PEDIATRIC ALLERGY, IMMUNOLOGY, AND PULMONOLOGY
Volume 25, Number 2, 2012

DOI: 10.1089/ped.2011.0116

RESEARCH METHOD

The Greater Cincinnati Pediatric Clinic Repository: A Novel Framework for Childhood Asthma and Allergy Research

Melinda Butsch Kovacic, Ph.D.,[1,2] Jocelyn M. Biagini Myers, Ph.D.,[1] Mark Lindsey, B.S.,[1] Tia Patterson, M.H.A.,[1] Sharon Sauter, M.S.,[1] Mark B. Ericksen, B.S.,[1] Patrick Ryan, Ph.D.,[3] Amal Assa'ad, M.D.,[4] Michelle Lierl, M.D.,[4] Thomas Fischer, M.D.,[4] Carolyn Kercsmar, M.D.,[5] Karen McDowell, M.D.,[5] Anne W. Lucky, M.D.,[6] Anita P. Sheth, M.D.,[6] Andrew D. Hershey, M.D.,[7] Richard M. Ruddy, M.D.,[8] Marc E. Rothenberg, M.D.,[4] and Gurjit K. Khurana Hershey, M.D., Ph.D.[1]

Background: Allergic disorders, including asthma, allergic rhinitis, atopic dermatitis, eosinophilic esophagitis, and food allergy, are a major global health burden. The study and management of allergic disorders is complicated by the considerable heterogeneity in both the presentation and natural history of these disorders. Biorepositories serve as an excellent source of data and biospecimens for delineating subphenotypes of allergic disorders, but such resources are lacking.

Methods: In order to define subphenotypes of allergic disease accurately, we established an infrastructure to link and efficiently utilize clinical and epidemiologic data with biospecimens into a single biorepository called the Greater Cincinnati Pediatric Clinic Repository (GCPCR). Children with allergic disorders as well as healthy controls are followed longitudinally at hospital clinic, emergency department, and inpatient visits. Subjects' asthma, allergy, and skin symptoms; past medical, family, social, diet, and environmental histories; physical activity; medication adherence; perceived quality of life; and demographics are ascertained. DNA is collected from all participants, and other biospecimens such as blood, hair, and nasal epithelial cells are collected on a subset.

Results: To date, the GCPCR has 6,317 predominantly Caucasian and African American participants, and 93% have banked DNA. This large sample size supports adequately powered genetic, epidemiologic, environmental, and health disparities studies of childhood allergic diseases.

Conclusions: The GCPCR is a unique biorepository that is continuously evaluated and refined to achieve and maintain rigorous clinical phenotype and biological data. Development of similar disease-specific repositories using common data elements is necessary to enable studies across multiple populations of comprehensively phenotyped patients.

Introduction

Allergic disorders are a major global health burden affecting up to 40% of the world population. The biologic underpinnings of these disorders, including asthma, allergic rhinitis, atopic dermatitis, eosinophilic esophagitis, and food allergy, involve common mechanistic pathways including innate and adaptive immunity, as well as mucosal barrier function. Individuals suffering from asthma, for example, often share similar clinical symptoms, but the disease is heterogeneous in terms of phenotypes and natural history.[1,2] This heterogeneity contributes to the difficulty in both studying and treating the disease. This is especially evident in children, where nearly two-thirds who currently have asthma reported at least one attack in the previous 12 months.[3] The National Asthma Education and Prevention Program recently issued its third Expert Panel Report outlining guidelines for the diagnosis and management of asthma, and emphasizes the importance of individualizing treatment plans for patients.[4] In order to design and develop therapies, the phenotypes of these patients need to be better defined and the mechanisms by which specific genes

Divisions of [1]Asthma Research, [2]Biostatistics and Epidemiology, [4]Allergy and Immunology, [5]Pulmonary Medicine, [6]Dermatology, [7]Neurology, and [8]Emergency Medicine, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio.
[3]Department of Environmental Health, University of Cincinnati, Cincinnati, Ohio.

and environmental factors confer their impact need to be elucidated. Delineation of phenotypes of asthma and other allergic disorders has been hindered by the lack of well-defined biorepositories of patients with these conditions. In order to meet this challenge, we established an infrastructure to link and efficiently utilize clinical and epidemiologic data with biospecimens into a single hospital clinic-based allergic disease biorepository appropriately called the Greater Cincinnati Pediatric Clinic Repository (GCPCR). The GCPCR currently serves as a distinct and valuable resource to bolster adequately powered clinical, epidemiologic, genetic, environmental, and, more recently, health disparities studies of asthma and allergic diseases in children and adolescents.

Materials and Methods

Early baseline survey

Recruitment into the GCPCR began in November, 2003 with the purpose of serving as a resource for primarily genetic research studies. At that time, procedures and data collection tools were implemented that would both standardize clinical care and facilitate scientific research of asthmatic and allergic children. Critical to integrating clinic flow with research was the use of a single dual-purpose questionnaire. To this end, the Initial Clinic Evaluation (ICE) was developed for children visiting the allergy/immunology clinics for the first time, and also served as a baseline research data tool for patients giving informed consent (Supplemental Table 1). Before its implementation, the ICE was evaluated and found to be acceptable and useful by families, clinical staff, and researchers alike, as it captured patients' asthma and allergy symptoms, past medical, family, social, and environmental histories. Importantly, the respiratory history questions were modified from the ISAAC Questionnaire,[5,6] and the symptom assessments were modified from the *Guidelines for the Diagnosis and Management of Asthma*.[7] A nurse or physician reviewed the completed questionnaire with the patient's parent/guardian as part of their medical evaluation. The questionnaire then became a component of each patient's Cincinnati Children's Hospital Medical Center (CCHMC) medical record. Clinical data as well as any samples collected by study coordinators during the visit from those patients providing consent were entered into the repository.

Informed consent

To facilitate consent, a coordinator was integrated into the clinics and approached patients during their clinical visit with no preference shown for age, gender, ethnicity, race, or disease outcome. If patients indicated interest in participation, the coordinator would provide a page-by-page explanation. Parents/guardians were then asked independently to sign a written consent form approved by the Cincinnati Children's Hospital Institutional Review Board. They were asked to (1) allow both their clinical and questionnaire-based information to be entered into a database for research purposes; (2) allow collection of buccal cells or saliva for nucleic acid isolation and use for genetic analyses; and (3) be contacted in the future to consider participation in future related or unrelated studies. All children were asked to assent, and children over the age of 8 years were asked to sign a written consent. A copy of the consent was given to the family, another copy was placed in the child's medical chart, and a third copy was kept by the GCPCR staff.

Expansion to a repository

In 2006, funding by the NIH became available to evaluate the genetic predisposition of asthma and allergic children. At this time, recruitment was expanded beyond the CCHMC Main Campus to include four other CCHMC allergy/immunology and pulmonary satellite sites with the goal of identifying additional children. Children diagnosed with atopic dermatitis were also recruited from CCHMC dermatology clinics (2006) and children with asthma were recruited from both the Emergency Department (2007) and the CCHMC Asthma Center (2008). At this time, the repository was established as the GCPCR. The Institutional Review Board (IRB) protocol and consent form were amended along the way to include collection of hair, blood, and nasal epithelial cells, as well as skin tape samples and transepidermal water loss measurements.

The addition of control children

In 2006, a subset of nearly 300 nonallergic children were enrolled, predominantly from the CCHMC outpatient dental clinic, to evaluate the associations of race and quality of life measured by the Pediatric Allergic Disease Quality of Life Questionnaire (PADQLQ), with allergic sensitization among healthy children with no family history of sensitization.[8] In this analysis, allergic disease-related quality of life was measured by the Pediatric Quality of Life Questionnaire (PADQLQ; calculated from 26 questions with a scale ranging from 0 to 6, with 0 being no symptoms and 6 being the worst allergy symptoms).[9] Importantly, no significant differences in quality of life scores were found between sensitized and nonsensitized subjects. In 2009, the GCPCR began actively enrolling nonallergic participants to meet the demands of ongoing genetic and exposure-based studies. To ensure that controls reflect the entire Cincinnati area, children were also recruited from the community at large using paper and online advertising media, as well as from outpatient clinics at CCHMC (i.e., plastics, headache, dental, and orthopedic clinics). Enrollment of control children included strict criteria to maximize differences compared to cases. Children were enrolled as nonallergic controls if they reported not having any personal or family history of asthma, and not having a personal history of hay fever, environmental allergies, food allergies, atopic dermatitis, or eczema. Many of the consented control children were asked to undergo environmental allergy skin prick testing (SPT) to a panel of 11 relevant environmental antigens indigenous to the Ohio valley (A.L.K. Laboratories Inc.) to determine sensitization. As before, parents of controls were asked to help their child complete a one-page modified version of the PADQLQ to rule out that they have experienced allergy-like symptoms in the past week.[9]

Current recruiting, sample handling, and data management process

In 2008, CCHMC implemented an interactive electronic medical record (EMR) system based on software developed by Epic Systems Corporation (Verona, WI) that has made recruitment and management of repository participants

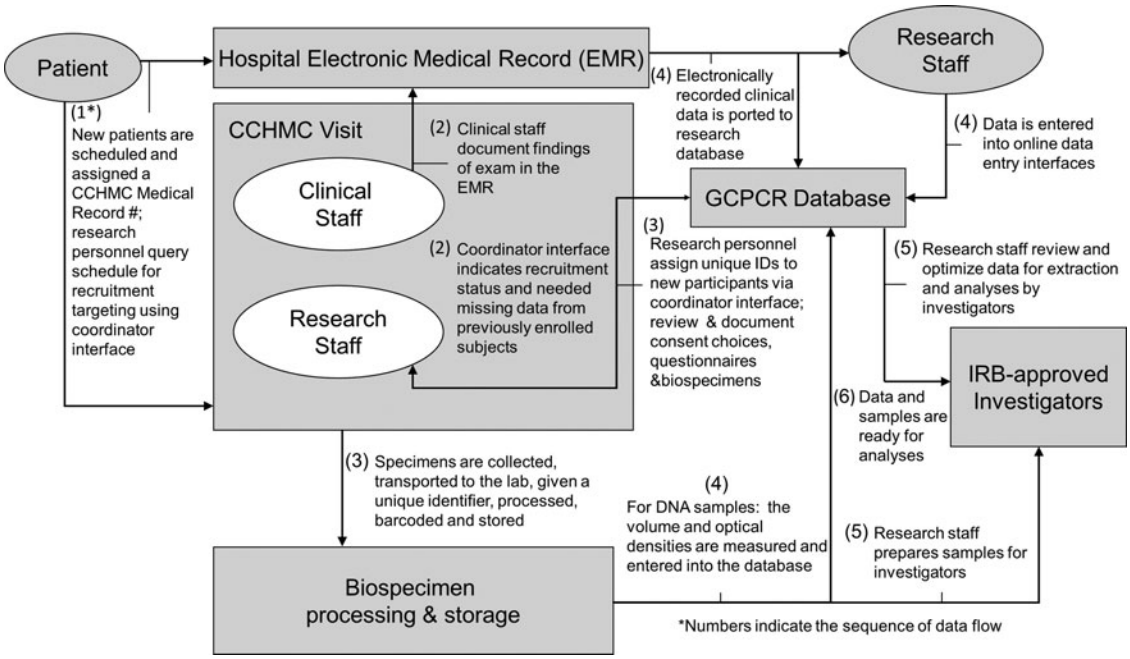

FIG. 1. Participant, data, and biological sample flow. CCHMC, Cincinnati Children's Hospital Medical Center; GCPCR, Greater Cincinnati Pediatric Clinic Repository; IRB, Institutional Review Board.

more efficient. All new patients to CCHMC clinics are scheduled through the Epic-based hospital system (Fig. 1, #1). Study recruitment personnel then prepare for recruitment in clinic by utilizing a custom-built web interface to identify (1) new patients that are eligible for participation in the repository; (2) previously consented repository participants who require additional sample collection or study questionnaires; or (3) participants who will need to complete follow-up questionnaires at their next clinic visit (Fig. 1, #2). During patients' CCHMC visits, clinical personnel examine the patients and document their results in the EMR (Fig. 1, #2). Either during the visit or prior to their departure, research personnel approach new patients, explain the goals of the repository, and obtain consent. Consented participants are assigned distinctive research participant identification numbers (PIDs). Patients (now repository participants) are then asked to complete repository-related questionnaires and provide requested biological samples (Fig. 1, #3). Biological specimens collected during the clinic visit are labeled with the PID and transported to the laboratory. Clinical data including physical exam findings, physician diagnoses, patients' weight and height, results of allergy SPTs, pulmonary function tests (PFT), chest and sinus x-rays, and other clinical and laboratory tests are extracted from participants' EMRs, and the relevant forms are entered electronically into a series of customized web-based interfaces (Fig. 1, #4). Data are transferred over a secure connection from the web browser to a server located in the CCHMC Data Center and placed in a mySQL database (Sun Microsystems, Inc.). Following DNA isolation in the laboratory, DNA samples are labeled with a unique bar-coded sample identification number (SID) that is linked to participants' PIDs and collection dates. SIDs and storage locations are automatically assigned by custom management software, and the labels are printed using its graphical interface. The concentration, quality, and volume of DNA for participant samples are determined by laboratory personnel and the corresponding data are directly entered into the sample management software after scanning the sample's barcode (Fig. 1, #3 and 4). Each time an aliquot is removed from the sample vial, laboratory staff scan the SID and record the volume of sample used, thereby ensuring that an accurate sample record is maintained.

IRB-approved investigators are able to request specific queries from the database directly by utilizing our internal website to complete a GCPCR Data and Sample Request in order to collaborate or obtain de-identified samples and data from this resource. In addition, using a customized Data Repository Query System, investigators will soon be able to view standardized reports detailing the repository's demographics, disease diagnoses, and recruitment progress in real time in order to capture a subset of variables and/or biospecimens to be obtained for further analysis (Fig. 1, #5) and then export the data for further analyses (Fig. 1, #6).

Demographic and other questionnaires

In addition to the baseline questionnaire, all GCPCR participants complete a one-page Demographic Data Questionnaire (DDQ) at their consenting visit. The DDQ captures age, gender, race, and ethnicity of the participant and participant's parents; annual household income; payment/insurance method; age; highest level of education of the subject's parents; and names and ages of the participant's siblings. Gender, race, and age data are verified using the hospitals' EMR. Beginning in 2008, asthmatic participants were also asked to complete modified versions of previously validated questionnaires related to individual and family-related quality of life in addition to their baseline questionnaire (Supplemental Table 1). These include section 3 of the Childhood Health Survey for Asthma (CHSA)[10] and the PADQLQ. Children with asthma were also asked to complete the Asthma Control Test (ACT).[11,12] Also in 2008,

regular follow-up of asthmatics began, and in 2009, regular follow-up of nonasthmatics began. Asthmatic and nonasthmatic participants returning to CCHMC for subsequent follow-up visits were also asked to complete new follow-up questionnaires (the Asthma Follow Up Questionnaire (AFUQ) or the Personal Health and Environment Questionnaire (PHEQ) respectively; Supplemental Table 1). The follow-up questionnaires mirror questions asked on the baseline questionnaire and further ascertain any health and exposure changes that have occurred since the participant's previous CCHMC visit. To further improve the follow-up of participants over time, the GCPCR staff is considering converting these questionnaires to an online format using the REDCap Software developed at Vanderbilt University.

GCPCR privacy and confidentiality

As with any repository in which individual donor identities are linked to biological specimens or data derived from specimens, participants are well-informed, privacy and confidentiality are maintained, and data and specimens are shared safely with other researchers. Some "best practices" were developed by the National Bioethics Advisory Commission and others[13] that provided us direction for: (1) limiting access to the codes that connect participants' identities and to their corresponding biological materials; (2) obtaining informed consent for the storage of biological materials for use in future research studies; (3) acquiring IRB approvals for the repository, as well as for the future research studies; and (4) instituting an advisory committee to protect the privacy of participants and monitor data and specimen and data usage. The repository housing the GCPCR and new protocols are reviewed by the GCPCR Advisory Committee made up of repository investigators and staff and informed by members of the Cincinnati Children's Institutional Review Board.

Disease definitions in the GCPCR

Asthma is diagnosed based on history, clinical examination, pulmonary function test results according to diagnostic criteria outlined by the American Thoracic Society, and the EPR-3 guidelines.[4,15] Children with allergic rhinitis are children with symptoms of the condition[16] that are sensitized to at least one aeroallergen based on allergy skin prick testing. Children with atopic dermatitis fulfilled the diagnostic criteria of Hanifin and Rajka.[17] Food allergy is diagnosed by history of an adverse reaction to the food and either a positive skin prick test or RAST to the food, or a positive oral food challenge. Repository questionnaires (Supplemental Table 1) and the EMR provide investigators with additional data including physician diagnoses, pulmonary function, and skin prick testing results, SCORAD measurements,[14] gastrointestinal pathology results (i.e., eosinophil counts), and transepidermal water loss measurements.

Results

Composition of the GCPCR

Since 2003, we have consented more than 6,317 participants. The rate of refusal of allergy or pulmonary clinic subjects approached in 2010 was less than 5%, with 70.5% of patients approached. Of those consented, more than 50% have consented to be recontacted in the future to participate in other prospective allergic disease studies. Of those recontacted for additional prospective studies, 60% have willingly participated. Of those early participants consented, 2,833 children completed the ICE and more than 1,171 individuals have completed the revised NVQ—our current baseline survey. At present, 2,054 (34%) of participants have completed at least one quality of life questionnaire and more than 904 (42% of all asthmatics) have completed at least one ACT. At this time, nearly 810 (13%) of participants (primarily children with asthma) have provided questionnaire-based data from two or more clinic visits. A summary of the consented participants' demographic data is provided in Table 1.

Quality control of the current baseline allergic disease questionnaire

In 2007, the ICE was expanded and revised after two phases of revision. In phase I, researchers, clinicians, nurses, staff, and parents/guardians were asked to identify topics of interest for the expansion and revision of the ICE. Based on their feedback, the questionnaire was expanded to an all-inclusive 18-page questionnaire called the ICE2 (Supplemental Table 1). To maintain continuity of data for the repository, the ICE2 linked similar or slightly modified questions from the original ICE. Answer choices were adapted to include discrete answers in place of open-ended questions or text fields whenever practical. The ICE2 also had a broader review of systems, more optimally captured patients' social and family histories, and history of environmental exposures (including an expanded tobacco smoke exposure section). Questions regarding food allergy symptoms, childhood diet history, and measures of physical activity, leisure time and sleep habits, and medication adherence were added. Parents/guardians of 39 children visiting the allergy and immunology clinics at CCHMC for the first time were mailed the ICE2 prior to their visit and were asked to complete a second ICE2 during their visit. Following completion, caregivers were interviewed by study staff to ascertain comprehension and acceptability (e.g., understandability, appropriateness of topics assessed, length of questionnaire) of the questionnaire. Approximately 20% of parents/guardians interviewed indicated that they had difficulty understanding questions or filling out the questionnaire. Ninety percent of parent/guardian of asthmatic children felt the questions were relevant to their child. The length of the ICE2 (18 pages) was too long for 85% of respondents. This feedback led to several iterations of revision of the questionnaire. Several questions/sections such as the review of systems were removed given the implementation of the EMR at CCHMC. The resulting New Patient Visit Questionnaire (NPVQ) and New Visit Questionnaire (NVQ) were then reevaluated in Phase II (Supplemental Table 1). As before, 40 caregivers were mailed the NPVQ and the same caregiver was asked to complete either the NPVQ or shorter NVQ during their child's scheduled clinic visit. Test–retest reliability was assessed using Spearman correlations. Internal consistency was evaluated with Cronbach's alpha of reliability. As a measure of validity, domain scores were compared to physician diagnoses.

Of the 40 caregivers asked to complete two baseline surveys (two NPVQs, two NVQs, or one of each) for quality control purposes, 63% were Caucasian, 30% were African American, and 8% were of another race or listed more than

Table 1. GCPCR Demographic Data Summary

	All GCPCR	*Asthma*	*Allergic rhinitis*	*Atopic dermatitis*	*Food allergy*	*Nonallergic control*	*Cincinnati*[b]	*United States*[c]
N	6,317	2,143	1,773	1,248	1,053	224	1,646,395	281,421,906
Sex								
Male	3,584	1,283	1,080	713	681	113	798,600	138,053,563
Female	2,733	860	693	535	372	111	847,795	143,368,343
Race								
Caucasian	3,778	1,080	968	609	684	157	1,375,267	194,552,774
African American	1,993	891	649	521	275	46	212,452	33,947,837
Other	430	141	137	89	82	14	40,959	17,615,477
Age								
<5 yrs	2,564	543	457	777	692	18		
5–12 yrs	2,726	1,214	990	397	284	104		
13–18 yrs	957	370	313	71	72	95		
>19 yrs	70	16	13	3	5	7		
Income[a]								
<$10K	823	351	240	179	110	45	8.8%	9.5%
$10–30K	1,065	407	334	187	129	33	24.2%	25.6%
$30–50K	695	248	227	105	113	22	22.8%	22.9%
$50–70K	586	148	157	113	105	24	16.5%	15.6%
>$70K	1,479	393	422	260	292	54	27.7%	26.4%

[a]Annual household income.
[b,c]Cincinnati metro region and US data based on 2000 Census Bureau Statistics.
GCPCR, Greater Cincinnati Pediatric Clinic Repository.

one race. Nearly 13% did not have their high school diploma or received their General Educational Development (GED) diploma, while 50% had earned a 4-year college or other graduate degree. Approximately 22% had a household income of less than $30K, but most (68%) had private health insurance. There were 20 caregivers (50%) parenting children that either had been or were later diagnosed with asthma; 20 of the children were male, and the ages of the children ranged from 3 to 18 years. According to the Patient Feedback Form, 80.6% of caregivers indicated that the questionnaires were relevant to their child's condition, while 43% still felt the questionnaire was either a little long or too long, particularly when they had been asked to complete the NPVQ. Questions present on both the NPVQ and the NVQ within five domains focused on asthma and respiratory symptoms, allergy symptoms, atopic dermatitis, family allergic disease history, and tobacco smoke exposure were used to develop domain scores (Supplemental Table 2). Reliability between questionnaires and within domains was examined. Test–retest reliability was assessed using Spearman correlations. Internal consistency was evaluated with Cronbach's alpha of reliability. We found a high level of agreement (Spearman $\sigma > 0.6$) and good internal consistency (Cronbach's $\alpha \geq 0.7$) for most domains (Table 2). For the seven question-based asthma history domain and respiratory symptom frequency domain (made up of symptom frequency questions for wheeze, cough, shortness of breath, and chest tightness), correlations between questionnaire scores were 0.94 and 0.83 respectively, indicating good reliability. As a measure of validity, domain scores were compared to physician diagnoses. Our results show that correlations were high for the asthma history and respiratory symptom frequency domains ($\sigma = 0.82$ and 0.83 for each score respectively), but lower for the allergic symptom domain ($\sigma = 0.25$). These findings are not surprising given the symptoms queried for asthma are specific to asthma, while allergic symptoms (itchy red eyes, runny nose, sneezing, itchy nose, and stuffy nose or congestion) queried are applicable to other conditions such as colds and respiratory infections. Reliability for the three question-based atopic dermatitis domain was high ($\sigma = 0.84$). However, validity was not measured from these questions, since children with atopic dermatitis symptoms are given the Atopic Dermatitis Questionnaire, which includes more detailed questions regarding location, triggers, and quality of life questions specific to skin symptoms, as well as the SCORAD index[14] as a measure of their disease severity.

We are in the process of revising and shortening the NVQ again given the results of this two-phase evaluation. Further adaptations to the revised form will occur following the publication of the new guidelines from the National Heart, Lung and Blood Institute and National Institute of Allergy and Infectious Diseases for standardized asthma outcomes in clinical research. In addition, we are in the early phases of planning conversion of our revised questionnaire and several of our other questionnaires to an online format in order to streamline data collection, data entry, and patient flow. An online format will also allow a skip logic to be implemented so that caregivers of children visiting for a food allergy with no related respiratory symptoms, for example, could skip questions irrelevant to their child and complete only sections pertinent to their child's symptoms. This would help alleviate caregivers' feelings that the questionnaire is too long, improve participants' overall experience with the study, and augment their willingness to participate during follow-up visits.

GCPCR disease frequencies

The most frequent diagnosis in the GCPCR is asthma ($n = 2,143$, 34%) followed by allergic rhinitis ($n = 1,773$, 28%),

Table 2. Baseline Questionnaire Reliability and Validity

Phase II primary questionnaire domains scores (score range)	# Contributing questions/ symptoms	Internal consistency within domain	% Agreement between two questionnaires	Question and domain validity versus physician diagnosis		
		Cronbach's alpha	Spearman correlation	Spearman correlation	+ Diagnosis median (Q1,Q3)	– Diagnosis median (Q1,Q3)
Asthma history score (0–7)	7	0.90	0.94**	0.82**	6.0 (4.5,7.0)	0.0 (0.0, 2.0)
Respiratory symptoms score (0–16)	4	0.75	0.83**	0.83**	7.0 (4.0, 8.0)	3.0 (1.0, 4.0)
Wheeze (0–4)	1	–	0.92**	0.48*	1.0 (1.0, 2.0)	0.0 (0.0, 1.0)
Cough (0–4)	1	–	0.75**	0.36*	2.0 (1.0, 3.0)	1.0 (0.0, 2.0)
Shortness of breath (0–4)	1	–	0.82**	0.46*	2.0 (1.5, 3.0)	0.0 (0.0, 1.0)
Chest tightness (0–4)	1	–	0.88**	0.57*	1.0 (0.0, 2.0)	0.0 (0.0, 0.0)
Allergic rhinitis symptoms score (0–5)	5	0.88	0.43*	0.25	4.0 (4.0, 5.0)	4.0 (2.0, 5.0)
Atopic dermatitis score (0–9)	3	0.92	0.84**	–	–	–
Family allergic disease score (0–12)	12	0.69	0.88**	–	–	–
Tobacco smoke exposure score (0–8)	8	0.71	0.90**	–	–	–

Twenty caregivers of asthmatics and 20 caregivers of nonasthmatics were asked to complete two baseline questionnaires (the 14-page New Patient Visit Questionnaire and the eight-page New Visit Questionnaire). One questionnaire was completed prior to their clinic visit and a 2nd during their visit. Sample size for each individual questionnaire domain or symptom may vary due to incomplete or missing data. See Table 2 for questions that make up each domain. Internal consistency within a single questionnaire was measured using the Cronbach's alpha statistic. Reliability between questionnaires and validity versus the physician diagnosis was measured using the Spearman correlation coefficient given the data were not normally distributed. + / – Diagnosis indicates the presence or absence of an asthma diagnosis for the Asthma History and Respiratory Symptoms Domain Scores and allergic rhinitis diagnosis for the Allergic Rhinitis Symptoms Domain Scores; Q1 = 25th percentile; Q3 = 75th percentile; $^*p < 0.05$; $^{**}p < 0.001$; – indicates the statistic was not computed.

Table 3. Top 10 Disease Diagnoses

Rank	Diagnosis	n =	%
1	Asthma	2,143	33.7
2	Allergic rhinitis	1,773	28.3
3	Atopic dermatitis	1,248	19.7
4	Food allergies	1,053	16.8
5	Eosinophilic esophagitis	221	3.5
6	Allergic conjunctivitis	181	2.9
7	Urticaria	108	1.7
8	Otitis media	71	1.1
9	Nonallergic rhinitis	64	1.0
10	Sinusitis	62	1.0

atopic dermatitis ($n = 1{,}248$, 20%), food allergy ($n = 1{,}053$, 17%), and eosinophilic esophagitis ($n = 221$, 4%; Table 3). Of 1,614 asthmatic children with self-reported symptom scores or pulmonary function testing results, approximately 23.5% and 27% were classified as having either moderate or severe asthma respectively. Interestingly, there are 1,175 (55%) asthmatic children with diagnoses of one or more of the other three classical allergic diseases we considered (atopic dermatitis/eczema, food allergy, allergic rhinitis) and 117 asthmatic children with all three comorbid disease diagnoses, enabling studies of various combinations of these disorders. For example, the phenotype of asthma in the context of atopic dermatitis may be distinct from asthma without atopic dermatitis.

Of the nonallergic controls tested, we found 45% were SPT positive (>3 mm wheal with erythema) to one or more allergens within 6 months of enrollment and an additional 6% tested positive at a subsequent clinic visit. As before, parents of controls were asked to help their child complete a one-page modified version of the PADQLQ to rule out that they have experienced allergy-like symptoms in the past week.[9] We identified 28% of a subset of our first 81 nonallergic controls in our analysis to be sensitized to at least one allergen. However, similar to our earlier published results,[8] when we compared the average PADQLQ scores between sensitized and non-sensitized subsets, we found no significant difference between the groups, 0.21 ± 0.36 versus 0.15 ± 0.27 ($p = 0.45$) respectively, indicating that both subsets of nonallergic controls were free of allergy symptoms.

Discussion

In summary, the GCPCR is a unique longitudinal biorepository that links clinical data and health outcomes to biologic samples and molecular data. It provides an infrastructure for better characterization of phenotypes so that observed heterogeneity in allergic diseases can be dissected. It is unique in that children with allergic disorders are recruited in parallel to healthy controls. Subjects' asthma, allergy, and skin symptoms; past medical, family, social, diet, and environmental histories; physical activity; medication adherence; perceived quality of life; and demographics are ascertained. DNA is collected from all participants, and other biospecimens such as blood, hair, and nasal epithelial cells are collected on a subset. The GCPCR is being utilized by several ongoing studies. It is an invaluable resource for clinical and translational research in

allergy as demonstrated by the broad range of research activities that the GCPCR supports, including genetics, environmental, obesity, health outcomes, health literacy, and health disparities studies.

Ongoing genetic association studies

The factors that contribute to development of asthma and other allergic diseases are diverse, but the greatest risk factor is a family history or genetic predisposition. Identifying genes associated with these complex diseases remains a fundamental challenge. Currently, the GCPCR has more than 6,317 participants predominantly of Caucasian and African American descent (Table 1). More than 93% of these participants have banked DNA samples (57% saliva, 43% buccal) collected for the purpose of genetic studies. Common biologic pathways underlie allergic diseases such as asthma, atopic dermatitis, and food allergy with the immune system and the epithelium being key links. Indeed, epithelial cells are increasingly recognized as critical participants in the pathogenesis of allergic inflammation. In fact, we previously reported that even healthy appearing skin of children with atopic dermatitis has a defective skin barrier function that is associated with disease severity[18] further supporting the likely role of genetic predisposition in this disease.

As such, NIH-funded studies are underway at CCHMC to elucidate the genes and biologic pathways that contribute allergic diseases such as atopic dermatitis and asthma. As the current standard of asthma treatment is ineffective in 40–70% of patients,[19] the need for novel and alternative therapies, such as those that target the epithelium, is high. Therefore, we recently designed a custom Illumina GoldenGate assay with 768 predominantly tagging SNPs within 52 genes that are located not only in pathways related to the epithelium, but also in those related to innate and adaptive immunity, inflammation, oxidative stress, microbe sensing, and protease inhibition. We genotyped 750 childhood asthmatics age 4 to 17 years old compared to 419 nonasthmatic/allergic children and 349 nonasthmatic/nonallergic controls using our custom assay in order to elucidate genetic associations with asthma.[20–22] A similar Illumina GoldenGate assay has been developed targeting genes specifically involved in atopic dermatitis. Future already funded studies will further investigate genetic associations of not only asthma and atopic dermatitis, but also these diseases independent of and in combination with allergic rhinitis and food allergies. Mucosal epithelial cells provide a critical physical defense barrier against allergens, pathogens, and other exposures, and also directly promote the development of Th2 responses. Each allergic disorder involves an epithelial or mucosal surface including the lung, gastrointestinal tract, or skin, but the epithelial pathways that contribute to disease at each site and possibly across sites have not been well investigated. The GCPCR will serve as a resource to make these unique studies possible.

Ongoing environmental exposure studies

Previous studies have indicated that environmental exposures dramatically influence the phenotype of allergic diseases.[23–25] For this reason, our questionnaires have detailed questions regarding primary and secondary tobacco smoke exposures, pet exposure, and early exposure to daycare. The smoking questions determine the type, the number of cigarettes per day, and the location of smoking of the participant, parents or guardians, and other members of the participant's household, as well as smoking habits of the biological mother during her pregnancy with the participant, the frequency of smoking in a car with the participant, and the average number of hours that the patient spends in the same area as someone who is smoking. Based on parental report of 850 of our asthmatic GCPCR recruits aged 5–18 years old, 33% indicated that their child lived with a smoker at the time of their child's enrollment. Of these subjects, 66% of mothers, 41% of fathers, and 29% of other adults living in the home reported smoking in the household. More than 19% of smoking households included both maternal and paternal smokers. Collection of hair samples in a subset of GCPCR participants has recently been implemented to enable future biologic measurement of nicotine and cotinine levels.

In addition to environmental tobacco smoke exposure, asthmatic and allergic children living in the Greater Cincinnati area are also greatly affected by traffic-related air pollution. Importantly, Greater Cincinnati has three converging major federal highways that cause traffic volume to be one of the largest in the nation. In fact, the level of exposure to truck related traffic is estimated to be one of the largest in the nation with 135,000–150,000 vehicles and 12,000–16,000 trucks per 24 hours traveling on each highway. Due to traffic, topography, and meteorological conditions in the Ohio Valley, air pollutants are concentrated in this area per an ongoing Environmental Protection Agency nonattainment rating. Elemental Carbon Attributable to Traffic (ECAT), a reliable surrogate for traffic related exposure, has previously been measured at 24 separate ambient air sampling sites in the greater Cincinnati area as part of the Cincinnati Childhood Allergy and Air Pollution Study (CCAAPS).[26] The average daily value of ECAT was determined for each sampling site for the years 2001–2005 and a land-use regression (LUR) approach was developed to relate the average daily values of ECAT at the sampling sites to surrounding traffic and land variables in a Geographic Information System (GIS). Estimates of truck-related traffic exposure at the home addresses of repository children are determined by combining ambient air sampling data with GIS data. As a result, the LUR model developed by CCAAPS can be used to estimate truck-related air pollution exposure in GCPCR participants. Using the same definition of high levels of exposure to traffic-related air pollution as CCAAPS ($>0.50\,\mu g/m^3$ per day), in 5,550 GCPCR participants, we found that 14% of the children would be considered highly exposed to DEP. The mean value was $0.38\,\mu g/m^3$, the range was $0.27–0.89\,\mu g/m^3$, and the standard deviation was 0.12. Children with asthma in our repository commonly live along one of Cincinnati's interstates (Fig. 2). Currently, there are two ongoing NIH-funded case-control studies evaluating oxidative stress-related biomarkers of both truck-related traffic exposure and secondhand smoke exposure side-by-side in the context of asthma. The studies' population bases are comprised of 64% of GCPCR participants with 330 enrollees undergoing analyses in the second year of the studies.

Ongoing disparity and health literacy studies

In the United States, the prevalence of asthma continues to increase, particularly among minority and low-income

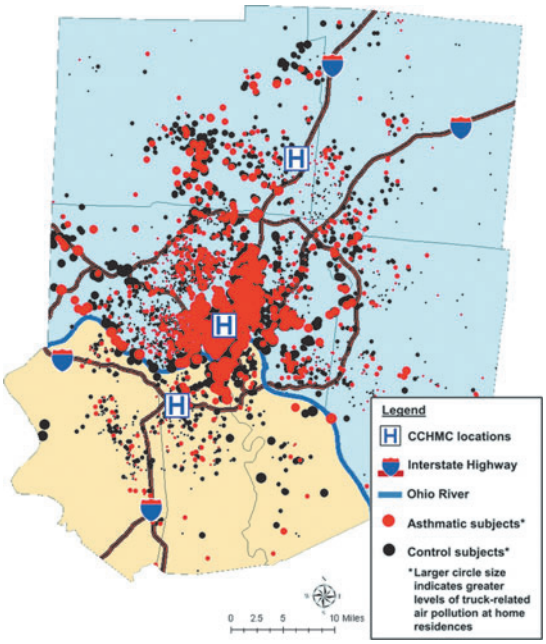

FIG. 2. Estimated traffic related air pollution at residential locations of GCPCR participants by disease status. Elemental analysis of ambient samples of PM2.5 was used to estimate levels of elemental carbon attributable to traffic (ECAT; a marker of traffic related air pollution) at 24 monitoring sites throughout the seven county Cincinnati Metropolitan area.[28] Geographic Information System software was used to obtain longitude and latitude coordinates from each repository participant's home address. A land-use regression model was used to derive estimates of ECAT exposure from these coordinates.[28]

groups, despite improvements in medications and treatment strategies. The GCPCR has more than1,993 African American participants and more than 1,800 participants in families with an income of less than $30,000 per year (Table 1). Geocoding maps of participants' primary residences suggest that African American asthmatics and nonasthmatics (Fig. 3) alike are clustered in or near the downtown Cincinnati area. The role of the parent or primary caregiver of children is critical in managing asthma and in influencing the child's health outcomes. This includes a variety of characteristics that contribute to a caregiver's ability to manage their child's asthma: access to medical care, insurance coverage, the ability to acquire prescription drugs, race, ethnicity, household income, family structure, residential location[27] and health literacy,[28] all of which have disproportionate effects on those in minority and low-income groups. While it is well established that patient education is essential for effective disease management, health literacy issues are rarely addressed and partially understood in research studies targeting low-income and/or minority patients. Minorities and low-income patients have a higher prevalence of low health literacy and utilize primary care and community health services more often than visiting a specialist. This is despite the literature that supports that specialists are better at treating the disease.[29] Recent reports also indicate that African Americans are less likely to seek the care of a specialist. Those that do seek a specialist's care, however, are among the sickest. While asthma outcomes are known to improve under specialty care, the impact of caregiver health literacy on asthma outcomes of minority children and/or children in low-income families visiting asthma and allergy specialists is not well established. In fact, in a currently enrolling study

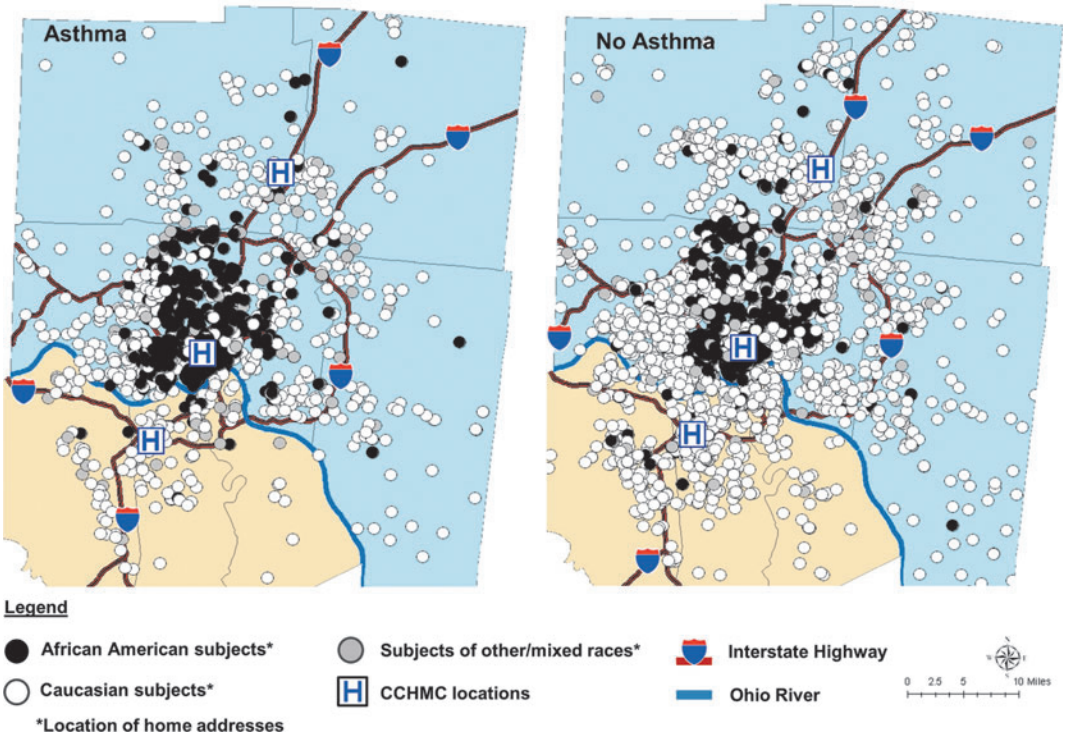

FIG. 3. Residential locations of GCPCR participants stratified by race and disease status. Geographic Information System software was used to obtain longitude and latitude coordinates from each repository participant's home address.

Table 4. Strengths of the GCPCR

- Large resource with a large proportion of African American participants (31% of all participants) in addition to Caucasians
- Extensive data collection on asthmatic and allergic individuals with clinical, demographic, and questionnaire-based data
- Numerous data sources for phenotyping from parent report, disease diagnosis by a pediatric specialist, longitudinal records, bronchial responsiveness, and skin-prick tests, etc.
- Numerous data sources for home environment exposures including ambient air pollution exposure, tobacco exposure, breast-feeding and early feeding habits, pet ownership, and daycare use during early life of the child
- Ability to obtain samples of saliva, nasal epithelial cells, mucus, blood, hair (already have on a subset)
- Nonallergic controls without personal and family history of allergic disease
- Captures children living in urban and rural environments
- Minimal study bias because cases and controls selected from the same population
- Genotyping on more than1,200 individuals including genotyping of AIMs
- Ability to collect samples on parents and siblings for family-based studies
- Customized database with tracking system of all biological and clinical data
- Questionnaires and data collection can be amended for changing interests and requirements
- Centralized recruitment strategy ensures a broad capture of clinic-based participants
- New scientific users able to choose participants based on a number of criteria
- Multiple studies all feed in to one repository
- Product of a multidisciplinary collaboration, which brings together the fields of pediatric pulmonary and allergy specialists, population health, immunology, genetics, and environmental epidemiology
- Fertile training ground for graduate students, postdoctoral fellows, and clinician fellows

utilizing the GCPCR, nearly 30% of 145 caregivers of 4–11-year-old asthmatic children screened using the Rapid Assessment of Adult Literacy in Medicine (REALM) had less than a high school reading level,[30] further suggesting a need examine social factors, in addition to biological factors, that promote asthma in highly susceptible populations and the role of poor health literacy as one of many factors that lead to disparate health outcomes.

Strengths and limitations of the GCPCR and its ongoing improvement

The GCPCR is a unique resource for allergic diseases in children that can and has been continuously adapted as new opportunities arise. Limitations of the cohort include current limited longitudinal data and demographic data from subjects enrolled early in course of the GCPCR. For example, in the early years, the GCPCR focused on recruiting children at their baseline visits to the specialist and consequently has limited longitudinal data on early participants. This is particularly problematic when evaluating younger children that may have subsequently developed asthma, for example, following their baseline visit. To remedy this issue, a data capture campaign was completed in 2008 to obtain missing data and to get updates on study participants. Nearly 45% of participants recontacted responded. In addition to the recapture campaign, follow-up questionnaires were developed at that time and study coordinators began to mail these questionnaires to previously consented participants with the goal of regular collection at subsequent clinic visits. Utilization of online questionnaires is now being considered and will likely permit more rapid questionnaire completion and previously consented participants would be able complete online questionnaires from home outside of or prior to their CCHMC visits. While complete validation of all the questionnaires used by the GCPCR is not possible given the ongoing clinical needs and magnitude of the information being collected, this is a future goal.

Another limitation of the study is also routed in the early years. Early on, only the race and ethnicity of the study participant were collected. In 2008, we began collecting information about race and ethnicity on both the participants and their parents using the revised guidelines established by the Office of Management and Budget at the NIH (2001) for studies recruiting minorities. These include two ethnic categories: Hispanic or Latino, and Not Hispanic or Latino; and five racial categories: American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander and White [NOT-OD-01-053]. In addition, we now also inquire about participants' siblings (full, half, step, and biologically unrelated) to enhance genetic and exposure-based studies. Genotyping is then used to corroborate relatedness as a second quality control measure, as well as correct gender assignment. Inclusion of a set of ancestry informative markers in genotyping studies along with principal component analysis allows us to discover misidentified ancestry, admixture and individuals that have a unique combination of ethnic backgrounds that make them outliers for population-specific studies. Recruitment of African Americans with asthma, as well as appropriate controls, is a priority of the GCPCR. The availability of these populations in the repository, as well as numerous other characteristics, makes the GCPCR a distinctive resource (Table 4).

Conclusions

In summary, the GCPCR is a well-designed and growing biorepository that has been constantly evaluated and refined to achieve and maintain rigorous clinical phenotype and biological data. It serves as the source of extensive longitudinal phenotype data and biologic samples for several new and ongoing studies and provides a unique resource for research collaborations. Continued development of repositories using common data elements will be necessary to enable studies across populations of extensively phenotyped patients.

Acknowledgments

The authors thank the physicians, nurses and staff of Cincinnati Children's Hospital Medical Center Allergy and Immunology clinics, Pulmonary Medicine clinics, Dermatology clinics, General and Community Pediatrics clinics, Headache Center clinics, Dental clinics, Orthopedic clinics, Plastic Surgery clinics, and the Emergency Department. We thank Drs. Sheela Geraghty, David Billmire, and Eric Wall for

their assistance. We thank all the patients and their families who participated in the GCPCR. This work was supported by the National Institutes of Health [U19A170235 to GKKH and R21016830, grant #RR026315-02 to MBK]. The project described was supported by the National Center for Research Resources, Grant KL2RR026315, and is now at the National Center for Advancing Translational Sciences, Grant KLTR000078. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Author Disclosure Statement

No conflicting financial interests exist.

References

1. Sly RM. Changing prevalence of allergic rhinitis and asthma. Ann Allergy Asthma Immunol 1999; 82:233–248; quiz 48–52.
2. Martinez FD, Wright AL, Taussig LM, Holberg CJ, Halonen M, Morgan WJ. Asthma and wheezing in the first six years of life. The Group Health Medical Associates. N Engl J Med 1995; 332:133–138.
3. Akinbami L. The state of childhood asthma, United States, 1980–2005. Adv Data 2006; 1–24.
4. National Asthma Education and Prevention Program (NAEPP). Bethesda (MD): National Heart L, and Blood Institute. Expert panel report 3: guidelines for the diagnosis and management of asthma 2007: 213–276.
5. ISAAC. Worldwide variation in prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and atopic eczema: ISAAC. The International Study of Asthma and Allergies in Childhood (ISAAC) Steering Committee. Lancet 1998; 351:1225–1232.
6. Asher MI, Keil U, Anderson HR, et al. International Study of Asthma and Allergies in Childhood (ISAAC): rationale and methods. Eur Respir J 1995; 8:483–491.
7. Expert, Panel, Report2. Guidelines for the Diagnosis and Management of Asthma; NIH Publication No. 97-4051. NIH, NHLBI. 1997.
8. Stevenson MD, Sellins S, Grube E, et al. Aeroallergen sensitization in healthy children: racial and socioeconomic correlates. J Pediatr 2007; 151:187–191.
9. Roberts G, Hurley C, Lack G. Development of a quality-of-life assessment for the allergic child or teenager with multisystem allergic disease. J Allergy Clin Immunol 2003; 111: 491–497.
10. Asmussen L, Olson LM, Grant EN, Fagan J, Weiss KB. Reliability and validity of the Children's Health Survey for Asthma. Pediatrics 1999; 104:e71.
11. Liu AH, Zeiger RS, Sorkness CA, et al. The Childhood Asthma Control Test: retrospective determination and clinical validation of a cut point to identify children with very poorly controlled asthma. J Allergy Clin Immunol 2010; 126:267–273, 73 e1.
12. Liu AH, Zeiger R, Sorkness C, et al. Development and cross-sectional validation of the Childhood Asthma Control Test. J Allergy Clin Immunol 2007; 119:817–825.
13. Eiseman E, Bloom G, Brower J, Clancy N, Olmsted S. Case Studies of Existing Human Tissue Repositories: "Best Practices" for a Biospecimen Resource for the Genomic and Proteomic Era. Santa Monica, CA: RAND Corporation 2003.
14. Kunz B, Oranje AP, Labreze L, Stalder JF, Ring J, Taieb A. Clinical validation and guidelines for the SCORAD index: consensus report of the European Task Force on Atopic Dermatitis. Dermatology 1997; 195:10–19.
15. Standardization of Spirometry, 1994 Update. American Thoracic Society. Am J Respir Crit Care Med 1995; 152:1107–1136.
16. Dykewicz MS, Hamilos DL. Rhinitis and sinusitis. J Allergy Clin Immunol 2010; 125:S103–115.
17. Paul WE, Zhu J. How are T(H)2-type immune responses initiated and amplified? Nat Rev Immunol 2010; 10:225–235.
18. Gupta J, Grube E, Ericksen MB, et al. Intrinsically defective skin barrier function in children with atopic dermatitis correlates with disease severity. J Allergy Clin Immunol 2008; 121: 725–730 e2.
19. Drazen JM, Silverman EK, Lee TH. Heterogeneity of therapeutic responses in asthma. Br Med Bull 2000; 56:1054–1070.
20. Baye TM, Butsch Kovacic M, Biagini Myers JM, et al. Differences in candidate gene association between European ancestry and African American asthmatic children. PLoS One 2011; 6:e16522.
21. Sherrill JD, Rothenberg ME. Genetic dissection of eosinophilic esophagitis provides insight into disease pathogenesis and treatment strategies. J Allergy Clin Immunol 2011; 128:23–32.
22. Sherrill JD, Gao PS, Stucke EM, et al. Variants of thymic stromal lymphopoietin and its receptor associate with eosinophilic esophagitis. J Allergy Clin Immunol 2010; 126:160–165 e3.
23. D'Amato G, Liccardi G, D'Amato M, Cazzola M. Respiratory allergic diseases induced by outdoor air pollution in urban areas. Monaldi Arch Chest Dis 2002; 57:161–163.
24. Gilmour MI, Jaakkola MS, London SJ, Nel AE, Rogers CA. How exposure to environmental tobacco smoke, outdoor air pollutants, and increased pollen burdens influences the incidence of asthma. Environ Health Perspect 2006; 114:627–633.
25. von Mutius E. The environmental predictors of allergic disease. J Allergy Clin Immunol. 2000; 105:9–19.
26. Ryan PH, Lemasters GK, Biswas P, et al. A comparison of proximity and land use regression traffic exposure models and wheezing in infants. Environ Health Perspect 2007; 115:278–284.
27. Chen AY, Escarce JJ. Family structure and the treatment of childhood asthma. Med Care 2008; 46:174–184.
28. DeWalt DA, Dilling MH, Rosenthal MS, Pignone MP. Low parental literacy is associated with worse asthma care measures in children. Ambul Pediatr 2007; 7:25–31.
29. Vollmer WM, O'Hollaren M, Ettinger KM, et al. Specialty differences in the management of asthma. A cross-sectional assessment of allergists' patients and generalists' patients in a large HMO. Arch Intern Med 1997; 157:1201–1208.
30. Davis TC, Long SW, Jackson RH, et al. Rapid estimate of adult literacy in medicine: a shortened screening instrument. Fam Med 1993; 25:391–395.

Address correspondence to:
Gurjit K. Khurana Hershey, M.D., Ph.D.
Division of Asthma Research
Cincinnati Children's Hospital Medical Center
3333 Burnet Ave., MLC 7037
Cincinnati, OH 45229

E-mail: gurjit.hershey@cchmc.org