PCORnet Antibiotics and Childhood Growth Study: Process for Cohort Creation and Cohort Description

*Jason P. Block, MD, MPH; L. Charles Bailey, MD, PhD; Matthew W. Gillman, MD, SM; Douglas Lunsford, BS; Janne Boone-Heinonen, PhD, MPH; Lauren P. Cleveland, MS, MPH; Jonathan Finkelstein, MD, MPH; Casie Horgan, MPH; Melanie Jay, MD, MS; Juliane S. Reynolds, MPH; Jessica Sturtevant, MS; Christopher B. Forrest, MD, PhD on behalf of the PCORnet Antibiotics Childhood Growth Study Group**

From the Division of Chronic Disease Research Across the Lifecourse (CoRAL) (Drs Block and Cleveland); Therapeutics Research and Infectious Disease Epidemiology Group (Ms Horgan, Ms Reynolds, and Ms Sturtevant), Department of Population Medicine, Harvard Pilgrim Health Care Institute; Department of Pediatrics (Dr Finkelstein), Harvard Medical School, Boston, Mass; Applied Clinical Research Center, Department of Pediatrics, Children's Hospital of Philadelphia (Drs Bailey and Forrest), Philadelphia, Pa; Environmental Influences on Child Health Outcomes (ECHO) Program, Office of the Director, National Institutes of Health (Dr Gillman), Bethesda, Md; North Fork School District (Mr Lunsford), Utica, Ohio; School of Public Health, Oregon Health & Science University/Portland State University (Dr Boone-Heinonen), Portland, Ore; and Department of Population Health, New York University School of Medicine (Dr Jay), New York, NY

The authors have no conflicts of interest to disclose.

Address correspondence to Jason Block, MD, MPH, Division of Chronic Disease Research Across the Lifecourse, Department of Population Medicine, Harvard Pilgrim Health Care Institute, Harvard Medical School, 401 Park Dr, Boston, MA 02215 (e-mail: jblock1@partners.org).

*Members of the PCORnet Antibiotics and Childhood Growth Study Group are listed in Appendix 1.

Received for publication October 11, 2017; accepted February 11, 2018.

ABSTRACT

OBJECTIVES: The National Patient-Centered Clinical Research Network (PCORnet) supports observational and clinical research using health care data. The PCORnet Antibiotics and Childhood Growth Study is one of PCORnet's inaugural observational studies. We sought to describe the processes used to integrate and analyze data from children across 36 participating institutions, and the cohort characteristics and prevalence of antibiotic use.

METHODS: We included children in the cohort if they had at least one same-day height and weight measured in each of 3 age periods: 1) before 12 months, 2) 12 to 30 months, and 3) after 24 months. We distributed statistical queries that each institution ran on its local version of the PCORnet common data model, with aggregate data returned for analysis. We defined overweight or obesity as age- and sex-specific body mass index ≥85th, obesity ≥95th percentile, and severe obesity ≥120% of 95th percentile.

RESULTS: A total of 681,739 children met the cohort inclusion criteria and were racially/ethnically diverse (24.9% black, 17.5% Hispanic). Before 24 months, 55.2% of children received at least one antibiotic prescription; 21.3% received a single antibiotic prescription, 14.3% received 4 or more, and 33.3% received a broad-spectrum antibiotic. Overweight and obesity prevalence was 27.6% at age 4 to <6 years (n = 362,044) and 36.2% at 9 to <11 years (n = 58,344).

CONCLUSIONS: The PCORnet antibiotics study is a large national longitudinal observational study in a diverse population that will examine the relationship between early antibiotic use and subsequent growth patterns in children.

KEYWORDS: antibiotics; body mass index; childhood growth; childhood obesity; electronic health records; research infrastructure

ACADEMIC PEDIATRICS 2018;■■:■■–■■

WHAT'S NEW

The National Patient-Centered Clinical Research Network (PCORnet) provides an unprecedented opportunity to conduct research using health care data. In 36 health care institutions, we assembled a large cohort to examine antibiotics and childhood growth. More than half of children received an antibiotic prescription before 2 years of age.

THE WIDESPREAD AVAILABILITY of health care data through electronic health records (EHRs) and other data sources provide unique opportunities to conduct pragmatic clinical trials and observational studies on a large scale. The National Patient-Centered Clinical Research Network (PCORnet) is a distributed research network that uses health care data to facilitate multisite clinical trials and observational research studies.[1–3] PCORnet has 13 clinical data research networks (CDRNs) that contribute health information for over 128 million patients. Within the CDRNs, data are organized in a common data model (CDM) that allows for standardization across institutions and the development of efficient and reusable tools to capture and analyze data.[4]

This type of data infrastructure is essential for patient-centered research that requires large sample sizes, such as studies of rare diseases or studies that require assessments of heterogeneity of treatment effects, with various types of exposures among specific subgroups. To help develop infrastructure for observational research in PCORnet, the Patient-Centered Outcome Research Institute (PCORI) funded 2 initial observational studies to explore diverse research questions and launch PCORnet into a research-ready data system.[3,5] One of these studies, the PCORnet Antibiotics and Childhood Growth Study, was the first effort in PCORnet to establish a large pediatric cohort across the network, and the first study in PCORnet to actively characterize prescribing data. Having access to a large pediatric cohort also will enable assessments of different types, timing, and doses of antibiotics on weight outcomes, which has been difficult to do with smaller studies.

Our objective was to evaluate the utility of this cohort for conducting comparative effectiveness research on medications and growth in young children. Here we describe the processes used to integrate, synchronize, and analyze data from children across 36 health care institutions organized in 10 CDRNs, and we describe the cohort characteristics, including antibiotic use before 24 months of age and prevalence of overweight and obesity from early to midchildhood.

METHODS

PARTICIPATING INSTITUTIONS AND THE PCORNET CDM

PCORI, created by the Affordable Care Act of 2010, is a funding agency that supports patient-centered comparative effectiveness research within 5 priority areas: "evaluating prevention, diagnosis, and treatment options; improving health systems; enhancing communication and dissemination of evidence; addressing disparities in health and health care; and improving comparative effectiveness research methods and data infrastructure."[6] PCORI created PCORnet to expand the data infrastructure available for comparative effectiveness research in a manner that incorporates the input of stakeholders.[2] In addition to the 13 participating CDRNs including data from nearly 100 health care systems, PCORnet has 20 People-Powered Research Networks that are focused on specific diseases or populations (contributing to both stakeholder engagement efforts and data) and 2 Health Plan Research Networks that are working to link health insurance claims data to PCORnet EHR data.

In PCORnet, data are organized by Network Partners. These Network Partners include data from either one contributing health care institution or, in the case of centralized Network Partners, from multiple institutions. In this study, 28 Network Partners participated, and these partners hold data for 36 institutions across 10 CDRNs, including integrated delivery systems, freestanding children's hospitals, and federally qualified health centers (Supplemental Table 1). To participate in the study, Network Partners had to meet data quality standards that were set forth by the PCORnet Coordinating Center. These included assessments of data model conformance, missing data in required tables and variables, and data plausibility in date and vital measure fields.[7] Required tables included enrollment, encounters, demographics, vital findings, diagnoses, and procedures. The study team additionally required that Network Partners had the capacity to create a pediatric cohort that met the study's inclusion criteria and that could identify antibiotic prescriptions. Of the 44 institutions initially planned for inclusion, we removed 8 from the study for the following reasons: did not meet Coordinating Center data quality standards (n = 1) or did not meet them by February 1, 2017 (n = 1); did not have access to outpatient prescription medications in their CDM (n = 2); were unable to map their prescribing data to RxNorm codes needed for the study by February 1, 2017 (n = 1); were unwilling to share individual-level data (n = 2); or chose not to participate because the site had a small pediatric population available in their CDM (n = 1). A team of stakeholders from participating CDRNs and 4 of the People-Powered Research Networks, including parents, providers, health system representatives, and patient advocates, closely informed the study conception and design, and provided ongoing feedback throughout the study.

The PCORnet CDM consists of 15 tables and over 100 variables available for research.[4] An in-depth assessment of data usability and consistency was necessary before conducting statistical analyses, a process called study-specific data characterization. For the PCORnet Antibiotics and Childhood Growth Study, this process included capturing site-level aggregate data on study-specific variables (eg, demographics, diagnoses, medications, vital signs). The study team analyzed this data to determine which sites met data quality eligibility requirements, while providing initial information on the cohort of interest.

DISTRIBUTED STATISTICAL NETWORK QUERIES

Sites extract data from their local EHR systems and other health care data repositories, such as insurance claims, and transform those data to meet CDM standards. The PCORnet distributed research network model addresses governance and privacy concerns by allowing institutions to maintain data locally, rather than create a network-wide centralized database. Queries written to conform to the CDM standards are distributed to Network Partners for local execution, resulting in the return of standardized output that can be aggregated with other partners. To produce statistical query packages for distribution, either for data characterization or study analyses, PCORnet follows a standard workflow, informed by the setup of the US Food and Drug Administration's Sentinel program.[8] The Sentinel program utilizes claims data from health insurers to examine drug safety across the United States. The workflow begins with the development of scientific specifications that describe the purpose of the query and the intended analyses, which serves as a blueprint for the programming team. The programmers then develop a SAS statistical query to capture relevant data from Network Partners or to conduct analyses (SAS Institute, Cary, NC). Study teams are also responsible for generating and reviewing codes for relevant variables used for the query. This study used International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM), and SNOMED-CT codes for diagnoses, and RxNorm and National Drug Code (NDC) codes for medications.