ABSTRACT

Introduction Preterm (PT) infants are at high likelihood for poor neurodevelopmental outcomes, including autism spectrum disorder (ASD), attention-deficit hyperactivity disorder (ADHD) and other neurodevelopmental disorders (NDDs), which could considerably impair the individuals’ functions throughout their whole life. The current cohort study aims to investigate adverse outcomes, especially NDDs, in PT children, and the related early aberrant brain developmental biomarkers.

Methods and analysis This is a prospective cohort study in Beijing, China. We plan to recruit 400 PT infants born at <37 weeks of gestational age (GA), and 200 full-term (FT) controls during the neonatal period (40 weeks corrected GA), then follow them up until they reach 6 years of age. This cohort is designed to assess neuropsychological functions, brain development, related environmental risk factors and the incidence of NDDs by using the following measures: (1) social, emotional, cognitive and sensorimotor functions; (2) MRI, electroencephalogram and functional near-infrared spectroscopy; (3) social economic status, maternal mental health and DNA methylation; and (4) symptoms and diagnosis of NDDs. Main data analyses will include comparing the neurodevelopment outcomes and brain developmental trajectories between PT and FT children using linear or logistic regressions and mixed-effects models. Regression analyses and machine learning will be used to identify early biological predictors and environmental risk or protective factors for later NDDs outcomes.

Ethics and dissemination Ethical approval has been obtained from the research ethics committee of Peking University Third Hospital (M2021087). This study is under review in the Chinese Clinical Trial Register. The study results from the current cohort will be disseminated and popularised through social media to participating parents, as well as parents who are giving care to PT children.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ Four hundred preterm (PT) infants and 200 full-term infants are being intensively followed for the first 6 years.
⇒ The current cohort evaluates diagnoses and symptom levels of all neurodevelopmental disorders defined by the Diagnostic Statistical Manual Fifth Edition since the very early life stages leveraging gold standard psychiatric interviews and widely accepted questionnaires.
⇒ The current cohort studies the developmental trajectories of PT children in multiple dimensions, including behavioural phenotypic information, biological manifestations and environmental factors.
⇒ Participants will be recruited and followed at only one site located in a metropolis, which might cause bias in terms of demographic characteristics.
⇒ Most assessments require offline participation, which might be inconvenient and could result in dropouts and missing data.

INTRODUCTION

Newborns with a gestational age (GA) of less than 37 weeks are considered as preterm (PT). According to the WHO, the PT birth rate has reached 10% worldwide, with nearly 15 million PT babies are born each year worldwide, and this number is still rising.1 Currently, although modern medicine has enabled more PT infants to survive at low GA and weight, long-term neurodevelopmental outcomes of PT infants still have not improved, with PT birth remaining a strong risk factor for neurodevelopmental disorders (NDDs).2-6

NDDs are a group of early-onset disorders with high prevalence and usually long courses, including attention-deficit hyperactivity disorder (ADHD), autism spectrum disorder (ASD), intellectual disability, communication disorder, specific learning disorder and motor disorder.7 PT children have an elevated likelihood for NDDs. The likelihood of developing ASD and ADHD was much higher in PT than in FT children, with prevalence of 6% versus 1% and 12.7% versus 5.9%, respectively.8-11 This likelihood tends to increase with the GA decrease. PT individuals also present worse cognitive performance than their FT counterparts, including poorer attention and lower IQ, from childhood to adulthood.12-13
As has been demonstrated, early exposure to the extrauterine environment significantly affects the early neurodevelopmental process, which might be the biological substrates and indicators for later poor neurodevelopmental outcomes and could aid in early diagnosis.20 21 On the other hand, individuals with NDDs often present biological abnormalities since the fetal stage, including abnormalities in the proliferation of neural progenitor cells, neuron generation and synapse formation, which have been found to be interrupted in PT infants and could ultimately lead to brain morphological and functional aberrations, such as the imbalance between excitatory and inhibitory functions.19 Previous studies have reported that vulnerabilities associated with the stage of brain development in PT children might be the neural substrate of poor neurodevelopmental outcomes, mainly caused by disrupted synaptogenesis and myelination.20 21

Currently, longitudinal cohorts focusing on brain development in PT children have been established, including the LaPreM,22 the Theirworld Edinburgh birth cohort,23 the PREBO cohort,24 the ePrime cohort and so on. On the other hand, researchers established prospective birth cohorts among infants with a family history of ASD including the autism Baby Siblings Research Consortium (BSRC),25 Genome, Environment, Microbiome and ASD and Health in Early Life (GEMINAH) study,26 Genome, Environment, Microbiome and Metabolome in Autism (GEMMA) study,26 European autism interventions–a multicentre study for developing new medications (EU-AIMS)27 and additional examples.

Despite the progress made by these cohorts, most of them are enrolled in Western countries and few have focused on the outcomes of NDDs in PT children.

Currently, the absence of early diagnostic tools for infants makes early intervention challenging. Establishing a longitudinal follow-up cohort to further delineate the trajectory of PT development, explore the relationships between aberrant brain development and NDDs and identify early modifiable risk factors would not only be beneficial in understanding the biological process underlying NDDs, but also shedding light on early diagnosis and intervention.

Therefore, the aims of this study are fourfold: (1) to provide a comprehensive and overall picture of early life development in PT children by comparing the developmental trajectories of neuropsychological functions (including social, emotional, cognitive and sensorimotor functions) and brain development (using MRI, electroencephalogram (EEG) and functional near-infrared spectroscopy (fNIRS)) from birth to 6 years old between PT and FT children; (2) to evaluate the outcomes of NDDs from 18 months to 6 years old, examine the differences between FT and PT children; (3) to explore the relationships between NDDs and early neuropsychological functions, brain development to identify early biomarkers for NDDs; and (4) to identify the patterns in which early life environmental factors affect neurodevelopment, including parental mental health, social economic status and breast feeding.

METHODS AND ANALYSES
Study design and setting
This study is a prospective longitudinal cohort study that examines the neuropsychological and brain development in PT children from birth to 6 years old. The current work was designed following the guidelines of STROBE.28

The current study was conducted at the Peking University Third Hospital and the Peking University Sixth hospital. Both sites locate in Beijing, China. The Children’s Healthcare Center at the Third Hospital serves over 10000 children per annum, with more than 8000 children currently under follow-up.

All the procedures, including participants recruitment, follow-up and the assessments will take place at the Children’s Healthcare Center at the Peking University Third Hospital. The assessments of NDDs diagnosis will be carried out by professional licensed psychiatric doctors from the Peking University Sixth Hospital. Recruitment began in August 2021 and is expected to be completed in December 2025. Follow-up visits will continue until December 2031.

Study participants
Inclusion criteria
The current study aims to recruit 400 PT infants (GA <37 weeks) and 200 full-term (FT) infants (GA >37 weeks).

Exclusion criteria
1. Congenital anomalies: (a) cerebral haemorrhage, periventricular leukomalacia, hypoxic ischaemic encephalopathy and brain structural defects; (b) pulmonary disease or bronchopulmonary dysplasia; (c) ventricular haemorrhage; (d) congenital malformations; (e) chromosomal abnormalities; (f) pia or ventricular surface abnormalities; (g) necrotising enterocolitis and complex feeding/nutritional disorders; (h) hearing or visual impairment; and (i) epilepsy.
2. Infants with contraindications to MRI.
3. FT control infants are additionally required to have no history and no first-degree relative with a diagnosis of NDDs.

Sample size
To ensure adequate statistical power, we estimated the sample size according to previous relevant studies. Sample estimation was completed by PASS 2021, with an alpha of 0.05, 1-beta of 0.90 and a follow-up rate of 80%.

1. The primary outcome is the different occurrences of NDDs between PT and FT infants. Previous studies have shown that the prevalence of ASD and ADHD was about 6% and 12.7% in PT infants, compared with 1.5% and 5.9% in the general population.8 9 11 29 To observe a significant difference in incidence between the two groups, a minimum sample size of 59:30 and 89:45 would be required.
2. The secondary outcome is the development of neuropsychological functions among PT children. Previous findings have demonstrated neurodevelopmental de-
lay in PT children. According to the results reported by Olsen et al 2022, a sufficient sample size would be 126:32, 74:38 and 42:22 in social–emotional, cognitive and language development, respectively.30

3. The development trajectories of the brain are also a main focus of the study. According to total brain volume changes during the first year of life in ASD and non-ASD children reported by Hazlett et al 2017, a minimum of 20 subjects with ASD are needed to obtain statistically significant results.31 The required number of premature infants was ASD cases (20)/incidence (7%)/insurance factor (0.8)=357.

4. As the current study is a multidimensional cohort study that investigates the outcomes at multiple time points, it is challenging to estimate an accurate sample size for building diagnostic models as well as examining complex non-linear relationships between multidimensional data.23 Therefore, we enlarged the sample size to 600, including 400 PT infants and 200 FT infants, in hope to provide sufficient data for multimodal integrative analysis.

Sample recruitment and procedure

The details of recruitment and study process are shown in figure 1. Participants recruitment take place in Peking University Third Hospital, using two methods: (1) mothers who deliver in the Department of Obstetrics will be enrolled after or before birth; and (2) PT and FT infants who visit the Healthcare Center at 40 weeks (±2 weeks) will be recruited as well. Research coordinators will provide eligible parents with a pamphlet containing all necessary information about the study, and will introduce the cohort to them face-to-face. All participants will be recruited only after fully informed and written consents retained.

The Children’s Healthcare Center of Peking University Third Hospital provides primary care on a regular schedule, which was formulated in accordance with the Chinese follow-up guidelines for PT infants (http://www.nhc.gov.cn/fys/mrgzdt/). According to the schedule, infants receive check-ups once a month before the age of 3 months, once every 2 months from 3 to 12 months old, once every 6 months after 1 year old and ends around the age of 3.

Time points of our cohort mainly overlapped with the regular visit before the age of 3, which provide convenience to participating families. Additionally, we will perform extra assessments at specific time points (details are shown in table 1 and figure 2), and keep following up at 4 and 6 years old (further details have been shown in the online supplemental materials).

Study retention

After enrolment, we will stay in contact with all participants through WeChat and messages. Coordinators will send out schedules before each time-point, and an official account will post popular science articles about caring for infants, especially PT infants, and address any healthcare issues raised by caregivers, which are supposed to increase the follow-up rate of the cohort. We will provide the results of the assessments to the parents, and a clinician will provide comprehensive clinical advice at the visit centre. Meanwhile, the families will have the option to receive an electronic explanation instead of visiting the care centre.

During the COVID-19 pandemic, if a family undergoes restrictions and could not finish the off-line assessment, questionnaires will be collected online under telephone instructions. If the restriction finished within 2 weeks, other assessments will be conducted after that. If not, this time-point will be marked as uncompleted.

Data collection

Table 1 and figure 2 summarise the assessment schedule, data collection methods, sample type/domain and the test or task. Data from cases and controls are collected using the same data collection instruments.

General information

Maternal and infant clinical information will be collected at baseline through hospital information system. Demographic information, including social economic status
(SES) and family history will be collected using questionnaires (for details see online supplemental materials).

**Neurodevelopment**
In the current study, neuropsychological functions will be assessed following the guidelines of research domain criteria (RDoC): social process, sensorimotor and cognitive system. In addition, language and communication abilities will also be assessed. Assessment tools have been selected according to age stages, the details are shown in table 1 and figure 2.

### Table 1 Details of the assessments in the cohort study

<table>
<thead>
<tr>
<th>Time points</th>
<th>Data type</th>
<th>Assessments tools/data content</th>
</tr>
</thead>
<tbody>
<tr>
<td>40 weeks</td>
<td>Clinical record and demographic information</td>
<td>HIS system and questionnaire</td>
</tr>
<tr>
<td></td>
<td>MRI, EEG, NIIRS</td>
<td>fmRI, T1, T2, dMRI, EEG, NIIRS</td>
</tr>
<tr>
<td></td>
<td>Neuropsychology</td>
<td>Bayley-III, ASQ-3, CNBS-R2016, ASQ-SE:2, face-to-face still face</td>
</tr>
<tr>
<td></td>
<td>Maternal mental health &amp; Breastfeeding</td>
<td>EPDS, BDI, BAII; feed-type and BSES-SF</td>
</tr>
<tr>
<td></td>
<td>Gene and epigenetics</td>
<td>Saliva sample, faecal sample</td>
</tr>
<tr>
<td>6 months</td>
<td>MRI, EEG, NIIRS</td>
<td>fmRI, T1, T2, dMRI, EEG, NIIRS</td>
</tr>
<tr>
<td></td>
<td>Neuropsychology and language</td>
<td>Bayley-III, ASQ-3, CNBS-R2016, ASQ-SE:2, CSBC, cup task and planning test; CSBC</td>
</tr>
<tr>
<td></td>
<td>Maternal mental health &amp; Breastfeeding</td>
<td>EPDS, BDI, BAII; feed-type and BSES-SF</td>
</tr>
<tr>
<td>12 months</td>
<td>MRI, EEG, NIIRS</td>
<td>fmRI, T1, T2, dMRI, EEG, NIIRS</td>
</tr>
<tr>
<td></td>
<td>Neuropsychology and language</td>
<td>Bayley-III, ASQ-3, CNBS-R2016, ASQ-SE:2, IBQ-R, face-to-face still face, Cup task and Planning test, VABS-III; CSBC</td>
</tr>
<tr>
<td>18 months</td>
<td>Neuropsychology and language</td>
<td>Bayley-III, ASQ-3, CNBS-R2016, ASQ-SE:2, LUI, cup task and planning test; LUI</td>
</tr>
<tr>
<td></td>
<td>Maternal mental health and breast feeding</td>
<td>EPDS, BDI, BAII; feed-type and BSES-SF</td>
</tr>
<tr>
<td>2 years</td>
<td>MRI, EEG, NIIRS</td>
<td>fmRI, T1, T2, dMRI, EEG, NIIRS</td>
</tr>
<tr>
<td>3 years</td>
<td>Neuropsychology</td>
<td>Bayley-III, ASQ-3, CNBS-R2016, ASQ-SE:2, VABS-III, Simon says task</td>
</tr>
<tr>
<td></td>
<td>Maternal mental health</td>
<td>EPDS, BDI, BAII</td>
</tr>
<tr>
<td></td>
<td>Symptoms of ASD and ADHD and other NDDs</td>
<td>ABC, CARS; CBCL, SDQ; YGTSS</td>
</tr>
<tr>
<td>4 years</td>
<td>Neuropsychology</td>
<td>Bayley-III, ASQ-3, CNBS-R2016, ASQ-SE:2, VABS-III, Simon says task</td>
</tr>
<tr>
<td></td>
<td>Maternal mental health</td>
<td>EPDS, BDI, BAII</td>
</tr>
<tr>
<td></td>
<td>Symptoms of ASD and ADHD and other NDDs</td>
<td>ABC, CARS; CBCL, SDQ, DIPA; YGTSS</td>
</tr>
<tr>
<td>6 years</td>
<td>Neuropsychology</td>
<td>CNBS-R2016, WSIC-IV, VABS-III, Stroop color-word test, TMT A and B, RCFT; DAP-IQ</td>
</tr>
<tr>
<td></td>
<td>Maternal mental health</td>
<td>EPDS, BDI, BAII</td>
</tr>
<tr>
<td></td>
<td>Symptoms of ASD and ADHD and other NDDs</td>
<td>ABC, CARS; CBCL, SDQ, DIPA, ADHD rating scale</td>
</tr>
<tr>
<td>3 years</td>
<td>NDDs diagnosis</td>
<td>Clinical interview, DIPA</td>
</tr>
<tr>
<td></td>
<td>Maternal mental health</td>
<td>EPDS, BDI, BAII</td>
</tr>
<tr>
<td></td>
<td>NDDs diagnosis</td>
<td>Clinical interview, DIPA, CRAT</td>
</tr>
<tr>
<td>4 years</td>
<td>NDDs diagnosis</td>
<td>Clinical interview, DIPA, CRAT</td>
</tr>
<tr>
<td>6 years</td>
<td>NDDs diagnosis</td>
<td>Clinical interview, KSADS-PL, CRAT</td>
</tr>
</tbody>
</table>

ABC, autism behaviour checklist; ADHD, attention-deficit hyperactivity disorder; ASD, autism spectrum disorder; ASQ-3, age and stages questionnaire-third edition; ASQ-SE:2, ages and stages questionnaire-social and emotion: second edition; BAI, Beck anxiety inventory; BDI, Beck depression inventory; BSES-SF, breastfeeding self-efficacy scale-short form; CARs, childhood autism rating scale; CSBC, child behaviour checklist; CNBS-R2016, children neuropsychological and behavioural scale revision 2016; CRAT, Chinese reading ability test; CSBC, communication and symbolic behaviour scale; DAP-IQ, draw-a-person intellectual ability test; DIPA, diagnostic infant and preschool assessment; ECBQ, early childhood behaviour questionnaire; EPDS, Edinburgh postnatal depression scale; fNIRS, functional near-infrared spectroscopy; HIS, hospital information system; IBQ, infant behaviour questionnaire; KSADS-PL, schedule for affective disorders and schizophrenia for school-age children-present and lifetime version; LUI, language use inventory; M-CHAT, modified-checklist for autism in toddlers; NDDs, neurodevelopmental disorders; RCFT, Rey complex figure test and recognition trial; SDQ, strengths and difficulties questionnaire; SES, social economic status; TMT, trail making test; VABS-III, Vineland adaptive behaviour scale-III; WSIC-IV, Wechsler intelligence scale for children, fourth edition; YGTSS, Yale global tic severity scale.
1. Standardised overall developmental assessments: several standardised tools will be used to assess the development in infancy, including the Bayley scales of infant and toddler development, third edition (Bayley-III), the ages and stages questionnaire-social and emotional: second edition; CNBS-R2016, children neuropsychological and behavioural scale revision 2016: CSBC; communication and symbolic behaviour scale; DAP-IQ, draw-a-person intellectual ability test; ECBQ, early childhood behaviour questionnaire; EEG, electroencephalogram; IBQ, infant Behaviour questionnaire; LUI, language use inventory; NDDs, neurodevelopmental disorders; RCFT, Rey complex figure test and recognition trial; SES, social economic status; TMT, trail making test; VABS, Vineland Adaptive Behaviour Scale-III; WSIC-IV, Wechsler intelligence scale for children, fourth edition.

2. General cognitive ability: (a) the Chinese version of the Wechsler intelligence scale for children, fourth edition.
3. Language and social communication assessments: previous studies have demonstrated the importance and validity of early screening for predicting later communication disorder, specific learning disorder and other NDDs. In the current cohort, we will comprehensively assess language ability and communication behaviour using the parent questionnaire of communication and symbolic behaviour scale (CSBC) (6 months, 12 months, 24 months) and language use inventory (LUI) (18 months, 24 months). The CSBS consists of 24 items, including three subscales: social, speech and symbolic composite. The LUI provides an assessment for pragmatic development, that is, the ability to use language in social situations.

4. Social–emotional function and temperature: (a) ages and stages questionnaire-social and emotion: second edition (ASQ:SE-2) will be used at 6 months, 12 months, 18 months and 24 months. ASQ:SE-2 is a brief caregivers-reported instrument for children aged 6–60 months; (b) infant behaviour questionnaire-revised (IBQ-R) very short form (12 months) and early childhood behaviour questionnaire (ECBQ) very short form (24 months) will be used to examine the temperature and early emotional competencies, which contain 37 and 36 items in three subscales (surgency, negative affect and effortful control), respectively; and (c) infants’ affective states will be coded during face-to-face still-face paradigm (40 weeks and 12 months).

5. Executive function: age-appropriate tasks were chosen to measure various executive processes, including working memory, inhibition and planning ability: (a) Cup task and Planning test (6, 12 and 18 months); (b) A-not-B with invisible displacement (24 months); (c) Simon says (36 and 48 months); (d) Stroop color-word test, Trail Making Test (TMT) A and B, and Rey complex figure test and recognition trial (RCFT) (6 years) (details of the tasks are shown in the online supplemental materials).

6. Adaptive function: Vineland Adaptive Behaviour Scale-III (VABS-III) will be used at 12 months, 18 months, 24 months, 3 years, 4 years and 6 years to assess the adaptive functions in five domains—communication, daily living skills, socialisation, motor skills and mal-adaptive behaviour.

### Brain development
The key aim of the current cohort is to depict brain developmental trajectories in PT infants during early life stage, and to explore the relationship between early brain development and later diagnosis of NDDs. To fulfil this aim, we plan to evaluate brain development both functionally and structurally using MRI, EEG and fNIRS technologies at five time points (Table 1 and figure 2).

1. Brain MRI: MRI will be performed with a 3T MRI machine (Magneton Tim Trio; Siemens, Erlangen, Germany), multimodal data will be acquired (T1w, T2w, resting-state fMRI (rsfMRI) and diffusion-weighted MRI (dMRI)). The detailed imaging protocols are shown in Table 2.

For infants, sleep scan will be tried for several times, if failed, sedation will be used before scan. At the time

---

**Table 2: Neuroimaging protocol**

<table>
<thead>
<tr>
<th>Modality</th>
<th>Protocols</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1-weighted 3D BRAVO scan</td>
<td>sagittal scan, TR=8.2 ms, TE=3.2 ms, FOV=256 × 256 mm², turning Angle=9°, frequency width=31.25 Hz/pixel, plane resolution=1 ×1 mm², layer thickness=1 mm, 192 layers in total</td>
<td>5 min and 4 s</td>
</tr>
<tr>
<td>T2-weighted dual-echo fast spin-echo scan</td>
<td>Axial scanning, TR=3000 ms, TE1=36 ms, TE2=162 ms, FOV=256 × 256 mm², rotation Angle=90°, Plane resolution=1 ×1 mm², layer thickness=1 mm, 192 layers in total</td>
<td>8 min and 24 s</td>
</tr>
<tr>
<td>rs-MRI</td>
<td>Axial scan with BOLD sensitive GRE-EPI sequence, TR=2000 ms, TE=20 ms, FOV=256 × 256 mm², turning Angle=90°, plane resolution=2 × 2 mm², layer thickness=3 mm, layer spacing=0.6 mm, 41 layers, including 240 time points</td>
<td>8 min</td>
</tr>
<tr>
<td>Diffusion MRI</td>
<td>Axial scan, gradient echo EPI sequence, TR=9000 ms, TE=89.4 ms, FOV=256 × 256 mm², layer thickness=2 mm, Matrix=128 × 128, plane resolution=2 × 2 mm², 75 layers, B=1000 s/mm², 64 gradient weighted directions of spherical distribution, 32 B=1000 s/mm², 32 B=3000 s/mm², 12 B0 images</td>
<td>10 min and 48 s</td>
</tr>
<tr>
<td></td>
<td>FOV, field-of-view; TE, echo time; TR, repetition time.</td>
<td></td>
</tr>
</tbody>
</table>
point of 6 years, awake scan will be scheduled, if necessary, a training session will be performed.

MRI data will be visually checked during scan for excessive motion, insufficient coverage and ghosting. If the data are poor-quality, several tries will be made. T1w, T2w and dMRI images will be transferred to a DICOM workstation and examined by at least one trained radiologist, parents could obtain the reports conveniently through self-service printing in the hospital. If there were any brain anomalies reported, another outpatient visit will be assigned to give further clinical advice to the family.

2. EEG and fNIRS: EEG and fNIRS data will be acquired in (a) a resting state and (b) during emotional face task, details are shown in online supplemental materials.

Genes and epigenetics
The current cohort plans to collect saliva samples at baseline (corrected GA of 40W). Saliva DNA extraction kits will be used for DNA extraction. DNA samples will be stored at −80°C and quality controlled before genotyping. The quality control (QC) standards include the main band of the sample being clear and more than 10 kb, no obvious degradation, a total amount of more than 1 μg and the standard A260/A280 being in the range of 1.7–2.1. DNA methylation (DNAm) will be measured by HumanMethylation450 Bead Chip.

Environmental risk factors
1. Breast feeding: breast feeding is an important early life environmental factor. Recent umbrella review and meta-analysis have suggested that breastfed infants tend to show better intelligence performance and lower ASD prevalence than those who are formula fed. In the cohort study, we assessed feed type and BF self-efficacy using the following methods: (i) according to the proportion of breast milk in total infant food intake, feeding type was divided into six categories—exclusive BF (=100%), almost exclusive BF (=100%), high proportion BF (≥80%), medium proportion BF (≥20%), low proportion BF (>0%, <20%) and token BF (≥0%); (ii) the Chinese version of the breast-feeding self-efficacy scale short form (BSES-SF).

2. Paternal mental health: the importance of maternal mental health for parent–infant attachment and child development has been widely noticed. Moreover, mothers experiencing a PT birth or nursing a PT infant are at high risk for mental health issues. In the cohort study, the following self-report questionnaires will be used at each time-point to assess symptoms of anxiety and depression among mothers—Edinburgh postnatal depression scale (EPDS), Beck depression inventory (BDI) and Beck anxiety inventory (BAI). If any score reaches the cut-off line, the coordinators will inform and communicate with the participants, if necessary, an outpatient visit to a psychiatrist will be assigned.

Outcomes of NDDs
1. ASD symptoms:
   a. Autism behaviour checklist (ABC) is a widely used questionnaire that consist of 57 items in five sub-scales: sensory, relating, motor, language and self-help (18months, 24months, 3years, 4years and 6years).
   b. Modified-checklist for Autism in toddlers (M-CHAT) will be used at 18months and 24months. M-CHAT was developed for early screening, and showed good sensitivity and specificity (0.96 and 0.60).
   c. Childhood autism rating scale (CARS) will be used at 24months, 3years, 4years and 6years. CARS is one of the most widely used questionnaires for assessing ASD symptoms. The CARS consists of 15 items on a four-point Likert scale. A total score of 30 or above indicates ASD diagnosis.

2. ADHD symptoms:
   a. Two versions of the Child behaviour checklist (CBCL) will be used to assess ADHD symptoms and emotional-behavioural problems at 2years and 3years (CBCL for ages 2–3), 4years and 6years (CBCL). CBCL 2–3 is a questionnaire consisting of 99 items in six subscales: emotionally reactive, anxious/depressed, somatic complaints, withdrawn, attention problems, aggressive behaviour and sleep problems. CBCL contains 118 items in nine empirically based syndromes: withdrawn, somatic complaints, anxious/depressed, social difficulties, thought difficulties, attention difficulties, sex difficulties, delinquent behaviour and aggressive behaviour.
   b. Strengths and difficulties questionnaire (SDQ) will be used at 24months, 3years, 4years and 6years. SDQ is a screening tool for ADHD and emotional-behavioural problems. It consists of 25 items including four difficulty subscales (emotional symptoms, behavioural problems, hyperactive-attention deficit, peer relationship problems) and a strengths scale (prosocial behaviour).
   c. ADHD rating scale (6 years old), a norm-referenced checklist that measures ADHD symptoms, which includes 18 items in two subscales: inattention and impulsivity/hyperactivity.

3. Motor disorder: Yale Global Tic Severity Scale (YGTSS) will be used to assess the severity of tic symptoms in five domains—total motor tic score, total verbal tic score, total tic score (motor+verbal), overall impairment rating, and global severity score.

4. Diagnosis: Based on the results of the aforementioned NDDs questionnaires, families of children who surpass the diagnostic cut-off of one of the questionnaires will be arranged a psychiatric interview. Clinical diagnosis of ASD (24 months and later), ADHD (48 months and later) and other NDDs will be made by a licensed paediatric psychiatrist after interviewing the
parent/caregiver and the child according to Diagnostic Statistical Manual Fifth Edition (DSM-5).

In this study, the diagnosis of NDDs will be verified using the semistructured interview of DIPA (after 2 years and before 6 years) or KSADS-PL (Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version) (6 years). The Diagnostic infant and preschool assessment (DIPA) is a semistructured interview tool for caregivers for children from 9 months to 6 years old. It covers symptoms in 13 diagnostic categories, including ADHD, post-traumatic stress disorder, separation anxiety disorder and other mental disorders. DIPA has been translated into Chinese and verified the validity and reliability. Autism Diagnostic Observation Schedule-Second Edition (ADOS-2) will be used to aid in ASD diagnosis, which is the most widely accepted standard for identifying children with ASD. Children who meet the criterion of specific learning disorder in DSM-5 will also undergo a set of comprehensive assessments using the Chinese reading ability test (CRAT).

Data management

Data management system

Data management will be conducted through an electronic system called h6 world (https://www.h6world.cn/). The system enables both participants and doctors/researchers to access the data, which benefits participant retention and real-time monitoring. The procedures of data management are as follows: (1) after enrolment, participants will be assigned a unique research ID code and an electronic CRF (eCRF) will be created using this code. The eCRF will contain all items of questionnaires and results of assessments in the entire follow-up schedule, and the location of storage for neuroimaging data and biosamples, which will also be registered in the eCRF; (2) at each time-point, questionnaires will be collected through eCRF, after finished, a report for each questionnaire will be generated automatically; (3) results of neuropsychological assessments will be entered into the eCRF by the coordinators; and (4) neuroimaging data will be stored and backed-up in a secure data storage facility using the ID code. Biological samples (saliva) will be frozen at −80°C. Neuroimaging data and biosamples will be registered in the h6 world system after collection and storage.

Data monitoring and quality assurance

Before undergoing assessments, coordinators will communicate with the parents about instructions for each assessment and questionnaire, and inform them of the use methods of electronic methods. After completion, each questionnaire will be evaluated by at least two coordinators/doctors for the completeness and reliability. If a questionnaire is of low quality, a telephone or face-to-face interview will be performed to complete and revise the questionnaire. QC methods for neuroimaging data are described in the Statistical analysis section. In addition, monitors will choose a random sample of the collected data and check the quality per month.

Statistical analysis

General statistical analysis methods

Neuropsychological functions, outcomes of NDDs and brain development profiles at each time point will be compared between PT and FT children, and between children with and without NDDs, using linear or logistic regression. The statistical significance level was set as p<0.05. To test early risk factors and signs, multivariate logistic regression will be used to select important risk factors including both biological and environmental determinants for later NDDs diagnosis or poor neurodevelopmental outcomes (online supplemental table S1), (online supplemental figure S1) (further details are shown in the online supplemental materials).

Developmental trajectory analysis

Generalised Linear Mixed Model will be used to compare the differences in developmental trajectories, including neuropsychological functions and neuroimaging between PT and FT infants, and between children with and without NDDs.

As previous studies have indicated, the developmental trajectories of infants are heterogeneous, several subgroups might exist. We intend to use the group-based trajectory modelling method in Proc Traj, SAS V.9.1 software to classify the subjects according to the trajectories. Bayesian Information Criterion (BIC) and mean post-test grouping probability (AvePP) will be used to test the model fit.

Neuroimaging data analysis

Details of analytic methods during preprocessing the MRI, EEG and fNIRS data are shown in the online supplemental materials.

Genetic and epigenetic data analysis

Details of analytic methods are shown in the online supplemental materials.

Patient and public involvement

The design of the current study did not involve participants or parents directly. However, the research coordinators will maintain contact with parents/caregivers through email, telephone or social media, and we will collect feedback during study procedure. Messages with necessary information, including appointment time and location will be sent before each follow-up visit.

Generally, PT children require special attention from their parents in daily life, which could be difficult and cause significant pressure to parents. The current cohort hopes to provide comprehensive healthcare not only for the children, but also for their parents, especially in terms of mental health. The cohort will screen and monitor parental mental health and provides further clinical consultants to all participating parents.
ETHICS AND DISSEMINATION

Ethical approval has been obtained from the Peking University Third Hospital medical science research ethics committee (M2021087). The ethical approval was obtained from one site only because all the procedures will be conducted at the Third Hospital. Although psychiatric interviews will be conducted by doctors from the Sixth Hospital, they will also take place at the Third Hospital.

The study results from the current cohort will be disseminated and popularised through social media to participating parents, as well as parents who are giving care to PT children.

Author affiliations

1Child and Adolescent Mental Health, Peking University Sixth Hospital, Beijing, China
2Department of Pediatrics, Peking University Third Hospital, Beijing, China
3Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands
4State Key Laboratory of Cognitive Neuroscience and Learning and International Digital Group/McGovern Institute for Brain Research; Center for Collaboration and Innovation in Brain and Learning Sciences, Beijing Normal University, Beijing, China

Contributors YZ designed the study and drafted the manuscript, LY and YH conceived the study, read and revised the manuscript and gave the final approval for the version to be published. XT gave necessary advice about the study plan and further supervised the manuscript, YL conceived and designed the study and contributed to the concept and management of the research. DW, XG, RX and NW contributed to planning the study. All authors read and approved the final version of the manuscript.

Funding Beijing Municipal Science and Technology Commission (Z18110001518005), National Natural Science Foundation of China (grant numbers: 81671358), the Natural science foundation of Beijing municipality (M22018).

Competing interests None declared.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Consent obtained from parent(s)/guardian(s).

Provenance and peer review Not commissioned; externally peer reviewed.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iDs

Yilu Zhao http://orcid.org/0000-0003-3219-3188
Xuping Gao http://orcid.org/0000-0003-3813-2036

REFERENCES


Supplementary Materials

Supplementary methods:

1. Collection of general information

Maternal and infant clinical information will be collected at baseline according to the hospital information system (HIS), and recorded in the eCRF by research coordinators. Socio-economic information will be collected using e-CRF, including family income, maternal and paternal educational attainment, maternal and paternal occupation, maternal and paternal age, parity (total live births including index birth), smoking during pregnancy, drinking alcohol during pregnancy, rural/urban residence. Family histories will also be collected using e-CRF. In each question, blanks or several options will be given. The demographic variables and the descriptions and categories of them are listed in the Table 1.

Table 1. The demographics information collected in the current cohort.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description or Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gestational age (GA)</td>
<td>1) Full-term (&gt;37 weeks), 2) extremely preterm (&lt;28 weeks), 3) very preterm (28 to 32 weeks), 4) moderate to late preterm (32 to 37 weeks)</td>
</tr>
<tr>
<td>Birth weight (BW)</td>
<td>1) Small for gestational age (SGA): birthweight below the 10th percentile for gestational age; 2) Appropriate for gestational age (AGA): birthweight between the 10th and 90th percentiles for gestational age; 3) Large for gestational age (LGA): birthweight above the 90th percentile for gestational age</td>
</tr>
<tr>
<td>Birth length (cm) and head circumference</td>
<td>The exact number will be extract from medical records.</td>
</tr>
<tr>
<td>Birth complications</td>
<td>Any birth complications have been diagnosed and recorded.</td>
</tr>
<tr>
<td>Perinatal complications</td>
<td>Any perinatal complications have been diagnosed and recorded.</td>
</tr>
<tr>
<td>Family income</td>
<td>Total pretax per-month incomes from all sources during the calendar year before the birth, categorized in 1) less than 3000 RMB (&lt;$435), 2) 3001-6000 RMB ($435-$867); 3) 6001-10000 RMB ($867-$1449); 4) 100001-30000 RMB ($1450-$4349); 5) 30001-50000 RMB ($4350-$7246); 6) more than 50001 RMB (&gt;7246)</td>
</tr>
<tr>
<td>Maternal and paternal age</td>
<td>&lt;20, 20-24, 25-34, 35-39, &gt;40</td>
</tr>
<tr>
<td>Maternal and Paternal occupation</td>
<td>Categorized into 1) Science and technology professionals; 2) Administrative personnel; 3) Education and cultural work; 4) Medical and health; 5) Business work; 6) Military personnel; 7) Social service work; 8) industrial workers; 9) Agriculture, forestry, animal husbandry, and fishery labor; 10) Self-employed; 22) Unemployed; 12) Retired; 13) Others</td>
</tr>
<tr>
<td>Parity</td>
<td>1) Primiparous; 2) 2–4 births; 3) &gt;5 births</td>
</tr>
<tr>
<td>Smoking during pregnancy</td>
<td>Any smoking or none</td>
</tr>
<tr>
<td>Drinking alcohol during pregnancy</td>
<td>Any alcohol use during pregnancy or none</td>
</tr>
<tr>
<td>rural/urban residence</td>
<td>1) Urban area; 2) County town; 3) Township; 4) Rural area</td>
</tr>
<tr>
<td>Family histories</td>
<td>Medical histories up to three generations of the children’s relatives for the following diseases: 1) mental disorders; 2) neurodevelopmental disorders; 3) dementia; 4) epilepsy; 5) Congenital developmental malformation; 6) hypertension; 7) obesity; 8) diabetes mellitus; 9) cardiovascular diseases.</td>
</tr>
</tbody>
</table>

2. Tasks for assessing the executive function (EF):

1) Cup and Planning task:
Our experiments will follow the procedure described by Feng and the colleagues [1]. During the Cup task, infants will sit on their mother's lap and complete a task of finding a toy on a tabletop under a series of instructions. The difficulty of the task varies with the number of cups hiding the toy and the "delay time" from hiding to allowing the child to start searching. The task begins with the easiest one-cup trial. An operator hides a toy under an opaque cup and ensures that the infant is watching. The delay time is divided into three levels, starting at 0 seconds, increasing to 4 seconds, and finally to 10 seconds. Three trials are conducted for each time delay. If an infant successfully completes nine trials or does not receive the toy in three consecutive trials at a certain delay time, the next task continues. In the two-cup and three-cup tasks, the toy is hidden under two or three opaque cups and divided into four delay time levels (0, 2, 4, and 10 seconds), with nine trials for each delay time. The toy is initially hidden in location A (under the left or right cup), with no delay. After the infant successfully attempts three times at location A, the toy is moved to another location B (under the other cup). When the infant completes all 36 trials in the two-cup task or fails three consecutive trials, the three-cup test continues. The three-cup test also follows the same transfer method, with the middle cup serving as interference, and the toy not hidden underneath, but only transferred under the left and right cups. When the three-cup test also meets similar termination conditions, the experiment ends. Score will be coded into three categories: working memory, inhibitory control of dominant stimuli, and inhibitory control of interference. A). The score for working memory is based on the participant's memory of location A in the two-cup and three-cup tasks. The infant receives a score of "1" each time they remember the location of the toy (i.e., retrieve the toy from the correct location). The total number of correct trials is then calculated. Therefore, the total score for working memory ranges from 0 to 56, with a higher score indicating better performance and
better working memory. B). The score for inhibitory control of dominant stimuli is based on the infant's performance in finding the correct hiding place in the two-cup and three-cup tasks and finding the toy when moving from location A to location B. The method is as follows: ① if the infant finds the toy at location B and makes a correct response to three consecutive tests at location A, they receive a score of "3"; ② if the infant finds the toy at location B and makes correct responses to two correct locations, they receive a score of "2"; ③ if the infant finds the toy at location B and makes a correct response to one location in one trial, they receive a score of "1". Then, the total score is divided by the total number of tests performed at position B. Thus, the total score ranges from 0 to 3, and a higher score indicates stronger inhibitory control for dominant response. C). The score for inhibiting external stimulus interference is calculated in the 1-cup, 2-cup, and 3-cup tasks, based on the infant's ability to inhibit interference by the cup covering the toy. When the infant reaches for the hidden toy and continues to play with the cup for more than 2 seconds instead of retrieving the toy, they receive a score of 1. The final score is obtained by dividing the total score by the number of times the infant reaches for the correct position, with a range of 0-1. A higher score indicates weaker inhibitory control for external stimulus interference.

In the Planning task, the infants need to clear obstacles to retrieve a toy, consisting of three different difficulty levels. First, in the first step of the test, the toy is placed in the farthest corner of a cloth, and the infant cannot reach it directly. Instead, they need to use an appropriate strategy, which involves pulling the corner of the cloth closer to them to retrieve the toy. Second, a transparent barrier is placed in front of the cloth, and the infant must move it aside before pulling the cloth to retrieve the toy. Third, a string is tied to the toy and the other end is attached to the farthest corner of the cloth. The infant needs to perform three steps to retrieve the toy: moving the transparent...
barrier, pulling the cloth to reach the string, and finally pulling the string to retrieve the toy. Each level is repeated three times, and if the infant succeeds in retrieving the toy at least once out of three attempts, the next level is performed. If the infant fails to retrieve the toy in any of the three attempts, the test is terminated. Finally, a score is given based on the infant's performance in each step. For example, if the infant is not aware of the task or does not pull the cloth, they score zero. If the infant pulls the cloth but does not retrieve the toy, they score one. If the infant pulls the cloth and retrieves the toy, they score two. The maximum scores for the first, second, and third steps are six, twelve, and eighteen, respectively. A higher score indicates better planning ability.

2) A-not-B with invisible displacement
The A-not-B task was conducted according to the procedure described by Bell and colleagues [2]. The child was seated on the caregiver’s or experimenter's lap facing a small table. An attractive toy was placed under an opaque cup in the center of the table. The experimenter tapped the center of the table and said, "Look!" to attract the child's attention to the cup. The experimenter then lifted the cup, allowing the child to retrieve the toy. The cup was then replaced in the center of the table, and the experimenter repeated the "Look!" and cup-lifting sequence for a total of four trials. Next, the cup was moved to a new location (B), which was equidistant from the original location (A) and the child's midline. The experimenter then placed a barrier on the table in front of the cup, preventing the child from seeing its location. The experimenter tapped the barrier and said, "Find it!" to cue the child to search for the toy. After a 5-second delay, during which the experimenter looked straight ahead and kept the child's attention centered, the experimenter lifted the barrier and the child was allowed to search for the toy. The child's first reach toward either the original location (A) or the new location (B) was recorded as a correct or incorrect response. If
the child made two consecutive correct responses to the same location, the cup was moved to the opposite location in a pattern of AAB. The task continued until the child completed 12 trials (6 per location).

3) Simon says:
The Simon says task is a cognitive task that assesses inhibitory control in children. The task involves giving instructions to children to perform certain actions, but only if the instructions are preceded by the phrase "Simon says." [3]. In this particular study, children were instructed to obey instructions given by a nice horse and to ignore instructions given by a mean cow followed the procedure described by Cuevas and Bell [4]. The task consisted of ten test trials, with half of them being inhibition trials involving the mean cow/bull and the other half being control trials involving the nice horse/pig. The children's responses were coded as correct or incorrect, and their performance was measured as the proportion of correct responses on inhibition trials.

4) Stroop color-word test, Trail Making Test (TMT) A and B, and Rey complex figure test and recognition trial (RCFT)
Stroop color-word test: In the first part, participants are asked to read out loud a list of color words printed in black ink. In the second part, participants are asked to name the ink color of a list of color words that are printed in a different ink color than the written word (e.g., the word "red" printed in blue ink). In the third part, participants are asked to name the ink color of a list of color words that are incongruent with the written word (e.g., the word "red" printed in green ink) [5].
TMT: The test consists of two parts. Part A requires the individual to connect numbered circles in sequential order as quickly as possible. Part B requires the individual to connect circles alternating between numbers and letters (1-A-2-B-3-C, etc.) in sequential order as quickly as possible. The time taken to complete each part is recorded and used as a measure of cognitive function [6].
RCFT: Children will be instructed to remember the RCFT figure within 30 seconds and subsequently draw what they remembered immediately (without intervening distraction) and then a 20-30 minutes' delay (other tests were conducted during this delay) [7]. Detail and structure scores were rated using 36-point and 6-point, respectively. The 36-point scoring system was initially devised by Osterrieth [8] and adapted by Booth [9], which could reflect both local and global VSWM. In the 36-point system, each figure is scored using 18 features: two points are given if the feature is placed correctly, and one point given if it is incomplete or placed poorly. The 6-point scoring system was devised by Binder [10], evaluated the accuracy and completeness for constructing five configural elements and the addition of the base rectangle, which was used to evaluate the organizational strategies during RCFT task and global processing.

2. EEG and fNRIS data collection:

The bioelectric amplifier was provided by the Symtop Instrument Co, Beijing, China. Three Ag-AgCl– disk electrodes were attached to the scalp for noninvasive registration by using Ten20 conductive paste (DO Weaver and Co, Aurora, CO). 1) At the baseline (40w), EEG data will be collected from ten electrodes at F3, F4, C3, C4, T3, T4, P3, P4, O1 and O2 (including frontal, occipital, temporal, parietal lobe and central area. According to the international 10-20 system), the reference electrodes will be placed over the mastoids (A1, A2), and the ground electrode on the forehead.

2) Since 6 months of age, EEG data will be collected from all the 19 electrodes (10-20 system).

1) Resting state: Resting state data will be acquired during relaxed wakefulness and sleep state. After the electrodes and sensor pads were placed, infants will be coaxed to fully relax by their parents or caregivers.
2) Emotional face task: Previous studies have demonstrated that facial expressions are important social signals since very early life stages, for instance, newborns showed different responses to happy and fearful expressions [11]. In the current task, we plan to test the biological response to different facial expressions using EEG and fNIRS. We chose 10 images of Chinese females posing happy, angry, fearful, and aversive facial expressions and 40 images of neural expressions as stimulus. Each section contains 10 neutral and 10 emotional images, which will be presented in a random order in the center of the screen. 2 seconds inter-trial interval (ITI) will be given. Each image will be presented three times, i.e., each section contains 60 tests. The order of the eight sections will be randomly assigned. To attract the attention of infants, white balls or squares will appear in the screen before the presentation of images, the presentation time is 400 ms. And the facial expression images will be presented for 800 ms, with an inter-stimulus interval (ISI) of 500-800 ms. Jitter time will be set within 100-300 ms [12]. We will focus on several ERP components which have been indicated underlying the socio-emotional process in infants, including N170, P300, P1, N290 [13-15].

3) Quality control: Channels will be referenced off-line to the average of the left and right mastoids [(A1 + A2) / 2], sampled at 250 Hz, and band-pass filtered (0.1–30 Hz). Using automated artifact detection, individual channels that exhibit an ERP amplitude exceeding 200 mV will be rejected, and any trial that contains more than 15% bad channels will be excluded. The bad channels will then be interpolated using a spherical spline algorithm to replace the missing data. Any epoch contaminated by eye blinks, eye movements, or muscle potentials exceeding ±150 μV at any electrode will be excluded. Following the automated artifact detection process, each trial will undergo a meticulous manual inspection for ocular and electromyographic artifacts. Manual inspection allows us to detect subtle artifacts that may have been overlooked.
by the automated procedure. Our team of experienced technicians will inspect each trial in detail, marking any identified artifacts for removal.

3. Statistical analysis
The current study aims to identify the early determinants of later outcomes using multidimensional data across different time-points using a various of statistical methods. The main outcomes and determinants have been listed in SFigure 1. The outcomes include NDDs, neuropsychological function, and brain development, while the independent variables consist of biological, phenotypical, and environmental data. Distinct multi-variate regression models will be used to investigate the effects of the interested determinants on each outcome. When the MRI imaging indexes were using as dependent variable, additional confounding variables will be controlled, including sedation and head motion.
To investigate the developmental trajectories across the interested outcomes. GLMM and GBTM will be used, as mentioned by the main text.
Multiple imputation by chained equations (MICE) will be conducted to impute the missingness. Inverse probability weighting and stabilized inverse probability weighting models will be utilized to address the bias caused by attrition.
Supplementary figure 1. Main outcomes and predictors in the current study.

4. MRI data analysis:

1) Brain structure: Whole brain three-dimensional data: using SPM12 (www.fil.ion.ucl.ac.uk/spm/software/spm12/) software quality, pretreatment, image nonuniformity correction. The Optimized voxel-based morphometry (OVBM) method was used to establish the DARTEL template for the whole brain and different tissues (gray matter and white matter). The volume volume, cortical surface area and cortical thickness of individual whole brain and ROI were calculated.

2) Diffusion MRI: DTI Studio software was used for data quality control and preprocessing. It mainly includes automatic detection of artifacts, correction of head
movement and eddy current deformation, model reconstruction based on dMRI images, calculation of total white matter fractional anisotropy (FA), mean diffusivity (MD), axial diffusivity (AD), and radial diffusivity (RD) images. Then, based on previous studies, white matter fibers related to social function were selected as sites of interest (including internal capsule, corpus callosum, superior longitudinal fasciculus, inferior longitudinal fasciculus, fornix, arcuate fasciculus, and uncinate fasciculus). Multi-shell tracing-imaging technology was used to reconstruct the target white matter fiber tracts and calculate the values of each index.

3) Brain function (rs-fMRI): SPM12 and Gretna software were used to complete the pretreatment, and then regional homogeneity (ReHo) was calculated at the individual level. Kendall's coefficient concordance (KCC) is used to represent the local consistency of the selected voxel and its neighboring voxel time series. The mean amplitude of brain low-frequency oscillatory blood oxygenation level dependent (BOLD) signal in a short period of time was calculated (low-frequency amplitude, Amplitude of low-frequency fluctuation, ALFF). The following seed point functional connection analysis and graph theory topological attribute analysis will be carried out by SeeCAT and Gretna software packages.

5. EEG/fNIRS data analysis

1) EEG: ICA method will be used to remove noise components such as ocular in the data. Then the EEG data will be preprocessed, including filtering, segmentation and baseline correction. For event-related EEG, ICA will be used to isolated the independent components related to emotion as mentioned above and calculate the differences of time-frequency energy, phase locking and other indicators among different emotional conditions.

2) fNIRS: The fNIRS data will be analyzed by Homer software (Hemo-Dynamic
Response (Homer)) and PMI toolkit (Photon Migration Imaging) developed by the Photon Imaging Laboratory of Harvard University.

6. Genetic and epigenetic analysis

Quality control of genetic and epigenetic analysis will be performed following standardized procedures [16, 17].

Genome-wide association analysis (GWAS) will be carried using PLINK 1.9 and xx. Regression analysis will be performed using each main outcome as specified in the next section [18]. Effects of the potential confounding factors will be controlled, including demographic characteristics, GA, sex, birthweight. A significant genome-wide P threshold will be set at 5E-08, indicative P threshold as 1E-05.

Generalized estimating equation (GEE) models with an identity link function and an independence correlation structure will be utilized to evaluate the associations of main outcomes with CpG-specific DNA methylation [19]. An FDR < 0.05 was considered statistically significant. For each CpG associated with interested outcomes with FDR < 0.05, we evaluated the interaction effects of the main outcomes and genotype of SNPs within ±1 Mb window of the CpG on the CpG’s methylation level [20].

7. Detailed follow-up procedure

1) After recruitment, the research coordinators will register the basic information for the family, and will schedule the date for the next regular visit and the extra assessment.

2) The research coordinators will keep in contact with the families through WeChat and messages. And They will send an online link for parent-reported questionnaires two days before the visit. Completion of the questionnaires will be checked in-person during the visit.

3) Before age 3, expert-administered questionnaires and other assessments (except for
the MRI, EEG, and fNIRS) will be scheduled for one to two days prior to the visit or on the morning of the same day as the regular visit. Adequate rest breaks will be provided. During the regular visit, clinicians will explain the reports of the extra assessments and give further advice.

4) After the child turns three, we will contact parents to schedule follow-up visits at their convenience. After follow-up visit, we will send electronic reports and brief explanations to families through social media, and arrange clinical visits as needed.

5) MRI, EEG, and fNIRS assessments will be assigned the same day, within one week prior to the regular visit. EEG and fNIRS will be conducted before MRI. MRI scans will be conducted during the infant's regular nap time.

8. Details of follow-up schedule
References:


[18] S. Purcell et al., “PLINK: A Tool Set for Whole-Genome Association and
