BMJ Open

Protocol for a longitudinal twin birth cohort study to unravel the complex interplay between early-life environmental and genetic risk factors in health and disease: the Chongqing Longitudinal Twin Study (LoTiS)

Chao Tong,[1,2] Li Wen,[1,2] Yinyin Xia,[2,3] Pamela Leong,[4,5,6] Lan Wang,[7] Xin Fan,[8] Ting-li Han,[9] Jeffrey M Craig,[5] Philip Baker,[5,10] Richard Saffery,[4,6] Hongbo Qi[1,2]

Check for updates

For numbered affiliations see end of article.

Correspondence to
Professor Hongbo Qi;
qihongbo@sina.com

ABSTRACT

Introduction Non-communicable diseases (NCD) now represent the major burden of adverse health in most countries. It is clear that much of the risk of such conditions begins very early in life, potentially in utero. Given their complex aetiology, an understanding of the origins of NCD requires an in-depth analysis of the interplay between genetic variation and environment, preferably over time. For decades, twin studies have played a key role in understanding such traits. Their strength lies in the ability to disentangle genetic and environmental factors that contribute to a phenotype. This is done by comparing genetically identical monozygotic (MZ) with dizygotic twins, who share on average 50% of genetic variation, or by comparing MZ twins within a pair. This study aims to determine the relative contributions of genes and environment to early-onset intermediate phenotypes related to later adult onset disease (such as growth and neurodevelopment) and to identify specific biomarkers and time points for emergence of phenotypes from infancy, largely independent of underlying genetic factors.

Methods/design The Chongqing Longitudinal Twin Study (*LoTiS*) will recruit 300 women pregnant with twins, enriched for MZ pregnancies, with follow-up to 3 years of age. Data collection will be undertaken at key time points in gestation (×3), at delivery and postnatally (×9). Maternal and infant biospecimens including blood, urine, hair, nails and buccal swabs along with measures such as fetal scans and body measurements will be collected. Additional information from questionnaires and medical records includes pregnancy, diet, sociodemographics, maternal stress, and infant growth and neurodevelopment.

Ethics and dissemination This study has been approved by the Ethics Committee of Chongqing Medical University (record no: 201530) and has been registered with the Chinese Clinical Trial Registry (registry no: ChiCTR-OOC-16008203). Results of the recruitment and all subsequent analyses will be submitted for publication in peer-reviewed journals.

Strengths and limitations of this study

► The wealth of longitudinal data and biospecimens collected as part of this study will enable an unparalleled investigation of the factors contributing to early-life health and disease.

► *LoTiS* (Longitudinal twin study) will be the largest dedicated longitudinal study yet undertaken of twins commencing prior to birth.

► The nature of cohort studies means that only limited paternal data will be available.

► A focus on Chinese participants may limit the generalisability of some findings to other populations.

► Some biospecimens, such as placenta, will not be fractionated to purified cell populations with implications for downstream analyses. Others, such as stool for microbiome analysis, will not be collected.

Trial registration number ChiCTR-OOC-16008203; Results.

BACKGROUND
Rising burden of non-communicable diseases

Non-communicable diseases (NCDs), comprising mainly cardiovascular disease, cancers, diabetes and chronic lung disease, caused 36 million deaths in 2008, with nearly 80% in low-income and middle-income countries.[1] Factors such as increasing income, mechanisation, urbanisation and improved access to processed food have rapidly changed public nutrition and lifestyle in developing countries, potentially contributing to a greater risk of chronic disease.[2] While infectious diseases were the leading cause of death in the last century, potentially

preventable NCD, including obesity and cardiovascular disease, asthma, allergy and autoimmune disease, cancer and mental health problems, will be the major killers in the 21st century. For example, the prevalence of self-reported NCD in Guangzhou, China, was 16.0% in 2009, higher than the 13.5% in 1999, according to the Regional Planning authority (2001–2005). The spectrum of NCD also changed during that time, with diabetes rising from the sixth most common disease to the second, replacing coronary heart disease.[3]

Fetal programming and Developmental Origins of Health and Disease

Many NCD in adults have their origins in early life. The majority are complex with multiple inter-related factors operating at many levels over time. Little is understood about how and when risk factors translate into chronic conditions, but it is clear that the emerging epidemic, particularly in developing countries, cannot be explained by genetic variation in isolation. Nevertheless, genetic variation is undoubtedly the foundation on which additional environmental-induced risk is established. Mounting evidence suggests that environmental cues such as maternal nutrition during pregnancy and early postnatal life play a key role. Pivotal to this emerging paradigm was the observation that individuals born with low birth weight were at higher risk of cardiovascular disease.[4] This idea has evolved into a broader hypothesis spanning many exposures and outcomes, the Developmental Origins of Health and Disease (DOHaD). DOHaD is supported by many direct animal studies and fewer human observational studies. The paradigm emphasises the likelihood that early-life exposures may influence developmental pathways and induce long-lasting changes in physiology, manifesting as changes in immune, metabolic, neurodevelopmental and cardiovascular health,[5] although the underlying mechanisms remain to be fully determined.

Despite this, it is increasingly clear that molecular/cellular processes act as the mediators of the influence of environmental exposures on the underlying genome, particularly early in development. However, gaining a full understanding of the interaction between exposure and genetic variation, molecular mediators and phenotypic outcomes is complex, generally necessitating large sample sizes studied longitudinally. This is primarily to adequately take into account the effects of dynamic non-genetic variation and the widespread genetic heterogeneity in human populations. The potential of twin studies to overcome many of these issues is now widely appreciated.

Utility of twin studies

Twin studies have played an important role in our understanding of individual variation for over a century. The comparison between monozygotic (MZ) and dizygotic (DZ) twins allows an estimation of the heritability of traits, which describes the proportion of the total variance that is attributable to variance in genetic

factors within any given population. Further, using well-established analytical approaches, the relative contributions of genetic, shared and non-shared environmental factors to a phenotype of interest can also be estimated. A particularly powerful approach to studying environmental influence on phenotype lies in the within-in-pair analysis of MZ twins, discordant for a phenotype of interest. Interestingly, despite being genetically identical, concordance rates for most complex diseases are well below 100%, indicating a role of non-shared environmental and/or stochastic factors in modulating phenotype.

Longitudinal twin studies now offer the opportunity to improve our understanding of the causal pathways underpinning a range of early-life phenotypes, particularly factors regulating mechanistic processes (eg, epigenetics, transcriptomics, metabolomics) thought to influence cell function and physiology. The strategic use of identical twins discordant for complex disease continues to reveal the molecular and cellular processes underpinning the disease process. Of particular note, we and others have demonstrated the importance of non-shared environments of twins in utero, particularly the effects of maternal and nutrient supply line factors (eg, placental structure, umbilical cord insertion point) on neonatal epigenetic profile. These effects appear tissue specific and locus specific, emphasising that responsiveness to environmental exposures in utero cannot be generalised.[6] Such complexity in responsiveness to environmental exposures in utero has implications for any study investigating the developmental origins of health and disease. The longitudinal study of twins, commencing prior to birth, provides a unique opportunity in a modest sample size to explore the relative role of environmental and genetic factors to a range of early complex phenotypes.

To date, there have been few attempts at establishing longitudinal twin birth cohorts. Our current study, the Peri/Postnatal Epigenetic Twin Study (PETS), was the first dedicated longitudinal twin birth cohort study internationally and was established to allow the investigation of factors contributing to epigenetic variation in humans at birth and beyond.[7] However, although PETS has revealed novel insights into DOHaD phenomena, participants are almost uniformly of European ancestry and were recruited from a population with a 'Westernised' lifestyle. As such, the generalisability of findings to non-European, low/middle-income nations remains unclear. Additionally, PETS did not collect longitudinal maternal samples for measurement of environmental pollutants, important in any modern cohort aimed to explore DOHaD causal pathways. In order to address these issues, we have recently established the Chinese Longitudinal Twin Study (*LoTiS*), commencing prior to birth.

AIMS

1. To establish a cohort of 300 Chinese twin pairs and their mothers recruited early in pregnancy with
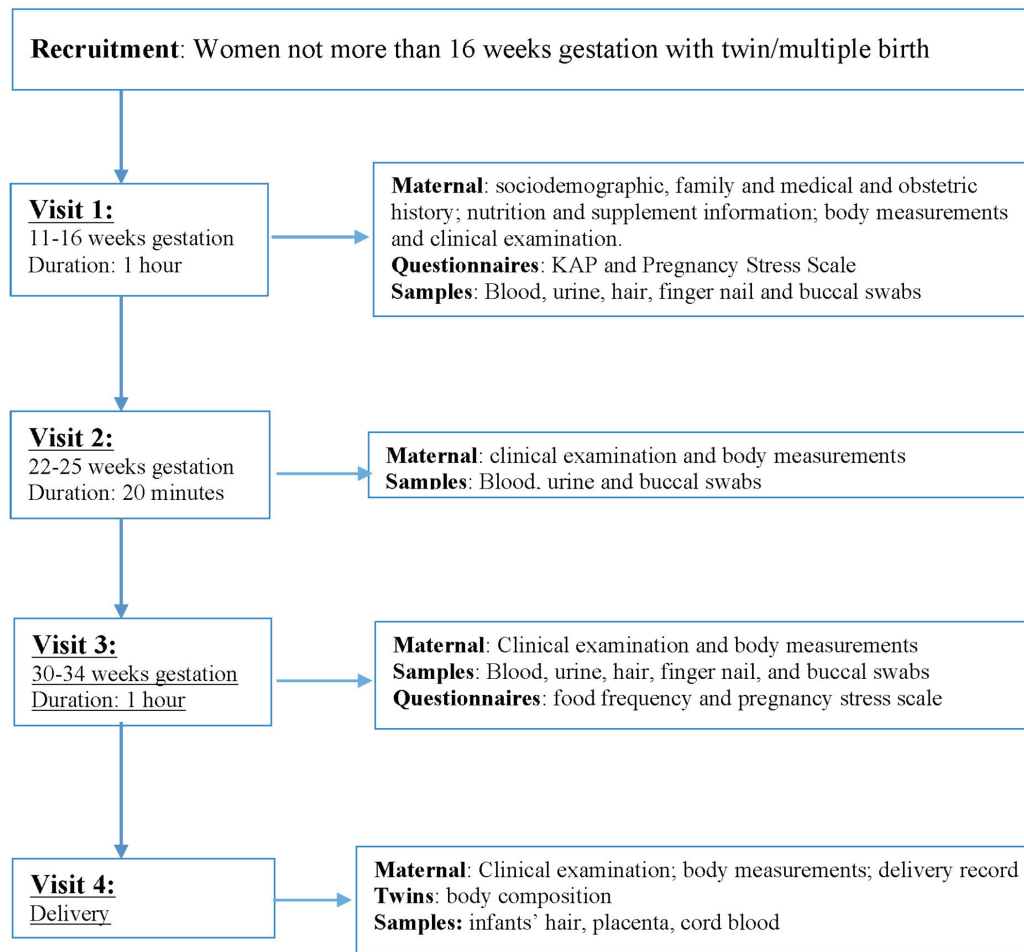
Figure 1 Flowchart: recruitment to delivery. KAP, knowledge-attitude-practice.

longitudinal assessments of exposure and outcome measures and biospecimens.

2. To define the relative contribution of genes and environment (shared and non-shared) to early-onset illness or intermediate phenotypes related to adult onset disease.

3. To identify specific biomarkers and time points for emergence of phenotypes from infancy, largely independent of underlying genetic factors.

METHODS/DESIGN

LoTiS is a pregnancy cohort based in Chongqing Medical University (CQMU), China. The approach for assembly of the cohort is based on the successful PETS cohort that has been ongoing since 2006.[7] As with PETS, women pregnant with twins will be recruited. We aim to recruit 300 women over a 2-year period who are between 11 and 16 weeks' gestation (figure 1) and attending the Outpatients Department of Obstetrics, First Affiliated Hospital of CQMU, or Department of Obstetrics, Chongqing Health Centre for Women and Children. We will follow infants through to 3 years of age. All prenatal visits will be conducted at the First Affiliated Hospital of CQMU and where possible, mothers will deliver at this hospital also.

Postnatal visits will be conducted either at the Hospital and Maternal Child Care Centre, Chongqing, or at CQMU. All questionnaires are derived from calibrated scales and measures and will be completed in face-to-face interviews by trained interviewers. Given the size and duration of the initial phase of the study, we will focus on quantitative outcomes (continuous variables) often showing discordance within young twin pairs. This includes measures of growth and adiposity, body composition, allergic and neurodevelopmental outcomes all thought to involve environmental modulation of underlying genetic risk. Exposure, outcome and molecular data generated during this study will be analysed using standard twin modelling to partition variance into genetic, shared and non-shared environments.

STUDY POPULATION AND RECRUITMENT

Mothers pregnant with twins will be recruited from the Outpatients Department of Obstetrics at the First Affiliated Hospital of CQMU and Department of Obstetrics, Chongqing Health Center for Women and Children, between November 2015 and December 2017 while attending antenatal assessments with an obstetrician

(figure 1). Women approached in the outpatients department by the project manager (PM) will be provided with information related to the study. Special emphasis will be made in regard to the longitudinal nature of the study and the need to provide multiple biospecimens, as well as to respond to a series of questionnaires throughout the duration of the study. Women will also be informed that some biospecimens are optional, requiring a separate consent, and that their decision whether to provide the optional samples will affect their participation in the study. Following an opportunity to ask any questions related to the study, women will be given the option to consent to participate. Women who meet the inclusion criteria for the study and consent to participate will then receive an appointment to attend the outpatients department for their first study visit.

INCLUSION AND EXCLUSION CRITERIA
Inclusion criteria
► Pregnant women attending the Outpatients Department of Obstetrics at the First Affiliated Hospital of CQMU and Department of Obstetrics at Chongqing Health Center for Women and Children.
► Gestational age of infants is not more than 16 weeks.
► Capacity to provide written, informed consent.

Exclusion criteria
► Women with a single pregnancy or multiples higher than twins (triplets, etc).
► Women more than 16 weeks' pregnant.
► Women with a diagnosis of any of the following in pregnancy will be withdrawn from the study by the investigators and the reason will be noted:
  – death of one/both infants;
  – twin to twin transfusion syndrome;
  – twin anaemia-polycythaemia sequence;
  – twin reversed arterial perfusion sequence;
  – malformation of one/both infants.
► Women who deliver at a birthing centre other than the First Affiliated Hospital of CQMU or Chongqing Health Center for Women and Children will be excluded from the study at this point due to the inability for researchers to obtain clinical birth related information and samples.

DATA COLLECTION
Maternal information
Information regarding past and present medical and health history, the current pregnancy and sociodemographic factors (age, education, occupation, families and lifestyle factors such as smoking and alcohol consumption history) are collected from pregnant women over three prenatal visits to the outpatients department of the CQMU or Department of Obstetrics at Chongqing Health Center for Women and Children. The first of these visits will occur when the women are between the 11 and 16 weeks' gestation period, with the second and third visits taking place between the 22th–25th and 30th–34th week of gestation, respectively. A clinical examination will be conducted at each visit and a range of body measurements (weight, height, waist circumference and blood pressure) recorded. Biological samples collected over this time include blood, urine, hair samples, finger nail clippings and buccal (cheek) swabs. Results of liver and kidney function tests and fetal ultrasounds will be recorded. Questionnaires seeking information about maternal stress during pregnancy and dietary habits (including supplements) will also be completed.

Twins' information: prenatal
At the 15th–16th week gestation ultrasound scan, twin 1 and twin 2 will be assigned according to location (including placentation site). Measures of estimated fetal weight, biparietal diameter, head circumference, abdominal circumference and umbilical insertion point will also be recorded. Before delivery, an ultrasound will be performed to confirm the final intrauterine positions of twin 1 and twin 2. Where there is clear evidence of opposite sex DZ twins at any ultrasound, these participants may be excluded from further participation in the study as needed. Further, following birth, PCR-STR (short tandem satellite repeats) will be performed to identify the zygosity in all same-sex twins. The same sex DZ twins can then also be excluded as a strategy to enrich the twin cohort for MZ twin pairs. During the delivery, before the placentas are delivered, the obstetrician on duty marks the umbilical cords by clamping a different number of haemostatic forceps onto each according to twin number. The placentas are then correlated with the newborns after determining birth weight.

Birth information
At the time of delivery, the PM will attend the birth. Detailed information is collected from the medical records related to the pregnancy, delivery and any medical complications. This includes features of amniotic fluid (eg, volume), umbilical cords (length and NuA (cord around neck)) and the placenta/s (size and weight). The sex of the infants, their birth weight and birth length will be recorded along with Apgar scores and whether admission to the neonatal unit was required, the reason and length of stay. Hair samples from each twin will also be collected along with placental tissue samples and cord blood (figure 1). Exceptions may occur when an emergency delivery with insufficient notice to the PM to arrive and/or when the birth occurs during the night. Where the PM has been unable to attend the birth at the First Affiliated Hospital of CQMU or Chongqing Health Center for Women and Children, such as an emergency delivery or the birth has occurred during the night, information pertaining to the birth will be extracted from the clinical records, along with any information about the infants (eg, body measurements and Apgar scores) within 3 days (before hospital discharge) and samples, such as infant hair will be collected. Hospital discharge

information will also be recorded which includes information about infant feeding.

Should the birth occur at a hospital other than the First Affiliated Hospital of CQMU or Chongqing Health Center for Women and Children, the family will be excluded from the study at this point due to the complete lack of clinical birth related information and samples.

Twins' information: postnatal

Mothers and twins will attend the First Affiliated Hospital of CQMU or Chongqing Health Center for Women and Children for regular visits until they reach 3 years of age (figure 2). In order to maximise retention in the study postnatally, several incentives will be offered including (1) free registration at paediatric clinics, (2) free phenotyping such as allergy screening and neurocognitive testing, (3) reimbursement of transportation costs and (4) access to 'green channel' in the hospital, abolishing waiting times.

A total of eight postnatal visits are planned, with the first commencing 6 weeks following birth (figure 2). However, the timing of the first three postnatal visits (all
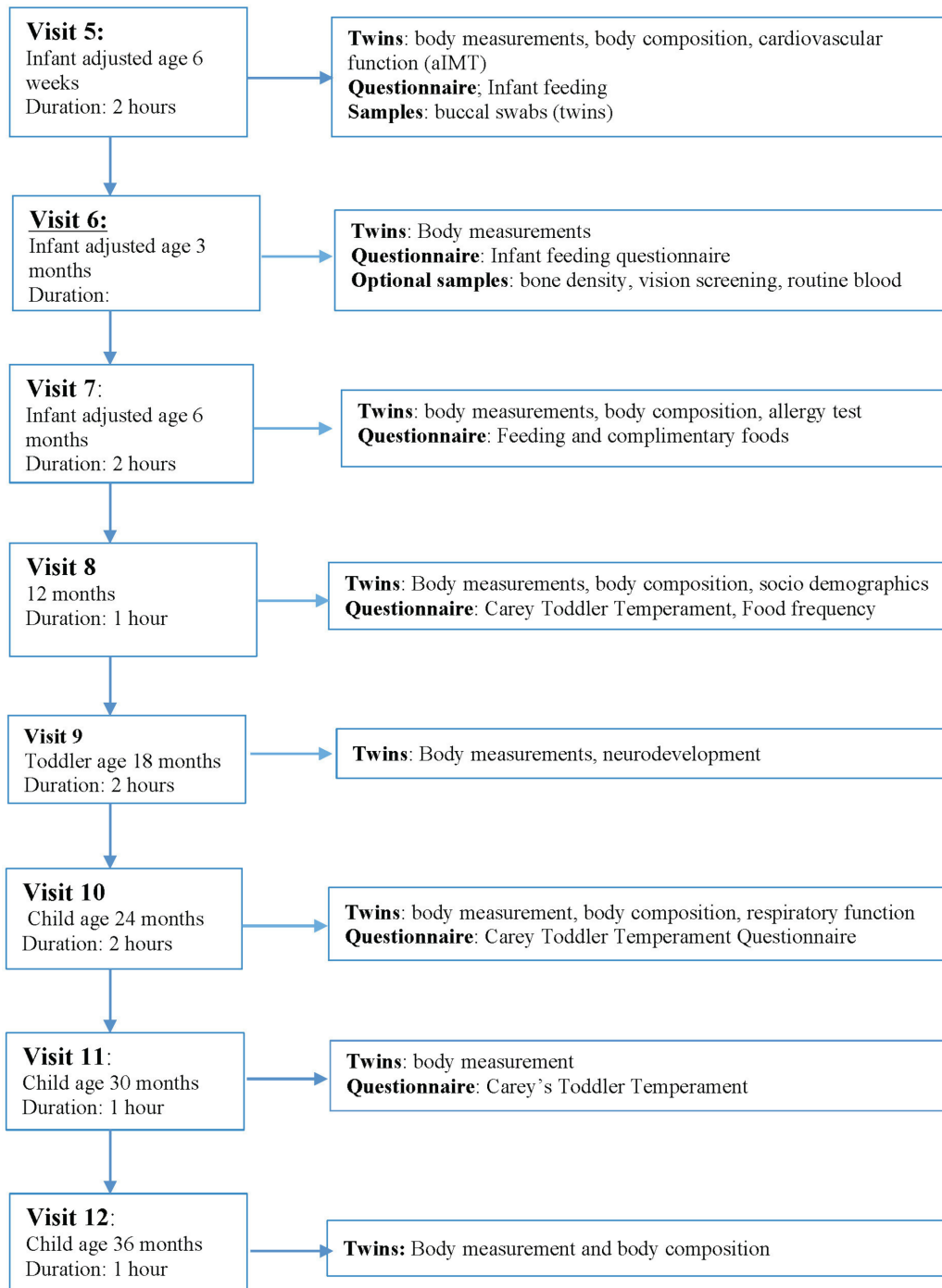

Figure 2 Flowchart: postnatal visits. aIMT, abdominal aortic intima-media thickness.

of which occur during the first 12 months) are based on a predicted gestational age of 40 weeks and adjusted accordingly. For example, infants born at 36 weeks' gestation will undertake their first postnatal visit 10 weeks following their birth, rather than the 6 weeks for a child who was born at 40 weeks' gestation. In addition to questionnaires about infant feeding and food preferences, assessments will be undertaken related to child development, temperament, sleep patterns, vision, cardiovascular function, body measurements, body composition, allergies, bone density and neurodevelopment. Samples collected will include blood and buccal swabs and, where possible, maternal urine.

QUESTIONNAIRES

All questionnaires used are from validated scales and measures. The questionnaires have been translated into Mandarin by a qualified translator. Information is collected from the mother during face-to-face interviews by a trained researcher (PM) in Mandarin and transferred to the database on the same day.

Pregnancy Stress Scale

This validated Pregnancy Stress Scale[8] is widely used to assess stress levels among Mandarin-speaking pregnant women and is a practical tool for use by nurses and healthcare workers. The questionnaire will be completed by the pregnant women at two time points: 11–16 weeks and 30–34 weeks.

There are 36 items in the scale grouped into five categories:

1. Stress regarding the pregnancy, labour and delivery (nine items). Questions include whether the labour will be difficult and whether the baby will be delivered safely.
2. Stress from baby care and changing family relationships (nine items). The types of questions asked in this section are about concerns around being able to breast feed the baby and about caring for the baby.
3. Stress from maternal role identification (eight items) ask questions about whether there are concerns about how the baby may look, what the baby may weigh and about the baby's sex.
4. Stress from social support seeking (four items). These questions are more related to whether there are concerns about access to appropriate support such as finding baby-sitters and deciding who will care for baby.
5. Stress from altered appearance and function (six items). This final section assesses maternal concerns about weight gain during pregnancy, and mobility due to altered body shape.

Gestational nutrition knowledge, attitudes and behaviour

The Gestational nutrition, knowledge, attitudes and behaviour questionnaire[9] is used to assess knowledge about basic nutrition, attitudes and behaviours. There are 20 questions in total, 10 of which are used to assess nutritional knowledge. Examples include whether they think that the more expensive and refined foods are more beneficial to health; whether they are aware of gestational BMI; which foods are the best sources of iron (with choices); best sources of calcium (with choices); best sources of vitamin C (with choices) and what kind of food deficiency in early pregnancy will increase the occurrence of neural tube defects (with choices). Examples of the five questions assessing attitudes to nutrition include whether the women will be looking for information about nutrition during pregnancy; whether their diets should be based on meals and supplemented by supplements and whether eating foods they dislike, but good for baby, will be a problem. The final section of this questionnaire asks about behaviours in relation to nutrition. There are five questions and these ask about nutritional supplements taken during pregnancy, and which of the food groups were eaten daily (with choices) and about the type and duration of exercise during the pregnancy. This information will be collected between the 11 and 16 weeks' gestation.

Brief Infant Sleep Questionnaire

Infant sleep patterns are assessed using the Brief Infant Sleep Questionnaire,[10] which is a reliable assessment tool for assessing and screening for infant sleep problems. The twins are assessed at 6 months of age and information collected include such sleep measures as nocturnal sleep duration, daytime sleep duration, the number of times the infant wakes during the night, the duration of wakefulness during the night, nocturnal sleep onset time, location of sleep and preferred body position.

Bayley III Scales of Infant Development—Chinese Revision

The Bayley Scales of Infant Development—Chinese Revision[11] is a standardised assessment of motor (fine and gross), language (receptive and expressive) and cognitive development of infants, toddlers and preschool children (ages 2–30 months). It is administered by a trained professional and consists of a series of a range of play tasks. This assessment is undertaken at 18 months of age.

Carey Toddler Temperament Questionnaire

The Carey Toddler Temperament Questionnaire[12] is used to determine the sensitivity of the child to external stimulus and the intensity of their response. A range of questions are asked about emotional reactions to food, animals, strangers, concentration levels, ability to socialise and interact with others, mood swings, ability to cooperate with others and share toys. This assessment is conducted at 30 months of age.

Food Frequency Questionnaire (FFQ)

FFQ is used to assess the dietary intake, nutritional status and dietary patterns in relation to pregnancy outcomes. Information about the frequency and amount of food items consumed in the previous 3 months is recorded in face-to-face interviews by a trained researcher at two time

points: 11–16 weeks' and 30–34 weeks' gestation. Ninety-five questions relate to the consumption of breads, cereals, noodle meals, vegetables, bean curd, fruits, poultry, meat dishes, biscuits, pastries, cakes, snacks, milk and dairy foods. Fifteen questions relate to dietary behaviours.

Twenty-four-hour food recall questionnaire

The 24-hour food recall questionnaire[13 14] is used to record detailed food intake within the 24-hour period up to the most recent meal. This questionnaire is designed to identify dietary patterns and related behaviours. A trained researcher, equipped with standardised utensils and pictures to guide the women in relation to portion sizes, conducts the questionnaire in a face-to-face interview. The women are asked to recall information about their consumption of food. Questions relate to the name and amount of the different foods consumed, and also the place where each was consumed. This assessment is conducted along with FFQ.

Nutrition Questionnaire for Children

The Nutrition Questionnaire for Children[15] consists of three parts. The first is feeding survey with a total of 25 questions. Examples include the following: What is the feeding pattern of your child in the first 6 months? What is the time of the introduction of solids? Is food specifically prepared for your child? How many times does your child feed each day? The second part is a 24-hour food recall questionnaire, and the third part is a Food Frequency Questionnaire which has been specifically designed for LoTiS infants. This assessment is to be conducted separately for all twins at 12 months of age.

Medication use

We are collecting all information of medication use around the time of pregnancy along with a history of diseases at the time of recruitment. Medication prescription during pregnancy is also being recorded.

SAMPLES
Maternal blood

Peripheral blood will be collected at each of the three prenatal visits by venepuncture for assessment of maternal complex lipids, micronutrients and endocrine status, collection of live lymphocytes and for DNA/RNA extraction. A total of 15 mL of blood will be collected at each visit: 6 mL in a 10 mL coagulant tube (for serum) and 6 mL in a 10 mL EDTA anticoagulant tube (plasma, leucocytes and red cells), with the remaining 3 mL in a 5 mL heparin anticoagulant tube (for plasma). Samples will be transported to the laboratory within 2 hours of collection for processing and storage as serum, plasma, red cells and leucocytes.

Cord blood and placental tissues

Umbilical cord blood will be collected from the umbilical vein into a 10 mL EDTA anticoagulant tube, a 5 mL heparin anticoagulant tube and a 10 mL procoagulant tube. A 4 mm skin biopsy punch will be used to randomly sample 20 pieces (2 mm×2 mm×4 mm per piece) of placenta tissue from each of the fetal and maternal surfaces. The samples will be immediately placed in a precooled (4°C) physiological saline solution for transportation to the laboratory for cold storage at −80°C within 30 min of collection.

Oral glucose tolerance test

An oral glucose tolerance test will be undertaken at the second prenatal visit when the women are between 22 and 25 weeks' pregnant to test for gestational diabetes. Measures recorded are fasting glucose levels, 60 min postprandial blood glucose level and 120 min postprandial blood glucose level.

Urine samples

A midstream urine sample is to be collected at each maternal prenatal visit to provide information about environmental exposures (such as bisphenol A). Cleaning wipes will be used to clean vulva before collection. A 15 mL sterile centrifuge tube will be used to collect 10 mL of urine. Following transportation to the laboratory, the samples will be centrifuged at 3000 rpm for 10 min and then aliquoted into 250 µL QR code tubes, and stored in −80°C within 30 min of processing.

Hair

Maternal hair will be collected for metabolomic and toxin measures at the first prenatal visit to assess both acute and chronic exposure to a range of minerals and toxins. The sample comprise approximately 30 hairs cut close to the scalp. Following the twins' births, approximately three tufts of infant hair will also be cut close to the scalp and stored for future analysis. Foil paper will be used to wrap around hairs at the time of collection with labelling to identify the scalp proximal end. The samples will be transported in cold storage to the laboratory within 2 hours of collection with long-term storage at −20°C.

Nails

Maternal fingernail samples are to be collected in the first and third prenatal visit for measure of chronic environmental exposures (such as heavy metals). Following collection, nails will be immediately wrapped in foil and then transported to the laboratory within 2 hours of collection for immediate storage at −20°C.

Buccal swabs

Buccal swabs will be used to collect cheek cells at all three maternal prenatal visits. DNA will be extracted from the swabs for later genetic and/or epigenetic testing. At the age of 6 weeks, cheek cells will be collected in the same manner from each twin. Prior to collection, participants will be asked to refrain from consuming food and liquids for 30 min. To collect the cells, a sterile oral swab will be used to rub the oral mucosa of each cheek about 10 times. The swabs will then be immediately stored in sterile

tubes, and conserved in a −80°C freezer. The DNA will be extracted from the swabs within a month of sample collection and again stored at −80°C.

ASSESSMENTS
Body measurements
Maternal body measurements will be performed at all three prenatal visits to assess weight gain and fetal development during pregnancy. For the first visit, height, weight and blood pressure will be recorded. For subsequent visits, weight, blood pressure, fundal height, abdominal circumference and fetal heart rate will be recorded. Infant body measurements are to be taken at birth and repeated at each of the eight follow-up visits by trained paediatric nurses to assess growth and development. A stadiometer (Beideneng, Shanghai, China) will be used to assess maternal height and weight, and corresponding infant information will be recorded on a growth curve. A standard measuring tape will be used to determine head circumference, chest circumference, abdominal circumference, and a digital body fat calliper will be used to determine skin fold thickness.

Body composition
Information about the twins' body composition will be collected at birth, and infant ages 6 weeks, 6 months, and 1, 2 and 3 years. This information will be used in conjunction with body measurements to gain insights into body composition. Parameters of total body water, fat mass, fat free mass, extracellular fluid, intracellular fluid and body mass index are collected by body composition analyser (Imp SFB7, Australia).

Delivery record
Information from the delivery record will be collected. This will include details about the method of delivery (caesarean or vaginal) and any preoperative and postoperative diagnosis that was made along with the size and weight of the placenta/s, the length of the umbilical cords, amniotic fluid volume and character. The infants' gestational age at delivery, their sex, weight, length and Apgar scores will also be recorded.

Respiratory function
Respiratory function will be measured by a trained professional when the infants are 2 years of age using tidal breathing analysis. A pulmonary function test apparatus (CareFusion, Germany) will be used to determine lung function parameters of respiratory rate, tidal volume, inspiratory time, expiratory time, time to reach peak tidal expiratory flow, expired volume at peak tidal expiratory flow, the ratio of inspiratory time to expiratory time and the ratio of time to reach peak tidal expiratory flow to total expiratory volume.

Cardiovascular function
When the twins are 6 weeks of age, an experienced sonographer will measure their cardiovascular function via transabdominal ultrasound. Both aortic intima media thickness and aortic diameter will be obtained using a commercially available system (Voluson E8, GE Healthcare, Austria) which is equipped with a C4-8 transducer (frequency 4–8 MHz).

Liver and kidney functions
Maternal liver and kidney functions are assessed during early (11–16 weeks) and late (30–34 weeks) pregnancy. Serum is collected from the women following an 8-hour period of fasting. The sera are examined at the clinical laboratories of the First Affiliated Hospital of CQMU or Chongqing Health Center for Women and Children. Measures will include total protein, albumin, globulin, total bile acid, aspartate aminotransferase, alanine aminotransferase, urea and creatinine.

Ultrasound scans
Fetal ultrasound scans will be obtained at each prenatal visit according to standard guidelines. A trained medical sonographer using a commercially available system (Voluson E8, GE Healthcare) equipped with a transducer (frequency 5–10 MHz) will conduct this assessment. Measures taken will include estimated fetal weight, biparietal diameter, head circumference, abdominal circumference and umbilical insertion point. Before delivery, a bedside ultrasound will be performed to confirm the final intrauterine position of twin 1 and twin 2.

SAMPLE STORAGE AND BIOBANK
All samples except hair and nails will be saved in QR code preprinted 500 µL cyrovials (Micronic, The Netherlands). The hair and nail samples will be wrapped in foil at point of collection, and saved in resealable bar-coded bags. Prior to storage in a biobank facility, each sample will be scanned by a Tracxer code reader (Micronics) enabling the sample and its location to be tracked in the database (Medscinet AB, Sweden).

Hair and nail samples will be stored in −20°C freezers (Zhongke MeiLing, China). Placental tissues will be preserved in liquid nitrogen tanks (Locator 4 Plus). All other samples (urine, serum, DNA) will be stored in −80°C freezers (907, Thermo Fisher Scientific, American).

The −80°C freezers (907, Thermo Fisher Scientific, American) have temperature sensors (PT100) and a monitoring system (Smart-Vue, Thermo Fisher Scientific, France) as well as a backup liquid nitrogen cooling system (DPL452, Taylor Wharton, USA).

DATABASE
The database used to store all information was designed and implemented by Medscinet AB. This company has been used to build similar data collection and storage bases for other population-based studies. Data will be stored in both Mandarin and English languages for extraction or annotation in either language.

DATA ANALYSIS

The extensive longitudinal data and sampling of biospecimens planned in Chongqing LoTiS affords considerable scope for addressing a range of research questions, from observational mapping of growth trajectories through epidemiological investigation of a range of maternal factors (between pairs), or non-shared environmental factors (within pairs) on offspring outcome. A key part of future studies will be the generation of a range of '-omics' type data, particularly genomics, epigenomics, transcriptomics and metabolomics. Generation and analysis of this will initially be carried out using standard approaches applicable to each type of data to yield a 'clean' dataset for each individual. Subsequent standard twin modelling approaches can then be used to partition the variation of specific phenotypic traits or '-omics' outputs into constituent variance components of additive genetic, non-additive genetic, shared environmental (eg, maternal factors) and non-shared environmental influences (eg, placental features). Between pair comparisons of similarity measures (eg, intraclass correlation coefficient) are a simple way to assess the likelihood of a genetic influence on any specific phenotype. Additional logistic regression approaches, incorporating twin relatedness into each model, will also be employed as required.

ETHICS AND DISSEMINATION

Data from the study will be published in medical research journals, and any molecular data will be published in appropriate public repositories.

DISCUSSION

Numerous studies have revealed that genetically identical mice reared in the same environmental conditions show normally distributed variation in most phenotypes, most likely driven in part by non-genetic processes.[16] Longitudinal measurement of such mechanisms in twins can provide a unique opportunity to identify genes sensitive to the environmental factors that are associated with complex disease.

LoTiS is specifically designed to combine the power of twin study design with longitudinal sampling of exposure, biospecimens and phenotypic outcomes predictive of later NCD, commencing prior to birth. Being based in China offers considerable importance both in terms of providing a non-European-based genetic profile, but also to address the unique nature of exposure/outcomes inherent to developing economies undergoing rapid Westernisation in terms of exposures and incidence of a range of NCD.

Whereas a comparison between DZ and MZ twins will reveal the relative contribution of genetic and environmental variation to early-life phenotypes, the capacity to control for genetic variation using within MZ twin pair comparisons will prove particularly valuable in helping to unravel the complexity of gene–environment–epigenetic interactions and their contribution to human health and disease.

Author affiliations
[1]Department of Obstetrics, The First Affiliated Hospital of Chongqing Medical University, Chongqing, China
[2]Canada-China-New Zealand Joint Laboratory of Maternal and Fetal Medicine, Chongqing Medical University, Chongqing, China
[3]Department of Occupational and Environmental Hygiene, School of Public Health and Management, Chongqing Medical University, Chongqing, China
[4]Cancer, Disease and Developmental Epigenetics, Murdoch Children's Research Institute, Royal Children's Hospital, Melbourne, Victoria, Australia
[5]Early Life Epigenetics, Murdoch Children's Research Institute, Royal Children's Hospital, Melbourne, Victoria, Australia
[6]Department of Paediatrics, University of Melbourne, Melbourne, Victoria, Australia
[7]Department of Obstetrics, Chongqing Health Centre for Women and Children, Chongqing, China
[8]Department of Pediatrics, Chongqing Health Centre for Women and Children, Chongqing, China
[9]Liggins Institute, University of Auckland, Auckland, New Zealand
[10]College of Medicine, Biological Sciences and Psychology, University of Leicester, Leicester, UK

Competing interests None declared.

Patient consent Parental/guardian consent obtained.

Ethics approval This study has been approved by the Ethics Committee of Chongqing Medical University (record no: 201530) and has been registered with the Chinese Clinical Trial Registry (registry no ChiCTR-OOC-16008203).

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. Alwan A, Maclean DR, Riley LM, et al. Monitoring and surveillance of chronic non-communicable diseases: progress and capacity in high-burden countries. *Lancet* 2010;376:1861–8.
2. Schmidt MI, Duncan BB, Azevedo e Silva G, et al. Chronic non-communicable diseases in Brazil: burden and current challenges. *Lancet* 2011;377:1949–61.
3. Pan B, Chen X, Wu X, et al. Prevalence of noncommunicable diseases and their risk factors in Guangzhou, China. *Prev Chronic Dis* 2014;11:E49.

4. Barker DJ, Bull AR, Osmond C, *et al*. Fetal and placental size and risk of hypertension in adult life. *BMJ* 1990;301:259–62.

5. Gluckman PD. Epigenetics and metabolism in 2011: epigenetics, the life-course and metabolic disease. *Nat Rev Endocrinol* 2011;8:74–6.

6. Loke YJ, Galati JC, Morley R, *et al*. Association of maternal and nutrient supply line factors with DNA methylation at the imprinted IGF2/H19 locus in multiple tissues of newborn twins. *Epigenetics* 2013;8:1069–79.

7. Saffery R, Morley R, Carlin JB, *et al*. Cohort profile: The peri/post-natal epigenetic twins study. *Int J Epidemiol* 2012;41:55–61.

8. Chen CH. Revision and validation of a scale to assess pregnancy stress. *J Nurs Res* 2015;23:25–32.

9. Qian-qian H, Cong L, W-l W. Investigation on nutritional knowledge, attitude and behavior of middle and late pregnant women. *Acta Universitatis Medicinalis Anhui* 2010;46:481–4.

10. Sadeh A. A brief screening questionnaire for infant sleep problems: validation and findings for an Internet sample. *Pediatrics* 2004;113:e570–7.

11. Yi S-r, Luo X-r, Yang Z-w, *et al*. Bayley scales of infant development-Chinese revision[J]. *Chin J Clin Psychol* 1993;1:71–5.

12. Xing S, Zhou Q, Archer M, *et al*. Infant temperamental reactivity, maternal and grandparental sensitivity: Differential susceptibility for behavior problems in China. *Early Hum Dev* 2016;101:99–105.

13. Chong MF, Chia AR, Colega M, *et al*. Maternal Protein Intake during Pregnancy Is Not Associated with Offspring Birth Weight in a Multiethnic Asian Population. *J Nutr* 2015;145:1303–10.

14. Chong MF-F, Chia A-R, Colega M, *et al*. Maternal Protein Intake during Pregnancy Is Not Associated with Offspring Birth Weight in a Multiethnic Asian Population. *J Nutr* 2015;145:1303–10.

15. E-h Ma. *Correlation analysis between nutrition and intelligence development of children aged 2-4 in Huining*: Lanzhou University, 2009.

16. Gärtner K. A third component causing random variability beside environment and genotype. A reason for the limited success of a 30 year long effort to standardize laboratory animals? *Lab Anim* 1990;24:71–7.