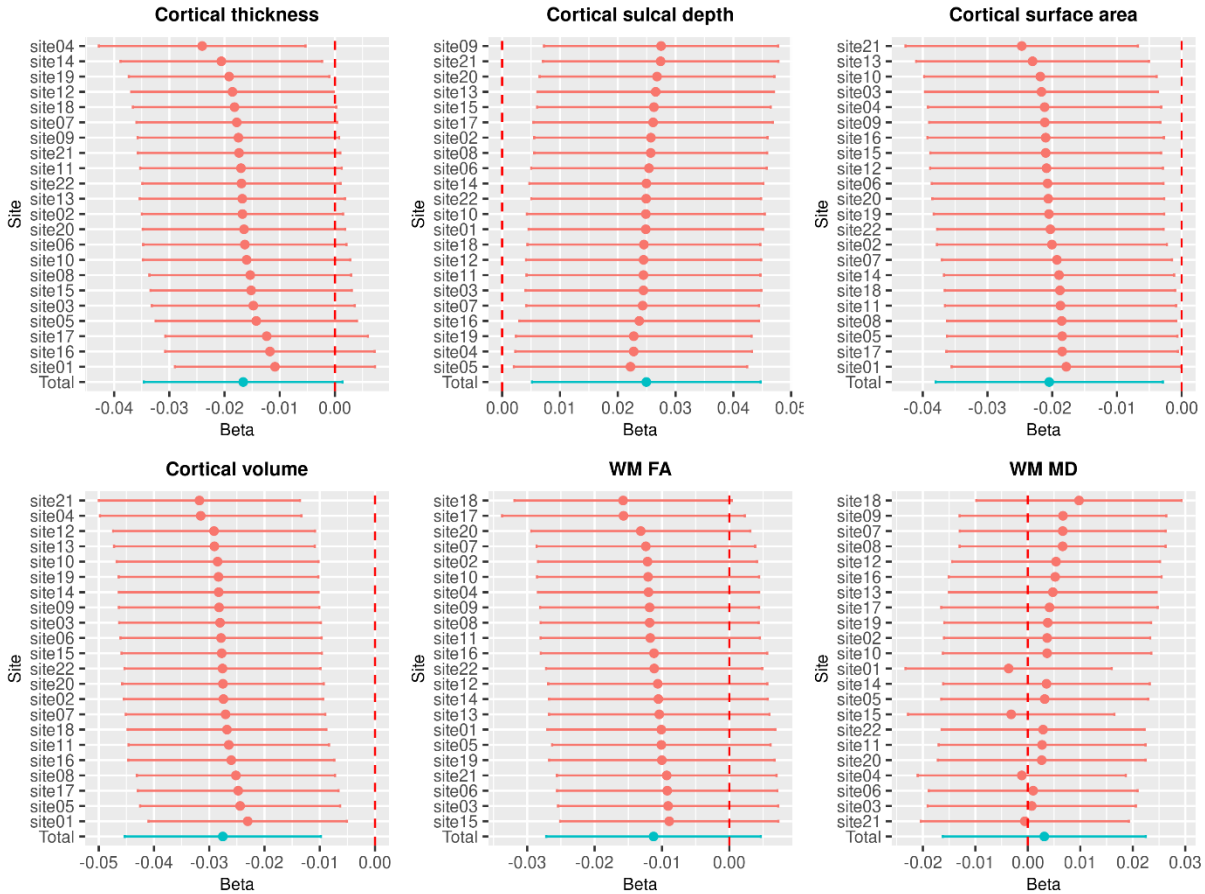


Figure S11. Leave-one-out analysis testing the association between general brain measures and MDD diagnosis reported by caregivers. The x-axes represent standardised effect sizes. The y-axes represent the analysis leaving the given site out. The green dots represent the effect sizes by using the whole sample. The error bars represent 95% confidence intervals.

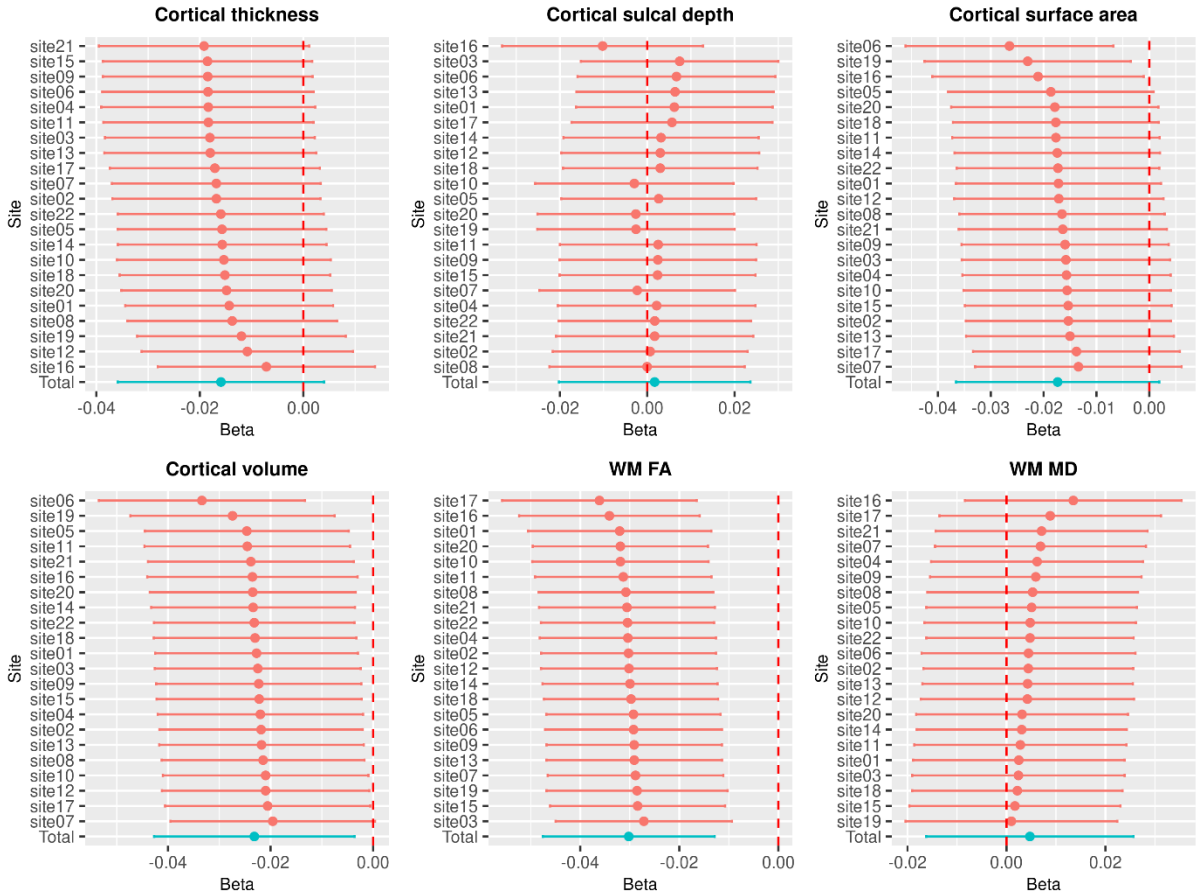


Figure S12. Leave-one-out analysis testing the association between general brain measures and MDD diagnosis reported by children. The x-axes represent standardised effect sizes. The y-axes represent the analysis leaving the given site out. The green dots represent the effect sizes by using the whole sample. The error bars represent 95% confidence intervals.

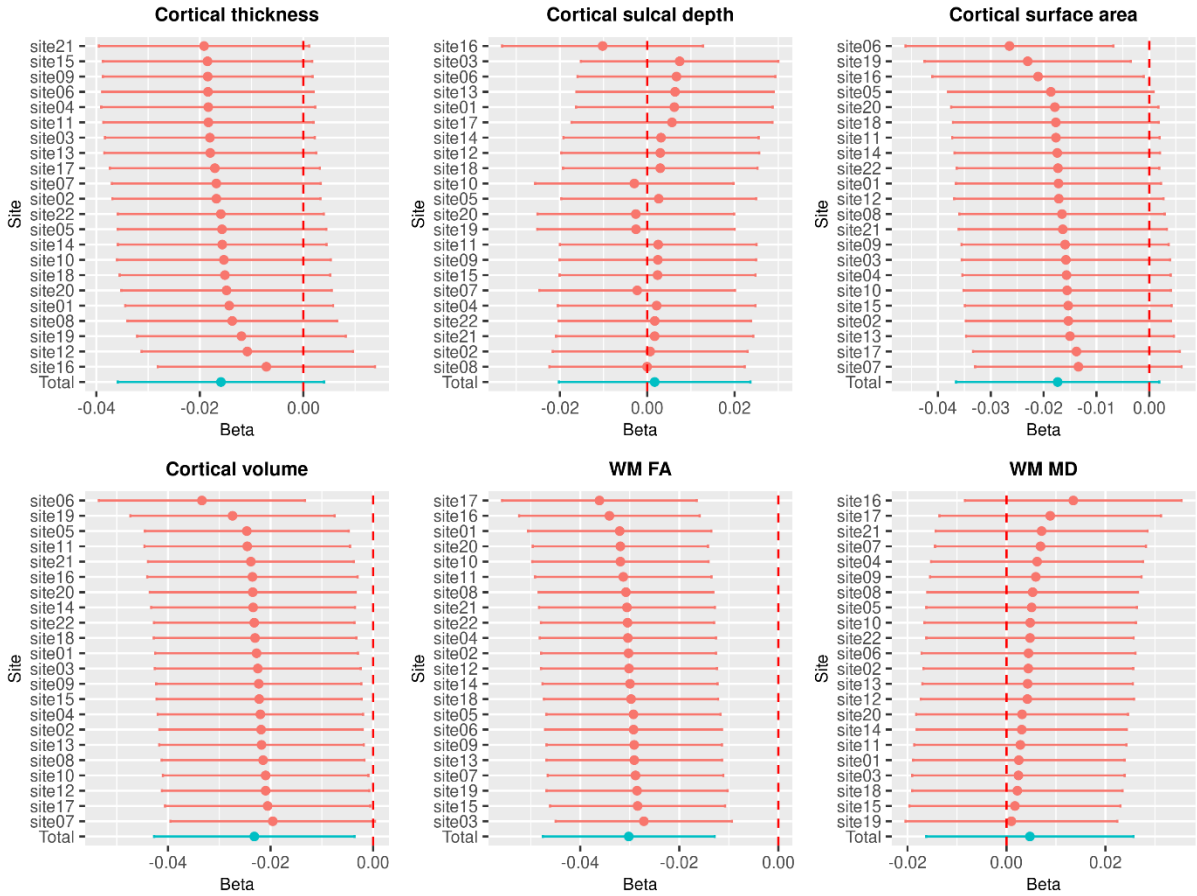


Figure S13. Associations between MDD, depressive symptoms and general measures of brain structures, controlling for MRI manufacturer. X-axes represent standardised effect sizes with error bars represent 95% confidence intervals, and y-axes represent each general measure of brain structure. Panel a shows the results for MDD/depressive symptoms reported by caregivers on children, and panel b shows the results for symptoms reported by children themselves.

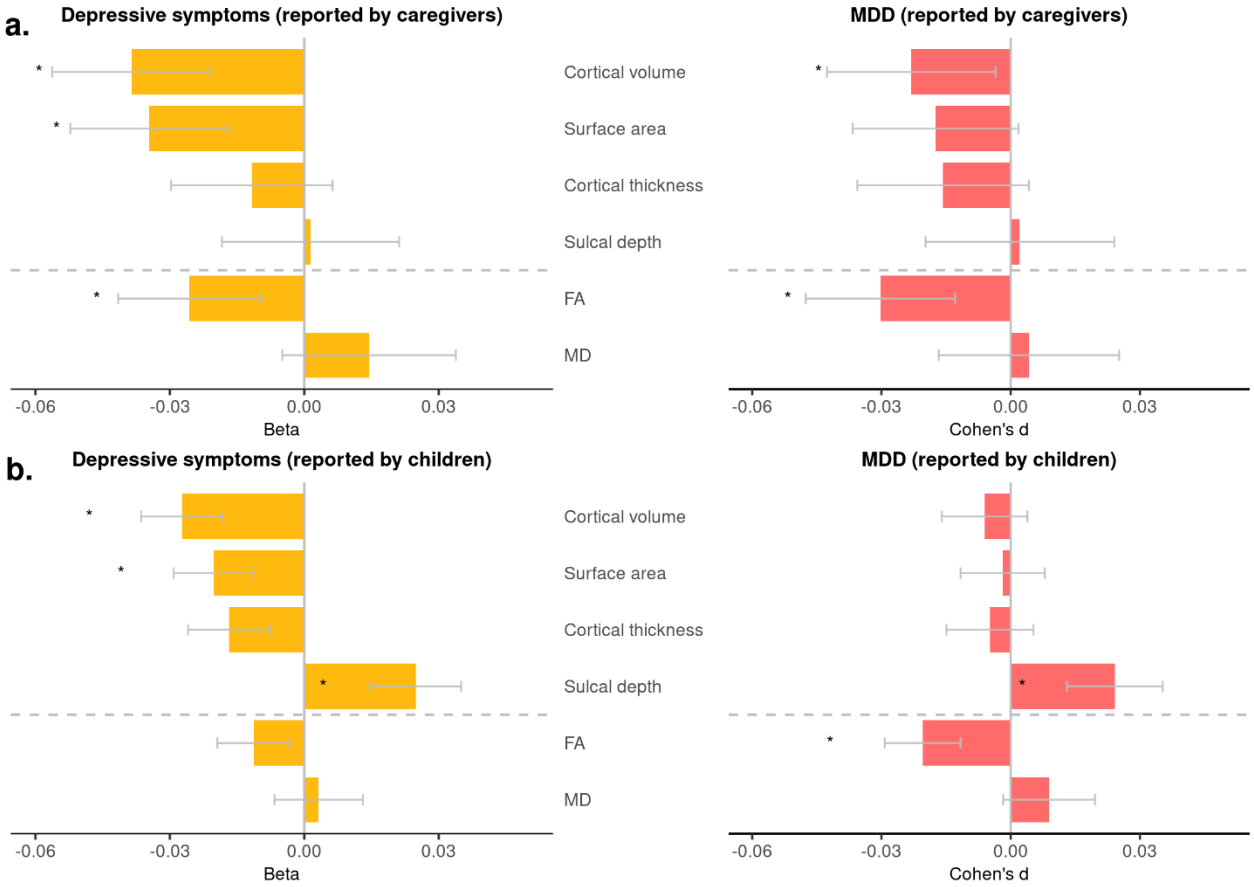


Figure S14. Correlations of effect sizes and p-values between the main model (not controlling for MRI manufacturer) and the secondary model controlling for MRI manufacturer. X-axes represent statistics of the main model and the Y-axes represent statistics of the secondary model. The left panel shows the correlation of standardised effect sizes (regression coefficient/Cohen's d depending on which independent variable was used – MDD/depressive symptoms), and the right panel shows the correlation of p-values. In the right panel, the grey dashed line shows the threshold of nominal significance ($p < 0.05$).

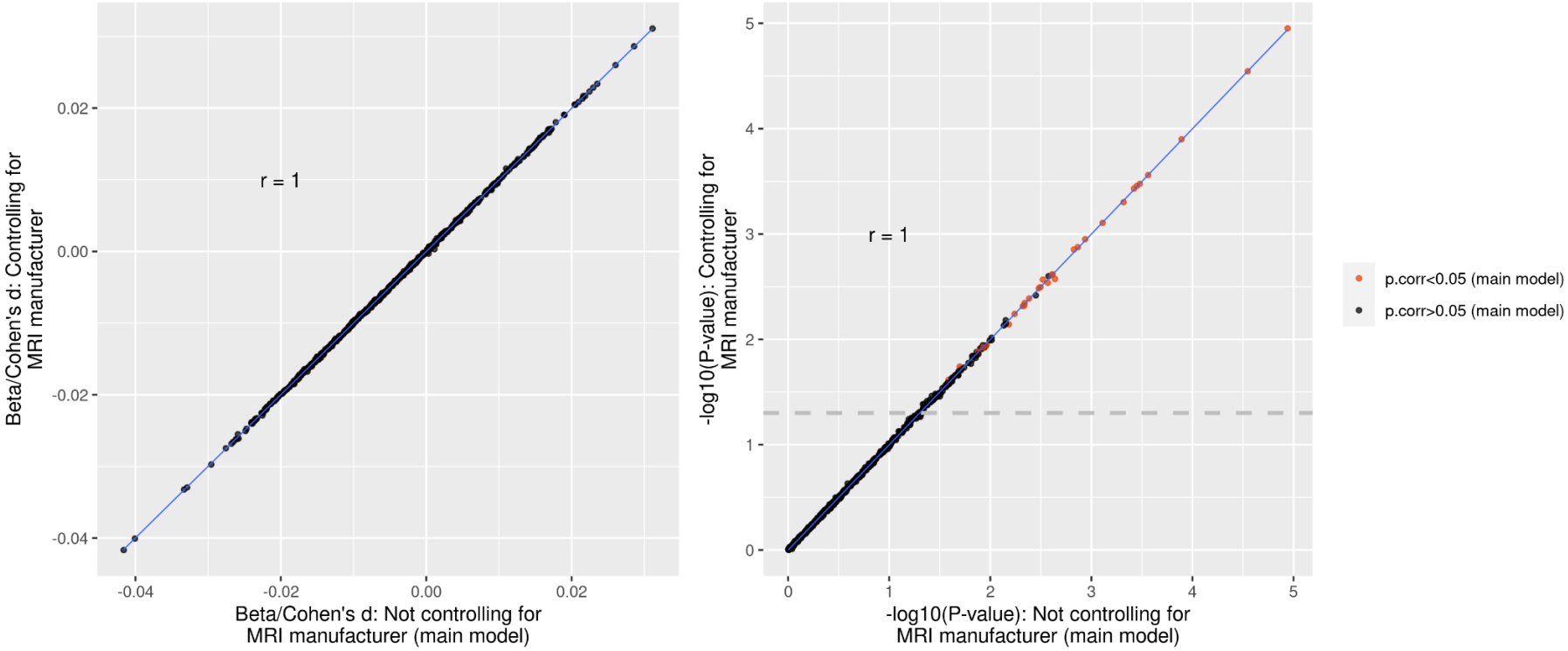

r = 1
Beta/Cohen's d: Controlling for MRI manufacturer
Beta/Cohen's d: Not controlling for MRI manufacturer (main model)
0.02
0.00
-0.02
-0.04
-0.04
-0.02
0.00
0.02
r = 1
-log10(P-value): Controlling for MRI manufacturer
-log10(P-value): Not controlling for MRI manufacturer (main model)
0
1
2
3
4
5
p.corr<0.05 (main model)
p.corr>0.05 (main model)

Figure S15. Associations between MDD, depressive symptoms and general measures of brain structures, controlling for medication. X-axes represent standardised effect sizes with error bars representing 95% confidence intervals, and y-axes represent each general measure of brain structure. Panel a shows the results for MDD/depressive symptoms reported by caregivers on children, and panel b shows the results for symptoms reported by children themselves.

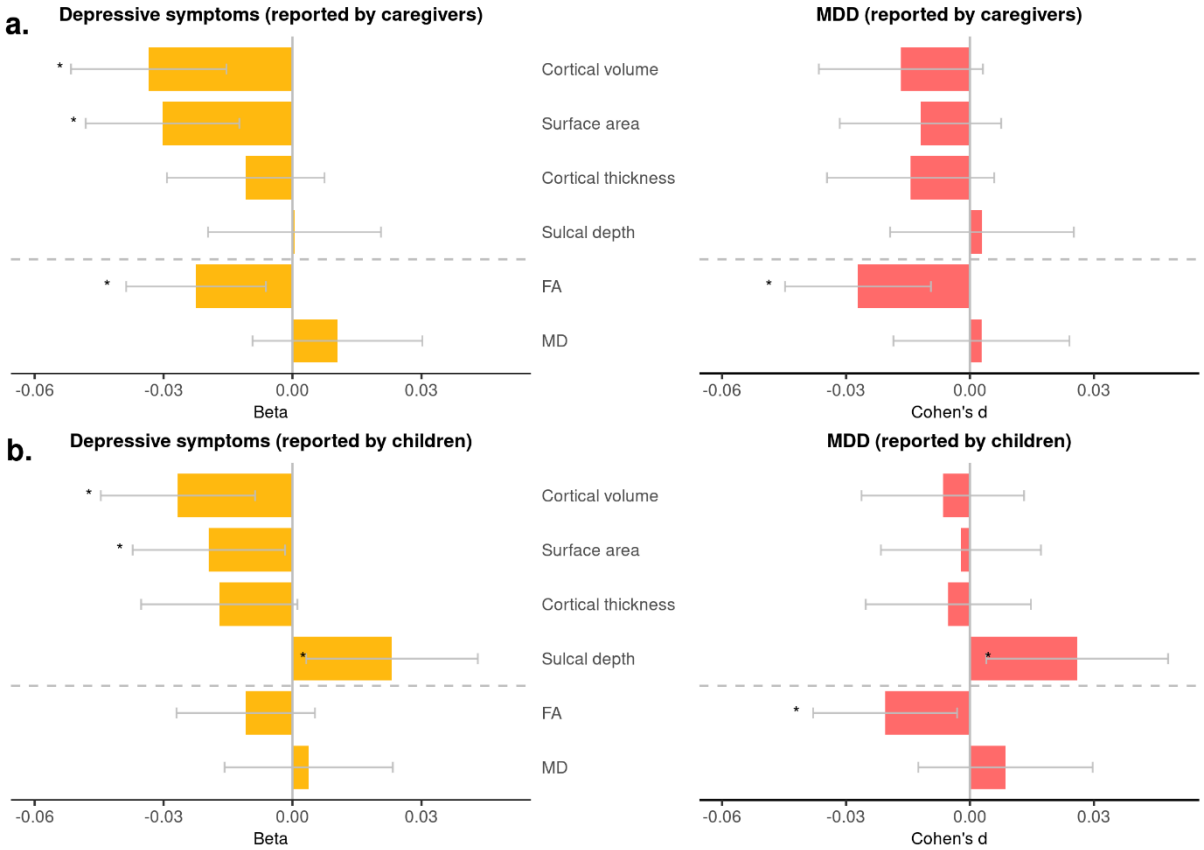


Figure S16. Correlations of effect sizes and p-values between the main model (not controlling for medication) and the secondary model controlling for medication. X-axes represent statistics of the main model and the Y-axes represent statistics of the secondary model. The left panel shows the correlation of standardised effect sizes (regression coefficient/Cohen's d depending on which independent variable was used – MDD/depressive symptoms), and the right panel shows the correlation of p-values. In the right panel, the grey dashed line shows the threshold of nominal significance ($p < 0.05$).

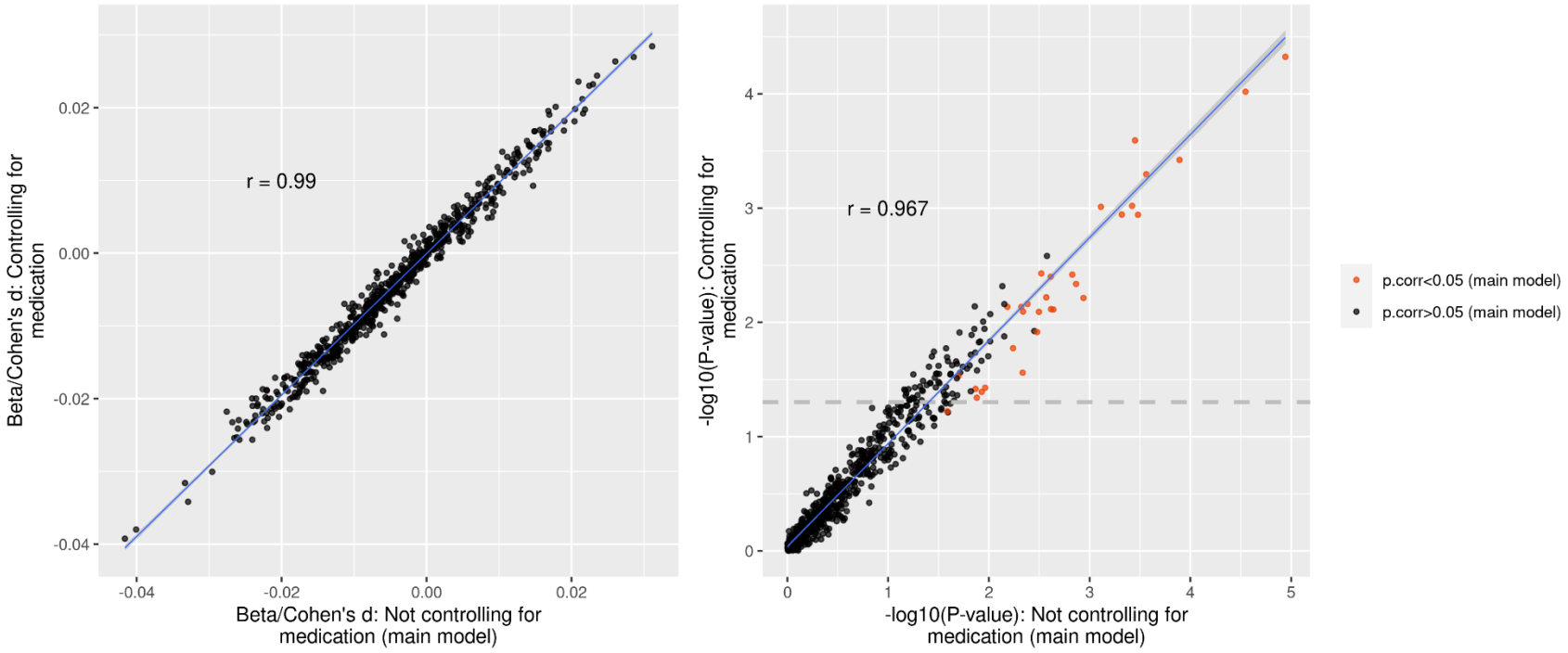


Figure S17. Associations between the average rating between caregiver and child reports of depressive symptoms and general measures of brain structures. X-axes represent standardised effect sizes, and y-axes represent each general measure of brain structure. Error bars represent 95% confidence intervals. Significant associations are highlighted with an asterisk.

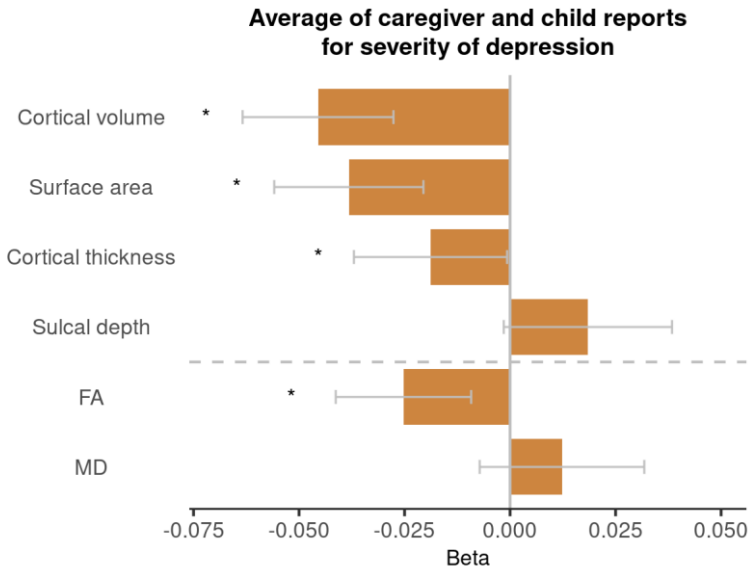


Figure S18. P-value plots for associations between average of caregiver and child report of severity of MDD and measures for single brain regions. X axes represent measures for brain structural measures, and y axes represent -log10 transformed p-values. Solid dots represent significant associations after FDR-correction. Pink dashed line represents nominally significance threshold. Those regions with the label '(+)' were new associations found with average DS but not with parent-reported DS. Regions with the label '(-)' were associations that were found with parent-reported DS but not with average DS.

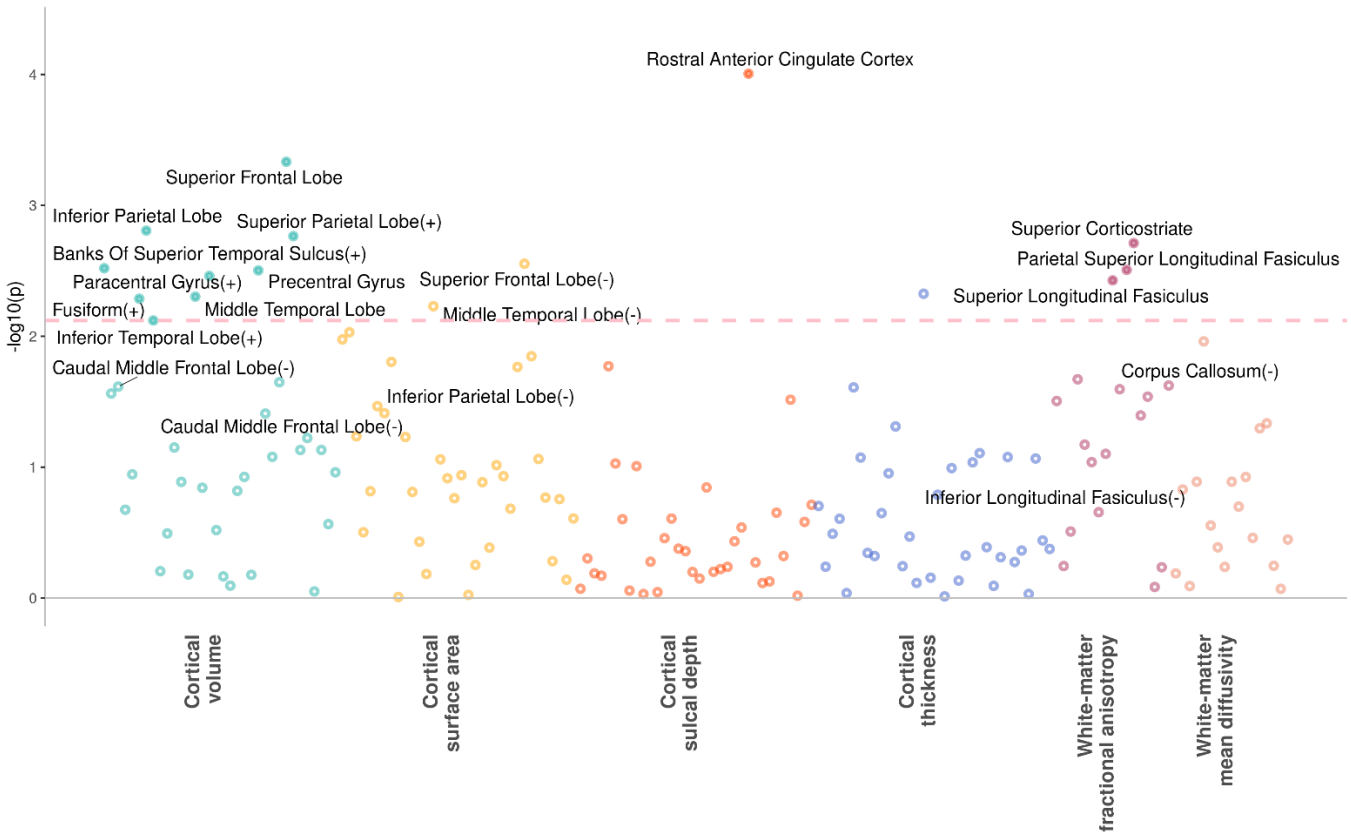

Rostral Anterior Cingulate Cortex
Superior Frontal Lobe
Inferior Parietal Lobe
Superior Parietal Lobe(+)
Banks Of Superior Temporal Sulcus(+)
Paracentral Gyrus(+)
Precentral Gyrus
Superior Frontal Lobe(-)
Fusiform(+)
Middle Temporal Lobe
Middle Temporal Lobe(-)
Inferior Temporal Lobe(+)
Caudal Middle Frontal Lobe(-)
Inferior Parietal Lobe(-)
Caudal Middle Frontal Lobe(-)
Superior Corticostriate
Parietal Superior Longitudinal Fasiculus
Superior Longitudinal Fasiculus
Corpus Callosum(-)
Inferior Longitudinal Fasiculus(-)
-log10(p)
0
1
2
3
4
Cortical volume
Cortical surface area
Cortical sulcal depth
Cortical thickness
White-matter fractional anisotropy
White-matter mean diffusivity

Figure S19. Associations between average of child and caregiver reports of depressive symptoms and regional measures of brain structures. Regional brain measures that were found significantly associated with caregiver report of depressive symptoms were chosen for this analysis and thus multiple comparison correction was conducted withing the tested associations only. X-axes represent standardised effect sizes, and y-axes represent each regional brain structure. Error bars represent 95% confidence intervals. Significant associations are highlighted with an asterisk.

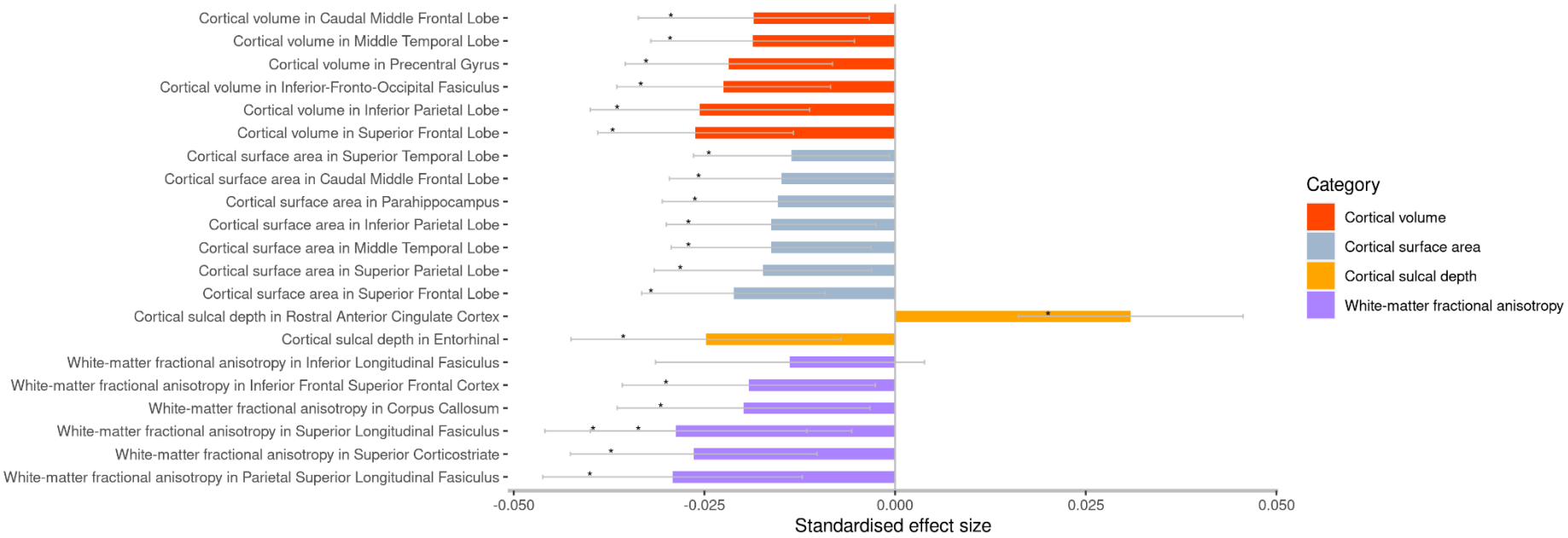


Figure S20. Reporter difference for each KSADS item. The X-axis represents the proportion of types of reporter comparisons (Caregiver > Child: caregiver=1 and child=0; Child > Caregiver: child=1 and caregiver=0; Caregiver = Child: caregiver=child=1 or caregiver=child=0). The y-axis represents each KSADS item used for deriving severity of MDD.

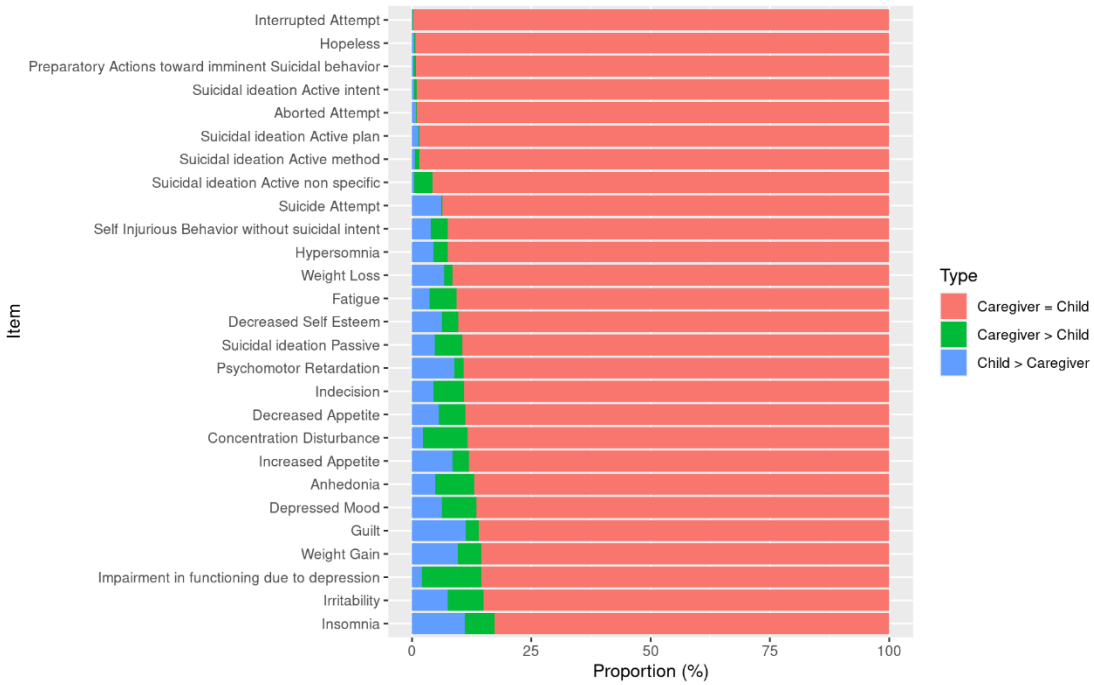


Figure S21. Correlations of effect sizes and p-values of using the subsample (N=3733) with post-processing QC data available and for those who passed QC in this subsample (N=3509). X-axes represent statistics of the main model and the Y-axes represent statistics of the secondary model. The left panel shows the correlation of standardised effect sizes (regression coefficient/Cohen's d depending on which independent variable was used – MDD/depressive symptoms), and the right panel shows the correlation of p-values. In the right panel, the grey dashed line shows the threshold of nominal significance ($p < 0.05$).

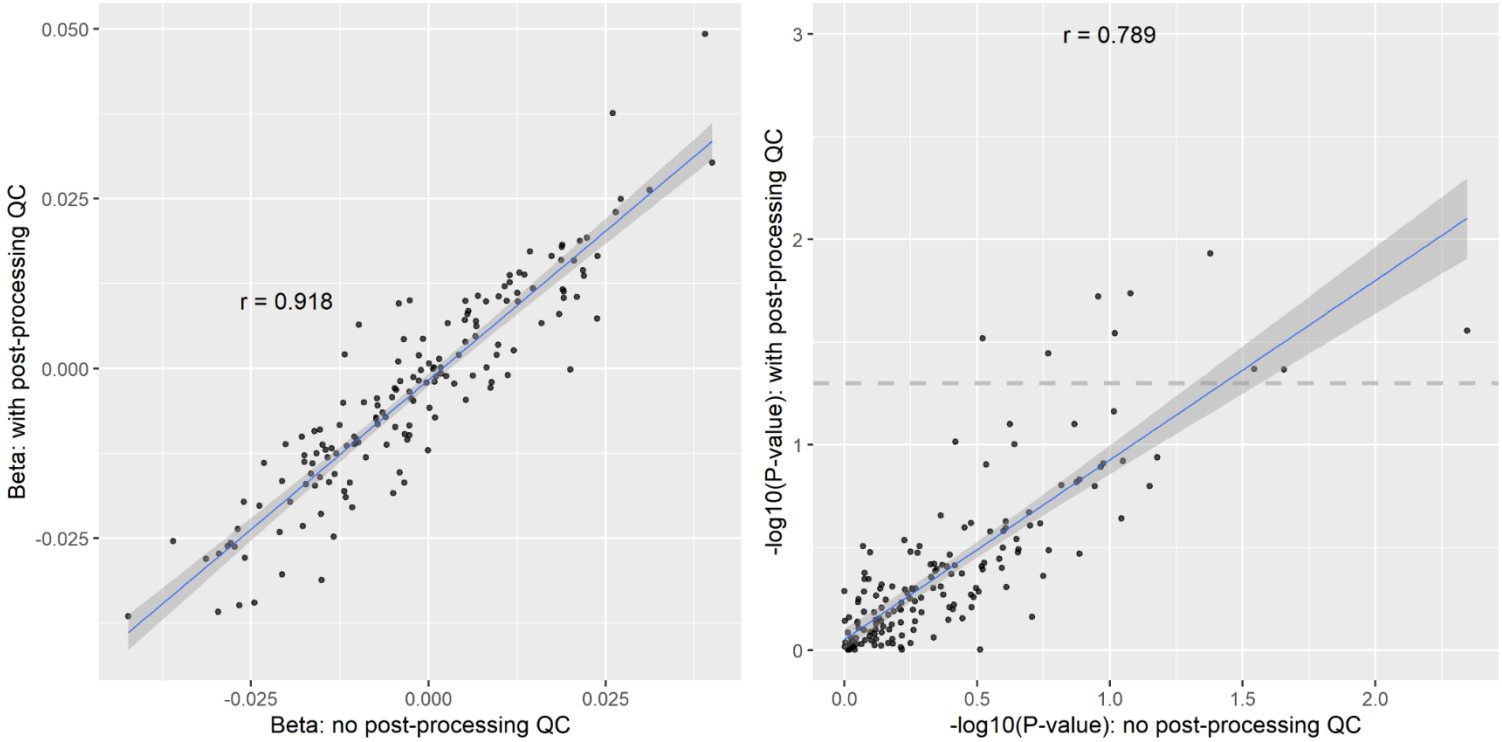


Figure S22. Associations between the absolute difference between caregiver and child reports of depressive symptoms and general measures of brain structures. X-axes represent standardised effect sizes, and y-axes represent each general measure of brain structure. Error bars represent 95% confidence intervals. Significant associations are highlighted with an asterisk.

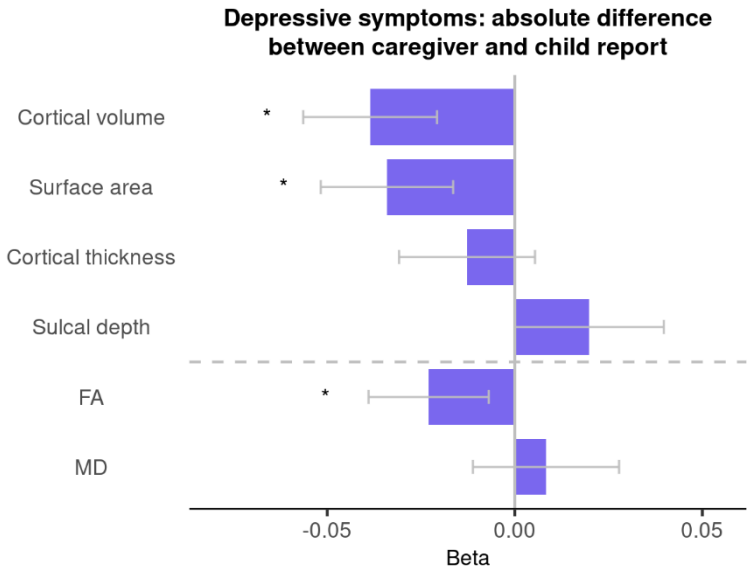


Figure S23. Associations between the absolute difference between child and caregiver reports of depressive symptoms regional measures of brain structures. Regional brain measures that were found significantly associated with caregiver report of depressive symptoms were chosen for this analysis. X-axes represent standardised effect sizes, and y-axes represent each regional brain structure. Error bars represent 95% confidence intervals. Significant associations are highlighted with an asterisk.

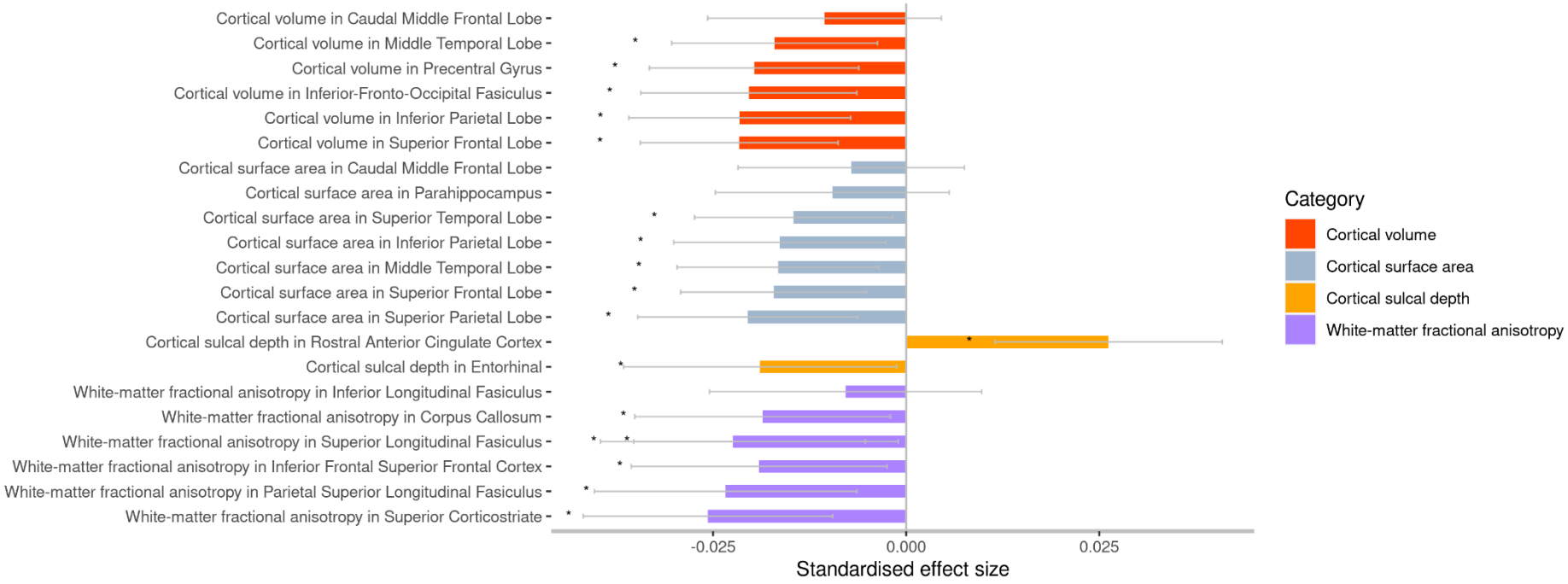


Figure S24. Distribution of CBCL-measured DS under each MDD severity category classified using KSADS (both KSADS and CBCL measures were reported by caregivers). X axis represents CBCL-measured DS. Each row's y axis represents distribution density. The vertical lines in each distribution represents the mean of the given distribution.

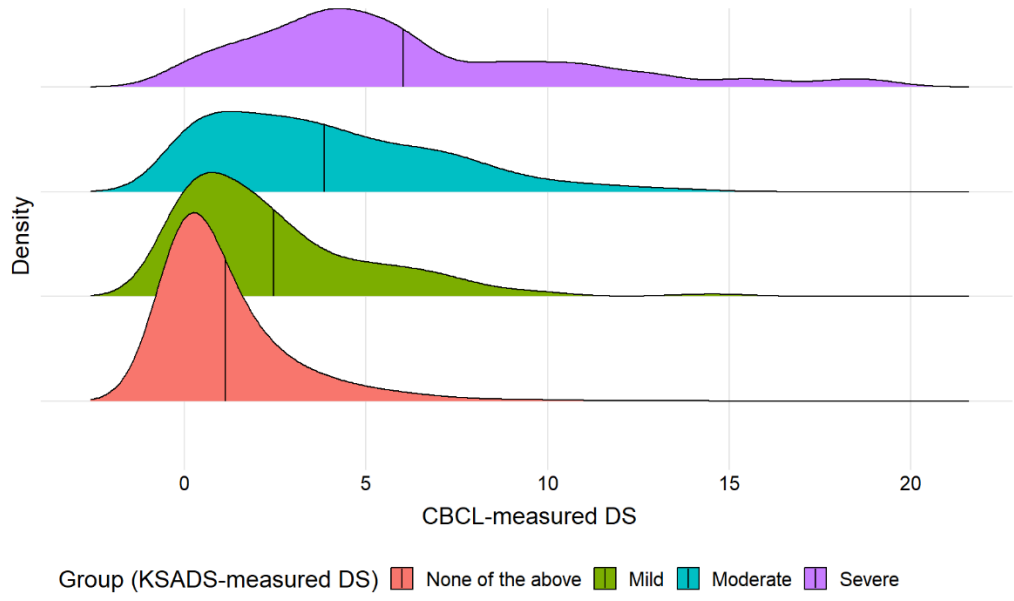


Figure S25. Associations between CBCL DSM-5-oriented score of depression and general measures of brain structures. X-axes represent standardised effect sizes, and y-axes represent each general measure of brain structure. Error bars represent 95% confidence intervals. Significant associations are highlighted with an asterisk.

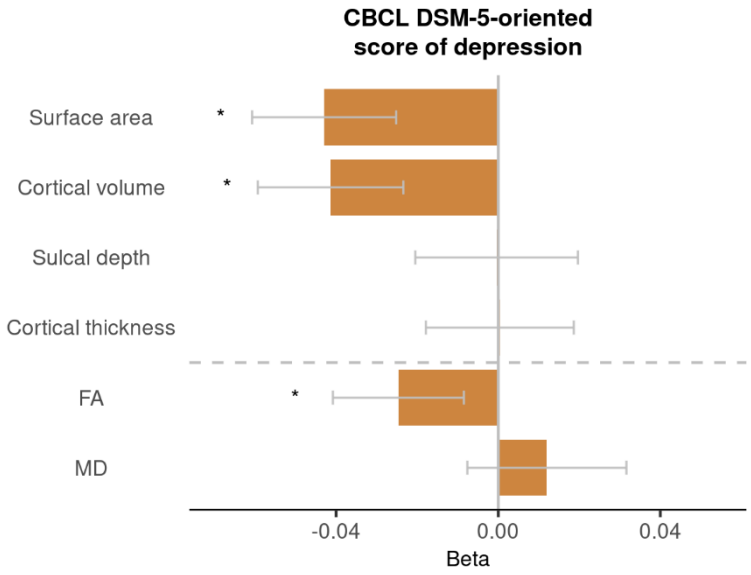


Figure S26. P-value plots for associations between depressive symptoms assessed by CBCL (the Child Behaviour Checklist, reported by caregivers) and measures for single brain regions. X axes represent measures for brain structural measures, and y axes represent -log10 transformed p-values. Panel (a) shows the results for depressive symptoms reported by caregivers on children, and panel (b) shows the results for symptoms reported by children themselves. Solid dots represent variables associated with depressive symptoms after FDR-correction. For clarity, threshold for significance after FDR-correction is shown as the pink dashed lines.

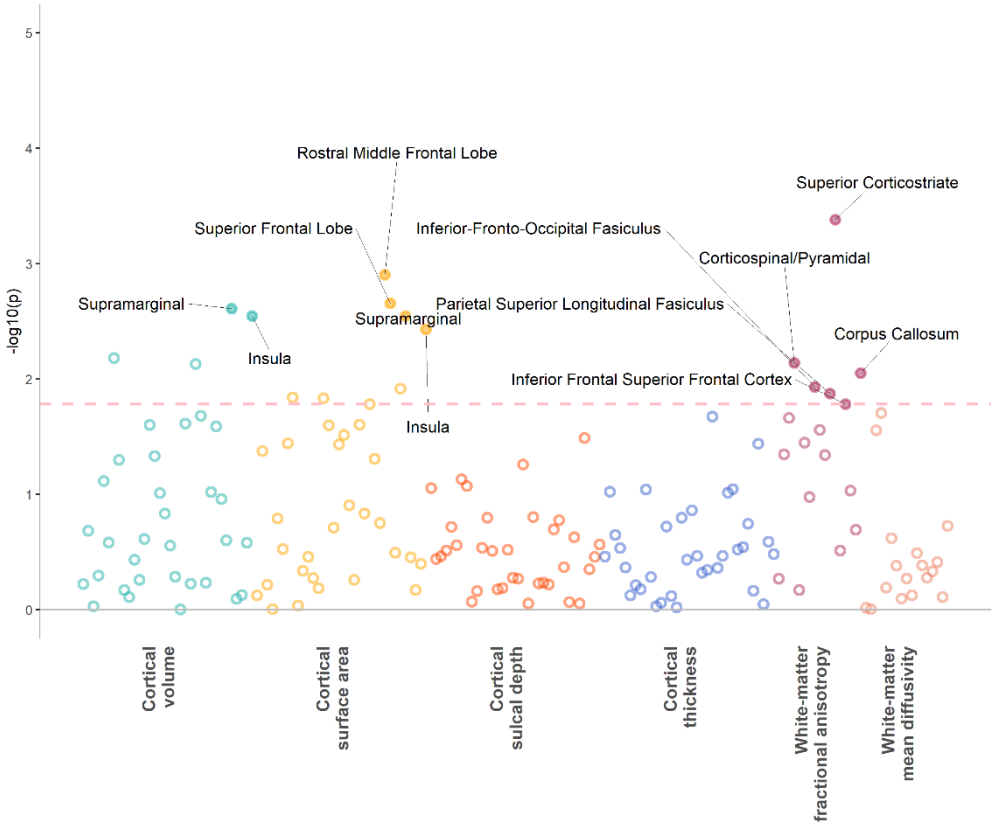


Figure S27. Associations between the caregiver report and general measures of brain structures after accounting for comorbidity of Bipolar I, Bipolar II, ADHD, Psychosis and Conduct disorder. X-axes represent standardised effect sizes, and y-axes represent each general measure of brain structure. Error bars represent 95% confidence intervals. Significant associations are highlighted with an asterisk.

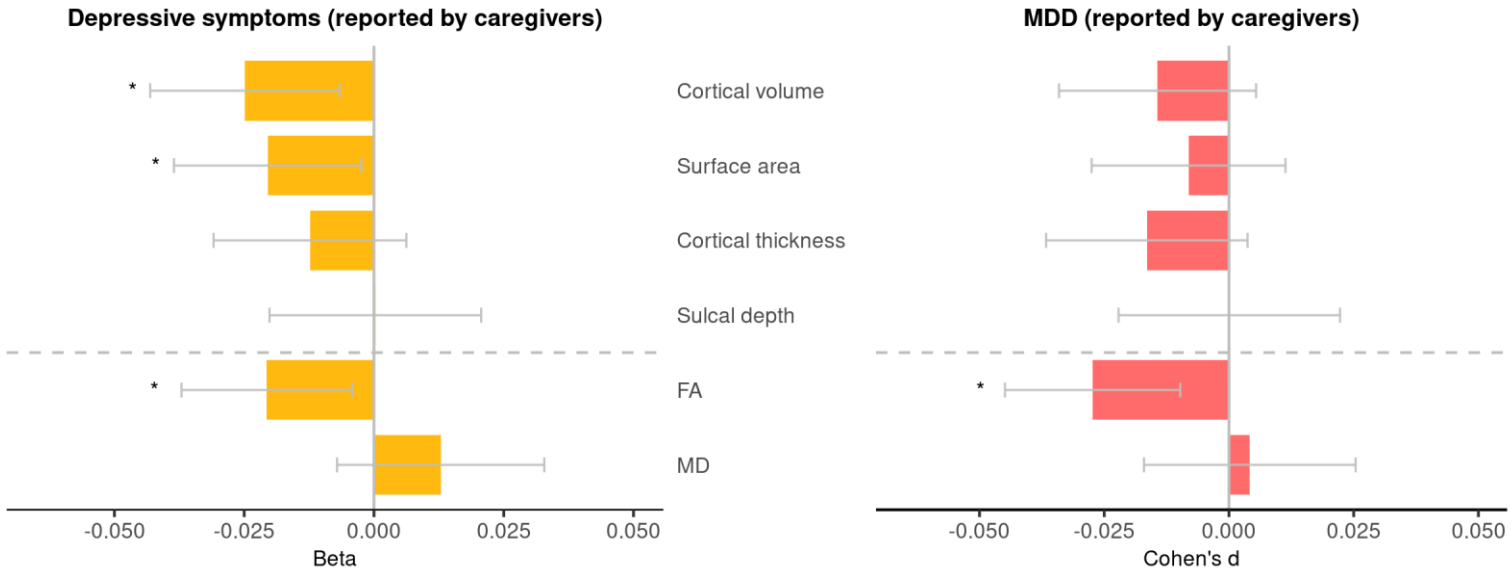
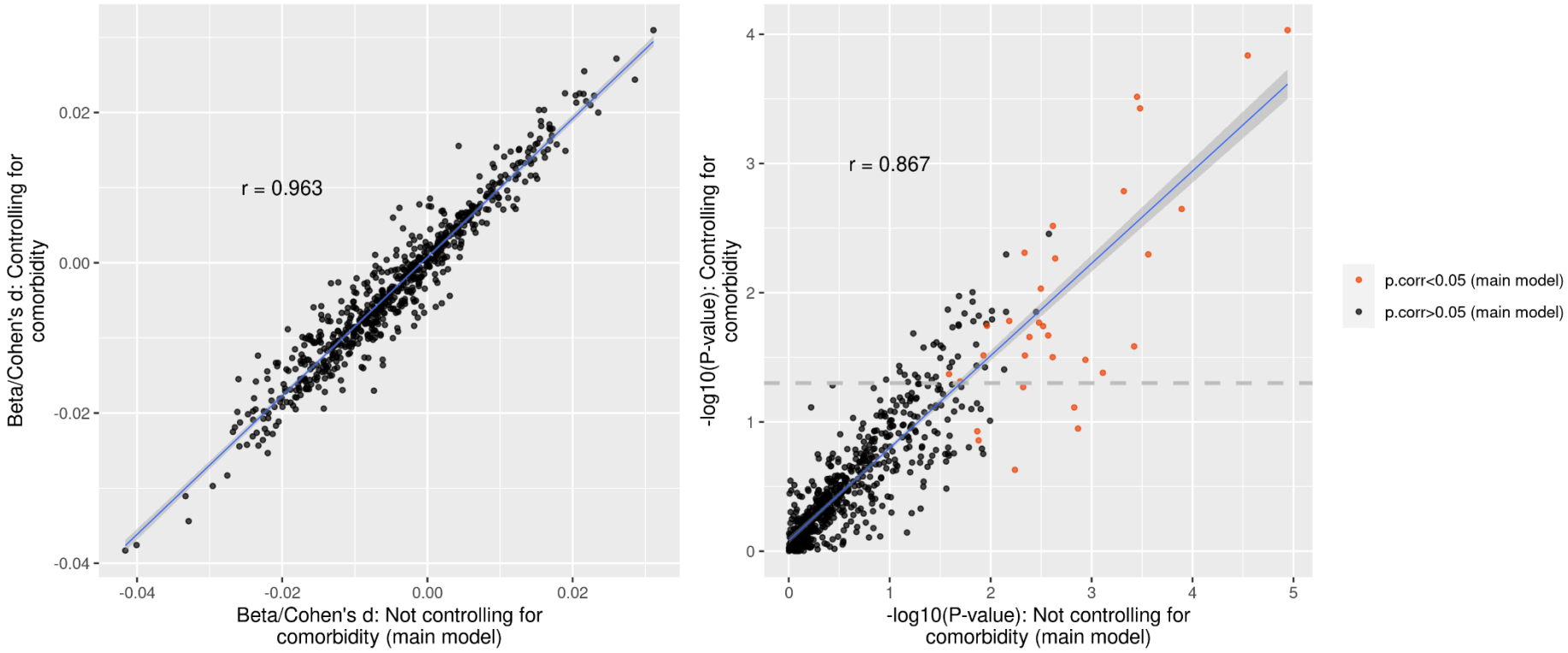


Figure S28. Correlations of effect sizes and p-values for the main model and a secondary model accounting for comorbidity of Bipolar I, Bipolar II, ADHD, Psychosis and Conduct disorder. X-axes represent statistics of the main model and the Y-axes represent statistics of the secondary model. The left panel shows the correlation of standardised effect sizes (regression coefficient/Cohen's d depending on which independent variable was used – MDD/depressive symptoms), and the right panel shows the correlation of p-values. In the right panel, the grey dashed line shows the threshold of nominal significance ($p < 0.05$).


r = 0.963
Beta/Cohen's d: Controlling for comorbidity
Beta/Cohen's d: Not controlling for comorbidity (main model)
0.02
0.00
-0.02
-0.04
r = 0.867
-log10(P-value): Controlling for comorbidity
-log10(P-value): Not controlling for comorbidity (main model)
0
1
2
3
4
5
p.corr<0.05 (main model)
p.corr>0.05 (main model)

Supplementary Data 1. Results for associations between MDD/Depressive symptoms and global neuroimaging measures. MDD.caregiver = MDD based on reports by caregivers, Depressive symptoms.p = depressive symptoms based on reports by caregivers, MDD.child = MDD based on reports by children and Depressive symptoms.y = depressive symptoms based on reports by children.

Supplementary Data 2. Results for associations between MDD/Depressive symptoms and cortical thickness. MDD.caregiver = MDD based on reports by caregivers, Depressive symptoms.p = depressive symptoms based on reports by caregivers, MDD.child = MDD based on reports by children and Depressive symptoms.y = depressive symptoms based on reports by children.

Supplementary Data 3. Results for associations between MDD/Depressive symptoms and cortical surface area. MDD.caregiver = MDD based on reports by caregivers, Depressive symptoms.p = depressive symptoms based on reports by caregivers, MDD.child = MDD based on reports by children and Depressive symptoms.y = depressive symptoms based on reports by children.

Supplementary Data 4. Results for associations between MDD/Depressive symptoms and cortical sulcal depth. MDD.caregiver = MDD based on reports by caregivers, Depressive symptoms.p = depressive symptoms based on reports by caregivers, MDD.child = MDD based on reports by children and Depressive symptoms.y = depressive symptoms based on reports by children.

Supplementary Data 5. Results for associations between MDD/Depressive symptoms and cortical volume. MDD.caregiver = MDD based on reports by caregivers, Depressive symptoms.p = depressive symptoms based on reports by caregivers, MDD.child = MDD based on reports by children and Depressive symptoms.y = depressive symptoms based on reports by children.

Supplementary Data 6. Results for associations between MDD/Depressive symptoms and white-matter fractional anisotropy. MDD.caregiver = MDD based on reports by caregivers, Depressive symptoms.p = depressive symptoms based on reports by caregivers, MDD.child = MDD based on reports by children and Depressive symptoms.y = depressive symptoms based on reports by children.

Supplementary Data 7. Results for associations between MDD/Depressive symptoms and white-matter mean diffusivity. MDD.caregiver = MDD based on reports by caregivers, Depressive symptoms.p = depressive symptoms based on reports by caregivers, MDD.child = MDD based on reports by children and Depressive symptoms.y = depressive symptoms based on reports by children.

Appendix 2: Supplementary Information for Chapter 4

The role of brain structure in the association between pubertal timing and depression risk in an early adolescent sample (the ABCD Study®): A registered report

MacSweeney et al.

Main analyses

Hypothesis 1 — Earlier pubertal timing is associated with later depression symptoms

The complete model output, including incidence rate ratios, standard errors, and p-values for the predictor and all covariates across the base and fully adjusted models are reported in Table S1 (females) and Table S2 (males).

Females: Effect of pubertal timing on youth depression

	Base			**Full**		
Predictors	*IRR*	*SE*	*P-Value*	*IRR*	*SE*	*P-Value*
(Intercept)	0.821	0.032	**<0.001**	0.636	0.172	0.093
Pubertal timing	1.313	0.041	**<0.001**	1.220	0.040	**<0.001**
Age	1.111	0.034	**0.001**	1.097	0.034	**0.002**
Race: Black	0.630	0.073	**<0.001**	0.723	0.088	**0.007**
Race: Asian	0.854	0.187	0.471	1.032	0.222	0.884
Race: AIAN/NHPI	0.961	0.360	0.915	0.884	0.328	0.739
Race: Other	1.299	0.213	0.110	1.497	0.246	**0.014**
Race: Mixed	1.217	0.116	**0.040**	1.138	0.107	0.169
BMI				1.105	0.036	**0.002**
Household income: $5,000-$11,999				1.094	0.371	0.792
Household income: $12,000-$15,999				0.850	0.297	0.642

Household income: $16,000-$24,999	1.379	0.430	0.303
Household income: $25,000-$34,999	1.107	0.331	0.735
Household income: $35,000-$49,999	1.339	0.383	0.307
Household income: $50,000-$74,999	1.300	0.364	0.349
Household income: $75,000-$99,999	1.245	0.349	0.433
Household income: $100,000-$199,999	1.338	0.367	0.288
Household income: >$200,000	1.255	0.358	0.427
Parent depressive symptoms	1.500	0.046	**<0.001**

Random Effects

σ^2	0.81	0.81
τ_{00}	1.29 $_{\text{rel_family_id}}$	1.12 $_{\text{rel_family_id}}$
	0.00 $_{\text{site_id_y1}}$	0.00 $_{\text{site_id_y1}}$
N	21 $_{\text{site_id_y1}}$	21 $_{\text{site_id_y1}}$
	2155 $_{\text{rel_family_id}}$	2105 $_{\text{rel_family_id}}$
Observations	2491	2426
Marginal R^2 / Conditional R^2	0.111 / NA	0.252 / NA

Table S1 — Females: Base and Fully adjusted models with associated statistics for effect of earlier pubertal timing on later depressive symptoms. Note: IRR = incidence rate ratio. AIAN/NHPI = AIAN/NHPI = American Indian/Alaska Native/Native Hawaiian and other Pacific Islander.

Males: Effect of pubertal timing on youth depression

	Base			Full		
Predictors	*IRR*	*SE*	*P-Value*	*IRR*	*SE*	*P-Value*
(Intercept)	0.738	0.027	**<0.001**	0.864	0.178	0.478
Pubertal timing	1.088	0.033	**0.006**	1.045	0.032	0.151
Age	1.041	0.031	0.167	1.025	0.029	0.397
Race: Black	0.685	0.083	**0.002**	0.671	0.083	**0.001**
Race: Asian	0.586	0.134	**0.019**	0.735	0.167	0.175
Race: AIAN/NHPI	0.950	0.419	0.908	1.046	0.443	0.915
Race: Other	1.096	0.186	0.590	1.036	0.173	0.834
Race: Mixed	1.165	0.118	0.130	1.087	0.106	0.391
BMI				1.126	0.033	**<0.001**
Household income: $5,000-$11,999				0.677	0.198	0.182
Household income: $12,000-$15,999				0.722	0.232	0.311
Household income: $16,000-$24,999				1.051	0.257	0.838
Household income: $25,000-$34,999				0.909	0.218	0.692
Household income: $35,000-$49,999				0.989	0.222	0.962
Household income: $50,000-$74,999				0.857	0.186	0.476

Household income: $75,000-$99,999		0.839	0.182	0.419
Household income: $100,000-$199,999		0.819	0.174	0.347
Household income: >$200,000		0.840	0.189	0.436
Parent depressive symptoms		1.562	0.045	**<0.001**
Random Effects				
σ^2	0.87	0.87		
τ_{00}	1.28 $_{\text{rel_family_id}}$	1.04 $_{\text{rel_family_id}}$		
	0.00 $_{\text{site_id_y1}}$	0.00 $_{\text{site_id_y1}}$		
N	21 $_{\text{site_id_y1}}$	21 $_{\text{site_id_y1}}$		
	2412 $_{\text{rel_family_id}}$	2369 $_{\text{rel_family_id}}$		
Observations	2752	2703		
Marginal R^2 / Conditional R^2	0.030 / NA	0.224 / NA		

Table S2 — Males: Base and Fully adjusted models with associated statistics for effect of earlier pubertal timing on later depressive symptoms. Note: IRR = incidence rate ratio. AIAN/NHPI = AIAN/NHPI = American Indian/Alaska Native/Native Hawaiian and other Pacific Islander.

Exploratory analyses

Hypothesis 1: Gonadal and adrenal pubertal timing measures

Independent models for gonadal and adrenal pubertal timing are presented in Tables S3 & S4 for females and S5 & S6 for males. We also ran a model that included both gonadal and adrenal timing to examine whether one aspect of pubertal development was associated with youth depression above and beyond the other. These results are presented in Tables S7 (female) and S8 (males).

Females: Effect of adrenal pubertal timing on youth depression

	Base			**Full**		
Predictors	*IRR*	*SE*	*P-Value*	*IRR*	*SE*	*P-Value*
(Intercept)	0.814	0.032	**<0.001**	0.654	0.176	0.114
Adrenal pubertal timing	1.259	0.040	**<0.001**	1.179	0.038	**<0.001**
Age	1.114	0.034	**<0.001**	1.095	0.033	**0.003**
Race: Black	0.667	0.077	**<0.001**	0.734	0.089	**0.011**
Race: Asian	0.881	0.193	0.563	1.066	0.229	0.766
Race: AIAN/NHPI	1.023	0.382	0.951	0.900	0.333	0.775
Race: Other	1.388	0.227	**0.045**	1.545	0.253	**0.008**
Race: Mixed	1.248	0.119	**0.020**	1.151	0.108	0.133
BMI				1.131	0.036	**<0.001**
Household income: $5,000-$11,999				1.074	0.364	0.834

Household income: $12,000-$15,999	0.843	0.294	0.626
Household income: $16,000-$24,999	1.372	0.427	0.310
Household income: $25,000-$34,999	1.089	0.325	0.775
Household income: $35,000-$49,999	1.308	0.373	0.347
Household income: $50,000-$74,999	1.273	0.355	0.387
Household income: $75,000-$99,999	1.210	0.337	0.495
Household income: $100,000-$199,999	1.290	0.352	0.351
Household income: >$200,000	1.197	0.340	0.528
Parent depressive symptoms	1.501	0.046	**<0.001**

Random Effects

σ^2	0.81	0.81
τ_{00}	1.29 $_{\text{rel_family_id}}$	1.11 $_{\text{rel_family_id}}$
	0.00 $_{\text{site_id_y1}}$	0.00 $_{\text{site_id_y1}}$
N	21 $_{\text{site_id_y1}}$	21 $_{\text{site_id_y1}}$
	2155 $_{\text{rel_family_id}}$	2105 $_{\text{rel_family_id}}$
Observations	2491	2426
Marginal R^2 / Conditional R^2	0.092 / NA	0.246 / NA

Table S3 — Females: Adrenal pubertal timing and youth depression: Base and Fully adjusted models with associated statistics. Note: IRR = incidence rate ratio. AIAN/NHPI = AIAN/NHPI = American Indian/Alaska Native/Native Hawaiian and other Pacific Islander.

Females: Effect of gonadal pubertal timing on youth depression

	Base			Full		
Predictors	*IRR*	*SE*	*P-Value*	*IRR*	*SE*	*P-Value*
(Intercept)	0.817	0.032	**<0.001**	0.628	0.170	0.085
Gonadal pubertal timing	1.281	0.039	**<0.001**	1.191	0.038	**<0.001**
Age	1.110	0.034	**0.001**	1.096	0.034	**0.003**
Race: Black	0.646	0.075	**<0.001**	0.737	0.089	**0.012**
Race: Asian	0.834	0.183	0.410	1.004	0.218	0.985
Race: AIAN/NHPI	0.951	0.358	0.894	0.871	0.324	0.710
Race: Other	1.267	0.209	0.150	1.467	0.242	**0.020**
Race: Mixed	1.214	0.117	**0.044**	1.133	0.107	0.184
BMI				1.114	0.036	**0.001**
Household income: $5,000-$11,999				1.105	0.376	0.769
Household income: $12,000-$15,999				0.870	0.305	0.691
Household income: $16,000-$24,999				1.391	0.435	0.291
Household income: $25,000-$34,999				1.131	0.340	0.682
Household income: $35,000-$49,999				1.351	0.388	0.295
Household income: $50,000-$74,999				1.315	0.369	0.330

Household income: $75,000-$99,999	1.258	0.353	0.413
Household income: $100,000-$199,999	1.348	0.371	0.277
Household income: >$200,000	1.266	0.362	0.409
Parent depressive symptoms	1.506	0.047	**<0.001**

Random Effects

σ^2	0.81	0.81
τ_{00}	1.30 $_{\text{rel_family_id}}$	1.13 $_{\text{rel_family_id}}$
	0.00 $_{\text{site_id_y1}}$	0.00 $_{\text{site_id_y1}}$
N	21 $_{\text{site_id_y1}}$	21 $_{\text{site_id_y1}}$
	2155 $_{\text{rel_family_id}}$	2105 $_{\text{rel_family_id}}$
Observations	2491	2426
Marginal R^2 / Conditional R^2	0.099 / NA	0.247 / NA

Table S4 — Females: Gonadal pubertal timing and youth depression: Base and Fully adjusted models with associated statistics. Note: IRR = incidence rate ratio. AIAN/NHPI = AIAN/NHPI = American Indian/Alaska Native/Native Hawaiian and other Pacific Islander.

Males: Effect of adrenal pubertal timing on youth depression

	Base			**Full**		
Predictors	*IRR*	*SE*	*P-Value*	*IRR*	*SE*	*P-Value*
(Intercept)	0.737	0.027	**<0.001**	0.848	0.176	0.425
Adrenal pubertal timing	1.105	0.033	**0.001**	1.055	0.032	0.078
Age	1.042	0.031	0.162	1.025	0.029	0.388

Race: Black	0.682	0.082	**0.002**	0.670	0.083	**0.001**
Race: Asian	0.591	0.135	**0.021**	0.737	0.168	0.180
Race: AIAN/NHPI	0.966	0.427	0.939	1.051	0.446	0.907
Race: Other	1.110	0.188	0.538	1.043	0.174	0.802
Race: Mixed	1.164	0.118	0.134	1.087	0.106	0.392
BMI				1.125	0.033	**<0.001**
Household income: $5,000-$11,999				0.695	0.203	0.214
Household income: $12,000-$15,999				0.733	0.236	0.333
Household income: $16,000-$24,999				1.082	0.265	0.747
Household income: $25,000-$34,999				0.924	0.222	0.741
Household income: $35,000-$49,999				1.010	0.227	0.966
Household income: $50,000-$74,999				0.872	0.190	0.530
Household income: $75,000-$99,999				0.855	0.186	0.473
Household income: $100,000-$199,999				0.833	0.177	0.391
Household income: >$200,000				0.854	0.192	0.482
Parent depressive symptoms				1.560	0.045	**<0.001**

Random Effects

σ^2	0.87	0.87
τ_{00}	1.28 $_{rel_family_id}$	1.04 $_{rel_family_id}$
	0.00 $_{site_id_y1}$	0.00 $_{site_id_y1}$
ICC		0.54
N	21 $_{site_id_y1}$	21 $_{site_id_y1}$
	2412 $_{rel_family_id}$	2369 $_{rel_family_id}$
Observations	2752	2703
Marginal R^2 / Conditional R^2	0.033 / NA	0.116 / 0.597

Table S5 — Males: Adrenal pubertal timing and youth depression: Base and Fully adjusted models with associated statistics. Note: IRR = incidence rate ratio. AIAN/NHPI = AIAN/NHPI = American Indian/Alaska Native/Native Hawaiian and other Pacific Islander.

Males: Effect of gonadal pubertal timing on youth depression

	Base			**Full**		
Predictors	*IRR*	*SE*	*P-Value*	*IRR*	*SE*	*P-Value*
(Intercept)	0.734	0.027	**<0.001**	0.862	0.178	0.472
Gonadal pubertal timing	1.044	0.032	0.154	1.021	0.031	0.487
Age	1.042	0.031	0.165	1.024	0.029	0.407
Race: Black	0.714	0.086	**0.005**	0.684	0.084	**0.002**
Race: Asian	0.581	0.133	**0.018**	0.732	0.167	0.171
Race: AIAN/NHPI	0.950	0.419	0.908	1.030	0.436	0.945
Race: Other	1.107	0.188	0.549	1.041	0.174	0.810
Race: Mixed	1.177	0.119	0.106	1.094	0.106	0.357
BMI				1.133	0.033	**<0.001**

Household income: $5,000-$11,999	0.680	0.199	0.187
Household income: $12,000-$15,999	0.723	0.232	0.312
Household income: $16,000-$24,999	1.051	0.256	0.839
Household income: $25,000-$34,999	0.913	0.219	0.703
Household income: $35,000-$49,999	0.990	0.222	0.964
Household income: $50,000-$74,999	0.858	0.186	0.479
Household income: $75,000-$99,999	0.838	0.182	0.417
Household income: $100,000-$199,999	0.818	0.173	0.342
Household income: >$200,000	0.838	0.188	0.431
Parent depressive symptoms	1.563	0.045	**<0.001**

Random Effects

σ^2	0.87	0.87
τ_{00}	1.28 $_{rel_family_id}$	1.04 $_{rel_family_id}$
	0.00 $_{site_id_y1}$	0.00 $_{site_id_y1}$
N	21 $_{site_id_y1}$	21 $_{site_id_y1}$
	2412 $_{rel_family_id}$	2369 $_{rel_family_id}$
Observations	2752	2703

Marginal R^2 / Conditional R^2	0.025 / NA	0.223 / NA

Table S6 — Males: Gonadal pubertal timing and youth depression: Base and Fully adjusted models with associated statistics. Note: IRR = incidence rate ratio. AIAN/NHPI = AIAN/NHPI = American Indian/Alaska Native/Native Hawaiian and other Pacific Islander.

Females: Effect of gonadal pubertal timing on youth depression, controlling for adrenal pubertal timing

		Base	
Predictors	*IRR*	*SE*	*P-Value*
(Intercept)	0.821	0.032	**<0.001**
Gonadal pubertal timing	1.194	0.044	**<0.001**
Adrenal pubertal timing	1.132	0.044	**0.001**
Age	1.111	0.034	**0.001**
Race: Black	0.629	0.073	**<0.001**
Race: Asian	0.851	0.186	0.463
Race: AIAN/NHPI	0.973	0.364	0.941
Race: Other	1.294	0.212	0.117
Race: Mixed	1.216	0.116	**0.041**

Random Effects

σ^2	0.81
$\tau_{00\ rel_family_id}$	1.29
$\tau_{00\ site_id_y1}$	0.00
ICC	0.61
$N_{site_id_y1}$	21
$N_{rel_family_id}$	2155

Observations	2491
Marginal R^2 / Conditional R^2	0.046 / 0.632

Table S7 — Females: Effect of gonadal pubertal timing on youth depression controlling for adrenal pubertal timing. Base model with associated statistics. Note: IRR = incidence rate ratio. AIAN/NHPI = AIAN/NHPI = American Indian/Alaska Native/Native Hawaiian and other Pacific Islander.

Males: Effect of gonadal timing on youth depression, controlling for adrenal pubertal timing

	Base		
Predictors	*IRR*	*SE*	*P-Value*
(Intercept)	0.737	0.027	**<0.001**
Gonadal pubertal timing	0.994	0.034	0.850
Adrenal pubertal timing	1.109	0.038	**0.003**
Age	1.042	0.031	0.161
Race: Black	0.684	0.083	**0.002**
Race: Asian	0.591	0.135	**0.022**
Race: AIAN/NHPI	0.959	0.425	0.925
Race: Other	1.113	0.189	0.528
Race: Mixed	1.165	0.118	0.132

Random Effects

σ^2	0.87
$\tau_{00\ rel_family_id}$	1.28
$\tau_{00\ site_id_y1}$	0.00
$N_{site_id_y1}$	21
$N_{rel_family_id}$	2412

Observations	2752
Marginal R^2 / Conditional R^2	0.033 / NA

Table S8 — Males: Effect of gonadal pubertal timing on youth depression controlling for adrenal pubertal timing. Base model with associated statistics. Note: IRR = incidence rate ratio. AIAN/NHPI = AIAN/NHPI = American Indian/Alaska Native/Native Hawaiian and other Pacific Islander.

Sensitivity Analyses

Hypothesis 1: Multiple imputation of missing outcome and covariate data

For H1, multiple imputation by chained equations (MICE) was undertaken using the "mice" package in R (Buuren & Groothuis-Oudshoorn, 2011) to impute missing data for youth depression at year 3 (outcome), BMI at year 1 and parental mood at year 2. In our final sample (Females: N= 2533; Males: N = 2792) which included participants with complete puberty data and who had attended the year 3 follow-up appointment, there was no missing data for site, age, sex, or the population weighting score variables. We did not impute data for race/ethnicity so participants with missing data (N = 30 (females); N = 26 (males)) for this variable were not included in the imputation analysis. Further, we did not impute household income as this data was only collected at baseline and we were unable to find a suitable auxiliary variable (e.g., highest parental education) as participants with missing household income data were also missing parental education data.

Auxiliary variables were only included if they predicted the variable being imputed or missingness in this variable (to reduce the bias of variables being "missing not at random") or if they had <40% missing data. The auxiliary variables included were: youth depression at baseline, year 1 and year 2 (measured via the CBCL withdrawn/depressed subscale); youth anxiety at baseline, year 1, year 2, and year 3; parent depression at baseline; and BMI at baseline. One hundred imputed datasets were created. Effect sizes from each imputed dataset were then pooled using Rubin's rule. We note that missingness in our final sample was very low in both our base models (females = 42/2533 (1.66%); males = 40/2792 (1.34%)) and fully adjusted models (females = 107/2533 (4.22%); males = 89/2792 (3.19%).

Similar effect sizes were found when missing data was imputed and pooled for both females and males, as shown in Table S9 and S10 respectively.

Females: Pooled effect sizes for the association between pubertal timing on youth depression using MICE

	Base			Full		
Predictors	*Beta Estimate*	*SE*	*P-Value*	*Beta Estimate*	*SE*	*P-Value*
(Intercept)	-0.193	0.039	**<0.001**	-0.448	0.258	0.083
Pubertal timing	0.273	0.031	**<0.001**	0.196	0.032	**<0.001**
Age	0.106	0.031	**0.001**	0.092	0.030	**0.002**
Race: Black	-0.463	0.116	**<0.001**	-0.343	0.118	**0.004**
Race: Asian	-0.162	0.215	0.45	0.043	0.207	0.835
Race: AIAN/NHPI	-0.046	0.375	0.902	-0.188	0.367	0.609
Race: Other	0.213	0.162	0.189	0.298	0.160	0.062
Race: Mixed	0.186	0.096	0.052	0.109	0.093	0.241
BMI				0.101	0.032	**0.001**
Household income: $5,000-$11,999				0.117	0.325	0.719
Household income: $12,000-$15,999				-0.108	0.327	0.742
Household income: $16,000-$24,999				0.345	0.299	0.248
Household income: $25,000-$34,999				0.172	0.286	0.548
Household income: $35,000-$49,999				0.286	0.274	0.297
Household income: $50,000-$74,999				0.257	0.268	0.337

Household income: $75,000-$99,999	0.225	0.268	0.402
Household income: $100,000-$199,999	0.290	0.262	0.269
Household income: >$200,000	0.198	0.274	0.470
Parent depressive symptoms	0.408	0.030	**<0.001**

Table S9 — Females: Effect of pubertal timing on youth depression with imputed missing data. Base and fully adjusted models with associated statistics (pooled effect sizes using MICE). Note: AIAN/NHPI = AIAN/NHPI = American Indian/Alaska Native/Native Hawaiian and other Pacific Islander.

Males: Pooled effect sizes for the association between pubertal timing on youth depression using MICE

	Base			**Full**		
Predictors	*Beta Estimate*	*SE*	*P-Value*	*Beta Estimate*	*SE*	*P-Value*
(Intercept)	-0.302	0.037	**<0.001**	-0.220	0.202	0.276
Pubertal timing	0.087	0.030	**0.004**	0.048	0.030	0.111
Age	0.041	0.029	0.166	0.022	0.028	0.430
Race: Black	-0.378	0.121	**0.002**	-0.377	0.121	**0.002**
Race: Asian	-0.547	0.228	0.017	-0.272	0.219	0.215
Race: AIAN/NHPI	-0.104	0.437	0.812	-0.007	0.420	0.987
Race: Other	0.073	0.169	0.667	0.048	0.164	0.771
Race: Mixed	0.151	0.101	0.134	0.074	0.096	0.443
BMI				0.119	0.029	**<0.001**
Household income: $5,000-$11,999				-0.441	0.287	0.125
Household income: $12,000-$15,999				-0.318	0.317	0.316

Household income: $16,000-$24,999	0.140	0.240	0.559
Household income: $25,000-$34,999	0.003	0.234	0.989
Household income: $35,000-$49,999	0.072	0.220	0.744
Household income: $50,000-$74,999	-0.084	0.213	0.693
Household income: $75,000-$99,999	-0.100	0.213	0.639
Household income: $100,000-$199,999	-0.118	0.208	0.571
Household income: >$200,000	-0.094	0.220	0.668
Parent depressive symptoms	0.442	0.028	**<0.001**

Table S10 — Males: Effect of pubertal timing on youth depression with imputed missing data. Base and fully adjusted models with associated statistics (pooled effect sizes using MICE). Note: AIAN/NHPI = AIAN/NHPI = American Indian/Alaska Native/Native Hawaiian and other Pacific Islander.

Hypothesis 1: Controlling for earlier youth depression

We examined the association between earlier pubertal timing and the potential change (or rather worsening) of depressive symptoms between timepoints (i.e., Year 1 and Year 3) by including Year 1 youth depressive symptoms as an additional covariate in our base and fully adjusted models.

Females: Effect of pubertal timing on youth depression controlling for earlier youth depression

	Base			**Full**		
Predictors	*IRR*	*SE*	*P-Value*	*IRR*	*SE*	*P-Value*
(Intercept)	1.000	0.034	1.000	0.635	0.152	0.058

Pubertal timing	1.173	0.032	**<0.001**	1.173	0.034	**<0.001**
Age	1.064	0.029	**0.021**	1.076	0.030	**0.009**
Race: Black	0.761	0.076	**0.006**	0.809	0.087	**0.050**
Race: Asian	0.950	0.178	0.783	1.015	0.200	0.938
Race: AIAN/NHPI	0.971	0.313	0.928	1.000	0.336	0.999
Race: Other	1.047	0.149	0.745	1.153	0.172	0.340
Race: Mixed	1.089	0.090	0.300	1.061	0.090	0.487
Youth depressive symptoms(Y 1)	1.613	0.036	**<0.001**	1.581	0.038	**<0.001**
Household income: $5,000-$11,999				1.479	0.440	0.187
Household income: $12,000-$15,999				1.139	0.349	0.671
Household income: $16,000-$24,999				1.501	0.415	0.141
Household income: $25,000-$34,999				1.035	0.275	0.898
Household income: $35,000-$49,999				1.382	0.351	0.202
Household income: $50,000-$74,999				1.328	0.330	0.252
Household income: $75,000-$99,999				1.257	0.312	0.357
Household income: $100,000-$199,999				1.367	0.332	0.199
Household income: >$200,000				1.366	0.346	0.219

Parent depressive symptoms			1.287	0.037	**<0.001**
Random Effects					
σ^2	0.81		0.81		
τ_{00}	0.80 $_{rel_family_id}$		0.80 $_{rel_family_id}$		
	0.00 $_{site_id_y1}$		0.00 $_{site_id_y1}$		
N	21 $_{site_id_y1}$		21 $_{site_id_y1}$		
	2155 $_{rel_family_id}$		2137 $_{rel_family_id}$		
Observations	2491		2469		
Marginal R^2 / Conditional R^2	0.264 / NA		0.337 / NA		

Table S11— Females: Effect of pubertal timing on youth depression controlling for earlier youth depression (at Year 1). Note: IRR = incidence rate ratio. AIAN/NHPI = AIAN/NHPI = American Indian/Alaska Native/Native Hawaiian and other Pacific Islander.

Males: Effect of pubertal timing on youth depression controlling for earlier youth depression

	Base			**Full**		
Predictors	*IRR*	*SE*	*P-Value*	*IRR*	*SE*	*P-Value*
(Intercept)	0.527	0.020	**<0.001**	0.531	0.096	**<0.001**
Pubertal timing	1.052	0.028	0.057	1.053	0.028	0.056
Age	1.003	0.026	0.902	0.998	0.026	0.944
Race: Black	0.751	0.078	**0.006**	0.797	0.087	**0.037**
Race: Asian	0.750	0.149	0.147	0.830	0.168	0.357
Race: AIAN/NHPI	1.039	0.382	0.916	1.179	0.434	0.655
Race: Other	1.018	0.148	0.905	1.060	0.156	0.692
Race: Mixed	1.131	0.097	0.152	1.098	0.094	0.274

Youth depressive symptoms (Y 1)	1.394	0.016	**<0.001**	1.348	0.017	**<0.001**
Household income: $5,000-$11,999				0.815	0.207	0.421
Household income: $12,000-$15,999				0.732	0.208	0.273
Household income: $16,000-$24,999				1.017	0.217	0.936
Household income: $25,000-$34,999				1.014	0.211	0.945
Household income: $35,000-$49,999				1.176	0.229	0.407
Household income: $50,000-$74,999				0.962	0.181	0.836
Household income: $75,000-$99,999				0.994	0.188	0.976
Household income: $100,000-$199,999				1.039	0.191	0.836
Household income: >$200,000				1.025	0.200	0.901
Parent depressive symptoms				1.273	0.033	**<0.001**

Random Effects

σ^2	0.87	0.87
τ_{00}	0.70 $_{rel_family_id}$	0.65 $_{rel_family_id}$
	0.00 $_{site_id_y1}$	0.00 $_{site_id_y1}$
N	21 $_{site_id_y1}$	21 $_{site_id_y1}$

	2412 rel_family_id	2393 rel_family_id
Observations	2752	2733
Marginal R^2 / Conditional R^2	0.291 / NA	0.331 / NA

Table S12— Males: Effect of pubertal timing on youth depression controlling for earlier youth depression (at Year 1). Note: IRR = incidence rate ratio. AIAN/NHPI = AIAN/NHPI = American Indian/Alaska Native/Native Hawaiian and other Pacific Islander.

Hypothesis 1: Population weight raked propensity score

For H1, we also compared models with and without a population weighting score included as a weight in our generalised linear mixed model in Table S13 (females) and Table S14 (males). The inclusion of a population weighting score seemed to inflate the effect sizes and standard errors of minority race/ethnicity groups (e.g., AIAN/NHPI), which may be due to the small number of individuals (0.85% (females)/0.57% (males)) who reported this as their race/ethnicity in ABCD. Given that our main research questions did not examine race interaction effects, we decided to remove the population weighting score from our main analyses and report it as sensitivity analyses instead.

Females: Effect of pubertal timing on youth depression

	Base			**Full**			**Full (weighted)**		
Predictors	*IRR*	*SE*	*P-Value*	*IRR*	*SE*	*P-Value*	*IRR*	*SE*	*P-Value*
(Intercept)	0.821	0.032	**<0.001**	0.636	0.172	0.093	0.002	0.002	**<0.001**
Pubertal timing	1.313	0.041	**<0.001**	1.220	0.040	**<0.001**	1.217	0.004	**<0.001**
Age	1.111	0.034	**0.001**	1.097	0.034	**0.002**	1.222	0.004	**<0.001**
Race: Black	0.630	0.073	**<0.001**	0.723	0.088	**0.007**	0.257	0.138	**0.011**
Race: Asian	0.854	0.187	0.471	1.032	0.222	0.884	0.942	0.914	0.951
Race: AIAN/NHPI	0.961	0.360	0.915	0.884	0.328	0.739	299.294	175.494	**<0.001**
Race: Other	1.299	0.213	0.110	1.497	0.246	**0.014**	3.100	0.121	**<0.001**
Race: Mixed	1.217	0.116	**0.040**	1.138	0.107	0.169	1.598	0.024	**<0.001**
BMI				1.105	0.036	**0.002**	1.238	0.004	**<0.001**
Household income:				1.094	0.371	0.792	6.451	9.889	0.224

$5,000-$11,999						
Household income: $12,000-$15,999	0.850	0.297	0.642	4.560	5.604	0.217
Household income: $16,000-$24,999	1.379	0.430	0.303	4.523	5.558	0.219
Household income: $25,000-$34,999	1.107	0.331	0.735	0.069	0.087	**0.033**
Household income: $35,000-$49,999	1.339	0.383	0.307	8.615	10.584	0.080
Household income: $50,000-$74,999	1.300	0.364	0.349	36.266	44.236	**0.003**
Household income: $75,000-$99,999	1.245	0.349	0.433	0.408	0.498	0.462
Household income: $100,000-$199,999	1.338	0.367	0.288	3.847	4.667	0.267
Household income: >$200,000	1.255	0.358	0.427	1.976	2.491	0.589

Parent depressive symptoms				1.500	0.046	**<0.001**	1.144	0.016	**<0.001**
Random Effects									
σ^2	0.81			0.81			4.76		
τ_{00}	1.29 $_{rel_family_id}$			1.12 $_{rel_family_id}$			43.19 $_{rel_family_id}$		
	0.00 $_{site_id_y1}$			0.00 $_{site_id_y1}$			0.00 $_{site_id_y1}$		
ICC							0.90		
N	21 $_{site_id_y1}$			21 $_{site_id_y1}$			21 $_{site_id_y1}$		
	2155 $_{rel_family_id}$			2105 $_{rel_family_id}$			2105 $_{rel_family_id}$		
Observations	2491			2426			2426		
Marginal R^2 / Conditional R^2	0.111 / NA			0.252 / NA			0.063 / 0.907		

Table S13 — Females: Base model, fully adjusted model and fully adjusted model with population propensity score weight included as a weight in the model with associated statistics. Note: IRR = incidence rate ratio. AIAN/NHPI = AIAN/NHPI = American Indian/Alaska Native/Native Hawaiian and other Pacific Islander.

Males: Effect of pubertal timing on youth depression

	Base			**Full**			**Full (weighted)**		
Predictors	*IRR*	*SE*	*P-Value*	*IRR*	*SE*	*P-Value*	*IRR*	*SE*	*P-Value*
(Intercept)	0.738	0.027	**<0.001**	0.864	0.178	0.478	0.003	0.003	**<0.001**
Pubertal timing	1.088	0.033	**0.006**	1.045	0.032	0.151	1.011	0.003	**<0.001**
Age	1.041	0.031	0.167	1.025	0.029	0.397	1.061	0.003	**<0.001**
Race: Black	0.685	0.083	**0.002**	0.671	0.083	**0.001**	0.518	0.288	0.237
Race: Asian	0.586	0.134	**0.019**	0.735	0.167	0.175	0.012	0.012	**<0.001**
Race: AIAN/NHPI	0.950	0.419	0.908	1.046	0.443	0.915	0.012	0.030	0.074
Race: Other	1.096	0.186	0.590	1.036	0.173	0.834	1.599	0.074	**<0.001**

Race: Mixed	1.165	0.118	0.130	1.087	0.106	0.391	87.971	19.418	**<0.001**
BMI				1.126	0.033	**<0.001**	1.245	0.004	**<0.001**
Household income: $5,000-$11,999				0.677	0.198	0.182	0.055	0.077	**0.040**
Household income: $12,000-$15,999				0.722	0.232	0.311	0.121	0.188	0.173
Household income: $16,000-$24,999				1.051	0.257	0.838	1.760	1.816	0.584
Household income: $25,000-$34,999				0.909	0.218	0.692	1.695	1.748	0.609
Household income: $35,000-$49,999				0.989	0.222	0.962	2.744	2.830	0.328
Household income: $50,000-$74,999				0.857	0.186	0.476	0.101	0.104	**0.026**
Household income: $75,000-$99,999				0.839	0.182	0.419	6.749	6.927	0.063
Household income: $100,000-$199,999				0.819	0.174	0.347	0.624	0.640	0.645
Household income: >$200,000				0.840	0.189	0.436	0.970	1.047	0.978
Parent depressive symptoms				1.562	0.045	**<0.001**	1.302	0.014	**<0.001**
Random Effects									
σ^2	0.87			0.87			5.53		
τ_{00}	1.28 $_{rel_family_id}$			1.04 $_{rel_family_id}$			43.82 $_{rel_family_id}$		
	0.00 $_{site_id_y1}$			0.00 $_{site_id_y1}$			0.28 $_{site_id_y1}$		
ICC							0.89		
N	21 $_{site_id_y1}$			21 $_{site_id_y1}$			21 $_{site_id_y1}$		

	2412 $_{rel_family_id}$	2369 $_{rel_family_id}$	2369 $_{rel_family_id}$
Observations	2752	2703	2703
Marginal R^2 / Conditional R^2	0.030 / NA	0.224 / NA	0.089 / 0.899

Table S14 — Males: Base model, fully adjusted model and fully adjusted model with population propensity score weight included as a weight in the model with associated statistics. Note: IRR = incidence rate ratio. AIAN/NHPI = AIAN/NHPI = American Indian/Alaska Native/Native Hawaiian and other Pacific Islander.

Appendix 3: Supplementary Information for Chapters 6 & 7

Chapter 6: Irritability in Adolescent Depression — Pilot Study Methods

Chapter 7: Exploring Dynamic Functional Brain Networks in Adolescent Depression Using a Co-produced Novel Irritability Task

Irritability study: Recruitment poster

Have you experienced mental health difficulties related to sadness and low mood?

We are seeking young people aged 16-19 years who have experienced depressive symptoms to take part in a research study that seeks to understand adolescent depression.

The study involves having an fMRI scan of your brain while doing some computerised tasks and completing some other questionnaires outside the scanner. This study is being organised by researchers at the University of Edinburgh.

You will be required to attend the Clinical Research Imaging Centre (CRIC) at the Royal Infirmary of Edinburgh for a two-hour session. Travel expenses will be reimbursed.

If you are interested in finding out more, email: **evaimag@ed.ac.uk** or scan the QR code to see our Participant Information Sheet.

THE UNIVERSITY *of* EDINBURGH

SCAN ME

Irritability study: Participant information sheet

THE UNIVERSITY of EDINBURGH

Title of project: Development of novel neuroimaging markers for the detection of adolescent depression

Information Sheet for participants

You are invited to take part in a research study. To help you decide whether or not to take part, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully. Talk to others about the study if you wish. Contact us if there is anything that is not clear, or if you would like more information. Take time to decide whether or not you wish to take part.

What is the study about?

Major Depressive Disorder (MDD) is among the most prevalent of all psychiatric conditions, and often leads to difficulties in personal, familial, and social life. Depression in half of all adults has a starting point in adolescence, signposting this period as critical for the origin and formation of neurobiological features of the condition. Numerous brain processes may be involved in different thinking styles, many of which have not yet been identified. It is therefore important to first identify differences in thinking styles and brain activation and how they may form part of the picture of risk. This study will therefore allow us to observe what happens within the brain while completing tasks assessing thinking and cognition, allowing for a better understanding of neurobiological features of these actions.

Why have I been invited to take part?

You have been invited to take part in this research because you have previously agreed to be re-contacted for future research studies and because you have contacted us with an interest to take part.

To take part in this research study, you should be aged 16-19 years and fluent in English.

There are certain criteria, such as diagnosis of a known developmental or genetic condition, that may exclude you from taking part in the study. The researchers will discuss these details with you before your participation.

Do I have to take part?

No, it is up to you to decide whether or not to take part. If you do decide to take part, you will be given this information sheet to keep and be asked to sign a consent form. If you decide to take part, you are still free to withdraw at any time and without giving a reason.

What will happen if I take part?

You will be contacted by e-mail to discuss the study in more detail and ask any questions you may have about the study. If you are happy to take part in the study, you will be sent a short questionnaire to determine if you are eligible to take part in the study. If you are eligible and are still happy to take part, you will need to sign our study consent form.

You will be invited to an appointment to have a Magnetic Resonance Imaging (MRI) scan. An MRI scanner is a machine that takes pictures of your brain. At the appointment, the scanner will take pictures of your brain while you are carrying out a few tasks, and this will help us understand what is happening in the brains of people when undertaking these tasks. You will be in the scanner for around 60 minutes.

Here is a picture of an MRI scanner similar to the one you will use (*Credits: University of Edinburgh*).

After the scan, you will be asked to complete a few further tasks that measure cognitive abilities. These tasks will look at your memory, attention, and mood.

The appointment will take place at the research imaging facility of the University of Edinburgh, which is based at the Queen's Medical Research Institute (QMRI) in the Royal Infirmary Edinburgh Campus. After this appointment, there will be no further visits required.

The appointment would take a minimum of 120 minutes and a maximum of 150 minutes. We will work together to set appointment that will not clash with school or university times.

You will be able to take as many breaks as you want should you feel tired. Refreshments will be provided in case you need water throughout the procedure.

What are the possible disadvantages or risks associated with taking part?

Questionnaires and cognitive measures: The questionnaires and measures involved in this study have previously been used in research. You will fill in questionnaires about your behaviour and cognition, which can sometimes make you more aware of your mood and difficulties. This may sometimes lead to making you feel upset. If this is the case, you may find it helpful to speak to your parents or general practitioner (GP).

Additionally, if any answers on the completed questionnaires tell us something that makes us worried about your wellbeing or that of others, we will discuss this with you at the in-person assessment and advise you regarding whom to contact. With your permission, we will also give this information to your GP in case you need further support.

MRI Scanning: There are virtually no risks associated with having an MRI brain scan. MRI does not involve any exposure to radiation. The MRI uses powerful magnets to take pictures of the brain, so it is very important that you have no metal on or in your body. We will check this very carefully with you before you go into the MRI scanner. The radiographer will be able to tell you whether this would prevent you from taking part in the study.

The noise during the scan is very loud, so we will provide you with earplugs or headphones to make you feel comfortable. Some people also report discomfort from lying in the scanner, having their head movements restrained, or due to feeling claustrophobic. You will be given a safety button that you can press at any time if you want to stop the scan. If pressed, the researchers will talk to you straight away, and will immediately withdraw you from the scanner if requested.

In a small number of cases, irregularities can show up in the MRI scan. Most often, these are slight variations from typical brain structure that have no clinical meaning. Very occasionally,

we find something that has a clinical meaning. In this case, we will give this information to your GP, who will discuss this with you and advise you regarding whom to contact for further assessment.

What are the benefits of taking part?

There are no direct benefits to you taking part in this study, but the results from this study will help us understand more about the brain processes involved in cognitive tasks which are biased in adolescents with depression, and might help to improve the healthcare of patients in the future. Your travel expenses to attend the appointment will also be covered up to a reasonable amount.

Will I get feedback about the study?

Individual results will not be provided to study participants routinely or on request. This is because it is more helpful to look at responses at a group-level rather than on an individual level. Summaries of any main research findings will be made available online for the general public at https://www.eva-edinburgh.com/ and published papers will be made available online at https://www.ed.ac.uk/psychiatry/research.

Can I withdraw from the study if I don't want to take part anymore?

You can withdraw from the current study at any time without any reason. To withdraw from the study, you will need to inform the researchers involved. Their contact information is included at the end of this form. If you withdraw from the study because you don't wish to continue with it, we will still use the information we have collected from you when we are preparing the results of the study unless you tell us that you don't want your information to be included.

How will my privacy be protected?

All the information you provide for the current study is confidential and will not be shared with anyone outside of the research team. You will not write your name on any of the questionnaires or tasks, but will instead be given a non-identifiable participant ID code. All information collected will therefore be kept anonymous. All the information collected will be stored safely (in encrypted files or within the secure university network). We will review all data you have provided for destruction every five years, and once the study is complete only an anonymised

master copy of the data will exist in archives. The findings of this study will be written up and may be published in academic journals or presented at conferences, but names or any other identifying information will never be disclosed.

Before the appointment you will be asked your GP's name. A clinical radiologist or a neurologist will review your brain scan and prepare a clinical report that will be available on the NHS electronic systems. Your GP may receive the routine clinical report after your MRI scan appointment. In the unlikely event that any irregularities show up in the MRI scan, we will give this information to your GP.

Who is organising this research?

The current research is funded by Wellcome Trust, and is being organised by Dr Heather Whalley and Dr Liana Romaniuk, both research scientists in Psychiatry; Niamh MacSweeney, a PhD student in Psychiatry; and Dr Stella Chan, a Clinical Psychologist. The study is sponsored by the University of Edinburgh.

Who has reviewed the research?

This research study has been looked at by an independent group of people called a Research Ethics Committee. A favourable ethical opinion has been obtained by the ACCORD Medical Research Ethics Committee (AMREC).

What if there is a problem?

If you have any concerns about the study and your participation in it, please contact the researchers involved, whose contact information is at the end of this form, and they will do their best to answer your questions. You can also contact the independent contact, who is someone not involved with the project.

What will I do now?

If you are happy to participate in the study, please e-mail the research assistant (contact details below: evaimag@ed.ac.uk) to indicate your interest. You will then be sent a short questionnaire that will assess your eligibility to take part in the study, along with a consent form specific to this. Once you send this back, you will be contacted by a researcher who will let you know if you are the right fit for the study and provide you with more details.

If you are the right fit for the study, you will also be asked to schedule your appointment. You will be asked to sign the consent form at your appointment.

Contact details

Dr Heather Whalley

Senior Research Fellow in Psychiatry, Centre for Clinical Brain Sciences, The University of Edinburgh

Email: heather.whalley@ed.ac.uk

Tel: 0131 537 6767

Dr Liana Romaniuk

Research Fellow, Division of Psychiatry, Centre for Clinical Brain Sciences, University of Edinburgh

Email: liana.romaniuk@ed.ac.uk

Tel: 0131 537 6767

Dr Stella Chan

Reader in Clinical Psychology, School of Health in Social Science, The University of Edinburgh

Email: stella.chan@ed.ac.uk

Tel: 0131 651 3935

Niamh MacSweeney

PhD Student, Division of Psychiatry, Centre for Clinical Brain Sciences, The University of Edinburgh

Email: evaimag@ed.ac.uk

Tel: 0131 537 6687

Independent contact not involved with the project - Prof. Stephen Lawrie

Head of the Division of Psychiatry, Centre for Clinical Brain Science, The University of Edinburgh

Email: s.lawrie@ed.ac.uk

Tel: 0131 537 6671

Complaints

If you wish to make a complaint about the study please contact the University of Edinburgh Research Governance Team: resgov@accord.scot

Irritability study: Participant consent form

THE UNIVERSITY of EDINBURGH

Title of project: Development of novel neuroimaging markers for the detection of adolescent depression
Consent form

Participant number:

Essential consent items: *please note that you must consent to all of these items if you wish to take part.*

	Initial in the box to give consent
I confirm that I have read and understood the participant information sheet for the above research project (version 7, 23 November 2020) and the data protection sheet (15 November 2019) and have had the opportunity to consider the information and to have my questions answered to my satisfaction.	
I have been informed of the discomforts and risks that I may experience as part of this study.	
I understand that my participation is voluntary and that I am free to withdraw from the study at any time without providing a reason, and without my medical care and/or legal rights being affected.	
I understand that relevant sections of data collected during the study may be looked at by individuals from the Sponsor (University of Edinburgh) or other regulatory authorities where it is relevant to my taking part in this research. I give permission for these individuals to have access to my data.	
I understand that all data provided will be stored safely for 5 years and that after 5 years, this data will be reviewed for possible destruction.	
I agree to my General Practitioner being informed of my participation in the study.	
I understand my scans will be viewed by a doctor, and that my GP will be informed of my participation in the study and may be provided with a routine clinical report after the MRI scan appointment.	
I give permission for the research team to contact my GP should they be worried about any answers provided on the questionnaires.	
I agree to take part in the above research project.	

***Optional consent items**: please note that you can take part in the study even if you don't consent to these items. You may leave these items blank and discuss them with the researcher.*

	YES	**NO**
Optional: I give permission for the researchers to re-contact me (using the contact information provided) with information about possible future research studies. I understand that there is no obligation to take part in any future research studies.		
Optional: I agree for my brain scan to be used in future ethically-approved studies and shared with the research community on the University of Edinburgh public data repository DataShare.		

Print your name clearly: _____
Your signature: _____
Date: _____

Name of researcher: _____
Researcher's signature: _____
Date: _____

Irritability study: Ethics approval letter

THE UNIVERSITY of EDINBURGH

Edinburgh Medical School
Research Ethics Committee (EMREC)
emrec@ed.ac.uk

Dr Heather Whalley
University of Edinburgh
Royal Edinburgh Hospital
Morningside Place
Edinburgh
EH10 5HF

28th July 2020

Dear Dr Whalley

Study Title: Novel neuroimaging markers in adolescent depression

REC Reference: 19-HV-061

Many thanks for submitting your amendments to EMREC and confirming the minor change to the Information Sheet.

Ethical opinion

We can now give a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation.

Ethical review of research sites

NHS Sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS R&D office prior to the start of the study.

Non-NHS Sites

For non-NHS sites, site management permission should be obtained in accordance with the procedures of the relevant host organisation. Sponsors are not required to notify the Committee of approvals from host organisations.

The University of Edinburgh is a charitable body registered in Scotland, with registration number SC005336.

THE UNIVERSITY of EDINBURGH

Management permission must be obtained from each host organisation prior to the start of the study at the site concerned.

It is responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

Documents reviewed

The final list of documents reviewed and approved by the Committee is as follows:

Document	*Version*	*Date*
IRAS Amendment SA_01 26 June 2020	N/A	26/06/2020
AMREC Decision Letter	N/A	15/04/2020
CR007-T02 Non-CTIMP Protocol MCB	V6	24/06/2020
Demographic information MCB	V5	24/06/2020
Consent Form MCB	V5	24/06/2020
Information Sheet MCB	V5	24/06/2020
Brief Resilience Scale	N/A	24/06/2020
Affective Reactivity Scale	N/A	24/06/2020
Rumination Scale	N/A	24/06/2020
Neuroticism Subscale of EPQ	N/A	24/06/2020
Multidimensional Scale of Perceived Social Support	N/A	24/06/2020

With the Committee's best wishes for the success of this project.

19-HV-061 Please quote this number on all correspondence

Yours sincerely,

Sarah Cunningham-Burley.

Sarah Cunningham-Burley

Chair, EMREC
Dean, Molecular, Genetic and Population Health Sciences

The University of Edinburgh is a charitable body registered in Scotland, with registration number SC005336.

Irritability task development

As detailed in Chapter 6 of this thesis, we developed a novel fMRI task targeting irritability. The mean irritability rating for each of the 51 scenarios generated by the young people (N = 61, aged 16-18 years) who co-developed our task can be found in Table S1. The 18 most highly rated irritating scenarios were used as stimuli in our fMRI task

Scenario	Mean irritability rating
1. I find it irritating when I am talking to someone, and they are being rude to me.	4.525
2. I find it irritating when I see people being rude to someone else.	4.459
3. I find it irritating when people lie.	4.328
4. I find it irritating when people don't listen to me in a conversation, or talk over me.	4.279
5. I find it irritating when people tell me to do something, even though I was going to do it.	4.279
6. I find it irritating when my parents do not take me seriously.	4.230
7. I find it irritating when I am being patronised.	4.230
8. I find it irritating when I get paid child-rate pay, but I am expected to pay adult-rate tickets on public transport.	4.213
9. I find it irritating when people don't listen to me in a conversation, especially when I am trying to be helpful.	4.197
10. I find it irritating when I do not perform as well as I can in any task.	4.066
11. I find it irritating when I am expected to do things that I don't know how to do.	3.984
12. I find it irritating when I realise that my parents or adults are not telling me the truth about something.	3.934
13. I find it irritating when I am asked to do chores several times, even though I already said I will do them.	3.934
14. I find it irritating when older people say that I am acting in a stereotypical teenager way.	3.934

15. I find it irritating when my parents listen in on my conversations and then become angry because they did not hear the entire conversation.	3.885
16. I find it irritating when my parents tell me to do something multiple times.	3.869
17. I find it irritating when I cannot motivate myself in a situation.	3.770
18. I find it irritating when I panic in a situation, and I can't calm down easily.	3.721
19. I find it irritating when someone eats with their mouth open.	3.689
20. I find it irritating when there are technical issues with services I use (e.g., WiFi)	3.672
21. I find it irritating that I am expected to know what I want to do in life - I am only a teenager.	3.656
22. I find it irritating when others touch my things.	3.590
23. I find it irritating when my siblings take my clothes, toys, or belongings without asking.	3.590
24. I find it irritating when I am in the middle of a game or conversation, and people come into my room to distract or ask questions, as I need my privacy.	3.574
25. I find it irritating when my parents tell me what I should be studying.	3.557
26. I find it irritating when people say that I am always on my phone.	3.541
27. I find it irritating when my parents order me around.	3.525
28. I find it irritating when I am being forced to socialise, especially when I am not in the mood to do so.	3.508
29. I find it irritating when I don't manage to get any studying done.	3.492
30. I find it irritating when I am the only one of my siblings that is asked to do housework.	3.459
31. I find it irritating when I am concentrated or in a mood and someone comes to talk to me.	3.426
32. I find billionaires and wealth inequality irritating.	3.410

33. I find it irritating when I encounter slow walkers.	3.361
34. I find it irritating when people don't listen to my advice.	3.328
35. I find it irritating when I am on my bike and someone drives too close to me, endangering my life.	3.246
36. I find it irritating when people shout while they are speaking.	3.246
37. I find it irritating when there is disruption to my daily routine.	3.213
38. I find tedious things irritating.	3.164
39. I find it irritating that I am expected not to react to younger siblings because I am the older sibling.	3.148
40. I find it irritating when I am put in a situation where I need to make awkward conversations with people.	3.115
41. I find it irritating when I am hungry.	3.115
42. I find it irritating when people continually ask me questions.	3.098
43. I find it irritating when people do not take decisions based on logic and fail to protect their own feelings.	3.098
44. I find it irritating when I am ignored by people when working at a restaurant and bring drinks / food over to their table.	3.016
45. I find it irritating when I get bored.	2.967
46. I find it irritating when other people are late.	2.967
47. I find heat irritating.	2.967
48. I find it irritating when older people judge me for wearing ripped jeans.	2.885
49. I find it irritating when I am expected to do work without being provided with a reason.	2.869
50. I find housework irritating.	2.721
51. I find it irritating when my siblings ask me to make dinner for them, when I am already making dinner for myself.	2.328

Table S2 — Mean irritability rating for 51 scenarios that were derived by young people and rated by an independent sample of youth (N = 61).

Imaging quality control and analysis

Irritability condition quality assessment statistics from HALFpipe

Subject	Aroma noise fraction	FD mean	FD percentage	Mean tSNR
sub1036	0.6226	0.2180	7.3930	57.3911
sub1034	0.6078	0.1047	0.0000	78.7679
sub1033	0.4444	0.1052	0.0000	80.2471
sub1032	0.5686	0.1653	1.9455	65.5935
sub1031	0.5472	0.1566	1.5564	61.4244
sub1030	0.6500	0.1455	1.1673	64.7302
sub1029	0.5682	0.1052	0.0000	78.2633
sub1028	0.7843	0.1411	1.5564	61.4103
sub1027	0.2564	0.0783	0.0000	89.6807
sub1026	0.4286	0.0718	0.0000	91.6286
sub1025	0.5098	0.1321	1.1673	65.6119
sub1024	0.4255	0.0988	0.0000	87.3443
sub1023	0.4348	0.1324	0.0000	72.5388
sub1022	0.5556	0.1615	1.9455	65.8956
sub1021	0.3953	0.1066	0.0000	85.3053
sub1019	0.6897	0.1494	2.3346	62.9730
sub1018	0.5714	0.0741	0.0000	83.8705
sub1017	0.4894	0.1238	0.0000	84.6500
sub1015	0.4583	0.1507	0.3891	77.2757
sub1014	0.6486	0.1352	0.0000	74.0896
sub1013	0.5636	0.1962	5.4475	51.7724
sub1012	0.4792	0.1030	0.7782	68.3823
sub1011	0.4545	0.0959	1.5564	69.6137
sub1007	0.5217	0.1318	0.3891	63.4936
sub1006	0.5000	0.0906	0.0000	71.6633
sub1004	0.4082	0.1003	0.0000	79.5157
sub1003	0.3659	0.0789	0.0000	94.8052
sub1002	0.4419	0.0800	0.0000	80.6975
sub1001	0.5227	0.1089	0.3891	77.0080

Table S2 — Irritability condition quality assessment statistics from HALFpipe. FD = Framewise displacement; tSNR = temporal signal-to-noise ratio.

Resting state condition quality assessment statistics from HALFpipe

Subject	Aroma noise fraction	FD mean	FD percentage	Mean tSNR
sub1036	0.7907	0.2147	5.8366	48.2617
sub1034	0.5849	0.0911	0.0000	79.3496
sub1033	0.5435	0.1051	0.0000	72.4813
sub1032	0.4898	0.1460	0.3891	62.3415
sub1031	0.5778	0.1354	1.1673	58.5040
sub1030	0.5250	0.1438	0.7782	68.6001
sub1029	0.5349	0.1127	0.0000	76.9799
sub1028	0.6000	0.1018	0.7782	69.7860
sub1027	0.4615	0.0615	0.0000	86.1619
sub1026	0.3617	0.0730	0.0000	88.9376
sub1025	0.6863	0.1474	0.3891	60.8922
sub1024	0.4419	0.0907	0.0000	84.9890
sub1023	0.4681	0.1393	0.0000	70.4654
sub1022	0.5854	0.1324	0.0000	71.7844
sub1021	0.5000	0.0904	0.3891	80.9363
sub1019	0.6087	0.1553	3.1128	60.2315
sub1018	0.4634	0.1035	0.0000	82.4164
sub1017	0.5349	0.1389	0.0000	84.9121
sub1015	0.5000	0.1604	0.0000	72.3395
sub1014	0.4444	0.1124	0.0000	70.2412
sub1013	0.4912	0.1789	3.5019	54.9970
sub1012	0.4468	0.1178	0.0000	64.4444
sub1011	0.4912	0.1374	4.2802	57.8438
sub1007	0.5417	0.1271	0.7782	66.8450
sub1006	0.5682	0.1106	0.3891	67.8699
sub1004	0.4444	0.0889	0.7782	85.7341
sub1003	0.4500	0.0922	0.0000	79.1167
sub1002	0.5208	0.1025	0.0000	65.6336
sub1001	0.4000	0.1088	0.0000	76.4660

Table S3 — Resting state condition quality assessment statistics from HALFpipe. FD = Framewise displacement; tSNR = temporal signal-to-noise ratio.

Regions removed from AAL120 brain parcellation

No.	Anatomical description	AAL label
101, 102	Lobule IV, V of cerebellar hemisphere	Cerebellum_4_5
103, 104	Lobule VI of cerebellar hemisphere	Cerebellum_6
105, 106	Lobule VIIB of cerebellar hemisphere	Cerebellum_7b
107, 108	Lobule VIII of cerebellar hemisphere	Cerebellum_8
109, 110	Lobule IX of cerebellar hemisphere	Cerebellum_9
111, 112	Lobule X of cerebellar hemisphere	Cerebellum_10
113	Lobule I, II of vermis	Vermis_1_2
114	Lobule III of vermis	Vermis_3
115	Lobule IV, V of vermis	Vermis_4_5
116	Lobule VI of vermis	Vermis_6
117	Lobule VII of vermis	Vermis_7
118	Lobule VIII of vermis	Vermis_8
119	Lobule IX of vermis	Vermis_9
120	Lobule X of vermis	Vermis_10

Table S4 — The 20 brain regions removed from the AAL120 brain parcellation prior to LEiDA analysis due to NaN values across participants. This resulted in a 100-region brain parcellation.

10 References

Abdi, H. (2007). The Bonferonni and Šidák Corrections for Multiple Comparisons. In N. Salkind (Ed.), *Encyclopedia of Measurement and Statistics* (pp. 103–107). Sage. https://doi.org/10.4135/9781412952644

Abreu, A. P., & Kaiser, U. B. (2016). Pubertal development and regulation. *The Lancet Diabetes & Endocrinology*, *4*(3), 254–264. https://doi.org/10.1016/S2213-8587(15)00418-0

Achenbach, T. M. (2006). As others see us clinical and research implications of cross-informant correlations for psychopathology. *Current Directions in Psychological Science*, *15*(2), 94–98. https://doi.org/10.1111/j.0963-7214.2006.00414.x

Achenbach, T. M. (2011). Child Behavior Checklist. In J. S. Kreutzer, J. DeLuca, & B. Caplan (Eds.), *Encyclopedia of Clinical Neuropsychology* (pp. 546–552). Springer. https://doi.org/10.1007/978-0-387-79948-3_1529

Achenbach, T. M., & Rescorla, L. A. (2004). *Manual for the ASEBA school-age forms & profiles*. University of Vermont, Research Center for Children.

Achterbergh, L., Pitman, A., Birken, M., Pearce, E., Sno, H., & Johnson, S. (2020). The experience of loneliness among young people with depression: A qualitative meta-synthesis of the literature. *BMC Psychiatry*, *20*(1), 415. https://doi.org/10.1186/s12888-020-02818-3

Aghajani, M., Veer, I. M., Van Lang, N. D. J., Meens, P. H. F., Van Den Bulk, B. G., Rombouts, S. A. R. B., Vermeiren, R. R. J. M., & Van Der Wee, N. J. (2014). Altered white-matter architecture in treatment-naive adolescents with clinical depression. *Psychological Medicine*, *44*(11), 2287–2298. https://doi.org/10.1017/S0033291713003000

Agler, R., & De Boeck, P. (2017). On the Interpretation and Use of Mediation: Multiple

Perspectives on Mediation Analysis. *Frontiers in Psychology*, *8*, 1984. https://doi.org/10.3389/fpsyg.2017.01984

Alakörkkö, T., Saarimäki, H., Glerean, E., Saramäki, J., & Korhonen, O. (2017). Effects of spatial smoothing on functional brain networks. *European Journal of Neuroscience*, *46*(9), 2471–2480. https://doi.org/10.1111/ejn.13717

Albert, D., Chein, J., & Steinberg, L. (2013). The Teenage Brain: Peer Influences on Adolescent Decision Making. *Current Directions in Psychological Science*, *22*(2), 114–120. https://doi.org/10.1177/0963721412471347

Alfaro-Almagro, F., Jenkinson, M., Bangerter, N. K., Andersson, J. L. R., Griffanti, L., Douaud, G., Sotiropoulos, S. N., Jbabdi, S., Hernandez-Fernandez, M., Vallee, E., Vidaurre, D., Webster, M., McCarthy, P., Rorden, C., Daducci, A., Alexander, D. C., Zhang, H., Dragonu, I., Matthews, P. M., … Smith, S. M. (2018). Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank. *NeuroImage*, *166*, 400–424. https://doi.org/10.1016/j.neuroimage.2017.10.034

Allen, E. A., Damaraju, E., Plis, S. M., Erhardt, E. B., Eichele, T., & Calhoun, V. D. (2014). Tracking Whole-Brain Connectivity Dynamics in the Resting State. *Cerebral Cortex*, *24*(3), 663–676. https://doi.org/10.1093/cercor/bhs352

Allen, N. B., & Dahl, R. E. (2015). Multi-Level Models of Internalizing Disorders and Translational Developmental Science: Seeking Etiological Insights that can Inform Early Intervention Strategies. *Journal of Abnormal Child Psychology*, *43*(5), 875–883. https://doi.org/10.1007/s10802-015-0024-9

Alloy, L. B., Hamilton, J. L., Hamlat, E. J., & Abramson, L. Y. (2016). Pubertal development, emotion regulatory styles, and the emergence of sex differences in internalizing

disorders and symptoms in adolescence. *Clinical Psychological Science*, *4*(5), 867–881. https://doi.org/10.1177/2167702616643008

Alonso Martínez, S., Deco, G., Ter Horst, G. J., & Cabral, J. (2020). The Dynamics of Functional Brain Networks Associated With Depressive Symptoms in a Nonclinical Sample. *Frontiers in Neural Circuits*, *14*, 570583. https://doi.org/10.3389/fncir.2020.570583

American Psychiatric Association. (2013a). *Diagnostic and statistical manual of mental disorders (5th edition)*.

American Psychiatric Association. (2013b). *Diagnostic and Statistical Manual of Mental Disorders: DSM-5.* (5th ed.). American Psychiatric Association.

Amlien, I. K., Fjell, A. M., Tamnes, C. K., Grydeland, H., Krogsrud, S. K., Chaplin, T. A., Rosa, M. G. P., & Walhovd, K. B. (2016). Organizing principles of human cortical development—Thickness and area from 4 to 30 years: Insights from comparative primate neuroanatomy. *Cerebral Cortex*, *26*(1), 257–267. https://doi.org/10.1093/cercor/bhu214

Amos, R., Manalastas, E. J., White, R., Bos, H., & Patalay, P. (2020). Mental health, social adversity, and health-related outcomes in sexual minority adolescents: A contemporary national cohort study. *The Lancet. Child & Adolescent Health*, *4*(1), 36–45. https://doi.org/10.1016/S2352-4642(19)30339-6

Anderson, E. R., & Mayes, L. C. (2010). Race/ethnicity and internalizing disorders in youth: A review. *Clinical Psychology Review*, *30*(3), 338–348. https://doi.org/10.1016/j.cpr.2009.12.008

Andrews, J. L., Ahmed, S. P., & Blakemore, S.-J. (2021). Navigating the Social Environment in Adolescence: The Role of Social Brain Development. *Biological Psychiatry*, *89*(2), 109–

118. https://doi.org/10.1016/j.biopsych.2020.09.012

Andrews-Hanna, J. R., Smallwood, J., & Spreng, R. N. (2014). The default network and self-generated thought: Component processes, dynamic control, and clinical relevance. *Annals of the New York Academy of Sciences*, *1316*(1), 29–52. https://doi.org/10.1111/nyas.12360

Arnone, D., McIntosh, A. M. M., Ebmeier, K. P. P., Munafò, M. R. R., & Anderson, I. M. M. (2012). Magnetic resonance imaging studies in unipolar depression: Systematic review and meta-regression analyses. *European Neuropsychopharmacology*, *22*(1), 1–16. https://doi.org/10.1016/j.euroneuro.2011.05.003

Arroll, B., Goodyear-Smith, F., Crengle, S., Gunn, J., Kerse, N., Fishman, T., Falloon, K., & Hatcher, S. (2010). Validation of PHQ-2 and PHQ-9 to screen for major depression in the primary care population. *Annals of Family Medicine*, *8*(4), 348–353. https://doi.org/10.1370/afm.1139

Auchter, A. M., Hernandez Mejia, M., Heyser, C. J., Shilling, P. D., Jernigan, T. L., Brown, S. A., Tapert, S. F., & Dowling, G. J. (2018). A description of the ABCD organizational structure and communication framework. *Developmental Cognitive Neuroscience*, *32*(November 2017), 8–15. https://doi.org/10.1016/j.dcn.2018.04.003

Avenevoli, S., Blader, J. C., & Leibenluft, E. (2015). Irritability in Youth: An Update. *Journal of the American Academy of Child & Adolescent Psychiatry*, *54*(11), 881–883. https://doi.org/10.1016/j.jaac.2015.08.012

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature*, *533*(7604), Article 7604. https://doi.org/10.1038/533452a

Barch, D. M., Albaugh, M. D., Avenevoli, S., Chang, L., Clark, D. B., Glantz, M. D., Hudziak, J. J.,

Jernigan, T. L., Tapert, S. F., Yurgelun-Todd, D., Alia-Klein, N., Potter, A. S., Paulus, M. P., Prouty, D., Zucker, R. A., & Sher, K. J. (2018a). Demographic, physical and mental health assessments in the adolescent brain and cognitive development study: Rationale and description. *Developmental Cognitive Neuroscience*, *32*(September 2017), 55–66. https://doi.org/10.1016/j.dcn.2017.10.010

Barch, D. M., Albaugh, M. D., Avenevoli, S., Chang, L., Clark, D. B., Glantz, M. D., Hudziak, J. J., Jernigan, T. L., Tapert, S. F., Yurgelun-Todd, D., Alia-Klein, N., Potter, A. S., Paulus, M. P., Prouty, D., Zucker, R. A., & Sher, K. J. (2018b). Demographic, physical and mental health assessments in the adolescent brain and cognitive development study: Rationale and description. In *Developmental Cognitive Neuroscience* (Vol. 32, pp. 55–66). Elsevier Ltd. https://doi.org/10.1016/j.dcn.2017.10.010

Barch, D. M., Albaugh, M. D., Baskin-Sommers, A., Bryant, B. E., Clark, D. B., Dick, A. S., Feczko, E., Foxe, J. J., Gee, D. G., Giedd, J., Glantz, M. D., Hudziak, J. J., Karcher, N. R., LeBlanc, K., Maddox, M., McGlade, E. C., Mulford, C., Nagel, B. J., Neigh, G., … Xie, L. (2021). Demographic and mental health assessments in the adolescent brain and cognitive development study: Updates and age-related trajectories. *Developmental Cognitive Neuroscience*, *52*, 101031. https://doi.org/10.1016/j.dcn.2021.101031

Barendse, M. E. A., Byrne, M. L., Flournoy, J. C., McNeilly, E. A., Guazzelli Williamson, V., Barrett, A.-M. Y., Chavez, S. J., Shirtcliff, E. A., Allen, N. B., & Pfeifer, J. H. (2021). Multimethod assessment of pubertal timing and associations with internalizing psychopathology in early adolescent girls. *Journal of Psychopathology and Clinical Science*, *131*(1), 14. https://doi.org/10.1037/abn0000721

Bassett, D. S., & Sporns, O. (2017). Network neuroscience. *Nature Neuroscience*, *20*(3), Article

3. https://doi.org/10.1038/nn.4502

Baum, G. L., Roalf, D. R., Cook, P. A., Ciric, R., Rosen, A. F. G., Xia, C., Elliott, M. A., Ruparel, K., Verma, R., Tunc, B., Gur, R. C., Gur, R. E., Bassett, D. S., & Satterthwaite, T. D. (2018). The impact of in-scanner head motion on structural connectivity derived from diffusion MRI. *NeuroImage*, *173*, 275–286. https://doi.org/10.1016/j.neuroimage.2018.02.041

Beard, C., Hsu, K. J., Rifkin, L. S., Busch, A. B., & Björgvinsson, T. (2016). Validation of the PHQ-9 in a psychiatric sample. *Journal of Affective Disorders*, *193*, 267–273. https://doi.org/10.1016/j.jad.2015.12.075

Beaulieu, C. (2002). The basis of anisotropic water diffusion in the nervous system – a technical review. *NMR in Biomedicine*, *15*(7–8), 435–455. https://doi.org/10.1002/nbm.782

Becht, A. I., & Mills, K. L. (2020). Modeling Individual Differences in Brain Development. *Biological Psychiatry*, *88*(1), 63–69. https://doi.org/10.1016/j.biopsych.2020.01.027

Benjamini, Y., & Yekutieli, D. (2001). The Control of the False Discovery Rate in Multiple Testing under Dependency. *The Annals of Statistics*, *29*(4), 1165–1188.

Bennett, R. H., Somandepalli, K., Roy, A. K., & Di Martino, A. (2017). The Neural Correlates of Emotional Lability in Children with Autism Spectrum Disorder. *Brain Connectivity*, *7*(5), 281–288. https://doi.org/10.1089/brain.2016.0472

Bertha, E. A., & Balázs, J. (2013). Subthreshold depression in adolescence: A systematic review. *European Child & Adolescent Psychiatry*, *22*(10), 589–603. https://doi.org/10.1007/s00787-013-0411-0

Bessette, K. L., Nave, A. M., Caprihan, A., & Stevens, M. C. (2014). White matter abnormalities

in adolescents with major depressive disorder. *Brain Imaging and Behavior*, *8*(4), 531–541. https://doi.org/10.1007/s11682-013-9274-8

Bethlehem, R. a. I., Seidlitz, J., White, S. R., Vogel, J. W., Anderson, K. M., Adamson, C., Adler, S., Alexopoulos, G. S., Anagnostou, E., Areces-Gonzalez, A., Astle, D. E., Auyeung, B., Ayub, M., Bae, J., Ball, G., Baron-Cohen, S., Beare, R., Bedford, S. A., Benegal, V., … Alexander-Bloch, A. F. (2022). Brain charts for the human lifespan. *Nature*, *604*(7906), Article 7906. https://doi.org/10.1038/s41586-022-04554-y

Bhanji, J. P., & Delgado, M. R. (2014). The social brain and reward: Social information processing in the human striatum. *WIREs Cognitive Science*, *5*(1), 61–73. https://doi.org/10.1002/wcs.1266

Biro, F. M., Greenspan, L. C., Galvez, M. P., Pinney, S. M., Teitelbaum, S., Windham, G. C., Deardorff, J., Herrick, R. L., Succop, P. A., Hiatt, R. A., Kushi, L. H., & Wolff, M. S. (2013). Onset of breast development in a longitudinal cohort. *Pediatrics*, *132*(6), 1019–1027. https://doi.org/10.1542/peds.2012-3773

Biro, F. M., Pinney, S. M., Huang, B., Baker, E. R., Walt Chandler, D., & Dorn, L. D. (2014). Hormone Changes in Peripubertal Girls. *The Journal of Clinical Endocrinology & Metabolism*, *99*(10), 3829–3835. https://doi.org/10.1210/jc.2013-4528

Biswal, B. B., Mennes, M., Zuo, X.-N., Gohel, S., Kelly, C., Smith, S. M., Beckmann, C. F., Adelstein, J. S., Buckner, R. L., Colcombe, S., Dogonowski, A.-M., Ernst, M., Fair, D., Hampson, M., Hoptman, M. J., Hyde, J. S., Kiviniemi, V. J., Kötter, R., Li, S.-J., … Milham, M. P. (2010). Toward discovery science of human brain function. *Proceedings of the National Academy of Sciences*, *107*(10), 4734–4739. https://doi.org/10.1073/pnas.0911855107

Biswal, B., Zerrin Yetkin, F., Haughton, V. M., & Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar mri. *Magnetic Resonance in Medicine*, *34*(4), 537–541. https://doi.org/10.1002/mrm.1910340409

Bjork, J. M., & Pardini, D. A. (2015). Who are those 'risk-taking adolescents'? Individual differences in developmental neuroimaging research. *Developmental Cognitive Neuroscience*, *11*, 56–64. https://doi.org/10.1016/j.dcn.2014.07.008

Black, M. M., & Hurley, K. M. (2014). Investment in early childhood development. *The Lancet*, *384*(9950), 1244–1245. https://doi.org/10.1016/S0140-6736(14)60607-3

Blakemore, S. J. (2008). The social brain in adolescence. *Nature Reviews Neuroscience*, *9*(4), 267–277. https://doi.org/10.1038/nrn2353

Blakemore, S. J. (2018). Avoiding Social Risk in Adolescence. *Current Directions in Psychological Science*, *27*(2), 116–122. https://doi.org/10.1177/0963721417738144

Blakemore, S.-J., & Mills, K. L. (2014). Is Adolescence a Sensitive Period for Sociocultural Processing? *Annual Review of Psychology*, *65*(1), 187–207. https://doi.org/10.1146/annurev-psych-010213-115202

Blanton, R. E., Cooney, R. E., Joormann, J., Eugène, F., Glover, G. H., & Gotlib, I. H. (2012). Pubertal stage and brain anatomy in girls. *Neuroscience*, *217*, 105–112. https://doi.org/10.1016/J.NEUROSCIENCE.2012.04.059

Bordini, B., & Rosenfield, R. L. (2011). Normal pubertal development: The endocrine basis of puberty. *Pediatrics in Review*, *32*(6), 223–229.

Botvinik-Nezer, R., Holzmeister, F., Camerer, C. F., Dreber, A., Huber, J., Johannesson, M., Kirchler, M., Iwanir, R., Mumford, J. A., Adcock, R. A., Avesani, P., Baczkowski, B. M., Bajracharya, A., Bakst, L., Ball, S., Barilari, M., Bault, N., Beaton, D., Beitner, J., …

Schonberg, T. (2020). Variability in the analysis of a single neuroimaging dataset by many teams. *Nature*, *582*(7810), Article 7810. https://doi.org/10.1038/s41586-020-2314-9

Boyce, W. T., Sokolowski, M. B., & Robinson, G. E. (2020). Genes and environments, development and time. *Proceedings of the National Academy of Sciences*, *117*(38), 23235–23241. https://doi.org/10.1073/pnas.2016710117

Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic Status and Child Development. *Annual Review of Psychology*, *53*(1), 371–399. https://doi.org/10.1146/annurev.psych.53.100901.135233

Brasil, H. H. A., & Bordin, I. A. (2010). Convergent validity of K-SADS-PL by comparison with CBCL in a Portuguese speaking outpatient population. *BMC Psychiatry*, *10*(83). https://doi.org/10.1186/1471-244X-10-83

Brooks-Gunn, J., Graber, J. A., & Paikoff, R. L. (1994). Studying links between hormones and negative affect: Models and measures. *Journal of Research on Adolescence*, *4*(4), 469–486. https://doi.org/10.1207/s15327795jra0404_2

Brooks-Gunn, J., Petersen, A. C., & Eichorn, D. (1985). The study of maturational timing effects in adolescence. *Journal of Youth and Adolescence*, *14*(3), 149–161. https://doi.org/10.1007/BF02090316

Brotman, M. A., Kircanski, K., Stringaris, A., Pine, D. S., & Leibenluft, E. (2017). Irritability in Youths: A Translational Model. *American Journal of Psychiatry*, *174*(6), 520–532. https://doi.org/10.1176/appi.ajp.2016.16070839

Broyd, S. J., Demanuele, C., Debener, S., Helps, S. K., James, C. J., & Sonuga-Barke, E. J. S. (2009). Default-mode brain dysfunction in mental disorders: A systematic review.

Neuroscience and Biobehavioral Reviews, *33*(3), 279–296. https://doi.org/10.1016/j.neubiorev.2008.09.002

Burgund, E. D., Kang, H. C., Kelly, J. E., Buckner, R. L., Snyder, A. Z., Petersen, S. E., & Schlaggar, B. L. (2002). The Feasibility of a Common Stereotactic Space for Children and Adults in fMRI Studies of Development. *NeuroImage*, *17*(1), 184–200. https://doi.org/10.1006/nimg.2002.1174

Burleson Daviss, W., Birmaher, B., Melhem, N. A., Axelson, D. A., Michaels, S. M., & Brent, D. A. (2006). Criterion validity of the Mood and Feelings Questionnaire for depressive episodes in clinic and non-clinic subjects. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *47*(9), 927–934. https://doi.org/10.1111/j.1469-7610.2006.01646.x

Buuren, S. van, & Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, *45*, 1–67. https://doi.org/10.18637/jss.v045.i03

Cabral, J., Kringelbach, M. L., & Deco, G. (2017). Functional connectivity dynamically evolves on multiple time-scales over a static structural connectome: Models and mechanisms. *NeuroImage*, *160*, 84–96. https://doi.org/10.1016/j.neuroimage.2017.03.045

Cabral, J., Vidaurre, D., Marques, P., Magalhães, R., Silva Moreira, P., Miguel Soares, J., Deco, G., Sousa, N., & Kringelbach, M. L. (2017). Cognitive performance in healthy older adults relates to spontaneous switching between states of functional connectivity during rest. *Scientific Reports*, *7*(1), 5135. https://doi.org/10.1038/s41598-017-05425-7

Calhoun, V. D., Miller, R., Pearlson, G., & Adalı, T. (2014). The Chronnectome: Time-Varying

Connectivity Networks as the Next Frontier in fMRI Data Discovery. *Neuron*, *84*(2), 262–274. https://doi.org/10.1016/j.neuron.2014.10.015

Callaghan, B. L., & Tottenham, N. (2016). The Stress Acceleration Hypothesis: Effects of early-life adversity on emotion circuits and behavior. *Current Opinion in Behavioral Sciences*, *7*, 76–81. https://doi.org/10.1016/j.cobeha.2015.11.018

Cano-López, J. B., García-Sancho, E., Fernández-Castilla, B., & Salguero, J. M. (2022). Empirical Evidence of the Metacognitive Model of Rumination and Depression in Clinical and Nonclinical Samples: A Systematic Review and Meta-Analysis. *Cognitive Therapy and Research*, *46*(2), 367–392. https://doi.org/10.1007/s10608-021-10260-2

Casey, B. J., Cannonier, T., Conley, M. I., Cohen, A. O., Barch, D. M., Heitzeg, M. M., Soules, M. E., Teslovich, T., Dellarco, D. V., Garavan, H., Orr, C. A., Wager, T. D., Banich, M. T., Speer, N. K., Sutherland, M. T., Riedel, M. C., Dick, A. S., Bjork, J. M., Thomas, K. M., … Dale, A. M. (2018a). The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. *Developmental Cognitive Neuroscience*, *32*(January), 43–54. https://doi.org/10.1016/j.dcn.2018.03.001

Casey, B. J., Cannonier, T., Conley, M. I., Cohen, A. O., Barch, D. M., Heitzeg, M. M., Soules, M. E., Teslovich, T., Dellarco, D. V., Garavan, H., Orr, C. A., Wager, T. D., Banich, M. T., Speer, N. K., Sutherland, M. T., Riedel, M. C., Dick, A. S., Bjork, J. M., Thomas, K. M., … Dale, A. M. (2018b). The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. In *Developmental Cognitive Neuroscience* (Vol. 32, pp. 43–54). Elsevier Ltd. https://doi.org/10.1016/j.dcn.2018.03.001

Casey, B. J., Getz, S., & Galvan, A. (2008). The adolescent brain. *Developmental Review*, *28*(1), 62–77. https://doi.org/10.1016/j.dr.2007.08.003

Chaarani, B., Kan, K.-J., Mackey, S., Spechler, P. A., Potter, A., Banaschewski, T., Millenet, S., Bokde, A. L. W., Bromberg, U., Büchel, C., Cattrell, A., Conrod, P. J., Desrivières, S., Flor, H., Frouin, V., Gallinat, J., Gowland, P., Heinz, A., Ittermann, B., … Stedman, A. (2020). Neural Correlates of Adolescent Irritability and Its Comorbidity With Psychiatric Disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*, *59*(12), 1371–1379. https://doi.org/10.1016/j.jaac.2019.11.028

Chen, C.-H., Gutierrez, E. D., Thompson, W., Panizzon, M. S., Jernigan, T. L., Eyler, L. T., Fennema-Notestine, C., Jak, A. J., Neale, M. C., Franz, C. E., Lyons, M. J., Grant, M. D., Fischl, B., Seidman, L. J., Tsuang, M. T., Kremen, W. S., & Dale, A. M. (2012). Hierarchical Genetic Organization of Human Cortical Surface Area. *Science*, *335*(6076), 1634–1636. https://doi.org/10.1126/science.1215330

Chen, J. E., Rubinov, M., & Chang, C. (2017). Methods and considerations for dynamic analysis of fMRI data. *Neuroimaging Clinics of North America*, *27*(4), 547–560. https://doi.org/10.1016/j.nic.2017.06.009

Cheng, T. W., Magis-Weinberg, L., Guazzelli Williamson, V., Ladouceur, C. D., Whittle, S. L., Herting, M. M., Uban, K. A., Byrne, M. L., Barendse, M. E. A., Shirtcliff, E. A., & Pfeifer, J. H. (2021). A Researcher's Guide to the Measurement and Modeling of Puberty in the ABCD Study® at Baseline. *Frontiers in Endocrinology*, *12*, 471. https://doi.org/10.3389/fendo.2021.608575

Chorot, P., Valiente, R. M., Magaz, A. M., Santed, M. A., & Sandin, B. (2017). Perceived parental child rearing and attachment as predictors of anxiety and depressive disorder symptoms in children: The mediational role of attachment. *Psychiatry Research*, *253*, 287–295. https://doi.org/10.1016/J.PSYCHRES.2017.04.015

Chumlea, W. C., Schubert, C. M., Roche, A. F., Kulin, H. E., Lee, P. A., Himes, J. H., & Sun, S. S. (2003). Age at menarche and racial comparisons in US girls. *Pediatrics*, *111*(1), 110–113. https://doi.org/10.1542/peds.111.1.110

Clayborne, Z. M., Varin, M., & Colman, I. (2019). Systematic Review and Meta-Analysis: Adolescent Depression and Long-Term Psychosocial Outcomes. *Journal of the American Academy of Child & Adolescent Psychiatry*, *58*(1), 72–79. https://doi.org/10.1016/j.jaac.2018.07.896

Cole, M. W., Reynolds, J. R., Power, J. D., Repovs, G., Anticevic, A., & Braver, T. S. (2013). Multi-task connectivity reveals flexible hubs for adaptive task control. *Nature Neuroscience*, *16*(9), 1348–1355. https://doi.org/10.1038/nn.3470

Colich, N. L., & McLaughlin, K. A. (2022). Accelerated pubertal development as a mechanism linking trauma exposure with depression and anxiety in adolescence. *Current Opinion in Psychology*, *46*, 101338. https://doi.org/10.1016/j.copsyc.2022.101338

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. *Psychological Bulletin*, *146*(9), 721–764. https://doi.org/10.1037/bul0000270

Collishaw, S., Hammerton, G., Mahedy, L., Sellers, R., Owen, M. J., Craddock, N., Thapar, A. K., Harold, G. T., Rice, F., & Thapar, A. (2016). Mental health resilience in the adolescent offspring of parents with depression: A prospective longitudinal study. *The Lancet Psychiatry*, *3*(1), 49–57. https://doi.org/10.1016/S2215-0366(15)00358-2

Conley, C. S., Rudolph, K. D., & Bryant, F. B. (2012). Explaining the Longitudinal Association between Puberty and Depression: Sex Differences in the Mediating Effects of Peer

Stress. *Developmental Psychopathology*, *24*(2), 691–701. https://doi.org/10.1017/S0954579412000259

Corbetta, M., Patel, G., & Shulman, G. L. (2008). The Reorienting System of the Human Brain: From Environment to Theory of Mind. *Neuron*, *58*(3), 306–324. https://doi.org/10.1016/j.neuron.2008.04.017

Costello, E. J., & Angold, A. (1988). Scales to assess child and adolescent depression: Checklists, screens, and nets. *Journal of the American Academy of Child and Adolescent Psychiatry*, *27*(6), 726–737. https://doi.org/10.1097/00004583-198811000-00011

Cox, S. R., Ritchie, S. J., Tucker-Drob, E. M., Liewald, D. C., Hagenaars, S. P., Davies, G., Wardlaw, J. M., Gale, C. R., Bastin, M. E., & Deary, I. J. (2016). Ageing and brain white matter structure in 3,513 UK Biobank participants. *Nature Communications*, *7*, 1–34. https://doi.org/10.1038/ncomms13629

Crone, E. A., & Dahl, R. E. (2012). Understanding adolescence as a period of social-affective engagement and goal flexibility. *Nature Reviews. Neuroscience*, *13*(9), 636–650. https://doi.org/10.1038/nrn3313

Cullen, K. R., Brown, R., Schreiner, M. W., Eberly, L. E., Klimes-Dougan, B., Reigstad, K., Hill, D., Lim, K. O., & Mueller, B. A. (2019). White matter microstructure relates to lassitude but not diagnosis in adolescents with depression. *Brain Imaging and Behavior*, 1–14. https://doi.org/10.1007/s11682-019-00078-8

Dahl, R. E., Allen, N. B., Wilbrecht, L., & Suleiman, A. B. (2018). Importance of investing in adolescence from a developmental science perspective. *Nature*, *554*(7693), 441–450. https://doi.org/10.1038/nature25770

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical Surface-Based Analysis: I. Segmentation and Surface Reconstruction. *NeuroImage*, *9*(2), 179–194. https://doi.org/10.1006/nimg.1998.0395

Dalgleish, T. (2004). The emotional brain. *Nature Reviews Neuroscience*, *5*(7), 582–589. https://doi.org/10.1038/nrn1432

Daly, M. (2022). Prevalence of Depression Among Adolescents in the U.S. From 2009 to 2019: Analysis of Trends by Sex, Race/Ethnicity, and Income. *Journal of Adolescent Health*, *70*(3), 496–499. https://doi.org/10.1016/j.jadohealth.2021.08.026

De Los Reyes, A. (2011). Introduction to the special section: More than measurement error: Discovering meaning behind informant discrepancies in clinical assessments of children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, *40*(1), 1–9. https://doi.org/10.1080/15374416.2011.533405

De Los Reyes, A., Augenstein, T. M., Wang, M., Thomas, S. A., Drabick, D. A. G., Burgers, D. E., & Rabinowitz, J. (2015). The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychological Bulletin*, *141*(4), 858–900. https://doi.org/10.1037/a0038498

De Los Reyes, A., Thomas, S. A., Goodman, K. L., & Kundey, S. M. A. (2013). Principles Underlying the Use of Multiple Informants' Reports. *Annual Review of Clinical Psychology*, *9*(1), 123–149. https://doi.org/10.1146/annurev-clinpsy-050212-185617

Deco, G., & Kringelbach, M. L. (2016). Metastability and Coherence: Extending the Communication through Coherence Hypothesis Using A Whole-Brain Computational Perspective. *Trends in Neurosciences*, *39*(3), 125–135. https://doi.org/10.1016/j.tins.2016.01.001

Defoe, I. N., Dubas, J. S., Figner, B., & Van Aken, M. A. G. (2015). A meta-analysis on age differences in risky decision making: Adolescents versus children and adults. *Psychological Bulletin*, *141*(1), 48–84. https://doi.org/10.1037/a0038088

Dennison, M., Whittle, S., Yücel, M., Vijayakumar, N., Kline, A., Simmons, J., & Allen, N. B. (2013). Mapping subcortical brain maturation during adolescence: Evidence of hemisphere- and sex-specific longitudinal changes. *Developmental Science*, *16*(5), 772–791. https://doi.org/10.1111/desc.12057

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., Albert, M. S., & Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, *31*(3), 968–980. https://doi.org/10.1016/j.neuroimage.2006.01.021

Destrieux, C., Fischl, B., Dale, A., & Halgren, E. (2010). Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *NeuroImage*, *53*(1), 1–15. https://doi.org/10.1016/j.neuroimage.2010.06.010

Deveney, C. M., Connolly, M. E., Haring, C. T., Bones, B. L., Reynolds, R. C., Kim, P., Pine, D. S., & Leibenluft, E. (2013). Neural mechanisms of frustration in chronically irritable children. *The American Journal of Psychiatry*, *170*(10), 1186–1194. https://doi.org/10.1176/appi.ajp.2013.12070917

Dick, A. S., Lopez, D. A., Watts, A. L., Heeringa, S., Reuter, C., Bartsch, H., Fan, C. C., Kennedy, D. N., Palmer, C., Marshall, A., Haist, F., Hawes, S., Nichols, T. E., Barch, D. M., Jernigan, T. L., Garavan, H., Grant, S., Pariyadath, V., Hoffman, E., … Thompson, W. K. (2021). Meaningful Associations in the Adolescent Brain Cognitive Development Study.

NeuroImage, 118262. https://doi.org/10.1016/j.neuroimage.2021.118262

Dorn, L. D., & Biro, F. M. (2011). Puberty and Its Measurement: A Decade in Review. *Journal of Research on Adolescence*, *21*(1), 180–195. https://doi.org/10.1111/j.1532-7795.2010.00722.x

Dorn, L. D., Dahl, R. E., Woodward, H. R., & Biro, F. (2006). Defining the Boundaries of Early Adolescence: A User's Guide to Assessing Pubertal Status and Pubertal Timing in Research With Adolescents. *Applied Developmental Science*, *10*(1), 30–56. https://doi.org/10.1207/s1532480xads1001_3

Dorn, L. D., Nottelmann, E. D., Inoff-Germain, G., Susman, E. J., & Chrousos, G. P. (1990). Perceptions of Puberty: Adolescent, Parent, and Health Care Personnel. *Developmental Psychology*, *26*(2), 322–329.

Ducharme, S., Albaugh, M. D., Nguyen, T. V., Hudziak, J. J., Mateos-Pérez, J. M., Labbe, A., Evans, A. C., Karama, S., Ball, W. S., Byars, A. W., Schapiro, M., Bommer, W., Carr, A., German, A., Dunn, S., Rivkin, M. J., Waber, D., Mulkern, R., Vajapeyam, S., … O'Neill, J. (2016). Trajectories of cortical thickness maturation in normal brain development—The importance of quality control procedures. *NeuroImage*, *125*, 267–279. https://doi.org/10.1016/j.neuroimage.2015.10.010

Edde, M., Leroux, G., Altena, E., & Chanraud, S. (2021). Functional brain connectivity changes across the human life span: From fetal development to old age. *Journal of Neuroscience Research*, *99*(1), 236–262. https://doi.org/10.1002/jnr.24669

Eg, J., Bilenberg, N., Costello, E. J., & Wesselhoeft, R. (2018). Self- and parent-reported depressive symptoms rated by the mood and feelings questionnaire. *Psychiatry Research*, *268*, 419–425. https://doi.org/10.1016/j.psychres.2018.07.016

Ernst, M., Torrisi, S., Balderston, N., Grillon, C., & Hale, E. A. (2015). FMRI Functional Connectivity Applied to Adolescent Neurodevelopment. *Annual Review of Clinical Psychology*, *11*, 361–377. https://doi.org/10.1146/annurev-clinpsy-032814-112753

Eshel, N., & Leibenluft, E. (2020). New Frontiers in Irritability Research—From Cradle to Grave and Bench to Bedside. *JAMA Psychiatry*, *77*(3), 227. https://doi.org/10.1001/jamapsychiatry.2019.3686

Esteban, O., Birman, D., Schaer, M., Koyejo, O. O., Poldrack, R. A., & Gorgolewski, K. J. (2017). MRIQC: Advancing the automatic prediction of image quality in MRI from unseen sites. *PLOS ONE*, *12*(9), e0184661. https://doi.org/10.1371/journal.pone.0184661

Esteban, O., Ciric, R., Finc, K., Blair, R. W., Markiewicz, C. J., Moodie, C. A., Kent, J. D., Goncalves, M., DuPre, E., Gomez, D. E. P., Ye, Z., Salo, T., Valabregue, R., Amlien, I. K., Liem, F., Jacoby, N., Stojić, H., Cieslak, M., Urchs, S., … Gorgolewski, K. J. (2020). Analysis of task-based functional MRI data preprocessed with fMRIPrep. *Nature Protocols*, *15*(7), Article 7. https://doi.org/10.1038/s41596-020-0327-3

Fair, D. A., Cohen, A. L., Power, J. D., Dosenbach, N. U. F., Church, J. A., Miezin, F. M., Schlaggar, B. L., & Petersen, S. E. (2009). Functional Brain Networks Develop from a "Local to Distributed" Organization. *PLOS Computational Biology*, *5*(5), e1000381. https://doi.org/10.1371/journal.pcbi.1000381

Fareri, D. S., Gabard-Durnam, L., Goff, B., Flannery, J., Gee, D. G., Lumian, D. S., Caldera, C., & Tottenham, N. (2015). Normative development of ventral striatal resting state connectivity in humans. *NeuroImage*, *118*, 422–437. https://doi.org/10.1016/j.neuroimage.2015.06.022

Farinha, M., Amado, C., Morgado, P., & Cabral, J. (2022). Increased Excursions to Functional

Networks in Schizophrenia in the Absence of Task. *Frontiers in Neuroscience*, *16*, 821179. https://doi.org/10.3389/fnins.2022.821179

Feaster, D. J., Mikulich-Gilbertson, S., & Brincks, A. M. (2011). Modeling site effects in the design and analysis of multisite trials. *The American Journal of Drug and Alcohol Abuse*, *37*(5), 383–391. https://doi.org/10.3109/00952990.2011.600386

Feldstein Ewing, S. W., Bjork, J. M., & Luciana, M. (2018). Implications of the ABCD study for developmental neuroscience. *Developmental Cognitive Neuroscience*, *32*, 161–164. https://doi.org/10.1016/j.dcn.2018.05.003

Feldstein Ewing, S. W., Dash, G. F., Thompson, W. K., Reuter, C., Diaz, V. G., Anokhin, A., Chang, L., Cottler, L. B., Dowling, G. J., LeBlanc, K., Zucker, R. A., Tapert, S. F., Brown, S. A., & Garavan, H. (2022). Measuring retention within the adolescent brain cognitive development (ABCD)SM study. *Developmental Cognitive Neuroscience*, *54*, 101081. https://doi.org/10.1016/j.dcn.2022.101081

Fergusson, D. M., & Woodward, L. J. (2002). Mental health, educational, and social role outcomes of adolescents with depression. *Archives of General Psychiatry*, *59*(3), 225–231. https://doi.org/10.1001/archpsyc.59.3.225

Ferrucci, L., Gonzalez-Freire, M., Fabbri, E., Simonsick, E., Tanaka, T., Moore, Z., Salimi, S., Sierra, F., & de Cabo, R. (2020). Measuring biological aging in humans: A quest. *Aging Cell*, *19*(2), e13080. https://doi.org/10.1111/acel.13080

Ferschmann, L., Bos, M. G. N., Herting, M. M., Mills, K. L., & Tamnes, C. K. (2022). Contextualizing adolescent structural brain development: Environmental determinants and mental health outcomes. *Current Opinion in Psychology*, *44*, 170–176. https://doi.org/10.1016/j.copsyc.2021.09.014

Figueroa, C. A., Cabral, J., Mocking, R. J. T., Rapuano, K. M., van Hartevelt, T. J., Deco, G., Expert, P., Schene, A. H., Kringelbach, M. L., & Ruhé, H. G. (2019). Altered ability to access a clinically relevant control network in patients remitted from major depressive disorder. *Human Brain Mapping*, *40*(9), 2771–2786. https://doi.org/10.1002/hbm.24559

Finn, E. S. (2021). Is it time to put rest to rest? *Trends in Cognitive Sciences*, *25*(12), 1021–1032. https://doi.org/10.1016/j.tics.2021.09.005

Fischl, B., & Dale, A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Sciences*, *97*(20), 11050–11055. https://doi.org/10.1073/pnas.200033797

Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., & Dale, A. M. (2002). Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron*, *33*(3), 341–355. https://doi.org/10.1016/s0896-6273(02)00569-x

Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical Surface-Based Analysis: II: Inflation, Flattening, and a Surface-Based Coordinate System. *NeuroImage*, *9*(2), 195–207. https://doi.org/10.1006/nimg.1998.0396

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D. H., Busa, E., Seidman, L. J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., & Dale, A. M. (2004). Automatically parcellating the human cerebral cortex. *Cerebral Cortex (New York, N.Y.: 1991)*, *14*(1), 11–22. https://doi.org/10.1093/cercor/bhg087

Foulkes, Lucy. (2022). *What Mental Illness Really Is... (And what it isn't)*. Vintage.

Frangou, S., Modabbernia, A., Williams, S. C. R., Papachristou, E., Doucet, G. E., Agartz, I., Aghajani, M., Akudjedu, T. N., Albajes-Eizagirre, A., Alnæs, D., Alpert, K. I., Andersson, M., Andreasen, N. C., Andreassen, O. A., Asherson, P., Banaschewski, T., Bargallo, N., Baumeister, S., Baur-Streubel, R., … Dima, D. (2022). Cortical thickness across the lifespan: Data from 17,075 healthy individuals aged 3–90 years. *Human Brain Mapping*, *43*(1), 431–451. https://doi.org/10.1002/hbm.25364

Fried, E. I. (2015). Problematic assumptions have slowed down depression research: Why symptoms, not syndromes are the way forward. *Frontiers in Psychology*, *6*, 309–309. https://doi.org/10.3389/fpsyg.2015.00309

Fried, E. I., Flake, J. K., & Robinaugh, D. J. (2022). Revisiting the theoretical and methodological foundations of depression measurement. *Nature Reviews Psychology*, 1–11. https://doi.org/10.1038/s44159-022-00050-2

Fried, E. I., & Nesse, R. M. (2015). Depression is not a consistent syndrome: An investigation of unique symptom patterns in the STAR∗D study. *Journal of Affective Disorders*, *172*, 96–102. https://doi.org/10.1016/j.jad.2014.10.010

Friston, K. J. (2011). Functional and Effective Connectivity: A Review. *Brain Connectivity*, *1*(1), 13–36. https://doi.org/10.1089/brain.2011.0008

Friston, Karl. J., Ashburner, J., Frith, C. D., Poline, J.-B., Heather, J. D., & Frackowiak, R. S. J. (1995). Spatial registration and normalization of images. *Human Brain Mapping*, *3*(3), 165–189. https://doi.org/10.1002/hbm.460030303

Fry, A., Littlejohns, T. J., Sudlow, C., Doherty, N., Adamska, L., Sprosen, T., Collins, R., & Allen, N. E. (2017). Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. *American Journal of*

Epidemiology, *186*(9), 1026–1034. https://doi.org/10.1093/aje/kwx246

Fulwiler, C. E., King, J. A., & Zhang, N. (2012). Amygdala–orbitofrontal resting-state functional connectivity is associated with trait anger. *NeuroReport*, *23*(10), 606–610.

Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science*, *2*, 156–168. https://doi.org/10.1177/2515245919847202

Gaffrey, M. S., Barch, D. M., Luby, J. L., & Petersen, S. E. (2021). Amygdala Functional Connectivity Is Associated With Emotion Regulation and Amygdala Reactivity in 4- to 6-Year-Olds. *Journal of the American Academy of Child & Adolescent Psychiatry*, *60*(1), 176–185. https://doi.org/10.1016/j.jaac.2020.01.024

Galván, A. (2021). Adolescent Brain Development and Contextual Influences: A Decade in Review. *Journal of Research on Adolescence*, *31*(4), 843–869. https://doi.org/10.1111/jora.12687

Garavan, H., Bartsch, H., Conway, K., Decastro, A., Goldstein, R. Z., Heeringa, S., Jernigan, T., Potter, A., Thompson, W., & Zahs, D. (2018a). Recruiting the ABCD sample: Design considerations and procedures. *Developmental Cognitive Neuroscience*, *32*(August 2017), 16–22. https://doi.org/10.1016/j.dcn.2018.04.004

Garavan, H., Bartsch, H., Conway, K., Decastro, A., Goldstein, R. Z., Heeringa, S., Jernigan, T., Potter, A., Thompson, W., & Zahs, D. (2018b). Recruiting the ABCD sample: Design considerations and procedures. *Developmental Cognitive Neuroscience*, *32*, 16–22. https://doi.org/10.1016/j.dcn.2018.04.004

Gardner, W., Mulvey, E. P., & Shaw, E. C. (1995). Regression analyses of counts and rates: Poisson, overdispersed Poisson, and negative binomial models. *Psychological Bulletin*,

118(3), 392–404. https://doi.org/10.1037/0033-2909.118.3.392

Gavrilescu, M., Shaw, M. E., Stuart, G. W., Eckersley, P., Svalbe, I. D., & Egan, G. F. (2002). Simulation of the Effects of Global Normalization Procedures in Functional MRI. *NeuroImage*, *17*(2), 532–542. https://doi.org/10.1006/nimg.2002.1226

Ge, X., Conger, R. D., & Elder Jr., G. H. (2001). Pubertal transition, stressful life events, and the emergence of gender differences in adolescent depressive symptoms. *Developmental Psychology*, *37*(3), 404–417. https://doi.org/10.1037/0012-1649.37.3.404

Ge, X., & Natsuaki, M. N. (2009). In search of explanations for early pubertal timing effects on developmental psychopathology. *Current Directions in Psychological Science*, *18*(6), 327–331. https://doi.org/10.1111/j.1467-8721.2009.01661.x

Gelman, A., & Hill, J. (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.

Ghosh, S. S., Kakunoori, S., Augustinack, J., Nieto-Castanon, A., Kovelman, I., Gaab, N., Christodoulou, J. A., Triantafyllou, C., Gabrieli, J. D. E., & Fischl, B. (2010). Evaluating the validity of volume-based and surface-based brain image registration for developmental cognitive neuroscience studies in children 4 to 11years of age. *NeuroImage*, *53*(1), 85–93. https://doi.org/10.1016/j.neuroimage.2010.05.075

Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., Paus, T., Evans, A. C., & Rapoport, J. L. (1999). Brain development during childhood and adolescence: A longitudinal MRI study. *Nature Neuroscience*, *2*(10), Article 10. https://doi.org/10.1038/13158

Gillan, Claire. M., & Rutledge, Robb. B. (2021). Smartphones and the Neuroscience of Mental Health. *Annual Review of Neuroscience*, *44*, 129–151.

https://doi.org/10.1146/annurev-neuro-101220-014053

Glerean, E., Salmi, J., Lahnakoski, J. M., Jääskeläinen, I. P., & Sams, M. (2012). Functional Magnetic Resonance Imaging Phase Synchronization as a Measure of Dynamic Functional Connectivity. *Brain Connectivity*, *2*(2), 91–101. https://doi.org/10.1089/brain.2011.0068

Global Burden of Diseases. (2018). Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: A systematic analysis for the Global Burden of Disease Study 2017. *Lancet (London, England)*, *392*(10159), 1789–1858. https://doi.org/10.1016/S0140-6736(18)32279-7

Goddings, A.-L., Beltz, A., Peper, J. S., Crone, E. A., & Braams, B. R. (2019). Understanding the Role of Puberty in Structural and Functional Development of the Adolescent Brain. *Journal of Research on Adolescence*, *29*(1), 32–53. https://doi.org/10.1111/jora.12408

Goddings, A.-L., Mills, K. L., Clasen, L. S., Giedd, J. N., Viner, R. M., & Blakemore, S.-J. (2014). The influence of puberty on subcortical brain development. *NeuroImage*, *88*, 242–251. https://doi.org/10.1016/j.neuroimage.2013.09.073

Goddings, A.-L., Roalf, D., Lebel, C., & Tamnes, C. K. (2021). Development of white matter microstructure and executive functions during childhood and adolescence: A review of diffusion MRI studies. *Developmental Cognitive Neuroscience*, *51*, 101008. https://doi.org/10.1016/j.dcn.2021.101008

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., Nugent, T. F., Herman, D. H., Clasen, L. S., Toga, A. W., Rapoport, J. L., & Thompson, P. M. (2004). Dynamic mapping of human cortical development during childhood through early

adulthood. *Proceedings of the National Academy of Sciences*, *101*(21), 8174–8179. https://doi.org/10.1073/pnas.0402680101

Gorgolewski, K. J., Auer, T., Calhoun, V. D., Craddock, R. C., Das, S., Duff, E. P., Flandin, G., Ghosh, S. S., Glatard, T., Halchenko, Y. O., Handwerker, D. A., Hanke, M., Keator, D., Li, X., Michael, Z., Maumet, C., Nichols, B. N., Nichols, T. E., Pellman, J., … Poldrack, R. A. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data*, *3*(1), Article 1. https://doi.org/10.1038/sdata.2016.44

Graber, J. A. (2013). Pubertal timing and the development of psychopathology in adolescence and beyond. *Hormones and Behavior*, *64*(2), 262–269. https://doi.org/10.1016/J.YHBEH.2013.04.003

Grasby, K. L., Jahanshad, N., Painter, J. N., Colodro-Conde, L., Bralten, J., Hibar, D. P., Lind, P. A., Pizzagalli, F., Ching, C. R. K., McMahon, M. A. B., Shatokhina, N., Zsembik, L. C. P., Thomopoulos, S. I., Zhu, A. H., Strike, L. T., Agartz, I., Alhusaini, S., Almeida, M. A. A., Alnæs, D., … Medland, S. E. (2020). The genetic architecture of the human cerebral cortex. *Science*, *367*(6484). https://doi.org/10.1126/science.aay6690

Grüning, B., Chilton, J., Köster, J., Dale, R., Soranzo, N., van den Beek, M., Goecks, J., Backofen, R., Nekrutenko, A., & Taylor, J. (2018). Practical Computational Reproducibility in the Life Sciences. *Cell Systems*, *6*(6), 631–635. https://doi.org/10.1016/j.cels.2018.03.014

Gur, R. E., Moore, T. M., Rosen, A. F. G., Barzilay, R., Roalf, D. R., Calkins, M. E., Ruparel, K., Scott, J. C., Almasy, L., Satterthwaite, T. D., Shinohara, R. T., & Gur, R. C. (2019). Burden of Environmental Adversity Associated With Psychopathology, Maturation, and Brain Behavior Parameters in Youths. *JAMA Psychiatry*, *76*(9), 966–975.

https://doi.org/10.1001/jamapsychiatry.2019.0943

Gusnard, D. A., Akbudak, E., Shulman, G. L., & Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: Relation to a default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(7), 4259–4264. https://doi.org/10.1073/pnas.071043098

Gutierrez-Barragan, D., Basson, M. A., Panzeri, S., & Gozzi, A. (2019). Infraslow State Fluctuations Govern Spontaneous fMRI Network Dynamics. *Current Biology*, *29*(14), 2295-2306.e5. https://doi.org/10.1016/j.cub.2019.06.017

Hadders-Algra, M. (2018). Early human motor development: From variation to the ability to vary and adapt. *Neuroscience & Biobehavioral Reviews*, *90*, 411–427. https://doi.org/10.1016/j.neubiorev.2018.05.009

Hagler, D. J., Ahmadi, M. E., Kuperman, J., Holland, D., McDonald, C. R., Halgren, E., & Dale, A. M. (2009). Automated white-matter tractography using a probabilistic diffusion tensor atlas: Application to temporal lobe epilepsy. *Human Brain Mapping*, *30*(5), 1535–1547. https://doi.org/10.1002/hbm.20619

Hagler, D. J., Hatton, SeanN., Cornejo, M. D., Makowski, C., Fair, D. A., Dick, A. S., Sutherland, M. T., Casey, B. J., Barch, D. M., Harms, M. P., Watts, R., Bjork, J. M., Garavan, H. P., Hilmer, L., Pung, C. J., Sicat, C. S., Kuperman, J., Bartsch, H., Xue, F., … Dale, A. M. (2019). Image processing and analysis methods for the Adolescent Brain Cognitive Development Study. *NeuroImage*, *202*, 116091. https://doi.org/10.1016/j.neuroimage.2019.116091

Hagler, D. J., Thompson, W. K., Chen, C.-H., Reuter, C., Akshoomoff, N., & Brown, T. T. (2022). Do aggregate, multimodal structural neuroimaging measures replicate regional

developmental differences observed in highly cited cellular histological studies? *Developmental Cognitive Neuroscience*, *54*, 101086. https://doi.org/10.1016/j.dcn.2022.101086

Hakulinen, C., Elovainio, M., Pulkki-Råback, L., Virtanen, M., Kivimäki, M., & Jokela, M. (2015). Personality and Depressive Symptoms: Individual Participant Meta-Analysis of 10 Cohort Studies. *Depression and Anxiety*, *32*(7), 461–470. https://doi.org/10.1002/da.22376

Hallquist, M. N., & Wiley, J. F. (2018). MplusAutomation: An R Package for Facilitating Large-Scale Latent Variable Analyses in Mplus. *Structural Equation Modeling : A Multidisciplinary Journal*, *25*(4), 621–638. https://doi.org/10.1080/10705511.2017.1402334

Hamilton, J. L., Hamlat, E. J., Stange, J. P., Abramson, L. Y., & Alloy, L. B. (2014). Pubertal timing and vulnerabilities to depression in early adolescence: Differential pathways to depressive symptoms by sex. *Journal of Adolescence*, *37*(2), 165–174. https://doi.org/10.1016/j.adolescence.2013.11.010

Hamilton, J. P., Farmer, M., Fogelman, P., & Gotlib, I. H. (2015). Depressive Rumination, the Default-Mode Network, and the Dark Matter of Clinical Neuroscience. *Biological Psychiatry*, *78*(4), 224–230. https://doi.org/10.1016/j.biopsych.2015.02.020

Hamlat, E. J., McCormick, K. C., Young, J. F., & Hankin, B. L. (2020). Early pubertal timing predicts onset and recurrence of depressive episodes in boys and girls. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, jcpp.13198-jcpp.13198. https://doi.org/10.1111/jcpp.13198

Hamlat, E. J., Snyder, H. R., Young, J. F., & Hankin, B. L. (2019). Pubertal Timing as a

Transdiagnostic Risk for Psychopathology in Youth. *Clinical Psychological Science*, *7*(3), 411–429. https://doi.org/10.1177/2167702618810518

Handwerker, D. A., Roopchansingh, V., Gonzalez-Castillo, J., & Bandettini, P. A. (2012). Periodic changes in fMRI connectivity. *NeuroImage*, *63*(3), 1712–1719. https://doi.org/10.1016/j.neuroimage.2012.06.078

Hankin, B. L. (2006). Adolescent depression: Description, causes, and interventions. *Epilepsy and Behavior*, *8*(1), 102–114. https://doi.org/10.1016/j.yebeh.2005.10.012

Hankin, B. L. (2015). Depression from childhood through adolescence: Risk mechanisms across multiple systems and levels of analysis. *Current Opinion in Psychology*, *4*, 13–20. https://doi.org/10.1016/j.copsyc.2015.01.003

Hankin, B. L., & Abramson, L. Y. (1999). Development of gender differences in depression: Description and possible explanations. *Annals of Medicine*, *31*(6), 372–379. https://doi.org/10.3109/07853899908998794

Harald, B., & Gordon, P. (2012). Meta-review of depressive subtyping models. *Journal of Affective Disorders*, *139*(2), 126–140. https://doi.org/10.1016/j.jad.2011.07.015

Havelock, J. C., Auchus, R. J., & Rainey, W. E. (2004). The Rise in Adrenal Androgen Biosynthesis: Adrenarche. *Seminars in Reproductive Medicine*, *22*(04), 337–347. https://doi.org/10.1055/s-2004-861550

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical Mediation Analysis in the New Millennium. *Communication Monographs*, *76*(4), 408–420. https://doi.org/10.1080/03637750903310360

Hellyer, P. J., Scott, G., Shanahan, M., Sharp, D. J., & Leech, R. (2015). Cognitive Flexibility through Metastable Neural Dynamics Is Disrupted by Damage to the Structural

Connectome. *Journal of Neuroscience*, *35*(24), 9050–9063. https://doi.org/10.1523/JNEUROSCI.4648-14.2015

Henderson, S. E., Johnson, A. R., Vallejo, A. I., Katz, L., Wong, E., & Gabbay, V. (2013). A preliminary study of white matter in adolescent depression: Relationships with illness severity, anhedonia, and irritability. *Frontiers in Psychiatry*, *4*(NOV), 152. https://doi.org/10.3389/fpsyt.2013.00152

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature*, *466*(7302), 29–29. https://doi.org/10.1038/466029a

Herting, M. M., Johnson, C., Mills, K. L., Vijayakumar, N., Dennison, M., Liu, C., Goddings, A.-L., Dahl, R. E., Sowell, E. R., Whittle, S., Allen, N. B., & Tamnes, C. K. (2018). Development of subcortical volumes across adolescence in males and females: A multisample study of longitudinal changes. *NeuroImage*, *172*, 194–205. https://doi.org/10.1016/j.neuroimage.2018.01.020

Herting, M. M., Maxwell, E. C., Irvine, C., & Nagel, B. J. (2012). The impact of sex, puberty, and hormones on white matter microstructure in adolescents. *Cerebral Cortex*, *22*(9), 1979–1992. https://doi.org/10.1093/cercor/bhr246

Herting, M. M., & Sowell, E. R. (2017). Puberty and structural brain development in humans. *Frontiers in Neuroendocrinology*, *44*(2017), 122–137. https://doi.org/10.1016/j.yfrne.2016.12.003

Herting, M. M., Uban, K. A., Robledo Gonzalez, M., Baker, F. C., Kan, E. C., Thompson, W. K., Granger, D. A., Albaugh, M. D., Anokhin, A. P., Bagot, K. S., Banich, M. T., Barch, D. M., Baskin-Sommers, A., Breslin, F. J., Casey, B. J., Chaarani, B., Chang, L., Clark, D. B., Cloak, C. C., … Edu, K. (2021). Correspondence Between Perceived Pubertal

Development and Hormone Levels in 9-10 Year-Olds From the Adolescent Brain Cognitive Development Study. *Article*, *11*, 1–1. https://doi.org/10.3389/fendo.2020.549928

Herzberg, M. P., & Gunnar, M. R. (2020). Early life stress and brain function: Activity and connectivity associated with processing emotion and reward. *NeuroImage*, *209*, 116493. https://doi.org/10.1016/j.neuroimage.2019.116493

Ho, T. C., & King, L. S. (2021). Mechanisms of neuroplasticity linking early adversity to depression: Developmental considerations. *Translational Psychiatry*, *11*(1), Article 1. https://doi.org/10.1038/s41398-021-01639-6

Holland, D., Kuperman, J. M., & Dale, A. M. (2010). Efficient correction of inhomogeneous static magnetic field-induced distortion in Echo Planar Imaging. *NeuroImage*, *50*(1), 175–183. https://doi.org/10.1016/j.neuroimage.2009.11.044

Holleran, L., Kelly, S., Alloza, C., Agartz, I., Andreassen, O. A., Arango, C., Banaj, N., Calhoun, V., Cannon, D., Carr, V., Corvin, A., Glahn, D. C., Gur, R., Hong, E., Hoschl, C., Howells, F. M., James, A., Janssen, J., Kochunov, P., … Donohoe, G. (2020). The relationship between white matter microstructure and general cognitive ability in patients with schizophrenia and healthy participants in the ENIGMA consortium. *American Journal of Psychiatry*. https://doi.org/10.1176/appi.ajp.2019.19030225

Hollocks, M. J., Lerh, J. W., Magiati, I., Meiser-Stedman, R., & Brugha, T. S. (2019). Anxiety and depression in adults with autism spectrum disorder: A systematic review and meta-analysis. *Psychological Medicine*, *49*(4), 559–572. https://doi.org/10.1017/S0033291718002283

Holm, M. C., Leonardsen, E. H., Beck, D., Dahl, A., Kjelkenes, R., Lange, A.-M. G. de, & Westlye,

L. T. (2022). *Linking brain maturation and puberty during early adolescence using longitudinal brain age prediction in the ABCD cohort* (p. 2022.05.16.22275146). medRxiv. https://doi.org/10.1101/2022.05.16.22275146

Horn, A. (2016). *MNI T1 6thGen NLIN to MNI 2009b NLIN ANTs transform* [Data set]. figshare. https://doi.org/10.6084/m9.figshare.3502238.v2

Howard, D. M., Adams, M. J., Clarke, T.-K., Hafferty, J. D., Gibson, J., Shirali, M., Coleman, J. R. I., Hagenaars, S. P., Ward, J., Wigmore, E. M., Alloza, C., Shen, X., Barbu, M. C., Xu, E. Y., Whalley, H. C., Marioni, R. E., Porteous, D. J., Davies, G., Deary, I. J., … McIntosh, A. M. (2019). Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nature Neuroscience*, *22*(3), 343–352. https://doi.org/10.1038/s41593-018-0326-7

Hu, S., Pruessner, J. C., Coupé, P., & Collins, D. L. (2013). Volumetric analysis of medial temporal lobe structures in brain development from childhood to adolescence. *NeuroImage*, *74*, 276–287. https://doi.org/10.1016/J.NEUROIMAGE.2013.02.032

Humphreys, K. L., Gleason, M. M., Drury, S. S., Miron, D., Nelson, C. A., Fox, N. A., & Zeanah, C. H. (2015). Effects of institutional rearing and foster care on psychopathology at age 12 years in Romania: Follow-up of an open, randomised controlled trial. *The Lancet. Psychiatry*, *2*(7), 625–634. https://doi.org/10.1016/S2215-0366(15)00095-4

Huttenlocher, P. R., & Dabholkar, A. S. (1997). Regional differences in synaptogenesis in human cerebral cortex. *The Journal of Comparative Neurology*, *387*(2), 167–178. https://doi.org/10.1002/(sici)1096-9861(19971020)387:2<167::aid-cne1>3.0.co;2-z

Iacono, W. G., Heath, A. C., Hewitt, J. K., Neale, M. C., Banich, M. T., Luciana, M. M., Madden, P. A., Barch, D. M., & Bjork, J. M. (2018). The utility of twins in developmental cognitive

neuroscience research: How twins strengthen the ABCD research design. *Developmental Cognitive Neuroscience*, *32*, 30–42. https://doi.org/10.1016/j.dcn.2017.09.001

Iraji, A., Faghiri, A., Lewis, N., Fu, Z., Rachakonda, S., & Calhoun, V. D. (2021). Tools of the trade: Estimating time-varying connectivity patterns from fMRI data. *Social Cognitive and Affective Neuroscience*, *16*(8), 849–874. https://doi.org/10.1093/scan/nsaa114

Ivie, E. J., Pettitt, A., Moses, L. J., & Allen, N. B. (2020). A meta-analysis of the association between adolescent social media use and depressive symptoms. *Journal of Affective Disorders*, *275*, 165–174. https://doi.org/10.1016/j.jad.2020.06.014

Jahanshad, N., Kochunov, P. V, Sprooten, E., Mandl, R. C., Nichols, T. E., Almasy, L., Blangero, J., Brouwer, R. M., Curran, J. E., Zubicaray, G. I. De, Duggirala, R., Fox, P. T., Hong, L. E., Landman, B. A., Martin, N. G., Mcmahon, K. L., Medland, S. E., Mitchell, B. D., Olvera, R. L., … Glahn, D. C. (2013). NeuroImage Multi-site genetic analysis of diffusion images and voxelwise heritability analysis: A pilot project of the ENIGMA – DTI working group. *NeuroImage*, *81*, 455–469. https://doi.org/10.1016/j.neuroimage.2013.04.061

James, S. L., Abate, D., Abate, K. H., Abay, S. M., Abbafati, C., Abbasi, N., Abbastabar, H., Abd-Allah, F., Abdela, J., Abdelalim, A., Abdollahpour, I., Abdulkader, R. S., Abebe, Z., Abera, S. F., Abil, O. Z., Abraha, H. N., Abu-Raddad, L. J., Abu-Rmeileh, N. M. E., Accrombessi, M. M. K., … Murray, C. J. L. (2018). Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: A systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*, *392*(10159), 1789–1858. https://doi.org/10.1016/S0140-6736(18)32279-7

Jami, E. S., Hammerschlag, A. R., Ip, H. F., Allegrini, A. G., Benyamin, B., Border, R., Diemer, E. W., Jiang, C., Karhunen, V., Lu, Y., Lu, Q., Mallard, T. T., Mishra, P. P., Nolte, I. M., Palviainen, T., Peterson, R. E., Sallis, H. M., Shabalin, A. A., Tate, A. E., … Middeldorp, C. M. (2022). Genome-wide Association Meta-analysis of Childhood and Adolescent Internalizing Symptoms. *Journal of the American Academy of Child & Adolescent Psychiatry*, *61*(7), 934–945. https://doi.org/10.1016/j.jaac.2021.11.035

Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., & Smith, S. M. (2012). FSL. *NeuroImage*, *62*(2), 782–790. https://doi.org/10.1016/j.neuroimage.2011.09.015

Jernigan, T. L., Brown, T. T., Hagler, D. J., Akshoomoff, N., Bartsch, H., Newman, E., Thompson, W. K., Bloss, C. S., Murray, S. S., Schork, N., Kennedy, D. N., Kuperman, J. M., McCabe, C., Chung, Y., Libiger, O., Maddox, M., Casey, B. J., Chang, L., Ernst, T. M., … Dale, A. M. (2016). The Pediatric Imaging, Neurocognition, and Genetics (PING) Data Repository. *NeuroImage*, *124*, 1149–1154. https://doi.org/10.1016/j.neuroimage.2015.04.057

Jobson, D. D., Hase, Y., Clarkson, A. N., & Kalaria, R. N. (2021). The role of the medial prefrontal cortex in cognition, ageing and dementia. *Brain Communications*, *3*(3), fcab125. https://doi.org/10.1093/braincomms/fcab125

Jones, D. K., Knösche, T. R., & Turner, R. (2013). White matter integrity, fiber count, and other fallacies: The do's and don'ts of diffusion MRI. *NeuroImage*, *73*, 239–254. https://doi.org/10.1016/j.neuroimage.2012.06.081

Jones, S. A., Morales, A. M., & Nagel, B. J. (2019). Resilience to risk for psychopathology: The role of white matter microstructural development in adolescence. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *4*(2), 180–189.

https://doi.org/10.1016/j.bpsc.2018.08.006

Kaiser, R. H., Andrews-Hanna, J. R., Wager, T. D., & Pizzagalli, D. A. (2015). Large-scale network dysfunction in major depressive disorder: A meta-analysis of resting-state functional connectivity. *JAMA Psychiatry*, *72*(6), 603–611. https://doi.org/10.1001/jamapsychiatry.2015.0071

Kaiser, R. H., Kang, M. S., Lew, Y., Van Der Feen, J., Aguirre, B., Clegg, R., Goer, F., Esposito, E., Auerbach, R. P., Hutchison, R. M., & Pizzagalli, D. A. (2019). Abnormal frontoinsular-default network dynamics in adolescent depression and rumination: A preliminary resting-state co-activation pattern analysis. *Neuropsychopharmacology*, *44*(9), Article 9. https://doi.org/10.1038/s41386-019-0399-3

Kaiser, R. H., Snyder, H. R., Goer, F., Clegg, R., Ironside, M., & Pizzagalli, D. A. (2018). Attention Bias in Rumination and Depression: Cognitive Mechanisms and Brain Networks. *Clinical Psychological Science: A Journal of the Association for Psychological Science*, *6*(6), 765–782. https://doi.org/10.1177/2167702618797935

Kaiser, R. H., Whitfield-Gabrieli, S., Dillon, D. G., Goer, F., Beltzer, M., Minkel, J., Smoski, M., Dichter, G., & Pizzagalli, D. A. (2016). Dynamic Resting-State Functional Connectivity in Major Depression. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *41*(7), 1822–1830. https://doi.org/10.1038/npp.2015.352

Kang, H. C., Burgund, E. D., Lugar, H. M., Petersen, S. E., & Schlaggar, B. L. (2003). Comparison of functional activation foci in children and adults using a common stereotactic space. *NeuroImage*, *19*(1), 16–28. https://doi.org/10.1016/S1053-8119(03)00038-7

Kaplowitz, P. B., Slora, E. J., Wasserman, R. C., Pedlow, S. E., & Herman-Giddens, M. E. (2001).

Earlier Onset of Puberty in Girls: Relation to Increased Body Mass Index and Race. *Pediatrics*, *108*(2), 347–353. https://doi.org/10.1542/peds.108.2.347

Karahanoğlu, F. I., & Van De Ville, D. (2015). Transient brain activity disentangles fMRI resting-state dynamics in terms of spatially and temporally overlapping networks. *Nature Communications*, *6*(1), Article 1. https://doi.org/10.1038/ncomms8751

Karcher, N. R., Barch, D. M., Avenevoli, S., Savill, M., Huber, R. S., Simon, T. J., Leckliter, I. N., Sher, K. J., & Loewy, R. L. (2018). Assessment of the Prodromal Questionnaire–Brief Child Version for Measurement of Self-reported Psychoticlike Experiences in Childhood. *JAMA Psychiatry*, *75*(8), 853–861. https://doi.org/10.1001/jamapsychiatry.2018.1334

Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., Williamson, D., & Ryan, N. (1997). Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): Initial Reliability and Validity Data. *Journal of the American Academy of Child & Adolescent Psychiatry*, *36*(7), 980–988. https://doi.org/10.1097/00004583-199707000-00021

Keiding, N., & Louis, T. A. (2016). Perils and potentials of self-selected entry to epidemiological studies and surveys. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *179*(2), 319–376. https://doi.org/10.1111/rssa.12136

Kelly, A. B., Mason, W. A., Chmelka, M. B., Herrenkohl, T. I., Kim, M. J., Patton, G. C., Hemphill, S. A., Toumbourou, J. W., & Catalano, R. F. (2016). Depressed Mood During Early to Middle Adolescence: A Bi-national Longitudinal Study of the Unique Impact of Family Conflict. *Journal of Youth and Adolescence*, *45*(8), 1604–1613. https://doi.org/10.1007/s10964-016-0433-2

Kempton, M. J., Salvador, Z., Munafò, M. R., Geddes, J. R., Simmons, A., Frangou, S., & Williams, S. C. R. (2011). Structural neuroimaging studies in major depressive disorder: Meta-analysis and comparison with bipolar disorder. *Archives of General Psychiatry*, *68*(7), 675–690. https://doi.org/10.1001/archgenpsychiatry.2011.60

Kessler, R. C., McLaughlin, K. A., Green, J. G., Gruber, M. J., Sampson, N. A., Zaslavsky, A. M., Aguilar-Gaxiola, S., Alhamzawi, A. O., Alonso, J., Angermeyer, M., Benjet, C., Bromet, E., Chatterji, S., Girolamo, G. de, Demyttenaere, K., Fayyad, J., Florescu, S., Gal, G., Gureje, O., … Williams, D. R. (2010). Childhood adversities and adult psychopathology in the WHO World Mental Health Surveys. *The British Journal of Psychiatry*, *197*(5), 378–385. https://doi.org/10.1192/bjp.bp.110.080499

Khan, S. S., Singer, B. D., & Vaughan, D. E. (2017). Molecular and physiological manifestations and measurement of aging in humans. *Aging Cell*, *16*(4), 624–633. https://doi.org/10.1111/acel.12601

Khazanov, G. K., & Ruscio, A. M. (2016). Is low positive emotionality a specific risk factor for depression? A meta-analysis of longitudinal studies. *Psychological Bulletin*, *142*(9), 991–1015. https://doi.org/10.1037/bul0000059

Knoll, L. J., Magis-Weinberg, L., Speekenbrink, M., & Blakemore, S. J. (2015). Social Influence on Risk Perception During Adolescence. *Psychological Science*, *26*(5), 583–592. https://doi.org/10.1177/0956797615569578

Koolschijn, P. C. M. P., Peper, J. S., & Crone, E. A. (2014). The influence of sex steroids on structural brain maturation in adolescence. *PLoS ONE*, *9*(1), e83929. https://doi.org/10.1371/journal.pone.0083929

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9. *Journal of General Internal*

Medicine, *16*(9), 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Kryza-Lacombe, M., Kiefer, C., Schwartz, K. T. G., Strickland, K., & Wiggins, J. L. (2020). Attention shifting in the context of emotional faces: Disentangling neural mechanisms of irritability from anxiety. *Depression and Anxiety*, *37*(7), 645–656. https://doi.org/10.1002/da.23010

Kundu, P., Benson, B. E., Rosen, D., Frangou, S., Leibenluft, E., Luh, W.-M., Bandettini, P. A., Pine, D. S., & Ernst, M. (2018). The Integration of Functional Brain Activity from Adolescence to Adulthood. *Journal of Neuroscience*, *38*(14), 3559–3570. https://doi.org/10.1523/JNEUROSCI.1864-17.2018

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*, 1–26. https://doi.org/10.18637/jss.v082.i13

Kwong, A. S. F., Morris, T. T., Pearson, R. M., Timpson, N. J., Rice, F., Stergiakouli, E., & Tilling, K. (2021). Polygenic risk for depression, anxiety and neuroticism are associated with the severity and rate of change in depressive symptoms across adolescence. *Journal of Child Psychology and Psychiatry*, *62*(12), 1462–1474. https://doi.org/10.1111/jcpp.13422

Kwong, A. S., Pearson, R. M., Adams, M. J., Northstone, K., Tilling, K., Smith, D., Fawns-Ritchie, C., Bould, H., Warne, N., & Zammit, S. (2021). Mental health before and during the COVID-19 pandemic in two longitudinal UK population cohorts. *The British Journal of Psychiatry*, *218*(6), 334–343.

Lai, C. H. (2013). Gray matter volume in major depressive disorder: A meta-analysis of voxel-based morphometry studies. *Psychiatry Research - Neuroimaging*, *211*(1), 37–46.

https://doi.org/10.1016/j.pscychresns.2012.06.006

Lebel, C., & Deoni, S. (2018). The development of brain white matter microstructure. *NeuroImage*, *182*, 207–218. https://doi.org/10.1016/j.neuroimage.2017.12.097

Lebel, C., Treit, S., & Beaulieu, C. (2019). A review of diffusion MRI of typical white matter development from early childhood to young adulthood. *NMR in Biomedicine*, *32*(4), e3778. https://doi.org/10.1002/nbm.3778

Lee, K. S., Hagan, C. N., Hughes, M., Cotter, G., McAdam Freud, E., Kircanski, K., Leibenluft, E., Brotman, M. A., & Tseng, W.-L. (2022). Systematic Review and Meta-analysis: Task-based fMRI Studies in Youths With Irritability. *Journal of the American Academy of Child & Adolescent Psychiatry*. https://doi.org/10.1016/j.jaac.2022.05.014

Leemans, A., & Jones, D. K. (2009). The B-matrix must be rotated when correcting for subject motion in DTI data. *Magnetic Resonance in Medicine*, *61*(6), 1336–1349. https://doi.org/10.1002/mrm.21890

Leibenluft, E., Cohen, P., Gorrindo, T., Brook, J. S., & Pine, D. S. (2006). Chronic Versus Episodic Irritability in Youth: ACommunity-Based, Longitudinal Study of Clinical and Diagnostic Associations. *Journal of Child and Adolescent Psychopharmacology*, *16*(4), 456–466. https://doi.org/10.1089/cap.2006.16.456

Leibenluft, E., & Stoddard, J. (2013). The developmental psychopathology of irritability. *Development and Psychopathology*, *25*(4 0 2), 1473–1487. https://doi.org/10.1017/S0954579413000722

LeMoult, J., Humphreys, K. L., Tracy, A., Hoffmeister, J.-A., Ip, E., & Gotlib, I. H. (2020). Meta-analysis: Exposure to Early Life Stress and Risk for Depression in Childhood and Adolescence. *Journal of the American Academy of Child and Adolescent Psychiatry*,

59(7), 842–855. https://doi.org/10.1016/j.jaac.2019.10.011

Lenroot, R. K., Gogtay, N., Greenstein, D. K., Wells, E. M., Wallace, G. L., Clasen, L. S., Blumenthal, J. D., Lerch, J., Zijdenbos, A. P., Evans, A. C., Thompson, P. M., & Giedd, J. N. (2007). Sexual dimorphism of brain developmental trajectories during childhood and adolescence. *NeuroImage*, *36*(4), 1065–1073. https://doi.org/10.1016/j.neuroimage.2007.03.053

Lerch, J. P., van der Kouwe, A. J. W., Raznahan, A., Paus, T., Johansen-Berg, H., Miller, K. L., Smith, S. M., Fischl, B., & Sotiropoulos, S. N. (2017). Studying neuroanatomy using MRI. *Nature Neuroscience*, *20*(3), Article 3. https://doi.org/10.1038/nn.4501

Levis, B., Benedetti, A., & Thombs, B. D. (2019). Accuracy of Patient Health Questionnaire-9 (PHQ-9) for screening to detect major depression: Individual participant data meta-analysis. *BMJ*, *365*, l1476. https://doi.org/10.1136/bmj.l1476

LeWinn, K. Z., Connolly, C. G., Wu, J., Drahos, M., Hoeft, F., Ho, T. C., Simmons, A. N., & Yang, T. T. (2014). White Matter Correlates of Adolescent Depression: Structural Evidence for Frontolimbic Disconnectivity. *Journal of the American Academy of Child & Adolescent Psychiatry*, *53*(8), 899-909.e7. https://doi.org/10.1016/J.JAAC.2014.04.021

Lewis, K. J. S., Mars, B., Lewis, G., Rice, F., Sellers, R., Thapar, A. A. K., Craddock, N., Collishaw, S., & Thapar, A. A. K. (2012). Do parents know best? Parent-reported vs. Child-reported depression symptoms as predictors of future child mood disorder in a high-risk sample. *Journal of Affective Disorders*, *141*(2–3), 233–236. https://doi.org/10.1016/j.jad.2012.03.008

Li, G., Nie, J., Wang, L., Shi, F., Lin, W., Gilmore, J. H., & Shen, D. (2013). Mapping region-

specific longitudinal cortical surface expansion from birth to 2 years of age. *Cerebral Cortex*, *23*(11), 2724–2733. https://doi.org/10.1093/cercor/bhs265

Liu, X., Chang, C., & Duyn, J. (2013). Decomposition of Spontaneous Brain Activity into Distinct fMRI Co-activation Patterns. *Frontiers in Systems Neuroscience*, *7*. https://www.frontiersin.org/articles/10.3389/fnsys.2013.00101

Liuzzi, M. T., Kryza-Lacombe, M., Christian, I. R., Palumbo, D. E., Amir, N., & Wiggins, J. L. (2020). Neural and behavioral correlates of inhibitory control in youths with varying levels of irritability. *Journal of Affective Disorders*, *273*, 567–575. https://doi.org/10.1016/j.jad.2020.04.049

Long, Y., Cao, H., Yan, C., Chen, X., Li, L., Castellanos, F. X., Bai, T., Bo, Q., Chen, G., Chen, N., Chen, W., Cheng, C., Cheng, Y., Cui, X., Duan, J., Fang, Y., Gong, Q., Guo, W., Hou, Z., … Liu, Z. (2020). Altered resting-state dynamic functional brain networks in major depressive disorder: Findings from the REST-meta-MDD consortium. *NeuroImage: Clinical*, *26*, 102163. https://doi.org/10.1016/j.nicl.2020.102163

Lopez-Leon, S., Janssens, A. C., Gonzalez-Zuloeta Ladd, A. M., Del-Favero, J., Claes, S. J., Oostra, B. A., & van Duijn, C. M. (2008). Meta-analyses of genetic studies on major depressive disorder. *Mol Psychiatry*, *13*(8), 772–785. https://doi.org/10.1038/sj.mp.4002088

Lord, L.-D., Expert, P., Atasoy, S., Roseman, L., Rapuano, K., Lambiotte, R., Nutt, D. J., Deco, G., Carhart-Harris, R. L., Kringelbach, M. L., & Cabral, J. (2019). Dynamical exploration of the repertoire of brain networks at rest is modulated by psilocybin. *NeuroImage*, *199*, 127–142. https://doi.org/10.1016/j.neuroimage.2019.05.060

Lund, J. S., Boothe, R. G., & Lund, R. D. (1977). Development of neurons in the visual cortex

(area 17) of the monkey (Macaca nemestrina): A Golgi study from fetal day 127 to postnatal maturity. *The Journal of Comparative Neurology*, *176*(2), 149–188. https://doi.org/10.1002/cne.901760203

Lyall, A. E., Shi, F., Geng, X., Woolson, S., Li, G., Wang, L., Hamer, R. M., Shen, D., & Gilmore, J. H. (2015). Dynamic Development of Regional Cortical Thickness and Surface Area in Early Childhood. *Cerebral Cortex*, *25*(8), 2204–2212. https://doi.org/10.1093/cercor/bhu027

Maciejewski, D., Hillegers, M., & Penninx, B. (2018). Offspring of parents with mood disorders: Time for more transgenerational research, screening and preventive intervention for this high-risk population. *Current Opinion in Psychiatry*, *31*(4), 349–357. https://doi.org/10.1097/YCO.0000000000000423

MacSweeney, N., Bowman, S., & Kelly, C. (2019). More than just characters in a story: Effective and meaningful involvement of young people in mental health research. *Journal of Public Mental Health*, *18*(1), 14–16. https://doi.org/10.1108/JPMH-07-2018-0053

Malhi, G. S., & Mann, J. J. (2018). Depression. *Lancet (London, England)*, *392*(10161), 2299–2312. https://doi.org/10.1016/S0140-6736(18)31948-2

Mancini, A., Magnotto, J. C., & Abreu, A. P. (2022). Genetics of pubertal timing. *Best Practice & Research Clinical Endocrinology & Metabolism*, *36*(1), 101618. https://doi.org/10.1016/j.beem.2022.101618

Maoz, H., Goldstein, T., Goldstein, B. I., Axelson, D. A., Fan, J., Hickey, M. B., Monk, K., Sakolsky, D., Diler, R. S., Brent, D., Kupfer, D. J., & Birmaher, B. (2014). The Effects of Parental Mood on Reports of Their Children's Psychopathology. *Journal of the American Academy of Child & Adolescent Psychiatry*, *53*(10), 1111-1122.e5.

https://doi.org/10.1016/j.jaac.2014.07.005

Marchitelli, R., Paillère-Martinot, M.-L., Bourvis, N., Guerin-Langlois, C., Kipman, A., Trichard, C., Douniol, M., Stordeur, C., Galinowski, A., Filippi, I., Bertschy, G., Weibel, S., Granger, B., Limosin, F., Cohen, D., Martinot, J.-L., & Artiges, E. (2022). Dynamic Functional Connectivity in Adolescence-Onset Major Depression: Relationships With Severity and Symptom Dimensions. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *7*(4), 385–396. https://doi.org/10.1016/j.bpsc.2021.05.003

Marek, S., & Dosenbach, N. U. F. (2018). The frontoparietal network: Function, electrophysiology, and importance of individual precision mapping. *Dialogues in Clinical Neuroscience*, *20*(2), 133–140.

Marek, S., Hwang, K., Foran, W., Hallquist, M. N., & Luna, B. (2015). The Contribution of Network Organization and Integration to the Development of Cognitive Control. *PLOS Biology*, *13*(12), e1002328. https://doi.org/10.1371/journal.pbio.1002328

Marek, S., Tervo-Clemmens, B., Calabro, F. J., Montez, D. F., Kay, B. P., Hatoum, A. S., Donohue, M. R., Foran, W., Miller, R. L., Hendrickson, T. J., Malone, S. M., Kandala, S., Feczko, E., Miranda-Dominguez, O., Graham, A. M., Earl, E. A., Perrone, A. J., Cordova, M., Doyle, O., … Dosenbach, N. U. F. (2022). Reproducible brain-wide association studies require thousands of individuals. *Nature*, *603*(7902), Article 7902. https://doi.org/10.1038/s41586-022-04492-9

Marino, C., Andrade, B., Campisi, S. C., Wong, M., Zhao, H., Jing, X., Aitken, M., Bonato, S., Haltigan, J., Wang, W., & Szatmari, P. (2021). Association Between Disturbed Sleep and Depression in Children and Youths: A Systematic Review and Meta-analysis of Cohort Studies. *JAMA Network Open*, *4*(3), e212373.

https://doi.org/10.1001/jamanetworkopen.2021.2373

Marshall, W. A., & Tanner, J. M. (1969). Variations in pattern of pubertal changes in girls. *Archives of Disease in Childhood*, *44*(235), 291–303. https://doi.org/10.1136/adc.44.235.291

Marshall, W. A., & Tanner, J. M. (1970). Variations in the Pattern of Pubertal Changes in Boys. *Archives of Disease in Childhood*, *45*(13). https://doi.org/10.1136/adc.45.239.13

Marusak, H. A., Calhoun, V. D., Brown, S., Crespo, L. M., Sala-Hamrick, K., Gotlib, I. H., & Thomason, M. E. (2017). Dynamic functional connectivity of neurocognitive networks in children. *Human Brain Mapping*, *38*(1), 97–108. https://doi.org/10.1002/hbm.23346

Matta, T. H., Flournoy, J. C., & Byrne, M. L. (2018). Making an unknown unknown a known unknown: Missing data in longitudinal neuroimaging studies. *Developmental Cognitive Neuroscience*, *33*, 83–98. https://doi.org/10.1016/j.dcn.2017.10.001

McAdams, T. A., Rijsdijk, F. V., Neiderhiser, J. M., Narusyte, J., Shaw, D. S., Natsuaki, M. N., Spotts, E. L., Ganiban, J. M., Reiss, D., Leve, L. D., Lichtenstein, P., & Eley, T. C. (2015). The relationship between parental depressive symptoms and offspring psychopathology: Evidence from a children-of-twins study and an adoption study. *Psychological Medicine*, *45*(12), 2583–2594. https://doi.org/10.1017/S0033291715000501

McAnarney, E. (1992). *Textbook of Adolescent Medicine*. Saunders.

McIntosh, A. M., Sullivan, P. F., & Lewis, C. M. (2019). Uncovering the Genetic Architecture of Major Depression. *Neuron*, *102*(1), 91–103. https://doi.org/10.1016/j.neuron.2019.03.022

McNeilly, E. A., Saragosa-Harris, N. M., Mills, K. L., Dahl, R. E., & Magis-Weinberg, L. (2022). Reward sensitivity and internalizing symptoms during the transition to puberty: An examination of 9-and 10-year-olds in the ABCD Study. *Developmental Cognitive Neuroscience*, *58*, 101172. https://doi.org/10.1016/j.dcn.2022.101172

Meinzer, M. C., Pettit, J. W., & Viswesvaran, C. (2014). The co-occurrence of attention-deficit/hyperactivity disorder and unipolar depression in children and adolescents: A meta-analytic review. *Clinical Psychology Review*, *34*(8), 595–607. https://doi.org/10.1016/j.cpr.2014.10.002

Mendle, J., Harden, K. P., Brooks-Gunn, J., & Graber, J. A. (2010). Development's tortoise and hare: Pubertal timing, pubertal tempo, and depressive symptoms in boys and girls. *Developmental Psychology*, *46*(5), 1341.

Milham, M. P., Craddock, R. C., & Klein, A. (2017). Clinically useful brain imaging for neuropsychiatry: How can we get there? In *Depression and Anxiety* (Vol. 34, Issue 7, pp. 578–587). Blackwell Publishing Inc. https://doi.org/10.1002/da.22627

Mills, K. L., Goddings, A. L., Herting, M. M., Meuwese, R., Blakemore, S. J., Crone, E. A., Dahl, R. E., Güroğlu, B., Raznahan, A., Sowell, E. R., & Tamnes, C. K. (2016). Structural brain development between childhood and adulthood: Convergence across four longitudinal samples. *NeuroImage*, *141*, 273–281. https://doi.org/10.1016/j.neuroimage.2016.07.044

Mills, K. L., Goddings, A.-L., Clasen, L. S., Giedd, J. N., & Blakemore, S.-J. (2014). The Developmental Mismatch in Structural Brain Maturation during Adolescence. *Developmental Neuroscience*, *36*(3–4), 147–160. https://doi.org/10.1159/000362328

Mills, K. L., Siegmund, K. D., Tamnes, C. K., Ferschmann, L., Wierenga, L. M., Bos, M. G. N.,

Luna, B., Li, C., & Herting, M. M. (2021). Inter-individual variability in structural brain development from late childhood to young adulthood. *NeuroImage*, *242*, 118450. https://doi.org/10.1016/j.neuroimage.2021.118450

Moore, S. E., Norman, R. E., Suetani, S., Thomas, H. J., Sly, P. D., & Scott, J. G. (2017). Consequences of bullying victimization in childhood and adolescence: A systematic review and meta-analysis. *World Journal of Psychiatry*, *7*(1), 60–76. https://doi.org/10.5498/wjp.v7.i1.60

Mulraney, M. A., Melvin, G. A., & Tonge, B. J. (2014). Psychometric properties of the Affective Reactivity Index in Australian adults and adolescents. *Psychological Assessment*, *26*(1), 148–155. https://doi.org/10.1037/a0034891

Naicker, K., Galambos, N. L., Zeng, Y., Senthilselvan, A., & Colman, I. (2013). Social, demographic, and health outcomes in the 10 years following adolescent depression. *Journal of Adolescent Health*, *52*(5), 533–538. https://doi.org/10.1016/j.jadohealth.2012.12.016

Natu, V. S., Gomez, J., Barnett, M., Jeska, B., Kirilina, E., Jaeger, C., Zhen, Z., Cox, S., Weiner, K. S., Weiskopf, N., & Grill-Spector, K. (2019). Apparent thinning of human visual cortex during childhood is associated with myelination. *Proceedings of the National Academy of Sciences*, *116*(41), 20750–20759. https://doi.org/10.1073/pnas.1904931116

Nelson, C. A., Zeanah, C. H., Fox, N. A., Marshall, P. J., Smyke, A. T., & Guthrie, D. (2007). Cognitive recovery in socially deprived young children: The Bucharest Early Intervention Project. *Science (New York, N.Y.)*, *318*(5858), 1937–1940. https://doi.org/10.1126/science.1143921

Neufang, S., Specht, K., Hausmann, M., Gunturkun, O., Herpertz-Dahlmann, B., Fink, G. R., &

Konrad, K. (2009). Sex Differences and the Impact of Steroid Hormones on the Developing Human Brain. *Cerebral Cortex*, *19*(2), 464–473. https://doi.org/10.1093/cercor/bhn100

Newman, M. E. J. (2006). Finding community structure in networks using the eigenvectors of matrices. *Physical Review E. Newman, M.E. (2006). Finding Community Structure in Networks Using the Eigenvectors of Matrices. Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics.*, *74*(3), 036104. https://doi.org/10.1103/PhysRevE.74.036104

Nielsen, A. N., Wakschlag, L. S., & Norton, E. S. (2021). Linking irritability and functional brain networks: A transdiagnostic case for expanding consideration of development and environment in RDoC. *Neuroscience & Biobehavioral Reviews*, *129*, 231–244. https://doi.org/10.1016/j.neubiorev.2021.07.022

Noble, K. G., McCandliss, B. D., & Farah, M. J. (2007). Socioeconomic gradients predict individual differences in neurocognitive abilities. *Developmental Science*, *10*(4), 464–480. https://doi.org/10.1111/j.1467-7687.2007.00600.x

Norbom, L. B., Ferschmann, L., Parker, N., Agartz, I., Andreassen, O. A., Paus, T., Westlye, L. T., & Tamnes, C. K. (2021). New insights into the dynamic development of the cerebral cortex in childhood and adolescence: Integrating macro- and microstructural MRI findings. *Progress in Neurobiology*, *204*, 102109. https://doi.org/10.1016/j.pneurobio.2021.102109

Norman, R. E., Byambaa, M., De, R., Butchart, A., Scott, J., & Vos, T. (2012). The Long-Term Health Consequences of Child Physical Abuse, Emotional Abuse, and Neglect: A Systematic Review and Meta-Analysis. *PLOS Medicine*, *9*(11), e1001349.

https://doi.org/10.1371/journal.pmed.1001349

Nunes, A., Schnack, H. G., Ching, C. R. K., Agartz, I., Akudjedu, T. N., Alda, M., Alnæs, D., Alonso-Lana, S., Bauer, J., Baune, B. T., Bøen, E., Bonnin, C. del M., Busatto, G. F., Canales-Rodríguez, E. J., Cannon, D. M., Caseras, X., Chaim-Avancini, T. M., Dannlowski, U., Díaz-Zuluaga, A. M., … Hajek, T. (2020). Using structural MRI to identify bipolar disorders – 13 site machine learning study in 3020 individuals from the ENIGMA Bipolar Disorders Working Group. *Molecular Psychiatry*, *25*(9), 2130–2143. https://doi.org/10.1038/s41380-018-0228-9

O'Brien, L. M., Ziegler, D. A., Deutsch, C. K., Frazier, J. A., Herbert, M. R., & Locascio, J. J. (2011). Statistical adjustments for brain size in volumetric neuroimaging studies: Some practical implications in methods. *Psychiatry Research*, *193*(2), 113–122. https://doi.org/10.1016/j.pscychresns.2011.01.007

Opel, N., Redlich, R., Dohm, K., Zaremba, D., Goltermann, J., Repple, J., Kaehler, C., Grotegerd, D., Leehr, E. J., Böhnlein, J., Förster, K., Meinert, S., Enneking, V., Sindermann, L., Dzvonyar, F., Emden, D., Leenings, R., Winter, N., Hahn, T., … Dannlowski, U. (2019). Mediation of the influence of childhood maltreatment on depression relapse by cortical structure: A 2-year longitudinal observational study. *The Lancet Psychiatry*, *6*(4), 318–326. https://doi.org/10.1016/S2215-0366(19)30044-6

Patel, Y., Shin, J., Drakesmith, M., Evans, J., Pausova, Z., & Paus, T. (2020). Virtual histology of multi-modal magnetic resonance imaging of cerebral cortex in young men. *NeuroImage*, *218*, 116968. https://doi.org/10.1016/j.neuroimage.2020.116968

Patel, Y., Shin, J., Gowland, P. A., Pausova, Z., Paus, T., & IMAGEN consortium. (2019). Maturation of the Human Cerebral Cortex During Adolescence: Myelin or Dendritic

Arbor? *Cerebral Cortex*, *29*(8), 3351–3362. https://doi.org/10.1093/cercor/bhy204

Pawliczek, C. M., Derntl, B., Kellermann, T., Gur, R. C., Schneider, F., & Habel, U. (2013). Anger under Control: Neural Correlates of Frustration as a Function of Trait Aggression. *PLoS ONE*, *8*(10), e78503. https://doi.org/10.1371/journal.pone.0078503

Peper, J. S., Brouwer, R. M., Schnack, H. G., van Baal, G. C., van Leeuwen, M., van den Berg, S. M., Delemarre-Van de Waal, H. A., Boomsma, D. I., Kahn, R. S., & Hulshoff Pol, H. E. (2009). Sex steroids and brain structure in pubertal boys and girls. *Psychoneuroendocrinology*, *34*(3), 332–342. https://doi.org/10.1016/j.psyneuen.2008.09.012

Peper, J. S., de Reus, M. A., van den Heuvel, M. P., & Schutter, D. J. L. G. (2015). Short fused? Associations between white matter connections, sex steroids, and aggression across adolescence. *Human Brain Mapping*, *36*(3), 1043–1052. https://doi.org/10.1002/hbm.22684

Perlman, S. B., Jones, B. M., Wakschlag, L. S., Axelson, D., Birmaher, B., & Phillips, M. L. (2015). Neural substrates of child irritability in typically developing and psychiatric populations. *Developmental Cognitive Neuroscience*, *14*, 71–80. https://doi.org/10.1016/j.dcn.2015.07.003

Peters, S., Van Duijvenvoorde, A. C. K., Koolschijn, P. C. M. P., & Crone, E. A. (2016). Longitudinal development of frontoparietal activity during feedback learning: Contributions of age, performance, working memory and cortical thickness. *Developmental Cognitive Neuroscience*, *19*, 211–222. https://doi.org/10.1016/j.dcn.2016.04.004

Petersen, A. C., Crockett, L., Richards, M., & Boxer, A. (1988). A self-report measure of

pubertal status: Reliability, validity, and initial norms. *Journal of Youth and Adolescence*, *17*(2), 117–133. https://doi.org/10.1007/BF01537962

Peverill, M., Dirks, M. A., Narvaja, T., Herts, K. L., Comer, J. S., & McLaughlin, K. A. (2021). Socioeconomic status and child psychopathology in the United States: A meta-analysis of population-based studies. *Clinical Psychology Review*, *83*, 101933. https://doi.org/10.1016/j.cpr.2020.101933

Pfefferbaum, A., Rohlfing, T., Pohl, K. M., Lane, B., Chu, W., Kwon, D., Nolan Nichols, B., Brown, S. A., Tapert, S. F., Cummins, K., Thompson, W. K., Brumback, T., Meloy, M. J., Jernigan, T. L., Dale, A., Colrain, I. M., Baker, F. C., Prouty, D., De Bellis, M. D., … Sullivan, E. V. (2016). Adolescent Development of Cortical and White Matter Structure in the NCANDA Sample: Role of Sex, Ethnicity, Puberty, and Alcohol Drinking. *Cerebral Cortex*, *26*(10), 4101–4121. https://doi.org/10.1093/cercor/bhv205

Pfeifer, J. H., & Allen, N. B. (2021). Puberty Initiates Cascading Relationships Between Neurodevelopmental, Social, and Internalizing Processes Across Adolescence. *Biological Psychiatry*, *89*(2), 99–108. https://doi.org/10.1016/j.biopsych.2020.09.002

Pfeifer, J. H., & Peake, S. J. (2012). Self-development: Integrating cognitive, socioemotional, and neuroimaging perspectives. *Developmental Cognitive Neuroscience*, *2*(1), 55–69. https://doi.org/10.1016/j.dcn.2011.07.012

Philipp-Muller, A., Lee, S. W. S., & Petty, R. E. (2022). Why are people antiscience, and what can we do about it? *Proceedings of the National Academy of Sciences*, *119*(30), e2120755119. https://doi.org/10.1073/pnas.2120755119

Pinheiro, J., Bates, D., DebRoy, S., & Sarkar, D. (2007). nlme: Linear and Nonlinear Mixed Effects Models. *R Package Version 3*, 1–97.

Pinquart, M., & Shen, Y. (2011). Depressive symptoms in children and adolescents with chronic physical illness: An updated meta-analysis. *Journal of Pediatric Psychology*, *36*(4), 375–384. https://doi.org/10.1093/jpepsy/jsq104

Pohl, K. M., Sullivan, E. V., Rohlfing, T., Chu, W., Kwon, D., Nichols, B. N., Zhang, Y., Brown, S. A., Tapert, S. F., Cummins, K., Thompson, W. K., Brumback, T., Colrain, I. M., Baker, F. C., Prouty, D., De Bellis, M. D., Voyvodic, J. T., Clark, D. B., Schirda, C., … Pfefferbaum, A. (2016). Harmonizing DTI measurements across scanners to examine the development of white matter microstructure in 803 adolescents of the NCANDA study. *NeuroImage*, *130*, 194–213. https://doi.org/10.1016/j.neuroimage.2016.01.061

Poldrack, R. A., Baker, C. I., Durnez, J., Gorgolewski, K. J., Matthews, P. M., Munafò, M. R., Nichols, T. E., Poline, J.-B., Vul, E., & Yarkoni, T. (2017). Scanning the horizon: Towards transparent and reproducible neuroimaging research. *Nature Reviews Neuroscience*, *18*(2), Article 2. https://doi.org/10.1038/nrn.2016.167

Popescu, S. G., Glocker, B., Sharp, D. J., & Cole, J. H. (2021). Local Brain-Age: A U-Net Model. *Frontiers in Aging Neuroscience*, *13*. https://www.frontiersin.org/articles/10.3389/fnagi.2021.761954

Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods*, *15*(3), 209. https://doi.org/10.1037/a0020141

Pruim, R. H. R., Mennes, M., van Rooij, D., Llera, A., Buitelaar, J. K., & Beckmann, C. F. (2015). ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data. *NeuroImage*, *112*, 267–277. https://doi.org/10.1016/j.neuroimage.2015.02.064

Quek, Y.-H., Tam, W. W. S., Zhang, M. W. B., & Ho, R. C. M. (2017). Exploring the association between childhood and adolescent obesity and depression: A meta-analysis. *Obesity Reviews*, *18*(7), 742–754. https://doi.org/10.1111/obr.12535

Rabany, L., Brocke, S., Calhoun, V. D., Pittman, B., Corbera, S., Wexler, B. E., Bell, M. D., Pelphrey, K., Pearlson, G. D., & Assaf, M. (2019). Dynamic functional connectivity in schizophrenia and autism spectrum disorder: Convergence, divergence and classification. *NeuroImage: Clinical*, *24*, 101966. https://doi.org/10.1016/j.nicl.2019.101966

Racine, N., McArthur, B. A., Cooke, J. E., Eirich, R., Zhu, J., & Madigan, S. (2021). Global Prevalence of Depressive and Anxiety Symptoms in Children and Adolescents During COVID-19: A Meta-analysis. *JAMA Pediatrics*, *175*(11), 1142–1150. https://doi.org/10.1001/jamapediatrics.2021.2482

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences*, *98*(2), 676–682. https://doi.org/10.1073/pnas.98.2.676

Rakesh, D., Cropley, V., Zalesky, A., Vijayakumar, N., Allen, N. B., & Whittle, S. (2021). Neighborhood disadvantage and longitudinal brain-predicted-age trajectory during adolescence. *Developmental Cognitive Neuroscience*, *51*, 101002. https://doi.org/10.1016/j.dcn.2021.101002

Rakesh, D., Seguin, C., Zalesky, A., Cropley, V., & Whittle, S. (2021). Associations Between Neighborhood Disadvantage, Resting-State Functional Connectivity, and Behavior in the Adolescent Brain Cognitive Development Study: The Moderating Role of Positive Family and School Environments. *Biological Psychiatry: Cognitive Neuroscience and*

Neuroimaging, *6*(9), 877–886. https://doi.org/10.1016/j.bpsc.2021.03.008

Rakesh, D., & Whittle, S. (2021). Socioeconomic status and the developing brain – A systematic review of neuroimaging findings in youth. *Neuroscience & Biobehavioral Reviews*, *130*, 379–407. https://doi.org/10.1016/j.neubiorev.2021.08.027

Rakic, P., Bourgeois, J. P., Eckenhoff, M. F., Zecevic, N., & Goldman-Rakic, P. S. (1986). Concurrent overproduction of synapses in diverse regions of the primate cerebral cortex. *Science (New York, N.Y.)*, *232*(4747), 232–235. https://doi.org/10.1126/science.3952506

Rao, W.-W., Zong, Q.-Q., Zhang, J.-W., An, F.-R., Jackson, T., Ungvari, G. S., Xiang, Y., Su, Y.-Y., D'Arcy, C., & Xiang, Y.-T. (2020). Obesity increases the risk of depression in children and adolescents: Results from a systematic review and meta-analysis. *Journal of Affective Disorders*, *267*, 78–85. https://doi.org/10.1016/j.jad.2020.01.154

Rausch, E., Racz, S. J., Augenstein, T. M., Keeley, L., Lipton, M. F., Szollos, S., Riffle, J., Moriarity, D., Kromash, R., & De Los Reyes, A. (2017). A Multi-Informant Approach to Measuring Depressive Symptoms in Clinical Assessments of Adolescent Social Anxiety Using the Beck Depression Inventory-II: Convergent, Incremental, and Criterion-Related Validity. *Child & Youth Care Forum*, *46*(5), 661–683. https://doi.org/10.1007/s10566-017-9403-4

Raznahan, A., Shaw, P., Lalonde, F., Stockman, M., Wallace, G. L., Greenstein, D., Clasen, L., Gogtay, N., & Giedd, J. N. (2011). How Does Your Cortex Grow? *Journal of Neuroscience*, *31*(19), 7174–7177. https://doi.org/10.1523/JNEUROSCI.0054-11.2011

Raznahan, A., Shaw, P. W., Lerch, J. P., Clasen, L. S., Greenstein, D., Berman, R., Pipitone, J., Chakravarty, M. M., & Giedd, J. N. (2014). Longitudinal four-dimensional mapping of

subcortical anatomy in human development. *Proceedings of the National Academy of Sciences*, *111*(4), 1592–1597. https://doi.org/10.1073/pnas.1316911111

Reynolds, S., Carrey, N., Jaworska, N., Langevin, L. M., Yang, X. R., & MacMaster, F. P. (2014). Cortical thickness in youth with major depressive disorder. *BMC Psychiatry*, *14*(1), 83. https://doi.org/10.1186/1471-244X-14-83

Rice, F., Sellers, R., Hammerton, G., Eyre, O., Bevan-Jones, R., Thapar, A. K., Collishaw, S., Harold, G. T., & Thapar, A. (2017). Antecedents of New-Onset Major Depressive Disorder in Children and Adolescents at High Familial Risk. *JAMA Psychiatry*, *74*(2), 153–160. https://doi.org/10.1001/jamapsychiatry.2016.3140

Richardson, H., Lisandrelli, G., Riobueno-Naylor, A., & Saxe, R. (2018). Development of the social brain from age three to twelve years. *Nature Communications*, *9*(1), Article 1. https://doi.org/10.1038/s41467-018-03399-2

Richardson, L. P., McCauley, E., Grossman, D. C., McCarty, C. A., Richards, J., Russo, J. E., Rockhill, C., & Katon, W. (2010). Evaluation of the Patient Health Questionnaire-9 Item for detecting major depression among adolescents. *Pediatrics*, *126*(6), 1117–1123. https://doi.org/10.1542/peds.2010-0852

Roberts, J. A., Gollo, L. L., Abeysuriya, R. G., Roberts, G., Mitchell, P. B., Woolrich, M. W., & Breakspear, M. (2019). Metastable brain waves. *Nature Communications*, *10*(1), Article 1. https://doi.org/10.1038/s41467-019-08999-0

Rolls, E. T., Joliot, M., & Tzourio-Mazoyer, N. (2015). Implementation of a new parcellation of the orbitofrontal cortex in the automated anatomical labeling atlas. *NeuroImage*, *122*, 1–5. https://doi.org/10.1016/j.neuroimage.2015.07.075

Romer, D. (2010). Adolescent risk taking, impulsivity, and brain development: Implications for

prevention. *Developmental Psychobiology*, *52*(3), 263–276. https://doi.org/10.1002/dev.20442

Roy, A. K., Bennett, R., Posner, J., Hulvershorn, L., Castellanos, F. X., & Klein, R. G. (2018). Altered Intrinsic Functional Connectivity of the Cingulate Cortex in Children with Severe Temper Outbursts. *Development and Psychopathology*, *30*(2), 571–579. https://doi.org/10.1017/S0954579417001080

Rutter, M. (2010). Gene-environment interplay. *Depression and Anxiety*, *27*, 1–4. https://doi.org/10.1002/da.20641

Sakoğlu, Ü., Pearlson, G. D., Kiehl, K. A., Wang, Y. M., Michael, A. M., & Calhoun, V. D. (2010). A method for evaluating dynamic functional network connectivity and task-modulation: Application to schizophrenia. *Magnetic Resonance Materials in Physics, Biology and Medicine*, *23*(5), 351–366. https://doi.org/10.1007/s10334-010-0197-8

Saragosa-Harris, N. M., Chaku, N., MacSweeney, N., Guazzelli Williamson, V., Scheuplein, M., Feola, B., Cardenas-Iniguez, C., Demir-Lira, E., McNeilly, E. A., Huffman, L. G., Whitmore, L., Michalska, K. J., Damme, K. S., Rakesh, D., & Mills, K. L. (2022). A practical guide for researchers and reviewers using the ABCD Study and other large longitudinal datasets. *Developmental Cognitive Neuroscience*, *55*, 101115. https://doi.org/10.1016/j.dcn.2022.101115

Satterthwaite, T. D., Vandekar, S., Wolf, D. H., Ruparel, K., Roalf, D. R., Jackson, C., Elliott, M. A., Bilker, W. B., Calkins, M. E., Prabhakaran, K., Davatzikos, C., Hakonarson, H., Gur, R. E., & Gur, R. C. (2014). Sex Differences in the Effect of Puberty on Hippocampal Morphology. *Journal of the American Academy of Child & Adolescent Psychiatry*, *53*(3), 341-350.e1. https://doi.org/10.1016/J.JAAC.2013.12.002

Sawyer, S. M., Azzopardi, P. S., Wickremarathne, D., & Patton, G. C. (2018). The age of adolescence. *The Lancet Child & Adolescent Health*, *2*(3), 223–228. https://doi.org/10.1016/S2352-4642(18)30022-1

Scheinost, D., Dadashkarimi, J., Finn, E. S., Wambach, C. G., MacGillivray, C., Roule, A. L., Niendam, T. A., Pine, D. S., Brotman, M. A., Leibenluft, E., & Tseng, W.-L. (2021). Functional connectivity during frustration: A preliminary study of predictive modeling of irritability in youth. *Neuropsychopharmacology*, *46*(7), 1300–1306. https://doi.org/10.1038/s41386-020-00954-8

Schlossberger, N. M., Turner, R. A., & Irwin, C. E. (1992). Validity of self-report of pubertal maturation in early adolescents. *The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine*, *13*(2), 109–113. https://doi.org/10.1016/1054-139x(92)90075-m

Schmaal, L., Hibar, D. P., Sämann, P. G., Hall, G. B., Baune, B. T., Jahanshad, N., Cheung, J. W., Van Erp, T. G. M., Bos, D., Ikram, M. A., Vernooij, M. W., Niessen, W. J., Tiemeier, H., Hofman, A., Wittfeld, K., Grabe, H. J., Janowitz, D., Bülow, R., Selonke, M., … Veltman, D. J. (2017). Cortical abnormalities in adults and adolescents with major depression based on brain scans from 20 cohorts worldwide in the ENIGMA Major Depressive Disorder Working Group. *Molecular Psychiatry*, *22*(6), 900–909. https://doi.org/10.1038/mp.2016.60

Schmaal, L., Pozzi, E., C Ho, T., van Velzen, L. S., Veer, I. M., Opel, N., Van Someren, E. J. W., Han, L. K. M., Aftanas, L., Aleman, A., Baune, B. T., Berger, K., Blanken, T. F., Capitão, L., Couvy-Duchesne, B., R Cullen, K., Dannlowski, U., Davey, C., Erwin-Grabner, T., … Veltman, D. J. (2020). ENIGMA MDD: seven years of global neuroimaging studies of

major depression through worldwide data sharing. *Translational Psychiatry*, *10*(1), 172. https://doi.org/10.1038/s41398-020-0842-6

Schmaal, L., Veltman, D. J., van Erp, T. G. M., Sämann, P. G., Frodl, T., Jahanshad, N., Loehrer, E., Tiemeier, H., Hofman, A., Niessen, W. J., Vernooij, M. W., Ikram, M. A., Wittfeld, K., Grabe, H. J., Block, A., Hegenscheid, K., Völzke, H., Hoehn, D., Czisch, M., … Hibar, D. P. (2016). Subcortical brain alterations in major depressive disorder: Findings from the ENIGMA Major Depressive Disorder working group. *Molecular Psychiatry*, *21*(6), Article 6. https://doi.org/10.1038/mp.2015.69

Sebastian, C., Viding, E., Williams, K. D., & Blakemore, S.-J. (2010). Social brain development and the affective consequences of ostracism in adolescence. *Brain and Cognition*, *72*(1), 134–145. https://doi.org/10.1016/j.bandc.2009.06.008

Ségonne, F., Dale, A. M., Busa, E., Glessner, M., Salat, D., Hahn, H. K., & Fischl, B. (2004). A hybrid approach to the skull stripping problem in MRI. *NeuroImage*, *22*(3), 1060–1075. https://doi.org/10.1016/j.neuroimage.2004.03.032

Sendi, M. S. E., Dini, H., Sui, J., Fu, Z., Qi, S., Riva-Posse, P., Abbott, C., Mayberg, H., & Calhoun, V. (2021). Dynamic Functional Connectivity Predicts Treatment Response to Electroconvulsive Therapy in Major Depressive Disorder. *Biological Psychiatry*, *89*(9), S169–S170. https://doi.org/10.1016/j.biopsych.2021.02.434

Serafini, G., Pompili, M., Borgwardt, S., Houenou, J., Geoffroy, P. A., Jardri, R., Girardi, P., & Amore, M. (2014). Brain changes in early-onset bipolar and unipolar depressive disorders: A systematic review in children and adolescents. *European Child and Adolescent Psychiatry*, *23*(11), 1023–1041. https://doi.org/10.1007/s00787-014-0614-z

Shad, M. U., Muddasani, S., & Rao, U. (2012). Gray matter differences between healthy and depressed adolescents: A voxel-based morphometry study. *Journal of Child and Adolescent Psychopharmacology*, *22*(3), 190–197. https://doi.org/10.1089/cap.2011.0005

Shen, X., Adams, M. J., Ritakari, T. E., Cox, S. R., McIntosh, A. M., & Whalley, H. C. (2019a). White matter microstructure and its relation to longitudinal measures of depressive symptoms in mid- and late life. *Biological Psychiatry*, *86*(10), 759–768. https://doi.org/10.1016/j.biopsych.2019.06.011

Shen, X., Adams, M. J., Ritakari, T. E., Cox, S. R., McIntosh, A. M., & Whalley, H. C. (2019b). White Matter Microstructure and Its Relation to Longitudinal Measures of Depressive Symptoms in Mid- and Late Life. *Biological Psychiatry*, *86*(10), 759–768. https://doi.org/10.1016/j.biopsych.2019.06.011

Shen, X., MacSweeney, N., Chan, S. W. Y., Barbu, M. C., Adams, M. J., Lawrie, S. M., Romaniuk, L., McIntosh, A. M., & Whalley, H. C. (2021). Brain structural associations with depression in a large early adolescent sample (the ABCD study®). *EClinicalMedicine*, *42*, 101204. https://doi.org/10.1016/j.eclinm.2021.101204

Shen, X., Reus, L. M., Cox, S. R., Adams, M. J., Liewald, D. C., Bastin, M. E., Smith, D. J., Deary, I. J., Whalley, H. C., & McIntosh, A. M. (2017). Subcortical volume and white matter integrity abnormalities in major depressive disorder: Findings from UK Biobank imaging data. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-05507-6

Sheridan, M. A., Mukerji, C. E., Wade, M., Humphreys, K. L., Garrisi, K., Goel, S., Patel, K., Fox, N. A., Zeanah, C. H., Nelson, C. A., & McLaughlin, K. A. (2022). Early deprivation alters structural brain development from middle childhood to adolescence. *Science*

Advances, *8*(40), eabn4316. https://doi.org/10.1126/sciadv.abn4316

Shirtcliff, E. A., Dahl, R. E., & Pollak, S. D. (2009). Pubertal development: Correspondence between hormonal and physical development. *Child Development*, *80*(2), 327–337. https://doi.org/10.1111/j.1467-8624.2009.01263.x

Shonkoff, J. P., Garner, A. S., Committee on Psychosocial Aspects of Child and Family Health, Committee on Early Childhood, Adoption, and Dependent Care, & Section on Developmental and Behavioral Pediatrics. (2012). The lifelong effects of early childhood adversity and toxic stress. *Pediatrics*, *129*(1), e232-246. https://doi.org/10.1542/peds.2011-2663

Simmonds, D. J., Hallquist, M. N., Asato, M., & Luna, B. (2014). Developmental stages and sex differences of white matter and behavioral development through adolescence: A longitudinal diffusion tensor imaging (DTI) study. *NeuroImage*, *92*, 356–368. https://doi.org/10.1016/j.neuroimage.2013.12.044

Skranes, J., Løhaugen, G. C. C., Martinussen, M., Håberg, A., Brubakk, A. M., & Dale, A. M. (2013). Cortical surface area and IQ in very-low-birth-weight (VLBW) young adults. *Cortex*, *49*(8), 2264–2271. https://doi.org/10.1016/j.cortex.2013.06.001

Solmi, M., Radua, J., Olivola, M., Croce, E., Soardo, L., Salazar de Pablo, G., Il Shin, J., Kirkbride, J. B., Jones, P., Kim, J. H., Kim, J. Y., Carvalho, A. F., Seeman, M. V., Correll, C. U., & Fusar-Poli, P. (2021). Age at onset of mental disorders worldwide: Large-scale meta-analysis of 192 epidemiological studies. *Molecular Psychiatry*, 1–15. https://doi.org/10.1038/s41380-021-01161-7

Solmi, M., Radua, J., Olivola, M., Croce, E., Soardo, L., Salazar de Pablo, G., Il Shin, J., Kirkbride, J. B., Jones, P., Kim, J. H., Kim, J. Y., Carvalho, A. F., Seeman, M. V., Correll, C. U., &

Fusar-Poli, P. (2022). Age at onset of mental disorders worldwide: Large-scale meta-analysis of 192 epidemiological studies. *Molecular Psychiatry*, *27*(1), Article 1. https://doi.org/10.1038/s41380-021-01161-7

Sonkusare, S., Breakspear, M., & Guo, C. (2019). Naturalistic Stimuli in Neuroscience: Critically Acclaimed. *Trends in Cognitive Sciences*, *23*(8), 699–714. https://doi.org/10.1016/j.tics.2019.05.004

Stein, A., Pearson, R. M., Goodman, S. H., Rapa, E., Rahman, A., McCallum, M., Howard, L. M., & Pariante, C. M. (2014). Effects of perinatal mental disorders on the fetus and child. *The Lancet*, *384*(9956), 1800–1819. https://doi.org/10.1016/S0140-6736(14)61277-0

Steinberg, L. (2008). A social neuroscience perspective on adolescent risk-taking. *Developmental Review*, *28*(1), 78–106. https://doi.org/10.1016/j.dr.2007.08.002

Stiles, J., & Jernigan, T. L. (2010). The Basics of Brain Development. *Neuropsychology Review*, *20*(4), 327–348. https://doi.org/10.1007/s11065-010-9148-4

Stirling, K., Toumbourou, J. W., & Rowland, B. (2015). Community factors influencing child and adolescent depression: A systematic review and meta-analysis. *The Australian and New Zealand Journal of Psychiatry*, *49*(10), 869–886. https://doi.org/10.1177/0004867415603129

Stoddard, J., Hsu, D., Reynolds, R. C., Brotman, M. A., Ernst, M., Pine, D. S., Leibenluft, E., & Dickstein, D. P. (2015). Aberrant amygdala intrinsic functional connectivity distinguishes youths with bipolar disorder from those with severe mood dysregulation. *Psychiatry Research: Neuroimaging*, *231*(2), 120–125. https://doi.org/10.1016/j.pscychresns.2014.11.006

Stoddard, J., Tseng, W.-L., Kim, P., Chen, G., Yi, J., Donahue, L., Brotman, M. A., Towbin, K. E.,

Pine, D. S., & Leibenluft, E. (2017). Association of Irritability and Anxiety With the Neural Mechanisms of Implicit Face Emotion Processing in Youths With Psychopathology. *JAMA Psychiatry*, *74*(1), 95. https://doi.org/10.1001/jamapsychiatry.2016.3282

Stringaris, A., Goodman, R., Ferdinando, S., Razdan, V., Muhrer, E., Leibenluft, E., & Brotman, M. A. (2012). The Affective Reactivity Index: A concise irritability scale for clinical and research settings. *Journal of Child Psychology and Psychiatry*, *53*(11), 1109–1117. https://doi.org/10.1111/j.1469-7610.2012.02561.x

Stringaris, A., Maughan, B., Copeland, W. S., Costello, E. J., & Angold, A. (2013). Irritable Mood as a Symptom of Depression in Youth: Prevalence, Developmental, and Clinical Correlates in the Great Smoky Mountains Study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *52*(8), 831–840. https://doi.org/10.1016/j.jaac.2013.05.017

Stringaris, A., Zavos, H., Leibenluft, E., Maughan, B., & Eley, T. C. (2012). Adolescent irritability: Phenotypic associations and genetic links with depressed mood. *The American Journal of Psychiatry*, *169*(1), 47–54. https://doi.org/10.1176/appi.ajp.2011.10101549

Sullivan, P. F., & Geschwind, D. H. (2019). Defining the Genetic, Genomic, Cellular, and Diagnostic Architectures of Psychiatric Disorders. *Cell*, *177*(1), 162–183. https://doi.org/10.1016/j.cell.2019.01.015

Sullivan, P. F., Neale, M. C., & Kendler, K. S. (2000). Genetic epidemiology of major depression: Review and meta-analysis. *American Journal of Psychiatry*, *157*(10), 1552–1562. https://doi.org/10.1176/appi.ajp.157.10.1552

Tagliazucchi, E., Balenzuela, P., Fraiman, D., & Chialvo, D. (2012). Criticality in Large-Scale

Brain fMRI Dynamics Unveiled by a Novel Point Process Analysis. *Frontiers in Physiology*, *3*. https://www.frontiersin.org/articles/10.3389/fphys.2012.00015

Tamnes, C. K., Bos, M. G. N., van de Kamp, F. C., Peters, S., & Crone, E. A. (2018). Longitudinal development of hippocampal subregions from childhood to adulthood. *Developmental Cognitive Neuroscience*, *30*, 212–222. https://doi.org/10.1016/j.dcn.2018.03.009

Tamnes, C. K., Herting, M. M., Goddings, A. L., Meuwese, R., Blakemore, S. J., Dahl, R. E., Güroğlu, B., Raznahan, A., Sowell, E. R., Crone, E. A., & Mills, K. L. (2017). Development of the cerebral cortex across adolescence: A multisample study of inter-related longitudinal changes in cortical volume, surface area, and thickness. *Journal of Neuroscience*, *37*(12), 3402–3412. https://doi.org/10.1523/JNEUROSCI.3302-16.2017

Tamnes, C. K., Roalf, D. R., Goddings, A.-L., & Lebel, C. (2018). Diffusion MRI of white matter microstructure development in childhood and adolescence: Methods, challenges and progress. *Developmental Cognitive Neuroscience*, *33*, 161–175. https://doi.org/10.1016/j.dcn.2017.12.002

Telzer, E. H., McCormick, E. M., Peters, S., Cosme, D., Pfeifer, J. H., & van Duijvenvoorde, A. C. K. (2018). Methodological considerations for developmental longitudinal fMRI research. *Developmental Cognitive Neuroscience*, *33*, 149–160. https://doi.org/10.1016/j.dcn.2018.02.004

Thapar, A., Collishaw, S., Pine, D. S., & Thapar, A. K. (2012). Depression in adolescence. *Lancet*, *379*(9820), 1056–1067. https://doi.org/10.1016/S0140-6736(11)60871-4

Thapar, A., Eyre, O., Patel, V., & Brent, D. (2022). Depression in young people. *The Lancet*. https://doi.org/10.1016/S0140-6736(22)01012-1

Thompson, P. (2019). Enigma, Big Data, and Neuroimaging Genetics In 50,000 People From 35 Countries: Challenges and Lessons Learned. *European Neuropsychopharmacology*. https://doi.org/10.1016/j.euroneuro.2017.06.131

Thompson, P. M., Stein, J. L., Medland, S. E., Hibar, D. P., Vasquez, A. A., Renteria, M. E., Toro, R., Jahanshad, N., Schumann, G., Franke, B., Wright, M. J., Martin, N. G., Agartz, I., Alda, M., Alhusaini, S., Almasy, L., Almeida, J., Alpert, K., Andreasen, N. C., … Drevets, W. (2014). The ENIGMA Consortium: Large-scale collaborative analyses of neuroimaging and genetic data. *Brain Imaging and Behavior*, *8*(2), 153–182. https://doi.org/10.1007/s11682-013-9269-5

Tian, Y. E., Cropley, V., Maier, A. B., Lautenschlager, N. T., Breakspear, M., & Zalesky, A. (2022). *Biological aging of human body and brain systems* (p. 2022.09.03.22279337). medRxiv. https://doi.org/10.1101/2022.09.03.22279337

Timpson, N. J., Greenwood, C. M. T., Soranzo, N., Lawson, D. J., & Richards, J. B. (2018). Genetic architecture: The shape of the genetic contribution to human traits and disease. *Nature Reviews Genetics*, *19*(2), Article 2. https://doi.org/10.1038/nrg.2017.101

Tisdall, M. D., Reuter, M., Qureshi, A., Buckner, R. L., Fischl, B., & van der Kouwe, A. J. W. (2016). Prospective motion correction with volumetric navigators (vNavs) reduces the bias and variance in brain morphometry induced by subject motion. *NeuroImage*, *127*, 11–22. https://doi.org/10.1016/j.neuroimage.2015.11.054

Tomova, L., Andrews, J. L., & Blakemore, S.-J. (2021). The importance of belonging and the avoidance of social risk taking in adolescence. *Developmental Review*, *61*, 100981. https://doi.org/10.1016/j.dr.2021.100981

Townsend, L., Kobak, K., Kearney, C., Milham, M., Andreotti, C., Escalera, J., Alexander, L., Gill, M. K., Birmaher, B., Sylvester, R., Rice, D., Deep, A., & Kaufman, J. (2020). Development of Three Web-Based Computerized Versions of the Kiddie Schedule for Affective Disorders and Schizophrenia Child Psychiatric Diagnostic Interview: Preliminary Validity Data. *Journal of the American Academy of Child and Adolescent Psychiatry*, *59*(2), 309–325. https://doi.org/10.1016/j.jaac.2019.05.009

Truelove-Hill, M., Erus, G., Bashyam, V., Varol, E., Sako, C., Gur, R. C., Gur, R. E., Koutsouleris, N., Zhuo, C., Fan, Y., Wolf, D. H., Satterthwaite, T. D., & Davatzikos, C. (2020). A Multidimensional Neural Maturation Index Reveals Reproducible Developmental Patterns in Children and Adolescents. *Journal of Neuroscience*, *40*(6), 1265–1275. https://doi.org/10.1523/JNEUROSCI.2092-19.2019

Tseng, W.-L., Deveney, C. M., Stoddard, J., Kircanski, K., Frackman, A. E., Yi, J. Y., Hsu, D., Moroney, E., Machlin, L., Donahue, L., Roule, A., Perhamus, G., Reynolds, R. C., Roberson-Nay, R., Hettema, J. M., Towbin, K. E., Stringaris, A., Pine, D. S., Brotman, M. A., & Leibenluft, E. (2019). Brain Mechanisms of Attention Orienting Following Frustration: Associations With Irritability and Age in Youths. *American Journal of Psychiatry*, *176*(1), 67–76. https://doi.org/10.1176/appi.ajp.2018.18040491

Tseng, W.-L., Moroney, E., Machlin, L., Roberson-Nay, R., Hettema, J. M., Carney, D., Stoddard, J., Towbin, K. A., Pine, D. S., Leibenluft, E., & Brotman, M. A. (2017). Test-retest reliability and validity of a frustration paradigm and irritability measures. *Journal of Affective Disorders*, *212*, 38–45. https://doi.org/10.1016/j.jad.2017.01.024

Tseng, W.-L., Thomas, L. A., Harkins, E., Pine, D. S., Leibenluft, E., & Brotman, M. A. (2016). Neural correlates of masked and unmasked face emotion processing in youth with

severe mood dysregulation. *Social Cognitive and Affective Neuroscience*, *11*(1), 78–88. https://doi.org/10.1093/scan/nsv087

Tuch, D. S., Reese, T. G., Wiegell, M. R., & Van J. Wedeen. (2003). Diffusion MRI of Complex Neural Architecture. *Neuron*, *40*(5), 885–895. https://doi.org/10.1016/S0896-6273(03)00758-X

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., & Joliot, M. (2002). Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain. *NeuroImage*, *15*(1), 273–289. https://doi.org/10.1006/nimg.2001.0978

Uddin, L. Q., Nomi, J. S., Hebert-Seropian, B., Ghaziri, J., & Boucher, O. (2017). Structure and function of the human insula. *Journal of Clinical Neurophysiology : Official Publication of the American Electroencephalographic Society*, *34*(4), 300–306. https://doi.org/10.1097/WNP.0000000000000377

Ullsperger, J. M., & Nikolas, M. A. (2017). A meta-analytic review of the association between pubertal timing and psychopathology in adolescence: Are there sex differences in risk? *Psychological Bulletin*, *143*(9), 903–938. https://doi.org/10.1037/bul0000106

United Nations. (2017). *World Population Prospects 2017*. UN Department for Economic and Social Affairs (UN DESA), Population Division.

Urošević, S., Collins, P., Muetzel, R., & Lim, K. O. (2014). Pubertal status associations with reward and threat sensitivities and subcortical brain volumes during adolescence. *Brain and Cognition*, *89*, 15–26. https://doi.org/10.1016/J.BANDC.2014.01.007

van Duijvenvoorde, A. C. K., Achterberg, M., Braams, B. R., Peters, S., & Crone, E. A. (2016). Testing a dual-systems model of adolescent brain development using resting-state

connectivity analyses. *NeuroImage*, *124*, 409–420. https://doi.org/10.1016/j.neuroimage.2015.04.069

van Duijvenvoorde, A. C. K., Westhoff, B., de Vos, F., Wierenga, L. M., & Crone, E. A. (2019). A three-wave longitudinal study of subcortical–cortical resting-state connectivity in adolescence: Testing age- and puberty-related changes. *Human Brain Mapping*, *40*(13), 3769–3783. https://doi.org/10.1002/hbm.24630

van Velzen, L. S., Kelly, S., Isaev, D., Aleman, A., Aftanas, L. I., Bauer, J., Baune, B. T., Brak, I. V., Carballedo, A., Connolly, C. G., Couvy-Duchesne, B., Cullen, K. R., Danilenko, K. V., Dannlowski, U., Enneking, V., Filimonova, E., Förster, K., Frodl, T., Gotlib, I. H., … Schmaal, L. (2020). White matter disturbances in major depressive disorder: A coordinated analysis across 20 international cohorts in the ENIGMA MDD working group. *Molecular Psychiatry*, *25*(7), 1511–1525. https://doi.org/10.1038/s41380-019-0477-2

Vanes, L., & Nosarti, C. (2022). Network Topology and Psychopathology Following Very Preterm Birth. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *7*(4), 349–351. https://doi.org/10.1016/j.bpsc.2021.12.011

Vértes, P. E., & Bullmore, E. T. (2015). Annual Research Review: Growth connectomics – the organization and reorganization of brain networks during normal and abnormal development. *Journal of Child Psychology and Psychiatry*, *56*(3), 299–320. https://doi.org/10.1111/jcpp.12365

Vidal-Pineiro, D., Parker, N., Shin, J., French, L., Grydeland, H., Jackowski, A. P., Mowinckel, A. M., Patel, Y., Pausova, Z., Salum, G., Sørensen, Ø., Walhovd, K. B., Paus, T., Fjell, A. M., & the Alzheimer's Disease Neuroimaging Initiative and the Australian Imaging

Biomarkers and Lifestyle flagship study of ageing. (2020). Cellular correlates of cortical thinning throughout the lifespan. *Scientific Reports*, *10*(1), 21803. https://doi.org/10.1038/s41598-020-78471-3

Vidal-Pineiro, D., Wang, Y., Krogsrud, S. K., Amlien, I. K., Baaré, W. F., Bartres-Faz, D., Bertram, L., Brandmaier, A. M., Drevon, C. A., Düzel, S., Ebmeier, K., Henson, R. N., Junqué, C., Kievit, R. A., Kühn, S., Leonardsen, E., Lindenberger, U., Madsen, K. S., Magnussen, F., … Fjell, A. (2021). Individual variations in 'brain age' relate to early-life factors more than to longitudinal brain change. *ELife*, *10*, e69995. https://doi.org/10.7554/eLife.69995

Vidal-Ribas, P., & Stringaris, A. (2021). How and Why Are Irritability and Depression Linked? *Child and Adolescent Psychiatric Clinics of North America*, *30*(2), 401–414. https://doi.org/10.1016/j.chc.2020.10.009

Vijayakumar, N., Allen, N. B., Youssef, G., Dennison, M., Yücel, M., Simmons, J. G., & Whittle, S. (2016). Brain development during adolescence: A mixed-longitudinal investigation of cortical thickness, surface area, and volume. *Human Brain Mapping*, *37*(6), 2027–2038. https://doi.org/10.1002/hbm.23154

Vijayakumar, N., Op de Macks, Z., Shirtcliff, E. A., & Pfeifer, J. H. (2018). Puberty and the human brain: Insights into adolescent development. *Neuroscience and Biobehavioral Reviews*, *92*, 417–436. https://doi.org/10.1016/j.neubiorev.2018.06.004

Vijayakumar, N., Youssef, G., Allen, N. B., Anderson, V., Efron, D., Mundy, L., Patton, G., Simmons, J. G., Silk, T., & Whittle, S. (2021). The effects of puberty and its hormones on subcortical brain development. *Comprehensive Psychoneuroendocrinology*, *7*, 100074. https://doi.org/10.1016/j.cpnec.2021.100074

Vohryzek, J., Deco, G., Cessac, B., Kringelbach, M. L., & Cabral, J. (2020). Ghost Attractors in Spontaneous Brain Activity: Recurrent Excursions Into Functionally-Relevant BOLD Phase-Locking States. *Frontiers in Systems Neuroscience*, *14*. https://www.frontiersin.org/article/10.3389/fnsys.2020.00020

Volker, D., Zijlstra-Vlasveld, M. C., Brouwers, E. P. M., Homans, W. A., Emons, W. H. M., & van der Feltz-Cornelis, C. M. (2016). Validation of the Patient Health Questionnaire-9 for Major Depressive Disorder in the Occupational Health Setting. *Journal of Occupational Rehabilitation*, *26*, 237–244. https://doi.org/10.1007/s10926-015-9607-0

Vulser, H., Martinot, M. L. P., Artiges, E., Miranda, R., Penttilä, J., Grimmer, Y., Van Noort, B. M., Stringaris, A., Struve, M., Fadai, T., Kappel, V., Goodman, R., Tzavara, E., Massaad, C., Banaschewski, T., Barker, G. J., Bokde, A. L. W., Bromberg, U., Brühl, R., … Lemaitre, H. (2018). Early variations in white matter microstructure and depression outcome in adolescents with subthreshold depression. *American Journal of Psychiatry*, *175*(12), 1255–1264. https://doi.org/10.1176/appi.ajp.2018.17070825

Waller, L., Erk, S., Pozzi, E., Toenders, Y. J., Haswell, C. C., Büttner, M., Thompson, P. M., Schmaal, L., Morey, R. A., Walter, H., & Veer, I. M. (2022). ENIGMA HALFpipe: Interactive, reproducible, and efficient analysis for resting-state and task-based fMRI data. *Human Brain Mapping*, *43*(9), 2727–2742. https://doi.org/10.1002/hbm.25829

Weathersby, F. L., King, J. B., Fox, J. C., Loret, A., & Anderson, J. S. (2019). Functional connectivity of emotional well-being: Overconnectivity between default and attentional networks is associated with attitudes of anger and aggression. *Psychiatry Research: Neuroimaging*, *291*, 52–62. https://doi.org/10.1016/j.pscychresns.2019.08.001

White, N., Roddey, C., Shankaranarayanan, A., Han, E., Rettmann, D., Santos, J., Kuperman, J., & Dale, A. (2010). PROMO: Real-time prospective motion correction in MRI using image-based tracking. *Magnetic Resonance in Medicine*, *63*(1), 91–105. https://doi.org/10.1002/mrm.22176

White, N. S., Leergaard, T. B., D'Arceuil, H., Bjaalie, J. G., & Dale, A. M. (2013). Probing tissue microstructure with restriction spectrum imaging: Histological and theoretical validation. *Human Brain Mapping*, *34*(2), 327–346. https://doi.org/10.1002/hbm.21454

Whitmore, L. B., & Mills, K. L. (2021). Co-creating developmental science. *Infant and Child Development*, e2273. https://doi.org/10.1002/icd.2273

Whittle, S., Yücel, M., Lorenzetti, V., Byrne, M. L., Simmons, J. G., Wood, S. J., Pantelis, C., & Allen, N. B. (2012). Pituitary volume mediates the relationship between pubertal timing and depressive symptoms during adolescence. *Psychoneuroendocrinology*, *37*(7), 881–891. https://doi.org/10.1016/j.psyneuen.2011.10.004

Wierenga, L. M., Bos, M. G. N., Schreuders, E., vd Kamp, F., Peper, J. S., Tamnes, C. K., & Crone, E. A. (2018). Unraveling age, puberty and testosterone effects on subcortical brain development across adolescence. *Psychoneuroendocrinology*, *91*, 105–114. https://doi.org/10.1016/j.psyneuen.2018.02.034

Wierenga, L. M., Langen, M., Oranje, B., & Durston, S. (2014). Unique developmental trajectories of cortical thickness and surface area. *NeuroImage*, *87*, 120–126. https://doi.org/10.1016/j.neuroimage.2013.11.010

Wiggins, J. L., Brotman, M. A., Adleman, N. E., Kim, P., Oakes, A. H., Reynolds, R. C., Chen, G., Pine, D. S., & Leibenluft, E. (2016). Neural Correlates of Irritability in Disruptive Mood

Dysregulation and Bipolar Disorders. *The American Journal of Psychiatry*, *173*(7), 722–730. https://doi.org/10.1176/appi.ajp.2015.15060833

Williams, K. D., Cheung, C. K. T., & Choi, W. (2000). Cyberostracism: Effects of being ignored over the Internet. *Journal of Personality and Social Psychology*, *79*, 748–762. https://doi.org/10.1037/0022-3514.79.5.748

Wilson, S., & Dumornay, N. M. (2022). Rising Rates of Adolescent Depression in the United States: Challenges and Opportunities in the 2020s. *Journal of Adolescent Health*, *70*(3), 354–355. https://doi.org/10.1016/j.jadohealth.2021.12.003

Winkler, A. M., Kochunov, P., Blangero, J., Almasy, L., Zilles, K., Fox, P. T., Duggirala, R., & Glahn, D. C. (2010). Cortical thickness or grey matter volume? The importance of selecting the phenotype for imaging genetics studies. *NeuroImage*, *53*(3), 1135–1146. https://doi.org/10.1016/j.neuroimage.2009.12.028

World Health Organisation. (2012). Depression: A Global Crisis. In *World Federation for Mental Health*.

Wyshak, G., & Frisch, R. E. (1982). Evidence for a secular trend in age of menarche. *New England Journal of Medicine*, *306*, 1033–1035.

Xu, J., Lyu, H., Li, T., Xu, Z., Fu, X., Jia, F., Wang, J., & Hu, Q. (2019). Delineating functional segregations of the human middle temporal gyrus with resting-state functional connectivity and coactivation patterns. *Human Brain Mapping*, *40*(18), 5159–5171. https://doi.org/10.1002/hbm.24763

Yendiki, A., Koldewyn, K., Kakunoori, S., Kanwisher, N., & Fischl, B. (2014). Spurious group differences due to head motion in a diffusion MRI study. *NeuroImage*, *88*, 79–90. https://doi.org/10.1016/j.neuroimage.2013.11.027

Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zöllei, L., Polimeni, J. R., Fischl, B., Liu, H., & Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, *106*(3), 1125–1165. https://doi.org/10.1152/jn.00338.2011

Yun, H. J., Im, K., Yang, J.-J., Yoon, U., & Lee, J.-M. (2013). Automated Sulcal Depth Measurement on Cortical Surface Reflecting Geometrical Properties of Sulci. *PLOS ONE*, *8*(2), e55977. https://doi.org/10.1371/journal.pone.0055977

Zalesky, A., Fornito, A., Cocchi, L., Gollo, L. L., & Breakspear, M. (2014). Time-resolved resting-state brain networks. *Proceedings of the National Academy of Sciences*, *111*(28), 10341–10346. https://doi.org/10.1073/pnas.1400181111

Zanolie, K., Ma, I., Bos, M. G. N., Schreuders, E., Vandenbroucke, A. R. E., van Hoorn, J., van Duijvenvoorde, A. C. K., Wierenga, L., Crone, E. A., & Güroğlu, B. (2022). Understanding the Dynamics of the Developing Adolescent Brain Through Team Science. *Frontiers in Integrative Neuroscience*, *16*. https://www.frontiersin.org/article/10.3389/fnint.2022.827097

Zanto, T. P., & Gazzaley, A. (2013). Fronto-parietal network: Flexible hub of cognitive control. *Trends in Cognitive Sciences*, *17*(12). https://doi.org/10.1016/j.tics.2013.10.001

Zhang, H., Schneider, T., Wheeler-Kingshott, C. A., & Alexander, D. C. (2012). NODDI: Practical in vivo neurite orientation dispersion and density imaging of the human brain. *NeuroImage*, *61*(4), 1000–1016. https://doi.org/10.1016/j.neuroimage.2012.03.072

Zhuang, J., Hrabe, J., Kangarlu, A., Xu, D., Bansal, R., Branch, C. A., & Peterson, B. S. (2006). Correction of eddy-current distortions in diffusion tensor images using the known

directions and strengths of diffusion gradients. *Journal of Magnetic Resonance Imaging*, *24*(5), 1188–1193. https://doi.org/10.1002/jmri.20727