Non-CTIMP Study Protocol

PREterm birth as a determinant of Neurodevelopment and COGnition in children: mechanisms and causal evidence

The PRENCOG study

PRENCOG
PRETERM NEURODEVELOPMENT COGNITION

Co-Sponsors	The University of Edinburgh and/or Lothian Health Board ACCORD The Queen's Medical Research Institute 47 Little France Crescent Edinburgh EH16 4TJ
Protocol authors	Investigators: James P Boardman, Rebecca M Reynolds, Hilary Richardson, Cheryl Battersby, Simon R Cox, Riccardo Marioni, Thanasis Tsanas, Heather Whalley, David Batty. Collaborators: Louise Arseneault, Ruth Andrew, Mark Bastin, Lauren Ingledow, Lee Murphy, Sarah Stock Early career researcher: Manuel Blesa Cabez Programme manager: Jill Hall
Funder	UKRI Medical Research Council
Funding Reference Number	MR/X003434/1
Chief Investigator	Professor James Peter Boardman
Sponsor number	AC23050
REC Number	23-SS-0067
Project registration	IRAS 325276
Version Number and Date	Version 1.1, 01/11/23

CONTENTS

LIST OF ABBREVIATIONS

ACCORD	Academic and Clinical Central Office for Research & Development - Joint office for The University of Edinburgh and Lothian Health Board
ADHD	Attention deficit hyperactivity disorder
BOLD	Blood oxygen level dependent
CI	Chief investigator
CpG	Cytosine-phosphate-guanine (*5′—C—phosphate—G—3′*)
CRF	Case report form
CRP	C-reactive protein
DNAm	DNA methylation
EoP	Encephalopathy of prematurity
FA	Fractional anisotropy
GA	Gestational age
GCP	Good Clinical Practice
HC	Hierarchical complexity
HIPAA	Health Insurance Portability and Accountability Act
HPA	Hypothalamic-pituitary-adrenal
ICH	International Conference on Harmonisation
ICVF	intracellular volume fraction of water
IQ	Intelligence quotient
LC-MS/MS	Liquid chromatography tandem mass spectrometry
MD	Mean diffusivity
MRI	Magnetic resonance imaging
MSN	Morphometric similarity network
MTsat	Magnetisation transfer imaging
NDI	Neurite density index
NICU	Neonatal intensive care unit
NNRD	National Neonatal Research Database
NODDI	Neurite orientation dispersion and density imaging

NPD	National Pupil Database
ODI	Orientation dispersion index
PAG	Parental advisory group
PI	Principal investigator
PSMD	Peak-width skeletonised
PTB	Preterm birth
PTB-RF	Preterm birth-associated risk factor
QA	Quality assurance
REC	Research Ethics Committee
RT-PCR	Reverse transcription polymerase chain reaction
SAIL	Secure Anonymised Information Linkage
ScotXeD	Scottish Exchange of Data
SEN	Special Edu
SES	Socioeconomic status
(S)IMD	(Scottish) Index of Multiple Deprivation
SLC	Speech, language or communication
SOP	Standard Operating Procedure
SQA	Scottish Qualifications Authority
TEBC	Theirworld Edinburgh Birth Cohort

1 INTRODUCTION

1.1 BACKGROUND

Globally, preterm birth (PTB), defined as birth at less than 37 weeks of gestation, affects 15 million pregnancies per annum, with country prevalence between 5 and 18%[1]. Over the past two decades, the survival rate of children born preterm has improved due to advances in perinatal medicine but outcomes remain challenging: 10-15% of children born very preterm (<32 weeks) develop cerebral palsy, 30-50% develop an intellectual disability, and this population is at increased risk of problems with socialisation, behaviour, language, low educational attainment, autism, and attention deficit hyperactivity disorder. Adults who were born preterm are more likely to experience a mood disorder, age-related cognitive impairment, schizophrenia, and cardiometabolic disease[2]. PTB accounts for one of the highest numbers of disability-adjusted life years of any single childhood condition[3]**. There are no effective treatments for improving brain health after preterm birth, which brings into sharp focus the need to identify protective factors and intervention targets.**

The neurobiological basis for adverse neurological, cognitive and psychiatric outcomes following preterm birth is related to cerebral white matter injury and subsequent dysmaturational processes in white matter and neuroaxonal structures collectively termed the 'encephalopathy of prematurity' (EoP)[4]. With features of EoP apparent on magnetic resonance imaging (MRI), this has become an important assessment modality for investigating determinants of brain health in preterm infants[5 6].

Our premise, based on studies showing that adverse outcomes are not inevitable[2 7], is that it is not PTB *per se* that has a deleterious effect on brain development but rather, it is multiple, often interacting PTB-associated risk factors (PTB-RFs). These are biological, psychosocial, and social/infrastructural, and can affect parent or child, or be shared, for instance: maternal/infant stress, infection/inflammation, suboptimal infant nutrition, co-morbidities of PTB, and socioeconomic deprivation (Figure 1).

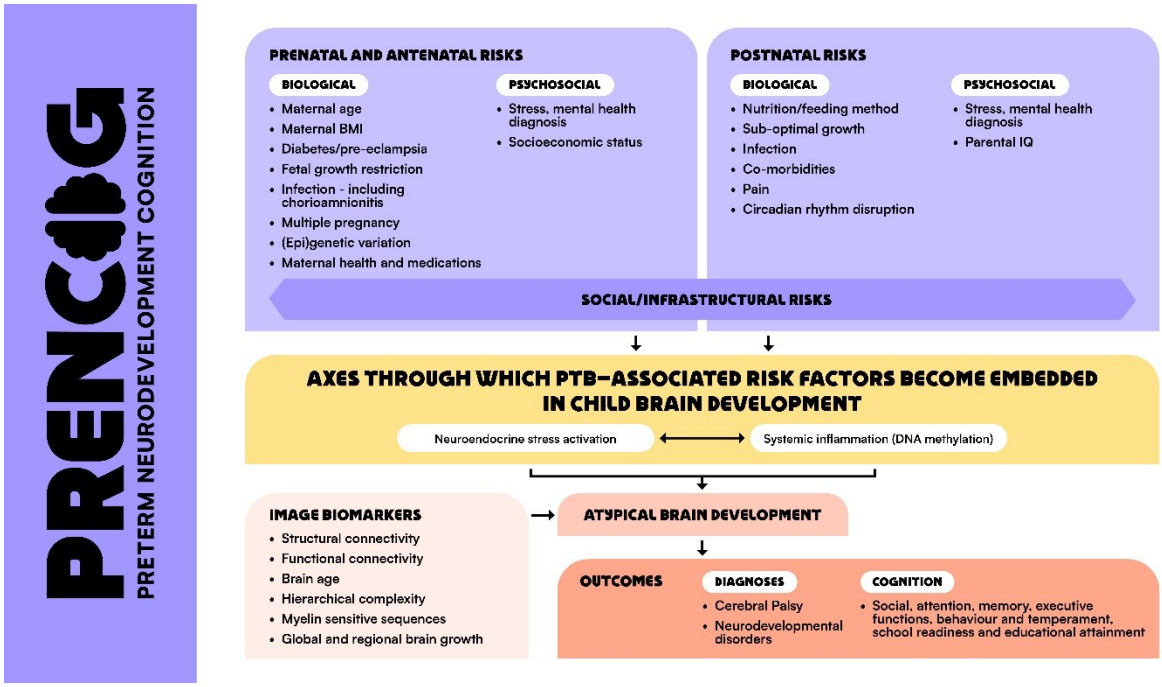

Figure 1. PTB-associated risk factors linked with altered cognition in children, proposed biological pathways that transmit risk to atypical brain development/outcome, and image biomarkers for delineating upstream pathways and predicting risk and resilience

1.2 RATIONALE FOR STUDY

To intervene against the negative effects of PTB on child development requires a quantitative understanding of PTB as a complex multi-dimensional risk exposure and new knowledge about *how* PTB-RFs modify brain development**.**

1.2.1 The perinatal stress environment and outcomes after preterm birth: Prenatal exposure to maternal stress affects 10–35% of children worldwide and is associated with adverse neuropsychiatric outcomes[8]. Adaptation of the maternal hypothalamic-pituitary-adrenal (HPA) axis with consequent variation in transfer of glucocorticoids to the developing fetus appears to be a key mechanism linking maternal stress to offspring neurodevelopment[9-11].

Our recent studies suggest this could be an important axis for embedding PTB-RFs in brain development. First, maternal hair cortisol concentrations during pregnancy are associated with newborn amygdala architecture across the whole GA range, indicating that HPA axis activation links the prenatal stress environment to a key neural substrate of socioemotional development in childhood[12]. Second, alterations in placental expression of genes regulating cortisol regeneration and placental transfer consistent with increased fetal glucocorticoid exposure occur in association with lower maternal socioeconomic status[13]. Third, maternal consumption of glycyrrhizin (a potent inhibitor of placental 11β-hydroxysteroid dehydrogenase type 2, the "barrier" to maternal glucocorticoids), is associated with adverse neurodevelopmental and neuropsychiatric outcomes in children[14]. Fourth, extremely preterm infants (<28 weeks) tend to have blunted cortisol reactivity to vaccination at 4 months, suggesting low GA (or a co-exposure such as repeated painful experiences during neonatal intensive care) programmes HPA axis adaptation. Fifth, neonatal hair glucocorticoids are a marker of both prenatal and postnatal physiological stressors in preterm infants[15]. Finally, chronic HPA axis activation is a plausible mechanistic link between early life stress, altered brain morphology and major depression in adulthood[16]. Based on these studies, we propose that atypical HPA axis activity is triggered by PTB-RFs and is an axis through which multidimensional exposures become embedded in the brain development of preterm infants.

1.2.2 Systemic inflammation and encephalopathy of prematurity: Early studies revealed that neurodevelopmental outcomes are worse if infants are exposed to co-morbidities of preterm birth characterised by systemic inflammation, for example, chorioamnionitis, bloodstream infection, and necrotising enterocolitis[17 18]. This is because inflammation alters oligodendrocyte precursor responses, increases proliferation and death, and impairs maturation into myelin-forming oligodendrocytes[19]. The consequent hypomyelination deprives axons of metabolic/trophic support and insulation for electrical impulse conduction, resulting in EoP.

Emerging evidence indicates that PTB is associated with sustained inflammation[20 21]. Specific mediators of the adaptive and immune responses to preterm birth and its co-morbidities have been linked to MRI features of EoP[22]; however, there are inconsistencies in the broader literature associating inflammation with neurodevelopment, in part, because of the absence of standard peripheral biomarkers of low-level systemic chronic inflammation in neonates, and partly because study designs have relied on a single (or low frequency) measurement of selected proteins that are highly phasic, maturation-dependent, and subject to swift and rapid concentration changes in plasma. There are no effective markers of low-level systemic chronic inflammation. We propose that DNA methylation (DNAm) profiles may provide a more accurate reflection of inflammatory exposure. This assertion is based on three key observations: specific DNAm profiles exist in inflammatory diseases[23 24], they predict levels of inflammatory proteins and neuroinflammation-related outcomes including brain structure and cognition in children and adults[25-29], and DNAm proxies have greater longitudinal stability and stronger associations with cognition than serum measures[30 31]. These observations are of particular interest because age- and birth weight- related differences in DNAm are present across a large number of CpGs[32-34]. Recently, we have shown that PTB is associated with profound and widely distributed changes in the methylome (saliva), and these are linked to MRI features of EoP[35].

1.2.3 Neural mechanisms of cognitive development in children born preterm: The neuroprofile of preterm children includes atypical social cognition[36], attention, memory[2], and language[37]. An open question is the extent to which these phenotypes are domain-specific, or whether domain-general traits underpin these challenges. For example, resolving whether social difficulties are due to reduced ability

to reason about other peoples' minds versus problems with different aspects of cognition - such as language or attention - is critical for developing mechanistically-informed interventions to promote social development. These alternative accounts cannot be distinguished behaviourally because both lead to impaired performance on behavioural tasks. Functional magnetic resonance imaging (fMRI), however, has potential to clarify the cognitive and neural mechanisms of impairments because different brain regions support different cognitive processes. Continuing the exemplar of social cognition, adults and children recruit a network of brain regions when engaging in social reasoning including bilateral temporal parietal junction (TPJ), precuneus, and medial prefrontal cortex. We have shown that this network becomes increasingly functionally specialised during early childhood[38] and that children who perform better on behavioural Theory of Mind tasks, an indicator of reasoning about others' minds, have more functionally mature, selective responses in the right TPJ[39 40] and more correlated activity between brain regions within this cortical network[38]. If children born preterm have domain-specific atypicalities in social cognition, these atypicalities should be accompanied by atypical development in social brain regions.

Recent fMRI studies provide evidence for early disruptions in functional connectivity (measured during rest) in infants born preterm[41] and, importantly, early functional connectivity in preterm infants predicts motor and cognitive outcomes at age two and can be promoted with interventions (e.g., playing music during neonatal care)[42]. To our knowledge, there have not been task-based fMRI studies of cognition in early childhood (i.e., ages 2-5 years) following preterm birth, in part because of methodological challenges of carrying out fMRI in young children. We have shown that fMRI is feasible with children as young as three years old with the use of strategies that prepare children for MRI scans and child-friendly movie-viewing fMRI experiments[38], and that movie-viewing fMRI experiments provide a particularly sensitive neural marker of individual differences in cognition, relative to measures collected during rest[43 44]. In preparation for this application, we secured a Wellcome Trust Institutional Strategic Support Fund award to set-up task-based fMRI in 5-year-olds, so we are ready to implement the protocols in our proposed study.

Elucidation of the pathways that link PTB-RFs with abnormal brain development and outcome is essential for developing neuroprotective strategies based on re-programming stress response systems[45], immunomodulation, and cognitive training.

Hypotheses

Work package 1:
1) Specific preterm birth associated risk factors (PTB-RFs), which may be clinical/demographic, psychosocial (e.g. maternal stress), or social/infrastructural (e.g. deprivation), modify the relationship between GA at birth and neurodevelopmental and/or educational outcomes

Work package 2:
2i) Atypical activation of the HPA axis leads to EoP, indexed on MRI by dysmaturity (altered chronological brain age), reduced connectome complexity, and biomarkers of hypomyelination

2ii) DNAm proxies of chronic systemic inflammation are present in preterm infants at term equivalent age and are associated with MRI features of EoP.
2iii) The effect of PTB-RFs identified in WP1 on brain development are mediated by alterations in the neonatal HPA axis and/or chronic systemic inflammation.

Work package 3*:*
3i) Preterm children have domain-specific delays in cognition at school age
3ii) Structure-function networks underpinning cognition are altered in preterm children
3iii) PTB-RFs identified in WP1 predict structural and functional maturity at 5-years

This protocol describes a programme of work packages. Permissions and approvals for each work package will be managed separately. New approvals will be sought for WP1 which will be led and managed by investigators based at Imperial College London. New approvals will also be sought for WP2 which will be led and managed by investigators based at the University of Edinburgh. WP3 already has approvals in place and is managed by University of Edinburgh investigators (Research Ethics

Committee No: 16/SS/0154, NHS Lothian R&D No: 2016/0255). Permissions are not required for WP4. Detailed information about these arrangements can be found in section 12.1 Ethical Conduct.

2 STUDY OBJECTIVES

2.1 OBJECTIVES

The overarching aim of this programme is to determine the biological, psychosocial and socioeconomic PTB-RFs that lead to adverse neurodevelopmental, cognitive, and educational outcomes in children born preterm and to identify the neuroendocrine and epigenetic axes that embed these risks in brain development. This aim will be met through 5 objectives:

2.1.1. Investigate the contributions of multidimensional PTB-RFs to neurodevelopmental and educational outcomes of children born preterm. We will create novel linkages between the National Neonatal Research Database, which contains routine data for ~100,000 infants born in the UK with gestational age (GA) <32 weeks between 01/01/2008-31/12/2019, with education records and neighbourhood deprivation indices. This will determine the relative contributions of demographic, clinical and socioeconomic PTB-RFs to neurodevelopmental and educational outcomes of British children.

2.1.2. Identify the biological axes underlying abnormal brain development in preterm infants. We will characterise brain dysmaturation associated with PTB using neonatal magnetic resonance imaging and use this to investigate the relationship between a) Hypothalamic pituitary-adrenal (HPA) axis activity, and; b) low-level chronic, systemic inflammation indexed by DNA methylation (DNAm), and brain development.

2.1.3. Identify the neural substrates associated with impaired cognitive and behavioural outcomes of preterm children at school age. We will use multiple neuroimaging modalities to identify structural and functional brain networks that predict outcomes (e.g., social development, attention) at 5 years. The contribution of potentially modifiable factors to childhood outcomes, including PTB-RFs identified in the first objective, will be investigated.

2.1.4. Identify neural markers to assess efficacy in future trials of neuroprotective therapies. Our studies linking HPA axis activity, DNAm markers of inflammation and quantitative MRI measures of brain development with outcomes will provide imaging biomarkers that can be used to design and test the efficacy of potential treatments in future studies.

2.1.5. Involve parents and survivors of preterm birth in research. We will engage and involve parents and advocacy groups (e.g., the Adult Preemie Advocacy Network) to guide the design, delivery and dissemination of the programme and to plan follow-on work that arises from it[46 47]. We will appoint a virtual parental advisory group (PAG) of 6-8 parents together. The PAG will co-design parent and public-facing materials and advise on dissemination plans. We will commission an artist to work with the PAG to develop pictorial/visual/animated ways of communicating.

We will with the public engagement officer at the MRC CRH, Ms Ginnie Clark, to run engagement events that may include:

- Co-ordinated patient information days to support current patients and their support networks
- Panel events at Science Festivals combining experts and past patients/parents of patients, a policy maker (Member of Scottish Parliament), and a health journalist.
- Awareness raising initiatives inviting target audiences, such as Instagram Lives (via UoE Medical School), Twitter Q&A, a MRC Festival event.
- The creation of a scientific animation to illustrate the subject, research insights, and offer accessible and digestible information to families requiring further support.
- Influencing Policy makers: Invite MP/MSP to learn more about our research into Preterm Birth and the possible outcomes, and how families can be better supported.
- Participation in the University of Edinburgh public lecture series.

2.2 ENDPOINTS

The programme is organised in 4 work packages, each with specific endpoints:

2.1.1 Endpoint WP1

We will define, at population level, the weighted contributions of multidimensional PTB-RFs to neurodevelopmental outcomes and real-world educational performance of children born preterm. Identifying the role of societal as well as clinical/demographic factors could be of particular importance to resource-poor settings.

2.1.2 Endpoint WP2

We will identify targets in neuroendocrine stress and immune pathways that lead to atypical brain development, and elucidate the PTB-RFs that activate those pathways.

2.2.3 Endpoint WP3

We will define the functional and structural neural substrates of critical cognitive functions in preterm children, determine the extent to which behavioural atypicalities are domain-specific versus domain-general, and, using perinatal data, characterize factors that explain the neural bases of cognitive impairment and resilience at 5-years (WP3).

2.2.4 Endpoint WP4

We will engage and involve parents and advocacy to guide the design, delivery and dissemination of the programme and to plan follow-on work that arises from it (WP4).

3 STUDY DESIGN

WP1 is a national population-based cohort study: >100,000 infants born in the UK with gestational age (GA) <32 weeks between 01/01/2008 & 31/12/2019 (National Neonatal Research Database, NNRD). Records will be linked between three databases: NNRD, National Pupil Database and the ScotXed Pupil Census.

WP2 is an exposure-based short-term cohort study: 300 mother-infant dyads (200 preterm with GA <32 weeks [exposed], 100 term with GA >37 weeks [non-exposed comparators]), recruited from the Edinburgh Royal Infirmary. The preterm group will be identified from infants born to women who present with threatened preterm birth and for whom delivery is planned or expected at less than 32 weeks GA. Term controls will be identified from the labour ward or antenatal clinic.

WP3 is an exposure-based longitudinal cohort study: 300 children (200 preterm, 100 term) are participants in a long-term cohort study of the effect of preterm birth on brain growth[48].

WP4 is a participatory project: it will engage and involve a parent advisory group (n=6-8) and the Adult Preemie Advocacy Network.

4 STUDY POPULATION

4.1 NUMBER OF PARTICIPANTS

WP1: >100,000 infants born in the UK with gestational age (GA) <32 weeks between 01/01/2008 & 31/12/2019 (National Neonatal Research Database, NNRD).

WP2: 300 mother-infant dyads: 200 preterm deliveries with GA <32 weeks (exposed, cases); 100 term deliveries with GA >37 weeks (non-exposed, comparators).

WP3: 300 children (200 preterm, 100 comparators).

4.2 INCLUSION CRITERIA

WP1: All neonates with GA< 32 weeks born in Great Britain and admitted to a neonatal unit. Data are available from England and Wales from 2012, and from Scotland from 2014.

WP2: Exposed: 300 preterm infants born at <32 weeks of gestational age (GA). Non-exposed comparators: 100 term infants born at >37 weeks of GA*.

Preterm infants are included if a mother booked her pregnancy and delivered at Simpson Centre for Reproductive Health (SCRH, the study centre), or if a mother booked her pregnancy at a hospital outside the study centre but was transferred to it with her baby *in utero* due to planned or expected birth <32 weeks. Preterm infants who are transferred to SCRH *ex utero* for intensive care are not included.

*GA is estimated based on first trimester ultrasound.

WP3: 300 participants (200 preterm infants and 100 comparators) in Theirworld Edinburgh Birth Cohort will be invited for MRI and cognitive assessments at age 5 years[48].

4.3 EXCLUSION CRITERIA

WP1: Neonatal death is the only exclusion criterion.

WP2: i) Preterm infants who are transferred to the study centre postnatally for intensive care; ii) Infants with congenital anomalies: structural or functional anomalies (e.g., metabolic disorders) that occur during intrauterine life and can be identified prenatally, at birth or later in life (WHO definition); iii) Infants with a contraindication to MRI at 3Tesla; iv) Preterm infants who are transferred to SCRH ex utero for intensive care are not included.

WP3: i) Children with congenital structural or functional anomalies (e.g., metabolic disorders) diagnosed after the neonatal period; ii) Children with acquired brain injury (e.g., traumatic brain injury, post meningitis); iii) Children with contraindication to MRI at 3Tesla.

4.4 CO-ENROLMENT

WP1: Not applicable.

WP2: The SCRH is an academic perinatal medicine centre that hosts observational research studies, and it is a recruiting centre for randomised controlled trials of therapies designed to improve the outcome of preterm infants and their mothers. Parents/carers of TEBC participants are encouraged to consider entry into such studies if eligible. Co-enrolment is informed by POL008 Co-enrolment Policy, produced by the Academic and Clinical Central Office for Research and Development (ACCORD), which is a partnership between the University of Edinburgh and NHS Lothian Health Board. Co-enrolment will be recorded.

WP3: Co-enrolment for 5-year-old TEBC participants will be managed in line with the ACCORD Co-enrolment policy in WP2 in accordance with existing approvals from National Research Ethics Service (NRES) 16/SS/0154 and NHS Lothian Research and Development (2016/0255).

5 PARTICIPANT SELECTION AND ENROLMENT

5.1 IDENTIFYING PARTICIPANTS

WP1: Not applicable.

WP2: Women who present to the Simpson Centre for Reproductive Health with threatened preterm labour and for whom delivery is planned or expected at less than 32 weeks GA. The comparator group (non-exposed term infants) will be born to women who attend the Simpson Centre for Reproductive Health for antenatal care or delivery at >37 weeks GA. Potential participants will be identified using NHS systems: maternity TRAK and the neonatal electronic patient record. This will be restricted to clinically qualified members of the research team with honorary contracts for work in NHS Lothian Women's and Children's services, and members of the direct clinical care team. Both groups are trained in NHS Lothian's Information Governance procedures. The initial approach to potential participants will be by a member of the direct care team.
No person without existing permission to access NHS Lothian maternity electronic systems will be granted permission to identify participants. Only clinically qualified members of the research team who hold honorary contracts with NHS Lothian Women's and Children's services will identify potential participants. They will view NHS systems from within the NHS firewall within the RIE. They will not share the details of potential participants identified in the screening process with anyone other than the direct healthcare team. Clinically qualified members of the research team are not part of the direct care team.

WP3: Not applicable because participants already enrolled in a longitudinal study, TEBC.

5.2 CONSENTING PARTICIPANTS

WP1: Not applicable.

WP2: Consent to enter the study must be sought from each participant only after a full explanation has been given, an information leaflet offered and time allowed for consideration. Signed participant consent will be obtained in two stages for the preterm group: first, for data collection from antenatal period to the first week of postnatal life, and; second for data and samples over the rest of the neonatal period to the end of the study. Signed participant consent for all aspects of the study will be obtained in one stage for the comparator group. The right of the participant to refuse to participate without giving reasons will be respected, and will not affect the treatment they receive. All participants are free to withdraw at any time from the assessment protocol treatment without giving reasons and without prejudicing further treatment.

Consent to link to maternal routine samples before discard will be sought.

Consent to link participants with routinely collected data in health, social care and education, and consent to recontact for follow-on studies subject to additional funding will be sought.

Informed consent may only be taken by a member of the research team with training in GCP and procedures for research involving children and young people.

WP3: Consent procedures for data acquisition at 5-years have been approved by National Research Ethics Service (NRES) 16/SS/0154.

5.2.1 Withdrawal of Study Participants

Participants are free to withdraw from the study at any point without necessarily providing reasons why they chose to withdraw, or a participant can be withdrawn by the Investigator. If withdrawal occurs, the primary reason for withdrawal will be documented in the participant’s case report form if possible. To safeguard rights, the minimum personally-identifiable information possible will be collected.

6 STUDY ASSESSMENTS

6.1 STUDY ASSESSMENTS

WP1: Not applicable.

WP2: See Table 1.

WP3: See Table 2.

Table 1. WP2 assessments

Data collection point	Age	Data collection method	Sample type / domain	Test / task	Time implication for participants	Notes
1	**Antenatal**	**administrative /electronic records & interview**	Medical, deomgraphic, SES	Ethnic background and language spoken at home; parents' education and employment; family income; family structure, housing, neighbourhood quality, parents' mental health, social network and support	45 mins	
				History and exposures: life events, prescribed medications, alcohol, smoking, substances, pregnancy complications		
2	**Birth**	**administrative /electronic records, questionnaire & tissue**	Medical	Peripartum history and exposures, mother and infant	30 mins	
				Anthropometry		
			Placenta	Structured histopathology rating and storage mRNA levels of glucocorticoid related genes	n/a	Collect and store
			Umbilical cord blood	i) 2ml umbilical cord blood; ii) dried blood spot for storage	n/a - sampling linked to clinical episodes	i) Endogenous glucocorticoids and metabolites (glucocorticoid release); glucocorticoid receptor in cord blood leukocytes (glucocorticoid signalling). ii) Inflammatory markers and DNA (collect and store)
			Hair, infant	Overall glucocorticoid secretion	15 mins	
			Hair, maternal	Overall glucocorticoid secretion	15 mins	
			Saliva	Methylome	10 mins	Term controls
3	**Neonatal**	**tissue**	Dried blood spot	Inflammatory markers and DNA	n/a - sampling linked to clinical episodes	Collect and store
		tissue	Saliva	Methylome	10 mins	Preterms at term equivalent age
		tissue	Hair, infant	Overall glucocorticoid secretion	15 minutes	Preterms at term equivalent age
		biosample	Faeces	Microbiome	n/a	Collect and store: early (7-14 days) stool (cases and controls) and pre-discharge (cases)
		administrative /electronic records & direct observation	Medical	Anthropometry	n/a - linked to clinical episodes	
			Co-morbidities and exposures	Co-morbidities of preterm birth, medications, feed type and method; health status of control group.	n/a - sampling linked to clinical episodes	
			Parent IQ	National Adult Reading Test	15 mins	
		MRI	Brain structure and connectivity	sMRI, dMRI	90 minutes	Morphometric similarity netwroks, hierarchical complexity, magnetisation transfer imaging
		administrative /electronic records & questionnaire	Demographics & medical	Update perinatal history	15mins	
				Edinburgh post-natal depression scale	5 mins	
				Parenting daily hassles	10 mins	
				WHO-QOL	10 mins	
				Adult temperament questionnaire	10 mins	

Table 2. WP3 assessments

Age	Data collection method	Sample type / domain	Test / task	Time implication for participants	Notes
5 years	tissue: saliva	Epigenetics	Methylome	10mins	
		HPA axis	Cortisol	Timed: waking, 30 minutes after waking and before bed	
	direct observation	Anthrompometry	Growth	5 mins	All
		Blood pressure	Hypertension	5 mins	All
		Ophthalmology	Refraction and acuity	20 mins	All
		Social cognition	Parent-child play; *Theory of Mind booklet task	30 mins	
		Executive function	Inhibition (Early Childhood Inhibition Touchscreen Task), Prohibition, and Spatial Working Memory	20 mins	All - direct assessment of EF
		Exploratory play	Novel toy task	5 mins	
		Reading	Woodcock-Johnson IV subscales	30 mins	All - direct assessment of reading
		IQ	Mullen Scales of Early Learning	40 mins	All - direct assessment of IQ
	*MRI	Brain structure, connectivity, and function	sMRI, dMRI, fMRI	60 mins	
	eye-tracking	Social cognition	Face scanning	20 mins	Existing evidence that these distinguish between children born preterm and at term in infancy. Longitudinal profile by analysing with data from earlier timepoints
			Face pop-out		
			Social preferential looking		
		Processing / learning speed	Novel cue-target association (Posner task)		
	parent questionnaire	Medical and demographics	Diagnoses, health service utilisation	20 mins	
		Temperament / general	Strengths and Difficulties Questionnaire	5 mins	Capture sub-diagnostic traits and school readiness
		Language	Children's Communication Checklist	15 mins	Capture language including pragmatics and usage
		Autism	Social Communication Questionnaire: Current	5 mins	Capture social cognitive development
		ADHD	DUPaul ADHD rating scale	5 mins	Capture sub-diagnostic ADHD traits
		Executive function	Behaviour Rating Inventory for Executive Function (BRIEF)	15 mins	Parent-report measure of executive function
		Quality of life	WHO-QOL	10 mins	
		COVID impact	Centre for Developing Brain Questionnaire	15 mins	Parent report of pandemic impact on child and family
		Feedback	Feedback form monitoring statisfaction with research project	5 mins	Evaluate acceptability of research setting, measures, info etc.
		Visual perception	Cerebral Visual Impairment Inventory	20 mins	
	teacher questionnaire	School milestones	CAIDS-Q	5 mins	Capture sub-diagnostic traits and school readiness
		Temperament / general	Strengths and Difficulties Questionnaire	5 mins	
	parent interview	Milestones	Record development against basic milestones, check diagnostic status	5 mins	Provide context for other measures
		General development	Vineland adaptive behaviour scales: parent rating forms	20 mins	

6.2 LONG TERM FOLLOW UP ASSESSMENTS

WP1: Not applicable.

WP2: Study assessments end in the neonatal period (timepoint 3, see Table 1). Consent will be sought to re-contact for follow-on studies, subject to funding.

WP3: Participants are involved in a longitudinal study with separate arrangements for long term follow-up assessments.

6.3 STORAGE AND ANALYSIS OF SAMPLES

Three categories of imaging/biological samples will be collected:

Neuroimaging (neonatal and a five-year brain MRI). Neuroimaging data will be stored on a dedicated General-purpose computing on graphics processing units (GPGPU) server, accessible to members of the research team. Neuroimaging summary statistics will be stored in *.CSV format and unprocessed neuroimaging data in standard NIFTI (*.nii) format, organised according to the BIDS specification**.** Processing software will be shared using UoE Research Services Version control - GitLab platform.

HPA axis activity (umbilical cord blood and maternal and neonatal hair). Laboratory analyses of corticosteroids and their precursors and metabolites in plasma (2ml) and hair (>0.3cm 2cm from neonates, up to 3cm from mothers) and saliva (at 5 years on waking, 30 minutes after waking and bed-time) will be conducted at the University of Edinburgh Clinical Research Facility Mass Spectrometry Core. We have developed a robust method for steroid extraction from plasma (100µL) and tissues[49], with quantification of cortisol and related corticosteroids including cortisone, as well as dexamethasone and its metabolites, simultaneously by liquid chromatography tandem mass spectrometry (LC-MS/MS), using a Sciex QTRAP® 6500 (Warrington, UK) operated in positive ion electrospray ionisation with a Waters Acquity™ UPLC system (Manchester, UK)[50]. We will include recently characterised DNAm signatures of HPA axis-immune activity such as FKBP5 (FK506-binding protein 51)[51 52]. Umbilical cord blood and placental tissue (WP2) will be stored in the Edinburgh Reproductive Tissue Biobank (ERTBB) prior to analysis in Edinburgh Clinical Research Facility Mass Spectrometry Core. Any residual material will be returned to ERTBB for storage and future use subjects to new ethical approval. Maternal and infant hair will be stored in MRC CRH at room temperature before transfer to University of Edinburgh Clinical Research Facility Mass Spectrometry Core (Professor James Boardman is shared custodian with Professor Rebecca Reynolds (Co-Investigator).

DNA methylation profile (saliva)**.** DNA from saliva will be extracted using prepIT.L2P reagent (DNA Genotek, Ontario, Canada). DNA will be bisulfite converted and methylation measured using Illumina HumanMethylationEPIC BeadChip (Illumina, San Diego, CA, USA) at the Edinburgh Clinical Research Facility Genetics Core (ECRF), Edinburgh, UK; the epigenetic measure of chronic inflammation will be calculated for each participant[27]. Transcriptions will be .CSV format. Raw data is .IDAT file format. All formats and software platforms support sharing and long-term validity. Remaining DNA will be stored at ECRF for possible future analysis subject to new ethical approvals.

Dried blood spots (umbilical cord blood). Blood spots will be collected using Schleicher and Schuell 903 filterpaper (6 x 3.2mm spots per subject). Cards will be stored at -20°C in the MRC CRH and analysed in batch, subject to funding[21 22 48].

Placenta. Samples will be stored at -80°C in the Edinburgh Reproductive Tissue BioBank (ERTBB); Professor James Boardman is shared custodian with Professor Rebecca Reynolds (Co-Investigator).

Gut microbiota (neonatal faeces)**.** The gut microbiome plays a role in human health and disease including child development[53], and it is modified by age at birth, sex, mode of delivery, antibiotic exposure, and feed type[54 55]. The microbiome may mediate interactions of the gut-brain axis[56 57]. Faecal samples will be collected at two time points: early stool, 7 to 14 days after birth (all infants in WP2) and

stool pre-discharge from hospital (for preterm group). Samples will be processed and stored at -80°C in the MRC CRH for later analyses, subject to funding.

Our intention is that the data derived by this programme will be available for use by the research and policy communities in perpetuity. **(i) Neuroimaging**; raw neuroimaging data acquired in Edinburgh are archived indefinitely. The image processing server is linked by automatic daily back-up to the off-site Edinburgh Compute and Data Facility (ECDF). The compute component of ECDF, Eddie, has built-in redundancy with error check and "self-healing,' and is itself backed-up using 2 tape systems. **(ii) HPA axis**; data will be stored securely under password protection using UoE DataShare, which is backed-up and maintained through UoE IT systems. **(iii) DNAm**; Raw .IDAT files will be stored on GEO indefinitely.

We have costed for University of Edinburgh storage with guaranteed backup and resilience for all new data generated. In the long-term, the University of Edinburgh provides a DataVault data retention service to archive golden copy research data (https://www.ed.ac.uk/information-services/research-support/research-data-service/after/datavault), which is also appropriate for sensitive data types; no charges apply to ≤100GB, which the phenotypic dataset will not exceed. A golden/master copy of new data will be migrated onto this service as part of the process of data curation at the end of the project.

HPA axis and DNAm data will be stored on UoE Research Data Management servers, supported by Information Services. The data in the RDM file-store is automatically replicated to an off-site disaster facility and is backed up with a 60-day retention period, with 10 days of file history online.

7 DATA COLLECTION

WP1. The National Neonatal Research Database contains quality-assured data (the Neonatal Data Set, an NHS Information Standard; DAPB1595)[58]. Information security management at the University of Edinburgh (Edinburgh Parallel Computer Centre, EPCC) is externally accredited as ISO 27001 compliant (Certificate number #276767-2018-AIS-GBR-UKAS). Records of infants will be linked to the National Pupil Database (NPD) in an NHS Trusted Research Environment (e.g. iCARE, the Imperial NHS Trust TRE) and ScotXed/SQA Pupil-Census in Scotland using date of birth, postcode and sex. De-identified datasets will be analysed using secure UoE or Imperial computers by members of the research team.

The NNRD contains data from >100,000 infants in the required gestational range and 2-year outcome data are available for >60% (>70% in most recent years). Record linkage and data analysis will be carried out within the research team.

WP2 and 3. Contact data and research data will be stored separately, linked by a unique participant identifier. Electronic data will be collected and managed using REDCap (a secure, web-based application) and stored on a University of Edinburgh server that is hosted and managed by the Edinburgh Clinical Trials Unit. This protocol is in line with the University's Data Policy and the Data Protection Act of 1998. Data will be entered to a REDCap database with an SOP to govern data completeness, accuracy and validity.

Neuroimaging. Full details of the neonatal imaging protocol are given in a published open-access paper[48] and the 5-year protocol has been developed through a completed pilot study. All MRI is acquired using a brain optimised research scanner within a tight quality-controlled environment at Edinburgh Imaging. Neuroimaging data will be collected by brain research radiographers and subject to Edinburgh Imaging QC procedures. Structural MRI data are reported according to a structured report by a paediatric radiologist, including for incidental findings.

HPA axis activity. We will collect the following samples: i) endogenous glucocorticoids and metabolites (*glucocorticoid release*) in umbilical cord blood; ii) glucocorticoid receptor in cord blood leukocytes (*glucocorticoid signalling*); iii) maternal and infant hair cortisol at birth and infant hair cortisol at term

equivalent age (*overall glucocorticoid secretion*). Sample collection will conform to an SOP, and analysis will use published, validated methods for measuring metabolites from hair and blood at the UoE Clinical Research Facility Mass Spectrometry Core [49 50]. The 5-year saliva cortisol samples (on waking, 30 minutes after waking and bed-time) are being collected as part of the TEBC protocol.

DNAm. Saliva DNAm will be sampled from preterm infants at term equivalent age to reflect the allostatic load of perinatal inflammation. Saliva will be collected in Oragene OG-575 Assisted Collection kits, by DNA Genotek, and according to an SOP. DNAm analysis using the Illumina Human MethylationEPIC BeadChip is established in the Edinburgh Clinical Research Facility Genetics Core.

Questionnaire and records. Demographic and clinical information will be extracted from the maternal and infant record by a member of the research team. The tools to assess cognition, behaviour, well-being and family circumstances are listed in Tables 1 and 2. Hard copies will be stored securely in locked filing cabinets in a locked room to which only study researchers have access.

The programme manager will oversee data collection, ensuring that time-points are adhered to for each participant. Participants will be contacted by post and / or telephone to arrange appointments and to prompt if questionnaires have not been returned, to maximise completeness of data collection.

7.1 Source Data Documentation

WP1, source data are: NNRD, the National Pupil Database, and the ScotXeD Pupil Census. Record-linkage will take place in an NHS Trusted Research Environment

WP2 and **WP3,** source data are: NHS Lothian maternity and infant TRAK, Infant Badgernet, questionnaires. Some source data will be captured in the Case Report Form. Source data is defined as all information in original records and certified copies of original records or clinical findings, observations, or other activities in a clinical study necessary for the reconstruction and evaluation of the study. Source data are contained in source documents.

7.2 Case Report Forms

Paper case report forms will be used to collect demographic, SES, and medical information from records and questionnaires. Data will be transcribed to REDCap with an SOP to govern data completeness, accuracy and validity.

8 DATA MANAGEMENT

8.1 Personal Data

WP1: The National Neonatal Research Database contains quality-assured data (the Neonatal Data Set, an NHS Information Standard; SCCI1595). Records of infants will be linked using date of birth, postcode and sex. De-identified datasets will be stored and analysed on a secure university system.

WP2: Personal data will include name, CHI number of mother and infant, address, email, telephone number, date of birth of infant, ethnicity and health data. Participants will be issued with a unique identifier at the point of recruitment and this shall be used to label all research data including biological samples. Contact data and research data will be stored separately, linked by the unique identifier.

WP3: Personal data will include name, CHI number of mother and infant, address, email, telephone number, date of birth of infant, ethnicity and health data. Participants will be issued with a unique identifier at the point of recruitment and this shall be used to label all research data including biological samples. Contact data and research data will be stored separately, linked by the unique identifier.

8.2 Data Information Flow

WP1: Records from the National Neonatal Research database will be linked to the National Pupil Database (NPD) in an NHS Trusted Research Environment (e.g. SAIL Databank) and ScotXeD/SQA Pupil-Census in Scotland.

WP2: All research data will be collected by a member of the research team. The study coordinator will oversee data collection, ensuring that time-points are adhered to for each participant. Participants will be contacted by post and/or telephone to arrange appointments and to prompt if questionnaires have not been returned, to maximise completeness of data collection. No personal data will flow after assignment of the unique identifier.

The MRI community has developed methods for anonymizing MRI data – namely, ensuring that identifying information (e.g., birth date, sex) is unattached from shared data, and "de-facing" the data. There are several algorithms for defacing MRI data (e.g., FreeSurfer's mri_deface). We will employ best practices for deidentifying and anonymizing MRI data prior to sharing. We will also ensure that participants understand the way that their data will be shared and obtain their consent. The risk of identification from MRI is low.

WP3: All research data will be collected by a member of the research team. The study coordinator will oversee data collection, ensuring that time-points are adhered to for each participant. Participants will be contacted by post and/or telephone to arrange appointments and to prompt if questionnaires have not been returned, to maximise completeness of data collection. No personal data will flow after assignment of the unique identifier, with the exception of tasks that are video recorded (Theory of Mind booklet, parent-child play, novel toy task, EF tasks). The data will be stored securely on a password protected file on the University of Edinburgh DataStore facility, accessible only by members of the research team.

The MRI community has developed methods for anonymizing MRI data – namely, ensuring that identifying information (e.g., birth date, sex) is unattached from shared data, and "de-facing" the data. There are several algorithms for defacing MRI data (e.g., FreeSurfer's mri_deface). We will employ best practices for deidentifying and anonymizing MRI data prior to sharing. We will also ensure that participants understand the way that their data will be shared and obtain their consent. The risk of identification from MRI is low.

8.3 Data Storage

WP1: The National Neonatal Research Database contains quality-assured data (the Neonatal Data Set, an NHS Information Standard; DAB1595).

WP2: Electronic data (including personal data) will be collected and managed using REDCap (a secure, web-based application) and stored on a University of Edinburgh server. This is in line with the University's Data Policy and the Data Protection Act of 1998.

Neuroimaging data will be stored on a dedicated GPGPU server, accessible to members of the research team. HPA axis and DNAm data will be stored on UoE Research Data Management servers, supported by Information Services. The data in the RDM file-store is automatically replicated to an off-site disaster facility and is backed up with a 60-day retention period, with 10 days of file history online.

Hard copies will be stored securely in locked filing cabinets in a locked room within University of Edinburgh premises (MRC Centre for Reproductive Health) to which only study researchers have access.

WP3: Electronic data (including personal data) will be collected and managed using REDCap (a secure, web-based application) and stored on a University of Edinburgh server. For all tasks that are video recorded, we will generate spreadsheets (using participant IDs) that hold coded data, that are then also

entered into REDCap (a secure, HIPAA-compliant web-based application) and stored on the University of Edinburgh DataStore server. Information security management at the University of Edinburgh (Edinburgh Parallel Computer Centre, EPCC) is externally accredited as ISO 27001 compliant (Certificate number #276767-2018-AIS-GBR-UKAS). Electronic data will be collected and managed using REDCap. This is in line with the University's Data Policy and the Data Protection Act of 1998.

Neuroimaging data will be stored on a dedicated GPGPU server, accessible to members of the research team.

Hard copies will be stored securely in locked filing cabinets in a locked room within University of Edinburgh premises (MRC Centre for Reproductive Health) to which only study researchers have access.

8.4 Data Retention

The data generated by this programme will be available for use by the research and policy communities in perpetuity.

WP2: (i) Neuroimaging. Raw human neuroimaging data acquired in Edinburgh are archived indefinitely in safe, long term storage in Edinburgh Compute and Data Facility (ECDF). The image processing server is linked by automatic daily back-up to the off-site ECDF. The compute component of ECDF, Eddie, has built-in redundancy with error check and "self-healing,' and is itself backed-up using 2 tape systems. **(ii) HPA axis**. Data will be stored securely under password protection using UoE DataShare, which is backed-up and maintained through UoE IT systems. **(iii) DNAm**. Raw .IDAT files will be stored on GEO indefinitely.

We have costed for University of Edinburgh storage with guaranteed backup and resilience for all new data generated. In the long-term, the University of Edinburgh provides a DataVault data retention service to archive golden copy research data, which is also appropriate for sensitive data types; no charges apply to ≤100GB, which the phenotypic dataset will not exceed. A golden/master copy of new data will be migrated onto this service as part of the process of data curation at the end of the project. Pseudo anonymised data will be stored in DataVault for 10 years, at the end of that period the data will be anonymised.

WP3: Neuroimaging. Raw human neuroimaging data acquired in Edinburgh are archived indefinitely. The image processing server is linked by automatic daily back-up to the off-site Edinburgh Compute and Data Facility (ECDF). The compute component of ECDF, Eddie, has built-in redundancy with error check and "self-healing,' and is itself backed-up using 2 tape systems.

We have costed for University of Edinburgh storage with guaranteed backup and resilience for all new data generated. In the long-term, the University of Edinburgh provides a DataVault data retention service to archive golden copy research data, which is also appropriate for sensitive data types; no charges apply to ≤100GB, which the phenotypic dataset will not exceed. A golden/master copy of new data will be migrated onto this service as part of the process of data curation at the end of the project.

8.5 Disposal of Data

No data will be disposed.

8.6 External Transfer of Data

WP1. Not applicable

WP2. Data collected or generated by the study (including personal data) will not be transferred to any external individuals or organisations outside of the Sponsoring organisation(s), without completion of a data access agreement between Edinburgh University and a recipient scientist in a receiving institution. Requests for access to de-identified, anonymized data will be governed by the PRENCOG Data Access and Collaboration Policy, according to the following principles:
Suitability for sharing. The data proposed for collection is suitable for sharing and will be useful to cognitive and clinical neuroscientists, image scientists, and neurobiologists. It provides a rich characterization of brain function and structure in early childhood with novel contextual information – enabling these data to address a wide variety of questions beyond those described in this proposal.
Discovery by potential users of the research data. (i) Neuroimaging will be retained within the University of Edinburgh and software will be shared via a web portal with GitHub integration. **(ii) HPA axis.** De-identified data will be made available in publications and through repositories such as Edinburgh DataShare and FigShare. **(iii) DNAm.** DNAm will be placed on Gene Expression Omnibus (GEO). A website will be developed to describe the project, link to related activity, of Pi and Co-Is, and signpost outputs. We will publish a description of the project to include design, governance and data access arrangements. We will use social media to promote the resource.
Governance of access. The study management group will review data requests from potential users, who are required to complete data access agreements. Once requests are approved a Data Access Agreement is signed, supported by the UoE Research Office and the corresponding office of the recipient's institution.
The study team's exclusive use of the data. Acquisition of data will continue until the end of the project. Primary publications may take 9-12 months after the end of the project. Once published, data will be shared immediately. Unpublished data will be shared after 3 years.
Restrictions or delays to sharing, with planned actions to limit such restrictions. Participants will be asked to consent to anonymous data-sharing. Since all MRI, HPA axis data and DNAm data are anonymised at source, there should be not be issues relating to identifiability. There will be some delay while we publish primary results, although as described above this will be completed as soon as possible.
Regulation of responsibilities of users. External users will be bound by data access agreements developed by the study management group. The data access agreement will specify the user's responsibilities concerning the received data and will confirm that the data will not be used for purposes other than those detailed in the data access agreement.

WP3. Data collected or generated by the study (including personal data) will not be transferred to any external individuals or organisations outside of the Sponsoring organisation(s), without completion of a data access agreement between Edinburgh University and a recipient scientist in a receiving institution. Requests for access to de-identified, anonymized data will be governed by the PRENCOG Data Access and Collaboration Policy, according to the following principles:
Suitability for sharing. The data proposed for collection is suitable for sharing and will be useful to cognitive and clinical neuroscientists, image scientists, and neurobiologists. It provides a rich characterization of brain function and structure in early childhood with novel contextual information – enabling these data to address a wide variety of questions beyond those described in this proposal.
Discovery by potential users of the research data. Neuroimaging will be made available in a suitable image bank and software will be shared via a web portal with GitHub integration. A website will be developed to describe the project, link to related activity, of Pi and Co-Is, and signpost outputs. We will publish a description of the project to include design, governance and data access arrangements. We will use social media to promote the resource.
Governance of access. The study management group will review data requests from potential users, who are required to complete data access agreements. Once requests are approved a Data Access Agreement is signed, supported by the UoE Research Office and the corresponding office of the recipient's institution.
The study team's exclusive use of the data. Acquisition of data will continue until the end of the project. Primary publications may take 9-12 months after the end of the project. Once published, data will be shared immediately. Unpublished data will be shared after 3 years.
Restrictions or delays to sharing, with planned actions to limit such restrictions. Participants will be asked to consent to anonymous data-sharing. Since all MRI data are anonymised at source, there

should be not be issues relating to identifiability. There will be some delay while we publish primary results, although as described above this will be completed as soon as possible.
Regulation of responsibilities of users. External users will be bound by data access agreements developed by the study management group. The data access agreement will specify the user's responsibilities concerning the received data and will confirm that the data will not be used for purposes other than those detailed in the data access agreement.

8.7 Data Controller

The University of Edinburgh and NHS Lothian are joint data controllers along with any other entities involved in delivering the study that may be a data controller in accordance with applicable laws (e.g. the site).

8.8 Data Breaches

Any data breaches will be reported to the University of Edinburgh (dpo@ed.ac.uk) and NHS Lothian (Lothian.DPO@nhslothian.scot.nhs.uk) Data Protection Officers who will onward report to the relevant authority according to the appropriate timelines if required.

9 STATISTICS AND DATA ANALYSIS

9.1 SAMPLE SIZE CALCULATION

WP1. The NNRD contains data from >100,000 infants in the required gestational range and 2-year outcome data are available for >60% (>70% in most recent years). Data completeness is high for key perinatal variables e.g., 90% for postcode at time of birth (used to derive deprivation index) and 93% for maternal age, and is lower for others e.g., maternal ethnicity 78%. We will investigate whether there are systematic patterns in missingness and we will analyse whether data imputation improves prediction models (univariate imputations using the mean/median values, and joint variable information-theoretic probabilistic methods for joint [conditional] distributions of variables).

Sensitivity analyses will be used to assess the impact of missing outcome data. In addition to applying statistical hypothesis testing (to assess p-values and statistical significance), we will also explore different statistical machine learning methods, including robust training and model validation using standard approaches (e.g. K-fold cross validation).

WP2. Sample size for groupwise comparisons of image data using biological variables is based on properties of the chosen EoP image phenotypes[59 60], and term and preterm differences we have observed in predictor variables (i) hair cortisol concentrations[15]: 401pg/mg(262-615) versus 82pg/mg(55-169), respectively; and (ii) group differences in DNAm CRP scores.

WP3. There are no prior fMRI studies of young children born preterm (2-6 years) from which to estimate the size of the effect of preterm birth on aspects of their functional brain responses. Given this, we performed an indicative power analysis to determine the sensitivity to detect individual differences in two measures of social brain development in 3- to 5-year-olds (n=65)[38]. In this sample, the effect size of age on inter-region correlations among social brain regions, controlling for participant motion, was $r_s(62)=.48$, $p=7.1\times10^{-5}$, and the effect size of age on the functional maturity of this cortical network, controlling for motion, was $r_s(62)=.40$, $p=.001$. These are medium-large effects. Sensitivity analyses with n=246 children and 90% power could detect effects larger than $r=.20$. The sample will enable sub-set analyses based on PTB-RFs identified in WP1. An indicative power estimate for subset analyses based on effect size of age between 3 and 5 years and assuming $\alpha = .05$ and 90% power, is that fMRI data from n=42 or n=60 children, respectively, would detect the two effects above.

9.2 PROPOSED ANALYSES

WP1. We will assess data quality before analysis, including internal consistency, outliers and missing data. Depending on the proportion of missing data, variables will be either excluded or completed with multiple imputation/other data imputation techniques which may be appropriate. Sensitivity analyses will be used to assess the impact of missing outcome data. We will minimise type 1 errors by applying a full analytics machine learning approach with standard approaches for model training and validation, rather than relying solely on statistical hypothesis testing and arbitrary p-value thresholding.

We will first investigate unadjusted associations between the 2 independent categorical variables (i.e. GA, (S)IMD quintile) and dependent variables (moderate to severe neurodevelopmental impairment at 2 years and academic attainment, defined in 4.1.3.3). We will use univariable statistical learning models to estimate the probability of having an adverse outcome: model 1, GA at birth; model 2, (S)IMD quintile (possible interactions between (S)IMD and GA at birth ([S]IMD x GA) will be tested). Then, a multivariable regression that includes the independent variables and significant interaction terms will be used to investigate associations after mutual adjustment for potential confounding variables/effect modifiers, including parental ethnicity and others listed in 4.1.3. Results will be reported as odds ratios with 95% CIs.

We will apply machine learning methods including feature selection and statistical mapping algorithms. Indicatively, we will use both filter feature selection approaches and embedded feature selection, e.g. the random forest feature importance scores to determine the most parsimonious model comprising a robust feature subset with maximally predictive performance using standard model validation approaches (e.g. k-fold cross-validation). We have previously shown that this approach determines predictors of language impairment in preterm infants with high accuracy[61]. Here, we envisage it will determine the PTB-RFs that jointly contribute, in descending order of importance, to neurodevelopment and educational outcome measures. We will apply partial dependence plots to gain further insights into how specific changes in variable values affect outcomes, and we will explore two-dimensional data representations (e.g., applying t-distributed stochastic neighbour embedding) to visualize data both using the original high-dimensional space and using feature subsets determined using feature selection algorithms. This approach will provide a clear demonstration of the effect of PTB-RFs inherently as part of the statistical model used, including direction of effect as a function of different outcome values, thus accounting for mediation.

WP2. Image processing will be carried out using established pipelines for morphometric similarity networks (MSNs)[59], hierarchical complexity (HC)[60] and using in-house techniques for MTsat**.** There are two statistical approaches. The first will use multivariate statistics in a predictive framework to test whether PTB-RFs are associated with MSNs (brain age), HC (connectome architecture), and MTsat (a marker of myelination). The second will use mediation analyses within a structural equation modelling framework to investigate the role of neuroendocrine stress activation and chronic inflammation as mediators of PTB-RFs on features of EoP[27]. This simultaneously characterises associations among HPA axis activation/DNAm, PTB-RFs, and brain features, and specifically tests the hypothesis that stress and/or chronic inflammation partly and significantly mediate associations between PTB-RFs and brain development.

WP3. fMRI data will be pre-processed using fMRIPrep[62]. A general-linear model will be used to analyse BOLD activity of each participant. Individual functional regions of interest (ROIs) will be defined as the 80 voxels within targeted search spaces (for social, attention, and language networks) that have the highest contrast values (i.e., are most similar to the adult timecourse), using a split-half approach such that extraction of response time courses and ROI definition occur in independent data. Functional maturity of each ROI will be calculated as the correlation between each child's (split-half) timecourse and the average adult timecourse; inter-region correlations will be calculated as the average correlation between timecourses across ROIs within each network. These measures capture region-specific functional maturity and the functional coherence of each network[38 43]. To ensure unpredicted effects do not go undetected in such a valuable dataset, additional measures of brain development, including the response magnitude to domain-relevant scenes will be explored. Whole-brain random effects analyses will be conducted to identify any regions that are differentially recruited during domain-relevant scenes

relative to control scenes[38] in preterm children as compared to term-born children. We will explore the impact of preterm birth on response time courses for attention and language, as well as social, networks. These analyses will be corrected for multiple comparisons by estimating the false-positive rate via 5,000 Monte Carlo permutations[63].

Diffusion images will be pre-processed (realigned, co-registered to the anatomical, and normalised) to reduce distortion, and submitted to eddy-current and motion correction analyses using QSIPrep[64]. Diffusion data will be analysed with a tensor model. Tensor model metrics (fractional anisotropy [FA] and mean diffusivity [MD]) are well studied and can be interpreted within a large literature on white matter development. Specific tracts will be identified based on previous work on white matter development and particular cognitive functions (for example, for social cognition, we will focus on the arcuate fasciculus, inferior fronto-occipital fasciculus, and corpus callosum)[65]. Average FA and MD will be calculated within tracts that connect individually-defined functional ROIs (as defined above), per network. Additionally, metrics with increased biological specificity from more advanced biophysical models of the diffusion signal will be explored. For example, the NODDI model enables calculating the intracellular volume fraction of water (ICVF) and orientation dispersion index (ODI)[66 67].

Structural images will be used as anatomical reference for the diffusion and fMRI data. Structural images will be processed using FreeSurfer performing cortical reconstruction and volumetric segmentation[68]. This step will be run inside the QSIPrep and fMRIPrep processes.

Linear and mixed effects regressions will be used. Analyses with neuroimaging measures will include participant motion and age as covariates. When multiple statistical tests within a family are conducted, Bonferroni corrections will be applied.

In neuroimaging analyses, we will control type 1 error while minimising the risk of type 2 error by use of family-wise error rate[69] and threshold-free cluster enhancement[70].

10 ADVERSE EVENTS

We do not anticipate risk from any of the biosample collections or questionnaires.

The MRI scanner makes a loud knocking noise during the scan, so ear muffs are used to prevent discomfort due to noise, and to encourage infants to sleep. We will use established procedures described ensuring infant safety and physiological stability during imaging[48]. The infant will have continuous monitoring of vital signs (heart rate and oxygen saturation) with an electronic monitor. The attending nurse / doctor will record observations every 5 minutes until 1 hour after the infant has woken up, and the scan will be stopped if there are any abnormalities in monitoring. Full neonatal resuscitation facilities are available on site. SOPs for ensuring safety in the MRI environment are already in place in Edinburgh Imaging facility and will be followed.

Five-year old participants will be acclimated to the MRI environment and trained to stay very still (i.e., <2mm motion) during a mock scan[71]. Five-year old participants will use in-ear headphones to listen to the soundtrack of movies played during structural scans and to hear the researchers operating the scan in the control room. The researchers will communicate with children approximately every five minutes during the scan; children will respond by speaking aloud. The in-ear headphones reduce the MRI noise to safe levels; soft pads will offer additional hearing protection and help to stabilise children's heads. An additional radiographer/member of the research team will stand near the child's feet and the bore of the MRI machine and monitor the child during the scan. If/when the child moves, this researcher will pat the child's leg, as a reminder to stay still.

11 OVERSIGHT ARRANGEMENTS

11.1 INSPECTION OF RECORDS

Investigators and institutions involved in the study will permit study related monitoring and audits on behalf of the Sponsor, REC review, and regulatory inspection(s). In the event of audit or monitoring, the Investigator agrees to allow the representatives of the Sponsor direct access to all study records and source documentation. In the event of regulatory inspection, the Investigator agrees to allow inspectors direct access to all study records and source documentation.

11.2 STUDY MONITORING AND AUDIT

The ACCORD Sponsor Representative will assess the study to determine if a study specific risk assessment is required.

If required, a study specific risk assessment will be performed by representatives of the Sponsor(s), ACCORD monitors and the QA group, in accordance with ACCORD governance and sponsorship SOPs. Input will be sought from the Chief Investigator or designee. The outcomes of the risk assessment will form the basis of the monitoring plans and audit plans.

If considered necessary, ACCORD clinical trial monitors, or designees, will perform monitoring activities in accordance with the study monitoring plan. This will involve on-site visits and remote monitoring activities as necessary. ACCORD QA personnel, or designees, will perform study audits in accordance with the study audit plan. This will involve investigator site audits, study management audits and facility (including 3rd parties) audits as necessary (delete where not required).

12 GOOD CLINICAL PRACTICE

12.1 ETHICAL CONDUCT

The study will be conducted in accordance with the principles of the International Conference on Harmonisation Tripartite Guideline for Good Clinical Practice (ICH GCP).

Before the study can commence, all necessary approvals will be obtained and any conditions of approvals will be met.

The programme will be sponsored jointly by the University of Edinburgh and NHS Lothian.

Permissions and approvals for each work package will be managed separately:

WP1: The National Research Ethics Service has approved the National Neonatal Research Database as a research database (Research Ethics Committee No: 16/LO/1093). We will seek ethics approval for WP1 from the Health Research Authority for linkage, and the National Research Ethics Service, NHS Research and Development Office, and the Public Benefit and Privacy Panel for Health and Social Care (Scotland). The process will be led and managed by investigators based at Imperial College London.

WP2: We will seek ethics approval for WP2 from the National Research Ethics Service and NHS Lothian Research and Development Office. The process will be led and managed by investigators based at the University of Edinburgh.

WP3: The National Research Ethics Service has approved WP3 ('Theirworld Edinburgh Birth Cohort', Research Ethics Committee No: 16/SS/0154, NHS Lothian R&D 2016/0255). The process is led and managed by investigators based at the University of Edinburgh.

WP4: Permissions are not required for WP4.

12.2 INVESTIGATOR RESPONSIBILITIES

The Investigator is responsible for the overall conduct of the study at the site and compliance with the protocol and any protocol amendments. In accordance with the principles of ICH GCP, the following areas listed in this section are also the responsibility of the Investigator. Responsibilities may be delegated to an appropriate member of study site staff.

Delegated tasks must be documented on a Delegation Log and signed by all those named on the list prior to undertaking applicable study-related procedures.

12.2.1 Informed Consent

The Investigator is responsible for ensuring informed consent is obtained before any study specific procedures are carried out. The decision of a participant to participate in clinical research is voluntary and should be based on a clear understanding of what is involved.

Participants must receive adequate oral and written information – appropriate Participant Information and Informed Consent Forms will be provided. The oral explanation to the participant will be performed by the Investigator or qualified delegated person, and must cover all the elements specified in the Participant Information Sheet and Consent Form.

The participant must be given every opportunity to clarify any points they do not understand and, if necessary, ask for more information. The participant must be given sufficient time to consider the information provided. It should be emphasised that the participant may withdraw their consent to participate at any time without loss of benefits to which they otherwise would be entitled.

The participant will be informed and agree to their medical records being inspected by regulatory authorities and representatives of the Sponsor(s).

The Investigator or delegated member of the study team and the participant will sign and date the Informed Consent Form(s) to confirm that consent has been obtained. The original will be signed in the Investigator Site File (ISF). The participant will receive a copy of the signed consent form and a copy will be filed in the participant's medical notes.

12.2.2 Study Site Staff

The Investigator must be familiar with the protocol and the study requirements. It is the Investigator's responsibility to ensure that all staff assisting with the study are adequately informed about the protocol and their study related duties.

12.2.3 Data Recording

The Principal Investigator is responsible for the quality of the data recorded in the CRF.

12.2.4 Investigator Documentation

The Principal Investigator will ensure that the required documentation is available in local Investigator Site files (ISFs).

12.2.5 GCP Training

For non-CTIMP (i.e. non-drug) studies all researchers are encouraged to undertake GCP training in order to understand the principles of GCP. This is not a mandatory requirement unless deemed so by the Sponsor. GCP training status for all investigators should be indicated in their respective CVs.

12.2.6 Data Protection Training

All University of Edinburgh employed researchers and study staff will complete the Data Protection Training through Learn.

NHS Lothian employed researchers and study staff will comply with NHS Lothian mandatory Information Governance Data Protection training through LearnPro.

Non-NHS Lothian staff that have access to NHS Lothian systems will familiarise themselves and abide by all NHS Lothian IT policies, as well as employer policies.

12.2.7 Information Security Training

All University of Edinburgh employed researchers, students and study staff will complete the Information Security Essentials modules through Learn and will have read the minimum and required reading setting out ground rules to be complied with.

NHS Lothian employed researchers and study staff will comply with NHS Lothian mandatory Information Governance IT Security training through LearnPro.

Non-NHS Lothian staff that have access to NHS Lothian systems will familiarise themselves and abide by all NHS Lothian IT policies, as well as employer policies.

12.2.8 Confidentiality

All laboratory specimens, evaluation forms, reports, and other records must be identified in a manner designed to maintain participant confidentiality. All records must be kept in a secure storage area with limited access. Clinical information will not be released without the written permission of the participant. The Investigator and study site staff involved with this study may not disclose or use for any purpose other than performance of the study, any data, record, or other unpublished information, which is confidential or identifiable, and has been disclosed to those individuals for the purpose of the study. Prior written agreement from the Sponsor or its designee must be obtained for the disclosure of any said confidential information to other parties.

We will refer incidental findings that may be clinically actionable to relevant NHS services and inform the child's GP, for example, unexpected abnormalities on MRI. If risk for maternal depression or childhood behavioural disorders that may require support are identified on screening questionnaires (Edinburgh Postnatal Depression Scale, DuPaul ADHD scale, Social Communication Questionnaire, Children's Communication Checklist-2), we will inform the participant and sign-post to relevant services. We will adhere to NHS Lothian safeguarding policy in the event that abuse to child or parent is suspected during the course of research activity.

12.2.9 Data Protection

All Investigators and study site staff involved with this study must comply with the requirements of the appropriate data protection legislation (including the General Data Protection Regulation and Data Protection Act) with regard to the collection, storage, processing and disclosure of personal information.

Computers used to collate the data will have limited access measures via user names and passwords.

Published results will not contain any personal data that could allow identification of individual participants.

12.3 STUDY MANAGEMENT GROUP

The Principal Investigator will take overall responsibility for the programme. The study will be coordinated by a Study Management Group, consisting of the Principal Investigator, a minimum of two co-investigators, and the Programme Manager.

The Programme Manager will oversee administrative aspects including compliance with research governance, budget management, and coordination of patient visits will oversee the study and will be

accountable to the Principal Investigator. S/he will be responsible for checking the CRFs for completeness, plausibility and consistency. Any queries will be resolved by the Investigator or delegated member of the trial team. S/he will: promote and market PRENCOG through design and implementation outreach strategies and monitor effectiveness of these; assist the PI in communication with ethics and R & D committees to ensure timely project approval; engage with CRH finance officer to coordinate invoices and payments; ensure compliance with regulations by designing, producing and regularly updating PRENCOG materials including the protocol, SOPs and data collection forms, and ensuring all collaborators have materials available.

A Delegation Log will be prepared, detailing the responsibilities of each member of staff working on the study.

The co-investigators will hold monthly meetings to assess progress against milestones in year 1; collaborators will be invited to attend quarterly. In years 2-5, investigator meetings may be quarterly.

12.4 SCIENTIFIC ADVISORY BOARD

We will appoint an advisory board of international experts spanning paediatric neuroimaging, developmental cognitive neuroscience, epigenetics, neuroendocrinology, and education. The advisory board will meet once per year, remotely.

13 STUDY CONDUCT RESPONSIBILITIES

13.1 PROTOCOL AMENDMENTS

Any changes in research activity, except those necessary to remove an apparent, immediate hazard to the participant in the case of an urgent safety measure, must be reviewed and approved by the Principal Investigator.

Proposed amendments will be submitted to the Sponsor for classification, review and authorisation.

Amendments to the protocol must be submitted in writing to the appropriate REC and local R&D for approval prior to implementation and prior to participants being enrolled into the amended protocol.

13.2 MANAGEMENT OF PROTOCOL NON-COMPLIANCE

13.2.1 Protocol Waivers

Prospective protocol deviations, i.e. protocol waivers, will not be approved by the Sponsors and therefore will not be implemented, except where necessary to eliminate an immediate hazard to study participants. If this necessitates a subsequent protocol amendment, this should be submitted to the REC and local R&D for review and approval if appropriate.

13.2.2 Management of Deviations and Violations

Deviations and violations are non-compliance events discovered after the event has occurred. Protocol deviations will be recorded in a protocol deviation log and logs will be submitted to the Sponsors every 3 months. Each protocol violation will be reported to the Sponsor within 3 days of becoming aware of the violation.

Deviation logs will be maintained for each site in multi-centre studies.

Deviation logs/violation forms will be transmitted via email to QA@accord.scot. Only forms in a pdf format will be accepted by ACCORD via email. Forms may also be submitted by hand to the office. Where missing information has not been sent to ACCORD after an initial report, ACCORD will contact

the Investigator and request the missing information. The Investigator must respond to these requests in a timely manner.

13.3 SERIOUS BREACH REQUIREMENTS

A serious breach is a breach which is likely to effect to a significant degree:

(a) the safety or physical or mental integrity of the participants of the study; or

(b) the scientific value of the study.

If a potential serious breach is identified by the Chief investigator, Principal Investigator or delegates, the Sponsor(s) (qa@accord.scot) must be notified within 24 hours. It is the responsibility of the Sponsor(s) to assess the impact of the breach on the scientific value of the study, to determine whether the incident constitutes a serious breach and report to research ethics committees as necessary.

13.4 STUDY RECORD RETENTION

All study documentation will be kept for a minimum of 10 years from the protocol defined end of study point. When the minimum retention period has elapsed, study documentation will be destroyed with permission from the Sponsor.

13.5 END OF STUDY

The end of study is defined as the last participant's last visit.

The Investigators and/or the Sponsor(s) have the right at any time to terminate the study for clinical or administrative reasons.

The end of the study will be reported to the REC, and R&D Office(s) and Sponsor(s) within 90 days, or 15 days if the study is terminated prematurely. The Investigators will inform participants of the premature study closure and ensure that the appropriate follow up is arranged for all participants involved. End of study notification will be reported to the Sponsor(s) via email to researchgovernance@ed.ac.uk.

A summary report of the study will be provided to the REC within 1 year of the end of the study.

13.6 CONTINUATION OF TREATMENT FOLLOWING THE END OF STUDY

Not applicable.

13.7 INSURANCE AND INDEMNITY

The Sponsor(s) are responsible for ensuring proper provision has been made for insurance or indemnity to cover their liability and the liability of the Chief Investigator and staff.

The following arrangements are in place to fulfil the Sponsor(s)' responsibilities:

- The Protocol has been designed by the Chief Investigator and researchers employed by the University and collaborators. The University has insurance in place (which includes no-fault compensation) for negligent harm caused by poor protocol design by the Chief Investigator and researchers employed by the University.
- Sites participating in the study will be liable for clinical negligence and other negligent harm to individuals taking part in the study and covered by the duty of care owed to them by the sites concerned. The Sponsor(s) require individual sites participating in the study to arrange for their own insurance or indemnity in respect of these liabilities.
- Sites which are part of the United Kingdom's National Health Service will have the benefit of NHS Indemnity.
- Sites out with the United Kingdom will be responsible for arranging their own indemnity or insurance for their participation in the study, as well as for compliance with local law applicable to their participation in the study.

14 AUTHORSHIP POLICY

The ICMJE criteria for authorship will govern all authorship decisions. PRENCOG co-investigators and collaborators will be shown all manuscripts arising from the study prior to submission so that they can consider whether they meet ICMJE criteria and accept or decline responsibility for co-authorship.

15 REFERENCES

1. Chawanpaiboon S, Vogel JP, Moller AB, et al. Global, regional, and national estimates of levels of preterm birth in 2014: a systematic review and modelling analysis. *Lancet Glob Health* 2019;7(1):e37-e46. doi: 10.1016/s2214-109x(18)30451-0 [published Online First: 2018/11/06]
2. Johnson S, Marlow N. Early and long-term outcome of infants born extremely preterm. *Arch Dis Child* 2017;102(1):97-102. doi: 10.1136/archdischild-2015-309581 [published Online First: 2016/08/12]
3. Blencowe H, Vos T, Lee AC, et al. Estimates of neonatal morbidities and disabilities at regional and global levels for 2010: introduction, methods overview, and relevant findings from the Global Burden of Disease study. *Pediatr Res* 2013;74 Suppl 1:4-16. doi: 10.1038/pr.2013.203
4. Volpe JJ. Dysmaturation of Premature Brain: Importance, Cellular Mechanisms, and Potential Interventions. *Pediatr Neurol* 2019;95:42-66. doi: 10.1016/j.pediatrneurol.2019.02.016 [published Online First: 2019/04/13]
5. Batalle D, Edwards AD, O'Muircheartaigh J. Annual Research Review: Not just a small adult brain: understanding later neurodevelopment through imaging the neonatal brain. *J Child Psychol Psychiatry* 2018;59(4):350-71. doi: 10.1111/jcpp.12838 [published Online First: 2017/11/07]
6. Boardman JP, Counsell SJ. Invited Review: Factors associated with atypical brain development in preterm infants: insights from magnetic resonance imaging. *Neuropathol Appl Neurobiol* 2020;46(5):413-21. doi: 10.1111/nan.12589 [published Online First: 2019/11/21]
7. Twilhaar ES, Wade RM, de Kieviet JF, et al. 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(4):361-67. doi: 10.1001/jamapediatrics.2017.5323 [published Online First: 2018/02/21]
8. Van den Bergh BRH, van den Heuvel MI, Lahti M, et al. Prenatal developmental origins of behavior and mental health: The influence of maternal stress in pregnancy. *Neurosci Biobehav Rev* 2020;117:26-64. doi: 10.1016/j.neubiorev.2017.07.003 [published Online First: 2017/08/02]
9. Moisiadis VG, Matthews SG. Glucocorticoids and fetal programming part 2: Mechanisms. *Nat Rev Endocrinol* 2014;10(7):403-11. doi: 10.1038/nrendo.2014.74 [published Online First: 20140527]
10. Buss C, Davis EP, Shahbaba B, et al. Maternal cortisol over the course of pregnancy and subsequent child amygdala and hippocampus volumes and affective problems. *Proc Natl Acad Sci U S A* 2012;109(20):E1312-9. doi: 10.1073/pnas.1201295109 [published Online First: 20120423]
11. Graham AM, Rasmussen JM, Entringer S, et al. Maternal Cortisol Concentrations During Pregnancy and Sex-Specific Associations With Neonatal Amygdala Connectivity and Emerging Internalizing Behaviors. *Biol Psychiatry* 2019;85(2):172-81. doi: 10.1016/j.biopsych.2018.06.023 [published Online First: 20180707]

12. Stoye DQ, Blesa M, Sullivan G, et al. Maternal cortisol is associated with neonatal amygdala microstructure and connectivity in a sexually dimorphic manner. *Elife* 2020;9 doi: 10.7554/eLife.60729 [published Online First: 2020/11/25]
13. Raikkonen K, O'Reilly JR, Pesonen AK, et al. Associations between maternal level of education and occupational status with placental glucocorticoid regeneration and sensitivity. *Clin Endocrinol (Oxf)* 2014;81(2):175-82. doi: 10.1111/cen.12412 [published Online First: 2014/01/22]
14. Raikkonen K, Martikainen S, Pesonen AK, et al. Maternal Licorice Consumption During Pregnancy and Pubertal, Cognitive, and Psychiatric Outcomes in Children. *Am J Epidemiol* 2017;185(5):317-28. doi: 10.1093/aje/kww172 [published Online First: 2017/02/06]
15. Stoye DQ, Sullivan G, Galdi P, et al. Perinatal determinants of neonatal hair glucocorticoid concentrations. *Psychoneuroendocrinology* 2021;128:105223. doi: 10.1016/j.psyneuen.2021.105223 [published Online First: 2021/04/21]
16. Green C, Stolicyn A, Harris MA, et al. Hair glucocorticoids are associated with childhood adversity, depressive symptoms and reduced global and lobar grey matter in Generation Scotland. *Transl Psychiatry* 2021;11(1):523. doi: 10.1038/s41398-021-01644-9 [published Online First: 20211012]
17. Stoll BJ, Hansen NI, Adams-Chapman I, et al. Neurodevelopmental and growth impairment among extremely low-birth-weight infants with neonatal infection. *JAMA* 2004;292(19):2357-65. doi: 10.1001/jama.292.19.2357 [published Online First: 2004/11/18]
18. Wu YW. Systematic review of chorioamnionitis and cerebral palsy. *MentRetardDevDisabilResRev* 2002;8(1):25-29.
19. Fleiss B, Gressens P. Neuroprotection of the preterm brain. *Handb Clin Neurol* 2019;162:315-28. doi: 10.1016/B978-0-444-64029-1.00015-1
20. Dammann O, Allred EN, Fichorova RN, et al. Duration of Systemic Inflammation in the First Postnatal Month Among Infants Born Before the 28th Week of Gestation. *Inflammation* 2016;39(2):672-7. doi: 10.1007/s10753-015-0293-z
21. Sullivan G, Galdi P, Borbye-Lorenzen N, et al. Preterm Birth Is Associated With Immune Dysregulation Which Persists in Infants Exposed to Histologic Chorioamnionitis. *Front Immunol* 2021;12:722489. doi: 10.3389/fimmu.2021.722489 [published Online First: 2021/09/14]
22. Sullivan G, Galdi P, Cabez MB, et al. Interleukin-8 dysregulation is implicated in brain dysmaturation following preterm birth. *Brain Behav Immun* 2020;90:311-18. doi: 10.1016/j.bbi.2020.09.007 [published Online First: 2020/09/14]
23. Ligthart S, Marzi C, Aslibekyan S, et al. DNA methylation signatures of chronic low-grade inflammation are associated with complex diseases. *Genome Biol* 2016;17(1):255. doi: 10.1186/s13059-016-1119-5 [published Online First: 2016/12/14]
24. Ventham NT, Kennedy NA, Adams AT, et al. Integrative epigenome-wide analysis demonstrates that DNA methylation may mediate genetic risk in inflammatory bowel disease. *Nat Commun* 2016;7:13507. doi: 10.1038/ncomms13507 [published Online First: 20161125]
25. Stenvinkel P, Karimi M, Johansson S, et al. Impact of inflammation on epigenetic DNA methylation - a novel risk factor for cardiovascular disease? *J Intern Med* 2007;261(5):488-99. doi: 10.1111/j.1365-2796.2007.01777.x
26. Freytag V, Carrillo-Roa T, Milnik A, et al. A peripheral epigenetic signature of immune system genes is linked to neocortical thickness and memory. *Nat Commun* 2017;8:15193. doi: 10.1038/ncomms15193 [published Online First: 20170426]
27. Conole ELS, Stevenson AJ, Maniega SM, et al. DNA Methylation and Protein Markers of Chronic Inflammation and Their Associations With Brain and Cognitive Aging.

Neurology 2021 doi: 10.1212/WNL.0000000000012997 [published Online First: 2021/11/19]
28. McCartney DL, Hillary RF, Conole ELS, et al. Blood-based epigenome-wide analyses of cognitive abilities. *medRxiv* 2021:2021.05.24.21257698. doi: 10.1101/2021.05.24.21257698
29. Raffington L, Tanksley P, Vinnik L, et al. Socially stratified DNA-methylation profiles are associated with disparities in child and adolescent mental health. *medRxiv* 2021:2021.09.17.21263582. doi: 10.1101/2021.09.17.21263582
30. Stevenson AJ, McCartney DL, Harris SE, et al. Trajectories of inflammatory biomarkers over the eighth decade and their associations with immune cell profiles and epigenetic ageing. *Clin Epigenetics* 2018;10(1):159. doi: 10.1186/s13148-018-0585-x [published Online First: 20181220]
31. Stevenson AJ, McCartney DL, Hillary RF, et al. Characterisation of an inflammation-related epigenetic score and its association with cognitive ability. *Clin Epigenetics* 2020;12(1):113. doi: 10.1186/s13148-020-00903-8 [published Online First: 2020/07/29]
32. Spiers H, Hannon E, Schalkwyk LC, et al. Methylomic trajectories across human fetal brain development. *Genome Res* 2015;25(3):338-52. doi: 10.1101/gr.180273.114 [published Online First: 2015/02/05]
33. Kupers LK, Monnereau C, Sharp GC, et al. Meta-analysis of epigenome-wide association studies in neonates reveals widespread differential DNA methylation associated with birthweight. *Nat Commun* 2019;10(1):1893. doi: 10.1038/s41467-019-09671-3 [published Online First: 2019/04/25]
34. McCartney DL, Min JL, Richmond RC, et al. Genome-wide association studies identify 137 genetic loci for DNA methylation biomarkers of aging. *Genome Biol* 2021;22(1):194. doi: 10.1186/s13059-021-02398-9 [published Online First: 2021/07/01]
35. Wheater ENW, Galdi P, McCartney DL, et al. DNA methylation in relation to gestational age and brain dysmaturation in preterm infants. *Brain Commun* 2022;4(2):fcac056. doi: 10.1093/braincomms/fcac056 [published Online First: 20220308]
36. Dean B, Ginnell L, Boardman JP, et al. Social cognition following preterm birth: A systematic review. *Neurosci Biobehav Rev* 2021;124:151-67. doi: 10.1016/j.neubiorev.2021.01.006 [published Online First: 2021/02/02]
37. Barre N, Morgan A, Doyle LW, et al. Language abilities in children who were very preterm and/or very low birth weight: a meta-analysis. *J Pediatr* 2011;158(5):766-74.e1. doi: 10.1016/j.jpeds.2010.10.032 [published Online First: 2010/12/15]
38. Richardson H, Lisandrelli G, Riobueno-Naylor A, et al. Development of the social brain from age three to twelve years. *Nat Commun* 2018;9(1):1027. doi: 10.1038/s41467-018-03399-2 [published Online First: 2018/03/14]
39. Gweon H, Dodell-Feder D, Bedny M, et al. Theory of mind performance in children correlates with functional specialization of a brain region for thinking about thoughts. *Child Dev* 2012;83(6):1853-68. doi: 10.1111/j.1467-8624.2012.01829.x [published Online First: 2012/08/02]
40. Richardson H, Koster-Hale J, Caselli N, et al. Reduced neural selectivity for mental states in deaf children with delayed exposure to sign language. *Nat Commun* 2020;11(1):3246. doi: 10.1038/s41467-020-17004-y [published Online First: 20200626]
41. Eyre M, Fitzgibbon SP, Ciarrusta J, et al. The Developing Human Connectome Project: typical and disrupted perinatal functional connectivity. *Brain* 2021;144(7):2199-213. doi: 10.1093/brain/awab118

42. Lordier L, Meskaldji DE, Grouiller F, et al. Music in premature infants enhances high-level cognitive brain networks. *Proc Natl Acad Sci U S A* 2019;116(24):12103-08. doi: 10.1073/pnas.1817536116 [published Online First: 20190528]
43. Richardson H. Development of brain networks for social functions: Confirmatory analyses in a large open source dataset. *Dev Cogn Neurosci* 2019;37:100598. doi: 10.1016/j.dcn.2018.11.002 [published Online First: 2018/12/14]
44. Finn ES, Bandettini PA. Movie-watching outperforms rest for functional connectivity-based prediction of behavior. *Neuroimage* 2021;235:117963. doi: 10.1016/j.neuroimage.2021.117963 [published Online First: 20210402]
45. McLaughlin KA, Sheridan MA, Tibu F, et al. Causal effects of the early caregiving environment on development of stress response systems in children. *Proc Natl Acad Sci U S A* 2015;112(18):5637-42. doi: 10.1073/pnas.1423363112 [published Online First: 2015/04/23]
46. Beange I, Kirkham EJ, Fletcher-Watson S, et al. Using a knowledge exchange event to assess study participants' attitudes to research in a rapidly evolving research context. *Wellcome Open Res* 2020;5:24. doi: 10.12688/wellcomeopenres.15651.2 [published Online First: 2020/09/15]
47. Ginnell L, Boardman JP, Reynolds RM, et al. Parent priorities for research and communication concerning childhood outcomes following preterm birth. *Wellcome Open Res* 2021;6:151. doi: 10.12688/wellcomeopenres.16863.2 [published Online First: 2021/11/11]
48. Boardman JP, Hall J, Thrippleton MJ, et al. Impact of preterm birth on brain development and long-term outcome: protocol for a cohort study in Scotland. *BMJ Open* 2020;10(3):e035854. doi: 10.1136/bmjopen-2019-035854 [published Online First: 2020/03/07]
49. Laforest S, Pelletier M, Denver N, et al. Estrogens and Glucocorticoids in Mammary Adipose Tissue: Relationships with Body Mass Index and Breast Cancer Features. *J Clin Endocrinol Metab* 2020;105(4) doi: 10.1210/clinem/dgz268
50. Stirrat LI, Just G, Homer NZM, et al. Glucocorticoids are lower at delivery in maternal, but not cord blood of obese pregnancies. *Sci Rep* 2017;7(1):10263. doi: 10.1038/s41598-017-10266-5 [published Online First: 20170831]
51. Wiechmann T, Roh S, Sauer S, et al. Identification of dynamic glucocorticoid-induced methylation changes at the FKBP5 locus. *Clin Epigenetics* 2019;11(1):83. doi: 10.1186/s13148-019-0682-5 [published Online First: 20190523]
52. Czamara D, Dieckmann L, Roh S, et al. Betamethasone administration during pregnancy is associated with placental epigenetic changes with implications for inflammation. *Clin Epigenetics* 2021;13(1):165. doi: 10.1186/s13148-021-01153-y [published Online First: 20210826]
53. Vaher K, Bogaert D, Richardson H, et al. Microbiome-gut-brain axis in brain development, cognition and behavior during infancy and early childhood. *Developmental Review* 2022;66:101038. doi: https://doi.org/10.1016/j.dr.2022.101038
54. Cong X, Xu W, Janton S, et al. Gut Microbiome Developmental Patterns in Early Life of Preterm Infants: Impacts of Feeding and Gender. *PLoS One* 2016;11(4):e0152751. doi: 10.1371/journal.pone.0152751 [published Online First: 2016/04/26]
55. Jakobsson HE, Abrahamsson TR, Jenmalm MC, et al. Decreased gut microbiota diversity, delayed Bacteroidetes colonisation and reduced Th1 responses in infants delivered by caesarean section. *Gut* 2014;63(4):559-66. doi: 10.1136/gutjnl-2012-303249 [published Online First: 2013/08/09]
56. Benakis C, Brea D, Caballero S, et al. Commensal microbiota affects ischemic stroke outcome by regulating intestinal gammadelta T cells. *Nat Med* 2016 doi: 10.1038/nm.4068 [published Online First: 2016/03/29]

57. Dinan TG, Cryan JF. Regulation of the stress response by the gut microbiota: implications for psychoneuroendocrinology. *Psychoneuroendocrinology* 2012;37(9):1369-78. doi: 10.1016/j.psyneuen.2012.03.007 [published Online First: 2012/04/10]
58. Modi N, Ashby D, Battersby C, et al. Developing routinely recorded clinical data from electronic patient records as a national resource to improve neonatal health care: the Medicines for Neonates research programme. Southampton (UK)2019.
59. Galdi P, Blesa M, Stoye DQ, et al. Neonatal morphometric similarity mapping for predicting brain age and characterizing neuroanatomic variation associated with preterm birth. *Neuroimage Clin* 2020;25:102195. doi: 10.1016/j.nicl.2020.102195 [published Online First: 2020/02/12]
60. Blesa M, Galdi P, Cox SR, et al. Hierarchical Complexity of the Macro-Scale Neonatal Brain. *Cereb Cortex* 2020 doi: 10.1093/cercor/bhaa345 [published Online First: 2020/12/07]
61. Valavani E, Blesa M, Galdi P, et al. Language function following preterm birth: prediction using machine learning. *Pediatr Res* 2021 doi: 10.1038/s41390-021-01779-x [published Online First: 2021/10/13]
62. Esteban O, Markiewicz CJ, Blair RW, et al. fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nat Methods* 2019;16(1):111-16. doi: 10.1038/s41592-018-0235-4 [published Online First: 20181210]
63. Zhang H, Schneider T, Wheeler-Kingshott CA, et al. NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain. *Neuroimage* 2012;61(4):1000-16. doi: 10.1016/j.neuroimage.2012.03.072 [published Online First: 2012/04/10]
64. Cieslak M, Cook PA, He X, et al. QSIPrep: an integrative platform for preprocessing and reconstructing diffusion MRI data. *Nat Methods* 2021;18(7):775-78. doi: 10.1038/s41592-021-01185-5 [published Online First: 20210621]
65. Grosse Wiesmann C, Schreiber J, Singer T, et al. White matter maturation is associated with the emergence of Theory of Mind in early childhood. *Nat Commun* 2017;8:14692. doi: 10.1038/ncomms14692 [published Online First: 20170321]
66. Richardson H, Saxe R. Development of predictive responses in theory of mind brain regions. *Dev Sci* 2020;23(1):e12863. doi: 10.1111/desc.12863 [published Online First: 20190618]
67. Eklund A, Nichols TE, Knutsson H. Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proc Natl Acad Sci U S A* 2016;113(28):7900-5. doi: 10.1073/pnas.1602413113 [published Online First: 2016/07/01]
68. Fischl B. FreeSurfer. *Neuroimage* 2012;62(2):774-81. doi: 10.1016/j.neuroimage.2012.01.021 [published Online First: 20120110]
69. Nichols T, Hayasaka S. Controlling the familywise error rate in functional neuroimaging: a comparative review. *StatMethods MedRes* 2003;12(5):419-46.
70. Smith SM, Nichols TE. Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage* 2009;44(1):83-98.
71. de Bie HM, Boersma M, Wattjes MP, et al. Preparing children with a mock scanner training protocol results in high quality structural and functional MRI scans. *Eur J Pediatr* 2010;169(9):1079-85. doi: 10.1007/s00431-010-1181-z [published Online First: 20100313]