

STUDY PROTOCOL

Open Access



# The Children's Respiratory and Environmental Workgroup (CREW) birth cohort consortium: design, methods, and study population

James E. Gern<sup>1,15\*</sup>, Daniel J. Jackson<sup>1</sup>, Robert F. Lemanske Jr<sup>1</sup>, Christine M. Seroogy<sup>1</sup>, Umberto Tachinardi<sup>1</sup>, Mark Craven<sup>1</sup>, Stephen Y. Hwang<sup>2</sup>, Carol M. Hamilton<sup>2</sup>, Wayne Huggins<sup>2</sup>, George T. O'Connor<sup>3</sup>, Diane R. Gold<sup>4</sup>, Rachel Miller<sup>5</sup>, Meyer Kattan<sup>5</sup>, Christine C. Johnson<sup>6</sup>, Dennis Ownby<sup>6</sup>, Edward M. Zoratti<sup>6</sup>, Robert A. Wood<sup>7</sup>, Cynthia M. Visness<sup>8</sup>, Fernando Martinez<sup>9</sup>, Anne Wright<sup>9</sup>, Susan Lynch<sup>10</sup>, Carole Ober<sup>11</sup>, Gurjit K. Khurana Hershey<sup>12</sup>, Patrick Ryan<sup>12</sup>, Tina Hartert<sup>13</sup> and Leonard B. Bacharier<sup>14</sup>

## Abstract

**Background:** Single birth cohort studies have been the basis for many discoveries about early life risk factors for childhood asthma but are limited in scope by sample size and characteristics of the local environment and population. The Children's Respiratory and Environmental Workgroup (CREW) was established to integrate multiple established asthma birth cohorts and to investigate asthma phenotypes and associated causal pathways (endotypes), focusing on how they are influenced by interactions between genetics, lifestyle, and environmental exposures during the prenatal period and early childhood.

**Methods and results:** CREW is funded by the NIH Environmental influences on Child Health Outcomes (ECHO) program, and consists of 12 individual cohorts and three additional scientific centers. The CREW study population is diverse in terms of race, ethnicity, geographical distribution, and year of recruitment. We hypothesize that there are phenotypes in childhood asthma that differ based on clinical characteristics and underlying molecular mechanisms. Furthermore, we propose that asthma endotypes and their defining biomarkers can be identified based on personal and early life environmental risk factors. CREW has three phases: 1) to pool and harmonize existing data from each cohort, 2) to collect new data using standardized procedures, and 3) to enroll new families during the prenatal period to supplement and enrich extant data and enable unified systems approaches for identifying asthma phenotypes and endotypes.

**Conclusions:** The overall goal of CREW program is to develop a better understanding of how early life environmental exposures and host factors interact to promote the development of specific asthma endotypes.

**Keywords:** Asthma, Birth cohort, Longitudinal study, Development, Allergy, Environment, Children

\* Correspondence: [gern@medicine.wisc.edu](mailto:gern@medicine.wisc.edu)

<sup>1</sup>University of Wisconsin School of Medicine and Public Health, Madison, WI 53706, USA

<sup>15</sup>Department of Pediatrics, School of Medicine and Public Health, University of Wisconsin-Madison, Clinical Science Center-K4/918, 600 Highland Ave, Madison, WI 53792-9988, USA

Full list of author information is available at the end of the article



## Background

The environment during the prenatal period and in early life is a major contributor to the risk of developing childhood asthma. Birth cohort studies from single research centers have identified individual factors that affect the risk for developing childhood asthma, including exposure to allergens, pollutants, bacteria, psychosocial stress, viral respiratory illnesses and patterns of microbial colonization. Despite such advances, further progress in understanding the root causes of asthma have been hampered by the relatively small size of previous studies that typically lack diverse population and broad assessments of early life exposures. Thus, single cohorts lack the resources to integrate multiple types of complex data to identify the individual and combined effects of environmental factors in asthma development, asthma phenotypes and endotypes. Furthermore, broad and diverse populations and environmental assessments are needed to determine and compare interactions with genetics, effect sizes, mediation and additive effects and dose response relationships, and to identify geographical and temporal differences in exposures and how they impact asthma development across the US. Another notable limitation of individual cohorts with modest sample size is the inability to define prenatal and early life risk factors for infrequent outcomes that are of special clinical significance, such as severe asthma and susceptibility to acute exacerbations. Furthermore, differences in data collection and sample analyses make it challenging to pool, harmonize and integrate findings from different cohorts.

To help overcome these challenges, investigators leading 12 asthma birth cohorts and three additional scientific

centers across the U.S. have established the Children's Respiratory and Environmental Workgroup (CREW) consortium. Collectively, the CREW consortium will identify phenotypes of childhood asthma to develop an understanding of which early life environmental influences are causally associated with these different types of asthma and when their influences are relevant, and to identify targets for future efforts aimed at preventing childhood asthma.

## Methods

### Study overview

CREW is funded as part of the NIH Environmental Influences on Child Health Outcomes (ECHO) program [1]. As such, in addition to pooling data and samples among CREW cohorts, CREW data also will be shared with the larger ECHO program. ECHO consists of 34 initial collaborative awards to birth cohorts or consortia that are investigating effects of early life exposures on five health outcomes in childhood: 1) airway disease (e.g. CREW), 2) obesity, 3) perinatal outcomes, 4) neurocognitive development, and 5) positive health. Enrollment for ECHO is currently underway and the total study population will include about 50,000 families, including the new families recruited during the prenatal period by all ECHO awardees.

CREW consists of 12 individual U.S. birth cohorts and three scientific centers (Table 1). Each of the cohorts was originally established with a major focus on identifying prenatal and/or early life exposures that influence the development of asthma. The initial enrollment for CREW cohorts included nearly 9000 newborns, and

**Table 1** CREW Member Cohorts\* and Study Populations

	Initial Enrollment	Location	Population	Recruitment	Cauc/WH	AA	Hispanic	Asian	Other/Missing
CAS	835	Suburban Detroit	General	1987–1989	752 (90%)	22 (3%)	20 (2%)	26 (3%)	15 (2%)
CCAAPS	762	Cincinnati	High risk	2001–2003	579 (76%)	158 (21%)	0 (0%)	2 (0.3%)	23 (3%)
CCCEH	727	Manhattan and Bronx	General	1998–2006	0	291 (40%)	436 (60%)	0	0
COAST	289	Madison	High risk, suburban	1998–2000	243 (87%)	12 (4%)	8 (3%)	0 (0%)	7 (2%)
EHAAS	505	Boston	High risk	1994–1996	370 (74%)	64 (12%)	30 (6%)	28 (6%)	7 (1%)
IIS	482	Tucson	General	1997–2003	280 (58%)	24 (5%)	125 (26%)	0 (0%)	53 (11%)
INSPIRE	1952	Middle Tennessee	General	2012–2014	1269 (65%)	351 (18%)	176 (9%)	0 (0%)	156 (8%)
MAAP	120	Metro Detroit	General	2014–2016	85 (71%)	22 (18%)	0 (0%)	4 (3%)	9 (8%)
TCRS	1246	Tucson	General	1980–1984	822 (66%)	50 (4%)	312 (25%)	0 (0%)	62 (5%)
URECA	609	NYC, St Louis, Baltimore, Boston	High risk, urban	2004–2006	22 (4%)	429 (70%)	116 (19%)	0 (0%)	42 (7%)
WHEALS	1258	Metro Detroit	General	2003–2007	325 (26%)	787 (63%)	81 (6%)	30 (2%)	35 (3%)
WISC	212	Rural WI	Rural: farm and non-farm	2013–2019	204 (96%)	4 (2%)	2 (1%)	2 (1%)	0 (0%)
Total	8997				4951 (55%)	2214 (25%)	1306 (15%)	92 (1%)	409 (4%)

\*Abbreviations: CAS, Children's Asthma Study; CCAAPS, Cincinnati Childhood Allergy and Air Pollution Study; CCCEH, Columbia Center for Children's Environmental Health Cohort; COAST, Childhood Origins of Asthma study; EHAAS; IIS, Infant Immune Study; INSPIRE, Infant Susceptibility to Pulmonary Infections and Asthma Following RSV Exposure study; MAAP, Microbes, Allergy, Asthma and Pets study; TCRS, Tucson Children's Respiratory Study; URECA, Urban Environment and Childhood Asthma study; WHEALS, Wayne County Health Environment Allergy and Asthma Longitudinal Study; WISC, Wisconsin Infant Study Cohort

approximately 6000 of whom are still being followed. Study designs with enrollment criteria for the individual cohorts have been previously described [2–17]. The investigators of CREW member cohorts have a broad variety of expertise related to early life environmental exposures, allergy and immunology, mechanisms of asthma pathogenesis, geospatial technology, pulmonology, epidemiology, virology, biostatistics, biomedical informatics, genetics, epigenetics and transcriptomics. In addition, there are three scientific centers to provide core services for analyses of the microbiome (University of California, San Francisco), genetics and epigenetics (University of Chicago) and geospatial data analyses (Harvard School of Public Health). The administrative center for CREW is at the University of Wisconsin-Madison, and Rho Inc. (Chapel Hill, NC) is the Coordinating Center for CREW.

CREW has three main hypotheses (Fig. 1). First, longitudinal analyses of individual characteristics, including patterns of respiratory symptoms, aeroallergen-specific IgE, lung function, sex and ethnicity can be used to identify respiratory phenotypes, including those that are highly enriched for asthma. We further propose that these phenotypes will have closer associations with etiologic factors (e.g. environmental exposures, genetics), compared to a general yes/no definition of childhood asthma due to reduction in misclassification. Second, host factors (e.g. genetics) interact with environmental exposures during the prenatal period and early childhood, leading in some cases to epigenetic alterations, and ultimately to specific changes in immune and lung development that increase the risk for specific asthma endotypes. Finally, molecular biomarkers can be identified to provide a precise and comprehensive identification of childhood asthma endotypes.

To address these hypotheses and goals, the CREW consortium has designed three protocols. The first protocol (Phase I) pools extant data from the 12 cohorts to address

specific hypotheses. Because the studies were designed by different groups with their own unique objectives, extensive data harmonization will be required. The second (Phase II) protocol will collect new data on current participants from 10 of the cohorts using standardized procedures and questionnaires and will develop standardized outcomes within CREW (and ECHO). Because most of the early life specimens for current CREW participants already have been collected, a third protocol (Phase III, or “NewCREW”) initiating a new multisite cohort will allow standardized collection of prenatal and early life environmental and immunologic samples to measure factors that could influence or be related to the development of allergic diseases and asthma.

### Phase I: leveraging extant data

#### Data sharing agreements

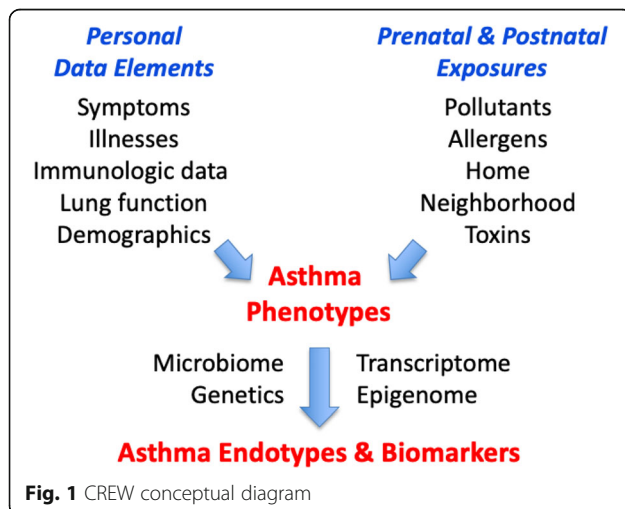
For Phase I, each participating institution reviewed and approved a consortium and data sharing agreement that enables within-group (including the ECHO data analysis center) sharing of previously collected data under a waiver of consent. This agreement was approved by Human Subjects Committees at each of the participating institutions.

Each of the CREW member cohorts has collected data and samples related to personal and environmental characteristics that could affect allergic diseases and asthma, which varied to some degree among cohorts. This information includes clinical questionnaires and biological measures that characterized environmental exposures (e.g. allergens, pollutants and microbes), family and personal history of allergic diseases, respiratory infections and illnesses, immune development, genomic data, and microbial colonization.

The individual datasets have been stripped of protected health information (PHI) as defined by HIPAA, except for those indirect identifiers permitted in a limited data set, such as dates of birth, dates of data and sample acquisition, and town/city, state and zip code. The current Data Sharing Protocol also allows for sharing of this pooled limited dataset with the ECHO Data Analysis Center to enable collaborative research studies with any of the 72 other ECHO cohorts.

#### Data harmonization and analysis projects

Demographic and other exposure and outcome data have been and will continue to be compiled by the individual cohorts and sent to the University of Wisconsin (UW). A Publications Committee has been established to review and approve analysis concepts and resulting publications. Data harmonization is a critical function that allows merging data from the 12 cohorts. The harmonization process requires understanding the original data definitions provided by cohort investigators and data managers, and mapping them into a common



**Fig. 1** CREW conceptual diagram

CREW definition developed by informaticians. This procedure is performed by data scientists at UW working closely with cohort investigators and data managers.

### **Geospatial approach**

Exposures to environmental toxicants and community characteristics during the prenatal period and early childhood have a significant impact on allergic disease and asthma. CREW investigators have developed distributed methods to derive environmental exposures and community characteristics, including measures of disparities, for all CREW participants based on residential locations. A containerized software platform was developed to enable individual CREW sites to geocode their participants' birth record addresses and link these to community characteristics [18]. Census tract-level information regarding race, ethnicity, education, housing, and other indicators of disadvantage/advantage was obtained by collaborators at Harvard Center for Geographic Analysis (Cambridge, MA) and the University of Arizona using the decennial U.S. census in 1980, 1990, 2000, and 2010. These data, combined with a TIGER/Line address range geocoder, census tract boundaries, and analysis code were containerized and distributed to individual CREW sites. Using this software, each study site can geocode their participants' birth record address and link these to census-tract level information for further analysis.

### **Phase II: standardized data collection**

The CREW Phase II protocol includes standardized measures of exposure assessments, immunologic studies, lung function testing and outcome measures. The measurements will be obtained at least once per life stage interval (1–4 years, 5–7 years, 8–10 years, 11–14 years, 15–17 years and 18 years and over) to standardize measurements while accommodating different schedules of the cohorts.

### **Enrollment criteria and recruitment procedures**

Participating families from 10 of the CREW member cohorts are eligible for Phase II study enrollment. Exclusion criteria include past or current medical problems that could pose additional risks from participation in the study, interfere with the participant's ability to comply with study requirements or otherwise adversely impact the quality or interpretation of the data. Recruitment techniques varied by study site, but study coordinators generally provided information about CREW and an invitation to participate during usual cohort activities by phone, email, mail, or at in-person study visits. For cohorts lacking upcoming activities, sites used mailings, secure web-based communication or phone calls to provide information and an invitation to join CREW. For recruitment of adult participants, sites had the flexibility to contact them first via mail or email to gauge interest in participating in CREW. This included a

short questionnaire to assess current allergies and asthma, and an invitation to join the new study.

Individual site Human Subjects Committees approved the CREW Phase II protocol. Families or adult participants provide informed consent and children between the ages of 7 and 18 years provide assent before being enrolled in CREW and ECHO.

### **Questionnaires**

Questionnaires will be administered to assess the clinical, environmental and medication history of the participants (Table 2). Questionnaire administration will include paper questionnaires, online questionnaires, electronic case report forms (eCRF), and interviewer-administered questionnaires. The study's primary source of information on the occurrence and frequency of wheezing is the Airways Outcomes questionnaire. This questionnaire includes questions from the ISAAC study [19], supplemented by additional questions related to food allergy, asthma medication use, anaphylaxis and allergic rhinitis. The Airways Outcomes questionnaire (Online data supplement) will also be administered by other ECHO cohorts, to enable standardized outcome assessments for allergic diseases and asthma across all ECHO cohorts. Additional questionnaires will include Tanner staging, parental or self-assessment of atopic dermatitis (POEM [20, 21]), and demographic information, among others (Table 3).

### **Physical examination**

Physical examination will include anthropometrics and measurement of heart rate and blood pressure.

### **Blood samples**

Peripheral blood will be being collected as fasting morning samples when possible, especially between ages 8–17 years to enable measurement of sex hormones, insulin and glucose. Approximately 15 mL will be obtained at each annual clinic visit. These samples will be transported to the center's laboratory on the day of collection and at room temperature. At the center-specific laboratories, blood will be processed and aliquots of serum and plasma will be cryopreserved at  $-80^{\circ}\text{C}$ . In addition, peripheral blood mononuclear cells will be separated from heparinized blood and cryopreserved in aliquots using pre-chilled alcohol reservoirs. A complete blood count and differential will be performed, and plasma will be analyzed for allergen-specific IgE (Uni-CAP, ThermoFisher). DNA from peripheral blood samples will be stored for studies of epigenetic regulation.

### **Airway samples**

Nasal wash, blow or lavage specimens will be collected using techniques that are tailored to the age of the child. For young children (birth to age 3–4 years), we will collect a nasal wash specimen by instilling 2 ml of buffered

**Table 2** Study Activities in the CREW Standardized Data Collection Protocol\*

	Age (in years)					
	1–4	5–7	8–10	11–13	14–17	> 18
Visit window	+ 12 mon	+ 12 mon	+ 12 mon	+ 12 mon	+ 12 mon	+ 12 mon
Time	2.0 h	3.0 h	3.0 h	3.0 h	3.0 h	3.0 h
Essential procedures						
Consent/Assent	x	x	x	x	x	x
Eligibility	x	x	x	x	x	x
Distribute home collection kits	x	x	x	x	x	x
PE and anthropometrics	x	x	x	x	x	x
Questionnaires	x	x	x	x	x	x
Blood	x	x	x	x	x	x
Urine	x	x	x	x	x	x
Stool	x	x	x	o	x	o
Nasal wash/blow/lavage	w	b	l	l	l	l
Home dust collection	x					
Spirometry			x	x	x	
Spirometry w/ reversibility		x	o	o	o	x
eNO		x	x	x	x	x
AE collection	x	x	x	x	x	x
Optional procedures						
Nasal brushing	o			o		o
Prick skin tests	o	o	o	o	o	o
Phone call (Follow up)	x	x	x	x	x	x

Abbreviations: w, nasal wash; b, nasal blow; l, nasal lavage; o, optional

**Table 3** Standardized Questionnaires and Timing of Administration

Questionnaire	Age in years					
	1–4	5–7	8–10	11–13	14–17	> 18
Demographics						
- Addresses	x	x	x	x	x	x
- Household census						
- SES						
Asthma control						
- ACT or Pediatric ACT		x	x	x	x	x
Self-Tanner Stage						
			x	x	x	
Health history						
- Allergic disease and asthma	x	x	x	x	x	x
- Allergic rhinitis						
- Atopic dermatitis						
- Food allergies						
- Perinatal (if not prospectively recorded)						
POEM (for atopic dermatitis)	x	x	x	x	x	x



saline into the nose using a modified bulb syringe, followed by gentle suctioning is applied to recover the sample [22]. For capable preschool children, we will use a nasal blow technique [23]. The procedure involves wetting the nose with normal saline spray, followed by blowing the nose into a collection tube.

For adults and children over the age of 8 years, a nasal lavage will be performed. Following a nasal blow while in a sitting position, saline is introduced into the one nostril, and nasal lavage fluid is passively recovered from the other nostril. This is repeated on the other side.

For each of these nasal samples, separate aliquots will be prepared for analyses of virology, microbiome, host cell gene expression, proteins and other analytes.

Nasal epithelial cells will be obtained brushing the inferior turbinate of one nasal passage using a nylon flocked swab. The epithelial cells will be a source for mRNA and DNA (epigenetics) isolation.

#### **Other samples (stool, urine, saliva)**

The participant will be asked to provide stool samples to examine fecal bacterial and fungal content. The sample can be collected during the clinic visit or at home using the collection kit and directions. If the specimen is collected at home, it can be mailed to the laboratory or brought to the clinic on the day of the study visit. Clean catch urine will be collected for measurement of chemical exposures, inflammation or biomarkers that might be related to asthma. Home collection kits will be distributed during the consenting visit. Pre-packaged collection kits and directions will be given to the participant and/or parent/guardian for collection of stool and urine. The study staff will review the directions with the participant and/or parent guardian.

#### **Allergy skin testing**

Prick skin testing to identify aeroallergen sensitization using a Multi-Test PC device (Lincoln Diagnostics, Decatur, IL) is an optional procedure at each of the life stages. Indoor and outdoor allergens that were selected for local prevalence (Table 4, Greer Laboratories). Wheal sizes are measured after 15 min. The wheal's longest length and width (measured perpendicular to the length at its midpoint) are measured to the nearest millimeter and averaged to give a mean wheal size. A positive reaction is defined as a mean wheal size at least 3 mm larger than the saline control.

#### **Total and allergen-specific IgE**

Total and allergen-specific IgE antibodies are measured by fluoroenzyme immunoassay (ImmunoCAP, ThermoFisher, Uppsala, Sweden). The panel of specific IgE antibodies vary by age and in some cases region of the country. For

**Table 4** CREW skin test antigens

Test #	Extract
1	Control, 50% (V/V) glycerin
2	Histamine 10 mg/ml
3	Mouse epithelia
4	Dog epithelia
5	Cat hair (standardized)
6	<i>D. farinae</i> + <i>D. pteronyssinus</i> mix
7	American/German cockroach mix
8	<i>Penicillium notatum</i> / <i>Penicillium chrysogenum</i>
9	<i>Alternaria tenuis</i>
10	<i>Cladosporium herbarum</i>
11	<i>Aspergillus</i> mix
12	Ragweed mix (giant and short)
13	Eastern 6 tree mix (American beech, eastern cottonwood, red oak, red/river birch, shagbark hickory, white ash)
14	K-O-T grass mix (Kentucky blue/june, orchard, timothy)
15	Maple/box elder mix
16	3 weed mix (cocklebur, lamb's quarter, pigweed)

detection of sensitization in early childhood, the specific IgE panel includes egg white, milk, peanut and a multiple aeroallergen screen (Phadiotop, ThermoFisher). If the Phadiotop is positive, a panel of aeroallergen-specific IgE (cat, dog, mold mix, and house dust mite mix) is measured. For adults and children over the age of 5 years, aeroallergen testing includes birch or oak, ragweed, and Timothy grass; and *Dermatophagoides farinae*, *Dermatophagoides pteronyssinus*, dog epithelium, cat dander/epithelium, German cockroach, mouse urine protein, and *Alternaria alternata*. Measurement of IgE antibodies to selected allergens will be performed at each life stage.

#### **Lung function and exhaled nitric oxide (eNO)**

To provide standardized protocols for lung function measurement, equipment, and overreading of studies, CREW has established a Pulmonary Function Core. Equipment was selected for spirometry (Jaeger Sentry Suite, Mettawa IL) and measurement of eNO (NIOX VERO®, Circassia Pharmaceuticals Inc., Morrisville, NC). Computer software is loaded on spirometers at each site with onsite training available to facilitate secure transmission of studies to the central server. CREW coordinators attended a 2-stage training session for introductory and advanced spirometry. All lung function studies will be overread by an experienced respiratory technician in Madison.

Beginning at age 5–7 years, children will perform a test for measurement of eNO. Prior to spirometry, exhaled NO is measured employing a technique modified after Silkoff et al. [24] and following American Thoracic

Society guidelines for eNO assessment [25]. In brief, this technique utilizes a resistive device that provides a constant low expiratory flow rate and ensures velum closure.

Spirometry will be conducted during clinic visits at each life stage beginning at 5–7 years according to ATS guidelines [26]. Participants will be asked to hold albuterol for 4 h prior to this procedure, and long-acting beta agonists for 12 h before the visit. If the participant feels they need to take their medication because their symptoms get worse, they will be instructed to do so and call the study staff.

Reversibility following administration of albuterol (4 puffs) will be performed at ages 5–7 years of age and in adults ( $\geq$  age 18 years) and is optional for intermediate age groups. After completion of spirometry as described above, albuterol via MDI with spacer will be administered. Fifteen minutes later the spirometry will be repeated. In females  $\geq$  age 12 years, urine pregnancy tests will be conducted to ensure that and this reversibility test will not be performed in pregnant woman.

#### **Investigator training**

Face-to-face and web-based training sessions were held for sample collection and processing and to review questionnaires and recruitment procedures. The initial face-to-face training occurred in December 2017 at the University of Wisconsin-Madison, and included study coordinator and lab representatives from each of the clinical centers. Sessions were designed to cover clinical procedures relating to collection and processing of biological samples, as well as pulmonary function testing, exhaled nitric oxide testing, and consenting. After completion of the in-person training, we conducted additional web-based training sessions during bi-weekly coordinator calls. These sessions are recorded for new staff to review in the future for certification. Web-based training is ongoing and occurs regularly as issues and concerns arise.

CREW employs a comprehensive manual of procedures that cover all aspects of study management and protocol implementation. The training program includes a mixed methodology approach, an accountability component, and pre-recorded training modules to ensure standardization of the training process. To assess training effectiveness, skills are evaluated with written exams, by observation, or certification with a qualified trainer, and by on-going quality control evaluations. Training completion is tracked using a web application (Moodle, <https://moodle.org>), which is monitored regularly by the Coordinating Center.

#### **Phase III: recruitment of additional families**

The CREW phase III protocol (“NewCREW”) will enroll approximately 400 additional pregnant women and their infants during the prenatal period at four CREW centers. The purpose of this protocol is to collect early life exposure, biospecimens, and outcomes data for the CREW

and ECHO protocols. There will be an emphasis on collection of maternal specimens during the prenatal period and extensive pre-, peri- and postnatal environmental sampling. The sampling strategy was designed to supplement and enrich existing data in CREW and ECHO and will provide specimens that are optimized for analyses of early life environmental exposures (e.g. microbial exposures and colonization), immune development (gene transcription, metabolites and proteins), epigenetics (nasal cells) and outcomes (e.g. wheeze, growth, allergic sensitization). Outcomes data will be collected for allergic diseases and asthma, and other ECHO outcomes including perinatal outcomes, neurocognitive development, obesity and positive health.

#### **Collaboration with the ECHO data collection protocol**

The primary objective of the ECHO-wide Cohort Data Collection Protocol is to facilitate the creation of a data platform consisting of essential data elements from all participating ECHO awardees, plus recommended data elements from subsets of cohorts. This protocol standardizes new data collection across the ECHO Program and specifies what data elements cohorts should collect (new data) and share (both existing data and new data) across the life course, from the prenatal period through adolescence, to create the ECHO-wide Cohort. CREW is participating in the ECHO-wide Cohort Data Collection Protocol in both Phases II and III by collecting essential and many recommended data elements. This includes collection of additional biological specimens, such as hair and toenails, for evaluation of chemical exposures.

#### **Data management**

##### **Central database**

All 3 phases of CREW will receive support from the CREW informatics team in data collection, data management, and the creation of a central database (CREW Asthma eLab) for the study. Asthma eLab was initially developed by investigators at the University of Manchester for use in the STELAR birth cohort consortium [27]. It serves as a secure database, a forum for investigator interaction, and a platform for analyses and manuscript development. With assistance from the NIH, the STELAR group transferred an updated version of Asthma eLab to the UW-Madison, and are helping with software customization, installation and training. STELAR and UW investigators collaborated to create the state-of-the-art Fast Health Interoperability Resources (FHIR) profiles needed for the Asthma eLab platform. These profiles will enable CREW data, research collaborations, data harmonization and data analyses to occur within eLab.

### **Data handling and security**

We have developed workflows and infrastructure to support secure data collection for different parts of the CREW project. Electronic clinical research forms (CRFs) were designed and implemented in REDCap to document consent for the Phase II CREW protocol, and a dashboard was implemented to track recruitment. Using a similar approach, about 30 CRFs were implemented to support data collection for the Phase II protocol. An Information Governance Policy was written and approved by the CREW Steering Committee to provide detailed rules for data access, handling and security.

### **Quality control**

Quality monitoring procedures for new data collection activities have been established for CREW and start with comprehensive training and certification of all staff before performing any study procedures. All CREW study data are subject to comprehensive built-in edit checks and manual queries as part of the REDCap data entry system. The CREW Coordinating Center employs a suite of statistical data checks to examine data for outliers and anomalies. Standard and custom reports and dashboard visualization displays (e.g. recruitment status, sample explorer tool) are used to monitor data quality and study performance and will be available to all CREW research staff via the CREW password-protected website (RhoPORTAL). These procedures will be implemented with each of the CREW protocols. In addition, we will implement and adhere to all quality monitoring procedures that will be established for the ECHO Data Collection Protocol.

### **Computation and statistical analyses**

#### ***The CREW Bioinformatics and Biostatistics Core (BBC)***

The BBC was formed to conduct and oversee key computational and statistical analysis tasks. These include establishing and operating the CREW e-Lab system, as described in the previous section. Analytic tasks include identification of the predominant asthma phenotypes and endotypes in the subject population using the high-dimensional, heterogeneous data (symptoms, lung function measures, immunologic variables, etc.). In addition, the BBC will develop models and statistical approaches to identify relationships among key variables in the complex data sets, including measures and load of exposure, epigenetic modifications, transcriptional profiling, metabolomics, protein expression and asthma phenotypes. These models will include longitudinal measurements, assessment of age dependency and consideration of covariates, effect modifiers and mediators.

Because there is a wide range of methodologies that can be applied to these tasks, and the CREW team encompasses a broad range of data-science expertise, we anticipate that we will investigate, develop, and evaluate a

variety of approaches for each task. In broad terms, some of the methodologies we will investigate include unsupervised learning, supervised learning, causal modeling, spatial modeling, and hypothesis testing. The BBC is comprised of biomedical bioinformaticists and statisticians from each site of our consortium. This team approach has the advantages of including investigators who are intimately familiar with the data sets while bringing a broad range of analytic expertise to the group.

### **Results**

#### **Phase I: leveraging extant data**

##### ***Study population***

The study participants of the CREW consortium represent a diverse national sample of children and their families recruited over the past 30 years into the 12 member cohorts. The total combined study population is 8997 at the time of birth (Table 1). The study population is quite diverse in terms of current age (ages < 1 through 36 years), date of recruitment (1980–2018), race and ethnicity. There is also considerable variation in the geographic locations of study participants, with representation from the East Coast (Baltimore, Boston, New York City), Midwest (Cincinnati, Detroit, Madison, Marshfield), South (Nashville, St. Louis) and West (Tucson). In addition to including large numbers of urban participants living close to academic medical centers, the CREW study population also includes suburban (CAS, COAST, EHAAS, WHEALS, INSPIRE) and rural (INSPIRE, WISC) children. The study designs include those selected to represent a general population, high risk cohorts on the basis of parental history of allergies or asthma, and stratified populations based on pet ownership (MAAP) or residence on a dairy farm (WISC). Collectively, the CREW consortium represents the largest and most diverse U.S. longitudinal cohort of children evaluated for asthma (Table 1).

##### ***Available data elements***

The ECHO Data Analysis Center (Research Triangle Inc. and Johns Hopkins University) and CREW Bioinformatics and Bioinformatics Core have surveyed the CREW member cohorts to identify common data elements to help direct the development of collaborative studies, listed by age group in Fig. 2. All cohorts have collected address information that will be translated into geocodes to be used to estimate a broad variety of exposures. During infancy and early childhood, multiple cohorts collected information on exposures to diet, allergens, tobacco smoke and pets. Immunologic testing was conducted in 11 cohorts in early life. Other areas of common data collection included early life viral respiratory infections (8 cohorts), allergen exposure (8 cohorts), obesity (8 cohorts), microbial



exposures and/or colonization (7 cohorts) and psychosocial stress (5 cohorts). All cohorts collected information on allergic outcomes and asthma.

### **Geocoding**

All 12 study cohorts completed distributive analyses to link addresses to geocodes. A total of 8978 birth record addresses were available for participants enrolled in CREW cohorts and, of these, 98.1% ( $n = 8810$ ) were geocoded with sufficient precision for additional analyses. The CREW workgroup has also worked closely with the ECHO Data Analysis Center (Johns Hopkins University and RTI) to apply our approach to the entire ECHO consortium to facilitate similar analyses across all ECHO cohorts.

Future work in CREW will extend these methods to additional spatiotemporally variable physical (e.g. air pollution, greenspace, temperature) and social exposures. In addition, we will link the derived environmental exposures and community characteristics to longitudinal wheeze phenotypes, asthma, and other childhood respiratory health outcomes to examine the associations of environmental exposures and community characteristics with these outcomes.

### **Inventory of samples available for analysis**

An inventory of stored samples (including serum, cells, DNA) in CREW cohorts has been completed. An explorer tool has been developed by the CREW Coordinating Center to enable CREW investigators to interactively query available samples by sample type, cohort, and age of child at collection (Fig. 3). This tool will be used to identify samples that can be used for “omics” analyses, allergic outcomes (allergen-specific IgE), biomarkers and exposure assessments. Importantly, this will provide an excellent resource for identifying sample sets for collaborative ECHO analyses.

### **Analysis projects and publications**

One of the main goals of CREW is to identify respiratory phenotypes based on longitudinal patterns of wheeze and other variables closely related to asthma such as allergic sensitization and lung function. Initial phenotyping of study participants based on longitudinal analysis of patterns of wheezing illnesses through the first decade of life is in progress. Analysis projects can be proposed by any CREW investigator, and are then reviewed and approved by the Publications and Steering Committees. Investigators from each cohort then have the final decision as to whether to participate in an analysis proposal. The ten currently approved analysis projects are listed in Table 5.

Because there is no universally accepted “gold standard” definition for the diagnosis of asthma in early childhood and definitions varied by study, we are exploring several different ways to classify asthma, including parental report

of doctor-diagnosed asthma, and doctor-diagnosed asthma with either symptoms or controller therapy in the past 12 months, which can work across all our cohorts.

### **Phase II: standardized data collection**

Participants who are currently enrolled or were previously enrolled in 10 of the CREW cohorts are being invited to participate in the standardized prospective CREW visits. The response by study participants has generally been positive, with < 5% refusals overall and 78% of the recruitment target being reached. Study visits began in August 2018, and the schedule of activities is shown in Table 3.

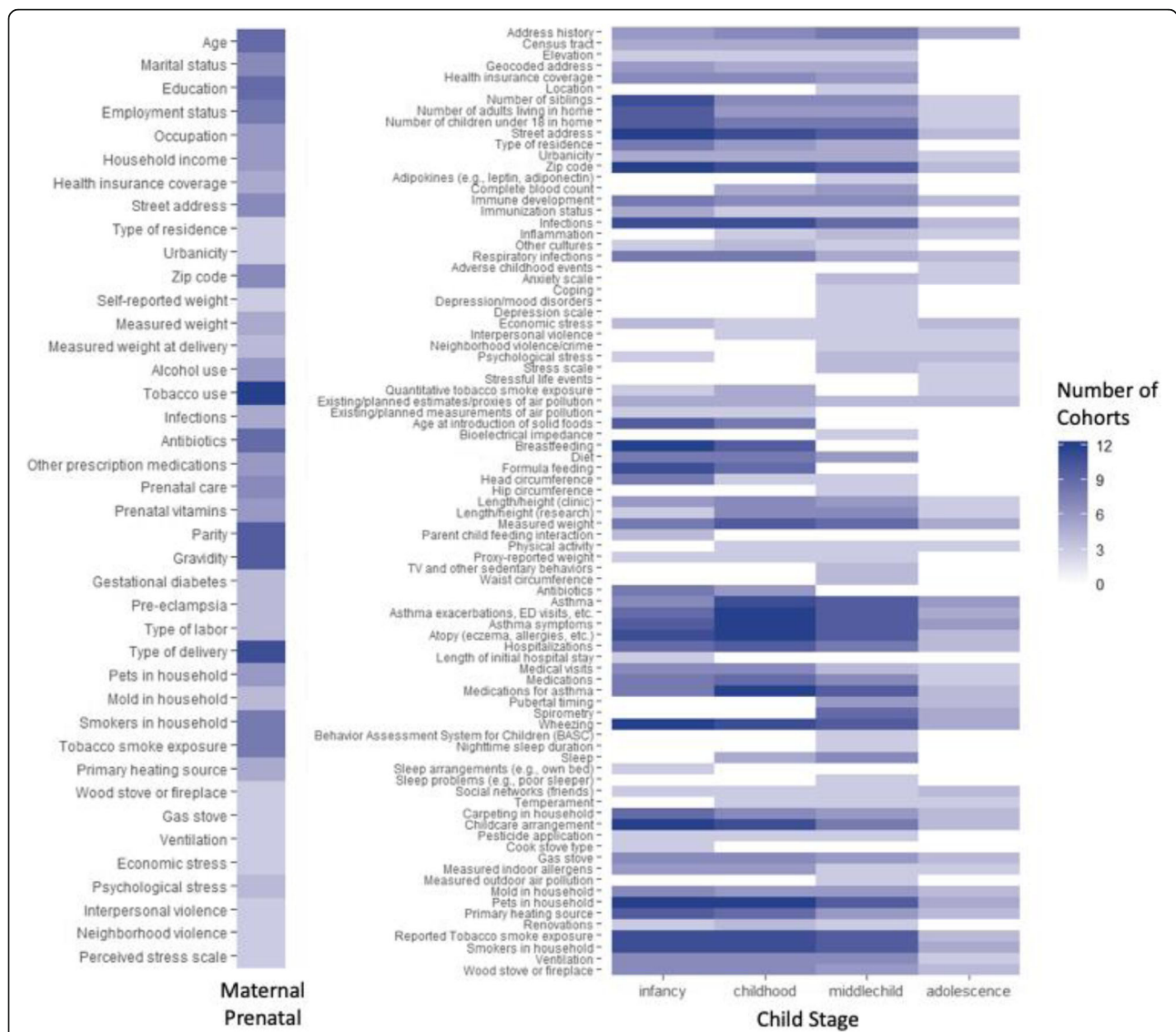
### **Phase III: NewCREW**

The NewCREW phase III protocol that will enroll approximately 400 additional pregnant women and their infants during the prenatal period at four CREW centers is being drafted. Enrollment is scheduled to begin in 2019.

### **Discussion**

The CREW consortium is integrating data from 12 US birth and infant cohort studies to identify prenatal and early life environmental factors that modify the risk for developing allergic diseases and asthma. These cohorts represent a number of US demographics, with good representation of white, African American and Hispanic children and their families. Each cohort has previously identified risk factors for the development of allergies and asthma and provided insights into asthma pathogenesis. Pooling data from individual cohorts to create the CREW cohort presents further opportunities, including the ability to test for risk factors for less common but clinically important outcomes such as severe childhood asthma, combined effects of multiple risk factors, geographical differences and mediating effects. Recruitment for some cohorts started in the 1980s while others are still recruiting today, and the several decade time span of study participation will enable us to test whether the risk factors for childhood asthma have changed over time in concert with marked increases in asthma prevalence. Notably, childhood asthma is a syndrome consisting of multiple asthma phenotypes that are difficult to discern using cross-sectional observations. By assessing a large and diverse cohort over time, the ultimate goal of CREW is to identify phenotypes of asthma and determine whether there are specific relationships between prenatal and early life risk factors that contribute to individual asthma phenotypes.

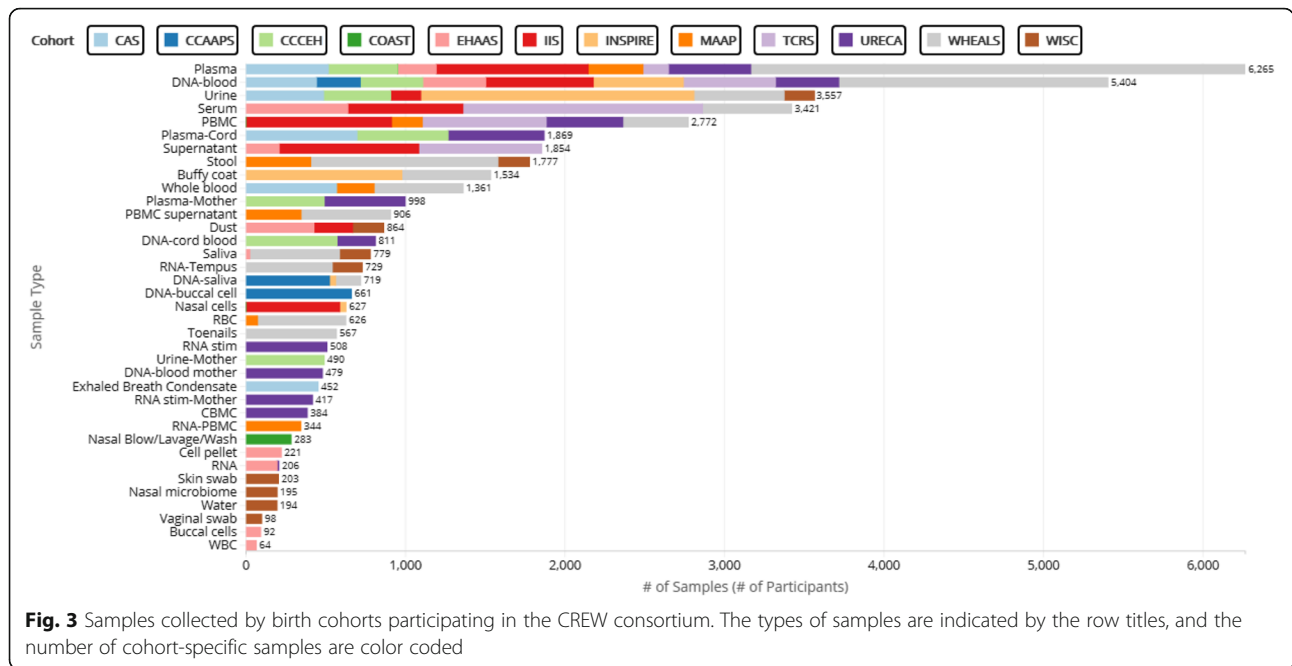
While this is the first US birth cohort consortium, several European consortia have been established to compare study designs and findings, standardize data collection, and perform meta-analyses or pooled analyses from multiple individual cohorts [28]. Examples include the Global Allergy and Asthma European Network (GA<sup>2</sup>LEN, established 2004) which brought



**Fig. 2** Data elements common to multiple cohorts. The number of cohorts that collected each data element is shown according to maternal prenatal exposure (a) and childhood life stage (b); infancy 0–11 mo., early childhood 1–4 years, middle childhood 5–11 years, adolescence 12–18 years)

together 25 research institutes in 16 countries to assess exposure variables, and allergic and respiratory outcome parameters [29]. Several other consortia have formed since 2009 to combine resources of multiple cohorts. These include Environmental Health Risks in European Birth Cohorts (ENRIECO), the Developing a Child Cohort Research Strategy for Europe (CHICOS) consortium, the Mechanisms of the Development of ALLergy (MeDALL), and the Study Team for Early Life Asthma Research (STELAR). CREW represents a unique opportunity to pool data from US cohorts. While productivity from individual cohorts has been high in the US, infrastructures to support collaborative research among allergy and asthma birth

cohorts have been lacking. In recognition of this limitation, the NIAID and NHLBI sponsored the Birth Cohorts in Asthma and Allergic Diseases workshop in 2013 together with investigators from the MeDALL consortium [30]. This workshop developed strategies for harmonization of existing asthma birth cohort data, discussed study designs and standardized methods for new data collection, and identified areas of unmet needs. The conference attendees, which included many current CREW and ECHO investigators, concluded that “establishing criteria for asthma phenotypes that are mechanistically-based and reflect biological entities, including degrees of severity, is the highest priority”. The CREW consortium is designed



to address this knowledge gap by pooling existing data, enriching these data through additional analyses on stored samples, and by enrolling new families into a protocol designed to optimize sampling procedures for systems analysis (genetics and epigenetics, microbial genomics, transcriptomics and metabolomics). Our goal is to analyze these data sets using predictive modeling techniques to identify endotypes of asthma related to specific environmental and personal risk

factors. Pooling of existing data sets within CREW rather than performing meta-analyses should be helpful with respect to ease of analysis, greater statistical power, and availability of advanced statistical techniques that require a single data set.

CREW cohorts have additional opportunities for collaborative research within the U.S. and internationally. As a component of ECHO, standardized questionnaires and data elements have been developed to measure early

**Table 5** Approved CREW analysis proposals

Title	Participating Cohorts											Total	
	CAS	CCAAPS	CCCEH	COAST	EHAAS	IIS	INSPIRE	MAAP	TCRS	URECA	WHEELS		WISC
Childhood research definitions of asthma outcomes	X	X	X	X	X	X			X	X	X		9
The natural history of pediatric asthma	X	X	X	X	X	X			X	X	X		9
Role of CDHR3 and 17q21 genetic variants as determinants of wheezing phenotypes in early life	X	X	X	X	X	X	X	X	X	X	X	X	12
Metabolomic profiling to identify pathways and endotypes of atopic and non-atopic childhood wheeze							X	X					3
Estimating wheezing severity in young children using a latent variable approach with the ISAAC wheezing module*					X		X						2
Vertical transmission of vaginal microbiota to the infant gut								X					3
Fine mapping of the 17q21 region and asthma	X	X	X	X	X		X		X	X	X		10
Regional and individual level socioeconomic characteristics of the CREW consortium	X	X	X	X	X	X	X	X	X	X	X	X	12
RSV and RV respiratory viral wheezing illness and 17q12–21 genotype				X			X		X	X		X	5

\*Published [32]

life exposures that could affect multiple aspects of child health (perinatal outcomes, neurocognitive development, obesity and respiratory health). In addition, ECHO includes an integrated outcome entitled “positive health”, which incorporates measures of physical, mental and social wellbeing. Detailed assessment of these outcomes will enable CREW investigators to test for interactions among these health conditions and allergies and asthma in a large sample size (estimated 50,000 participants). Approximately two thirds of ECHO cohorts have prospectively collected information about respiratory outcomes, and all will complete standardized allergy and respiratory outcome assessments. In addition, about a quarter of the cohorts had focused on asthma prior to ECHO, and the ECHO Airways Outcome Working Group has a rich extant data set of predictors and outcome assessments for allergic diseases and asthma. Furthermore, adaptation and collaborative development of the STELAR Asthma eLab for use in CREW will promote research efforts between US and European cohorts.

Limitations of the CREW Consortium are mainly related to the different study designs and methods used to collect early life data. As a result, data will need to be harmonized to enable pooling for group analyses. Because the majority of the children are school-aged or older, most of the early life data in CREW already have been collected. Nonetheless, some opportunities exist for enriching data sets with the use of retrospective questionnaires. For example, retrospective analyses of pet ownership during early childhood correlates well with prospectively collected data [31]. In addition, use of time-adjusted geospatial technology will enable enrichment of early life exposure data related to pollutants and a wide range of social and economic variables.

In conclusion, CREW represents the first large scale collaborative study involving U.S. birth cohorts focused on allergic diseases and asthma in children. Pooling of extant data in CREW and using standardized outcomes will provide many opportunities to identify early life environmental exposures that are linked to specific endotypes of asthma. As part of the ECHO program, CREW will also have access to data from a much larger sample size and be able to ask questions about the relationship between asthma and other health outcomes in children.

#### Abbreviations

BBC: Biostatistics/Bioinformatics core; CAS: Childhood Asthma Study; CHICOS: Developing a Child Cohort Research Strategy for Europe; COAST: Childhood Origins of Asthma Study; CREW: Children’s Respiratory and Environmental Workgroup; CRF, eCRF: Case report form, electronic case report form; ECHO: Environmental influences on Child Health Outcomes; EHAAS: The Epidemiology of Home Allergens and Asthma Study; eNO: Exhaled nitric oxide; ENRIECO: Environmental Health Risks in European Birth Cohorts; FHIR: Fast Health Interoperability Resources; GA<sup>2</sup>LEN: Global Allergy and Asthma European Network; HIPAA: Health Insurance Portability and Accountability Act; INSPIRE: Infant Susceptibility to Pulmonary Infections and Asthma following RSV Exposure; ISAAC: International Study of Asthma

and Allergies in Children; MAAP: Microbes, Allergy, Asthma and Pets; MDI: Metered dose inhaler; MeDALL: Mechanisms of the Development of Allergy; NHLBI: National Heart Lung and Blood Institute; NIAID: National Institute for Allergy and Infectious Disease; PHI: Protected health information; STELAR: Study Team for Early Life Asthma Research; UW: University of Wisconsin; WHEALS: Wayne County Health, Environment, Allergy and Asthma Longitudinal Study; WISC: Wisconsin Infant Study Cohort

#### Acknowledgements

The Children’s Environment and Respiratory Workgroup (CREW) is a collaboration of the following cohorts, institutions, investigators and staff (principal investigators are indicated by an asterisk): Columbia Center for Children’s Environmental Health (CCCEH): Rachel Miller\*, Howard Andrews, Julie Herbstman, Lori Hoepner, Frederica Perera, Matthew Perzanowski, Xinhua Liu, Judyth Ramirez, Janelle Rivera, Deliang Tang, Kylie Wheelock, Jaqueline Jezioro. Microbes, Allergy, Asthma and Pets (MAAP): Henry Ford Health System, Detroit, MI: E Zoratti\*, CC Johnson, A Sitarik, S Havstad, K Woodcroft, A Levin, G Wegienka, B Davidson, S Finazzo, K Bobbitt, E Mann, S Bellemore, S Zhang, A Wahlman, K Jones; Augusta University, Augusta, GA: D Ownby; University of Michigan, Ann Arbor, MI: N Lukacs; University of California, San Francisco, CA: S Lynch, H Boushey. Tucson Children’s Respiratory Study (TCRS): Anne L. Wright\*, Fernando D. Martinez\*, Wayne Morgan, Debra A. Stern, Dean Billheimer, Brian Hallmark, Paloma Beamer, Nathan Lothrop, Lydia De La Ossa, Silvia Lopez, Marilyn Halonen, Amber Spangenberg, David Spies. Infant Immune Study (IS): Anne L. Wright\*, Fernando D. Martinez\*, Wayne Morgan, Debra A. Stern, Dean Billheimer, Brian Hallmark, Paloma Beamer, Nathan Lothrop, Heidi Erickson, Marilyn Halonen, Amber Spangenberg, David Spies. Wisconsin Infant Study Cohort (WISC): Marshfield Clinic Research Institute: Casper Bendixsen, Kathrine Barnes, Tara Johnson, Terry Foss, Elizabeth Armagost, Vicki Moon, Tammy Koepel, Erin Donnerbauer, Wayne Frome, Brent Olson. University of Wisconsin-Madison: Christine M. Seroogy\*, James E. Gern\*, Samantha Fye, Irene Ong, Deborah Chasman, Rose Vrtis, Heather Floerke, Amy Dresen, Yury Bochkov, Tressa Pappas, Kristine Grindle, Michael Evans, Ronald Gangnon. Childhood Origins of Asthma Study (COAST): Robert F. Lemanske, Jr.\*, Daniel J. Jackson\*, James E. Gern, Carole Ober, Anne Marie Singh, Ronald E. Gangnon, Michael D. Evans, Victoria Rajamanickam, Christopher Tisler, Lisa Salazar, Susan Doyle, Yury Bochkov, Rebecca Brockman-Schneider, Rose Vrtis, Kristine Grindle, Tressa Pappas, Elizabeth Anderson, Kathy Roberg, Kirsten Carlson-Dakes, Mark DeVries, Douglas DaSilva, Ronald Sorkness, Lance Mikus, Julia Bach. Urban Environment and Childhood Asthma Study (URECA): Johns Hopkins University, Baltimore, MD: R Wood\*, E Matsui, H Lederman, F Witter, S Leimenstoll, D Scott, M Cootauco, P Jones. Boston University School of Medicine, Boston, MA: G O’Connor\*, W Cruikshank, M Sandel, A Lee-Parritz, C Jordan, E Gjerasi, P Price-Johnson, L Gagalis, L Wang, N Gonzalez, M Tuzova. Harvard Medical School, Boston, MA – D Gold, R Wright. Columbia University, New York, NY: M Kattan\*, C Lamm, N Whitney, P Yaniv, M Pierce, Jaqueline Jezioro. Mount Sinai School of Medicine, New York, NY: H Sampson, R Sperling, N Rivers. Washington University School of Medicine, St Louis, MO: G Bloomberg\*, L Bacharier\*, Y Sadovsky, E Tesson, C Koerkenmeier, R Sharp, K Ray, J Durrange, I Bauer, A Freie, V Morgan. Statistical and Clinical Coordinating Center - Rho, Inc., Chapel Hill, NC: C Visness\*, P Zook, M Yaeger, J Martin, A Calatroni, K Jaffee, W Taylor, R Budrevich, H Mitchell. Scientific Coordination and Administrative Center, University of Wisconsin, Madison, WI: W Busse\*, J Gern\*, P Heinritz, C Sorkness, K Hernandez, Y. Bochkov, K Grindle, A Dresen, T Pappas, M. Renneberg, B. Stoffel. National Institute of Allergy and Infectious Diseases, Bethesda, MD: P Gergen, A Togias, E Smartt, K Thompson. Cincinnati Childhood Allergy and Air Pollution Study (CCAAPS): Gurgit K. Khurana Hershey\*, Patrick H. Ryan\*, Jocelyn M. Biagini Myers\*, Grace K. LeMasters\*, Kristi Curtsinger, Liza Murrison\*, Lisa Martin, Jeffrey W. Burkle, Christopher Wolfe, Zachary Flege, David Morgan, Kristina Keidel, Krista Tensing, Taylor Groeschel.



The Infant Susceptibility to Pulmonary Infections and Asthma Following RSV Exposure (INSPIRE):

Vanderbilt University Medical Center, Nashville, TN: Andrew Abreo MD+, Niek Achten MD+, Alyssa Bednarek BS, Steven M. Brunwasser PhD\*, Teresa M. Chipps BS, Alexandra Connolly BS, Kaitlin Costello BA, Marian Dorst BA+, William D. Dupont PhD\*, Roxanne Filardo-Collins RN, BSN, Rebecca Gammell BA+, Tebeb Gebretsadiq MPH\*, Kayla Goodman, Tina V. Hartert MD, MPH\*\*, Emma Larkin PhD+, Jessica Levine RN, Zhouwen Liu MS, Christian Lynch MPH, Megan Mccollum MS, Patricia Minton RN, AE-C+, Paul E. Moore MD, Sarah Osmundson MD\*, R. Stokes Peebles MD, Christian Rosas-Salazar MD, MPH, Cosby Stone Jr. MD, MPH+, Theresa Rogers RN+, Pat Russell RN, BSN, Kedir Turi PhD\*, Kim B. Woodward MSN, APRN, CPNP-PC, Pingsheng Wu PhD. Vanderbilt Technologies for Advanced Genomics, Nashville, TN: Suman R. Das PhD, Seesandra Rajagopala PhD, Meghan H. Shilts MHS, MS. Vanderbilt Infectious Diseases, Nashville, TN: James D. Chappell MD. J. Craig Venter Institute, Rockville, MD: Rebecca Halpin BS, Karen E. Nelson PhD, Seth Schobel PhD, Jyoti Shankar MBBS, PhD, Andrey Tovchigrechko PhD, Shibu Yooseph PhD.

University of Wisconsin, Madison, WI: Zhengzheng Tang PhD.

Emory University, Atlanta, GA: Larry J. Anderson MD, Tatiana Chirkova PhD, Samadhan J. Jadhao PhD, Martin L. Moore PhD.

The Epidemiology of Home Allergens and Asthma Study (EHAAS): Diane R. Gold, Soma Datta, Sharon O'Toole, Conner Fleurat, Leanna Farnham.

Wayne County Health, Environment, Allergy and Asthma Longitudinal Study (WHEALS): CC Johnson\*, G Wegienka, S Havstad, E Zoratti, A Cassidy-Bushrow, A Levin, H Kim, K Woodcroft, A Sitarik, C Joseph, LK Williams, C Barone, K Bobbitt, S Zhang, J Campbell, K Bourgeois, M Aubuchon, J Ezell, K Jones; D Ownby.

Childhood Allergy Study (CAS): D Ownby,\* CC Johnson,\* C Joseph, E Zoratti,\* G Wegienka, S Havstad, K Woodcroft, E Peterson, S Hensley Alford, J McCullough, C Strauchman Boyer, S Blocki, G Birg, N Akkerman, K Wells, S Zhang, C Nicholas, A Jones, G Stouffer.

Administrative Center, University of Wisconsin-Madison: James E. Gern,\* Gina Crisafi, Dorothy Floerke, Rick Kelley.

Coordinating Center, Rho, Inc. Federal Systems Division: Cynthia Visness\*, Melissa Yaeger, Samara Dixon, Kathy Collier.

Biomedical Informatics and Biostatistical Core (BBC), University of Wisconsin-Madison: Umberto Tachinardi\*, Mark Craven\*, Eneida Mendonca, Lisa Gress, Adam Nunez, Laura Ladick.

University of Manchester: Philip Couch, Camille Johnson, John Ainsworth, Victoria Turner.

Imperial College, London UK: Adnan Custovic.

Genetics Core, University of Chicago, Chicago, IL: Carole Ober\*, Daniel Nicolae. Geospatial Core, Harvard School of Public Health, Boston, MA: Diane Gold\*, Heike Gibson, Brent Coull, Antonella Zanobetti, Weeberb Requia, Joel Schwartz, Itai Kloog, Qian Di, Peter James, Marcia Jimenez Pescador, Jaime Hart.

Microbiome Core, University of California-San Francisco: Susan Lynch\*, Kathrine McCauley.

This manuscript is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

#### Authors' contributions

JEG, DJJ, RFL, CMS, UT, MC, PR, GTO, DRG, RM, MK, CCJ, DO, EMZ, RAW, CMV, FM, AW, SL, CO, GKKH, PR, TH and LBB made substantial contributions to the conception or design of the work reported, and participated in study design and the acquisition of reported data. SYH, CMH, WH, CMV analyzed the reported data. JEG wrote the initial draft of the manuscript. All authors reviewed and/or critically revised the manuscript for important intellectual content and provide final approval of the version to be published. All authors agree both to be personally accountable for the author's own contributions and to ensure that questions related to the accuracy or integrity of any part of the work, even ones in which the author was not personally involved, are appropriately investigated, resolved, and the resolution documented in the literature.

#### Funding

CREW is funded by HHS/NIH grant 5UG3OD023282. Additional support was provided by individual cohorts' grants/contracts:

Columbia University: P01ES09600, R01 ES008977, P30ES09089, R01 ES013163, R827027.

Tucson Children's Respiratory Study (TCRS): NHLBI 132523.

Infant Immune Study (IIS): HL-56177.

Childhood Origins of Asthma Study (COAST): P01 HL070831, U10 HL064305, R01 HL061879.

Urban Environment and Childhood Asthma Study (URECA): NO1-AI-25496, NO1-AI-25482, HHSN272200900052C, HHSN272201000052I, NCR/NIH RR00052, M01RR00533, 1UL1RR025771, M01RR00071, 1UL1RR024156, UL1TR001079, 5UL1RR024992-02, NCATS/NIH UL1TR000040.

Cincinnati Childhood Allergy and Air Pollution Study (CCAAPS): R01 ES11170, R01 ES019890.

The Epidemiology of Home Allergens and Asthma Study (EHAAS): R01 AI035786.

Wayne County Health, Environment, Allergy and Asthma Longitudinal Study (WHEALS): R01 AI050681, R56 AI050681, R01 AI061774, R21 AI059415, K01 AI070606, R21 AI069271, R01 HL113010, R21 ES022321, P01 AI089473, R21 AI080066, R01 AI110450, R01 HD082147 and the Fund for Henry Ford Health System.

Childhood Allergy Study (CAS): R01 AI024156, R03 HL067427, R01 AI051598, Blue Cross Foundation Johnson, and the Fund for Henry Ford Hospital.

Microbes, Allergy, Asthma and Pets (MAAP): P01 AI089473 and Fund for Henry Ford Hospital.

Infant Susceptibility to Pulmonary Infections and Asthma following RSV Exposure (INSPIRE): NIH/NIAID U19 AI 095227, NIH/NCATS UL1 TR 002243, NIH/NIAID K24 AI 077930, NIH/NHLBI R21 HD 087864, NIH/NHLBI X01 HL 134583.

Wisconsin Infant Study Cohort (WISC): U19 AI104317, NCATS UL1TR000427, the charitable donors to the Marshfield Clinic Health System Foundation, and the Upper Midwest Agricultural Safety and Health Center (UMASH) U54 OH010170.

RTI International, Research Triangle Park, North Carolina, USA: Funding from NIH for the Environmental influences of Child Health Outcomes (ECHO) Cohort Data Analysis Center (U24OD023382).

#### Availability of data and materials

The datasets used during the current study are available from the corresponding author on reasonable request.

#### Ethics approval and consent to participate

A data sharing protocol and a data use agreement were approved by the local institutional review boards for each participating cohort. All parents or legal guardians provided written, informed consent for the child's original cohort participation.

#### Consent for publication

Not applicable.

#### Competing interests

All authors report grants from NIH during the conduct of study. J.E. Gern reports personal fees from PREP Biopharm Inc., Regeneron, MedImmune, and Ena Pharmaceuticals, as well as personal fees and stock options from Meissa Vaccines Inc. outside the submitted work. In addition, J.E. Gern has a patent "Methods of Propagating Rhinovirus C in Previously Unsusceptible Cell Lines" issued, and a patent "Adapted Rhinovirus C" pending. D. J Jackson reports grants from GlaxoSmithKline, as well as personal fees from Novartis, Vifor Pharma, Pfizer, Boehringer Ingelheim, and Commense outside the submitted work. L.B. Bacharier reports personal fees from GlaxoSmithKline, Genentech/Novartis, Merck, DBV Technologies, Teva, Boehringer Ingelheim Sanofi/Regeneron, Vectura, Circassia, and AstraZeneca outside the submitted work. R.A. Wood reports personal fees from AAAAI in the form of a presidential stipend and fees for work as a JACI Associate Editor, as well as royalties from Up To Date, outside the selected work. In addition, R.A. Wood grants from Aimmune, DBV, Astellas, Sanofi, Regeneron, and HAL Allergy outside the selected work. S. Lynch reports personal fees from Siolta Therapeutics outside the submitted work. R.F. Lemanske reports grants as part of the Clinical and Translational Science Award from NIH, Childhood Origins of Asthma (COAST), and AsthmaNet outside the submitted work. R.F. Lemanske also reports personal fees from LSU, University of Kentucky, and Thermo Fischer outside the submitted work. In addition, R.F. Lemanske reports royalties from Elsevier and Up To Date outside the submitted work. T. Hartert reports grants from the World Health Organization outside the submitted work. F. Martinez reports personal fees for consultancy from Commense Inc. and Copeval outside the submitted work. In addition, F. Martinez reports



grants from Johnson & Johnson and Vifor Pharma outside the submitted work. G.T. O'Connor reports personal fees from AstraZeneca for consultancy, as well as a grant from Janssen Pharmaceuticals outside the submitted work. M. Kattan reports personal fees from Novartis Pharma outside the submitted work. M. Craven, D.R. Gold, U. Tachinardi, C.M. Seroogy, E.M. Zoratti, P. Ryan, D. Ownby, C. Ober, R. Miller, S.Y. Hwang, W. Huggins, C.C. Johnson, G.K.K. Hershey, C.M. Visness, A. Wright, and C.M. Hamilton have nothing to disclose outside the submitted work.

#### Author details

<sup>1</sup>University of Wisconsin School of Medicine and Public Health, Madison, WI 53706, USA. <sup>2</sup>RTI International, East Cornwallis Road, Post Office Box 12194, Raleigh, Research Triangle Park, NC 27709-2194, USA. <sup>3</sup>Boston University School of Medicine, 72 E Concord St, Boston, MA 02118, USA. <sup>4</sup>Channing Laboratory, Brigham and Women's Hospital, Boston, MA 02115, USA. <sup>5</sup>Columbia University, Vagelos College of Physicians and Surgeons, New York, NY 10032, USA. <sup>6</sup>Henry Ford Health System, Detroit, MI 48202, USA. <sup>7</sup>Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. <sup>8</sup>Rho Federal Systems Division, Inc., Chapel Hill, NC 27517, USA. <sup>9</sup>University of Arizona, Tucson, AZ 85721, USA. <sup>10</sup>University of California, San Francisco, CA 94143, USA. <sup>11</sup>University of Chicago, Chicago, IL 60637, USA. <sup>12</sup>University of Cincinnati, Cincinnati, OH 45220, USA. <sup>13</sup>Vanderbilt University School of Medicine, Nashville, TN 37232, USA. <sup>14</sup>Washington University School of Medicine, St. Louis, MO 63110, USA. <sup>15</sup>Department of Pediatrics, School of Medicine and Public Health, University of Wisconsin-Madison, Clinical Science Center-K4/918, 600 Highland Ave, Madison, WI 53792-9988, USA.

Received: 1 February 2019 Accepted: 3 June 2019

Published online: 10 June 2019

#### References

- NIH ECHO Program. ECHO, Environmental influences on Child Health Outcomes; A program supported by the NIH 2017 [Available from: <http://echochildren.org/>].
- Lemanske RF Jr. The childhood origins of asthma (COAST) study. *Pediatr Allergy Immunol*. 2002;13(Suppl 15):38–43.
- Taussig LM, Wright AL, Holberg CJ, Halonen M, Morgan WJ, Martinez FD. Tucson children's respiratory study: 1980 to present. *J Allergy Clin Immunol*. 2003;111(4):661–75.
- Gold DR, Burge HA, Carey V, Milton DK, Platts-Mills T, Weiss ST. Predictors of repeated wheeze in the first year of life: the relative roles of cockroach, birth weight, acute lower respiratory illness, and maternal smoking. *Am J Respir Crit Care Med*. 1999;160(1):227–36.
- Oddy WH, Halonen M, Martinez FD, Lohman IC, Stern DA, Kurzius-Spencer M, et al. TGF-beta in human milk is associated with wheeze in infancy. *J Allergy Clin Immunol*. 2003;112(4):723–8.
- Taussig LM, Wright AL, Morgan WJ, Harrison HR, Ray CG. The Tucson Children's respiratory study. I. Design and implementation of a prospective study of acute and chronic respiratory illness in children. *Am J Epidemiol*. 1989;129(6):1219–31.
- Donohue KM, Al-alem U, Perzanowski MS, Chew GL, Johnson A, Divjan A, et al. Anti-cockroach and anti-mouse IgE are associated with early wheeze and atopy in an inner-city birth cohort. *J Allergy Clin Immunol*. 2008;122(5):914–20.
- Perera FP, Illman SM, Kinney PL, Whyatt RM, Kelvin EA, Shepard P, et al. The challenge of preventing environmentally related disease in young children: community-based research in new York City. *Environ Health Perspect*. 2002;110(2):197–204.
- LeMasters GK, Wilson K, Levin L, Biagini J, Ryan P, Lockey JE, et al. High prevalence of aeroallergen sensitization among infants of atopic parents. *J Pediatr*. 2006;149(4):505–11.
- Ryan PH, LeMasters G, Biagini J, Bernstein D, Grinshpun SA, Shukla R, et al. Is it traffic type, volume, or distance? Wheezing in infants living near truck and bus traffic. *J Allergy Clin Immunol*. 2005;116(2):279–84.
- Neaville WA, Tisler C, Anklam K, Gilbertson-White S, Hamilton R, Adler K, et al. Developmental cytokine response profiles and the clinical and immunologic expression of atopy in infancy. *Jaci*, vol. 112; 2003. p. 740–6.
- Gern JE, Visness CM, Gergen PJ, Wood RA, Bloomberg GR, O'Connor GT, et al. The urban environment and childhood asthma (URECA) birth cohort study: design, methods, and study population. *BMC Pulm Med*. 2009;9:17.
- Barlow-Anacker A, Bochkov Y, Gern J, Seroogy CM. Neonatal immune response to rhinovirus A16 has diminished dendritic cell function and increased B cell activation. *PLoS One*. 2017;12(10):e0180664.
- Ownby DR, Johnson CC, Peterson EL. Exposure to dogs and cats in the first year of life and risk of allergic sensitization at 6 to 7 years of age. *JAMA*. 2002;288(8):963–72.
- Ownby DR, Johnson CC, Peterson EL. Maternal smoking does not influence cord serum IgE or IgD concentrations. *J Allergy Clin Immunol*. 1991;88(4):555–60.
- Aichbhaumik N, Zoratti EM, Strickler R, Wegienka G, Ownby DR, Havstad S, et al. Prenatal exposure to household pets influences fetal immunoglobulin E production. *Clin Exp Allergy*. 2008;38(11):1787–94.
- Havstad S, Johnson CC, Kim H, Levin AM, Zoratti EM, Joseph CL, et al. Atopic phenotypes identified with latent class analyses at age 2 years. *J Allergy Clin Immunol*. 2014;134(3):722–7 e2.
- Brokamp C, Wolfe C, Lingren T, Harley J, Ryan P. Decentralized and reproducible geocoding and characterization of community and environmental exposures for multisite studies. *J Am Med Inform Assoc*. 2017.
- Strachan D, Sibbald B, Weiland S, Ait-Khaled N, Anabwani G, Anderson HR, et al. Worldwide variations in prevalence of symptoms of allergic rhinoconjunctivitis in children: the international study of asthma and allergies in childhood (ISAAC). *Pediatric Allergy & Immunology*. 1997;8(4):161–76.
- Charman CR, Venn AJ, Ravenscroft JC, Williams HC. Translating patient-oriented eczema measure (POEM) scores into clinical practice by suggesting severity strata derived using anchor-based methods. *Br J Dermatol*. 2013;169(6):1326–32.
- Ridd MJ, Gaunt DM, Guy RH, Redmond NM, Garfield K, Hollinghurst S, et al. Comparison of patient (POEM), observer (EASI, SASSAD, TIS) and corneometry measures of emollient effectiveness in children with eczema: findings from the COMET feasibility trial. *Br J Dermatol*. 2018.
- Lemanske RF Jr, Jackson DJ, Gangnon RE, Evans MD, Li Z, Shult PA, et al. Rhinovirus illnesses during infancy predict subsequent childhood wheezing. *J Allergy Clin Immunol*. 2005;116(3):571–7.
- Kloepfer KM, Olenec JP, Lee WM, Liu G, Vrtis RF, Roberg KA, et al. Increased H1N1 infection rate in children with asthma. *Am J Respir Crit Care Med*. 2012;185(12):1275–9.
- Silkoff PE, McClean PA, Slutsky AS, Furlott HG, Hoffstein E, Wakita S, et al. Marked flow-dependence of exhaled nitric oxide using a new technique to exclude nasal nitric oxide. *Am J Respir Crit Care Med*. 1997;155(1):260–7.
- American Thoracic Society. Recommendations for standardized procedures for the on-line and off-line measurement of exhaled lower respiratory nitric oxide and nasal nitric oxide in adults and children-1999. This official statement of the American Thoracic Society was adopted by the ATS Board of directors, July 1999. *Am J Respir Crit Care Med*. 1999;160(6):2104–17.
- Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, et al. Standardisation of spirometry. *Eur Respir J*. 2005;26(2):319–38.
- Custovic A, Ainsworth J, Arshad H, Bishop C, Buchan I, Cullinan P, et al. The study team for early life asthma research (STELAR) consortium 'Asthma e-lab': team science bringing data, methods and investigators together. *Thorax*. 2015;70(8):799–801.
- Bousquet J, Anto J, Sunyer J, Nieuwenhuijsen M, Vrijheid M, Keil T, et al. Pooling birth cohorts in allergy and asthma: European Union-funded initiatives - a MeDALL, CHICOS, ENRIECO, and GA (2) LEN joint paper. *Int Arch Allergy Immunol*. 2013;161(1):1–10.
- Frew AJ. GA2LEN—the global allergy and asthma European network. *Clin Exp Allergy*. 2005;35(2):122–5.
- Bousquet J, Gern JE, Martinez FD, Anto JM, Johnson CC, Holt PG, et al. Birth cohorts in asthma and allergic diseases: report of a NIAID/NHLBI/MeDALL joint workshop. *J Allergy Clin Immunol*. 2014;133(6):1535–46.
- Nicholas C, Wegienka G, Havstad S, Ownby D, Johnson CC, Zoratti E. How accurately do young adults recall childhood pets? A validation study. *Am J Epidemiol*. 2009;170(3):388–92.
- Brunwasser SM, Gebretsadik T, Gold DR, Turi KN, Stone CA Jr, Datta S, et al. A new model of wheezing severity in young children using the validated ISAAC wheezing module: a latent variable approach with validation in independent cohorts. *PLoS One*. 2018;13(4):e0194739.

#### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.