Original Study

Boston Birth Cohort profile: rationale and study design

Colleen Pearson[1], Tami Bartell[2], Guoying Wang[3], Xiumei Hong[3], Serena A. Rusk[3,*] [iD], LingLing Fu[1], Sandra Cerda[4], Blandine Bustamante-Helfrich[5], Wendy Kuohung[6], Christina Yarrington[6], William G. Adams[1], Xiaobin Wang[3,7]

[1] Department of Pediatrics, Boston University School of Medicine and Boston Medical Center, Boston, MA, USA, [2] Patrick M. Magoon Institute for Healthy Communities, Ann & Robert H. Lurie Children's Hospital of Chicago, Chicago, IL, USA, [3] Center on the Early Life Origins of Disease, Department of Population, Family and Reproductive Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA, [4] Department of Pathology, Boston University School of Medicine and Boston Medical Center, Boston, MA, USA, [5] University of the Incarnate Word School of Osteopathic Medicine, San Antonio, TX, USA, [6] Department of Obstetrics and Gynecology, Boston University School of Medicine and Boston Medical Center, Boston, MA, USA, [7] Department of Pediatrics, Johns Hopkins School of Medicine, Baltimore, MD, USA.

* Corresponding author. Address: Xiaobin Wang, Center on the Early Life Origins of Disease, Johns Hopkins University Bloomberg School of Public Health, 615 N. Wolfe Street, E4132, Baltimore, MD 21205. Tel: 410-955-5824. E-mail address: xwang82@jhu.edu

Abstract

In 1998, the Boston Birth Cohort (BBC) was initiated at Boston Medical Center (BMC) in response to persistently high rates of preterm birth (PTB, defined as birth before 37 weeks of gestation) in the US population and the longstanding profound PTB disparity among Black, Indigenous, and people of color (BIPOC). The BBC encompasses two linked study protocols: The PTB Study serves as the baseline recruitment in the BBC. It aims to address fundamental questions about the causes and consequences of PTB. The study oversamples preterm babies using a case/control study design, in which cases are defined as mothers who deliver a preterm and/or low birthweight baby (<2500 grams regardless of gestational age). Controls are enrolled at a 2:1 control/case ratio and matched by maternal age (±5 years), self-reported race and ethnicity, and date of delivery (± 7 days for case delivery). From inception, it was designed as a comprehensive gene-environmental study of PTB. As a natural extension, the Children's Health Study, under a separate but linked Institutional Review Board protocol, is a longitudinal follow-up study of the participants who were recruited at birth in the PTB Study and who continue pediatric care at BMC. This linked model allows for investigation of early life origins of pediatric and chronic disease in a prospective cohort design. The BBC is one of the largest and longest National Institutes of Health–funded prospective birth cohort studies in the United States, consisting of 8733 mother-child dyads enrolled in the PTB Study at birth, and of those, 3592 children have been enrolled in the Children's Health Study, with a median follow-up of 14.5 years. The BBC mirrors the urban, underresourced, and underrepresented BIPOC population served by BMC. A high proportion of BBC children were born prematurely and had chronic health conditions (e.g., asthma, obesity, and elevated blood pressure) in childhood. The BBC's long-term goal has been to build a large, comprehensive database (epidemiological, clinical, and multiomics) and biospecimen repository to elucidate early life origins of pediatric and chronic diseases and identify modifiable upstream factors (e.g., psychosocial, environmental, and nutritional) to improve health across the life course for BIPOC mothers and children.

Keywords: Boston Birth Cohort, Longitudinal study, Genetics, Environment, Early life origins of chronic diseases

Introduction

This article focuses on the rationale, study design, and scope of data collection of the Boston Birth Cohort (BBC). In a subsequent article, we will highlight major scientific discoveries in the BBC as the result of the past 20+ years of research efforts across multiple organs/systems and scientific disciplines.

Why was the BBC cohort established?

The BBC was initiated in 1998 at Boston Medical Center (BMC) in response to persistently high rates of preterm birth

(PTB; defined as birth before 37 weeks of gestation) in the overall US population and the longstanding profound PTB disparity among Black, Indigenous, and people of color (BIPOC).

PTB is a major cause of neonatal and infant morbidity and mortality with lifetime health consequences for children both in the US and worldwide. In 1998, when the BBC was initiated, the rate of PTB in the US rose to 11.6%, following a rise from 11.0% to 11.4% from 1996 to 1997.[1] In each year since, nearly a half-million babies have been born preterm.[2] Despite intensive research and intervention efforts, the PTB rate has remained high and, at present, affects one in 10 babies in the US and one in seven BIPOC women.[3] Worldwide, 11% of infants are born preterm, and 1 million infants die annually due to PTB.[1,4]

PTB represents a longstanding and profound health disparity in the US. In 1998 and still today, 20 years later, the rate of PTB among BIPOC women (14%) has been about 50% higher than

the rate among White women (9%).[5] In 2021, for the first time in 6 years, the US PTB rate declined slightly from 10.2% in 2019 to 10.1% in 2020, except among BIPOC mothers whose rate increased from 14.25% to 14.36%.[6]

One major obstacle in preventing PTB is our incomplete understanding of its causes and underlying mechanisms. This knowledge gap has been partly explained by the fact that PTB is a complex trait determined by multiple environmental and genetic factors,[7] and because of the significant heterogeneity and complexity of biological pathways that could lead to PTB.[8,9]

Early studies on PTB were largely focused on identifying socio-demographic, environmental, and clinical risk factors, of which a number have been recognized, including maternal race and ethnicity,[10–12] age,[13–15] education,[16,17] income,[18,19] nativity,[20] malnutrition,[21,22] stress,[23–25] social support,[26,27] smoking,[28–30] and air pollution.[31–34] A history of prior PTB[35,36] and a woman having been born prematurely herself appear to be the strongest risk factors identified to date.[37] An early study showed that the combination of these two factors significantly increased the risk of PTB or low birthweight (LBW).[38] However, these risk factors still only explain a small fraction of PTB in the overall population and do not appear to drive the PTB disparity among BIPOC women.

While the importance of genetic susceptibility in PTB has been recognized and multiple genetic variants have been linked to risk of PTB, neither candidate genes nor genomewide association studies (GWAS) have been able to pinpoint genetic loci that could help to explain a substantial portion of PTB.[39–41] However, lack of representation of BIPOC in genetic studies can lead to misclassification of benign *vs.* pathologic genetic variants and create or exacerbate health disparities in historically underserved populations.[42] From a health equity perspective, even though BIPOC women have the highest rates of PTB, the majority of US birth cohorts that have included GWAS have primarily been conducted in White European populations.[43–45] Meanwhile, the realization that traditional genetic factors cannot explain the development of heterogeneous, multifactorial health outcomes like PTB has shifted interest toward gene-environment interactions (G×E). Despite their promise, there is a lack of adequately powered studies on the role of G×E interactions in PTB among all populations. Studies conducted in the BBC have provided strong evidence to support this direction.[12,38,46–50]

A more profound realization is that the lack of birth cohort studies conducted among historically underrepresented and understudied BIPOC may largely explain our limited understanding of the causes and health disparity of PTB. The BBC's focus on BIPOC women and children and the interconnections between environmental exposures and genetics are crucial steps toward understanding the etiology of PTB and addressing the persistent disparity of PTB.

What did the BBC initially aim to investigate?

As illustrated in Figure 1, from when it was first conceptualized in 1998, the BBC was designed to address some fundamental questions about the causes of PTB. By considering and collecting data on both genetic and environmental risk factors (including those now also referred to as social determinants of health [SDOH]), the BBC planned and allowed for the analyses of the impact of multilevel, multidimensional biopsychosocial-environmental risk factors and their complex interplay on PTB.

Health disparities framework underlying research in the BBC

In 2015, the BBC adapted the research framework of the National Institute on Minority Health and Health Disparities [51] to more thoroughly address the complex and multifaceted nature of minority health and health disparities in BIPOC through the examination of contributions across different domains of influence within a life course perspective (biological, behavioral, nutritional, environmental, and healthcare) and at different levels of influence (individual, family/interpersonal, and community/societal) within those domains. Table 1 highlights the domains examined in the BBC, which have been updated, expanded, and adjusted over time.

What are the long-term goals of the BBC?

A long-term goal for the BBC has been to build an adequately powered, comprehensive database, and biospecimen
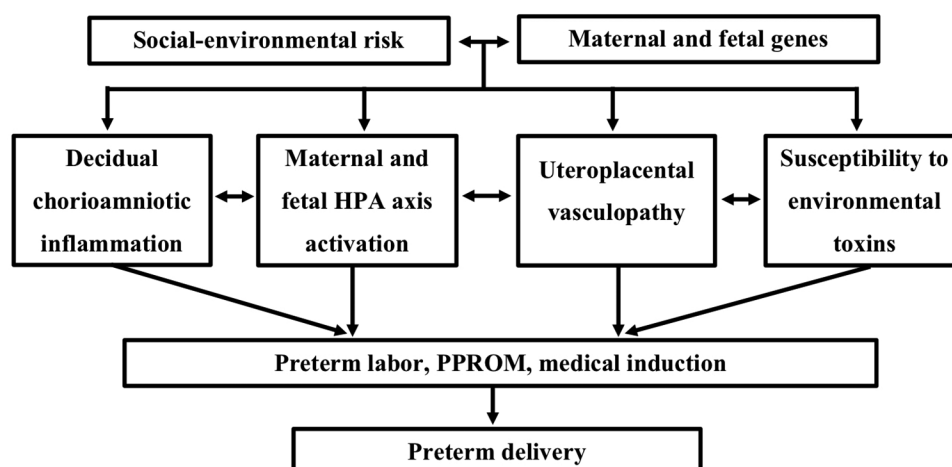

Figure 1. Early conceptual framework for studying preterm birth pathways in the BBC. Adapted from Wang et al [48]. BBC, Boston Birth Cohort; HPA, hypothalamus pituitary adrenal gland; PROM, premature rupture of membranes.

Table 1

Overview of the scope of data collection in the Boston Birth Cohort.

Domains of influence	Levels of influence		
	Individual	Family/interpersonal	Community/society
Biological	Genomewide genotypes; metabolome, RNA, biomarkers of folate, B12, vitamin D	Maternal and child health	Immunization, prenatal screening (AFP, hCG, and estriol)
Behavioral	Smoking, alcohol, opioids, marijuana, cocaine, stress, and mental health	Marital status, household composition, support of partner, family, and friends; smokers in the household; and mental health and sleep/wake patterns	Witness of violence, rate of crime, smoking, alcohol, drug use, and incarceration
Nutritional	Dietary patterns, beverage consumption, receipt of WIC, and food stamps	Cultural food preference	Grocery stores, farmer's market access, fast food restaurants, health statistics by zip codes, rates of obesity and diabetes
Environment	Exposure to ambient and indoor air pollution, fuel type, old housing, indoor pests, household pets, carpet, and molds	Family income, country of origin, immigration, ability to speak English, receipt of public assistance, WIC, employment, and access to technology	Disinvested and marginalized neighborhoods, race and ethnicity (self-reported), green space, built environment, public schools, parks, highways, transportation availability, and neighborhood crime rates
Health care	Prenatal and delivery care; lab tests, BMI, BP, diagnoses, medications, and services	Type of insurance, type and frequency of clinical visits, age, parity, and child with special health care needs	Type and completeness of recommended preventive care, and zip code level morbidity and mortality
Sources of data	Questionnaires, medical records, study visits, lab tests, and biosample/biomarkers	Questionnaires, medical records, lab tests, study visits, and biosample/biomarkers	Census data, vital statistics, police crime reports, GPS, and air monitors

WIC, women, infant and children's supplemental nutrition program; AFP, alpha-fetoprotein; hCG, human chorionic gonadotropin; GPS, global positioning system; RNA, ribonucleic acid; BMI, body mass index; BP blood pressure.

repository that allows for investigation of the causes and long-term health consequences of PTB, and for elucidation of early life origins of pediatric and chronic diseases and the health disparity among BIPOC mother-child dyads enrolled in the BBC. Ultimately, the findings from the BBC will inform clinical and public health practice and policies, including evidence-based primary prevention of PTB and early interventions for children born preterm to mitigate adverse long-term health consequences.

How was the BBC designed?

As our understanding of PTB and its impact on life-long health evolved along with the emergence of cutting-edge biomedical advances in science and technology, so did our vision for the cohort. As a result, the BBC encompasses two linked study protocols: the PTB Study and the Children's Health Study, as illustrated in Figure 2.

The Preterm Birth Study Study serves as the first time point of recruitment and represents baseline recruitment in the BBC. We use a case/control study design, in which cases are defined as mothers who deliver a preterm (<37 weeks gestation) and/or LBW baby (<2500 grams regardless of gestational age). Controls are enrolled at a 2:1 control/case ratio and matched by maternal age (±5 years), self-reported race and ethnicity, and date of delivery (± 7 days for case delivery).

The PTB Study was designed to perform a comprehensive genetic-epidemiologic analysis of PTB and to dissect biological pathways leading to PTB by examining the role of maternal versus fetal genes in each pathway to increase our understanding of genetic and environmental determinants of PTB.[46,48–50,52–57] Over time, we have expanded our research from maternal/fetal genomics to epigenomics[58–60] and metabolomics,[61,62] and have identified deoxyribonucleic acid methylation alterations and metabolites associated with PTB.

The Children's Health Study

As a natural evolution of the BBC PTB Study, the Children's Health Study, initiated under a separate but linked Institutional Review Board (IRB) protocol, aims to conduct a longitudinal follow-up study of the mother-child pairs who were recruited at birth in the PTB Study and who continue pediatric care at BMC. Early studies produced evidence that while PTB manifests at birth, its effects continue across the life course.[63–66] These findings are in line with the theory developed by David Barker that chronic degenerative conditions of adulthood may be triggered by circumstances decades earlier including PTB and
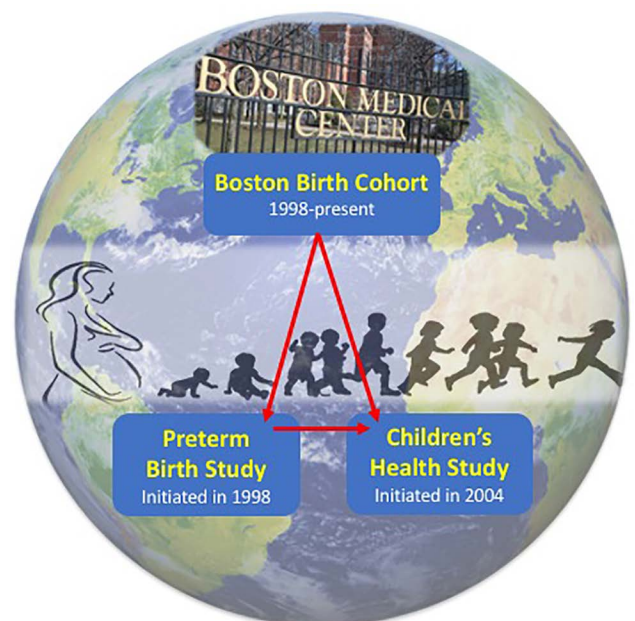

Figure 2. Illustration of the Boston Birth Cohort with two linked IRB-approved protocols at the Boston Medical Center. IRB, Institutional Review Board.

LBW.[67,68] That theory, now named the Developmental Origins of Health and Disease [69] and otherwise known as early life epigenetic programming, points to exposures in utero or during early childhood that program dysfunctions that appear later in life.[70–73] These observations motivated further expansion of the BBC in 2004 to conduct postnatal follow-up of the BBC study population as the Children's Health Study. This allows for investigation of early life origins of pediatric and chronic disease in a prospective cohort design.[63–65,74–89]

Who are the BBC participants?

All participants in the BBC are recruited at BMC in Boston, MA, the largest safety-net hospital in New England. BMC serves approximately 65% of Boston residents who are concentrated in neighborhoods with the highest level of health disparities. More than half of BMC patients reside in disinvested, economically depleted neighborhoods and receive health care coverage through government payers such as Medicaid, the Health Safety Net, and Medicare.[90]

What are the study procedures of the BBC?

Recruitment to the Preterm Birth Study

We identify potential participants for the PTB Study by tracking daily births and admissions to the postpartum floor at BMC. The inclusion and exclusion criteria include the following: any woman admitted to the postpartum floor who delivers a singleton live infant and meets our case (gestational age <37 weeks or lowbirthweight if ≤ 2500 grams and non lowbirthweight is >2500 grams) or control (full-term ≥37 weeks with birthweight ≥2500 grams) definition. A small percentage of women (<1%) with pregnancies that involve in vitro fertilization, multiple gestation, chromosomal abnormalities, or major birth defects; or women who delivered as the result of major trauma are excluded from the study. Women who do not understand English, Spanish, or Haitian Creole, the three primary languages of the BMC patient population for which we have language-specific IRB-approved consent forms, are also excluded.

Multilingual and culturally diverse Research Assistants (RAs) meet with eligible mothers at the maternal bedside during her postpartum hospital admission 24 hours after a vaginal delivery and 48–72 hours after Cesarean section to assess interest in participation. This timing was chosen to allow for maternal recuperation from labor and delivery (L&D). After determining eligibility, the RA explains the study in-depth and obtains written informed consent. Once consent is obtained, the RA administers a questionnaire interview and collects a maternal blood sample. The RA also retrieves cord blood and placenta that were obtained by the L&D nurses at delivery and stored in a designated refrigerator in L&D for study collection.

Recruitment to the Children's Health Study

Beginning when a PTB Study child is at least 6 months of age, we invite the mother-child pair to expand their study participation to the Children's Health Study to be longitudinally followed from birth (through linkage) to age 21. We delay recruitment until the infant is at least 6 months old to allow the mothers time to establish pediatric care and adjust to the demands of caring for a newborn. The inclusion/exclusion criteria include the following: past participation in the PTB Study and biological mother. Biological mothers who have lost permanent legal custody of their children are excluded because they are unable to provide consent for their children.

RAs approach mothers at BMC during check-in to their own or their child's medical appointments to assess interest in participation and determine eligibility. After an in-depth discussion of the study protocol, written informed consent is obtained from the mother at recruitment and the child at age-determined intervals as required by the BMC IRB for minimal risk studies involving children. Biospecimen collection and protocol measurements also are completed at age-determined intervals. When procedural changes are made, all participants are reconsented. In 2020, the IRB eliminated child assent and replaced it with Parental Permission Consent.

How long and how often have BBC participants been followed?

Recruitment into the PTB Study is ongoing. Enrollment in the Children's Health Study is dynamic; we are both actively recruiting and prospectively following children born over the past 2 decades. Due to the rolling enrollment in the Children's Health Study, the length of follow-up varies, with a median length of 14.5 years. Many mothers who participated in the PTB Study but did not initially establish pediatric care at BMC in infancy return to BMC for care in childhood and are eligible for recruitment.

Relevant to both the PTB Study and the Children's Health Study, a growing number of children who were recruited during our earliest enrollment period are now aged 18+ years (i.e., reproductive age range). Beginning in June 2019 through an IRB amendment, we query if a newly recruited mother-infant dyad in the PTB Study is a third-generation BBC participant. These mothers may "opt-in" to allow yearly electronic health record (EHR) surveillance for herself and her newborn until age 18. We also received IRB approval to reconsent the mother and the age 18+ child to continue participation over their lifetime, with an emphasis on yearly health data collection from the EHR. We hope to follow the children into adulthood and the mothers into middle and late adulthood.

As of June 30, 2021, 8733 mother-newborn pairs have been enrolled in the PTB Study, including 4134 (47.34%) BIPOC; 2479 (28.30%) Hispanic; 1065 (12.20%) non-hispanic white; 214 (2.45%) Asian American Pacific Islander; and 841 (9.63%) who did not identify with a race or ethnicity category. Race and ethnicity were self-reported by the mothers during the questionnaire interview. To enhance statistical power to study PTB, the BBC oversampled infants who were born preterm (3106 [35.57%] of the overall BBC). Of the 8733 mother-infant dyads, 3592 children from the PTB Study have been enrolled into the Children's Health Study. Key characteristics of the mother-infant dyads recruited at birth and subsequently recruited into the Children's Health Study are presented in Table 2.

The BBC has received initial and annual continuation review and approval by the IRBs of BMC and Johns Hopkins University. Because we ask several sensitive questions during the maternal interview, in addition to standard Health Insurance Portability and Accountability Act protections, we obtained additional protections of confidentiality. Initially, we obtained

Table 2

Characteristics of mother-child dyads in the Boston Birth Cohort by the two linked study protocols: Preterm Birth Study and Children's Health Study.

	Preterm Birth Study	Children's Health Study
N*	8623 mother/infant dyads	3394 children
Maternal characteristics at delivery		
Maternal age, years	28 (23–33)	28 (23–33)
Null parity, %	43	43
Asian American Pacific Islander,%*	2.2	1.7
African American and non-Hispanic Black, %*	32.5	39
Cabo Verdean, %*	4.3	4.9
Haitian, %*	14.9	19.8
Hispanic, %*	28.3	22
Multiethnic, %*	0.6	0.3
Non-hispanic White %*	11.9	7.4
Did not identify race or ethnicity, %*	5.3	4.9
Education (college or above), %	36	36
Married, %	35	33
US born, %	38	39
High stress during pregnancy, %	19	19
Smoking during pregnancy, %	19	18
Alcohol consumption, %	9	8
Prepregnancy overweight or obesity, %	48	52
Pre- or gestational diabetes, %	11	12
Cesarean section, %	33	36
Age at menarche, years	13 (12–14)	13 (12–14)
Child characteristics		
Child age, years	–	14 (12–18)
Boy, %	50	50
Preterm birth (<37 weeks), %	27	29
Low birthweight (<2500 g), %	26	28

Data are presented as % or median (interquartile range).

*N represents the complete cleaned dataset used for current analysis and not total accrual.

*Self-reported maternal race and ethnicity were defined as a socially constructed identity, not a biological or genetic category.

a Certificate of Protection under Massachusetts General Law, C11, Section 24A: Reduction of Morbidity and Mortality within the Commonwealth; this protection has since been replaced with a Certificate of Confidentiality granted by the National Institutes of Health CC-HO-11-15. RAs emphasize these added protections to assure mothers of the high level of confidentiality protections to minimize concerns about disclosing sensitive information for research purposes.

What variables have been measured in the BBC?

Data collection in the BBC

The BBC has a rich data repository that is unprecedented in studies conducted with participants who have been historically underrepresented in research.

Our data contain key phenotypes, genetic, microbial, and biological samples collected at multiple time points from birth through early adulthood. Current data collection in the Children's Health Study focuses on the children's transition from mid-childhood to adolescence and young adulthood and on the mothers' transition to middle age. A summary of the data collection scope and timing intervals is shown below in Table 3 for the PTB Study and in Table 4 for the Children's Health Study. A complete list of questionnaires, medical record abstractions, and variable measurements is available in the Supplemental Data, http://links.lww.com/PN9/A7. **The Preterm Birth Study.** In-person maternal questionnaires administered 24–72 hours after delivery, collect baseline information on maternal environmental exposures pre- and

post-conception and the mother's previous birth outcomes. Newborn gestational age is calculated based on the last menstrual period or early ultrasound (<20 weeks gestation). Maternal health is assessed through maternal self-report and EHR abstraction before and during pregnancy for the presence of hypertensive disorders (chronic/gestational hypertension, preeclampsia, eclampsia, and hemolysis elevated liver enzymes and low platelets syndrome); diabetes (diabetes mellitus types 1&2 and gestational diabetes); intrauterine infection, genitourinary infection, past reproductive history and outcomes, and use of illicit substances (blood and urine toxicology screening during pregnancy and at delivery on both mother and neonate) including alcohol use (ever, during pregnancy) and smoking (ever and in each trimester of pregnancy). A modified food frequency questionnaire that focuses on consumption of fruits, vegetables, fiber-rich foods, dairy, pasta, breads, cereal, meat, eggs, nuts, seeds, tofu, shellfish, fish, coffee, teas, juices, and soda is administered to assess average daily intake during pregnancy. To assess stress in pregnancy, we administer the Perceived Stress Scale-10, a 10-question classic stress assessment used because of its brevity, reliability, and validity across language groups.[91] We follow the mother and infant's postpartum course through EHR review until both are discharged from hospital care. The RA abstracts clinical variables from the EHR.

The Children's Health Study. The RA administers yearly questionnaires to assess child health with an emphasis on nutrition, asthma and allergies, development, and pubertal timing. We administer three developmental screening questionnaires: the Modified Checklist for Autism in Toddlers is administered beginning when the child is age 16–30 months. The Social Responsiveness Scale (SRS)-2 (Pre-school version) is administered at age 30–47 months

Table 3

Data collection and measurements in the Preterm Birth Study on mother and newborn.

Biospecimens	Measurements
Maternal venous blood	Plasma leptin, adiponectin, plasma folate and B12, plasma metabolome, Genome-wide DNA methylation, Genome-wide genotype, RBC metal levels (Pb, Hg, Cd, Mn, Se), archived DNA, RNA, and plasma samples
Cord blood	Plasma leptin, adiponectin, plasma folate and B12, inflammatory markers, plasma metabolome, Genome-wide DNA methylation, Genome-wide genotype, plasma vitamin D, archived DNA, RNA, and plasma samples
Placenta maternal surface	Gross findings, placental weight, and histopathologic findings
Maternal oral swabs	Archived swabs for future oral microbiome
Placenta fetal surface and distal section of umbilical cord	Histopathologic findings
Newborn meconium	Archived for future gut microbiome analysis
Electronic health records	Data collection
Mother (1 year prior to pregnancy and throughout pregnancy until discharge from hospital care)	Height, weight, BMI, BP, labs, blood type and Rh, number of prenatal visits, ultrasounds with biometry, prenatal care site, clinical presentation at delivery, presence of pregnancy complications (DM, GDM, HELLP, Preeclampsia, Eclampsia, Abruption, GHTN, CHTN, placenta previa, and incompetent cervix), mode of delivery, induction of labor, gravidity, parity, number SAB's and TAB's, infertility and treatment, urine cultures and dips, presence of sexually transmitted disease, fetal fibronectin, 24-hour urine protein levels, amniocentesis, preterm labor and management, chronic health conditions, length of rupture of membranes to delivery, medications, placenta pathology reports, cigarette smoking, alcohol and illicit drug use including toxicology screening, food and drug allergies, address, zip code, and date of birth
Newborn (until discharge from hospital care)	Length, head circumference, date and time of delivery, *sex assigned at birth, labs, blood type and Rh, weight, gestational age by LMP and early US and/or Dubowitz score, New Ballard Score, Apgar scores, head ultrasounds, presence of birth defects, TORCH infections, length of hospitalization, and medical complications (NEC, IVH, and PDA)

Maternal questionnaire interview	Data collection
In-person interview at bedside during hospitalization	Fagerström Test for nicotine dependence, Perceived Stress Scale (PSS), epidemiology questionnaires; home environment, prenatal care site, number of prenatal visits attended and missed, food frequency questionnaire, maternal and paternal allergies and asthma history, medications and vitamin and herbal supplements, chronic health conditions, reproductive history, previous birth outcomes and pregnancy complications, age of menarche, smoking, alcohol and illicit drug use history and during pregnancy, maternal and paternal height and weight, maternal and maternal parents nativity, years living in the US, physical activity, employment, self-identified race and ethnicity, highest level of education, social support, income, and public assistance

*Sex assigned at birth is defined as a biologic construct.

BMI, body mass index; BP, blood pressure; Cd, cadmium; CHTN, chronic hypertension; DM, diabetes mellitus; DNA, deoxyribonucleic acid; GDM, gestational diabetes mellitus; GHTN, gestational hypertension; HELLP, Hemolysis, Elevated Liver enzymes and Low Platelets; Hg, hemoglobin; IVH, intraventricular hemorrhage; LMP, last menstrual period; Mn, manganese; NEC, necrotizing enterocolitis; Pb, lead; PDA, patent ductus arteriosus; RBC, red blood cell; RNA, ribonucleic acid; SAB, spontaneous abortion; Se, selenium; TAB, therapeutic abortion; TORCH, toxoplasmosis, others (syphilis, hepatitis B), rubella, cytomegalovirus, herpes simplex.

followed by the SRS-2 (School Age) at age ≥48 months. The Social Communication Questionnaire (SCQ) is administered at age ≥30 months. The Autism Diagnostic Observation Schedule, Second Addition (ADOS-2), is a semi-structured, standardized assessment instrument that includes a number of play-based activities and social situations and is designed to obtain information in the areas of communication, reciprocal social interactions, and restricted and repetitive behaviors associated with a diagnosis of autism spectrum disorder. The ADOS-2 is completed on children who failed the SCQ with a research score of 11 or greater and is administered and scored by an ADOS-trained Developmental and Behavioral pediatrician. A self-administered Pubertal Development questionnaire is administered at age 6 years and is conducted yearly. We measure the body composition of both mother and child using the Tanita Body Composition Analyzer BF-350 scale (Tanita Corporation of America, Arlington Heights, IL). The RA also measures maternal and child waist circumference. Waist circumference measurements are taken with a soft cloth tape measure in cm just above the hip bone in a horizontal plane around the abdomen at the level of the highest point of the iliac crest. Before reading the tape measure, the RA ensures that the tape is snug but does not compress the skin and is parallel to the floor. The measurement is made at the end of a normal expiration. Measurements are recorded at 2 separate intervals after asking the mother and/or child to take a breath and exhale. Spirometry and impulse oscillometry (IOS) and pulmonary function tests (PFTs) are performed on a subset of children beginning at age 3 and 6 years, respectively. Due to funding limitations, we eliminated PFT and IOS administration in 2015. We also performed Radioallergosorbent (RAST) testing. Diagnostic testing for specific immunoglobulin E (IgE) to eight common food allergens (milk, peanut, egg, walnut, shellfish, fish, soy, and wheat) was completed by Quest Diagnostics, a Clinical Laboratory Improvement Amendments Certified Lab (Chantilly, VA). RAST results with specific IgE >0.35 kU/L were reported to both pediatric primary care physicians and mothers and entered into the EHR. Due to reduced funding, we eliminated RAST testing in 2016. We schedule childhood study blood draws to align with physician-ordered blood draws and piggyback on our study blood collection by providing additional study tubes to the pediatric phlebotomist for research analysis. Children provide a yearly urine specimen when able. Maternal and child EHRs are reviewed each year to assess clinical variables including weight, height, blood pressure, lab work, sick visits, hospitalizations, diagnostic tests, medications, and International Classification of Disease (ICD) codes. In 2017, BMC incorporated an SDOH screening tool called THRIVE. Responses are entered in the EHR dashboard that automatically assigns ICD-10 codes. We can access responses to questions about homelessness, food and housing insecurity, inability to afford medications, lack of transportation, educational aspirations, utility bill concerns, and caregiving issues, as documented by providers during routine outpatient healthcare visits outside of the study encounters. With the integration of community health care data into Epic's EpicCare

Table 4

Data collection, measurements, and timing in the Children's Health Study*.

Type of data collection	Yearly	1–2 year	3–5 year	6–9 year	10–14 year	15–21 year
Epidemiology questionnaires; home environment	✓					
Geocodes and EPA air monitoring data	✓					
Covid-19 hardship questionnaires (2020)						
Food frequency questionnaire	✓					
Child clinical EHR data abstraction	✓					
	From birth onward					
Mother clinical EHR data abstraction	✓					
THRIVE survey (social determinants of health)	✓					
M-Chat		✓				
SCQ, SRS-2		✓	✓	✓	✓	✓
ADOS-2 (on children who failed the SCQ at 11>)			✓	✓		
Child asthma, allergy questionnaire interview	✓					
Children's Sleep Habits Questionnaire (modified)	✓					
Child pubertal staging questionnaire, self-administered				✓	✓	✓
Mother waist circumference and body composition	✓					
Child waist circumference and body composition				✓	✓	✓
Child radioallergosorbent test		✓	✓	✓	✓	✓
Child urine			✓	✓	✓	✓
Child IOS and PFTs			✓	✓	✓	✓
Child height, weight, BMI, and BP at each well-child visit	✓					
Child age of menarche, menstrual disorders in girls				✓	✓	✓
Child pubertal progression in boys and girls questionnaire, self-administered				✓	✓	✓
Physician diagnosis of pubertal disorders				✓	✓	✓
Child blood sample collection and archive		✓	✓	✓	✓	✓

*Data from the Children's Health Study is linked to data from the Preterm Birth Study (Table 3).
ADOS-2, Autism Diagnostic Observation Schedule-second addition; BMI, body mass index; BP, blood pressure; Covid-19, coronavirus disease 2019; EPA, Environmental Protection Agency; IOS, impulse oscillometry; M-Chat, Modified Checklist for Autism in Toddlers; PFTs Pulmonary Function Tests; SCQ, Social Communication Questionnaire; SRS, Social Responsiveness Scale; WC, waist circumference.

(Care Everywhere), we can access BMC-affiliated community health center healthcare utilization outside of BMC to complete collection of clinical data, thereby reducing any missing outside healthcare utilization from our rolling recruitment strategy.

Biospecimen acquisition and processing in the BBC

Trained RAs draw maternal venous blood at the bedside. We collect a maximum of 10 mL ethylenediaminetetraacetic acid maternal and cord blood. L&D nursing staff collect study cord blood at delivery from all newborns through venous umbilical cord milking. For ribonucleic acid (RNA) analysis and stabilization, we use PAXgene Blood RNA tubes and collect an additional 2.5 mL of maternal whole blood and cord blood on a subset of mothers and newborns (incubated for at least 2 hours at room temperature). Maternal venous, cord, and childhood blood is centrifuged (1430 g at 0°C for 13 minutes) and fractionated into plasma, white blood cells, and red blood cells. Fresh placentas are sectioned in 2 × 2 cm on both maternal and fetal sides containing decidua basalis tissue, membranes, and a small piece of the distal umbilical cord, and stored in –20° freezers. For any recruited participant whose placenta is examined by the hospital pathologist for gross and histopathological review, we code the pathological findings into predominantly eight categories, in line with the classification proposed by Redline.[92] Using Fisher Scientific Sterile Transport Swabs, the maternal oral cavity is swabbed in six locations (buccal, tongue dorsum and lateral, hard palate, sublingual, and gingiva). Meconium is collected from infant waste.

In childhood, we collect venous blood (K2EDTA) at five time points: age 9–12 months, 2–3 years, 4–5 years, 6–7 years, and 10–15 years. RAST blood is collected in 6-mL red-top (plain, nonserum separator) tubes. To ensure maximum efficiency of sample use, each fractionated sample is divided into small aliquots and biobanked. Children provide a yearly urine sample of at least 10 mL collected in bisphenol A-free sterile specimen containers as a sensitive and noninvasive source for measurement of a range of environmental pollutants and renal function. All biospecimens are stored in –80° freezers pending analysis.

What health conditions were examined in the BBC?

We have examined maternal and child health outcomes along the following organs/systems. This list will expand with ongoing investigations in the BBC.

- Prenatal, perinatal, and birth outcomes
- Allergy, asthma, and upper and lower airway conditions
- Cardiometabolic outcomes
- Neurodevelopmental outcomes
- Pubertal developmental outcomes

A more detailed description of phenotypes and a list of the BBC publications by organs/systems will be summarized in a subsequent publication.

Research infrastructure of the BBC

We intentionally chose BMC as a single recruitment site for a multitude of reasons. BMC is the medical home to our participants anchored within a large academic medical research center. This allows for optimization of resources and uniform and comprehensive repeated clinical data collection within a single medical system and EHR platform.

The conduct of clinical studies also requires robust research infrastructure support, which we can leverage through additional clinical resources, specifically the presence of a study office centrally located in the pediatric clinic along with a continuous EHR query system that began in 2002. Our foremost strategy has been to eliminate, as much as possible, the logistical barriers for mothers to participate in the study. We align study visits with regularly scheduled medical visits across multiple clinical outpatient departments, thereby reducing maternal and child burden by eliminating the need to travel to BMC specifically for study visits.

Along with primary care clinics, the multiple specialty care clinics and patient-focused programs at BMC increase the number and frequency of outpatient visits attended overall. This allows for multiple opportunities to engage with participants outside of yearly well-child visits and offers flexibility for the mother and child to complete the protocol across multiple encounters, further reducing their time burden while maximizing data collection. With a single recruitment site and central location for comprehensive pediatric health care services including inpatient, emergency department, primary care, and subspecialty clinics, we can be nimble and pivot our research team to a specific clinic where participants have the greatest number of appointments, thereby focusing our resources to maximize initial recruitment and engage and retain recruited participants.

We have determined that the most important retention strategy is to be flexible and cognizant of maternal time constraints to complete the study procedures in a single session. We include an option for "deferment of procedures" in the informed consent to allow participants to complete questionnaires and measurements over several encounters. The RA can assess in real-time if a specific clinical care visit has been long or if the mother is rushed and defer approaching her for study activities until the next visit. This approach accounts for our considerably small withdrawal rate and highlights the advantages of our integrated study presence in the pediatric clinic.

EHR and Clinical Data Warehouse

BMC established an EHR query system in 2002. With IRB approval and in collaboration with BMC Information Technology, we developed a weekly computerized query report using patient appointment scheduling to identify participants eligible for or recruited into the Children's Health Study. By providing the medical records of recruited PTB Study mothers and children, a query can be run each week to output those participants with a scheduled appointment in the pediatric or other ancillary clinics. The data manager further refines the appointment list against the recruitment log file to provide a weekly recruitment list. Additionally, within the EHR system, we conduct real-time monitoring for canceled appointments and, conversely, for when a participant has "checked in" for their appointment.

Clinical variables are obtained from the Clinical Data Warehouse (CDW), a repository of BMC's patient EHR data including legacy (data from various data systems that are no longer in active clinical use) and community health center data. We provide the medical record numbers of recruited participants in the PTB Study and Children's Health Study to the CDW Senior Data Manager with updates on recruitment at 6-month intervals.

Engagement with participating mothers, children, and providers

Due to concerns about low literacy among BMC mothers, most questionnaires (except for pubertal self-report unless requested by the mother) are read aloud by the RAs. This reduces time and missing data. Since our transition to the research electronic data capture (REDCap) system, we can simplify the presentation of the questionnaires by using the program's branching logic and voice recording the questions allowing a playback feature for mothers, if needed.

Our study office is equipped with a PlayStation, a television, children's movies, handheld electronics, and I-pads. In general, performs child measurements while the second RA completes maternal questionnaires and measurements. The RAs are also familiar to providers and administrative staff and able to assess wait times, allowing some study procedures to begin while a participant is waiting for their medical visit to start.

We also periodically conduct information sessions with clinical providers about the Children's Health Study's research goals, as we understand clinical staff are instrumental in their assistance as trusted sources of information about research participation.

A study information board hangs in the pediatric clinic hallway with information about the study, pictures of the Pulmonary Function testing and Bioelectric Impedance Analysis equipment along with photographs of RAs, their names, what languages they speak, and a brief biography. As a result, many study mothers and children come to the study office to ask if they are due for any measurements since they are in the clinic and waiting for or finishing their appointment. This is also helpful information for clinicians who rotate through the pediatric outpatient clinics.

We provide a small gift card incentive after protocol completion in the PTB Study and at each yearly completed study visit in the Children's Health Study. Taxi vouchers are provided if a participant arrives before their medical appointment to complete study procedures specifically for study purposes.

Impact of COVID-19 on the BBC

In March 2020, all in-person non-COVID-19 and safety-related research activities were suspended at BMC. With IRB approval, we administered questionnaires to BBC participants electronically using REDCap or over the telephone, adding two newly developed COVID-19 hardship questionnaires. The first COVID-19 questionnaire was designed to assess the acute impact experienced within the first 6 months of the stay-at-home order issued by the City of Boston on March 10, 2020. This questionnaire assessed COVID-19 positivity rates and testing rates, household composition, missed healthcare appointments, experiences of food insecurity, and significant events (job loss, school closures, death, or hospitalization of family members due to COVID-19 positivity). As COVID-19 pandemic restrictions continued in Boston, we administered a second COVID-19 questionnaire to assess the longer-term impact of the pandemic, especially regarding health care visits, school disruptions, caregiver stress, use of telemedicine, loss of employment, housing and food insecurity, and attitudes toward COVID-19 vaccines. Additionally, through weekly EHR surveillance, we queried the incident risk of COVID-19

positive mothers and children in the BBC and followed their health outcomes and symptomology.[74]

In essence, COVID-19 not only changed the way the study team interacted with the cohort but also impacted the cohort's health outcomes, which glaringly highlighted the consequences of health inequities among populations living in low-resource, marginalized communities who make up a large proportion of the essential workforce. The BBC is well-positioned to examine the intersectionality and joint impact of biological factors, SDOH, and COVID-19 on the long-term physical, social, and emotional health of the mothers and children in the BBC, which, in turn, will inform clinical and public health preparedness and policies to minimize adverse consequences due to current and future pandemics.[74]

What are the strengths and limitations of the BBC?

Unique strengths of the BBC

In addition to the common strengths associated with prospective birth cohort design, such as temporal, dose-response relationships and the ability to study comorbidities, control for time-dependent covariables, and conduct longitudinal data analyses, the BBC has the following unique strengths:

- Largest, longitudinal birth cohort study of US urban BIPOC women and their children from low-resource communities who bear a disproportionally high burden of social adversities and disparate health outcomes including PTB, but who have been severely underrepresented in PTB research.
- Participants were and continue to be recruited from a single-site medical center using uniform systems including a long-standing, single hospital EHR across time and essential clinical, laboratory, and computational infrastructures to ensure that clinical data are comprehensive and complete.
- By utilizing a wide range of informatics, a diverse study staff, and a study design that embeds study visits alongside clinical visits, through the two linked studies of the BBC, we have successfully recruited and retained participants from birth to age 21 and are approved to reconsent for lifetime follow-up.

- We have a comprehensive collection and archive of extensive epidemiological and clinical data including SDOH and maternal, placental, newborn, and childhood biospecimens.
- Application of cutting-edge biomedical advances in science and technology in the fields of genetics, epigenetics, and metabolomics while also addressing SDOH among BIPOC women and their children living in urban and low-resourced communities. While the collection of genetic, epigenetic, and metabolomic data is not unique, it is rare for such data to be obtained in an understudied and underrepresented BIPOC population. In doing so, the BBC helps to ensure that all women and their children can equally benefit from scientific and technological advancements.

Due to exponential advances made in science and technology since 1998, we now use a life-course-biopsychosocial framework,[93,94] which can for the first-time integrate key contributors of PTB including genome, epigenome, metabolome, health inequities, and environmental factors with maternal, placental, and fetal triads in a single study. This positions the BBC to apply current and future biomarkers, identified in the populations most likely to benefit from them, to prevent PTB and improve health outcomes. Figure 3 displays the current investigational model for applying the BBC data collection.

Limitations, challenges, and opportunities of the BBC

Both a strength and a limitation of the BBC are that it is conducted in a predominately BIPOC population; while findings from the BBC may not be generalizable to other populations, the findings apply to those most impacted by PTB and its long-term consequences. A potential limitation is that, although mothers have preferred having questionnaires read aloud to them, reading questions aloud may have inadvertently produced social desirability bias in their responses.

Furthermore, the unique features of the BBC stated above overcome many of the limitations of most existing birth cohort studies, which struggle with limited sample sizes, fragmented data, longitudinal retention, and lack of multilevel multidimensional
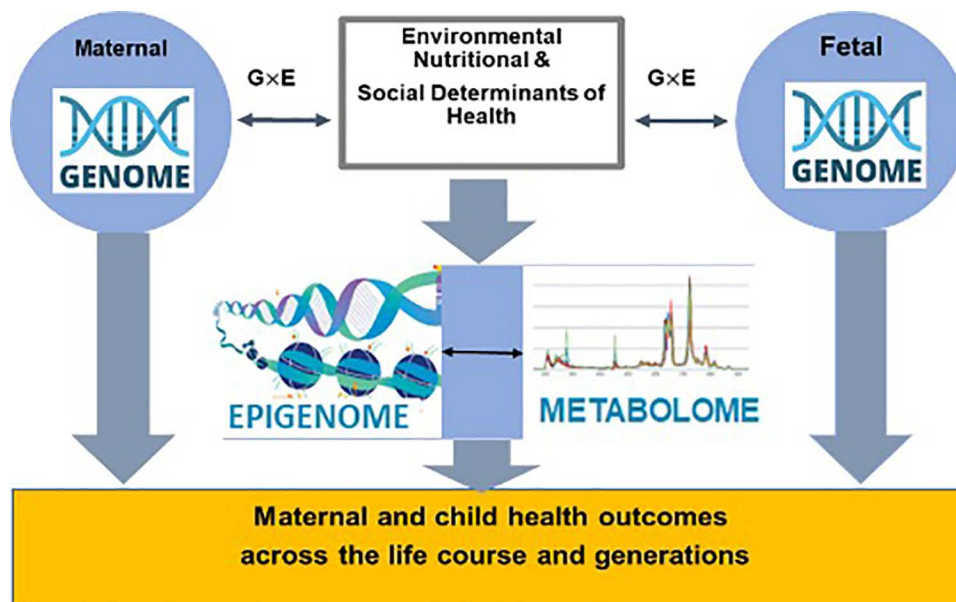

Figure 3. Updated conceptual framework for studying PTB pathways in the BBC. BBC, Boston Birth Cohort; PTB, preterm birth.

data to bridge biology with SDOH. Missing data and attrition/loss to follow-up are still threats to this and any longitudinal studies. There is also the challenge of describing such a dynamic long-term study like the BBC, largely due to the complexity that underlies its seamless operation. This includes significant ongoing IRB policy changes related to retention and tracking that have impacted the cohort in important ways.

The BBC is also uniquely positioned to address the lack of BIPOC representation in biomedical research, and even more so in research that uses advanced biotechnology such as multi-omics. The BBC, with its extensive biorepository, along with extensive and longitudinal collection of epidemiological and clinical data, offers an unprecedented opportunity to gain a deeper understanding of the underlying causes and mechanisms of the persistent disparity in PTB; to provide critically needed evidence to inform public policy, social reform, health services organization and delivery, and clinical and public health programs; and, ultimately, to improve individual and population health for BIPOC and for all society.

Acknowledgments

Ethical Conduct of Research

The Institutional Review Boards (IRBs) of Boston Medical Center and Johns Hopkins Bloomberg School of Public Health approved both study protocols of the BBC. We continually assess RA training, informed consent administration observations, retrain periodically, and perform numerous data quality checks.

Availability of supporting data

Our goal is to make the BBC data available to investigators, while at the same time respecting the confidentiality of our participants and ensuring the integrity of the data. A subset of the data including maternal and child genotypes and phenotypes from the Preterm Birth Study is available under controlled access request at: Genome-Wide Association Study of Preterm Birth—NCBI (https://www.ncbi.nlm.nih.gov/). Investigators interested in learning more about the BBC and how to obtain BBC data can contact us at BBCcontact@bmc.org. Our team will review specific data requests during monthly team meetings. For general inquiries, we will respond as quickly as possible.

Clinicaltrials.gov registration

The BBC is registered under clinicaltrials.gov NCT03228875.

Funding

Author contributions

XW designed and initiated the BBC and has been the principal investigator of the BBC since its inception; CP, GW, XH, SR, LF, SC, BB, WK, CY, WGA, and XW conducted research; and CP, TB, and XW wrote the article. All authors have contributed to and reviewed the article and approved the final article for submission. CP, WGA, and XW had primary responsibility for final content.

Conflict of interest

None declared.

References

[1] Chawanpaiboon S, Vogel JP, Moller AB, et al. Global, regional, and national estimates of levels of preterm birth in 2014: a systematic review and modelling analysis. Lancet Glob Health 2019;7(1):e37–e46. doi:10.1016/S2214-109X(18)30451-0.

[2] March of Dimes, PMNCH, Save the Children, WHO. Born Too Soon: The Global Action Report on Preterm Birth. 2012. https://www.marchofdimes.org/materials/born-too-soon-the-global-action-report-on-preterm-.pdf. [Accessed April 20, 2022]

[3] Ventura SJ, Martin JA, Curtin SC, Mathews TJ, Park MM. Births: Final data for 1998. Natl Vital Stat Rep 2000;48(3):1–100. https://stacks.cdc.gov/view/cdc/85761.

[4] Blencowe H, Cousens S, Oestergaard MZ, et al. National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications. Lancet 2012;379(9832):2162–2172. doi:10.1016/S0140-6736(12)60820-4.

[5] Centers for Disease Control and Prevention. Preterm Birth. 2021. https://www.cdc.gov/reproductivehealth/maternalinfanthealth/pretermbirth.htm#:~:text=Preterm%20birth%20is%20when%20a,2019%20to%2010.1%25%20in%202020. [Accessed April 20, 2022].

[6] 2021 March of Dimes Report Card. 2021. https://www.marchofdimes.org/materials/March-of-Dimes-2021-Full-Report-Card.pdf

[7] Hodgson EJ, Lockwood CJ. Chapter 2 - Preterm birth: a complex disease. In: Berghella V, ed. Preterm birth: prevention and management. 2010:8–16.

[8] Goldenberg RL, Culhane JF, Iams JD, et al. Epidemiology and causes of preterm birth. Lancet 2008;371(9606):75–84. doi:10.1016/S0140-6736(08)60074-4.

[9] Brou L, Almli LM, Pearce BD, et al. Dysregulated biomarkers induce distinct pathways in preterm birth. BJOG 2012;119(4):458–473. doi:10.1111/j.1471-0528.2011.03266.x.

[10] Almeida J, Becares L, Erbetta K, et al. Racial/Ethnic inequities in low birth weight and preterm birth: the role of multiple forms of stress. Matern Child Health J 2018;22(8):1154–1163. doi:10.1007/s10995-018-2500-7.

[11] Culhane JF, Goldenberg RL. Racial disparities in preterm birth. Semin Perinatol 2011;35(4):234–239. doi:10.1053/j.semperi.2011.02.020.

[12] Tsai HJ, Yu Y, Zhang S, et al. Association of genetic ancestry with preterm delivery and related traits among African American mothers. Am J Obstet Gynecol 2009;201(1):94 e1–94 10. doi:10.1016/j.ajog.2009.02.027.

[13] Schempf AH, Branum AM, Lukacs SL, et al. Maternal age and parity-associated risks of preterm birth: differences by race/ethnicity. Paediatr Perinat Epidemiol 2007;21(1):34–43. doi:10.1111/j.1365-3016.2007.00785.x.

[14] Keiser AM, Salinas YD, DeWan AT, et al. Risks of preterm birth among non-Hispanic black and non-Hispanic white women: effect modification by maternal age. Paediatr Perinat Epidemiol 2019;33(5):346–356. doi:10.1111/ppe.12572.

[15] Aldous MB, Edmonson MB. Maternal age at first childbirth and risk of low birth weight and preterm delivery in Washington State. JAMA 1993;270(21):2574–2577. doi:10.1001/jama.1993.03510210060028.

[16] El-Sayed AM, Galea S. Temporal changes in socioeconomic influences on health: maternal education and preterm birth. Am J Public Health 2012;102(9):1715–1721. doi:10.2105/AJPH.2011.300564.

[17] Jansen PW, Tiemeier H, Jaddoe VW, et al. Explaining educational inequalities in preterm birth: the generation r study. Arch Dis Child Fetal Neonatal Ed 2009;94(1):F28–F34. doi:10.1136/adc.2007.136945.

[18] Olson ME, Diekema D, Elliott BA, et al. Impact of income and income inequality on infant health outcomes in the United States. Pediatrics 2010;126(6):1165–1173. doi:10.1542/peds.2009-3378.

[19] DeFranco EA, Lian M, Muglia LA, et al. Area-level poverty and preterm birth risk: a population-based multilevel analysis. BMC Publ Health 2008;8:316. doi:10.1186/1471-2458-8-316.

[20] Elo IT, Vang Z, Culhane JF. Variation in birth outcomes by mother's country of birth among non-Hispanic black women in the United States. Matern Child Health J 2014;18(10):2371–2381. doi:10.1007/s10995-014-1477-0.

[21] Mason JB, Saldanha LS, Ramakrishnan U, et al. Opportunities for improving maternal nutrition and birth outcomes: synthesis of country experiences. Food Nutr Bull 2012;33(2 Suppl):S104–S137. doi:10.1177/15648265120332S107.

[22] Olapeju B, Saifuddin A, Wang G, et al. Maternal postpartum plasma folate status and preterm birth in a high-risk US population. Public Health Nutr 2019;22(7):1281–1291. doi:10.1017/S1368980018003221.

[23] Dole N, Savitz DA, Hertz-Picciotto I, et al. Maternal stress and preterm birth. Am J Epidemiol 2003;157(1):14–24. doi:10.1093/aje/kwf176.

[24] Wadhwa PD, Entringer S, Buss C, et al. The contribution of maternal stress to preterm birth: issues and considerations. Clin Perinatol 2011;38(3):351–384. doi:10.1016/j.clp.2011.06.007.

[25] Rosenthal L, Lobel M. Explaining racial disparities in adverse birth outcomes: unique sources of stress for Black American women. Soc Sci Med 2011;72(6):977–983. doi:10.1016/j.socscimed.2011.01.013.

[26] Hetherington E, Doktorchik C, Premji SS, et al. Preterm birth and social support during pregnancy: a systematic review and meta-analysis. Paediatr Perinat Epidemiol 2015;29(6):523–535. doi:10.1111/ppe.12225.

[27] Surkan PJ, Dong L, Ji Y, et al. Paternal involvement and support and risk of preterm birth: findings from the Boston birth cohort. J Psychosom Obstet Gynaecol 2019;40(1):48–56. doi:10.1080/0167482X.2017.1398725.

[28] van den Berg G, van Eijsden M, Vrijkotte TG, et al. Educational inequalities in perinatal outcomes: the mediating effect of smoking and environmental tobacco exposure. PLoS One 2012;7(5):e37002. doi:10.1371/journal.pone.0037002.

[29] Kyrklund-Blomberg NB, Granath F, Cnattingius S. Maternal smoking and causes of very preterm birth. Acta Obstet Gynecol Scand 2005;84(6):572–577. doi:10.1111/j.0001-6349.2005.00848.x.

[30] Soneji S, Beltran-Sanchez H. Association of maternal cigarette smoking and smoking cessation with preterm birth. JAMA Netw Open 2019;2(4):e192514. doi:10.1001/jamanetworkopen.2019.2514.

[31] Li Q, Wang YY, Guo Y, et al. Effect of airborne particulate matter of 2.5 μm or less on preterm birth: a national birth cohort study in China. Environ Int 2018;121(Pt 2):1128–1136. doi:10.1016/j.envint.2018.10.025.

[32] Wang YY, Li Q, Guo Y, et al. Association of long-term exposure to airborne particulate matter of 1 μm or less with preterm birth in China. JAMA Pediatr 2018;172(3):e174872. doi:10.1001/jamapediatrics.2017.4872.

[33] Guo LQ, Chen Y, Mi BB, et al. Ambient air pollution and adverse birth outcomes: a systematic review and meta-analysis. J Zhejiang Univ Sci B 2019;20(3):238–252. doi:10.1631/jzus.B1800122.

[34] Bekkar B, Pacheco S, Basu R, et al. Association of air pollution and heat exposure with preterm birth, low birth weight, and Stillbirth in the US: a systematic review. JAMA Netw Open 2020;3(6):e208243. doi:10.1001/jamanetworkopen.2020.8243.

[35] Goldenberg RL, Culhane JF. Prepregnancy health status and the risk of preterm delivery. Arch Pediatr Adolesc Med 2005;159(1):89–90. doi:10.1001/archpedi.159.1.89.

[36] Kistka ZA, Palomar L, Lee KA, et al. Racial disparity in the frequency of recurrence of preterm birth. Am J Obstet Gynecol 2007;196(2):131 e1–131 e6. doi:10.1016/j.ajog.2006.06.093.

[37] Wilcox AJ, Skjaerven R, Lie RT. Familial patterns of preterm delivery: maternal and fetal contributions. Am J Epidemiol 2008;167(4):474–479. doi:10.1093/aje/kwm319.

[38] Wang X, Zuckerman B, Coffman GA, et al. Familial aggregation of low birth weight among whites and blacks in the United States. N Engl J Med 1995;333(26):1744–1749. doi:10.1056/NEJM199512283332606.

[39] Knijnenburg TA, Vockley JG, Chambwe N, et al. Genomic and molecular characterization of preterm birth. Proc Natl Acad Sci USA 2019;116(12):5819–5827. doi:10.1073/pnas.1716314116.

[40] Zhang G, Jacobsson B, Muglia LJ. Genetic associations with spontaneous preterm birth. N Engl J Med 2017;377(24):2401–2402. doi:10.1056/NEJMc1713902.

[41] Zhang G, Feenstra B, Bacelis J, et al. Genetic associations with gestational duration and spontaneous preterm birth. N Engl J Med 2017;377(12):1156–1167. doi:10.1056/NEJMoa1612665.

[42] Rebbeck TR, Bridges JFP, Mack JW, et al. A framework for promoting diversity, equity, and inclusion in genetics and genomics research. JAMA Health Forum 2022;3(4):e220603–e220603. doi:10.1001/jamahealthforum.2022.0603.

[43] Oken E, Baccarelli AA, Gold DR, et al. Cohort profile: project viva. Int J Epidemiol 2015;44(1):37–48. doi:10.1093/ije/dyu008.

[44] Almquist YB, Grotta A, Vagero D, et al. Cohort profile update: the Stockholm birth cohort study (SBC). Int J Epidemiol 2020;49(2):367–367e. doi:10.1093/ije/dyz185.

[45] Elliott J, Shepherd P. Cohort profile: 1970 British Birth Cohort (BCS70). Int J Epidemiol 2006;35(4):836–843. doi:10.1093/ije/dyl174.

[46] Hong X, Hao K, Ji H, et al. Genome-wide approach identifies a novel gene-maternal pre-pregnancy BMI interaction on preterm birth. Nat Commun 2017;8:15608. doi:10.1038/ncomms15608.

[47] Hong X, Bartell TR, Wang X. Gaining a deeper understanding of social determinants of preterm birth by integrating multi-omics data. Pediatr Res 2021;89(2):336–343. doi:10.1038/s41390-020-01266-9.

[48] Wang X, Zuckerman B, Pearson C, et al. Maternal cigarette smoking, metabolic gene polymorphism, and infant birth weight. JAMA 2002;287(2):195–202. doi:10.1001/jama.287.2.195.

[49] Yu Y, Tsai HJ, Liu X, et al. The joint association between F5 gene polymorphisms and maternal smoking during pregnancy on preterm delivery. Hum Genet 2009;124(6):659–668. doi:10.1007/s00439-008-0589-2.

[50] Tsai HJ, Liu X, Mestan K, et al. Maternal cigarette smoking, metabolic gene polymorphisms, and preterm delivery: new insights on GxE interactions and pathogenic pathways. Hum Genet 2008;123(4):359–369. doi:10.1007/s00439-008-0485-9.

[51] National Institute on Minority Health and Health Disparities. NIMHD Research Framework. Available from: ttps://nimhd.nih.gov/research-Framework. [Accessed April 25, 2022].

[52] Hao K, Wang X, Niu T, et al. A candidate gene association study on preterm delivery: application of high-throughput genotyping technology and advanced statistical methods. Hum Mol Genet 2004;13(7):683–691. doi:10.1093/hmg/ddh091.

[53] Wang L, Feng Y, Zhang Y, et al. Prolylcarboxypeptidase gene, chronic hypertension, and risk of preeclampsia. Am J Obstet Gynecol 2006;195(1):162–171. doi:10.1016/j.ajog.2006.01.079.

[54] Wang L, Wang X, Laird N, et al. Polymorphism in maternal LRP8 gene is associated with fetal growth. Am J Hum Genet 2006;78(5):770–777. doi:10.1086/503712.

[55] Li J, Hong X, Mesiano S, et al. Natural selection has differentiated the progesterone receptor among human populations. Am J Hum Genet 2018;103(1):45–57. doi:10.1016/j.ajhg.2018.05.009.

[56] Hong X, Rosenberg AZ, Zhang B, et al. Joint associations of maternal-fetal APOL1 genotypes and maternal country of origin with preeclampsia risk. Am J Kidney Dis 2021;77(6):879–888.e1. doi:10.1053/j.ajkd.2020.10.020.

[57] Hong X, Surkan PJ, Zhang B, et al. Genome-wide association study identifies a novel maternal gene x stress interaction associated with spontaneous preterm birth. Pediatr Res 2021;89(6):1549–1556. doi:10.1038/s41390-020-1093-1.

[58] Hong X, Sherwood B, Ladd-Acosta C, et al. Genome-wide DNA methylation associations with spontaneous preterm birth in US blacks: findings in maternal and cord blood samples. Epigenetics 2018;13(2):163–172. doi:10.1080/15592294.2017.1287654.

[59] Xu R, Hong X, Zhang B, et al. DNA methylation mediates the effect of maternal smoking on offspring birthweight: a birth cohort study of multi-ethnic US mother-newborn pairs. Clin Epigenetics 2021;13(1):47. doi:10.1186/s13148-021-01032-6.

[60] Zhang B, Hong X, Ji H, et al. Maternal smoking during pregnancy and cord blood DNA methylation: new insight on sex differences and effect

modification by maternal folate levels. Epigenetics 2018;13(5):505–518. doi:10.1080/15592294.2018.1475978.

[61] Zhang M, Buckley JP, Liang L, et al. A metabolome-wide association study of in utero metal and trace element exposures with cord blood metabolome profile: findings from the Boston Birth Cohort. Environ Int 2022;158:106976. doi:10.1016/j.envint.2021.106976.

[62] Hong X, Zhang B, Liang L, et al. Postpartum plasma metabolomic profile among women with preeclampsia and preterm delivery: implications for long-term health. BMC Med 2020;18(1):277. doi:10.1186/s12916-020-01741-4.

[63] He H, Butz A, Keet CA, et al. Preterm birth with childhood asthma: the role of degree of prematurity and asthma definitions. Am J Respir Crit Care Med 2015;192(4):520–523. doi:10.1164/rccm.201503-0522LE.

[64] Wang G, Divall S, Radovick S, et al. Preterm birth and random plasma insulin levels at birth and in early childhood. JAMA 2014;311(6):587–596. doi:10.1001/jama.2014.1.

[65] Kumar R, Yu Y, Story RE, et al. Prematurity, chorioamnionitis, and the development of recurrent wheezing: a prospective birth cohort study. J Allergy Clin Immunol 2008;121(4):878–84.e6. doi:10.1016/j.jaci.2008.01.030.

[66] Raghavan R, Helfrich BB, Cerda SR, et al. Preterm birth subtypes, placental pathology findings, and risk of neurodevelopmental disabilities during childhood. Placenta 2019;83:17–25. doi:10.1016/j.placenta.2019.06.374.

[67] Barker DJ, Godfrey KM, Gluckman PD, et al. Fetal nutrition and cardiovascular disease in adult life. Lancet 1993;341(8850):938–941. doi:10.1016/0140-6736(93)91224-a.

[68] Barker DJP. *Mothers, Babies, and Disease in Later Life*. London: BMJ publishing group London; 1994.

[69] Gluckman PD, Hanson MA, Mitchell MD. Developmental origins of health and disease: reducing the burden of chronic disease in the next generation. Genome Med 2010;2(2):14. doi:10.1186/gm135.

[70] Wadhwa PD, Buss C, Entringer S, Swanson JM. Developmental origins of health and disease: brief history of the approach and current focus on epigenetic mechanisms. Semin Reprod Med 2009;27(5):358–368.

[71] Gillman MW. Developmental origins of health and disease. N Engl J Med 2005;353(17):1848–1850. doi:10.1056/NEJMe058187.

[72] Godfrey KM, Lillycrop KA, Burdge GC, et al. Epigenetic mechanisms and the mismatch concept of the developmental origins of health and disease. Pediatr Res 2007;61(7):5R5–10R. doi:10.1203/pdr.0b013e318045bedb.

[73] Heindel JJ, Balbus J, Birnbaum L, et al. Developmental origins of health and disease: integrating environmental influences. Endocrinology 2015;156(10):3416–3421. doi:10.1210/EN.2015-1394.

[74] Wang G, Foney DM, DiBari J, et al. A prospective cohort study on the intersectionality of obesity, chronic disease, social factors, and incident risk of COVID-19 in US low-income minority middle-age mothers. Int J Obes (Lond) 2021;45(12):2577–2584. doi:10.1038/s41366-021-00943-x.

[75] Zhang M, Liu T, Wang G, et al. In utero exposure to heavy metals and trace elements and childhood blood pressure in a U.S. urban, low-income, minority birth cohort. Environ Health Perspect 2021;129(6):67005. doi:10.1289/EHP8325.

[76] Anand NS, Ji Y, Wang G, et al. Maternal and cord plasma branched-chain amino acids and child risk of attention-deficit hyperactivity disorder: a prospective birth cohort study. J Child Psychol Psychiatry 2021;62(7):868–875. doi:10.1111/jcpp.13332.

[77] Raghavan R, Selhub J, Paul L, et al. A prospective birth cohort study on cord blood folate subtypes and risk of autism spectrum disorder. Am J Clin Nutr 2020;112(5):1304–1317. doi:10.1093/ajcn/nqaa208.

[78] Ji Y, Azuine RE, Zhang Y, et al. Association of cord plasma biomarkers of in utero acetaminophen exposure with risk of attention-deficit/hyperactivity disorder and autism spectrum disorder in

childhood. JAMA Psychiatry 2020;77(2):180–189. doi:10.1001/jamapsychiatry.2019.3259.

[79] Wang G, Sun Q, Liang L, et al. Inter-generational link of obesity in term and preterm births: role of maternal plasma acylcarnitines. Int J Obes (Lond) 2019;43(10):1967–1977. doi:10.1038/s41366-019-0417-x.

[80] Hong X, Liang L, Ji H, et al. Fetal lipidome and incident risk of food allergy: a prospective birth cohort study. Pediatr Allergy Immunol 2022;33(2):e13722. doi:10.1111/pai.13722.

[81] Ogunwole SM, Mwinnyaa G, Wang X, et al. Preeclampsia across pregnancies and associated risk factors: findings from a high-risk US birth cohort. J Am Heart Assoc 2021;10(17):e019612. doi:10.1161/JAHA.120.019612.

[82] Liu T, Zhang M, Guallar E, et al. Trace minerals, heavy metals, and preeclampsia: findings from the Boston birth cohort. J Am Heart Assoc 2019;8(16):e012436. doi:10.1161/JAHA.119.012436.

[83] Wang G, Liu X, Bartell TR, Pearson C, Cheng TL, Wang X. Vitamin D trajectories from birth to early childhood and elevated systolic blood pressure during childhood and adolescence. Hypertension 2019;74(2):421–430. doi:10.1161/HYPERTENSIONAHA.119.13120

[84] McGowan EC, Hong X, Selhub J, et al. Association between folate metabolites and the development of food allergy in children. J Allergy Clin Immunol Pract 2020;8(1):132–140.e5. doi:10.1016/j.jaip.2019.06.017.

[85] Hong X, Liang L, Sun Q, et al. Maternal triacylglycerol signature and risk of food allergy in offspring. J Allergy Clin Immunol 2019;144(3):729–737. doi:10.1016/j.jaci.2019.03.033.

[86] Ji Y, Riley AW, Lee LC, et al. Maternal biomarkers of acetaminophen use and offspring attention deficit hyperactivity disorder. Brain Sci 2018;8(7):127. doi:10.3390/brainsci8070127.

[87] Wang G, Hu FB, Mistry KB, et al. Association between maternal pre-pregnancy body mass index and plasma folate concentrations with child metabolic health. JAMA Pediatr 2016;170(8):e160845. doi:10.1001/jamapediatrics.2016.0845.

[88] Bousquet J, Gern JE, Martinez FD, et al. Birth cohorts in asthma and allergic diseases: report of a NIAID/NHLBI/MeDALL joint workshop. J Allergy Clin Immunol 2014;133(6):1535–1546. doi:10.1016/j.jaci.2014.01.018.

[89] Kumar R, Ouyang F, Story RE, et al. Gestational diabetes, atopic dermatitis, and allergen sensitization in early childhood. J Allergy Clin Immunol 2009;124(5):1031-8 e1-4. doi:10.1016/j.jaci.2009.06.052.

[90] Boston Medical Center. About BMC. 2022. Available from: https://www.bmc.org/about-bmc. [Accessed April 20, 2022].

[91] Baik SH, Fox RS, Mills SD, et al. Reliability and validity of the Perceived Stress Scale-10 in Hispanic Americans with English or Spanish language preference. J Health Psychol 2019;24(5):628–639. doi:10.1177/1359105316684938.

[92] Redline RW. Classification of placental lesions. Am J Obstet Gynecol 2015;213(4 Suppl):S21–S28. doi:10.1016/j.ajog.2015.05.056.

[93] Krieger N. 11 The science and epidemiology of racism and health: racial/ethnic categories, biological expressions of racism, and the embodiment of inequality—an ecosocial perspective. What's the Use of Race?: Modern Governance and the Biology of Difference; 2010:225.

[94] Krieger N. Methods for the scientific study of discrimination and health: an ecosocial approach. Am J Public Health 2012;102(5):936–944. doi:10.2105/AJPH.2011.300544.