# A lifespan perspective on brain-behavioural heterogeneity following very preterm birth

Hadaya, Laila

*Awarding institution:*
King's College London

# A lifespan perspective on brain-behavioural heterogeneity following very preterm birth


**Laila Hadaya**

Supervisors: Professor Chiara Nosarti & Dr Dafnis Batalle

Department of Child and Adolescent Psychiatry

Institute of Psychiatry Psychology and Neuroscience

King's College London

*Thesis submitted for the degree of Doctor of Philosophy in Developmental Neuroscience and Biological Psychiatry*

January 2024

# Contents

# List of figures

# List of supplementary figures

# List of tables

# List of supplementary tables

# Acknowledgements

I am deeply thankful for my supervisors Professor Chiara Nosarti and Dr Dafnis Batalle for their patience, guidance, and support over the years and for inspiring and motivating me both academically and professionally. I am beyond grateful for being a part of their stimulating and accommodating labs and for the invaluable experiences I was exposed to while working amongst world-leading researchers at the Child and Adolescent Psychiatry, Forensic and Neurodevelopmental Sciences, and Perinatal Imaging departments. I am also very appreciative of the clinical, radiography, and administrative teams at the Centre for the Developing Brain and the participant families for their substantial contributions to the research studies, and the Medical Research Council for funding the research. I am additionally very grateful for the transformative opportunity of working with exceptional collaborators at the Centre for Neuroimaging Sciences and the NIHR BRC Translational Bioinformatics Platform, especially František Váša and Konstantina Dimitrakopoulou, who were very knowledgeable, attentive, and inspirational mentors and played pivotal roles in my educational journey throughout this PhD. My deepest and warmest thanks are also extended to my colleagues-turned-friends, Dana, Lucy, Marguerite, and the CoDe-Neuro family, who taught me so much and were always there for moral support. Finally, my eternal love and gratitude go to my family and friends, both near and far. Thank you for the unconditional love and support and for cheering me on and keeping me sane and strong, I truly could not have done this without you.

# Abstract

**Background:** Very preterm birth (VPT; at ≤ 32 weeks' gestation) occurs during a highly critical stage of brain development, which makes the VPT new-born brain highly vulnerable to insult and long-lasting neurodevelopmental sequelae. Structural and functional brain alterations may be at least partly responsible for the behavioural difficulties described in VPT individuals across the lifespan. However, there is marked heterogeneity in the extent and presence of behavioural difficulties amongst VPT individuals, making it challenging to identify those vulnerable to developing mental health problems and cognitive difficulties. Hence, identifying underlying neurobiological markers of specific behavioural outcomes could help recognize those VPT individuals who may benefit from targeted support.

**Objective:** The overarching objective of this PhD thesis is to stratify the heterogeneity in behavioural outcomes exhibited by VPT individuals, explore structural and functional brain alterations which may characterise distinct behavioural subgroups, and investigate the influence of clinical and environmental factors. The thesis is organised into four experimental studies addressing the following specific aims:

- **Study #1:** to identify differences in neonatal structural brain volumes in subgroups of VPT born children screening negatively and positively for autism spectrum conditions and to explore the role of developmental delay in mediating and moderating childhood autism traits.
- **Study #2:** to parse heterogeneity in neonatal clinical and social risk and childhood behavioural outcomes using data-driven integrative consensus clustering and to explore differences in neonatal brain volumes and structural and functional connectivity between the resultant subgroups.
- **Study #3:** to compare resting state functional connectivity and structural volumes between groups of children stratified both in terms of their clinical characteristics (i.e., VPT and full-term (FT) birth) and their behavioural profiles identified using data-driven consensus clustering regardless of their gestational age at birth. Post-hoc analyses aimed to elucidate whether clinical and environmental factors differ within VPT or FT children belonging to distinct subgroups.
- **Study #4:** to use the same approach employed in Study #3 to investigate resting state functional connectivity in a sample of adults born VPT and FT.

**Methodology:** Distinct psychometric screening criteria (Study #1) and advanced data-driven clustering approaches (Studies #2-4) were used to parse behavioural heterogeneity at different time points throughout development: toddlerhood (Study #1), early childhood (Study #2), middle childhood (Study #3), and adulthood (Study #4). Magnetic resonance (MR) imaging data were collected and analysed using advanced whole-brain approaches such as Tensor Based Morphometry, Tract Based Spatial Statistics, graph theory metrics, and Network Based Statistic to quantify brain volumes and structural and functional connectivity patterns characterising behavioural heterogeneity.

**Study participants:** Participants were drawn from two cohort studies: (i) the Evaluation of Preterm Imaging Study (ePrime), which evaluated brain development using multi-modal MR imaging (at term equivalent age and at 7-12 years) and behavioural assessments (at 2, 4-7, and 7-12 years); and (ii) the University College Hospital London (UCHL) study, which studied adults (median age 30 years) using multi-modal MR imaging and behavioural assessments.

**Results:**

- **Study #1:** Smaller neonatal brainstem volumes and high levels of developmental delay were seen in one of two subgroups of VPT toddlers screening positively for autism according to different psychometric criteria, relative to those screening negatively. Developmental delay in this positively screening subgroup was seen to be mediating and moderating the onset of autism traits in childhood. Smaller neonatal cerebellar volumes differentiated between the two distinct subgroups of VPT children screening positively for autism, despite exhibiting similar extent of autism traits in early childhood. Together, results suggest the presence of distinct aetiological trajectories associated with autism traits.
- **Study #2:** In early childhood, three data-driven behavioural subgroups of VPT children were delineated: **(i) a '*Resilient*' subgroup** with favourable behavioural outcomes and a more cognitively stimulating home environment, **(ii) an '*At-risk*' subgroup** with behavioural difficulties and high neonatal clinical risk, and **(iii) an '*Intermediate*' subgroup** with intermediate behavioural outcomes, low neonatal clinical risk and a less cognitively stimulating home environment. Relative to the '*Intermediate*' subgroup, the '*Resilient*' subgroup displayed larger neonatal fronto-limbic regional volumes and functional connectivity and the '*At-risk*' subgroup showed widespread fronto-temporo-limbic white matter alterations. These findings highlight the value of studying neonatal patterns of functional and structural brain development as potential biomarkers of childhood

outcomes, as well as the importance of a supportive home environment to foster child development following VPT birth.

- **Study #3:** Results show evidence of widespread volumetric alterations and increased functional connectivity in VPT children relative to controls. In middle childhood, stratifying the sample into two data-driven behavioural subgroups, independently of birth status identified: **(i) a '*General Difficulties*' subgroup** displaying widespread decreases in functional connectivity and greater behavioural difficulties relative to a **(ii) '*General Resilience*' subgroup**. A three-subgroup solution was also explored, identifying: a **(I) '*Neurodevelopmental Difficulties*'** subgroup with socio-emotional and higher-order cognitive difficulties and reduced rostro-lateral prefrontal, brainstem, occipital, and cerebellar volumes, **and a (II) '*Psychiatric Difficulties*'** subgroup exhibiting psychiatric and executive function difficulties with reduced dorsolateral prefrontal and cerebellar volumes, relative to a **(III) '*Typical Development*' subgroup**. All brain differences, apart from cerebellar and occipital volumetric alterations, significantly differentiated between distinct data-driven behavioural subgroups after adjusting for preterm birth, highlighting the presence of VPT-specific neural alterations as well as unique neural patterns underlying behavioural difficulties in the general population, independently of birth status. Furthermore, VPT (but not FT) children belonging to Neurodevelopmental Difficulties or Psychiatric Difficulties subgroups displayed greater social risk relative to those in the Typical Development subgroup.
- **Study #4:** Complex widespread patterns of both increased and decreased functional connectivity were found in VPT compared to FT born adults in default mode, visual, and ventral attention networks. In adulthood, when VPT and FT born adults were stratified in terms of their behavioural profiles (irrespective of preterm birth), two data-driven subgroups were identified: **(i) an '*At-risk*' subgroup** with more behavioural difficulties and reduced functional connectivity in frontal opercular and insular areas relative to a **(ii) '*Resilient*' subgroup** with more favourable behavioural outcomes. These results indicate that functional connectivity between the default mode, ventral attention, and visual networks characterise clinically defined groups (VPT and FT) and are different from the connectivity patterns that characterise adults subdivided in terms of their behavioural profiles (irrespective of birth group), which are anchored in insular and frontal opercular regions. Moreover, social risk was found to be greater within adults born VPT belonging to At-risk relative to Resilient subgroups, while this relationship was not identified in those born FT.

**Conclusions:** Collectively, these findings indicate a long-lasting presence of neurodevelopmental heterogeneity within VPT and FT samples which seems to persist throughout the lifespan. Specific structural and functional alterations differentiating between distinct behavioural profiles across both VPT and FT samples are also identified; whereby alterations localised to fronto-temporo-limbic regions seem to be characteristic of behavioural difficulties in VPT and FT samples regardless of their birth status, while alterations to brain regions including visual and cerebellar areas may be characterising biomarkers of outcomes specifically in VPT samples. Implications of these findings suggest a potential benefit of using MRI to detect neurobiological markers of behavioural outcomes, which can in turn guide the implementation of personalised interventions for those at-risk of developing specific behavioural difficulties. Results presented here also highlight the importance of fostering an enriching environment to probe resilience against developing behavioural difficulties, particularly for those born VPT.

# Authorship declaration

I hereby declare that all the work carried out and presented in this thesis is my own, unless otherwise stated where acknowledgement is given to the authors. The work contributing towards this thesis was completed between February 2020 and January 2024 at the Centre for the Developing Brain, Department of Perinatal Imaging, St Thomas' Hospital, King's College London, London, UK and the Department of Child and Adolescent Psychiatry, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK.

This is a PhD thesis incorporating peer-reviewed publications (Study #1 and Study #2). This thesis also includes work which has been submitted to peer-reviewed journals for consideration or is in preparation for submission (Study #3 and Study #4) but does not include any content submitted for consideration for any other degree or qualification. Please note that co-author contributions do not affect the nature of this thesis being primarily my own work and that I am responsible for the methodological design, data collection, data cleaning, data pre-processing, data analysis, result interpretation, and writing incorporated within this thesis, with the exception of data collection for Studies #1, #2, and #4 and diffusion and functional MRI data pre-processing for Study #2 which were completed before the start of this PhD by previous researchers at the Centre for the Developing Brain and the Neurodevelopment and Mental Health Group, under the supervision of Professor Chiara Nosarti and Professor A David Edwards.

# Paper publications and conference attendance

## List of publications included in this thesis

**Study #1 – Chapter 3, published:**

**Hadaya, L.**, Vanes, L., Karolis, V., Kanel, D., Leoni, M., Happé, F., Edwards, A. D., Counsell, S. J., Batalle, D., & Nosarti, C. (2022). Distinct Neurodevelopmental Trajectories in Groups of Very Preterm Children Screening Positively for Autism Spectrum Conditions. Journal of Autism and Developmental Disorders. https://doi.org/10.1007/s10803-022-05789-4

> *Co-author contributions:* ***LH****, KD, MS, DB and CN contributed to the concept and design of the study.* ***LH****, KD, LDV, DK, SFM, OGG, SJC, MS, DB and CN contributed towards material preparation or data analysis and* ***LH****, KD, MS, DB and CN to the interpretation of data for the study. Data collection was completed by the ePrime study research team supervised by ADE and CN.* ***LH****, KD, MS, DB and CN drafted the manuscript and revised it critically for important intellectual content. All authors approved the final manuscript to be published.*

**Study #2 – Chapter 4, published:**

**Hadaya, L.**, Dimitrakopoulou, K., Vanes, L. D., Kanel, D., Fenn-Moltu, S., Gale-Grant, O., Counsell, S. J., Edwards, A. D., Saqi, M., Batalle, D., & Nosarti, C. (2023). Parsing brain-behavior heterogeneity in very preterm born children using integrated similarity networks. Translational Psychiatry, 13(1), Article 1. https://doi.org/10.1038/s41398-023-02401-w

> *Co-author contributions: All authors contributed to the study conception and design. Material preparation and data analysis were performed by* ***LH****, LV, VK, CN, DK, and ML. Data collection at baseline, at the 2-year- and 4–7-year-old follow up was completed by the e-Prime study research team supervised by ADE and CN. The first draft of the manuscript was written by* ***LH*** *and CN and was commented on by all authors of the manuscript. The final manuscript was read and approved by all authors.*

**Study #3 – Chapter 5, in preparation:**

**Hadaya, L.**, Váša, F., Kanel, D., Shi, L., Leoni, M., Dimitrakopoulou, K., Saqi, M, Edwards, A. D., Counsell, S. J., Leech, R., Batalle, D., & Nosarti, C. Exploring brain structure and function in clinical and data-driven groups of preterm and term children. [*Manuscript in preparation*].

*Co-author contributions:* ***LH****, FV, KD, MS, ADE, SJC, RL, DB, and CN contributed to the concept and design of the study.* ***LH****, FV, LS, DK, RL, DB, and CN contributed towards material preparation or data analysis. Data collection was completed by* ***LH****, DK, and ML. LH, FV, DB, and CN contributed towards the interpretation of data for the study.* ***LH****, FV, DB, and CN drafted the manuscript and revised it critically for important intellectual content. All authors approved the final manuscript to be published.*

**Study #4 – Chapter 6, submitted for publication and uploaded as pre-print:**

**Hadaya, L.,** Váša, F., Dimitrakopoulou, K., Saqi, M, Shergill, S.S., Edwards, A. D., Batalle, D., Leech, R., & Nosarti, C. (2024). Exploring functional connectivity in clinical and data-driven groups of preterm and term adults. (p. 2024.01.22.576651). bioRxiv. https://doi.org/10.1101/2024.01.22.576651

*Co-author contributions:* ***LH****, FV, KD, MS, ADE, SJC, RL, DB, and CN contributed to the concept and design of the study.* ***LH****, FV, LS, DK, RL, DB, and CN contributed towards material preparation or data analysis. Data collection was completed by* ***LH****, DK, and ML. LH, FV, DB, and CN contributed towards the interpretation of data for the study.* ***LH****, FV, DB, and CN drafted the manuscript and revised it critically for important intellectual content. All authors approved the final manuscript to be published.*

**Review article – Appendix C, published:**

**Hadaya, L**., & Nosarti, C. (2020). The neurobiological correlates of cognitive outcomes in adolescence and adulthood following very preterm birth. Seminars in Fetal & Neonatal Medicine, 25(3), 101117. https://doi.org/10.1016/j.siny.2020.101117

## List of additional studies published during period of PhD completion but are not for consideration here

Kanel, D., Vanes, L. D., Ball, G., **Hadaya, L.**, Falconer, S., Counsell, S. J., Edwards, A. D., & Nosarti, C. (2022). Neonatal amygdala resting-state functional connectivity and socio-emotional development in very preterm children. Brain Communications, 4(1), fcac009. https://doi.org/10.1093/braincomms/fcac009

Kanel, D., Vanes, L., Pecheva, D., **Hadaya, L.**, Falconer, S., Counsell, S., Edwards, D., & Nosarti, C. (2021). Neonatal white matter microstructure and emotional development during the pre-school years in children who were born very preterm. eNeuro, 8(5), ENEURO.0546-20.2021. https://doi.org/10.1523/ENEURO.0546-20.2021

Krajner, F., **Hadaya, L.**, McQueen, G., Sendt, K.-V., Gillespie, A., Avila, A., Lally, J., Hedges, E. P., Diederen, K., Howes, O. D., Barker, G. J., Lythgoe, D. J., Kempton, M. J., McGuire, P., MacCabe, J. H., & Egerton, A. (2022). Subcortical volume reduction and cortical thinning 3 months after switching to clozapine in treatment resistant schizophrenia. NPJ Schizophrenia, 8(1), 13. https://doi.org/10.1038/s41537-022-00230-2

Lautarescu, A., **Hadaya, L.**, Craig, M. C., Makropoulos, A., Batalle, D., Nosarti, C., Edwards, A. D., Counsell, S. J., & Victor, S. (2021). Exploring the relationship between maternal prenatal stress and brain structure in premature neonates. PloS One, 16(4), e0250413. https://doi.org/10.1371/journal.pone.0250413

Leoni, M., Vanes, L. D., **Hadaya, L.**, Kanel, D., Dazzan, P., Simonoff, E., Counsell, S. J., Happé, F., Edwards, A. D., & Nosarti, C. (2023). Exploring cognitive, behavioral and autistic trait network topology in very preterm and term-born children. Frontiers in Psychology, 14. https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1119196

Lovato, I., Vanes, L. D., Sacchi, C., Simonelli, A., **Hadaya, L.**, Kanel, D., Falconer, S., Counsell, S., Redshaw, M., Kennea, N., Edwards, A. D., & Nosarti, C. (2022). Early childhood temperamental trajectories following very preterm birth and their association with parenting style. Children, 9(4), 508. https://doi.org/10.3390/children9040508

Sun, Z., **Hadaya, L.**, Leoni, M., Dazzan, P., Simonoff, E., Counsell, S. J., Edwards, A. D., Nosarti, C., & Vanes, L. (2023). Comparing the emotional impact of the UK COVID-19 lockdown in very

preterm and full-term born children: A longitudinal study. Frontiers in Child and Adolescent Psychiatry, 2. https://www.frontiersin.org/articles/10.3389/frcha.2023.1193258

Vanes, L. D., **Hadaya, L.**, Kanel, D., Falconer, S., Ball, G., Batalle, D., Counsell, S. J., Edwards, A. D., & Nosarti, C. (2021). Associations Between Neonatal Brain Structure, the Home Environment, and Childhood Outcomes Following Very Preterm Birth. Biological Psychiatry Global Open Science, 1(2), 146–155. https://doi.org/10.1016/j.bpsgos.2021.05.002

Vanes, L., Fenn-Moltu, S., **Hadaya, L.**, Fitzgibbon, S., Cordero-Grande, L., Price, A., Chew, A., Falconer, S., Arichi, T., Counsell, S. J., Hajnal, J. V., Batalle, D., Edwards, A. D., & Nosarti, C. (2023). Longitudinal neonatal brain development and socio-demographic correlates of infant outcomes following preterm birth. Developmental Cognitive Neuroscience, 61, 101250. https://doi.org/10.1016/j.dcn.2023.101250

## List of conference attendance and poster presentations

2022 Flux Congress

2021 Flux (Virtual) Congress

2021 Paediatric Academic Societies (Virtual) Meeting

2021 Society for Research in Child Development (Virtual) Biennial Meeting

2020 Flux (Virtual) Congress

## List of abbreviations

ADHD = Attention Deficit/Hyperactive Disorder

ANTs = Advanced Normalization Tools

ASC = Autism Spectrum Conditions

AQ-10 = Autism Quotient

Bayley-III = Bayley Scales of Infant Development, Third Edition

BOLD = Blood Oxygen Level Dependent

BRIEF-2 = Behavior Rating Inventory of Executive Function, Second Edition

BRIEF-P = Behavior Rating Inventory of Executive Function pre-school version

CANTAB = Cambridge Neurophysiological Test Automated Battery

CAARMS = Comprehensive Assessment of At-Risk Mental States

CBQ = Very Short Form - Childhood Behavioral Questionnaire

COWAT = Controlled Oral Word Association Test

CPAP = Continuous Positive Airway Pressure

CPT = Continuous Performance Test

CSPS = Cognitive Stimulating Parenting Scale

DAWBA = Development and Well-Being Assessment

DMN = Default Mode Network

DSM = Diagnostic and Statistical Manual

d-MRI = Diffusion MRI

EmQue = Empathy Questionnaire

ERC = Emotion Regulation Checklist

ERT = Emotion Recognition Task

ePrime = Evaluation of Preterm Imaging

EPT = Extremely Preterm

FA = Fractional Anisotropy

FC = Functional Connectivity

FD = Framewise Displacement

FDR = False Discovery Rate

f-MRI = Functional MRI

FSL = FMRIB Software Library

FT = Full-term

FWER = Family Wise Error Rate

FWHM = Full Width Half Maximum

GA = Gestational Age

GHQ = General Health Questionnaire

GLM = General Linear Model

HCP-MMP = Human Connectome Project Multi-Modal Parcellation

HSCT = Hayling Sentence Completion Test

ICA = Independent Component Analysis

ICD = International Classification of Diseases

IED = Intra-Extra Dimensional Set Shift

IMD = Index of Multiple Deprivation

IQ = Intelligence Quotient

LMPT = Late-to-Moderate Preterm

M-CHAT = Modified Checklist for Autism in Toddlers

MD = Mean diffusivity

MOT = Motor Screening Task

MRI = Magnetic Resonance Imaging

NBS = Network Based Statistic

PAL = Paired Associates Learning

PARCA-R = Parent Report of Children's Abilities Revised

PCA = Principal Component Analysis

PMA = Post Menstrual Age (at Scan)

PDI = Peters Delusion Inventory

PQ-BC = Prodromal Questionnaire–Brief Child Version

ROI = Region-of-Interest

RFS = Role Functioning Scale

RRB = Restricted/Repetitive Behaviour

rsFC = Resting State Functional Connectivity

rs-fMRI = Resting State Functional MRI

SAS = Social Adjustment Scale

SCAS = Spence Children's Anxiety Scale

SCI = Social Communication/Interaction

SDQ = Strengths and Difficulties Questionnaire

s-MRI = Structural MRI

SNF = Similarity Network Fusion

SOC = Stockings of Cambridge

SRS-2 = Social Responsiveness Scale – Second Edition

SyN = Symmetric Normalisation

TBM = Tensor Based Morphometry

TBSS = Tract Based Spatial Statistics

TEA = Term Equivalent Age

TFCE = Threshold-Free Cluster Enhancement

TMCQ = Temperament in Middle Childhood Questionnaire

TMT-B = Trail Making Task B

TPN = Parenteral Nutrition

UCHL = University College Hospital London

VAN = Ventral Attention Network

VPT = Very Preterm Birth

WASI = Wechsler Abbreviated Scale of Intelligence

WISC-IV = Wechsler Scale Intelligence for Children

WPPSI-IV= Wechsler Preschool & Primary Scale of Intelligence

# CHAPTER 1 - Introduction

This chapter highlights the importance of the last trimester of gestation in terms of establishing typical developmental trajectories. It also describes the long-term neurodevelopmental consequences of preterm birth, which occurs during this critical developmental window, disrupting typical developmental processes and giving rise to an increased risk of perinatal brain injury, intricate structural and function brain alterations, developmental delays, and behavioural difficulties in childhood and later in life. The unfavourable outlooks of preterm birth are described to be particularly alarming as it is still not possible to predict which preterm born infants are likely to develop behavioural sequelae, making it challenging to devise preventative interventions or targeted treatment programmes for those at risk and probing imminent need to identify predictive markers. This chapter, therefore, discusses the benefits of using advanced data-driven and whole-brain Magnetic Resonance Imaging (MRI) analysis techniques to examine intricate structural and functional brain alterations as potential neurobiological markers of behavioural outcomes. Moreover, this chapter also touches on the fact that the 'preterm behavioural phenotype' is characterised by heterogeneous and complex behavioural profiles and neurodevelopmental trajectories, which make it challenging to characterise brain-behavioural relationships using traditional methodological frameworks. It then describes conceptual reformulations and innovative methodological approaches as appropriate alternative tools to stratify the heterogeneous and complex characteristics associated with behavioural outcomes prior to examining associations with neural alterations in preterm samples. Finally, this chapter is concluded by detailing the gaps in the literature in terms of characterising neurobiological markers of behavioural heterogeneity and listing the main objectives of this PhD thesis which aim to characterise nuanced brain-behavioural patterns following preterm birth.

## 1.1 The importance of the early environment for typical neurodevelopment

### 1.1.1 Brain development – from early gestation to adulthood

Brain development is a long-term and complex biological process, which starts early on during the first trimester of pregnancy, accelerating rapidly throughout the second and third trimesters and continuing to occur postnatally in childhood and adolescence, before steadily reaching maturation in early adulthood (Figure 1.1) (Thompson and Nelson, 2001; Nelson, Thomas and de Haan, 2006; Batalle, Edwards and O'Muircheartaigh, 2018). There are multiple actively occurring neurobiological processes during the gestational period, supporting rapid brain development by the full-term (FT) period, which typically occurs at 37-42 gestational weeks (Figure 1.1). Owing to these dynamic biological changes occurring in the intra-uterine environment, neonatal brain development exhibits considerable maturational growth by the time of FT birth, reaching ~25% the weight of a fully developed adult brain despite the neonate weighing only ~6% the body weight of an adult (Dekaban and Sadowsky, 1978; Grigorenko, 2017), highlighting the importance of the early developmental period.

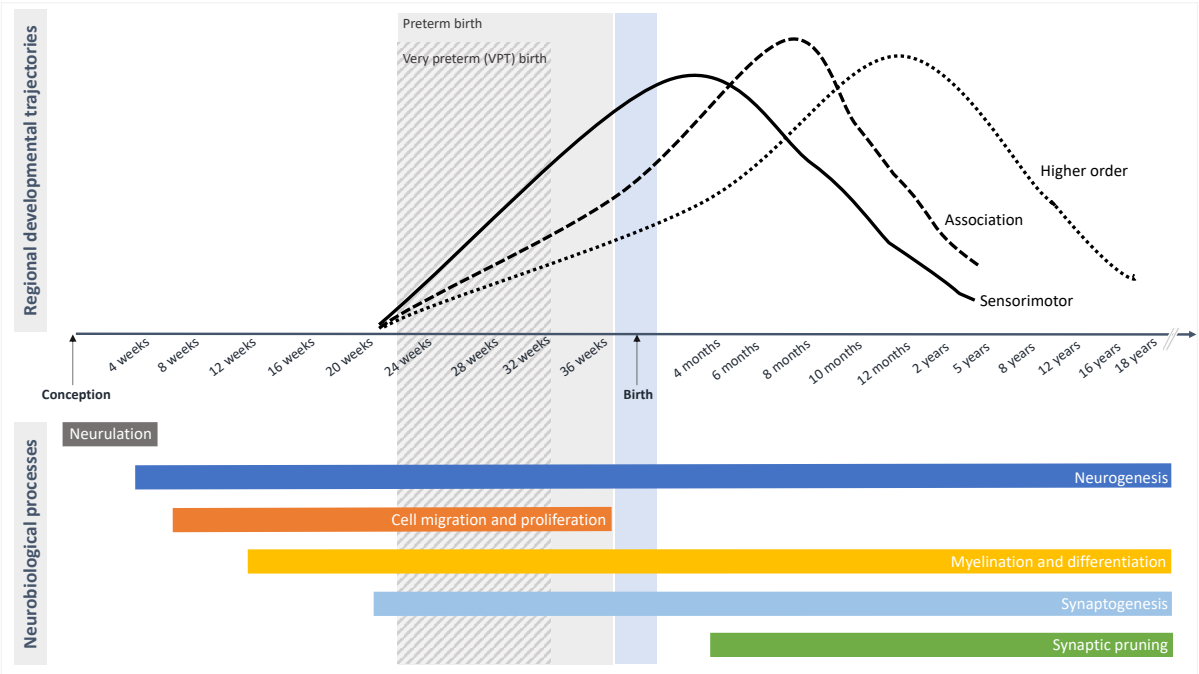



***Figure 1.1. Timeline of brain development from early gestation to adulthood.***

*The figure illustrates timelines of typical developmental trajectories and neurobiological processes and depicts the critical developmental window during which preterm birth occurs (at <37 weeks of gestation; shaded in grey) and disrupts these processes. Shading in striped grey denotes the period of very preterm (VPT) birth (at ≤32 weeks of gestation) which also includes extremely preterm birth (at ≤28 weeks of gestation). Figure adapted from (Thompson and Nelson, 2001; Uytun, 2018).*

The first neurobiological process, referred to as neurulation, kick-starts brain development by establishing neural tube hindbrain, midbrain, and forebrain formations (Purves *et al.*, 2001; Nikolopoulou *et al.*, 2017; Leibovitz, Lerman-Sagie and Haddad, 2022). These processes are subsequently followed by dramatic changes at the cellular level including the onset of neurogenesis and cell migration and proliferation in the first trimester (Tierney and Nelson, 2009; Leibovitz, Lerman-Sagie and Haddad, 2022). By the second trimester, myelination, differentiation, and synaptogenesis processes begin to occur, allowing for early cortical organisation and neural connectivity patterns to be established by the end of the second trimester, with myelination and synaptogenesis processes accelerating in the third trimester and only stabilising a couple years after birth (Figure 1.1) (Huttenlocher and Dabholkar, 1997; Tierney and Nelson, 2009; Leibovitz, Lerman-Sagie and Haddad, 2022). Postnatally, synaptogenesis and myelination neurobiological processes continue to drive steady increases in brain growth over the first few years of life before synaptic pruning processes start probing decreases in cortical volumes in middle childhood, whereby myelination and synaptic pruning processes persist long after the early and middle childhood periods (Figure 1.1) (Huttenlocher and Dabholkar, 1997; Ducharme *et al.*, 2016; Gennatas *et al.*, 2017).

Notably, however, these developmental trajectories do not occur in a uniform manner across all regions of the brain, and nor do they undergo linear developmental trajectories with age. Rather, distinct brain regions seem to undergo unique developmental trajectories. For instance, lower-order primary sensory and motor regions tend to be the first cortical areas to develop, followed by parietal and temporal association cortices, and higher-order cognitive regions such as the prefrontal cortex subsequently (Figure 1.1) (Thompson and Nelson, 2001; Gogtay *et al.*, 2004; Gilmore *et al.*, 2012; Grigorenko, 2017; Hodel, 2018; Fenchel *et al.*, 2020). The dynamic developmental trajectories followed by distinct regions are potentially guided by their functional properties, with studies showing evidence of age-dependent neural network maturation increasing in proportion with the development of cognitive and behavioural abilities in adolescence and early adulthood (Thompson and Nelson, 2001; Gogtay *et al.*, 2004; Nelson, Thomas and de Haan, 2006; Gu *et al.*, 2015; Grayson and Fair, 2017; Hodel, 2018; Baum *et al.*, 2020; Váša *et al.*, 2020; Akarca *et al.*, 2021; Tooley *et al.*, 2022).

### 1.1.2 The late gestational period – a critical developmental window

It is also crucial to note, however, that despite the observations discussed (in section 1.1.1) depicting brain development as a biologically complex process which does not reach maturation

until decades later in adulthood; it is in fact during the earliest developmental stages, even before birth, that the most fundamental structural and functional neurobiological foundations are established (Figure 1.1). Evidence from an increasing number of studies examining typical foetal intra-uterine brain development collectively emphasises the importance of the late gestational period in terms of supporting typical maturational trajectories (Clouchoux *et al.*, 2012; Wilson *et al.*, 2021, 2023; Moore *et al.*, 2023). Between 25 and 37 gestational weeks, major white matter tracts undergo transitional maturational changes (Wilson *et al.*, 2021, 2023), cortical and cerebellar volumes exhibit 2-4-fold increases (Clouchoux *et al.*, 2012), and functional connectivity resting state network organisations in higher-order and association cortices continue to mature (Karolis *et al.*, 2023; Moore *et al.*, 2023). Furthermore, these actively occurring early biological processes tend to have long-term influences on both brain and behavioural development (Davis *et al.*, 2011; Gale-Grant *et al.*, 2021). Whereby, even amongst infants born at FT (37-42 gestational weeks), those with longer gestation, seem to have greater neurodevelopmental advantages in terms of early (Gale-Grant *et al.*, 2021) and later (Davis *et al.*, 2011) brain development, as well as optimal behavioural linguistic, motor, and cognitive outcomes in toddlerhood (Espel *et al.*, 2014; Gale-Grant *et al.*, 2021).

### 1.1.3 Altered neurodevelopment – consequences of preterm birth

#### *1.1.3.1 Perinatal brain injury*

Preterm birth, which occurs during critical stages of brain development (Figure 1.1), results in severe and long-lasting neurodevelopmental consequences. It has also been associated with major perinatal brain injury, such as periventricular leukomalacia, focal necrotic cortical, subcortical, and white matter lesions, accompanying intraventricular-periventricular germinal matrix haemorrhage which ranges from minor to major, depending on the presence and extent of ventricular dilation and intraparenchymal haemorrhage (Volpe, 2009a, 2009b). More so, the risk of these lesions arising is even further heightened when birth occurs before 33 weeks of gestation, which is a period of preterm birth which is clinically defined as very preterm (VPT) birth and encompasses those born extremely preterm (EPT) at less than 28 weeks of gestation. The increased susceptibility to perinatal brain injury amongst VPT born infants is likely linked to the fact that the germinal matrix, a metabolically active embryonic region responsible for active neuronal proliferation, remains highly vascularized until around 32-33 weeks' gestation (Starr *et al.*, 2023). Disruptions to this highly vulnerable vascular system at this time, therefore, heighten vulnerability to brain injury in areas within the proximity of the germinal matrix, such as the caudate or thalamostriatal regions

(Volpe, 2009a, 2009b). Historically, perinatal brain injury was first characterised using cranial neuroimaging techniques such as ultrasonographic imaging (ultrasound), which is a suitable and relatively inexpensive tool capable of detecting perinatal brain injury relative to other approaches such as Magnetic Resonance Imaging (MRI) (Plaisier *et al.*, 2015; Burkitt *et al.*, 2019; Skiöld *et al.*, 2019). On the other hand, however, studies in preterm samples found ultrasound was less sensitive to subtle cerebral abnormalities which were otherwise detected using MRI (Plaisier *et al.*, 2015; Burkitt *et al.*, 2019).

MRI is a sophisticated neuroimaging tool capable of generating images with high resolution, superior soft tissue contrast, and three-dimensional visualisations of the brain. It is due to the vast advancements in magnetic resonance physics and the evolving understanding of neurobiology, that MRI has been able to generate images of the brain differentiating between distinct tissue types at an unprecedented level of detail (Berger, 2002). That is because MRI fundamentally captures the magnetic resonance of positively charged hydrogen atom protons, which are in fact highly prevalent in the body due to their presence in fat and water molecules (Figure 1.2). Moreover, as distinct brain tissue types (e.g., cerebrospinal fluid, cortical grey matter, subcortical deep grey matter, or white matter) contain different densities and concentrations of water and fat molecules, differing radio wave signals emitted from the distinct tissue type $T_1$ and $T_2$ relaxation processes are detected by the MRI receiver coils and used to construct $T_1$-weighted and $T_2$-weighted structural MRI (s-MRI) images, which can be used for a variety of purposes (Berger, 2002).

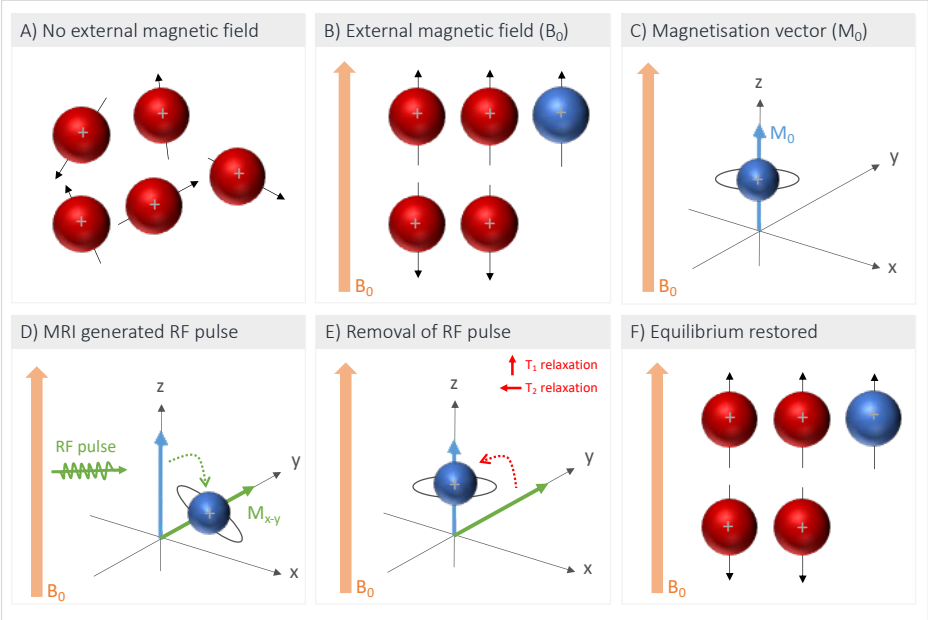



***Figure 1.2. Schematic illustrating basic principles of MRI.***

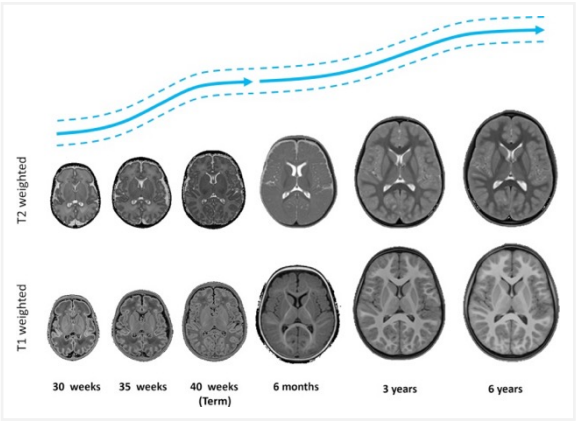
*A) Hydrogen atoms contain positively charged protons (red spheres) which spin freely in random directions in the human body. B) MRI introduces an external magnetic field $B_0$ (orange arrow), which rotates protons spinning along their axis to align with $B_0$ in one of two possible orientations: spin-up*

*(parallel to $B_0$) or spin-down (anti-parallel to $B_0$). C) The difference between these two alignment orientations (blue sphere) results in a small net magnetisation ($M_0$) (blue arrow). D) MRI introduces a second magnetic field orthogonal to $M_0$ resulting in a radio frequency (RF) pulse which induces a transfer of energy, rotating $M_0$ protons out of alignment, at an angle depending on the amplitude and duration of RF. E) After the removal of the RF pulse, two processes: longitudinal spin-lattice $T_1$ relaxation (i.e., time taken for magnetisation to reach 63% of the original $M_0$ magnetisation along the longitudinal plan) and transverse spin-spin $T_2$ relaxation (i.e., time taken magnetisation to decay to 37% its original value) allow realigned protons to return to F) equilibrium state. Figure adapted from (Mastrogiacomo* et al.*, 2019).*

With T1-weighted images having a higher sensitivity to fat and T2-weighted images being water-sensitive, the two contrasts can be used to capture distinct types of structural information. Fat-sensitive T1-weighted imaging emits stronger signal intensities for grey and white matter tissue types, making it an adequate sequence for capturing anatomical tissue type segmentations (Avants *et al.*, 2008a). On the other hand, water-sensitive T2-weighted imaging can visualise fluids and high-water content tissues with brighter signal intensity, making this contrast better suited for the identification of ischemic injury, lesions, and haemorrhages (Rutherford *et al.*, 2006; Batalle, Edwards and O'Muircheartaigh, 2018). It is also used in perinatal imaging for adequate characterisation of grey and white matter tissue contrasts. Unlike more mature child and adult brains, the young neonatal and infant brains are rich in water content at this early developmental window, making T1-weighted images not suitable for these purposes at early ages (Figure 1.3) (Rutherford *et al.*, 2006; Batalle, Edwards and O'Muircheartaigh, 2018).

***Figure 1.3. Developmental changes in T1- and T2-weighted image contrasts.***

*Axial MRI brain images demonstrating dynamic changes in T1-weighted and T2-weighted contrasts over the late gestational period and early childhood. Figure reproduced from (Batalle, Edwards and O'Muircheartaigh, 2018) – permitted for distribution under CC BY 4.0 license.*




### *1.1.3.2 Intricate structural and functional brain alterations*

Initially, the predominant use of MRI in preterm research was to generate qualitative ratings of perinatal brain injury, based on visual inspections of structural MRI (s-MRI) images (Inder *et al.*, 2003). However, with the increased adoption of advanced and quantitative neuroimaging analysis techniques (Fischl and Dale, 2000; Calhoun *et al.*, 2001; Avants and Gee, 2004; Smith *et al.*, 2006; Ashburner and Friston, 2009; Friston, 2011; Perrot, Rivière and Mangin, 2011) and the introduction of additional MRI modalities such as diffusion-weighted (d-MRI) and functional (f-

MRI) MRI, studies were able to characterise intricate structural and functional brain alterations associated with preterm birth, which appear to arise from the earliest stages of life and last into childhood and adulthood.

Studies in preterm born neonates show evidence of alterations to cortical folding (Lefèvre *et al.*, 2016), brain growth (Bouyssi-Kobar *et al.*, 2016; Lefèvre *et al.*, 2016), and white matter tract maturation (Wilson *et al.*, 2021) when compared to typically developing foetuses in the intra-uterine environment during the 2nd and 3rd trimesters. Moreover, altered intra-uterine functional connectivity has even been reported in preterm foetuses before preterm birth (Thomason *et al.*, 2017), highlighting an *in-utero* onset of the neurodevelopmental alterations associated with preterm birth. Structural and functional brain alterations seem to be persisting even beyond the early developmental periods, with studies in neonates at term-equivalent age (Doria *et al.*, 2010; Shimony *et al.*, 2016; Smyser *et al.*, 2016; Thompson *et al.*, 2018; Dubois *et al.*, 2019; Dimitrova, Arulkumaran, *et al.*, 2021; Dimitrova, Pietsch, *et al.*, 2021; Fenn-Moltu *et al.*, 2022), childhood (Nagy *et al.*, 2003; Kesler *et al.*, 2006; Lax *et al.*, 2013; Degnan *et al.*, 2015b; Thompson *et al.*, 2016, 2020; Lean *et al.*, 2017; Lemola *et al.*, 2017; Wehrle *et al.*, 2018; Cho *et al.*, 2022; Mossad *et al.*, 2022; Kvanta *et al.*, 2023), adolescence (Giménez *et al.*, 2006; Mullen *et al.*, 2011; Nosarti *et al.*, 2014; Wilke *et al.*, 2014; Rowlands *et al.*, 2016; Wehrle *et al.*, 2018; Zhou *et al.*, 2018; Thompson *et al.*, 2020), and adulthood (Eikenes *et al.*, 2011; White *et al.*, 2014; Jurcoane *et al.*, 2016; Papini *et al.*, 2016; Shang *et al.*, 2018) reporting similar structural and functional neurodevelopmental alterations in preterm born individuals relative to FT controls.

In terms of brain structure, advances in quantitative analytic approaches have allowed for measures of regional structural volumes, cortical folding, and cortical thickness to be calculated from s-MRI images (Fischl and Dale, 2000; Avants and Gee, 2004; Perrot, Rivière and Mangin, 2011). d-MRI is another powerful neuroimaging modality which captures micro-structural tissue properties by contrasting isotropic and anisotropic movement of water molecules. Isotropic diffusion occurs in regions where water molecules are free to move in random directions, such as cerebrospinal fluid, while the anisotropic movement of water occurs in structural properties such as white matter tracts where water molecules are restricted to certain directions along the tracts, making d-MRI an adequate tool for detecting perinatal hypoxic-ischemic injury or white matter tract microstructural integrity and structural connectivity (Forbes, Pipe and Bird, 2000; Wolf *et al.*, 2001; Rutherford *et al.*, 2006; Huisman, 2010). d-MRI can quantity white matter tissue microstructure by measuring isotropic diffusion using Mean Diffusivity (MD) and anisotropic movement along white matter fibres using Fractional Anisotropy (FA) properties, whereby smaller

MD and larger FA values would reflect more optimal myelination and microstructure integrity. f-MRI, on the other hand, uses Blood Oxygen Level Dependent (BOLD) contrasts to measure neural activity in the brain. It essentially measures changes in the magnetic field caused by red cell haemoglobin oxygenation and deoxygenation fluctuations, which are believed to be occurring in response to the increased regional metabolic changes in functionally activated cortical areas (Ogawa *et al.*, 1990; Glover, 2011). Since distinct cortical regions exhibiting simultaneous BOLD signal activation are considered to be functionally connected, f-MRI acquired in the absence of specific stimuli or tasks has been used to investigate the intrinsic functional architecture of the brain (Friston, 2011). This methodological approach is commonly referred to as 'resting state f-MRI' (rs-fMRI) or task-free f-MRI.

#### *1.1.3.3 Behavioural sequelae of preterm birth*

Preterm birth has also been associated with high rates of disabilities including cerebral palsy, blindness, deafness, and severe learning difficulties (Blencowe *et al.*, 2013; Pierrat *et al.*, 2021) as well as long-lasting behavioural sequelae associated with deficits in socio-emotional and cognitive processing which persist throughout the preterm lifespan (Johnson and Marlow, 2011; Nosarti *et al.*, 2012; Wolke, 2019; P. J. Anderson *et al.*, 2021). In addition, preterm samples tend to exhibit a 2-4-fold increased prevalence of neurodevelopmental and psychiatric conditions such as Autism Spectrum Conditions (ASC), Attention Deficit/Hyperactive Disorder (ADHD), and internalising disorders such as anxiety and depression, relative to the general population (Johnson *et al.*, 2010b; Nosarti *et al.*, 2012; P. J. Anderson *et al.*, 2021). Whilst there has been an observed decline in rates of major brain injury and severe impairments and disabilities over the past few decades, which are likely owing to the vast advancements in modern neonatal medicine (Platt *et al.*, 2007; Moore, Hennessy, *et al.*, 2012; Blencowe *et al.*, 2013), rates of developmental difficulties in cognitive, socio-emotional, sensory, motor, linguistic, and intelligence domains remain to be increasingly high in this clinical population. Preterm birth has also been found to negatively impact quality of life in multiple domains in adulthood: physical, cognitive, and mental health, romantic relationships, occupational performance, academic achievements, and social functioning (Cooke, 2004; Wolke, Chernova, *et al.*, 2013; Winstanley *et al.*, 2015; Kroll *et al.*, 2017; Mendonça, Bilgin and Wolke, 2019; Bolbocean *et al.*, 2023), which are potentially related to the quality of physical, cognitive, and mental health in adolescence or adulthood and clinical risk perinatally (Wolke, Chernova, *et al.*, 2013; Winstanley *et al.*, 2015; Kroll *et al.*, 2017).

In light of these observations, the high prevalence rates of preterm birth (~5-20% of births, depending on geographical region) draw serious medical and sociological concerns for healthcare professionals, policymakers, and preterm born individuals and their families (Chawanpaiboon *et al.*, 2019; Walani, 2020). This is especially alarming as it is still not possible to predict which preterm born infants are likely to develop behavioural difficulties, making it challenging to devise preventative interventions or targeted treatment programmes for those at risk. Therefore, highlighting the need to determine specific markers predictive of neurodevelopmental trajectories in preterm samples.

## 1.2 The complexity of brain-behavioural relationships in preterm samples

### 1.2.1 Intricate brain alterations underlie behavioural outcomes – the need for advanced neuroimaging tools

As the brain plays an integral role in maintaining and regulating behavioural, sensory, motor, socio-emotional, and cognitive processes, it is therefore plausible that the pervasive brain alterations occurring in preterm samples may explain the link between preterm birth and behavioural sequelae. In fact, perinatal brain injury and brain alterations in preterm samples are often localised to brain regions known to be involved in sensory, motor, cognitive, and socio-emotional processing, such as the caudate, cerebellum, brainstem, thalamus, hippocampus, and temporal, sensorimotor, and parieto-occipital cortices (Isaacs *et al.*, 2004; Kesler *et al.*, 2006; Volpe, 2009a, 2009b; Lax *et al.*, 2013; Ball *et al.*, 2016; Lean *et al.*, 2017). In turn, this suggests that these neural alterations may potentially act as useful markers predictive of specific behavioural outcomes. Supporting these claims, on one hand, studies have shown that perinatal brain injury in preterm neonates was able to predict severe cognitive deficits and cerebral palsy in toddlerhood (Burkitt *et al.*, 2019; Skiöld *et al.*, 2019). On the other hand, however, it is worth mentioning that it is still not possible to reliably predict the developmental, socio-emotional, or cognitive behavioural sequelae of preterm birth based on perinatal brain injury classifications (Isaacs *et al.*, 2004; Burkitt *et al.*, 2019).

There has been a conceptual shift in the field towards exploring whether the intricate structural and functional brain alterations, which may occur in the presence or even the absence of perinatal brain injury (Isaacs *et al.*, 2004; Bouyssi-Kobar *et al.*, 2016; Shimony *et al.*, 2016), can explain the onset of behavioural sequelae. It was only possible to formally investigate these notions relatively recently, following the development of advanced MRI tools capable of measuring and quantifying these intricate neural alterations. The earliest studies in preterm samples adopting these advanced

neuroimaging analysis approaches managed to identify brain-behavioural associations, which were not previously characterised. Namely, they identified structural alterations in specific pre-selected brain regions of interest (ROIs), such as the hippocampus, caudate, and temporal cortex, and their association with behavioural difficulties relating to intelligence and ADHD (Peterson *et al.*, 2000; Abernethy, Palaniappan and Cooke, 2002; Abernethy, Cooke and Foulder-Hughes, 2004; Nosarti *et al.*, 2005). These findings highlight the power of using MRI to detect specific brain changes likely to predict later outcomes. Arguably, however, a limitation of these traditional methodological designs lies in their approach of investigating relationships between the brain and behaviour outcomes in a "one-to-one" manner, often restricting brain analyses to isolated ROIs. This limitation may hinder the ability to fully capture the nuanced and intricate brain mechanisms that underlie the emergence of behavioural outcomes (Westlin *et al.*, 2023). That is because, whilst particular brain regions can be considered key players mediating certain functions, it is important to note that they do not work in isolation and are in fact embedded within a functionally and structurally connected network of multiple brain regions working in tandem to mediate behavioural outcomes (Friston, 2011). Accordingly, a single brain region may also be involved in multiple distinct neural networks supporting different behavioural processes (Westlin *et al.*, 2023). Therefore, relying on ROI analyses may be considered restrictive.

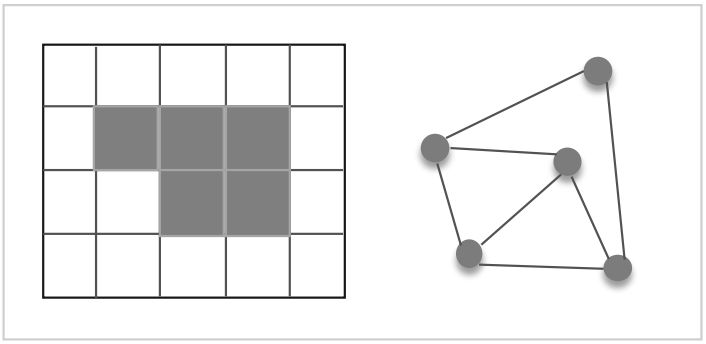
While ROI analyses may be advantageous in cases where research questions are more narrowly focused or when sufficient prior knowledge can inform ROI selection, the use of alternative and more advanced whole-brain and data-driven neuroimaging analysis approaches offer an enhanced ability to explore nuanced relationships without the confinements of specific a priori assumptions. They also reduce chances of confirmation bias and increase the possibility of identifying novel findings. As brain alterations in preterm samples and their relationships with behavioural sequalae are not yet fully established, departing from traditional ROI analyses may provide nuanced insights into brain-behavioural patterns in this population. In keeping with this argument, a study implementing whole-brain and voxel-wise exploratory analyses in a sample of preterm children characterised structural changes in the hippocampus and temporal cortex to be associated with intelligence scores, confirming expectations based on findings from previous ROI studies (Peterson *et al.*, 2000; Abernethy, Palaniappan and Cooke, 2002), whilst also identifying highly significant yet unanticipated correlations between intelligence and structural changes in the occipital lobe (Isaacs *et al.*, 2004).

The ability to explore brain-behavioural associations across the whole-brain, without compromising statistical power and integrity, has been made possible through the evolution of

computational advancements. Examples of these computationally demanding and advanced neuroimaging methods include the use of a data-driven Independent Component Analysis approach which detects distinct functionally connected neural networks by characterising spatially independent patterns of brain activity from f-MRI data (Calhoun *et al.*, 2001); mathematical graph-theory representations of f-MRI or d-MRI to study the brain as a neural network of functionally and structurally inter-connected regions (Zalesky, Fornito and Bullmore, 2010; Friston, 2011; Holiga *et al.*, 2019); Voxel Based Morphometry to quantify regional grey matter tissue density from s-MRI data (Ashburner and Friston, 2009); Tensor Based Morphometry (TBM) to measure relative deformations in structure size or shape from s-MRI or d-MRI data (Avants and Gee, 2004); and Tract Based Spatial Statistics (TBSS) to extract a white matter skeleton and quantify micro-structural properties along tracts (Smith *et al.*, 2006). These methods then apply mass-univariate statistical testing at the whole-brain level, where statistical modelling is performed for each brain region or voxel and permutation testing and appropriate Family Wise Error Rate (FWER) corrections are applied to reduce the incidence of Type 1 errors (Zalesky, Fornito and Bullmore, 2010; Jenkinson *et al.*, 2012). Cluster-based statistical methods, aiming to reduce Type 2 error rates, such as Threshold-Free Cluster Enhancement (TFCE) and Network Based Statistic (NBS), are also incorporated within the described statistical inference protocols (Smith and Nichols, 2009; Zalesky, Fornito and Bullmore, 2010). TFCE and NBS methods, respectively, allow for spatially or topologically related brain regions showing associations with behavioural outcomes to be collectively characterised as clusters or components without being increasingly penalised by the FWER corrections (Smith and Nichols, 2009; Zalesky, Fornito and Bullmore, 2010) (Figure 1.4).

***Figure 1.4. Spatial and topological clusters.***

*Cluster of 5 voxels within neighbouring spatial proximity is seen on the left and a component of 5 connected nodes within topological space is demonstrated on the right. Figure adapted from (Zalesky, 2012).*



Over the past decade, there has been an exponential increase in the number of preterm studies applying whole-brain and data-driven neuroimaging analysis approaches to characterise brain-behavioural relationships in preterm individuals, providing the field with novel insights (Constable *et al.*, 2013; Scheinost *et al.*, 2015; Rowlands *et al.*, 2016; Lean *et al.*, 2017; Choi *et al.*, 2018; Wheelock *et al.*, 2018, 2021; Vanes *et al.*, 2021, 2023). Studies identified neural alterations which were associated with behavioural outcomes in preterm samples, and found no evidence of these

relationships in FT samples (Constable *et al.*, 2013; Scheinost *et al.*, 2015; Rowlands *et al.*, 2016; Lean *et al.*, 2017; Wheelock *et al.*, 2018, 2021). These findings shed light on potential preterm-specific brain-behavioural associations and confirm the advantage of using such neuroimaging techniques to investigate brain-behavioural relationships.

Nevertheless, these studies also revealed some challenges associated with the characterisation of brain-behavioural relationships in preterm samples, which remain to be addressed. For instance, other findings indicated that some neural alterations in regions and networks involved in somatomotor, default mode network (DMN), or language processing (Rowlands *et al.*, 2016; Wheelock *et al.*, 2021), were associated with behavioural outcomes across both preterm and FT samples; demonstrating that some brain-behavioural patterns may be generalisable to both preterm and FT individuals, independently of clinical birth status. Moreover, additional findings made it apparent that brain alterations occurring in preterm samples may not necessarily only imply detrimental developmental consequences as has been intuitively expected and previously seen (Isaacs *et al.*, 2004; Nosarti *et al.*, 2005; Myers *et al.*, 2010; Rogers *et al.*, 2012, 2014; Olsen *et al.*, 2018; Wheelock *et al.*, 2018, 2021; Vanes *et al.*, 2021; Kanel *et al.*, 2022), but that these neural changes may also serve as neurobiological compensatory mechanisms which probe resilience against the onset of difficulties; whereby greater differences in neural patterns in preterm relative to FT individuals were associated with better preterm behavioural outcomes (Schafer *et al.*, 2009; Daamen *et al.*, 2014; Finke *et al.*, 2015; Scheinost *et al.*, 2015). Furthermore, recent studies also reported that preterm birth does not result in uniform brain changes across the preterm population. Rather, preterm born individuals seem to exhibit heterogeneous variations of neurodevelopmental alterations relative to typical developmental trajectories (Dimitrova, Arulkumaran, *et al.*, 2021; Dimitrova, Pietsch, *et al.*, 2021), with the magnitude of variability seemingly increasing by adulthood (Stoecklein *et al.*, 2020).

In essence, cumulative observations from these studies indicate that preterm birth results in heterogeneous neurodevelopmental trajectories which remain to be characterised. On one hand, the mechanistic intricacy and complexity of neural patterns in preterm samples are, at least partially, contributing to the difficulty of identifying intricate brain-behavioural relationships within this population. As discussed in this subsection, cutting-edge neuroimaging analysis techniques may help counteract some of these concerns. On the other hand, however, the ability to accurately ascertain brain-behavioural relationships is not only limited by the robustness of the tools used to capture brain changes but also by those used to measure behavioural outcomes.

Therefore, efforts to appropriately characterise behavioural outcomes in preterm populations are needed in studies aiming to investigate brain-behavioural associations.

### 1.2.2 Complex behavioural phenotypes – limitations of traditional frameworks

Similar to psychiatric populations, preterm born individuals exhibit complex phenotypic presentations of behavioural profiles, which may not be adequately characterised using traditional methodological approaches. This oversight in traditional methodological frameworks was first pinpointed by psychiatric researchers (Cuthbert and Insel, 2013; Cuthbert, 2014; Kotov *et al.*, 2017; Morris *et al.*, 2022), who have emphasised the need for reformulated conceptual frameworks and innovative methodological approaches that could be used to characterise complex behavioural phenotypes in psychiatric populations.

This section of the thesis describes phenotypic similarities between preterm and psychiatric populations in terms of their complex behavioural profiles (section 1.2.2). Subsequently, the following section (section 1.2.3) introduces conceptual reformulations and innovative methodological approaches which have been applied in psychiatric samples and have strengthened the ability to elucidate nuanced brain-behavioural associations. It is finally argued in section 1.2.4 that the use of those frameworks in preterm samples would help provide novel insights into the intricate brain-behavioural relationships exhibited by this population.

#### *1.2.2.1 Subthreshold and co-occurring behavioural difficulties*

An increasingly discussed criticism of current psychiatric research and clinical practices relates to the reliance on rigid taxonomies, defined by diagnostic manuals such as the Diagnostic and Statistical Manual (DSM) (American Psychiatric Association, 2013) or the International Classification of Diseases (ICD) (World Health Organization, 2019), to diagnose or describe psychiatric disorders. According to these guidelines, an individual must exceed a specific number of symptoms before receiving a psychiatric diagnosis. While the binary nature of these classifications helps clinicians identify those in need of treatment, it can, however, be problematic for individuals exhibiting subthreshold difficulties (i.e., presenting with a subset of symptoms of a specific condition which does not meet the diagnostic criteria). In this case, patients would not receive a diagnosis and therefore not receive treatment for their condition, which has in fact been associated with greater public health burdens and increased mortality rates (Horwath *et al.*, 1994; Cuijpers *et al.*, 2013; Biella *et al.*, 2019). Moreover, there are high prevalence rates of co-occurring psychiatric conditions which further suggests that current diagnostic classification approaches

inadequately capture phenotypic presentations (van Loo and Romeijn, 2015). For instance, mood disorders such as anxiety or depression tend to co-occur and seem to present with more severe and long-lasting behavioural difficulties when compared to either disorder alone (Beekman *et al.*, 2000; Schoevers *et al.*, 2005). Moreover, around 21% of people with ADHD tend to also have co-occurring ASC (Hollingdale *et al.*, 2020), and 45% of those with a psychiatric disorder meet criteria for another two or more co-occurring diagnoses within the span of a year (Kessler *et al.*, 2005). The presence of subthreshold and co-occurring behavioural difficulties, therefore, makes it challenging to identify which treatments patients are likely to be responsive to. Subthreshold and co-occurring behavioural difficulties are in fact also observed amongst preterm born individuals (Johnson and Marlow, 2011; Kim *et al.*, 2016; Johnson *et al.*, 2018), further emphasising the need to deviate away from clinical classifications of behavioural difficulties in this population.

#### *1.2.2.2 Transdiagnostic 'gate-way' mechanisms*

Pioneering psychiatric studies have also characterised specific 'transdiagnostic' behavioural mechanisms involving socio-emotional and cognitive processing, which appear to be acting as 'gate-way' mechanisms mediating the onset of distinct behavioural outcomes across multiple different neurodevelopmental and diagnostic groups (Fernandez, Jazaieri and Gross, 2016; Huang-Pollock *et al.*, 2017; Bathelt *et al.*, 2018; Astle *et al.*, 2019; Barkus and Badcock, 2019; Kushki *et al.*, 2019; Mareva, CALM team and Holmes, 2019; Dalgleish *et al.*, 2020; Santens *et al.*, 2020; Vaidya *et al.*, 2020; Wade *et al.*, 2020; Mareva *et al.*, 2023). However, these mechanisms and their integrated interplay are overlooked by traditional psychiatric frameworks, indicating another limitation with the use of traditional diagnostic screening tools to capture behavioural outcomes.

Socio-emotional processing is an umbrella term describing the multiple behavioural mechanisms required to appropriately process and respond to social and emotional stimuli (Ochsner, 2008), such as the ability to effectively modulate and control emotional reactivity (i.e., emotion regulation) (Thompson, Meyer and Jochem, 2008), accurately recognise emotional expressions (i.e., emotion recognition) (Ferretti and Papaleo, 2019), and effectively and appropriately communicate and interact with others (i.e., social communication and interaction) (Constantino *et al.*, 2003). Deficits in socio-emotional processing domains are associated with the presence of internalising behavioural difficulties, such as excessive worrying, rumination, perceived rejection, loneliness, or social withdrawal, as well as impairments in social functioning such as difficulties in forming and maintaining healthy relationships (Kanai *et al.*, 2012; Knight *et al.*, 2019; Leathem *et al.*, 2021; Høegh *et al.*, 2022).

Executive functions, on the other hand, refer to high-level cognitive mechanisms important for orchestrating one's ability to manage thoughts, behaviours, tasks, and goals. Executive function abilities provide individuals with the necessary cognitive tools needed to shift between tasks or mental states (i.e., shifting), inhibit irrelevant distracting stimuli or inappropriate impulsive behaviours (i.e., inhibition), mentally retain information that is readily accessible (i.e., working memory), manage emotions effectively (i.e., emotion control), adapt to new changes or have new perspectives (i.e., cognitive flexibility), and execute behaviours supporting effective time management, organisation, and problem-solving skills (i.e., organisation/planning) (Diamond, 2013). Crucially, executive functions support cognitive, behavioural, and emotional regulation processes, forming core mechanisms facilitating multi-faceted avenues of human functioning such as academic performance, learning, decision-making, task execution, reward processing, socialising, self-concept, and emotion regulation (Bailey *et al.*, 2018; Gioia, Isquith and Roth, 2018; Cortés Pascual, Moyano Muñoz and Quílez Robres, 2019; Mareva, CALM team and Holmes, 2019; Wade *et al.*, 2020; Salehinejad *et al.*, 2021; Mareva *et al.*, 2023).

Dissimilar to goal-oriented executive functions, intelligence refers to the cognitive process which has been historically described as the innate ability to process, interpret, and appropriately respond to novel and complex visual, verbal, and abstract information as well as the ability to learn from experiences (Neisser *et al.*, 1996; Duggan and Garcia-Barrera, 2015). While the two cognitive constructs have well-defined distinctions between them and may not necessarily show collinearity in preterm samples or the general population, existing evidence also suggests that the two cognitive domains are not completely independent and in fact interact with one another to support appropriate cognitive and socio-emotional behavioural processing (Nosarti *et al.*, 2007; Solomon, Buaminger and Rogers, 2011; Duncan, 2013; Opitz *et al.*, 2014; Duggan and Garcia-Barrera, 2015; Kroll *et al.*, 2017; Imanipour *et al.*, 2021). Simultaneously, intact cognitive processing is also dependent on optimal behavioural processing (Donati, Meaburn and Dumontheil, 2021). The symbiotic relationship between the multiple distinct cognitive and behavioural constructs indicates that a complex interplay of cognitive and socio-emotional processes mediates behavioural outcomes. This emphasises the importance of integrating measures of both cognitive and socio-emotional processing mechanisms when attempting to characterise behavioural outcomes.

#### *1.2.2.3 Within-group heterogeneity*

It is becoming increasingly apparent that individuals belonging to the same diagnostic (e.g., ASC, ADHD, Obsessive Compulsive Disorder, anxiety, or depression) or non-diagnostic (i.e.,

typically developing) group present with varying cognitive or socio-emotional behavioural traits, despite sharing the same diagnostic labels (Bathelt *et al.*, 2018; Astle *et al.*, 2019; Barkus and Badcock, 2019; Kushki *et al.*, 2019; Vaidya *et al.*, 2020; Mareva *et al.*, 2023). This heterogeneity in behavioural outcomes, therefore, provides additional evidence suggesting a need for reformulated methodological frameworks to study behavioural outcomes beyond the constraints of rigid diagnostic categories. Similarly, phenotypic heterogeneity in behavioural profiles between individuals belonging to a single clinical group has also been identified within samples of preterm born children. Some studies found that distinct subgroups of preterm children may present with elevated cognitive, psychiatric, and socio-emotional behavioural difficulties which may or may not exceed clinical thresholds, while others may exhibit domain-specific behavioural difficulties or even present with no behavioural difficulties at all (Poehlmann *et al.*, 2015; Ross *et al.*, 2016; Johnson *et al.*, 2018; Burnett *et al.*, 2019; Lean *et al.*, 2020; van Houdt *et al.*, 2020; Bogičević *et al.*, 2021). The prominence of phenotypic variations across preterm born individuals indicates the presence of multifinality within this population, whereby individuals exposed to a similar exposure, (i.e., preterm birth) may develop distinct outcomes (Cicchetti and Rogosch, 1996) (Figure 1.5).

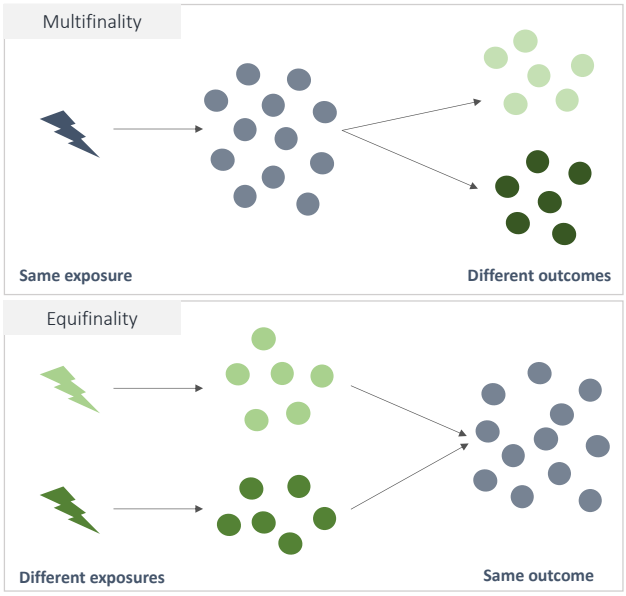



***Figure 1.5. Graphical illustrations of different types of behavioural heterogeneity exhibited by preterm populations.***

*Top panel describes heterogeneous outcomes across individuals, despite having a shared exposure (i.e., multifinality). Bottom panel describes exhibiting similar phenotypic outcomes despite following heterogeneous aetiological trajectories (i.e., equifinality).*

### *1.2.2.4 Aetiological heterogeneity*

On the other hand, there is an additional level of behavioural heterogeneity exhibited by these populations, which can be described by the concept of 'equifinality' in psychiatric research (Cicchetti and Rogosch, 1996). 'Equifinality' refers to the presence of similar phenotypic outcomes across individuals exposed to different risk factors (Figure 1.5), and it highlights the presence of various aetiological factors which may be contributing to the same behavioural difficulties. For

instance, as intact motor, auditory, and visual processing abilities are required for appropriate behavioural processing, it has been argued that sensorimotor or cognitive impairments in preterm samples and children with ASC may be contributing to the onset of observed socio-emotional and cognitive difficulties (Luyster *et al.*, 2011; Moore, Johnson, *et al.*, 2012; Blencowe *et al.*, 2013; Van Hus *et al.*, 2014; Rubenstein *et al.*, 2018; Vollmer and Stålnacke, 2019; Johansson *et al.*, 2023). In this regard, it is plausible that the presence of those developmental delays probes the onset of behavioural difficulties. On the other hand, however, studies have reported the presence of these deficits even in the absence of developmental delays (Rubenstein *et al.*, 2018; Vanes *et al.*, 2023), suggesting that the manifestation of common behavioural outcomes across distinct subgroups of preterm individuals may not be attributable to a single underlying mechanism.

Another example of equifinality of outcomes has been described by studies stratifying behavioural heterogeneity in samples of both preterm and FT children. Despite the two sets of children belonging to distinct clinical birth status groups (i.e., preterm vs FT birth), studies identified a range of distinct subgroups, exhibiting behavioural profiles with either generalised difficulties, domain-specific difficulties, or no difficulties, which were comprised of both preterm and FT children (Johnson *et al.*, 2018; Burnett *et al.*, 2019; Lean *et al.*, 2020). Although preterm children, relative to controls, were more likely to display suboptimal behavioural profiles, and the opposite ratio was true for profiles of no or limited behavioural difficulties; it is nonetheless apparent that preterm and FT born children may display the same behavioural profiles despite belonging to distinct clinical birth status groups (Johnson *et al.*, 2018; Burnett *et al.*, 2019; Lean *et al.*, 2020). However, the underlying mechanisms mediating the onset of distinct behavioural outcomes across these heterogeneous subgroups remain to be elucidated.

#### *1.2.2.5 Complex role of environmental and clinical factors*

The onset of these heterogeneous behavioural trajectories could potentially be explained by the presence of environmental or clinical risk factors. In both preterm samples and the general population, studies show that individuals exposed to adverse environments, such as early adversity, lower socio-economic status, or neighbourhood deprivation, are more likely to present with psychiatric disorders and behavioural difficulties (Khan *et al.*, 2019; Remes *et al.*, 2019; Lean *et al.*, 2020; Marsh, Dobson and Maddison, 2020), while those exposed to a socially and cognitively stimulating environment often present with better outcomes (Als *et al.*, 2004; Schoentgen, Gagliardi and Défontaines, 2020; Vanes *et al.*, 2021). Similarly, clinical risk factors, such as intra-uterine growth restriction, birth weight, gestational age, perinatal brain injury, or neonatal

physiological health, also impact neurodevelopmental trajectories negatively (Ment *et al.*, 2003; Levine *et al.*, 2015; Ross *et al.*, 2016; Brouwer *et al.*, 2017; Al-Haddad *et al.*, 2019; Stålnacke *et al.*, 2019; Durrant *et al.*, 2020; Bala *et al.*, 2023). Some clinical measures of post-natal growth were even found to be associated with certain behavioural difficulties related to cognitive and motor abilities, but not others (e.g. internalising and externalising behaviours), suggesting that clinical factors may impact developmental trajectories in a domain-specific manner (Dotinga *et al.*, 2016, 2019).

Moreover, current observations are making it increasingly apparent that individuals exposed to the same environmental or clinical factors may still develop dissimilar behavioural outcomes. Such that, it seems that these factors do not only influence behavioural trajectories independently, but they tend to moderate behavioural trajectories by interacting with one other (Kim-Cohen *et al.*, 2006; Luu *et al.*, 2009, 2011; Als *et al.*, 2012; Gunnar *et al.*, 2012; Wickremasinghe *et al.*, 2012; Benavente-Fernández *et al.*, 2019; Lovato *et al.*, 2022; Sun *et al.*, 2023). Whereby, the presence of an enriching environment may protect against the onset of psychiatric and cognitive difficulties in predisposed individuals with an elevated clinical risk, while adverse environments may further exacerbate behavioural difficulties in the presence of clinical risk (Figure 1.6). It is, therefore, imperative to consider the collective roles of environmental and clinical factors when aiming to characterise neurodevelopmental trajectories in such populations.

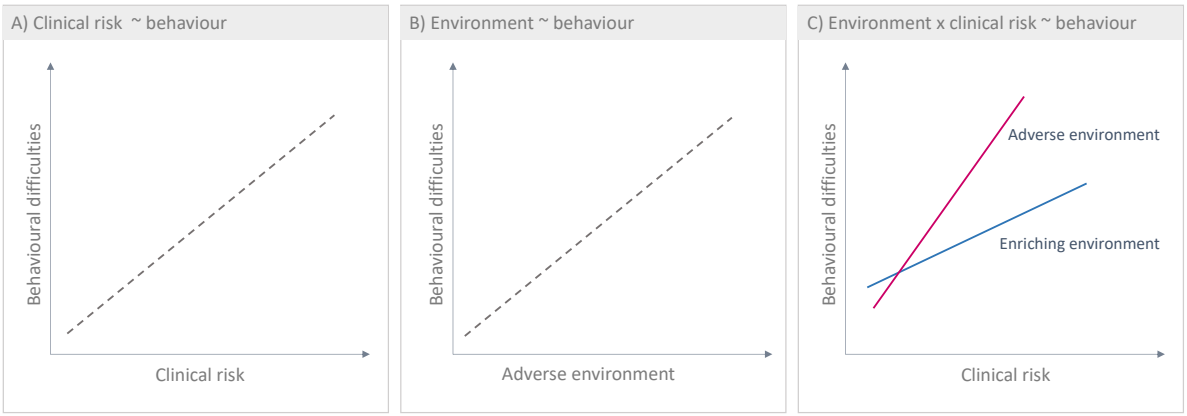



***Figure 1.6. Risk factor associations with behavioural outcomes.***

*The effect of A) clinical risk, B) environmental adversity, C) and the interaction between clinical risk and environmental adversity on behavioural difficulties.*

### 1.2.3 Methodological solutions – conceptual reformulations and innovative approaches

The complex and heterogeneous behavioural trajectories exhibited by psychiatric, preterm, and FT populations, underscore an important limitation associated with the use of traditional methodological designs, such as case-control or cohort study comparisons. That is because

grouping individuals solely based on clinical labels (e.g., either psychiatric diagnostic label or clinical birth status), erroneously assumes homogeneity within a group which in fact demonstrates intricate complexities and variations. Therefore, the use of those methods would in turn hinder the ability to capture mechanisms or factors predictive of distinct behavioural trajectories. Instead, conceptual reformulations and innovative methodological approaches are needed to adequately characterise these complex behavioural trajectories (Cuthbert and Insel, 2013; Cuthbert, 2014; Kotov *et al.*, 2017; Morris *et al.*, 2022).

#### *1.2.3.1 Measuring behavioural traits along a continuum*

In order to tackle concerns associated with the prevalence of subthreshold and co-occurring behavioural difficulties and transdiagnostic behavioural mechanisms which are not captured by psychiatric diagnostic classifications, behavioural traits could instead be understood along a continuum – ranging from better to worse outcomes. This could be achieved by using validated psychometric scales which measure behavioural constructs of psychopathology or cognition using dimensional summary scores rather than binary or categorical diagnostic labels. For instance, popular psychometric tools such as the Strengths and Difficulties Questionnaire (SDQ) (Goodman, 2001), Behaviour Rating Inventory of Executive Function (BRIEF) (Gioia, Isquith and Roth, 2018), Social Responsiveness Scale (SRS-2) (Constantino and Gruber, 2012), or Autism Quotient (AQ-10) (Booth *et al.*, 2013), traditionally used in clinics to indicate supra-threshold disorder presence, can also be used to capture severity of behavioural difficulties associated with general psychopathology, executive dysfunction, and social functioning along a continuum. Furthermore, subscale scores from each of those measures can also be extracted to measure behavioural subdomains such as internalising or externalising psychopathology using the SDQ (Goodman, 2001), measures of specific executive functions or emotional, cognitive, or behavioural regulation indices from the BRIEF (Gioia, Isquith and Roth, 2018), and social communication or repetitive behavioural difficulties from the SRS-2 (Constantino and Gruber, 2012). The use of other questionnaires such as the Child Behavioural Questionnaire (CBQ) (Rothbart *et al.*, 2001), Emotion Regulation Checklist (ERC) (Shields and Cicchetti, 1997), or Empathy Questionnaire (EmQue) (Rieffe, Ketelaar and Wiefferink, 2010), can also be used to characterise continuous measures of additional transdiagnostic behavioural constructs within the sphere of socio-emotional processing.

#### *1.2.3.2 Stratifying behavioural heterogeneity using reformulated conceptual psychometric criteria*

Another central concern that needs to be addressed relates to the prominent heterogeneity in behavioural profiles, aetiologies, and trajectories. Previous studies in preterm samples have succeeded in delineating behavioural heterogeneity using distinct psychometric criteria, whereby toddlers were stratified into three subgroups according to distinct scores on the Modified Checklist for Autism in Toddlers (M-CHAT): *critical positive*, *non-critical positive*, and *negative M-CHAT* subgroups (Luyster *et al.*, 2011; Moore, Johnson, *et al.*, 2012). In contrast to findings in the general population, which view both types of positive M-CHAT scoring criteria (i.e., *critical positive* and *non-critical positive*) to be indicative of an elevated likelihood of developing autism (Robins *et al.*, 2001), reports in preterm samples found one of those subgroups, the critical positive subgroup, to be exhibiting greater developmental impairments relative to the other non-critical positive subgroup (Luyster *et al.*, 2011; Moore, Johnson, *et al.*, 2012). These findings highlight the potential benefits of using the different M-CHAT scoring criteria approaches to identify subgroups of preterm individuals following distinct aetiological trajectories.

#### *1.2.3.3 Stratifying behavioural heterogeneity using data-driven methods and multi-dimensional input features*

Using data-driven stratification methods, such as latent class analyses or clustering techniques, to characterise behavioural heterogeneity, over the use of pre-defined classification approaches, such as the M-CHAT psychometric criteria, include added benefits of identifying novel and nuanced patterns of underlying heterogeneity which are based on multi-dimensional input variables. As cognitive, social, or emotional behavioural difficulties rarely occur in isolation and in fact mediate behavioural outcomes via a complex interplay betweenthem (Fernandez, Jazaieri and Gross, 2016; Huang-Pollock *et al.*, 2017; Bathelt *et al.*, 2018; Astle *et al.*, 2019; Barkus and Badcock, 2019; Kushki *et al.*, 2019; Mareva, CALM team and Holmes, 2019; Dalgleish *et al.*, 2020; Vaidya *et al.*, 2020; Wade *et al.*, 2020; Mareva *et al.*, 2023), considering the collective roles of multiple behavioural processes could provide additional insights when examining behavioural profiles.

Some dimensionality reduction techniques, such as factor analyses, have previously been used to integrate multiple behavioural measures of psychopathology, resulting in a unidimensional summary score measuring general psychopathology (referred to as the "p-factor"), which was regarded as a better predictor of disorder persistence and severity and had stronger correlations with etiological factors than individual internalising, externalising, or thought disorder construct factors (Lahey *et al.*, 2012; Caspi *et al.*, 2014; Caspi and Moffitt, 2018). However, studies combining multi-dimensional measures of psychopathology, socio-emotional processing, as well as cognitive

functions were able to characterise behavioural outcomes and behavioural heterogeneity with improved predictive validity (Bathelt *et al.*, 2018; Kushki *et al.*, 2019; Santamaría-García *et al.*, 2020; Vaidya *et al.*, 2020), further confirming the benefits of integrating multi-dimensional behavioural measures and highlighting the notion that cognitive or socio-emotional mechanisms are indeed not mediating the onset of behavioural difficulties independently of one another.

In order to incorporate information from multiple behavioural inputs and utilise it to stratify individuals into distinct data-driven subgroup categories, unsupervised machine learning classification approaches, such as clustering techniques, can be used. The most typically used clustering algorithm is K-means clustering, which defines cluster membership based on proximity to the cluster's centroid and usually searches for round-shaped clusters of similar size (MacQueen, 1967) (Figure 1.7). It randomly places initial cluster centroids and runs multiple iterations before the most optimal positioning of centroids is identified and used to classify membership, making this method sensitive to the randomly assigned initial placement of centroids (MacQueen, 1967). However, the development of relatively more recent clustering techniques, such as spectral clustering, has provided multi-fold benefits against the use of classical clustering algorithms such as K-means (Ng, Jordan and Weiss, 2001; von Luxburg, 2007; Ratle *et al.*, 2008; Jain, 2010; Rodriguez *et al.*, 2019). For instance, spectral clustering does not rely on centroid localisation, therefore, reducing biases introduced by initial centroid placement. Instead, it combines spectral and graph theoretical approaches to characterise cluster membership based on measures of nonlinear similarity (Ng, Jordan and Weiss, 2001). It calculates Eigenvectors of the input data Laplacian distance matrix, providing a lower-dimensionality graphic representation of data embedded based on weighted measures of similarity and, in turn, allowing for an enhanced ability to capture complex data structures without imposing pre-fixed assumptions about cluster shapes or sizes (Ng, Jordan and Weiss, 2001; von Luxburg, 2007; Ratle *et al.*, 2008; Jain, 2010; Rodriguez *et al.*, 2019) (Figure 1.7) (Figure 1.8).

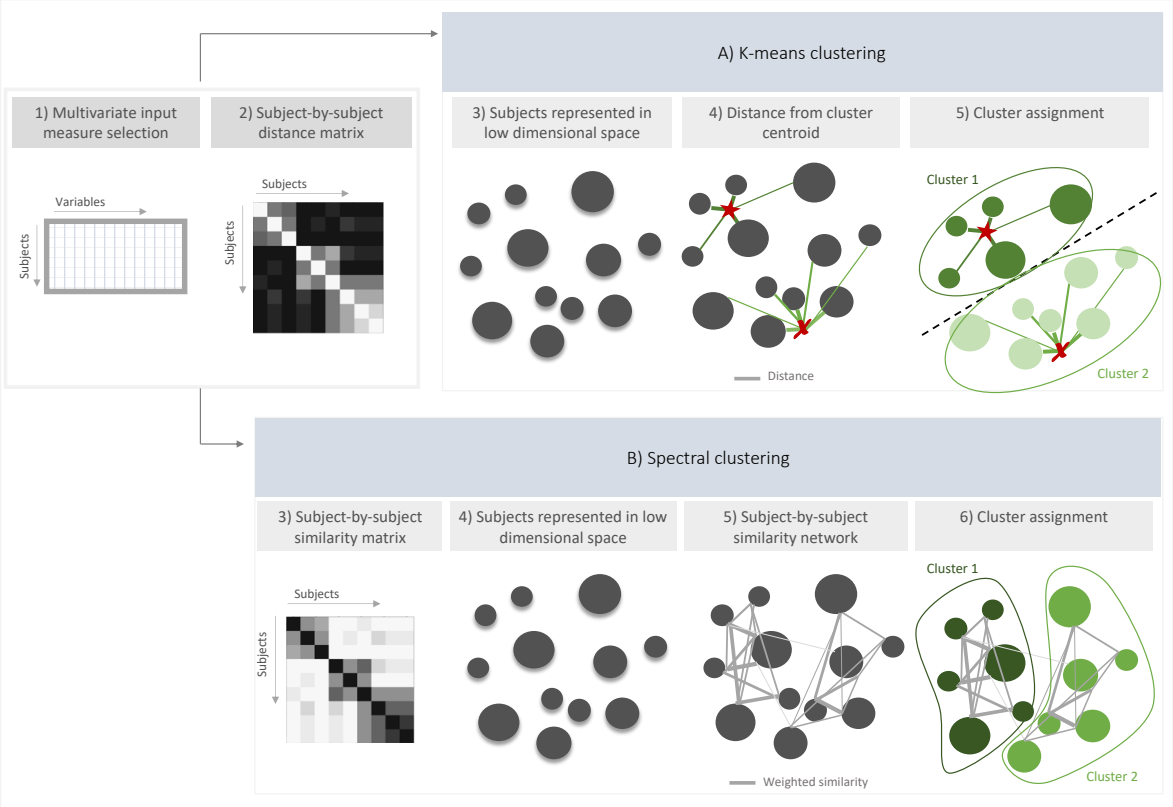



***Figure 1.7. Steps followed by K-means clustering and spectral clustering pipelines to stratify underlying heterogeneity based on multi-dimensional (multivariate) input data.***

*Both methods follow the same initial pre-processing steps (1 and 2), such that they both calculate a distance matrix based on multivariate input measures. Subsequently, A) K-means clustering algorithms partition the sample into clusters based on measures of subject proximity to the cluster centroid (i.e., distance), while B) spectral clustering, on the other hand, incorporates additional spectral and graph theoretical approaches, whereby nonlinear similarity between subjects is calculated and used to construct a weighted similarity network before delineating subjects into clusters based on measures of connectedness.*

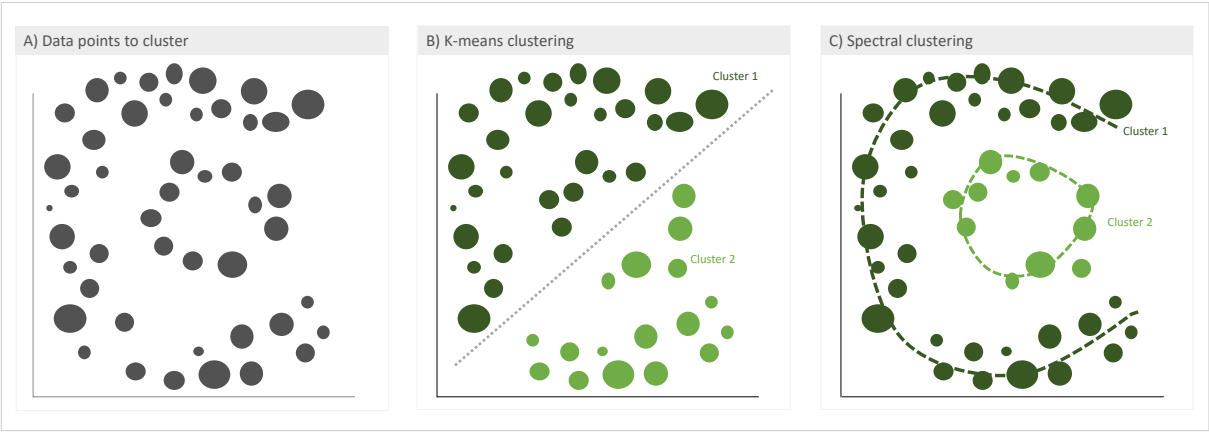



**Figure *1.8*. Clustering data points according to different clustering algorithms.**

*Spectral clustering can identify clusters with complex data structure shapes of different sizes which may not be captured by classic methods such as K-means clustering which tend to search for round-shaped clusters of relatively similar sizes.*

These exploratory data-driven approaches are often accompanied by concerns associated with overfitting, cluster assignment stability, and result accuracy which make it challenging to determine the validity and generalisability of the data-driven classifications (Monti *et al.*, 2003; Hawkins, 2004; Simpson, Armstrong and Jarman, 2010); however, recent advances in bioinformatics and data science have introduced novel 'consensus clustering' and 'integrative clustering' protocols which aim to alleviate these concerns. By collating outputs from multiple clustering iterations to estimate a 'consensus' clustering solution, previous 'consensus clustering' approaches have utilised either bootstrap sampling methods to cluster random subsamples of the dataset over several iterations (Figure 1.9) or ensemble learning methods to generate multiple clustering solutions based on different clustering algorithms or clustering model hyperparameter selections (Monti *et al.*, 2003; Simpson, Armstrong and Jarman, 2010; Chiu and Talhouk, 2018; Xu and Goodacre, 2018; Khan *et al.*, 2019; Markello *et al.*, 2020, 2020). Resultant bootstrap sampling or ensemble learning clustering solutions are then collated, and data-driven probabilistic metrics are used to generate the final 'consensus clustering' classification results (Figure 1.9).

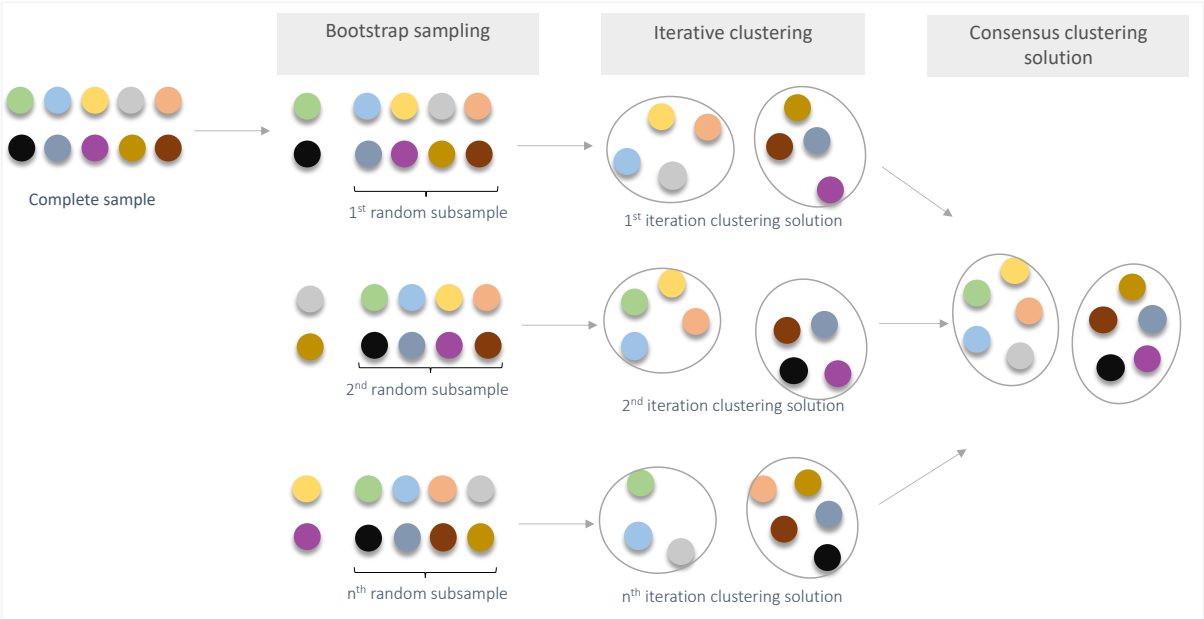



***Figure 1.9. Illustrative example of a consensus clustering approach.***

*Consensus clustering solution generated based on multiple clustering solutions generated from clustering random subsamples of the original dataset, over multiple iterations.*

Other advances in the field have developed cutting-edge 'integrative clustering' techniques which incorporate sophisticated data integration steps prior to the implementation of clustering algorithms, allowing for an enhanced ability to capture complex latent data structure patterns based on multivariate inputs from multiple distinct data types. Similarity Network Fusion (SNF) is a

novel integrative clustering approach which has been widely applied across multiple biomedical domains (Wang *et al.*, 2014; Cavalli *et al.*, 2017; Stefanik *et al.*, 2018; Raita *et al.*, 2020; Bhalla *et al.*, 2021; Hong *et al.*, 2021; Jacobs *et al.*, 2021; Markello *et al.*, 2021) and shows superior performance to previously existing integrative clustering techniques (Wang *et al.*, 2014). SNF uses nonlinear integration methods based on message-passing theory to iteratively update information from multiple similarity networks generated from distinct input data types. Finally, it generates a 'final fused similarity matrix' with a high signal-to-noise ratio which captures subject similarity based on multi-modal input data (Wang *et al.*, 2014) (Figure 1.10). This robust iterative integration process mitigates against noise by producing a final network which discards weak similarities and retains the ones which are present across multiple data types. In addition, these sophisticated approaches allow for the ability to cluster a sample based on high-dimensionality and multi-modal data types (e.g., neuroimaging, genomic, socio-demographic, and behavioural data).

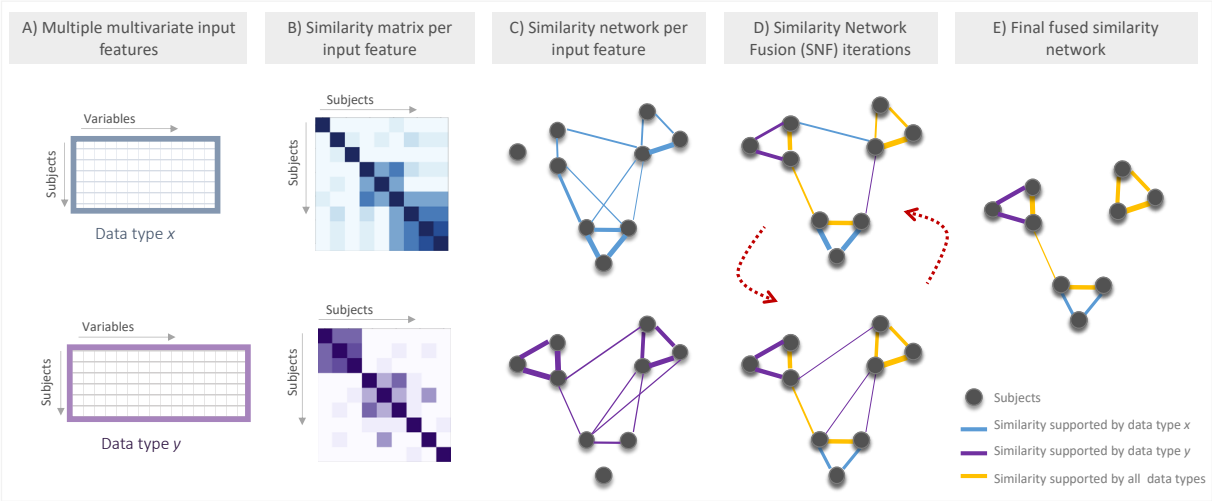



***Figure 1.10. Schematic illustrating how Similarity Network Fusion (SNF) integrates multi-modal data types.***

*A) Using multiple different data types from the same set of subjects, the SNF pipeline generates B) similarity matrices capturing pairwise similarities between each two subjects and uses this information to generate C) weighted similarity networks for each data type. D) Over multiple iterations, the SNF pipeline updates the network using information from the distinct similarity networks, making it increasingly more similar to the other similarity networks with each iteration. E) A final "fused" (i.e., integrated) similarity network is generated and can be used as an input in spectral clustering algorithms. Figure reproduced based on (Wang et al., 2014), with permission from copyright owners.*

### 1.2.4 Translatability of methodological solutions to preterm samples

There has been a rapid increase in the number of studies investigating intricate neurobiological markers of behavioural heterogeneity within psychiatric populations. Advanced neuroimaging analyses quantifying micro-structural, macro-structural, and structural and functional connectivity properties are seen to distinguish between transdiagnostic data-driven behavioural subgroups with greater differentiation than group-wise comparisons of traditional

diagnostic classification groupings (Bathelt *et al.*, 2018; Stefanik *et al.*, 2018; Kushki *et al.*, 2019; Vaidya *et al.*, 2020). By implementing sophisticated and powerful data-driven methods to characterise behavioural heterogeneity based on multiple input modalities, psychiatric studies have not only identified nuanced brain-behavioural relationships, but also complex relationships between multiple distinct behavioural processes and the roles of environmental and clinical risk factors (Stefanik *et al.*, 2018; Hong *et al.*, 2021; Jacobs *et al.*, 2021).

In preterm samples, however, studies mapping neurobiological markers of behavioural heterogeneity remain rather scarce (Ross *et al.*, 2016; Lean *et al.*, 2020; Bogičević *et al.*, 2021). In fact, of those few preterm studies using sophisticated data-driven stratification (Ross *et al.*, 2016; Lean *et al.*, 2020; Bogičević *et al.*, 2021) or distinct M-CHAT psychometric scoring criteria (Luyster *et al.*, 2011; Moore, Johnson, *et al.*, 2012) to delineate behavioural heterogeneity, only two studies measured perinatal brain injury using qualitative ratings (Ross *et al.*, 2016; Bogičević *et al.*, 2021) and one study assessed neonatal brain tissues using quantitative measures of cerebral spinal fluid, white matter, cerebellum, cortical, and deep nuclear grey matter volumes (Lean *et al.*, 2020). However, no studies have so far investigated differences between distinct behavioural subgroups using advanced neuroimaging methods capturing either white matter microstructural characteristics, functional connectivity, or volume differences at the whole-brain level in preterm samples. It also remains to be explored whether brain measures distinguishing between distinct behavioural subgroups are unique to preterm born individuals or may be generalisable to FT born individuals presenting with similar behavioural profiles. Moreover, whether the behavioural heterogeneity observed in preterm children (Johnson *et al.*, 2018; Burnett *et al.*, 2019; Lean *et al.*, 2020) lasts into adulthood and whether neural alterations associated with specific behavioural outcomes are independent of birth status, also remain to be elucidated.

Furthermore, in previous preterm studies, classic data-driven stratification techniques have been implemented to delineate behavioural heterogeneity based on multi-dimensional measures of behaviour (Luu *et al.*, 2011; Poehlmann *et al.*, 2015; Ross *et al.*, 2016; Johnson *et al.*, 2018; Burnett *et al.*, 2019; Lean *et al.*, 2020; van Houdt *et al.*, 2020; Bogičević *et al.*, 2021). In those studies, analyses investigating whether clinical or environmental factors differentiated between the distinct behavioural subgroups were typically explored post-hoc, helping elucidate the protective or detrimental roles specific clinical or environmental factors may play in mediating behavioural outcomes. However, the use of advanced techniques such as SNF to integrate clinical and environmental factors, as well as specific behavioural outcomes prior to clustering, could help in the discovery of more nuanced subtypes, as was seen in psychiatric samples (Hong *et al.*, 2021).

In conclusion, the use of these sophisticated and powerful data-driven methods to stratify behavioural heterogeneity based on inputs from multi-dimensional behavioural measures (e.g., socio-emotional, cognitive, psychiatric, etc.) and the implementation of advanced neuroimaging analysis techniques to measure intricate structural and functional brain changes can help tackle the challenges associated with identifying brain-behavioural associations in preterm samples which remain to be elucidated.

## 1.3 Thesis aims and hypotheses

The overarching objective of my PhD is to map neural markers underlying the neurodevelopmental heterogeneity observed in preterm samples. To address specific gaps in the literature, this thesis consists of four experimental studies, which aim to i) stratify behavioural heterogeneity in VPT samples by using distinct psychometric scoring criteria and advanced data-driven clustering approaches, ii) implement advanced neuroimaging analysis techniques to characterise neural markers associated with distinct behavioural outcomes, and iii) investigate the influence of clinical and environmental factors.

### 1.3.1 Experimental Study #1 – Using distinct M-CHAT psychometric scoring criteria to delineate longitudinal brain-behavioural heterogeneity in VPT toddlers

**Study aims**

**1)** To investigate whether brain-behavioural heterogeneity in VPT toddlers can be characterised by stratifying individuals according to distinct psychometric screening criteria for autism using the M-CHAT.

**2)** To explore whether developmental delay mediates or interacts with childhood autism traits in the distinct psychometric screening subgroups.

**Study hypotheses**

**1)** Neonatal structural volumes and childhood autism traits would differ between the distinct psychometric screening subgroups.

**2)** Developmental delay would be mediating and moderating the onset of autism traits in childhood in certain psychometric screening subgroups but not others.

### 1.3.2 Experimental Study #2 – Using data-driven integrative consensus clustering to parse longitudinal brain-behavioural heterogeneity in VPT children

**Study aims**

**1)** To parse heterogeneity in neonatal clinical and social risk and childhood behavioural outcomes using data-driven integrative consensus clustering techniques.

**2)** To explore differences in neonatal brain volumes and structural and functional connectivity between the distinct data-driven subgroups using advanced neuroimaging analysis approaches.

**Study hypotheses**

1) It would be possible to identify distinct subgroups of VPT children exhibiting nuanced clinical, environmental, and behavioural patterns.
2) That unique neonatal neural patterns would differentiate between the distinct subgroups.

### 1.3.3 Experimental Study #3 – Elucidating brain-behavioural heterogeneity in VPT and FT children using data-driven consensus clustering

**Study aims**

1) To use advanced neuroimaging analyses to compare resting state functional connectivity and structural volumes differentiating between groups of VPT and FT children stratified both in terms of:
    a. Clinical birth status – i.e., VPT vs FT birth.
    b. Data-driven behavioural subgroups identified using consensus clustering regardless of gestational age at birth.
2) To explore differences in clinical and environmental factors between the distinct data-driven behavioural subgroups.

**Study hypotheses**

1) Unique neural patterns would differentiate between the distinct data-driven behavioural subgroups, regardless of clinical birth status.
2) Clinical or environmental factors may probe risk or resilience to behavioural difficulties in some behavioural subgroups but not others.

### 1.3.4 Experimental Study #4 – Elucidating brain-behavioural heterogeneity in VPT and FT adults using data-driven consensus clustering

**Study #4 aims** to characterise resting state functional connectivity using a sample of VPT and FT adults to explore the same set of aims investigated in Study #3

**Study hypotheses** are the same as those described above for Study #3

# CHAPTER 2 - Methodology and datasets

## 2.1 Datasets and study design

Two cohort study datasets, designed to investigate long-term brain and behavioural sequelae of VPT birth were used to address the aims of this PhD. They are referred to here as 1) the Evaluation of Preterm Imaging (ePrime) study cohort and 2) the University College Hospital London (UCHL) study cohort and are described in detail in this chapter. Figure 2.1 depicts the clinical, environmental, brain, and behavioural data types collected in the two cohorts across the multiple follow-up timepoints.

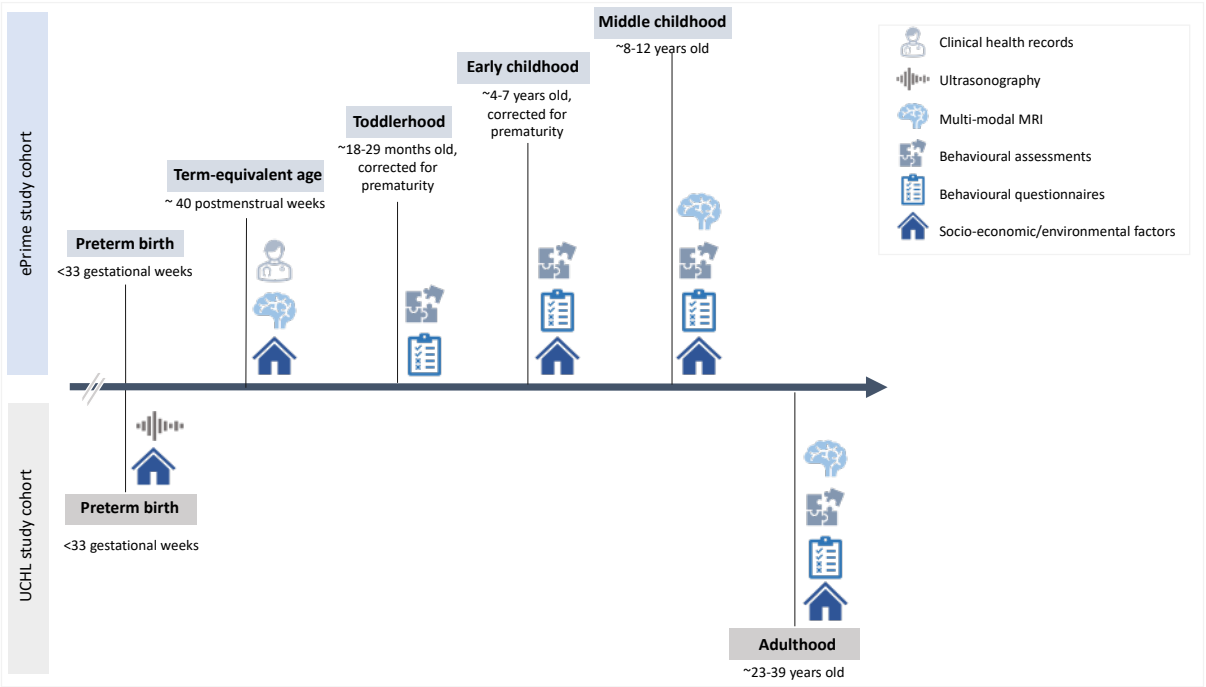



***Figure 2.1. Data types collected across the multiple follow-up timepoints.***

*Illustration depicting clinical, environmental, brain, and behavioural data types collected across the multiple follow-up timepoints of the ePrime and UCHL cohort studies which were used to investigate the aims of this PhD thesis.*

### 2.1.1 The ePrime study cohort

The ePrime study initially recruited 511 VPT born infants from 14 London hospitals between 16 April 2010 and 31 July 2013 (EudraCT 2009-011602-42) (Edwards *et al.*, 2018). Inclusion criteria includes infants born at or before 32 weeks of gestation, with non-inpatient mothers who were over 16 years of age. Infants with congenital malformation, prior MRI, metallic implants, parents who did not speak English, or parents who were subject to child protection

proceedings, were not eligible to take part in the study and were therefore excluded from enrolment.

**Ethical approval** for the study was granted by the Hammersmith and Queen Charlotte's Research Ethic Committee (09/H0707/98). Written and verbal informed consent were obtained from the parents or legal guardians of all participating infants.

To date, the ePrime study participants have been assessed at 4 distinct time points throughout their development: 1) infancy (term-equivalent age), 2) toddlerhood (2 years old), 3) early childhood (4-7 years old) and 4) middle childhood (7-12 years old) (Figure 2.1). Flowchart in Figure 2.2 summarizes participant sample sizes across the different timepoints.

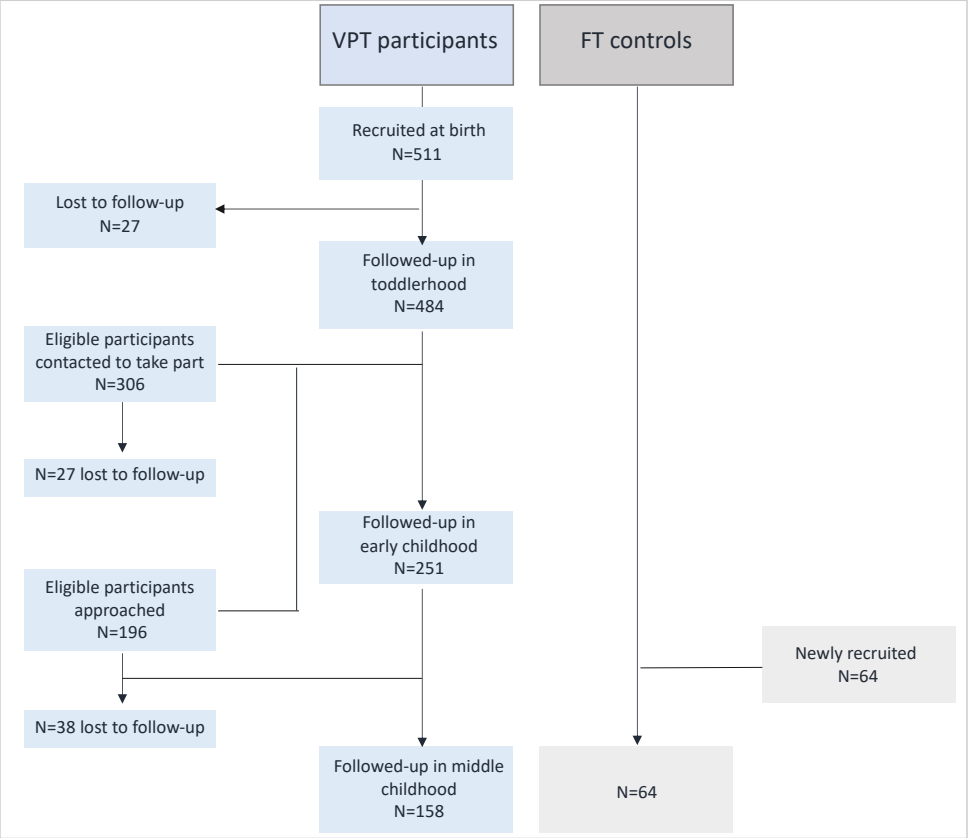



*Figure 2.2. Flowchart describing ePrime study cohort sample sizes.*

### *2.1.1.1 Term-equivalent age (infancy)*

511 participants recruited to take part in the ePrime study underwent multi-modal MRI at term-equivalent age (between 38-53 weeks postmenstrual age at scan) at Queen Charlotte's and Chelsea Hospital, London. Perinatal clinical and environmental measures were also collected.

**Multi-modal MRI acquisition.** Images were acquired on a 3-Tesla MR imaging system (Philips Medical Systems, Best, The Netherlands) using an 8-channel phased array head coil. T1- and T2-weighted high-resolution structural images were acquired with the following pulse sequence imaging parameters; 3D MPRAGE T1-weighted high resolution volume: repetition time (TR) = 17ms, echo time (TE) = 4.6 ms, flip angle = 13°, slice thickness = 0.8mm, field of view (FOV) = 210 x 210 $mm^2$, matrix = 256 x 256, voxel size=0.82 x 0.82 x 0.8 $mm^3$; T2-weighted fast spin echo imaging: TR = 8,670 ms, TE = 160 ms, flip angle = 90°, slice thickness = 1 mm, FOV = 220 x 220 $mm^2$, matrix = 256 x 256 voxel size = 0.86 x 0.86 x 1 $mm^3$. Single shot echo planar d-MRI was acquired in the transverse plane in 32 non-collinear directions: (TR = 8000 ms; TE = 49 ms, slice thickness = 2 mm, FOV = 224 x 224 $mm^2$, matrix = 128 x 128, voxel size=1.75 x 1.75 x 2 $mm^3$, b-value = 750 s/$mm^2$, SENSE factor 2). Resting state f-MRI was acquired using T2 gradient echo planar imaging: TR=1.5s, ET=45ms, flip angle= 90°, 256 volumes, slice thickness = 3.25 mm, in-plane resolution=2.5 x 2.5 $mm^2$, 22 slices, scan duration=6.4mins. Infants were scanned while asleep and earmuffs (Mini Muffs, Natus Medical Inc., San Carlos, CA, USA) and earplugs moulded from silicone-based putty (President Putty, Coltene Whaledent, Mahwah, NJ, USA) were placed in their ears for protection. The imaging sessions were supervised by an experienced paediatrician and 25-50 mg/kg of chloral hydrate was used for infants whose parents preferred sedation prior to the scan (87%).

**Perinatal clinical and demographic data.** The following clinical and demographic measures were collected from the infant's electronic health records using the Standardised Electronic Neonatal Database: multiple pregnancy, antenatal hypertension, premature rupture of membranes, urinary tract infection, gestational diabetes, oligohydramnios, polyhydramnios, drug abuse, in vitro fertilization, bacterial infection, mode of delivery, twin-to-twin transfusion, chorioamnionitis, intrauterine growth restriction, antenatal steroid administration, surfactant administration, treatment for patent ductus arteriosus, surgical treatment for necrotising enterocolitis, formula feeding, days on mechanical ventilation, days on continuous positive airway pressure, and days on parenteral nutrition, gestational age, sex, birth weight, feeding on maternal expressed breast milk, preeclampsia and pregnancy induced hypertension and placental abruption or antenatal haemorrhage. In order to generate a dimensional summary score of neonatal clinical severity, a Principal Component Analysis (PCA) was used to reduce the dimensionality of the described clinical measures. The PCA generated a summary score termed the "neonatal sickness index", whereby higher index scores reflect higher levels of neonatal sickness or clinical risk (Kanel *et al.*, 2021).

**Presence and degree of neonatal brain injury** was detected using MRI at term-equivalent age by an experienced perinatal neuroradiologist. Neonatal brain injury was classified according to the following three categories: major lesions (evidence of periventricular leukomalacia, parenchymal haemorrhagic infarction, or other major ischemic or haemorrhagic lesions), minor lesions (any other lesions), or no lesions (Inder *et al.*, 2003).

**Environmental factors.** Parental postcode at time of birth was also collected and was used to calculate the Index of Multiple Deprivation (IMD) (Department for Communities and Local Government, 2011; https://tools.npeu.ox.ac.uk/imd/), which is used as a proxy of environmental risk. IMD is a postcode-derived measure quantifying neighbourhood deprivation in the following seven domains: income, employment, education/skills/training, health, crime, housing and living environment.

#### *2.1.1.2 2-year-old follow-up (toddlerhood)*

497 participants were subsequently followed-up for behavioural assessments during toddlerhood (between 18 and 29 months).

**Behavioural measures.** Participating parents were asked to complete the following parent-rated questionnaires: the Parent Report of Children's Abilities Revised (PARCA-R) (Saudino *et al.*, 1998; Johnson *et al.*, 2004) which measures the toddler's verbal (sentence complexity and vocabulary) and non-verbal cognitive skills; and the M-CHAT (Robins *et al.*, 2001) which is a screening tool for autism. The toddlers underwent an assessment measuring fine and gross motor skills, cognition, and expressive and receptive linguistic development using the Bayley Scales of Infant Development, Third Edition (Bayley-III) (Bayley, 2006).

#### *2.1.1.3 4-7-year-old follow-up (early childhood)*

251 children from the ePrime cohort were subsequently followed up for behavioural assessments at St Thomas' Hospital, London. Participants who had consented to be approached for taking part in future research studies and were between the ages of 4 and 7 years before the early childhood follow-up study end date (i.e., 1st September 2019) were contacted to take part (N=306). Invitations were sent out by chronological order of age. Informed written consent was obtained from all participants. Ethical approval for follow-up was granted by the Stanmore Research Ethics Committee (14/LO/0677).

**Behavioural measures.** To measure psychopathology, socio-emotional processing and cognitive abilities, participating parents were asked to complete a series of questionnaires and children underwent a battery of assessments with trained researchers. Administered questionnaires and assessments are summarised in Table 2.1.

***Table 2.1. Behavioural measures administered across the different follow-up timepoints.***

| | | ePrime cohort | | UCHL cohort |
|---|---|---|---|---|
| | **Toddlerhood (2 years)** | **Early childhood (4-7-years)** | **Middle childhood (7-12 years)** | **Adulthood (23-39 years)** |
| **Behavioural outcomes** | | | | |
| *Psychopathology* | -- | SDQ | SDQ, DAWBA | GHQ-12, CAARMS |
| *Delusional ideation/prodromal symptoms* | -- | -- | PQ-BC | PDI |
| *Anxiety* | -- | -- | SCAS | -- |
| *Autistic traits* | M-CHAT | SRS-2 | SRS-2 | AQ10 |
| *Temperament* | -- | CBQ | TMCQ | -- |
| *Empathy* | -- | EmQue | -- | -- |
| *Emotion recognition* | -- | ERT | ERT | ERT |
| *Emotion regulation* | -- | -- | ERC | -- |
| *Socialisation* | -- | -- | Friends survey [a] | SAS, RFS |
| **Cognitive abilities** | | | | |
| *Executive function/attention* | -- | BRIEF-P | BRIEF-2 | TMT-B, HSCT, IED, SOC, PAL, MOT, CPT |
| *Verbal fluency* | | | | COWAT-FAS |
| *Intelligence* | -- | WPPSI | WISC | WASI |
| **Early development** | | | | |
| *Language, cognitive, and motor* | Bayley-III | -- | -- | -- |
| *Language and cognition* | PARCA-R | -- | -- | -- |

***Note***. *[a] Friend survey is an in-house child-friendly adaption of the adult GENSI survey (Stark and Krosnick, 2017). It quantifies the number of friends forming a child's network of friends, the closeness to each friend, and the friends network transitivity (i.e., inter-connectedness of their friendships). Measure abbreviations and references: AQ- 10= Autism Quotient (Allison, Auyeung and Baron-Cohen, 2012; Booth et al., 2013); Bayley-III= the Bayley Scales of Infant Development, Third Edition (Bayley, 2006); BRIEF-2 = Behavior Rating Inventory of Executive Function, Second Edition (Gioia,*

*Isquith and Roth, 2018); BRIEF-P= Behavior Rating Inventory of Executive Function pre-school version (Sherman and Brooks, 2010); CANTAB = Cambridge Neurophysiological Test Automated Battery; (CANTAB; CANTABeclipse version, 2003) (Fray, Robbins and Sahakian, 1996); CAARMS = Comprehensive Assessment of At-Risk Mental States (Yung et al., 2005); CBQ = Very Short Form - Childhood Behavioral Questionnaire (Rothbart et al., 2001); COWAT = Controlled Oral Word Association Test (Benton, Hamsher and Sivan, 1983); CPT = Continuous Performance Test (Conners et al., 2003); DAWBA= Development and Well-Being Assessment (Goodman et al., 2000); EmQue = Empathy Questionnaire (Rieffe, Ketelaar and Wiefferink, 2010); ERC=Emotion Regulation Checklist (Shields and Cicchetti, 1997); ERT = Emotion Recognition Task (Montagne et al., 2007); GHQ = General Health Questionnaire (Goldberg and Williams, 1991); HSCT= Hayling Sentence Completion Test (Burgess and Shallice, 1997); IED = Intra-Extra Dimensional Set Shift (CANTAB; CANTABeclipse version, 2003) (Fray, Robbins and Sahakian, 1996); M-CHAT= Modified Checklist for Autism in Toddlers (Robins et al., 2001); MOT = Motor Screening Task (MOT) (CANTAB; CANTABeclipse version, 2003) (Fray, Robbins and Sahakian, 1996); PAL= Paired Associates Learning (CANTAB; CANTABeclipse version, 2003) (Fray, Robbins and Sahakian, 1996); PARCA-R= Parent Report of Children's Abilities Revised (Saudino et al., 1998; Johnson et al., 2004); PDI= Peters Delusion Inventory (Peters et al., 2004); PQ-BC= Prodromal Questionnaire–Brief Child Version (Karcher et al., 2018); RFS=Role Functioning Scale (Goodman et al., 1993); SAS = Social Adjustment Scale (Weissman and Bothwell, 1976); SCAS=Spence Children's Anxiety Scale (Nauta et al., 2004); SDQ = Strengths and Difficulties Questionnaire (Goodman, 2001); SOC = Stockings of Cambridge; (CANTAB; CANTABeclipse version, 2003) (Fray, Robbins and Sahakian, 1996); SRS-2 = Social Responsiveness Scale – Second Edition (Constantino and Gruber, 2012); TMT-B = Trail Making Task B; (Tombaugh, 2004); TMCQ=Temperament in Middle Childhood Questionnaire (Simonds and Rothbart, 2006); WASI = Wechsler Abbreviated Scale of Intelligence (Wechsler, 1999); WISC-IV = Wechsler Scale Intelligence for Children (David Wechsler, 2012); WPPSI-IV= Wechsler Preschool & Primary Scale of Intelligence (D Wechsler, 2012).*

**Environmental factors**. Child residential postcode at time of follow-up was used to calculate IMD scores, which were used as a proxy for social risk in early childhood. Parents were also asked to complete an adapted version of the Cognitive Stimulating Parenting Scale (CSPS) (Wolke, Jaekel, *et al.*, 2013), and total CSPS scores were used to measure the cognitively stimulating environment at home.

#### *2.1.1.4 7-12-year-old follow-up (middle childhood)*

A later follow-up study investigating brain and behavioural development in middle childhood was launched in 2018. ePrime study participants who were between the ages of 7 and 12 years and had consented to be contacted about future research studies, were invited to take part in this follow-up study at the Centre for the Developing Brain, St Thomas' Hospital, London. To date, 196 ePrime participants have been contacted. Invitations were sent in chronological order of birth and recruitment is still ongoing for this study. Children meeting study exclusion criteria of having impairments affecting their capacity to complete assessments (i.e., severe learning difficulties, moderate to severe cerebral palsy, blindness, or deafness) or clinical history of neurological conditions such as meningitis or head injury, were excluded from enrolment.

A control group of full-term (FT) children was recruited from the community. In order to be eligible to take part, FT children had to be 7-12 years old, born at term (between 38-42 weeks of gestation), and attending mainstream school. FT children meeting any of the exclusion criteria mentioned above or having contraindications for MRI (such as claustrophobia or metal implants) were excluded from the study.

Participants taking part in this study underwent multi-modal MR imaging and behavioural assessments. Informed written parental consent and child assent were obtained from all

participants. Ethical approval for this study was granted by the Stanmore Research Ethics Committee (18/LO/0048), and the London South East Research Ethics Committee (19/LO/1940).

**Multi-modal MRI acquisition.** Images were acquired with a 32-channel head coil using a dedicated neonatal and paediatric scanner (Philips 3T Achieva system) placed in the Evelina Newborn Imaging Centre, Evelina London Children's Hospital. s-MRI 3D MPRAGE T1-weighted images were acquired using the following sequence parameters: TR =7.9ms, ET = 3.6 ms, TI = 900 ms; flip angle = 8°, FOV = 240 x 220 x 160 $mm^3$, voxel size = 1 mm isotropic, SENSE factor of 1.5 along the first phase encoding direction and 2 along the second direction. d-MRI was acquired in 2 shells (44 non-collinear directions with b-value = 750 $s/mm^2$ and 64 directions with b-value = 2500$s/mm^2$), multiband 4, voxel size = 1.75 mm isotropic. Multi-slice gradient echo EPI f-MRI (900 volumes, TR=1160 ms, TE=33 ms, flip angle = 60 degrees, acquisition matrix = 88 x 87 mm, acquisition voxel size = 2.5 x 2.5 x 2.5 $mm^3$, reconstruction voxel size = 1.9 x 1.9 x 2.5 $mm^3$, FOV = 220 x 220 x 35 $mm^3$, multiband = 4) was acquired, while children watched the '*Inscapes'* low-cognitive load and non-narrative movie paradigm depicting abstract moving shapes (https://www.headspacestudios.org/inscapes), which is found to reduce in-scanner head motion, enhance wakefulness compliance, and improve ability to detect intrinsic functional networks (Vanderwal *et al.*, 2015). Using the same multi-slice gradient echo EPI parameters, task-based f-MRI was acquired while the children were engaged in an emotion-processing face-matching task. During this task, children are presented with images of faces expressing different emotions (emotional stimuli) or images of shapes (neutral stimuli) and asked to choose items that match (Marusak, Carré and Thomason, 2013).

**Behavioural measures.** Psychopathology, socio-emotional processing, and cognitive abilities were assessed using parent-rated questionnaires or interviews and neuropsychological assessments administered to the child. Measures administered are summarised in Table 2.1.

**Environmental factors.** Child residential postcode at time of follow-up was used to calculate IMD scores, which were used as a proxy for social risk in early childhood.

### 2.1.2 UCHL study cohort

Between 1979 and 1985, 473 VPT infants born at <33 weeks of gestation were recruited from the Neonatal Unit at University College Hospital in London (Figure 2.3). Individuals were followed-up during adulthood at the age of 28-35 years and underwent multi-modal MRI and

behavioural assessments. Participants meeting study exclusion criteria including severe hearing and motor impairments or history of neurological complications (i.e., meningitis, head injury, cerebral infections) were excluded from enrolment. The UCHL study also conducted follow-up studies to collect behavioural data at 1, 4, 8, 14, and 18 years, and multi-modal MRI at 14 and 18, but data from those timepoints are not used in this thesis.

Controls born at FT (38-42 weeks' gestation) and weighing >2500 grams at birth were recruited from the community at the time of the follow-up study in adulthood (Figure 2.3). Exclusion criteria for the control group included any clinical complications at birth, prolonged gestation (>42 weeks), having severe hearing and motor impairments or meeting any of the above exclusion criteria.

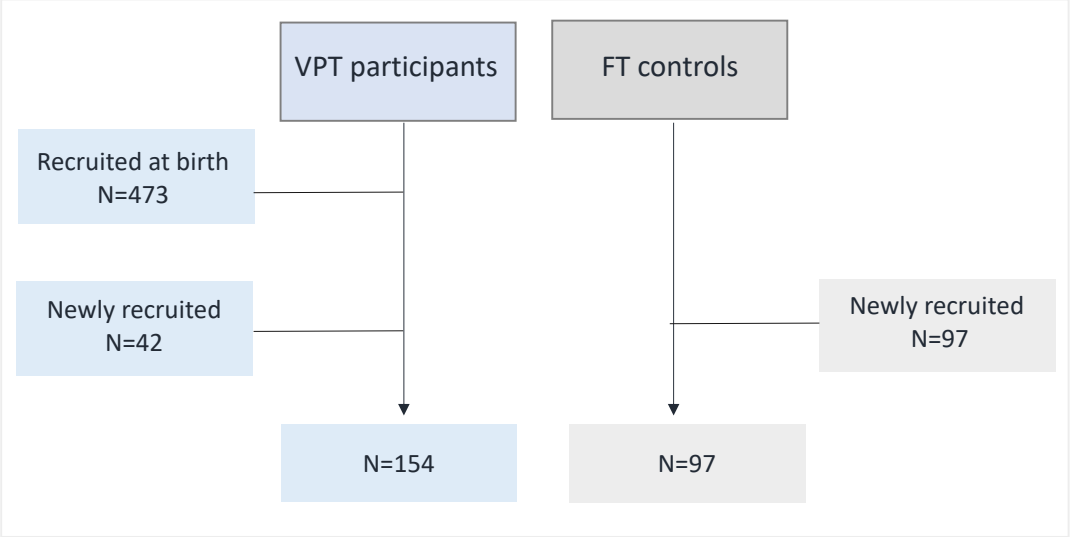



*Figure 2.3. Flowchart describing UCHL study cohort recruitment sample sizes.*

**Multi-modal MRI acquisition** was collected using a 3 Tesla Signa MR scanner (General Electric Healthcare) at the Maudsley Hospital, London. Gradient echo EPI resting state f-MRI was acquired while participants stared at a central cross on a screen for 8 minutes 32 s, using the following parameters 256 volumes, TR=2000 ms, TE=30 ms, flip angle=75 degrees, matrix=64x64, 37 non-contiguous slices of 2.4 mm thickness, 1.1 mm interslice gap, and 3.4 mm in-plane resolution. Structural fast spoiled gradient-echo (FSPGR) pulse sequence T1-weighted images were acquired using TR = 7.1 ms, TE = 2.8ms, matrix = 256 x 256 $mm^2$, voxel size = 1.1 mm isotropic.

**Behavioural measures**. A series of questionnaires, interviews, and cognitive assessments were administered to measure psychopathology, mental well-being, socio-emotional processing,

and cognitive abilities in the adults. Specific assessments administered are summarised in Table 2.1.

**Perinatal clinical data.** Perinatal brain injury was determined based on brain ultrasound scans at term-equivalent age. Brain injury categories were labelled according to the following classifications: minor injury (grade I – II periventricular haemorrhage without ventricular dilation), major injury (grade III – IV periventricular haemorrhage with ventricular dilation), or no injury (Bowerman *et al.*, 1984).

**Environmental factors**. Socio-economic status was classified according to participant occupation at the time of the follow-up study and parental occupation at time of birth. Following criteria defined by the Office of National Statistics, Standard Occupational Classification, 1980, socio-economic status was characterised by three levels: I (Higher managerial, administrative and professional occupations), II (Intermediate occupations, small employers and own account workers), and III (Routine and manual occupations – lower supervisory and technical and semi-routine and routine occupations).

## 2.2 Methodology

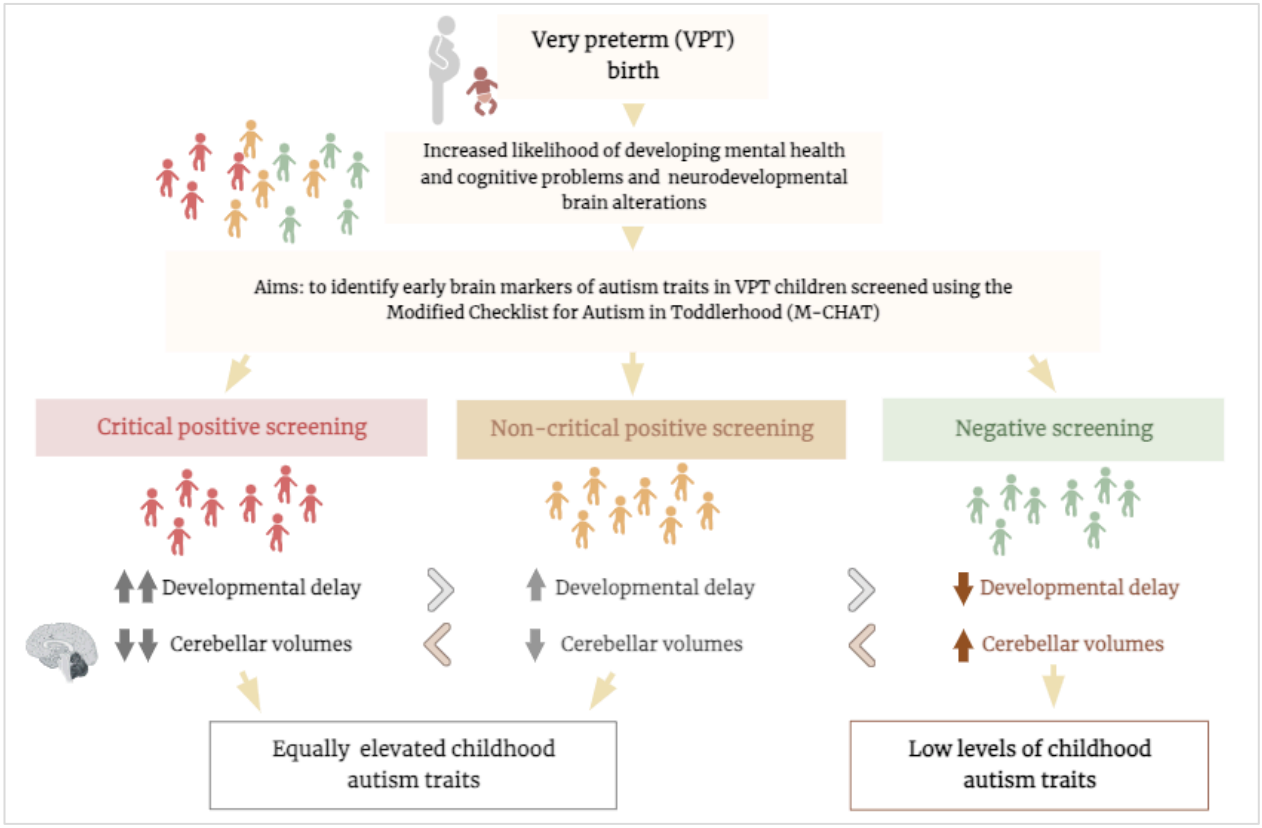
The methods used to investigate the specific aims set out in this PhD thesis are described in the Methods sections of the specific study chapters.

# CHAPTER 3 - Study #1: Using distinct M-CHAT psychometric scoring criteria to delineate longitudinal brain-behavioural heterogeneity in VPT toddlers



***Reference:*** *Hadaya, L., Vanes, L., Karolis, V., Kanel, D., Leoni, M., Happé, F., Edwards, A. D., Counsell, S. J., Batalle, D., & Nosarti, C. (2022). Distinct Neurodevelopmental Trajectories in Groups of Very Preterm Children Screening Positively for Autism Spectrum Conditions. Journal of Autism and Developmental Disorders. https://doi.org/10.1007/s10803-022-05789-4*

**Study #1 graphical abstract** *(created with BioRender.com):*




## 3.1 Abstract

Very preterm (VPT; < 33 weeks' gestation) toddlers screening positively for autism spectrum conditions (ASC) may display heterogenous neurodevelopmental trajectories. Here we studied neonatal brain volumes and childhood ASC traits evaluated with the Social Responsiveness Scale (SRS-2) in VPT-born toddlers (N = 371; median age 20.17 months) sub-divided into three groups based on their Modified-Checklist for Autism in Toddlers scores. These were: those screening positively failing at least 2 critical items (critical-positive); failing any 3 items, but less than 2 critical items (non-critical-positive); and screening negatively. Critical-positive scorers had smaller neonatal cerebellar volumes compared to non-critical-positive and negative scorers. However, both positive screening groups exhibited higher childhood ASC traits compared to the negative screening group, suggesting distinct aetiological trajectories associated with ASC outcomes.

## 3.2 Introduction

The parent-rated Modified Checklist for Autism in Toddlers (M-CHAT), assessing child skills and behaviours, was developed as a screening tool for autism spectrum conditions (ASC) (Robins *et al.*, 2001). ASC are characterised by two sets of core symptoms: (a) social communication and interaction deficits (SCI), which reflect difficulties in non-verbal social gestures, socio-emotional reciprocity and maintaining and developing social relationships, and (b) restricted interests and repetitive behaviours (RRBs), which include restricted and fixated interests, ritualised behaviours and altered sensitivity to sensory stimuli (American Psychiatric Association, 2013). According to the original M-CHAT scoring criteria, a positive M-CHAT screening is obtained when a child fails two or more 'critical' items within a set of six (e.g., "Does your child imitate you?", "Does your child take an interest in other children?"), or three or more items overall (Robins et al., 2001. However, research in low-risk toddlers has more recently led to the recommendation of abandoning these criteria in favour of a total number of items failed, as this approach has been shown to improve the tool's sensitivity to identify a later ASC diagnosis (Chlebowski *et al.*, 2013).

Studies in high-risk samples using the original screening criteria have shown that very preterm (VPT; < 32 weeks' gestation) and extremely preterm (EPT; < 28 weeks' gestation) born toddlers are more likely to screen positively on the M-CHAT (21–25%; Kuban et al., 2009; Limperopoulos et al., 2008), compared to full-term born toddlers (5.7%; Kleinman et al., 2008). These findings, together with those showing a higher prevalence of ASC diagnoses in children born VPT (7%) compared to those born at term (1.5%; Agrawal et al., 2018; Joseph et al., 2017), suggest that VPT children may be vulnerable to experiencing both subthreshold and clinical core ASC symptoms. However, in high-risk EPT/VPT toddlers the interpretability of the M-CHAT screening has been questioned (Luyster *et al.*, 2011; Moore, Johnson, *et al.*, 2012), as these children tend to display impaired social and communication skills, which are shared features of both the so-called "preterm behavioural phenotype" (Johnson and Marlow, 2011) and ASC traits (American Psychiatric Association, 2013). Moore et al. (2012) suggested that the two original M-CHAT positive scoring criteria may differentiate between EPT toddlers with and without neurodevelopmental disabilities, as they found that the stricter critical positive screening criteria were associated with more severe neurodevelopmental impairments compared to the more liberal non- critical criteria (Luyster *et al.*, 2011; Moore, Johnson, *et al.*, 2012). Given the increased risk of developmental delay following preterm birth (Blencowe *et al.*, 2013) and the frequent co-occurrence of developmental delay in ASC (Rubenstein *et al.*, 2018), the use of the initially

proposed different M-CHAT positive scoring criteria may therefore aid the identification of subgroups of EPT/VPT toddlers exhibiting distinct neurodevelopmental trajectories.

Widespread alterations in brain development associated with VPT birth (Volpe, 2009a), may at least partly contribute to the increased likelihood of ASC behaviours in VPT children. Structural reductions in volume and alterations in functional connectivity in temporal, prefrontal, limbic and cerebellar regions have been observed in VPT individuals in the neonatal period and beyond (Rogers *et al.*, 2012; Ball *et al.*, 2013, 2016; Healy *et al.*, 2013; Fenoglio, Georgieff and Elison, 2017; Kanel *et al.*, 2022). Alterations in these regions have also been implicated in key components of ASC symptomatology (Ha *et al.*, 2015; Alcalá-López *et al.*, 2018; Ciarrusta *et al.*, 2019; Gandhi and Lee, 2021) and in VPT neonates who develop ASC later in childhood (Ure *et al.*, 2016; Padilla *et al.*, 2017; Eklöf *et al.*, 2019). However, no study to date has explored whether different M-CHAT positive scoring criteria could be used to identify subgroups of VPT toddlers who differ in terms of early brain development and ASC behaviour later in childhood.

In order to address these questions, this study had two main aims: to explore whether distinct M-CHAT screening groups (critical positive, non-critical positive and negative), which have been previously studied in relation to neurodevelopmental impairments in EPT toddlers (Moore, Johnson, *et al.*, 2012), also differed in VPT toddlers in terms of (a) neonatal structural brain volumes and (b) ASC profiles later in child- hood. Exploratory analyses were further conducted to probe the role of developmental delay in shaping the childhood trajectory for ASC traits in the different screening groups, with the use of mediation and moderation analyses.

Our first hypothesis was that both M-CHAT positive screening groups (i.e., critical positive and non-critical positive) would display volumetric reductions at term-equivalent age in brain regions implicated in ASC symptomatology (e.g., temporal, prefrontal cortex and cerebellum) compared to the negative screening group. Our second hypothesis was that toddlers belonging to the two M-CHAT positive screening groups would display more ASC-type behaviours in childhood (age 4–7 years) than toddlers belonging to the negative screening group. Thirdly, exploratory analyses tested two competing hypotheses, namely that the critical positive scorers would either exhibit: (a) fewer ASC-like behaviours than the non-critical positive scorers, indicating that a critical positive screening may reflect developmental delay (Luyster *et al.*, 2011; Moore, Johnson, *et al.*, 2012), rather than persisting ASC behaviours, or (b) similar ASC-like behaviours to the non-critical positive scorers, indicating distinct trajectories leading to similar ASC behaviours (i.e., equifinality; Cicchetti & Rogosch, 1996).

## 3.3 Methods

### 3.3.1 Participants and study design

511 children born at 33 weeks' gestational age or less (median = 30 weeks; range = 23–32 weeks), between April 2010 and July 2013, were enrolled into the "Evaluation of Preterm Imaging" study (ePrime; EudraCT 2009-011602- 42; Edwards et al., 2018) from 14 neonatal units across Lon- don. Inclusion criteria were: birth at or less than 33 weeks' gestation; English-speaking parents not undergoing child protection proceedings; no magnetic resonance imaging (MRI) contraindications or major congenital malformations. Infants underwent multi-modal (T1-weighted, T2-weighted, diffusion and functional) MRI at term-equivalent age (38–44 weeks) and were followed-up for behavioural and cognitive assessments at 2 (N = 484; 95% of the initial sample) and 4–7 years (N = 251; 82% of those children approached for follow-up).

Complete M-CHAT follow-up data at 2 years were avail- able for 371 children (49.60% female; 23.18% born EPT) meeting MRI analysis inclusion criteria: i.e., postmenstrual age (PMA) at scan < 46 weeks, having no periventricular leukomalacia, parenchymal haemorrhagic infarction, or other major ischemic or haemorrhagic lesions detected on MRI or missing T2-weighted or motion corrupted images. 177 children had complete SRS-2 data at the subsequent 4–7-year follow-up (46.90% females; 25.42% born EPT). Sample characteristics are summarised in Table 3.1. The EPT and VPT born children within our cohort did not differ in severity of ASC traits or developmental delay (Table 3.SM1).

*Table 3.1. Sample characteristics.*

| Variables | Median (range) | |
|---|---|---|
| | **2-year follow-up (N = 371)** | **4–7-year follow-up (N = 177)** |
| **GA, weeks** | 30.29 (23.57–32.86) | 30.29 (24–32.86) |
| **IMD score at birth** | 17.71 (1.73–60.58) | 16.12 (1.73–59.16) |
| **PMA at scan** | 42.57 (37.86–44.86) | 42.57 (38.29–44.86) |
| **Neonatal sickness** [a] | −0.30 (−1.36–2.55) | −0.35 (−1.34–2.18) |
| **Corrected age at assessments** | 20.17 (18.37–29.33) months | 4.59 (4.18–7.17) years |

***Note.*** *Sample characteristics (median and range) for 2-year follow-up sample with complete M-CHAT and structural MRI data and for 4–7-year follow-up sample with complete M-CHAT and SRS-2 data. GA gestational age, IMD index multiple deprivation, PMA postmenstrual age* [a]*excluding one subject with incomplete clinical data.*

### 3.3.2 MR imaging data

#### *3.3.2.1 Data acquisition*

A 3-Tesla system (Philips Medical Systems, Best, The Netherlands) was used to acquire MR images using an 8-channel phased array head coil. A paediatrician supervised infant care during MR imaging. Pulse oximetry, temperature, and electrocardiography data were monitored throughout the session. Silicone-based putty (President Putty, Coltene Whaledent, Mahwah, NJ, USA) and neonatal earmuffs (MiniMuffs, Natus Medical Inc., San Carlos, CA, USA) were used for ear protection. Oral chloral hydrate (25–50 mg kg−1) was administered to infants whose parents chose sedation for the procedure (87%). High-resolution anatomical images were acquired with T2-weighted fast spin echo sequences (repetition time = 8,670 ms; echo time = 160 ms; flip angle = 90°, slice thickness = 1 mm, field of view = 220 × 220 $mm^2$, voxel size = 0.86 × 0.86 × 1 $mm^3$).

### 3.3.3 Tensor Based Morphometry

Following methods described in Vanes et al., (2021) and Lautarescu et al., (2021) T2-weighted (images and tissue type segmentations) were registered to a study-specific template using ANTS software Symmetric Normalisation algorithms (Avants et al., 2011). Resultant nonlinear transformation deformation tensor fields (warps) were used to calculate deformation tensor field gradients (log-Jacobian determinant maps) as a measure of relative brain volume. Greater log-Jacobian values represent the extent of contraction voxels undergo following registration (i.e., larger volumes), while smaller values represent volume reductions (Avants & Gee, 2004). Smoothing with 4 mm full-width half-maximum Gaussian filter was applied.

### 3.3.4 Perinatal socio-demographic and clinical data

#### *3.3.4.1 Perinatal clinical data*

With parental consent, the infant's electronic medical records were accessed using the Standardised Electronic Neonatal Database to collect perinatal socio-demographic and clinical data. Data capturing neonatal clinical risk were collected as part of the larger ePrime study (Edwards et al., 2018) as clinical risk can exacerbate the long-term sequelae of VPT birth (Volpe,

2009). A principal component analysis (PCA) summarised 28 perinatal clinical variables explaining 72% of their variance with a single component, which was labelled 'neonatal sickness index', as previously described in Kanel et al. (2021). The variables with the highest factor loadings were: GA, days on total parenteral nutrition, days on continuous positive airway pressure, days on mechanical ventilation and surfactant administration. Clinical variables were coded so that increased neonatal sickness index values indicate greater clinical risk.

#### *3.3.4.2 Perinatal environmental data*

An Index of Multiple Deprivation (IMD) score was computed from the infant's residential postcode at time of birth (Department for Communities and Local Government, 2011; https://tools.npeu.ox.ac.uk/imd/). The IMD summarises area-level information in 7 domains: income, employment, education, health, crime, housing and living environment. Higher IMD scores reflect increased deprivation in the neighbourhood, hence higher social risk.

### 3.3.5 Behavioural and cognitive measures

At the 2-year follow-up, toddlers were assessed with the parent-rated M-CHAT. Critical positive M-CHAT screening was defined by failing any 2 out of the 6 critical items: "Does your child take an interest in other children?", "Does your child ever use his/her index finger to point, to indicate interest in something?", "Does your child ever bring objects over to you to show you something?", "Does your child imitate you?", "Does your child respond to his/her name when you call?", "If you point at a toy across the room, does your child look at it?" (Robins et al., 2001). The definition used by Moore and colleagues (Moore, Johnson, *et al.*, 2012) was used to define 'non-critical' positive screening: failing any 3 or more items, but fewer than two critical items. Toddlers not meeting either of these criteria received a negative M-CHAT screening.

The following measures were used to assess infants' development at 2 years: the Bayley Scales of Infant Development, Third Edition (Bayley-III; Bayley, 2006), which evaluates expressive and receptive language, fine and gross motor skills and composite cognitive scores, and the Parent Report of Children's Abilities Revised (PARCA-R; (PARCA-R; Johnson et al., 2004; Saudino et al., 1998), which evaluates toddlers' vocabulary and sentence complexity and non-verbal cognitive skills.

To reduce the dimensionality of the behavioural outcome data, a PCA was performed. All Bayley-III and PARCA-R index scores were included in the model and the elbow-method was

used to determine the number of principal components explaining most of the variance in the data. A scree plot showing the percentage of variance explained by each principal component (i.e., eigenvalues) suggests an optimal number of 2 principal components (Figure SM 3.1), jointly explaining a cumulative 69% of total variance. Pearson correlations between each of the two resultant principal components and individual index scores were used to define each of the components. PC1 correlated negatively with all Bayley-III and PARCA-R items, resulting in a component summarising global (cognitive, language and motor) developmental delay, while PC2 correlated positively with language items (PARCA-R sentence complexity and vocabulary scores and Bayley-III expressive language scores) and showed negative correlations with gross and fine motor Bayley-III scores (Figure SM 3.2). The first principal component was labelled as a global 'developmental delay' index and the second as a 'language' index.

At the 4- to 7-year-old follow-up, the Social Responsiveness Scale, Second Edition (SRS-2; (Constantino and Gruber, 2012) was administered to measure core ASC symptoms in early childhood; it contains a Social Communication/Interaction (SCI) and a Restricted/Repetitive Behaviour (RRB) subscale. The SCI subscale indexes deficits in behaviours relating to social awareness, cognition, communication, and motivation, and the RRB subscale reflects the severity of restrictive and repetitive patterns of behaviours and interests (Constantino and Gruber, 2012). The SRS-2 shows good internal consistency (Cronbach's alpha = 0.92 and 0.93 for females and males, respectively) as well as construct, convergent and concurrent validity in 5–8-year-old children from the United Kingdom (Wigham *et al.*, 2012).

### 3.3.6 Statistical analyses

#### *3.3.6.1 Univariate phenotypic group differences*

Statistical analyses were conducted using R (version 3.6.1). Non-parametric Kruskal–Wallis tests compared continuous measures (developmental profiles at 2 years, socio-demographic and clinical profiles at birth and SRS-2 SCI and RRB scores at 4–7 years) between M-CHAT groups (*onewaytests* R package; Dag et al., 2018). For categorical variables (sex), Chi-squared test was used. Post-hoc pairwise comparisons were made for variables showing a significant effect of group ($p < 0.05$). Post-hoc pairwise between-group median differences (for continuous variables) or odds ratios (for categorical variables) were reported and post-hoc pairwise comparison p-values were corrected using False Discovery Rate (Benjamini and Hochberg, 1995). A generalised linear model with 10,000 permutations investigating the effect of M-CHAT group on SCI and RRB scores and

correcting for covarying effects of developmental delay, sex, IMD and neonatal sickness index, was also tested (p-permute; https://github.com/lucasfr/grouped_perm_glm).

#### *3.3.6.2 Childhood symptoms exceeding clinical cut-offs for autism*

Having a total SRS-2 T-score greater than or equal to 76 is considered to be clinically meaningful as it indicates a high likelihood of receiving an ASC diagnosis (Constantino and Gruber, 2012). We calculated the number of children scoring above the SRS-2 clinical cut-off within each M-CHAT group. Sample size calculations were then performed in order to ascertain whether the sample size was adequate for predictive validity analyses (Linden, 2020). The following measures were used as inputs in the sample size calculation: expected sensitivity/specificity (52%/84% respectively; Kim et al., 2016), prevalence in current sample (2%) and confidence interval for estimates (95%-CI with CI-width = 0.1).

#### *3.3.6.3 Mass-univariate group differences in brain volume*

Differences in voxel-wise volume (log-Jacobian) measures at term-equivalent age between the three M-CHAT screening groups were investigated using general linear models correcting for sex, PMA, IMD and neonatal sick- ness index. FMRIB Software Library (FSL)'s *randomise* function with 10,000 permutations per run was used for non-parametric permutation testing with Threshold-Free Cluster Enhancement and controlled for family-wise error rate. Significance was set at $p < 0.05$ per contrast, given the exploratory nature of the analysis.

Post-hoc analyses investigating associations between neonatal brain volumes showing between-group differences and ASC traits in childhood are described in the supplemental information (Table SM 3.2). We also explored associations between M-CHAT total items failed and neonatal whole-brain Jacobian values.

#### *3.3.6.4 Testing the role of developmental delay*

To test for a potential role of early developmental delay in explaining (mediating) or exacerbating (moderating) later group differences in core ASC symptoms, analyses using general linear models were conducted.

Specifically, where between-group differences in later ASC symptoms (SRS-2 SCI or RRB) at 4–7 years were observed, we tested whether these differences were significantly mediated by developmental delay at 2 years. In addition, to test whether developmental delay at 2 years shows

a differential relationship with later ASC symptoms in the separate M-CHAT groups, we tested for effects of develop- mental delay and M-CHAT screening, as well as their inter- action, on SRS-2 SCI and RRB scores. Both mediation and moderation analyses used sex, IMD, and neonatal sickness index as confounders. Mediation was tested via bootstrapping of the indirect effect (based on 5000 bootstrap samples) using the R 'mediation' package (Tingley *et al.*, 2014). To adjust for multiple comparisons due to two separate outcome variables (SRS-2 RRB and SCI), 97.5%-confidence intervals (97.5%-CIs) were generated. P-values with a corrected significance threshold of $p < 0.05/2$ (i.e., 0.025) were estimated from non-parametric permutation testing with 10,000 per- mutations (p-permute; https://github.com/lucasfr/grouped_ perm_glm).

## 3.4 Results

### 3.4.1 Comparing M-CHAT groups on socio-demographic, clinical and developmental outcomes

Median scores and F-statistics and p-values comparing M-CHAT group socio-demographic and clinical outcomes are summarised in Table 3.2 and developmental profiles and Bayley-III and PARCA-R composite scores in Table 3.3. The three groups did not differ in corrected age at M-CHAT assessment, PMA at scan, GA at birth, birthweight, neonatal sickness index or language development (Tables 3.2; Table 3.3).

***Table 3.2. Socio-demographic and clinical profiles for M-CHAT groups.***

| Variable | Median (Interquartile range) | | | F-statistic; p-value |
|---|---|---|---|---|
| | Negative (N = 267; 143 female) | Non-critical positive (N = 77; 33 female) | Critical positive (N = 27; 8 female) | |
| **Corrected age at 2 years, months** | 20.20 (0.67) | 20.13 (0.70) | 20.03 (0.37) | F = 2.11; p = 0.310 |
| **Corrected age at 4–7 years, years** | 4.59 (0.58) | 4.67 (0.90) | 4.59 (0.91) | F = 4.78; p = 0.092 |
| **PMA at scan, weeks** | 42.57 (2.00) | 42.71 (2.14) | 42.57 (1.50) | F = 5.27; p = 0.072 |
| **IMD score at birth** | 16.54 (17.00) | 19.92 (15.71) | 25.87 (15.79) | F = 7.63; p = 0.022* |
| **GA, weeks** | 30.29 (3.50) | 30.86 (4.14) | 28.86 (3.36) | F = 3.58; p = 0.167 |
| **Birthweight, grams** | 1315 (570.00) | 1270 (650.00) | 1040 (485.00) | F = 3.27; p = 0.196 |
| **Neonatal sickness index[a]** | −0.36 (1.71) | −0.45 (1.49) | 0.55 (1.59) | F = 3.91; p = 0.142 |

***Note.*** *GA gestational age at birth, IMD index multiple deprivation, PMA postmenstrual age. *p < 0.05. [a] excluding one subject with incomplete clinical data.*

*Table 3.3. Developmental profiles and Bayley-II and PARCA-R composite scores for M-CHAT groups.*

| Variable | Median (Interquartile range) | | | Statistic; p-value |
|---|---|---|---|---|
| | Negative (N = 267) | Non-critical positive (N = 77) | Critical positive (N = 27) | |
| **Developmental profiles at 2 years[b]** | | | | |
| *Developmental delay* | −0.57 (2.67) | 0.88 (2.25) | 2.86 (2.76) | F = 57.40; p < 0.001*** |
| *Language* | −0.02 (1.28) | −0.09 (1.28) | 0.28 (1.54) | F = 2.33; p = 0.313 |
| **Bayley-III composite scores at 2 years[b]** | | | | |
| *Cognitive* | 95.00 (18.75) | 90.00 (20.00) | 82.50 (20.00) | F = 27.28, p < 0.001*** |
| *Language* | 97.00 (20.00) | 83.00 (20.00) | 69.50 (14.00) | F = 45.36, p < 0.001*** |
| *Motor* | 100.00 (9.00) | 94.00 (12.00) | 82.00 (18.00) | F = 45.14, p < 0.001*** |
| **PARCA-R composite scores at 2 years[b]** | | | | |
| *Vocabulary* | 19.00 (18.05) | 10.00 (11.25) | 3.00 (5.75) | F = 34.52; p < 0.001*** |
| *Sentence complexity* | 5.00 (6.00) | 3.00 (4.25) | 4.00 (2.00) | F = 36.21; p < 0.001*** |

***Note.*** *IMD index multiple deprivation, PARCA-R parent report of children's abilities-revised. ***p < 0.001. [b]excluding 31 subjects with incomplete developmental data at 2-years.*

Variables showing a significant group effect were investigated for pairwise group differences and median difference and post-hoc p-values for between-group differences are reported in Table 3.4. In summary, social risk (IMD scores) was lower in negative M-CHAT scorers than critical positive scorers, but did not differ between other groups. Of the three groups, negative M-CHAT scorers had the lowest developmental delay scores (indicating better language, cognitive and motor scores), the critical positive scorers showed the greatest developmental delay and non-critical positive scorers showed intermediate developmental delay scores. There was an overall difference in male-to-female ratios between the different M-CHAT sub-groups (Chi-squared = 7.38; p = 0.025), although all pairwise comparisons were not statistically significant (p > 0.05; M-CHAT negative group compared to non-critical and critical groups; odds ratio = 1.54 and 2.74, p = 0.147 and 0.053, respectively; non-critical group compared to the critical group; odds ratio = 1.78, p = 0.226). The proportion of females in the M-CHAT negative, non-critical positive and critical positive groups were 53.56%, 42.86% and 29.63% respectively.

*Table 3.4. Post-hoc pairwise differences (between the three M-CHAT screening groups) for variables with a significant effect of M-CHAT group.*

| Variable | Median difference (p-value) | | |
|---|---|---|---|
| | Negative vs non-critical positive | Negative vs critical positive | Non-critical positive vs critical positive |
| **IMD score at birth** | −3.38 (p = 0.646) | −9.33 (p = 0.020)* | −5.95 (p = 0.039)* |
| **Developmental delay** | 5.00 (p = 0.003)** | −3.43 (p < 0.001)*** | −1.98 (p <0.001)*** |
| **Bayley-III: *Cognitive*** | 14.00 (p < 0.001)*** | 12.50 (p < 0.001)*** | 7.50 (p = 0.006)** |
| **Bayley-III: *Language*** | 6.00 (p < 0.001)*** | 27.50 (p < 0.001)*** | 13.50 (p < 0.001)*** |
| **Bayley-III: *Motor*** | 9.00 (p = 0.001)*** | 18.00 (p < 0.001)*** | 12.00 (p < 0.001)*** |
| **PARCA-R: *Vocabulary*** | 2.00 (p < 0.001)*** | 16.00 (p < 0.001)*** | 7.00 (p = 0.003)** |
| ***PARCA-R: Sentence complexity*** | −1.45 (p <0.001)*** | 4.00 (p < 0.001)*** | 2.00 (p = 0.010)** |

***Note.*** *Between-group statistics (median differences for variables with significant effects of M-CHAT group) and pairwise comparison p-values are reported for variables showing significant effects of M-CHAT group. IMD index multiple deprivation, PARCA-R parent report of children's abilities-revised. *p < 0.05; **p < 0.010; ***p < 0.001*

### 3.4.2 Differences in brain volume at term-equivalent age between M-CHAT groups

Voxel-wise group comparisons of relative brain volume (correcting for sex, PMA, IMD and neonatal sickness index) showed that critical positive scorers had reduced regional volume in the bilateral deep cerebellar nuclei, middle cerebellar peduncles and midbrain and medulla regions of the brainstem compared to negative scorers (Figure 3.1A). Critical positive scorers also showed volume reductions in an overlapping region in the right cerebellar nuclei compared to the non-critical positive group (Figure 3.1B). Coloured T-statistic maps of regions showing significant differences between critical and negative scorers are depicted in Figure 3.1A and between critical and non-critical scorers in Figure 3.1B, where T-statistic values ranging from 1.70 to 4.70 are denoted by the colour bar. Non-parametric permutation tests with Threshold-Free Cluster Enhancement controlling family-wise error rate were used to identify between-group differences ($p < 0.05$).

There were no significant associations between regional cerebellar volumes and ASC traits at 4–7 years of age in any of the three groups (Table SM 3.2). Furthermore, when investigating the association between M-CHAT total items failed and neonatal whole-brain Jacobians values, we found no significant correlations ($p > 0.05$).

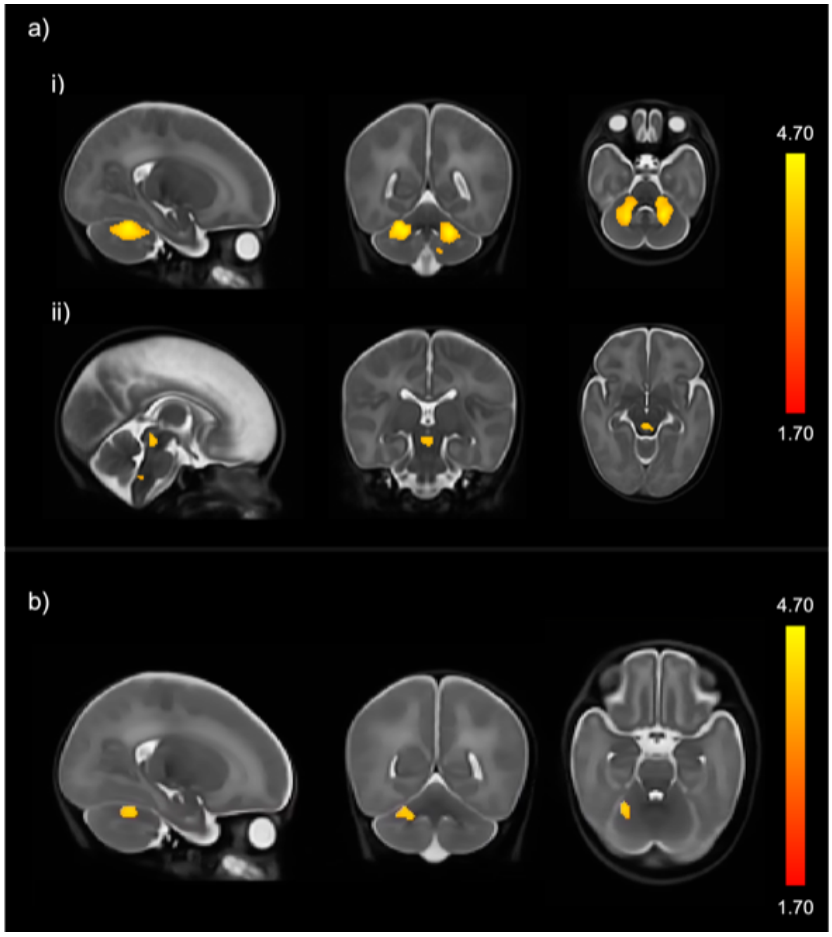



*Figure 3.1. Study-specific brain template overlaid with coloured T-statistics map of brain regions significantly smaller in the M-CHAT critical positive group compared to (a) the M-CHAT negative group and (b) the M-CHAT non-critical positive group.*

### 3.4.3 ASC traits in childhood

A significant effect of group on SRS-2 SCI and RRB was observed (Table 3.5). Pairwise comparisons showed that both M-CHAT (critical and non-critical) positive groups had higher SCI and RRB scores compared to the negative group; however, SCI and RRB scores did not differ between the two positive groups (Table 3.6; Figure 3.2A). These findings did not change after adjusting for sex, IMD, neonatal sickness index and developmental delay.

*Table 3.5. ASC traits at 4–7 years in the M-CHAT screening groups.*

| Variable | Median (Interquartile range) | | | F-statistic; p-value |
|---|---|---|---|---|
| | Negative (N = 130) | Non-critical positive (N = 32) | Critical positive (N = 15) | |
| **SRS-2 SCI** | 45.00 (9.50) | 49.50 (10.00) | 55.00 (17.00) | F = 17.69; p<0.001*** |
| **SRS-2 RRB** | 4.00 (5.00) | 5.50 (7.25) | 11.00 (12.50) | F = 14.02; p< 0.001*** |

***Note.*** *RRB restricted interests and repetitive behaviours, SCI social communication/interaction, SRS-2 social responsiveness scale, second edition. ***p < 0.001.*

*Table 3.6. Post-hoc pairwise differences (between the three M-CHAT screening groups) for ASC traits with a significant effect of M-CHAT group.*

| Variable | Median difference (p-value) | | |
|---|---|---|---|
| | Negative vs non-critical positive | Negative vs critical positive | Non-critical positive vs critical positive |
| **SRS-2 SCI** | −4.50 (p = 0.006) ** | −10.00 (p = 0.001) *** | −5.50 (p = 0.133) |
| **SRS-2 RRB** | −1.50 (p = 0.020) * | −7.00 (p = 0.005) ** | −5.50 (p =0.122) |

***Note.*** *Between-group statistics (median differences for variables with significant effects of M-CHAT group) and pairwise comparison p-values are reported for SRS-2 ASC trait outcomes showing significant effects of M-CHAT group. RRB restricted interests and repetitive behaviours, SCI social communication/interaction, SRS-2 social responsiveness scale, second edition. *p < 0.05; **p < 0.01; ***p < 0.001.*

5 children out 177 (2.8%) had SRS-2 scores exceeding clinical cut-offs for autism (i.e., having SRS-2 total T-scores greater than or equal to 76), where 2 belonged to the non- critical positive group and 3 belonged to the critical positive group. Formal predictive validity analyses were not performed, as sample size analyses estimated a larger sample (N = 480) would be needed to carry them out.

### 3.4.4 Mediating and moderating effects of developmental delay on asc traits

#### *3.4.4.1 Mediation analyses*

Due to the significant differences observed in both SRS-2 SCI and RRB childhood scores between negative scorers and the two positive groups, we tested whether pairwise group

differences were at least partially accounted for by developmental delay. Developmental delay significantly partially mediated differences in SCI when comparing negative to critical (indirect effect 97.5%-CI = 1.69, 8.46; p < 0.001) and non-critical positive groups (indirect effect 97.5%-CI = 0.22, 2.65; p = 0.005; Figure 3.2Bi). Proportion mediated (Prop.med) was 0.18 for M-CHAT negative vs non-critical positive group, and 0.38 for M-CHAT negative vs critical positive group.

Developmental delay also significantly partially mediated group differences in RRB when comparing the negative to the critical positive (indirect effect 97.5%-CI = 1.29, 8.92; p = 0.002; Prop.med = 0.36), but not to the non-critical positive group (indirect effect 97.5%-CI = -0.39, 2.29; p = 0.138; Prop.med = 0.18; Figure 3.2Bii). Mediation analyses for the two positive groups were not conducted, as these did not differ significantly in SCI or RRB scores.

#### *3.4.4.2 Moderation Analyses*

A linear model regressing SCI scores on M-CHAT grouping, developmental delay, and their interaction (M-CHAT × developmental delay), controlling for sex, IMD and neonatal sickness index, found no significant interaction, F(2, 159) = 2.73, p = 0.069; p-permute = 0.074, indicating that the effect of developmental delay on SCI scores was similar in the three M-CHAT groups.

In contrast, a model regressing RRB scores on M-CHAT grouping, developmental delay, and their interaction (M-CHAT × developmental delay), controlling for sex, IMD, and neonatal sickness index, revealed a significant overall interaction, F(2,159) = 6.73, p = 0.002; p-permute = 0.003. Re-coding each group as the reference category showed this was due to a significant interaction when comparing the critical positive group to both negative and non-critical positive groups (Table 3.7). The M-CHAT critical positive group had a stronger (positive) association between developmental delay and RRB scores compared to both negative and non-critical positive groups (Figure 3.2C).

*Table 3.7. M-CHAT x developmental delay interaction on SRS-2 RRB scores.*

| Interaction term | Beta | SE | T-statistic | 97.5%-CI | Permutation p-value |
|---|---|---|---|---|---|
| **M-CHAT (non-critical positive vs negative) × developmental delay** | −2.00 | 1.00 | −2.01 | (−4.25, 0.26) | 0.047 |
| **M-CHAT (critical positive vs negative) × developmental delay** | 2.95 | 1.11 | 2.66 | (0.44, 5.46) | 0.013* |
| **M-CHAT (critical positive vs non-critical positive) × developmental delay** | 4.95 | 1.35 | 3.66 | (1.89, 8.00) | 0.001*** |

***Note.*** *Table summarising, beta, standard error (SE), T-statistic, 97.5% confidence intervals (97.5%-CI) and non-parametric permutation testing p-values for effect of interaction terms between M-CHAT group and developmental delay on RRB scores. RRB restricted interests and repetitive behaviours, SE standard error, SRS-2 social responsiveness scale, second edition. *p < 0.025; ***p < 0.001*

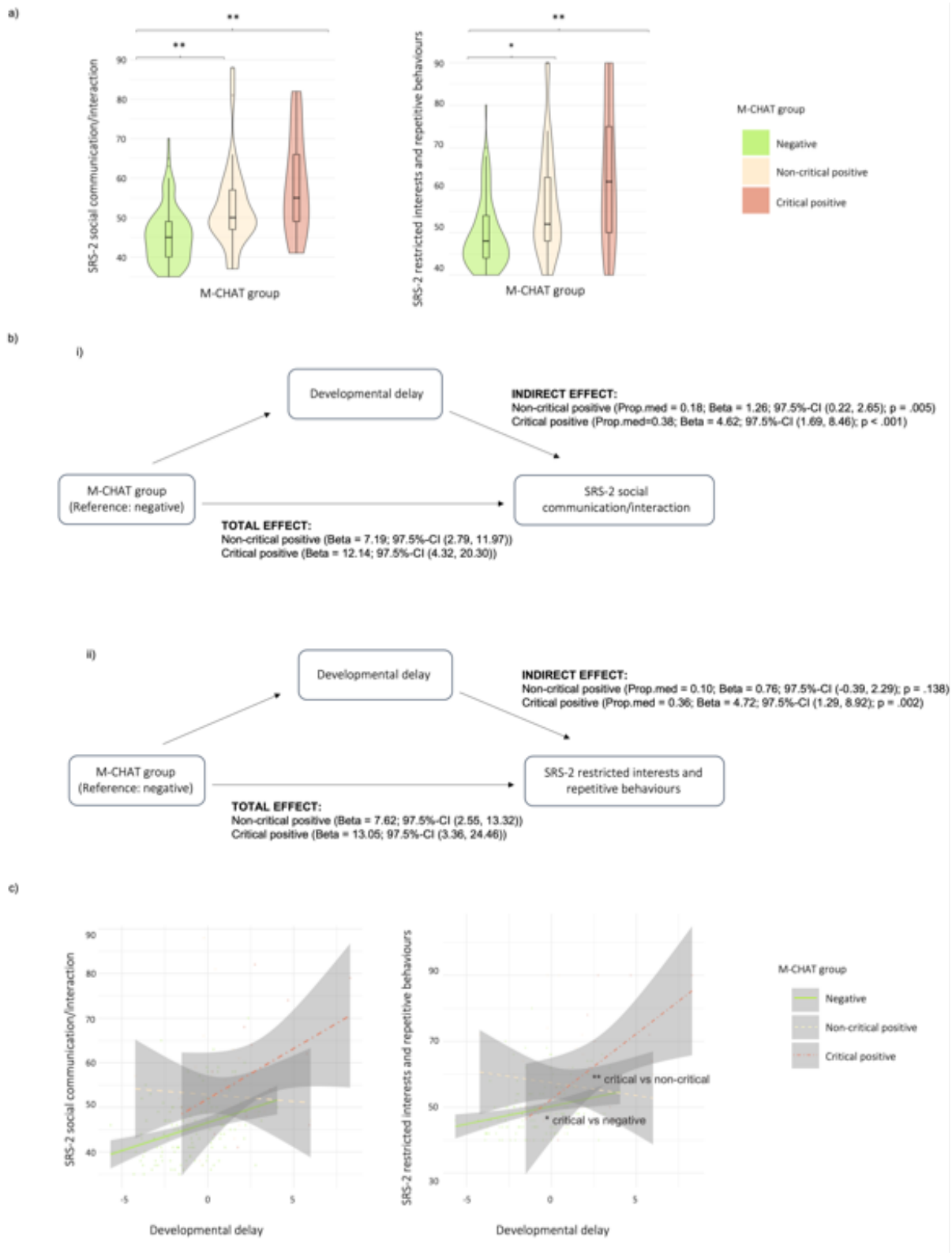



*Figure 3.2. a) SRS-2 SCI/RRB median differences between M-CHAT screening groups, b) the mediating effect of developmental delay on the relationship between M-CHAT and SCI/RRB and c) the moderating effect of the M-CHAT group × developmental delay interaction on SCI.*

## 3.5 Discussion

This study investigated neonatal brain volumes and ASC traits in childhood in VPT children sub-divided into three groups, based on their M-CHAT screening outcomes (negative, non-critical positive and critical positive). Addressing our first aim, we found that the three groups exhibited differences in structural brain volumes at term-equivalent age, indicating distinct early biological phenotypes. The critical positive scorers displayed smaller volumes in cerebellar and brainstem regions compared to negative scorers, and smaller regional cerebellar volumes compared to non-critical positive scorers. Addressing our second aim, we found that while both positive groups showed higher ASC core symptom scores (RRB and SCI) relative to negative scorers, there were no significant differences between the two positive groups. However, the critical positive scorers showed greater developmental delay compared to the other two groups. Taken together our findings suggest that the two M-CHAT positive groups do not differ in the severity of childhood ASC traits and we speculate that they may be following distinct aetiological trajectories leading to similar ASC traits in childhood (i.e., equifinality; Cicchetti & Rogosch, 1996).

The early differences in regional brain volumes found between the positive M-CHAT groups, provide evidence for potentially distinct biological mechanisms underlying later ASC outcomes in a subset of VPT children. The critical positive M-CHAT group showed reduced relative volumes within regions of the right cerebellar nuclei compared to the non-critical positive group, and more widespread reductions in bilateral cerebellar nuclei and brainstem (medulla oblongata and midbrain) volumes compared to the negative group. The cerebellum is known to play a critical role in coordinating motor, sensory and cognitive abilities, which are also impacted in ASC (Wang, Kloth and Badura, 2014). Cerebellar alterations have been associated with ASC symptomatology/traits both in animal and human studies. Cellular cerebellar pathology has been linked to increased ASC-like behaviours in mice (Tsai *et al.*, 2012), smaller white matter volume in the cerebellum has been described in adults with ASC (Toal *et al.*, 2010) and number and density of Purkinje cells has been shown post-mortem to be altered in individuals with ASC (Wegiel *et al.*, 2010, 2014). In VPT samples, cerebellar volume reductions in childhood (Ure *et al.*, 2016) and increased cerebellar haemorrhagic injury in infancy (Limperopoulos *et al.*, 2007) were displayed in those with an ASC diagnosis or those screening positively on the M-CHAT. In both studies, VPT children with ASC diagnoses (Ure *et al.*, 2016) and with cerebellar injury (Limperopoulos *et al.*, 2007) had a high prevalence of developmental delay. Similar to the results of the aforementioned studies, which show cerebellar volume reductions in groups of children with increased developmental delay, we also found that the group exhibiting the most severe developmental delay

(i.e., M-CHAT critical positive group) had smaller cerebellar volumes relative to the non-critical positive and negative groups.

The brainstem, which in this study showed reduced regional volumes in the M-CHAT critical positive relative to the M-CHAT negative group, is an early phylogenetic region of the brain known to be important for primitive functions such as arousal, respiration, and physiological regulation, although there is some evidence of its role in self-regulatory behaviours (Geva & Feldman, 2008; Geva et al., 2014). Of particular relevance to the current findings, Geva et al. (2013) showed that brainstem functioning in VPT infants was associated with social integration abilities assessed using modulation of gaze in response to social stimuli at 4 months. Furthermore, white matter reductions in the brainstem have been observed in adults with ASC compared to controls (Toal *et al.*, 2010) and early histological work investigating brainstem injury, specifically in the motor cranial nerve nuclei, suggest that early alterations to this brain region may contribute to the onset of autism later in life (Rodier *et al.*, 1996, 1997; Rodier, 2002). The cerebellar nuclei and brainstem (medulla oblongata and midbrain) interact with one another to facilitate sensory, motor and regulatory processes (Watson *et al.*, 2013). The olivary complex in the medulla sends fibres to the cerebellar nuclei allowing for integration of motor and sensory information and has been found to be altered post-mortem in individuals with ASC (Wegiel *et al.*, 2013). Interactions between the midbrain and the olivary-cerebellar complex have been discussed in the context of processes relating to "survival networks", which involve behavioural (social, motor and sensory) regulation in response to emotional and environmental stimuli (Watson *et al.*, 2013), which are core processes in ASC symptomatology. In light of these findings, we tentatively speculate that the regional brain alterations we observed in the M-CHAT critical positive compared to the negative group may represent a biological mechanism contributing to the increased RRB and SCI behaviours seen in this group.

Findings showing neonatal regional brain volume reductions as well as increased developmental delay observed in critical compared to non-critical positive scorers, despite the two groups showing similar childhood ASC traits (SCI/RRB), probed us to further investigate developmental delay in relation to ASC traits in the different groups. Results showed that developmental delay had both an explanatory (i.e., mediating) effect, as well as an exacerbating role (i.e., moderating effect) specific to RRB scores, in the critical positive group (but not SCI scores). These results suggest that VPT toddlers meeting the critical positive M-CHAT criteria may, therefore, represent an aetiologically distinct subgroup of children whose developmental difficulties increase their likelihood of developing RRB symptoms. Differences in RRB traits

between preterm and term-born children have been previously explained by differences in IQ (Johnson *et al.*, 2010a), further supporting the notion that developmental delay may contribute to elevated childhood RRB traits. However, it is worth noting that in our study RRB traits were only partially explained by developmental delay, as the higher childhood RRB scores in M-CHAT critical positive scorers compared to negative and non-critical positive scorers were significant after correcting for developmental delay.

The two M-CHAT positive screening groups did not dif- fer in SCI scores, but had elevated SCI scores relative to the negative screening group, which were significant even after correcting for developmental delay. This indicates that developmental delay at least partially contributes to the SCI difficulties seen in both M-CHAT positive groups, which is in line with observations in children with ASC (Hirosawa *et al.*, 2020). However, developmental delay in the current study did not moderate the relationship between M-CHAT group and SCI difficulties, suggesting that the effect of developmental delay on subsequent SCI outcomes was similar in all three groups. These results motivate future studies to investigate which additional biological and/or environmental factors could be driving similar SCI outcomes in the two positive groups, who showed distinct neurodevelopmental profiles early in life.

This study's findings tentatively suggest that the M-CHAT in VPT toddlers represents a useful tool to identify individuals with an increased likelihood of displaying ASC traits in childhood. This is firstly supported by findings showing increased developmental difficulties in both M-CHAT positive groups compared to the negative group, as well as higher median RRB and SCI scores, even after accounting for developmental delay. Secondly, as all children scoring above SRS-2 clinical cut-off thresholds (N = 5, or 2.8% of the sample) belonged to both M-CHAT positive groups, this study suggests that the tool has high sensitivity in VPT cohorts. Finally, although most positive scorers did not exceed the SRS-2 clinical cut-off score for ASC, they did exhibit subthreshold socio-emotional difficulties which are reportedly common amongst VPT children (Johnson and Marlow, 2011).

This study has several limitations, the main being that ASC diagnoses were not systematically evaluated at childhood assessment (4–7 years), although a current follow-up study is now collecting these data at 8–9 years. Moreover, sample size analyses showed we did not have an adequate number of participants to perform formal predictive validity analyses, as the number of children in our sample exceeding SRS-2 clinical cut-off scores were very few. Another limitation of this study is that the results presented are not generalisable to children with major brain lesions, who

are likely to have more severe developmental impairments later in life (Volpe, 2009a), but were not included in the current analyses. Future studies could therefore focus on better understanding the relationship between developmental delay following major brain injury and later ASC behaviours/traits. In addition, other neuroimaging modalities measuring brain functional and structural connectivity were not investigated, and future studies could use a multi-modal approach to provide greater insight into the biological underpinnings associated with the distinct pathways to increased likelihood of developing ASC following VPT birth. Furthermore, while in this paper we consider separate M-CHAT groups, it is plausible that the three groups may lie on a continuum. The non-critical positive scorers' developmental outcomes were in fact intermediate between the two other groups, with the negative scorers showing the best outcomes and the critical positive scorers showing the poorest outcomes.

In summary, our results highlight the distinct early developmental and neurobiological characteristics in M-CHAT critical versus non-critical positive scorers, despite them presenting with similar childhood ASC-symptom profiles. Our results also further highlight the importance of interpreting M-CHAT screenings in combination with other developmental measures when assessing VPT toddlers. Identifying biomarkers and developmental trajectories of later ASC outcomes could guide clinicians and researchers to devise personalised interventions aimed at supporting children's development based on their distinct phenotypic presentations preceding the onset of ASC symptoms.

## 3.6 Supplemental Information

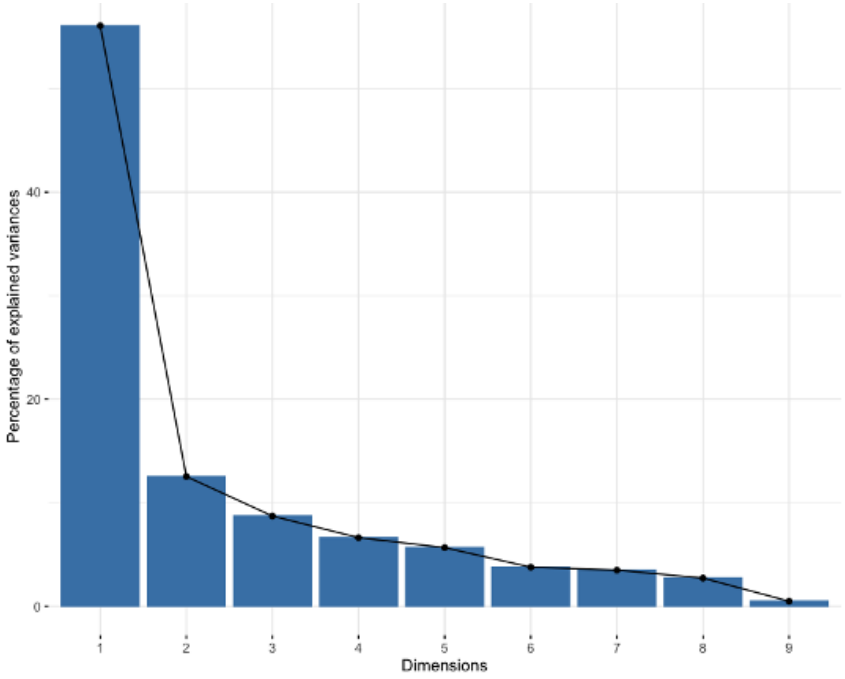



*Figure SM 3.1. Scree plot showing the percentage of variance explained by each principal component.*

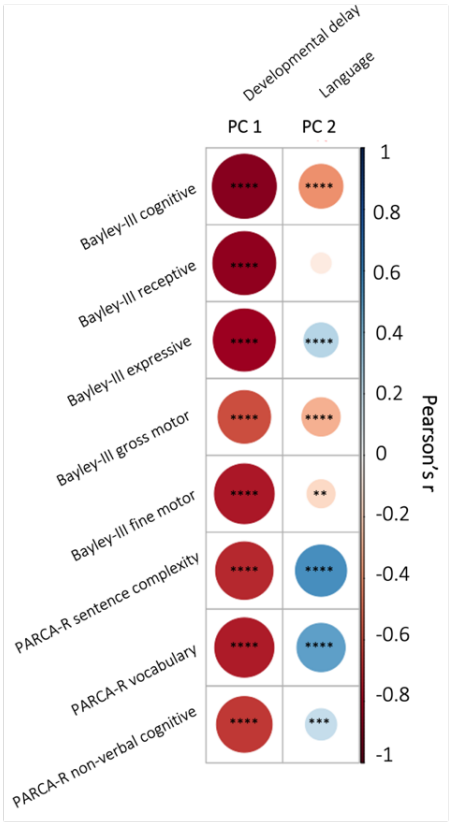



*Figure SM 3.2. Principal component loadings on Bayley-III and PARCA-R subscale scores.*

Figure SM 3.2 shows Pearson correlations between each of the two principal components (PC1 and PC2) with Bayley-III cognitive, receptive, expressive, gross motor and fine motor scores and PARCA-R vocabulary, sentence complexity and non-verbal cognitive scores. Positive correlations are indicated in blue and negative correlations in red. Higher correlation coefficients are visualized by darker shades of red/blue and larger circles. **** = $p < .0001$; *** = $p < .001$; ** = $p < .01$; * = $p < .05$.

***Table SM 3.1. ASC traits at 4-7 years and developmental delay at 2-years old in EPT vs VPT born children.***

| Variable | Median (Interquartile range) | | F-statistic; p-value |
|---|---|---|---|
| | EPT | VPT | |
| **SRS-2 SCI** | 48.00 (15.00) | 46.00 (10.00) | F = 1.92; p = .166 |
| **SRS-2 RRB** | 50.00 (16.00) | 48.00 (10.5) | F = 1.32; p = .252 |
| **Developmental delay** | -0.69 (4.05) | -0.23 (2.78) | F < .001; p = .986 |

***Notes.*** *Abbreviations: EPT = extremely preterm. RRB = Restricted Interests and Repetitive Behaviours. SCI = Social Communication/Interaction. SRS-2: Social Responsiveness Scale, Second Edition. VPT = very preterm.*

***Post-hoc analyses investigating associations between neonatal brain volumes and ASC traits in childhood.*** In order to investigate whether the identified anatomical findings aligned with later high risk for autism, we investigated associations between values reflecting neonatal brain volume (mean Jacobian values extracted from the cerebellum cluster showing a significant difference in the critical group compared to both non-critical and negative subgroups) and later ASC traits (SCI and RRB at 4-7 years; Table SM 3.2). We also conducted the same analyses in the three M-CHAT subgroups (i.e., negative, non-critical positive and critical positive scorers) separately. We found no significant associations between regional cerebellar volume and ASC traits in the entire cohort or in any of the groups (Table SM 3.2).

*Table SM 3.2. Spearman R coefficients and p-values for correlations between mean neonatal cerebellar log-Jacobian values and childhood SCI/RRB scores.*

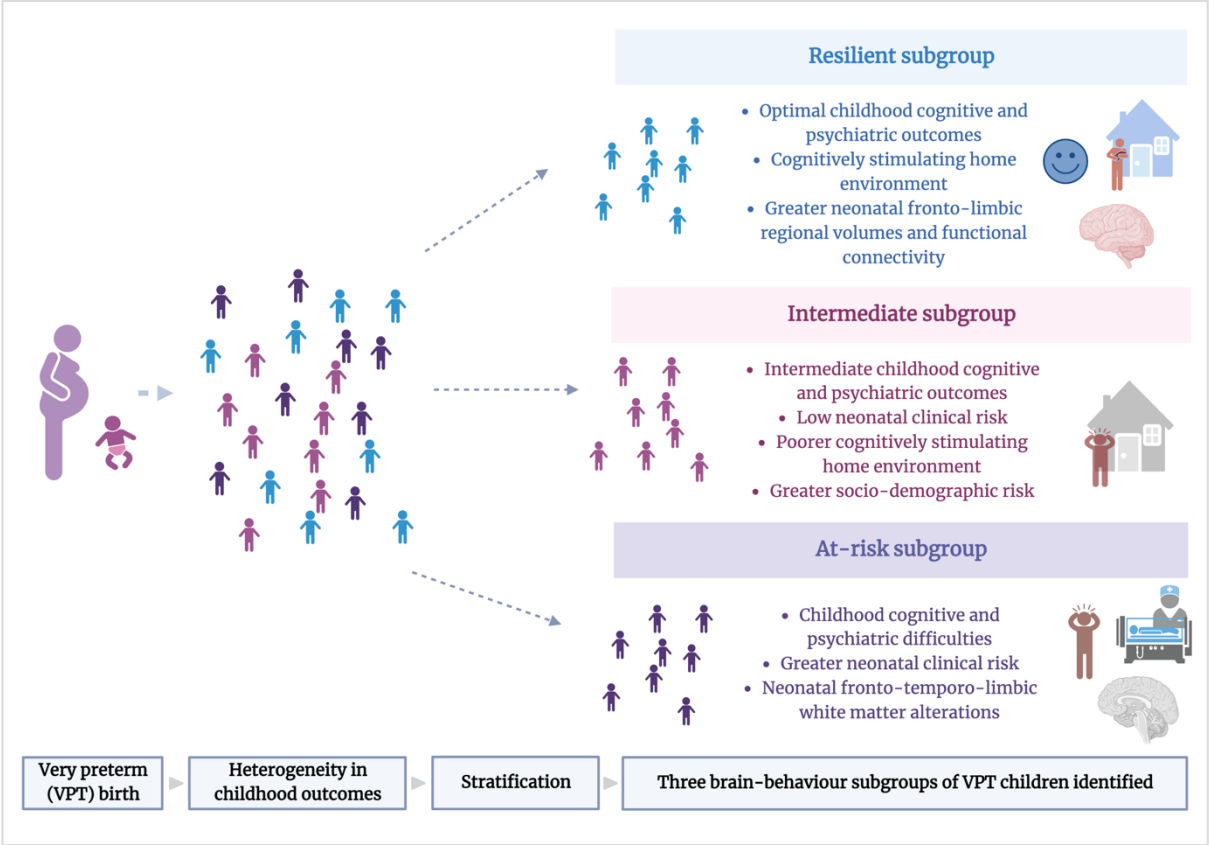
| M-CHAT group | Correlation with SCI (Spearman R; p-value) | Correlation with RRB (Spearman R; p-value) |
|---|---|---|
| **All participants** | R = - .12; p = .115 | R = -.14; p = .070 |
| **Negative scorers** | R = - .06; p = .539 | R = -.12; p = .175 |
| **Non-critical positive scorers** | R = -.01, p = .973 | R = 0.09; p = .630 |
| **Critical positive scorers** | R = -.21; p = .449 | R = -.20; p = .483 |

# CHAPTER 4 - Study #2: Using data-driven integrative consensus clustering to parse longitudinal brain-behavioural heterogeneity in VPT children



***Reference:** Hadaya, L., Dimitrakopoulou, K., Vanes, L. D., Kanel, D., Fenn-Moltu, S., Gale-Grant, O., Counsell, S. J., Edwards, A. D., Saqi, M., Batalle, D., & Nosarti, C. (2023). Parsing brain-behavior heterogeneity in very preterm born children using integrated similarity networks. Translational Psychiatry, 13(1), Article 1. https://doi.org/10.1038/s41398-023-02401-w*

**Study #2 graphical abstract** *(created with BioRender.com):*




## 4.1 Abstract

Very preterm birth (VPT; ≤32 weeks' gestation) is associated with altered brain development and cognitive and behavioral difficulties across the lifespan. However, heterogeneity in outcomes among individuals born VPT makes it challenging to identify those most vulnerable to neurodevelopmental sequelae. Here, we aimed to stratify VPT children into distinct behavioral subgroups and explore between-subgroup differences in neonatal brain structure and function. 198 VPT children (98 females) previously enrolled in the Evaluation of Preterm Imaging Study (EudraCT 2009-011602-42) underwent Magnetic Resonance Imaging at term-equivalent age and neuropsychological assessments at 4–7 years. Using an integrative clustering approach, we combined neonatal socio-demographic, clinical factors and childhood socio-emotional and executive function outcomes, to identify distinct subgroups of children based on their similarity profiles in a multidimensional space. We characterized resultant subgroups using domain-specific outcomes (temperament, psychopathology, IQ and cognitively stimulating home environment) and explored between- subgroup differences in neonatal brain volumes (voxel-wise Tensor-Based-Morphometry), functional connectivity (voxel-wise degree centrality) and structural connectivity

(Tract-Based-Spatial-Statistics). Results showed two- and three-cluster data-driven solutions. The two-cluster solution comprised a 'resilient' subgroup (lower psychopathology and higher IQ, executive function and socio-emotional scores) and an 'at-risk' subgroup (poorer behavioral and cognitive outcomes). No neuroimaging differences between the resilient and at-risk subgroups were found. The three-cluster solution showed an additional third 'intermediate' subgroup, displaying behavioral and cognitive outcomes intermediate between the resilient and at-risk subgroups. The resilient subgroup had the most cognitively stimulating home environment and the at-risk subgroup showed the highest neonatal clinical risk, while the intermediate subgroup showed the lowest clinical, but the highest socio-demographic risk. Compared to the intermediate subgroup, the resilient subgroup displayed larger neonatal insular and orbitofrontal volumes and stronger orbitofrontal functional connectivity, while the at-risk group showed widespread white matter microstructural alterations. These findings suggest that risk stratification following VPT birth is feasible and could be used translationally to guide personalized interventions aimed at promoting children's resilience.

## 4.2 Introduction

Very preterm birth (VPT; ≤32 weeks' gestation) is associated with an increased likelihood of developing cognitive and behavioral difficulties across the lifespan (Johnson *et al.*, 2010b; Treyvaud *et al.*, 2013; Mathewson *et al.*, 2017; Twilhaar *et al.*, 2018; van Houdt *et al.*, 2019). Efforts to conceptualize these difficulties have proposed a "preterm behavioral phenotype", characterized by problems in emotional and social processing, and inattention (Johnson and Marlow, 2011). However, while some VPT children display a behavioral profile reflecting a preterm phenotype, others follow typical developmental trajectories (Burnett *et al.*, 2019; Lean *et al.*, 2020; van Houdt *et al.*, 2020). Such behavioral heterogeneity following VPT birth presents a challenge for building risk prediction models (Crilly, Haneuse and Litt, 2021), as multiple causes may lead to the same outcome and as a single mechanism may lead to multiple outcomes (Cicchetti and Rogosch, 1996).

Several endogenous and exogenous factors contribute to a child's behavioral development and a complex interplay between environmental, clinical, and neurobiological features could result in co-occurring neurodevelopmental, cognitive and behavioral difficulties following VPT birth (Volpe, 2009a). These factors are often non-independent and their combination (e.g., neurobiological and socio-demographic variables) may result in improved prediction of functional outcomes (Wickremasinghe *et al.*, 2012). For instance, both socio-demographic deprivation and increased neonatal clinical risk have been associated with neurodevelopmental as well as behavioral

difficulties in VPT children. These encompass executive and socio-emotional functions (Levine *et al.*, 2015; Brouwer *et al.*, 2017; Benavente-Fernández *et al.*, 2019), which could be considered as gateway mechanisms that shape behavioral outcomes, as they are subserved by brain networks relating to both bottom-up stimulus processing and top-down behavioral control (Luo *et al.*, 2014). Impairments in these domains have in fact been associated with later academic and mental health difficulties (Treyvaud *et al.*, 2013; Woodward *et al.*, 2017).

Previous studies have attempted to stratify outcome heterogeneity in preterm children using clustering and latent-class analyses (Burnett *et al.*, 2019; Lean *et al.*, 2020; van Houdt *et al.*, 2020)·(Poehlmann *et al.*, 2015; Johnson *et al.*, 2018). These studies typically used cognitive and behavioral measures as input features, and then compared subgroups in terms of specific clinical and environmental risk factors that were not used in the stratification analyses (i.e., out- of-model). Some found differences in neonatal clinical profiles between subgroups of preterm children (Johnson *et al.*, 2018) and others showed that familial characteristics, such as parental education, maternal distress, and cognitively stimulating parenting, differentiated resilient subgroups from those exhibiting behavioral difficulties (Lean *et al.*, 2020; van Houdt *et al.*, 2020). Here, instead, we chose to include input measures of known risk factors (i.e., clinical and environmental variables) alongside in-model cognitive and behavioral measures, in order to delineate the complex interplay between different risk factors and behavioral outcome measures; thus increasing the likelihood of discovering nuanced subtypes of preterm children who exhibit similar behavioral outcomes, but with possibly different underlying correlates (i.e., equifinality) (Cicchetti and Rogosch, 1996).

A growing body of research, investigating specific factors associated with later behavioral outcomes, is studying the early neural signatures that may shape an individual's neurodevelopmental trajectory. Alterations in brain volumes (Rogers *et al.*, 2012; Cismaru *et al.*, 2016), white matter microstructure (Kanel *et al.*, 2021; Lee *et al.*, 2021), and functional connectivity (Ramphal et al., 2020; Rogers et al., 2017) at birth in regions and networks subserving social, emotional and attentional processes, have been associated with later behavioral difficulties in VPT samples. Differences between latent subgroups of VPT children and infants have been previously studied in relation to qualitative measures of brain abnormalities and/or high grade brain injury based on neonatal Magnetic Resonance Imaging (MRI), as well as quantitative differences in brain tissue volumes (Ross *et al.*, 2016; Lean *et al.*, 2020; Bogičević *et al.*, 2021). However, it remains to be explored whether distinct multidimensional subgroups of VPT children could also be characterized by localized differences in early brain development using advanced quantitative measures of brain structure and function, such as log-Jacobians, tract based spatial statistics and

degree centrality, which have previously been used in neonatal samples (Ball *et al.*, 2010, 2017; Fenn-Moltu *et al.*, 2022). Conducting analyses at the whole-brain and voxel-wise level, allows for an enhanced spatial localization of potential structural and functional between-subgroup differences, thus extending previous research (Ross *et al.*, 2016; Lean *et al.*, 2020; Bogičević *et al.*, 2021).

The main aim of this study was to parse brain-behavior heterogeneity in VPT children, by identifying subgroups with similar environmental, clinical and behavioral profiles and examining between-subgroup differences in structural and functional brain features at term-equivalent age. Firstly, we implemented an integrative clustering approach (Similarity Network Fusion; SNF) (Wang *et al.*, 2014) to stratify VPT children into distinct subgroups based on three data types: (i) neonatal clinical and socio- demographic variables, (ii) childhood socio-emotional outcomes and (iii) executive function measures. The advantage of this approach is that it integrates sample-similarity networks built from each distinct data type and constructs a final integrated network, which contains common and complementary information from the different data types. This is then used to stratify the sample into distinct subgroups using clustering (Wang *et al.*, 2014). We also investigated whether resultant subgroups differed in outcomes that were not used in stratification analyses (i.e., out-of-model variables); in order to provide external validation (Stefanik *et al.*, 2018; Hong *et al.*, 2021; Jacobs *et al.*, 2021). Finally, we explored between-subgroup differences in regional brain volume and structural and functional connectivity at term-equivalent age. We hypothesized that there would be distinct subgroups of VPT children characterized by unique neonatal neural signatures.

## 4.3 Methods

### 4.3.1 Study design

**Participants.** Five hundred and eleven infants born VPT were recruited from 14 neonatal units in London in 2010–2013 and entered the Evaluation of Preterm Imaging Study (ePrime; EudraCT 2009-011602-42) (Edwards *et al.*, 2018). Infants with congenital malformation, prior MRI, metallic implants, whose parents did not speak English or were subject to child protection proceedings were not eligible for participation in the study.

Participants underwent multi-modal MRI at 38–53 weeks post-menstrual age (PMA) on a 3-Tesla MR imaging system (Philips Medical Systems, Best, The Netherlands) located on the neonatal intensive care unit at Queen Charlotte's and Chelsea Hospital, London, using an 8-

channel phased array head coil. For data acquisition and imaging parameters see Supplemental Information. Infants whose parents chose sedation for the procedure (87%) received oral chloral hydrate (25–50 mg/kg).

In total, 251 participants (including 29 sets of multiple pregnancy children) were followed-up between the age of 4 and 7 years at the Center for the Developing Brain, St Thomas' Hospital, London. This was a convenience sample corresponding to 82% of 306 participants who were past their fourth birthday by the study end date, September 1st 2019, and had consented to be contacted for future research. Invitations for follow- up were sent in chronological order of birth.

Ethical approval was granted by the Hammersmith and Queen Charlotte's Research Ethic Committee (09/H0707/98) and the Stanmore Research Ethics Committee (14/LO/0677). Informed consent was obtained from all participants.

**Clinical and socio-demographic data.** We used Principal Component Analysis (PCA) to select neonatal clinical variables of interest from a set of 28 available variables. These were: gestational age (GA) at birth, number of days on mechanical ventilation, number of days on continuous positive airway pressure (CPAP) and number of days on parenteral nutrition (TPN), which loaded onto a single component explaining 72% of the variance in the data. This component was labeled 'neonatal sickness index'. Please refer to our previous work (Kanel *et al.*, 2021) and Supplemental Information for more details on the PCA analysis.

Socio-demographic risk was evaluated using a postcode derived measure of deprivation in England, the Index of Multiple Deprivation 2010 (IMD; http://tools.npeu.ox.ac.uk/imd/), whereby higher IMD scores reflect greater deprivation. The IMD combines neighborhood-specific information about seven domains of deprivation: income, employment, education/skills/training, health, crime, housing and living environment. The IMD was collected at the term-equivalent age. Continuous IMD scores were used in the integrative-clustering and evaluation of subgroup profile analyses. IMD quintiles are provided when reporting sample characteristics (Table 4.1) for ease of interpretability.

**Childhood assessment.** Intelligence quotient (IQ) was evaluated using the Wechsler Preschool and Primary Scale for Intelligence (WPPSI-IV) (D Wechsler, 2012) and executive function using the preschool version of the parent-rated Behavior Rating Inventory of Executive Function (BRIEF-P) (Sherman and Brooks, 2010). Socio-emotional processing was evaluated using the Empathy Questionnaire (EmQue) (Rieffe, Ketelaar and Wiefferink, 2010) and the Social

Responsiveness Scale, Second Edition (SRS-2) (Constantino and Gruber, 2012). Psychopathology was assessed using the Strengths and Difficulties Questionnaire (SDQ) (Goodman, 2001), temperament using the Child Behavioral Questionnaire - Very Short Form (CBQ) (Rothbart *et al.*, 2001) and cognitively stimulating home environment using an adapted version of the Cognitive Stimulating Parenting Scale (CSPS) (Wolke, Jaekel, *et al.*, 2013).

**Exclusions.** Twenty-seven participants were excluded due to incomplete childhood outcome data, 17 due to major brain lesions (periventricular leukomalacia, parenchymal hemorrhagic infarction, or other ischemic or hemorrhagic lesions), detected on neonatal T2-weighted MRI images at term by an experienced perinatal neuroradiologist, and 5 participants due to missing T2-weighted MRI images, hence the inability to evaluate the presence of major lesions (Figure SM 4.1).

*Table 4.1. Socio-demographic and clinical participant data.*

| | | Integrative clustering sample *n* = 198 | Diffusion MRI TBSS analysis sample *n* = 166 | Structural MRI log-Jacobian analysis sample *n* = 165 | rs- fMRI degree centrality analysis sample *n* = 129 |
|---|---|---|---|---|---|
| **Corrected age at assessment, years** | Median | 4.63 | 4.60 | 4.59 | 5.63 |
| | Range | 4.18–7.17 | 4.18–7.17 | 4.18–7.17 | 4.18–7.17 |
| **PMA, weeks** | Median | 42.57 | 42.43 | 42.57 | 42.43 |
| | Range | 38.29–52.86 | 38.29–44.86 | 38.29–44.86 | 38.29–44.86 |
| **Sex, male:female** | n= | 100:98 | 88:78 | 86:79 | 68:61 |
| **Self-reported maternal ethnicity** | n (%) | | | | |
| *Asian* | | 50 (25.3%) | 44 (26.5%) | 43 (26.1%) | 34 (26.4%) |
| *Black/African/Caribbean/ Black British* | | 30 (15.2%) | 23 (13.9%) | 25 (15.2%) | 15 (11.6%) |
| *Mixed/Multiple ethnic groups* | | 3 (1.5%) | 3 (1.8%) | 3 (1.8%) | 3 (2.33%) |
| *White* | | 112 (56.6%) | 93 (56.0%) | 91 (55.2%) | 75 (58.1%) |
| **Self-reported paternal ethnicity** | *n* (%) | | | | |
| *Asian* | | 34 (17.2%) | 29 (17.5%) | 27 (16.4%) | 23 (17.8%) |
| *Black/African/Caribbean/ Black British* | | 23 (11.6%) | 19 (11.5%) | 20 (12.1%) | 14 (10.9%) |
| *Mixed/Multiple ethnic groups* | | 2 (1.0%) | 1 (0.6%) | 1 (0.6%) | 0 (0.0%) |
| *White* | | 95 (48.0%) | 80 (48.2%) | 79 (47.9%) | 63 (48.8%) |
| **Neonatal IMD, quintiles** | *n* (%) | | | | |
| *1 (least deprived)* | | 49 (24.8%) | 40 (24.1%) | 38 (23.0%) | 30 (23.3%) |
| *2* | | 37 (18.7%) | 31 (18.7%) | 32 (19.4%) | 25 (19.4%) |
| *3* | | 44 (22.2%) | 39 (0.6%) | 38 (23.0%) | 30 (23.3%) |
| *4* | | 48 (24.2%) | 39 (23.5%) | 38 (23.0%) | 31 (24.0%) |
| *5 (most deprived)* | | 20 (10.1%) | 17 (10.2%) | 19 (11.5%) | 13 (10.1%) |
| **GA at birth, weeks** | Median | 30.14 | 30.29 | 30.14 | 30.14 |
| | Range | 23.86–32.86 | 24.00–32.86 | 24.00–32.86 | 24.00–32.86 |
| **Neonatal clinical risk** | *n*= | | | | |
| *Days TPN, ratio 0:1:2* | | 68:98:32 | 62:78:26 | 63:77:25 | 49:61:19 |
| *Days CPAP, ratio 0:1:2* | | 33:125:40 | 30:107:29 | 31:103:31 | 23:82:24 |
| *Days ventilation, ratio 0:1:2* | | 101:74:23 | 92:59:15 | 92:58:15 | 72:46:11 |

***Note:*** *Table describing sample socio-demographic and clinical characteristics for the integrative clustering and MRI analyses. Neonatal clinical risk categories (0, 1 and 2) respectively correspond to zero days, more than zero days, but less than the top quintile, and within the top quintile. IMD quintiles 1–5 respectively correspond to the least deprived quintile (1) to the most deprived quintile (5). Ethnicity was grouped according to the Office of National Statistics classifications 2016 (see Supplementary Information). CPAP continuous positive airway pressure, GA gestational age at birth, IMD Index of Multiple Deprivation, PMA post-menstrual age at scan, rs-fMRI resting-state functional MRI, TBSS Tract Based Spatial Statistics, TPN total parenteral nutrition.*

### 4.3.2 Data integration and clustering

Analyses were conducted in R (version 3.6.1). Using SNF, three data types were integrated: (Type 1) neonatal socio-demographic and clinical variables: IMD at birth, GA, days on ventilation, days on TPN and days on CPAP. (Type 2) childhood socio-emotional outcomes: EmQue subscale raw scores - emotion contagion, attention to others' emotions, prosocial behaviors and SRS-2 total raw score. (Type 3) childhood executive function: BRIEF-P raw subscale scores - inhibit, shift, emotional control, working memory and plan/organize.

Prior to integration, participants with in-model outlier values greater than 3 times the interquartile range were excluded. A total of 198 children were included in the SNF analyses. Zero-inflated neonatal clinical risk variables (days ventilation, days TPN and days CPAP) were converted into ordinal categorical variables with three levels: (Level 0: zero days; Level 1: greater than zero and not within the top quintile; Level 2: within the top quintile). For the mixed data type (numeric and categorical data; data type 1), Gower's standardization based on the range was applied using the daisy function from cluster R package (Maechler *et al.*, 2021) and for numeric only matrices (data types 2 and 3), variables were standardized to have a mean value of 0 and a standard deviation of 1 using the *standardNormalization* function from SNFtool R package (Wang *et al.*, 2018).

An adaptation of the ExecuteSNF.CC function (Xu *et al.*, 2017) was used for the data integration and clustering steps. Dissimilarity Gower distance (for the mixed data type) and Euclidean distance (for numeric data types) matrices were calculated and used to create similarity matrices using the SNFtool R package's *affinityMatrix* function (Wang *et al.*, 2018). This was followed by an integration of the similarity matrices using SNFtool's *SNF* algorithm resulting in a 'fused similarity matrix' (Wang *et al.*, 2018). The integrative clustering process can be summarized into two steps:

Step 1: SNF method has two main hyperparameters, K and alpha. K (i.e., neighborhood size) indicates the number of neighbors of a node to consider when the similarity networks are being generated and alpha is an edge weighting parameter determining the weight of edges between nodes in the networks. We tried 30 combinations of K and alpha hyperparameters {K = 10, 15, 20, 25, 30; alpha = 0.3, 0.4, 0.5, 0.6, 0.7, 0.8}, similar to the approach followed in (Markello *et al.*, 2021). The K-alpha hyperparameter values were chosen based on the ranges recommended in the SNFtool R package, 10–30 for K and 0.3–0.8 for alpha (Wang *et al.*, 2014, 2018). Consensus clustering, using *ConsensusClusterPlus* function (Wilkerson and Hayes, 2010), was then applied to each fused similarity matrix, corresponding to a K-alpha combination, where spectral clustering

was run 1000 times with 80% of the population randomly subsampled for each clustering run and a single consensus clustering result obtained from hierarchical clustering. Step 2: Next, out of the 30 clustering results produced in step 1, the one with the highest average silhouette width score was retained. Steps 1 and 2 were repeated 1000 times in a bootstrap approach, after selecting and pre-processing the three data matrices of 80% of the sample set. The 1000 resultant retained clustering outputs were then fed to the diceR R package's *consensus_combine* function (Chiu and Talhouk, 2018) which implements hierarchical clustering on the consensus matrix and generates the final consensus clustering. Figure 4.1 summarizes the data-integration and clustering steps and the code used can be accessed here: https://github.com/lailahadaya/preterm- ExecuteSNF.CC. Further details can also be found in Supplemental Information.

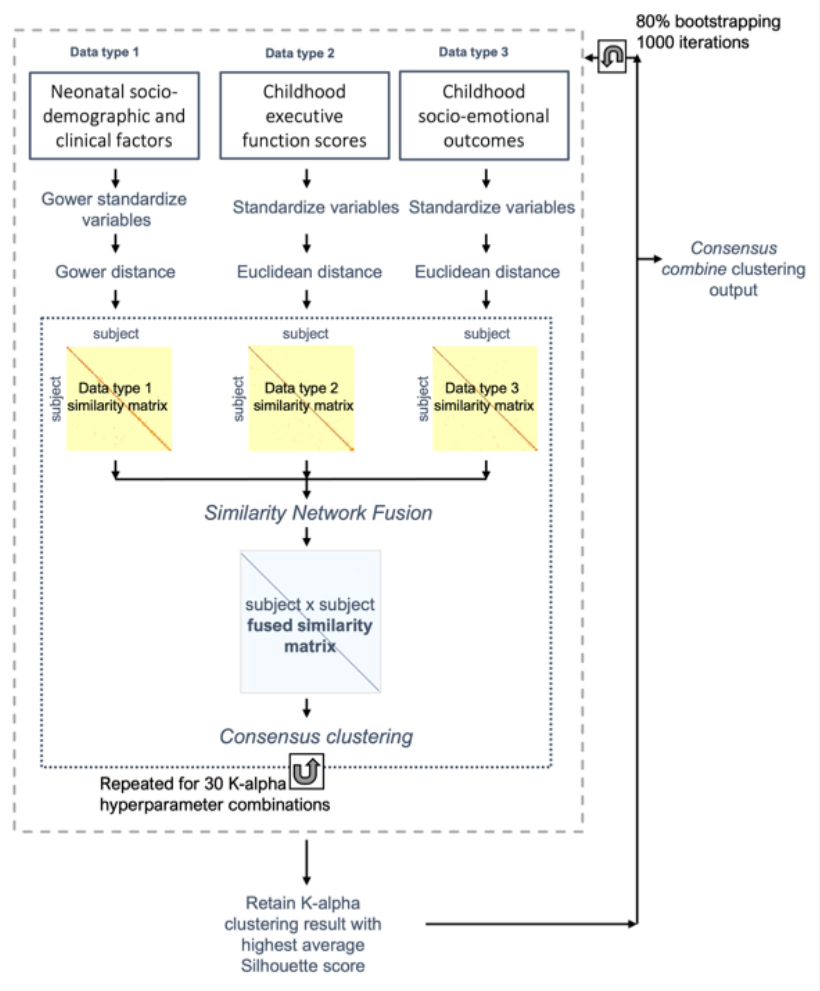



***Figure 4.1. Data integration and clustering pipeline.***

*Figure summarizing the data pre-processing (variable normalization), data integration and clustering pipeline executed in order to obtain the final consensus cluster assignment.*

Before implementing steps 1 and 2, it was essential to determine the number of clusters. For this, we used the SNFtool R package's *estimateNumberOfClustersGivenGraph* function (Wang *et al.*, 2018) to calculate Eigengap and Rotation Cost heuristics for each K-alpha combination (Figure SM 4.2). This process suggested C = 2, C = 3 and C = 4 as the optimal number of clusters. Consensus matrices and silhouette scores were generated and compared for these three potential clustering solutions (Figure SM 4.2). Resultant subgroups from C = 2 and C = 3 were chosen to be evaluated for phenotypic differences, as their silhouette scores and consensus matrices gave better values in comparison to those of C = 4 (Figure SM 4.2). More details on the estimation of cluster numbers can be found in Supplemental Information. An alluvial plot was used to illustrate the transition of subject subgroup classification between the two-cluster and three-cluster solutions (Figure SM 4.3).

### 4.3.3 Evaluation of subgroup profiles

Resultant subgroups were characterized based on in-model and out-of- model variables. For the out-of-model features, subgroups were compared in terms of psychiatric symptoms (SDQ internalizing, externalizing problems and total scores), temperament (CBQ negative affectivity, surgency and effortful control scores), cognitive abilities (WPPSI full-scale IQ), and cognitive stimulation at home (CSPS score). Details on selection of in-model and out-of-model variables can be found in Supplemental Information and Figures SM 4.4 and SM 4.5.

For numeric measures, between-subgroup differences were assessed using non-parametric one-way tests: Mann-Whitney when C = 2 or Kruskal Wallis when C = 3 (Dag, Dolgun and Konar, 2018). Shapiro-Wilk test was used to assess normality. For categorical variables, Chi-squared test was used to evaluate differences in proportions of individuals in each group when count per cell was >5 and Fischer's Exact test was used otherwise. To compare differences between the ordinal neonatal clinical variables with 3 categories (Levels 0, 1 and 2) and the non-ordinal subgroups from C = 2 and C = 3, the Extended Cochran-Armitage Test was used. We also ran supplementary post-hoc analyses investigating subgroup differences in clinical variables not included as in-model variables (please see Supplementary Information for more details).

Results with $p < 0.05$ were considered to be statistically significant. To correct for multiple comparisons the False Discovery Rate method was used. The same statistical analyses were repeated using general linear models correcting for potential confounders (age and sex) and 5000 permutation test iterations (França, Ge and Batalle, 2022). Effect sizes for non-normally distributed variables were measured using Wilcoxon Glass Rank Biserial Correlation (gr) for

measuring differences between two groups and Epsilon Squared for three groups. For continuous normally distributed variables, Cohen's F was used and Cramer's V for categorical variables.

### 4.3.4 Exploring neonatal brain differences between subgroups

Tract Based Spatial Statistics (TBSS) was used to assess white matter microstructure at the voxel-level using fractional anisotropy (FA) and mean diffusivity (MD) maps (Smith *et al.*, 2006). FA approximates the directional profile of water diffusion in each voxel and MD measures the average movement of water molecules within a voxel. Higher FA and lower MD values reflect more optimal white matter myelination and microstructure. For diffusion MRI (d-MRI) image pre-processing and TBSS protocol details please refer to Supplemental Information.

Structural MRI (s-MRI) log-Jacobian determinant maps were calculated to quantify regional brain volumes (greater log-Jacobian values reflect larger relative structural volumes), using Tensor Based Morphometry, following methods described in our previous work (Lautarescu *et al.*, 2021; Vanes *et al.*, 2021) and in Supplemental Information.

Resting-state functional MRI (rs-fMRI) data were pre-processed as in our previous work (Ball *et al.*, 2016); for more details see Supplemental Information. Functional connectivity was evaluated using a measure of weighted degree centrality at the voxel-level (i.e., the sum of the correlations between the time-series of each voxel and all other voxels within a gray matter mask of the brain) (Holiga *et al.*, 2019; Fenn-Moltu *et al.*, 2022). Edges with a correlation coefficient below a threshold of 0.2 were excluded and the degree centrality values for each voxel in the gray matter mask were z-scored and used in subsequent between-subgroup analyses. Whilst other functional network measures are available (i.e., participation coefficient and within module-z (Power *et al.*, 2013), we opted to study degree centrality as we recently showed this to be disrupted in preterm born neonates (Fenn-Moltu *et al.*, 2022). Furthermore, degree centrality is a good voxel-wise summary measure of connectivity strength, which is reliable and correlates with relevant phenotypes, such as age and sex (Zuo *et al.*, 2012). It has been used to study typical cognitive function (Heuvel and Sporns, 2013) and has recently been shown to be a reproducible metric to detect atypical functional connectivity patterns in neurodevelopmental disorders (Holiga *et al.*, 2019).

The number of children included in the different modality-specific MRI analyses slightly differed due data availability: TBSS (n = 166), log-Jacobian determinant maps (n = 165) and degree

centrality (n = 129); please see Table SM 4.1. Exclusions for specific MRI analyses are depicted in Figure SM 4.1.

Between-subgroup differences were investigated in the whole-brain at the voxel-level in terms of: log-Jacobian determinants, TBSS metrics (FA and MD) and degree centrality. FMRIB Software Library (FSL) (Jenkinson *et al.*, 2012) randomise function was used to implement non-parametric permutation methods for statistical inference. This method was used to model each contrast of interest for each voxel, i.e., a general linear model (GLM) correcting for PMA at scan and sex. rs-fMRI models also included motion estimates (standardized DVARS) as a covariate. Family Wise Error (FWE) rate with Threshold-Free Cluster Enhancement (TFCE) was applied to correct for multiple comparisons over the multiple voxels, while enhancing "cluster-like" structures of voxels without defining them as binary components (Smith and Nichols, 2009). Statistics were calculated using random permutation tests with 10000 permutations. Given the exploratory nature of our analysis, we did not correct for multiple contrasts tested (i.e., log-Jacobians, TBSS FA and MD, degree centrality). We show results significant at $p < 0.05$ FWE-corrected per contrast. Mean values from clusters of modality-specific voxels showing significant between-subgroup differences were extracted to calculate Cohen's F effect sizes.

### 4.3.5 Sensitivity analyses

There were 29 sets of children born from multiple pregnancy events in our sample. In order to account for multiple pregnancy confounding, we conducted additional sensitivity analyses including only one child from each set of multiple pregnancy siblings.

## 4.4 Results

### 4.4.1 Participant characteristics

Participants' socio-demographic and clinical characteristics are shown in Table 4.1. Compared to participants who completed the follow-up assessment (n = 251; median GA = 29.24 weeks; median IMD at birth=19.48), individuals who were not assessed (n = 259; median GA = 29.27 weeks; median IMD at birth = 21.40) did not differ in GA (gr = 0.01; p = 0.807), but had greater neonatal socio-demographic deprivation (gr = 0.11; p = 0.028). Compared to the initial baseline cohort (n = 511; median GA = 30.00 weeks; median IMD at birth=18.19), participants who were studied here (n = 198) had slightly older GA (median GA = 30.14 weeks; gr = −0.13; p = 0.009) and relative socio-demographic advantage (median IMD score at birth=15.58, gr = 0.11; p = 0.027).

### 4.4.2 Two-cluster solution subgroup profiles

When stratifying the sample into two clusters and comparing them in terms of in-model variables, subgroup 1 (termed here the 'resilient' subgroup) showed significantly better socio-communication (i.e., lower SRS-2 scores) and executive function abilities (i.e., lower BRIEF-P emotion control, inhibit, shift, working memory and plan/organize scores), lower emotion contagion (EmQue) scores, and higher prosocial actions scores (EmQue) during childhood, than subgroup 2 (termed here the 'at-risk' subgroup); all ps $< 0.05$, after FDR correction. The resilient subgroup had lower neonatal clinical risk compared to the at-risk subgroup, with a greater proportion of children receiving no neonatal mechanical ventilation and a smaller proportion of children receiving prolonged neonatal CPAP (both ps $< 0.05$, after FDR correction). Subgroups did not differ in terms of days on TPN in the neonatal period ($p > 0.05$).

Differences in out-of-model variables included lower psychopathology scores (SDQ internalizing and externalizing problems) and negative affectivity scores (CBQ) as well as higher effortful control (CBQ), IQ and cognitive stimulation at home (CSPS) during childhood in the resilient compared to the at-risk subgroup; all ps $< 0.05$, after FDR correction (Figure 4.2; Table SM 4.2).

The two subgroups showed no significant differences in log-Jacobian determinant values, degree centrality or white matter microstructural characteristics (all ps $> 0.05$). Resultant subgroups also did not show differences in sex, age at assessment or PMA at scan (Figure 4.2; Table SM 4.2).

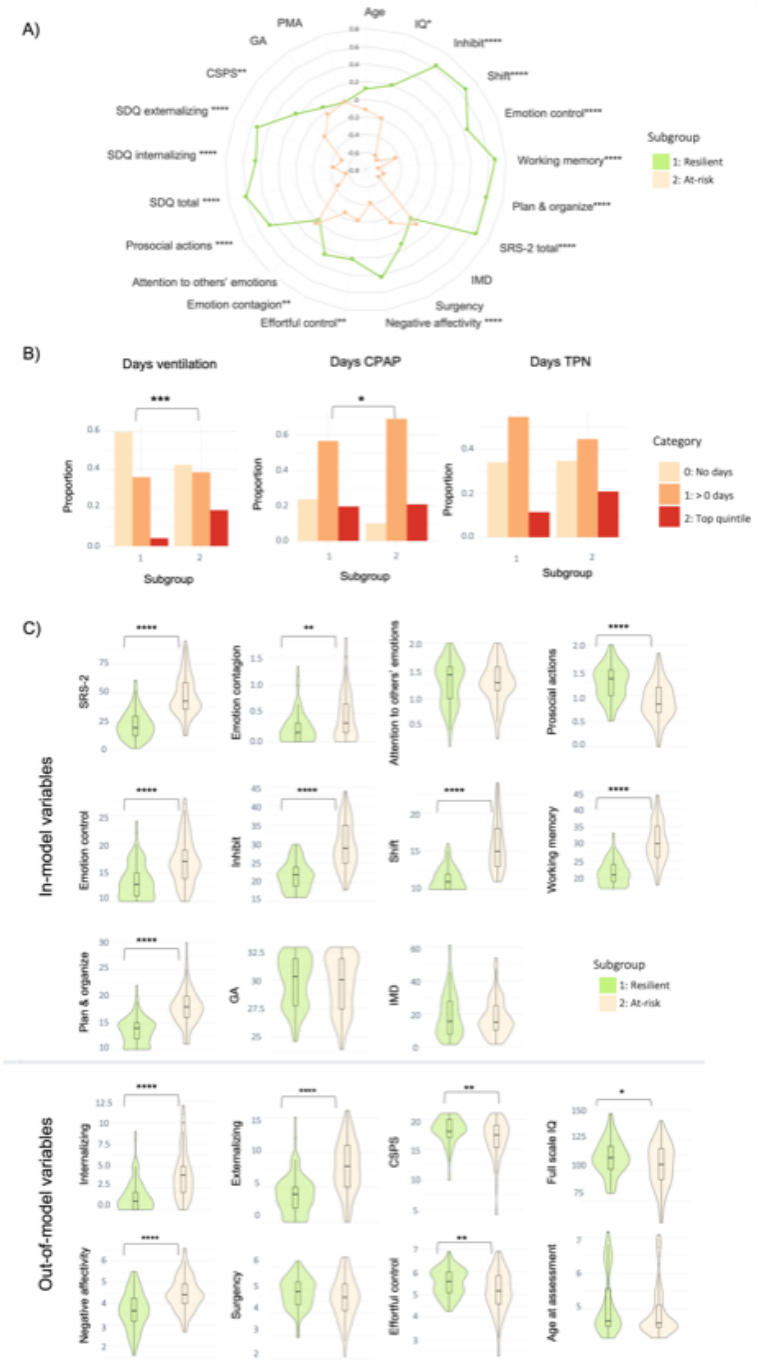



***Figure 4.2. Two-cluster solution subgroup profiles.***

*A) Radar plot showing the two-cluster solution subgroup profiles using z-scores for subgroup 1 (i.e., resilient subgroup; green) and subgroup 2 (i.e., at-risk subgroup; beige). For visual illustrative purposes, scales which usually indicate poorer outcomes have been reversed so that larger z-scores on behavioral variables indicate better outcomes. B) Bar plots for clinical risk variables (days on TPN, days on mechanical ventilation and days on CPAP, left to right, respectively) for each of the two subgroups. Plots represent the proportion of children belonging to each clinical risk category within a subgroup, where category 0 represents the lowest clinical risk (light beige; no days of clinical intervention), category 1 represents medium clinical risk (orange; more than 0 days of intervention but less than the top quintile), and category 2 represents the highest clinical risk (red; within the top quintile). C) Violin plots showing differences between the subgroups in terms of in-model and out-of-model variables. Significant differences are marked with bars between the subgroups. *=$p < 0.05$; **=$p < 0.01$; ***=$p < 0.001$, ****=$p < 0.0001$.*

### 4.4.3 Three-cluster solution subgroup profiles

To increase subtyping resolution and explore latent heterogeneity not captured by a two-subgroup partitioning, the sample was further stratified into 3 subgroups. Two of the three resulting clusters largely reflected profiles similar to those from C = 2.

The first was a 'resilient' subgroup (subgroup 1) with favorable childhood socio-communicative (SRS-2), empathy (EmQue) and executive function (BRIEF-P) outcomes in terms of in-model variables; low childhood psychopathology (SDQ internalizing and externalizing problems) and negative affectivity scores (CBQ) and high effortful control scores (CBQ), IQ and cognitive stimulation at home (CPSP) in terms of out-of-model variables. The second was an 'at-risk' subgroup (subgroup 2), with the poorest outcomes in terms of in-model variables (childhood socio-communication (SRS-2), empathy and executive function (BRIEF- P) scores), as well as out-of-model childhood psychopathology (SDQ), effortful control (CBQ) and negative affectivity measures (CBQ), combined with the highest neonatal clinical risk (Figure 4.3; Table SM 4.3).

A third subgroup (labelled 'intermediate') emerged, which had poorer in-model and out-of-model childhood cognitive and behavioral scores when compared to the resilient subgroup, but better scores when compared to those of the at-risk subgroup. The intermediate subgroup also had the lowest neonatal clinical risk compared to both resilient and at-risk subgroups (Figure 4.3; Table SM 4.3). The transition of subject classifications from the two- to the three-cluster solution is illustrated in an alluvial plot (Figure SM 4.3).

In terms of environmental factors, the resilient subgroup had higher levels of childhood cognitive stimulation at home (CSPS) in comparison to both at-risk and intermediate subgroups, while the intermediate subgroup had higher neonatal socio-demographic risk (IMD) in comparison to both at-risk and resilient subgroups. All ps $< 0.05$ after FDR correction. The three subgroups did not differ in sex, age at assessment or PMA at scan.

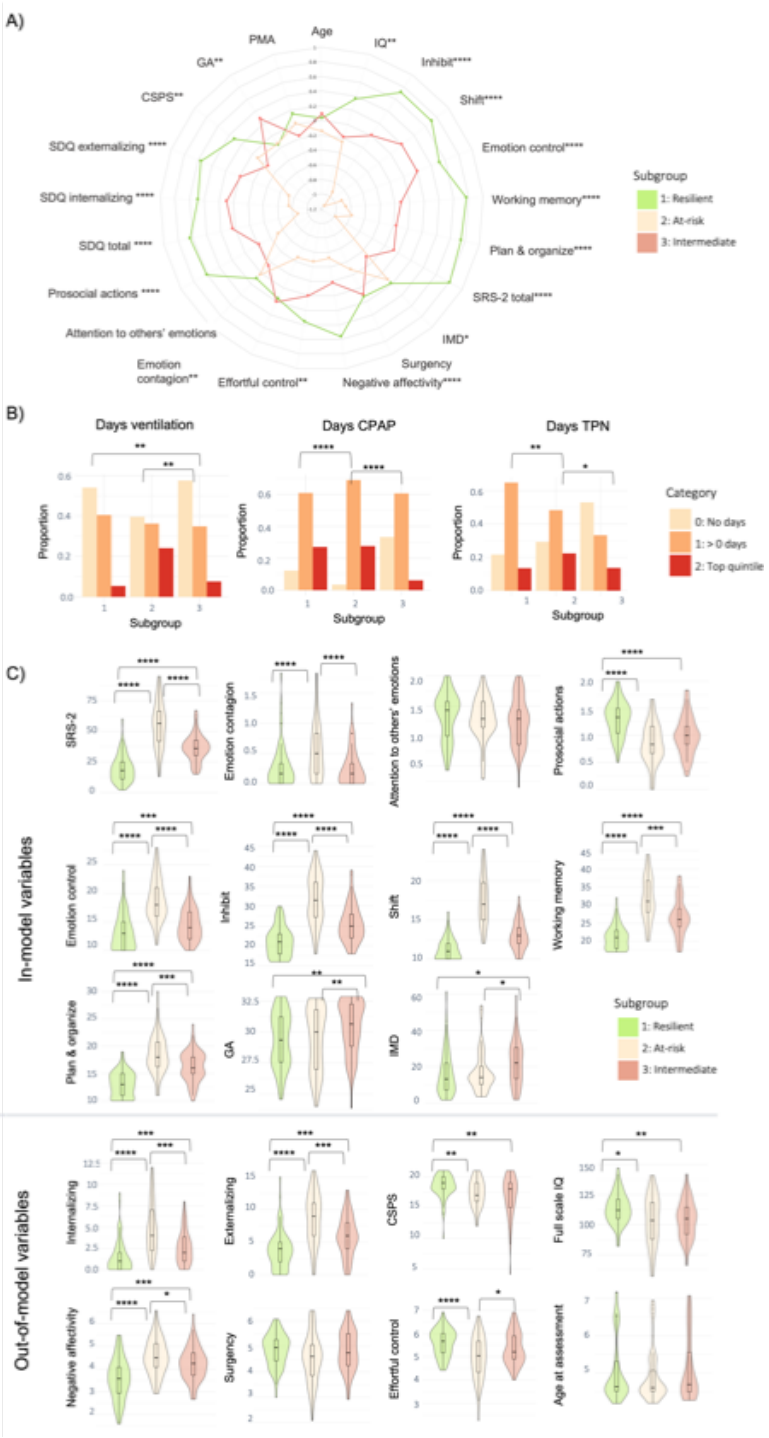



***Figure 4.3. Three-cluster solution subgroup profiles.***

*A) Radar plot showing the three-cluster solution subgroup profiles using z-scores. For visual illustration purposes, scales which usually indicate poorer outcomes have been reversed, so that larger z-scores on behavioral variables indicate better outcomes. Subgroup 1 (resilient) is marked in green, subgroup 2 (at-risk) in beige and subgroup 3 (intermediate outcomes but lowest clinical risk) in pink. B) Bar plots for clinical risk variables (days on TPN, days on mechanical ventilation and days on CPAP, left to right, respectively) for each of the three subgroups. Plots represent the proportion of children belonging to each clinical risk category within a subgroup, where category 0 represents the lowest clinical risk (light beige; no days of clinical intervention), category 1 represents medium clinical risk (orange; more than 0 days of intervention but less than the top quintile), and category 2 represents the highest clinical risk (red; within the top quintile). C) Violin plots showing differences in in-model and out-of-model measures at the group-wise level. Significant differences are marked with bars between the subgroups. *=p < 0.05; **=p < 0.01; ***=p < 0.001, ****=p < 0.0001.*

In terms of brain imaging markers at term, the resilient subgroup displayed larger relative volumes (i.e., greater log-Jacobian determinant values) in the left insula and bilateral orbitofrontal cortices (Figure 4.4A; Table SM 4.4) and higher degree centrality in an overlapping region in the left orbitofrontal cortex (Figure 4.4B; Table SM 4.5) compared to the intermediate subgroup. The intermediate subgroup, compared to the at-risk subgroup, showed increased FA in several areas of the white matter skeleton, including the fornix, corpus callosum, corticospinal tract, inferior longitudinal, inferior fronto-occipital and uncinate fasciculi (Figure 4.4Ci; Table SM 4.4), as well as lower MD in the fornix and body of the corpus callosum (Figure 4.4Cii; Table SM 4.4). The resilient and at-risk subgroups did not differ in any brain measures ($p > 0.05$).

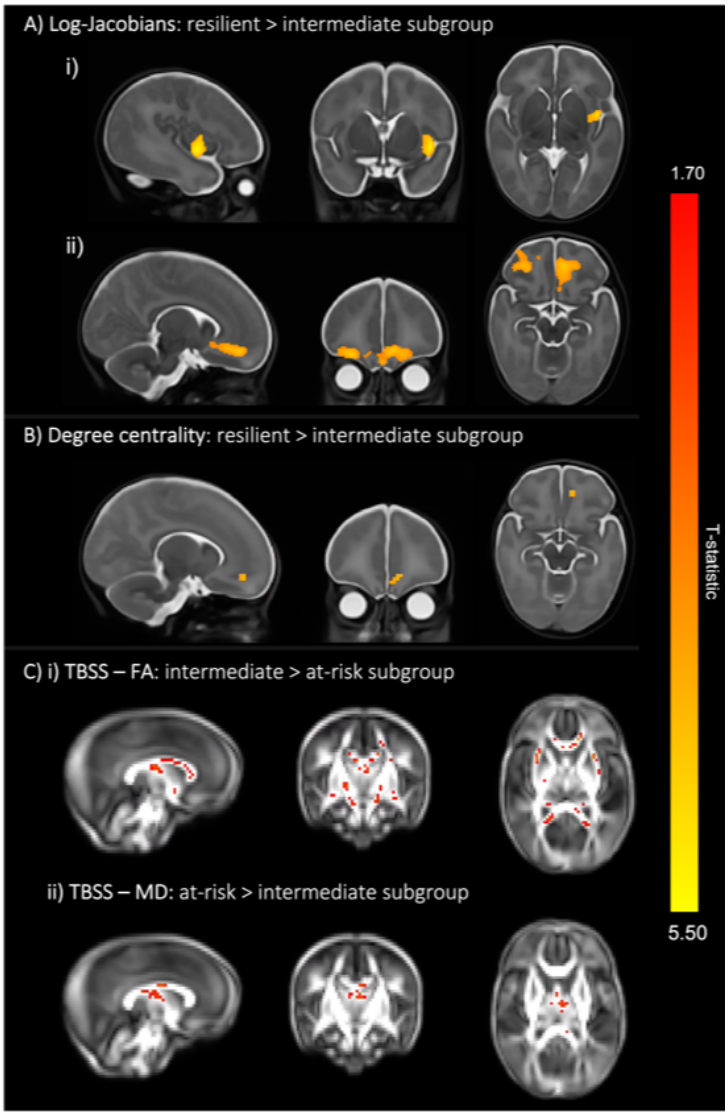



***Figure 4.4. Three-cluster solution brain differences at term-equivalent age.***

*A) Colored voxels indicate regions with significantly larger log- Jacobian determinant values in the resilient subgroup (subgroup 1) compared to the intermediate subgroup (subgroup 3) in i) left insula and the ii) bilateral orbitofrontal cortices ($p < 0.05$). GLM included sex and PMA at scan as covariates and TFCE and FWE corrections were applied. B) Voxels showing significantly larger degree centrality values in the resilient subgroup (subgroup) 1 compared to the intermediate subgroup (subgroup 3) are seen in an overlapping left orbitofrontal region at $p < 0.05$. GLM included sex, PMA at scan and motion (standardized DVARS) as covariates; TFCE and FWE were applied. C) Colored voxels represent white matter regions showing i) significantly higher FA values in the intermediate subgroup compared to the at-risk subgroup and ii) significantly higher MD values in the at-risk subgroup compared to the intermediate subgroup ($p < 0.05$). T-statistic values are represented in the color bar, where red colored voxels indicate smaller T-statistic values and yellow voxels indicate higher T-statistic values ranging between 1.70 and 5.50.*

### 4.4.4 Sensitivity analyses

Sensitivity analyses including only one sibling, selected at random from each multiple pregnancy set, revealed similar results (Table SM 4.5; Table SM 4.6; Figure SM 4.6; Figure SM 4.7), although the difference in neonatal functional connectivity between the resilient and intermediate groups was no longer significant ($p = 0.08$). In addition, the resilient subgroup displayed larger neonatal relative volume of the right insula compared to the intermediate subgroup. For more details, please refer to Supplemental Information.

## 4.5 Discussion

Using an integrative clustering approach, we identified subgroups of VPT children with distinct neurodevelopmental profiles. We described a two-cluster solution, showing a resilient subgroup with comparably favorable childhood behavioral and cognitive outcomes and increased cognitive stimulation at home, and a second, at-risk subgroup, with poorer childhood behavioral and cognitive outcomes and high neonatal clinical risk. We also described a three-cluster solution, showing two subgroups largely characterized by the profiles observed in the two-cluster solution, as well as a newly emerging third intermediate subgroup, with a childhood behavioral and cognitive profile intermediate between the resilient and the at-risk subgroups. Nuanced differences in socio-demographic, neonatal clinical and early brain measures appeared upon comparing subgroups from the three-cluster solution. Notably, the resilient subgroup displayed larger fronto-limbic brain regions and increased functional connectivity at term compared to the intermediate subgroup. The at-risk subgroup showed widespread white matter microstructural alterations in fronto-temporo-limbic tracts compared to the intermediate subgroup. Furthermore, the resilient subgroup had a more cognitively stimulating childhood home environment compared to the at-risk and intermediate subgroups, while the intermediate subgroup had the lowest clinical risk. Together, these findings highlight the potential value of neonatal structural and functional brain measures as useful biomarkers of later childhood outcomes in distinct VPT subgroups, as well as the importance of a supportive home environment to foster child development.

In the at-risk subgroup from the two-cluster solution, poorer childhood socio-emotional, executive function, IQ, mental health and temperament outcomes may have been driven by a combination of both higher clinical risk at birth and a less stimulating childhood home environment, when compared to the resilient subgroup. Previous studies in VPT children have shown cognitively stimulating parenting to be positively associated with improved socio-emotional processing and cognitive outcomes at 2 years of age (Treyvaud *et al.*, 2012) and reduced

psychopathology and executive function difficulties at 4–7 (Vanes *et al.*, 2021). A cognitively stimulating home environment also differentiated between psychiatric profiles at 5 (Lean *et al.*, 2020). Moreover, increased neonatal clinical risk in the at-risk subgroup is consistent with previous findings, showing that perinatal medical complications following VPT birth may lead to increased behavioral and developmental problems (Neubauer, Voss and Kattner, 2008; Levine *et al.*, 2015; Brouwer *et al.*, 2017). The resilient and at-risk subgroups, however, did not differ in any of the neonatal brain measures investigated, suggesting that there may be additional non-measured variables underlying different childhood outcomes that need further investigation, such as alterations in pro-inflammatory immunomarkers (Pariante, 2016; J. Anderson *et al.*, 2021) and/or microbiome assembly (Clapp *et al.*, 2017; Mukhopadhyay *et al.*, 2022), which are reportedly associated with increased behavioral difficulties.

To further parse heterogeneity in VPT children, we also explored a three-cluster solution. These analyses showed that two subgroups mostly reflected the profiles seen in the two-cluster solution: 1) a resilient subgroup with high levels of childhood cognitive stimulation at home and 2) an at-risk subgroup with high levels of neonatal clinical risk. A third subgroup with intermediate childhood behavioral and cognitive profiles also emerged, in which childhood psychopathology, temperament and cognitive outcomes were poorer than those observed in the resilient subgroup, but more favorable than those observed in the at-risk subgroup. Intriguingly, the intermediate subgroup exhibited the lowest neonatal clinical risk compared to the other two subgroups, with a greater proportion of infants receiving no neonatal mechanical ventilation, CPAP or TPN and with higher median GA at birth. However, the intermediate subgroup also had higher environmental risk, namely reduced childhood cognitively stimulating home environment, compared to the resilient subgroup, and higher neonatal socio-demographic deprivation, compared to both the at-risk and resilient subgroups. These findings suggest that developmental outcomes may not be understood by exploring a single causal pathway and are best studied in a multidimensional space; for example, clinical risk, which has been linearly correlated with developmental outcomes in previous studies (Neubauer, Voss and Kattner, 2008; Levine *et al.*, 2015), ought to be investigated together with other factors that may influence development, i.e., environmental risk.

The at-risk compared to the intermediate subgroup showed widespread alterations in white matter microstructure (lower FA and higher MD) in the fornix, corpus callosum, corticospinal tract, inferior longitudinal, inferior fronto-occipital and uncinate fasciculi. The at-risk subgroup had also the highest neonatal clinical risk, hence the observed white matter changes are likely to

be associated with preterm-related neonatal complications (Volpe, 2009a; Back, 2017; Lee, 2017). White matter alterations in fronto-temporo-limbic tracts, including those observed here, have been previously associated with poorer cognitive outcomes (Mooshagian, 2008; Thompson *et al.*, 2012; van Duinkerken *et al.*, 2012; Huang *et al.*, 2015; Vollmer *et al.*, 2017; Chen *et al.*, 2020). They have also been implicated in emotion processing (Crespi *et al.*, 2014; Herbet, Zemmoura and Duffau, 2018; Wier *et al.*, 2019) and psychiatric disorders, including depression and schizophrenia (Lamar *et al.*, 2013; Von Der Heide *et al.*, 2013; Hung *et al.*, 2017). The intermediate subgroup, conversely, had the lowest neonatal clinical risk, and higher FA/lower MD values in fronto-temporo-limbic tracts compared to the at-risk subgroup. These findings led us to speculate that having relative fewer neonatal clinical complications, and hence fewer preterm-related white matter alterations, may contribute to these children's more favorable cognitive, socio-emotional and behavioral outcomes, compared to the at-risk subgroup.

Children in the resilient subgroup exhibited higher prosocial behavior and empathy, as well as fewer childhood externalizing and internalizing symptoms and executive function difficulties, compared to the intermediate and at-risk subgroups. They also showed lower childhood negative affectivity scores, referring to the expression of dysregulated negative emotions and increased sensitivity in response to surrounding stimuli (Rothbart *et al.*, 2003; Rothbart, 2004). While the resilient group showed no significant brain differences compared to the at-risk subgroup, we speculate that the combination of two protective factors, an enriching home environment and lower neonatal clinical risk, may have contributed to attenuating the expression of the behavioral and cognitive difficulties associated with VPT birth. These findings also support the idea of multifinality, whereby individuals with no overt brain differences at term may display distinct behavioral outcomes later in childhood. Compared to the intermediate subgroup, however, the resilient subgroup displayed larger relative volumes in the left insular and bilateral orbitofrontal cortices and increased functional connectivity in an overlapping left orbitofrontal region at term, years before the behavioral and cognitive childhood outcomes were assessed. These findings could be interpreted in terms of a more advanced maturation of the fronto-limbic network in the resilient subgroup, as orbitofrontal functional connectivity and insular cortical microstructure and morphology have been positively associated with GA at birth and PMA at scan (Toulmin *et al.*, 2015; Mouka *et al.*, 2019; Dimitrova, Pietsch, *et al.*, 2021). However, as several other brain areas are undergoing rapid neurodevelopmental changes at the time our participants underwent MRI, including somatosensory, occipital, temporal, parietal and other areas of frontal cortex (Dimitrova, Pietsch, *et al.*, 2021), we speculate the orbitofrontal cortex and the insula may be preferentially discriminating between the intermediate and the resilient subgroup, in the context of the brain-

wide analysis approaches employed here, because they play critical functional roles in the cognitive and behavioral outcomes we studied. The orbitofrontal cortex is involved in the top-down regulation of goal-oriented executive functions and socio- emotional processing, reward-guided learning and decision making (Rempel-Clower, 2007; Rushworth *et al.*, 2011; McTeague *et al.*, 2020); the insula is important for regulating internal processes, including emotional responses to external stimuli (Uddin *et al.*, 2017). Structural alterations in the orbitofrontal cortex and insula, which are structurally connected (Uddin *et al.*, 2011), have been associated with emotion dysregulation (Petrovic *et al.*, 2016) and with higher externalizing behaviors (Tanzer *et al.*, 2021).

The orbitofrontal cortex is sensitive to environmental stimuli, such as early life stress (Li *et al.*, 2013; Hodel, 2018). Individuals with a history of physical abuse (Hanson *et al.*, 2010) and VPT infants exposed to painful procedures (Ranger *et al.*, 2013) both show reduced orbitofrontal volumes in childhood. Furthermore, alterations in orbitofrontal connectivity and gyrification have been associated with social processing impairments in VPT children (Fischi-Gómez *et al.*, 2015) and with executive function difficulties in extremely preterm (EPT; < 28 weeks' gestation) adolescents (Ganella *et al.*, 2014), respectively. Smaller insular volumes have been associated with worse emotion regulation skills (Giuliani *et al.*, 2011) and weaker insular functional connectivity with decreased empathic responses (Fan *et al.*, 2011). In the late preterm period, the insula becomes a key hub region (Gao *et al.*, 2011) and a major source of transient bursting events that support brain maturation (Arichi *et al.*, 2017). A more mature fronto-limbic network may have therefore supported a favorable development of emotion regulation capacity, cognition, and behavior (Liu *et al.*, 2010; Guo *et al.*, 2015), resulting in the resilient subgroup exhibiting lower externalizing and internalizing symptoms, increased empathy, emotion regulation abilities and executive function skills in childhood.

This study demonstrates that it is possible to parse heterogeneity in VPT children in a meaningful way. We show that protective brain maturational patterns in the neonatal period may contribute to a more resilient behavioral profile in childhood. This is encouraging, as the preterm brain is susceptible to neuroplastic changes in response to behavioral and environmental interventions, both early in life and later in childhood (DeMaster *et al.*, 2019). For example, neuroplastic changes have been observed following 'supportive-touch' (i.e., skin-to-skin contact or breastfeeding) (Maitre *et al.*, 2017), maternal sensitivity training (Milgrom *et al.*, 2010), visual stimulus cues of the mother's face (Gee *et al.*, 2014), parental praise (Matsudaira *et al.*, 2016) or music interventions in the neonatal intensive care unit (Lordier *et al.*, 2019). Such methods could, therefore, be used in the future to strengthen fronto-limbic circuitry to boost children's resilience.

Furthermore, our findings suggest that enriching environments may promote resilience towards more favorable behavioral outcomes. This could be done by increasing parental awareness about the importance of cognitive stimulation at home. Our findings also show that the subgroup of children with the highest neonatal clinical risk exhibit the poorest outcomes, highlighting the need to develop and implement targeted interventions for the most clinically vulnerable VPT children.

It is worth noting that the median outcome scores (IQ, BRIEF-P, SRS-2 and SDQ) for our three subgroups were within normative ranges and below clinical thresholds, even for the at-risk subgroup. Subthreshold psychiatric symptoms have been reported in other at-risk subgroups of VPT children (Lean *et al.*, 2020; van Houdt *et al.*, 2020), and have also been associated with an increased risk of developing psychiatric disorders later in life (Briggs-Gowan *et al.*, 2003). In this context, subthreshold psychiatric symptoms may represent transdiagnostic traits that would remain undetected, and therefore untreated, if considered in a purely clinically diagnostic context, highlighting the importance of addressing psychopathology dimensionally (Insel *et al.*, 2010; Cuthbert, 2014).

Strengths of this study include a fairly large sample size and a rich longitudinal dataset with clinical data from birth, neonatal multi-modal MRI at term and behavioral follow-up in early childhood. However, a limitation of this study is that the VPT participants included in our analyses (n = 198) had a relative socio-demographic advantage and older gestational age at birth than the initial baseline cohort (n = 511), which may limit the generalizability of our findings to a portion of the socio-demographic and gestational age spectrum. In addition, the lack of a full-term group and the exclusion of children with major brain lesions in the integrative-clustering analyses may have also limited the variability in our data, and in turn contributed to the failure to identify a more impaired subgroup here. Future studies must take extra caution when interpreting such results and make increased efforts to recruit more diverse participant samples.

Additional limitations to consider include the use of parental reports for most child behavioral measures, except IQ, which could lead to common method variance bias (Podsakoff *et al.*, 2003) and result in underreporting of psychopathology (Mathai, Anderson and Bourne, 2004). The lack of information on familial cognitive outcomes and psychiatric history, which are heritable traits (McGrath *et al.*, 2014), prevents us from estimating trait heritability. Moreover, the small to moderate effect sizes reported for neonatal brain differences between subgroups may limit their immediate clinical meaningfulness or translatability into clinical practice. However, the fact that these brain differences only emerged after subdividing the sample into more refined and

homogenous phenotypic subgroups (C = 3 vs C = 2), highlights the benefit of using advanced clustering approaches such as SNF. We speculate that these effects may be diluted in the two-cluster solution due to the presence of individuals within both (at-risk and resilient) subgroups having profiles that are more similar to an intermediate subgroup profile (please see Figure SM 4.3).

Sensitivity analyses including one sibling only from each twin/triplet set mostly replicated the main findings, showing similar early brain patterns as well as cognitive, neonatal clinical, social, and childhood behavioral profiles for both two- and three-cluster solutions, suggesting that the effects seen here are not biased by the presence of multiple pregnancy siblings in the main analyses. While the functional connectivity results were no longer significant in the sensitivity analyses, we speculate this may be due to a loss in power associated with the reduced sample size.

In summary, using an integrative clustering approach, we were able to stratify VPT children into distinct multidimensional subgroups. A subgroup of VPT children at risk of experiencing behavioral and cognitive difficulties was characterized by high neonatal clinical complications and white matter microstructural alterations at term, whereas a resilient subgroup, with comparably favorable childhood behavioral outcomes, was characterized by increased childhood cognitive stimulation at home and larger and functionally more connected fronto-limbic brain regions at term. These results highlight a potential application of precision psychiatry, to enable meaningful inferences to be made at the individual level. Patterns of fronto-limbic brain maturation may be used as image-based biomarkers of outcomes in VPT children, while promoting enriching environments may foster more optimal behavioral outcomes. Risk stratification following VPT birth could, therefore, guide personalized behavioral interventions aimed at supporting healthy development in vulnerable children.

## 4.6 Supplemental Information

### 4.6.1 MRI acquisition parameters

Images were acquired with the following parameters: T2-weighted fast-spin echo sequences: TR = 8670 ms, TE = 160 ms, flip angle = 90°, slice thickness = 1 mm, FOV = 220 x 220 mm, matrix = 256 x 256, voxel size = 0.86x0.86x1 mm. Single-shot echo-planar diffusion MRI was acquired in the transverse plane in 32 non-collinear directions: (TR = 8000 ms; TE = 49 ms, slice thickness = 2 mm, FOV = 224x224 mm, matrix = 128x128, voxel size = 1.75x1.75x2 mm, b-value = 750 s/mm$^2$, SENSE factor 2). T2-gradient echo-planar imaging at rest: TR= 1.5 s, ET = 45 ms, flip angle = 90°, 256 volumes, slice thickness= 3.25 mm, in-plane resolution= 2.5x2.5 mm, 22 slices, scan duration= 6.4 mins. Infants received MRI while asleep and 87% were sedated (prior to MRI) with 25-50 mg/kg chloral hydrate. Neonatal earmuffs (MiniMuffs, Natus Medical Inc., San Carlos, CA, USA) and earplugs molded from silicone-based putty (President Putty, Coltene Whaledent, Mahwah, NJ, USA) were used for ear protection.

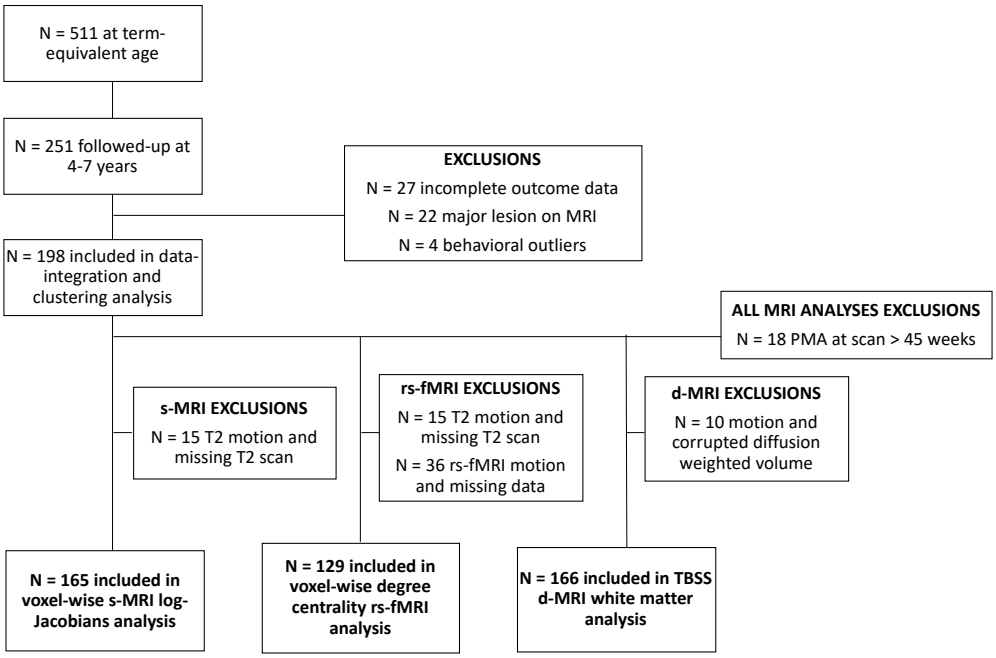



*Figure SM 4.1. Study sample flowchart.*

### 4.6.2 Ethnicity classifications

Ethnicity in Table 4.1 was grouped according to the Office of National Statistics classifications (2016). The groups were characterised accordingly: White (English/Welsh/Scottish/Northern Irish/British, Irish, Any other White background);

Mixed/Multiple ethnic groups (White and Black Caribbean, White and Black African, White and Asian, Any other Mixed/Multiple ethnic background); Asian/Asian British (Indian, Pakistani, Bangladeshi, Chinese, Any other Asian background); Black/African/Caribbean/Black British (African, Caribbean, Any other Black/African/Caribbean background); Other ethnic group (Arab, Any other ethnic group).

### 4.6.3 Data integration and clustering pipeline

An adaptation of the *ExecuteSNF.CC* function (Xu *et al.*, 2017) was used to generate distance and similarity matrices and run both integration and consensus clustering algorithms on a combination of mixed (type 1) and numeric (type 2 and 3) data types. Code can be accessed here: https://github.com/lailahadaya/preterm-ExecuteSNF.CC

***Data integration.*** Similarity Network Fusion (SNF) is a message passing theory method, which updates the final fused matrix over a series of iterations (specified by hyperparameter T; default T=20), increasing the signal-to-noise ratio by updating the final fused matrix with each iteration and discarding weak and inconsistent connections (edges) between subjects (nodes) across data types and maintaining stronger or consistent edges across data types with a neighborhood size of K (Wang *et al.*, 2014, 2018).

***Clustering.*** The "consensus clustering" resultant subgroups obtained in step 2 (called *consensus clusters*) were drawn from an agglomerative hierarchical clustering approach, which partitioned individuals into subgroups based on the pairwise *consensus value* (i.e., the proportion of times each two individuals co-clustered across the several clustering attempts) (Wilkerson and Hayes, 2010; Liu and Shang, 2018). Silhouette width scores were used to measure clustering quality for each subject/node within the network, with values ranging from -1 to 1, where larger values indicate a subject is closer to other subjects within the same cluster than to subjects in other clusters and values closer to -1 indicate misclassification.

### 4.6.4 Estimation of number of clusters

To estimate the optimal number of clusters (between C=2 and C=10), Eigengap and Rotation Cost heuristics were run for each combination of K-alpha hyperparameters. A two-cluster solution (C=2) was estimated to be the optimal number of clusters 60/60 times (30/30 times as the *best* solution using Eigengap and 30/30 times as the *best* solution using Rotation Cost), followed by C=3 (25/60 times; 13/30 times as the *second best* solution using Eigengap and 12/30 times as the *second best* using Rotation Cost) and C=4 (31/60 times; 16/30 times as the *second best*

solution using Eigengap and 15/30 times as the *second best* solution using Rotation Cost) (Figure SM 4.2A).

The clustering steps were repeated for C=2, C=3, C=4, consensus matrices were constructed (Figure SM 4.2B; Figure SM 4.2C) and Silhouette width scores were calculated using the consensus matrices (Figure SM 4.2D). The sample was clustered into C=2 and C=3 subgroups and analyzed for phenotypic differences, as the average Silhouette width scores were highest and had the least number of negative Silhouette width values (Figure SM 4.2D) and as consensus matrices also reflected higher proportions of co-clustering over multiple runs for both C=2 and C=3 compared to C=4.


***Figure SM 4.2. Estimating the optimal number of clusters.***

*A) the percentage of times each number was selected as the most optimal number of clusters (best; pale blue), and second most optimal number of clusters (second best; blue) for the 30 combinations of K-alpha hyperparameters using Eigengap and Rotation Cost. B) Consensus matrices were calculated after clustering into C=2, C=3 and C=4 respectively. Each consensus matrix (the sum of times each pair of subjects co-clustered divided by the sum of times each pair of subjects were both present in a given iteration's subsample) reflects the proportion of times each pair of subjects co-clustered. Darker blue represents higher co-clustering proportions and lighter colors reflect lower proportions. Final clustering results are shown for the four subgroups (green =*

*subgroup 1; beige = subgroup 2; pale pink = subgroup 3; teal = subgroup 4). C) Left: mean consensus pairwise values (mean proportion of times each pair of subjects co-clustered) for each number of clusters (C=2, C=3 and C=4), Right: mean pairwise consensus values for pairs of subjects co-clustering in each subgroup for C=2, C=3 and C=4. D) Silhouette width values for each subgroup after clustering into C=2, C=3, C=4 (left to right respectively).*

As seen in Figure SM 4.3, the majority of participants in subgroup 1 from the two-cluster solution (bottom left; C=2) were preserved into subgroup 1 from the three-cluster solution (bottom right; C=3). Similarily, a large proportion of participants from subgroup 2 in C=2 (top left) were preserved into subgroup 2 from C=3 (top right). The third subgroup (subgroup 3) from C=3 consists of participants that moved from both C=2 subgroup 1 and subgroup 2. Participants in the two subgroups from the two-cluster solution (C=2 subgroup 1 and subgroup 2) that moved into a cluster with a corresponding behavioral profile in the three-cluster solution (C=3 subgroup 1 and subgroup 2) are represented by grey lines, while red lines seen represent participants that moved into the newly emerging subgroup 3.

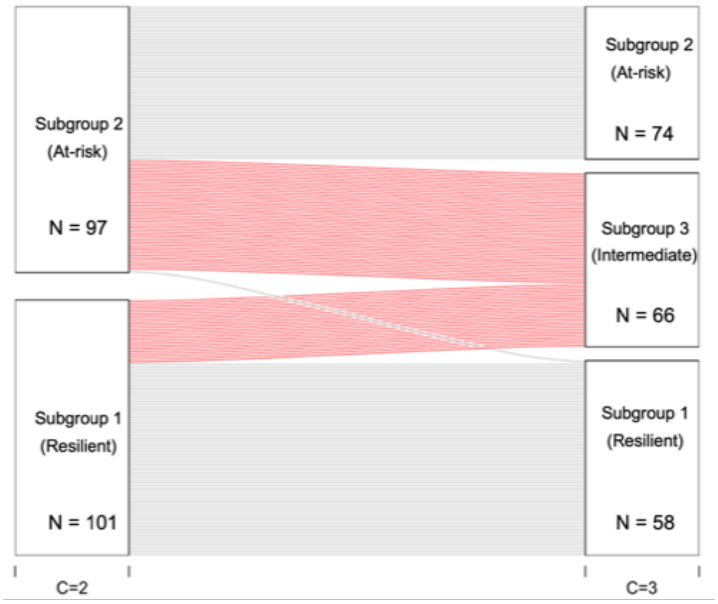



***Figure SM 4.3. Alluvial plot showing the transition of subject assignment from C=2 subgroups to C=3 subgroups.***

### 4.6.5 Selection of in-model and out-of-model variables

Given the prevalence of socio-emotional and executive function difficulties in VPT children, as well as the importance of both clinical and environmental risk factors in shaping these outcomes, relevant variables from these domains were included in the data integration and clustering model, in order to parse heterogeneity in the sample. EmQue empathy subscales and SRS-2 scores were used in-model as measures of socio-emotional processing, which is critical for social communication and interaction (Rieffe, Ketelaar and Wiefferink, 2010; Constantino and

Gruber, 2012; Montagna and Nosarti, 2016). The BRIEF subscales (inhibit, shift, emotional control, working memory and plan/organize) were used as measures of executive function, behavioral self-regulation and cognitive control (Sherman and Brooks, 2010). Four clinical variables (days TPN, days CPAP, days ventilation and GA at birth) were included in the model based on our previous principal component analysis in an overlapping cohort sample, to obtain a clinical summary score derived from 28 maternal and infant clinical variables, as described in (Kanel *et al.*, 2021). Clinical variables were: multiple pregnancy, antenatal hypertension, premature rupture of membranes, urinary tract infection, gestational diabetes, oligohydramnios, polyhydramnios, drug abuse, in vitro fertilization, bacterial infection, mode of delivery, twin-to-twin transfusion, chorioamnionitis, intrauterine growth restriction, antenatal steroid administration, surfactant administration, treatment for patent ductus arteriosus, surgical treatment for necrotising enterocolitis, formula feeding, days on mechanical ventilation, days on continuous positive airway pressure, and days on parenteral nutrition, gestational age, sex, birth weight, feeding on maternal expressed breast milk, preeclampsia and pregnancy induced hypertension and placental abruption or antenatal haemorrhage.

A measure of neighborhood deprivation (Index of Multiple Deprivation) was also used in-model as a measure of socio-demographic risk. Out-of-model cognitive, behavioral and environmental variables were selected in order to provide external validation of the resultant subgroup profiles. These variables were IQ, CSPS total score, temperamental traits (CBQ negative affectivity, surgency and effortful control scores), which reflect the ability to regulate emotions and behaviors in responses to emotional stimuli (Rothbart, 2004) and SDQ internalizing and externalizing symptom scores.

A correlation plot summarising Spearman Rho correlation coefficients between in-model and out-of-model variables is shown below (Figure SM 4.4). The correlations between the out-of-model variables with the in-model variables range from weak to moderate.

### 4.6.6 Post-hoc analysis – clustering based on neonatal socio-demographic and clinical risk factors only

In order to further demonstrate the benefit of using a cluster solution generated by an integrative clustering approach in comparison to that from clustering only one data type, we ran a post-hoc analysis where we clustered the cohort based on only their neonatal socio-demographic and clinical risk factors (i.e., single data type). As expected, results changed substantially (see Figure SM 4.5). The alluvial plots indicate the allocation of each subject (where each subject is represented

by a red line) within subgroups created after clustering with all data types (left) or with only the socio-demographic and clinical risk data type (right). We also investigated the resultant subgroups for differences in childhood outcomes and found no significant differences for C=2 and C=3 (all *p*s>0.05). These results further highlight the importance of integrating heterogenous data types which capture information from different domains.

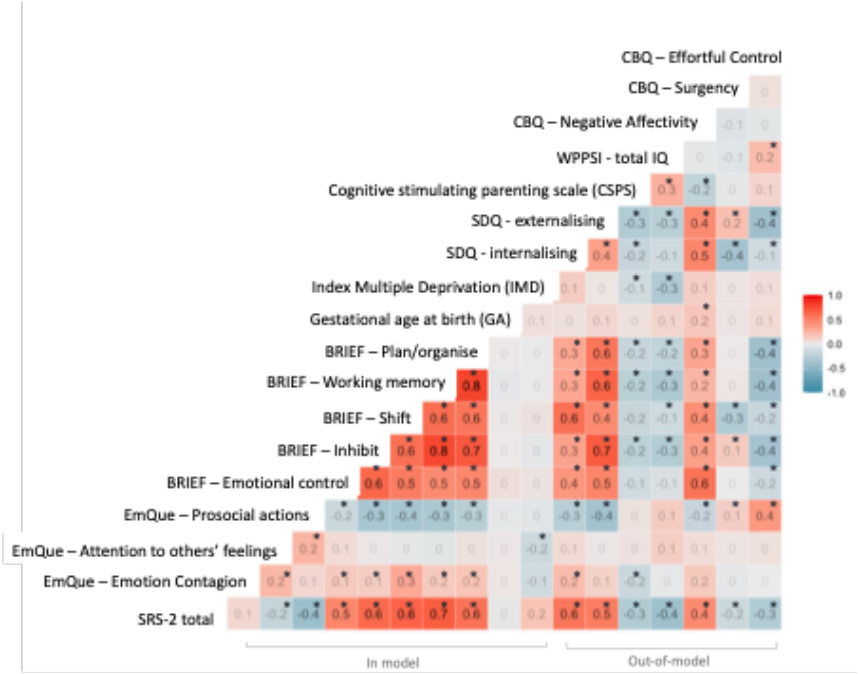



***Figure SM 4.4. Correlation plot of in-model and out-of-model variables.***

*Numbers denote Spearman R values for each pair of variables, which are depicted with a color gradient (see colorbar). P-values are denoted by asterisks whereby *=p<0.05 (uncorrected).*

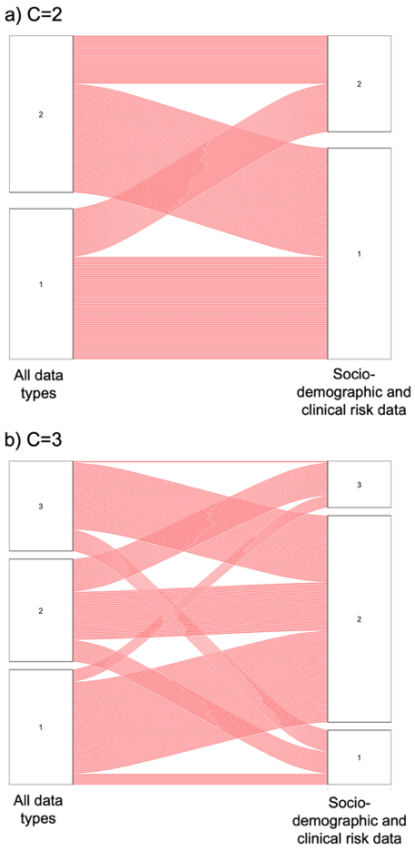



*Figure SM 4.5. Alluvial plot indicating allocation of participants in subgroups after implementing integrative clustering based on all data types (left) and clustering of only the socio-demographic and clinical risk data type (right) for a) two clusters (C=2) and b) three clusters (C=3).*

### 4.6.7 MRI pre-processing and analyses

***Diffusion MRI image pre-processing and Tract Based Spatial Statistics.*** Diffusion MRI (d-MRI) pre-processing steps are described in our previous work (Pecheva *et al.*, 2017; Kanel *et al.*, 2021). To briefly summarize, images were visually checked for motion artefacts and corrupted volumes were excluded. All d-MRI datasets that were included in the analysis had a total of 5 volumes or fewer excluded. BET (version 2.1; http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/BET) (Smith, 2002; Jenkinson *et al.*, 2012)(Jenkinson *et al.*, 2012) was used to extract non-brain tissue from the data and *eddy_current* to correct for eddy current artefacts (Andersson and Sotiropoulos, 2016). FSL's *dtifit* was used to fit the tensor model (FMRIB, Oxford, http://fsl.fmrib.ox.ac.uk).

Image registration was performed using DTI-TK and integrated within the TBSS pipeline to produce a population specific DTI template. From this template a mean FA map was derived and then thinned by perpendicular non-maximum suppression to create a mean FA skeleton. A FA threshold of $\geq 0.15$ was used to limit the inclusion of voxels with high inter-subject variability

and non-white matter voxels. FA and MD maps were projected onto this skeleton prior to voxel-wise statistical analysis.

***Tensor Based Morphometry processing.*** Two input modalities (T2-weighted images and T2-weighted image tissue type segmentations) were registered to a study-specific T2-weighted template using the ANTS software multi-modal Symmetric Normalisation (SyN) (Avants *et al.*, 2011) algorithm as described in Lautarescu et al. (Lautarescu *et al.*, 2021).

Deformation tensor field gradients (i.e., log Jacobian determinant maps) were then computed from the resultant T2-weighted deformation tensor fields (i.e., warps) of the non-linear transformations. Jacobian determinant map values reflect the degree of contraction or expansion a voxel undergoes following the non-linear transformation from native to template space (Avants and Gee, 2004). The logarithm Jacobian maps were smoothed (4 mm FWHM Gaussian filter) to improve the signal to noise ratio. Data were resampled from 0.5 to 1 $mm^3$ isotropic voxel size to decrease computation and memory load.

***Functional MRI image pre-processing and motion-correction.*** Functional images were pre-processed as in Ball et al. (Ball *et al.*, 2016). In summary, images with visible motion artefacts were excluded after visual inspection. Single-subject independent component analysis (ICA) with automatic dimensionality estimation was applied to each individual's dataset using FSL MELODIC (Beckmann and Smith, 2004), following removal of the first 6 volumes (to allow for T1 equilibration), motion correction using MCFLIRT, and high-pass filtering (125 s cutoff, 0.008 Hz). Following ICA, FSL FIX (Salimi-Khorshidi *et al.*, 2014) was applied for automatic denoising and artefact removal. A population-specific neonatal template with tissue priors was used for standard-space masking (Serag *et al.*, 2012), and the FIX algorithm was trained on hand-classified fMRI datasets from 40 preterm infants aged 28-44 weeks, collected on the same scanner (including both low-motion and high-motion subjects; for more details see supplemental materials of (Ball *et al.*, 2016). After components were classified as either signal or noise, the unique variance of each noise component, as well as the full variance of the motion parameters and derivatives, were regressed out of the data (Satterthwaite *et al.*, 2013; Griffanti *et al.*, 2014). Standardized DVARS, a framewise data quality index (Power *et al.*, 2012), was calculated before and after applying FIX and significantly improved after FIX clean-up ($t(315)=9.01$, $p<0.001$). Finally, FSL Motion Outliers was applied to each dataset to identify remaining volumes that were corrupted by large motion, and subjects with more than two standard deviations above the mean number of corrupted

volumes were removed. This resulted in a final fMRI sample of 298 infants. Of these, 129 with complete follow-up data were included in further analysis.

Cleaned functional images from this sample were resampled to 2 $mm^3$ isotropic voxels, registered to the study-specific T2-weighted template using boundary-based registration, and smoothed with a 4 mm full-width half-maximum Gaussian kernel.

*Table SM 4.1. Subgroup sample sizes for each imaging modality.*

| Cluster solution | Subgroup | d-MRI sample | s-MRI sample | rs-fMRI sample | Total sample |
|---|---|---|---|---|---|
| **C=2** | *Resilient* | 80 | 82 | 63 | 97 |
| | *At-risk* | 86 | 83 | 66 | 101 |
| **C=3** | *Resilient* | 60 | 61 | 48 | 74 |
| | *At-risk* | 52 | 48 | 38 | 58 |
| | *Intermediate* | 54 | 56 | 43 | 66 |
| **Total sample** | | 166 | 165 | 129 | 198 |

*Table SM 4.2. Two-cluster solution profiles using in-model and out-of-model variables.*

| | Variables | Subgroup 1 | Subgroup 2 | p-value | Effect size |
|---|---|---|---|---|---|
| **In-model variables** | ***SRS-2*** | 22.00 (5.00) | 43.00 (5.75) | <0.001 | -0.79 |
| | ***EmQue: Emotion Contagion*** | 0.17 (0.33) | 0.33 (0.67) | 0.001 | -0.26 |
| | ***EmQue: Attention to others' emotions*** | 1.43 (0.57) | 1.29 (0.43) | 0.933 | 0.01 |
| | ***EmQue: Prosocial actions*** | 1.33 (0.50) | 0.83 (0.50) | <0.001 | 0.53 |
| | ***BRIEF-P: Emotion control*** | 13 (4.00) | 16.50 (4.43) | <0.001 | -0.52 |
| | ***BRIEF-P: Inhibit*** | 22.00 (5.00) | 28.50 (8.75) | <0.001 | -0.73 |
| | ***BRIEF-P: Shift*** | 11.00 (2.00) | 15.00 (5.00) | <0.001 | -0.81 |
| | ***BRIEF-P: Working memory*** | 21.00 (5.00) | 29.00 (8.75) | <0.001 | -0.81 |
| | ***BRIEF-P: Plan & organize*** | 14.00 (3.00) | 17.50 (3.00) | <0.001 | -0.73 |

| | | | | | |
|---|---|---|---|---|---|
| | ***GA*** | 30.29 (4.00) | 29.93 (4.43) | 0.590 | 0.04 |
| | ***IMD at birth*** | 15.99 (20.68) | 15.15 (14.36) | 0.888 | 0.01 |
| | ***Days TPN*** *(0:1:2), n=* | 33:53:11 | 35:45:21 | 0.153 ($X^2$=3.76) | V=0.14 |
| | ***Days ventilation*** *(0:1:2), n=* | 58:35:4 | 43:39:10 | 0.002 | V=0.25 |
| | ***Days CPAP*** *(0:1:2), n=* | 23:55:19 | 10:70:21 | 0.031 ($X^2$=6.94) | V=0.18 |
| **Out-of-model variables** | ***SDQ total*** | 6.00 (5.00) | 11.00 (5.75) | <0.001 | -0.73 |
| | ***SDQ: Internalizing*** | 1.00 (3.00) | 4.00 (3.75) | <0.001 | -0.56 |
| | ***SDQ: Externalizing*** | 4.00 (3.00) | 8.00 (5.00) | <0.001 | -0.61 |
| | ***$CSPS*** | 18.00 (3.00) | 17.00 (3.53) | 0.003 | 0.24 |
| | ***$$WPPSI: Full scale IQ*** | 110.00 (17.50) | 104.50 (26.00) | 0.026 | 0.18 |
| | ***CBQ: Negative affectivity*** | 3.67 (1.10) | 4.42 (00.96) | <0.001 | -0.49 |
| | ***CBQ: Surgency*** | 5.00 (0.92) | 4.67 (1.31) | 0.057 | 0.16 |
| | ***CBQ: Effortful control*** | 5.50 (0.92) | 5.17 (1.23) | 0.002 | 0.25 |
| | ***Corrected age at assessment: years*** | 4.67 (0.82) | 4.59 (0.58) | 0.074 | 0.15 |
| | ***IMD at assessment*** | 13.75 (19.58) | 15. 88 (16.51) | 0.499 | -0.057 |
| | ***Sex*** *(M:F), n=* | 50:47 | 50:51 | 0.885 ($X^2$=0.02) | V=0.02 |
| | ***Total,*** *n (%)* | 97 (48.99%) | 101 (51.01%) | / | |

***Note:*** *Median (IQR) is provided unless otherwise stated. P-values refer to results from Mann-Whitney non-parametric test. Effect sizes reported are Glass Rank Biserial Correlation unless otherwise stated, where Cramer's V is reported for categorical variables. Days TPN (0:1:2), days ventilation (0:1:2) and days CPAP (0:1:2), correspond to the ratio of the three clinical risk categories: 0, 1 and 2. Respectively, they correspond to zero days, more than zero days but less than the top quintile, and within the top quintile. $=one missing participant; $$=two missing participants. Chi-squared test was used for categorical comparisons. Abbreviations: BRIEF-P = Behavior Rating Inventory of Executive Function pre-school version; CBQ = Childhood Behavioral Questionnaire; CSPS = Cognitively Stimulating Parenting Scale; CPAP = continuous positive airway pressure. EmQue = Empathy Questionnaire; GA = gestational age; IMD = Index of Multiple Deprivation; IQR = interquartile range; PMA = post menstrual age at scan; SDQ = Strengths and Difficulties Questionnaire; SRS-2 = Social Responsiveness Scale – Second Edition; TPN = total parenteral nutrition.*

*Table SM 4.3. Three-cluster solution profiles using in-model and out-of-model variables.*

| | Variables | Subgroup 1 | Subgroup 2 | Subgroup 3 | p-value | Effect size |
|---|---|---|---|---|---|---|
| **In-model Variables** | ***SRS-2*** | 18.00 (14) | 56.50 (24.50) | 36.00 (12.50) | <0.001 | 0.55 |
| | ***EmQue: Emotion Contagion*** | 0.17 (0.33) | 0.50 (0.67) | 0.17 (0.33) | 0.003 | 0.08 |
| | ***EmQue: Attention to others' emotions*** | 1.43 (0.57) | 1.29 (0.43) | 1.29 (0.57) | 0.186 | 0.01 |
| | ***EmQue: Prosocial actions*** | 1.33 (0.46) | 0.83 (0.50) | 1.00 (0.33) | <0.001 | 0.24 |
| | ***BRIEF-P: Emotion control*** | 13.00 (5.00) | 18.00 (5.00) | 14.00 (4.75) | <0.001 | 0.34 |
| | ***BRIEF-P: Inhibit*** | 21.00 (5.00) | 31.50 (8.75) | 25.00 (6.00) | <0.001 | 0.44 |
| | ***BRIEF-P: Shift*** | 11.00 (2.00) | 17.00 (4.75) | 13.00 (2.00) | <0.001 | 0.59 |
| | ***BRIEF-P: Working memory*** | 21.00 (5.00) | 31.00 (8.75) | 26.00 (5.00) | <0.001 | 0.48 |
| | ***BRIEF-P: Plan & organize*** | 13.00 (4.00) | 18.00 (4.50) | 16.00 (3.00) | <0.001 | 0.42 |
| | ***GA*** | 29.36 (3.64) | 30.00 (4.79) | 30.64 (3.39) | 0.003 | 0.04 |
| | ***IMD at birth*** | 13.23 (14.77) | 14.12 (9.86) | 22.08 (17.11) | 0.018 | 0.06 |
| | ***Days TPN*** *(0:1:2), n=* | 16:48:10 | 17:28:13 | 35:22:9 | <0.001 ($X^2$=19.64) | V=0.22 |
| | ***Days ventilation*** *(0:1:2), n=* | 40:30:4 | 23:21:14 | 38:23:5 | 0.014 | V=0.19 |
| | ***Days CPAP*** *(0:1:2), n=* | 9:45:20 | 2:40:16 | 22:40:4 | <0.001 | V=0.27 |
| **Out-of-model variables** | ***SDQ total*** | 5.00 (4.00) | 13.00 (5.75) | 8.00 (5.00) | <0.001 | 0.40 |
| | ***SDQ: Internalizing*** | 1.00 (2.00) | 4.00 (4.75) | 2.00 (2.83) | <0.001 | 0.28 |
| | ***SDQ: Externalizing*** | 4.00 (3.00) | 9.00 (5.00) | 6.00 (4.00) | <0.001 | 0.28 |
| | ***CSPS*** | 19.00 (2.00) | 17.00 (3.00) | 18.00 (4.00) | 0.006 | 0.28 |
| | ***WPPSI: Full scale IQ*** | 112.00 (16.00) | 103.50 (30.25) | 105.00 (22.50) | 0.007 | 0.07 |
| | ***CBQ: Negative affectivity*** | 3.58 (1.15) | 4.50 (1.04) | 4.25 (0.98) | <0.001 | 0.20 |
| | ***CBQ: Surgency*** | 4.96 (0.83) | 4.61 (1.21) | 4.75 (1.25) | 0.071 | 0.03 |
| | ***CBQ: Effortful control*** | 5.67 (0.83) | 5.04 (1.33) | 5.21 (0.99) | 0.002 | 0.10 |

| | | | | | |
|---|---|---|---|---|---|
| ***Corrected age at assessment: years*** | 4.63 (0.80) | 4.60 (0.56) | 4.69 (1.07) | 0.78 | 0.01 |
| ***IMD at assessment*** | 12.02 (14.44) | 18.73 (17.18) | 15.15 (17.99) | 0.133 | 0.034 |
| ***Sex*** *(M:F), n=* | 36:38 | 29:29 | 35:31 | 0.871 ($X^2$=0.28) | V=0.04 |
| **Total**, n (%) | 74 (37.37%) | 58 (29.29%) | 66 (33.33%) | / | |

***Note:*** *Median (IQR) is provided unless otherwise stated. P-values refer to results from Kruskal-Wallis non-parametric test. Effect sizes reported are Epsilon Squared unless otherwise stated where Cramer's V is reported for categorical variables. Days TPN (0:1:2), days ventilation (0:1:2) and days CPAP (0:1:2), correspond to the ratio of the three clinical risk categories: 0,1 and 2. Respectively, they correspond to zero days, more than zero days but less than the top quintile and within the top quintile. $=one missing participants; $$=two missing participants. Chi-squared test was used for categorical comparisons when subject count per cell was >5 and Fisher's Exact when cell count was 5 or less. Abbreviations: BRIEF-P = Behavior Rating Inventory of Executive Function preschool version; CBQ = Childhood Behavioral Questionnaire; CSPS = Cognitively Stimulating Parenting Scale; CPAP = continuous positive airway pressure. EmQue = Empathy Questionnaire; GA = gestational age; IMD = Index of Multiple Deprivation; IQR = interquartile range; PMA = post menstrual age at scan; SDQ = Strengths and Difficulties Questionnaire; SRS-2 = Social Responsiveness Scale – Second Edition; TPN = total parenteral nutrition.*

***Table SM 4.4. Effect sizes, number of significant voxels and p-values for brain regions showing significant differences between subgroups.***

| Brain measure | Contrast | Region | Number of voxels | p-value | Cohen's F effect size |
|---|---|---|---|---|---|
| **Log-Jacobian determinant** | ***Resilient > intermediate*** | Left insular | 553 | 0.01 | 0.52 |
| | ***Resilient > intermediate*** | Left orbitofrontal | 1652 | 0.01 | 0.46 |
| | ***Resilient > intermediate*** | Right orbitofrontal regions | 897 | 0.01-0.05 | 0.48 |
| **Degree centrality** | ***Resilient > intermediate*** | Left orbitofrontal | 11 | 0.04 | 0.45 |
| **TBSS – FA** | ***Intermediate > at-risk*** | Fornix, CC, CST, ILF, IFO and UF | 1370 | 0.01 | 0.08 |
| **TBSS – MD** | ***Intermediate < at-risk*** | Fornix and CC body | 93 | 0.04 | 0.05 |

***Note.*** *Abbreviations: CC= corpus callosum, CST= corticospinal tract, FA = fractional anisotropy, IFO = inferior fronto-occipital fasciculus, ILF=inferior longitudinal fasciculus, MD = mean diffusivity, TBSS = tract based spatial statistics, UF = uncinate fasciculus.*

### 4.6.8 Sensitivity analyses

In order to account for multiple pregnancy confounding, we conducted sensitivity analyses including only one child, at random, from each set of multiple pregnancy siblings. Results for the in-model and out-of-model cognitive, behavioral, clinical and socio-demographic risk variables remained similar to results including children born from multiple pregnancy for both C=2 (Table

SM 4.5) and C=3 (Table SM 4.6). Sensitivity analysis results showed that the log-Jacobian brain differences observed between the subgroups were largely preserved (Figure SM 4.6), with the resilient subgroup having larger brain volumes in the left insula and bilateral orbitofrontal cortices compared to the intermediate group (Figure SM 4.6A) and the intermediate group showing higher FA and lower MD in white matter tracts compared to the at-risk subgroup (Figure SM 4.6C). Results of sensitivity analysis also showed an additional result that was not observed in the full sample: the resilient subgroup displayed larger brain volumes in the right insula compared to the intermediate subgroup. However, the functional connectivity degree centrality was no longer significant, with $p$=0.08 (Figure SM 4.6B). We believe this may have been due to a loss in power, as a result of the smaller sample size.

***Table SM 4.5. Results of sensitivity analysis, including only one child from each set of multiple pregnancy siblings, of the two-cluster solution profiles using in-model and out-of-model variables.***

| | Variables | Subgroup 1 | Subgroup 2 | p-value |
|---|---|---|---|---|
| **In-model variables** | ***SRS-2*** | 22.00 (5) | 43.00 (5.75) | <0.001 |
| | ***EmQue: Emotion Contagion*** | 0.17 (0.33) | 0.33 (0.67) | 0.005 |
| | ***EmQue: Attention to others' emotions*** | 1.43 (0.57) | 1.29 (0.43) | 0.658 |
| | ***EmQue: Prosocial actions*** | 1.33 (0.50) | 0.83 (0.50) | <0.001 |
| | ***BRIEF-P: Emotion control*** | 13 (4.00) | 16.50 (4.43) | <0.001 |
| | ***BRIEF-P: Inhibit*** | 22.00 (5.00) | 28.50 (8.75) | <0.001 |
| | ***BRIEF-P: Shift*** | 11.00 (2.00) | 15.00 (5.00) | <0.001 |
| | ***BRIEF-P: Working memory*** | 21.00 (5.00) | 29.00 (8.75) | <0.001 |
| | ***BRIEF-P: Plan & organize*** | 14.00 (3.00) | 17.50 (3.00) | <0.001 |
| | ***GA*** | 30.29 (4.00) | 29.93 (4.43) | 0.315 |
| | ***IMD*** | 15.99 (20.68) | 15.15 (14.36) | 0.968 |
| | ***Days TPN*** *(0:1:2), n=* | 27:43:11 | 31:38:21 | 0.197 ($X^2$=3.25) |
| | ***Days ventilation*** *(0:1:2), n=* | 49:28:4 | 36:36:18 | 0.004 |
| | ***Days CPAP*** *(0:1:2), n=* | 18:48:15 | 9:61:20 | 0.091 ($X^2$=4.80) |
| **Out-of-model variables** | ***SDQ total*** | 6.00 (5.00) | 11.00 (5.75) | <0.001 |
| | ***SDQ: Internalising*** | 1.00 (3.00) | 4.00 (3.75) | <0.001 |
| | ***SDQ: Externalising*** | 4.00 (3.00) | 8.00 (5.00) | <0.001 |

| | | | |
|---|---|---|---|
| [$]***CSPS*** | 18.00 (3.00) | 17.00 (3.53) | 0.006 |
| [$$]***WPPSI: Full scale IQ*** | 110.00 (17.50) | 104.50 (26.00) | 0.019 |
| ***CBQ: Negative affectivity*** | 3.67 (1.10) | 4.42 (00.96) | 0.058 |
| ***CBQ: Surgency*** | 5.00 (0.92) | 4.67 (1.31) | <0.001 |
| ***CBQ: Effortful control*** | 5.50 (0.92) | 5.17 (1.23) | 0.012 |
| ***Corrected age at assessment: years*** | 4.67 (0.82) | 4.59 (0.58) | 0.082 |
| ***IMD at assessment*** | 13.75 (19.65) | 15.75 (16.76) | 0.553 |
| ***Sex*** *(M:F), n=* | 43:38 | 43:47 | 0.589 ($X^2$=0.29) |
| ***Total****, n (%)* | 81 (47.37%) | 90 (52.63%) | / |

***Note:*** *Median (IQR) is provided unless otherwise stated. P-values refer to results from Mann-Whitney non-parametric test. Days TPN (0:1:2), days ventilation (0:1:2) and days CPAP (0:1:2), correspond to the ratio of the three clinical risk categories: 0,1 and 2. Respectively, they correspond to zero days, more than zero days but less than the top quintile and within the top quintile. [$]=one missing participant; [$$]=two missing participants. Chi-squared test was used for categorical comparisons when subject count per cell was >5 and Fisher's Exact when cell count was 5 or less. Abbreviations: BRIEF-P = Behavior Rating Inventory of Executive Function preschool version; CBQ = Childhood Behavioral Questionnaire; CSPS = Cognitively Stimulating Parenting Scale; CPAP = continuous positive airway pressure. EmQue = Empathy Questionnaire; GA = gestational age; IMD = Index of Multiple Deprivation; IQR = interquartile range; PMA = post menstrual age at scan; SDQ = Strengths and Difficulties Questionnaire; SRS-2 = Social Responsiveness Scale – Second Edition; TPN = total parenteral nutrition.*

*Table SM 4.6. Results of sensitivity analysis, including only one child from each set of multiple pregnancy siblings, of the three-cluster solution profiles using in-model and out-of-model variables.*

| | Variables | Subgroup 1 | Subgroup 2 | Subgroup 3 | p-value |
|---|---|---|---|---|---|
| **In-model variables** | ***SRS-2*** | 18.50 (13.75) | 36.00 (11.75) | 57.00 (25.00) | <0.001 |
| | ***EmQue: Emotion Contagion*** | 0.17 (0.43) | 0.17 (0.46) | 0.50 (0.67) | 0.003 |
| | ***EmQue: Attention to others' emotions*** | 1.43 (0.43) | 1.29 (0.57) | 1.29 (0.43) | 0.187 |
| | ***EmQue: Prosocial actions*** | 1.33 (0.33) | 1.00 (0.33) | 0.83 (0.50) | <0.001 |
| | ***BRIEF-P: Emotion control*** | 13.00 (5.00) | 14.00 (4.00) | 18.00 (5.00) | <0.001 |
| | ***BRIEF-P: Inhibit*** | 21.00 (5.50) | 25.00 (7.00) | 30.00 (9.00) | <0.001 |
| | ***BRIEF-P: Shift*** | 11.00 (2.00) | 13.00 (2.00) | 17.00 (4.50) | <0.001 |
| | ***BRIEF-P: Working memory*** | 21.00 (4.00) | 26.00 (5.75) | 30.00 (8.50) | <0.001 |
| | ***BRIEF-P: Plan & organize*** | 13.00 (3.00) | 16.50 (3.00) | 17.00 (4.00) | <0.001 |
| | ***GA*** | 29.79 (3.68) | 30.50 (3.29) | 29.86 (4.71) | 0.003 |
| | ***IMD,*** | 12.88 (15.34) | 23.00 (17.43) | 14.40 (9.62) | 0.018 |
| | ***Days TPN*** *(0:1:2), n=* | 13:39:10 | 15:23:13 | 30:19:9 | 0.003 ($X^2$=16.41) |
| | ***Days ventilation*** *(0:1:2), n =* | 34:24:4 | 19:18:14 | 32:22:4 | 0.013 |
| | **Days CPAP** (0:1:2), n= | 7:39:16 | 2:34:15 | 18:36:4 | 0.001 |
| **Out-of-model variables** | **SDQ total** | 5.00 (5.00) | 8.17 (5.75) | 13.00 (6.50) | <0.001 |
| | **SDQ: Internalizing** | 1.00 (2.00) | 2.00 (3.00) | 5.00 (4.00) | <0.001 |
| | **SDQ: Externalizing** | 3.50 (3.00) | 6.00 (3.75) | 9.00 (5.50) | <0.001 |
| | **CSPS** | 18.90 (2.00) | 17.84 (4.00) | 17.00 (3.00) | 0.006 |
| | **WPPSI: Full scale IQ** | 112.00 (16.00) | 105.50 (22.25) | 100.00 (30.00) | 0.007 |
| | **CBQ: Negative affectivity** | 3.58 (1.15) | 4.25 (0.80) | 4.55 (1.08) | <0.001 |
| | **CBQ: Surgency** | 5.00 (0.73) | 4.75 (1.25) | 4.58 (1.30) | 0.071 |
| | **CBQ: Effortful control** | 5.67 (0.83) | 5.17 (0.92) | 5.08 (1.42) | <0.001 |

| | | | | |
|---|---|---|---|---|
| **Corrected age at assessment: years** | 4.60 (0.50) | 4.69 (0.94) | 4.60 (0.49) | 0.778 |
| **IMD at assessment** | 12.02 (15.26) | 18.49 (17.88) | 15.29 (18.66) | 0.203 |
| ***Sex (M:F), n=*** | 32:30 | 24:27 | 30:28 | 0.859 ($X^2$=0.30) |
| ***Total, n (%)*** | 62 | 51 | 58 | / |

***Note*****:** *Median (IQR) is provided unless otherwise stated. P-values refer to results from Kruskal-Wallis nonparametric test. Days TPN (0:1:2), days ventilation (0:1:2) and days CPAP (0:1:2), correspond to the ratio of the three clinical risk categories: 0,1 and 2. Respectively, they correspond to zero days, more than zero days but less than the top quintile and within the top quintile. $=one missing participants; $$=two missing participants. Chi-squared test was used for categorical comparisons when subject count per cell was >5 and Fisher's Exact when cell count was 5 or less. Abbreviations: BRIEF-P = Behavior Rating Inventory of Executive Function preschool version; CBQ = Childhood Behavioral Questionnaire; CSPS = Cognitively Stimulating Parenting Scale; CPAP = continuous positive airway pressure. EmQue = Empathy Questionnaire; GA = gestational age; IMD = Index of Multiple Deprivation; IQR = interquartile range; PMA = post menstrual age at scan; SDQ = Strengths and Difficulties Questionnaire; SRS-2 = Social Responsiveness Scale – Second Edition; TPN = total parenteral nutrition.*

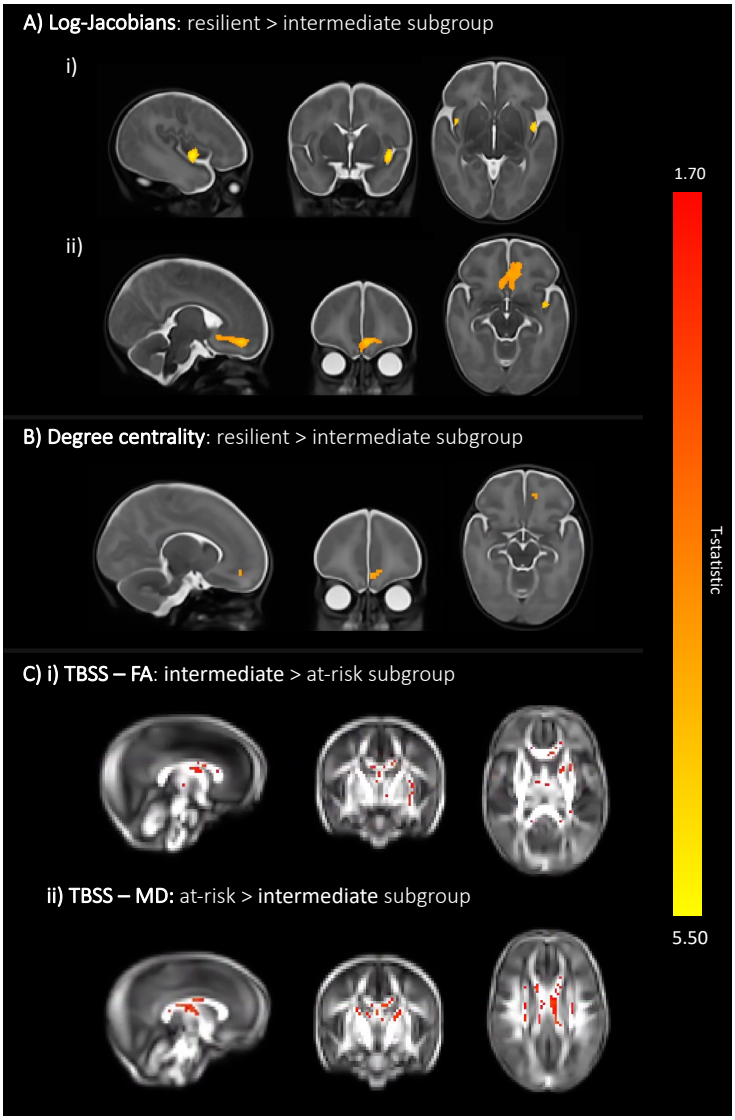



***Figure SM 4.6. Brain differences at term-equivalent age of the three cluster-solution, including only one child from each set of multiple pregnancy siblings.***

*A) Compared to subgroup 3, subgroup 1 had voxels in the i) left and right insula and the ii) left and right orbitofrontal cortex with significantly larger log-Jacobian determinant values at $p < .05$ (i.e., larger relative volumes in these areas). Sex and PMA were included as covariates and TFCE and FWE corrections were applied. B) an overlapping left orbitofrontal region had voxels showing larger degree centrality (i.e., higher functional connectivity with all other voxels in the grey matter mask) values in subgroup 1 compared to subgroup 3 ($p=0.08$). Sex and PMA were included as covariates and FWE and TFCE corrections were applied. T-statistic values are represented in the color bar, where red colored voxels indicate lower T-statistic values and more yellow voxels indicate higher T-statistic values. C) Colored voxels indicate regions of white matter with: i) significantly higher FA values in the intermediate subgroup compared to the at-risk subgroup and ii) significantly higher MD values in the at-risk subgroup compared to the intermediate subgroup ($p<0.05$). T-statistic values are represented in the color bar, where red colored voxels indicate smaller T-statistic values and yellow voxels indicate higher T-statistic values, ranging between 1.70 and 5.50.*

For C=2, out of the 24 sibling sets with more than one sibling included in the clustering analyses, four sibling sets (Figure SM 4.7; sibling sets: D, F, P and R) do not cluster together, while the remaining 20 sibling sets do. As for C=3, 5 out of the 24 sibling sets do not co-cluster (Figure SM 4.7; sibling sets: D, J, P, S and W). Although siblings tend to group together, we think it is unlikely they may be driving the subgroup profiles described in the main analyses, as we find similar results in the sensitivity analyses which only include one sibling from each set. Moreover, sibling sets are evenly dispersed across the different subgroups and do not cluster into a single subgroup, which suggests that children born from multiple pregnancies do not tend to co-cluster with one another.

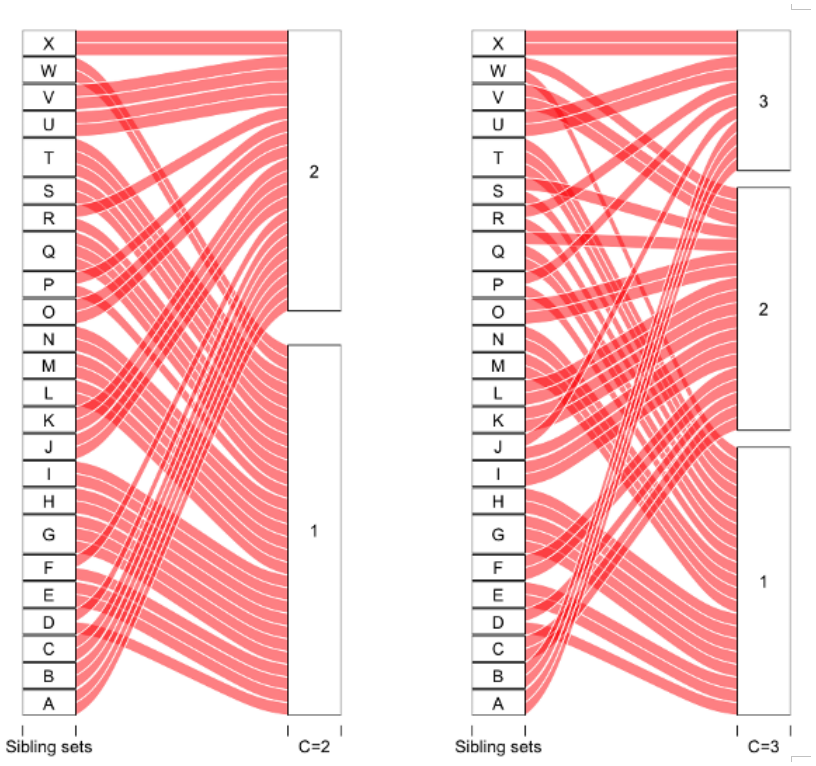



***Figure SM 4.7. Alluvial plots for sibling sets.***

*Alluvial plots indicating which subgroup each sibling within the sibling sets clusters into. Results for the two-cluster solution (C=2) are on the left and the three-cluster solution (C=3) on the right.*

### 4.6.9 Investigating differences between subgroup profiles after adjusting for confounders

To explore whether possible confounders (age and sex) altered our results we re-ran comparisons between clusters for non-MRI out-of-model (SDQ internalizing , SDQ externalizing, CSPS, full scale IQ, CBQ surgency, CBQ effortful control and CBQ negative affectivity) and in-model measures (SRS-2 total, EmQue emotion contagion, EmQue Attention to others' emotions, EmQue Prosocial actions, BRIEF-P Emotion control, BRIEF-P Inhibit, BRIEF-P Shift, BRIEF-P Working memory, BRIEF-P Plan and organize). We found that results remained similar and survived FDR correction. However, CBQ surgency scores were lower in the at-risk compared to the resilient subgroups in the C=2 cluster solution only, after adjusting for age and sex. This was also the case for C=3, whereby the at-risk subgroup showed significantly lower CBQ surgency scores compared to both resilient and intermediate subgroups. A meta-analysis investigating gender differences in temperament found that boys show increased surgency levels compared to girls (Else-Quest *et al.*, 2006). Furthermore, differences between the at-risk and intermediate subgroups in EmQue prosocial subscale scores emerged after adjusting for age and sex. All results survived FDR correction.

It is also worth noting that controlling for age in the post-hoc analyses is particularly relevant for the BRIEF-P raw score results. The BRIEF-P has been validated for use in children up to age 5 years 11 months, but to ensure consistency in the assessment protocol, it was also administered to the small proportion of older children in our sample (15% aged 6; 3% aged 7). To account for this, we used the BRIEF-P raw scores and not the age-adjusted T-scores, which may be appropriate for within-sample analyses. Furthermore, while BRIEF-P scores vary with age in the general population (Sherman and Brooks, 2010), we did not observe this pattern in our sample (R values for the correlation between the five BRIEF-P subscales and participants' age at assessment were between -0.130 and 0.001; all $p > 0.05$).

### 4.6.10 Supplementary post-hoc analyses investigating subgroup differences in clinical variables other than those used in-model

Supplementary post-hoc analyses investigated between-subgroup differences in clinical variables which were not included as in-model variables. The out-of-model clinical variables did not significantly differ between subgroups after FDR correction (Table SM 4.7 and Table SM 4.8).

*Table SM 4.7. Two-cluster solution results: out-of-model clinical variables.*

| | Subgroup 1 (Resilient) | Subgroup 2 (At-risk) | Uncorrected p-value | FDR corrected p-value |
|---|---|---|---|---|
| **Multiple pregnancy, ratio (no: yes: unknown)** | 57: 23: 17 | 73: 12: 16 | 0.068 | 0.499 |
| **Antenatal hypertension, ratio (no: yes)** | 95:2 | 95:6 | 0.280 | 0.943 |
| **Premature rupture of membranes, ratio (no: yes: unknown)** | 60: 19: 18 | 68: 16: 17 | 0.703 | 0.943 |
| **Urinary tract infection, ratio (no: yes: unknown)** | 78: 1: 18 | 82: 2: 17 | 0.943 | 0.943 |
| **Gestational diabetes, ratio (no: yes: unknown)** | 76: 3: 18 | 79: 5: 17 | 0.848 | 0.943 |
| **Oligohydramnios, ratio (no: yes: unknown)** | 76: 3: 18 | 78: 6: 17 | 0.648 | 0.943 |
| **Polyhydramnios, ratio (no: yes: unknown)** | 79: 0: 18 | 83: 1: 17 | 0.925 | 0.943 |
| **Maternal drug abuse, ratio (no: unknown)** | 79:18 | 84:17 | 0.895 | 0.943 |
| **In vitro fertilization, ratio (no: yes: unknown)** | 70: 9: 18 | 76: 8: 17 | 0.881 | 0.943 |
| **Bacterial infection, ratio (no: yes)** | 94:3 | 99:2 | 0.678 | 0.943 |
| **Mode of delivery, ratio (elective: emergency: vaginal: unknown)** | 8: 53: 32: 4 | 9: 53: 32: 7 | 0.867 | 0.943 |
| **Twin-to-twin transfusion, ratio (no: yes: unknown)** | 78: 3: 16 | 84: 1: 16 | 0.599 | 0.943 |
| **Chorioamnionitis, ratio (no: yes: unknown)** | 78: 1: 18 | 82: 2: 17 | 0.943 | 0.943 |
| **Intrauterine growth restriction, ratio (no: yes: unknown)** | 70: 9: 18 | 77: 7: 17 | 0.767 | 0.943 |
| **Surfactant administered, ratio (no: yes)** | 53: 44 | 39: 62 | **0.034** | 0.498 |
| **Patent ductus arteriosus treatment, ratio (no: yes)** | 93: 4 | 92: 9 | 0.252 | 0.943 |
| **Surgery for necrotising enterocolitis, ratio (no: yes)** | 97: 0 | 96: 5 | 0.060 | 0.499 |
| **Formula feeding, ratio (no: yes)** | 55: 42 | 44: 46 | 0.861 | 0.943 |
| **Birth weight in grams, median (iqr)** | 1280 (660) | 1270 (505) | 0.503 | 0.943 |
| **Feeding on maternal expressed breast milk, ratio (no: yes)** | 11: 86 | 15: 86 | 0.603 | 0.943 |
| **Preeclampsia and induced hypertension, ratio (no: yes: unknown)** | 72: 7: 18 | 72: 12: 17 | 0.532 | 0.943 |
| **Placental abruption or antenatal hemorrhage, ratio (no: yes: unknown)** | 68: 18: 11 | 66: 15: 20 | 0.242 | 0.943 |

***Note.*** *Chi-squared test was used for categorical comparisons when subject count per cell was >5 and Fisher's Exact when cell count was 5 or less. Mann-Whitney non-parametric test was used for the only continuous variable: birth weight in grams.*

*Table SM 4.8. Three-cluster solution results: out-of-model clinical variables.*

| | Subgroup 1 (Resilient) | Subgroup 2 (At-risk) | Subgroup 3 (Intermediate) | Uncorrected p-value | FDR corrected p-value |
|---|---|---|---|---|---|
| **Multiple pregnancy, ratio (no: yes: unknown)** | 42: 18: 14 | 37: 9: 12 | 51: 8: 7 | 0.104 | 0.458 |
| **Antenatal hypertension, ratio (no: yes)** | 73:1 | 57:1 | 60:6 | 0.058 | 0.425 |
| **Premature rupture of membranes, ratio (no: yes: unknown)** | 44: 15: 15 | 36: 10: 12 | 48: 10: 8 | 0.510 | 0. 680 |
| **Urinary tract infection, ratio (no: yes: unknown)** | 58: 1: 15 | 45: 1: 12 | 57: 1: 8 | 0.641 | 0.705 |
| **Gestational diabetes, ratio (no: yes: unknown)** | 56: 3: 15 | 43: 3: 12 | 56: 2: 8 | 0.597 | 0.691 |
| **Oligohydramnios, ratio (no: yes: unknown)** | 56: 3: 15 | 42: 4: 12 | 56: 2: 8 | 0.473 | 0.680 |
| **Polyhydramnios, ratio (no: yes: unknown)** | 59: 0: 15 | 46: 0: 12 | 57: 1: 8 | 0.370 | 0.680 |
| **Maternal drug abuse, ratio (no: unknown)** | 59: 15 | 46: 12 | 58: 8 | 0.349 | 0.680 |
| **In vitro fertilization, ratio (no: yes: unknown)** | 51: 8: 15 | 42: 4: 12 | 53: 5: 8 | 0.550 | 0.680 |
| **Bacterial infection, ratio (no: yes)** | 72: 2 | 56: 2 | 65: 1 | 0.861 | 0.861 |
| **Mode of delivery, ratio (elective: emergency: vaginal: unknown)** | 6: 42: 22: 4 | 3: 28: 21: 6 | 8: 36: 21: 1 | 0.343 | 0.680 |
| **Twin-to-twin transfusion, ratio (no: yes: unknown)** | 58: 2: 14 | 46: 0: 12 | 58: 2: 6 | 0.205 | 0.680 |
| **Chorioamnionitis, ratio (no: yes: unknown)** | 58: 1: 15 | 46: 0: 12 | 56: 2: 8 | 0.438 | 0.680 |
| **Intrauterine growth restriction, ratio (no: yes: unknown)** | 53: 6: 15 | 43: 3: 12 | 51: 7: 8 | 0.556 | 0.680 |
| **Surfactant administered, ratio (no: yes)** | 34: 40 | 20: 38 | 38: 28 | **0.036** | 0.396 |
| **Patent ductus arteriosus treatment, ratio (no: yes)** | 71: 3 | 55: 3 | 59: 7 | 0.296 | 0.680 |
| **Surgery for necrotising enterocolitis, ratio (no: yes)** | 74: 0 | 55: 3 | 64: 2 | 0.104 | 0.458 |
| **Formula feeding, ratio (no: yes)** | 41: 33 | 30: 28 | 39: 27 | 0.712 | 0.746 |

| | | | | | |
|---|---|---|---|---|---|
| **Birth weight in grams, median (iqr)** | 1200 (620) | 1245 (478) | 1375 (505) | 0.0170 | 0.374 |
| **Feeding on maternal expressed breast milk, ratio (no: yes)** | 7: 67 | 9: 49 | 10: 56 | 0.497 | 0.680 |
| **Preeclampsia and induced hypertension, ratio (no: yes: unknown)** | 55: 4: 15 | 39: 7: 12 | 50: 8: 8 | 0.374 | 0.680 |
| **Placental abruption or antenatal hemorrhage, ratio (no: yes: unknown)** | 49: 15: 10 | 36: 10: 12 | 49: 8: 9 | 0.486 | 0.680 |

***Note.*** *Chi-squared test was used for categorical comparisons when subject count per cell was >5 and Fisher's Exact when cell count was 5 or less. Mann-Whitney non-parametric test was used for the only continuous variable: birth weight in grams.*

# CHAPTER 5 - Study #3: Elucidating brain-behavioural heterogeneity in VPT and FT children using data-driven consensus clustering

***Reference:*** *Hadaya, L., Váša, F., Kanel, D., Shi, L., Leoni, M., Dimitrakopoulou, K., Saqi, M, Edwards, A. D., Counsell, S. J., Leech, R., Batalle, D., & Nosarti, C. Exploring brain structure and function in clinical and data-driven groups of preterm and term children. [Manuscript in preparation].*

## 5.1 Abstract

**Importance**: Very preterm (VPT), compared to full-term (FT) born individuals, are at higher risk of developing behavioural difficulties and structural and functional brain alterations. Delineating biomarkers of behavioural sequelae could help guide intervention strategies, but the behavioural heterogeneity exhibited by both VPT and FT individuals complicates the ability to detect brain-behavioural associations. Data-driven approaches clustering individuals based on behavioural outcomes, regardless of clinical labels, may help identify homogenous subgroups with improved clinical translatability.

**Objective**: To identify structural and functional brain differences between children grouped according to i) clinical birth status (VPT vs FT) and ii) data-driven behavioural subgroups identified using consensus clustering (regardless of birth status).

**Design:** VPT children from the BIPP study (EudraCT 2009-011602-42) and FT controls underwent behavioural assessments and Magnetic Resonance Imaging.

**Setting**: Longitudinal cohort study.

**Participants**: 7–12-year-old convenience sample of 117 VPT and 56 FT children born at <33 and 37-42 weeks' gestation, respectively.

**Main Outcomes and Measures**: Structural volumes and intrinsic functional connectivity (FC) measured using Tensor Based Morphometry and Network Based Statistic, respectively.

**Results**: Relative to controls, VPT children displayed widespread volumetric alterations and increased FC in default mode, somatomotor, ventral attention, and language network areas. Grouping children into two data-driven behavioural subgroups identified a "*General Difficulties*" subgroup displaying widespread FC reductions and behavioural difficulties, relative to a "*General Resilience*" subgroup. Exploring three data-driven subgroups identified, a "*Neurodevelopmental Difficulties*" subgroup displaying socio-emotional and higher-order cognitive difficulties with reduced parahippocampal, rostro-lateral prefrontal, brainstem, occipital, and cerebellar volumes relative to a "*Typical Development*" subgroup, and a "*Psychiatric Difficulties*" subgroup exhibiting psychiatric and executive function difficulties with reduced dorsolateral prefrontal and cerebellar volumes relative to a "*Typical Development*" subgroup. Of the identified brain measures differentiating between data-driven behavioural subgroups, only brainstem, prefrontal, and FC alterations were significant after adjusting for birth status.

**Conclusions and Relevance**: While some brain-behavioural phenotypes involving cerebellar or occipital alterations may not arise independently of VPT birth, others involving unique FC patterns or prefrontal and brainstem volumes seem to be generalisable to a wider population of children independently of birth status.

**Key points:**

- **Question**: Are there differences in brain structure and function between children subdivided according to 1) clinical (VPT or FT birth) and 2) data-driven behavioural subgroup characterisations, irrespective of birth status?
- **Findings**: Some volumetric and functional alterations differentiating between VPT and FT children regionally overlapped with those differing between the distinct data-driven behavioural subgroups. However, unique brain-behavioural phenotypes involving specific functional and volumetric alterations localised to brainstem and prefrontal regions were present independently of clinical birth status.
- **Meaning**: Neural patterns differentiating between data-driven behavioural subgroups independently of birth status may represent generalisable biomarkers or neural mechanisms underlying specific behavioural phenotypes.

## 5.2 Introduction

Very preterm (VPT; < 33 weeks' gestation) birth occurs at a critical stage of gestation during which rapid neurobiological maturational processes are occurring. As a result, typical neurodevelopmental trajectories are interrupted and long-lasting structural and functional alterations arise in brain regions important for behavioural processing, including subcortical, prefrontal, occipital, and limbic areas (Stoecklein *et al.*, 2020; Dimitrova, Arulkumaran, *et al.*, 2021; França *et al.*, 2023; Ji *et al.*, 2023). Those born VPT, compared to their full-term (FT) born peers, are at an increased risk of developing behavioural sequelae relating to neurodevelopmental and psychiatric difficulties in cognitive and socio-emotional processing (Johnson and Marlow, 2011; Kroll *et al.*, 2017; P. J. Anderson *et al.*, 2021).

Detecting neurobiological alterations associated with specific behavioural sequelae in VPT populations could help guide early intervention strategies aiming to reduce behavioural risk in vulnerable individuals. However, neurodevelopmental trajectories in preterm samples are complex, making it challenging to delineate biomarkers predictive of behavioural sequelae. For instance, while the elevated behavioural risk in VPT populations may be explained by the observed

long-lasting structural and functional brain changes (Rogers *et al.*, 2012, 2014, 2017; Kanel *et al.*, 2021, 2022; Wheelock *et al.*, 2021); these alterations may also reflect neural adaptations supporting optimal outcomes (Schafer *et al.*, 2009; Scheinost *et al.*, 2015; Choi *et al.*, 2018; Wheelock *et al.*, 2018, 2021; Vanes *et al.*, 2021). Furthermore, behavioural profiles exhibited by VPT samples are not homogeneous. Whereby, distinct subsets of VPT children exhibit varying behavioural profiles characterised by either generalised behavioural difficulties, behavioural difficulties in specific behavioural subdomains, or no behavioural difficulties (Poehlmann *et al.*, 2015; Ross *et al.*, 2016; Johnson *et al.*, 2018; Burnett *et al.*, 2019; Lean *et al.*, 2020; van Houdt *et al.*, 2020; Bogičević *et al.*, 2021; Hadaya *et al.*, 2023). In fact, the observed behavioural heterogeneity is also shared with FT samples, with data-driven studies reporting subsets of both preterm and FT children co-clustering into the same behavioural subgroups despite their differences in clinical birth status (Johnson *et al.*, 2018; Burnett *et al.*, 2019; Lean *et al.*, 2020). Moreover, while some brain changes involving alterations to occipital or cerebellar regions were found to be specifically associated with behavioural outcomes in preterm samples and not in those born FT (Constable *et al.*, 2013; Rowlands *et al.*, 2016; Lean *et al.*, 2017; Wheelock *et al.*, 2018, 2021), other findings identified neurobiological correlates of behavioural outcomes involving somatomotor, default mode (DMN), and language networks which were exhibited by both preterm and FT individuals (Rowlands *et al.*, 2016; Wheelock *et al.*, 2021). Together these findings highlight the importance of departing from traditional methodological approaches, such as in case-control comparisons, which examine sets of individuals as comparable homogeneous groups depending on clinical labels (e.g., birth status).

Studies in psychiatric samples have applied advanced data-driven stratification approaches to delineate homogeneous subgroups of individuals exhibiting similar behavioural outcomes despite belonging to different clinical groups (e.g., Attention Deficit/Hyperactive, Autism Spectrum Conditions, and typical development), identifying unique brain-behavioural phenotypes characteristic of the distinct transdiagnostic subgroups with more effective clinical translatability and predictive validity (Stefanik *et al.*, 2018; Vaidya *et al.*, 2020; Jacobs *et al.*, 2021; Jung and Kim, 2023). Previous data-driven stratification studies in preterm children have identified specific neurobiological markers of behavioural heterogeneity (Ross *et al.*, 2016; Lean *et al.*, 2020; Bogičević *et al.*, 2021; Hadaya *et al.*, 2023); yet it remains to be investigated whether neural alterations characterising behavioural subgroups are present independently of birth status.

Here, we aim to delineate behavioural heterogeneity in a *"trans-clinical"* sample of both VPT and FT born children at 7–12 years and identify brain-behavioural associations present

independently of clinical birth status. First, to confirm previous findings reporting neurodevelopmental alterations in preterm children relative to controls (Degnan *et al.*, 2015a; Zhou *et al.*, 2018; P. J. Anderson *et al.*, 2021; Mossad *et al.*, 2022), we investigate brain and behavioural differences between VPT and FT clinical birth status groups. We then apply a rigorous consensus clustering stratification approach to delineate behavioural heterogeneity in a sample of VPT and FT children and explore whether structural volumetric and intrinsic functional connectivity (FC) patterns differ between the distinct behavioural subgroups. Finally, as current observations indicate a combined influence of neurobiological, clinical, and environmental factors on trajectories of behavioural sequelae in preterm samples (Wickremasinghe *et al.*, 2012; Ross *et al.*, 2016; Benavente-Fernández *et al.*, 2019; Lean *et al.*, 2020; Bogičević *et al.*, 2021; Vanes *et al.*, 2021, 2023; Hadaya *et al.*, 2023), we also investigate between-subgroup differences in clinical and environmental factors.

## 5.3 Methods

### 5.3.1 Study design

**Sample.** 511 VPT born infants recruited from neonatal intensive care units in London between 2010 and 2013 (EudraCT 2009-011602-42; Edwards et al, 2018) received multi-modal MRI at term-equivalent age (TEA), subsequent behavioural assessment in toddlerhood (2 years) (Hadaya *et al.*, 2022) and early childhood (4-7 years) (Hadaya *et al.*, 2023), and behavioural assessment and multi-modal MRI in middle childhood (7-12 years; median 9 years) (Leoni *et al.*, 2023). At the middle-childhood follow-up, a FT born (i.e., at 37-42 gestational weeks) control group of age-matched children attending mainstream schools, having no MRI contraindications (e.g., claustrophobia or metallic implants), and not meeting study exclusion criteria (i.e., having severe learning difficulties, moderate/severe cerebral palsy, blindness, or deafness impairments affecting capacity to complete assessments, a history of neurological conditions or head injury) was newly recruited from the community. Informed and written parental consent and child assent were obtained for all participants. Ethical approval was granted by Stanmore Research Ethics Committee (18/LO/0048) and London South East Research Ethics Committee (19/LO/1940).

**Socio-demographic and perinatal clinical measures.** Index of Multiple Deprivation 2019 (IMD; http://tools.npeu.ox.ac.uk/imd/) scores, measuring neighbourhood deprivation, were generated using participant residential postcode. For VPT participants, various clinical measures were collected at TEA from Standardised Electronic Neonatal Database medical records and used to generate a "*neonatal sickness*" index score as discussed in previous work (Kanel *et al.*, 2021; Hadaya

*et al.*, 2022, 2023). VPT neonatal brain lesion classifications (i.e., major lesions: periventricular leukomalacia, parenchymal haemorrhagic infarction, or other ischemic or haemorrhagic lesions; minor lesions: any other lesions; or no lesions) were rated by experienced perinatal neuroradiologists based on structural T2-weighted MRI scans at TEA.

**Behavioural measures.** Executive function was measured using the Behavior Rating Inventory of Executive Function, Second Edition (BRIEF-2) (Gioia *et al.*, 2000). Verbal comprehension, perceptual reasoning, working memory, and processing speed intelligence subdomains were quantified using the Weschler Intelligence Scale for Children, Fourth Edition (WISC-IV) (David Wechsler, 2012). Socio-emotional processing, temperament, psychopathology, anxiety, and autism traits were estimated using the emotion recognition task (ERT) (Montagne *et al.*, 2007), Emotion Regulation Checklist (ERC) (Shields and Cicchetti, 1997), Temperament in Middle Childhood Questionnaire (TMCQ) (Simonds and Rothbart, 2006), Strengths and Difficulties Questionnaire (SDQ) (Goodman, 2001); Social Responsiveness Scale, Second Edition (SRS-2) (Constantino and Gruber, 2012), and Spence Children's Anxiety Scale (SCAS) (Spence, 1998), respectively. Details on each of the behavioural measurements administered can be found in Table SM 5.1.

**Structural T1-weighted and functional MRI (f-MRI) data** were acquired at the Evelina Newborn Imaging Centre, Evelina London Children's Hospital (London, UK), with a dedicated neonatal and paediatric scanner (Philips 3T Achieva system) using a 32-channel head coil. To protect hearing, paediatric earplugs and noise-cancelling headphones were used. During f-MRI acquisition, children watched *'Inscapes'* (https://www.headspacestudios.org/inscapes), a low-demand, non-verbal, and non-social movie paradigm featuring moving abstract shapes, which was developed to decrease head motion, improve wakefulness compliance, and increase the ability of capturing intrinsic non-evoked FC patterns (Vanderwal *et al.*, 2015). See Supplemental Information for acquisition parameters, image pre-processing, and motion correction protocols followed.

Following Tensor Based Morphometry approaches previously described in (Hadaya *et al.*, 2022, 2023) and Supplemental Information, we calculated voxel-wise log-Jacobian determinant maps to quantify structural brain volumes, whereby larger values indicate greater regional brain volumes relative to the rest of the brain (Avants and Gee, 2004). FC matrices were computed from f-MRI data parcellated into 358 cortical and 16 subcortical bilateral regions (Fischl, 2012; Glasser *et al.*, 2016). Pearson's correlation coefficients between the time-series of each pair of regions were calculated, followed by exclusions of weak correlations (i.e., retaining edge coefficients with r

values ≥ 0.2), and Fisher Z-transformations (Buckner *et al.*, 2009; Zalesky *et al.*, 2016; Fenn-Moltu *et al.*, 2022).

### 5.3.2 Consensus clustering

Behavioural measures from 153 participants (Table SM 5.1; Table SM 5.2; Figure SM 5.1) were used as input features in a rigorous consensus clustering pipeline adapted from (Wilkerson and Hayes, 2010; Wang *et al.*, 2014; Hadaya *et al.*, 2023), as described in (Hadaya *et al.*, 2024).

In summary, behavioural data were pre-processed (details in Supplemental Information) and standardised before being converted into a Euclidean distance matrix and then a similarity matrix which was clustered using a spectral clustering algorithm. The similarity matrix generation and clustering steps were repeated thirty times using different nearest neighbour (K) and edge weighting (alpha) hyperparameter combinations (Wang *et al.*, 2014, 2018), and retaining the clustering result with the highest silhouette width average score (Hadaya *et al.*, 2023, 2024). A bootstrap approach iteratively repeated this process 1,000 times using a random subsample (i.e., 80% of the sample) each time to counteract for possible effects of over-fitting. A final consensus clustering outcome was estimated from the 1000 bootstrap outputs (Chiu and Talhouk, 2018). Eigen-gap, Rotation Cost, consensus matrices, and Silhouette scores were used to determine C=2 and C=3 as the two most optimal numbers of clusters (see Supplemental Information and Figure SM 5.2 for more details).

### 5.3.3 Statistical analyses

**Between-group differences in behavioural, clinical, and sociodemographic measures**. Using False Discovery Rate (FDR) corrections, we investigated between-group differences in non-MRI continuous (nonparametric Wilcox Rank-Sum T-test and Kruskal Wallis) and categorical (Chi-squared or Fischer's Exact) variables. Nonparametric 5,000 permutation testing p-values adjusting for covariates (age, sex, and IMD) are also reported (França, Ge and Batalle, 2022). Sensitivity analyses explored whether VPT children included in our analyses (n=117) differed from those excluded from our analyses (n=41) (for reasons summarised in Figure SM 5.1). Post-hoc exploratory analyses investigated whether perinatal clinical or socio-demographic measures differentiated between VPT children belonging to the distinct data-driven subgroups, and whether FT children exhibited socio-demographic between-subgroup differences.

**Between-group differences in structural brain volumes**. Family Wise Error Rate (FWER) and Threshold Free Cluster Enhancement (TFCE) mass-univariate non-parametric

10,000 permutation testing adjusting for covariates (age, sex, and IMD), detected significant clusters of neighbouring voxels, using p-FWER-TFCE < 0.05 per contrast (Smith and Nichols, 2009; Jenkinson *et al.*, 2012).

**Between-group differences in FC** were explored using Network Based Statistic (NBS), a method which implements edgewise mass-univariate testing and subsequent breadth-first search approaches to identify significant NBS components, only using edges exceeding a pre-defined threshold (Zalesky, Fornito and Bullmore, 2010). To establish component significance (p-FWER<0.05 per contrast), we ran 1,000 permutation tests adjusting for covariates age, sex, IMD, and FD at three p-NBS-Thresholds (0.05, 0.01, and 0.001), using the NBR R package (Gracia-Tabuenca and Alcauter, 2020), as described in our previous work (Hadaya *et al.*, 2024). To identify component "hub" regions with a high degree of connectivity, we calculated the percentages of edges connected to each node as a proportion of total component edges. We also labelled component cortical nodes according to previously defined visual, somatomotor, ventral attention (VAN), dorsal attention, limbic, frontoparietal, and DMN intrinsic connectivity networks (Yeo *et al.*, 2011), considered the 16 subcortical nodes as an eighth subcortical network, and measured within- and between-network connectivity degree and strength using publicly available code accessible here: https://github.com/frantisekvasa/functional_network_development/blob/master/nspn.fmri.R (Váša *et al.*, 2020). Sørensen-Dice coefficients measured nodal and edgewise similarity between NBS components (Sørensen, 1948).

To identify significant effects present independently of birth status, sensitivity analyses investigated structural and FC neuroimaging features differentiating between data-driven behavioural subgroups after adjusting for birth status as an additional covariate. Post-hoc two-way group interaction analyses investigated whether structural and FC differences between VPT and FT children varied according to data-driven behavioural subgrouping.

## 5.4 Results

### 5.4.1 Sample characteristics and behavioural outcomes

VPT and FT sample characteristics are summarised in Table 5.1. The VPT, compared to the FT group, had poorer IQ, effortful control, autism traits, and externalising symptom scores, but no between-group differences in other behavioural measures examined (Table 5.2; Table SM 5.3; Figure 5.1A) or in in-scanner head motion (Table SM 5.4). No differences in clinical measures or

IMD were found between VPT children included and those excluded from our analyses; however, those included exhibited relatively more optimal socio-emotional and cognitive behavioural outcomes (Table SM 5.5).

***Table 5.1. Socio-demographic and clinical measures in VPT and FT groups.***

| | VPT | FT | p-value |
|---|---|---|---|
| **GA at birth, weeks** | 29.86 (4.00) | 40.14 (1.50) | <0.001 |
| **Sex, n (%)** | | | 0.519 |
| *Male* | 62 (52.99%) | 26 (46.43%) | |
| *Female* | 55 (47.01%) | 30 (53.57%) | |
| **Age at assessment, years** | 9.25 (1.17) | 8.87 (0.94) | 0.007 |
| **IMD rank** | 19575.00 (14506.00) | 21397.00 (18256.75) | 0.686 |
| **Ethnicity, n (%)** | | | 0.012 |
| *Asian/Asian British* | 14 (11.97%) | 2 (3.57%) | |
| *Black/African/Caribbean/Black British* | 9 (7.69%) | 0 (0.00%) | |
| *Mixed/Multiple ethnic groups* | 16 (13.68%) | 7 (12.50%) | |
| *White* | 59 (50.43%) | 42 (75%) | |
| *Other* | 2 (1.71%) | 0 (0.00%) | |
| **Neonatal brain lesions, n (%)** | | | n/a |
| *No lesions* | 42 (35.90%) | n/a | |
| *Minor lesions* | 68 (58.12%) | n/a | |
| **Total, n** | 117 | 56 | |

***Note.*** *Median (interquartile range) are reported unless stated otherwise, where number of participants (n) is reported alongside percentage (%). Missing data: 5 FT children and 18 VPT children had missing ethnicity data; 7 VPT children had missing neonatal brain lesion data.*

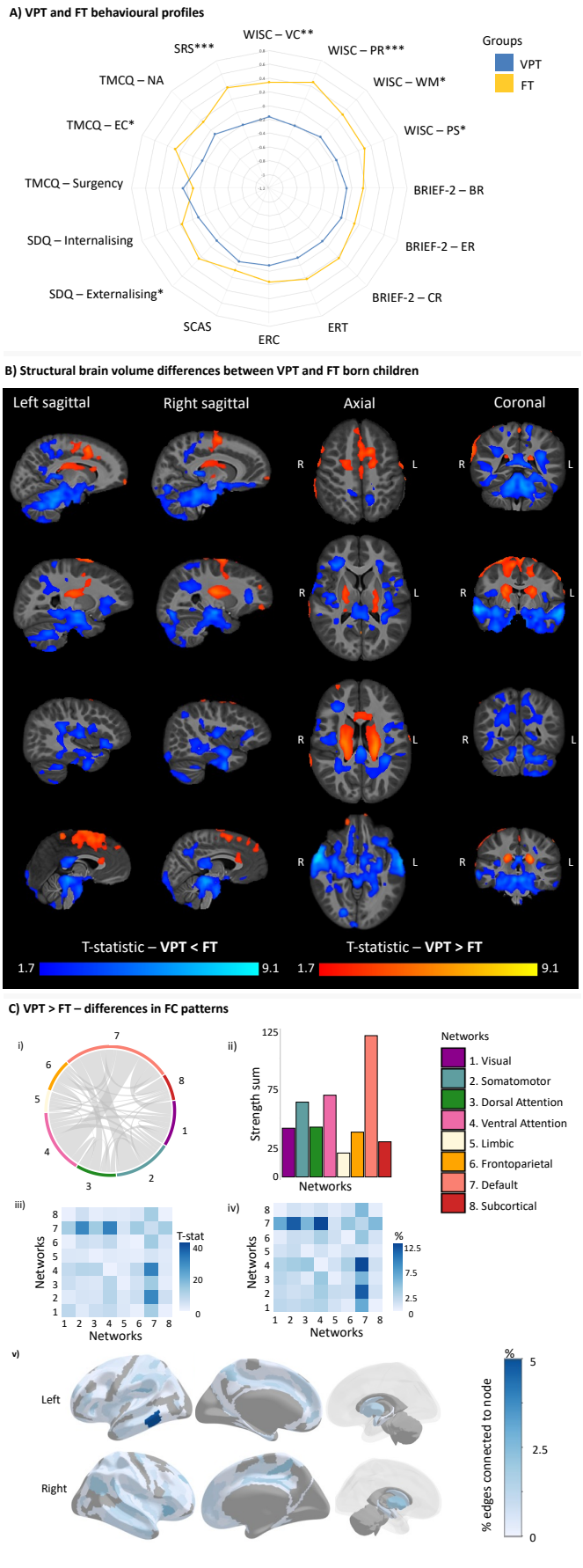



***Figure 5.1. Behavioural, structural brain volume, and FC pattern differences between VPT and FT born children, correcting for age, sex, IMD, (and FD in C).***

*A) Radar plots demonstrating VPT (in blue) and FT (in yellow) children's behavioural profiles using Z-scores; whereby, higher Z-scores reflect more optimal outcomes. To ease visual interpretability, Z-scores were inverted where necessary (i.e., for BRIEF-2, ERC, SCAS, SDQ, SRS-2, and TMCQ–NA). Measures with significant between-group differences are marked accordingly: ***=p-FDR<0.001, **=p-FDR<0.01, *=p-FDR<0.05. B) T-statistic values for voxels with significantly smaller (in blue) or larger (in red) log-Jacobian values (i.e., relative brain volumes) in the VPT relative to the FT group at p-TFCE-FWER<0.05 per contrast. Left (L) and right (R) hemisphere orientations are labelled accordingly. C) Significant NBS component at 0.01 p-NBS-Threshold comprised of 324 nodes (86.63% of regions) and 709 edges (0.01% of all 69,751 possible connections), p-FWER=0.041, T-statistic=223.96. Plots showing NBS component i) within- and between-network connections, ii) total connectivity strength for each intrinsic network, iii) strength and iv) proportion of within- and between-network connectivity, v) percentage of edges connected to each node; regions plotted in grey were not part of the component.*

## 5.4.2 Data-driven behavioural subgroup characteristics and behavioural outcomes

### *5.4.2.1 Two-subgroup solution*

A **"*General Resilience*" subgroup** of children with more optimal behavioural outcomes, relative to a **"*General Difficulties*" subgroup** in all behavioural domains measured, except working memory, surgency, and emotion recognition (Table SM 5.6; Figure 5.2A), and less in-scanner head motion (Table SM 5.4); with VPT children forming 55.58% and 73.85% of either subgroup, respectively (Figure 5.3).

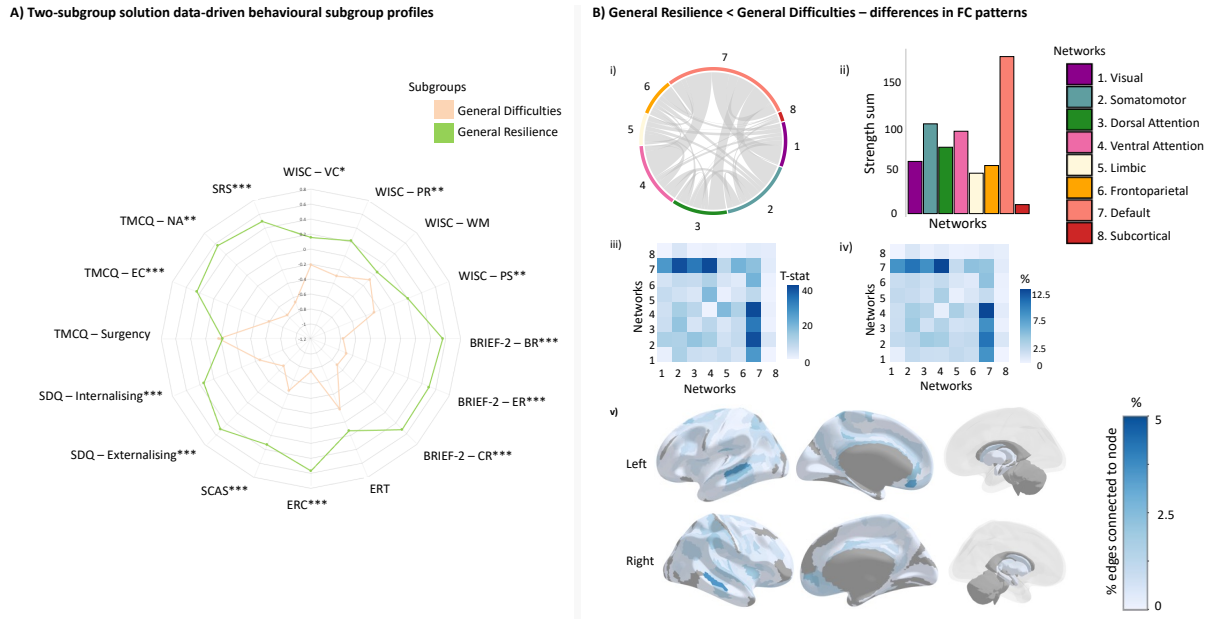



***Figure 5.2. Behavioural and FC pattern differences between the two-subgroup solution data-driven behavioural subgroups, correcting for age, sex, IMD, (and FD in C).***

*A) Radar plots demonstrating General Difficulties (in beige) and General Resilience (in green) subgroup behavioural profiles using Z-scores; whereby, higher Z-scores reflect more optimal outcomes. To ease visual interpretability, Z-scores were inverted where necessary (i.e., for BRIEF-2, ERC, SCAS, SDQ, SRS-2, and TMCQ–NA). Measures with significant between-subgroup differences are marked accordingly: ***=p-FDR<0.001, **=p-FDR<0.01, *=p-FDR<0.05. C) Significant NBS component at 0.01 p-NBS-Threshold comprised of 346 nodes (92.51% of regions) and 1037 edges (1.49% of all 69,751 possible connections); p-FWER=0.028; T-statistic=329.70. Plots showing NBS component i) within- and between-network connections, ii) total connectivity strength for each intrinsic network, iii) strength and iv) proportion of within- and between-network connectivity, v) percentage of edges connected to each node; regions plotted in grey were not part of the component.*

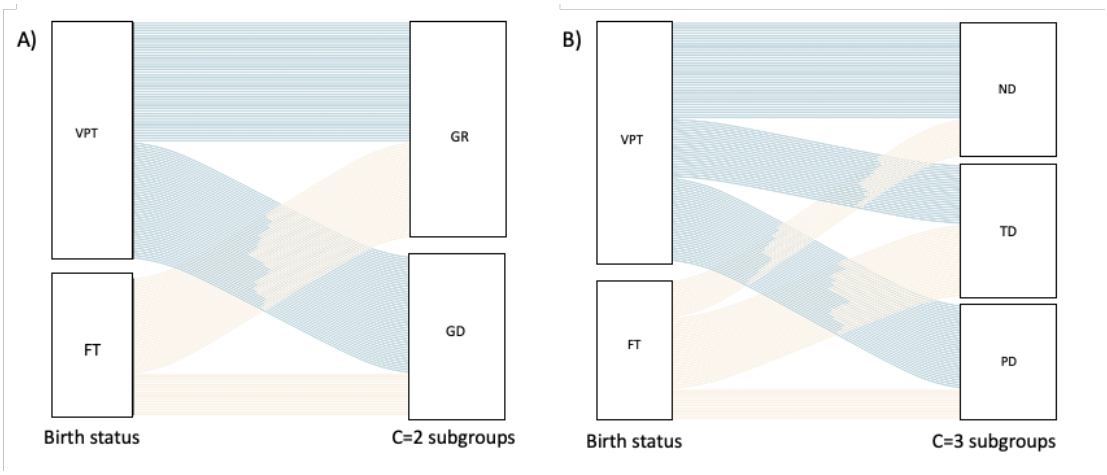



***Figure 5.3. Alluvial plot showing VPT (in blue) and FT (in beige) children clustering into distinct data-driven behavioural subgroups based on the A) two-subgroup solution (C=2) and the B) three-subgroup solution (C=3).***

*GR = General Resilience, GD = General Difficulties, ND = Neurodevelopmental Difficulties, TD = Typical Development, and PD = Psychiatric Difficulties.*

### *5.4.2.2 Three-subgroup solution*

The C=3 solution identified nuanced behavioural subgroups with stronger effect sizes than C=2 (Table SM 5.7; Figure 5.4A; Figure SM 5.3), characterising a **"*Psychiatric Difficulties*" (PD) subgroup,** with the highest scores on measures of autism traits, anxiety, internalising and externalising symptoms, executive function, emotion regulation, negative affectivity, and effortful control; a **"*Typically Developing*" (TD) subgroup**, with the most optimal intelligence, emotion recognition, autism traits, and effortful control and negative affectivity temperament outcomes, and a **"*Neurodevelopmental Difficulties*" (ND) subgroup**, showing more difficulties in higher-level working memory and perceptual reasoning intelligence subdomains, relative to the other two subgroups. The ND subgroup displayed further difficulties in verbal comprehension and processing speed intelligence subdomains, emotion recognition, and surgency (exacerbated relative to TD, but comparable to PD), and optimal executive function, emotion regulation, anxiety, and psychopathology outcomes (more optimal than PD, but comparable to TD) (Figure SM 5.4). VPT children formed 74.46%, 46.15%, and 66.67% of each subgroup, respectively. No between-subgroup differences in head-motion were seen (Table SM 5.4).

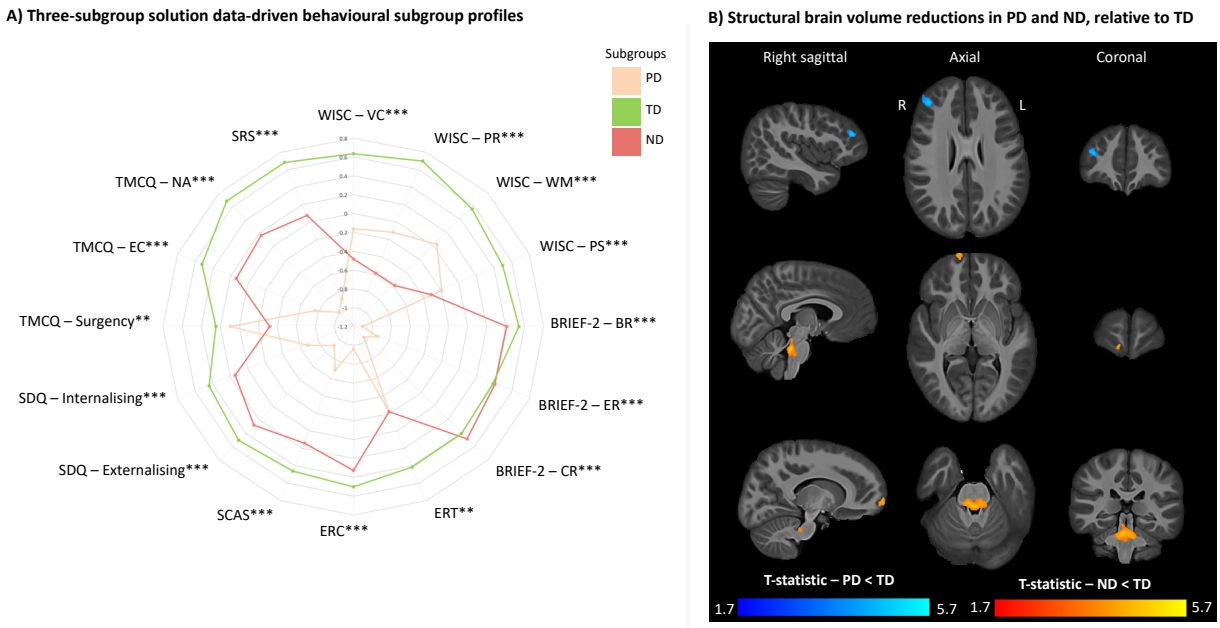



***Figure 5.4. Behavioural and structural brain volume differences between the three-subgroup solution data-driven subgroups, correcting for age, sex, IMD, (and birth status in B).***

*A) Radar plots demonstrating Psychiatric Difficulties (PD; in beige), Typical Development (TD; in green), and Neurodevelopmental Difficulties (ND; in pink) subgroup behavioural profiles using Z-scores; whereby, higher Z-scores reflect more optimal outcomes. To ease visual interpretability, Z-scores were inverted where necessary (i.e., for BRIEF-2, ERC, SCAS, SDQ, SRS-2, and TMCQ–NA). Measures with significant between-subgroup differences are marked accordingly: ***=p-FDR<0.001, **=p-FDR<0.01, *=p-FDR<0.05. B) T-statistic values for voxels with significantly smaller log-Jacobian values (i.e., relative brain volumes) in the PD (in blue) and ND (in red) subgroups relative to the TD subgroup, at p-TFCE-FWER<0.05 per contrast, after correcting for birth status. Left (L) and right (R) hemisphere orientations are labelled accordingly.*

### 5.4.3 Post-hoc exploratory analyses

IMD significantly differed between VPT children in the C=3 solution. VPT children in the ND subgroup had the highest levels of deprivation, followed by PD, and TD respectively (Table SM 5.8). No other clinical or socio-demographic measures examined differentiated between VPT children (Table SM 5.8), or FT children (Table SM 5.9) in the C=2, or C=3 subgroups.

### 5.4.4 Structural and functional alterations

#### *5.4.4.1 VPT vs FT*

**Structural alterations in VPT vs FT children.** Relative to FT born peers, VPT children displayed widespread volume reductions spanning temporal, amygdala, hippocampus, operculum, insula, cerebellum, brainstem, orbitofrontal, parietal, occipital, and corpus callosum splenium regions, and larger brain volumes localised in ventricular, frontal, cingulate, somatomotor, corpus callosum genu, and internal capsule posterior limb regions (Figure 5.1B).

**FC alterations in VPT > FT.** An NBS component significant at 0.01 (but not 0.05 or 0.001) p-NBS-Threshold, comprised of 324 nodes (86.63% of regions) and 709 edges (0.01% of all 69,751 possible connections) (p-FWER=0.041; T-statistic=223.96), displaying greater FC (i.e., hyperconnectivity) in the VPT children relative to controls was also identified. Regions of the DMN, VAN, and somatomotor networks, and a component 'hub' region localised to the left lateral inferior and middle temporal cortex and regions including insular, cingulate, prefrontal, perisylvian language, auditory association, and inferior frontal areas exhibited the highest levels of connectivity within the component (Figure 5.1C) (Table SM 5.10).

#### *5.4.4.2 Two-subgroup solution*

Structural brain volumes did not significantly differ between the C=2 subgroups (p-TFCE-FWER>0.05).

**FC alterations in General Resilience < General Difficulties**. Lower FC (i.e., hypoconnectivity) in the General Difficulties relative to the General Resilience subgroup was identified in an NBS component significant at 0.01 p-NBS-Threshold (but not 0.05 or 0.001), comprised of 346 nodes (92.51% of regions) and 1037 edges (1.49% of all 69,751 possible connections) (p-FWER=0.028; T-statistic=329.70), including areas of the DMN, VAN, and somatomotor networks (Figure 5.2B) and component hub regions in lateral temporal (auditory association), insula, and prefrontal areas (Table SM 5.11). Sensitivity analyses adjusting for birth status identified significant components (at 0.01 and 0.05 p-NBS-Thresholds) exhibiting similar connectivity patterns (see Figures SM 5.5 and SM 5.6 and Supplemental Information).

Post-hoc analyses report strong nodal (n=304; Sørensen-Dice=0.91), but poor edgewise similarity (n=36; Sørensen-Dice=0.04) between the behavioural (*General Resilience < General Difficulties)* and clinical (*VPT > FT*) NBS components, and no significant two-way group interactions (i.e., C=2 and birth status) (p-FWER>0.05).

#### *5.4.4.3 Three-subgroup solution*

**Structural alterations in PD < TD and ND < TD.** Relative to the TD subgroup, the PD displayed smaller right dorsolateral prefrontal and cerebellar volumes, and the ND exhibited localised reductions in brainstem, cerebellum, parahippocampal, and right occipital regions (Figure SM 5.7). After adjusting for birth status, *PD < TD* effects in the right dorsolateral prefrontal cortex, and *ND < TD* effects in a ventral region of the brainstem pons remained significant, and a new *ND < TD* effect in the right rostro-lateral prefrontal cortex was observed (Figure 5.4B).

No significant volumetric differences between ND and PD subgroups (p-FWER >0.05), two-way group (i.e., C=3 and birth status) interaction effects (p-FWER >0.05), or significant NBS components were identified (p-TFCE-FWER>0.05).

## 5.5 Discussion

Here we identify widespread volumetric changes spanning temporal, parietal, frontal, cerebellar, and brainstem regions, in addition to functional hyperconnectivity comprising regions of the DMN, VAN, and somatomotor networks, and language processing areas such as the middle and inferior temporal gyri (Braga *et al.*, 2020; Briggs *et al.*, 2021; Binding *et al.*, 2022); confirming previous findings reporting structural alterations (Kesler *et al.*, 2006; Lax *et al.*, 2013; Nosarti *et al.*, 2014; Lean *et al.*, 2017; Lemola *et al.*, 2017; Zhou *et al.*, 2018; Kvanta *et al.*, 2023) and hyperconnectivity patterns (Wilke *et al.*, 2014; Degnan *et al.*, 2015a; Wehrle *et al.*, 2018; Cho *et al.*, 2022; Mossad *et al.*, 2022) in VPT children and adolescents relative to FT born controls. While these neural alterations have been studied as potential mechanisms underlying the behavioural profiles observed in VPT samples (Johnson and Marlow, 2011; Kroll *et al.*, 2017; P. J. Anderson *et al.*, 2021), such as social and emotional processing, mental health and cognitive difficulties reported here (Skranes *et al.*, 2007; Nosarti *et al.*, 2008; Mossad *et al.*, 2022; Hadaya *et al.*, 2023); it is important to note that these alterations have also been studied as mechanisms potentially supporting optimal attention and inhibition control (Wheelock *et al.*, 2021), balance (Wheelock *et al.*, 2018), language (Scheinost *et al.*, 2015; Choi *et al.*, 2018), and cognitive abilities (Schafer *et al.*, 2009; Salvan *et al.*, 2014). The neural alterations observed here may, therefore, also reflect neural adaptations maintaining optimal executive function abilities, which were comparable across VPT and FT children in our sample and in previous studies reporting functional hyperconnectivity in late prematurity (Degnan *et al.*, 2015a). However, studies also identified overlapping brain changes underlying behavioural outcomes in both VPT and FT individuals (Rowlands *et al.*, 2016; Wheelock *et al.*, 2021), making it challenging to delineate preterm-specific brain-behavioural phenotypes. This task is further complicated by the behavioural heterogeneity seen within and between both VPT and FT samples; whereby, specific subsets of both FT and VPT children were co-clustering into subgroups displaying similar behavioural profiles of either elevated levels of difficulties or optimal outcomes (across both C=2 and C=3 data-driven solutions), similar to previous reports (Johnson *et al.*, 2018; Burnett *et al.*, 2019; Lean *et al.*, 2020; Hadaya *et al.*, 2023). Expanding on these findings, we identified specific neuroimaging features associated with distinct behavioural subgroups, independently of birth status; echoing our previous study in adults (Hadaya *et al.*, 2024).

Our C=2 solution identified a "*General Difficulties"* subgroup with suboptimal cognitive, socio-emotional, and psychiatric outcomes, also displaying functional hypoconnectivity anchored in regions of the DMN and VAN, relative to a "*General Resilience"* subgroup showing no behavioural difficulties. We speculate that the observed functional hypoconnectivity could potentially represent neural mechanisms underlying children's cognitive and behavioural difficulties, irrespective of their clinical birth status. Supporting this hypothesis, previous studies found that altered functional DMN connectivity in children and adolescents was associated with unidimensional measures of generalised psychopathology (Karcher *et al.*, 2021; Hong, Hwang and Lee, 2023) and was characteristic of transdiagnostic cognitive deficits (Bathelt *et al.*, 2018; Tong *et al.*, 2022). Furthermore, component hub nodes in superior temporal, insular, and cingulate areas of the VAN have also been implicated in neural processes supporting intelligence (Hebling Vieira *et al.*, 2021), scholastic performance (Chaddock-Heyman *et al.*, 2018), anxiety, and depression in children and adolescents (Sylvester *et al.*, 2013; Pannekoek *et al.*, 2014). Notably, these alterations differentiated between the behavioural subgroups, independently of clinical birth status, and were previously reported to show associations with attention and intelligence outcomes in both VPT and FT children (Rowlands *et al.*, 2016; Wheelock *et al.*, 2021). Together, this evidence suggests that while VPT birth can result in preterm-specific neurobiological alterations, distinct neural signatures may underlie behavioural outcomes in the general population, irrespective of clinical birth status. Moreover, despite the observation that the networks (e.g., DMN, VAN, somatomotor) and regions (e.g., temporal cortex, insula) which differed between the C=2 behavioural subgroups overlapped with those differing between VPT and FT children, we found that specific region-to-region connections contributing to the two FC components highly differed. Given the complexity of human behaviours, their neural mechanisms, and the influence of environmental, biological, and contextual factors on their interaction, it is not surprising that individual brain regions and networks could be involved in multiple neural patterns underlying distinct behavioural outcomes; whereby the communication between certain combinations of regions and networks is more important in mediating specific behavioural outcomes than the involvement of the regions and networks themselves (Kragel and LaBar, 2016; Westlin *et al.*, 2023).

Stratifying our sample into C=3 subgroups, we identified nuanced brain-behavioural phenotypes in middle childhood, that were also independent of birth status. Firstly, a "*PD subgroup*" exhibiting elevated socio-emotional processing difficulties, autistic traits, anxiety, internalising and externalising psychopathology, and lower intelligence scores, displayed volumetric reductions in the right dorsolateral prefrontal cortex relative to a "*TD subgroup*" exhibiting optimal outcomes. Considering these phenotypic characteristics, behavioural difficulties in the PD subgroup are likely

exacerbated or underpinned by the marked deficits in executive function skills, given that executive functions (Green, Johnson and Bretherton, 2014; Mareva, CALM team and Holmes, 2019; Benallie *et al.*, 2021) and the dorsolateral prefrontal cortex (Jones and Graff-Radford, 2021) are critical players in regulating interpersonal socio-communication, learning, and affective skills.

Secondly, an "*ND subgroup*" exhibiting cognitive and socio-emotional difficulties, displayed volumetric reductions in the right rostro-lateral prefrontal cortex and bilateral brainstem (pons) relative to the TD subgroup. We speculate that the marked deficits in working memory and abstract perceptual reasoning by the ND subgroup may be driving their socio-emotional difficulties. In fact, specific aspects of working memory and perceptual reasoning, such as information processing, storage, and inference generation, work in tandem to support adequate "pragmatic use" of language and generate appropriate behavioural responses to the complex, multi-modal, and abstract stimuli faced in emotional and social contexts (Bishop, 2000; Solomon, Buaminger and Rogers, 2011; Opitz *et al.*, 2014; Imanipour *et al.*, 2021). Moreover, the structurally altered rostro-lateral prefrontal cortex is a high dendritic density neo-cortical region acting as an integrative hub for higher-order subgoal cognitive processes (including working memory and perceptual reasoning) working together in pursuit of complex behavioural functions, such as socio-emotional processing (Braver and Bongiolatti, 2002; Amati and Shallice, 2007; Dumontheil, Burgess and Blakemore, 2008; Dumontheil, 2014; Hornick and Shetreet, 2022). Furthermore, the observed structural alterations in the brainstem overlap with cranial nerve nuclei and reticular formation subnuclei areas, which are involved in primitive mechanisms required for optimal socio-emotional functioning (Mangold and M Das, 2023), such as internal (e.g., somatomotor, autonomic, or visceral) and external (e.g., visual or auditory) signal integration and saliency filtering (Mangold and M Das, 2023). Further supporting this notion, increased connectivity between prefrontal and brainstem regions, stimulated by low-intensity exercise or deep-brain stimulation, has been reported to improve social cognition, pragmatic use of language, and increased feelings of well-being (Mazzone *et al.*, 2005; Zanini *et al.*, 2009; Dietrich and Audiffren, 2011; Ludyga, Ishihara and Kamijo, 2022). Therefore, suggesting that these hypoconnectivity patterns could act as biomarkers of cognitive deficits driving socio-emotional difficulties in children regardless of birth status.

Our findings also identified neurobiological alterations which were only present before adjusting for birth status. Relative to the TD subgroup, PD reductions were localised to the cerebellum, and ND volumetric reductions to occipital, cerebellum, and parahippocampal regions, suggesting that these regions respectively underlie PD and ND behavioural outcomes in VPT

samples specifically, and not across both birth status groups. Previous studies have also identified a VPT-specific preferential reliance on visual networks for attention processing (Lean *et al.*, 2017; Wheelock *et al.*, 2021), and cerebellar connectivity to support linguistic and motor abilities (Constable *et al.*, 2013; Wheelock *et al.*, 2018), relative to FT controls. In conclusion, our results collectively indicate that some brain-behavioural phenotypes may be preterm-specific, while others could be generalised to the whole childhood population, independently of birth status.

While it remains to be elucidated whether VPT and FT born individuals with similar brain-behavioural phenotypes have shared or distinct underlying risk factors, here we found that elevated neighbourhood deprivation was associated with ND behavioural outcomes for VPT, but not FT born children; supporting previous findings showing an elevated susceptibility towards developing behavioural difficulties in vulnerable populations, such as VPT individuals, exposed to adverse environments (Belsky and Pluess, 2009; Lean *et al.*, 2020; Vanes *et al.*, 2021; Hadaya *et al.*, 2023). However, additional factors (e.g., genetics, parenting, immunity, gut health, or nutrition) need to be explored in future studies in order to better characterise aetiological trajectories probing developmental risk or resilience (Dinan and Cryan, 2016; Pariante, 2016). It is also unclear why we found FC differences between the C=2 but not C=3 subgroups, despite C=3 showing stronger between-subgroup structural and behavioural effects. On one hand, this may be because neural connectivity patterns are not yet fully specialised at this age, during which rapid developmental changes occur (Chai *et al.*, 2014; Fuhrmann *et al.*, 2020; López-Vicente *et al.*, 2021). Alternatively, it is also possible that the smaller sample sizes in the C=3 comparisons were sufficiently powered to detect statistically significant effects in structural, but not FC analyses which are usually less powered (Button *et al.*, 2013).

## 5.6 Supplemental Information

### 5.6.1 Structural and f-MRI acquisition parameters details

Structural brain scans (3D MPRAGE T1-weighted images) were acquired using the following parameters: TR= 7.9ms, ET= 3.6ms, TI= 900ms; flip angle= 8°, field of view = 240x220x160mm$^3$, voxel size= 1mm isotropic, SENSE factor of 1.5 along the first phase encoding direction and 2 along the second direction. Children were watching a show of their choice as structural images were being acquired.

Multi-slice gradient echo EPI f-MRI was acquired using the following parameters: 900 volumes, TR=1160ms, TE=33ms, flip angle = 60 degrees, acquisition matrix = 88 x 87 mm, acquisition voxel size = 2.5 x 2.5 x 2.5 mm$^3$, reconstruction voxel size = 1.9 x 1.9 x 2.5 mm$^3$, FOV = 220 x 220 x 35 mm$^3$, multiband = 4; while children watched the '*Inscapes'* low-cognitive load and non-narrative movie paradigm (https://www.headspacestudios.org/inscapes) (Vanderwal *et al.*, 2015).

### 5.6.2 Tensor Based Morphometry

The Advanced Normalization Tools (ANTS) software Greedy Symmetric Normalisation (SyN) algorithm (Avants *et al.*, 2011) was applied to build a study-specific template using T1-weighted images from a subset of 87 participants (N=47 FT controls and 40 VPT children). T1-weighted scans were then registered to the study-specific template using the ANTS SyN algorithms to generate deformation tensor maps which were used to compute deformation tensor field gradients (i.e., log-Jacobian determinant maps). Computed maps were smoothed with a 4mm full-width half-maximum Gaussian filter.

### 5.6.3 f-MRI data pre-processing

The standardised fMRIPrep (20.1.1, RRID:SCR_016216) (Esteban *et al.*, 2019) rs-fMRI pre-processing pipeline was used to implement skull stripping, slice-time correction, boundary-based registration, and head motion estimation steps. After implementing fMRIPrep on f-MRI data, estimated measures of head motion (global signal and six motion parameters: three translation and three rotation parameters) were regressed out using the FMRIB Software Library (FSL) *fsl_regfilt* command (Jenkinson *et al.*, 2012) and AFNI software *3dBandpass* command (Cox, 1996) was used to apply a bandpass filter (0.01 – 0.1 Hz).

The Human Connectome Project Multi-Modal Parcellation; HCP-MMP (v1) (Glasser *et al.*, 2016) and subcortical FreeSurfer (Fischl, 2012) atlases were used to parcellate f-MRI data into 358 cortical and 16 subcortical bilateral regions, respectively. As both atlases include a hippocampal region, the FreeSurfer segmentations of this region were used and HCP-MMP hippocampal parcellations were discarded. Participants with excess in-scanner head motion (characterised by mean framewise displacement (FD) exceeding 0.15mm or maximum FD of 1mm), major perinatal brain lesions, or poor functional and anatomical scan alignment were excluded from further analyses (Figure SM 5.1).

The full description of the anatomical and functional data fMRIPrep pre-processing pipeline found below in italic grey text is extracted from the boilerplate automatically generated by fMRIPrep (released under the CC0 license):

*"Results included in this manuscript come from preprocessing performed using fMRIPrep 20.1.1* (Esteban *et al.*, 2019)*, (RRID:SCR_016216), which is based on Nipype 1.5.0* (Gorgolewski *et al.*, 2011)*, (RRID:SCR_002502).*

***Anatomical data pre-processing.*** *The T1-weighted (T1w) image was corrected for intensity non-uniformity (INU) with N4BiasFieldCorrection* (Tustison *et al.*, 2010)*, distributed with ANTs 2.2.0* (Avants *et al.*, 2008b)*, (RRID:SCR_004757), and used as T1w-reference throughout the workflow. The T1w-reference was then skull-stripped with a Nipype implementation of the antsBrainExtraction.sh workflow (from ANTs), using OASIS30ANTs as target template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using fast (FSL 5.0.9, RRID:SCR_002823) (Zhang, Brady and Smith, 2001). Brain surfaces were reconstructed using recon-all FreeSurfer 6.0.1, (RRID:SCR_001847), (Dale, Fischl and Sereno, 1999), and the brain mask estimated previously was refined with a custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived segmentations of the cortical gray-matter of Mindboggle (RRID:SCR_002438) (Klein et al., 2017). Volume-based spatial normalization to three standard spaces (MNIPediatricAsym:cohort-3, MyCustom, MNI152NLin2009cAsym) was performed through nonlinear registration with antsRegistration (ANTs 2.2.0), using brain-extracted versions of both T1w reference and the T1w template. The following templates were selected for spatial normalization: MNI's unbiased standard MRI template for pediatric data from the 4.5 to 18.5y age range [RRID:SCR_008796; TemplateFlow ID: MNIPediatricAsym:cohort-3], Study specific child template [TemplateFlow ID: MyCustom], ICBM 152 Nonlinear Asymmetrical template version 2009c [(Fonov et al., 2011), RRID:SCR_008796; TemplateFlow ID: MNI152NLin2009cAsym],*

***Functional data preprocessing***. *First, a reference volume and its skull-stripped version were generated using a custom methodology of fMRIPrep. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) are estimated before any spatiotemporal filtering using mcflirt (FSL 5.0.9) (Jenkinson et al., 2002). Susceptibility distortion correction (SDC) was omitted. The BOLD reference was then co-registered to the T1w reference using bbregister (FreeSurfer) which implements boundary-based registration (Greve and Fischl, 2009). Co-registration was configured with six degrees of freedom. The BOLD time-series (including slice-timing correction when applied) were resampled onto their original, native space by applying the transforms to correct for head-motion. These resampled BOLD time-series will be referred to as preprocessed BOLD in original space, or just preprocessed BOLD. The BOLD time-series were resampled into several standard spaces, correspondingly generating the following spatially-normalized, preprocessed BOLD runs: MNIPediatricAsym:cohort-3, MyCustom. First, a reference volume and its skull-stripped version were generated using a custom methodology of fMRIPrep. Several confounding time-series were calculated based on the preprocessed BOLD: framewise displacement (FD), DVARS and three region-wise global signals. FD was computed using two formulations following Power (absolute sum of relative motions, (Power et al., 2014) and Jenkinson (relative root mean square displacement between affines, (Jenkinson et al., 2002)). FD and DVARS are calculated for each functional run, both using their implementations in Nipype (following the definitions by (Power et al., 2014)). The three global signals are extracted within the CSF, the WM, and the whole-brain masks. Additionally, a set of physiological regressors were extracted to allow for component-based noise correction (CompCor, (Behzadi et al., 2007)). Principal components are estimated after high-pass filtering the preprocessed BOLD time-series (using a discrete cosine filter with 128s cut-off) for the two CompCor variants: temporal (tCompCor) and anatomical (aCompCor). tCompCor components are then calculated from the top 5% variable voxels within a mask covering the subcortical regions. This subcortical mask is obtained by heavily eroding the brain mask, which ensures it does not include cortical GM regions. For aCompCor, components are calculated within the intersection of the aforementioned mask and the union of CSF and WM masks calculated in T1w space, after their projection to the native space of each functional run (using the inverse BOLD-to-T1w transformation). Components are also calculated separately within the WM and CSF masks. For each CompCor decomposition, the k components with the largest singular values are retained, such that the retained components' time series are sufficient to explain 50 percent of variance across the nuisance mask (CSF, WM, combined, or temporal). The remaining components are dropped from consideration. The head-motion estimates calculated in the correction step were also placed within the corresponding confounds file. The confound time series derived from head motion estimates and global signals were expanded with the inclusion of temporal derivatives and quadratic terms for each (Satterthwaite et al., 2013). Frames that exceeded a threshold of 0.5 mm FD or 1.5 standardised DVARS were annotated as motion outliers. All resamplings can be performed with a single interpolation step by composing all the pertinent transformations (i.e. head-motion transform matrices, susceptibility distortion correction when available, and co-registrations to*

*anatomical and output spaces). Gridded (volumetric) resamplings were performed using antsApplyTransforms (ANTs), configured with Lanczos interpolation to minimize the smoothing effects of other kernels (Lanczos, 1964). Non-gridded (surface) resamplings were performed using mri_vol2surf (FreeSurfer).*

*Many internal operations of fMRIPrep use Nilearn 0.6.2 (Abraham et al., 2014)) (RRID:SCR_001362), mostly within the functional processing workflow. For more details of the pipeline, see the section corresponding to workflows in fMRIPrep's documentation."*

### 5.6.4 Consensus clustering

***Pre-processing behavioural data.*** Participants were excluded if they had major perinatal brain lesions (i.e., periventricular leukomalacia, parenchymal haemorrhagic infarction, or other ischemic or haemorrhagic lesions), more than 25% of behavioural data missing, or outlier data points exceeding median values by 3 times the interquartile range (Figure SM 5.1). Behavioural measures with data missing in less than 25% of the sample were included the analyses, and any missing data was imputed using K-nearest neighbour imputation. Age-normalised scores were used for the WISC, BRIEF-2 and SRS-2 scales. For the remaining behavioural variables (i.e., ERT, ERC, SDQ, TMCQ, and SCAS scores), scores were regressed against age and the residuals were used.

***Behavioural data used as input features in the clustering model.*** The following behavioural measures were used as input features in the consensus clustering pipeline: WISC-VC, WISC-PR, WISC-WM, WISC-PS, BRIEF-BR, BRIEF-ER and BRIEF-CR, ERT total, ERC total, SCAS total, SDQ – Internalising, SDQ – Externalising, TMCQ – Surgency, TMCQ – EC, TMCQ – NA, and SRS-2 total (details described in Table SM 5.1).

***Estimating the optimal number of clusters.*** Eigen-gap and Rotation Cost heuristics were used to estimate the "best" and "second best" number of clusters based on thirty different hyperparameter combinations, indicating C=2, C=3, and C=5 as the most optimal number of clusters (Figure SM 5.2A). We, therefore, ran the complete consensus clustering pipeline using C=2, C=3, and C=5 (selected based on Eigen-gap and Rotation Cost estimations) and used corresponding consensus matrices, consensus values, and Silhouette scores to determine C=2 and C=3 as the most optimal number of clusters to be used for subsequent phenotypic evaluation (Figure SM 5.2).

### 5.6.5 C=2 FC alterations in General Resilience < General Difficulties – post-hoc analyses

Sensitivity analyses adjusting for birth status identified a significant component at p-NBS-Threshold = 0.01 (p-FWER=0.021; T-statistic= 362.65) comprised of 350 nodes (93.58% of regions) and 1088 edges (i.e., 1.56% of connections), with very similar connectivity patterns to the component identified in the main analysis before adjusting for birth status (Figure 5.7). Subsequent post-hoc analyses investigating similarity between *General Resilience* < *General Difficulties* NBS components identified before (Figure 5.7) and after (Figure SM 5.4) removing effects of birth status, report a high number of overlapping nodes (n=343 shared nodes; Sørensen-Dice=0.99) and edges (n=913 shared edges; Sørensen-Dice=0.86), suggesting that FC pattern differences between the two data-driven subgroups are occurring independently of clinical birth status (i.e., VPT or FT birth).

The described effects in Figure 5.7 were detected at the p-NBS-Threshold = 0.01. No other significant components were identified at 0.05, or 0.001 p-NBS-Thresholds (p-FWER>0.05); however, at p-NBS-Threshold = 0.05, a *General Difficulties* < *General Resilience* component with effects trending towards significance (p-FWER=0.053). This component exhibited largely similar FC patterns to those identified at p-NBS-Threshold = 0.01 (Nodal Sørensen-Dice coefficient = 1.00; edgewise Sørensen-Dice coefficient = 0.89); and notably, NBS component effects were significant after removing effects of birth status (p-FWER=0.041) (Figure SM 5.5). This observation further supports the notion that FC patterns differentiating between the C=2 subgroups are occurring independently of clinical birth status.

### 5.6.6 Supplementary tables

***Table SM 5.1. Behavioural assessment description.***

| Behaviour | Measurement | Measurement description |
|---|---|---|
| **Intelligence** | *WISC – VC* | Age-normalised composite score computed based on performance on 3 subtests from the WISC-IV: Similarities, Vocabulary, and Comprehension. **Measures the child's verbal intellectual abilities**, encompassing acquired verbal knowledge and the use and comprehension of language to reason and problem-solve. Higher scores indicate better outcomes. |
| | *WISC – PR* | Age-normalised composite score computed based on performance on 3 subtests: Block Design, Picture Concepts, and Matrix Reasoning. **Measures the ability to interpret and organise information using non-verbal visuo-spatial reasoning skills.** Higher scores indicate better outcomes. |
| | *WISC – WM* | Age-normalised composite score computed based on performance on 2 subtests: Digit Span and Letter-Number Sequencing. **Measures the ability to temporarily retain and process information in conscious awareness to complete tasks.** Higher scores indicate better outcomes. |
| | *WISC – PS* | Age-normalised composite score computed based on performance on 2 subtests: Coding and Symbol Search. **Measures efficiency in processing information using visual stimuli and graphomotor skills to solve tasks.** Higher scores indicate better outcomes. |
| **Executive function** | *BRIEF-2 – BR* | Age-normalised T-score calculated based on Inhibit and Self-Monitor subscale scores. **Measures regulatory and monitoring processes supporting appropriate behavioural self-regulation.** Higher scores indicate more difficulties in this domain. Scores ≥ 76 indicate clinical relevance. |
| | *BRIEF-2 – ER* | Age-normalised T-score calculated based on Shift and Emotion Control subscale scores. **Measures the ability to set shift or adapt to changes in the surroundings** (e.g., environment, people, plans, or demands), supporting appropriate flexibility and emotion regulation. Higher scores indicate more difficulties in this domain. Scores ≥ 67 indicate clinical relevance. |
| | *BRIEF-2 – CR* | Age-normalised T-score calculated based on Initiate, Working Memory, Plan/Organise, Task-Monitor, and Organisation of Materials subscale scores. **Measures the ability to effectively manage and control cognitive processes** supporting active problem-solving in various contexts as well as the ability to complete tasks. Higher scores indicate more difficulties in this domain. Scores ≥ 77, indicate clinical relevance. |
| **Socio-emotional processing** | *ERT total* | A summative score measuring the total number of correct responses on a **task measuring child's ability to accurately recognise emotions expressed (happy, sad, surprise, anger, disgust, fear, and neutral)** from pictures of children of a similar age at two different intensities (50% and 100% levels of intensity). |
| | *ERC total* | A composite score measuring the **ability to self-regulate emotional expressions**, encompassing both emotion regulation and dysregulation (lability/negativity) processes. Higher scores indicate poorer adaptive emotion regulation (e.g., empathy, appropriate emotion expression, and self-awareness) and greater emotion dysregulation (e.g., mood lability, lack of flexibility). |

| | | |
|---|---|---|
| **Autism traits** | *SRS-2 total* | Age-normalised T-score **measuring autism trait presence and severity**, regarding difficulties in social awareness, social cognition, social communication, social motivation, and restricted interests and repetitive behaviours. Higher scores reflect greater severity of autism trait difficulties. T-scores ≥ 76 are considered to have clinical significance. |
| **Psychopathology** | *SDQ– Externalising* | **A measure of inattention, hyperactivity, and conduct problems.** Higher scores reflect more difficulties in this domain. |
| | *SDQ – Internalising* | **A measure of emotional and peer problems.** Higher scores reflect more difficulties in this domain. |
| **Anxiety** | SCAS total | A summative score measuring severity of anxiety symptoms, based on separation anxiety, social phobia, obsessive compulsive, panic/agoraphobia, physical injury and generalised anxiety symptoms. Higher scores indicate greater severity of symptoms. |
| **Temperament** | TMCQ – NA | **Temperament trait measuring the tendency to express negative emotions** (i.e., feeling sad, fearful, angry, irritable, or uncomfortable) in response to internal and external stimuli **with difficulties soothing.** Higher scores indicate higher levels of negative reactivity. |
| | TMCQ – EC | Temperament trait measuring behavioural dimensions such as attention focusing, activation control inhibitory control, low-intensity pleasure, and perceptual sensitivity, which support the child's **ability to actively regulate behaviours.** Higher scores indicate better ability to self-regulate. |
| | TMCQ – Surgency | Temperament trait measuring behavioural dimensions relating to **activity levels, impulsivity, high-intensity pleasure, and extraversion (i.e., less shyness).** Higher scores indicate higher levels of surgency/extraversion. |

***Abbreviations.*** *BRIEF-2 = Behavior Rating Inventory of Executive Function, Second Edition; BRIEF-2 – BR = behavioural regulation; BRIEF-2 – CR = cognitive regulation; BRIEF-2 – ER = emotion regulation. ERC = Emotion Regulation Checklist. ERT = Emotion Recognition Task. SCAS = Spence Children's Anxiety Scale. SDQ = Strengths and Difficulties Questionnaire. TMCQ = Temperament in Middle Childhood Questionnaire; TMCQ – NA = negative affectivity; TMCQ – EC = effortful control. WISC = Weschler Intelligence Scale for Children, Fourth Edition; PR = perceptual reasoning; PS = processing speed, VC = verbal comprehension; WM = working memory.*

*Table SM 5.2. Behavioural profiles of VPT and FT samples included in consensus clustering analyses (n=153).*

| | FT (n=56) | VPT (n=97) | p-value | p-FDR | Adj. p-FDR | Effect size |
|---|---|---|---|---|---|---|
| **WISC – VC** | 109.00 (19.00) | 104.00 (18.00) | 0.017 | 0.068 | **0.036** | **0.232** |
| **WISC – PR** | 116.00 (19.00) | 104.00 (21.00) | <0.001 | <0.001 | **<0.001** | **0.348** |
| **WISC – WM** | 104.00 (14.00) | 99.00 (13.00) | 0.009 | 0.048 | **0.048** | **0.253** |
| **WISC – PS** | 112.00 (24.00) | 103.00 (18.00) | 0.026 | 0.069 | 0.064 | 0.216 |
| **BRIEF-2 – BR** | 46.50 (13.00) | 50.00 (13.00) | 0.140 | 0.204 | 0.370 | -0.143 |
| **BRIEF-2 – ER** | 51.00 (14.50) | 52.00 (15.00) | 0.293 | 0.335 | 0.370 | -0.102 |
| **BRIEF-2 – CR** | 49.50 (14.50) | 52.00 (11.00) | 0.128 | 0.204 | 0.283 | -0.148 |
| **ERT total** | 41.50 (3.25) | 41.00 (6.00) | 0.041 | 0.094 | 0.064 | 0.199 |
| **ERC total** | 52.50 (6.25) | 54.00 (6.00) | 0.136 | 0.198 | 0.316 | -0.145 |
| **SCAS total** | 13.00 (9.25) | 13.00 (9.00) | 0.903 | 0.903 | 0.980 | 0.012 |
| **SDQ – Externalising** | 4.50 (3.50) | 5.00 (3.00) | 0.060 | 0.120 | 0.096 | -0.182 |
| **SDQ – Internalising** | 3.00 (3.00) | 4.00 (4.00) | 0.358 | 0.382 | 0.370 | -0.089 |
| **TMCQ – Surgency** | 3.13 (0.54) | 3.25 (0.73) | 0.171 | 0.228 | 0.316 | -0.133 |
| **TMCQ – EC** | 3.51 (0.55) | 3.43 (0.60) | 0.022 | 0.069 | **0.036** | **0.222** |
| **TMCQ – NA** | 2.11 (0.66) | 2.17 (0.83) | 0.268 | 0.330 | 0.316 | -0.108 |
| **SRS-2 total** | 48.00 (10.25) | 54.00 (12.00) | 0.001 | 0.008 | **0.016** | **-0.328** |

***Note.*** *Median (interquartile range) are reported. Adj. FDR p-values correspond to FDR-corrected p-values after adjusting for covariates (sex, age, and IMD). Wilcoxon Glass Rank Biserial Correlation was used to estimate effect sizes. P-values and effect sizes in bold are significant (Adj. p-FDR < 0.05). Abbreviations. BRIEF-2 = Behavior Rating Inventory of Executive Function, Second Edition; BRIEF-2 – BR = behavioural regulation; BRIEF-2 – CR = cognitive regulation; BRIEF-2 – ER = emotion regulation. ERC = Emotion Regulation Checklist. ERT = Emotion Recognition Task. SCAS = Spence Children's Anxiety Scale. SDQ = Strengths and Difficulties Questionnaire. TMCQ = Temperament in Middle Childhood Questionnaire; TMCQ – NA = negative affectivity; TMCQ – EC = effortful control. WISC = Weschler Intelligence Scale for Children, Fourth Edition; PR = perceptual reasoning; PS = processing speed, VC = verbal comprehension; WM = working memory.*

*Table SM 5.3. VPT and FT behavioural outcome measures.*

| | VPT (n=117) | FT (n=56) | p-value | p-FDR | Adj. p-FDR | Effect size |
|---|---|---|---|---|---|---|
| **WISC – VC** | 104.00 (21.00) | 109.00 (19.00) | 0.005 | 0.020 | **0.005** | **-0.092** |
| **WISC – PR** | 104.00 (23.00) | 116.00 (19.00) | <0.001 | <0.001 | **<0.001** | **-0.151** |
| **WISC – WM** | 99.00 (16.00) | 104.00 (14.00) | 0.003 | 0.016 | **0.016** | **-0.064** |
| [a]**WISC – PS** | 103.00 (18.00) | 112.00 (24.00) | 0.007 | 0.022 | **0.016** | **0.119** |
| [b]**BRIEF-2 – BR** | 50.00 (13.00) | 46.00 (13.00) | 0.078 | 0.125 | 0.1785 | -0.168 |
| [b]**BRIEF-2 – ER** | 52.00 (17.00) | 51.00 (15.00) | 0.202 | 0.245 | 0.243 | -0.121 |
| [b]**BRIEF-2 – CR** | 54.00 (17.00) | 49.00 (15.00) | 0.034 | 0.068 | 0.062 | -0.202 |
| [c]**ERT total** | 41.00 (6.00) | 41.50 (3.25) | 0.058 | 0.103 | 0.058 | 0.179 |
| [d]**ERC total** | 3.54 (0.46) | 3.44 (0.41) | 0.088 | 0.128 | 0.179 | -0.167 |
| [e]**SCAS total** | 13.00 (11.00) | 13.00 (9.50) | 0.808 | 0.808 | 0.403 | 0.026 |
| [f]**SDQ – Externalising** | 5.00 (4.00) | 4.50 (3.50) | 0.03 | 0.068 | **0.048** | **-0.051** |
| [f]**SDQ – Internalising** | 4.00 (5.00) | 3.00 (3.00) | 0.151 | 0.201 | 0.179 | -0.060 |
| [g]**TMCQ – Surgency** | 3.24 (0.79) | 3.13 (0.54) | 0.252 | 0.269 | 0.371 | -0.101 |
| [g]**TMCQ – EC** | 3.42 (0.60) | 3.51 (0.55) | 0.02 | 0.053 | **0.021** | **0.223** |
| [g]**TMCQ – NA** | 2.17 (0.87) | 2.11 (0.66) | 0.214 | 0.245 | 0.179 | -0.120 |
| [g]**SRS-2 total** | 54.00 (14.50) | 48.00 (10.25) | <0.001 | <0.001 | **<0.001** | **-0.356** |

*Note. Median (interquartile range) are reported. Adj. FDR p-values correspond to FDR-corrected p-values after adjusting for covariates (sex, age, and IMD). Wilcoxon Glass Rank Biserial Correlation was used to estimate effect sizes. P-values and effect sizes in bold are significant (Adj. p-FDR < 0.05). Missing data:* [a, b, c, d, e, f, g] *respectively correspond to n=1, 5, 4, 15, 18, 7, and 13 missing data points. Abbreviations. BRIEF-2 = Behavior Rating Inventory of Executive Function, Second Edition; BRIEF-2 – BR = behavioural regulation; BRIEF-2 – CR = cognitive regulation; BRIEF-2 – ER = emotion regulation. ERC = Emotion Regulation Checklist. ERT = Emotion Recognition Task. SCAS = Spence Children's Anxiety Scale. SDQ = Strengths and Difficulties Questionnaire. TMCQ = Temperament in Middle Childhood Questionnaire; TMCQ – NA = negative affectivity; TMCQ – EC = effortful control. WISC = Weschler Intelligence Scale for Children, Fourth Edition; PR = perceptual reasoning; PS = processing speed, VC = verbal comprehension; WM = working memory.*

*Table SM 5.4. Between-group differences in in-scanner head motion.*

| | FD, mm | p-value |
|---|---|---|
| **Clinical birth status** | | 0.609 |
| *VPT* | 0.06 (0.03) | |
| *FT* | 0.06 (0.02) | |
| **C=2 data-driven subgroups** | | 0.024 |
| *General Difficulties* | 0.07 (0.03) | |
| *General Resilience* | 0.06 (0.02) | |
| **C=3 data-driven subgroups** | | 0.163 |
| *Psychiatric Difficulties* | 0.07 (0.03) | |
| *Typical Development* | 0.06 (0.03) | |
| *Neurodevelopmental Difficulties* | 0.06 (0.02) | |

***Note.*** *In-scanner head motion measured using mean Framewise Displacement (FD). Median and (interquartile range) are reported. Appropriate non-parametric statistical tests were used to determine significant between-group effects. P<0.05 is considered significant.*

*Table SM 5.5. Comparing profiles of VPT children included and excluded from the VPT vs FT analyses.*

| | VPT included (n=117) | VPT excluded (n=41) | p-value |
|---|---|---|---|
| **GA at birth, weeks** | 29.86 (4.00) | 30.00 (3.43) | 0.634 |
| [a] **Neonatal sickness** | 0.02 (1.61) | 0.24 (1.99) | 0.355 |
| **Age at assessment, years** | 9.25 (1.17) | 8.92 (1.00) | 0.348 |
| **IMD, rank** | 19575.00 (14506.00) | 18982.00 (13220.00) | 0.699 |
| [b] **WISC – VC** | 104.00 (21.00) | 98.00 (19.00) | 0.050 |
| [c] **WISC – PR** | 104.00 (23.00) | 100.00 (16.00) | **0.020** |
| [b] **WISC – WM** | 99.00 (16.00) | 89.50 (16.75) | **0.001** |
| [d] **WISC – PS** | 103.00 (18.00) | 91.00 (22.00) | **<0.001** |
| [e] **BRIEF-2 – BR** | 50.00 (13.00) | 55.00 (19.75) | **0.041** |
| [e] **BRIEF-2 – ER** | 52.00 (17.00) | 57.00 (19.75) | **0.025** |
| [e] **BRIEF-2 – CR** | 54.00 (17.00) | 58.50 (10.75) | 0.060 |
| [f] **ERT total** | 41.00 (6.00) | 38.00 (5.00) | **0.018** |
| [g] **ERC total** | 3.54 (0.46) | 3.62 (0.42) | 0.249 |
| [h] **SCAS total** | 13.00 (11.00) | 20.00 (11.50) | **0.002** |
| [i] **SDQ – Externalising** | 5.00 (4.00) | 6.00 (3.00) | 0.475 |
| [i] **SDQ – Internalising** | 4.00 (5.00) | 5.00 (4.00) | **0.006** |
| [j] **TMCQ – Surgency** | 3.24 (0.79) | 2.99 (0.74) | 0.224 |
| [i] **TMCQ – EC** | 3.42 (0.60) | 3.18 (0.69) | 0.388 |
| [j] **TMCQ – NA** | 2.17 (0.87) | 2.65 (0.86) | **0.025** |
| [k] **SRS-2 total** | 54.00 (14.50) | 60.00 (21.00) | 0.096 |

**Note.** *Median (interquartile range) are reported. FDR p-values correspond to FDR-corrected p-values. Missing data: [a, b, c, d, e, f, g, h, i, j, k] respectively correspond to n=2, 1, 4, 5, 9, 8, 17, 22, 11, 18, and 23 missing data points. Abbreviations. BRIEF-2 = Behavior Rating Inventory of Executive Function, Second Edition; BRIEF-2 – BR = behavioural regulation; BRIEF-2 – CR = cognitive regulation; BRIEF-2 – ER = emotion regulation. ERC = Emotion Regulation Checklist. ERT = Emotion Recognition Task. SCAS = Spence Children's Anxiety Scale. SDQ = Strengths and Difficulties Questionnaire. TMCQ = Temperament in Middle Childhood Questionnaire; TMCQ – NA = negative affectivity; TMCQ – EC = effortful control. WISC = Weschler Intelligence Scale for Children, Fourth Edition; PR = perceptual reasoning; PS = processing speed, VC = verbal comprehension; WM = working memory.*

*Table SM 5.6. Two-subgroup solution behavioural profiles.*

| | Subgroup 1: General Difficulties (n=65) | Subgroup 2: General Resilience (n=88) | p-value | p-FDR | Adj. p-FDR | Effect size |
|---|---|---|---|---|---|---|
| **Socio-demographic and clinical measures** | | | | | | |
| *Age at assessment, years* | 8.92 (1.25) | 9.17 (0.77) | 0.264 | 0.292 | n/a | -0.106 |
| *IMD, rank* | 18705.00 (15890.00) | 22741.00 (16539.00) | 0.374 | 0.393 | n/a | -0.084 |
| *Sex, M:F (%M)* | 39:26 (60.00%) | 43:45 (48.86%) | 0.230 | 0.268 | n/a | V=0.110 |
| *Birth status, VPT:FT (%VPT)* | 48:17 (73.85%) | 49:39 (55.68%) | 0.033 | **0.046** | **n/a** | **V=0.186** |
| **Behavioural measures** | | | | | | |
| *WISC – VC* | 102.00 (19.00) | 108.00 (16.75) | 0.011 | **0.018** | **0.028** | **-0.240** |
| *WISC – PR* | 104.00 (18.00) | 112.00 (21.00) | <0.001 | **<0.001** | **0.004** | **-0.335** |
| *WISC – WM* | 99.00 (16.00) | 102.00 (16.75) | 0.117 | 0.145 | 0.414 | -0.148 |
| *WISC – PS* | 100.00 (15.00) | 109.00 (21) | 0.002 | **0.004** | **0.003** | **-0.293** |
| *BRIEF-2 – BR* | 57.00 (13.00) | 45.00 (7.25) | <0.001 | **<0.001** | **<0.001** | **0.784** |
| *BRIEF-2 – ER* | 61.00 (14.00) | 47.00 (9.25) | <0.001 | **<0.001** | **<0.001** | **0.712** |
| *BRIEF-2 – CR* | 58.00 (12.00) | 47.00 (8.25) | <0.001 | **<0.001** | **<0.001** | **0.720** |
| *ERT total* | 40.00 (6.00) | 42.00 (6.00) | 0.074 | 0.097 | 0.059 | -0.169 |
| *ERC total* | 3.71 (0.42) | 3.29 (0.33) | <0.001 | **<0.001** | **<0.001** | **0.777** |
| *SCAS total* | 16.00 (14.00) | 11.9 (8.00) | <0.001 | **<0.001** | **<0.001** | **0.406** |
| *SDQ – Externalising* | 7.00 (5.00) | 4.00 (3.00) | <0.001 | **<0.001** | **<0.001** | **0.692** |
| *SDQ – Internalising* | 5.00 (4.00) | 3.00 (2.00) | <0.001 | **<0.001** | **<0.001** | **0.455** |
| *TMCQ – Surgency* | 3.16 (0.72) | 3.21 (0.63) | 0.738 | 0.738 | 0.749 | 0.032 |
| *TMCQ – EC* | 3.18 (0.63) | 3.59 (0.42) | <0.001 | **<0.001** | **<0.001** | **-0.584** |
| *TMCQ – NA* | 2.69 (0.72) | 1.87 (0.52) | <0.001 | **<0.001** | **<0.001** | **0.773** |
| *SRS-2 total* | 60 (16.00) | 48.00 (10.00) | <0.001 | **<0.001** | **<0.001** | **0.702** |

***Note.*** *Median (interquartile range) reported unless stated otherwise, where ratios of male to female (M:F), and VPT to FT (VPT:FT) children are reported alongside percentage (%). Adj. FDR p-value corresponds to the p-value after adjusting for covariates (sex, age, and IMD) and correcting for multiple comparisons with FDR. Effect sizes are calculated using Wilcoxon Glass Rank Biserial Correlation, unless otherwise stated. Cohen's V (V) effect size was used for categorical variables. P-values and effect sizes in bold are significant (Adj. p-FDR < 0.05). Abbreviations. BRIEF-2 = Behavior Rating Inventory of Executive Function, Second Edition; BRIEF-2 – BR = behavioural regulation; BRIEF-2 – CR = cognitive regulation; BRIEF-2 – ER = emotion regulation. ERC = Emotion Regulation Checklist. ERT = Emotion Recognition Task. SCAS = Spence Children's Anxiety Scale. SDQ = Strengths and Difficulties Questionnaire. TMCQ = Temperament in Middle Childhood Questionnaire; TMCQ – NA = negative affectivity; TMCQ – EC = effortful control. WISC = Weschler Intelligence Scale for Children, Fourth Edition; PR = perceptual reasoning; PS = processing speed, VC = verbal comprehension; WM = working memory.*

*Table SM 5.7. Three-subgroup solution behavioural profiles.*

| | Subgroup 1: PD (n=46) | Subgroup 2: TD (n=53) | Subgroup 3: ND (n=54) | p-value | p-FDR | Adj. p-FDR | Effect size |
|---|---|---|---|---|---|---|---|
| **Socio-demographic and clinical measures** | | | | | | | |
| *Age at assessment, years* | 8.92 (1.21) | 9.00 (0.83) | 9.29 (0.81) | 0.352 | 0.352 | n/a | -0.12 |
| *IMD, rank* | 18937.00 (17051.00) | 24196.00 (17267.00) | 19319.50 (16556.00) | 0.045 | 0.053 | n/a | -0.148 |
| *Sex, M:F (%M)* | 28:18 (60.87%) | 30:23 (56.60%) | 24:30 (44.44%) | 0.224 | 0.236 | n/a | V=0.140 |
| *Birth status, VPT:FT (%VPT)* | 34:12 (73.91%) | 24:29 (45.28%) | 39:15 (72.22%) | 0.003 | **0.004** | **n/a** | **V=0.274** |
| **Behavioural measures** | | | | | | | |
| *WISC – VC* | 104.00 (18) | 114.00 (18.00) | 100.00 (13.00) | <0.001 | **<0.001** | **<0.001** | **-0.487** |
| *WISC – PR* | 104.00 (14.25) | 119.00 (15.00) | 100.00 (16.00) | <0.001 | **<0.001** | **<0.001** | **-0.523** |
| *WISC – WM* | 102.00 (16.00) | 107.00 (14.00) | 94.00 (11.00) | <0.001 | **<0.001** | **<0.001** | **-0.345** |
| *WISC – PS* | 100.00 (17.25) | 115.00 (15.00) | 103.00 (15.00) | 0.002 | **0.003** | **<0.001** | **-0.405** |
| *BRIEF-2 – BR* | 60.00 (11.00) | 45.00 (8.00) | 47.00 (6.75) | <0.001 | **<0.001** | **<0.001** | **0.878** |
| *BRIEF-2 – ER* | 62.50 (10.25) | 48.00 (10.00) | 47.00 (8.00) | <0.001 | **<0.001** | **<0.001** | **0.773** |
| *BRIEF-2 – CR* | 64.00 (9.75) | 50.00 (11.00) | 47.50 (8.75) | <0.001 | **<0.001** | **0.001** | **0.824** |
| *ERT total* | 40.00 (6.00) | 42.2 (4.00) | 39.50 (4.75) | 0.001 | **0.002** | **<0.001** | **-0.368** |
| *ERC total* | 3.75 (0.46) | 3.29 (0.38) | 3.38 (0.29) | <0.001 | **<0.001** | **<0.001** | **0.792** |
| *SCAS total* | 18.00 (13.75) | 9.00 (9.00) | 12.50 (7.50) | <0.001 | **<0.001** | **<0.001** | **0.585** |
| *SDQ – Externalising* | 8.00 (4.75) | 3.00 (3.00) | 5.00 (3.00) | <0.001 | **<0.001** | **<0.001** | **0.737** |
| *SDQ – Internalising* | 6.00 (3.75) | 3.00 (1.00) | 3.00 (2.90) | <0.001 | **<0.001** | **0.007** | **0.621** |
| *TMCQ – Surgency* | 3.19 (0.80) | 3.31 (0.59) | 3.10 (0.67) | 0.003 | **0.004** | **<0.001** | **-0.057** |
| *TMCQ – EC* | 3.09 (0.75) | 3.61 (0.40)) | 3.44 (0.47) | <0.001 | **<0.001** | **<0.001** | **-0.663** |
| *TMCQ – NA* | 2.84 (0.61) | 1.82 (0.53) | 2.10 (0.48) | <0.001 | **<0.001** | **<0.001** | **0.877** |
| *SRS-2 total* | 63.00 (20.75) | 46.00 (8.00) | 53.00 (8.00) | <0.001 | **<0.001** | **<0.001** | **0.819** |

Note. *Median (interquartile range) reported unless stated otherwise, where ratios of male to female (M:F), and VPT to FT (VPT:FT) children are reported alongside percentage (%). Adj. FDR p-value corresponds to the p-value after adjusting for covariates (sex, age, and IMD) and correcting for multiple comparisons with FDR. Effect sizes are calculated using Wilcoxon Glass Rank Biserial Correlation, unless otherwise stated. Cramer's V (V) effect size was used for categorical variables. P-values and effect sizes in bold are significant (Adj. p-FDR < 0.05). Abbreviations. BRIEF-2 = Behavior Rating Inventory of Executive Function, Second Edition; BRIEF-2 – BR = behavioural regulation; BRIEF-2 – CR = cognitive regulation; BRIEF-2 – ER = emotion regulation. ERC = Emotion Regulation Checklist. ERT = Emotion Recognition Task. SCAS = Spence Children's Anxiety Scale. SDQ = Strengths and Difficulties Questionnaire. TMCQ = Temperament in Middle Childhood Questionnaire; TMCQ – NA = negative affectivity; TMCQ – EC = effortful control. WISC = Weschler Intelligence Scale for Children, Fourth Edition; PR = perceptual reasoning; PS = processing speed, VC = verbal comprehension; WM = working memory.*

***Table SM 5.8. VPT children – post-hoc exploratory analyses investigating C=2 and C=3 between-subgroup differences in clinical and socio-demographic measures.***

| | C=2 | | | C=3 | | | |
|---|---|---|---|---|---|---|---|
| | GD (n=48) | GR (n=49) | p-value | PD (n=34) | TD (n=24) | ND (n=39) | p-value |
| **Clinical measures** | | | | | | | |
| *GA, weeks* | 30.21 (4.00) | 29.71 (4.14) | 0.217 | 30.71 (3.68) | 30.07 (4.14) | 29.29 (4.14) | 0.183 |
| *Neonatal sickness index* | -0.04 (1.49) | -0.14 (1.82) | 0.764 | -0.37 (1.49) | -0.35 (1.63) | 0.24 (1.80) | 0.738 |
| *Neonatal brain lesions (minor: none), n* | 33: 15 | 26: 22 | 0.208 | 23: 11 | 12: 12 | 24: 14 | 0.382 |
| **Socio-demographic measures** | | | | | | | |
| *IMD, rank* | 18738.00 (14592.00) | 22232.00 (15228.00) | 0.754 | 20630.50 (16232.00) | 24537.50 (13354.00) | 16117.00 (17024.00) | 0.045 |
| *Sex (M:F), n* | 31: 17 | 25: 24 | 0.216 | 24: 10 | 13: 11 | 19: 20 | 0.155 |

**Note.** *Median (interquartile range) reported unless otherwise stated where number of participants (n) is reported. Abbreviations. GA = Gestational Age at birth, GD = General Difficulties, GR = General Resilience, IMD = Index Multiple Deprivation, ND = Neurodevelopmental Difficulties, PD = Psychiatric Difficulties, TD = Typical Development.*

***Table SM 5.9. FT children – post-hoc exploratory analyses investigating C=2 and C=3 between-subgroup differences in socio-demographic measures.***

| | C=2 | | | C=3 | | | |
|---|---|---|---|---|---|---|---|
| | GD (n=17) | GR (n=39) | p-value | PD (n=12) | TD (n=29) | ND (n=15) | p-value |
| **IMD, rank** | 18118.00 (13719.00) | 22741.00 (18427.50) | 0.269 | 11710.50 (10099.50) | 22741.00 (18466.00) | 22842.00 (10822.00) | 0.119 |
| **Sex (M:F), n** | 8: 9 | 18: 21 | 1.000 | 4: 8 | 17: 12 | 5: 10 | 0.166 |

***Note.*** *Abbreviations. GD = General Difficulties, GR = General Resilience, IMD = Index Multiple Deprivation; IQR = interquartile range; n = number of participants; ND = Neurodevelopmental Difficulties, PD = Psychiatric Difficulties, TD = Typical Development.*

*Table SM 5.10. VPT > FT – nodes with the highest number of connections within the significant NBS component.*

| Brain region | HCP-MMP atlas region | Number of edges | Percentage of edges |
|---|---|---|---|
| **Lateral temporal cortex (inferior temporal gyrus)** | left Area TE1 posterior | 33 | 4.654443 |
| **Insular cortex** | right Anterior Ventral Insular Area | 16 | 2.2567 |
| **Subcortex** | right Putamen | 15 | 2.115656 |
| **Posterior opercular cortex** | left Area PFcm | 15 | 2.115656 |
| **Paracentral Lobular and Mid Cingulate Cortex** | right Area 5m ventral | 14 | 1.974612 |
| **Posterior cingulate cortex** | right Area 31a | 13 | 1.833568 |
| **Posterior cingulate cortex** | right Area 23d | 12 | 1.692525 |
| **Posterior cingulate cortex** | right Area dorsal 23 a+b | 12 | 1.692525 |
| **Dorsolateral prefrontal cortex** | right Area 9 Posterior | 12 | 1.692525 |
| **Superior temporal gyrus (auditory association cortex)** | right Area TA2 | 12 | 1.692525 |
| **Inferior parietal cortex** | right Area PGi | 12 | 1.692525 |
| **Posterior cingulate cortex** | left PreCuneus Visual Area | 11 | 1.551481 |
| **Inferior parietal cortex** | left Area PFt | 11 | 1.551481 |
| **Temporo-Parieto-Occipital Junction** | right PeriSylvian Language Area | 11 | 1.551481 |
| **Premotor cortex** | right Dorsal area 6 | 11 | 1.551481 |
| **Auditory association cortex** | right Auditory 5 Complex | 11 | 1.551481 |
| **Frontal opercular cortex** | left Area Frontal Opercular 5 | 10 | 1.410437 |
| **Anterior Cingulate and Medial Prefrontal Cortex** | right Area Posterior 24 prime | 10 | 1.410437 |
| **Insular cortex** | right Para-Insular Area | 10 | 1.410437 |
| **Subcortex** | left Putamen | 9 | 1.269394 |
| **Subcortex** | left Pallidum | 9 | 1.269394 |
| **Subcortex** | right Pallidum | 9 | 1.269394 |
| **Superior parietal cortex** | left Medial Area 7A | 9 | 1.269394 |
| **Superior parietal cortex** | left Area 7PC | 9 | 1.269394 |
| **Anterior Cingulate and Medial Prefrontal Cortex** | left Area Posterior 24 prime | 9 | 1.269394 |
| **Anterior cingulate and medial prefrontal cortex** | left Area p32 | 9 | 1.269394 |

| | | | |
|---|---|---|---|
| **Auditory association cortex** | right Area STSd posterior | 9 | 1.269394 |
| **Lateral temporal cortex** | right Area TE1 anterior | 9 | 1.269394 |
| **Insular cortex** | right Insular Granular Complex | 9 | 1.269394 |
| **Anterior cingulate and medial prefrontal cortex** | left Area dorsal 32 | 8 | 1.12835 |
| **Inferior frontal cortex** | left Area IFSa | 8 | 1.12835 |
| **Dorsolateral prefrontal cortex** | left Area 9 anterior | 8 | 1.12835 |
| **Inferior parietal cortex** | left Area PFm Complex | 8 | 1.12835 |
| **Insular cortex** | left Insular Granular Complex | 8 | 1.12835 |
| **Paracentral Lobular and Mid Cingulate Cortex** | right Area 6mp | 8 | 1.12835 |
| **Inferior frontal cortex** | right Area 44 | 8 | 1.12835 |
| **Posterior cingulate cortex** | right Frontal Opercular Area 2 | 8 | 1.12835 |
| **Auditory association cortex** | right Area STSv posterior | 8 | 1.12835 |

*Table SM 5.11. Two subgroup-solution General Difficulties < General Resilience – nodes with the highest number of connections within the significant NBS component.*

| Brain region | HCP-MMP atlas region | Number of edges | Percentage of edges |
|---|---|---|---|
| **Auditory association cortex (Superior temporal sulcus)** | right Area STSv posterior | 35 | 3.37512054 |
| **Auditory association cortex (Superior temporal)** | left Area STSd posterior | 33 | 3.182256509 |
| **Anterior cingulate and medial prefrontal cortex** | left Area s32 | 31 | 2.989392478 |
| **Auditory association cortex (Superior temporal)** | left Area STSv posterior | 22 | 2.121504339 |
| **Anterior cingulate and medial prefrontal cortex** | right Area p32 | 21 | 2.025072324 |
| **Paracentral lobular and mid cingulate cortex** | left Area 5L | 20 | 1.928640309 |
| **Posterior opercular cortex** | left Area OP1/SII | 20 | 1.928640309 |
| **Posterior cingulate cortex** | left Area 31a | 20 | 1.928640309 |
| **Anterior cingulate and medial prefrontal cortex** | left Area a24 | 19 | 1.832208293 |
| **Dorsolateral prefrontal cortex** | left Superior 6-8 Transitional Area | 19 | 1.832208293 |
| **Superior parietal cortex** | left Area Lateral IntraParietal dorsal | 18 | 1.735776278 |
| **Insular cortex** | right Middle Insular Area | 18 | 1.735776278 |
| **Insular cortex** | right Area 52 | 17 | 1.639344262 |
| **Inferior parietal cortex** | right Area PF Complex | 17 | 1.639344262 |
| **Posterior cingulate cortex** | left Retrosplenial Complex | 16 | 1.542912247 |
| **Superior temporal gyrus (auditory association cortex)** | right Area TA2 | 16 | 1.542912247 |
| **Auditory association cortex** | right Auditory 5 Complex | 16 | 1.542912247 |
| **Superior parietal cortex** | left Area Lateral IntraParietal ventral | 15 | 1.446480231 |
| **Paracentral Lobular and Mid Cingulate Cortex (Supplementary motor area)** | left Area 6mp | 15 | 1.446480231 |
| **Anterior cingulate and medial prefrontal cortex** | left Area 33 prime | 15 | 1.446480231 |
| **Anterior cingulate and medial prefrontal cortex** | right Area a24 | 15 | 1.446480231 |

| | | | |
|---|---|---|---|
| **Dorsal stream visual cortex** | left Area V3A | 14 | 1.350048216 |
| **Primary somatosensory complex** | right Area 2 | 14 | 1.350048216 |
| **Anterior cingulate and medial prefrontal cortex** | right Area Posterior 24 prime | 14 | 1.350048216 |
| **Superior parietal cortex** | right Anterior IntraParietal | 14 | 1.350048216 |
| **Early auditory cortex** | right ParaBelt Complex | 14 | 1.350048216 |
| **Lateral temporal cortex** | right Area TE1 Middle | 14 | 1.350048216 |
| **Orbital and polar frontal cortex** | left posterior OFC Complex | 13 | 1.253616201 |
| **Inferior frontal cortex** | right Area IFJp | 13 | 1.253616201 |
| **Orbital and polar frontal cortex** | right Area 13l | 13 | 1.253616201 |
| **Inferior parietal cortex** | right Area PGi | 13 | 1.253616201 |
| **Superior parietal cortex** | left Medial IntraParietal Area | 12 | 1.157184185 |
| **Auditory association cortex** | left Auditory 5 Complex | 12 | 1.157184185 |
| **Inferior parietal cortex** | left Area IntraParietal 2 | 12 | 1.157184185 |
| **Medial temporal cortex (parahippocampal)** | left ParaHippocampal Area 2 | 12 | 1.157184185 |
| **Anterior cingulate and medial prefrontal cortex** | right Area 33 prime | 12 | 1.157184185 |
| **Posterior opercular cortex** | right Frontal Opercular Area 4 | 12 | 1.157184185 |
| **Anterior Cingulate and Medial Prefrontal Cortex** | right Area 25 | 12 | 1.157184185 |
| **Anterior cingulate and medial prefrontal cortex** | right Area s32 | 12 | 1.157184185 |
| **Early auditory cortex** | right Auditory 4 Complex | 12 | 1.157184185 |
| **Posterior cingulate cortex** | left Parieto-Occipital Sulcus Area 2 | 11 | 1.06075217 |
| **Posterior cingulate cortex** | left Parieto-Occipital Sulcus Area 1 | 11 | 1.06075217 |
| **Posterior cingulate cortex** | left Area dorsal 23 a+b | 11 | 1.06075217 |
| **Orbital and polar frontal cortex** | left Orbital Frontal Complex | 11 | 1.06075217 |
| **Premotor cortex** | left Area 6 anterior | 11 | 1.06075217 |
| **Dorsolateral prefrontal cortex** | left Inferior 6-8 Transitional Area | 11 | 1.06075217 |
| **Frontal opercular cortex** | left Frontal Opercular Area 3 | 11 | 1.06075217 |
| **Lateral temporal cortex** | left Area TE1 posterior | 11 | 1.06075217 |

| | | | |
|---|---|---|---|
| **Anterior cingulate and medial prefrontal cortex** | left Area anterior 32 prime | 11 | 1.06075217 |
| **Superior parietal cortex** | right Lateral Area 7P | 11 | 1.06075217 |
| **Orbital and polar frontal cortex** | right Area 47s | 11 | 1.06075217 |
| **Inferior parietal cortex** | right Area IntraParietal 2 | 11 | 1.06075217 |

### 5.6.7 Supplementary figures

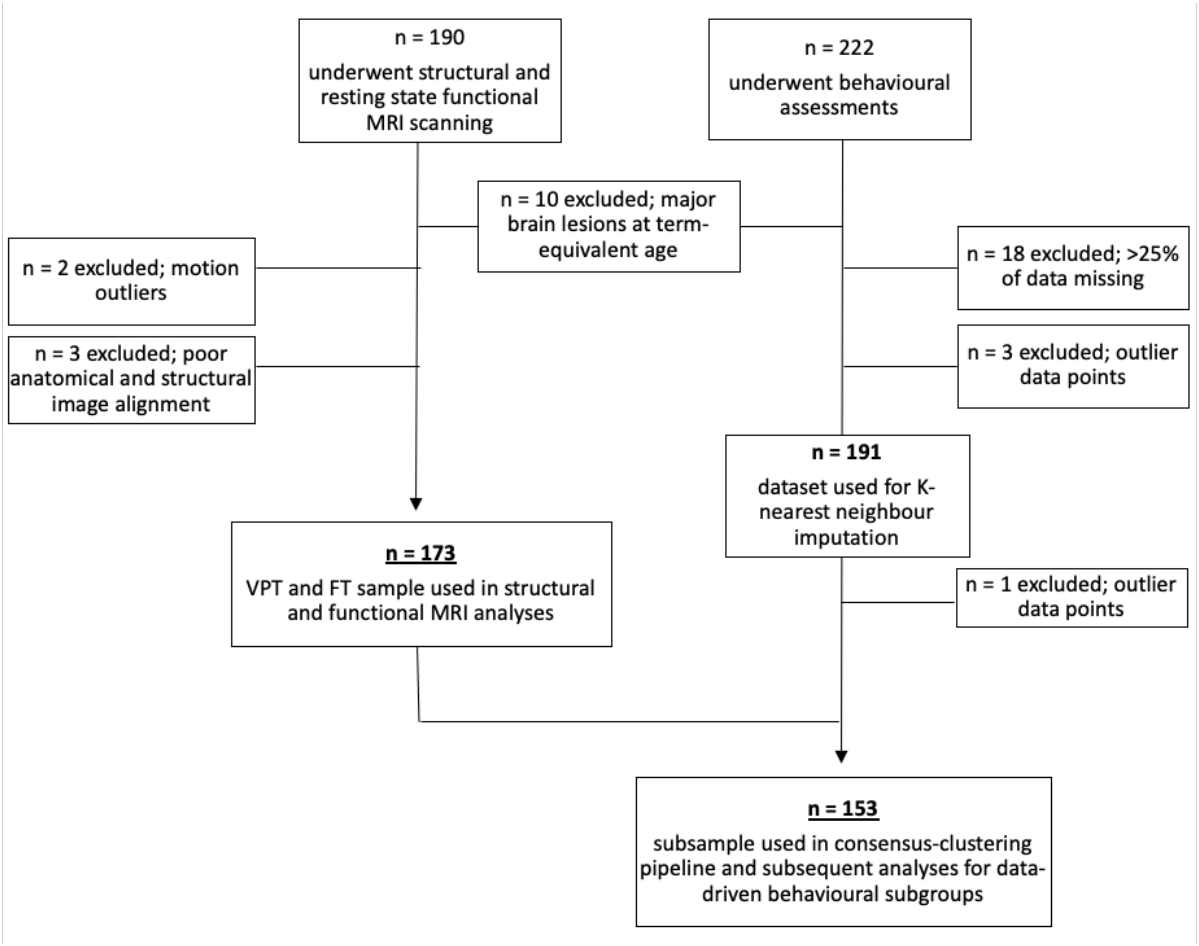



*Figure SM 5.1. Participants' selection flow diagram.*

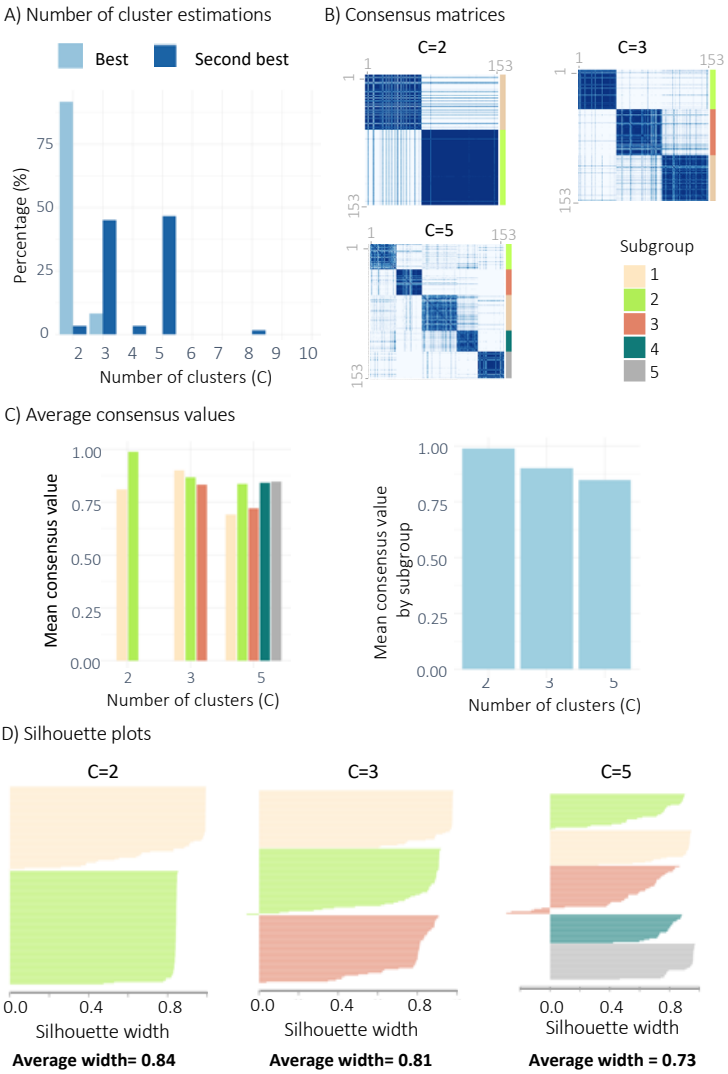



***Figure SM 5.2. Estimating the optimal number of clusters.***

*A) Percentage of times across the 30 combinations of K-alpha parameters each number of clusters (C) was selected as the best (pale blue) or second best (blue) using Eigengap and Rotation Cost. B) Consensus matrices depicting consensus values measuring proportion of times each pair of subjects co-clustered into the same subgroup over the 1000 iterations with darker blue colours indicating higher proportions of co-clustering. C) Mean consensus values for C=2, C=3, and C=5 subgrouping solutions and for each subgroup within the C=2, C=3, and C=5 subgrouping solutions (left and right, respectively). D) Silhouette width values for each subgroup within the different number of clusters runs: C=2, C=3 and C=5.*

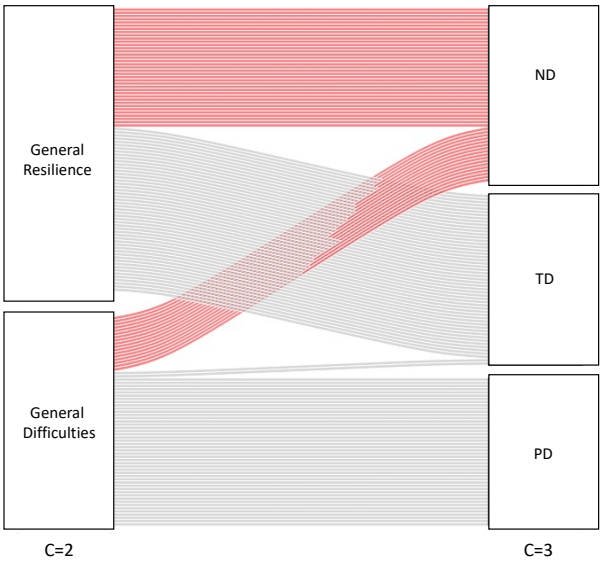



***Figure SM 5.3. Alluvial plot showing participant transitions from C=2 subgroups to the C=3 subgroups.***

*Majority of children clustering into the C=3 TD subgroup also cluster into the General Resilience subgroup from the C=2 solution and all children clustering into the C=3 PD subgroup also cluster into the C=2 General Difficulties subgroup (transitions marked in grey), while those clustering into the ND subgroup (transitions marked in red) cluster into both C=2 subgroups. Abbreviations: ND = Neurodevelopmental Difficulties; PD = Psychiatric Difficulties; TD = Typical Development.*

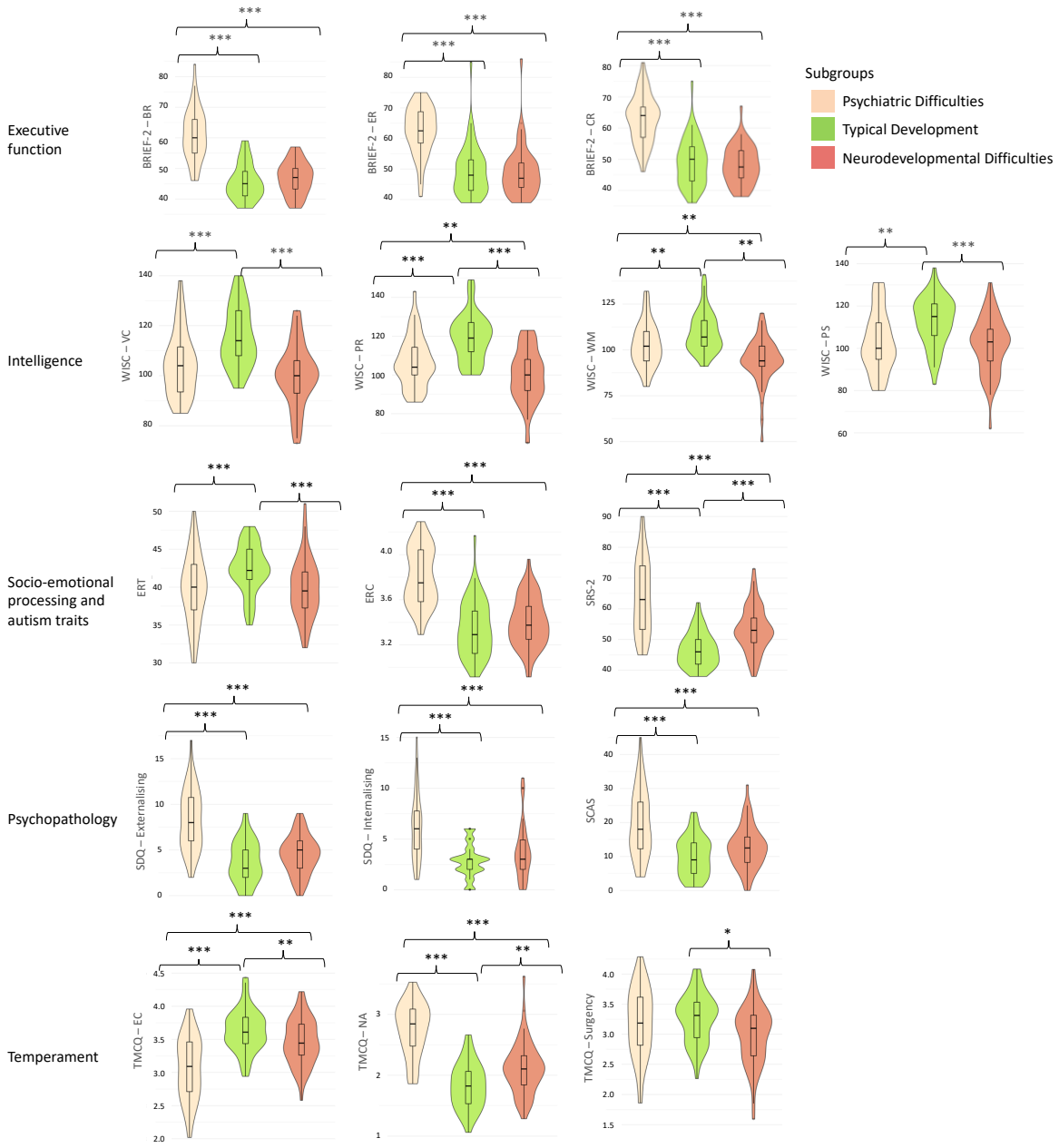



***Figure SM 5.4. Three-subgroup solution – between-subgroup differences in behavioural outcomes.***

*Reported FDR-corrected p-values (***<0.001, **<0.01, *<0.05) indicate significant between-subgroup differences after adjusting for covariates (age, sex, and IMD).*

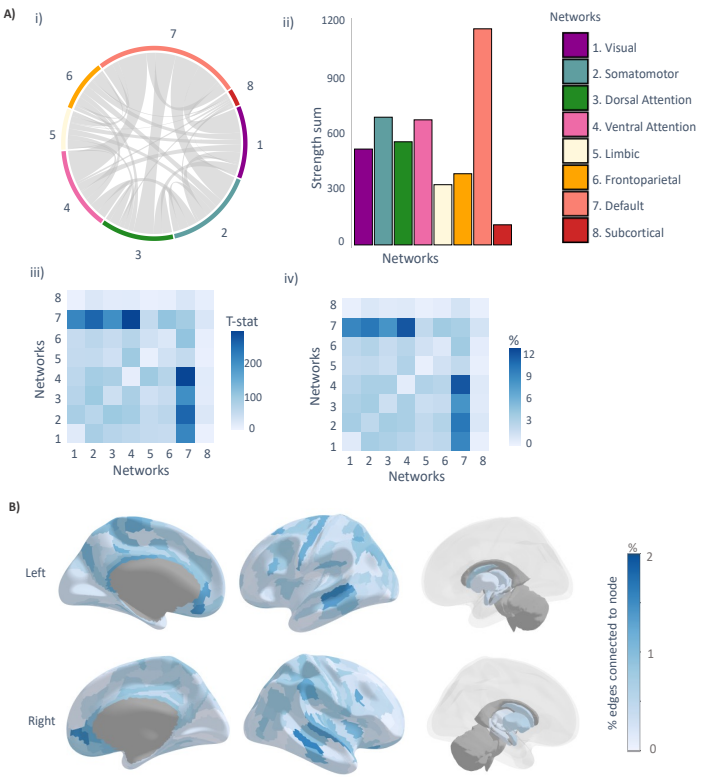



***Figure SM 5.5. General Resilience < General Difficulties – NBS FC component results after correcting for birth status (at p-NBS-Threshold = 0.01).***

*A) Intrinsic FC network i) within- and between-network connectivity, ii) total connection strength (sum of T-statistic values), within- and between-network connectivity iii) strength and iv) frequency, and B) regional FC within the component.*

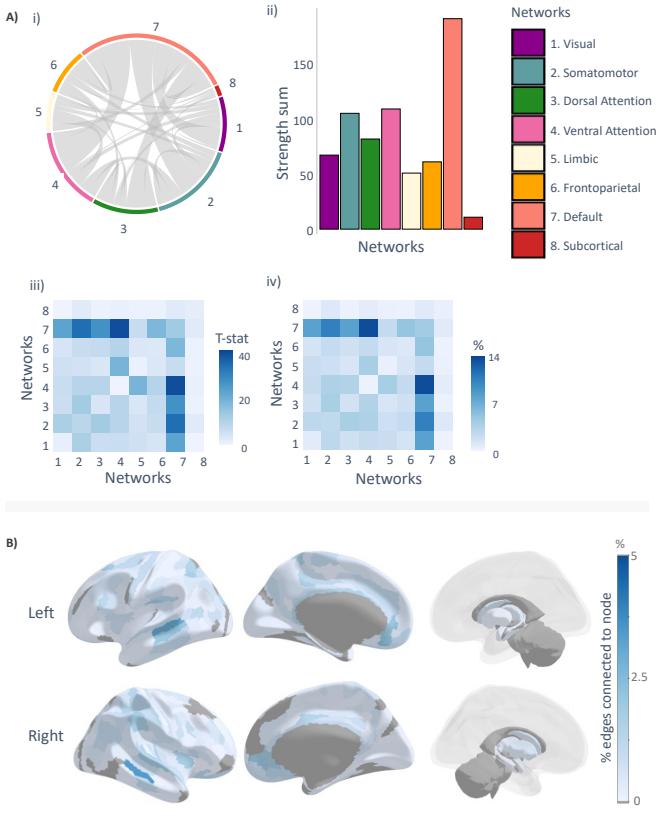



***Figure SM 5.6. General Resilience < General Difficulties – NBS FC component results after correcting for birth status (at p-NBS-Threshold = 0.05).***

*A) Intrinsic FC network i) within- and between-network connectivity, ii) total connection strength (sum of T-statistic values), within- and between-network connectivity iii) strength and iv) frequency, and B) regional FC within the component.*

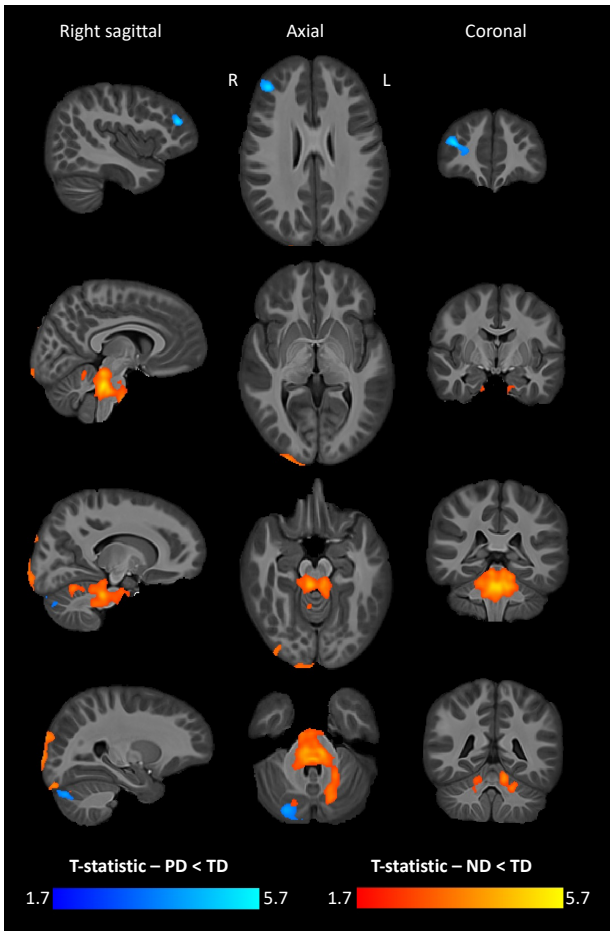



***Figure SM 5.7. Structural brain volume differences between the three-subgroup solution data-driven subgroups, correcting for age, sex, and IMD.***

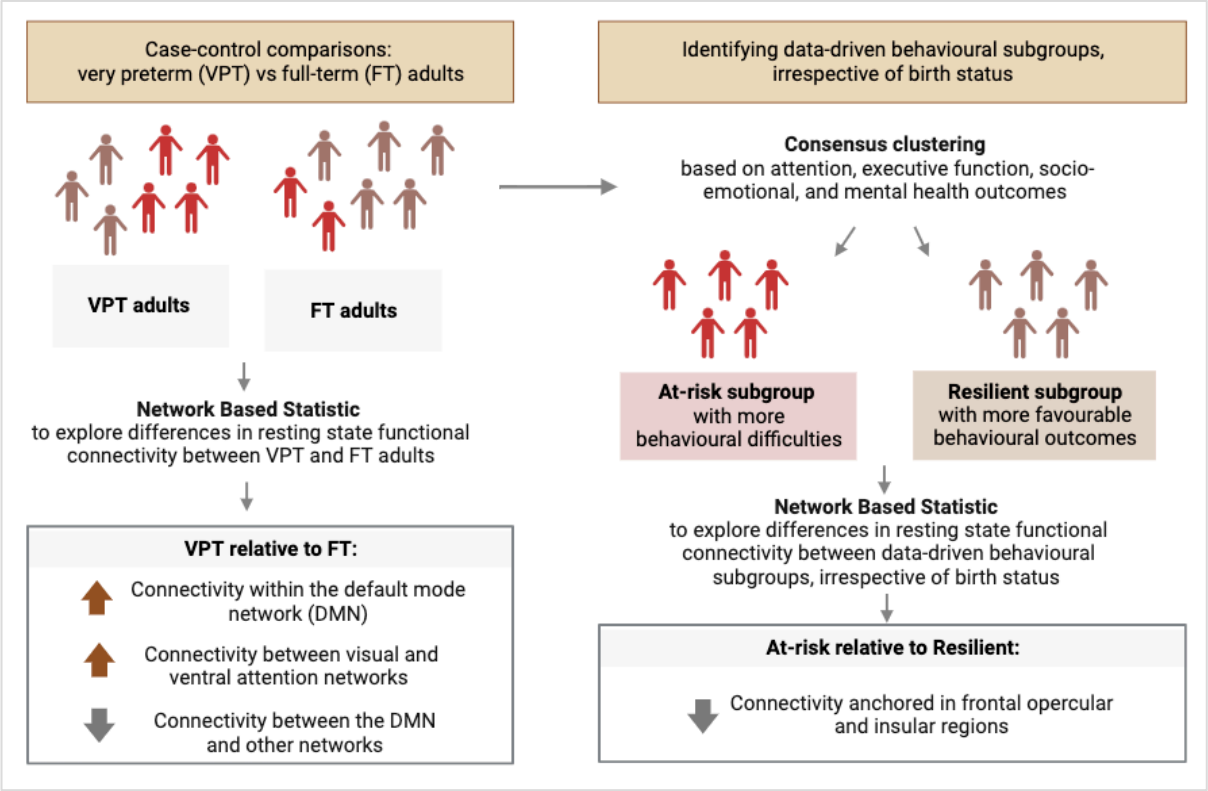
*T-statistic values for voxels with significantly smaller log-Jacobian values (i.e., relative brain volumes) in the PD (in blue) and ND (in red) subgroups relative to the TD subgroup, at p-TFCE-FWER<0.05 per contrast, before correcting for birth status. Left (L) and right (R) hemisphere orientations are labelled accordingly*

# CHAPTER 6 - Study #4: Elucidating brain-behavioural heterogeneity in VPT and FT adults using data-driven consensus clustering



***Reference****: Hadaya, L., Váša, F., Dimitrakopoulou, K., Saqi, M, Shergill, S.S., Edwards, A. D., Batalle, D., Leech, R., & Nosarti, C. (2024). Exploring functional connectivity in clinical and data-driven groups of preterm and term adults. (p. 2024.01.22.576651). bioRxiv. https://doi.org/10.1101/2024.01.22.576651*

**Study #4 graphical abstract** *(created with BioRender.com):*




## 6.1 Abstract

**Background:** Adults born very preterm (i.e., at <33 weeks' gestation) are more susceptible to long-lasting structural and functional brain alterations and cognitive and socio-emotional difficulties, compared to full-term controls. However, behavioural heterogeneity within very preterm and full-term individuals makes it challenging to find biomarkers of specific outcomes. To address these questions, we parsed brain-behaviour heterogeneity in participants subdivided according to their clinical birth status (very preterm vs full-term) and/or data-driven behavioural phenotype (regardless of birth status).

**Methods:** The Network Based Statistic approach was used to identify topological components of resting state functional connectivity differentiating between i) 116 very preterm and 83 full-term adults (43% and 57% female, respectively), and ii) data-driven behavioural subgroups identified using consensus clustering (n= 156, 46% female). Age, sex, socio-economic status, and in-scanner head motion were used as confounders in all analyses. Post-hoc two-way group interactions between clinical birth status and behavioural data-driven subgrouping classification labels explored whether functional connectivity differences between very preterm and full-term adults varied according to distinct behavioural outcomes.

**Results:** Very preterm compared to full-term adults had poorer scores in selective measures of cognitive and socio-emotional processing and displayed complex patterns of hyper- and hypo-connectivity in subsections of the default mode, visual, and ventral attention networks. Stratifying the study participants in terms of their behavioural profiles (irrespective of birth status), identified two data-driven subgroups: An "At-risk" subgroup, characterised by increased cognitive, mental health, and socio-emotional difficulties, displaying hypo-connectivity anchored in frontal opercular and insular regions, relative to a "Resilient" subgroup with more favourable outcomes. No significant interaction was noted between clinical birth status and behavioural data-driven subgrouping classification labels in terms of functional connectivity.

**Conclusions:** Functional connectivity differentiating between very preterm and full-term adults was dissimilar to functional connectivity differentiating between the data-driven behavioural subgroups. We speculate that functional connectivity alterations observed in very preterm relative to full-term adults may confer both risk and resilience to developing behavioural sequelae associated with very preterm birth, while the localised functional connectivity alterations seen in the "At-risk" subgroup relative to the "Resilient" subgroup may underlie less favourable behavioural outcomes in adulthood, irrespective of birth status.

## 6.2 Introduction

Very preterm birth (VPT; i.e., at <33 weeks' gestation) occurs during a rapid stage of brain development, making those born VPT vulnerable to neurological insult (Volpe, 2009a) and long-lasting difficulties in attention, executive function, and socio-emotional processing (Johnson and Marlow, 2011; Kroll *et al.*, 2017; P. J. Anderson *et al.*, 2021). Functional connectivity alterations in brain regions and networks important for cognitive and affective processing have also been reported in VPT samples across the lifespan, and have been studied amongst the possible biological mechanisms underlying the behavioural difficulties associated with VPT birth (Bäuml *et al.*, 2015; Papini *et al.*, 2016; Rogers *et al.*, 2017; Sylvester *et al.*, 2018; Ramphal *et al.*, 2020; Kanel *et al.*, 2022; Mueller *et al.*, 2022; Siffredi *et al.*, 2022). It is important to highlight, however, that not only have previous studies identified brain changes associated with behavioural difficulties in those born VPT, but have also characterised neural adaptions which support domain-specific performance (Daamen *et al.*, 2014; Finke *et al.*, 2015; Nosarti *et al.*, 2006, 2009; Schafer *et al.*, 2009). These findings, therefore, indicate that the functional reorganisation of the VPT brain has complex implications for outcomes, as it may probe both risk and resilience to behavioural difficulties.

Further complicating the understanding of brain-behavioural relationships in VPT populations, is the fact that those born preterm tend to exhibit heterogenous behavioural outcomes. Previous studies aiming to stratify this heterogeneity implemented latent profile analyses using behavioural measures from both preterm and FT born children (Johnson *et al.*, 2018; Burnett *et al.*, 2019; Lean *et al.*, 2020). Their results indicated that while those born preterm were more likely to present with psychiatric, cognitive, or socio-emotional difficulties, some preterm children displayed distinct profiles characterised by fewer or no behavioural difficulties. Moreover, while FT children predominantly exhibited more normo-typical behavioural profiles, some FT children displayed behavioural difficulties similar to those observed in preterm children (Johnson *et al.*, 2018; Burnett *et al.*, 2019; Lean *et al.*, 2020). Together, these findings indicate that VPT and FT groups exhibit both within- and between-group heterogeneity, which needs to be addressed in order to develop individually tailored and biologically specific interventions aimed at supporting healthy development (Cuthbert and Insel, 2013; Morris *et al.*, 2022). This can be achieved by, firstly, implementing data-driven stratification approaches to identify distinct subgroups of individuals exhibiting similar behavioural profiles, irrespective of their birth status, and secondly, by investigating brain correlates differentiating between the distinct data-driven behavioural subgroups.

Similarly, individuals belonging to distinct diagnostic and non-diagnostic psychiatric groups also exhibit within- and between-group heterogeneity in terms of phenotypic profiles. Recent studies implementing such approaches in psychiatric populations have successfully identified patterns of structural and functional connectivity characterising distinct data-driven behavioural subgroups irrespective of diagnostic labels (Bathelt *et al.*, 2018; Astle *et al.*, 2019; Siugzdaite *et al.*, 2020; Jones, the CALM Team and Astle, 2021; Mareva *et al.*, 2023; Vandewouw *et al.*, 2023). A small number of studies in VPT children followed similar methodological approaches and investigated the underlying brain changes differentiating within-group behavioural heterogeneity. Results of these studies showed that early brain insult (Ross *et al.*, 2016; Bogičević *et al.*, 2021) and structural and functional brain alterations (Lean *et al.*, 2020; Hadaya *et al.*, 2023) characterised the distinct subgroups. However, it remains to be explored whether the heterogeneity in behavioural outcomes seen within and between VPT and FT born individuals persists into adulthood, and if it does, whether resting state functional connectivity (rsFC) changes may be associated with distinct data-driven behavioural phenotypes, irrespective of gestational age at birth.

Our study firstly aimed to identify long-lasting neurodevelopmental alterations associated with VPT birth, by investigating differences in rsFC and behavioural outcomes between VPT and FT born adults. Secondly, our study aimed to delineate behavioural heterogeneity in VPT and FT born adults irrespective of gestational age at birth, by using a robust data-driven consensus clustering approach to stratify participants based on behavioural measures (executive function, attention, intelligence, socio-emotional processing, psychopathology, and autistic traits), and to explore whether resultant data-driven behavioural subgroups would exhibit differences in rsFC. Finally, post-hoc two-way group interactions between clinical (i.e., VPT vs FT birth) and behavioural (i.e., data-driven subgrouping) classification labels were used to explore whether rsFC pattern differences between VPT and FT adults, varied according to distinct behavioural outcomes.

## 6.3 Methods

### 6.3.1 Study design

***Participants.*** VPT infants (i.e., born at <33 weeks of gestation) were recruited at birth from the Neonatal Unit at University College London Hospital (London, UK) between 1979 and 1985. Enrolled participants received cranial ultrasonographic imaging several times during the first week of life and weekly until discharge from hospital (Stewart *et al.*, 1983) and were subsequently followed up in childhood at 1, 4 and 8 years of age (Stewart *et al.*, 1989; Roth *et al.*, 1994), adolescence (15 years), early (20 years), and middle adulthood (30 years) (Karolis *et al.*, 2017). Age-matched controls, born at FT (37-42 weeks of gestation), were recruited from the community in middle adulthood. Exclusion criteria for the controls were any clinical complications at birth (i.e., prolonged gestation at >42 weeks, low birth weight <2500g, receiving endotracheal mechanical ventilation). Exclusion criteria for both VPT and FT participants included severe hearing and motor impairments, or history of neurological complications (i.e., meningitis, head injury, cerebral infections). For this study, we used neuroimaging and behavioural data from the middle adulthood follow-up.

Research study practices were conducted in accordance with the Declaration of Helsinki. Ethical approval was granted by the South London and Maudsley Research and Ethics Committee and the Psychiatry, Nursing and Midwifery Research Ethics Subcommittee (PNM/12/13-10), King's College London. All participants were native English speakers. Written informed consent was obtained from all study participants and participant privacy rights were observed.

***Clinical and socio-demographic details***. Gestational age at birth and birth weight were collected from medical discharge notes for VPT participants. Participants born VPT were classified into three groups, according to cranial ultrasound diagnosis: no evidence of perinatal brain injury (no injury), grade I – II periventricular haemorrhage without ventricular dilation (minor injury) and grade III – IV periventricular haemorrhage with ventricular dilation (major injury) (Nosarti *et al.*, 2002).

For both VPT and FT groups, self-reported ethnicity was recorded according to the following groups: African, Afro-Caribbean, Caucasian/White, Indian Subcontinent, and Other. Socio-economic status was defined according to participants' self-reported occupation at the time of the study and parental occupation at birth. Occupations were categorised according to the Office of National Statistics, 1980 Standard Occupation Classification; I: Higher managerial, administrative, and professional occupations; II: Intermediate occupations, small employers, and own account workers; III: Routine and manual occupations – lower supervisory and technical and semi-routine and routine occupations.

***Cognitive assessments***. The following cognitive assessments were administered to measure language, executive attention, and general intelligence: Hayling Sentence Completion Test (HSCT) (Burgess and Shallice, 1997); Controlled Oral Word Association Test (COWAT-FAS) (Benton, Hamsher and Sivan, 1983); four subtests from the Cambridge Neurophysiological Test Automated Battery (CANTAB) 2003 eclipse version (Fray, Robbins and Sahakian, 1996): 1) Stockings of Cambridge (SOC), 2) Intra-Extra Dimensional Set Shift (IED), 3) Paired Associates Learning (PAL), and 4) Motor Screening Task (MOT); the Trail Making Task – B (TMT-B) (Tombaugh, 2004); Continuous Performance Test – 2nd edition (CPT)(Conners *et al.*, 2003); and Wechsler Abbreviated Scale of Intelligence (WASI) (Wechsler, 1999). Specific task descriptions are detailed in Table SM 6.1.

***Psychiatric and behavioural assessments***. General psychopathology was measured using the Comprehensive Assessment of At- Risk Mental States (CAARMS) (Yung *et al.*, 2005), a semi-structured clinical interview, which measures aspects of psychopathology relating to mania, depression, suicidality and self-harm, mood swings/lability, anxiety, obsessive compulsive disorder symptoms, dissociative symptoms and impaired tolerance to normal stress; scores on the general psychopathology subscale were used in our analyses. The self-administered General Health Questionnaire (GHQ-12) (Goldberg and Williams, 1991) was used to measure general well-being, Peters Delusion Inventory (PDI) (Peters *et al.*, 2004) to measure delusional ideation traits, Autism

Quotient (AQ-10) (Allison, Auyeung and Baron-Cohen, 2012; Booth *et al.*, 2013) to measure autism traits (i.e., social interaction, communication, attention switching, attention to detail and imagination), Social Adjustment Scale (SAS) (Weissman and Bothwell, 1976) to measure participants' satisfaction with their social situation, and Role Functioning Scale (RFS) (Goodman *et al.*, 1993) to measure individuals' ability to function in their daily life. The Emotion recognition task (ERT) (Montagne *et al.*, 2007) was administered to measure participants' ability to recognise expressed emotions (happiness, sadness, surprise, anger, disgust and fear), as described in our previous work (Papini *et al.*, 2016).

***Structural and functional magnetic resonance imaging (MRI) acquisition.*** MRI data were acquired at the Maudsley Hospital, London, using a 3 Tesla Signa MR scanner (General Electric Healthcare). Structural fast spoiled gradient-echo (FSPGR) pulse sequence T1-weighted images were collected using the following sequence parameters: TR=7.1 ms, TE=2.8 ms, matrix=256x256, voxel size=1.1 mm isotropic. Gradient echo EPI resting state functional MRI data were collected while participants stared at a central cross on a screen for 8 minutes 32 s, using the following parameters 256 volumes, TR=2000 ms, TE=30 ms, flip angle=75 degrees, matrix=64x64, 37 non-contiguous slices of 2.4 mm thickness, 1.1 mm interslice gap, and 3.4 mm in-plane resolution.

### 6.3.2 MRI data pre-processing

Resting state functional MRI data pre-processing was performed using fMRIPrep 20.1.1, RRID:SCR_016216 (Esteban *et al.*, 2019), which is based on Nipype 1.5.0, RRID:SCR_002502 (Gorgolewski *et al.*, 2011). In summary, steps included skull stripping, slice-time correction, co-registration to the T1w reference image using boundary-based registration (Greve and Fischl, 2009) and head motion estimation (i.e., global signal and six motion parameters: three translation and three rotation parameters). The complete pre-processing protocol is detailed in the Supplementary Material.

After pre-processing, data were de-noised by regressing out estimated motion confounders (i.e., global signal and six motion parameters: three translation and three rotation parameters) using the FMRIB Software Library (FSL) *fsl_regfilt* command (Jenkinson *et al.*, 2012). A band-pass filter (0.01 – 0.1 Hz) was applied to the data using the AFNI software *3dBandpass* command (Cox, 1996). Participants were excluded if they exhibited excessive in-scanner head motion (i.e., mean frame-wise displacement (FD) exceeding 0.4mm or a maximum FD exceeding 4mm) or had functional

MRI scans showing poor alignment with anatomical data. Sample sizes and participant exclusions are summarised in a flowchart in Figure SM 6.1.

### 6.3.3 Brain parcellation and rsFC estimation

Resting state functional MRI data were parcellated into bilaterally symmetric cortical regions using the Human Connectome Project Multi-Modal Parcellation; HCP-MMP (v1) atlas (Glasser *et al.*, 2016) and bilateral subcortical FreeSurfer regions (Fischl, 2012). The two bilateral hippocampal regions from the HCP-MMP atlas were excluded as these regions were included as part of the FreeSurfer subcortical segmentation, resulting in a total of 374 regions included in our analyses (i.e., 358 HCP-MMP atlas bilateral cortical regions and 16 FreeSurfer bilateral subcortical regions).

An average of the functional MRI blood oxygen level-dependent (BOLD) signal time series across all voxels in each parcellation was used to estimate the regional time series for each of the 374 brain regions. For each participant, rsFC matrices were calculated using Pearson's correlations between pairs of all 374 regional time series. A threshold of 0.2 was used to eliminate weak correlations (i.e., weights of edges with r values ≥ 0.2 were retained) and a Fisher Z-transformation was applied (Buckner *et al.*, 2009; Zalesky *et al.*, 2016; Fenn-Moltu *et al.*, 2022).

### 6.3.4 Consensus clustering

To partition participants (both VPT and FT; n=156) into data-driven behavioural subgroups, a consensus clustering pipeline (Figure 5.1) was implemented using the following 13 behavioural measures as input features: COWAT-FAS mean total words produced, SOC total number of problems solved, IED total errors adjusted, MOT mean reaction time, TMT-B time elapsed, CPT total reaction time, full-scale IQ, total PDI score, total AQ10 score, CAARMS total general psychopathology score, total GHQ score, ERT total number of correct responses and total SAS score (see Supplementary Material for data pre-processing and feature selection procedures).

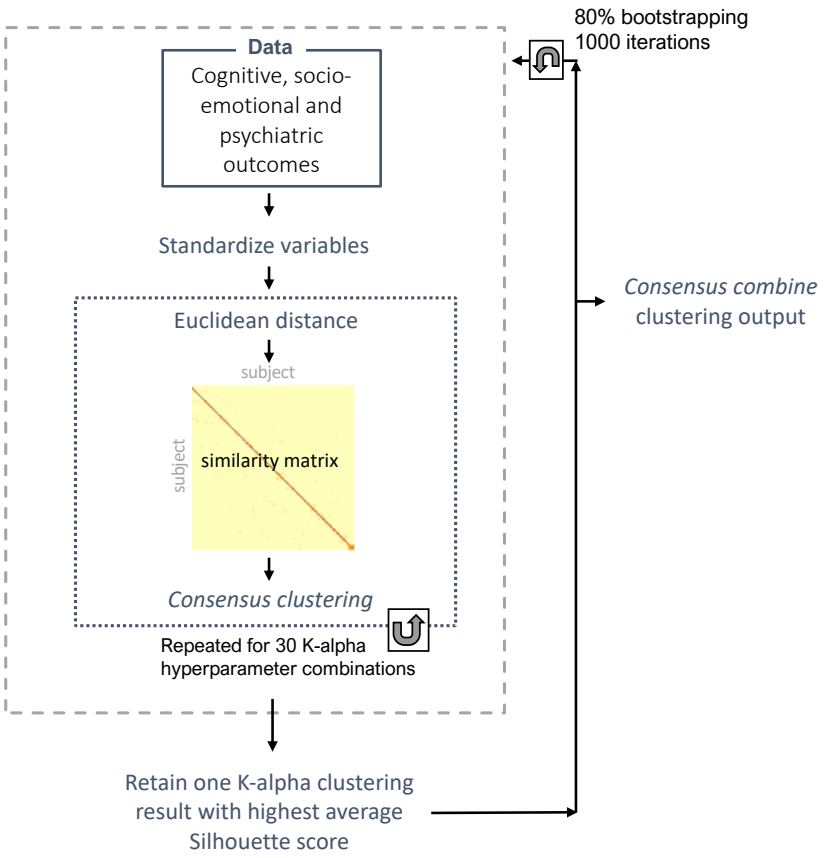



*Figure 6.1. Consensus clustering pipeline followed.*

Each variable was first standardised to have a mean = 0 and standard deviation = 1, and an Euclidean distance matrix of the input data was calculated. A similarity matrix (network) was then calculated from the distance matrix, using the *affinityMatrix* function (SNFtool R package) (Wang *et al.*, 2018), which utilises two hyperparameters: neighbourhood size (K) and alpha (edge weighting parameter) that help increase the signal to noise ratio and in turn improve result validity and reliability. K corresponds to the number of surrounding nodes to consider for each node in the similarity network and alpha determines a threshold for the strength of the edges in the similarity network (i.e., pairwise similarity between nodes within the sample). Greater K values result in more dense similarity networks and smaller values result in more sparse similarity networks, while greater alpha values result in weaker edges being retained and smaller alpha values result in similarity networks which retain edges with higher similarity. Thirty different K-alpha combinations were used to generate thirty similarity networks based on the following values: K = 10, 15, 20, 25, 30 and alpha = 0.3, 0.4, 0.5, 0.6, 0.7, 0.8. These values lie within the ranges recommended in the SNFtool package: 10–30 for K and 0.3–0.8 alpha (Wang *et al.*, 2018). Each of the resultant thirty similarity networks was successively inputted into the consensus clustering algorithm (*ConsensusClusterPlus* function, ConsensusClusterPlus R package) (Wilkerson and Hayes,

2010) which performs agglomerative hierarchical clustering following a nested bootstrapping (n=1000) spectral clustering for each of the thirty similarity networks. From the thirty resultant clustering outputs, the solution with the highest average silhouette width score was retained.

In order to improve the generalisability of our solution and avoid overfitting of hyperparameter selection, the steps described in the above paragraph were repeated 1,000 times where a randomised selection of 80% of the sample was used each time. The final resultant 1,000 clustering outputs were then fed into a hierarchical clustering function (*consensus_combine,* DiceR package) (Chiu and Talhouk, 2018), to output a final consensus clustering result based on the consensus matrix.

To determine the optimal number of clusters (C), Eigengap and Rotation Cost metrics were firstly used to estimate the best and second-best number of clusters (*estimateNumberOfClustersGivenGraph* function SNFtool R package) (Wang *et al.*, 2018) for each of the thirty K-alpha combinations, identifying C=2, C=3, and C=5 as the top three clustering solutions. We then ran the described consensus clustering pipeline three separate times, once for each of these solutions (C=2, C=3, and C=5), and subsequently calculated consensus matrices and silhouette scores for each cluster solution. Resultant consensus matrix and silhouette score outputs suggested an optimal number of clusters of C=2 (Figure SM 6.2); therefore, we evaluate subgroups obtained from the C=2 solution.

The consensus clustering pipeline implemented here is adapted from the integrative clustering method used in our previous work (Hadaya *et al.*, 2023), code: https://github.com/lailahadaya/preterm-ExecuteSNF.C), where we do not apply the data-integration step in the current study.

### 6.3.5 Statistical analyses

#### *6.3.5.1 Evaluation of clinical, socio-demographic, and behavioural profiles*

The non-parametric Wilcox Rank Sum T-test was used for continuous variables and Chi-squared or Fischer's Exact tests for categorical variables. Effect sizes were calculated using Wilcoxon Glass Rank Biserial Correlation for continuous variables and Cramer's V (V) for categorical variables. False Discovery Rate (FDR) was used to account for multiple comparison testing (Benjamini and Hochberg, 1995). Sensitivity analyses using non-parametric permutation testing (5000 permutations) adjusted for potential covariates (age, sex, and socio-economic status) (França, Ge and Batalle, 2022). P-values<0.05 were considered to be statistically significant.

#### *6.3.5.2 Between-group differences in rsFC at a topological network-level*

The Network Based Statistic (NBS), a cluster-based statistics approach, was applied (Zalesky, Fornito and Bullmore, 2010). NBS implements the following steps: 1) mass-univariate testing with a suitable statistical test of interest on all possible connections (i.e., edges), 2) next, only edges with p-values below a pre-defined threshold (p-NBS-Threshold) are maintained, 3) retained suprathreshold edges are then used to identify topologically connected structures (referred to as NBS 'components') present amongst the collection of suprathreshold edges using breadth-first search (Ahuja, Magnanti and Orlin, 1993), and finally, 4) permutation testing is used to assign a Family Wise Error Rate corrected p-value (p-FWER) for each identified component, based on the component's strength. NBS testing is derived from traditional cluster-based thresholding of statistical maps; however, rather than generating clusters of voxels with spatial proximity in physical space, NBS can be applied to graph-like structures to generate clusters with interconnected edges in topological space (Nichols and Holmes, 2002; Zalesky, Fornito and Bullmore, 2010). An advantage of using NBS, compared to an approach that controls for FWER at an edgewise basis (such as False-Discovery Rate), is that it can provide increased statistical power by detecting the effect of interest in a collection of connections which are collectively contributing to the effect of interest as opposed to uniquely contributing to the effect on an individual edgewise-level.

Selecting a threshold in NBS (described in step 2 above) is a relatively arbitrary choice, which can be determined by experimenting with a selection of conservative and stringent thresholds (Zalesky, Fornito and Bullmore, 2010). We ran NBS testing at three different p-value thresholds (i.e., p-NBS-Threshold = 0.05, 0.01, and 0.001) to identify relevant suprathreshold edges to be grouped into NBS components for further analysis. We implemented NBS testing with 1000 permutations using the NBR R package *nbr_lm* function (NBR) (Gracia-Tabuenca and Alcauter, 2020). Statistical models tested included the following covariates: mean FD (as a measure of in-scanner head motion), sex, age, and socio-economic status. The same sets of methods were implemented to identify differences in rsFC between 1) VPT and FT individuals and 2) data-driven behavioural subgroups.

NBS generates two resultant outputs: 1) component strength or intensity – i.e., the sum of test statistic (T-statistic) values from all edges within the significant component, and 2) component size or extent – i.e., the number of connections comprising the significant component. We also calculated the number of connections belonging to each node within the component as a

proportion of the total number of possible edges within that component and presented results graphically using the ggseg3d R package (Mowinckel and Vidal-Piñeiro, 2020). To measure within and between network connectivity, we labelled nodes according to seven previously defined intrinsic connectivity networks (i.e., visual, somatomotor, dorsal attention, ventral attention (VAN), limbic, frontoparietal, and default mode (DMN) networks) (Yeo *et al.*, 2011) and an eighth network comprised of 16 subcortical regions (Váša *et al.*, 2020) and calculated connectivity proportion and strength; code accessible at: https://github.com/frantisekvasa/functional_network_development/blob/master/nspn.fmri.R.

### 6.3.6 Post-hoc exploratory analyses

We estimated the extent of nodal and edgewise overlap between the NBS components characterising clinical (i.e., VPT vs FT birth) and data-driven behavioural subgrouping classifications using Tthe Sørensen-Dice similarity coefficient, which is calculated as the ratio of two times the number of overlapping features between two sets, over the total number of features present across both sets (Sørensen, 1948), with values ranging between 0 and 1.

Post-hoc exploratory NBS analyses investigated whether differences in rsFC between VPT and FT clinical groups varied according to distinct behavioural outcomes, using two-way group interactions between clinical and data-driven behavioural classification labels.

We also investigated differences in early clinical risk (i.e., gestational age at birth, birth weight, and perinatal brain injury) and socio-demographic measures between VPT adults belonging to the distinct data-driven behavioural subgroups, and in socio-demographic measures between FT adults in the distinct data-driven subgroups.

## 6.4 Results

### 6.4.1 VPT and FT groups

The socio-demographic and clinical profiles of VPT and FT adults are summarised in Table 6.1 and their behavioural outcomes in Table 6.2 and Figure 6.2A. In summary, adults born VPT had lower full-scale IQ (WASI), attention set shifting (CANTAB-IED), and emotion recognition (ERT) scores than adults born FT. Head motion during functional MRI acquisition was greater in the VPT (median FD = 0.15mm, range=0.07 – 0.40) than the FT group (median FD=0.12mm, range=0.05 – 0.35; $p<0.001$). Supplementary analyses show that VPT adults excluded from analyses (n=37) for reasons described in Figure SM 6.1, had relatively poorer of

cognitive and socio-emotional scores relative to those VPT adults included in the analyses (n=116) (Table SM 6.2).

*Table 6.1. Clinical and socio-demographic characteristics of study participants.*

| | VPT (n=116) | FT (n=83) | p-value |
|---|---|---|---|
| **Gestational age at birth**, median (range) weeks | 30.00 (24.00 – 32.00) | n/a | n/a |
| **Birth weight**, median (range) grams | 1345 (552 – 2390) | 3440 (2690 – 4990) | < 0.001 |
| **Sex**, n (%) | | | 0.082 |
| *Male* | 66 (56.90%) | 36 (43.37%) | |
| *Female* | 50 (43.10%) | 47 (56.63%) | |
| [a] **Ethnicity**, n (%) | | | 0.139 |
| *African* | 2 (1.72%) | 5 (6.02%) | |
| *Afro-Caribbean* | 2 (1.72%) | 4 (4.82%) | |
| *Caucasian/White* | 76 (65.52%) | 55 (66.27%) | |
| *Indian Subcontinent* | 8 (6.90%) | 2 (2.41%) | |
| *Other* | 4 (3.45%) | 6 (7.23%) | |
| [b] **Perinatal brain injury**, n (%) | | | n/a |
| *No injury* | 62 (53.45%) | n/a | |
| *Minor injury* | 27 (23.28%) | n/a | |
| *Major injury* | 26 (22.41%) | n/a | |
| [c] **Parental socio-economic status at birth**, n (%) | | | 0.106 |
| *I – II* | 43 (51.81%) | 38 (32.76%) | |
| *III* | 36 (43.37%) | 15 (12.93%) | |
| *IV – V* | 8 (9.63%) | 3 (2.59%) | |
| [c] **Participants' current socio-economic status**, n (%) | | | < 0.001 |
| *I – II* | 51 (43.97%) | 36 (43.37%) | |
| *III* | 41 (35.35%) | 26 (31.33%) | |
| *IV – V* | 6 (5.17%) | 0 (0.00%) | |
| *Student* | 1 (0.86%) | 16 (19.28%) | |
| *Unemployed* | 16 (13.8%) | 4 (4.82%) | |
| **Age at assessment**, median (range) years | 31.37 (23.346 – 39.33) | 28.73 (26.26 – 36.49) | < 0.001 |

***Note***. *[a] Ethnicity was self-reported. [b] Ultrasound scans were used to classify perinatal brain injury into three categories: no haemorrhage (no injury), grade I – II periventricular haemorrhage without ventricular dilation (minor injury) and grade III – IV periventricular haemorrhage with ventricular dilation (major injury). [c] Socio-economic status was categorised according to the Office of National Statistics, 1980 occupation classifications. I: Higher managerial, administrative and professional occupations; II: Intermediate occupations, small employers and own account workers; III: Routine and manual occupations – lower supervisory and technical and semi-routine and routine occupations. Missing data: 1 VPT and 1 FT had missing socio-economic status data; 24 VPT and 11 FT had missing ethnicity data; 1 VPT has missing perinatal brain injury classification.*

*Table 6.2. Behavioural outcomes in VPT and FT adults.*

| | VPT (n=116) | FT (n=83) | p-value | FDR p-value | Adj. FDR p-value | Effect size |
|---|---|---|---|---|---|---|
| **COWAT, total words** | 13.00 (5.75) | 14.00 (5.25) | 0.052 | 0.166 | 0.115 | -0.042 |
| **CANTAB – SOC, problems solved** | 9.00 (2.75) | 10.00 (2.00) | 0.063 | 0.166 | 0.106 | -0.064 |
| **CANTAB – IED, total errors adjusted** | 15.00 (25.50) | 10.50 (14.65) | 0.002 | 0.007 | 0.021 | 0.184 |
| **TMT-B, time to finish task** | 73.50 (40.50) | 71.30 (39.05) | 0.081 | 0.175 | 0.068 | -0.093 |
| **CPT, total reaction time for correct responses** | 417.50 (59.15) | 414.00 (54.40) | 0.921 | 0.921 | 0.936 | -0.009 |
| **WASI – full scale IQ** | 106 .00(13.75) | 113.50 (12.25) | <0.001 | <0.001 | <0.001 | 0.088 |
| **CANTAB – MOT, reaction time** | 691.00 (200.80) | 734.00 (196.90) | 0.307 | 0.399 | 0.456 | 0.062 |
| **PDI, total score** | 21.50 (50.25) | 18.00 (39.25) | 0.406 | 0.480 | 0.530 | 0.002 |
| **AQ10, total score** | 2.00 (2.44) | 3.00 (2.32) | 0.198 | 0.322 | 0.257 | 0.121 |
| **CAARMS, general psychopathology score** | 2.00 (5.50) | 2.00 (4.00) | 0.232 | 0.335 | 0.220 | -0.111 |
| **GHQ, total score** | 10.00 (6.00) | 10.00 (7.00) | 0.891 | 0.921 | 0.943 | 0.070 |
| **ERT, total correct** | 56.60 (11.15) | 62.00 (9.45) | <0.001 | <0.001 | <0.001 | 0.358 |
| **SAS, total score** | 1.58 (0.45) | 1.69 (0.53) | 0.127 | 0.236 | 0.4021 | 0.136 |

***Note.*** *Median (interquartile range) reported. "Adj. FDR p-value" corresponds to the p-value after adjusting for covariates (sex, age, socio-economic status) and correcting for multiple comparisons with FDR. Effect sizes are calculated using Wilcoxon Glass Rank Biserial Correlation. Missing data: [a] FT n=7, VPT n=22; [b] FT n=21, VPT n=22; [c] FT n=21, VPT n=19; [d] FT n=12, VPT n=17; [e] FT n=5, VPT n=9. Abbreviations. AQ10 = Autism Quotient; CANTAB = Cambridge Neurophysiological Test Automated Battery; CAARMS = Comprehensive Assessment of At- Risk Mental States; COWAT = Controlled Oral Word Association Test; CPT = Continuous Performance Test; ERT = Emotion Recognition Task; FT= full-term; GHQ = General Health Questionnaire; IED = Intra-Extra Dimensional Set Shift; MOT = Motor Screening Task; PDI = Peters Delusion Inventory; SAS = Social Adjustment Scale; SOC = Stockings of Cambridge; TMT-B = Trail Making Task B; VPT = very preterm; WASI = Wechsler Abbreviated Scale of Intelligence.*

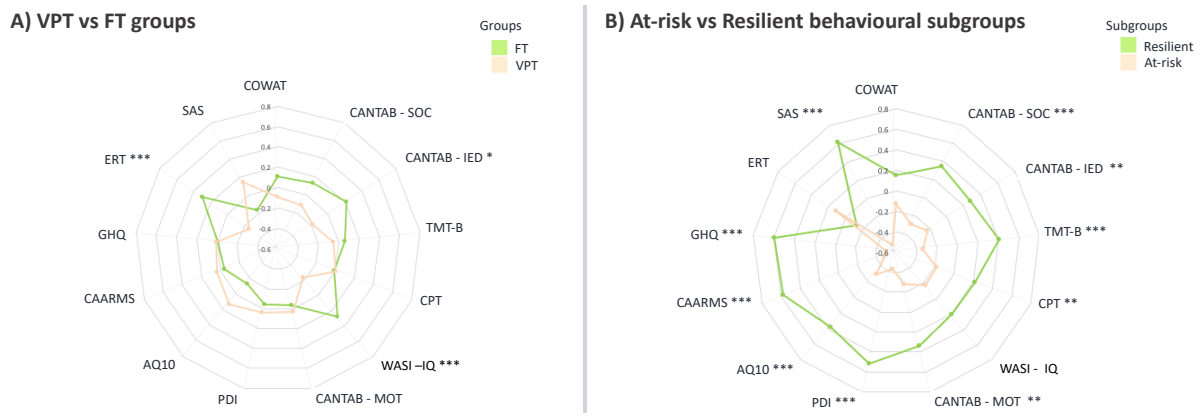



***Figure 6.2. Radar plots showing differences in behavioural profiles between A) VPT and FT adults and B) At-risk and Resilient data-driven behavioural subgroups.***

*Z-scores were computed for each group and plotted accordingly. For visual illustrative purposes, values for scales indicating poorer outcomes were reversed, so that larger Z-scores here indicate generally more optimal outcomes. *=p<0.05; **=p<0.01; ***p<0.001. Refer to Table 6.2 legend for behavioural measure abbreviations and Table SM6.1 for descriptions.*

### 6.4.2 Data-driven behavioural subgroups

Two data-driven behavioural subgroups were identified and labelled as 'At-risk' and 'Resilient', based on their observed phenotypic profiles (Table 6.3; Figure 6.2B).

*Table 6.3. At-risk and Resilient behavioural subgroup profiles.*

| | Subgroup 1 – Resilient (n=71) | Subgroup 2 – At-risk (n=85) | p-value | FDR p-value | Adj. FDR p-value | Effect size |
|---|---|---|---|---|---|---|
| **Age at assessment, years** | 29.83 (4.16) | 30.22 (4.47) | 0.972 | 0.972 | n/a | -0.004 |
| **Framewise Displacement, mm** | 0.13 (0.07) | 0.13 (0.06) | 0.654 | 0.690 | 0.575 | -0.042 |
| **COWAT, total words** | 14.00 (5.50) | 13.00 (4.00) | 0.071 | 0.097 | 0.117 | 0.168 |
| **CANTAB – SOC, problems solved** | 10.00 (2.00) | 9.00 (2.00) | < 0.001 | < 0.001 | < 0.001 | 0.371 |
| **CANTAB – IED, total errors adjusted** | 10.00 (11.00) | 18.00 (26.60) | 0.002 | 0.004 | 0.002 | -0.289 |
| **TMT-B, time to finish task** | 61.00 (25.20) | 78.00 (39.00) | < 0.001 | < 0.001 | < 0.001 | -0.428 |
| **CPT, total reaction time for correct responses** | 406.00 (51.30) | 421.00 (61.40) | 0.005 | 0.009 | 0.008 | -0.260 |
| **WASI – full scale IQ** | 112.00 (15.50) | 108.00 (14.00) | 0.038 | 0.059 | 0.008 | 0.194 |
| **CANTAB – MOT, reaction time** | 675.00 (171.50) | 741.00 (255.00) | < 0.001 | < 0.001 | < 0.001 | -0.341 |
| **PDI, total score** | 13.00 (16.50) | 41.80 (45.00) | < 0.001 | < 0.001 | < 0.001 | -0.596 |
| **AQ10, total score** | 2.00 (1.92) | 3.00 (2.71) | < 0.001 | < 0.001 | < 0.001 | -0.385 |
| **CAARMS, general psychopathology score** | 0.00 (2.00) | 4.60 (4.20) | < 0.001 | < 0.001 | < 0.001 | -0.654 |
| **GHQ, total score** | 8.00 (2.00) | 13.00 (6.00) | < 0.001 | < 0.001 | < 0.001 | -0.663 |
| **ERT, total correct** | 58.40 (12.60) | 60.00 (9.00) | 0.112 | 0.142 | 0.132 | -0.148 |
| **SAS, total score** | 1.44 (0.26) | 1.81 (0.50) | < 0.001 | < 0.001 | < 0.001 | -0.691 |
| **Birth status, n (%)** | | | 0.558 | 0.623 | n/a | V = 0.060 |
| *VPT* | 41 (57.75%) | 44 (51.767%) | | | | |
| *FT* | 30 (42.25%) | 41 (48.24%) | | | | |
| **Sex, n (%)** | | | 0.169 | 0.200 | n/a | V = 0.123 |
| *Male* | 43 (60.56%) | 41 (48.24%) | | | | |
| *Female* | 28 (39.44%) | 44 (51.77%) | | | | |
| [a] **Participants' current socio-economic status, n (%)** | | | < 0.001 | 0.001 | n/a | V = 0.365 |
| *I – II* | 46 (64.79%) | 30 (35.29%) | | | | |
| *III* | 21 (29.58%) | 31 (36.47%) | | | | |
| *IV – V* | 0 (0.00%) | 2 (2.35%) | | | | |
| *Student* | 1 (1.41%) | 11 (12.94%) | | | | |
| *Unemployed* | 2 (2.82%) | 11 (12.94%) | | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| [a] **Parental socio-economic status at birth, n (%)** | | | 0.055 | 0.080 | n/a | V = 0.208 |
| *I – II* | 44 (61.97%) | 33 (38.82%) | | | | |
| *III* | 16 (22.53%) | 30 (35.29%) | | | | |
| *IV – V* | 5 (7.04%) | 6 (7.06%) | | | | |

***Note.*** *Median (interquartile range) reported unless stated otherwise where number of participants (n) is reported alongside percentage (%). "Adj. FDR p-value" corresponds to the p-value after adjusting for covariates (sex, age, socio-economic status) and correcting for multiple comparisons with FDR. Effect sizes are calculated using Wilcoxon Glass Rank Biserial Correlation, unless otherwise stated. Cramer's V (V) effect size was used for categorical variables.* [a] *defined in Table 6.1. Abbreviations: as defined in Table 6.2.*

In summary, the At-risk subgroup had less optimal executive function and attention scores probing spatial planning, attentional set shifting, visuo-motor coordination, comprehension abilities, sustained attention and response inhibition (CANTAB – SOC, MOT and IED, the TMT-B, and CPT), compared to the Resilient subgroup. The At-risk subgroup also had less optimal social adjustment, mental wellbeing, and psychiatric scores (PDI, CAARMS, GHQ, and SAS), and increased autistic traits (AQ-10 scores), compared to the Resilient subgroup. The two subgroups showed no differences in full-scale IQ (WASI), emotion recognition (ERT), or phonemic verbal fluency (COWAT). However, the At-risk subgroup had a higher proportion of individuals with lower own socio-economic status compared to the Resilient subgroup. Parental socio-economic status did not differ between the subgroups.

52% of the VPT adults in our sample clustered into the At-risk subgroup, and the remaining 48% into the Resilient subgroup (Figure 6.3). Upon examining VPT adults only, there were no significant differences between the At-risk and Resilient subgroups in terms of perinatal clinical measures (i.e., gestational age, birth weight, or perinatal brain injury) (Table SM 6.3; Figure SM 6.3). In terms of parental socio-economic status, there were no differences between At-risk and Resilient subgroups within VPT or FT adults (Table SM 6.3 and Table SM 6.4, respectively). As for participants' own socio-economic status, only those born VPT displayed differences between the data-driven behavioural subgroups, where more VPT individuals with higher managerial, administrative, and professional occupations belonged the Resilient subgroup compared to the At-Risk subgroup (Table SM 6.3). However, socio-economic status for those born FT did not differ significantly between the two data-driven subgroups (Table SM 6.4).

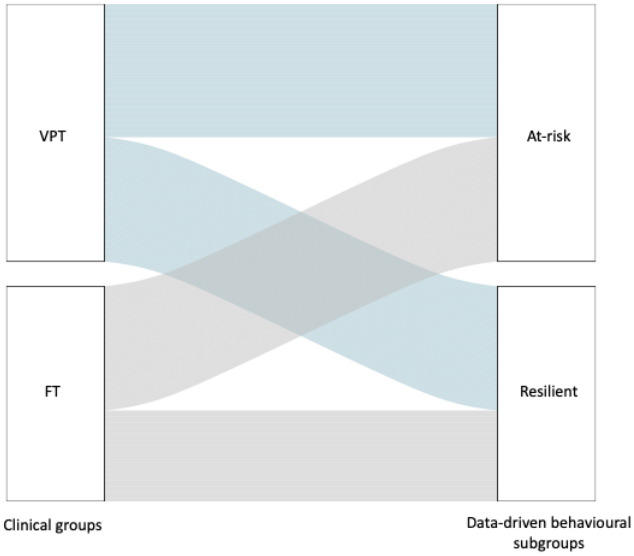



*Figure 6.3. Alluvial plot showing VPT (in blue) and FT (in grey) adults clustering into the At-risk and Resilient data-driven behavioural subgroups.*

### 6.4.3 Between-group differences in rsFC

We report NBS analyses using p-NBS-Threshold values powered to detect a significant effect, whilst also reducing component size (i.e., not p = 0.05) (Table SM 6.5). Main results reported here are from one-tailed NBS analyses using p-NBS-Threshold = 0.01, and additional sensitivity analyses investigating rsFC using a more stringent threshold (p-NBS-Threshold = 0.001) are reported in Supplementary Material.

**VPT < FT.** NBS results showed weaker rsFC in the VPT group compared to the FT group (i.e., VPT < FT) in one component comprising 360 nodes (i.e., 96.25% of all regions) and 1467 edges (i.e., 2.10% of the 69,751 possible connections), with a component strength of 616.04 (p-FWER value = 0.007). Regions included in this component were widespread across the brain (Figure 6.4A; Table SM 6.6). Nodes with the highest number of connections within the component (i.e., component 'hub' regions) were predominantly localised to superior temporal gyrus, inferior and superior parietal cortex, inferior frontal, orbitofrontal, anterior cingulate and medial prefrontal cortex, inferior premotor, a lateral occipital/posterior temporal visual area, dorsolateral prefrontal cortex, medial and lateral temporal, and posterior cingulate cortex. Component within- and between-network connectivity was highest in the DMN (Figure 6.5A).

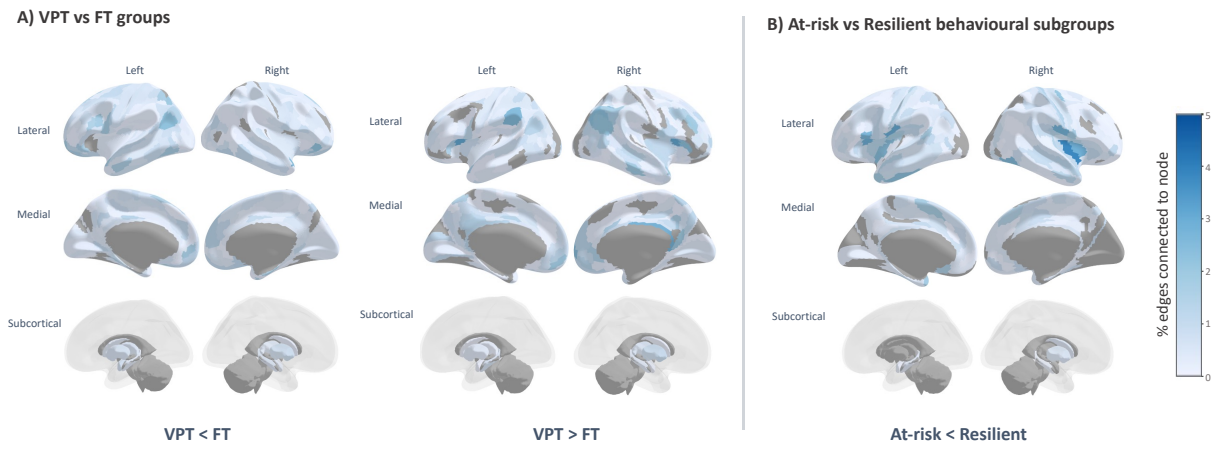



***Figure 6.4. Percentage of edges connected to each region (i.e., node) within the significant NBS components for A) VPT vs FT groups and B) At-risk vs Resilient behavioural subgroups.***

*Main analysis results from NBS modelling using a p-NBS-Threshold of 0.01, 1000 permutations, and linear models correcting for covariates (age, sex, in-scanner head motion, and socio-economic status). Darker colours (blue) denote higher percentages of edges and lighter colours (light blue and white) denote lower percentages, with areas marked in grey indicating regions that are not forming part of the NBS component.*

**VPT > FT.** NBS results also showed greater rsFC in the VPT group compared to the FT group (i.e., VPT > FT) in one component comprising 340 nodes (i.e., 90.91% of regions), 962 edges (i.e., 1.37% of possible connections) and component strength of 358.03 (p-FWER value < 0.001). 'Hub' regions within this component were less widespread across the brain and localised within posterior opercular cortex, posterior cingulate cortex, inferior parietal cortex, right orbitofrontal cortex, bilateral anterior cingulate and medial prefrontal cortex, superior temporal gyrus (auditory association cortex), dorsolateral prefrontal cortex, right lateral temporal cortex, right temporo-parietal-occipital junction, and medial superior parietal cortex (Figure 6.4A; Table SM 6.7). The highest number of connections found in the component were within the DMN itself, followed by a moderate number of widespread connections in the VAN, and especially between the VAN and the visual network.

A total of 326 nodes (i.e., 87.17% of regions) were present in both VPT < FT and VPT > FT components; however, the sets of edges connecting nodes within each component were mutually exclusive with no overlapping edges.

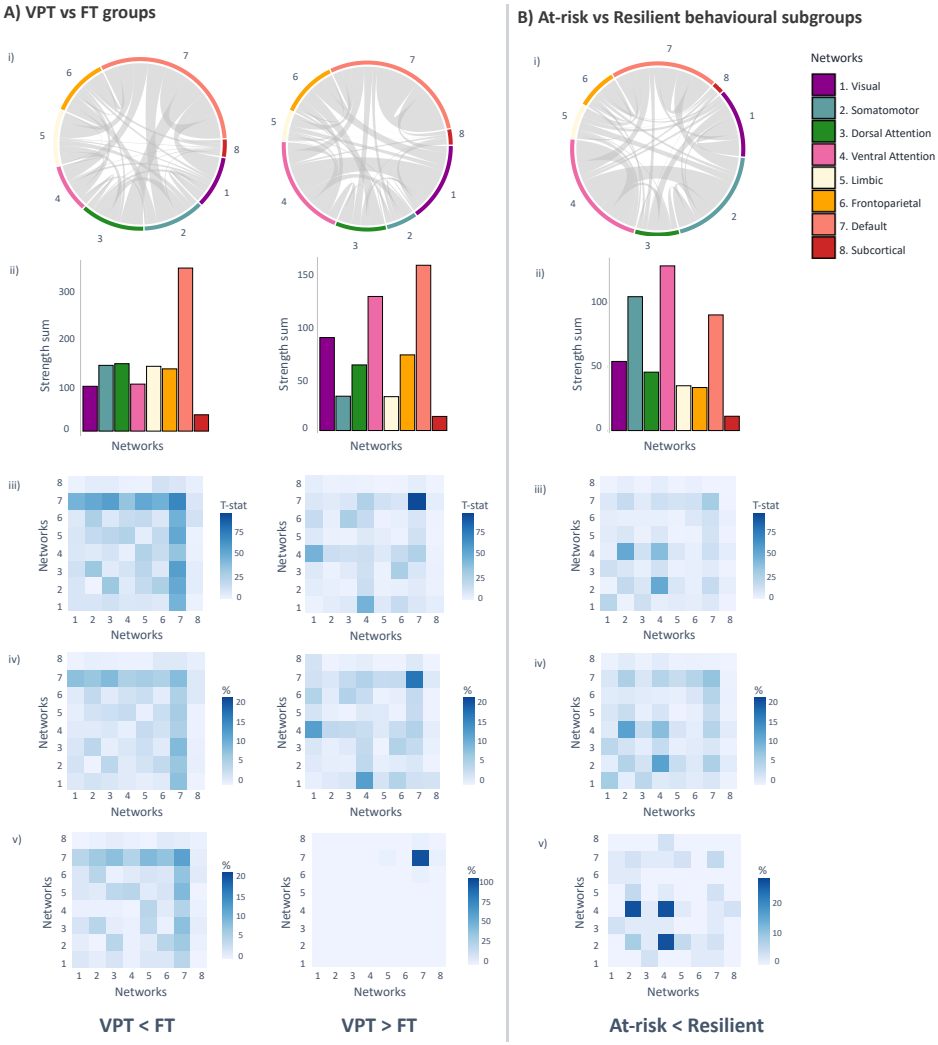



***Figure 6.5. Within- and between-network connectivity of the significant NBS components in A) VPT vs FT groups and B) At-risk vs Resilient behavioural subgroups.***

*Results from main NBS analyses using a p-NBS-threshold of 0.01: i) circle plots illustrating within- and between-network connections within the significant component only; ii) bar plots showing the sum of T-statistic strength values within the significant NBS component belonging to the different intrinsic connectivity networks (i.e., seven Yeo networks and an eighth network of subcortical regions), and iii) within- and between-network connectivity strength (T-statistic sum). Heatmaps showing total number of within- and between-network connections as a percentage of the total number of connections forming the significant component: iv) at p-NBS-threshold = 0.01, and v) p-NBS-threshold = 0.001.*

**At-Risk < Resilient.** Contrasts testing for lower rsFC in the At-risk compared to the Resilient subgroup identified one significant NBS component with 337 nodes (i.e., 90.11% of regions), 832 edges (i.e., 1.19% of possible connections) and a strength sum of 309.04 (p-FWER = 0.019). Hub regions with the highest number of connections within the component were predominantly located in insular, frontal opercular, and posterior opercular cortex (Figure 6.4B; Table SM 6.8). Other hub regions were found in the left inferior frontal cortex, lateral temporal cortex, right temporo-occipital visual area, left temporo-parieto-occipital junction, anterior cingulate, medial prefrontal cortex, left supplementary motor area, primary somatosensory cortex,

and the superior temporal sulcus (auditory association cortex) (Figure 6.4B; Table SM 6.8). Component within- and between-network connectivity were most pronounced between the VAN and somatomotor networks, and within the VAN (Figure 6.5B).

**At-Risk > Resilient.** No significant NBS components were detected when testing for higher rsFC in the At-risk compared to the Resilient subgroup.

Confirming the robustness of the observed effects from analyses using a p-NBS-threshold of 0.01, sensitivity NBS analyses using a more stringent p-NBS-threshold of 0.001 reported significant components with greater sparsity (Table SM 6.9), but largely similar rsFC patterns (Figure 6.4; Figure 6.5Av; Figure 6.5Bv).

Post-hoc exploratory analyses investigating the interaction between clinical (VPT vs FT) groups and data-driven behavioural subgroups (At-risk vs Resilient) on rsFC did not identify significant components (p-FWER > 0.05) at any p-NBS-Threshold examined (0.05, 0.01, and 0.001). Similarity index calculations indicated that the At-risk < Resilient component had a high number of nodes, which were also part of the VPT < FT component (n=325; Sørensen-Dice = 0.93) and the VPT > FT component (n=304; Sørensen-Dice = 0.90), but very few edges overlapped with either clinical component; n=9 edges (Sørensen-Dice = 0.01) and n=22 edges (Sørensen-Dice = 0.03), respectively.

## 6.5 Discussion

In this study, we compared rsFC between groups of adults stratified in terms of (i) their clinical characteristics (i.e., VPT and FT birth) as well as (ii) their behavioural profiles identified using data-driven consensus clustering, regardless of their gestational age at birth. In VPT compared to FT adults, we identified complex preterm-specific patterns of both increased and decreased intrinsic rsFC predominately characterised by hypo-connectivity between the DMN and other networks examined and hyper-connectivity within the DMN and between the VAN and the visual network. When VPT and FT born adults were stratified in terms of their data-driven behavioural profiles, irrespective of gestational age at birth, we identified an 'At-risk' subgroup with more behavioural difficulties and reduced rsFC anchored in frontal opercular and insular areas of the VAN, relative to a 'Resilient' subgroup with more favourable behavioural outcomes.

In summary, our results indicate that there are complex and widespread long-lasting preterm-specific rsFC alterations, which we speculate may confer both risk and resilience to the behavioural sequelae associated with VPT birth. That is, while these rsFC alterations may partly

explain the behavioural difficulties specific to those born VPT in cognitive and socio-emotional processing observed here, they may also aid the preservation of optimal outcomes in other behavioural domains where no between-group differences were noted (e.g., psychiatric difficulties, sustained attention, planning or phonemic verbal fluency). On the other hand, localised functional hypo-connectivity anchored in insular and frontal opercular regions observed in our study may characterise participants with unfavourable compared to favourable cognitive and behavioural outcomes, irrespective of birth status.

### 6.5.1 Differences in rsFC and behavioural outcomes between VPT and FT born adults

We identified complex patterns of both hypo- and hyper-connectivity predominantly located in the DMN, VAN, and visual networks in VPT compared to FT participants. Such rsFC alterations are evident in adulthood and may represent the neurobiological architecture underlying the attentional, cognitive, and socio-emotional processing difficulties associated with VPT birth, commonly referred to as the "preterm behavioural phenotype" (Johnson and Marlow, 2011). However, in our cohort, VPT relative to FT born adults only differed in selected dimensions that have been studied as part of the "preterm behavioural phenotype"; they had lower full-scale IQ, difficulties in rule learning, attentional set shifting abilities (measured by the CANTAB IED), and emotion recognition.

VPT adults, compared to controls, displayed functional hypo-connectivity between the DMN and the visual, somatomotor, dorsal attention, limbic and frontoparietal networks, as well as hyper-connectivity within the DMN itself. In line with our findings, patterns of both hyper- and hypo-connectivity in the DMN have been previously reported in VPT born children and adults (Bäuml *et al.*, 2015; Degnan *et al.*, 2015b; Wheelock *et al.*, 2021; Mossad *et al.*, 2022), suggesting that functional DMN connectivity alterations may be characteristic of VPT samples. Functional DMN connectivity emerges during the third trimester of gestation, a critical period of brain development during which VPT infants are born, and previous studies have reported structural and functional brain alterations at term-equivalent age in regions belonging to the DMN (Doria *et al.*, 2010; Smyser *et al.*, 2010, 2016; Sa de Almeida *et al.*, 2021; Scheinost *et al.*, 2022). Extending beyond preterm populations, functional alterations in the DMN have been described in several psychiatric conditions, including schizophrenia, anxiety, and mood disorders (Buckner, 2013; Doucet *et al.*, 2020), suggesting that the DMN rsFC alterations observed in VPT individuals may represent neurobiological changes which could contribute to the behavioural difficulties associated with VPT birth.

On the other hand, alterations in DMN rsFC have also been studied as adaptive neural mechanisms; for instance, maintaining attentional capture (i.e., less distractibility) in male veterans (Poole *et al.*, 2016). Such findings suggest that functional reorganisation of the DMN may also reflect compensatory biological alterations supporting selective cognitive and behavioural processing in VPT individuals; in this context referring to the behavioural outcomes where no between-group differences were noted in our study sample, including spatial planning (CANTAB – SOC), coordination (MOT), cognitive flexibility (TMT-B), phonemic verbal fluency (COWAT), sustained attention (CPT), social adaptation (SAS), prodromal symptoms (PDI), autism traits (AQ10) and general psychopathology (CAARMS and GHQ). This finding emphasises the notion that complex neurobiological alterations following VPT birth may confer both risk and resilience to the long-term consequences of VPT birth. Further supporting this point, we also identified patterns of hyper-connectivity in the VPT relative to the FT group in the VAN, a "circuit-breaker" network which disengages during tasks requiring focused attention and activates to redirect attention towards external task-irrelevant stimuli (Corbetta and Shulman, 2002; Vossel, Geng and Fink, 2014). Notably, the highest proportion of connections were between the VAN and the visual network, which may reflect adaptive functional reorganisation in the VPT group. In a previous study, stronger rsFC changes in visual and attention networks were associated with fewer attention deficits in visual short-term memory storage in VPT relative to FT adults (Finke *et al.*, 2015). Another study found that attention processing was selectively supported by VAN and visual network connectivity in VPT born children, and by dorsal attention, frontoparietal, and cingulo-opercular network connectivity in FT controls (Wheelock *et al.*, 2021). The authors argued that VPT children may have a greater reliance on visually stimulated "bottom-up" neural processes to maintain attention mechanisms, which is in line with their previous findings showing poorer attention abilities in VPT children with reduced volumes in regions of the visual network (Lean *et al.*, 2017).

We also identified that component 'hub' regions (i.e., those with a high percentage of connections within the component) with higher rsFC in the VPT group relative to the FT group, were localised to brain regions previously identified as nodes of a 'rich-club' network (i.e., the sub-network of highly connected brain regions which are also highly connected to one another), important for efficient integration and transfer of information between systems (van den Heuvel and Sporns, 2013; Grayson *et al.*, 2014). We previously reported stronger rich-club network structural connectivity and weaker peripheral connectivity in an overlapping sample of VPT adults compared to FT controls, and argued that increased resources in the VPT brain may be

preferentially allocated to the rich-club network in order to maintain efficient information exchange across the brain (Karolis *et al.*, 2016).

### 6.5.2 Differences in rsFC and behavioural outcomes between data-driven behavioural subgroups

Considering the neurodevelopmental heterogeneity exhibited within and between those born VPT and FT, it remains to be established whether we can use rsFC to characterise the prevalent behavioural difficulties observed in VPT individuals (Nosarti *et al.*, 2012; P. J. Anderson *et al.*, 2021). Aiming to address this question, we stratified VPT and FT adults into data-driven behavioural subgroups and investigated specific rsFC alterations which may differentiate between them. We identified two data-driven behavioural subgroups, irrespective of birth status (VPT and FT): an 'At-risk' subgroup with more executive function, attention, socio-emotional, and psychiatric difficulties, compared to a 'Resilient' subgroup, with more favourable behavioural outcomes. Notably, the behavioural differences observed between data-driven subgroups were more pronounced than those observed between VPT and FT adults.

We also identified underlying rsFC differences characterising the distinct data-driven behavioural subgroups, where the At-risk, compared to the Resilient subgroup, displayed hypo-connectivity within the VAN and between the VAN and the somatomotor network. Specifically, the predominant connectivity patterns forming this component were anchored in frontal opercular and insular regions of the brain, which play an integral role in detecting bottom-up salient information from the environment and switching between networks to produce appropriate cognitive control, socio-emotional, and interoceptive somatomotor responses (Menon and Uddin, 2010; Deen, Pitskel and Pelphrey, 2011; Higo *et al.*, 2011; Loitfelder *et al.*, 2015; Uddin *et al.*, 2017; Quirmbach and Limanowski, 2022). Our findings here are in line with previous studies showing structural and functional alterations in insular and opercular regions in adults experiencing mental health difficulties (Horn *et al.*, 2010; Yin *et al.*, 2018) and executive dysfunction (Hausman *et al.*, 2022). Furthermore, studies investigating rsFC across multiple psychiatric groups identified transdiagnostic patterns of hypo-connectivity in lower-order networks, such as the somatomotor network, as well as higher order networks, such as the VAN (Baker *et al.*, 2019; Li *et al.*, 2021). The rsFC patterns identified here characterised data-driven behavioural subgroups irrespective of gestational age at birth (VPT and FT), indicating that these specific neural mechanisms may represent biomarkers of behavioural outcomes in the general population which are not unique to VPT individuals. We also found no significant interaction effects between birth group (VPT vs

FT) and data-driven behavioural subgroups (At-risk vs Resilient) on rsFC and very little overlap in rsFC between the clinical and behavioural components identified by NBS, which may further support our speculation that the differences in rsFC between the data-driven subgroups may be characterising behavioural outcomes independently of gestational age at birth. However, future studies with larger samples, and hence greater statistical power, may further investigate the possible influence of VPT (vs FT) birth on the relationship between rsFC alterations and behavioural outcomes.

Our post-hoc analyses aimed to explore whether specific enriching factors, or lack of certain social or clinical risk factors, protected the VPT adults belonging to the Resilient subgroup from developing an At-risk behavioural profile. In contrast to previous studies in VPT children, we found that perinatal clinical risk was not higher in VPT adults who belonged to an At-risk (vs Resilient) subgroup (Poehlmann *et al.*, 2015; Hadaya *et al.*, 2023). Social risk, on the other hand, may be specifically related to the difficulties observed in the VPT At-risk subgroup, which contained more VPT adults from more socially disadvantaged backgrounds compared to the Resilient subgroup, while this relationship was not observed in FT adults. These findings as well as previous studies in children (Ross *et al.*, 2016; Lean *et al.*, 2020; Vanes *et al.*, 2021; Hadaya *et al.*, 2023) could be interpreted within a "differential susceptibility" framework, which posits that vulnerable individuals (e.g., those born VPT) are particularly sensitive to environmental influences, where negative or positive factors (such as social (dis)advantage) can promote either worse or more optimal outcomes, respectively (Belsky, Bakermans-Kranenburg and van IJzendoorn, 2007). Therefore, VPT adults in the At-risk subgroup may have experienced a "double-hit" of being born VPT as well as being socio-economically disadvantaged. Nonetheless, it is worth noting, that socio-economic status in our sample only partially explained behavioural outcomes, as our main behavioural and rsFC results remained significant after adjusting for this covariate. It is therefore plausible that additional unmeasured environmental or hereditary factors (e.g., parental mental health or cognitively stimulating home environment) (Lean *et al.*, 2020; Vanes *et al.*, 2021; Hadaya *et al.*, 2023) may have contributed to the behavioural outcomes observed in the distinct subgroups.

This study has several strengths, which include the use of a large sample of both VPT and FT born controls, the implementation of rigorous consensus clustering methods to obtain data-driven behavioural subgroups, as well as the use of fMRIPrep, a robust automated resting state functional MRI pre-processing pipeline which promotes pre-processing transparency and aims to alleviate hurdles related to reproducibility in functional MRI analyses (Pernet and Poline, 2015; Esteban *et al.*, 2019). We also acknowledge several limitations to our study. After excluding

participants with excessive head motion, behavioural outliers, missing data, or poor alignment of functional MRI data, supplementary analyses showed that the subsample of VPT adults used in our analyses had relatively better cognitive and socio-emotional processing outcomes in comparison to VPT adults excluded from the analyses. This may limit the generalisability of our results to cohorts of low-risk VPT adults with relatively favourable behavioural outcomes. It may also explain why our two data-driven behavioural subgroups have similar proportions of VPT and FT born individuals, which is not in line with previous studies in children that have reported higher ratios of VPT to FT individuals belonging to At-risk subgroups and lower ratios to Resilient subgroups (Burnett *et al.*, 2019; Lean *et al.*, 2020). On the other hand, our results may be reflective of the increased rates of mental health difficulties with increasing age, which may not yet be apparent in childhood (Otto *et al.*, 2021; Solmi *et al.*, 2022). Future studies with more representative samples of VPT adults could help elucidate these potentially inconsistent findings. Another possible limitation is that we did not include known risk factors (such as socio-economic status, parenting or clinical measures) in the clustering model, which may have increased the difficulty in identifying nuanced subgroups exhibiting 'equifinal' trajectories (i.e., those with similar behavioural outcomes but distinct underlying risk factors) (Cicchetti and Rogosch, 1996; Hadaya *et al.*, 2023). However, to our knowledge, this is the first study to parse behavioural heterogeneity in VPT adults; therefore, we decided to follow an approach similar to those implemented in the vast majority of studies in VPT children, where individual-level behavioural variables were included as inputs to the clustering model and risk factors were explored post-hoc (Poehlmann *et al.*, 2015; Ross *et al.*, 2016; Johnson *et al.*, 2018; Burnett *et al.*, 2019; Lean *et al.*, 2020; van Houdt *et al.*, 2020; Bogičević *et al.*, 2021).

In summary, this study shows that there are complex patterns of rsFC alterations which are specifically associated with VPT birth in adult life. We speculate that these alterations may reflect neural adaptations conferring both risk and resilience to the long-term sequelae of VPT birth. We also identify distinct rsFC alterations in insular and frontal opercular regions in a data-driven At-risk relative to a Resilient behavioural subgroup, irrespective of birth status (VPT vs FT), indicating that these neurobiological changes may reflect biomarkers of behavioural outcomes in the general population that are not unique to those born VPT.

## 6.6 Supplemental Information

### 6.6.1 Anatomical and functional MRI data pre-processing with fMRIPrep

The full description of the anatomical and functional data pre-processing pipeline below is extracted from the boilerplate automatically generated by fMRIPrep (released under the CC0 license).

***Anatomical data pre-processing.*** The T1-weighted (T1w) image was corrected for intensity non-uniformity (INU) with N4BiasFieldCorrection (Tustison *et al.*, 2010), distributed with ANTs 2.2.0 (RRID:SCR_004757, (Avants *et al.*, 2008b)), and used as T1w-reference throughout the workflow. The T1w-reference was then skull-stripped with a Nipype implementation of the antsBrainExtraction.sh workflow (from ANTs), using OASIS30ANTs as target template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using fast (FSL 5.0.9, RRID:SCR_002823, (Zhang, Brady and Smith, 2001)). Brain surfaces were reconstructed using recon-all (FreeSurfer 6.0.1, RRID:SCR_001847, (Dale, Fischl and Sereno, 1999)), and the brain mask estimated previously was refined with a custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived segmentations of the cortical gray-matter of Mindboggle (RRID:SCR_002438, (Klein *et al.*, 2017)). Volume-based spatial normalization to one standard space (MNI152NLin2009cAsym) was performed through nonlinear registration with antsRegistration (ANTs 2.2.0), using brain-extracted versions of both T1w reference and the T1w template. The following template was selected for spatial normalization: ICBM 152 Nonlinear Asymmetrical template version 2009c [(Fonov *et al.*, 2011), RRID:SCR_008796; TemplateFlow ID: MNI152NLin2009cAsym].

***Functional data pre-processing.*** For each of the 1 BOLD runs found per subject (across all tasks and sessions), the following preprocessing was performed. First, a reference volume and its skull-stripped version were generated using a custom methodology of fMRIPrep. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) are estimated before any spatiotemporal filtering using mcflirt (FSL 5.0.9, (Jenkinson *et al.*, 2002). BOLD runs were slice-time corrected using 3dTshift from AFNI 20160207 (Cox and Hyde 1997, RRID:SCR_005927). Susceptibility distortion correction (SDC) was omitted. The BOLD reference was then co-registered to the T1w reference using bbregister (FreeSurfer) which implements boundary-based registration (Greve and Fischl, 2009). Co-registration was configured with six degrees of freedom. The BOLD time-series (including slice-timing correction when applied) were resampled onto their original, native space by applying

the transforms to correct for head-motion. These resampled BOLD time-series will be referred to as preprocessed BOLD in original space, or just preprocessed BOLD. The BOLD time-series were resampled into standard space, generating a preprocessed BOLD run in MNI152NLin2009cAsym space. First, a reference volume and its skull-stripped version were generated using a custom methodology of fMRIPrep. Several confounding time-series were calculated based on the preprocessed BOLD: framewise displacement (FD), DVARS and three region-wise global signals. FD was computed using two formulations following Power (absolute sum of relative motions, (Power *et al.*, 2014)) and Jenkinson (relative root mean square displacement between affines, (Jenkinson *et al.*, 2002)). FD and DVARS are calculated for each functional run, both using their implementations in Nipype (following the definitions by (Power *et al.*, 2014)). The three global signals are extracted within the CSF, the WM, and the whole-brain masks. Additionally, a set of physiological regressors were extracted to allow for component-based noise correction (CompCor, (Behzadi *et al.*, 2007)). Principal components are estimated after high-pass filtering the preprocessed BOLD time-series (using a discrete cosine filter with 128s cut-off) for the two CompCor variants: temporal (tCompCor) and anatomical (aCompCor). tCompCor components are then calculated from the top 5% variable voxels within a mask covering the subcortical regions. This subcortical mask is obtained by heavily eroding the brain mask, which ensures it does not include cortical GM regions. For aCompCor, components are calculated within the intersection of the aforementioned mask and the union of CSF and WM masks calculated in T1w space, after their projection to the native space of each functional run (using the inverse BOLD-to-T1w transformation). Components are also calculated separately within the WM and CSF masks. For each CompCor decomposition, the k components with the largest singular values are retained, such that the retained components time series are sufficient to explain 50 percent of variance across the nuisance mask (CSF, WM, combined, or temporal). The remaining components are dropped from consideration. The head-motion estimates calculated in the correction step were also placed within the corresponding confounds file. The confound time series derived from head motion estimates and global signals were expanded with the inclusion of temporal derivatives and quadratic terms for each (Satterthwaite *et al.*, 2013). Frames that exceeded a threshold of 0.5 mm FD or 1.5 standardised DVARS were annotated as motion outliers. All resamplings can be performed with a single interpolation step by composing all the pertinent transformations (i.e. head-motion transform matrices, susceptibility distortion correction when available, and co-registrations to anatomical and output spaces). Gridded (volumetric) resamplings were performed using antsApplyTransforms (ANTs), configured with Lanczos interpolation to

minimize the smoothing effects of other kernels (Lanczos, 1964). Non-gridded (surface) resamplings were performed using mri_vol2surf (FreeSurfer).

Many internal operations of fMRIPrep use Nilearn 0.6.2 (Abraham *et al.*, 2014), RRID:SCR_001362), mostly within the functional processing workflow. For more details of the pipeline, see the section corresponding to workflows in fMRIPrep's documentation (https://fmriprep.org/en/latest/workflows.html).

### 6.6.2 Behavioural data pre-processing and consensus clustering feature selection

Behavioural variables with data missing in >25% of the sample (i.e., total RFS scores) or with outlier data point values in >5% of the sample (i.e., PAL task total adjusted errors score and HSCT total scores) were excluded from the analyses. Participants were excluded from the analyses if they had more than 25% of behavioural data missing or had outlier data points on any of the included behavioural measures. Any remaining missing data were imputed using the K-nearest neighbour method. The final sample used in the consensus clustering pipeline excluded participants not included in the functional connectivity analysis and those with further outlier data points on any of the included behavioural measures (Figure SM 6.1). Outliers are defined as data points with values exceeding the median by three times the interquartile range or more.

### 6.6.3 Supplementary figures

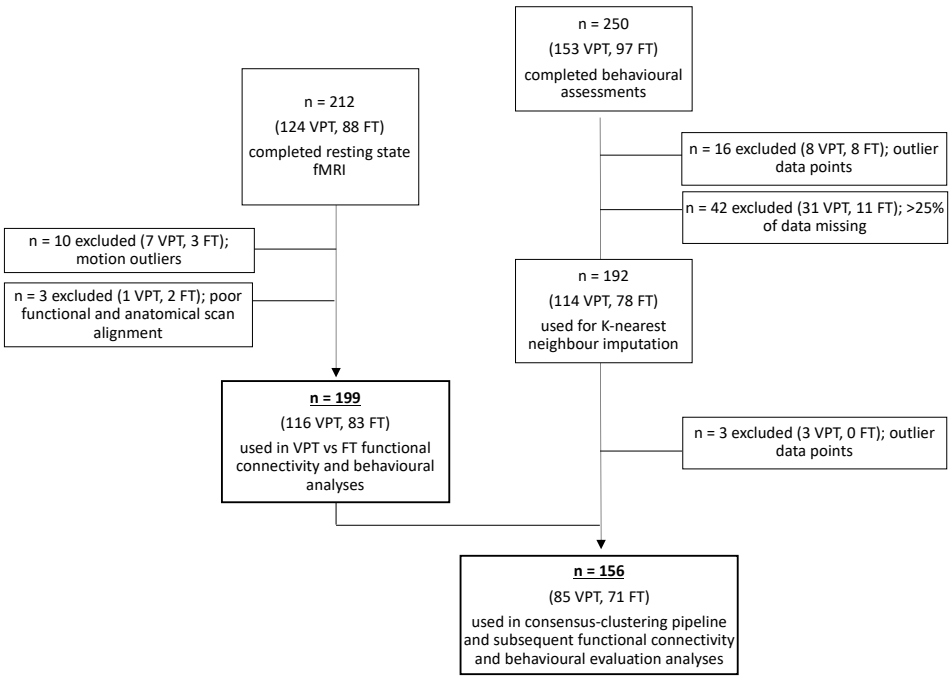



*Figure SM 6.1. Participants' selection flow diagram.*

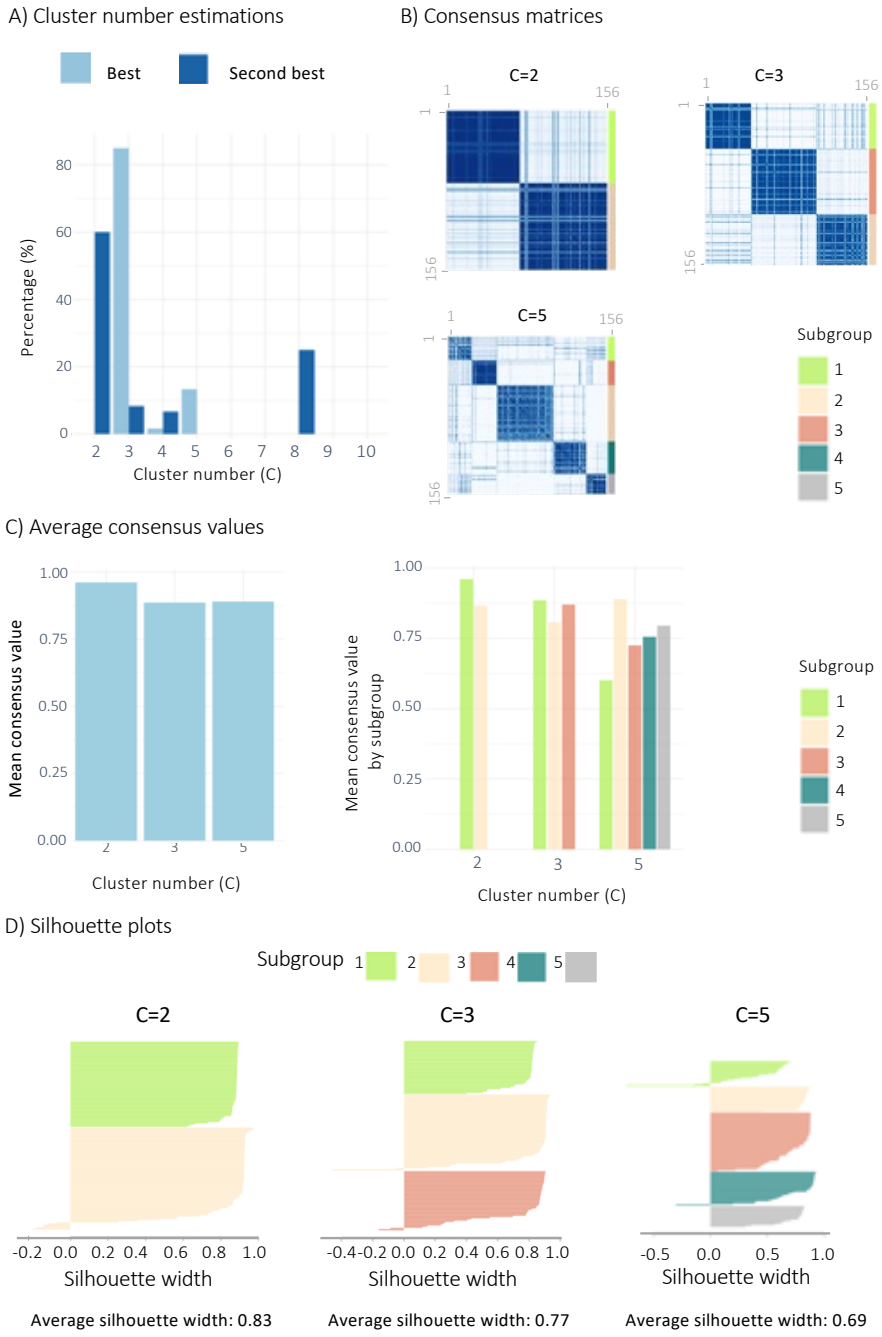



***Figure SM 6.2. Estimating the optimal number of clusters.***

*A) Number of times (in percent) each number of clusters was selected as the best (pale blue) or second best (blue) using Eigengap and Rotation Cost for the 30 combinations of K-alpha parameters, with C=2 and C=3 being the most frequently estimated best and second best number of clusters, followed by C=5. B) Consensus matrices from C=2, C=3 and C=5 showing consensus values (i.e., proportion of times each pair of subjects co-clustered into the same cluster over the 1,000 iterations; darker blue indicates higher proportions of co-clustering. C) Mean consensus values for each number of clusters (C=2, C=3 and C=5) and for each subgroup within the different number of clusters runs (left and right, respectively) with the highest values belonging to C=2. D) Silhouette width values for each subgroup within the different number of clusters runs: C=2, C=3 and C=5, whereby C=2 was also displaying the highest values.*

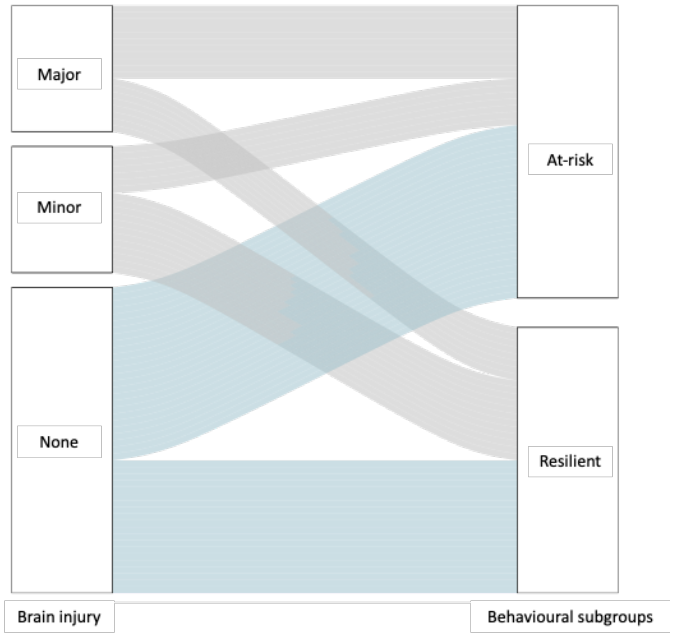



*Figure SM 6.3. Alluvial plots showing VPT individuals with no brain injury (in blue) and minor or major brain injury (in grey) clustering into the At-risk and Resilient data-driven behavioural subgroups.*

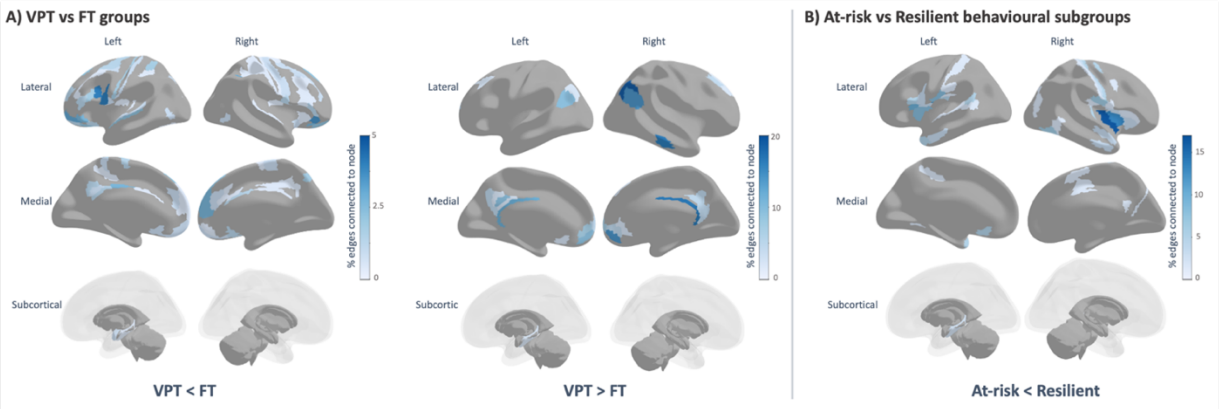



***Figure SM 6.4. Percentage of edges connected to each region within the significant NBS components at 0.001 p- NBS-Threshold.***

*Darker colours (blue) denote higher percentages and lighter colours (light blue and white) denote lower percentages, with areas marked in grey indicating regions that are not forming part of the NBS component.*

### 6.6.4 Supplementary tables

***Table SM 6.1. Cognitive assessment descriptions.***

| Assessment | Assessment description | Construct measured by assessment |
|---|---|---|
| **Hayling Sentence Completion Test (HSCT)** | Participants are asked to complete sentences correctly and incorrectly by providing semantically related or unrelated words. | This task **measures both response initiation speed and response inhibition**. Total scaled score is measured using response latency and response errors. |
| **Controlled Oral Word Association Test (COWAT-FAS)** | Participants are given 60 seconds to list as many words (excluding proper names, numbers, or words with different tenses/endings) as possible beginning with a given letter (i.e., F, A, and S). | This task **measures phonemic verbal fluency**. The total score of phonemic fluency is calculated as the sum of total words produced for each letter. |
| **Stockings of Cambridge (SOC) task** | Participants are required to rearrange three coloured stimuli to match the pattern of the stimuli displayed on the screen in the minimum number of moves possible. | This task **measures spatial planning abilities**. A score of 'Problems Solved in Minimum Moves' is calculated. |
| **Intra-Extra Dimensional Set Shift (IED) task** | Participants are presented with stimuli and are required to learn a rule and select correct responses. The task involves categorising visual stimuli into sets (i.e., the visual discrimination of shapes vs lines) and being able to flexibly respond to changes in stimuli (i.e., shifting attention). The rule changes are intra-dimensional at first (i.e., the shapes or lines are still the relevant stimuli) and then become extra-dimensional (i.e., the shapes or lines are no longer the relevant stimuli). | This task **measures attentional set shifting**. The number of errors made are calculated and adjusted for task stages completed by adding errors for stages not completed. Outcome measure is a 'Total Errors Adjusted' score. |
| **Paired Associates Learning (PAL) task** | Multiple patterns appear on a screen in a random order and then disappear. These patterns reappear, and the participant is asked to recall where on the screen that pattern was originally displayed. | **Measures episodic visuo-spatial associative memory**. The total number of adjusted errors is used a summary score. |
| **Motor Screening Task (MOT)** | A cross appears on the screen and participants are asked to press on it as quickly and as accurately as possible. | This task **measures visuo-motor coordination and comprehension abilities**. The mean reaction time is used as a summary score. |
| **The Trail Making Task – B (TMT-B)** | The task requires participants to draw lines between circles randomly distributed on a paper (whereby each circle contains a letter or a number). They participant is asked to connect the circles sequentially by alternating between numbers and letters. | This task **measures visual scanning, attentional set shifting and cognitive flexibility**. The time required to complete the task is measured and used as a summary score. |
| **Continuous Performance Test (CPT)** | Participants are presented with a series of continuously changing visual stimuli and the participant is required to respond by pressing a button when a stimulus is presented on the screen and to refrain from responding when a "non-target" stimulus (e.g., "X") is presented. | This task **measures sustained attention and response inhibition**. The total reaction time for correct responses is used a summary score. |
| **Wechsler Abbreviated Scale of Intelligence (WASI)** | Participants complete a series of tasks measuring visuo-spatial and problem-solving, abstract verbal and non-verbal reasoning, verbal expression, semantic knowledge, and verbal comprehension, | The scaled total IQ score is used as a summary **measure of general intelligence.** |

*Table SM 6.2. Clinical, socio-demographic, and behavioural profiles of included VPT sample (n=116), relative to VPT excluded from whole sample (n=37).*

| Variable | VPT excluded (n=37) | VPT included (n=116) | p-value |
|---|---|---|---|
| **Age at assessment, years** | 31.62 (3.39) | 31.37 (3.71) | 0.115 |
| **Gestational age at birth, weeks** | 28.00 (4.00) | 30.00 (3.00) | 0.024 |
| **Birth weight, grams** | 1183.00 (253.00) | 1345.00 (512.00) | 0.453 |
| [a] **Perinatal brain injury (None: Minor: Major), n** | 8:7:13 | 49:22:22 | 0.038 |
| [b] **Participants' current socio-economic status (I – II: III: I-V: Student: Unemployed), n** | 16: 7: 5: 1: 8 | 51: 41: 6: 1: 16 | 0.094 |
| [b] **Parental socio-economic status at birth (I – II: III: I-V), n** | 15:9:3 | 43:36:8 | 0.761 |
| **Sex M:F, n** | 24:13 | 66:50 | 0.506 |
| [c] **COWAT, total words** | 10.00 (7.25) | 13.00 (5.75) | 0.042 |
| [c] **CANTAB – SOC, problems solved** | 8.60 (3.00) | 9.00 (2.75) | 0.228 |
| [c] **CANTAB – IED, total errors adjusted** | 20.60 (41.75) | 15.00 (25.50) | 0.342 |
| [c] **TMT-B, time to finish task** | 100.50 (51.80) | 73.50 (40.50) | 0.007 |
| [c] **CPT, total reaction time for correct responses** | 430.90 (84.75) | 417.50 (59.15) | 0.226 |
| [c] **WASI – full scale IQ** | 100.70 (24.65) | 106 .00(13.75) | 0.097 |
| [c] **CANTAB – MOT, reaction time** | 717.50 (242.65) | 691.00 (200.80) | 0.990 |
| [d] **PDI, total score** | 24.50 (43.50) | 21.50 (50.25) | 0.680 |
| [e] **AQ10, total score** | 3.31 (3.00) | 2.00 (2.44) | 0.031 |
| [f] **CAARMS, general psychopathology score** | 2.50 (9.25) | 2.00 (5.50) | 0.133 |
| [g] **GHQ, total score** | 12.00 (6.00) | 10.00 (6.00) | 0.594 |
| [c] **ERT, total correct** | 53.50 (10.25) | 56.60 (11.15) | 0.092 |
| [c] **SAS, total score** | 1.49 (0.56) | 1.58 (0.45) | 0.867 |

***Note.*** *Median (interquartile range) reported unless number of participants (n) is reported. [a] Perinatal brain injury rated from ultrasound scans: no haemorrhage (none), grade I – II periventricular haemorrhage without ventricular dilation (minor injury) and grade III – IV periventricular haemorrhage with ventricular dilation (major injury). [b] Socio-economic status occupation classifications. I: Higher managerial, administrative and professional occupations; II: Intermediate occupations, small employers and own account workers; III: Routine and manual occupations – lower supervisory and technical and semi-routine and routine occupations. Missing data for VPT excluded and included respectively: [c](n=9, n=22), [d](n=1, n=22), [e](n=10, n=19), [f](n=1, n=17), [g](n=0, n=9).*

*Table SM 6.3. VPT only – At-risk and Resilient behavioural subgroup clinical and socio-demographic profiles.*

| Variable | VPT Subgroup 1 – Resilient | VPT Subgroup 2 – At-risk | p-value | FDR p-value |
|---|---|---|---|---|
| **Age at assessment, years** | 31.03 (3.34) | 30.50 (3.60) | 0.635 | 0.635 |
| **Gestational age at birth, weeks** | 30.00 (4.00) | 30.00 (3.25) | 0.187 | 0.262 |
| **Birth weight, grams** | 1382.50 (479.50) | 1201.50 (498.75) | 0.035 | 0.081 |
| **[a] Perinatal brain injury, n (%)** | | | 0.303 | 0.416 |
| *None* | 20 (48.78%) | 26 (59.09%) | | |
| *Minor* | 12 (29.27%) | 7 (15.91%) | | |
| *Major* | 8 (19.51%) | 11 (25%) | | |
| **[b] Participants' current socio-economic status, n (%)** | | | | |
| *I – II* | 28 (68.29%) | 16 (36.36%%) | 0.007 | 0.047 |
| *III* | 12 (29.27%) | 17 (38.64%) | | |
| *IV – V* | 0 (0.00%) | 2 (4.55%) | | |
| *Student* | 0 (0.00%) | 1 (2.27%) | | |
| *Unemployed* | 1 (2.44%%) | 8 (18.18%) | | |
| **[b] Parental socio-economic status at birth, n (%)** | | | 0.177 | 0.262 |
| *I – II* | 23 (56.10%) | 16 (36.36%) | | |
| *III* | 12 (29.27%) | 20 (45.46%) | | |
| *IV – V* | 4 (9.76%) | 4 (9.09%) | | |
| **Sex, n (%)** | | | 0.029 | 0.081 |
| *Male* | 30 (73.17%) | 21 (47.73%) | | |
| *Female* | 11 (26.83%) | 23 (52.27%) | | |
| **Total, n** | 41 | 44 | | |

*Note. Median (interquartile range) reported unless stated otherwise where number of participants (n) is reported alongside percentage (%). [a] Brain ultrasound scans were used to rate perinatal brain injury into three categories: no haemorrhage (no injury), grade I – II periventricular haemorrhage without ventricular dilation (minor injury) and grade III – IV periventricular haemorrhage with ventricular dilation (major injury). [b] Socio-economic status was categorised according to the Office of National Statistics, 1980 occupation classifications. I: Higher managerial, administrative and professional occupations; II: Intermediate occupations, small employers and own account workers; III: Routine and manual occupations – lower supervisory and technical and semi-routine and routine occupations.*

*Table SM 6.4. FT only – At-risk and Resilient behavioural subgroup socio-demographic profiles.*

| Variable | FT Subgroup 1 – Resilient | FT Subgroup 2 – At-risk | p-value | FDR p-value |
|---|---|---|---|---|
| **Age at assessment, years** | 28.82 (3.36) | 29.48 (5.31) | 0.61 | 0.811 |
| **[a] Participants' current socio-economic status, n (%)** | | | 0.035 | 0.138 |
| *I – II* | 18 (60.00%) | 14 (34.15%) | | |
| *III* | 9 (30.00%) | 14 (34.15%) | | |
| *IV – V* | 0 (0.00%) | 0 (0.00%) | | |
| *Student* | 1 (3.33%) | 10 (24.39%) | | |
| *Unemployed* | 1 (3.33%) | 3 (7.31%) | | |
| **[a] Parental socio-economic status at birth, n (%)** | | | 0.217 | 0.435 |
| *I – II* | 21 (70%) | 17 (41.46%) | | |
| *III* | 4 (13.33%) | 10 (24.39%) | | |
| *IV – V* | 1 (3.33%) | 2 (4.88%) | | |
| **Sex, n (%)** | | | 0.831 | 0.831 |
| *Male* | 13 (43.33%) | 20 (48.78%) | | |
| *Female* | 17 (56.67%) | 21 (51.22%) | | |
| **Total, n** | 30 | 41 | | |

***Note.*** *Median (interquartile range) reported unless stated otherwise where number of participants (n) is reported alongside percentage (%). [a] Socio-economic status was categorised according to the Office of National Statistics, 1980 occupation classifications. I: Higher managerial, administrative and professional occupations; II: Intermediate occupations, small employers and own account workers; III: Routine and manual occupations – lower supervisory and technical and semi-routine and routine occupations.*

*Table SM 6.5. Experimenting with p-value thresholds for NBS using two-tailed statistical testing.*

| Two-tailed effect of interest | p-value threshold | Number of identified components | Number of significant components | Significant component size | Significant component strength | FWE p-value |
|---|---|---|---|---|---|---|
| **VPT vs FT** | 0.05 | 1 | 1 | 5244 | 2347.93 | 0.004 |
| | 0.01 | 1 | 1 | 1332 | 509.11 | 0.001 |
| | 0.001 | 22 | 1 | 153 | 50.08 | <0.001 |
| **At-risk vs Resilient** | 0.05 | 1 | 0 | n/a | n/a | n/a |
| | 0.01 | 3 | 1 | 693 | 232.11 | 0.013 |
| | 0.001 | 29 | 1 | 27 | 12.27 | 0.013 |

*Table SM 6.6. VPT < FT – nodes with the highest number of connections within the significant NBS component.*

| Brain region | HCP-MMP atlas region | Number of edges | Percentage of edges |
|---|---|---|---|
| **Superior temporal gyrus (auditory association cortex)** | right Area STGa | 36 | 2.453988 |
| **Inferior parietal cortex** | left Area PGi | 33 | 2.249489 |
| **Inferior frontal cortex** | right Area 47l (47 lateral) | 32 | 2.181322 |
| **Superior parietal cortex** | left Medial IntraParietal Area | 29 | 1.976823 |
| **Orbitofrontal cortex** | right Area 47m | 27 | 1.840491 |
| **Inferior frontal cortex** | left Area IFJa | 26 | 1.772324 |
| **Orbitofrontal cortex** | left Area anterior 10p | 26 | 1.772324 |
| **Anterior cingulate and medial prefrontal cortex** | left Area p32 | 25 | 1.704158 |
| **Inferior premotor** | left Rostral Area 6 | 25 | 1.704158 |
| **Superior temporal gyrus (auditory association cortex)** | left Area STGa | 25 | 1.704158 |
| **Lateral occipital/posterior temporal visual area** | left Area PH | 25 | 1.704158 |
| **Superior parietal cortex** | left Area Lateral IntraParietal ventral | 24 | 1.635992 |
| **Orbitofrontal cortex** | left Area 47l (47 lateral) | 24 | 1.635992 |
| **Orbitofrontal cortex** | left Area 11l | 24 | 1.635992 |
| **Orbitofrontal cortex** | left Orbital Frontal Complex | 24 | 1.635992 |
| **Lateral temporal cortex** | left Area TE1 anterior | 23 | 1.567825 |
| **Dorsolateral prefrontal cortex** | left Superior Frontal Language Area | 22 | 1.499659 |
| **Inferior frontal cortex** | left Area IFSp | 22 | 1.499659 |
| **medial temporal lobe** | left ParaHippocampal Area 2 | 22 | 1.499659 |
| **Medial temporal lobe** | right Entorhinal Cortex | 22 | 1.499659 |
| **Lateral temporal cortex** | right Area TG Ventral | 21 | 1.431493 |
| **Posterior cingulate cortex** | left Area 23d | 20 | 1.363327 |
| **Dorsolateral prefrontal cortex** | left Area 8B Lateral | 19 | 1.29516 |
| **Inferior parietal cortex** | left Area PGs | 19 | 1.29516 |
| **Dorsolateral prefrontal cortex** | right Area 8B Lateral | 19 | 1.29516 |
| **Orbitofrontal cortex** | right posterior OFC Complex | 19 | 1.29516 |

| | | | |
|---|---|---|---|
| **Dorsal stream visual cortex** | left Ventral Area 6 | 18 | 1.226994 |
| **Anterior cingulate and medial prefrontal cortex** | right Area 9 Middle | 18 | 1.226994 |
| **Inferior parietal cortex** | right Area IntraParietal 2 | 18 | 1.226994 |
| **Orbitofrontal cortex** | right Orbital Frontal Complex | 17 | 1.158828 |
| **Lateral occipital/posterior temporal visual area** | right Area PH | 17 | 1.158828 |
| **Anterior cingulate and medial prefrontal cortex** | right Area 25 | 17 | 1.158828 |
| **Anterior cingulate and medial prefrontal cortex** | left Area 10r | 16 | 1.090661 |
| **Orbitofrontal cortex** | left Area anterior 47r | 16 | 1.090661 |
| **Orbitofrontal cortex** | left Area 13l | 16 | 1.090661 |
| **Orbitofrontal cortex** | right Area 11l | 16 | 1.090661 |
| **Superior temporal gyrus (auditory association cortex)** | right Area TA2 | 16 | 1.090661 |

*Table SM 6.7. VPT > FT – nodes with the highest number of connections within the significant NBS component.*

| Brain region | HCP-MMP atlas region | Number of edges | Percentage of edges |
|---|---|---|---|
| **Posterior opercular cortex** | left Frontal Opercular Area 4 | 30 | 3.118503 |
| **Posterior opercular cortex** | right Frontal Opercular Area 4 | 30 | 3.118503 |
| **Posterior cingulate cortex** | right Complex | 26 | 2.702703 |
| **Inferior parietal cortex** | left Area PF Complex | 24 | 2.494802 |
| **Inferior parietal cortex** | right Area IntraParietal 2 | 24 | 2.494802 |
| **Orbitofrontal cortex** | right Area 10d | 23 | 2.390852 |
| **Inferior parietal cortex** | right Area PGs | 22 | 2.286902 |
| **Posterior opercular cortex** | right Area Frontal Opercular 5 | 22 | 2.286902 |
| **Inferior parietal cortex** | right Area PGi | 20 | 2.079002 |
| **Anterior Cingulate and Medial Prefrontal Cortex** | left Area 10r | 18 | 1.871102 |
| **Posterior cingulate cortex** | right Parieto-Occipital Sulcus Area 1 | 18 | 1.871102 |
| **Dorsolateral prefrontal cortex** | right Area posterior 9-46v | 18 | 1.871102 |
| **Superior temporal gyrus (auditory association cortex)** | right Area STGa | 18 | 1.871102 |
| **Lateral temporal cortex** | right Area TE1 Middle | 18 | 1.871102 |
| **Superior temporal gyrus (auditory association cortex)** | left Area STGa | 16 | 1.663202 |
| **Inferior parietal cortex** | right Area PFm Complex | 16 | 1.663202 |
| **Posterior opercular cortex** | left Area Frontal Opercular 5 | 15 | 1.559252 |
| **Posterior cingulate cortex** | right Area 31pd | 15 | 1.559252 |
| **Anterior Cingulate and Medial Prefrontal Cortex** | right Anterior 24 prime | 14 | 1.455301 |
| **Anterior Cingulate and Medial Prefrontal Cortex** | right Area a24 | 14 | 1.455301 |
| **Posterior cingulate cortex** | left Area dorsal 23 a+b | 13 | 1.351351 |
| **Posterior cingulate cortex** | left Area 31p ventral | 13 | 1.351351 |
| **Superior parietal cortex (medial)** | left Medial Area 7A | 13 | 1.351351 |
| **Anterior Cingulate and Medial Prefrontal Cortex** | left Area Posterior 24 prime | 13 | 1.351351 |
| **Anterior Cingulate and Medial Prefrontal Cortex** | right Area 10r | 13 | 1.351351 |

| | | | |
|---|---|---|---|
| **Dorsolatetal prefrontal cortex** | right Area 8C | 13 | 1.351351 |
| **Posterior opercular cortex** | right Area 43 | 13 | 1.351351 |
| **Posterior cingulate cortex** | left Retrosplenial Complex | 12 | 1.247401 |
| **Inferior parietal cortex** | left Area PGi | 12 | 1.247401 |
| **Superior parietal cortex (medial)** | right Medial Area 7A | 12 | 1.247401 |
| **Anterior Cingulate and Medial Prefrontal Cortex** | right Area 9 Middle | 12 | 1.247401 |
| **Temporo-parietal-occipital junction** | right TemporoParietoOccipital Junction 2 | 12 | 1.247401 |
| **Orbitofrontal cortex** | right Area posterior 10p | 12 | 1.247401 |

*Table SM 6.8. At-risk < Resilient – nodes with the highest number of connections within the significant NBS component.*

| Brain region | HCP-MMP atlas region | Number of edges | Percentage of edges |
|---|---|---|---|
| **Insular cortex** | right Posterior Insular Area 2 | 32 | 3.846154 |
| **Frontal opercular cortex** | left Area OP4/PV | 25 | 3.004808 |
| **Frontal opercular cortex** | left Frontal Opercular Area 4 | 24 | 2.884615 |
| **Insular cortex** | right Middle Insular Area | 24 | 2.884615 |
| **Frontal opercular cortex** | left Frontal Opercular Area 2 | 23 | 2.764423 |
| **Posterior opercular cortex** | right Area OP4/PV | 23 | 2.764423 |
| **Inferior frontal cortex** | left Area 44 | 22 | 2.644231 |
| **Lateral occipital/posterior temporal visual area** | right Area PH | 21 | 2.524038 |
| **Insular cortex** | left Posterior Insular Area 2 | 20 | 2.403846 |
| **Lateral temporal cortex** | left Area TE2 anterior | 19 | 2.283654 |
| **Superior premotor cortex** | right Dorsal area 6 | 18 | 2.163462 |
| **Lateral temporal cortex** | left Area TG dorsal | 15 | 1.802885 |
| **Posterior opercular cortex** | right Area 43 | 15 | 1.802885 |
| **Temporo-parieto-occipital junction** | left Superior Temporal Visual Area | 14 | 1.682692 |
| **Supplementary motor area** | left Supplementary and Cingulate Eye Field | 14 | 1.682692 |
| **Insular cortex** | left Insular Granular Complex | 14 | 1.682692 |
| **Lateral temporal cortex** | right Area TF | 14 | 1.682692 |
| **Posterior opercular cortex** | left Area OP2-3/VS | 13 | 1.5625 |
| **Anterior Cingulate and Medial Prefrontal Cortex** | left Area 25 | 13 | 1.5625 |
| **Primary somatosensory cortex** | right Primary Sensory Cortex | 13 | 1.5625 |
| **Posterior opercular cortex** | right Area OP1/SII | 12 | 1.442308 |
| **Insular cortex** | right Area 52 | 12 | 1.442308 |
| **Primary somatosensory cortex** | left Primary Sensory Cortex | 11 | 1.322115 |
| **Insular cortex** | left Middle Insular Area | 11 | 1.322115 |
| **Frontal opercular cortex** | left Frontal Opercular Area 3 | 11 | 1.322115 |
| **Superior temporal sulcus (auditory association cortex)** | left Area STSd anterior | 11 | 1.322115 |

| | | | |
|---|---|---|---|
| **Lateral temporal cortex** | left Area TF | 11 | 1.322115 |
| **Anterior Cingulate and Medial Prefrontal Cortex** | right Area 8BM | 11 | 1.322115 |
| **Posterior opercular cortex** | right Area OP2-3/VS | 11 | 1.322115 |
| **Frontal opercular cortex** | right Frontal Opercular Area 3 | 11 | 1.322115 |
| **Posterior opercular cortex** | left Area PFcm | 10 | 1.201923 |
| **Posterior opercular cortex** | left Frontal Opercular Area 1 | 10 | 1.201923 |
| **Frontal opercular cortex** | right Frontal Opercular Area 2 | 10 | 1.201923 |
| **Superior temporal sulcus (auditory association cortex)** | right Area STSd anterior | 10 | 1.201923 |
| **Lateral temporal cortex** | right Area TE1 anterior | 10 | 1.201923 |

*Table SM 6.9. Sensitivity analyses – NBS component results at p-NBS-threshold = 0.001.*

| | Edges, n (% of all possible connections) | Nodes, n (% of all regions) | Component strength, T-stat | FWE p-value |
|---|---|---|---|---|
| ***VPT < FT*** | 221 (0.32%) | 179 (47.86%) | 77.89 | 0.001 |
| ***VPT > FT*** | 42 (0.060%) | 29 (7.75%) | 16.36 | 0.001 |
| ***At-risk < Resilient*** | 60 (0.086%) | 52 (13.90%) | 24.84 | 0.006 |

# CHAPTER 7 - Integrative discussion and conclusions

The objective of this PhD thesis was to characterise specific neural markers associated with distinct behavioural outcomes in VPT samples across the lifespan. To address this main aim, four studies examined in this thesis i) stratified behavioural heterogeneity in VPT (and FT) samples by using distinct psychometric scoring criteria and advanced data-driven clustering approaches, ii) implemented advanced whole-brain neuroimaging analyses to map structural and functional brain changes associated with distinct behavioural outcomes, and iii) investigated the role of clinical and environmental factors in distinct behavioural subgroups. In the first section of this chapter (section 7.1), the specific aims addressed in each study are revisited and summaries of the findings are also provided. In the subsequent sections (sections 7.2-7.5), an integrative discussion presents four main themes highlighting the key interpretations of the study findings presented in this thesis. Finally, this chapter concludes by reporting the implications, strengths, limitations, and future directions of this work (section 7.6).

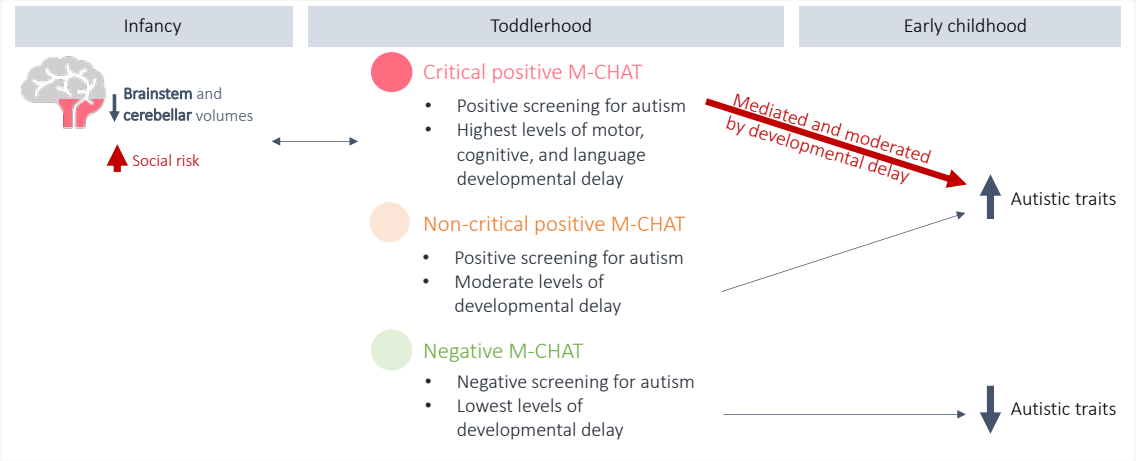
## 7.1 Review of main aims and study findings

### 7.1.1 **Experimental Study #1** – Using distinct M-CHAT psychometric scoring criteria to delineate longitudinal brain-behavioural heterogeneity in VPT toddlers

***Aim 1:*** *"To investigate whether brain-behavioural heterogeneity in VPT toddlers can be characterised by stratifying individuals according to distinct psychometric screening criteria for autism using the M-CHAT"*

***Aim 2:*** *"To explore whether developmental delay mediates or interacts with childhood autism traits in the distinct psychometric screening subgroups"*

**Graphical summary of findings:**

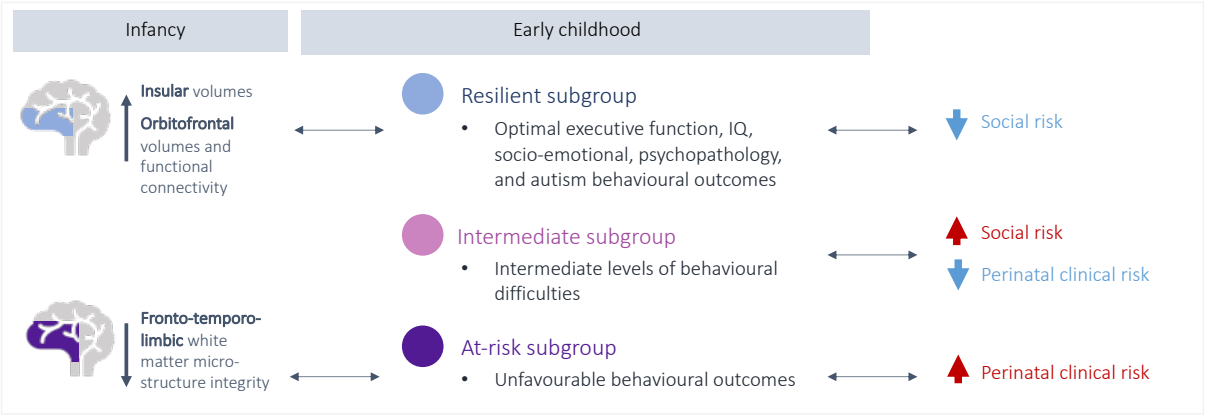



### 7.1.2 **Experimental Study #2** – Using data-driven integrative consensus clustering to parse longitudinal brain-behavioural heterogeneity in VPT children

***Aim 1:*** *"To parse heterogeneity in neonatal clinical and social risk and childhood behavioural outcomes using data-driven integrative consensus clustering techniques"*

***Aim 2:*** *"To explore differences in neonatal brain volumes and structural and functional connectivity between the distinct data-driven subgroups using advanced neuroimaging analysis approaches"*

**Graphical summary of findings:**




### 7.1.3 **Experimental Study #3** – Elucidating brain-behavioural heterogeneity in VPT and FT children using data-driven consensus clustering

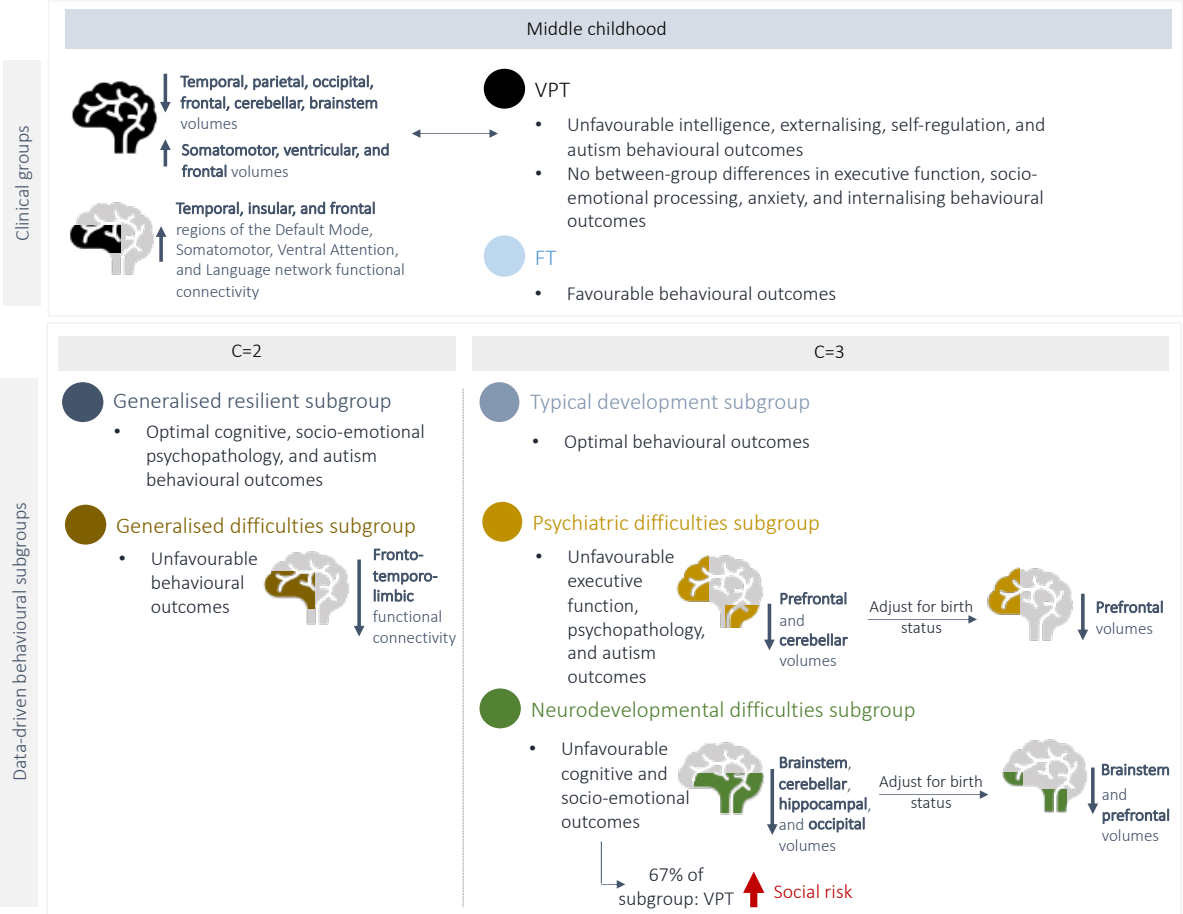
***Aim 1:*** ***"****To use advanced neuroimaging analyses to compare resting state functional connectivity and structural volumes differentiating between groups of VPT and FT children stratified both in terms of:*

- *a. Clinical birth status – i.e., VPT vs FT birth*

- *b. Data-driven behavioural subgroups identified using consensus clustering regardless of gestational age at birth"*

***Aim 2:*** *"To explore differences in clinical and social factors between the distinct data-driven behavioural subgroups"*

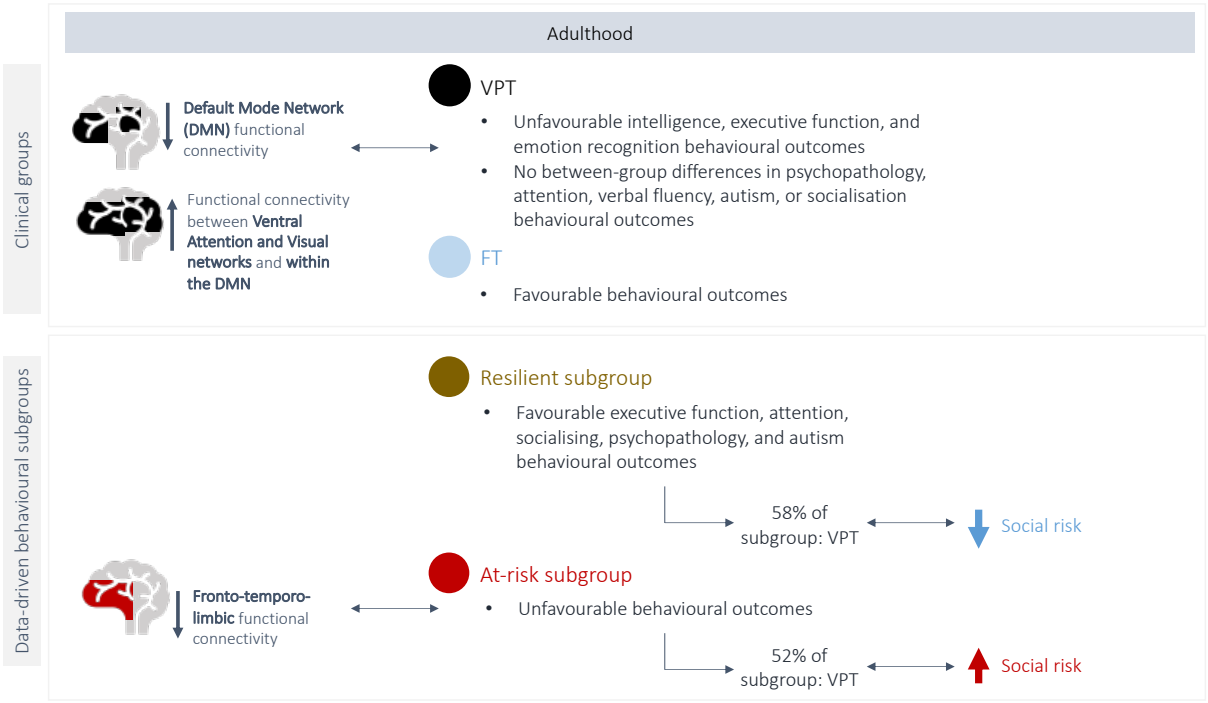
**Graphical summary of findings:**




### 7.1.4 **Experimental Study #4** – Elucidating brain-behavioural heterogeneity in VPT and FT adults using data-driven consensus clustering

***Aims:*** *"To characterise resting state functional connectivity using a sample of VPT and FT adults to explore the same set of aims investigated in Study #3"*

**Graphical summary of findings:**




## 7.2 Theme 1: Neurodevelopmental profiles unique to VPT individuals relative to FT controls

Group-wise differences in both brain and behavioural profiles differentiated between VPT and FT groups in childhood (Study #3) and adulthood (Study #4).

### 7.2.1 Behavioural differences between VPT and FT individuals in childhood and adulthood

Studies #3 and #4 provide evidence of preterm-specific behavioural sequelae in both childhood and adulthood, demonstrating the presence of long-lasting behavioural difficulties across the preterm lifespan. Behavioural outcomes that differed in VPT relative to FT groups included specific aspects of the previously described 'preterm behavioural phenotype' – a behavioural profile characterised by social, emotional, and attention processing difficulties (Johnson and Marlow, 2011). Namely, VPT individuals exhibited behavioural profiles encompassing difficulties in intelligence, socio-emotional processing, and externalising (i.e., hyperactivity and inattention) behaviours which have also been previously reported in VPT samples (Spittle *et al.*, 2009; Bora *et al.*, 2011; Eves *et al.*, 2021). On the other hand, however, VPT children and adults examined here did not exhibit additional commonly occurring difficulties often associated with preterm birth. Namely, difficulties between VPT and FT individuals relating to executive function abilities, psychopathology, and emotional processing and regulation

(Aarnoudse-Moens *et al.*, 2009; Spittle *et al.*, 2009; Kroll *et al.*, 2017; P. J. Anderson *et al.*, 2021) were not observed in the samples studied here.

There are a few possible justifications potentially explaining the lack of consistency of findings across studies. For instance, the samples studied here seem to be comprised of a relatively low-risk subpopulation of VPT born individuals. Sensitivity analyses performed in Study #3 and Study #4 show that VPT individuals included in the final analyses present with relatively more optimal socio-emotional and cognitive behavioural outcomes when compared to VPT individuals excluded from the final analyses. Moreover, the lack of consensus, even amongst existing studies (Spittle *et al.*, 2009; Bora *et al.*, 2011), in terms of specific behavioural domains that may be affected following preterm birth, could reflect the underlying behavioural heterogeneity exhibited by both preterm and FT samples. It is possible that the phenotypic heterogeneity may mask true underlying effects in group-level comparisons (e.g., VPT vs FT individuals). T**hese results highlight the importance of characterising behavioural heterogeneity across samples in order to accurately evaluate the behavioural profiles associated with VPT birth.**

### 7.2.2 Structural and functional brain differences between VPT and FT individuals in childhood and adulthood

In terms of neurobiological changes, VPT children (Study #3) and adults (Study #4) displayed widespread alterations relative to their FT born peers which were largely in line with previous literature (Kesler *et al.*, 2006; Lax *et al.*, 2013; Nosarti *et al.*, 2014; Wilke *et al.*, 2014; Bäuml *et al.*, 2015; Degnan *et al.*, 2015b; Lean *et al.*, 2017; Lemola *et al.*, 2017; Wehrle *et al.*, 2018; Zhou *et al.*, 2018; Cho *et al.*, 2022; Mossad *et al.*, 2022; Kvanta *et al.*, 2023). In middle childhood, structural changes in the VPT group localised to temporal, limbic, cerebellar, brainstem, orbitofrontal, occipital, and parietal regions displayed volumetric reductions while changes in frontal, cingulate, somatomotor, corpus callosal, and ventricular regions exhibited volumetric expansions compared to relative structural volumes in FT children. Overlapping regions also displayed altered functional connectivity both in childhood (Study #3) and in adulthood (Study #4) in VPT relative to FT groups, encompassing intrinsic networks such as the DMN, VAN, somatomotor network, and language network.

During childhood (Study #3), patterns of increased functional connectivity (i.e., hyper-connectivity) in DMN, VAN, somatomotor, and language networks were prominent in VPT relative to FT individuals, which is in line with previous studies in VPT and late-to-moderate preterm (LMPT) children, reporting hyper-connectivity in overlapping networks and regions at 6

(Cho *et al.*, 2022; Mossad *et al.*, 2022), 9-13 (Degnan *et al.*, 2015a), and 10-16 years (Wilke *et al.*, 2014; Wehrle *et al.*, 2018) relative to FT controls. In contrast to findings in childhood, patterns of reduced functional connectivity (i.e., hypo-connectivity) in the identified networks were reported in adulthood (Study #4) (White *et al.*, 2014; Bäuml *et al.*, 2015). These patterns of hypo-connectivity have in fact also been reported at earlier developmental stages at 3-5 years (Damaraju *et al.*, 2010; Choi *et al.*, 2018), but appear to be even more prominent in adulthood (White *et al.*, 2014; Bäuml *et al.*, 2015). At first glance these results appear to be conflicting in terms of the directionality of effects; such that, VPT individuals demonstrate hyper-connectivity in childhood but hypo-connectivity in adulthood. However, it is in fact likely that these observations may reflect the dynamic developmental changes that the brain undergoes across the lifespan (Figure 1.1). Previous experimental evidence shows that preterm birth alters typical developmental trajectories, whereby FT (but not VPT) children showed increased functional connectivity from age 6 to 8 (Mossad *et al.*, 2022), while preterm (but not FT) children displayed an increase between 8 and 16 years (Rowlands *et al.*, 2016), demonstrating evidence of preterm children undergoing differential maturation trajectories when compared to FT peers. Similarly, structural changes across both groups also display differential developmental trajectories. For instance, frontal, parietal and temporal cortical thinning, which occurs by 7 years in FT children, is delayed in preterm children and occurs between 7 and 12 years (Mürner-Lavanchy *et al.*, 2014). **Therefore, this thesis finds evidence supporting previous claims stating that preterm birth results in long-term alterations to structural and functional brain development when compared to FT controls.**

### 7.2.3 Brain-behavioural differences between VPT and FT individuals in childhood and adulthood

Based on findings from Study #3, Study #4, and previous studies, it is speculated here that neurobiological alterations in VPT individuals may, on one hand, reflect altered neural mechanisms which probe the onset of unfavourable behavioural outcomes, given they are localised to regions and networks implicated in socio-emotional and cognitive behavioural processing (Skranes *et al.*, 2007; Nosarti *et al.*, 2008; Rogers *et al.*, 2012, 2014, 2017; Papini *et al.*, 2016; Kanel *et al.*, 2021, 2022; Wheelock *et al.*, 2021, 2021; Mossad *et al.*, 2022). However, on the other hand, these alterations may also act as compensatory neural adaptations which support optimal behavioural processing in VPT samples. For example, specific functional connectivity alterations in networks including the VAN, DMN, and somatomotor network were seen to support specific domains of attention (Wheelock *et al.*, 2021), balance (Wheelock *et al.*, 2018), and language processing in task-based f-MRI (Schafer *et al.*, 2009; Finke *et al.*, 2015) in VPT samples, but not in FT peers. Moreover, DMN

hyper-connectivity in VPT children was found to be associated with better social functioning and working memory outcomes (Mossad *et al.*, 2022) and better inhibition control at 12 years (Wheelock *et al.*, 2021). A previous study in LMPT children reporting no differences in executive function abilities between LMPT and FT groups also found evidence of DMN hyper-connectivity in the preterm relative to the FT group (Degnan *et al.*, 2015a). **Therefore, these findings suggest that while some of these neural alterations may be associated with a risk of developing some behavioural difficulties, other neural adaptations are potentially supporting optimal outcomes in VPT samples.**

## 7.3 Theme 2: Neurodevelopmental heterogeneity across the lifespan

### 7.3.1 Behavioural heterogeneity across the VPT lifespan

Despite belonging to the same birth status group, VPT individuals examined in this thesis displayed distinct behavioural phenotypes in toddlerhood (Study #1), early childhood (Study #2), middle childhood (Study #3), and adulthood (Study #4); therefore, indicating evidence of multifinality of outcomes (as described in Figure 1.5) across the lifespan.

#### *7.3.1.1 Rates of optimal behavioural outcomes in VPT samples decline over time*

In line with findings in Study #1, previous studies implementing data-driven stratification approaches in LMPT (Johnson *et al.*, 2018) and VPT (Ross *et al.*, 2016) toddlers also found that the majority of preterm toddlers studied were more likely to belong to subgroups characterised by optimal or favourable outcomes on measures of cognitive, motor, and linguistic developmental and autistic traits (see Table 7.1; Table 7.2). On the other hand, findings from Studies #3 and #4 in older age groups in childhood and adulthood show that at least half (>50%) of the VPT children and adults demonstrate profiles with unfavourable behavioural outcomes (Table 7.1), corroborating findings from previous studies in childhood samples (Table 7.2). **These findings could be interpreted as suggesting that the rates of optimal behavioural outcomes in VPT samples decline with increasing age (Table 7.1; Table 7.2).** Endorsing this notion, studies modelling longitudinal behavioural trajectories in preterm samples find that VPT toddlers are eight times more likely to follow declining language development trajectories relative to their FT peers between the ages of 2 and 13 years (Nguyen *et al.*, 2018), with almost half (42%) of VPT individuals exhibiting poor attention outcomes at 13 years which either remained stable from 8 years of age or declined further over time (Bogičević *et al.*, 2021). This decline in performance with increasing

age may be a result of the increased social and cognitive demands required of children and adults as they grow older.

*Table 7.1. Proportion of individuals clustering into distinct behavioural subgroups – summary of results from Studies #1-4.*

| | Optimal outcome subgroups, n (% of sample) | Unfavourable outcome subgroups, n (% of sample) | |
|---|---|---|---|
| **Toddlerhood, VPT only** | | | |
| *C=3 (Study #1)* | ***Negative M-CHAT***, n= 130 (74%) | ***Non-critical M-CHAT***, n= 32 (18%) | ***Critical M-CHAT***, n= 15 (9%) |
| **Early childhood, VPT only** | | | |
| *C=2 (Study #2)* | ***Resilient***, n=97 (49%) | ***At-risk***, n= 101 (51%) | / |
| *C=3 (Study #2)* | ***Resilient***, n=74 (37%) | ***At-risk***, n=58 (29%) | ***Intermediate***, n=66 (33%) |
| **Middle childhood, VPT and FT** | | | |
| *C=2 (Study #3)* | ***General Resilience***, n=88 (51% VPT; 70% FT) | ***General Difficulties***, n=65 (50% VPT; 30% FT) | |
| *C=3 (Study #3)* | ***Typically Developing***, n=53 (25% VPT; 52% FT) | ***Psychiatric Difficulties***, n=46 (35% VPT; 21% FT) | ***Neurodevelopmental Difficulties***, n=54 (4% VPT; 27% FT) |
| **Adulthood, VPT and FT** | | | |
| *C=2 (Study #4)* | ***Resilient***, n=71 (48% VPT; 42% FT) | ***At-risk***, n =85 (52% VPT; 58% FT) | |

***Note***. *Subgroup, n (%) describes sample size and percentages of preterm and FT samples clustering into each subgroup.*

*Table 7.2. Data-driven stratification studies in preterm samples.*

| Sample (reference) | Input features | Subgroups, (% of sample) | Method | Out-of-model features |
|---|---|---|---|---|
| **Toddlerhood** | | | | |
| *LMPT + FT, 24 months (Johnson et al., 2018)* | Dichotomous measures: psychopathology, autism, socio-emotional processing, development (cognition and language), eating difficulties | **C=3:** ***Optimal outcomes*** (67% LMPT; 84% FT); ***Unfavourable outcomes*** (26% LMPT; 16% FT); ***Preterm behavioural phenotype*** (7% LMPT) | Latent Class Analysis | **Perinatal clinical risk, demographic information, social risk** |
| *VPT only, 18 months (Ross et al., 2016)* | Cognition, language | **C=4:** ***Optimal*** (17%); ***Average*** (54%); ***Language delay*** (21%); ***Global delay*** (8.5%) | Cluster analysis | **Perinatal brain injury, social risk,** demographic information, **perinatal clinical risk**, **psychopathology** |
| ***Early childhood*** | | | | |
| *Preterm only, 6 years (Poehlmann et al., 2015)* | Psychopathology, peer relations, sleep, socialising, cognition (executive function), learning | **C=3:** ***Resilient*** (31%); ***At-risk*** (57%); ***Clinical difficulties*** (12%) | Latent Profile Analysis | **Perinatal clinical risk, dysfunctional parenting**, social risk, demographic information, maternal mental health, early cognition, **early regulation** |
| *VPT + FT, 2 and 5 years (Lean et al., 2020)* | Psychopathology, autism, early (cognitive, motor, language) development | **C=4:** ***Typical development*** (27% VPT; 65% FT); ***At-risk – mild*** (45% VPT; 23% FT); ***At-risk – severe*** (13% VPT; 10% FT%); ***Inattentive/ hyperactive*** (15% VPT; 3% FT) | Latent Profile Analysis | **Maternal mental health, familial dysfunction**, **social risk (FT only)** |
| *VPT only, 4-7 years (Study #2)* | Cognition, psychopathology, autism, socio-emotional processing, perinatal clinical risk, social risk | **C=2:** ***Resilient*** (49%); ***At-risk*** (51%)<br><br>**C=3:** ***Resilient*** (37%); ***At-risk*** (29%); ***Intermediate*** (33%) | Integrative consensus clustering | **s-MRI, f-MRI, d-MRI, cognition, psychopathology, socio-emotional processing, perinatal clinical risk, social risk, environment**, demographic information |
| **Middle childhood** | | | | |
| *EPT + FT,7-8 years (Burnett et al., 2019)* | Psychopathology, peer relations | **C=4:** ***Preterm behavioural phenotype*** (20% VPT; 12% FT); ***Minimal difficulties*** (55% of VPT; 74% FT); ***Global difficulties*** (8% VPT; 3% FT); ***Elevated behavioural difficulties except in peer relations*** (16% VPT; 11% FT) | Latent profile analysis | **Cognitive and academic profiles** |
| *VPT, 8-12 years (van Houdt et al., 2020)* | Psychopathology, peer relations, cognition (executive function) | **C=2:** ***Minimal difficulties*** (76% VPT); ***High difficulties*** (24% VPT) | Hierarchical clustering | **Cognitive (IQ),** perinatal clinical risk, **social risk** |

| | | | | |
|---|---|---|---|---|
| *VPT + FT, 7 and 13 years* (Bogičević et al., 2021) | Attention | **C=3, 7yo:** ***Optimal/average*** (47% VPT; 83% FT); ***Difficulties*** (32% VPT; 17% FT); ***Elevated difficulties*** (21% VPT)<br><br>**C=3, 13yo:** ***Optimal/average*** (45% VPT; 70% FT); ***Difficulties*** (44% VPT; 28% FT); ***Elevated difficulties*** (11% VPT; 2% FT) | Latent profile analysis | **Perinatal brain injury,** perinatal clinical risk, **social risk**, environment |
| *VPT + FT, 7-12 years (Study #3)* | Cognition, psychopathology, autism, socio-emotional processing | **C=2:** ***General Resilience*** (51% VPT; 70% FT); ***General Difficulties*** (50% VPT; 30% FT)<br><br>**C=3:** ***Typically Developing*** (25% VPT; 52% FT); ***Psychiatric Difficulties*** (35% VPT; 21% FT); ***Neurodevelopmental Difficulties*** (40% VPT; 27% FT) | Consensus clustering | **s-MRI, f-MRI,** perinatal clinical risk, **social risk (VPT only),** demographic information |
| **Adulthood** | | | | |
| *VPT + FT, 23-39 years (Study #4)* | Cognition, psychopathology, socio-emotional processing | **C=2:** ***Resilient*** (48% VPT; 42% FT); ***At-risk*** (52% VPT; 58% FT) | Consensus clustering | **f-MRI,** perinatal clinical risk, **social risk (VPT only),** demographic information |

***Note.*** *Subgroup, (%) describes percentages of preterm and FT samples clustering into each subgroup. Out-of-model features highlighted in bold indicate measures displaying statistically significant between-subgroup differences.*

#### *7.3.1.2 Limitations of relying on early developmental delay to predict outcome*

Whilst some literature suggests there may be a decline in rates of optimal behavioural outcomes in preterm toddlers with increasing age (Table 7.1; Table 7.2), it is important to note that measures of developmental delay examined in toddlerhood often capture sensorimotor, cognitive, and language abilities which, despite having been associated with an increased risk of behavioural difficulties in some individuals (Luyster *et al.*, 2011; Moore, Johnson, *et al.*, 2012; Blencowe *et al.*, 2013; Van Hus *et al.*, 2014; Vollmer and Stålnacke, 2019; Johansson *et al.*, 2023), do not necessarily represent a risk factor for emerging behavioural difficulties in all individuals (Rubenstein *et al.*, 2018; Durrant *et al.*, 2020; Vanes *et al.*, 2023). For example, Study #1 shows that while developmental delay in the critical positive M-CHAT subgroup in toddlerhood was moderating and mediating elevated levels of autistic behaviours later in childhood, this was not the case in the non-critical positive M-CHAT subgroup. Here, developmental delay was only weakly mediating the onset of equally elevated levels of autistic behaviours later in childhood. Furthermore, across both positive M-CHAT subgroups in Study #1, developmental delay did not fully mediate the relationship between positive subgroup membership and later autistic traits. In addition, (Lean *et al.*, 2020) previously identified distinct behavioural subgroups exhibiting heterogeneous phenotypes of unfavourable childhood behavioural outcomes despite having similar levels of developmental delay in toddlerhood. Another study found that rates of change in developmental delay profiles were better predictors of later outcomes than outcomes at any given moment in time (Durrant *et al.*, 2020). **Together, these findings indicate that later behavioural outcomes are not fully explainable using measures of early developmental delay and that additional unmeasured risk factors (such as genetics, environmental factors, clinical factors, other behavioural measures, or neural alterations), are likely contributing to the onset of such behavioural difficulties. They also highlight the prominent presence of behavioural heterogeneity in preterm samples, which appears to be long-lasting, dynamic, and associated with complex and distinct aetiological trajectories** (Luu *et al.*, 2009, 2011; Nguyen *et al.*, 2018; Bogičević *et al.*, 2021)**.**

#### *7.3.1.3 Rates of sub-threshold behavioural difficulties are consistently elevated across the VPT lifespan*

Another conclusion drawn based on findings from this thesis and previously published studies suggests that, amongst VPT toddlers and children displaying unfavourable outcomes, more preterm individuals tend to cluster into subgroups exhibiting sub-threshold behavioural difficulties than profiles characterised by severe or supra-threshold levels of behavioural sequelae (Poehlmann

*et al.*, 2015; Ross *et al.*, 2016; Johnson *et al.*, 2018; Burnett *et al.*, 2019; Lean *et al.*, 2020; van Houdt *et al.*, 2020; Bogičević *et al.*, 2021) (Table 7.1; Table 7.2). Those subgroups tend to demonstrate profiles characteristic of either a "preterm behavioural phenotype", or of behavioural difficulties in specific cognitive, psychopathological, or socio-emotional sub-domains. Findings in Study #1-3 are in line with previous studies reporting the presence of these patterns in toddlers and children born VPT; however, findings from Study #4, provide evidence of these findings in adulthood for the first time. **The prominence of sub-threshold difficulties in VPT samples emphasises the need to implement research studies and clinical trials using reformulated conceptual and methodological frameworks which deviate away from group-wise and case-control comparisons and move towards data-driven stratification approaches.**

### 7.3.2 Behavioural heterogeneity in VPT and FT samples, regardless of clinical birth status

**Results from Study #3 confirm previous findings in preterm children which show that behavioural heterogeneity is not unique to preterm samples and in fact also characterises FT children** (Johnson *et al.*, 2018; Burnett *et al.*, 2019; Lean *et al.*, 2020; Bogičević *et al.*, 2021)**.** Whereby, ~14-48% of FT children exhibit unfavourable behavioural profiles resembling those observed in preterm samples (Johnson *et al.*, 2018; Burnett *et al.*, 2019; Lean *et al.*, 2020; Bogičević *et al.*, 2021) (Table 7.2). **Results from Study #4 extend existing evidence in children to suggest that behavioural heterogeneity across both FT and VPT individuals is also present in adulthood.** However, in adulthood, a greater proportion of FT individuals displayed an unfavourable behavioural profile (58%) in comparison to the lower rates observed in childhood (~14-48%). One possible explanation for these time-dependent changes between childhood and adulthood could involve the age of onset of several major psychiatric conditions being in adolescence and early adulthood (Otto *et al.*, 2021; Solmi *et al.*, 2022). However, the results of this study alone may not be generalisable to the wider population and suggest a need for more studies investigating mental health and cognitive trajectories over protracted time periods following VPT birth.

## 7.4 Theme 3: Neurobiological markers of behavioural heterogeneity

Specific brain changes presented here at term-equivalent age in longitudinal studies (Studies #1 and #2) in VPT samples, as well as in cross-section studies in childhood (Studies #3) and adulthood (Studies #4) in VPT and FT samples, were found to be associated with distinct behavioural outcomes, suggesting the potential utility of MRI to detect neurobiological markers predictive of heterogeneous behavioural outcomes across the lifespan.

### 7.4.1 Neurobiological markers of behavioural heterogeneity in VPT and FT samples, regardless of birth status

#### *7.4.1.1 Fronto-temporo-limbic brain alterations as potential markers of outcome*

Structural and functional alterations spanning fronto-temporo-limbic regions, such as the insula, temporal cortex, and orbitofrontal cortex, were found to be characteristic of generalised behavioural difficulties spanning socio-emotional, psychopathology, autism, and cognitive behavioural domains across the lifespan (Studies #2-4). Namely, alterations in white matter micro-structural characteristics in infancy (Studies #2), smaller regional volumes in infancy and childhood (Studies #2-3), and reduced functional connectivity in infancy, childhood, and adulthood (Studies #2-4) were distinguishing data-driven subgroups displaying unfavourable behavioural outcomes from those exhibiting optimal behavioural phenotypes. **Taken together, these findings suggest that fronto-temporo-limbic neural alterations underlie the presence of generalised behavioural difficulties.**

#### *7.4.1.2 Brainstem alterations as potential markers of outcome*

On the other hand, volume reductions localised to the brainstem, identified in Study #1 in infancy and Study #3 in childhood, differentiated subgroups demonstrating domain-specific difficulties: i.e., a critical-positive M-CHAT subgroup displaying high levels of cognitive, language, and motor developmental delays in toddlerhood and autism traits in childhood (Study #1) and a Neurodevelopmental Difficulties subgroup exhibiting elevated socio-emotional and cognitive behavioural difficulties in childhood (Study #3), relative to subgroups displaying optimal outcomes. **Together, these findings indicate that structural alterations in the brainstem may act as a neurobiological marker for specific behavioural phenotypes involving socio-emotional difficulties and autistic-like behaviours which are likely mediated by cognitive difficulties or developmental delay.**

#### *7.4.1.3 Brain alterations as potential markers of outcome in both VPT and FT samples*

The observed structural and functional alterations in fronto-temporo-limbic and brainstem areas were identified using analyses with samples of either only VPT individuals (Study #1-2) or both VPT and FT individuals (Studies #3-4). **This suggests that those biomarkers identified by the present analyses may be independent of birth status and that alterations to fronto-temporo-limbic and brainstem regions may reflect neural signatures for specific behavioural difficulties that are generalisable to both VPT and FT clinical groups.** Further

confirming this speculation, results from Study #3 found that region-specific functional connectivity alterations and volumetric reductions were significantly differentiating between distinct behavioural subgroups after statistically accounting for birth status. Moreover, specific region-to-region connections contributing to the significant functional connectivity components distinguishing between the distinct behavioural subgroups demonstrated negligible overlap with those differentiating between VPT and FT groups (Study #3-4).

### 7.4.2 Neurobiological markers of behavioural heterogeneity which may be specific to VPT samples

According to findings in Studies #1-4, it can be speculated that **neural changes to cerebellar and visual processing brain regions may be acting as neurobiological markers underlying behavioural outcomes specifically in those born VPT.** Volumetric reductions in the cerebellum of VPT infants at term-equivalent age (Study #1) differentiated the critical positive M-CHAT subgroup, displaying elevated levels of developmental delay and childhood autistic traits, from the non-critical positive M-CHAT subgroup displaying lower levels of developmental delay and the negative M-CHAT subgroup exhibiting optimal behavioural outcomes. Furthermore, smaller cerebellar and occipital volumes were seen to differentiate between VPT and FT children exhibiting unfavourable behavioural profiles characterised by autistic traits, psychopathology, executive function deficits, and socio-emotional processing difficulties, relative to those displaying optimal behavioural outcomes (Study #3). Importantly, however, the effects observed in cerebellar and visual areas in Study #3 were no longer significant after statistically adjusting for birth status (i.e., VPT vs FT birth), implying that these brain-behavioural patterns may be intrinsically associated with birth status. Furthermore, VPT and FT adults exhibited increased functional connectivity between the VAN and the visual network, which is argued here to be representing a functionally adaptive neural re-organisation which may be supporting optimal behavioural processing. Supporting these speculations, previous studies in preterm samples have also reported evidence of brain-behavioural associations involving visual and cerebellar regions in preterm samples which were not detected in FT samples, such that neural alterations in those regions were mitigating against unfavourable social and cognitive functioning outcomes in preterm individuals (Lean *et al.*, 2017; Wheelock *et al.*, 2018, 2021).

## 7.5 Theme 4: The role of environmental and clinical factors in explaining brain-behavioural heterogeneity

Overall, across Studies #1-4, social risk was higher in subgroups exhibiting unfavourable behavioural outcomes, relative to those displaying optimal outcomes. However, social risk was only found to be different between VPT individuals belonging to distinct behavioural subgroups, but not in FT individuals (Studies #3-4). **This may suggest that VPT individuals exposed to environmental risk experience a 'double-hit'; whereby, their preterm birth acts as the 'first hit' posing an initial vulnerability to developing behavioural difficulties, while the environmental adversity acts as a 'second hit' which heightens their risk of developing difficulties** when compared to their peers who did not experience adversity or were born FT.

Notably, however, social risk does not appear to fully account for between-subgroup differences in behavioural outcomes, as these remained significant even after adjusting for social risk (Studies #1-4). This suggests that there is potentially a complex myriad of factors interacting with one another to mediate distinct behavioural outcomes. In fact, results in Study #2 depicted a complex interplay between clinical and environmental factors in VPT samples. In this study, a particularly innovative approach was used to incorporate clinical and environmental data as input features in the clustering model alongside a combination of behavioural measures. As a result of which, nuanced relationships were delineated, and findings emphasised the importance of an enriching environment in nurturing healthy behavioural development. Lower social risk also appeared to outweigh the detrimental effects of high perinatal clinical risk in one subgroup with optimal outcomes relative to another subgroup with unfavourable outcomes which was characterised by greater social risk and lower perinatal clinical risk (Study #2), supporting findings seen previously in (Poehlmann *et al.*, 2015).

Moreover, whilst Study #2 and other previous studies (Ross *et al.*, 2016; Johnson *et al.*, 2018) found that perinatal clinical risk was significantly greater in children with unfavourable outcomes relative to those exhibiting more favourable outcomes in early childhood, a larger number of studies (Lean *et al.*, 2020; van Houdt *et al.*, 2020; Bogičević *et al.*, 2021), including Studies #1, 3 and 4, found that perinatal clinical risk did not differ between the distinct subgroups. While at first glance, discrepancies between Study #2 in early childhood and Studies #3-4 in middle childhood and adulthood may suggest that perinatal clinical risk no longer poses an increased risk for behavioural difficulties beyond early childhood; on the other hand, however, the true effect of perinatal clinical risk on behavioural outcomes could have been obscured by the use of different

methodological approaches. Studies #3-4 did not include risk factors as input features in the clustering model, unlike Study #2 which did and managed to delineate nuanced patterns differentiating between distinct subgroups. **The complex dynamics between clinical and environmental factors may help explain the inconsistencies in terms of findings reporting an association between elevated perinatal clinical risk and unfavourable behavioural outcomes in preterm samples.**

Finally, in the work presented in this thesis, the severity of perinatal brain injury in VPT children or adults did not differ between distinct behavioural subgroups (Studies #3-4), confirming previous findings reporting poor predictive validity of perinatal brain injury in predicting onset of later behavioural difficulties (Isaacs *et al.*, 2004; Burkitt *et al.*, 2019). **These findings further highlight the advantages of using advanced neuroimaging tools such as whole-brain MRI analyses to characterise intricate structural and functional brain changes potentially underlying behavioural sequelae.**

## 7.6 Implications, limitations, and future directions

### 7.6.1 Clinical and real-world implications

The work presented in this PhD thesis provides insights into the intricate interplay between brain and behavioural alterations following VPT and FT birth, which can offer valuable implications for clinical practitioners and researchers alike. For instance, it has been consistently observed across the four experimental studies presented here, that neurobiological alterations to specific brain regions such as the brainstem, fronto-temporo-limbic, cerebellar, or visual areas may confer an individual's elevated risk of developing specific behavioural difficulties later in life. **This suggests that utilising MRI to detect neurobiological markers predictive of behavioural outcomes, could potentially help identify high-risk individuals in need of targeted treatment plans early in life.**

**Once individuals vulnerable to behavioural deficits are identified, personalised care could be achieved by devising targeted interventions which can improve social, cognitive, and emotional behavioural outcomes.** This could be achieved by stimulating brain activity in implicated regions such as the brainstem and prefrontal cortex through non-invasive interventions including low-intensity exercise or deep-brain stimulation (Mazzone *et al.*, 2005; Zanini *et al.*, 2009; Dietrich and Audiffren, 2011; Ludyga, Ishihara and Kamijo, 2022). Similarly, environmental interventions exposing individuals to music or maternal stimuli in the format of parental praise,

skin-to-skin contact, breastfeeding, or visual stimuli of the mother's face have also been seen to enhance neuroplastic changes in the brain, suggesting that they may promote fronto-temporo-limbic circuitry and in turn foster optimal behavioural outcomes (Milgrom *et al.*, 2010; Gee *et al.*, 2014; Matsudaira *et al.*, 2016; Maitre *et al.*, 2017; Lordier *et al.*, 2019). Previous reports have also shown that transdiagnostic behavioural processes, such as emotion regulation or executive functions, are in fact malleable traits which can be altered through behavioural interventions, such as cognitive training, moderate-intensity exercise, or neurofeedback training (Stamenova and Levine, 2019; Yu, Tseng and Lin, 2020; Wang *et al.*, 2023).

Findings from this thesis suggest that brainstem and fronto-temporo-limbic neurobiological changes are likely arising independently of birth status, indicating that intervention strategies may be generalisable and beneficial to both VPT and FT born individuals exhibiting those neural alterations. Furthermore, based on results presented in this thesis, it has been speculated that neural alterations localised to cerebellar and visual processing areas may be specifically altered in VPT samples and are probably not generalisable to FT populations. This, therefore, suggests that there may be a potential benefit of introducing interventions which promote visual processing to facilitate attention, learning, and communication abilities in those born VPT who are exhibiting cerebellar or visual brain alterations, as it may indicate that they are at an elevated risk of developing socio-emotional and cognitive difficulties (Alimović and Mejaski-Bosnjak, 2011; Bobek and Tversky, 2016; Burstein, Zevin and Geva, 2021). In addition, the work presented in this thesis provides insights into the importance of fostering an enriching environment in order to promote healthy development and optimal behavioural processing throughout the lifespan, which is especially critical in VPT samples displaying elevated clinical risk.

Taken together, results indicate the importance of taking into account neurobiological, clinical, as well as environmental risk factors to characterise improved prognostic trajectories. **Therefore, it can be concluded that identifying the neural markers of specific behavioural outcomes and elucidating the role of additional clinical and environmental risk factors could help enhance diagnostic prediction and in turn guide the implementation of targeted treatment plans for at-risk individuals.**

### 7.6.2 Strengths

This thesis presents a series of studies implementing rigorous methodological approaches in order to capture novel and nuanced brain-behavioural associations in VPT samples, whilst also acknowledging the involvement of clinical and environmental factors and delineating

neurodevelopmental profiles which may be exhibited by both VPT and FT samples. It presents the first studies to elucidate intricate neurobiological changes associated with distinct behavioural profiles in VPT samples, while also identifying brain-behavioural phenotypes which are generalisable to both VPT and FT samples, regardless of birth status.

Essentially, the use of both behavioural profile stratification and advanced whole-brain neuroimaging analysis approaches in this thesis represent a fundamental and critical key strength relative to previous brain-behavioural studies in VPT samples, which often rely on traditional methods to capture behavioural changes at the group level or use qualitative measures of perinatal brain injury to characterise brain alterations underlying behavioural heterogeneity. Additional strengths associated with this thesis are related to the inclusion of large samples, longitudinal analyses, and investigations of brain-behavioural heterogeneity across different time points over the lifespan. Furthermore, there are several notable strengths associated with this thesis which reinforce the methodological rigour and result interpretability. Firstly, advanced neuroimaging analysis approaches incorporating mass-univariate statistical testing and robust corrections for multiple comparisons were used to examine brain-behavioural associations at the whole-brain level with a reduced risk of Type 1 (false positive) or Type 2 (false negative) errors arising. Secondly, all f-MRI analyses included in this thesis rigorously accounted for head motion during scanning, by statistically adjusting for head motion in all studies and implementing a novel 'task-free' movie paradigm in Study #3, which was specifically designed to improve compliance and reduce motion in children during f-MRI scanning. Thirdly, the use of data-driven stratification approaches which incorporated robust consensus clustering methodology (in Studies #2-4) and an integration step (in Study #2), represents a particular key strength of the studies examined here, as these steps effectively address concerns related to overfitting, accuracy, and validity. Together, the described strengths add depth and credibility to the findings, further solidifying the impact and relevance of this work within the field.

### 7.6.3 Limitations

Despite the numerous strengths discussed, it is, nonetheless, of particular importance to also acknowledge several limitations. Firstly, a significant limitation relates to the generalisability of findings. Results are likely not generalisable to all VPT individuals, as samples included in the final analyses comprised VPT individuals with a relatively lower risk of severe behavioural and developmental difficulties. Exclusion criteria in Studies #3 and #4 resulted in the removal of children and adults exhibiting excessive head motion during f-MRI scanning or those who did not

successfully complete MRI scanning. This may have inadvertently resulted in a final sample that is not representative of the broader spectrum of VPT or FT individuals, as those at a higher risk of anxiety or neurodevelopmental difficulties are more likely to experience challenges during MRI procedures (Sarji *et al.*, 1998; Kong *et al.*, 2014; Pardoe, Kucharsky Hiess and Kuzniecky, 2016; Caballero, Mistry and Torres, 2020). Moreover, the use of samples from larger longitudinal studies, such as the ePrime and UCHL cohorts, introduces a potential bias associated with the retained samples at later follow-up time points, as individuals with higher clinical and social risk are less likely to return for assessments (as seen in Study #2). **These limitations underscore the importance of considering the specific characteristics of the included population when interpreting and applying the thesis findings to broader VPT cohorts.**

Secondly, another limitation concerns the comparability of findings from Studies #1-3 and Study #4. That is because the former and latter are using distinct cohort study datasets (ePrime and UCHL respectively), which include samples of VPT and FT individuals from two different generations. The rapid advancement in neonatal medicine over the past few decades suggests that the ePrime study cohort may comprise a healthier sample of VPT individuals (Studies #1-3), relative to those in the UCHL study (Study #4). The high rates of individuals with perinatal brain injury in the UCHL cohort (47%) confirm this speculation. **These limitations make it challenging to accurately draw longitudinal projections to contemporary samples based on findings from the UCHL cohort.**

An additional bias may also be associated with the comparability of Study #3 with studies taking place during a different point in time, as the duration of the COVID-19 pandemic overlapped with the period during which data for this study were collected. The COVID-19 pandemic was associated with increases in behavioural difficulties, which affected both FT and VPT children (Ng and Ng, 2022; Theberath *et al.*, 2022; Sun *et al.*, 2023). In fact, a recent study using a subsample of the children examined in Study #3 found that VPT (but not FT) children with elevated pre-existing socio-emotional difficulties displayed greater levels of emotional problems during the pandemic, despite both groups showing no significant differences in socio-emotional/emotional behavioural difficulties at either time point (Sun *et al.*, 2023). **This suggests that there is a selective risk impacting some individuals but not others; in turn, emphasising the importance of considering the role of the pandemic as an external factor potentially biasing interpretations upon comparing results from Study #3 to others in the field.**

### 7.6.4 Future directions

Several important challenges need to be addressed in the future in order to further develop the positive impact of the findings discussed in this thesis. Firstly, implementing methods that model longitudinal behavioural trajectories, such as latent growth modelling, could help elucidate subject-specific developmental trajectories and in turn estimate the proportions of individuals with stable, improving, or deteriorating outcomes over time. These approaches not only address limitations related to cross-sectional study comparability (as discussed in the limitations section above) but also offer a nuanced understanding of individual developmental trajectories. Whilst this has been previously explored in preterm samples, those studies often examined trajectories across a specific behavioural domain, such as language or attention development (Luu *et al.*, 2009; Nguyen *et al.*, 2018; Bogičević *et al.*, 2021). However, there is a need to implement such approaches using multi-dimensional measures of behaviour spanning across multiple behavioural domains in preterm samples. Furthermore, as structural and functional features of the brain do not influence behavioural outcomes in isolation, there is promise in using novel techniques which integrate information from multiple MRI modalities to comprehensively characterise brain features. By holistically quantifying brain alterations using these advanced multi-modal neuroimaging analysis methods, a more detailed exploration of neurobiological markers associated with behavioural outcomes could be examined (Calhoun and Sui, 2016; Ball *et al.*, 2017).

## 7.7 Conclusions

This thesis provides novel findings in terms of specific structural and functional brain alterations underlying behavioural heterogeneity captured using multi-dimensional measures of cognitive, psychopathology, and socio-emotional processing in VPT and FT samples across the lifespan. Neurobiological alterations to specific brain areas such as the fronto-temporo-limbic and brainstem regions were seen to be associated with behavioural outcomes related to cognitive, psychopathology, autism, and socio-emotional processing difficulties. Namely, fronto-temporo-limbic alterations were associated with generalised difficulties spanning most behavioural domains examined, while brainstem alterations tended to be associated with difficulties relating to socio-emotional processing and cognitive, language, and sensorimotor developmental delays. Interestingly, these brain-behavioural relationships were not unique to VPT samples but were also seen in certain subsets of FT born individuals. This indicates the presence of brain-behavioural profiles which are generalisable to specific subgroups comprised of both VPT and FT individuals, irrespective of birth status. On the other hand, however, it is speculated here, that unique brain

changes localised to visual processing and cerebellar areas are more likely to represent neural changes that are specific to VPT samples. These findings suggest a potential benefit of using MRI to detect neurobiological markers of behavioural outcomes, which can in turn guide the implementation of personalised behavioural, environmental, and clinical interventions for those at-risk of developing specific behavioural difficulties. Results also highlight the crucial role of fostering an enriching environment to probe resilience against developing behavioural difficulties, particularly for those born VPT. The rigorous data-driven stratification pipelines and advanced whole-brain analyses applied here strengthen result validity and accuracy, allowing for nuanced brain-behavioural associations in VPT and FT samples to be identified. Future research studies can extend the impact of these findings by implementing longitudinal behavioural modelling and multi-modal neuroimaging analyses.

# References

Aarnoudse-Moens, C.S.H. *et al.* (2009) 'Meta-analysis of neurobehavioral outcomes in very preterm and/or very low birth weight children', *Pediatrics*, 124(2), pp. 717–728. Available at: https://doi.org/10.1542/peds.2008-2816.

Abernethy, L.J., Cooke, R.W.I. and Foulder-Hughes, L. (2004) 'Caudate and hippocampal volumes, intelligence, and motor impairment in 7-year-old children who were born preterm', *Pediatric Research*, 55(5), pp. 884–893. Available at: https://doi.org/10.1203/01.PDR.0000117843.21534.49.

Abernethy, L.J., Palaniappan, M. and Cooke, R.W.I. (2002) 'Quantitative magnetic resonance imaging of the brain in survivors of very low birth weight', *Archives of Disease in Childhood*, 87(4), pp. 279–283. Available at: https://doi.org/10.1136/adc.87.4.279.

Abraham, A. *et al.* (2014) 'Machine learning for neuroimaging with scikit-learn', *Frontiers in Neuroinformatics*, 8. Available at: https://www.frontiersin.org/articles/10.3389/fninf.2014.00014.

Agrawal, S. *et al.* (2018) 'Prevalence of Autism Spectrum Disorder in preterm infants: A Meta-analysis', *Pediatrics*, 142(3), p. e20180134. Available at: https://doi.org/10.1542/peds.2018-0134.

Ahuja, R., Magnanti, T. and Orlin, J. (1993) *Network flows: Theory, Algorithms, and Applications*. Prentice Hall.

Akarca, D. *et al.* (2021) 'A generative network model of neurodevelopmental diversity in structural brain organization', *Nature Communications*, 12(1), p. 4216. Available at: https://doi.org/10.1038/s41467-021-24430-z.

Alcalá-López, D. *et al.* (2018) 'Computing the social brain connectome across systems and states', *Cerebral Cortex*, 28(7), pp. 2207–2232. Available at: https://doi.org/10.1093/cercor/bhx121.

Al-Haddad, B.J.S. *et al.* (2019) 'The fetal origins of mental illness', *American Journal of Obstetrics and Gynecology*, 221(6), pp. 549–562. Available at: https://doi.org/10.1016/j.ajog.2019.06.013.

Alimović, S. and Mejaski-Bosnjak, V. (2011) 'Stimulation of functional vision in children with perinatal brain damage', *Collegium Antropologicum*, 35 Suppl 1, pp. 3–9.

Allison, C., Auyeung, B. and Baron-Cohen, S. (2012) 'Toward brief "red flags" for autism screening: The short autism spectrum quotient and the short quantitative checklist in 1,000 cases and 3,000 controls', *Adolescent Psychiatry*, 51(2), p. 18. Available at: https://doi.org/10.1016/j.jaac.2011.11.003

Als, H. *et al.* (2004) 'Early experience alters brain function and structure', *Pediatrics*, 113(4), pp. 846–857. Available at: https://doi.org/10.1542/peds.113.4.846.

Als, H. *et al.* (2012) 'NIDCAP improves brain function and structure in preterm infants with severe intrauterine growth restriction', *Journal of Perinatology: Official Journal of the California Perinatal Association*, 32(10), pp. 797–803. Available at: https://doi.org/10.1038/jp.2011.201.

Amati, D. and Shallice, T. (2007) 'On the emergence of modern humans', *Cognition*, 103(3), pp. 358–385. Available at: https://doi.org/10.1016/j.cognition.2006.04.002.

American Psychiatric Association (2013) 'Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, DSM-5.' Washington, DC: American Psychiatric Association. Available at: https://doi.org/10.1176/appi.books.9780890425596.

Anderson, J. *et al.* (2021) 'Immune profiling of cord blood from preterm and term infants reveals distinct differences in pro-inflammatory responses', *Frontiers in Immunology*, 12, p. 777927. Available at: https://doi.org/10.3389/fimmu.2021.777927.

Anderson, P.J. *et al.* (2021) 'Psychiatric disorders in individuals born very preterm / very low-birth weight: An individual participant data (IPD) meta-analysis', *EClinicalMedicine*, 42, p. 101216. Available at: https://doi.org/10.1016/j.eclinm.2021.101216.

Andersson, J.L.R. and Sotiropoulos, S.N. (2016) 'An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging', *NeuroImage*, 125, pp. 1063–1078. Available at: https://doi.org/10.1016/j.neuroimage.2015.10.019.

Arichi, T. *et al.* (2017) 'Localization of spontaneous bursting neuronal activity in the preterm human brain with simultaneous EEG-fMRI', *eLife*. Edited by S. Kastner, 6, p. e27814. Available at: https://doi.org/10.7554/eLife.27814.

Ashburner, J. and Friston, K.J. (2009) 'Voxel Based Morphometry', in L.R. Squire (ed.) *Encyclopedia of Neuroscience*. Oxford: Academic Press, pp. 471–477. Available at: https://doi.org/10.1016/B978-008045046-9.00306-5.

Astle, D.E. *et al.* (2019) 'Remapping the cognitive and neural profiles of children who struggle at school', *Developmental Science*, 22(1), p. e12747. Available at: https://doi.org/10.1111/desc.12747.

Avants, B. and Gee, J.C. (2004) 'Geodesic estimation for large deformation anatomical shape averaging and interpolation', *NeuroImage*, 23 Suppl 1, pp. S139-150. Available at: https://doi.org/10.1016/j.neuroimage.2004.07.010.

Avants, B.B. *et al.* (2008a) 'Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain', *Medical Image Analysis*, 12(1), pp. 26–41. Available at: https://doi.org/10.1016/j.media.2007.06.004.

Avants, B.B. *et al.* (2008b) 'Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain', *Medical Image Analysis*, 12(1), pp. 26–41. Available at: https://doi.org/10.1016/j.media.2007.06.004.

Avants, B.B. *et al.* (2011) 'A Reproducible Evaluation of ANTs Similarity Metric Performance in Brain Image Registration', *NeuroImage*, 54(3), pp. 2033–2044. Available at: https://doi.org/10.1016/j.neuroimage.2010.09.025.

Back, S.A. (2017) 'White matter injury in the preterm infant: pathology and mechanisms', *Acta neuropathologica*, 134(3), pp. 331–349. Available at: https://doi.org/10.1007/s00401-017-1718-6.

Bailey, B.A. *et al.* (2018) 'The role of executive functioning and academic achievement in the academic self-concept of children and adolescents referred for neuropsychological assessment', *Children*, 5(7), p. 83. Available at: https://doi.org/10.3390/children5070083.

Baker, J.T. *et al.* (2019) 'Functional connectomics of affective and psychotic pathology', *Proceedings of the National Academy of Sciences*, 116(18), pp. 9050–9059. Available at: https://doi.org/10.1073/pnas.1820780116.

Bala, J.J. *et al.* (2023) 'Association between 5-min Apgar score and attention deficit hyperactivity disorder: a Scotland-wide record linkage study of 758,423 school children', *BMC Psychiatry*, 23(1), p. 794. Available at: https://doi.org/10.1186/s12888-023-05217-6.

Ball, G. *et al.* (2010) 'An optimised tract-based spatial statistics protocol for neonates: applications to prematurity and chronic lung disease', *NeuroImage*, 53(1), pp. 94–102. Available at: https://doi.org/10.1016/j.neuroimage.2010.05.055.

Ball, G. *et al.* (2013) 'The influence of preterm birth on the developing thalamocortical connectome', *Cortex*, 49(6), pp. 1711–1721. Available at: https://doi.org/10.1016/j.cortex.2012.07.006.

Ball, G. *et al.* (2016) 'Machine-learning to characterise neonatal functional connectivity in the preterm brain', *Neuroimage*, 124(Pt A), pp. 267–275. Available at: https://doi.org/10.1016/j.neuroimage.2015.08.055.

Ball, G. *et al.* (2017) 'Multimodal image analysis of clinical influences on preterm brain development', *Annals of Neurology*, 82(2), pp. 233–246. Available at: https://doi.org/10.1002/ana.24995.

Barkus, E. and Badcock, J.C. (2019) 'A transdiagnostic perspective on social anhedonia', *Frontiers in Psychiatry*, 10. Available at: https://www.frontiersin.org/articles/10.3389/fpsyt.2019.00216.

Batalle, D., Edwards, A.D. and O'Muircheartaigh, J. (2018) 'Annual research review: Not just a small adult brain: Understanding later neurodevelopment through imaging the neonatal brain', *Journal of Child Psychology and Psychiatry*, 59(4), pp. 350–371. Available at: https://doi.org/10.1111/jcpp.12838.

Bathelt, J. *et al.* (2018) 'Data-driven subtyping of executive function-related behavioral problems in children', *Journal of the American Academy of Child and Adolescent Psychiatry*, 57(4), pp. 252-262.e4. Available at: https://doi.org/10.1016/j.jaac.2018.01.014.

Baum, G.L. *et al.* (2020) 'Development of structure–function coupling in human brain networks during youth', *Proceedings of the National Academy of Sciences*, 117(1), pp. 771–778. Available at: https://doi.org/10.1073/pnas.1912034117.

Bäuml, J.G. *et al.* (2015) 'Correspondence between aberrant intrinsic network connectivity and gray-matter volume in the ventral brain of preterm born adults', *Cerebral Cortex*, 25(11), pp. 4135–4145. Available at: https://doi.org/10.1093/cercor/bhu133.

Bayley, N. (2006) 'Bayley Scales of Infant and Toddler Development 3rd edn.', *San Antonio, TX: Harcourt Assessment Inc.*

Beckmann, C.F. and Smith, S.M. (2004) 'Probabilistic independent component analysis for functional magnetic resonance imaging', *IEEE transactions on medical imaging*, 23(2), pp. 137–152. Available at: https://doi.org/10.1109/TMI.2003.822821.

Beekman, A.T. *et al.* (2000) 'Anxiety and depression in later life: Co-occurrence and communality of risk factors', *The American Journal of Psychiatry*, 157(1), pp. 89–95. Available at: https://doi.org/10.1176/ajp.157.1.89.

Behzadi, Y. *et al.* (2007) 'A component based noise correction method (CompCor) for BOLD and perfusion based fMRI', *NeuroImage*, 37(1), pp. 90–101. Available at: https://doi.org/10.1016/j.neuroimage.2007.04.042.

Belsky, J., Bakermans-Kranenburg, M.J. and van IJzendoorn, M.H. (2007) 'For better and for worse: differential susceptibility to environmental influences', *Current Directions in Psychological Science*, 16(6), pp. 300–304. Available at: https://doi.org/10.1111/j.1467-8721.2007.00525.x.

Belsky, J. and Pluess, M. (2009) 'Beyond diathesis stress: Differential susceptibility to environmental influences', *Psychological Bulletin*, 135(6), pp. 885–908. Available at: https://doi.org/10.1037/a0017376.

Benallie, K.J. *et al.* (2021) 'Executive functioning in children with ASD + ADHD and ASD + ID: A systematic review', *Research in Autism Spectrum Disorders*, 86, p. 101807. Available at: https://doi.org/10.1016/j.rasd.2021.101807.

Benavente-Fernández, I. *et al.* (2019) 'Association of socioeconomic status and brain injury with neurodevelopmental outcomes of very preterm children', *JAMA Network Open*, 2(5), pp. e192914–e192914. Available at: https://doi.org/10.1001/jamanetworkopen.2019.2914.

Benjamini, Y. and Hochberg, Y. (1995) 'Controlling the False Discovery Rate: A practical and powerful approach to multiple testing', *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1), pp. 289–300.

Benton, A.L., Hamsher, d. S.K. and Sivan, A.B. (1983) *Controlled Oral Word Association Test*. Available at: https://psycnet.apa.org/doiLanding?doi=10.1037%2Ft10132-000.

Berger, A. (2002) 'Magnetic resonance imaging', *BMJ : British Medical Journal*, 324(7328), p. 35.

Bhalla, S. *et al.* (2021) 'Patient similarity network of newly diagnosed multiple myeloma identifies patient subgroups with distinct genetic features and clinical implications', *Science Advances*, 7(47), p. eabg9551. Available at: https://doi.org/10.1126/sciadv.abg9551.

Biella, M.M. *et al.* (2019) 'Subthreshold depression needs a prime time in old age psychiatry? a narrative review of current evidence', *Neuropsychiatric Disease and Treatment*, 15, pp. 2763–2772. Available at: https://doi.org/10.2147/NDT.S223640.

Binding, L.P. *et al.* (2022) 'Structure and function of language networks in temporal lobe epilepsy', *Epilepsia*, 63(5), pp. 1025–1040. Available at: https://doi.org/10.1111/epi.17204.

Bishop, D.V.M. (2000) 'Pragmatic language impairment: A correlate of SLI, a distinct subgroup, or part of the autistic continuum?', in *Speech and Language Impairments in Children*. Psychology Press.

Blencowe, H. *et al.* (2013) 'Preterm birth–associated neurodevelopmental impairment estimates at regional and global levels for 2010', *Pediatric Research*, 74(1), pp. 17–34. Available at: https://doi.org/10.1038/pr.2013.204.

Bobek, E. and Tversky, B. (2016) 'Creating visual explanations improves learning', *Cognitive Research: Principles and Implications*, 1, p. 27. Available at: https://doi.org/10.1186/s41235-016-0031-6.

Bogičević, L. *et al.* (2021) 'Individual attention patterns in children born very preterm and full term at 7 and 13 years of age', *Journal of the International Neuropsychological Society*, 27(10), pp. 970–980. Available at: https://doi.org/10.1017/S1355617720001411.

Bolbocean, C. *et al.* (2023) 'Health-related quality-of-life outcomes of very preterm or very low birth weight adults: evidence from an individual participant data meta-analysis', *PharmacoEconomics*, 41(1), pp. 93–105. Available at: https://doi.org/10.1007/s40273-022-01201-2.

Booth, T. *et al.* (2013) 'Brief report: An evaluation of the AQ-10 as a brief screening instrument for asd in adults', *Journal of Autism and Developmental Disorders*, 43(12), pp. 2997–3000. Available at: https://doi.org/10.1007/s10803-013-1844-5.

Bora, S. *et al.* (2011) 'Emotional and behavioural adjustment of children born very preterm at early school age', *Journal of Paediatrics and Child Health*, 47(12), pp. 863–869. Available at: https://doi.org/10.1111/j.1440-1754.2011.02105.x.

Bouyssi-Kobar, M. *et al.* (2016) 'Third trimester brain growth in preterm infants compared with in utero healthy fetuses', *Pediatrics*, 138(5), p. e20161640. Available at: https://doi.org/10.1542/peds.2016-1640.

Bowerman, R.A. *et al.* (1984) 'Natural history of neonatal periventricular/intraventricular hemorrhage and its complications: sonographic observations', *AJR. American journal of roentgenology*, 143(5), pp. 1041–1052. Available at: https://doi.org/10.2214/ajr.143.5.1041.

Braga, R.M. *et al.* (2020) 'Situating the left-lateralized language network in the broader organization of multiple specialized large-scale distributed networks', *Journal of Neurophysiology*, 124(5), pp. 1415–1448. Available at: https://doi.org/10.1152/jn.00753.2019.

Braver, T.S. and Bongiolatti, S.R. (2002) 'The role of frontopolar cortex in subgoal processing during working memory', *NeuroImage*, 15(3), pp. 523–536. Available at: https://doi.org/10.1006/nimg.2001.1019.

Briggs, R.G. *et al.* (2021) 'The unique fiber anatomy of middle temporal gyrus default mode connectivity', *Operative Neurosurgery*, 21(1), pp. E8–E14. Available at: https://doi.org/10.1093/ons/opab109.

Briggs-Gowan, M.J. *et al.* (2003) 'Persistence of psychiatric disorders in pediatric settings', *Journal of the American Academy of Child and Adolescent Psychiatry*, 42(11), pp. 1360–1369. Available at: https://doi.org/10.1097/01.CHI.0000084834.67701.8a.

Brouwer, M.J. *et al.* (2017) 'Preterm brain injury on term-equivalent age MRI in relation to perinatal factors and neurodevelopmental outcome at two years', *PLOS ONE*, 12(5), p. e0177128. Available at: https://doi.org/10.1371/journal.pone.0177128.

Buckner, R.L. *et al.* (2009) 'Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to alzheimer's disease', *Journal of Neuroscience*, 29(6), pp. 1860–1873. Available at: https://doi.org/10.1523/JNEUROSCI.5062-08.2009.

Buckner, R.L. (2013) 'The brain's default network: origins and implications for the study of psychosis', *Dialogues in Clinical Neuroscience*, 15(3), pp. 351–358. https://doi.org/10.31887/DCNS.2013.15.3/rbuckner

Burgess, P.W. and Shallice, T. (1997) *The Hayling and Brixton tests*. Bury St Edmunds: Thames Valley Test Company.

Burkitt, K. *et al.* (2019) 'Comparison of cranial ultrasound and MRI for detecting BRAIN injury in extremely preterm infants and correlation with neurological outcomes at 1 and 3 years', *European Journal of Pediatrics*, 178(7), pp. 1053–1061. Available at: https://doi.org/10.1007/s00431-019-03388-7.

Burnett, A.C. *et al.* (2019) 'Exploring the "preterm behavioral phenotype" in children born extremely preterm', *Journal of developmental and behavioral pediatrics: JDBP*, 40(3), pp. 200–207. Available at: https://doi.org/10.1097/DBP.0000000000000646.

Burstein, O., Zevin, Z. and Geva, R. (2021) 'Preterm birth and the development of visual attention during the first 2 years of life', *JAMA Network Open*, 4(3), p. e213687. Available at: https://doi.org/10.1001/jamanetworkopen.2021.3687.

Button, K.S. *et al.* (2013) 'Power failure: why small sample size undermines the reliability of neuroscience', *Nature Reviews Neuroscience*, 14(5), pp. 365–376. Available at: https://doi.org/10.1038/nrn3475.

Caballero, C., Mistry, S. and Torres, E.B. (2020) 'Age-dependent statistical changes of involuntary head motion signatures across autism and controls of the ABIDE repository', *Frontiers in Integrative Neuroscience*, 14. Available at: https://www.frontiersin.org/articles/10.3389/fnint.2020.00023.

Calhoun, V.D. *et al.* (2001) 'A method for making group inferences from functional MRI data using independent component analysis', *Human Brain Mapping*, 14(3), p. 140. Available at: https://doi.org/10.1002/hbm.1048.

Calhoun, V.D. and Sui, J. (2016) 'Multimodal fusion of brain imaging data: A key to finding the missing link(s) in complex mental illness', *Biological Psychiatry. Cognitive Neuroscience and Neuroimaging*, 1(3), pp. 230–244. Available at: https://doi.org/10.1016/j.bpsc.2015.12.005.

Caspi, A. *et al.* (2014) 'The p Factor: One general psychopathology factor in the structure of psychiatric disorders?', *Clinical psychological science : a journal of the Association for Psychological Science*, 2(2), pp. 119–137. Available at: https://doi.org/10.1177/2167702613497473.

Caspi, A. and Moffitt, T.E. (2018) 'All for one and one for all: mental disorders in one dimension', *The American Journal of Psychiatry*, 175(9), pp. 831–844. Available at: https://doi.org/10.1176/appi.ajp.2018.17121383.

Cavalli, F.M.G. *et al.* (2017) 'Intertumoral heterogeneity within medulloblastoma subgroups', *Cancer Cell*, 31(6), pp. 737-754.e6. Available at: https://doi.org/10.1016/j.ccell.2017.05.005.

Chaddock-Heyman, L. *et al.* (2018) 'Scholastic performance and functional connectivity of brain networks in children', *PLoS ONE*, 13(1), p. e0190073. Available at: https://doi.org/10.1371/journal.pone.0190073.

Chai, X.J. *et al.* (2014) 'Selective development of anticorrelated networks in the intrinsic functional organization of the human brain', *Journal of Cognitive Neuroscience*, 26(3), pp. 501–513. Available at: https://doi.org/10.1162/jocn_a_00517.

Chawanpaiboon, S. *et al.* (2019) 'Global, regional, and national estimates of levels of preterm birth in 2014: a systematic review and modelling analysis', *The Lancet Global Health*, 7(1), pp. e37–e46. Available at: https://doi.org/10.1016/S2214-109X(18)30451-0.

Chen, H.-F. *et al.* (2020) 'Microstructural disruption of the right inferior fronto-occipital and inferior longitudinal fasciculus contributes to WMH-related cognitive impairment', *CNS neuroscience & therapeutics*, 26(5), pp. 576–588. Available at: https://doi.org/10.1111/cns.13283.

Chiu, D.S. and Talhouk, A. (2018) 'diceR: an R package for class discovery using an ensemble driven approach', *BMC bioinformatics*, 19(1), p. 11. Available at: https://doi.org/10.1186/s12859-017-1996-y.

Chlebowski, C. *et al.* (2013) 'Large-scale use of the modified checklist for autism in low-risk toddlers', *Pediatrics*, 131(4), pp. e1121–e1127. Available at: https://doi.org/10.1542/peds.2012-1525.

Cho, H.J. *et al.* (2022) 'Altered functional connectivity in children born very preterm at school age', *Scientific Reports*, 12(1), p. 7308. Available at: https://doi.org/10.1038/s41598-022-11184-x.

Choi, E.J. *et al.* (2018) 'Language network function in young children born very preterm', *Frontiers in Human Neuroscience*, 12, p. 512. Available at: https://doi.org/10.3389/fnhum.2018.00512.

Ciarrusta, J. *et al.* (2019) 'Social brain functional maturation in newborn infants with and without a family history of autism spectrum disorder', *JAMA Network Open*, 2(4), p. e191868. Available at: https://doi.org/10.1001/jamanetworkopen.2019.1868.

Cicchetti, D. and Rogosch, F.A. (1996) 'Equifinality and multifinality in developmental psychopathology', *Development and Psychopathology*, 8(4), pp. 597–600. Available at: https://doi.org/10.1017/S0954579400007318.

Cismaru, A.L. *et al.* (2016) 'Altered amygdala development and fear processing in prematurely born infants', *Frontiers in Neuroanatomy*, 10, p. 55. Available at: https://doi.org/10.3389/fnana.2016.00055.

Clapp, M. *et al.* (2017) 'Gut microbiota's effect on mental health: The gut-brain axis', *Clinics and Practice*, 7(4), p. 987. Available at: https://doi.org/10.4081/cp.2017.987.

Clouchoux, C. *et al.* (2012) 'Normative fetal brain growth by quantitative in vivo magnetic resonance imaging', *American Journal of Obstetrics and Gynecology*, 206(2), p. 173.e1–8. Available at: https://doi.org/10.1016/j.ajog.2011.10.002.

Conners, C. *et al.* (2003) 'Continuous Performance Test Performance in a Normative Epidemiological Sample', *Journal of abnormal child psychology*. Available at: https://doi.org/10.1023/A:1025457300409.

Constable, R.T. *et al.* (2013) 'A left cerebellar pathway mediates language in prematurely-born young adults', *NeuroImage*, 64, p. 371. Available at: https://doi.org/10.1016/j.neuroimage.2012.09.008.

Constantino, J.N. *et al.* (2003) 'Validation of a brief quantitative measure of autistic traits: comparison of the social responsiveness scale with the autism diagnostic interview-revised', *Journal of Autism and Developmental Disorders*, 33(4), pp. 427–433. Available at: https://doi.org/10.1023/a:1025014929212.

Constantino, J.N. and Gruber, C.P. (2012) *Social Responsiveness Scale, Second Edition (SRS-2): manual*. 2nd edn. Los Angeles, CA: Western Psychological Services.

Cooke, R.W.I. (2004) 'Health, lifestyle, and quality of life for young adults born very preterm', *Archives of Disease in Childhood*, 89(3), pp. 201–206. Available at: https://doi.org/10.1136/adc.2003.030197.

Corbetta, M. and Shulman, G.L. (2002) 'Control of goal-directed and stimulus-driven attention in the brain', *Nature Reviews Neuroscience*, 3(3), pp. 201–215. Available at: https://doi.org/10.1038/nrn755.

Cortés Pascual, A., Moyano Muñoz, N. and Quílez Robres, A. (2019) 'The relationship between executive functions and academic performance in primary education: Review and meta-analysis', *Frontiers in Psychology*, 10. Available at: https://www.frontiersin.org/articles/10.3389/fpsyg.2019.01582.

Cox, R.W. (1996) 'AFNI: software for analysis and visualization of functional magnetic resonance neuroimages', *Computers and Biomedical Research, an International Journal*, 29(3), pp. 162–173. Available at: https://doi.org/10.1006/cbmr.1996.0014.

Crespi, C. *et al.* (2014) 'Microstructural white matter correlates of emotion recognition impairment in Amyotrophic Lateral Sclerosis', *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 53, pp. 1–8. Available at: https://doi.org/10.1016/j.cortex.2014.01.002.

Crilly, C.J., Haneuse, S. and Litt, J.S. (2021) 'Predicting the outcomes of preterm neonates beyond the neonatal intensive care unit: What are we missing?', *Pediatric Research*, 89(3), pp. 426–445. Available at: https://doi.org/10.1038/s41390-020-0968-5.

Cuijpers, P. *et al.* (2013) 'Differential mortality rates in major and subthreshold depression: meta-analysis of studies that measured both', *The British Journal of Psychiatry: The Journal of Mental Science*, 202(1), pp. 22–27. Available at: https://doi.org/10.1192/bjp.bp.112.112169.

Cuthbert, B.N. (2014) 'The RDoC framework: facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology', *World Psychiatry*, 13(1), pp. 28–35. Available at: https://doi.org/10.1002/wps.20087.

Cuthbert, B.N. and Insel, T.R. (2013) 'Toward the future of psychiatric diagnosis: the seven pillars of RDoC', *BMC Medicine*, 11(1), p. 126. Available at: https://doi.org/10.1186/1741-7015-11-126.

Daamen, M. *et al.* (2014) 'Working memory in preterm-born adults: Load-dependent compensatory activity of the posterior default mode network', *Human Brain Mapping*, 36(3), pp. 1121–1137. Available at: https://doi.org/10.1002/hbm.22691.

Dag, O., Dolgun, A. and Konar, N., Meric (2018) 'onewaytests: An R package for one-way tests in independent groups designs', *The R Journal*, 10(1), p. 175. Available at: https://doi.org/10.32614/RJ-2018-022.

Dale, A.M., Fischl, B. and Sereno, M.I. (1999) 'Cortical surface-based analysis: I. segmentation and surface reconstruction', *NeuroImage*, 9(2), pp. 179–194. Available at: https://doi.org/10.1006/nimg.1998.0395.

Dalgleish, T. *et al.* (2020) 'Transdiagnostic approaches to mental health problems: Current status and future directions', *Journal of Consulting and Clinical Psychology*, 88(3), pp. 179–195. Available at: https://doi.org/10.1037/ccp0000482.

Damaraju, E. *et al.* (2010) 'Resting-state functional connectivity differences in premature children', *Frontiers in Systems Neuroscience*, 4. Available at: https://www.frontiersin.org/articles/10.3389/fnsys.2010.00023.

Davis, E.P. *et al.* (2011) 'Children's brain development benefits from longer gestation', *Frontiers in Psychology*, 2, p. 1. Available at: https://doi.org/10.3389/fpsyg.2011.00001.

Deen, B., Pitskel, N.B. and Pelphrey, K.A. (2011) 'Three systems of insular functional connectivity identified with cluster analysis', *Cerebral Cortex*, 21(7), pp. 1498–1506. Available at: https://doi.org/10.1093/cercor/bhq186.

Degnan, A.J. *et al.* (2015a) 'Alterations of resting state networks and structural connectivity in relation to the prefrontal and anterior cingulate cortices in late prematurity', *NeuroReport*, 26(1), p. 22. Available at: https://doi.org/10.1097/WNR.0000000000000296.

Degnan, A.J. *et al.* (2015b) 'Altered structural and functional connectivity in late preterm preadolescence: an anatomic seed-based study of resting state networks related to the posteromedial and lateral parietal cortex', *PLOS ONE*, 10(6), p. e0130686. Available at: https://doi.org/10.1371/journal.pone.0130686.

Dekaban, A.S. and Sadowsky, D. (1978) 'Changes in brain weights during the span of human life: Relation of brain weights to body heights and body weights', *Annals of Neurology*, 4(4), pp. 345–356. Available at: https://doi.org/10.1002/ana.410040410.

DeMaster, D. *et al.* (2019) 'Nurturing the preterm infant brain: Leveraging neuroplasticity to improve neurobehavioral outcomes', *Pediatric Research*, 85(2), pp. 166–175. Available at: https://doi.org/10.1038/s41390-018-0203-9.

Diamond, A. (2013) 'Executive Functions', *Annual review of psychology*, 64, pp. 135–168. Available at: https://doi.org/10.1146/annurev-psych-113011-143750.

Dietrich, A. and Audiffren, M. (2011) 'The reticular-activating hypofrontality (RAH) model of acute exercise', *Neuroscience and Biobehavioral Reviews*, 35(6), pp. 1305–1325. Available at: https://doi.org/10.1016/j.neubiorev.2011.02.001.

Dimitrova, R., Arulkumaran, S., *et al.* (2021) 'Phenotyping the preterm brain: characterizing individual deviations from normative volumetric development in two large infant cohorts', *Cerebral Cortex*, 31(8), pp. 3665–3677. Available at: https://doi.org/10.1093/cercor/bhab039.

Dimitrova, R., Pietsch, M., *et al.* (2021) 'Preterm birth alters the development of cortical microstructure and morphology at term-equivalent age', *NeuroImage*, 243, p. 118488. Available at: https://doi.org/10.1016/j.neuroimage.2021.118488.

Dinan, T.G. and Cryan, J.F. (2016) 'Mood by microbe: towards clinical translation', *Genome Medicine*, 8, p. 36. Available at: https://doi.org/10.1186/s13073-016-0292-1.

Donati, G., Meaburn, E. and Dumontheil, I. (2021) 'Internalising and externalising in early adolescence predict later executive function, not the other way around: a cross-lagged panel

analysis', *Cognition & Emotion*, 35(5), pp. 986–998. Available at: https://doi.org/10.1080/02699931.2021.1918644.

Doria, V. *et al.* (2010) 'Emergence of resting state networks in the preterm human brain', *Proceedings of the National Academy of Sciences*, 107(46), pp. 20015–20020. Available at: https://doi.org/10.1073/pnas.1007921107.

Dotinga, B.M. *et al.* (2016) 'Longitudinal growth and neuropsychological functioning at age 7 in moderate and late preterms', *Pediatrics*, 138(4), p. e20153638. Available at: https://doi.org/10.1542/peds.2015-3638.

Dotinga, B.M. *et al.* (2019) 'Longitudinal growth and emotional and behavioral problems at age 7 in moderate and late preterms', *PLoS ONE*, 14(1), p. e0211427. Available at: https://doi.org/10.1371/journal.pone.0211427.

Doucet, G.E. *et al.* (2020) 'Transdiagnostic and disease-specific abnormalities in the default-mode network hubs in psychiatric disorders: A meta-analysis of resting-state functional imaging studies', *European Psychiatry*, 63(1), p. e57. Available at: https://doi.org/10.1192/j.eurpsy.2020.57.

Dubois, J. *et al.* (2019) 'The dynamics of cortical folding waves and prematurity-related deviations revealed by spatial and spectral analysis of gyrification', *NeuroImage*, 185, pp. 934–946. Available at: https://doi.org/10.1016/j.neuroimage.2018.03.005.

Ducharme, S. *et al.* (2016) 'Trajectories of cortical thickness maturation in normal brain development--The importance of quality control procedures', *NeuroImage*, 125, pp. 267–279. Available at: https://doi.org/10.1016/j.neuroimage.2015.10.010.

Duggan, E.C. and Garcia-Barrera, M.A. (2015) 'Executive functioning and intelligence', in *Handbook of intelligence: Evolutionary theory, historical perspective, and current concepts*. New York, NY, US: Springer Science + Business Media, pp. 435–458. Available at: https://doi.org/10.1007/978-1-4939-1562-0_27.

van Duinkerken, E. *et al.* (2012) 'Diffusion tensor imaging in type 1 diabetes: decreased white matter integrity relates to cognitive functions', *Diabetologia*, 55(4), pp. 1218–1220. Available at: https://doi.org/10.1007/s00125-012-2488-2.

Dumontheil, I. (2014) 'Development of abstract thinking during childhood and adolescence: The role of rostrolateral prefrontal cortex', *Developmental Cognitive Neuroscience*, 10, p. 57. Available at: https://doi.org/10.1016/j.dcn.2014.07.009.

Dumontheil, I., Burgess, P.W. and Blakemore, S.-J. (2008) 'Development of rostral prefrontal cortex and cognitive and behavioural disorders', *Developmental Medicine and Child Neurology*, 50(3), pp. 168–181. Available at: https://doi.org/10.1111/j.1469-8749.2008.02026.x.

Duncan, J. (2013) 'The structure of cognition: Attentional episodes in mind and brain', *Neuron*, 80(1), pp. 35–50. Available at: https://doi.org/10.1016/j.neuron.2013.09.015.

Durrant, C. *et al.* (2020) 'Developmental trajectories of infants born at less than 30 weeks' gestation on the Bayley-III Scales', *Archives of Disease in Childhood - Fetal and Neonatal Edition*, 105(6), pp. 623–627. Available at: https://doi.org/10.1136/archdischild-2019-317810.

Edwards, A.D. *et al.* (2018) 'Effect of MRI on preterm infants and their families: a randomised trial with nested diagnostic and economic evaluation', *Archives of Disease in Childhood - Fetal and*

*Neonatal Edition*, 103(1), pp. F15–F21. Available at: https://doi.org/10.1136/archdischild-2017-313102.

Eikenes, L. *et al.* (2011) 'Young adults born preterm with very low birth weight demonstrate widespread white matter alterations on brain DTI', *NeuroImage*, 54(3), pp. 1774–1785. Available at: https://doi.org/10.1016/j.neuroimage.2010.10.037.

Eklöf, E. *et al.* (2019) 'Reduced structural brain asymmetry during neonatal life is potentially related to autism spectrum disorders in children born extremely preterm', *Autism Research: Official Journal of the International Society for Autism Research*, 12(9), pp. 1334–1343. Available at: https://doi.org/10.1002/aur.2169.

Else-Quest, N.M. *et al.* (2006) 'Gender differences in temperament: a meta-analysis', *Psychological Bulletin*, 132(1), pp. 33–72. Available at: https://doi.org/10.1037/0033-2909.132.1.33.

Espel, E.V. *et al.* (2014) 'Longer gestation among children born full term influences cognitive and motor development', *PloS One*, 9(11), p. e113758. Available at: https://doi.org/10.1371/journal.pone.0113758.

Esteban, O. *et al.* (2019) 'fMRIPrep: a robust preprocessing pipeline for functional MRI', *Nature Methods*, 16(1), pp. 111–116. Available at: https://doi.org/10.1038/s41592-018-0235-4.

Eves, R. *et al.* (2021) 'Association of very preterm birth or very low birth weight with intelligence in adulthood: An individual participant data meta-analysis', *JAMA pediatrics*, 175(8), p. e211058. Available at: https://doi.org/10.1001/jamapediatrics.2021.1058.

Fan, Y. *et al.* (2011) 'Is there a core neural network in empathy? An fMRI based quantitative meta-analysis', *Neuroscience & Biobehavioral Reviews*, 35(3), pp. 903–911. Available at: https://doi.org/10.1016/j.neubiorev.2010.10.009.

Fenchel, D. *et al.* (2020) 'Development of microstructural and morphological cortical profiles in the neonatal brain', *Cerebral Cortex*, 30(11), pp. 5767–5779. Available at: https://doi.org/10.1093/cercor/bhaa150.

Fenn-Moltu, S. *et al.* (2022) 'Development of neonatal brain functional centrality and alterations associated with preterm birth', *Cerebral Cortex*, p. bhac444. Available at: https://doi.org/10.1093/cercor/bhac444.

Fenoglio, A., Georgieff, M.K. and Elison, J.T. (2017) 'Social brain circuitry and social cognition in infants born preterm', *Journal of Neurodevelopmental Disorders*, 9, p. 27. Available at: https://doi.org/10.1186/s11689-017-9206-9.

Fernandez, K.C., Jazaieri, H. and Gross, J.J. (2016) 'Emotion regulation: A transdiagnostic perspective on a new RDoC domain', *Cognitive Therapy and Research*, 40(3), pp. 426–440. Available at: https://doi.org/10.1007/s10608-016-9772-2.

Ferretti, V. and Papaleo, F. (2019) 'Understanding others: Emotion recognition in humans and other animals', *Genes, Brain and Behavior*, 18(1), p. e12544. Available at: https://doi.org/10.1111/gbb.12544.

Finke, K. *et al.* (2015) 'Visual attention in preterm born adults: Specifically impaired attentional sub-mechanisms that link with altered intrinsic brain networks in a compensation-like mode', *NeuroImage*, 107, pp. 95–106. Available at: https://doi.org/10.1016/j.neuroimage.2014.11.062.

Fischi-Gómez, E. *et al.* (2015) 'Structural brain connectivity in school-age preterm infants provides evidence for impaired networks relevant for higher order cognitive skills and social cognition', *Cerebral Cortex*, 25(9), pp. 2793–2805. Available at: https://doi.org/10.1093/cercor/bhu073.

Fischl, B. (2012) 'FreeSurfer', *NeuroImage*, 62(2), pp. 774–781. Available at: https://doi.org/10.1016/j.neuroimage.2012.01.021.

Fischl, B. and Dale, A.M. (2000) 'Measuring the thickness of the human cerebral cortex from magnetic resonance images', *Proceedings of the National Academy of Sciences of the United States of America*, 97(20), pp. 11050–11055. Available at: https://doi.org/10.1073/pnas.200033797.

Fonov, V. *et al.* (2011) 'Unbiased average age-appropriate atlases for pediatric studies', *NeuroImage*, 54(1), pp. 313–327. Available at: https://doi.org/10.1016/j.neuroimage.2010.07.033.

Forbes, K.P.N., Pipe, J.G. and Bird, R. (2000) 'Neonatal hypoxic-ischemic encephalopathy: detection with diffusion-weighted MR imaging', *AJNR: American Journal of Neuroradiology*, 21(8), pp. 1490–1496.

França, L., Ge, Y. and Batalle, D. (2022) 'p-testR'. Zenodo. Available at: https://doi.org/10.5281/zenodo.7051925.

França, L.G.S. *et al.* (2023) 'Neonatal brain dynamic functional connectivity: impact of preterm birth and association with early childhood neurodevelopment'. bioRxiv, p. 2022.11.16.516610. Available at: https://doi.org/10.1101/2022.11.16.516610.

Fray, P.J., Robbins, T.W. and Sahakian, B.J. (1996) 'Neuorpsychiatyric applications of CANTAB', *International Journal of Geriatric Psychiatry*, 11(4), pp. 329–336. Available at: https://doi.org/10.1002/(SICI)1099-1166(199604)11:4<329::AID-GPS453>3.0.CO;2-6.

Friston, K.J. (2011) 'Functional and effective connectivity: a review', *Brain Connectivity*, 1(1), pp. 13–36. Available at: https://doi.org/10.1089/brain.2011.0008.

Fuhrmann, D. *et al.* (2020) 'A hierarchical watershed model of fluid intelligence in childhood and adolescence', *Cerebral Cortex*, 30(1), pp. 339–352. Available at: https://doi.org/10.1093/cercor/bhz091.

Gale-Grant, O. *et al.* (2021) 'Effects of gestational age at birth on perinatal structural brain development in healthy term-born babies', *Human Brain Mapping*, 43(5), pp. 1577–1589. Available at: https://doi.org/10.1002/hbm.25743.

Gandhi, T. and Lee, C.C. (2021) 'Neural mechanisms underlying repetitive behaviors in rodent models of autism spectrum disorders', *Frontiers in Cellular Neuroscience*, 14, p. 463. Available at: https://doi.org/10.3389/fncel.2020.592710.

Ganella, E.P. *et al.* (2014) 'Abnormalities in orbitofrontal cortex gyrification and mental health outcomes in adolescents born extremely preterm and/or at an extremely low birth weight', *Human Brain Mapping*, 36(3), pp. 1138–1150. Available at: https://doi.org/10.1002/hbm.22692.

Gao, W. *et al.* (2011) 'Temporal and spatial evolution of brain network topology during the first two years of life', *PLOS ONE*, 6(9), p. e25278. Available at: https://doi.org/10.1371/journal.pone.0025278.

Gee, D.G. *et al.* (2014) 'Maternal buffering of human amygdala–prefrontal circuitry during childhood but not adolescence', *Psychological science*, 25(11), pp. 2067–2078. Available at: https://doi.org/10.1177/0956797614550878.

Gennatas, E.D. *et al.* (2017) 'Age-related effects and sex differences in gray matter density, volume, mass, and cortical thickness from childhood to young adulthood', *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 37(20), pp. 5065–5073. Available at: https://doi.org/10.1523/JNEUROSCI.3550-16.2017.

Geva, R. *et al.* (2013) 'Neonatal brainstem dysfunction risks infant social engagement', *Social Cognitive and Affective Neuroscience*, 8(2), pp. 158–164. Available at: https://doi.org/10.1093/scan/nsr082.

Geva, R. *et al.* (2014) 'Neonatal brainstem dysfunction after preterm birth predicts behavioral inhibition', *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 55(7), pp. 802–810. Available at: https://doi.org/10.1111/jcpp.12188.

Geva, R. and Feldman, R. (2008) 'A neurobiological model for the effects of early brainstem functioning on the development of behavior and emotion regulation in infants: implications for prenatal and perinatal risk', *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 49(10), pp. 1031–1041. Available at: https://doi.org/10.1111/j.1469-7610.2008.01918.x.

Gilmore, J. *et al.* (2012) 'Longitudinal development of cortical and subcortical gray matter from birth to 2 years', *Cerebral cortex*, 22(11). Available at: https://doi.org/10.1093/cercor/bhr327.

Giménez, M. *et al.* (2006) 'White matter volume and concentration reductions in adolescents with history of very preterm birth: a voxel-based morphometry study', *NeuroImage*, 32(4), pp. 1485–1498. Available at: https://doi.org/10.1016/j.neuroimage.2006.05.013.

Gioia, G.A. *et al.* (2000) 'Behavior Rating Inventory of Executive Function®, Second Edition (BRIEF®2)', p. 2.

Gioia, G.A., Isquith, P.K. and Roth, R.M. (2018) 'Behavior Rating Inventory for Executive Function', in J.S. Kreutzer, J. DeLuca, and B. Caplan (eds) *Encyclopedia of Clinical Neuropsychology*. Cham: Springer International Publishing, pp. 532–538. Available at: https://doi.org/10.1007/978-3-319-57111-9_1881.

Giuliani, N.R. *et al.* (2011) 'Emotion regulation and brain plasticity: expressive suppression use predicts anterior insula volume', *NeuroImage*, 58(1), pp. 10–15. Available at: https://doi.org/10.1016/j.neuroimage.2011.06.028.

Glasser, M.F. *et al.* (2016) 'A multi-modal parcellation of human cerebral cortex', *Nature*, 536(7615), pp. 171–178. Available at: https://doi.org/10.1038/nature18933.

Glover, G.H. (2011) 'Overview of functional magnetic resonance imaging', *Neurosurgery clinics of North America*, 22(2), pp. 133–139. Available at: https://doi.org/10.1016/j.nec.2010.11.001.

Gogtay, N. *et al.* (2004) 'Dynamic mapping of human cortical development during childhood through early adulthood', *Proceedings of the National Academy of Sciences*, 101(21), pp. 8174–8179. Available at: https://doi.org/10.1073/pnas.0402680101.

Goldberg, D.P. and Williams, P. (1991) *A user's guide to the General Health Questionnaire*. Institute of Psychiatry, University of London: NFER-NELSON.

Goodman, R. *et al.* (2000) 'The Development and Well-Being Assessment: Description and initial validation of an integrated assessment of child and adolescent psychopathology', *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 41(5), pp. 645–655.

Goodman, R. (2001) 'Psychometric properties of the Strengths and Difficulties Questionnaire (SDQ)', *Journal of the American Academy of Child and Adolescent Psychiatry*, 40, pp. 1337–45. Available at: https://doi.org/10.1097/00004583-200111000-00015.

Goodman, S.H. *et al.* (1993) 'Assessing levels of adaptive functioning: The Role Functioning Scale', *Community Mental Health Journal*, 29(2), pp. 119–131. Available at: https://doi.org/10.1007/BF00756338.

Gorgolewski, K. *et al.* (2011) 'Nipype: A flexible, lightweight and extensible neuroimaging data processing framework in python', *Frontiers in Neuroinformatics*, 5. Available at: https://www.frontiersin.org/articles/10.3389/fninf.2011.00013.

Gracia-Tabuenca, Z. and Alcauter, S. (2020) 'NBR: Network-based R-statistics for (unbalanced) longitudinal samples'. bioRxiv, p. 2020.11.07.373019. Available at: https://doi.org/10.1101/2020.11.07.373019.

Grayson, D.S. *et al.* (2014) 'Structural and functional rich club organization of the brain in children and adults', *PLoS ONE*, 9(2), p. e88297. Available at: https://doi.org/10.1371/journal.pone.0088297.

Grayson, D.S. and Fair, D.A. (2017) 'Development of large-scale functional networks from birth to adulthood: A guide to the neuroimaging literature', *NeuroImage*, 160, pp. 15–31. Available at: https://doi.org/10.1016/j.neuroimage.2017.01.079.

Green, B.C., Johnson, K.A. and Bretherton, L. (2014) 'Pragmatic language difficulties in children with hyperactivity and attention problems: an integrated review', *International Journal of Language & Communication Disorders*, 49(1), pp. 15–29. Available at: https://doi.org/10.1111/1460-6984.12056.

Greve, D.N. and Fischl, B. (2009) 'Accurate and robust brain image alignment using boundary-based registration', *NeuroImage*, 48(1), pp. 63–72. Available at: https://doi.org/10.1016/j.neuroimage.2009.06.060.

Griffanti, L. *et al.* (2014) 'ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging', *NeuroImage*, 95, pp. 232–247. Available at: https://doi.org/10.1016/j.neuroimage.2014.03.034.

Grigorenko, E.L. (2017) 'Brain development: the effect of interventions on children and adolescents', in *Child and Adolescent Health and Development. 3rd edition*. The International Bank for Reconstruction and Development / The World Bank. Available at: https://doi.org/10.1596/978-1-4648-0423-6_ch10.

Gu, S. *et al.* (2015) 'Emergence of system roles in normative neurodevelopment', *Proceedings of the National Academy of Sciences of the United States of America*, 112(44), pp. 13681–13686. Available at: https://doi.org/10.1073/pnas.1502829112.

Gunnar, M.R. *et al.* (2012) 'The BDNF Val66Met polymorphism moderates early deprivation effects on attention problems', *Development and psychopathology*, 24(4), pp. 1215–1223. Available at: https://doi.org/10.1017/S095457941200065X.

Guo, W. *et al.* (2015) 'Decreased insular connectivity in drug-naive major depressive disorder at rest', *Journal of Affective Disorders*, 179, pp. 31–37. Available at: https://doi.org/10.1016/j.jad.2015.03.028.

Ha, S. *et al.* (2015) 'Characteristics of brains in autism spectrum disorder: structure, function and connectivity across the lifespan', *Experimental Neurobiology*, 24(4), pp. 273–284. Available at: https://doi.org/10.5607/en.2015.24.4.273.

Hadaya, L. *et al.* (2022) 'Distinct neurodevelopmental trajectories in groups of very preterm children screening positively for autism spectrum conditions', *Journal of Autism and Developmental Disorders*. Available at: https://doi.org/10.1007/s10803-022-05789-4.

Hadaya, L. *et al.* (2023) 'Parsing brain-behavior heterogeneity in very preterm born children using integrated similarity networks', *Translational Psychiatry*, 13(1), pp. 1–14. Available at: https://doi.org/10.1038/s41398-023-02401-w.

Hadaya, L. *et al.* (2024) 'Exploring functional connectivity in clinical and data-driven groups of preterm and term adults'. bioRxiv, p. 2024.01.22.576651. Available at: https://doi.org/10.1101/2024.01.22.576651.

Hanson, J.L. *et al.* (2010) 'Early stress is associated with alterations in the orbitofrontal cortex: a tensor-based morphometry investigation of brain structure and behavioral risk', *The Journal of Neuroscience*, 30(22), pp. 7466–7472. Available at: https://doi.org/10.1523/JNEUROSCI.0859-10.2010.

Hausman, H.K. *et al.* (2022) 'Cingulo-opercular and frontoparietal control network connectivity and executive functioning in older adults', *GeroScience*, 44(2), pp. 847–866. Available at: https://doi.org/10.1007/s11357-021-00503-1.

Hawkins, D.M. (2004) 'The problem of overfitting', *Journal of Chemical Information and Computer Sciences*, 44(1), pp. 1–12. Available at: https://doi.org/10.1021/ci0342472.

Healy, E. *et al.* (2013) 'Preterm birth and adolescent social functioning-alterations in emotion-processing brain areas', *The Journal of Pediatrics*, 163(6), pp. 1596–1604. Available at: https://doi.org/10.1016/j.jpeds.2013.08.011.

Hebling Vieira, B. *et al.* (2021) 'A deep learning based approach identifies regions more relevant than resting-state networks to the prediction of general intelligence from resting-state fMRI', *Human Brain Mapping*, 42(18), pp. 5873–5887. Available at: https://doi.org/10.1002/hbm.25656.

Herbet, G., Zemmoura, I. and Duffau, H. (2018) 'Functional anatomy of the inferior longitudinal fasciculus: from historical reports to current hypotheses', *Frontiers in Neuroanatomy*, 12, p. 77. Available at: https://doi.org/10.3389/fnana.2018.00077.

van den Heuvel, M.P. and Sporns, O. (2013) 'An anatomical substrate for integration among functional networks in human cortex', *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 33(36), pp. 14489–14500. Available at: https://doi.org/10.1523/JNEUROSCI.2128-13.2013.

Heuvel, M.P. van den and Sporns, O. (2013) 'Network hubs in the human brain', *Trends in Cognitive Sciences*, 17(12), pp. 683–696. Available at: https://doi.org/10.1016/j.tics.2013.09.012.

Higo, T. *et al.* (2011) 'Distributed and causal influence of frontal operculum in task control', *Proceedings of the National Academy of Sciences of the United States of America*, 108(10), pp. 4230–4235. Available at: https://doi.org/10.1073/pnas.1013361108.

Hirosawa, T. *et al.* (2020) 'Different associations between intelligence and social cognition in children with and without autism spectrum disorders', *PLoS ONE*, 15(8), p. e0235380. Available at: https://doi.org/10.1371/journal.pone.0235380.

Hodel, A.S. (2018) 'Rapid infant prefrontal cortex development and sensitivity to early environmental experience', *Developmental review : DR*, 48, pp. 113–144. Available at: https://doi.org/10.1016/j.dr.2018.02.003.

Høegh, M.C. *et al.* (2022) 'Affective lability and social functioning in severe mental disorders', *European Archives of Psychiatry and Clinical Neuroscience*, 272(5), pp. 873–885. Available at: https://doi.org/10.1007/s00406-022-01380-1.

Holiga, Š. *et al.* (2019) 'Patients with autism spectrum disorders display reproducible functional connectivity alterations', *Science Translational Medicine*, 11(481). Available at: https://doi.org/10.1126/scitranslmed.aat9223.

Hollingdale, J. *et al.* (2020) 'Autistic spectrum disorder symptoms in children and adolescents with attention-deficit/hyperactivity disorder: a meta-analytical review', *Psychological Medicine*, 50(13), pp. 2240–2253. Available at: https://doi.org/10.1017/S0033291719002368.

Hong, J., Hwang, J. and Lee, J.-H. (2023) 'General psychopathology factor (p-factor) prediction using resting-state functional connectivity and a scanner-generalization neural network', *Journal of Psychiatric Research*, 158, pp. 114–125. Available at: https://doi.org/10.1016/j.jpsychires.2022.12.037.

Hong, S.-J. *et al.* (2021) 'Decomposing complex links between the childhood environment and brain structure in school-aged youth', *Developmental Cognitive Neuroscience*, 48, p. 100919. Available at: https://doi.org/10.1016/j.dcn.2021.100919.

Horn, D. *et al.* (2010) 'Glutamatergic and resting-state functional connectivity correlates of severity in major depression – the role of pregenual anterior cingulate cortex and anterior insula', *Frontiers in Systems Neuroscience*, 4. Available at: https://www.frontiersin.org/articles/10.3389/fnsys.2010.00033.

Hornick, S. and Shetreet, E. (2022) 'Pragmatic inferences: Neuroimaging of ad-hoc implicatures', *Journal of Neurolinguistics*, 64, p. 101090. Available at: https://doi.org/10.1016/j.jneuroling.2022.101090.

Horwath, E. *et al.* (1994) 'What are the public health implications of subclinical depressive symptoms?', *The Psychiatric Quarterly*, 65(4), pp. 323–337. Available at: https://doi.org/10.1007/BF02354307.

van Houdt, C.A. *et al.* (2019) 'Executive function deficits in children born preterm or at low birthweight: a meta-analysis', *Developmental Medicine and Child Neurology*, 61(9), pp. 1015–1024. Available at: https://doi.org/10.1111/dmcn.14213.

van Houdt, C.A. *et al.* (2020) 'Subtypes of behavioral functioning in 8–12 year old very preterm children', *Early Human Development*, 142, p. 104968. Available at: https://doi.org/10.1016/j.earlhumdev.2020.104968.

Huang, X. *et al.* (2015) 'Cognitive impairments associated with corpus callosum infarction: a ten cases study', *International Journal of Clinical and Experimental Medicine*, 8(11), pp. 21991–21998.

Huang-Pollock, C. *et al.* (2017) 'Is poor working memory a transdiagnostic risk factor for psychopathology?', *Journal of abnormal child psychology*, 45(8), pp. 1477–1490. Available at: https://doi.org/10.1007/s10802-016-0219-8.

Huisman, T.A.G.M. (2010) 'Diffusion-weighted and diffusion tensor imaging of the brain, made easy', *Cancer Imaging*, 10(1A), pp. S163–S171. Available at: https://doi.org/10.1102/1470-7330.2010.9023.

Hung, Y. *et al.* (2017) 'Impaired frontal-limbic white matter maturation in children at risk for major depression', *Cerebral Cortex*, 27(9), pp. 4478–4491. Available at: https://doi.org/10.1093/cercor/bhw250.

Huttenlocher, P.R. and Dabholkar, A.S. (1997) 'Regional differences in synaptogenesis in human cerebral cortex', *The Journal of Comparative Neurology*, 387(2), pp. 167–178. Available at: https://doi.org/10.1002/(sici)1096-9861(19971020)387:2<167::aid-cne1>3.0.co;2-z.

Imanipour, S. *et al.* (2021) 'Deficits in working memory and theory of mind may underlie difficulties in social perception of children with ADHD', *Neurology Research International*, 2021, p. 3793750. Available at: https://doi.org/10.1155/2021/3793750.

Inder, T.E. *et al.* (2003) 'Defining the nature of the cerebral abnormalities in the premature infant: a qualitative magnetic resonance imaging study', *The Journal of Pediatrics*, 143(2), pp. 171–179. Available at: https://doi.org/10.1067/S0022-3476(03)00357-3.

Insel, T. *et al.* (2010) 'Research domain criteria (RDoC): toward a new classification framework for research on mental disorders', *The American Journal of Psychiatry*, 167(7), pp. 748–751. Available at: https://doi.org/10.1176/appi.ajp.2010.09091379.

Isaacs, E.B. *et al.* (2004) 'Brain morphometry and IQ measurements in preterm children', *Brain: A Journal of Neurology*, 127(Pt 12), pp. 2595–2607. Available at: https://doi.org/10.1093/brain/awh300.

Jacobs, G.R. *et al.* (2021) 'Integration of brain and behavior measures for identification of data-driven groups cutting across children with ASD, ADHD, or OCD', *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, 46(3), pp. 643–653. Available at: https://doi.org/10.1038/s41386-020-00902-6.

Jain, A.K. (2010) 'Data clustering: 50 years beyond K-means', *Pattern Recognition Letters*, 31(8), pp. 651–666. Available at: https://doi.org/10.1016/j.patrec.2009.09.011.

Jenkinson, M. *et al.* (2002) 'Improved optimization for the robust and accurate linear registration and motion correction of brain images', *NeuroImage*, 17(2), pp. 825–841. Available at: https://doi.org/10.1016/s1053-8119(02)91132-8.

Jenkinson, M. *et al.* (2012) 'FSL', *NeuroImage*, 62(2), pp. 782–790. Available at: https://doi.org/10.1016/j.neuroimage.2011.09.015.

Ji, W. *et al.* (2023) 'Preterm birth associated alterations in brain structure, cognitive functioning and behavior in children from the ABCD dataset', *Psychological Medicine*, pp. 1–10. Available at: https://doi.org/10.1017/S0033291723001757.

Johansson, M. *et al.* (2023) 'Different aspects of visual perception are important for 12-year social functioning depending on gestational age', *Acta Paediatrica*, 112(7), pp. 1537–1547. Available at: https://doi.org/10.1111/apa.16794.

Johnson, S. *et al.* (2004) 'Validation of a parent report measure of cognitive development in very preterm infants', *Developmental Medicine & Child Neurology*, 46(6), pp. 389–397. Available at: https://doi.org/10.1017/S0012162204000635.

Johnson, S. *et al.* (2010a) 'Autism spectrum disorders in extremely preterm children', *The Journal of Pediatrics*, 156(4), pp. 525-531.e2. Available at: https://doi.org/10.1016/j.jpeds.2009.10.041.

Johnson, S. *et al.* (2010b) 'Psychiatric disorders in extremely preterm children: longitudinal finding at age 11 years in the EPICure study', *Journal of the American Academy of Child and Adolescent Psychiatry*, 49(5), pp. 453-463.e1.

Johnson, S. *et al.* (2018) 'Differentiating the preterm phenotype: distinct profiles of cognitive and behavioral development following late and moderately preterm birth', *The Journal of Pediatrics*, 193, pp. 85-92.e1.

Johnson, S. and Marlow, N. (2011) 'Preterm birth and childhood psychiatric disorders', *Pediatric Research*, 69(5 Pt 2), pp. 11R–8R. Available at: https://doi.org/10.1203/PDR.0b013e318212faa0.

Jones, D.T. and Graff-Radford, J. (2021) 'Executive dysfunction and the prefrontal cortex', *Continuum*, 27(6), pp. 1586–1601. Available at: https://doi.org/10.1212/CON.0000000000001009.

Jones, J.S., the CALM Team and Astle, D.E. (2021) 'A transdiagnostic data-driven study of children's behaviour and the functional connectome', *Developmental Cognitive Neuroscience*, 52, p. 101027. Available at: https://doi.org/10.1016/j.dcn.2021.101027.

Joseph, R.M. *et al.* (2017) 'Prevalence and associated features of autism spectrum disorder in extremely low gestational age newborns at age 10 years', *Autism research : official journal of the International Society for Autism Research*, 10(2), pp. 224–232. Available at: https://doi.org/10.1002/aur.1644.

Jung, B. and Kim, H. (2023) 'The validity of transdiagnostic factors in predicting homotypic and heterotypic continuity of psychopathology symptoms over time', *Frontiers in Psychiatry*, 14, p. 1096572. Available at: https://doi.org/10.3389/fpsyt.2023.1096572.

Jurcoane, A. *et al.* (2016) 'White matter alterations of the corticospinal tract in adults born very preterm and/or with very low birth weight', *Human Brain Mapping*, 37(1), pp. 289–299. Available at: https://doi.org/10.1002/hbm.23031.

Kanai, R. *et al.* (2012) 'Brain structure links loneliness to social perception', *Current Biology*, 22(20), pp. 1975–1979. Available at: https://doi.org/10.1016/j.cub.2012.08.045.

Kanel, D. *et al.* (2021) 'Neonatal white matter microstructure and emotional development during the pre-school years in children who were born very preterm', *eNeuro*, 8(5), p. ENEURO.0546-20.2021. Available at: https://doi.org/10.1523/ENEURO.0546-20.2021.

Kanel, D. *et al.* (2022) 'Neonatal amygdala resting-state functional connectivity and socio-emotional development in very preterm children', *Brain Communications*, 4(1), p. fcac009. Available at: https://doi.org/10.1093/braincomms/fcac009.

Karcher, N.R. *et al.* (2018) 'Assessment of the Prodromal Questionnaire–Brief Child Version for measurement of self-reported psychoticlike experiences in childhood', *JAMA Psychiatry*, 75(8), pp. 853–861. Available at: https://doi.org/10.1001/jamapsychiatry.2018.1334.

Karcher, N.R. *et al.* (2021) 'Associations between resting state functional connectivity and a hierarchical dimensional structure of psychopathology in middle childhood', *Biological psychiatry. Cognitive neuroscience and neuroimaging*, 6(5), pp. 508–517. Available at: https://doi.org/10.1016/j.bpsc.2020.09.008.

Karolis, V.R. *et al.* (2016) 'Reinforcement of the brain's rich-club architecture following early neurodevelopmental disruption caused by very preterm birth', *Cerebral Cortex*, 26(3), pp. 1322–1335. Available at: https://doi.org/10.1093/cercor/bhv305.

Karolis, V.R. *et al.* (2017) 'Volumetric grey matter alterations in adolescents and adults born very preterm suggest accelerated brain maturation', *NeuroImage*, 163, pp. 379–389. Available at: https://doi.org/10.1016/j.neuroimage.2017.09.039.

Karolis, V.R. *et al.* (2023) 'Maturational networks of human fetal brain activity reveal emerging connectivity patterns prior to ex-utero exposure', *Communications Biology*, 6. Available at: https://doi.org/10.1038/s42003-023-04969-x.

Kesler, S.R. *et al.* (2006) 'Increased temporal lobe gyrification in preterm children', *Neuropsychologia*, 44(3), pp. 445–453. Available at: https://doi.org/10.1016/j.neuropsychologia.2005.05.015.

Kessler, R.C. *et al.* (2005) 'Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication', *Archives of General Psychiatry*, 62(6), pp. 593–602. Available at: https://doi.org/10.1001/archpsyc.62.6.593.

Khan, A. *et al.* (2019) 'Environmental pollution is associated with increased risk of psychiatric disorders in the US and Denmark', *PLOS Biology*, 17(8), p. e3000353. Available at: https://doi.org/10.1371/journal.pbio.3000353.

Kim, S.H. *et al.* (2016) 'Predictive validity of the Modified Checklist for Autism in Toddlers (M-CHAT) born very preterm', *The Journal of Pediatrics*, 178, pp. 101-107.e2. Available at: https://doi.org/10.1016/j.jpeds.2016.07.052.

Kim-Cohen, J. *et al.* (2006) 'MAOA , maltreatment, and gene–environment interaction predicting children's mental health: new evidence and a meta-analysis', *Molecular Psychiatry*, 11(10), p. 903. Available at: https://doi.org/10.1038/sj.mp.4001851.

Klein, A. *et al.* (2017) 'Mindboggling morphometry of human brains', *PLOS Computational Biology*, 13(2), p. e1005350. Available at: https://doi.org/10.1371/journal.pcbi.1005350.

Kleinman, J.M. *et al.* (2008) 'The Modified Checklist for Autism in Toddlers: A follow-up study investigating the early detection of autism spectrum disorders', *Journal of autism and developmental disorders*, 38(5), pp. 827–839. Available at: https://doi.org/10.1007/s10803-007-0450-9.

Knight, L.K. *et al.* (2019) 'Convergent neural correlates of empathy and anxiety during socioemotional processing', *Frontiers in Human Neuroscience*, 13. Available at: https://www.frontiersin.org/articles/10.3389/fnhum.2019.00094.

Kong, X. *et al.* (2014) 'Individual differences in impulsivity predict head motion during magnetic resonance imaging', *PLoS ONE*, 9(8), p. e104989. Available at: https://doi.org/10.1371/journal.pone.0104989.

Kotov, R. *et al.* (2017) 'The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies', *Journal of Abnormal Psychology*, 126(4), pp. 454–477. Available at: https://doi.org/10.1037/abn0000258.

Kragel, P.A. and LaBar, K.S. (2016) 'Decoding the nature of emotion in the brain', *Trends in Cognitive Sciences*, 20(6), pp. 444–455. Available at: https://doi.org/10.1016/j.tics.2016.03.011.

Kroll, J. *et al.* (2017) 'Real-life impact of executive function impairments in adults who were born very preterm', *Journal of the International Neuropsychological Society: JINS*, 23(5), pp. 381–389. Available at: https://doi.org/10.1017/S1355617717000169.

Kuban, K.C.K. *et al.* (2009) 'Positive screening on the Modified Checklist for Autism in Toddlers (M-CHAT) in extremely low gestational age newborns', *The Journal of Pediatrics*, 154(4), pp. 535-540.e1. Available at: https://doi.org/10.1016/j.jpeds.2008.10.011.

Kushki, A. *et al.* (2019) 'Examining overlap and homogeneity in ASD, ADHD, and OCD: a data-driven, diagnosis-agnostic approach', *Translational Psychiatry*, 9(1), pp. 1–11. Available at: https://doi.org/10.1038/s41398-019-0631-2.

Kvanta, H. *et al.* (2023) 'Extreme prematurity and perinatal risk factors related to extremely preterm birth are associated with complex patterns of regional brain volume alterations at 10 years of age: a voxel-based morphometry study', *Frontiers in Neurology*, 14, p. 1148781. Available at: https://doi.org/10.3389/fneur.2023.1148781.

Lahey, B.B. *et al.* (2012) 'Is there a general factor of prevalent psychopathology during adulthood?', *Journal of Abnormal Psychology*, 121(4), pp. 971–977. Available at: https://doi.org/10.1037/a0028355.

Lamar, M. *et al.* (2013) 'Prefrontal vulnerabilities and whole brain connectivity in aging and depression', *Neuropsychologia*, 51(8), pp. 1463–1470. Available at: https://doi.org/10.1016/j.neuropsychologia.2013.05.004.

Lanczos, C. (1964) 'Evaluation of Noisy Data', *Journal of the Society for Industrial and Applied Mathematics Series B Numerical Analysis*, 1(1), pp. 76–85. Available at: https://doi.org/10.1137/0701007.

Lautarescu, A. *et al.* (2021) 'Exploring the relationship between maternal prenatal stress and brain structure in premature neonates', *PloS One*, 16(4), p. e0250413. Available at: https://doi.org/10.1371/journal.pone.0250413.

Lax, I.D. *et al.* (2013) 'Neuroanatomical consequences of very preterm birth in middle childhood', *Brain Structure and Function*, 218(2), pp. 575–585. Available at: https://doi.org/10.1007/s00429-012-0417-2.

Lean, R.E. *et al.* (2017) 'Attention and regional gray matter development in very preterm children at age 12 years', *Journal of the International Neuropsychological Society: JINS*, 23(7), pp. 539–550. Available at: https://doi.org/10.1017/S1355617717000388.

Lean, R.E. *et al.* (2020) 'Maternal and family factors differentiate profiles of psychiatric impairments in very preterm children at age 5-years', *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 61(2), pp. 157–166. Available at: https://doi.org/10.1111/jcpp.13116.

Leathem, L.D. *et al.* (2021) 'Socioemotional mechanisms of loneliness in subclinical psychosis', *Schizophrenia Research*, 238, pp. 145–151. Available at: https://doi.org/10.1016/j.schres.2021.10.002.

Lee, H.J. *et al.* (2021) 'The cingulum in very preterm infants relates to language and social-emotional impairment at 2 years of term-equivalent age', *NeuroImage. Clinical*, 29, p. 102528. Available at: https://doi.org/10.1016/j.nicl.2020.102528.

Lee, Y.A. (2017) 'White Matter Injury of Prematurity: Its Mechanisms and Clinical Features', *Journal of Pathology and Translational Medicine*, 51(5), pp. 449–455. Available at: https://doi.org/10.4132/jptm.2017.07.25.

Lefèvre, J. *et al.* (2016) 'Are developmental trajectories of cortical folding comparable between cross-sectional datasets of fetuses and preterm newborns?', *Cerebral Cortex*, 26(7), pp. 3023–3035. Available at: https://doi.org/10.1093/cercor/bhv123.

Leibovitz, Z., Lerman-Sagie, T. and Haddad, L. (2022) 'Fetal brain development: Regulating processes and related malformations', *Life*, 12(6). Available at: https://doi.org/10.3390/life12060809.

Lemola, S. *et al.* (2017) 'Effects of gestational age on brain volume and cognitive functions in generally healthy very preterm born children during school-age: A voxel-based morphometry study', *PloS One*, 12(8), p. e0183519. Available at: https://doi.org/10.1371/journal.pone.0183519.

Leoni, M. *et al.* (2023) 'Exploring cognitive, behavioral and autistic trait network topology in very preterm and term-born children', *Frontiers in Psychology*, 14. Available at: https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1119196.

Levine, T.A. *et al.* (2015) 'Early childhood neurodevelopment after intrauterine growth restriction: a systematic review', *Pediatrics*, 135(1), pp. 126–141. Available at: https://doi.org/10.1542/peds.2014-1143.

Li, C. *et al.* (2021) 'Transdiagnostic time-varying dysconnectivity across major psychiatric disorders', *Human Brain Mapping*, 42(4), pp. 1182–1196. Available at: https://doi.org/10.1002/hbm.25285.

Li, G. *et al.* (2013) 'Mapping region-specific longitudinal cortical surface expansion from birth to 2 years of age', *Cerebral Cortex*, 23(11), pp. 2724–2733. Available at: https://doi.org/10.1093/cercor/bhs265.

Limperopoulos, C. *et al.* (2007) 'Does cerebellar injury in premature infants contribute to the high prevalence of long-term cognitive, learning, and behavioral disability in survivors?', *Pediatrics*, 120(3), pp. 584–593. Available at: https://doi.org/10.1542/peds.2007-1041.

Limperopoulos, C. *et al.* (2008) 'Positive screening for autism in ex-preterm infants: prevalence and risk factors', *Pediatrics*, 121(4), pp. 758–765. Available at: https://doi.org/10.1542/peds.2007-2158.

Linden, A. (2020) 'DIAGSAMPSI: Stata module for computing sample size for a single diagnostic test with a binary outcome.'

Liu, S. and Shang, X. (2018) 'Hierarchical similarity network fusion for discovering cancer subtypes', in F. Zhang et al. (eds) *Bioinformatics Research and Applications*. Cham: Springer International Publishing (Lecture Notes in Computer Science), pp. 125–136. Available at: https://doi.org/10.1007/978-3-319-94968-0_11.

Liu, Z. *et al.* (2010) 'Decreased regional homogeneity in insula and cerebellum: a resting-state fMRI study in patients with major depression and subjects at high risk for major depression', *Psychiatry Research*, 182(3), pp. 211–215. Available at: https://doi.org/10.1016/j.pscychresns.2010.03.004.

Loitfelder, M. *et al.* (2015) 'Functional connectivity changes and executive and social problems in neurofibromatosis type I', *Brain Connectivity*, 5(5), pp. 312–320. Available at: https://doi.org/10.1089/brain.2014.0334.

van Loo, H.M. and Romeijn, J.-W. (2015) 'Psychiatric comorbidity: fact or artifact?', *Theoretical Medicine and Bioethics*, 36(1), pp. 41–60. Available at: https://doi.org/10.1007/s11017-015-9321-0.

López-Vicente, M. *et al.* (2021) 'Developmental changes in dynamic functional connectivity from childhood into adolescence', *Frontiers in Systems Neuroscience*, 15, p. 724805. Available at: https://doi.org/10.3389/fnsys.2021.724805.

Lordier, L. *et al.* (2019) 'Music in premature infants enhances high-level cognitive brain networks', *Proceedings of the National Academy of Sciences of the United States of America*, 116(24), pp. 12103–12108. Available at: https://doi.org/10.1073/pnas.1817536116.

Lovato, I. *et al.* (2022) 'Early childhood temperamental trajectories following very preterm birth and their association with parenting style', *Children*, 9(4), p. 508. Available at: https://doi.org/10.3390/children9040508.

Ludyga, S., Ishihara, T. and Kamijo, K. (2022) 'The nervous system as a pathway for exercise to improve social cognition', *Exercise and Sport Sciences Reviews*, 50(4), pp. 203–212. Available at: https://doi.org/10.1249/JES.0000000000000300.

Luo, Y. *et al.* (2014) 'Emotion perception and executive control interact in the salience network during emotionally charged working memory processing', *Human Brain Mapping*, 35(11), pp. 5606–5616. Available at: https://doi.org/10.1002/hbm.22573.

Luu, T.M. *et al.* (2009) 'Trajectories of receptive language development from 3 to 12 years of age for very preterm children', *Pediatrics*, 124(1), pp. 333–341. Available at: https://doi.org/10.1542/peds.2008-2587.

Luu, T.M. *et al.* (2011) 'Evidence for catch-up in cognition and receptive vocabulary among adolescents born very preterm', *Pediatrics*, 128(2), pp. 313–322. Available at: https://doi.org/10.1542/peds.2010-2655.

von Luxburg, U. (2007) 'A tutorial on spectral clustering', *Statistics and Computing*, 17(4), pp. 395–416. Available at: https://doi.org/10.1007/s11222-007-9033-z.

Luyster, R.J. *et al.* (2011) 'The Modified Checklist for Autism in Toddlers in extremely low gestational age newborns: individual items associated with motor, cognitive, vision and hearing limitations', *Paediatric and Perinatal Epidemiology*, 25(4), pp. 366–376. Available at: https://doi.org/10.1111/j.1365-3016.2010.01187.x.

MacQueen, J. (1967) 'Some methods for classification and analysis of multivariate observations', in. Available at: https://www.semanticscholar.org/paper/Some-methods-for-classification-and-analysis-of-MacQueen/ac8ab51a86f1a9ae74dd0e4576d1a019f5e654ed.

Maechler, M. *et al.* (2021) 'cluster: Cluster Analysis Basics and Extensions. R package version 2.1.2 — For new features, see the "Changelog" file (in the package source), https://CRAN.R-project.org/package=cluster.'

Maitre, N.L. *et al.* (2017) 'The dual nature of early-life experience on somatosensory processing in the human infant brain', *Current biology: CB*, 27(7), pp. 1048–1054. Available at: https://doi.org/10.1016/j.cub.2017.02.036.

Mangold, S.A. and M Das, J. (2023) 'Neuroanatomy, reticular formation', in *StatPearls*. Treasure Island (FL): StatPearls Publishing. Available at: http://www.ncbi.nlm.nih.gov/books/NBK556102/.

Mareva, S. *et al.* (2023) 'Transdiagnostic profiles of behaviour and communication relate to academic and socioemotional functioning and neural white matter organisation', *Journal of Child Psychology and Psychiatry*, 64(2), pp. 217–233. Available at: https://doi.org/10.1111/jcpp.13685.

Mareva, S., CALM team and Holmes, J. (2019) 'Transdiagnostic associations across communication, cognitive, and behavioural problems in a developmentally at-risk population: a network approach', *BMC pediatrics*, 19(1), p. 452. Available at: https://doi.org/10.1186/s12887-019-1818-7.

Markello, R.D. *et al.* (2020) 'Integrated morphometric, molecular, and clinical characterization of Parkinson's disease pathology', *bioRxiv*, p. 2020.03.05.979526. Available at: https://doi.org/10.1101/2020.03.05.979526.

Markello, R.D. *et al.* (2021) 'Multimodal phenotypic axes of Parkinson's disease', *npj Parkinson's Disease*, 7(1), pp. 1–12. Available at: https://doi.org/10.1038/s41531-020-00144-9.

Marsh, S., Dobson, R. and Maddison, R. (2020) 'The relationship between household chaos and child, parent, and family outcomes: a systematic scoping review', *BMC Public Health*, 20(1), p. 513. Available at: https://doi.org/10.1186/s12889-020-08587-8.

Marusak, H.A., Carré, J.M. and Thomason, M.E. (2013) 'The stimuli drive the response: An fMRI study of youth processing adult or child emotional face stimuli', *NeuroImage*, 83, pp. 679–689. Available at: https://doi.org/10.1016/j.neuroimage.2013.07.002.

Mastrogiacomo, S. *et al.* (2019) 'Magnetic Resonance Imaging of hard tissues and hard tissue engineered bio-substitutes', *Molecular Imaging and Biology*, 21(6), pp. 1003–1019. Available at: https://doi.org/10.1007/s11307-019-01345-2.

Mathai, J., Anderson, P. and Bourne, A. (2004) 'Comparing psychiatric diagnoses generated by the strengths and difficulties questionnaire with diagnoses made by clinicians', *Australian & New*

*Zealand Journal of Psychiatry*, 38(8), pp. 639–643. Available at: https://doi.org/10.1080/j.1440-1614.2004.01428.x.

Mathewson, K.J. *et al.* (2017) 'Mental health of extremely low birth weight survivors: A systematic review and meta-analysis', *Psychological Bulletin*, 143(4), pp. 347–383. Available at: https://doi.org/10.1037/bul0000091.

Matsudaira, I. *et al.* (2016) 'Parental praise correlates with posterior insular cortex gray matter volume in children and adolescents', *PLOS ONE*, 11(4), p. e0154220. Available at: https://doi.org/10.1371/journal.pone.0154220.

Mazzone, P. *et al.* (2005) 'Implantation of human pedunculopontine nucleus: a safe and clinically relevant target in Parkinson's disease', *Neuroreport*, 16(17), pp. 1877–1881. Available at: https://doi.org/10.1097/01.wnr.0000187629.38010.12.

McGrath, J.J. *et al.* (2014) 'The association between family history of mental disorders and general cognitive ability', *Translational Psychiatry*, 4(7), p. e412. Available at: https://doi.org/10.1038/tp.2014.60.

McTeague, L.M. *et al.* (2020) 'Identification of common neural circuit disruptions in emotional processing across psychiatric disorders', *American Journal of Psychiatry*, 177(5), pp. 411–421. Available at: https://doi.org/10.1176/appi.ajp.2019.18111271.

Mendonça, M., Bilgin, A. and Wolke, D. (2019) 'Association of preterm birth and low birth weight with romantic partnership, sexual intercourse, and parenthood in adulthood: a systematic review and meta-analysis', *JAMA network open*, 2(7), p. e196961. Available at: https://doi.org/10.1001/jamanetworkopen.2019.6961.

Menon, V. and Uddin, L.Q. (2010) 'Saliency, switching, attention and control: a network model of insula function', *Brain Structure & Function*, 214(5–6), pp. 655–667. Available at: https://doi.org/10.1007/s00429-010-0262-0.

Ment, L.R. *et al.* (2003) 'Change in cognitive function over time in very low-birth-weight infants', *JAMA*, 289(6), pp. 705–711. Available at: https://doi.org/10.1001/jama.289.6.705.

Milgrom, J. *et al.* (2010) 'Early sensitivity training for parents of preterm infants: impact on the developing brain', *Pediatric Research*, 67(3), pp. 330–335. Available at: https://doi.org/10.1203/PDR.0b013e3181cb8e2f.

Montagna, A. and Nosarti, C. (2016) 'Socio-emotional development following very preterm birth: pathways to psychopathology', *Frontiers in Psychology*, 7, p. 80. Available at: https://doi.org/10.3389/fpsyg.2016.00080.

Montagne, B. *et al.* (2007) 'The Emotion Recognition Task: a paradigm to measure the perception of facial emotional expressions at different intensities', *Perceptual and Motor Skills*, 104(2), pp. 589–598. Available at: https://doi.org/10.2466/pms.104.2.589-598.

Monti, S. *et al.* (2003) 'Consensus clustering: A resampling-based method for class discovery and visualization of gene expression microarray data', *Machine Learning*, 52(1), pp. 91–118. Available at: https://doi.org/10.1023/A:1023949509487.

Moore, J.W. *et al.* (2023) 'Gradient organisation of functional connectivity within resting state networks is present from 25 weeks gestation in the human fetal brain', *eLife*, 12. Available at: https://doi.org/10.7554/eLife.90536.

Moore, T., Hennessy, E.M., *et al.* (2012) 'Neurological and developmental outcome in extremely preterm children born in England in 1995 and 2006: the EPICure studies', *BMJ (Clinical research ed.)*, 345, p. e7961. Available at: https://doi.org/10.1136/bmj.e7961.

Moore, T., Johnson, S., *et al.* (2012) 'Screening for autism in extremely preterm infants: problems in interpretation', *Developmental Medicine and Child Neurology*, 54(6), pp. 514–520. Available at: https://doi.org/10.1111/j.1469-8749.2012.04265.x.

Mooshagian, E. (2008) 'Anatomy of the corpus callosum reveals its function', *The Journal of Neuroscience*, 28(7), pp. 1535–1536. Available at: https://doi.org/10.1523/JNEUROSCI.5426-07.2008.

Morris, S.E. *et al.* (2022) 'Revisiting the seven pillars of RDoC', *BMC medicine*, 20(1), p. 220. Available at: https://doi.org/10.1186/s12916-022-02414-0.

Mossad, S.I. *et al.* (2022) 'Very preterm brain at rest: longitudinal social-cognitive network connectivity during childhood', *Social Cognitive and Affective Neuroscience*, 17(4), pp. 377–386. Available at: https://doi.org/10.1093/scan/nsab110.

Mouka, V. *et al.* (2019) 'Functional and structural connectivity of the brain in very preterm babies: relationship with gestational age and body and brain growth', *Pediatric Radiology*, 49(8), pp. 1078–1084. Available at: https://doi.org/10.1007/s00247-019-04412-6.

Mowinckel, A.M. and Vidal-Piñeiro, D. (2020) 'Visualization of brain statistics with R packages ggseg and ggseg3d', *Advances in Methods and Practices in Psychological Science*, 3(4), pp. 466–483. Available at: https://doi.org/10.1177/2515245920928009.

Mueller, M. *et al.* (2022) 'Amygdala subnuclei volumes, functional connectivity, and social-emotional outcomes in children born very preterm', *Cerebral Cortex Communications*, 3(3), p. tgac028. Available at: https://doi.org/10.1093/texcom/tgac028.

Mukhopadhyay, S. *et al.* (2022) 'Preterm infants at low risk for early-onset sepsis differ in early fecal microbiome assembly', *Gut Microbes*, 14(1), p. 2154091. Available at: https://doi.org/10.1080/19490976.2022.2154091.

Mullen, K.M. *et al.* (2011) 'Preterm birth results in alterations in neural connectivity at age 16 years', *NeuroImage*, 54(4), pp. 2563–2570. Available at: https://doi.org/10.1016/j.neuroimage.2010.11.019.

Mürner-Lavanchy, I. *et al.* (2014) 'Delay of cortical thinning in very preterm born children', *Early Human Development*, 90(9), pp. 443–450. Available at: https://doi.org/10.1016/j.earlhumdev.2014.05.013.

Myers, E.H. *et al.* (2010) 'Functional connectivity to a right hemisphere language center in prematurely born adolescents', *NeuroImage*, 51(4), pp. 1445–1452. Available at: https://doi.org/10.1016/j.neuroimage.2010.03.049.

Nagy, Z. *et al.* (2003) 'Preterm children have disturbances of white matter at 11 years of age as shown by diffusion tensor imaging', *Pediatric Research*, 54(5), pp. 672–679. Available at: https://doi.org/10.1203/01.PDR.0000084083.71422.16.

Nauta, M.H. *et al.* (2004) 'A parent-report measure of children's anxiety: psychometric properties and comparison with child-report in a clinic and normal sample', *Behaviour Research and Therapy*, 42(7), pp. 813–839. Available at: https://doi.org/10.1016/S0005-7967(03)00200-6.

Neisser, U. *et al.* (1996) 'Intelligence: Knowns and unknowns', *American Psychologist*, 51(2), pp. 77–101. Available at: https://doi.org/10.1037/0003-066X.51.2.77.

Nelson, C.A., Thomas, K.M. and de Haan, M. (2006) 'Neural bases of cognitive development', in *Handbook of child psychology: Cognition, perception, and language, Vol. 2, 6th ed.* Hoboken, NJ, US: John Wiley & Sons, Inc., pp. 3–57.

Neubauer, A.-P., Voss, W. and Kattner, E. (2008) 'Outcome of extremely low birth weight survivors at school age: the influence of perinatal parameters on neurodevelopment', *European Journal of Pediatrics*, 167(1), pp. 87–95. Available at: https://doi.org/10.1007/s00431-007-0435-x.

Ng, A., Jordan, M. and Weiss, Y. (2001) 'On Spectral Clustering: Analysis and an algorithm', in *Advances in Neural Information Processing Systems*. MIT Press. Available at: https://proceedings.neurips.cc/paper_files/paper/2001/hash/801272ee79cfde7fa5960571fee36b9b-Abstract.html.

Ng, C.S.M. and Ng, S.S.L. (2022) 'Impact of the COVID-19 pandemic on children's mental health: A systematic review', *Frontiers in Psychiatry*, 13, p. 975936. Available at: https://doi.org/10.3389/fpsyt.2022.975936.

Nguyen, T.-N.-N. *et al.* (2018) 'Language trajectories of children born very preterm and full term from early to late childhood', *The Journal of Pediatrics*, 202, pp. 86-91.e1. Available at: https://doi.org/10.1016/j.jpeds.2018.06.036.

Nichols, T.E. and Holmes, A.P. (2002) 'Nonparametric permutation tests for functional neuroimaging: a primer with examples', *Human Brain Mapping*, 15(1), pp. 1–25. Available at: https://doi.org/10.1002/hbm.1058.

Nikolopoulou, E. *et al.* (2017) 'Neural tube closure: cellular, molecular and biomechanical mechanisms', *Development*, 144(4), p. 552. Available at: https://doi.org/10.1242/dev.145904.

Nosarti, C. *et al.* (2002) 'Adolescents who were born very preterm have decreased brain volumes', *Brain*, 125(7), pp. 1616–1623. Available at: https://doi.org/10.1093/brain/awf157.

Nosarti, C. *et al.* (2005) 'Hyperactivity in adolescents born very preterm is associated with decreased caudate volume', *Biological Psychiatry*, 57(6), pp. 661–666. Available at: https://doi.org/10.1016/j.biopsych.2004.12.003.

Nosarti, C. *et al.* (2006) 'Altered functional neuroanatomy of response inhibition in adolescent males who were born very preterm', *Developmental Medicine and Child Neurology*, 48(4), pp. 265–271. Available at: https://doi.org/10.1017/S0012162206000582.

Nosarti, C. *et al.* (2007) 'Impaired executive functioning in young adults born very preterm', *Journal of the International Neuropsychological Society: JINS*, 13(4), pp. 571–581. Available at: https://doi.org/10.1017/S1355617707070725.

Nosarti, C. *et al.* (2008) 'Grey and white matter distribution in very preterm adolescents mediates neurodevelopmental outcome', *Brain: A Journal of Neurology*, 131(Pt 1), pp. 205–217. Available at: https://doi.org/10.1093/brain/awm282.

Nosarti, C. *et al.* (2009) 'Neural substrates of letter fluency processing in young adults who were born very preterm: Alterations in frontal and striatal regions', *NeuroImage*, 47(4), pp. 1904–1913. Available at: https://doi.org/10.1016/j.neuroimage.2009.04.041.

Nosarti, C. *et al.* (2012) 'Preterm birth and psychiatric disorders in young adult life', *Archives of General Psychiatry*, 69(6), pp. E1-8. Available at: https://doi.org/10.1001/archgenpsychiatry.2011.1374.

Nosarti, C. *et al.* (2014) 'Preterm birth and structural brain alterations in early adulthood', *NeuroImage: Clinical*, 6, pp. 180–191. Available at: https://doi.org/10.1016/j.nicl.2014.08.005.

Ochsner, K.N. (2008) 'The social-emotional processing stream: Five core constructs and their translational potential for schizophrenia and beyond', *Biological psychiatry*, 64(1), pp. 48–61. Available at: https://doi.org/10.1016/j.biopsych.2008.04.024.

Ogawa, S. *et al.* (1990) 'Brain magnetic resonance imaging with contrast dependent on blood oxygenation.', *Proceedings of the National Academy of Sciences*, 87(24), pp. 9868–9872. Available at: https://doi.org/10.1073/pnas.87.24.9868.

Olsen, A. *et al.* (2018) 'Preterm birth leads to hyper-reactive cognitive control processing and poor white matter organization in adulthood', *NeuroImage*, 167, pp. 419–428. Available at: https://doi.org/10.1016/j.neuroimage.2017.11.055.

Opitz, P.C. *et al.* (2014) 'Fluid cognitive ability is a resource for successful emotion regulation in older and younger adults', *Frontiers in Psychology*, 5, p. 609. Available at: https://doi.org/10.3389/fpsyg.2014.00609.

Otto, C. *et al.* (2021) 'Mental health and well-being from childhood to adulthood: design, methods and results of the 11-year follow-up of the BELLA study', *European Child & Adolescent Psychiatry*, 30(10), pp. 1559–1577. Available at: https://doi.org/10.1007/s00787-020-01630-4.

Padilla, N. *et al.* (2017) 'Poor brain growth in extremely preterm neonates long before the onset of autism spectrum disorder symptoms', *Cerebral Cortex*, 27(2), pp. 1245–1252. Available at: https://doi.org/10.1093/cercor/bhv300.

Pannekoek, J.N. *et al.* (2014) 'Aberrant resting-state functional connectivity in limbic and salience networks in treatment--naïve clinically depressed adolescents', *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 55(12), pp. 1317–1327. Available at: https://doi.org/10.1111/jcpp.12266.

Papini, C. *et al.* (2016) 'Altered resting-state functional connectivity in emotion-processing brain regions in adults who were born very preterm', *Psychological Medicine*, 46(14), pp. 3025–3039. Available at: https://doi.org/10.1017/S0033291716001604.

Pardoe, H.R., Kucharsky Hiess, R. and Kuzniecky, R. (2016) 'Motion and morphometry in clinical and nonclinical populations', *NeuroImage*, 135, pp. 177–185. Available at: https://doi.org/10.1016/j.neuroimage.2016.05.005.

Pariante, C.M. (2016) 'Neuroscience, mental health and the immune system: overcoming the brain-mind-body trichotomy', *Epidemiology and Psychiatric Sciences*, 25(2), pp. 101–105. Available at: https://doi.org/10.1017/S204579601500089X.

Pecheva, D. *et al.* (2017) 'A tract-specific approach to assessing white matter in preterm infants', *Neuroimage*, 157, pp. 675–694. Available at: https://doi.org/10.1016/j.neuroimage.2017.04.057.

Pernet, C. and Poline, J.-B. (2015) 'Improving functional magnetic resonance imaging reproducibility', *GigaScience*, 4, p. 15. Available at: https://doi.org/10.1186/s13742-015-0055-8.

Perrot, M., Rivière, D. and Mangin, J. (2011) 'Cortical sulci recognition and spatial normalization', *Medical image analysis*, 15(4). Available at: https://doi.org/10.1016/j.media.2011.02.008.

Peters, E. *et al.* (2004) 'Measuring delusional ideation: the 21-item Peters et al. Delusions Inventory (PDI)', *Schizophrenia Bulletin*, 30(4), pp. 1005–1022. Available at: https://doi.org/10.1093/oxfordjournals.schbul.a007116.

Peterson, B.S. *et al.* (2000) 'Regional brain volume abnormalities and long-term cognitive outcome in preterm infants', *JAMA*, 284(15), pp. 1939–1947. Available at: https://doi.org/10.1001/jama.284.15.1939.

Petrovic, P. *et al.* (2016) 'Significant grey matter changes in a region of the orbitofrontal cortex in healthy participants predicts emotional dysregulation', *Social Cognitive and Affective Neuroscience*, 11(7), pp. 1041–1049. Available at: https://doi.org/10.1093/scan/nsv072.

Pierrat, V. *et al.* (2021) 'Neurodevelopmental outcomes at age 5 among children born preterm: EPIPAGE-2 cohort study', *The BMJ*, 373, p. n741. Available at: https://doi.org/10.1136/bmj.n741.

Plaisier, A. *et al.* (2015) 'Serial cranial ultrasonography or early MRI for detecting preterm brain injury?', *Archives of Disease in Childhood. Fetal and Neonatal Edition*, 100(4), pp. F293-300. Available at: https://doi.org/10.1136/archdischild-2014-306129.

Platt, M.J. *et al.* (2007) 'Trends in cerebral palsy among infants of very low birthweight (<1500 g) or born prematurely (<32 weeks) in 16 European centres: a database study', *Lancet*, 369(9555), pp. 43–50. Available at: https://doi.org/10.1016/S0140-6736(07)60030-0.

Podsakoff, P.M. *et al.* (2003) 'Common method biases in behavioral research: A critical review of the literature and recommended remedies', *The Journal of Applied Psychology*, 88(5), pp. 879–903. Available at: https://doi.org/10.1037/0021-9010.88.5.879.

Poehlmann, J. *et al.* (2015) 'Risk and resilience in preterm children at age 6', *Development and psychopathology*, 27(3), pp. 843–858. Available at: https://doi.org/10.1017/S095457941400087X.

Poole, V.N. *et al.* (2016) 'Intrinsic functional connectivity predicts individual differences in distractibility', *Neuropsychologia*, 86, pp. 176–182. Available at: https://doi.org/10.1016/j.neuropsychologia.2016.04.023.

Power, J.D. *et al.* (2012) 'Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion', *NeuroImage*, 59(3), pp. 2142–2154. Available at: https://doi.org/10.1016/j.neuroimage.2011.10.018.

Power, J.D. *et al.* (2013) 'Evidence for hubs in human functional brain networks', *Neuron*, 79(4), pp. 798–813. Available at: https://doi.org/10.1016/j.neuron.2013.07.035.

Power, J.D. *et al.* (2014) 'Methods to detect, characterize, and remove motion artifact in resting state fMRI', *NeuroImage*, 84, pp. 320–341. Available at: https://doi.org/10.1016/j.neuroimage.2013.08.048.

Purves, D. *et al.* (2001) 'Formation of the major brain subdivisions', in *neuroscience. 2nd edition*. sinauer associates. Available at: https://www.ncbi.nlm.nih.gov/books/NBK10954/.

Quirmbach, F. and Limanowski, J. (2022) 'A Crucial Role of the Frontal Operculum in Task-Set Dependent Visuomotor Performance Monitoring', *eNeuro*, 9(2). Available at: https://doi.org/10.1523/ENEURO.0524-21.2021.

Raita, Y. *et al.* (2020) 'Integrated-omics endotyping of infants with rhinovirus bronchiolitis and risk of childhood asthma', *Journal of Allergy and Clinical Immunology*, 0(0). Available at: https://doi.org/10.1016/j.jaci.2020.11.002.

Ramphal, B. *et al.* (2020) 'Brain connectivity and socioeconomic status at birth and externalizing symptoms at age 2 years', *Developmental Cognitive Neuroscience*, 45, p. 100811. Available at: https://doi.org/10.1016/j.dcn.2020.100811.

Ranger, M. *et al.* (2013) 'Neonatal pain-related stress predicts cortical thickness at age 7 years in children born very preterm', *PLoS ONE*, 8(10), p. e76702. Available at: https://doi.org/10.1371/journal.pone.0076702.

Ratle, F. *et al.* (2008) 'Advanced clustering methods for mining chemical databases in forensic science', *Chemometrics and Intelligent Laboratory Systems*, 90(2), pp. 123–131. Available at: https://doi.org/10.1016/j.chemolab.2007.09.001.

Remes, O. *et al.* (2019) 'Association between area deprivation and major depressive disorder in British men and women: a cohort study', *BMJ Open*, 9(11), p. e027530. Available at: https://doi.org/10.1136/bmjopen-2018-027530.

Rempel-Clower, N.L. (2007) 'Role of orbitofrontal cortex connections in emotion', *Annals of the New York Academy of Sciences*, 1121(1), pp. 72–86. Available at: https://doi.org/10.1196/annals.1401.026.

Rieffe, C., Ketelaar, L. and Wiefferink, C.H. (2010) 'Assessing empathy in young children: Construction and validation of an Empathy Questionnaire (EmQue)', *Personality and Individual Differences*, 49(5), pp. 362–367. Available at: https://doi.org/10.1016/j.paid.2010.03.046.

Robins, D.L. *et al.* (2001) 'The Modified Checklist for Autism in Toddlers: an initial study investigating the early detection of autism and pervasive developmental disorders', *Journal of Autism and Developmental Disorders*, 31(2), pp. 131–144. Available at: https://doi.org/10.1023/a:1010738829569.

Rodier, P.M. *et al.* (1996) 'Embryological origin for autism: developmental anomalies of the cranial nerve motor nuclei', *The Journal of Comparative Neurology*, 370(2), pp. 247–261. Available at: https://doi.org/10.1002/(SICI)1096-9861(19960624)370:2<247::AID-CNE8>3.0.CO;2-2.

Rodier, P.M. *et al.* (1997) 'Linking etiologies in humans and animal models: studies of autism', *Reproductive Toxicology (Elmsford, N.Y.)*, 11(2–3), pp. 417–422. Available at: https://doi.org/10.1016/s0890-6238(97)80001-u.

Rodier, P.M. (2002) 'Converging evidence for brain stem injury in autism', *Development and Psychopathology*, 14(3), pp. 537–557. Available at: https://doi.org/10.1017/s0954579402003085.

Rodriguez, M.Z. *et al.* (2019) 'Clustering algorithms: A comparative approach', *PLoS ONE*, 14(1), p. e0210236. Available at: https://doi.org/10.1371/journal.pone.0210236.

Rogers, C.E. *et al.* (2012) 'Regional cerebral development at term relates to school-age social-emotional development in very preterm children', *Journal of the American Academy of Child and Adolescent Psychiatry*, 51(2), pp. 181–191. Available at: https://doi.org/10.1016/j.jaac.2011.11.009.

Rogers, C.E. *et al.* (2014) 'Altered gray matter volume and school age anxiety in children born late preterm', *The Journal of Pediatrics*, 165(5), pp. 928–935. Available at: https://doi.org/10.1016/j.jpeds.2014.06.063.

Rogers, C.E. *et al.* (2017) 'Neonatal amygdala functional connectivity at rest in healthy and preterm infants and early internalizing symptoms', *Journal of the American Academy of Child and Adolescent Psychiatry*, 56(2), pp. 157–166. Available at: https://doi.org/10.1016/j.jaac.2016.11.005.

Ross, G.S. *et al.* (2016) 'Using cluster analysis to provide new insights into development of very low birthweight (VLBW) premature infants', *Early Human Development*, 92(Complete), pp. 45–49. Available at: https://doi.org/10.1016/j.earlhumdev.2015.11.005.

Roth, S.C. *et al.* (1994) 'Relation between neurodevelopmental status of very preterm infants at one and eight years', *Developmental Medicine and Child Neurology*, 36(12), pp. 1049–1062. Available at: https://doi.org/10.1111/j.1469-8749.1994.tb11808.x.

Rothbart, M.K. *et al.* (2001) 'Investigations of temperament at three to seven years: the Children's Behavior Questionnaire', *Child Development*, 72(5), pp. 1394–1408. Available at: https://doi.org/10.1111/1467-8624.00355.

Rothbart, M.K. *et al.* (2003) 'Developing mechanisms of temperamental effortful control', *Journal of Personality*, 71(6), pp. 1113–1144. Available at: https://doi.org/10.1111/1467-6494.7106009.

Rothbart, M.K. (2004) 'Temperament and the Pursuit of an Integrated Developmental Psychology', *Merrill-Palmer Quarterly*, 50(4), pp. 492–505.

Rowlands, M.A. *et al.* (2016) 'Language at rest: A longitudinal study of intrinsic functional connectivity in preterm children', *NeuroImage. Clinical*, 11, pp. 149–157. Available at: https://doi.org/10.1016/j.nicl.2016.01.016.

Rubenstein, E. *et al.* (2018) 'Trends in documented co-occurring conditions in children with autism spectrum disorder, 2002–2010', *Research in developmental disabilities*, 83, pp. 168–178. Available at: https://doi.org/10.1016/j.ridd.2018.08.015.

Rushworth, M.F.S. *et al.* (2011) 'Frontal cortex and reward-guided learning and decision-making', *Neuron*, 70(6), pp. 1054–1069. Available at: https://doi.org/10.1016/j.neuron.2011.05.014.

Rutherford, M. *et al.* (2006) 'Magnetic resonance imaging in perinatal brain injury: clinical presentation, lesions and outcome', *Pediatric Radiology*, 36(7), pp. 582–592. Available at: https://doi.org/10.1007/s00247-006-0164-8.

Sa de Almeida, J. *et al.* (2021) 'Preterm birth leads to impaired rich-club organization and fronto-paralimbic/limbic structural connectivity in newborns', *NeuroImage*, 225, p. 117440. Available at: https://doi.org/10.1016/j.neuroimage.2020.117440.

Salehinejad, M.A. *et al.* (2021) 'Hot and cold executive functions in the brain: A prefrontal-cingular network', *Brain and Neuroscience Advances*, 5, p. 23982128211007769. Available at: https://doi.org/10.1177/23982128211007769.

Salimi-Khorshidi, G. *et al.* (2014) 'Automatic denoising of functional MRI data: combining independent component analysis and hierarchical fusion of classifiers', *NeuroImage*, 90, pp. 449–468. Available at: https://doi.org/10.1016/j.neuroimage.2013.11.046.

Salvan, P. *et al.* (2014) 'Road work on memory lane—Functional and structural alterations to the learning and memory circuit in adults born very preterm', *NeuroImage*, 102, pp. 152–161. Available at: https://doi.org/10.1016/j.neuroimage.2013.12.031.

Santamaría-García, H. *et al.* (2020) 'The role of social cognition skills and social determinants of health in predicting symptoms of mental illness', *Translational Psychiatry*, 10(1), pp. 1–13. Available at: https://doi.org/10.1038/s41398-020-0852-4.

Santens, E. *et al.* (2020) 'Effortful control – A transdiagnostic dimension underlying internalizing and externalizing psychopathology', *Neuropsychobiology*, 79(4), pp. 255–269. Available at: https://doi.org/10.1159/000506134.

Sarji, S.A. *et al.* (1998) 'Failed magnetic resonance imaging examinations due to claustrophobia', *Australasian Radiology*, 42(4), pp. 293–295. Available at: https://doi.org/10.1111/j.1440-1673.1998.tb00525.x.

Satterthwaite, T.D. *et al.* (2013) 'An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data', *NeuroImage*, 64, pp. 240–256. Available at: https://doi.org/10.1016/j.neuroimage.2012.08.052.

Saudino, K.J. *et al.* (1998) 'The validity of parent-based assessment of the cognitive abilities of 2-year-olds', *British Journal of Developmental Psychology*, 16(3), pp. 349–363. Available at: https://doi.org/10.1111/j.2044-835X.1998.tb00757.x.

Schafer, R.J. *et al.* (2009) 'Alterations in functional connectivity for language in prematurely born adolescents', *Brain*, 132(3), pp. 661–670. Available at: https://doi.org/10.1093/brain/awn353.

Scheinost, D. *et al.* (2015) 'Cerebral lateralization is protective in the very prematurely born', *Cerebral Cortex*, 25(7), pp. 1858–1866. Available at: https://doi.org/10.1093/cercor/bht430.

Scheinost, D. *et al.* (2022) 'Developmental trajectories of the default mode, executive control, and salience networks from the third trimester through the newborn period'. bioRxiv, p. 2022.09.27.509687. Available at: https://doi.org/10.1101/2022.09.27.509687.

Schoentgen, B., Gagliardi, G. and Défontaines, B. (2020) 'Environmental and cognitive enrichment in childhood as protective factors in the adult and aging brain', *Frontiers in Psychology*, 11. Available at: https://www.frontiersin.org/articles/10.3389/fpsyg.2020.01814.

Schoevers, R.A. *et al.* (2005) 'Depression and generalized anxiety disorder: Co-occurrence and longitudinal patterns in elderly patients', *The American Journal of Geriatric Psychiatry*, 13(1), pp. 31–39. Available at: https://doi.org/10.1097/00019442-200501000-00006.

Serag, A. *et al.* (2012) 'Construction of a consistent high-definition spatio-temporal atlas of the developing brain using adaptive kernel regression', *NeuroImage*, 59(3), pp. 2255–2265. Available at: https://doi.org/10.1016/j.neuroimage.2011.09.062.

Shang, J. *et al.* (2018) 'Decreased BOLD fluctuations in lateral temporal cortices of premature born adults', *Human Brain Mapping*, 39(12), pp. 4903–4912. Available at: https://doi.org/10.1002/hbm.24332.

Sherman, E.M.S. and Brooks, B.L. (2010) 'Behavior Rating Inventory of Executive Function-Preschool Version (BRIEF-P): Test review and clinical guidelines for use', *Child Neuropsychology*, 16, pp. 503–519. Available at: https://doi.org/10.1080/09297041003679344.

Shields, A. and Cicchetti, D. (1997) 'Emotion regulation among school-age children: The development and validation of a new criterion Q-sort scale', *Developmental Psychology*, 33(6), pp. 906–916. Available at: https://doi.org/10.1037/0012-1649.33.6.906.

Shimony, J.S. *et al.* (2016) 'Comparison of cortical folding measures for evaluation of developing human brain', *NeuroImage*, 125, pp. 780–790. Available at: https://doi.org/10.1016/j.neuroimage.2015.11.001.

Siffredi, V. *et al.* (2022) 'Large-scale brain network dynamics in very preterm children and relationship with socio-emotional outcomes: an exploratory study', *Pediatric Research*, pp. 1–9. Available at: https://doi.org/10.1038/s41390-022-02342-y.

Simonds, J. and Rothbart, M.K. (2006) 'Temperament in middle childhood questionnaire', *Downloaded in http://www. bowdoin. edu/~ sputnam/rothbart-temperament-questionnaires* [Preprint].

Simpson, T.I., Armstrong, J.D. and Jarman, A.P. (2010) 'Merged consensus clustering to assess and improve class discovery with microarray data', *BMC Bioinformatics*, 11(1), p. 590. Available at: https://doi.org/10.1186/1471-2105-11-590.

Siugzdaite, R. *et al.* (2020) 'Transdiagnostic brain mapping in developmental disorders', *Current Biology*, 30(7), pp. 1245-1257.e4. Available at: https://doi.org/10.1016/j.cub.2020.01.078.

Skiöld, B. *et al.* (2019) 'A novel scoring system for term-equivalent-age cranial ultrasound in extremely preterm infants', *Ultrasound in Medicine & Biology*, 45(3), pp. 786–794. Available at: https://doi.org/10.1016/j.ultrasmedbio.2018.11.005.

Skranes, J. *et al.* (2007) 'Clinical findings and white matter abnormalities seen on diffusion tensor imaging in adolescents with very low birth weight', *Brain: A Journal of Neurology*, 130(Pt 3), pp. 654–666. Available at: https://doi.org/10.1093/brain/awm001.

Smith, S.M. (2002) 'Fast robust automated brain extraction', *Human Brain Mapping*, 17(3), pp. 143–155. Available at: https://doi.org/10.1002/hbm.10062.

Smith, S.M. *et al.* (2006) 'Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data', *NeuroImage*, 31(4), pp. 1487–1505. Available at: https://doi.org/10.1016/j.neuroimage.2006.02.024.

Smith, S.M. and Nichols, T.E. (2009) 'Threshold-Free Cluster Enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference', *NeuroImage*, 44(1), pp. 83–98. Available at: https://doi.org/10.1016/j.neuroimage.2008.03.061.

Smyser, C.D. *et al.* (2010) 'Longitudinal analysis of neural network development in preterm infants', *Cerebral Cortex*, 20(12), pp. 2852–2862. Available at: https://doi.org/10.1093/cercor/bhq035.

Smyser, C.D. *et al.* (2016) 'Resting-state network complexity and magnitude are reduced in prematurely born infants', *Cerebral Cortex*, 26(1), pp. 322–333. Available at: https://doi.org/10.1093/cercor/bhu251.

Solmi, M. *et al.* (2022) 'Age at onset of mental disorders worldwide: large-scale meta-analysis of 192 epidemiological studies', *Molecular Psychiatry*, 27(1), pp. 281–295. Available at: https://doi.org/10.1038/s41380-021-01161-7.

Solomon, M., Buaminger, N. and Rogers, S.J. (2011) 'Abstract reasoning and friendship in high functioning preadolescents with autism spectrum disorders', *Journal of Autism and Developmental Disorders*, 41(1), pp. 32–43. Available at: https://doi.org/10.1007/s10803-010-1017-8.

Sørensen, T. (1948) *A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons*. Munksgaard.

Spence, S.H. (1998) 'A measure of anxiety symptoms among children', *Behaviour Research and Therapy*, 36(5), pp. 545–566. Available at: https://doi.org/10.1016/s0005-7967(98)00034-5.

Spittle, A.J. *et al.* (2009) 'Early emergence of behavior and social-emotional problems in very preterm infants', *Journal of the American Academy of Child and Adolescent Psychiatry*, 48(9), pp. 909–918. Available at: https://doi.org/10.1097/CHI.0b013e3181af8235.

Stålnacke, J. *et al.* (2019) 'A longitudinal model of executive function development from birth through adolescence in children born very or extremely preterm', *Child Neuropsychology*, 25(3), pp. 318–335. Available at: https://doi.org/10.1080/09297049.2018.1477928.

Stamenova, V. and Levine, B. (2019) 'Effectiveness of goal management training® in improving executive functions: A meta-analysis', *Neuropsychological Rehabilitation*, 29(10), pp. 1569–1599. Available at: https://doi.org/10.1080/09602011.2018.1438294.

Stark, T.H. and Krosnick, J.A. (2017) 'GENSI: A new graphical tool to collect ego-centered network data', *Social Networks*, 48, pp. 36–45. Available at: https://doi.org/10.1016/j.socnet.2016.07.007.

Starr, R. *et al.* (2023) 'Periventricular and intraventricular hemorrhage', in *StatPearls*. Treasure Island (FL): StatPearls Publishing. Available at: http://www.ncbi.nlm.nih.gov/books/NBK538310/.

Stefanik, L. *et al.* (2018) 'Brain-behavior participant similarity networks among youth and emerging adults with schizophrenia spectrum, autism spectrum, or bipolar disorder and matched controls', *Neuropsychopharmacology*, 43(5), pp. 1180–1188. Available at: https://doi.org/10.1038/npp.2017.274.

Stewart, A.L. *et al.* (1983) 'Ultrasound appearance of the brain in very preterm infants and neurodevelopmental outcome at 18 months of age.', *Archives of Disease in Childhood*, 58(8), pp. 598–604.

Stewart, A.L. *et al.* (1989) 'Relationship between neurodevelopmental status of very preterm infants at one and four years', *Developmental Medicine and Child Neurology*, 31(6), pp. 756–765. Available at: https://doi.org/10.1111/j.1469-8749.1989.tb04071.x.

Stoecklein, S. *et al.* (2020) 'Variable functional connectivity architecture of the preterm human brain: Impact of developmental cortical expansion and maturation', *Proceedings of the National Academy of Sciences*, 117(2), pp. 1201–1206. Available at: https://doi.org/10.1073/pnas.1907892117.

Sun, Z. *et al.* (2023) 'Comparing the emotional impact of the UK COVID-19 lockdown in very preterm and full-term born children: a longitudinal study', *Frontiers in Child and Adolescent Psychiatry*, 2. Available at: https://www.frontiersin.org/articles/10.3389/frcha.2023.1193258.

Sylvester, C.M. *et al.* (2013) 'Resting state functional connectivity of the ventral attention network in children with a history of depression or anxiety', *Journal of the American Academy of Child and Adolescent Psychiatry*, 52(12), pp. 1326-1336.e5. Available at: https://doi.org/10.1016/j.jaac.2013.10.001.

Sylvester, C.M. *et al.* (2018) 'Cortical functional connectivity evident after birth and behavioral inhibition at age two years', *The American journal of psychiatry*, 175(2), pp. 180–187. Available at: https://doi.org/10.1176/appi.ajp.2017.17010018.

Tanzer, M. *et al.* (2021) 'Cortical thickness of the insula and prefrontal cortex relates to externalizing behavior: Cross-sectional and prospective findings', *Development and Psychopathology*, pp. 1–11. Available at: https://doi.org/10.1017/S0954579420000619.

Theberath, M. *et al.* (2022) 'Effects of COVID-19 pandemic on mental health of children and adolescents: A systematic review of survey studies', *SAGE Open Medicine*, 10, p. 20503121221086712. Available at: https://doi.org/10.1177/20503121221086712.

Thomason, M.E. *et al.* (2017) 'Weak functional connectivity in the human fetal brain prior to preterm birth', *Scientific Reports*, 7(1), p. 39286. Available at: https://doi.org/10.1038/srep39286.

Thompson, D.K. *et al.* (2012) 'Corpus callosum alterations in very preterm infants: perinatal correlates and 2 year neurodevelopmental outcomes', *Neuroimage*, 59(4), pp. 3571–3581. Available at: https://doi.org/10.1016/j.neuroimage.2011.11.057.

Thompson, D.K. *et al.* (2016) 'Structural connectivity relates to perinatal factors and functional impairment at 7years in children born very preterm', *NeuroImage*, 134, pp. 328–337. Available at: https://doi.org/10.1016/j.neuroimage.2016.03.070.

Thompson, D.K. *et al.* (2018) 'Characterisation of brain volume and microstructure at term-equivalent age in infants born across the gestational age spectrum', *NeuroImage : Clinical*, 21, p. 101630. Available at: https://doi.org/10.1016/j.nicl.2018.101630.

Thompson, D.K. *et al.* (2020) 'Tracking regional brain growth up to age 13 in children born term and very preterm', *Nature Communications*, 11(1), p. 696. Available at: https://doi.org/10.1038/s41467-020-14334-9.

Thompson, R.A., Meyer, S. and Jochem, R. (2008) 'Emotion regulation', in M.M. Haith and J.B. Benson (eds) *Encyclopedia of Infant and Early Childhood Development*. San Diego: Academic Press, pp. 431–441. Available at: https://doi.org/10.1016/B978-012370877-9.00055-4.

Thompson, R.A. and Nelson, C.A. (2001) 'Developmental science and the media. Early brain development', *The American Psychologist*, 56(1), pp. 5–15. Available at: https://doi.org/10.1037/0003-066x.56.1.5.

Tierney, A.L. and Nelson, C.A. (2009) 'Brain development and the role of experience in the early years', *Zero to three*, 30(2), pp. 9–13.

Tingley, D. *et al.* (2014) '**mediation** : *R* Package for Causal Mediation Analysis', *Journal of Statistical Software*, 59(5). Available at: https://doi.org/10.18637/jss.v059.i05.

Toal, F. *et al.* (2010) 'Clinical and anatomical heterogeneity in autistic spectrum disorder: a structural MRI study', *Psychological Medicine*, 40(7), pp. 1171–1181. Available at: https://doi.org/10.1017/S0033291709991541.

Tombaugh, T.N. (2004) 'Trail Making Test A and B: normative data stratified by age and education', *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, 19(2), pp. 203–214. Available at: https://doi.org/10.1016/S0887-6177(03)00039-8.

Tong, X. *et al.* (2022) 'Transdiagnostic connectome signatures from resting-state fMRI predict individual-level intellectual capacity', *Translational Psychiatry*, 12(1), pp. 1–11. Available at: https://doi.org/10.1038/s41398-022-02134-2.

Tooley, U.A. *et al.* (2022) 'The age of reason: functional brain network development during childhood', *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 42(44), pp. 8237–8251. Available at: https://doi.org/10.1523/JNEUROSCI.0511-22.2022.

Toulmin, H. *et al.* (2015) 'Specialization and integration of functional thalamocortical connectivity in the human infant', *Proceedings of the National Academy of Sciences of the United States of America*, 112(20), pp. 6485–6490. Available at: https://doi.org/10.1073/pnas.1422638112.

Treyvaud, K. *et al.* (2012) 'Can the home environment promote resilience for children born very preterm in the context of social and medical risk?', *Journal of Experimental Child Psychology*, 112(3), pp. 326–337. Available at: https://doi.org/10.1016/j.jecp.2012.02.009.

Treyvaud, K. *et al.* (2013) 'Psychiatric outcomes at age seven for very preterm children: rates and predictors', *Journal of Child Psychology and Psychiatry*, 54(7), pp. 772–779. Available at: https://doi.org/10.1111/jcpp.12040.

Tsai, P.T. *et al.* (2012) 'Autistic-like behaviour and cerebellar dysfunction in Purkinje cell Tsc1 mutant mice', *Nature*, 488(7413), pp. 647–651. Available at: https://doi.org/10.1038/nature11310.

Tustison, N.J. *et al.* (2010) 'N4ITK: Improved N3 Bias Correction', *IEEE transactions on medical imaging*, 29(6), pp. 1310–1320. Available at: https://doi.org/10.1109/TMI.2010.2046908.

Twilhaar, E.S. *et al.* (2018) 'Cognitive outcomes of children born extremely or very preterm since the 1990s and associated risk factors: A meta-analysis and meta-regression', *JAMA pediatrics*, 172(4), pp. 361–367. Available at: https://doi.org/10.1001/jamapediatrics.2017.5323.

Uddin, L.Q. *et al.* (2011) 'Dynamic reconfiguration of structural and functional connectivity across core neurocognitive brain networks with development', *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 31(50), pp. 18578–18589. Available at: https://doi.org/10.1523/JNEUROSCI.4465-11.2011.

Uddin, L.Q. *et al.* (2017) 'Structure and function of the human insula', *Journal of clinical neurophysiology : official publication of the American Electroencephalographic Society*, 34(4), pp. 300–306. Available at: https://doi.org/10.1097/WNP.0000000000000377.

Ure, A.M. *et al.* (2016) 'Neonatal brain abnormalities associated with autism spectrum disorder in children born very preterm', *Autism Research: Official Journal of the International Society for Autism Research*, 9(5), pp. 543–552. Available at: https://doi.org/10.1002/aur.1558.

Uytun, M.C. (2018) 'Development period of prefrontal cortex', in *Prefrontal Cortex*. IntechOpen. Available at: https://doi.org/10.5772/intechopen.78697.

Vaidya, C.J. *et al.* (2020) 'Data-driven identification of subtypes of executive function across typical development, attention deficit hyperactivity disorder, and autism spectrum disorders', *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 61(1), pp. 51–61. Available at: https://doi.org/10.1111/jcpp.13114.

Van Hus, J.W. *et al.* (2014) 'Motor impairment in very preterm-born children: links with other developmental deficits at 5 years of age', *Developmental Medicine & Child Neurology*, 56(6), pp. 587–594. Available at: https://doi.org/10.1111/dmcn.12295.

Vanderwal, T. *et al.* (2015) 'Inscapes: A movie paradigm to improve compliance in functional magnetic resonance imaging', *NeuroImage*, 122, pp. 222–232. Available at: https://doi.org/10.1016/j.neuroimage.2015.07.069.

Vandewouw, M.M. *et al.* (2023) 'Identifying replicable subgroups in neurodevelopmental conditions using resting-state functional magnetic resonance imaging data', *JAMA Network Open*, 6(3), p. e232066. Available at: https://doi.org/10.1001/jamanetworkopen.2023.2066.

Vanes, L. *et al.* (2023) 'Longitudinal neonatal brain development and socio-demographic correlates of infant outcomes following preterm birth', *Developmental Cognitive Neuroscience*, 61, p. 101250. Available at: https://doi.org/10.1016/j.dcn.2023.101250.

Vanes, L.D. *et al.* (2021) 'Associations between neonatal brain structure, the home environment, and childhood outcomes following very preterm birth', *Biological Psychiatry: Global Open Science*, 1(2), pp. 146–155. Available at: https://doi.org/10.1016/j.bpsgos.2021.05.002.

Váša, F. *et al.* (2020) 'Conservative and disruptive modes of adolescent change in human brain functional connectivity', *Proceedings of the National Academy of Sciences*, 117(6), pp. 3248–3253. Available at: https://doi.org/10.1073/pnas.1906144117.

Vollmer, B. *et al.* (2017) 'Correlation between white matter microstructure and executive functions suggests early developmental influence on long fibre tracts in preterm born adolescents', *PloS One*, 12(6), p. e0178893. Available at: https://doi.org/10.1371/journal.pone.0178893.

Vollmer, B. and Stålnacke, J. (2019) 'Young adult motor, sensory, and cognitive outcomes and longitudinal development after very and extremely preterm birth', *Neuropediatrics*, 50(4), pp. 219–227. Available at: https://doi.org/10.1055/s-0039-1688955.

Volpe, J.J. (2009a) 'Brain injury in premature infants: a complex amalgam of destructive and developmental disturbances', *Lancet neurology*, 8(1), pp. 110–124. Available at: https://doi.org/10.1016/S1474-4422(08)70294-1.

Volpe, J.J. (2009b) 'The encephalopathy of prematurity – brain injury and impaired brain development inextricably intertwined', *Seminars in pediatric neurology*, 16(4), pp. 167–178. Available at: https://doi.org/10.1016/j.spen.2009.09.005.

Von Der Heide, R.J. *et al.* (2013) 'Dissecting the uncinate fasciculus: disorders, controversies and a hypothesis', *Brain*, 136(6), pp. 1692–1707. Available at: https://doi.org/10.1093/brain/awt094.

Vossel, S., Geng, J.J. and Fink, G.R. (2014) 'Dorsal and ventral attention systems', *The Neuroscientist*, 20(2), pp. 150–159. Available at: https://doi.org/10.1177/1073858413494269.

Wade, M. *et al.* (2020) 'Global deficits in executive functioning are transdiagnostic mediators between severe childhood neglect and psychopathology in adolescence', *Psychological Medicine*, 50(10), pp. 1687–1694. Available at: https://doi.org/10.1017/S0033291719001764.

Walani, S.R. (2020) 'Global burden of preterm birth', *International Journal of Gynaecology and Obstetrics: The Official Organ of the International Federation of Gynaecology and Obstetrics*, 150(1), pp. 31–33. Available at: https://doi.org/10.1002/ijgo.13195.

Wang, B. *et al.* (2014) 'Similarity network fusion for aggregating data types on a genomic scale', *Nature Methods*, 11(3), pp. 333–337. Available at: https://doi.org/10.1038/nmeth.2810.

Wang, B. *et al.* (2018) 'SNFtool: Similarity Network Fusion. R package version 2.3.0. https://CRAN.R-project.org/package=SNFtool'.

Wang, H. *et al.* (2023) 'Meta-analysis on the effects of moderate-intensity exercise intervention on executive functioning in children', *PLOS ONE*, 18(2), p. e0279846. Available at: https://doi.org/10.1371/journal.pone.0279846.

Wang, S.S.-H., Kloth, A.D. and Badura, A. (2014) 'The cerebellum, sensitive periods, and autism', *Neuron*, 83(3), pp. 518–532. Available at: https://doi.org/10.1016/j.neuron.2014.07.016.

Watson, T.C. *et al.* (2013) 'The olivo-cerebellar system and its relationship to survival circuits', *Frontiers in Neural Circuits*, 7, p. 72. Available at: https://doi.org/10.3389/fncir.2013.00072.

Wechsler, D. (1999) *Wechsler Abbreviated Scale of Intelligence*. Available at: https://psycnet.apa.org/doiLanding?doi=10.1037%2Ft15170-000.

Wechsler, David (2012) 'Wechsler Intelligence Scale for Children, Fourth Edition'. Available at: https://doi.org/10.1037/t15174-000.

Wechsler, D (2012) 'Wechsler Preschool and Primary Scale of Intelligence-fourth edition technical manual and interpretive manual. San Antonio,TX: Psychological Corporation. Copyright 2012. Pearson, Inc., and/or its affiliates.' Available at: http://www.pearsonclinical.com/.

Wegiel, Jerzy *et al.* (2010) 'The neuropathology of autism: defects of neurogenesis and neuronal migration, and dysplastic changes', *Acta Neuropathologica*, 119(6), pp. 755–770. Available at: https://doi.org/10.1007/s00401-010-0655-4.

Wegiel, Jerzy *et al.* (2013) 'Contribution of olivofloccular circuitry developmental defects to atypical gaze in autism', *Brain research*, 1512, pp. 106–122. Available at: https://doi.org/10.1016/j.brainres.2013.03.037.

Wegiel, Jerzy *et al.* (2014) 'Stereological study of the neuronal number and volume of 38 brain subdivisions of subjects diagnosed with autism reveals significant alterations restricted to the striatum, amygdala and cerebellum', *Acta Neuropathologica Communications*, 2, p. 141. Available at: https://doi.org/10.1186/s40478-014-0141-7.

Wehrle, F.M. *et al.* (2018) 'Altered resting-state functional connectivity in children and adolescents born very preterm short title', *NeuroImage. Clinical*, 20, pp. 1148–1156. Available at: https://doi.org/10.1016/j.nicl.2018.10.002.

Weissman, M.M. and Bothwell, S. (1976) 'Assessment of social adjustment by patient self-report', *Archives of General Psychiatry*, 33(9), pp. 1111–1115. Available at: https://doi.org/10.1001/archpsyc.1976.01770090101010.

Westlin, C. *et al.* (2023) 'Improving the study of brain-behavior relationships by revisiting basic assumptions', *Trends in cognitive sciences*, 27(3), pp. 246–257. Available at: https://doi.org/10.1016/j.tics.2022.12.015.

Wheelock, M.D. *et al.* (2018) 'Altered functional network connectivity relates to motor development in children born very preterm', *NeuroImage*, 183, pp. 574–583. Available at: https://doi.org/10.1016/j.neuroimage.2018.08.051.

Wheelock, M.D. *et al.* (2021) 'Functional connectivity network disruption underlies domain-specific impairments in attention for children born very preterm', *Cerebral Cortex*, 31(2), pp. 1383–1394. Available at: https://doi.org/10.1093/cercor/bhaa303.

White, T.P. *et al.* (2014) 'Dysconnectivity of neurocognitive networks at rest in very-preterm born adults', *NeuroImage. Clinical*, 4, pp. 352–365. Available at: https://doi.org/10.1016/j.nicl.2014.01.005.

Wickremasinghe, A. *et al.* (2012) 'Evaluation of the ability of neurobiological, neurodevelopmental and socio-economic variables to predict cognitive outcome in premature infants.', *Child: care, health and development*. Available at: https://doi.org/10.1111/j.1365-2214.2011.01281.x.

Wier, R. *et al.* (2019) 'Fronto-limbic white matter microstructure, behavior, and emotion regulation in survivors of pediatric brain tumor', *Journal of Neuro-Oncology*, 143(3), pp. 483–493. Available at: https://doi.org/10.1007/s11060-019-03180-5.

Wigham, S. *et al.* (2012) 'The reliability and validity of the Social Responsiveness Scale in a UK general child population', *Research in Developmental Disabilities*, 33(3), pp. 944–950. Available at: https://doi.org/10.1016/j.ridd.2011.12.017.

Wilke, M. *et al.* (2014) 'Specific impairment of functional connectivity between language regions in former early preterms', *Human Brain Mapping*, 35(7), pp. 3372–3384. Available at: https://doi.org/10.1002/hbm.22408.

Wilkerson, M.D. and Hayes, D.N. (2010) 'ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking', *Bioinformatics*, 26(12), pp. 1572–1573. Available at: https://doi.org/10.1093/bioinformatics/btq170.

Wilson, S. *et al.* (2021) 'Development of human white matter pathways in utero over the second and third trimester', *Proceedings of the National Academy of Sciences of the United States of America*, 118(20), p. e2023598118. Available at: https://doi.org/10.1073/pnas.2023598118.

Wilson, S. *et al.* (2023) 'Spatiotemporal tissue maturation of thalamocortical pathways in the human fetal brain', *eLife*, 12, p. e83727. Available at: https://doi.org/10.7554/eLife.83727.

Winstanley, A. *et al.* (2015) 'The subjective well-being of adults born preterm', *Journal of Research in Personality*, 59, pp. 23–30. Available at: https://doi.org/10.1016/j.jrp.2015.09.002.

Wolf, R.L. *et al.* (2001) 'Quantitative apparent diffusion coefficient measurements in term neonates for early detection of hypoxic-ischemic brain injury: Initial experience', *Radiology*, 218(3), pp. 825–833. Available at: https://doi.org/10.1148/radiology.218.3.r01fe47825.

Wolke, D., Jaekel, J., *et al.* (2013) 'Effects of sensitive parenting on the academic resilience of very preterm and very low birth weight adolescents', *The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine*, 53(5), pp. 642–647. Available at: https://doi.org/10.1016/j.jadohealth.2013.06.014.

Wolke, D., Chernova, J., *et al.* (2013) 'Self and parent perspectives on health-related quality of life of adolescents born very preterm', *The Journal of Pediatrics*, 163(4), pp. 1020-1026.e2. Available at: https://doi.org/10.1016/j.jpeds.2013.04.030.

Wolke, D. (2019) 'Is social inequality in cognitive outcomes increased by preterm birth–related complications?', *JAMA Network Open*, 2(5), pp. e192902–e192902. Available at: https://doi.org/10.1001/jamanetworkopen.2019.2902.

Woodward, L.J. *et al.* (2017) 'Preschool self regulation predicts later mental health and educational achievement in very preterm and typically developing children', *The Clinical Neuropsychologist*, 31(2), pp. 404–422. Available at: https://doi.org/10.1080/13854046.2016.1251614.

World Health Organization (2019) *International statistical classification of diseases and related health problems (11th ed.)*. Available at: https://icd.who.int/en.

Xu, T. *et al.* (2017) 'CancerSubtypes: an R/Bioconductor package for molecular cancer subtype identification, validation and visualization', *Bioinformatics*, 33(19), pp. 3131–3133. Available at: https://doi.org/10.1093/bioinformatics/btx378.

Xu, Y. and Goodacre, R. (2018) 'On splitting training and validation set: a comparative study of cross-validation, bootstrap and systematic sampling for estimating the generalization performance of supervised learning', *Journal of Analysis and Testing*, 2(3), pp. 249–262. Available at: https://doi.org/10.1007/s41664-018-0068-2.

Yeo, B.T. *et al.* (2011) 'The organization of the human cerebral cortex estimated by intrinsic functional connectivity', *Journal of Neurophysiology*, 106(3), pp. 1125–1165. Available at: https://doi.org/10.1152/jn.00338.2011.

Yin, Z. *et al.* (2018) 'Decreased functional connectivity in insular subregions in depressive episodes of bipolar disorder and major depressive disorder', *Frontiers in Neuroscience*, 12, p. 842. Available at: https://doi.org/10.3389/fnins.2018.00842.

Yu, S.-H., Tseng, C.-Y. and Lin, W.-L. (2020) 'A neurofeedback protocol for executive function to reduce depression and rumination: A controlled study', *Clinical Psychopharmacology and Neuroscience*, 18(3), pp. 375–385. Available at: https://doi.org/10.9758/cpn.2020.18.3.375.

Yung, A.R. *et al.* (2005) 'Mapping the onset of psychosis: the Comprehensive Assessment of At-Risk Mental States', *The Australian and New Zealand Journal of Psychiatry*, 39(11–12), pp. 964–971. Available at: https://doi.org/10.1080/j.1440-1614.2005.01714.x.

Zalesky, A. (2012) 'Reference Manual for NBS Connectome (v1.2)'.

Zalesky, A. *et al.* (2016) 'Connectome sensitivity or specificity: which is more important?', *NeuroImage*, 142, pp. 407–420. Available at: https://doi.org/10.1016/j.neuroimage.2016.06.035.

Zalesky, A., Fornito, A. and Bullmore, E.T. (2010) 'Network-based statistic: identifying differences in brain networks', *NeuroImage*, 53(4), pp. 1197–1207. Available at: https://doi.org/10.1016/j.neuroimage.2010.06.041.

Zanini, S. *et al.* (2009) 'Grammar improvement following deep brain stimulation of the subthalamic and the pedunculopontine nuclei in advanced Parkinson's disease: a pilot study', *Parkinsonism & Related Disorders*, 15(8), pp. 606–609. Available at: https://doi.org/10.1016/j.parkreldis.2008.12.003.

Zhang, Y., Brady, M. and Smith, S. (2001) 'Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm', *IEEE transactions on medical imaging*, 20(1), pp. 45–57. Available at: https://doi.org/10.1109/42.906424.

Zhou, L. *et al.* (2018) 'Brain gray and white matter abnormalities in preterm-born adolescents: A meta-analysis of voxel-based morphometry studies', *PLOS ONE*, 13(10), p. e0203498. Available at: https://doi.org/10.1371/journal.pone.0203498.

Zuo, X.-N. *et al.* (2012) 'Network centrality in the human functional connectome', *Cerebral Cortex*, 22(8), pp. 1862–1875. Available at: https://doi.org/10.1093/cercor/bhr269.

# Appendix A

A PDF copy of the published manuscript referred to in Chapter 3 can be found here.

***Reference:*** *Hadaya, L., Vanes, L., Karolis, V., Kanel, D., Leoni, M., Happé, F., Edwards, A. D., Counsell, S. J., Batalle, D., & Nosarti, C. (2022). Distinct Neurodevelopmental Trajectories in Groups of Very Preterm Children Screening Positively for Autism Spectrum Conditions. Journal of Autism and Developmental Disorders.* https://doi.org/10.1007/s10803-022-05789-4.

Journal of Autism and Developmental Disorders
https://doi.org/10.1007/s10803-022-05789-4

ORIGINAL PAPER

# Distinct Neurodevelopmental Trajectories in Groups of Very Preterm Children Screening Positively for Autism Spectrum Conditions

Laila Hadaya[1,2] · Lucy Vanes[1,2] · Vyacheslav Karolis[1,3] · Dana Kanel[1,2] · Marguerite Leoni[4] · Francesca Happé[4] · A. David Edwards[1] · Serena J. Counsell[1] · Dafnis Batalle[1,5] · Chiara Nosarti[1,2]

Accepted: 9 October 2022


**Abstract**
Very preterm (VPT; < 33 weeks' gestation) toddlers screening positively for autism spectrum conditions (ASC) may display heterogenous neurodevelopmental trajectories. Here we studied neonatal brain volumes and childhood ASC traits evaluated with the Social Responsiveness Scale (SRS-2) in VPT-born toddlers (N = 371; median age 20.17 months) sub-divided into three groups based on their Modified-Checklist for Autism in Toddlers scores. These were: those screening positively failing at least 2 critical items (*critical-positive*); failing any 3 items, but less than 2 critical items (*non-critical-positive*); and screening negatively. Critical-positive scorers had smaller neonatal cerebellar volumes compared to non-critical-positive and negative scorers. However, both positive screening groups exhibited higher childhood ASC traits compared to the negative screening group, suggesting distinct aetiological trajectories associated with ASC outcomes.

**Keywords** Autism spectrum conditions · Developmental delay · Very preterm birth · Structural MRI

## Introduction

The parent-rated Modified Checklist for Autism in Toddlers (M-CHAT), assessing child skills and behaviours, was developed as a screening tool for autism spectrum conditions (ASC) (Robins et al., 2001). ASC are characterised by two sets of core symptoms: (a) social communication and interaction deficits (SCI), which reflect difficulties in non-verbal social gestures, socio-emotional reciprocity and maintaining and developing social relationships, and (b) restricted interests and repetitive behaviours (RRBs), which include restricted and fixated interests, ritualised behaviours and altered sensitivity to sensory stimuli (American Psychiatric Association, 2013). According to the original M-CHAT scoring criteria, a positive M-CHAT screening is obtained when a child fails two or more 'critical' items within a set of six (e.g., "Does your child imitate you?", "Does your child take an interest in other children?"), or three or more items overall (Robins et al., 2001). However, research in low-risk toddlers has more recently led to the recommendation of abandoning these criteria in favour of a total number of items failed, as this approach has been shown to improve the tool's sensitivity to identify a later ASC diagnosis (Chlebowski et al., 2013).

Studies in high-risk samples using the original screening criteria have shown that very preterm (VPT; < 32 weeks' gestation) and extremely preterm (EPT; < 28 weeks' gestation) born toddlers are more likely to screen positively on the M-CHAT (21–25%; Limperopoulos et al., 2008; Kuban et al., 2009), compared to full-term born toddlers (5.7%; Kleinman et al., 2008). These findings, together with those showing a higher prevalence of ASC diagnoses in children

✉ Chiara Nosarti
chiara.nosarti@kcl.ac.uk

1 Centre for the Developing Brain, Department of Perinatal Imaging and Health, School of Biomedical Engineering and Imaging Sciences, King's College London, London SE1 7EH, UK

2 Department of Child and Adolescent Psychiatry, Institute of Psychiatry Psychology and Neuroscience, King's College London, 16 De Crespigny Park, London SE5 8AF, UK

3 Wellcome Centre for Integrative Neuroimaging, FMRIB, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford OX3 9DU, UK

4 Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry Psychology and Neuroscience, King's College London, London SE5 8AF, UK

5 Department of Forensic and Neurodevelopmental Sciences, Institute of Psychiatry Psychology and Neuroscience, King's College London, London SE5 8AF, UK

Published online: 23 October 2022

born VPT (7%) compared to those born at term (1.5%; Joseph et al., 2017; Agrawal et al., 2018), suggest that VPT children may be vulnerable to experiencing both subthreshold and clinical core ASC symptoms. However, in high-risk EPT/VPT toddlers the interpretability of the M-CHAT screening has been questioned (Luyster et al., 2011; Moore et al., 2012), as these children tend to display impaired social and communication skills, which are shared features of both the so-called "preterm behavioural phenotype" (Johnson & Marlow, 2011) and ASC traits (American Psychiatric Association, 2013). Moore et al. (2012) suggested that the two original M-CHAT positive scoring criteria may differentiate between EPT toddlers with and without neurodevelopmental disabilities, as they found that the stricter critical positive screening criteria were associated with more severe neurodevelopmental impairments compared to the more liberal non-critical criteria (Luyster et al., 2011; Moore et al., 2012). Given the increased risk of developmental delay following preterm birth (Blencowe et al., 2013) and the frequent co-occurrence of developmental delay in ASC (Rubenstein et al., 2018), the use of the initially proposed different M-CHAT positive scoring criteria may therefore aid the identification of subgroups of EPT/VPT toddlers exhibiting distinct neurodevelopmental trajectories.

Widespread alterations in brain development associated with VPT birth (Volpe, 2009), may at least partly contribute to the increased likelihood of ASC behaviours in VPT children. Structural reductions in volume and alterations in functional connectivity in temporal, prefrontal, limbic and cerebellar regions have been observed in VPT individuals in the neonatal period and beyond (Ball et al., 2013, 2016; Fenoglio et al., 2017; Healy et al., 2013; Kanel et al., 2022; Rogers et al., 2012). Alterations in these regions have also been implicated in key components of ASC symptomatology (Alcalá-López et al., 2018; Ciarrusta et al., 2019; Gandhi & Lee, 2021; Ha et al., 2015) and in VPT neonates who develop ASC later in childhood (Eklöf et al., 2019; Padilla et al., 2017; Ure et al., 2016). However, no study to date has explored whether different M-CHAT positive scoring criteria could be used to identify subgroups of VPT toddlers who differ in terms of early brain development and ASC behaviour later in childhood.

In order to address these questions, this study had two main aims: to explore whether distinct M-CHAT screening groups (critical positive, non-critical positive and negative), which have been previously studied in relation to neurodevelopmental impairments in EPT toddlers (Moore et al., 2012), also differed in VPT toddlers in terms of (a) neonatal structural brain volumes and (b) ASC profiles later in childhood. Exploratory analyses were further conducted to probe the role of developmental delay in shaping the childhood trajectory for ASC traits in the different screening groups, with the use of mediation and moderation analyses.

Our first hypothesis was that both M-CHAT positive screening groups (i.e., critical positive and non-critical positive) would display volumetric reductions at term-equivalent age in brain regions implicated in ASC symptomatology (e.g., temporal, prefrontal cortex and cerebellum) compared to the negative screening group. Our second hypothesis was that toddlers belonging to the two M-CHAT positive screening groups would display more ASC-type behaviours in childhood (age 4–7 years) than toddlers belonging to the negative screening group. Thirdly, exploratory analyses tested two competing hypotheses, namely that the critical positive scorers would either exhibit: (a) fewer ASC-like behaviours than the non-critical positive scorers, indicating that a critical positive screening may reflect developmental delay (Luyster et al., 2011; Moore et al., 2012), rather than persisting ASC behaviours, or (b) similar ASC-like behaviours to the non-critical positive scorers, indicating distinct trajectories leading to similar ASC behaviours (i.e., equifinality; Cicchetti & Rogosch, 1996).

## Methods

### Participants and Study Design

511 children born at 33 weeks' gestational age or less (median = 30 weeks; range = 23–32 weeks), between April 2010 and July 2013, were enrolled into the "Evaluation of Preterm Imaging" study (ePrime; EudraCT 2009-011602-42; Edwards et al., 2018) from 14 neonatal units across London. Inclusion criteria were: birth at or less than 33 weeks' gestation; English-speaking parents not undergoing child protection proceedings; no magnetic resonance imaging (MRI) contraindications or major congenital malformations. Infants underwent multimodal (T1-weighted, T2-weighted, diffusion and functional) MRI at term-equivalent age (38–44 weeks) and were followed-up for behavioural and cognitive assessments at 2 (N = 484; 95% of the initial sample) and 4–7 years (N = 251; 82% of those children approached for follow-up).

Complete M-CHAT follow-up data at 2 years were available for 371 children (49.60% female; 23.18% born EPT) meeting MRI analysis inclusion criteria: i.e., postmenstrual age (PMA) at scan < 46 weeks, having no periventricular leukomalacia, parenchymal haemorrhagic infarction, or other major ischemic or haemorrhagic lesions detected on MRI or missing T2-weighted or motion corrupted images. 177 children had complete SRS-2 data at the subsequent 4–7-year follow-up (46.90% females; 25.42% born EPT). Sample characteristics are summarised in Table 1. The EPT and VPT born children within our cohort did not differ in severity of ASC traits or developmental delay (Table SM1).

**Table 1** Sample characteristics

| Variables | Median (range) | |
|---|---|---|
| | 2-year follow-up (N = 371) | 4–7-year follow-up (N = 177) |
| GA, weeks | 30.29 (23.57–32.86) | 30.29 (24–32.86) |
| IMD score at birth | 17.71 (1.73–60.58) | 16.12 (1.73–59.16) |
| PMA at scan | 42.57 (37.86–44.86) | 42.57 (38.29–44.86) |
| Neonatal sickness[a] | −0.30 (−1.36–2.55) | −0.35 (−1.34–2.18) |
| Corrected age at assessment | 20.17 (18.37–29.33) months | 4.59 (4.18–7.17) years |

Sample characteristics (median and range) for 2-year follow-up sample with complete M-CHAT and structural MRI data and for 4–7-year follow-up sample with complete M-CHAT and SRS-2 data

*GA* gestational age, *IMD* index multiple deprivation, *PMA* postmenstrual age

[a]excluding one subject with incomplete clinical data

## MR Imaging Data

### Data Acquisition

A 3-Tesla system (Philips Medical Systems, Best, The Netherlands) was used to acquire MR images using an 8-channel phased array head coil. A paediatrician supervised infant care during MR imaging. Pulse oximetry, temperature, and electrocardiography data were monitored throughout the session. Silicone-based putty (President Putty, Coltene Whaledent, Mahwah, NJ, USA) and neonatal earmuffs (MiniMuffs, Natus Medical Inc., San Carlos, CA, USA) were used for ear protection. Oral chloral hydrate (25–50 mg $kg^{-1}$) was administered to infants whose parents chose sedation for the procedure (87%). High-resolution anatomical images were acquired with T2-weighted fast spin echo sequences (repetition time = 8,670 ms; echo time = 160 ms; flip angle = 90°, slice thickness = 1 mm, field of view = 220 × 220 $mm^2$, voxel size = 0.86 × 0.86 × 1 $mm^3$).

## Tensor Based Morphometry

Following methods described in Vanes et al. (2021) and Lautarescu et al. (2021) T2-weighted (images and tissue type segmentations) were registered to a study-specific template using ANTS software Symmetric Normalisation algorithms (Avants et al., 2011). Resultant nonlinear transformation deformation tensor fields (warps) were used to calculate deformation tensor field gradients (log-Jacobian determinant maps) as a measure of relative brain volume. Greater log-Jacobian values represent the extent of contraction voxels undergo following registration (i.e., larger volumes), while smaller values represent volume reductions (Avants & Gee, 2004). Smoothing with 4 mm full-width half-maximum Gaussian filter was applied.

## Perinatal Socio-Demographic and Clinical Data

### Perinatal Clinical Data

With parental consent, the infant's electronic medical records were accessed using the Standardised Electronic Neonatal Database to collect perinatal socio-demographic and clinical data. Data capturing neonatal clinical risk were collected as part of the larger ePrime study (Edwards et al., 2018), as clinical risk can exacerbate the long-term sequelae of VPT birth (Volpe, 2009). A principal component analysis (PCA) summarised 28 perinatal clinical variables explaining 72% of their variance with a single component, which was labelled 'neonatal sickness index', as previously described in Kanel et al. (2021), The variables with the highest factor loadings were: GA, days on total parenteral nutrition, days on continuous positive airway pressure, days on mechanical ventilation and surfactant administration. Clinical variables were coded so that increased neonatal sickness index values indicate greater clinical risk.

### Perinatal Environmental Data

An Index of Multiple Deprivation (IMD) score was computed from the infant's residential postcode at time of birth (Department for Communities and Local Government, 2011; https://tools.npeu.ox.ac.uk/imd/). The IMD summarises area-level information in 7 domains: income, employment, education, health, crime, housing and living environment. Higher IMD scores reflect increased deprivation in the neighbourhood, hence higher social risk.

## Behavioural and Cognitive Measures

At the 2-year follow-up, toddlers were assessed with the parent-rated M-CHAT. Critical positive M-CHAT screening was defined by failing any 2 out of the 6 critical items: "Does your child take an interest in other children?", "Does your child ever use his/her index finger to point, to indicate

interest in something?", "Does your child ever bring objects over to you to show you something?", "Does your child imitate you?", "Does your child respond to his/her name when you call?", "If you point at a toy across the room, does your child look at it?" (Robins et al., 2001). The definition used by Moore and colleagues (Moore et al., 2012) was used to define 'non-critical' positive screening: failing any 3 or more items, but fewer than two critical items. Toddlers not meeting either of these criteria received a negative M-CHAT screening.

The following measures were used to assess infants' development at 2 years: the Bayley Scales of Infant Development, Third Edition (Bayley-III; Bayley, 2006), which evaluates expressive and receptive language, fine and gross motor skills and composite cognitive scores, and the Parent Report of Children's Abilities Revised (PARCA-R; Johnson et al., 2004; Saudino et al., 1998), which evaluates toddlers' vocabulary and sentence complexity and non-verbal cognitive skills.

To reduce the dimensionality of the behavioural outcome data, a PCA was performed. All Bayley-III and PARCA-R index scores were included in the model and the elbow-method was used to determine the number of principal components explaining most of the variance in the data. A scree plot showing the percentage of variance explained by each principal component (i.e., eigenvalues) suggests an optimal number of 2 principal components (Supplementary Information eFig. SM1), jointly explaining a cumulative 69% of total variance. Pearson correlations between each of the two resultant principal components and individual index scores were used to define each of the components. PC1 correlated negatively with all Bayley-III and PARCA-R items, resulting in a component summarising global (cognitive, language and motor) developmental delay, while PC2 correlated positively with language items (PARCA-R sentence complexity and vocabulary scores and Bayley-III expressive language scores) and showed negative correlations with gross and fine motor Bayley-III scores (Supplementary Information eFig. SM2). The first principal component was labelled as a global 'developmental delay' index and the second as a 'language' index.

At the 4- to 7-year-old follow-up, the Social Responsiveness Scale, Second Edition (SRS-2; Constantino & Gruber, 2012) was administered to measure core ASC symptoms in early childhood; it contains a Social Communication/Interaction (SCI) and a Restricted/Repetitive Behaviour (RRB) subscale. The SCI subscale indexes deficits in behaviours relating to social awareness, cognition, communication, and motivation, and the RRB subscale reflects the severity of restrictive and repetitive patterns of behaviours and interests (Constantino & Gruber, 2012). The SRS-2 shows good internal consistency (Cronbach's alpha = 0.92 and 0.93 for females and males, respectively) as well as construct, convergent and concurrent validity in 5–8-year-old children from the United Kingdom (Wigham et al., 2012).

## Statistical Analyses

### Univariate Phenotypic Group Differences

Statistical analyses were conducted using R (version 3.6.1). Non-parametric Kruskal–Wallis tests compared continuous measures (developmental profiles at 2 years, socio-demographic and clinical profiles at birth and SRS-2 SCI and RRB scores at 4–7 years) between M-CHAT groups (*onewaytests* R package; Dag et al., 2018). For categorical variables (sex), Chi-squared test was used. Post-hoc pairwise comparisons were made for variables showing a significant effect of group ($p < 0.05$). Post-hoc pairwise between-group median differences (for continuous variables) or odds ratios (for categorical variables) were reported and post-hoc pairwise comparison p-values were corrected using False Discovery Rate (Benjamini & Hochberg, 1995). A generalised linear model with 10,000 permutations investigating the effect of M-CHAT group on SCI and RRB scores and correcting for covarying effects of developmental delay, sex, IMD and neonatal sickness index, was also tested (p-permute; https://github.com/lucasfr/grouped_perm_glm).

### Childhood Symptoms Exceeding Clinical Cut-Offs For Autism

Having a total SRS-2 T-score greater than or equal to 76 is considered to be clinically meaningful as it indicates a high likelihood of receiving an ASC diagnosis (Constantino & Gruber, 2012). We calculated the number of children scoring above the SRS-2 clinical cut-off within each M-CHAT group. Sample size calculations were then performed in order to ascertain whether the sample size was adequate for predictive validity analyses (Linden, 2020). The following measures were used as inputs in the sample size calculation: expected sensitivity/specificity (52%/84% respectively; Kim et al., 2016), prevalence in current sample (2%) and confidence interval for estimates (95%-CI with CI-width = 0.1).

### Mass-Univariate Group Differences in Brain Volume

Differences in voxel-wise volume (log-Jacobian) measures at term-equivalent age between the three M-CHAT screening groups were investigated using general linear models correcting for sex, PMA, IMD and neonatal sickness index. FMRIB Software Library (FSL)'s *randomise* function with 10,000 permutations per run was used for non-parametric permutation testing with Threshold-Free

Cluster Enhancement and controlled for family-wise error rate. Significance was set at $p < 0.05$ per contrast, given the exploratory nature of the analysis.

Post-hoc analyses investigating associations between neonatal brain volumes showing between-group differences and ASC traits in childhood are described in the supplemental information (Table SM2). We also explored associations between M-CHAT total items failed and neonatal whole-brain Jacobian values.

#### Testing the Role of Developmental Delay

To test for a potential role of early developmental delay in explaining (mediating) or exacerbating (moderating) later group differences in core ASC symptoms, analyses using general linear models were conducted.

Specifically, where between-group differences in later ASC symptoms (SRS-2 SCI or RRB) at 4–7 years were observed, we tested whether these differences were significantly mediated by developmental delay at 2 years. In addition, to test whether developmental delay at 2 years shows a differential relationship with later ASC symptoms in the separate M-CHAT groups, we tested for effects of developmental delay and M-CHAT screening, as well as their interaction, on SRS-2 SCI and RRB scores. Both mediation and moderation analyses used sex, IMD, and neonatal sickness index as confounders. Mediation was tested via bootstrapping of the indirect effect (based on 5000 bootstrap samples) using the R 'mediation' package (Tingley et al., 2014). To adjust for multiple comparisons due to two separate outcome variables (SRS-2 RRB and SCI), 97.5%-confidence intervals (97.5%-CIs) were generated. P-values with a corrected significance threshold of $p < 0.05/2$ (i.e., 0.025) were estimated from non-parametric permutation testing with 10,000 permutations (p-permute; https://github.com/lucasfr/grouped_perm_glm).

## Results

### Comparing M-CHAT Groups on Socio-Demographic, Clinical and Developmental Outcomes

Median scores and F-statistics and p-values comparing M-CHAT group socio-demographic and clinical outcomes are summarised in Table 2 and developmental profiles and Bayley-III and PARCA-R composite scores in Table 3. The three groups did not differ in corrected age at M-CHAT assessment, PMA at scan, GA at birth, birthweight, neonatal sickness index or language development (Tables 2; 3).

Variables showing a significant group effect were investigated for pairwise group differences and median difference and post-hoc p-values for between-group differences are reported in Table 4. In summary, social risk (IMD scores) was lower in negative M-CHAT scorers than critical positive scorers, but did not differ between other groups. Of the three groups, negative M-CHAT scorers had the lowest developmental delay scores (indicating better language, cognitive and motor scores), the critical positive scorers showed the greatest developmental delay and non-critical positive scorers showed intermediate developmental delay scores. There was an overall difference in male-to-female ratios between the different M-CHAT sub-groups (Chi-squared = 7.38; $p = 0.025$), although all pairwise comparisons were not statistically significant ($p > 0.05$; M-CHAT negative group compared to non-critical and critical groups; odds ratio = 1.54 and 2.74, $p = 0.147$ and 0.053, respectively;

**Table 2** Socio-demographic and clinical profiles for M-CHAT groups

| Variable | Median (Interquartile range) | | | F-statistic; p-value |
|---|---|---|---|---|
| | Negative (N = 267; 143 female) | Non-critical positive (N = 77; 33 female) | Critical positive (N = 27; 8 female) | |
| Socio-demographic variables | | | | |
| Corrected age at 2 years, months median (IQR) | 20.20 (0.67) | 20.13 (0.70) | 20.03 (0.37) | F = 2.11; p = 0.310 |
| Corrected age at 4–7 years, years | 4.59 (0.58) | 4.67 (0.90) | 4.59 (0.91) | F = 4.78; p = 0.092 |
| PMA at scan, weeks | 42.57 (2.00) | 42.71 (2.14) | 42.57 (1.50) | F = 5.27; p = 0.072 |
| IMD score at birth | 16.54 (17.00) | 19.92 (15.71) | 25.87 (15.79) | F = 7.63; p = 0.022* |
| GA, weeks | 30.29 (3.50) | 30.86 (4.14) | 28.86 (3.36) | F = 3.58; p = 0.167 |
| Birthweight, grams | 1315 (570.00) | 1270 (650.00) | 1040 (485.00) | F = 3.27; p = 0.196 |
| Neonatal sickness index[a] | −0.36 (1.71) | −0.45 (1.49) | 0.55 (1.59) | F = 3.91; p = 0.142 |

*GA* gestational age at birth, *IMD* index multiple deprivation, *PMA* postmenstrual age

*$p < 0.05$

[a]excluding one subject with incomplete clinical data

**Table 3** Developmental profiles and Bayley-II and PARCA-R composite scores for M-CHAT groups

| Variable | Median (Interquartile range) | | | Statistic; p-value |
|---|---|---|---|---|
| | Negative (N = 267) | Non-critical positive (N = 77) | Critical positive (N = 27) | |
| Developmental profiles at 2 years[b] | | | | |
| Developmental delay | −0.57 (2.67) | 0.88 (2.25) | 2.86 (2.76) | F = 57.40; p < 0.001*** |
| Language | −0.02 (1.28) | −0.09 (1.28) | 0.28 (1.54) | F = 2.33; p = 0.313 |
| Bayley-III composite scores at 2 years[b] | | | | |
| Cognitive | 95.00 (18.75) | 90.00 (20.00) | 82.50 (20.00) | F = 27.28, p < 0.001*** |
| Language | 97.00 (20.00) | 83.00 (20.00) | 69.50 (14.00) | F = 45.36, p < 0.001*** |
| Motor | 100.00 (9.00) | 94.00 (12.00) | 82.00 (18.00) | F = 45.14, p < 0.001*** |
| PARCA-R composite scores at 2 years[b] | | | | |
| Vocabulary | 19.00 (18.05) | 10.00 (11.25) | 3.00 (5.75) | F = 34.52; p < 0.001*** |
| Sentence complexity | 5.00 (6.00) | 3.00 (4.25) | 4.00 (2.00) | F = 36.21; p < 0.001*** |

*IMD* index multiple deprivation, *PARCA-R* parent report of children's abilities-revised

***p < 0.001

[b]excluding 31 subjects with incomplete developmental data at 2-years

**Table 4** Post-hoc pairwise differences (between the three M-CHAT screening groups) for variables with a significant effect of M-CHAT group

| Variable | Median difference (p-value) | | |
|---|---|---|---|
| | Negative vs non-critical positive | Negative vs critical positive | Non-critical positive vs critical positive |
| IMD score at birth | −3.38 (p = 0.646) | −9.33 (p = 0.020)* | −5.95 (p = 0.039)* |
| Developmental delay | 5.00 (p = 0.003)** | −3.43 (p < 0.001)*** | −1.98 (p < 0.001)*** |
| Bayley-III: Cognitive | 14.00 (p < 0.001)*** | 12.50 (p < 0.001)*** | 7.50 (p = 0.006)** |
| Bayley-III: Language | 6.00 (p < 0.001)*** | 27.50 (p < 0.001)*** | 13.50 (p < 0.001)*** |
| Bayley-III: Motor | 9.00 (p = 0.001)*** | 18.00 (p < 0.001)*** | 12.00 (p < 0.001)*** |
| PARCA-R: Vocabulary | 2.00 (p < 0.001)*** | 16.00 (p < 0.001)*** | 7.00 (p = 0.003)** |
| PARCA-R: Sentence complexity | −1.45 (p < 0.001)*** | 4.00 (p < 0.001)*** | 2.00 (p = 0.010)** |

Between-group statistics (median differences for variables with significant effects of M-CHAT group) and pairwise comparison p-values are reported for variables showing significant effects of M-CHAT group

*IMD* index multiple deprivation, *PARCA-R* parent report of children's abilities-revised

*p < 0.05; **p < 0.010; ***p < 0.001

non-critical group compared to the critical group; odds ratio = 1.78, p = 0.226). The proportion of females in the M-CHAT negative, non-critical positive and critical positive groups were 53.56%, 42.86% and 29.63% respectively.

## Differences In Brain Volume at Term-Equivalent Age Between M-CHAT Groups

Voxel-wise group comparisons of relative brain volume (correcting for sex, PMA, IMD and neonatal sickness index) showed that critical positive scorers had reduced regional volume in the bilateral deep cerebellar nuclei, middle cerebellar peduncles and midbrain and medulla regions of the brainstem compared to negative scorers (Fig. 1A). Critical positive scorers also showed volume reductions in an overlapping region in the right cerebellar nuclei compared to the non-critical positive group (Fig. 1B). Coloured T-statistic maps of regions showing significant differences between critical and negative scorers are depicted in Fig. 1A and between critical and non-critical scorers in Fig. 1B, where T-statistic values ranging from 1.70 to 4.70 are denoted by the colour bar. Non-parametric permutation tests with Threshold-Free Cluster Enhancement controlling family-wise error rate were used to identify between-group differences (p < 0.05).

There were no significant associations between regional cerebellar volumes and ASC traits at 4–7 years of age in

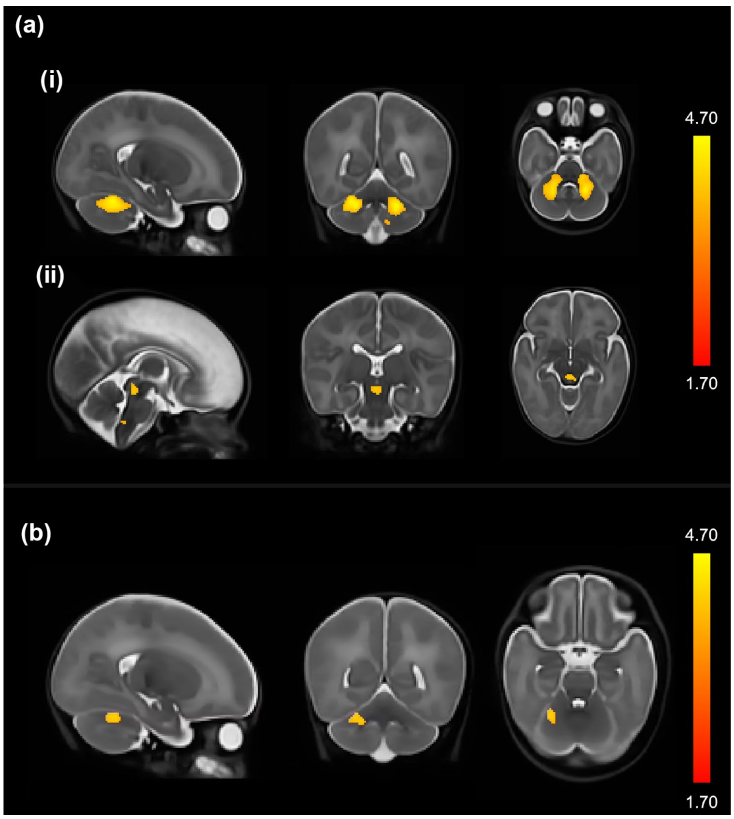



**Fig. 1** Study-specific brain template overlaid with coloured T-statistics map of brain regions significantly smaller in the M-CHAT critical positive group compared to **a** the M-CHAT negative group and **b** the M-CHAT non-critical positive group

any of the three groups (Table SM2). Furthermore, when investigating the association between M-CHAT total items failed and neonatal whole-brain Jacobians values, we found no significant correlations ($p > 0.05$).

## ASC Traits in Childhood

A significant effect of group on SRS-2 SCI and RRB was observed (Table 5). Pairwise comparisons showed that both M-CHAT (critical and non-critical) positive groups had higher SCI and RRB scores compared to the negative group; however, SCI and RRB scores did not differ between the two

**Table 5** ASC traits at 4–7 years in the M-CHAT screening groups

| Variable | Median (Interquartile range) | | | F-statistic; p-value |
|---|---|---|---|---|
| | Negative (N = 130) | Non-critical positive (N = 32) | Critical positive (N = 15) | |
| SRS-2 SCI | 45.00 (9.50) | 49.50 (10.00) | 55.00 (17.00) | $F = 17.69$; $p < 0.001$*** |
| SRS-2 RRB | 4.00 (5.00) | 5.50 (7.25) | 11.00 (12.50) | $F = 14.02$; $p < 0.001$*** |

*RRB* restricted interests and repetitive behaviours, *SCI* social communication/interaction, *SRS-2* social responsiveness scale, second edition

***$p < 0.001$

Appendix A



**Table 6** Post-hoc pairwise differences (between the three M-CHAT screening groups) for ASC traits with a significant effect of M-CHAT group

| Variable | Median difference (p-value) | | |
|---|---|---|---|
| | Negative vs non-critical | Negative vs critical | Non-critical vs critical |
| SRS-2 SCI | −4.50 (p = 0.006)** | −10.00 (p = 0.001)*** | −5.50 (p = 0.133) |
| SRS-2 RRB | −1.50 (p = 0.020)* | −7.00 (p = 0.005)** | −5.50 (p = 0.122) |

Between-group statistics (median differences for variables with significant effects of M-CHAT group) and pairwise comparison p-values are reported for SRS-2 ASC trait outcomes showing significant effects of M-CHAT group

*RRB* restricted interests and repetitive behaviours, *SCI* social communication/interaction, *SRS-2* social responsiveness scale, second edition

*p < 0.05; **p < 0.01; ***p < 0.001

positive groups (Table 6; Fig. 2A). These findings did not change after adjusting for sex, IMD, neonatal sickness index and developmental delay.

5 children out 177 (2.8%) had SRS-2 scores exceeding clinical cut-offs for autism (i.e., having SRS-2 total T-scores greater than or equal to 76), where 2 belonged to the non-critical positive group and 3 belonged to the critical positive group. Formal predictive validity analyses were not performed, as sample size analyses estimated a larger sample (N = 480) would be needed to carry them out.

## Mediating and Moderating Effects of Developmental Delay on ASC Traits

### Mediation Analyses

Due to the significant differences observed in both SRS-2 SCI and RRB childhood scores between negative scorers and the two positive groups, we tested whether pairwise group differences were at least partially accounted for by developmental delay. Developmental delay significantly partially mediated differences in SCI when comparing negative to critical (indirect effect 97.5%-CI = 1.69, 8.46; p < 0.001) and non-critical positive groups (indirect effect 97.5%-CI = 0.22, 2.65; p = 0.005; Fig. 2Bi). Proportion mediated (Prop.med) was 0.18 for M-CHAT negative vs non-critical positive group, and 0.38 for M-CHAT negative vs critical positive group.

Developmental delay also significantly partially mediated group differences in RRB when comparing the negative to the critical positive (indirect effect 97.5%-CI = 1.29, 8.92; p = 0.002; Prop.med = 0.36), but not to the non-critical positive group (indirect effect 97.5%-CI = -0.39, 2.29; p = 0.138; Prop.med = 0.18; Fig. 2Bii). Mediation analyses for the two positive groups were not conducted, as these did not differ significantly in SCI or RRB scores.

### Moderation Analyses

A linear model regressing SCI scores on M-CHAT grouping, developmental delay, and their interaction (M-CHAT × developmental delay), controlling for sex, IMD and neonatal sickness index, found no significant interaction, $F(2, 159) = 2.73$, $p = 0.069$; p-permute = 0.074, indicating that the effect of developmental delay on SCI scores was similar in the three M-CHAT groups.

In contrast, a model regressing RRB scores on M-CHAT grouping, developmental delay, and their interaction (M-CHAT × developmental delay), controlling for sex, IMD, and neonatal sickness index, revealed a significant overall interaction, $F(2,159) = 6.73$, $p = 0.002$; p-permute = 0.003. Re-coding each group as the reference category showed this was due to a significant interaction when comparing the critical positive group to both negative and non-critical positive groups (Table 7). The M-CHAT critical positive group had a stronger (positive) association between developmental delay and RRB scores compared to both negative and non-critical positive groups (Fig. 2C).

## Discussion

This study investigated neonatal brain volumes and ASC traits in childhood in VPT children sub-divided into three groups, based on their M-CHAT screening outcomes (negative, non-critical positive and critical positive). Addressing our first aim, we found that the three groups exhibited differences in structural brain volumes at term-equivalent age, indicating distinct early biological phenotypes. The critical positive scorers displayed smaller volumes in cerebellar and brainstem regions compared to negative scorers, and smaller regional cerebellar volumes compared to non-critical positive scorers. Addressing our second aim, we found that while both positive groups showed higher ASC core symptom scores (RRB and SCI) relative to negative scorers, there were no significant differences between the two positive groups. However, the critical positive scorers

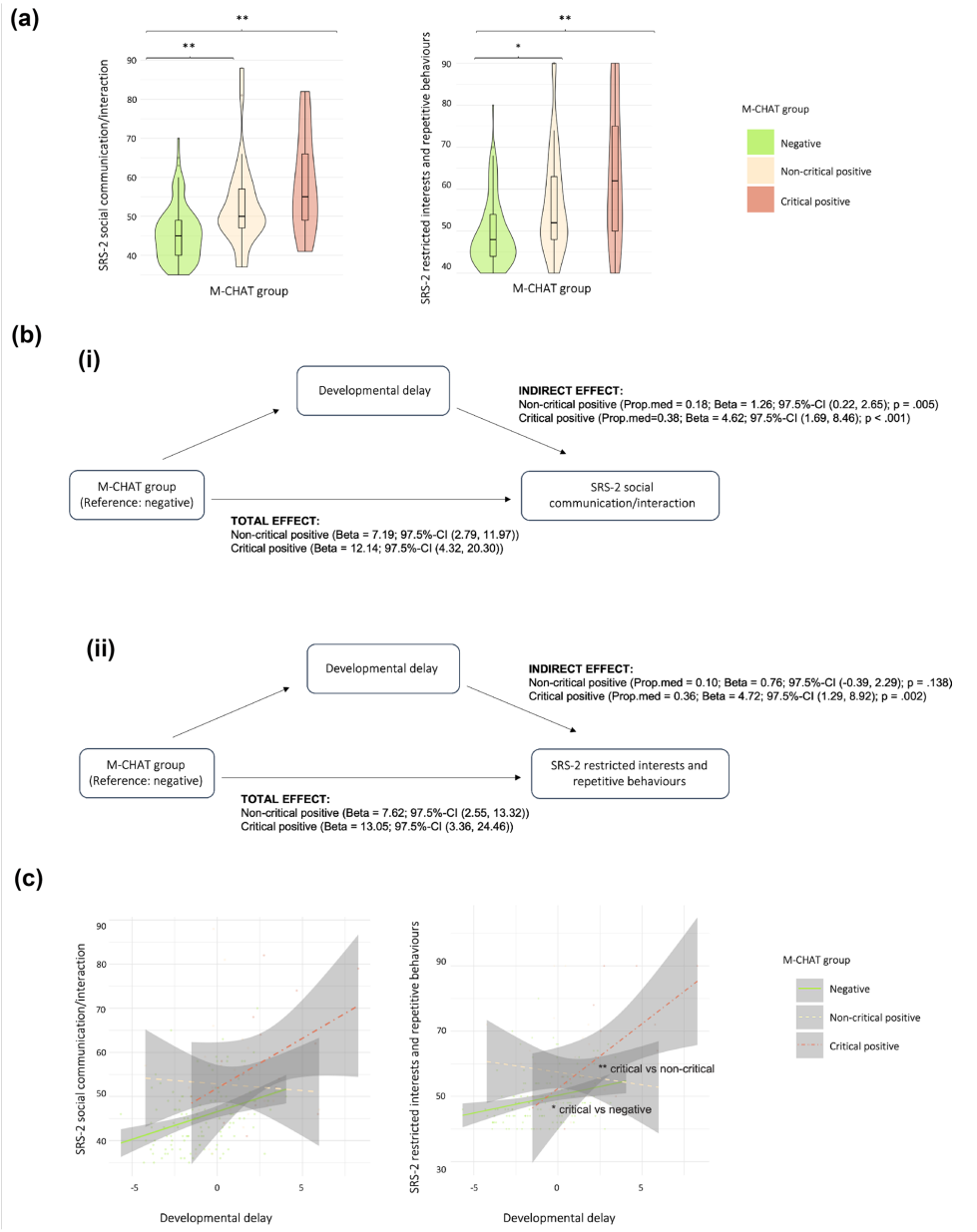



**Fig. 2** **a** SRS-2 SCI/RRB median differences between M-CHAT screening groups, **b** the mediating effect of developmental delay on the relationship between M-CHAT and SCI/RRB and **c** the moderating effect of the M-CHAT group × developmental delay interaction on SCI/RRB. *p <0.025; **p < 0.010

**Table 7** M-CHAT x developmental delay interaction on SRS-2 RRB scores

| Interaction term | Beta | SE | T-statistic | 97.5%-CI | Permutation p-value |
|---|---|---|---|---|---|
| M-CHAT (non-critical positive vs negative) × developmental delay | −2.00 | 1.00 | −2.01 | (−4.25, 0.26) | 0.047 |
| M-CHAT (critical positive vs negative) × developmental delay | 2.95 | 1.11 | 2.66 | (0.44, 5.46) | 0.013* |
| M-CHAT (critical positive vs non-critical positive) × developmental delay | 4.95 | 1.35 | 3.66 | (1.89, 8.00) | 0.001*** |

Table summarising, beta, standard error (SE), T-statistic, 97.5% confidence intervals (97.5%-CI) and non-parametric permutation testing p-values for effect of interaction terms between M-CHAT group and developmental delay on RRB scores

*RRB* restricted interests and repetitive behaviours, *SE* standard error, *SRS-2* social responsiveness scale, second edition

*p < 0.025; ***p < 0.001

showed greater developmental delay compared to the other two groups. Taken together our findings suggest that the two M-CHAT positive groups do not differ in the severity of childhood ASC traits and we speculate that they may be following distinct aetiological trajectories leading to similar ASC traits in childhood (i.e., equifinality; Cicchetti & Rogosch, 1996).

The early differences in regional brain volumes found between the positive M-CHAT groups, provide evidence for potentially distinct biological mechanisms underlying later ASC outcomes in a subset of VPT children. The critical positive M-CHAT group showed reduced relative volumes within regions of the right cerebellar nuclei compared to the non-critical positive group, and more widespread reductions in bilateral cerebellar nuclei and brainstem (medulla oblongata and midbrain) volumes compared to the negative group. The cerebellum is known to play a critical role in coordinating motor, sensory and cognitive abilities, which are also impacted in ASC (Wang et al., 2014). Cerebellar alterations have been associated with ASC symptomatology/traits both in animal and human studies. Cellular cerebellar pathology has been linked to increased ASC-like behaviours in mice (Tsai et al., 2012), smaller white matter volume in the cerebellum has been described in adults with ASC (Toal et al., 2010) and number and density of Purkinje cells has been shown *post-mortem* to be altered in individuals with ASC (Wegiel et al., 2010, 2014). In VPT samples, cerebellar volume reductions in childhood (Ure et al., 2016) and increased cerebellar haemorrhagic injury in infancy (Limperopoulos et al., 2007) were displayed in those with an ASC diagnosis or those screening positively on the M-CHAT. In both studies, VPT children with ASC diagnoses (Ure et al., 2016) and with cerebellar injury (Limperopoulos et al., 2007) had a high prevalence of developmental delay. Similar to the results of the aforementioned studies, which show cerebellar volume reductions in groups of children with increased developmental delay, we also found that the group exhibiting the most severe developmental delay (i.e., M-CHAT critical positive group) had smaller cerebellar volumes relative to the non-critical positive and negative groups.

The brainstem, which in this study showed reduced regional volumes in the M-CHAT critical positive relative to the M-CHAT negative group, is an early phylogenetic region of the brain known to be important for primitive functions such as arousal, respiration, and physiological regulation, although there is some evidence of its role in self-regulatory behaviours (Geva & Feldman, 2008; Geva et al., 2014). Of particular relevance to the current findings, Geva et al. (2013) showed that brainstem functioning in VPT infants was associated with social integration abilities assessed using modulation of gaze in response to social stimuli at 4 months. Furthermore, white matter reductions in the brainstem have been observed in adults with ASC compared to controls (Toal et al., 2010) and early histological work investigating brainstem injury, specifically in the motor cranial nerve nuclei, suggest that early alterations to this brain region may contribute to the onset of autism later in life (Rodier, 2002; Rodier et al., 1996, 1997). The cerebellar nuclei and brainstem (medulla oblongata and midbrain) interact with one another to facilitate sensory, motor and regulatory processes (Watson et al., 2013). The olivary complex in the medulla sends fibres to the cerebellar nuclei allowing for integration of motor and sensory information and has been found to be altered *post-mortem* in individuals with ASC (Wegiel et al., 2013). Interactions between the midbrain and the olivary-cerebellar complex have been discussed in the context of processes relating to "survival networks", which involve behavioural (social, motor and sensory) regulation in response to emotional and environmental stimuli (Watson et al., 2013), which are core processes in ASC symptomatology. In light of these findings, we tentatively speculate that the regional brain alterations we observed in the M-CHAT critical positive compared to the negative group may represent a biological mechanism contributing to the increased RRB and SCI behaviours seen in this group.

Findings showing neonatal regional brain volume reductions as well as increased developmental delay observed in critical compared to non-critical positive scorers, despite the two groups showing similar childhood ASC traits (SCI/RRB), probed us to further investigate developmental delay

in relation to ASC traits in the different groups. Results showed that developmental delay had both an explanatory (i.e., mediating) effect, as well as an exacerbating role (i.e., moderating effect) specific to RRB scores, in the critical positive group (but not SCI scores). These results suggest that VPT toddlers meeting the critical positive M-CHAT criteria may, therefore, represent an aetiologically distinct subgroup of children whose developmental difficulties increase their likelihood of developing RRB symptoms. Differences in RRB traits between preterm and term-born children have been previously explained by differences in IQ (Johnson et al., 2010), further supporting the notion that developmental delay may contribute to elevated childhood RRB traits. However, it is worth noting that in our study RRB traits were only partially explained by developmental delay, as the higher childhood RRB scores in M-CHAT critical positive scorers compared to negative and non-critical positive scorers were significant after correcting for developmental delay.

The two M-CHAT positive screening groups did not differ in SCI scores, but had elevated SCI scores relative to the negative screening group, which were significant even after correcting for developmental delay. This indicates that developmental delay at least partially contributes to the SCI difficulties seen in both M-CHAT positive groups, which is in line with observations in children with ASC (Hirosawa et al., 2020). However, developmental delay in the current study did not moderate the relationship between M-CHAT group and SCI difficulties, suggesting that the effect of developmental delay on subsequent SCI outcomes was similar in all three groups. These results motivate future studies to investigate which additional biological and/or environmental factors could be driving similar SCI outcomes in the two positive groups, who showed distinct neurodevelopmental profiles early in life.

This study's findings tentatively suggest that the M-CHAT in VPT toddlers represents a useful tool to identify individuals with an increased likelihood of displaying ASC traits in childhood. This is firstly supported by findings showing increased developmental difficulties in both M-CHAT positive groups compared to the negative group, as well as higher median RRB and SCI scores, even after accounting for developmental delay. Secondly, as all children scoring above SRS-2 clinical cut-off thresholds (N = 5, or 2.8% of the sample) belonged to both M-CHAT positive groups, this study suggests that the tool has high sensitivity in VPT cohorts. Finally, although most positive scorers did not exceed the SRS-2 clinical cut-off score for ASC, they did exhibit subthreshold socio-emotional difficulties which are reportedly common amongst VPT children (Johnson & Marlow, 2011).

This study has several limitations, the main being that ASC diagnoses were not systematically evaluated at childhood assessment (4–7 years), although a current follow-up study is now collecting these data at 8–9 years. Moreover, sample size analyses showed we did not have an adequate number of participants to perform formal predictive validity analyses, as the number of children in our sample exceeding SRS-2 clinical cut-off scores were very few. Another limitation of this study is that the results presented are not generalisable to children with major brain lesions, who are likely to have more severe developmental impairments later in life (Volpe, 2009), but were not included in the current analyses. Future studies could therefore focus on better understanding the relationship between developmental delay following major brain injury and later ASC behaviours/traits. In addition, other neuroimaging modalities measuring brain functional and structural connectivity were not investigated, and future studies could use a multi-modal approach to provide greater insight into the biological underpinnings associated with the distinct pathways to increased likelihood of developing ASC following VPT birth. Furthermore, while in this paper we consider separate M-CHAT groups, it is plausible that the three groups may lie on a continuum. The non-critical positive scorers' developmental outcomes were in fact intermediate between the two other groups, with the negative scorers showing the best outcomes and the critical positive scorers showing the poorest outcomes.

In summary, our results highlight the distinct early developmental and neurobiological characteristics in M-CHAT critical versus non-critical positive scorers, despite them presenting with similar childhood ASC-symptom profiles. Our results also further highlight the importance of interpreting M-CHAT screenings in combination with other developmental measures when assessing VPT toddlers. Identifying biomarkers and developmental trajectories of later ASC outcomes could guide clinicians and researchers to devise personalised interventions aimed at supporting children's development based on their distinct phenotypic presentations preceding the onset of ASC symptoms.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s10803-022-05789-4.

**Acknowledgments** The authors would like to thank the participating families of the ePrime study, and all research, radiography and clinical staff involved in the study.

**Authors Contribution** All authors contributed to the study conception and design. Material preparation and data analysis were performed by Laila Hadaya, Lucy Vanes, Vyacheslav Karolis, Chiara Nosarti, Dana Kanel and Marguerite Leoni. Data collection at baseline, at the 2-year- and 4–7-year-old follow up was completed by the e-Prime study research team supervised by David Edwards and Chiara Nosarti. The first draft of the manuscript was written by Laila Hadaya and Chiara Nosarti and was commented on by all authors of the manuscript. The final manuscript was read and approved by all authors.

**Funding** This work was financially supported by: The Medical Research Council, UK [Grant numbers: MR/K006355/1 and MR/S026460/1] and Action Medical Research and Dangoor Education [Grant number: GN2606]. Data analysed in this study was collected during independent research funded by the National Institute for Health Research (NIHR) Programme Grants for Applied Research Programme [Grant number: RP-PG-0707-10154] and supported by NIHR Biomedical Research Centre, Guy's and St Thomas' NHS Foundation Trust and King's College London, the NIHR Clinical Research Facility, Guy's and St Thomas' and the MRC Centre for Neurodevelopmental Disorders.

## Declarations

**Conflict of Interest** The authors have no financial or non-financial conflicts of interest to disclose.

**Ethics Approval** Ethical approval for the study was granted by the Hammersmith and Queen Charlotte's Research Ethic Committee (09/H0707/98) and the National Research Ethics Committee (14/LO/0677) and the study was conducted in accordance with ethical standards of the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards.

**Data Availability** Access to the dataset supporting this article can be made available upon request from the corresponding author.



## References

Agrawal, S., Rao, S. C., Bulsara, M. K., & Patole, S. K. (2018). Prevalence of autism spectrum disorder in preterm infants: a meta-analysis. *Pediatrics, 142*(3), e20180134. https://doi.org/10.1542/peds.2018-0134

Alcalá-López, D., Smallwood, J., Jefferies, E., Van Overwalle, F., Vogeley, K., Mars, R. B., Turetsky, B. I., Laird, A. R., Fox, P. T., Eickhoff, S. B., & Bzdok, D. (2018). Computing the social brain connectome across systems and states. *Cerebral Cortex, 28*(7), 2207–2232. https://doi.org/10.1093/cercor/bhx121

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders, Fifth Edition, DSM-5*. USA: American Psychiatric Association.

Avants, B., & Gee, J. C. (2004). Geodesic estimation for large deformation anatomical shape averaging and interpolation. *NeuroImage, 23*(Suppl 1), S139-150. https://doi.org/10.1016/j.neuroimage.2004.07.010

Avants, B. B., Tustison, N. J., Song, G., Cook, P. A., Klein, A., & Gee, J. C. (2011). A Reproducible evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage, 54*(3), 2033–2044. https://doi.org/10.1016/j.neuroimage.2010.09.025

Ball, G., Aljabar, P., Arichi, T., Tusor, N., Cox, D., Merchant, N., Nongena, P., Hajnal, J. V., Edwards, A. D., & Counsell, S. J. (2016). Machine-learning to characterise neonatal functional connectivity in the preterm brain. *NeuroImage, 124*(Pt A), 267–275. https://doi.org/10.1016/j.neuroimage.2015.08.055

Ball, G., Boardman, J. P., Aljabar, P., Pandit, A., Arichi, T., Merchant, N., Rueckert, D., Edwards, A. D., & Counsell, S. J. (2013). The influence of preterm birth on the developing thalamocortical connectome. *Cortex, 49*(6), 1711–1721. https://doi.org/10.1016/j.cortex.2012.07.006

Bayley, N. (2006). *Bayley scales of infant and toddler development* (3rd ed.). San Antonio: Harcourt Assessment Inc.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B, 57*(1), 289–300.

Blencowe, H., Lee, A. C., Cousens, S., Bahalim, A., Narwal, R., Zhong, N., Chou, D., Say, L., Modi, N., Katz, J., Vos, T., Marlow, N., & Lawn, J. E. (2013). Preterm birth–associated neurodevelopmental impairment estimates at regional and global levels for 2010. *Pediatric Research, 74*(1), 17–34. https://doi.org/10.1038/pr.2013.204

Chlebowski, C., Robins, D. L., Barton, M. L., & Fein, D. (2013). Large-scale use of the modified checklist for Autism in low-risk toddlers. *Pediatrics, 131*(4), e1121–e1127. https://doi.org/10.1542/peds.2012-1525

Ciarrusta, J., O'Muircheartaigh, J., Dimitrova, R., Batalle, D., Cordero-Grande, L., Price, A., Hughes, E., Steinweg, J. K., Kangas, J., Perry, E., Javed, A., Stoencheva, V., Akolekar, R., Victor, S., Hajnal, J., Murphy, D., Edwards, D., Arichi, T., & McAlonan, G. (2019). Social brain functional maturation in newborn infants with and without a family history of autism spectrum disorder. *JAMA Network Open, 2*(4), e191868. https://doi.org/10.1001/jamanetworkopen.2019.1868

Cicchetti, D., & Rogosch, F. A. (1996). Equifinality and multifinality in developmental psychopathology. *Development and Psychopathology, 8*(4), 597–600. https://doi.org/10.1017/S0954579400007318

Constantino, J. N., & Gruber, C. P. (2012). *Social responsiveness scale second edition (srs-2): manual*. Torrance: Western Psychological Services.

Dag, O., Dolgun, A., Konar, N., & Meric. (2018). onewaytests: An R package for one-way tests in independent groups designs. *The R Journal, 10*(1), 175. https://doi.org/10.32614/RJ-2018-022

Edwards, A. D., Redshaw, M. E., Kennea, N., Rivero-Arias, O., Gonzales-Cinca, N., Nongena, P., Ederies, M., Falconer, S., Chew, A., Omar, O., Hardy, P., Harvey, M. E., Eddama, O., Hayward, N., Wurie, J., Azzopardi, D., Rutherford, M. A., & Counsell, S. (2018). Effect of MRI on preterm infants and their families: a randomised trial with nested diagnostic and economic evaluation. *Archives of Disease in Childhood Fetal and Neonatal Edition, 103*(1), F15–F21. https://doi.org/10.1136/archdischild-2017-313102

Eklöf, E., Mårtensson, G. E., Ådén, U., & Padilla, N. (2019). Reduced structural brain asymmetry during neonatal life is potentially related to autism spectrum disorders in children born extremely preterm. *Autism Research, 12*(9), 1334–1343. https://doi.org/10.1002/aur.2169

Fenoglio, A., Georgieff, M. K., & Elison, J. T. (2017). Social brain circuitry and social cognition in infants born preterm. *Journal of Neurodevelopmental Disorders, 9*, 27. https://doi.org/10.1186/s11689-017-9206-9

Gandhi, T., & Lee, C. C. (2021). Neural mechanisms underlying repetitive behaviors in rodent models of autism spectrum disorders. *Frontiers in Cellular Neuroscience, 14*, 463. https://doi.org/10.3389/fncel.2020.592710

Geva, R., & Feldman, R. (2008). A neurobiological model for the effects of early brainstem functioning on the development of behavior and emotion regulation in infants: Implications for

prenatal and perinatal risk. *Journal of Child Psychology and Psychiatry, 49*(10), 1031–1041. https://doi.org/10.1111/j.1469-7610.2008.01918.x

Geva, R., Schreiber, J., Segal-Caspi, L., & Markus-Shiffman, M. (2014). Neonatal brainstem dysfunction after preterm birth predicts behavioral inhibition. *Journal of Child Psychology Psychiatry, 55*(7), 802–810. https://doi.org/10.1111/jcpp.12188

Geva, R., Sopher, K., Kurtzman, L., Galili, G., Feldman, R., & Kuint, J. (2013). Neonatal brainstem dysfunction risks infant social engagement. *Social Cognitive and Affective Neuroscience, 8*(2), 158–164. https://doi.org/10.1093/scan/nsr082

Ha, S., Sohn, I.-J., Kim, N., Sim, H. J., & Cheon, K.-A. (2015). Characteristics of brains in autism spectrum disorder: structure, function and connectivity across the lifespan. *Experimental Neurobiology, 24*(4), 273–284. https://doi.org/10.5607/en.2015.24.4.273

Healy, E., Reichenberg, A., Nam, K. W., Allin, M. P. G., Walshe, M., Rifkin, L., Murray, S. R. M., & Nosarti, C. (2013). Preterm birth and adolescent social functioning-alterations in emotion-processing brain areas. *The Journal of Pediatrics, 163*(6), 1596–1604. https://doi.org/10.1016/j.jpeds.2013.08.011

Hirosawa, T., Kontani, K., Fukai, M., Kameya, M., Soma, D., Hino, S., Kitamura, T., Hasegawa, C., An, K., Takahashi, T., Yoshimura, Y., & Kikuchi, M. (2020). Different associations between intelligence and social cognition in children with and without autism spectrum disorders. *PLoS ONE, 15*(8), e0235380. https://doi.org/10.1371/journal.pone.0235380

Johnson, S., Hollis, C., Kochhar, P., Hennessy, E., Wolke, D., & Marlow, N. (2010). Autism spectrum disorders in extremely preterm children. *The Journal of Pediatrics, 156*(4), 525-531.e2. https://doi.org/10.1016/j.jpeds.2009.10.041

Johnson, S., Marlow, D. M., Wolke, D., Davidson, L., Marston, L., O'Hare, A., Peacock, J., & Schulte, J. (2004). Validation of a parent report measure of cognitive development in very preterm infants. *Developmental Medicine and Child Neurology, 46*(6), 389–397. https://doi.org/10.1017/S0012162204000635

Johnson, S., & Marlow, N. (2011). Preterm birth and childhood psychiatric disorders. *Pediatric Research, 69*(5 Pt 2), 11R-R18. https://doi.org/10.1203/PDR.0b013e318212faa0

Joseph, R. M., O'Shea, T. M., Allred, E. N., Heeren, T., Hirtz, D., Paneth, N., Leviton, A., & Kuban, K. C. K. (2017). Prevalence and associated features of autism spectrum disorder in extremely low gestational age newborns at age 10 years. *Autism Research, 10*(2), 224–232. https://doi.org/10.1002/aur.1644

Kanel, D., Vanes, L. D., Ball, G., Hadaya, L., Falconer, S., Counsell, S. J., Edwards, A. D., & Nosarti, C. (2022). Neonatal amygdala resting-state functional connectivity and socio-emotional development in very preterm children. *Brain Communications, 4*(1), fcac009. https://doi.org/10.1093/braincomms/fcac009

Kanel, D., Vanes, L., Pecheva, D., Hadaya, L., Falconer, S., Counsell, S., Edwards, D., & Nosarti, C. (2021). Neonatal white matter microstructure and emotional development during the preschool years in children who were born very preterm. *Eneuro*. https://doi.org/10.1523/ENEURO.0546-20.2021

Kim, S. H., Joseph, R. M., Frazier, J. A., O'Shea, T. M., Chawarska, K., Allred, E. N., Leviton, A., & Kuban, K. K. (2016). Predictive validity of the Modified Checklist for Autism in Toddlers (M-CHAT) born very preterm. *The Journal of Pediatrics, 178*, 101–107. https://doi.org/10.1016/j.jpeds.2016.07.052

Kleinman, J. M., Robins, D. L., Ventola, P. E., Pandey, J., Boorstein, H. C., Esser, E. L., Wilson, L. B., Rosenthal, M. A., Sutera, S., Verbalis, A. D., Barton, M., Hodgson, S., Green, J., Dumont-Mathieu, T., Volkmar, F., Chawarska, K., Klin, A., & Fein, D. (2008). The modified checklist for autism in toddlers: a follow-up study investigating the early detection of autism spectrum disorders. *Journal of Autism and Developmental Disorders, 38*(5), 827–839. https://doi.org/10.1007/s10803-007-0450-9

Kuban, K. C. K., O'Shea, T. M., Allred, E. N., Tager-Flusberg, H., Goldstein, D. J., & Leviton, A. (2009). Positive screening on the modified checklist for autism in toddlers (M-CHAT) in extremely low gestational age newborns. *The Journal of Pediatrics, 154*(4), 535-540.e1. https://doi.org/10.1016/j.jpeds.2008.10.011

Lautarescu, A., Hadaya, L., Craig, M. C., Makropoulos, A., Batalle, D., Nosarti, C., Edwards, A. D., Counsell, S. J., & Victor, S. (2021). Exploring the relationship between maternal prenatal stress and brain structure in premature neonates. *PLoS ONE, 16*(4), e0250413. https://doi.org/10.1371/journal.pone.0250413

Limperopoulos, C., Bassan, H., Gauvreau, K., Robertson, R. L., Sullivan, N. R., Benson, C. B., Avery, L., Stewart, J., Md, J. S. S., Ringer, S. A., Volpe, J. J., & duPlessis, A. J. (2007). Does cerebellar injury in premature infants contribute to the high prevalence of long-term cognitive, learning, and behavioral disability in survivors? *Pediatrics, 120*(3), 584–593. https://doi.org/10.1542/peds.2007-1041

Limperopoulos, C., Bassan, H., Sullivan, N. R., Soul, J. S., Robertson, R. L., Moore, M., Ringer, S. A., Volpe, J. J., & du Plessis, A. J. (2008). Positive screening for autism in ex-preterm infants: prevalence and risk factors. *Pediatrics, 121*(4), 758–765. https://doi.org/10.1542/peds.2007-2158

Linden, A. (2020). DIAGSAMPSI: Stata module for computing sample size for a single diagnostic test with a binary outcome.

Luyster, R. J., Kuban, K. C. K., O'Shea, T. M., Paneth, N., Allred, E. N., & Leviton, A. (2011). The modified checklist for autism in toddlers in extremely low gestational age newborns: individual items associated with motor, cognitive, vision and hearing limitations. *Paediatric and Perinatal Epidemiology, 25*(4), 366–376. https://doi.org/10.1111/j.1365-3016.2010.01187.x

Moore, T., Johnson, S., Hennessy, E., & Marlow, N. (2012). Screening for autism in extremely preterm infants: problems in interpretation. *Developmental Medicine and Child Neurology, 54*(6), 514–520. https://doi.org/10.1111/j.1469-8749.2012.04265.x

Padilla, N., Eklöf, E., Mårtensson, G. E., Bölte, S., Lagercrantz, H., & Ådén, U. (2017). Poor brain growth in extremely preterm neonates long before the onset of autism spectrum disorder symptoms. *Cerebral Cortex, 27*(2), 1245–1252. https://doi.org/10.1093/cercor/bhv300

Robins, D. L., Fein, D., Barton, M. L., & Green, J. A. (2001). The modified checklist for autism in toddlers: an initial study investigating the early detection of autism and pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 31*(2), 131–144. https://doi.org/10.1023/a:1010738829569

Rodier, P. M. (2002). Converging evidence for brain stem injury in autism. *Development and Psychopathology, 14*(3), 537–557. https://doi.org/10.1017/s0954579402003085

Rodier, P. M., Ingram, J. L., Tisdale, B., & Croog, V. J. (1997). Linking etiologies in humans and animal models: studies of autism. *Reproductive Toxicology, 11*, 417–422. https://doi.org/10.1016/s0890-6238(97)80001-u

Rodier, P. M., Ingram, J. L., Tisdale, B., Nelson, S., & Romano, J. (1996). Embryological origin for autism: developmental anomalies of the cranial nerve motor nuclei. *The Journal of Comparative Neurology, 370*(2), 247–261.

Rogers, C. E., Anderson, P. J., Thompson, D. K., Kidokoro, H., Wallendorf, M., Treyvaud, K., Roberts, G., Doyle, L. W., Neil, J. J., & Inder, T. E. (2012). Regional cerebral development at term relates to school-age social-emotional development in very preterm children. *Journal of the American Academy of Child and Adolescent Psychiatry, 51*(2), 181–191. https://doi.org/10.1016/j.jaac.2011.11.009

Rubenstein, E., Schieve, L., Wiggins, L., Rice, C., Van Naarden Braun, K., Christensen, D., Durkin, M., Daniels, J., & Lee, L.-C. (2018). Trends in documented co-occurring conditions in children with

autism spectrum disorder, 2002–2010. *Research in Developmental Disabilities, 83*, 168–178. https://doi.org/10.1016/j.ridd.2018.08.015

Saudino, K. J., Dale, P. S., Oliver, B., Petrill, S. A., Richardson, V., Rutter, M., Simonoff, E., Stevenson, J., & Plomin, R. (1998). The validity of parent-based assessment of the cognitive abilities of 2-year-olds. *British Journal of Developmental Psychology, 16*(3), 349–363. https://doi.org/10.1111/j.2044-835X.1998.tb00757.x

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). mediation: *R* Package for causal mediation analysis. *Journal of Statistical Software*. https://doi.org/10.18637/jss.v059.i05

Toal, F., Daly, E. M., Page, L., Deeley, Q., Hallahan, B., Bloemen, O., Cutter, W. J., Brammer, M. J., Curran, S., Robertson, D., Murphy, C., Murphy, K. C., & Murphy, D. G. M. (2010). Clinical and anatomical heterogeneity in autistic spectrum disorder: a structural MRI study. *Psychological Medicine, 40*(7), 1171–1181. https://doi.org/10.1017/S0033291709991541

Tsai, P. T., Hull, C., Chu, Y., Greene-Colozzi, E., Sadowski, A. R., Leech, J. M., Steinberg, J., Crawley, J. N., Regehr, W. G., & Sahin, M. (2012). Autistic-like behaviour and cerebellar dysfunction in Purkinje cell Tsc1 mutant mice. *Nature, 488*(7413), 647–651. https://doi.org/10.1038/nature11310

Ure, A. M., Treyvaud, K., Thompson, D. K., Pascoe, L., Roberts, G., Lee, K. J., Seal, M. L., Northam, E., Cheong, J. L., Hunt, R. W., Inder, T., Doyle, L. W., & Anderson, P. J. (2016). Neonatal brain abnormalities associated with autism spectrum disorder in children born very preterm. *Autism Research, 9*(5), 543–552. https://doi.org/10.1002/aur.1558

Vanes, L. D., Hadaya, L., Kanel, D., Falconer, S., Ball, G., Batalle, D., Counsell, S. J., Edwards, A. D., & Nosarti, C. (2021). Associations between neonatal brain structure, the home environment, and childhood outcomes following very preterm birth. *Biological Psychiatry, 1*(2), 146–155. https://doi.org/10.1016/j.bpsgos.2021.05.002

Volpe, J. J. (2009). Brain injury in premature infants: a complex amalgam of destructive and developmental disturbances. *Lancet Neurology, 8*(1), 110–124. https://doi.org/10.1016/S1474-4422(08)70294-1

Wang, S.S.-H., Kloth, A. D., & Badura, A. (2014). The cerebellum, sensitive periods, and autism. *Neuron, 83*(3), 518–532. https://doi.org/10.1016/j.neuron.2014.07.016

Watson, T. C., Koutsikou, S., Cerminara, N. L., Flavell, C. R., Crook, J. J., Lumb, B. M., & Apps, R. (2013). The olivo-cerebellar system and its relationship to survival circuits. *Frontiers in Neural Circuits, 7*, 72. https://doi.org/10.3389/fncir.2013.00072

Wegiel, J., Flory, M., Kuchna, I., Nowicki, K., Ma, S. Y., Imaki, H., Wegiel, J., Cohen, I. L., London, E., Wisniewski, T., & Brown, W. T. (2014). Stereological study of the neuronal number and volume of 38 brain subdivisions of subjects diagnosed with autism reveals significant alterations restricted to the striatum, amygdala and cerebellum. *Acta Neuropathologica Communications, 2*, 141. https://doi.org/10.1186/s40478-014-0141-7

Wegiel, J., Kuchna, I., Nowicki, K., Imaki, H., Wegiel, J., Ma, S. Y., Azmitia, E. C., Banerjee, P., Flory, M., Cohen, I. L., London, E., Brown, W. T., Hare, C. K., & Wisniewski, T. (2013). Contribution of olivofloccular circuitry developmental defects to atypical gaze in autism. *Brain Research, 1512*, 106–122. https://doi.org/10.1016/j.brainres.2013.03.037

Wegiel, J., Kuchna, I., Nowicki, K., Imaki, H., Wegiel, J., Marchi, E., Ma, S. Y., Chauhan, A., Chauhan, V., Bobrowicz, T. W., de Leon, M., Louis, L. A. S., Cohen, I. L., London, E., Brown, W. T., & Wisniewski, T. (2010). The neuropathology of autism: defects of neurogenesis and neuronal migration, and dysplastic changes. *Acta Neuropathologica, 119*(6), 755–770. https://doi.org/10.1007/s00401-010-0655-4

Wigham, S., McConachie, H., Tandos, J., & Le Couteur, A. S. (2012). The reliability and validity of the Social Responsiveness Scale in a UK general child population. *Research in Developmental Disabilities, 33*(3), 944–950. https://doi.org/10.1016/j.ridd.2011.12.017

# Appendix B

A PDF copy of the published manuscript referred to in Chapter 4 can be found here.

***Reference:*** *Hadaya, L., Dimitrakopoulou, K., Vanes, L. D., Kanel, D., Fenn-Moltu, S., Gale-Grant, O., Counsell, S. J., Edwards, A. D., Saqi, M., Batalle, D., & Nosarti, C. (2023). Parsing brain-behavior heterogeneity in very preterm born children using integrated similarity networks. Translational Psychiatry, 13(1), Article 1.* https://doi.org/10.1038/s41398-023-02401-w

Translational Psychiatry www.nature.com/tp

ARTICLE OPEN

Check for updates

# Parsing brain-behavior heterogeneity in very preterm born children using integrated similarity networks

Laila Hadaya[1,2], Konstantina Dimitrakopoulou[3], Lucy D. Vanes[4], Dana Kanel[1,2], Sunniva Fenn-Moltu[1,5], Oliver Gale-Grant[1,5,6], Serena J. Counsell[1], A. David Edwards[1], Mansoor Saqi[3], Dafnis Batalle[1,5] and Chiara Nosarti[1,2]✉



Very preterm birth (VPT; ≤32 weeks' gestation) is associated with altered brain development and cognitive and behavioral difficulties across the lifespan. However, heterogeneity in outcomes among individuals born VPT makes it challenging to identify those most vulnerable to neurodevelopmental sequelae. Here, we aimed to stratify VPT children into distinct behavioral subgroups and explore between-subgroup differences in neonatal brain structure and function. 198 VPT children (98 females) previously enrolled in the Evaluation of Preterm Imaging Study (EudraCT 2009-011602-42) underwent Magnetic Resonance Imaging at term-equivalent age and neuropsychological assessments at 4–7 years. Using an integrative clustering approach, we combined neonatal socio-demographic, clinical factors and childhood socio-emotional and executive function outcomes, to identify distinct subgroups of children based on their similarity profiles in a multidimensional space. We characterized resultant subgroups using domain-specific outcomes (temperament, psychopathology, IQ and cognitively stimulating home environment) and explored between-subgroup differences in neonatal brain volumes (voxel-wise Tensor-Based-Morphometry), functional connectivity (voxel-wise degree centrality) and structural connectivity (Tract-Based-Spatial-Statistics). Results showed two- and three-cluster data-driven solutions. The two-cluster solution comprised a 'resilient' subgroup (lower psychopathology and higher IQ, executive function and socio-emotional scores) and an 'at-risk' subgroup (poorer behavioral and cognitive outcomes). No neuroimaging differences between the resilient and at-risk subgroups were found. The three-cluster solution showed an additional third 'intermediate' subgroup, displaying behavioral and cognitive outcomes intermediate between the resilient and at-risk subgroups. The resilient subgroup had the most cognitively stimulating home environment and the at-risk subgroup showed the highest neonatal clinical risk, while the intermediate subgroup showed the lowest clinical, but the highest socio-demographic risk. Compared to the intermediate subgroup, the resilient subgroup displayed larger neonatal insular and orbitofrontal volumes and stronger orbitofrontal functional connectivity, while the at-risk group showed widespread white matter microstructural alterations. These findings suggest that risk stratification following VPT birth is feasible and could be used translationally to guide personalized interventions aimed at promoting children's resilience.

*Translational Psychiatry* (2023)13:108 ; https://doi.org/10.1038/s41398-023-02401-w

## INTRODUCTION

Very preterm birth (VPT; ≤32 weeks' gestation) is associated with an increased likelihood of developing cognitive and behavioral difficulties across the lifespan [1–5]. Efforts to conceptualize these difficulties have proposed a "preterm behavioral phenotype", characterized by problems in emotional and social processing, and inattention [6]. However, while some VPT children display a behavioral profile reflecting a preterm phenotype, others follow typical developmental trajectories [7–9]. Such behavioral heterogeneity following VPT birth presents a challenge for building risk prediction models [10], as multiple causes may lead to the same outcome and as a single mechanism may lead to multiple outcomes [11].

Several endogenous and exogenous factors contribute to a child's behavioral development and a complex interplay between environmental, clinical, and neurobiological features could result in co-occurring neurodevelopmental, cognitive and behavioral difficulties following VPT birth [12]. These factors are often non-independent and their combination (e.g., neurobiological and socio-demographic variables) may result in improved prediction of functional outcomes [13]. For instance, both socio-demographic deprivation and increased neonatal clinical risk have been associated with neurodevelopmental as well as behavioral difficulties in VPT children. These encompass executive and socio-emotional functions [14–16], which could be considered as gateway mechanisms that shape behavioral outcomes, as they are

[1]Centre for the Developing Brain, Department of Perinatal Imaging and Health, Faculty of Life Sciences & Medicine, King's College London, London, UK. [2]Department of Child and Adolescent Psychiatry, Institute of Psychiatry Psychology and Neuroscience, King's College London, London, UK. [3]Translational Bioinformatics Platform, NIHR Biomedical Research Centre, Guy's and St. Thomas' NHS Foundation Trust and King's College London, London, UK. [4]Centre for Neuroimaging Sciences, Institute of Psychiatry Psychology and Neuroscience, King's College London, London, UK. [5]Department of Forensic and Neurodevelopmental Sciences, Institute of Psychiatry Psychology and Neuroscience, King's College London, London, UK. [6]MRC Centre for Neurodevelopmental Disorders, King's College London, London, UK. ✉email: chiara.nosarti@kcl.ac.uk

Received: 24 October 2022 Revised: 14 March 2023 Accepted: 17 March 2023
Published online: 03 April 2023

subserved by brain networks relating to both bottom-up stimulus processing and top-down behavioral control [17]. Impairments in these domains have in fact been associated with later academic and mental health difficulties [3, 18].

Previous studies have attempted to stratify outcome heterogeneity in preterm children using clustering and latent-class analyses [7–9, 19, 20]. These studies typically used cognitive and behavioral measures as input features, and then compared subgroups in terms of specific clinical and environmental risk factors that were not used in the stratification analyses (i.e., out-of-model). Some found differences in neonatal clinical profiles between subgroups of preterm children [20] and others showed that familial characteristics, such as parental education, maternal distress, and cognitively stimulating parenting, differentiated resilient subgroups from those exhibiting behavioral difficulties [8, 9]. Here, instead, we chose to include input measures of known risk factors (i.e., clinical and environmental variables) alongside in-model cognitive and behavioral measures, in order to delineate the complex interplay between different risk factors and behavioral outcome measures; thus increasing the likelihood of discovering nuanced subtypes of preterm children who exhibit similar behavioral outcomes, but with possibly different underlying correlates (i.e., equifinality) [11].

A growing body of research, investigating specific factors associated with later behavioral outcomes, is studying the early neural signatures that may shape an individual's neurodevelopmental trajectory. Alterations in brain volumes [21, 22], white matter microstructure [23, 24], and functional connectivity [25, 26] at birth in regions and networks subserving social, emotional and attentional processes, have been associated with later behavioral difficulties in VPT samples. Differences between latent subgroups of VPT children and infants have been previously studied in relation to qualitative measures of brain abnormalities and/or high grade brain injury based on neonatal Magnetic Resonance Imaging (MRI), as well as quantitative differences in brain tissue volumes [8, 27, 28]. However, it remains to be explored whether distinct multidimensional subgroups of VPT children could also be characterized by localized differences in early brain development using advanced quantitative measures of brain structure and function, such as log-Jacobians, tract based spatial statistics and degree centrality, which have previously been used in neonatal samples [29–31]. Conducting analyses at the whole-brain and voxel-wise level, allows for an enhanced spatial localization of potential structural and functional between-subgroup differences, thus extending previous research [8, 27, 28].

The main aim of this study was to parse brain-behavior heterogeneity in VPT children, by identifying subgroups with similar environmental, clinical and behavioral profiles and examining between-subgroup differences in structural and functional brain features at term-equivalent age. Firstly, we implemented an integrative clustering approach (Similarity Network Fusion; SNF) [32] to stratify VPT children into distinct subgroups based on three data types: (i) neonatal clinical and socio-demographic variables, (ii) childhood socio-emotional outcomes and (iii) executive function measures. The advantage of this approach is that it integrates sample-similarity networks built from each distinct data type and constructs a final integrated network, which contains common and complementary information from the different data types. This is then used to stratify the sample into distinct subgroups using clustering [32]. We also investigated whether resultant subgroups differed in outcomes that were not used in stratification analyses (i.e., out-of-model variables); in order to provide external validation [33–35]. Finally, we explored between-subgroup differences in regional brain volume and structural and functional connectivity at term-equivalent age. We hypothesized that there would be distinct subgroups of VPT children characterized by unique neonatal neural signatures.

## METHODS

### Study design

*Participants*. Five hundred and eleven infants born VPT were recruited from 14 neonatal units in London in 2010–2013 and entered the Evaluation of Preterm Imaging Study (ePrime; EudraCT 2009-011602-42) [36]. Infants with congenital malformation, prior MRI, metallic implants, whose parents did not speak English or were subject to child protection proceedings were not eligible for participation in the study.

Participants underwent multimodal MRI at 38–53 weeks post-menstrual age (PMA) on a 3-Tesla MR imaging system (Philips Medical Systems, Best, The Netherlands) located on the neonatal intensive care unit at Queen Charlotte's and Chelsea Hospital, London, using an 8-channel phased array head coil. For data acquisition and imaging parameters see Supplemental Information. Infants whose parents chose sedation for the procedure (87%) received oral chloral hydrate (25–50 mg/kg).

In total, 251 participants (including 29 sets of multiple pregnancy children) were followed-up between the age of 4 and 7 years at the Center for the Developing Brain, St Thomas' Hospital, London. This was a convenience sample corresponding to 82% of 306 participants who were past their fourth birthday by the study end date, September 1st 2019, and had consented to be contacted for future research. Invitations for follow-up were sent in chronological order of birth.

Ethical approval was granted by the Hammersmith and Queen Charlotte's Research Ethic Committee (09/H0707/98) and the Stanmore Research Ethics Committee (14/LO/0677). Informed consent was obtained from all participants.

*Clinical and socio-demographic data*. We used Principal Component Analysis (PCA) to select neonatal clinical variables of interest from a set of 28 available variables. These were: gestational age (GA) at birth, number of days on mechanical ventilation, number of days on continuous positive airway pressure (CPAP) and number of days on parenteral nutrition (TPN), which loaded onto a single component explaining 72% of the variance in the data. This component was labeled 'neonatal sickness index'. Please refer to our previous work [24] and Supplemental Information for more details on the PCA analysis.

Socio-demographic risk was evaluated using a postcode derived measure of deprivation in England, the Index of Multiple Deprivation 2010 (IMD; http://tools.npeu.ox.ac.uk/imd/), whereby higher IMD scores reflect greater deprivation. The IMD combines neighborhood-specific information about seven domains of deprivation: income, employment, education/skills/training, health, crime, housing and living environment. The IMD was collected at the term-equivalent age. Continuous IMD scores were used in the integrative-clustering and evaluation of subgroup profile analyses. IMD quintiles are provided when reporting sample characteristics (Table 1) for ease of interpretability.

*Childhood assessment*. Intelligence quotient (IQ) was evaluated using the Wechsler Preschool and Primary Scale for Intelligence (WPPSI-IV) [37] and executive function using the preschool version of the parent-rated Behavior Rating Inventory of Executive Function (BRIEF-P) [38]. Socio-emotional processing was evaluated using the Empathy Questionnaire (EmQue) [39] and the Social Responsiveness Scale, Second Edition (SRS-2) [40]. Psychopathology was assessed using the Strengths and Difficulties Questionnaire (SDQ) [41], temperament using the Child Behavioral Questionnaire - Very Short Form (CBQ) [42] and cognitively stimulating home environment using an adapted version of the Cognitive Stimulating Parenting Scale (CSPS) [43].

*Exclusions*. Twenty-seven participants were excluded due to incomplete childhood outcome data, 17 due to major brain lesions (periventricular leukomalacia, parenchymal hemorrhagic infarction, or other ischemic or hemorrhagic lesions), detected on neonatal T2-weighted MRI images at term by an experienced perinatal neuroradiologist, and 5 participants due to missing T2-weighted MRI images, hence the inability to evaluate the presence of major lesions (Fig. S1).

### Data integration and clustering

Analyses were conducted in R (version 3.6.1). Using SNF, three data types were integrated: (Type 1) neonatal socio-demographic and clinical variables: IMD at birth, GA, days on ventilation, days on TPN and days on CPAP. (Type 2) childhood socio-emotional outcomes: EmQue subscale raw scores - emotion contagion, attention to others' emotions, prosocial behaviors and SRS-2 total raw score. (Type 3) childhood executive function:

**Table 1.** Socio-demographic and clinical participant data.

| | | Integrative clustering sample $n = 198$ | Diffusion MRI TBSS analysis sample $n = 166$ | Structural MRI log-Jacobian analysis sample $n = 165$ | rs-fMRI degree centrality analysis sample $n = 129$ |
|---|---|---|---|---|---|
| Corrected age at assessment, years | Median | 4.63 | 4.60 | 4.59 | 5.63 |
| | Range | 4.18–7.17 | 4.18–7.17 | 4.18–7.17 | 4.18–7.17 |
| PMA, weeks | Median | 42.57 | 42.43 | 42.57 | 42.43 |
| | Range | 38.29–52.86 | 38.29–44.86 | 38.29–44.86 | 38.29–44.86 |
| Sex, male:female | n= | 100:98 | 88:78 | 86:79 | 68:61 |
| Self-reported maternal ethnicity | n (%) | | | | |
| Asian | | 50 (25.3%) | 44 (26.5%) | 43 (26.1%) | 34 (26.4%) |
| Black/African/Caribbean/ Black British | | 30 (15.2%) | 23 (13.9%) | 25 (15.2%) | 15 (11.6%) |
| Mixed/Multiple ethnic groups | | 3 (1.5%) | 3 (1.8%) | 3 (1.8%) | 3 (2.33%) |
| White | | 112 (56.6%) | 93 (56.0%) | 91 (55.2%) | 75 (58.1%) |
| Self-reported paternal ethnicity | *n* (%) | | | | |
| Asian | | 34 (17.2%) | 29 (17.5%) | 27 (16.4%) | 23 (17.8%) |
| Black/African/Caribbean/ Black British | | 23 (11.6%) | 19 (11.5%) | 20 (12.1%) | 14 (10.9%) |
| Mixed/Multiple ethnic groups | | 2 (1.0%) | 1 (0.6%) | 1 (0.6%) | 0 (0.0%) |
| White | | 95 (48.0%) | 80 (48.2%) | 79 (47.9%) | 63 (48.8%) |
| Neonatal IMD, quintiles | *n* (%) | | | | |
| 1 (least deprived) | | 49 (24.8%) | 40 (24.1%) | 38 (23.0%) | 30 (23.3%) |
| 2 | | 37 (18.7%) | 31 (18.7%) | 32 (19.4%) | 25 (19.4%) |
| 3 | | 44 (22.2%) | 39 (0.6%) | 38 (23.0%) | 30 (23.3%) |
| 4 | | 48 (24.2%) | 39 (23.5%) | 38 (23.0%) | 31 (24.0%) |
| 5 (most deprived) | | 20 (10.1%) | 17 (10.2%) | 19 (11.5%) | 13 (10.1%) |
| GA at birth, weeks | Median | 30.14 | 30.29 | 30.14 | 30.14 |
| | Range | 23.86–32.86 | 24.00–32.86 | 24.00–32.86 | 24.00–32.86 |
| Neonatal clinical risk | *n*= | | | | |
| Days TPN, ratio 0:1:2 | | 68:98:32 | 62:78:26 | 63:77:25 | 49:61:19 |
| Days CPAP, ratio 0:1:2 | | 33:125:40 | 30:107:29 | 31:103:31 | 23:82:24 |
| Days ventilation, ratio 0:1:2 | | 101:74:23 | 92:59:15 | 92:58:15 | 72:46:11 |

Note: Table describing sample socio-demographic and clinical characteristics for the integrative clustering and MRI analyses.
Neonatal clinical risk categories (0, 1 and 2) respectively correspond to zero days, more than zero days, but less than the top quintile, and within the top quintile. IMD quintiles 1–5 respectively correspond to the least deprived quintile (1) to the most deprived quintile (5). Ethnicity was grouped according to the Office of National Statistics classifications 2016 (see Supplementary Information).
*CPAP* continuous positive airway pressure, *GA* gestational age at birth, *IMD* Index of Multiple Deprivation, *PMA* post-menstrual age at scan, *rs-fMRI* resting-state functional MRI, *TBSS* Tract Based Spatial Statistics, *TPN* total parenteral nutrition.

BRIEF-P raw subscale scores - inhibit, shift, emotional control, working memory and plan/organize.

Prior to integration, participants with in-model outlier values greater than 3 times the interquartile range were excluded. A total of 198 children were included in the SNF analyses. Zero-inflated neonatal clinical risk variables (days ventilation, days TPN and days CPAP) were converted into ordinal categorical variables with three levels: (Level 0: zero days; Level 1: greater than zero and not within the top quintile; Level 2: within the top quintile). For the mixed data type (numeric and categorical data; data type 1), Gower's standardization based on the range was applied using the *daisy* function from cluster R package [44] and for numeric only matrices (data types 2 and 3), variables were standardized to have a mean value of 0 and a standard deviation of 1 using the *standardNormalization* function from SNFtool R package [45].

An adaptation of the *ExecuteSNF.CC* function [46] was used for the data integration and clustering steps. Dissimilarity Gower distance (for the mixed data type) and Euclidean distance (for numeric data types) matrices were calculated and used to create similarity matrices using the SNFtool R package's *affinityMatrix* function [45]. This was followed by an integration of the similarity matrices using SNFtool's *SNF* algorithm resulting in a 'fused similarity matrix' [45]. The integrative clustering process can be summarized into two steps:

Step 1: SNF method has two main hyperparameters, K and alpha. K (i.e., neighborhood size) indicates the number of neighbors of a node to consider when the similarity networks are being generated and alpha is an edge weighting parameter determining the weight of edges between nodes in the networks. We tried 30 combinations of K and alpha hyperparameters {K = 10, 15, 20, 25, 30; alpha = 0.3, 0.4, 0.5, 0.6, 0.7, 0.8}, similar to the approach followed in [47]. The K-alpha hyperparameter values were chosen based on the ranges recommended in the SNFtool R package, 10–30 for K and 0.3–0.8 for alpha [32, 45]. Consensus clustering, using *ConsensusClusterPlus* function [48], was then applied to each fused

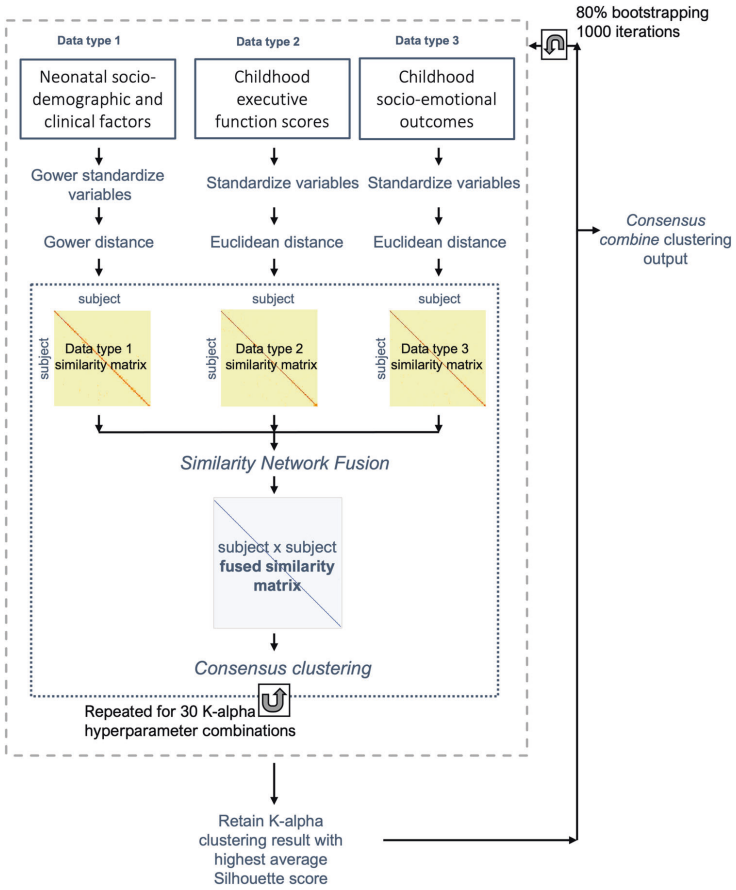



**Fig. 1 Data integration and clustering pipeline.** Figure summarizing the data pre-processing (variable normalization), data integration and clustering pipeline executed in order to obtain the final consensus cluster assignment.

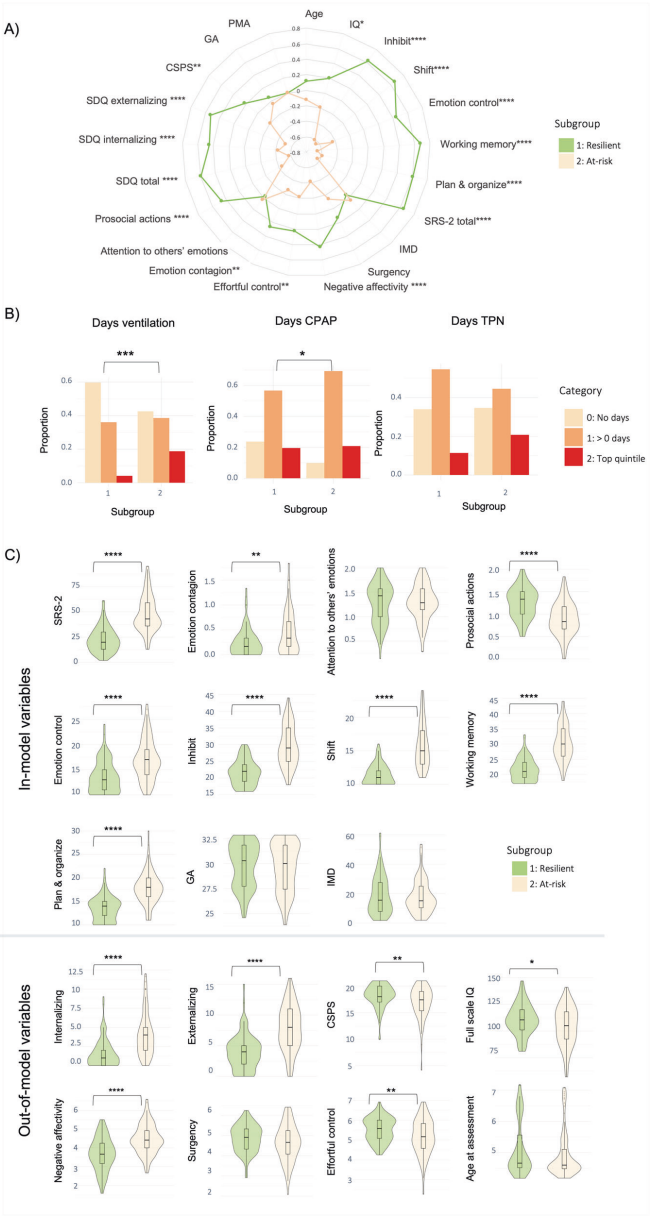
similarity matrix, corresponding to a K-alpha combination, where spectral clustering was run 1000 times with 80% of the population randomly subsampled for each clustering run and a single consensus clustering result obtained from hierarchical clustering. Step 2: Next, out of the 30 clustering results produced in step 1, the one with the highest average silhouette width score was retained. Steps 1 and 2 were repeated 1000 times in a bootstrap approach, after selecting and pre-processing the three data matrices of 80% of the sample set. The 1000 resultant retained clustering outputs were then fed to the diceR R package's *consensus_combine* function [49] which implements hierarchical clustering on the consensus matrix and generates the final consensus clustering. Figure 1 summarizes the data-integration and clustering steps and the code used can be accessed here: https://github.com/lailahadaya/preterm-ExecuteSNF.CC. Further details can also be found in Supplemental Information.

Before implementing steps 1 and 2, it was essential to determine the number of clusters. For this, we used the SNFtool R package's *estimateNumberOfClustersGivenGraph* function [45] to calculate Eigengap and Rotation Cost heuristics for each K-alpha combination (Fig. S2). This process suggested C = 2, C = 3 and C = 4 as the optimal number of clusters. Consensus matrices and silhouette scores were generated and compared for these three potential clustering solutions (Fig. S2). Resultant subgroups from C = 2 and C = 3 were chosen to be evaluated for phenotypic differences, as their silhouette scores and consensus matrices gave better values in comparison to those of C = 4 (Fig. S2). More details on the estimation of cluster numbers can be found in Supplemental Information. An alluvial plot was used to illustrate the transition of subject subgroup classification between the two-cluster and three-cluster solutions (Fig. S3).

### Evaluation of subgroup profiles

Resultant subgroups were characterized based on in-model and out-of-model variables. For the out-of-model features, subgroups were compared in terms of psychiatric symptoms (SDQ internalizing, externalizing problems and total scores), temperament (CBQ negative affectivity, surgency and effortful control scores), cognitive abilities (WPPSI full-scale IQ), and cognitive stimulation at home (CSPS score). Details on selection of in-model and out-of-model variables can be found in Supplemental Information and Figs. S4 and S5.

For numeric measures, between-subgroup differences were assessed using non-parametric one-way tests: Mann-Whitney when C = 2 or Kruskal Wallis when C = 3 [50]. Shapiro-Wilk test was used to assess normality. For categorical variables, Chi-squared test was used to evaluate differences in




proportions of individuals in each group when count per cell was >5 and Fischer's Exact test was used otherwise. To compare differences between the ordinal neonatal clinical variables with 3 categories (Levels 0, 1 and 2) and the non-ordinal subgroups from C = 2 and C = 3, the Extended Cochran-Armitage Test was used. We also ran supplementary post-hoc analyses investigating subgroup differences in clinical variables not included as in-model variables (please see Supplementary Information for more details).

Results with $p < 0.05$ were considered to be statistically significant. To correct for multiple comparisons the False Discovery Rate method was used. The same statistical analyses were repeated using general linear models correcting for potential confounders (age and sex) and 5000

**Fig. 2 Two-cluster solution subgroup profiles. A** Radar plot showing the two-cluster solution subgroup profiles using z-scores for subgroup 1 (i.e., resilient subgroup; green) and subgroup 2 (i.e., at-risk subgroup; beige). For visual illustrative purposes, scales which usually indicate poorer outcomes have been reversed so that larger z-scores on behavioral variables indicate better outcomes. **B** Bar plots for clinical risk variables (days on TPN, days on mechanical ventilation and days on CPAP, left to right, respectively) for each of the two subgroups. Plots represent the proportion of children belonging to each clinical risk category within a subgroup, where category 0 represents the lowest clinical risk (light beige; no days of clinical intervention), category 1 represents medium clinical risk (orange; more than 0 days of intervention but less than the top quintile), and category 2 represents the highest clinical risk (red; within the top quintile). **C** Violin plots showing differences between the subgroups in terms of in-model and out-of-model variables. Significant differences are marked with bars between the subgroups. *=$p < 0.05$; **=$p < 0.01$; ***=$p < 0.001$, ****=$p < 0.0001$.

permutation test iterations [51]. Effect sizes for non-normally distributed variables were measured using Wilcoxon Glass Rank Biserial Correlation (gr) for measuring differences between two groups and Epsilon Squared for three groups. For continuous normally distributed variables, Cohen's F was used and Cramer's V for categorical variables.

### Exploring neonatal brain differences between subgroups

Tract Based Spatial Statistics (TBSS) was used to assess white matter microstructure at the voxel-level using fractional anisotropy (FA) and mean diffusivity (MD) maps [52]. FA approximates the directional profile of water diffusion in each voxel and MD measures the average movement of water molecules within a voxel. Higher FA and lower MD values reflect more optimal white matter myelination and microstructure. For diffusion MRI (d-MRI) image pre-processing and TBSS protocol details please refer to Supplemental Information.

Structural MRI (s-MRI) log-Jacobian determinant maps were calculated to quantify regional brain volumes (greater log-Jacobian values reflect larger relative structural volumes), using Tensor Based Morphometry, following methods described in our previous work [53, 54] and in Supplemental Information.

Resting-state functional MRI (rs-fMRI) data were pre-processed as in our previous work;[55] for more details see Supplemental Information. Functional connectivity was evaluated using a measure of weighted degree centrality at the voxel-level (i.e., the sum of the correlations between the time-series of each voxel and all other voxels within a gray matter mask of the brain) [31, 56]. Edges with a correlation coefficient below a threshold of 0.2 were excluded and the degree centrality values for each voxel in the gray matter mask were z-scored and used in subsequent between-subgroup analyses. Whilst other functional network measures are available (i.e., participation coefficient and within module-z [57], we opted to study degree centrality as we recently showed this to be disrupted in preterm born neonates [31]. Furthermore, degree centrality is a good voxel-wise summary measure of connectivity strength, which is reliable and correlates with relevant phenotypes, such as age and sex [58]. It has been used to study typical cognitive function [59] and has recently been shown to be a reproducible metric to detect atypical functional connectivity patterns in neurodevelopmental disorders [56].

The number of children included in the different modality-specific MRI analyses slightly differed due data availability: TBSS ($n = 166$), log-Jacobian determinant maps ($n = 165$) and degree centrality ($n = 129$); please see Table S1. Exclusions for specific MRI analyses are depicted in Fig. S1.

Between-subgroup differences were investigated in the whole-brain at the voxel-level in terms of: log-Jacobian determinants, TBSS metrics (FA and MD) and degree centrality. FMRIB Software Library (FSL) [60] *randomise* function was used to implement non-parametric permutation methods for statistical inference. This method was used to model each contrast of interest for each voxel, i.e., a general linear model (GLM) correcting for PMA at scan and sex. rs-fMRI models also included motion estimates (standardized DVARS) as a covariate. Family Wise Error (FWE) rate with Threshold-Free Cluster Enhancement (TFCE) was applied to correct for multiple comparisons over the multiple voxels, while enhancing "cluster-like" structures of voxels without defining them as binary components [61]. Statistics were calculated using random permutation tests with 10000 permutations. Given the exploratory nature of our analysis, we did not correct for multiple contrasts tested (i.e., log-Jacobians, TBSS FA and MD, degree centrality). We show results significant at $p < 0.05$ FWE-corrected per contrast. Mean values from clusters of modality-specific voxels showing significant between-subgroup differences were extracted to calculate Cohen's F effect sizes.

### Sensitivity analyses

There were 29 sets of children born from multiple pregnancy events in our sample. In order to account for multiple pregnancy confounding, we conducted additional sensitivity analyses including only one child from each set of multiple pregnancy siblings.

## RESULTS

### Participant characteristics

Participants' socio-demographic and clinical characteristics are shown in Table 1. Compared to participants who completed the follow-up assessment ($n = 251$; median GA = 29.24 weeks; median IMD at birth=19.48), individuals who were not assessed ($n = 259$; median GA = 29.27 weeks; median IMD at birth = 21.40) did not differ in GA (gr = 0.01; $p = 0.807$), but had greater neonatal socio-demographic deprivation (gr = 0.11; $p = 0.028$). Compared to the initial baseline cohort (n = 511; median GA = 30.00 weeks; median IMD at birth=18.19), participants who were studied here ($n = 198$) had slightly older GA (median GA = 30.14 weeks; gr = −0.13; $p = 0.009$) and relative socio-demographic advantage (median IMD score at birth=15.58, $gr = 0.11$; $p = 0.027$).

### Two-cluster solution subgroup profiles

When stratifying the sample into two clusters and comparing them in terms of in-model variables, subgroup 1 (termed here the 'resilient' subgroup) showed significantly better socio-communication (i.e., lower SRS-2 scores) and executive function abilities (i.e., lower BRIEF-P emotion control, inhibit, shift, working memory and plan/organize scores), lower emotion contagion (EmQue) scores, and higher prosocial actions scores (EmQue) during childhood, than subgroup 2 (termed here the 'at-risk' subgroup); all $ps < 0.05$, after FDR correction. The resilient subgroup had lower neonatal clinical risk compared to the at-risk subgroup, with a greater proportion of children receiving no neonatal mechanical ventilation and a smaller proportion of children receiving prolonged neonatal CPAP (both $ps < 0.05$, after FDR correction). Subgroups did not differ in terms of days on TPN in the neonatal period ($p > 0.05$).

Differences in out-of-model variables included lower psychopathology scores (SDQ internalizing and externalizing problems) and negative affectivity scores (CBQ) as well as higher effortful control (CBQ), IQ and cognitive stimulation at home (CSPS) during childhood in the resilient compared to the at-risk subgroup; all $ps < 0.05$, after FDR correction (Fig. 2; Table S2).

The two subgroups showed no significant differences in log-Jacobian determinant values, degree centrality or white matter microstructural characteristics (all $ps > 0.05$). Resultant subgroups also did not show differences in sex, age at assessment or PMA at scan (Fig. 2; Table S2).

### Three-cluster solution subgroup profiles

To increase subtyping resolution and explore latent heterogeneity not captured by a two-subgroup partitioning, the sample was further stratified into 3 subgroups. Two of the three resulting clusters largely reflected profiles similar to those from C = 2.

The first was a 'resilient' subgroup (subgroup 1) with favorable childhood socio-communicative (SRS-2), empathy (EmQue) and executive function (BRIEF-P) outcomes in terms of in-model variables; low childhood psychopathology (SDQ internalizing and externalizing problems) and negative affectivity scores (CBQ) and high effortful control scores (CBQ), IQ and cognitive stimulation at home (CPSP) in terms of out-of-model variables. The second was an 'at-risk' subgroup (subgroup 2), with the poorest outcomes in terms of in-model variables (childhood socio-communication (SRS-2), empathy and executive function (BRIEF-P) scores), as well as out-of-model childhood psychopathology (SDQ), effortful control (CBQ) and negative affectivity measures (CBQ), combined with the highest neonatal clinical risk (Fig. 3; Table S3).

A third subgroup (labeled 'intermediate') emerged, which had poorer in-model and out-of-model childhood cognitive and behavioral scores when compared to the resilient subgroup, but better scores when compared to those of the at-risk subgroup. The intermediate subgroup also had the lowest neonatal clinical risk compared to both resilient and at-risk subgroups (Fig. 3; Table S3). The transition of subject classifications from the two- to the three-cluster solution is illustrated in an alluvial plot (Fig. S3).

In terms of environmental factors, the resilient subgroup had higher levels of childhood cognitive stimulation at home (CSPS) in comparison to both at-risk and intermediate subgroups, while the intermediate subgroup had higher neonatal socio-demographic risk (IMD) in comparison to both at-risk and resilient subgroups. All $ps < 0.05$ after FDR correction. The three subgroups did not differ in sex, age at assessment or PMA at scan.

In terms of brain imaging markers at term, the resilient subgroup displayed larger relative volumes (i.e., greater log-Jacobian determinant values) in the left insula and bilateral orbitofrontal cortices (Fig. 4A; Table S4) and higher degree centrality in an overlapping region in the left orbitofrontal cortex (Fig. 4B; Table S5) compared to the intermediate subgroup. The intermediate subgroup, compared to the at-risk subgroup, showed increased FA in several areas of the white matter skeleton, including the fornix, corpus callosum, corticospinal tract, inferior longitudinal, inferior fronto-occipital and uncinate fasciculi (Fig. 4Ci; Table S4), as well as lower MD in the fornix and body of the corpus callosum (Fig. 4Cii; Table S4). The resilient and at-risk subgroups did not differ in any brain measures ($p > 0.05$).

### Sensitivity analyses

Sensitivity analyses including only one sibling, selected at random from each multiple pregnancy set, revealed similar results (Table S5; Table S6; Fig. S6; Fig. S7), although the difference in neonatal functional connectivity between the resilient and intermediate groups was no longer significant ($p = 0.08$). In addition, the resilient subgroup displayed larger neonatal relative volume of the right insula compared to the intermediate subgroup. For more details, please refer to Supplemental Information.

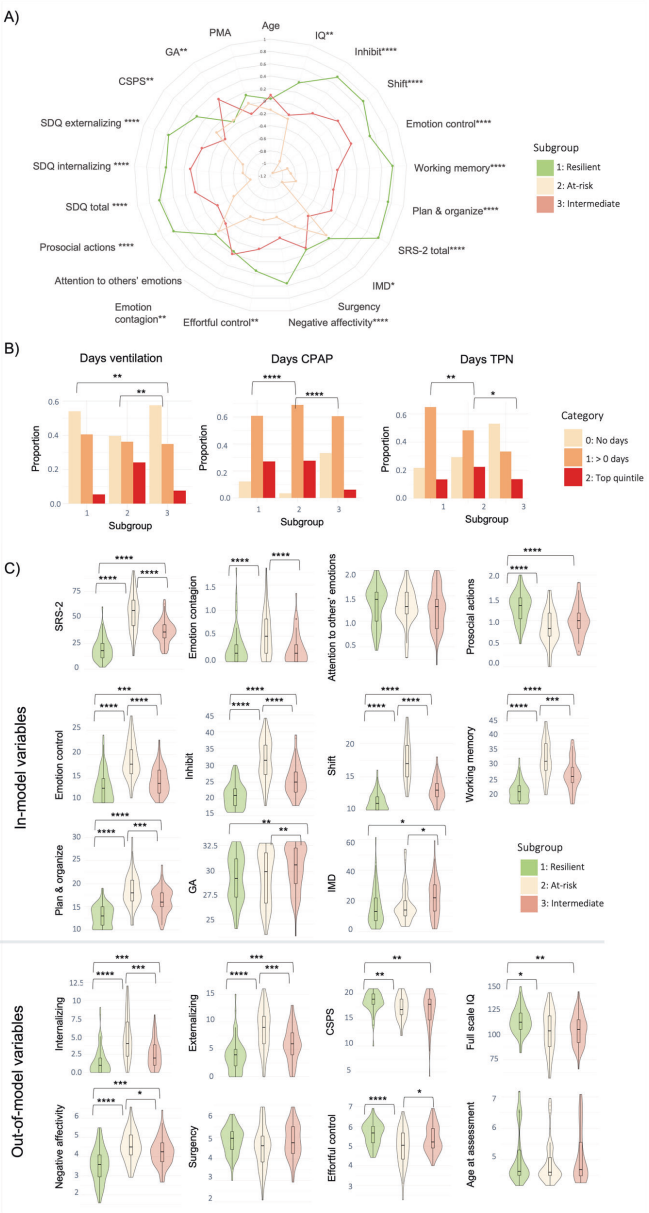
## DISCUSSION

Using an integrative clustering approach, we identified subgroups of VPT children with distinct neurodevelopmental profiles. We described a two-cluster solution, showing a resilient subgroup with comparably favorable childhood behavioral and cognitive outcomes and increased cognitive stimulation at home, and a second, at-risk subgroup, with poorer childhood behavioral and cognitive outcomes and high neonatal clinical risk. We also described a three-cluster solution, showing two subgroups largely characterized by the profiles observed in the two-cluster solution, as well as a newly emerging third intermediate subgroup, with a childhood behavioral and cognitive profile intermediate between the resilient and the at-risk subgroups. Nuanced differences in socio-demographic, neonatal clinical and early brain measures appeared upon comparing subgroups from the three-cluster solution. Notably, the resilient subgroup displayed larger fronto-limbic brain regions and increased functional connectivity at term compared to the intermediate subgroup. The at-risk subgroup showed widespread white matter microstructural alterations in fronto-temporo-limbic tracts compared to the intermediate subgroup. Furthermore, the resilient subgroup had a more cognitively stimulating childhood home environment compared to the at-risk and intermediate subgroups, while the intermediate subgroup had the lowest clinical risk. Together, these findings highlight the potential value of neonatal structural and functional brain measures as useful biomarkers of later childhood outcomes in distinct VPT subgroups, as well as the importance of a supportive home environment to foster child development.

In the at-risk subgroup from the two-cluster solution, poorer childhood socio-emotional, executive function, IQ, mental health and temperament outcomes may have been driven by a combination of both higher clinical risk at birth and a less stimulating childhood home environment, when compared to the resilient subgroup. Previous studies in VPT children have shown cognitively stimulating parenting to be positively associated with improved socio-emotional processing and cognitive outcomes at 2 years of age [62] and reduced psychopathology and executive function difficulties at 4–7 [54]. A cognitively stimulating home environment also differentiated between psychiatric profiles at 5 [8]. Moreover, increased neonatal clinical risk in the at-risk subgroup is consistent with previous findings, showing that perinatal medical complications following VPT birth may lead to increased behavioral and developmental problems [15, 16, 63]. The resilient and at-risk subgroups, however, did not differ in any of the neonatal brain measures investigated, suggesting that there may be additional non-measured variables underlying different childhood outcomes that need further investigation, such as alterations in pro-inflammatory immunomarkers [64, 65] and/or microbiome assembly [66, 67], which are reportedly associated with increased behavioral difficulties.

To further parse heterogeneity in VPT children, we also explored a three-cluster solution. These analyses showed that two subgroups mostly reflected the profiles seen in the two-cluster solution: 1) a resilient subgroup with high levels of childhood cognitive stimulation at home and 2) an at-risk subgroup with high levels of neonatal clinical risk. A third subgroup with intermediate childhood behavioral and cognitive profiles also emerged, in which childhood psychopathology, temperament and cognitive outcomes were poorer than those observed in the resilient subgroup, but more favorable than those observed in the at-risk subgroup. Intriguingly, the intermediate subgroup exhibited the lowest neonatal clinical risk compared to the other two subgroups, with a greater proportion of infants receiving no neonatal mechanical ventilation, CPAP or TPN and with higher median GA at birth. However, the intermediate subgroup also had higher environmental risk, namely reduced childhood cognitively stimulating home environment, compared to the resilient subgroup, and higher neonatal socio-demographic deprivation, compared to both the at-risk and resilient subgroups. These findings suggest that developmental outcomes may not be understood by exploring a single causal pathway and are best studied in a multidimensional space; for example, clinical risk, which has been linearly correlated with developmental outcomes in previous studies [16, 63], ought to be investigated together with other factors that may influence development, i.e., environmental risk.

The at-risk compared to the intermediate subgroup showed widespread alterations in white matter microstructure (lower FA and higher MD) in the fornix, corpus callosum, corticospinal tract, inferior longitudinal, inferior fronto-occipital and uncinate fasciculi. The at-risk subgroup had also the highest neonatal clinical risk, hence the observed white matter changes are likely to be




associated with preterm-related neonatal complications [12, 68, 69]. White matter alterations in fronto-temporo-limbic tracts, including those observed here, have been previously associated with poorer cognitive outcomes [70–75]. They have also been implicated in emotion processing [76–78] and psychiatric disorders, including depression and schizophrenia [79–81]. The intermediate subgroup, conversely, had the lowest neonatal clinical risk, and higher FA/lower MD values in fronto-temporo-limbic tracts compared to the at-risk subgroup. These findings led us to speculate that having relative fewer neonatal

**Fig. 3 Three-cluster solution subgroup profiles. A** Radar plot showing the three-cluster solution subgroup profiles using z-scores. For visual illustration purposes, scales which usually indicate poorer outcomes have been reversed, so that larger z-scores on behavioral variables indicate better outcomes. Subgroup 1 (resilient) is marked in green, subgroup 2 (at-risk) in beige and subgroup 3 (intermediate outcomes but lowest clinical risk) in pink. **B** Bar plots for clinical risk variables (days on TPN, days on mechanical ventilation and days on CPAP, left to right, respectively) for each of the three subgroups. Plots represent the proportion of children belonging to each clinical risk category within a subgroup, where category 0 represents the lowest clinical risk (light beige; no days of clinical intervention), category 1 represents medium clinical risk (orange; more than 0 days of intervention but less than the top quintile), and category 2 represents the highest clinical risk (red; within the top quintile). **C** Violin plots showing differences in in-model and out-of-model measures at the group-wise level. Significant differences are marked with bars between the subgroups. *=$p < 0.05$; **=$p < 0.01$; ***=$p < 0.001$, ****=$p < 0.0001$.

clinical complications, and hence fewer preterm-related white matter alterations, may contribute to these children's more favorable cognitive, socio-emotional and behavioral outcomes, compared to the at-risk subgroup.

Children in the resilient subgroup exhibited higher prosocial behavior and empathy, as well as fewer childhood externalizing and internalizing symptoms and executive function difficulties, compared to the intermediate and at-risk subgroups. They also showed lower childhood negative affectivity scores, referring to the expression of dysregulated negative emotions and increased sensitivity in response to surrounding stimuli [82, 83]. While the resilient group showed no significant brain differences compared to the at-risk subgroup, we speculate that the combination of two protective factors, an enriching home environment and lower neonatal clinical risk, may have contributed to attenuating the expression of the behavioral and cognitive difficulties associated with VPT birth. These findings also support the idea of multi-finality, whereby individuals with no overt brain differences at term may display distinct behavioral outcomes later in childhood.

Compared to the intermediate subgroup, however, the resilient subgroup displayed larger relative volumes in the left insular and bilateral orbitofrontal cortices and increased functional connectivity in an overlapping left orbitofrontal region at term, years before the behavioral and cognitive childhood outcomes were assessed. These findings could be interpreted in terms of a more advanced maturation of the fronto-limbic network in the resilient subgroup, as orbitofrontal functional connectivity and insular cortical microstructure and morphology have been positively associated with GA at birth and PMA at scan [84–86]. However, as several other brain areas are undergoing rapid neurodevelopmental changes at the time our participants underwent MRI, including somatosensory, occipital, temporal, parietal and other areas of frontal cortex [86], we speculate the orbitofrontal cortex and the insula may be preferentially discriminating between the intermediate and the resilient subgroup, in the context of the brain-wide analysis approaches employed here, because they play critical functional roles in the cognitive and behavioral outcomes we studied. The orbitofrontal cortex is involved in the top-down regulation of goal-oriented executive functions and socio-emotional processing, reward-guided learning and decision making [87–89]; the insula is important for regulating internal processes, including emotional responses to external stimuli [90]. Structural alterations in the orbitofrontal cortex and insula, which are structurally connected [91], have been associated with emotion dysregulation [92] and with higher externalizing behaviors [93].

The orbitofrontal cortex is sensitive to environmental stimuli, such as early life stress [94, 95]. Individuals with a history of physical abuse [96] and VPT infants exposed to painful procedures [97] both show reduced orbitofrontal volumes in childhood. Furthermore, alterations in orbitofrontal connectivity and gyrification have been associated with social processing impairments in VPT children [98] and with executive function difficulties in extremely preterm (EPT; < 28 weeks' gestation) adolescents [99], respectively. Smaller insular volumes have been associated with worse emotion regulation skills [100] and weaker insular functional connectivity with decreased empathic responses [101]. In the late preterm period, the insula becomes a key hub region [102] and a major source of transient bursting events that support brain maturation [103]. A more mature fronto-limbic network may have therefore supported a favorable development of emotion regulation capacity, cognition, and behavior [104, 105], resulting in the resilient subgroup exhibiting lower externalizing and internalizing symptoms, increased empathy, emotion regulation abilities and executive function skills in childhood.

This study demonstrates that it is possible to parse heterogeneity in VPT children in a meaningful way. We show that protective brain maturational patterns in the neonatal period may contribute to a more resilient behavioral profile in childhood. This is encouraging, as the preterm brain is susceptible to neuroplastic changes in response to behavioral and environmental interventions, both early in life and later in childhood [106]. For example, neuroplastic changes have been observed following 'supportive-touch' (i.e., skin-to-skin contact or breastfeeding') [107], maternal sensitivity training [108], visual stimulus cues of the mother's face [109], parental praise [110] or music interventions in the neonatal intensive care unit [111]. Such methods could, therefore, be used in the future to strengthen fronto-limbic circuitry to boost children's resilience. Furthermore, our findings suggest that enriching environments may promote resilience towards more favorable behavioral outcomes. This could be done by increasing parental awareness about the importance of cognitive stimulation at home. Our findings also show that the subgroup of children with the highest neonatal clinical risk exhibit the poorest outcomes, highlighting the need to develop and implement targeted interventions for the most clinically vulnerable VPT children.

It is worth noting that the median outcome scores (IQ, BRIEF-P, SRS-2 and SDQ) for our three subgroups were within normative ranges and below clinical thresholds, even for the at-risk subgroup. Subthreshold psychiatric symptoms have been reported in other at-risk subgroups of VPT children [9, 8], and have also been associated with an increased risk of developing psychiatric disorders later in life [112]. In this context, subthreshold psychiatric symptoms may represent transdiagnostic traits that would remain undetected, and therefore untreated, if considered in a purely clinically diagnostic context, highlighting the importance of addressing psychopathology dimensionally [113, 114].

Strengths of this study include a fairly large sample size and a rich longitudinal dataset with clinical data from birth, neonatal multi-modal MRI at term and behavioral follow-up in early childhood. However, a limitation of this study is that the VPT participants included in our analyses ($n = 198$) had a relative socio-demographic advantage and older gestational age at birth than the initial baseline cohort ($n = 511$), which may limit the generalizability of our findings to a portion of the socio-demographic and gestational age spectrum. In addition, the lack of a full-term group and the exclusion of children with major brain lesions in the integrative-clustering analyses may have also limited the variability in our data, and in turn contributed to the failure to identify a more impaired subgroup here. Future studies must take

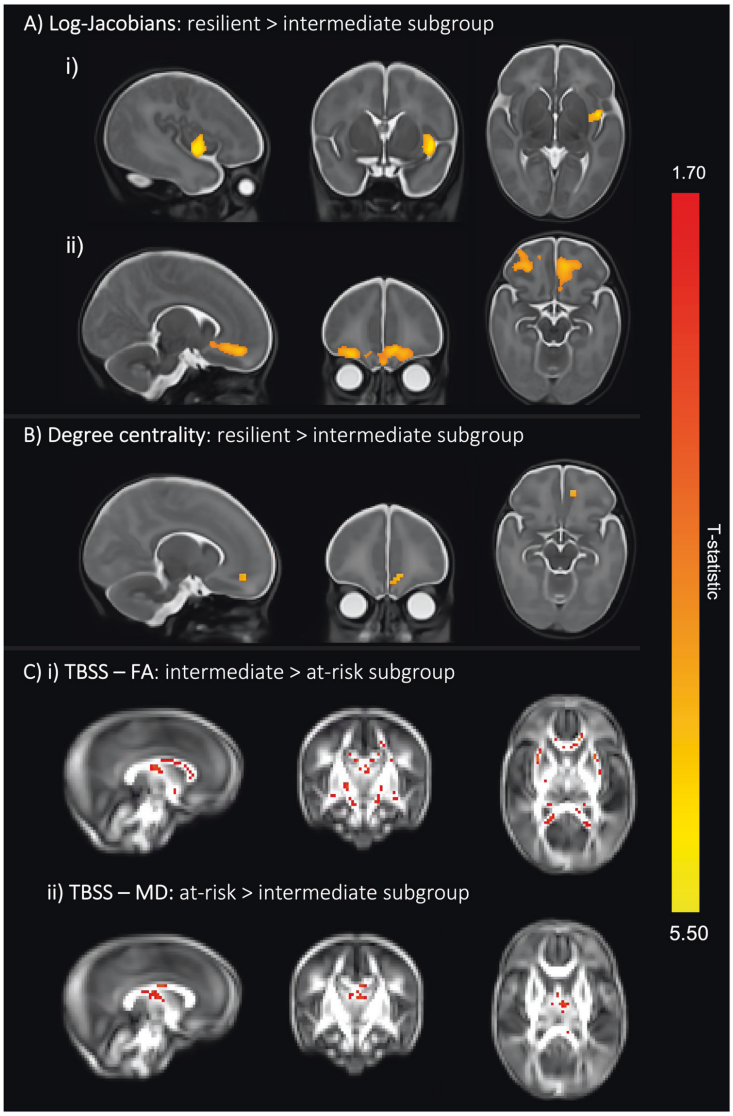



**Fig. 4 Three-cluster solution brain differences at term-equivalent age. A** Colored voxels indicate regions with significantly larger log-Jacobian determinant values in the resilient subgroup (subgroup 1) compared to the intermediate subgroup (subgroup 3) in i) left insula and the ii) bilateral orbitofrontal cortices ($p < 0.05$). GLM included sex and PMA at scan as covariates and TFCE and FWE corrections were applied. **B** Voxels showing significantly larger degree centrality values in the resilient subgroup (subgroup) 1 compared to the intermediate subgroup (subgroup 3) are seen in an overlapping left orbitofrontal region at $p < 0.05$. GLM included sex, PMA at scan and motion (standardized DVARS) as covariates; TFCE and FWE were applied. **C** Colored voxels represent white matter regions showing i) significantly higher FA values in the intermediate subgroup compared to the at-risk subgroup and ii) significantly higher MD values in the at-risk subgroup compared to the intermediate subgroup ($p < 0.05$). T-statistic values are represented in the color bar, where red colored voxels indicate smaller T-statistic values and yellow voxels indicate higher T-statistic values ranging between 1.70 and 5.50.

extra caution when interpreting such results and make increased efforts to recruit more diverse participant samples.

Additional limitations to consider include the use of parental reports for most child behavioral measures, except IQ, which could lead to common method variance bias [115] and result in underreporting of psychopathology [116]. The lack of information on familial cognitive outcomes and psychiatric history, which are heritable traits [117], prevents us from estimating trait heritability. Moreover, the small to moderate effect sizes reported for neonatal brain differences between subgroups may limit their immediate clinical meaningfulness or translatability into clinical practice. However, the fact that these brain differences only emerged after subdividing the sample into more refined and homogenous phenotypic subgroups ($C = 3$ vs $C = 2$), highlights the benefit of using advanced clustering approaches such as SNF. We speculate that these effects may be diluted in the two-cluster solution due to the presence of individuals within both (at-risk and resilient) subgroups having profiles that are more similar to an intermediate subgroup profile (please see Fig. S3).

Sensitivity analyses including one sibling only from each twin/triplet set mostly replicated the main findings, showing similar early brain patterns as well as cognitive, neonatal clinical, social, and childhood behavioral profiles for both two- and three-cluster solutions, suggesting that the effects seen here are not biased by the presence of multiple pregnancy siblings in the main analyses. While the functional connectivity results were no longer significant in the sensitivity analyses, we speculate this may be due to a loss in power associated with the reduced sample size.

In summary, using an integrative clustering approach, we were able to stratify VPT children into distinct multidimensional subgroups. A subgroup of VPT children at risk of experiencing behavioral and cognitive difficulties was characterized by high neonatal clinical complications and white matter microstructural alterations at term, whereas a resilient subgroup, with comparably favorable childhood behavioral outcomes, was characterized by increased childhood cognitive stimulation at home and larger and functionally more connected fronto-limbic brain regions at term. These results highlight a potential application of precision psychiatry, to enable meaningful inferences to be made at the individual level. Patterns of fronto-limbic brain maturation may be used as image-based biomarkers of outcomes in VPT children, while promoting enriching environments may foster more optimal behavioral outcomes. Risk stratification following VPT birth could, therefore, guide personalized behavioral interventions aimed at supporting healthy development in vulnerable children.

## DATA AVAILABILITY

Access to the dataset supporting this article can be made available upon request from the corresponding author. Code used to generate the results central to this paper can be accessed here: https://github.com/lailahadaya/preterm-ExecuteSNF.CC.

## REFERENCES

1. Johnson S, Hollis C, Kochhar P, Hennessy E, Wolke D, Marlow N. Psychiatric disorders in extremely preterm children: longitudinal finding at age 11 years in the EPICure study. J Am Acad Child Adolesc Psychiatry. 2010;49:453–463. e1
2. Mathewson KJ, Chow CHT, Dobson KG, Pope EI, Schmidt LA, Van Lieshout RJ. Mental health of extremely low birth weight survivors: A systematic review and meta-analysis. Psychol Bull. 2017;143:347–83.
3. Treyvaud K, Ure A, Doyle LW, Lee KJ, Rogers CE, Kidokoro H, et al. Psychiatric outcomes at age seven for very preterm children: rates and predictors. J Child Psychol Psychiatry. 2013;54:772–9.
4. Twilhaar ES, Wade RM, de Kieviet JF, van Goudoever JB, van Elburg RM, Oosterlaan J. Cognitive outcomes of children born extremely or very preterm since the 1990s and associated risk factors: a meta-analysis and meta-regression. JAMA Pediatr. 2018;172:361–7.
5. van Houdt CA, Oosterlaan J. van Wassenaer-Leemhuis AG, van Kaam AH, Aarnoudse-Moens CSH. Executive function deficits in children born preterm or at low birthweight: a meta-analysis. Dev Med Child Neurol. 2019;61:1015–24.
6. Johnson S, Marlow N. Preterm birth and childhood psychiatric disorders. Pediatr Res. 2011;69:11R–8R.
7. Burnett AC, Youssef G, Anderson PJ, Duff J, Doyle LW, Cheong JLY, et al. Exploring the 'preterm behavioral phenotype' in children born extremely preterm. J Dev Behav Pediatr JDBP. 2019;40:200–7.
8. Lean RE, Lessov-Shlaggar CN, Gerstein ED, Smyser TA, Paul RA, Smyser CD, et al. Maternal and family factors differentiate profiles of psychiatric impairments in very preterm children at age 5-years. J Child Psychol Psychiatry. 2020;61:157–66.
9. van Houdt CA, Oosterlaan J, Aarnoudse-Moens CSH, van Kaam AH, van Wassenaer-Leemhuis AG. Subtypes of behavioral functioning in 8–12 year old very preterm children. Early Hum Dev. 2020;142:104968.
10. Crilly CJ, Haneuse S, Litt JS. Predicting the outcomes of preterm neonates beyond the neonatal intensive care unit: What are we missing? Pediatr Res. 2021;89:426–45.
11. Cicchetti D, Rogosch FA. Equifinality and multifinality in developmental psychopathology. Dev Psychopathol. 1996;8:597–600.
12. Volpe JJ. Brain injury in premature infants: a complex amalgam of destructive and developmental disturbances. Lancet Neurol. 2009;8:110–24.
13. Wickremasinghe A, Hartman T, Voigt R, Katusic S, Weaver A, Colby C et al. Evaluation of the ability of neurobiological, neurodevelopmental and socioeconomic variables to predict cognitive outcome in premature infants. Child Care Health Dev. 2012. https://doi.org/10.1111/j.1365-2214.2011.01281.x.
14. Benavente-Fernández I, Synnes A, Grunau RE, Chau V, Ramraj C, Glass T, et al. Association of socioeconomic status and brain injury with neurodevelopmental outcomes of very preterm children. JAMA Netw Open. 2019;2:e192914–e192914.
15. Brouwer MJ, Kersbergen KJ, van Kooij BJM, Benders MJNL, van Haastert IC, Koopman-Esseboom C. Preterm brain injury on term-equivalent age MRI in relation to perinatal factors and neurodevelopmental outcome at two years. PLOS ONE. 2017;12:e0177128.
16. Levine TA, Grunau RE, McAuliffe FM, Pinnamaneni R, Foran A, Alderdice FA. Early childhood neurodevelopment after intrauterine growth restriction: a systematic review. Pediatrics. 2015;135:126–41.
17. Luo Y, Qin S, Fernández G, Zhang Y, Klumpers F, Li H. Emotion perception and executive control interact in the salience network during emotionally charged working memory processing. Hum Brain Mapp. 2014;35:5606–16.
18. Woodward LJ, Lu Z, Morris AR, Healey DM. Preschool self regulation predicts later mental health and educational achievement in very preterm and typically developing children. Clin Neuropsychol. 2017;31:404–22.
19. Poehlmann J, Gerstein ED, Burnson C, Weymouth L, Bolt DM, Maleck S, et al. Risk and resilience in preterm children at age 6. Dev Psychopathol. 2015;27:843–58.
20. Johnson S, Waheed G, Manktelow B, Field DJ, Marlow N, Draper ES, et al. Differentiating the preterm phenotype: distinct profiles of cognitive and behavioral development following late and moderately preterm birth. J Pediatr. 2018;193:85–92.e1
21. Cismaru AL, Gui L, Vasung L, Lejeune F, Barisnikov K, Truttmann A, et al. Altered amygdala development and fear processing in prematurely born infants. Front Neuroanat. 2016;10:55.
22. Rogers CE, Anderson PJ, Thompson DK, Kidokoro H, Wallendorf M, Treyvaud K, et al. Regional cerebral development at term relates to school-age social-emotional development in very preterm children. J Am Acad Child Adolesc Psychiatry. 2012;51:181–91.
23. Lee HJ, Kwon H, Kim JI, Lee JY, Lee JY, Bang S, et al. The cingulum in very preterm infants relates to language and social-emotional impairment at 2 years of term-equivalent age. NeuroImage Clin. 2021;29:102528.
24. Kanel D, Vanes L, Pecheva D, Hadaya L, Falconer S, Counsell S, et al. Neonatal white matter microstructure and emotional development during the pre-school years in children who were born very preterm. eNeuro. 2021;8:ENEURO.0546-20.2021.
25. Rogers CE, Sylvester CM, Mintz C, Kenley JK, Shimony JS, Barch DM, et al. Neonatal amygdala functional connectivity at rest in healthy and preterm infants and early internalizing symptoms. J Am Acad Child Adolesc Psychiatry. 2017;56:157–66.
26. Ramphal B, Whalen DJ, Kenley JK, Yu Q, Smyser CD, Rogers CE, et al. Brain connectivity and socioeconomic status at birth and externalizing symptoms at age 2 years. Dev Cogn Neurosci. 2020;45:100811.
27. Bogičević L, Pascoe L, Nguyen T-N-N, Burnett AC, Verhoeven M, Thompson DK, et al. Individual attention patterns in children born very preterm and full term at 7 and 13 years of age. J Int Neuropsychol Soc. 2021;27:970–80.
28. Ross GS, Foran LM, Barbot B, Sossin KM, Perlman JM. Using cluster analysis to provide new insights into development of very low birthweight (VLBW) premature infants. Early Hum Dev. 2016;92:45–49.

29. Ball G, Counsell SJ, Anjari M, Merchant N, Arichi T, Doria V, et al. An optimised tract-based spatial statistics protocol for neonates: applications to prematurity and chronic lung disease. NeuroImage. 2010;53:94–102.
30. Ball G, Aljabar P, Nongena P, Kennea N, Gonzalez-Cinca N, Falconer S, et al. Multimodal image analysis of clinical influences on preterm brain development. Ann Neurol. 2017;82:233–46.
31. Fenn-Moltu S, Fitzgibbon SP, Ciarrusta J, Eyre M, Cordero-Grande L, Chew A, et al. Development of neonatal brain functional centrality and alterations associated with preterm birth. Cereb Cortex. 2022: bhac444.
32. Wang B, Mezlini AM, Demir F, Fiume M, Tu Z, Brudno M, et al. Similarity network fusion for aggregating data types on a genomic scale. Nat Methods. 2014;11:333–7.
33. Jacobs GR, Voineskos AN, Hawco C, Stefanik L, Forde NJ, Dickie EW, et al. Integration of brain and behavior measures for identification of data-driven groups cutting across children with ASD, ADHD, or OCD. Neuropsychopharmacol Publ Am Coll Neuropsychopharmacol. 2021;46:643–53.
34. Stefanik L, Erdman L, Ameis SH, Foussias G, Mulsant BH, Behdinan T, et al. Brain-behavior participant similarity networks among youth and emerging adults with schizophrenia spectrum, autism spectrum, or bipolar disorder and matched controls. Neuropsychopharmacology. 2018;43:1180–8.
35. Hong S-J, Sisk LM, Caballero C, Mekhanik A, Roy AK, Milham MP, et al. Decomposing complex links between the childhood environment and brain structure in school-aged youth. Dev Cogn Neurosci. 2021;48:100919.
36. Edwards AD, Redshaw ME, Kennea N, Rivero-Arias O, Gonzales-Cinca N, Nongena P, et al. Effect of MRI on preterm infants and their families: a randomised trial with nested diagnostic and economic evaluation. Arch Dis Child—Fetal Neonatal Ed. 2018;103:F15–F21.
37. Wechsler D Wechsler Preschool and Primary Scale of Intelligence-fourth edition technical manual and interpretive manual. San Antonio,TX: Psychological Corporation. Copyright 2012. Pearson, Inc., and/or its affiliates. 2012. http://www.pearsonclinical.com/.
38. Sherman EMS, Brooks BL. Behavior Rating Inventory of Executive Function-Preschool Version (BRIEF-P): Test review and clinical guidelines for use. Child Neuropsychol. 2010;16:503–19.
39. Rieffe C, Ketelaar L, Wiefferink CH. Assessing empathy in young children: Construction and validation of an Empathy Questionnaire (EmQue). Personal Individ Differ. 2010;49:362–7.
40. Constantino JN, Gruber CP. Social Responsiveness Scale second edition (SRS-2): manual. 2nd ed. Western Psychological Services: Los Angeles, CA, 2012.
41. Goodman R. Psychometric properties of the strengths and difficulties questionnaire (SDQ). J Am Acad Child Adolesc Psychiatry. 2001;40:1337–45.
42. Rothbart MK, Ahadi SA, Hershey KL, Fisher P. Investigations of temperament at three to seven years: the Children's Behavior Questionnaire. Child Dev. 2001;72:1394–408.
43. Wolke D, Jaekel J, Hall J, Baumann N. Effects of sensitive parenting on the academic resilience of very preterm and very low birth weight adolescents. J Adolesc Health Publ Soc Adolesc Med. 2013;53:642–7.
44. Maechler M, Rousseeuw P, Struyf A, Hubert M, Hornik K cluster: cluster analysis basics and extensions. R package version 2.1.2 — For new features, see the 'Changelog' file (in the package source), https://CRAN.R-project.org/package=cluster. 2021.
45. Wang B, Mezlini, AM, Demir F, Fiume M, Tu Z, Brudno M *et al*. SNFtool: Similarity Network Fusion. R package version 2.3.0. https://CRAN.R-project.org/package=SNFtool. 2018.
46. Xu T, Le TD, Liu L, Su N, Wang R, Sun B, et al. CancerSubtypes: an R/Bioconductor package for molecular cancer subtype identification, validation and visualization. Bioinformatics. 2017;33:3131–3.
47. Markello RD, Shafiei G, Tremblay C, Postuma RB, Dagher A, Misic B. Multimodal phenotypic axes of Parkinson's disease. Npj Park Dis. 2021;7:1–12.
48. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics. 2010;26:1572–3.
49. Chiu DS, Talhouk A. diceR: an R package for class discovery using an ensemble driven approach. BMC Bioinforma. 2018;19:11.
50. Dag O, Dolgun A. Konar N Meric. onewaytests: An R package for one-way tests in independent groups designs. R J. 2018;10:175.
51. França L, Ge Y, Batalle D p-testR. 2022. https://doi.org/10.5281/zenodo.7051925.
52. Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, Mackay CE, et al. Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. NeuroImage. 2006;31:1487–505.
53. Lautarescu A, Hadaya L, Craig MC, Makropoulos A, Batalle D, Nosarti C, et al. Exploring the relationship between maternal prenatal stress and brain structure in premature neonates. PloS One. 2021;16:e0250413.
54. Vanes LD, Hadaya L, Kanel D, Falconer S, Ball G, Batalle D, et al. Associations between neonatal brain structure, the home environment, and childhood outcomes following very preterm birth. Biol Psychiatry Glob Open Sci. 2021;1:146–55.
55. Ball G, Aljabar P, Arichi T, Tusor N, Cox D, Merchant N, et al. Machine-learning to characterise neonatal functional connectivity in the preterm brain. Neuroimage. 2016;124:267–75.
56. Holiga Š, Hipp JF, Chatham CH, Garces P, Spooren W, D'Ardhuy XL *et al*. Patients with autism spectrum disorders display reproducible functional connectivity alterations. *Sci Transl Med* 2019; 11. https://doi.org/10.1126/scitranslmed.aat9223.
57. Power JD, Schlaggar BL, Lessov-Schlaggar CN, Petersen SE. Evidence for hubs in human functional brain networks. Neuron. 2013;79:798–813.
58. Zuo X-N, Ehmke R, Mennes M, Imperati D, Castellanos FX, Sporns O, et al. Network centrality in the human functional connectome. Cereb Cortex. 2012;22:1862–75.
59. van den Heuvel MP, Sporns O. Network hubs in the human brain. Trends Cogn Sci. 2013;17:683–96.
60. Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM. FSL. NeuroImage. 2012;62:782–90.
61. Smith SM, Nichols TE. Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. NeuroImage. 2009;44:83–98.
62. Treyvaud K, Inder TE, Lee KJ, Northam EA, Doyle LW, Anderson PJ. Can the home environment promote resilience for children born very preterm in the context of social and medical risk? J Exp Child Psychol. 2012;112:326–37.
63. Neubauer A-P, Voss W, Kattner E. Outcome of extremely low birth weight survivors at school age: the influence of perinatal parameters on neurodevelopment. Eur J Pediatr. 2008;167:87–95.
64. Anderson J, Thang CM, Thanh LQ, Dai VTT, Phan VT, Nhu BTH, et al. Immune Profiling of cord blood from preterm and term infants reveals distinct differences in pro-inflammatory responses. Front Immunol. 2021;12:777927.
65. Pariante CM. Neuroscience, mental health and the immune system: overcoming the brain-mind-body trichotomy. Epidemiol Psychiatr Sci. 2016;25:101–5.
66. Clapp M, Aurora N, Herrera L, Bhatia M, Wilen E, Wakefield S. Gut microbiota's effect on mental health: The gut-brain axis. Clin Pr. 2017;7:987.
67. Mukhopadhyay S, Lee J-J, Hartman E, Woodford E, Dhudasia MB, Mattei LM, et al. Preterm infants at low risk for early-onset sepsis differ in early fecal microbiome assembly. Gut Microbes. 2022;14:2154091.
68. Back SA. White matter injury in the preterm infant: pathology and mechanisms. Acta Neuropathol (Berl). 2017;134:331–49.
69. Lee YA. White matter injury of prematurity: its mechanisms and clinical features. J Pathol Transl Med. 2017;51:449–55.
70. Huang X, Du X, Song H, Zhang Q, Jia J, Xiao T, et al. Cognitive impairments associated with corpus callosum infarction: a ten cases study. Int J Clin Exp Med. 2015;8:21991–8.
71. Mooshagian E. Anatomy of the corpus callosum reveals its function. J Neurosci. 2008;28:1535–6.
72. Thompson DK, Inder TE, Faggian N, Warfield SK, Anderson PJ, Doyle LW, et al. Corpus callosum alterations in very preterm infants: perinatal correlates and 2 year neurodevelopmental outcomes. Neuroimage. 2012;59:3571–81.
73. Chen H-F, Huang L-L, Li H-Y, Qian Y, Yang D, Qing Z, et al. Microstructural disruption of the right inferior fronto-occipital and inferior longitudinal fasciculus contributes to WMH-related cognitive impairment. CNS Neurosci Ther. 2020;26:576–88.
74. van Duinkerken E, Schoonheim MM, IJzerman RG, Klein M, Ryan CM, Moll AC, et al. Diffusion tensor imaging in type 1 diabetes: decreased white matter integrity relates to cognitive functions. Diabetologia. 2012;55:1218–20.
75. Vollmer B, Lundequist A, Mårtensson G, Nagy Z, Lagercrantz H, Smedler A-C, et al. Correlation between white matter microstructure and executive functions suggests early developmental influence on long fibre tracts in preterm born adolescents. PloS One. 2017;12:e0178893.
76. Crespi C, Cerami C, Dodich A, Canessa N, Arpone M, Iannaccone S, et al. Microstructural white matter correlates of emotion recognition impairment in Amyotrophic Lateral Sclerosis. Cortex J Devoted Study Nerv Syst Behav. 2014;53:1–8.
77. Wier R, Aleksonis HA, Pearson MM, Cannistraci CJ, Anderson AW, Kuttesch JF, et al. Fronto-limbic white matter microstructure, behavior, and emotion regulation in survivors of pediatric brain tumor. J Neurooncol. 2019;143:483–93.
78. Herbet G, Zemmoura I, Duffau H. Functional anatomy of the inferior longitudinal fasciculus: from historical reports to current hypotheses. Front Neuroanat. 2018;12:77.
79. Hung Y, Saygin ZM, Biederman J, Hirshfeld-Becker D, Uchida M, Doehrmann O, et al. Impaired frontal-limbic white matter maturation in children at risk for major depression. Cereb Cortex. 2017;27:4478–91.
80. Lamar M, Charlton RA, Ajilore O, Zhang A, Yang S, Barrick TR, et al. Prefrontal vulnerabilities and whole brain connectivity in aging and depression. Neuropsychologia. 2013;51:1463–70.

81. Von Der Heide RJ, Skipper LM, Klobusicky E, Olson IR. Dissecting the uncinate fasciculus: disorders, controversies and a hypothesis. Brain. 2013;136:1692–707.
82. Rothbart MK, Ellis LK, Rueda MR, Posner MI. Developing mechanisms of temperamental effortful control. J Pers. 2003;71:1113–44.
83. Rothbart MK. Temperament and the pursuit of an integrated developmental psychology. Merrill-Palmer Q. 2004;50:492–505.
84. Toulmin H, Beckmann CF, O'Muircheartaigh J, Ball G, Nongena P, Makropoulos A, et al. Specialization and integration of functional thalamocortical connectivity in the human infant. Proc Natl Acad Sci USA. 2015;112:6485–90.
85. Mouka V, Drougia A, Xydis VG, Astrakas LG, Zikou AK, Kosta P, et al. Functional and structural connectivity of the brain in very preterm babies: relationship with gestational age and body and brain growth. Pediatr Radio. 2019;49:1078–84.
86. Dimitrova R, Pietsch M, Ciarrusta J, Fitzgibbon SP, Williams LZJ, Christiaens D, et al. Preterm birth alters the development of cortical microstructure and morphology at term-equivalent age. NeuroImage. 2021;243:118488.
87. McTeague LM, Rosenberg BM, Lopez JW, Carreon DM, Huemer J, Jiang Y, et al. Identification of common neural circuit disruptions in emotional processing across psychiatric disorders. Am J Psychiatry. 2020;177:411–21.
88. Rempel-Clower NL. Role of orbitofrontal cortex connections in emotion. Ann NY Acad Sci. 2007;1121:72–86.
89. Rushworth MFS, Noonan MP, Boorman ED, Walton ME, Behrens TE. Frontal cortex and reward-guided learning and decision-making. Neuron. 2011;70:1054–69.
90. Uddin LQ, Nomi JS, Hebert-Seropian B, Ghaziri J, Boucher O. Structure and function of the human insula. J Clin Neurophysiol Publ Am Electroencephalogr Soc. 2017;34:300–6.
91. Uddin LQ, Supekar KS, Ryali S, Menon V. Dynamic reconfiguration of structural and functional connectivity across core neurocognitive brain networks with development. J Neurosci J Soc Neurosci. 2011;31:18578–89.
92. Petrovic P, Ekman CJ, Klahr J, Tigerström L, Rydén G, Johansson AGM, et al. Significant grey matter changes in a region of the orbitofrontal cortex in healthy participants predicts emotional dysregulation. Soc Cogn Affect Neurosci. 2016;11:1041–9.
93. Tanzer M, Derome M, Morosan L, Salaminios G, Debban, M. Cortical thickness of the insula and prefrontal cortex relates to externalizing behavior: Cross-sectional and prospective findings. Dev Psychopathol. 2021;33:1437–47.
94. Hodel AS. Rapid infant prefrontal cortex development and sensitivity to early environmental experience. Dev Rev DR. 2018;48:113–44.
95. Li G, Nie J, Wang L, Shi F, Lin W, Gilmore JH, et al. Mapping region-specific longitudinal cortical surface expansion from birth to 2 years of age. Cereb Cortex NYN 1991. 2013;23:2724–33.
96. Hanson JL, Chung MK, Avants BB, Shirtcliff EA, Gee JC, Davidson RJ, et al. Early stress is associated with alterations in the orbitofrontal cortex: a tensor-based morphometry investigation of brain structure and behavioral risk. J Neurosci. 2010;30:7466–72.
97. Ranger M, Chau CMY, Garg A, Woodward TS, Beg MF, Bjornson B, et al. Neonatal pain-related stress predicts cortical thickness at age 7 years in children born very preterm. PLoS ONE. 2013;8:e76702.
98. Fischi-Gómez E, Vasung L, Meskaldji D-E, Lazeyras F, Borradori-Tolsa C, Hagmann P, et al. Structural brain connectivity in school-age preterm infants provides evidence for impaired networks relevant for higher order cognitive skills and social cognition. Cereb Cortex. 2015;25:2793–805.
99. Ganella EP, Burnett A, Cheong J, Thompson D, Roberts G, Wood S, et al. Abnormalities in orbitofrontal cortex gyrification and mental health outcomes in adolescents born extremely preterm and/or at an extremely low birth weight. Hum Brain Mapp. 2014;36:1138–50.
100. Giuliani NR, Drabant EM, Bhatnagar R, Gross JJ. Emotion regulation and brain plasticity: expressive suppression use predicts anterior insula volume. NeuroImage. 2011;58:10–15.
101. Fan Y, Duncan NW, de Greck M, Northoff G. Is there a core neural network in empathy? An fMRI based quantitative meta-analysis. Neurosci Biobehav Rev. 2011;35:903–11.
102. Gao W, Gilmore JH, Giovanello KS, Smith JK, Shen D, Zhu H, et al. Temporal and spatial evolution of brain network topology during the first two years of life. PLOS ONE. 2011;6:e25278.
103. Arichi T, Whitehead K, Barone G, Pressler R, Padormo F, Edwards AD, et al. Localization of spontaneous bursting neuronal activity in the preterm human brain with simultaneous EEG-fMRI. eLife. 2017;6:e27814.
104. Guo W, Liu F, Xiao C, Zhang Z, Liu J, Yu M, et al. Decreased insular connectivity in drug-naive major depressive disorder at rest. J Affect Disord. 2015;179:31–37.
105. Liu Z, Xu C, Xu Y, Wang Y, Zhao B, Lv Y, et al. Decreased regional homogeneity in insula and cerebellum: a resting-state fMRI study in patients with major depression and subjects at high risk for major depression. Psychiatry Res. 2010;182:211–5.
106. DeMaster D, Bick J, Johnson U, Montroy JJ, Landry S, Duncan AF. Nurturing the preterm infant brain: leveraging neuroplasticity to improve neurobehavioral outcomes. Pediatr Res. 2019;85:166–75.
107. Maitre NL, Key AP, Chorna OD, Slaughter JC, Matusz PJ, Wallace MT, et al. The dual nature of early-life experience on somatosensory processing in the human infant brain. Curr Biol CB. 2017;27:1048–54.
108. Milgrom J, Newnham C, Anderson PJ, Doyle LW, Gemmill AW, Lee K, et al. Early sensitivity training for parents of preterm infants: impact on the developing brain. Pediatr Res. 2010;67:330–5.
109. Gee DG, Gabard-Durnam L, Telzer EH, Humphreys KL, Goff B, Shapiro M, et al. Maternal buffering of human amygdala–prefrontal circuitry during childhood but not adolescence. Psychol Sci. 2014;25:2067–78.
110. Matsudaira I, Yokota S, Hashimoto T, Takeuchi H, Asano K, Asano M, et al. Parental praise correlates with posterior insular cortex gray matter volume in children and adolescents. PLOS ONE. 2016;11:e0154220.
111. Lordier L, Meskaldji D-E, Grouiller F, Pittet MP, Vollenweider A, Vasung L, et al. Music in premature infants enhances high-level cognitive brain networks. Proc Natl Acad Sci USA. 2019;116:12103–8.
112. Briggs-Gowan MJ, Owens PL, Schwab-Stone ME, Leventhal JM, Leaf PJ, Horwitz SM. Persistence of psychiatric disorders in pediatric settings. J Am Acad Child Adolesc Psychiatry. 2003;42:1360–9.
113. Cuthbert BN. The RDoC framework: facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology. World Psychiatry. 2014;13:28–35.
114. Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, et al. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. Am J Psychiatry. 2010;167:748–51.
115. Podsakoff PM, MacKenzie SB, Lee J-Y, Podsakoff NP. Common method biases in behavioral research: a critical review of the literature and recommended remedies. J Appl Psychol. 2003;88:879–903.
116. Mathai J, Anderson P, Bourne A. Comparing psychiatric diagnoses generated by the strengths and difficulties questionnaire with diagnoses made by clinicians. Aust NZJ Psychiatry. 2004;38:639–43.
117. McGrath JJ, Wray NR, Pedersen CB, Mortensen PB, Greve AN, Petersen L. The association between family history of mental disorders and general cognitive ability. Transl Psychiatry. 2014;4:e412.

## ACKNOWLEDGEMENTS

We would like to thank participant families and children that took part in this study as well as research radiographers, administrative staff and clinical personnel that made this work possible. This work was supported by the Medical Research Council (MRC), UK [grant numbers: MR/K006355/1 and MR/S026460/1], Action Medical Research and Dangoor Education [grant number: GN2606] and King's College London member of the MRC Doctoral Training Partnership in Biomedical Sciences [MR/N013700/1] and a MRC/Sackler Foundation grant [MR/P502108/1]. This study uses data acquired during independent research funded by the National Institute for Health Research (NIHR) Programme Grants for Applied Research Programme [grant number: RP-PG-0707-10154] and the research is supported by the NIHR Biomedical Research Center, Guy's and St Thomas' NHS Foundation Trust and King's College London, the NIHR Clinical Research Facility, Guy's and St Thomas'. The authors acknowledge use of the research computing facility at King's College London, *Rosalind* (https://rosalind.kcl.ac.uk), which is delivered in partnership with the National Institute for Health Research (NIHR) Biomedical Research Centers at South London & Maudsley and Guy's & St. Thomas' NHS Foundation Trusts, and part-funded by capital equipment grants from the Maudsley Charity [award 980] and Guy's & St. Thomas' Charity [TR130505]. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, King's College London, or the Department of Health and Social Care.

## AUTHOR CONTRIBUTIONS

LH, KD, MS, DB and CN contributed to the concept and design of the study. LH, KD, LDV, DK, SFM, OGG, SJC, MS, DB and CN contributed towards material preparation or data analysis and LH, KD, MS, DB and CN to the interpretation of data for the study. Data collection was completed by the ePrime study research team supervised by ADE and CN. LH, KD, MS, DB and CN drafted the manuscript and revised it critically for important intellectual content. All authors approved the final manuscript to be published.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s41398-023-02401-w.

**Correspondence** and requests for materials should be addressed to Chiara Nosarti.

**Reprints and permission information** is available at http://www.nature.com/reprints

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

# Appendix C

***Reference:*** *Hadaya, L., & Nosarti, C. (2020). The neurobiological correlates of cognitive outcomes in adolescence and adulthood following very preterm birth. Seminars in Fetal & Neonatal Medicine, 25(3), 101117. https://doi.org/10.1016/j.siny.2020.101117.*

Seminars in Fetal and Neonatal Medicine 25 (2020) 101117

Contents lists available at ScienceDirect

Seminars in Fetal and Neonatal Medicine

journal homepage: www.elsevier.com/locate/siny

# The neurobiological correlates of cognitive outcomes in adolescence and adulthood following very preterm birth

Laila Hadaya[a,b,1], Chiara Nosarti[b,1,*]

[a] Centre for the Developing Brain, Department of Perinatal Imaging and Health, Faculty of Life Science and Medicine, King's College London, London, United Kingdom
[b] Department of Child and Adolescent Psychiatry, Institute of Psychiatry Psychology and Neuroscience, King's College London, London, United Kingdom

ARTICLE INFO

*Keywords:*
Preterm birth
Executive function
Neuroimaging
Magnetic resonance imaging
Diffusion magnetic resonance imaging
Functional neuroimaging
Cognition
Intelligence
Adult
Adolescent

ABSTRACT

Very preterm birth (< 33 weeks of gestation) has been associated with alterations in structural and functional brain development in regions that are believed to underlie a variety of cognitive processes. While such alterations have been often studied in the context of cognitive vulnerability, early disruption to programmed developmental processes may also lead to neuroplastic and functional adaptations, which support cognitive performance. In this review, we will focus on executive function and intelligence as the main cognitive outcomes following very preterm birth in adolescence and adulthood in relation to their structural and functional neurobiological correlates. The neuroimaging modalities we review provide quantitative assessments of brain morphology, white matter macro and micro-structure, structural and functional connectivity and haemodynamic responses associated with specific cognitive operations. Identifying the neurobiological underpinning of the long-term sequelae associated with very preterm birth may guide the development and implementation of targeted neurobehaviourally-informed interventions for those at high risk.

## 1. Introduction

Individuals who were born very preterm (< 33 weeks of gestation) are at increased risk of experiencing long-term cognitive difficulties, encompassing executive function and general intelligence, compared to those born at term [1–3] Difficulties in these domains have been associated with real-life achievements in adulthood, including less satisfactory social relationships and lower academic and economic attainments [1,4].

The exact mechanisms underlying the cognitive sequelae associated with very preterm birth are poorly understood, but they are likely to involve impaired neurodevelopment, as the immature nervous system is vulnerable to exogenous and endogenous insults during the third trimester of gestation, due to its rapidly developing and complex characteristics [5]. Several studies have documented neurobiological alterations in very preterm neonates in brain regions and networks involved in high-order cognitive processes, that have been particularly clearly defined in the thalamocortical tracts [6] and have been shown to predict later neurocognitive function [6,7]. Similar patterns of alterations have been also described in adult survivors of very preterm birth and may at least partly account for their long-term cognitive difficulties [8–11]. Hence, there is an urgent need for research that can aid early identification of the neuroimaging signatures that could identify sub-groups of individuals who are at risk of cognitive sequelae, as these could guide the development and implementation of targeted neurobehaviourally-informed interventions early in life.

In this review we will discuss the results of non-invasive *in-vivo* magnetic resonance imaging (MRI) studies in very preterm adolescents and adults that have investigated brain morphology, white matter macro and micro-structure, structural and functional connectivity and haemodynamic responses associated with and potentially mediating intelligence and executive function outcomes.

We define intelligence as the "ability to understand complex ideas, to adapt effectively to the environment, to learn from experience, to engage in various forms of reasoning, to overcome obstacles by taking thought" [12]. Intelligence has been subdivided into "fluid intelligence" (cognitive processes applied in non-routine, novel or complex tasks) and "crystallised intelligence" (cognitive abilities related to previously acquired knowledge). Intelligence quotient (IQ) scores assess the performance of individuals on intelligence tests. Executive function, on the other hand, refers to cognitive processes responsible for the execution of goal-oriented tasks and include decision-making, planning, task-

* Corresponding author. Centre for the Developing Brain, Department of Perinatal Imaging and Health, Faculty of Life Science and Medicine, King's College London, London, United Kingdom.
*E-mail addresses:* laila.hadaya@kcl.ac.uk (L. Hadaya), chiara.nosarti@kcl.ac.uk (C. Nosarti).
[1] Address: Centre for the Developing Brain, 1st Floor South Wing, St Thomas Hospital, Westminster Bridge Rd, London, SE1 7 EH.

https://doi.org/10.1016/j.siny.2020.101117

Available online 15 May 2020

switching, inhibitory control, working memory, cognitive flexibility and verbal fluency [13].

Although intelligence and executive functions possess overlapping features and some studies have used the terms interchangeably, others have found that not all executive functions (for instance, the ability to inhibit prepotent responses and to shift mental sets) correlate with IQ (see Ref. [14] for review). We have previously reported lower executive function scores in very preterm adults compared to controls independently of IQ [1,15]. However, at the neuroanatomical level, it is challenging to disentangle regions specific to a single cognitive domain, as both executive function and intelligence are sub-served by a "multiple-demand" brain network, encompassing fronto-parietal cortical regions, basal ganglia, thalamus and cerebellum [16].

## 2. Grey matter morphology and cognitive outcomes

Structural MRI (s-MRI) is a technique that applies strong magnetic fields to different tissue types and creates high-resolution images of different parts of the body. s-MRI is able to clearly visualize brain anatomy at a resolution of about 1 cubic millimetre. MRI analyses are then used to quantify volumes and surface-based metrics of grey and white matter, which contain neuronal cell bodies, and myelinated axons that allow for the communication between different neuronal cells and brain structures, respectively.

Using s-MRI, several earlier studies explored structural grey matter alterations in specific regions-of-interest in very preterm individuals compared to controls (reviewed in Ref. [17]). However, only a few studies to date have investigated structural changes associated with cognitive outcomes. These have focused on brain regions that are known to be vulnerable to hypoxic-ischemic injury following preterm birth, such as the hippocampus and the cholinergic basal forebrain, which are also implicated in the cognitive sequelae of very preterm birth. In very preterm adults, Aanes and colleagues [18] reported positive associations between hippocampal volumes and memory scores, whereas Grothe and colleagues [19] showed that larger cholinergic basal forebrain volumes were associated with higher IQ. Moreover, basal forebrain volumes were found to mediate the relationship between IQ and neonatal clinical complications, measured as an index of respiratory support, feeding dependency and neurological status severity. These results have potential clinical relevance, suggesting that novel cholinergic pharmacological interventions that have improved cognitive functions in other samples, such as individuals with Alzheimer's disease [20], could be trialled to support the development of very preterm children with a history of neonatal clinical complications.

s-MRI has been further used to investigate between-group differences in the whole-brain. Overall, findings have described widespread grey matter volume and surface area alterations in very preterm adults and adolescents compared to controls in temporal, frontal and occipital cortices, as well as subcortical regions encompassing the caudate nucleus and the thalamus [10,21,22]. The volumes of numerous regions found to be smaller in very preterm individuals (compared to controls) have been associated with gestational age (i.e., the more pronounced the alterations, the younger gestational age) and with poorer cognitive outcomes that have been studied as contributing to the long-term sequelae of very preterm birth, e.g. executive function and intelligence [10,11,18,21].

However, most studies in very preterm adults have been cross-sectional, and therefore causality between brain alterations and specific outcomes cannot be inferred. Only a few studies to date have investigated the developmental trajectories of brain structure between adolescence and adulthood [23–25]. Rimol and colleagues [25] found that cortical thickness decreased from 15 to 20 years of age in a similar way in both very preterm individuals and controls, whereas Nam and colleagues [23] reported differential cortical maturational trajectories between the groups. They used cortical thickness measurements at mid-adolescence to predict alterations five years later and achieved a mean accuracy of 86.5% in identifying spatially discriminating features between very preterm individuals and controls in temporal, occipito-temporal, parietal and prefrontal cortices. Within these regions, longitudinal cortical changes were associated with performance on executive function tasks probing response inhibition and mental flexibility [23]. These findings suggest that similar approaches could be used to predict long lasting brain alterations in atypically developing samples and their associated cognitive functions could represent targets for intervention.

In the same cohort of study participants studied by Nam [23], Karolis and colleagues [24] included a third MRI scan at the age of 30 years and showed accelerated age-dependent grey matter volume maturation in very preterm individuals between mid-adolescence and adulthood. They also found that grey matter volume in lateral parieto-temporal cortices was not altered in very preterm individuals, and that this displayed a stronger association with IQ in the very preterm compared to control group, reflecting potential functional adaptation to support cognitive performance.

## 3. White matter macro- and microstructure and cognitive outcomes

s-MRI can also be used to investigate between-group changes in white matter macrostructure (i.e. size and volume). Similar to studies investigating grey matter, earlier reports used a region-of-interest approach to quantify white matter in regions that are vulnerable to injury associated with preterm birth. One of these regions is the corpus callosum, the largest bundle of commissural fibers in the brain and a pathway of crucial importance to coordinate interhemispheric integration and specialization required for complex multimodal cognitive operations. Consistent findings reported smaller callosal surface areas in very preterm individuals compared to controls, and associations between surface area reductions and poorer executive functions and IQ [26–28].

Using a whole-brain approach, we also previously described widespread white matter macrostructural alterations (both volume decreases and increases in very preterm individuals compared to controls) in temporal, parietal and frontal regions and in the major fasciculi [10,11]. In adulthood, we found that white matter volume reductions in a region beneath the left inferior gyrus accounted for 14% of the variance in participants' full-scale IQ, while smaller values in corpus callosum and thalamus explained 21% of the variance in executive functions [11].

Diffusion MRI (d-MRI) can be additionally used to quantify white matter macrostructure, as well as its microstructure. d-MRI measures the displacement of water molecules in tissue over time. Indices such as fractional anisotropy (FA) quantify the directional dependence of water molecule motion in each voxel, whereas mean (MD), axial (AD) and radial diffusivity (RD) reflect the average diffusion within a voxel, the magnitude of diffusion parallel to the principal axis of diffusion (considered to be along the axis of a white matter tract) and the magnitude of diffusion perpendicular to the primary direction of diffusion, respectively. Smaller fractional anisotropy or larger mean diffusivity values reflect poorer white matter microstructural integrity.

Lower FA in tracts connecting the fronto-to-temporal, fronto-to-parietal and temporal-to-occipital lobes and in projection fibres including the corticospinal and anterior thalamic radiations were described in very preterm adolescents compared to full-term controls. FA values, full-scale IQ and executive function scores (working memory and cognitive flexibility) were positively associated in the very preterm group [29]. Lower FA and higher MD in major white matter tracts were also associated with lower IQ scores in very preterm adults, but not in controls [30]. Our previous work showed that lower FA in the genu of the corpus callosum, which connects the right and left prefrontal regions that are centrally involved in executive processing, and in the right superior longitudinal fasciculus association fibre tract, correlated

with lower IQ in very preterm adults [31].

Using tractography, a three-dimensional d-MRI modelling technique that outlines white matter pathways connecting different brain regions, we previously reported reduced volume of the ventral cingulum, a tract connecting temporal to parietal cortices, in very preterm adults compared to controls, which was associated with lower IQ [32]. Similarly, other studies have described higher MD in the corticospinal tract [33] and lower volumes in the fornix and cingulum [34]. However, tract specific features have been seldomly studied in relation to intelligence and executive function.

In addition to investigating individual tracts, d-MRI allows us to study whole brain structural connectivity. This is a particularly informative approach, as the brain is heavily interconnected and comprised of collections of subnetworks or "modules", which allow for an integrative mode of information transfer across different cognitive, sensory and motor modalities [35]. In particular, there exists a set of regions known as the "rich-club", which are more densely connected with one another than they are with other regions outside the network. The "rich-club" is involved in integrative global functions including high order cognitive functions. We previously reported that very preterm adults exhibited a stronger "rich-club" architecture than controls, despite possessing fewer white matter resources [36]. To explore "rich-club" network robustness we further used a simulated lesion approach and deleted components one at a time from the connectivity matrix (i.e., nodes and edges). We found that 'lesioning' striatal-cortical connections produced greater alterations in global connectivity in the very preterm group compared to controls, reflecting their altered role in supporting a global exchange of information throughout the brain. Furthermore, network alterations had functional implications as they correlated with measures of information flow and rule learning.

Taken together, results of s-MRI and d-MRI studies highlight associations between alterations in brain morphology and white matter macro- and microstructure in the preterm brain and cognitive difficulties. Results of these studies therefore suggest that the brain areas associated with worse cognitive outcomes overlap with those which are vulnerable to anatomical alterations following very preterm birth.

It is important to note that although the majority of studies has investigated white and grey matter alterations separately following very preterm birth, these structural changes tend to co-occur. Early white matter injury in the immature brain has been associated with reduced cortical complexity and volume, suggesting synergistic mechanisms contributing to impaired maturation of oligodendroglia and of the developing grey matter [37]. Meng et al. [22] described lower FA in widespread brain regions including the cerebellum in very preterm adults, which was associated with reduced thalamus and striatum grey matter volumes and with lower full-scale IQ and younger gestational age. Corpus callosum FA reductions were also found to mediate reduced frontal lobe cortical thickness in very preterm adults [38]. Therefore, future studies should consider investigating the underlying neurobiological correlates of cognitive outcomes using multimodal approaches.

## 4. Intrinsic functional connectivity and cognitive outcomes

Functional-MRI (fMRI) quantifies neural activity by measuring a blood oxygen level dependent (BOLD) signal, which refers to the increase in blood flow to the local blood vessels that accompanies neural activity in the brain. The temporally related BOLD activations of distinct brain regions 'at rest' (i.e., resting-state fMRI (rs-fMRI)) measure the spontaneous, low frequency fluctuations in the BOLD signal to characterise synchronous activations in large-scale spatially distinct regions, in order to determine dissociable resting-state brain networks that intrinsically work together during rest.

The increase in global connectivity between networks in the brain has been associated with increased executive function abilities from adolescence to adulthood [39]. Networks involved in executive functions remain sparsely connected in childhood but develop over time [40]. Central to executive processing is the default mode network (DMN), which includes the medial prefrontal, posterior cingulate, parietal, lateral temporal cortices and hippocampus. The DMN is actively implicated in internal or introspective thoughts, but shows suppressed engagement during externally focused goal-directed behaviour. As cognitive demands increase, functional connectivity reconfiguration occurs between the DMN and other networks, such as the salience and fronto-parietal network [41].

Reduced connectivity between the DMN and other networks involved in high-order cognitive processing (i.e., salience and central executive networks) was described in very preterm compared to full-term born adults [42,43], and was more evident in individuals born at younger gestational ages [42]. White and colleagues (2014), did not observe direct associations between reduced between-network connectivity and cognitive outcome measures, despite very preterm adults displaying lower executive function scores [42]. However, perinatal risk factors (i.e. severity of neonatal brain injury and gestational age) moderated the relationship between between-network connectivity and executive function.

In the Bavarian Longitudinal Study, lower full-scale IQ was associated with decreased right fronto-parietal intrinsic functional connectivity (i.e., left middle temporal/lateral occipital gyrus) in very preterm adults [44]. However, this relationship disappeared after controlling for childhood mathematical abilities, which were in turn positively associated with decreased fronto-parietal intrinsic functional connectivity. As working memory is required to support mathematical operations and fronto-parietal networks have been associated with working memory processing [45], it could be speculated that early working memory impairments following very preterm birth may mediate general cognitive abilities in adulthood. Thus, working memory training interventions in childhood may improve IQ later in life [46].

## 5. Task based functional MRI and cognitive outcomes

During task-based fMRI, several images of the brain are acquired over a period of time, while participants are completing a task in the scanner, and BOLD signal fluctuations are recorded.

Several studies have investigated the haemodynamic responses associated with high-order cognitive processing in very preterm individuals. We previously probed inhibitory control, which refers to the capacity to repress inappropriate responses and is crucial for the development of executive function. We found that very preterm adolescents performed as well as their full-term peers on a motor response inhibition task. However, they exhibited a complex pattern of altered haemodynamic responses compared to controls, which included decreases in prefrontal (inferior frontal and anterior cingulate gyri) and subcortical areas (caudate nucleus, thalamus, globus pallidus) and in the bilateral cerebellum; and increases in prefrontal (insula, orbitofrontal and dorsolateral prefrontal cortices), bilateral temporal-occipital regions and caudate nucleus [47] The hyper-activation patterns were interpreted as neural mechanisms potentially compensating for the frontal and subcortical hypo-activation patterns, in order to support the successful completion of the inhibitory control task. We later used the same task in very preterm adults and showed altered haemodynamic responses, i.e., hyper-activation in the middle temporal and posterior cingulate gyri and precuneus [48], suggesting that functional alterations may represent long-lasting neural adaptations.

A closely related function to inhibitory control is executive attention, which refers to the ability to focus attention on a task-relevant feature during interference by task-irrelevant stimuli and has been shown to be selectively impaired in both preterm children and adults [49,50]. An fMRI investigation of executive attention by Daamen and colleagues (2015) in adults born very preterm failed to reveal any between-group differences in task-related fronto-cingulo-parietal activation. However, in this study, gestational age and clinical risk variables were associated with haemodynamic responses in dorsal anterior

cingulate and lateral occipital regions, suggesting that neonatal complications modulated executive attention processing [50].

Similar to what was observed in inhibitory control, fMRI findings using verbal fluency highlight potentially compensatory haemodynamic patterns that support letter fluency processing. Verbal fluency refers to the strategic retrieval of words and relies on executive functions, such as cognitive flexibility and shifting, as well as semantic categorisation. A typical verbal fluency task involves the overt generation of a word starting with a specific letter (i.e., phonological verbal fluency). Very preterm adults have been found to perform as well as controls on this type of tasks, although between-group haemodynamic response differences were observed [51,52]. In very preterm young adults compared to controls these were observed in the caudate nucleus, which is involved in the inhibition of irrelevant words and goal-directed language production [53], the anterior cingulate cortex, which is involved in decision making [54], as well as frontal, temporal and parietal regions and the thalamus and insula during the more demanding verbal fluency trials [51,52]. However, we found that very preterm adults performed worse than controls on the 'hard' letter trials and that poorer task performance was associated with decreased haemodynamic response suppression in the right sensorimotor cortex [55]. In this context, response suppression may represent an inhibitory mechanism to reduce distracting neural processes and to support visual attention, thus facilitating adequate performance on verbal fluency and working memory tasks [56]. Therefore, the decreased haemodynamic response suppression observed in very preterm adults may have been underlying their verbal fluency disadvantage.

Other fMRI studies in very preterm adults have probed working memory, which is crucial for optimal learning and development and has been frequently found to be compromised in younger very preterm cohorts (for a summary see Ref. [57]). During completion of a working memory task, very preterm adults showed DMN deactivation and fronto-cingulo-parietal and subcortical activation patterns that were similar to those observed in controls. However, during the more demanding working memory trials, very preterm adults exhibited greater parahippocampal and posterior DMN deactivations compared to controls, possibly reflecting compensatory processes to achieve satisfactory task performance [58]. We previously noted that working memory performance in very preterm adults was affected by perinatal clinical risk. Very preterm adults who had sustained severe perinatal brain injury (i.e., periventricular haemorrhage and ventricular dilation) performed worse than those with less severe injuries on more demanding working memory trials and displayed load-dependent decreases in haemodynamic responses in left inferior frontal gyrus [52]. In a different study, however, we observed that increased activation in the perisylvian cortex acted as a compensatory mechanism in very preterm adults who had suffered severe perinatal brain injury as this activation correlated with better working memory performance in this subgroup, but not in very preterm adults with less severe injuries and term-born controls [59]. This also indicates that structural alterations tend to accompany changes in functional activation.

Taken together, the results of structural and functional MRI studies suggest that very preterm birth could be regarded as a lifelong neurodevelopmental condition. However, the relationship between structural and functional alterations remains poorly understood as only a few studies to date have used more than one imaging modality (i.e. multimodal imaging) [34,55,59,60]. Multimodality represents a thorough approach in the search for potential biomarkers of cognitive deficits following very preterm birth, as it gains insight into the integration of structural and functional features associated with specific outcomes that may not be captured using single modalities. For instance, multimodal analysis has been performed in other neurodevelopmental disorders, such as attention deficit hyperactivity disorder, extracting multimodal components that share the same across-subject variation. This approach has revealed co-occurring and mechanistically related anatomical and functional alterations [61]. Such studies prompt future research to investigate neurobiological correlates of cognitive outcomes using multimodal imaging, in order to fully elucidate and understand the neurobiology underlying cognitive outcomes following very preterm birth.

## 6. Conclusions

Non-invasive MRI techniques have identified associations between alterations in brain morphology, white matter macro and micro-structure, structural and functional connectivity, functional haemodynamic responses and long-term cognitive outcomes following very preterm birth in adult life. However, there is more work to be done as neurobiological correlates have not yet shown to accurately predict cognitive outcomes. Advanced diffusion, functional and volumetric MRI offers the opportunity to improve significantly on routine neuroimaging and to define underlying neuroanatomical features associated with specific outcomes. Using machine learning and non-subjective clustering of imaging and clinical data, Ball and colleagues defined novel imaging features associated with antenatal and postnatal adversity and which predicted adverse outcome [62]. A promising approach is to combine longitudinal multimodal MRI with advanced predictive statistical models, in order to go beyond group-level comparison and achieve individual-level clinical predictions.

Identifying the underlying neurobiological correlates of the long-term sequelae associated with very preterm birth may guide the development and implementation of targeted neurobehaviourally-informed interventions for those at high risk and may facilitate the identification of new behavioural targets for improving cognitive outcomes.

### Practice points

- Very preterm adolescents and adults display structural and functional brain alterations in regions implicated in a variety of cognitive processes
- Some alterations have been associated with lower IQ and poorer executive function, while others may represent neuroplastic and functional adaptations, which support cognitive performance
- The brain areas associated with worse cognitive outcomes overlap with those which are vulnerable to perinatal brain injury and medical complications following very preterm birth

### Research directions

- Multimodal imaging studies are needed to better understand mechanistically related anatomical and functional alterations associated with specific cognitive outcomes
- Longitudinal studies are needed to investigate the structural and functional growth trajectories that precede the development of high-order cognitive impairments
- Advanced statistical analyses coupled with multimodal imaging are required to achieve individual-level clinical predictions

### Acknowledgements

This work was supported by the Medical Research Council (UK) (grant no: MR/S026460/1).

### References

[1] Kroll J, Karolis V, Brittain PJ, et al. Real-life impact of executive function impairments in adults who were born very preterm. J Int Neuropsychol Soc

2017;23:381–9.
[2] Johnson S, Marlow N. Early and long-term outcome of infants born extremely preterm. Arch Dis Child 2017;102:97–102.
[3] Skranes J, Løhaugen GCC. Reduction in general intelligence and executive function persists into adulthood among very preterm or very low birthweight children. Evid Base Ment Health 2016;19:e28.
[4] Jaekel J, Baumann N, Bartmann P, Wolke D. General cognitive but not mathematic abilities predict very preterm and healthy term born adults' wealth. PloS One 2019;14:e0212789.
[5] Volpe JJ. Brain injury in premature infants: a complex amalgam of destructive and developmental disturbances. Lancet Neurol 2009;8:110–24.
[6] Ball G, Pazderova L, Chew A, et al. Thalamocortical connectivity predicts cognition in children born preterm. Cereb Cortex N Y NY 2015;25:4310–8.
[7] Loh WY, Anderson PJ, Cheong JLY, et al. Neonatal basal ganglia and thalamic volumes: very preterm birth and 7-year neurodevelopmental outcomes. Pediatr Res 2017;82:970–8.
[8] Menegaux A, Meng C, Neitzel J, et al. Impaired visual short-term memory capacity is distinctively associated with structural connectivity of the posterior thalamic radiation and the splenium of the corpus callosum in preterm-born adults. Neuroimage 2017;150:68–76.
[9] Salvan P, Tournier JD, Batalle D, et al. Language ability in preterm children is associated with arcuate fasciculi microstructure at term. Hum Brain Mapp 2017;38:3836–47.
[10] Nosarti C, Giouroukou E, Healy E, et al. Grey and white matter distribution in very preterm adolescents mediates neurodevelopmental outcome. Brain J Neurol 2008;131:205–17.
[11] Nosarti C, Nam KW, Walshe M, et al. Preterm birth and structural brain alterations in early adulthood. NeuroImage Clin 2014;6:180–91.
[12] Neisser U, Boodoo G, Bouchard Jr. TJ, et al. Intelligence: knowns and unknowns. Am Psychol 1996;51:77–101.
[13] Diamond A. Executive functions. Annu Rev Psychol 2013;64:135–68.
[14] Duggan EC, Garcia-Barrera MA. Executive functioning and intelligence. Handbook of intelligence: evolutionary theory, historical perspective, and current concepts. New York, NY, US: Springer Science + Business Media; 2015. p. 435–58.
[15] Nosarti C, Giouroukou E, Micali N, Rifkin L, Morris RG, Murray RM. Impaired executive functioning in young adults born very preterm. J Int Neuropsychol Soc JINS 2007;13:571–81.
[16] Duncan J. The Structure of cognition: attentional episodes in mind and brain. Neuron 2013;80:35–50.
[17] de Kieviet JF, Zoetebier L, van Elburg RM, Vermeulen RJ, Oosterlaan J. Brain development of very preterm and very low-birthweight children in childhood and adolescence: a meta-analysis. Dev Med Child Neurol 2012;54:313–23.
[18] Aanes S, Bjuland KJ, Skranes J, Løhaugen GCC. Memory function and hippocampal volumes in preterm born very-low-birth-weight (VLBW) young adults. Neuroimage 2015;105:76–83.
[19] Grothe MJ, Scheef L, Bäuml J, et al. Reduced cholinergic basal forebrain integrity links neonatal complications and adult cognitive deficits after premature birth. Biol Psychiatr 2017;82:119–26.
[20] Ferreira-Vieira TH, Guimaraes IM, Silva FR, Ribeiro FM. Alzheimer's disease: targeting the cholinergic system. Curr Neuropharmacol 2016;14:101–15.
[21] Østgård HF, Sølsnes AE, Bjuland KJ, et al. Executive function relates to surface area of frontal and temporal cortex in very-low-birth-weight late teenagers. Early Hum Dev 2016;95:47–53.
[22] Meng C, Bäuml JG, Daamen M, et al. Extensive and interrelated subcortical white and gray matter alterations in preterm-born adults. Brain Struct Funct 2016;221:2109–21.
[23] Nam KW, Castellanos N, Simmons A, et al. Alterations in cortical thickness development in preterm-born individuals: implications for high-order cognitive functions. Neuroimage 2015;115:64–75.
[24] Karolis VR, Froudist-Walsh S, Kroll J, et al. Volumetric grey matter alterations in adolescents and adults born very preterm suggest accelerated brain maturation. Neuroimage 2017;163:379–89.
[25] Rimol LM, Bjuland KJ, Løhaugen GCC, et al. Cortical trajectories during adolescence in preterm born teenagers with very low birthweight. Cortex J Devoted Study Nerv Syst Behav 2016;75:120–31.
[26] Narberhaus A, Segarra D, Caldú X, et al. Gestational age at preterm birth in relation to corpus callosum and general cognitive outcome in adolescents. J Child Neurol 2007;22:761–5.
[27] Narberhaus A, Segarra D, Caldú X, et al. Corpus callosum and prefrontal functions in adolescents with history of very preterm birth. Neuropsychologia 2008;46:111–6.
[28] Nosarti C, Rushe TM, Woodruff PWR, Stewart AL, Rifkin L, Murray RM. Corpus callosum size and very preterm birth: relationship to neuropsychological outcome. Brain 2004;127:2080–9.
[29] Vollmer B, Lundequist A, Mårtensson G, et al. Correlation between white matter microstructure and executive functions suggests early developmental influence on long fibre tracts in preterm born adolescents. PloS One 2017;12. e0178893.
[30] Eikenes L, Løhaugen GC, Brubakk A-M, Skranes J, Håberg AK. Young adults born preterm with very low birth weight demonstrate widespread white matter alterations on brain DTI. Neuroimage 2011;54:1774–85.
[31] Allin MPG, Kontis D, Walshe M, et al. White matter and cognition in adults who were born preterm. PloS One 2011;6:e4525.
[32] Caldinelli C, Froudist-Walsh S, Karolis V, et al. White matter alterations to cingulum and fornix following very preterm birth and their relationship with cognitive functions. Neuroimage 2017;150:373–82.
[33] Jurcoane A, Daamen M, Scheef L, et al. White matter alterations of the corticospinal tract in adults born very preterm and/or with very low birth weight. Hum Brain Mapp 2016;37:289–99.
[34] Tseng C-EJ, Froudist-Walsh S, Brittain PJ, et al. A multimodal imaging study of recognition memory in very preterm born adults. Hum Brain Mapp 2017;38:644–55.
[35] van den Heuvel MP, Kersbergen KJ, de Reus MA, et al. The Neonatal connectome during preterm brain development. Cerebr Cortex 2015;25:3000–13.
[36] Karolis VR, Froudist-Walsh S, Brittain PJ, et al. Reinforcement of the brain's rich-club architecture following early neurodevelopmental disruption caused by very preterm birth. Cerebr Cortex 2016;26:1322–35.
[37] Dean JM, Bennet L, Back SA, McClendon E, Riddle A, Gunn AJ. What brakes the preterm brain? An arresting story. Pediatr Res 2014;75:227–33.
[38] Rimol LM, Botellero VL, Bjuland KJ, et al. Reduced white matter fractional anisotropy mediates cortical thickening in adults born preterm with very low birthweight. Neuroimage 2019;188:217–27.
[39] Stevens MC, Skudlarski P, Pearlson GD, Calhoun VD. Age-related cognitive gains are mediated by the effects of white matter development on brain network integration. Neuroimage 2009;48:738–46.
[40] Fair DA, Cohen AL, Dosenbach NUF, et al. The maturing architecture of the brain's default network. Proc Natl Acad Sci Unit States Am 2008;105:4028.
[41] Vatansever D, Menon DK, Manktelow AE, Sahakian BJ, Stamatakis EA. Default mode dynamics for global functional integration. J Neurosci 2015;35:15254–62.
[42] White TP, Symington I, Castellanos NP, et al. Dysconnectivity of neurocognitive networks at rest in very-preterm born adults. NeuroImage Clin 2014;4:352–65.
[43] Konishi M, McLaren DG, Engen H, Smallwood J. Shaped by the Past: the default mode network supports cognition that is independent of immediate perceptual input. PloS One 2015;10:e0132209.
[44] Bäuml JG, Meng C, Daamen M, et al. The association of children's mathematic abilities with both adults' cognitive abilities and intrinsic fronto-parietal networks is altered in preterm-born individuals. Brain Struct Funct 2017;222:799–812.
[45] Soreq E, Leech R, Hampshire A. Dynamic network coding of working-memory domains and working-memory processes. Nat Commun 2019;10:1–14.
[46] Løhaugen GCC, Antonsen I, Håberg A, et al. Computerized working memory training improves function in adolescents born at extremely low birth weight. J Pediatr 2011;158:555–61. e4.
[47] Nosarti C, Rubia K, Smith AB, et al. Altered functional neuroanatomy of response inhibition in adolescent males who were born very preterm. Dev Med Child Neurol 2006;48:265–71.
[48] Lawrence EJ, Rubia K, Murray RM, et al. The neural basis of response inhibition and attention allocation as mediated by gestational age. Hum Brain Mapp 2009;30:1038–50.
[49] Burnett AC, Anderson PJ, Lee KJ, et al. Trends in executive functioning in extremely preterm children across 3 birth eras. Pediatrics 2018;141:e20171958.
[50] Daamen M, Bäuml JG, Scheef, et al. Neural correlates of executive attention in adults born very preterm. NeuroImage Clin 2015;9:581–91.
[51] Nosarti C, Shergill SS, Allin MP, et al. Neural substrates of letter fluency processing in young adults who were born very preterm: alterations in frontal and striatal regions. Neuroimage 2009;47:1904–13.
[52] Kalpakidou AK, Allin MPG, Walshe M, et al. Functional neuroanatomy of executive function after neonatal brain injury in adults who were born very preterm. PloS One 2014;9:e113975.
[53] Ali N, Green DW, Kherif F, Devlin JT, Price CJ. The role of the left head of caudate in suppressing irrelevant words. J Cognit Neurosci 2010;22:2369–86.
[54] Lavin C, Melis C, Mikulan EP, Gelormini C, Huepe D, Ibanez A. The anterior cingulate cortex: an integrative hub for human socially-driven interactions. Front Neurosci 2013;7:64.
[55] Tseng C-EJ, Froudist-Walsh S, Kroll J, et al. Verbal fluency is affected by altered brain lateralization in adults who were born very preterm. eNeuro 2019;6. e0274-18.
[56] Tomasi D, Wang R, Wang G-J, Volkow ND. Functional connectivity and brain activation: a synergistic approach. Cerebr Cortex 2014;24:2619–29.
[57] Taylor HG, Clark CAC. Executive function in children born preterm: risk factors and implications for outcome. Semin Perinatol 2016;40:520–9.
[58] Daamen M, Bäuml JG, Scheef L, et al. Working memory in preterm-born adults: load-dependent compensatory activity of the posterior default mode network. Hum Brain Mapp 2015;36:1121–37.
[59] Froudist-Walsh S, Karolis V, Caldinelli C, et al. Very early brain damage leads to remodeling of the working memory system in adulthood: a combined fmri/tractography study. J Neurosci 2015;35:15787–99.
[60] Olsen A, Dennis EL, Evensen KAI, et al. Preterm birth leads to hyper-reactive cognitive control processing and poor white matter organization in adulthood. Neuroimage 2018;167:419–28.
[61] Wu Z-M, Llera A, Hoogman M, et al. Linked anatomical and functional brain alterations in children with attention-deficit/hyperactivity disorder. NeuroImage Clin 2019;23:101851.
[62] Ball G, Aljabar P, Nongena P, et al. Multimodal image analysis of clinical influences on preterm brain development. Ann Neurol 2017;82:233–46.