Journal of Clinical and Translational Science

www.cambridge.org/cts

Research Methods and Technology Brief Report

Cite this article: Ryan PH, Brokamp C, Blossom J, Lothrop N, Miller RL, Beamer PI, Visness CM, Zanobetti A, Andrews H, Bacharier LB, Hartert T, Johnson CC, Ownby D, Lemanske RF, Gibson H, Requia W, Coull B, Zoratti EM, Wright AL, Martinez FD, Seroogy CM, Gern JE, and Gold DR. A distributed geospatial approach to describe community characteristics for multisite studies. *Journal of Clinical and Translational Science* **5**: e86, 1–8. doi: 10.1017/cts.2021.7

Received: 5 October 2020
Revised: 8 January 2021
Accepted: 28 January 2021

Keywords:
Geospatial analysis; GIS; participant confidentiality; determinants of health; multisite studies

Address for correspondence:
P. Ryan, PhD, Cincinnati Children's Hospital Medical Center, 3333 Burnet Ave, Cincinnati, OH 45140, USA. Email: patrick.ryan@cchmc.org

© The Author(s), 2021. Published by Cambridge University Press on behalf of The Association for Clinical and Translational Science. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

A distributed geospatial approach to describe community characteristics for multisite studies

Patrick H. Ryan[1,2], Cole Brokamp[1,2], Jeff Blossom[3], Nathan Lothrop[4,5], Rachel L. Miller[6], Paloma I. Beamer[4,5], Cynthia M. Visness[7], Antonella Zanobetti[8], Howard Andrews[9], Leonard B. Bacharier[10], Tina Hartert[11], Christine C. Johnson[12], Dennis Ownby[13], Robert F. Lemanske[14], Heike Gibson[8], Weeberb Requia[8], Brent Coull[8], Edward M. Zoratti[12], Anne L. Wright[4,15], Fernando D. Martinez[4,15], Christine M. Seroogy[14], James E. Gern[14] and Diane R. Gold[8,16] on behalf of the CREW Consortium

[1]Department of Pediatrics, University of Cincinnati, College of Medicine, Cincinnati, OH, USA; [2]Division of Biostatistics and Epidemiology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA; [3]Center for Geographic Analysis, Harvard University, Cambridge, MA, USA; [4]Asthma and Airways Disease Research Center, University of Arizona, Tucson, AZ, USA; [5]Department of Community, Environment, and Policy, Mel and Enic Zuckerman College of Public Health, University of Arizona, Tucson, AZ, USA; [6]Division of Clinical Immunology, Icahn School of Medicine at Mount Sinai, New York, NY, USA; [7]Rho, Inc., Chapel Hill, NC, USA; [8]Department of Environmental Health, Harvard T. H. Chan School of Public Health, Boston, MA, USA; [9]Department of Biostatistics, Mailman School of Public Health, Columbia University, New York, NY, USA; [10]Washington University School of Medicine, Division of Pediatric Allergy, Immunology, and Pulmonary Medicine, St. Louis, MO, USA; [11]Vanderbilt University School of Medicine, Division of Allergy, Pulmonary, and Critical Care Medicine, Nashville, TN, USA; [12]Henry Ford Health System, Detroit, MI, USA; [13]Division of Allergy and Immunology, Augusta University, Augusta, GA, USA; [14]Department of Pediatrics, University of Wisconsin School of Medicine and Public Health, Madison, WI, USA; [15]Division of Pulmonary and Sleep Medicine, Department of Pediatrics, College of Medicine, University of Arizona, Tucson, AZ, USA and [16]Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA

Abstract

Understanding place-based contributors to health requires geographically and culturally diverse study populations, but sharing location data is a significant challenge to multisite studies. Here, we describe a standardized and reproducible method to perform geospatial analyses for multisite studies. Using census tract-level information, we created software for geocoding and geospatial data linkage that was distributed to a consortium of birth cohorts located throughout the USA. Individual sites performed geospatial linkages and returned tract-level information for 8810 children to a central site for analyses. Our generalizable approach demonstrates the feasibility of geospatial analyses across study sites to promote collaborative translational research.

Introduction

Maintaining patient privacy is a common challenge faced by researchers seeking to understand the relationship of place and health [1–4]. This issue can be especially problematic in multisite studies due to study protocols and confidentiality concerns that limit the sharing of geographic data. Existing approaches to geospatial analyses in multisite studies include the use of a central site to conduct analyses or the application of spatial techniques (e.g. changing geographic coordinates to protect confidentiality, i.e. "geomasking") to protect patient privacy [4]. However, the former method is limited by data use agreements (DUAs) and the latter may result in exposure misclassification when geomasked locations result in spatial misalignment.

Alternatively, study sites may perform geocoding and geospatial linkages independently before removing identifiable information for joint analyses. This decentralized approach, however, faces challenges in reproducibility and standardization due to geocoding methods, geographic information software (GIS), and expertise that varies across study sites. Here, we describe the application of a novel method to perform reproducible, standardized, and confidential geospatial analyses for multisite studies. Our approach extends a previously developed Decentralized Geomarker Assessment for Multi-site Studies (DeGAUSS) containerization platform to perform geocoding and extraction of polygon feature geospatial data over multiple time periods and large geographic areas [5]. As an example case, we use our approach to ascertaining US census tract-level information for participants enrolled in the Children's Respiratory and

Environmental Workgroup (CREW), a network of 12 birth cohorts each studying the development of allergy and asthma in childhood [6].

Materials and Methods

Study Population

Our approach was motivated by the CREW consortium, a network of birth cohorts recruited from 1980 to 2020. Information regarding participating cohorts, including eligibility criteria, study recruitment, and other methods, are published [6] and in Supplementary Table 1. Due to its large sample size and geographic distribution, CREW provides a unique platform to examine environmental and community factors that contribute to the disproportionately higher burden of asthma-related morbidity and mortality among disadvantaged communities [7–10]. A CREW data sharing protocol and DUA were approved by the local IRB for each cohort. However, the DUA allowed only limited datasets to be shared and prohibited the distribution of identifying information, including addresses or geocodes.

Distributed Geospatial Analysis

We extended our DeGAUSS software to enable all CREW sites, including those with limited geospatial expertise, to derive spatio-temporal US census tract-level information for their participants at birth. A key advantage to DeGAUSS is the use of a software containerization platform to wrap necessary software, system dependencies, and geospatial data in a stand-alone package that will work the same regardless of its host environment [11]. Previously, we created DeGUASS and applied this tool in the Electronic Medical Records and Genomics (eMERGE) network in a proof-of-concept study [5]. Whereas our prior DeGAUSS software included only a geocoder and code to link geospatial coordinates to nearby roadways and one census tract variable, the CREW consortium required significantly expanded geographical and temporal data to be included for analyses. Therefore, a new custom DeGAUSS container containing decennial US Census data (described below), census tract polygon boundary files for the 1980, 1990, 2000, and 2010 census, and R code (R Foundation for Statistical Computing: Vienna, Austria; 2014) was created to merge census tract-level data to the geocoded locations of CREW birth addresses. Additional details regarding DeGAUSS and the CREW container are provided in the supplementary materials and online [12,13].

A flow diagram depicting the distributed approach to geospatial analyses for the CREW consortium is provided in Figure 1. The DeGAUSS container image was created by C.B. at a single location and distributed to each cohort. The DeGAUSS software required cohort users to provide an input.csv file containing the geocoded coordinates of their participants' birth record address. Site end users also specified the census year to assign the appropriate tract boundary file and census data based on participants' year of birth. The output data file from the DeGAUSS container contained the original input data, including geocoded locations, and appended census tract information including the census tract FIPS code in which the birth record address was located and census variables. Site end users manually removed identifying information, including the geographic coordinates and census tract FIPS code, prior to returning the de-identified dataset to a central coordinating center.

US Census/American Community Survey Data

Longitudinal US Census data and boundary files for the years 1980, 1990, 2000, and 2010 were downloaded for the entire USA from Social Explorer (www.socialexplorer.com, New York City, NY: Social Explorer 2017; accessed 12/17–1/18) by year, requested variables, and tract level of geography. For 1980, certain census variables were only available from the National Historic GIS data service (Minneapolis, MN: NHGIS; accessed 11/17–1/18) and were updated to reflect variable calculation definitions used in later censuses. As the 2010 census did not include information regarding median household income, median gross rent, or median housing values, these data were downloaded from the 2008–2012 American Community Survey (ACS, https://www.census.gov/programs-surveys/acs/data.html; accessed 12/17–1/18). Additional information regarding the census and ACS variables downloaded and included in the DeGAUSS container for linkage to birth record addresses is available in the supplementary material.

Statistical Analyses

After sharing de-identified data with a central site, descriptive statistics and box-and-whisker plots for all census tract variables were calculated for the combined CREW consortium and for each cohort individually. Comparison of census tract-level to self-reported race, ethnicity, and household income was conducted by plotting the distribution of each census tract-level variable according to self-reported variable. Self-reported household income was compared to census tract median household income using each cohorts' income categories as collected by questionnaire. Additional details regarding self-reported race, ethnicity, and income information are available in the supplementary material.

Results

All cohorts ($n = 12$) completed the distributed analysis and returned de-identified data to the coordinating center. Collectively, 8997 participants were enrolled in CREW cohorts, and 98% ($n = 8810$) of these had birth record addresses geocoded with sufficient precision for linkage to a census tract.

A summary of population, race, ethnicity, and income data for the census tracts in which participants resided at birth is provided in Table 1. CREW participants resided in both low and high population density regions as reflected in the average tract population density (persons per km^2) that ranged from 148 for Wisconsin Infant Study Cohort (WISC) participants in rural Wisconsin to 45,772 for CCCEH participants in New York City (Table 1). Overall, CREW participants lived in tracts that were 67% White, 23% Black, 2% Asian, and 9% Other race. There was, however, variability in tract racial distribution both within and across cohorts (Table 1); eight cohorts enrolled participants from tracts where more than 75% of the population was White, while participants enrolled in URECA, WHEALS, and CCCEH resided in census tracts with populations having a greater proportion of Black or Other race. Most cohorts enrolled participants from tracts having a Hispanic population less than 10%, though the mean Hispanic population in tracts of CCCEH, IIS, TCRS, and URECA (Boston, MA and New York, NY sites) participants ranged from 21 to 57%. The overall mean of households in poverty was 15% in census tracts where CREW participants resided at birth but as shown in Figure 2, there was significant variability within and across cohorts. Additional information on census tract-level

https://doi.org/10.1017/cts.2021.7 Published online by Cambridge University Press

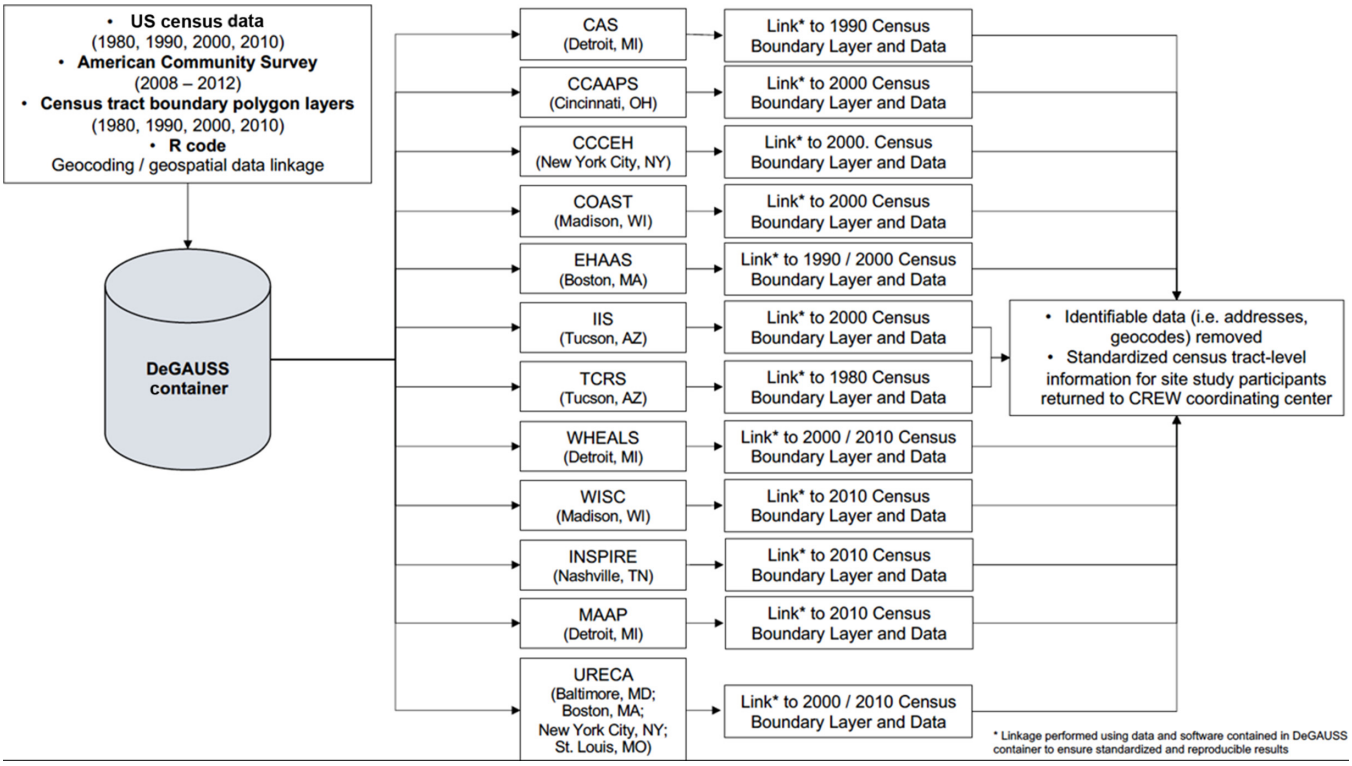


Fig. 1. Flow diagram of distributed geospatial analyses for the Children's Respiratory and Environmental Workgroup (CREW) consortium. CAS, Childhood Allergy and Asthma Study; COAST, Childhood Origins of Asthma Study; CCAAPS, Cincinnati Childhood Allergy and Air Pollution Study; CCCEH, Columbia Center for Children's Environmental Health; EHAAS, Epidemiology of Home Allergens and Asthma Study; INSPIRE, Infant Susceptibility to Pulmonary Infections and Asthma Following RSV Exposure; IIS, Infant Immune Study; MAAP, Microbes, Allergy, Asthma, and Pets; TCRS, Tucson Children's Respiratory Study; URECA, Urban Environment and Childhood Asthma; WHEALS, Wayne County Health, Environment, Allergy, and Asthma Longitudinal Study; WISC, Wisconsin Infant Study Cohort.

median household income, percentage (%) of households in poverty, % occupied housing, and median housing value as indicators of neighborhood income and housing is provided in the supplementary materials.

The distribution of tract-level race data (%White, %Black, % Asian, %Other) according to participants self-reported race is presented in Figure 3. Overall, participants who reported being White or Asian race lived in census tracts with majority White populations, whereas participants who reported Black race resided in census tracts with greater variability in shares of White and Black populations (Figure 3). Additional comparisons of individual-level to neighborhood-level income and ethnicity are provided in the supplementary material.

Discussion

Environmental exposures and the community in which they occur are significant causes of human disease, including asthma [7,14–17]. Disentangling the environmental and social exposures that contribute to health disparities necessitates geographically and culturally diverse studies. The methods described here make the characterization of community characteristics in a confidential yet standardized and reproducible manner more feasible for researchers and policy-makers. Importantly, our method is generalizable to additional types of geographic data, including polygon and point data, allowing other studies to customize and incorporate geocoding and geospatial analyses into their approach.

Our distributed geospatial approach offers some important advantages compared to existing methods, including reduced exposure misclassification, maintaining participant confidentiality, and reducing the need for geospatial expertise at each study site [1]. One existing and alternative method to maintain subject confidentiality is the alteration of participants' geographic coordinates. Referred to as "geomasking" or "jittering," this approach involves either a random shift in the location of subjects or a systematic transformation of the locations known only to the researchers [4]. However, this method may introduce errors or biases introduced due to the displacement of the participants' actual location, particularly for analyses requiring an exact location. For example, the amount of geomasking required to make geographic datasets de-identified according to HIPAA standards may result in incorrect census tracts being linked to individual subjects resulting in exposure misclassification. An alternative approach to incorporating geospatial information into multisite studies is to obtain IRB approval and DUAs to share subjects' identifiable information with a central site for geocoding and analysis. Challenges with this strategy include hesitation on the part of institutions to share identifiable information (e.g. addresses). Performing geospatial linkages at individual sites is another approach but may produce non-standardized and non-reproducible results due to the use of varying geocoding platforms, software, and dataset.

Our DeGAUSS method overcomes these limitations because the geospatial data and software are developed at a central site, ensuring that all individual study sites run the same software on identically constructed datasets. Importantly, our method also accounts for spatial, temporal, and informational changes in census tract boundaries and census data over the study period.

https://doi.org/10.1017/cts.2021.7 Published online by Cambridge University Press

https://doi.org/10.1017/cts.2021.7 Published online by Cambridge University Press

Table 1. *Summary of census tract-level population, race, ethnicity, and income for CREW participants at birth record address*

			Census variable (mean, SD)							
			Population	Race				Ethnicity	Income	
Cohort	Participants with deocoded birth address	Census year	Population density (/km^2)	% White	% Black	% Asian	% Other	% Hispanic	Median household income (2012 USD)	% Below poverty level
CREW	8810	1980/1990/2000/2010	5734 (13,897)	67 (34)	23 (32)	2 (3)	9 (14)	13 (21)	53,930 (26,368)	15 (14)
CAS	825	1990	1457 (970)	93 (13)	4 (13)	2 (2)	1 (1)	1 (1)	73,449 (26,708)	6 (5)
CCAAPS	762	2000	1346 (938)	79 (28)	17 (28)	1 (2)	2 (1)	1 (1)	64,895 (27,632)	10 (15)
CCCEH	724	2000	45,772 (1956)	16 (9)	41 (28)	2 (2)	42 (21)	57 (27)	31,171 (7308)	34 (9)
COAST	280	2000	877 (882)	89 (11)	4 (5)	3 (6)	4 (3)	4 (4)	71,090 (20,443)	5 (6)
EHAAS	499	1990/2000	3944 (3936)	83 (23)	9 (19)	5 (4)	3 (6)	5 (7)	86,218 (33,980)	8 (8)
IIS	481	2000	1380 (880)	79 (12)	4 (2)	3 (1)	15 (11)	23 (17)	57,774 (23,415)	9 (8)
INSPIRE	1921	2010	628 (692)	73 (25)	20 (23)	2 (2)	6 (5)	7 (7)	47,606 (20,237)	16 (12)
MAAP	141	2010	1165 (700)	82 (19)	11 (17)	4 (5)	3 (2)	4 (3)	69,042 (27,253)	9 (7)
TCRS	1142	1980	1172 (813)	85 (14)	3(4)	1 (1)	10 (12)	21 (21)	51,535 (17,849)	12 (8)
URECA - Baltimore	163	2000/2010	6480 (3468)	16 (22)	79 (26)	1 (2)	4 (5)	3 (7)	33,202 (12,999)	28 (13)
URECA – Boston	143	2000/2010	8244 (4544)	30 (26)	46 (29)	5 (6)	20 (11)	25 (17)	44,680 (20,190)	23 (11)
URECA - New York	118	2000/2010	33,137 (14,409)	26 (15)	41 (19)	3 (3)	30 (13)	51 (20)	30,967 (17,645)	31 (12)
URECA - St. Louis	178	2000/2010	2375 (1292)	21 (25)	76 (27)	1 (2)	3 (8)	2 (3)	31,467 (10,297)	28 (11)
WHEALS	1248	2000/2010	2376 (1198)	41 (38)	53 (41)	2 (4)	5 (6)	4 (10)	49,092 (24,995)	20 (15)
WISC	185	2010	148 (301)	96 (3)	1 (1)	1 (1)	3 (3)	3 (3)	50,871 (9775)	9 (5)

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

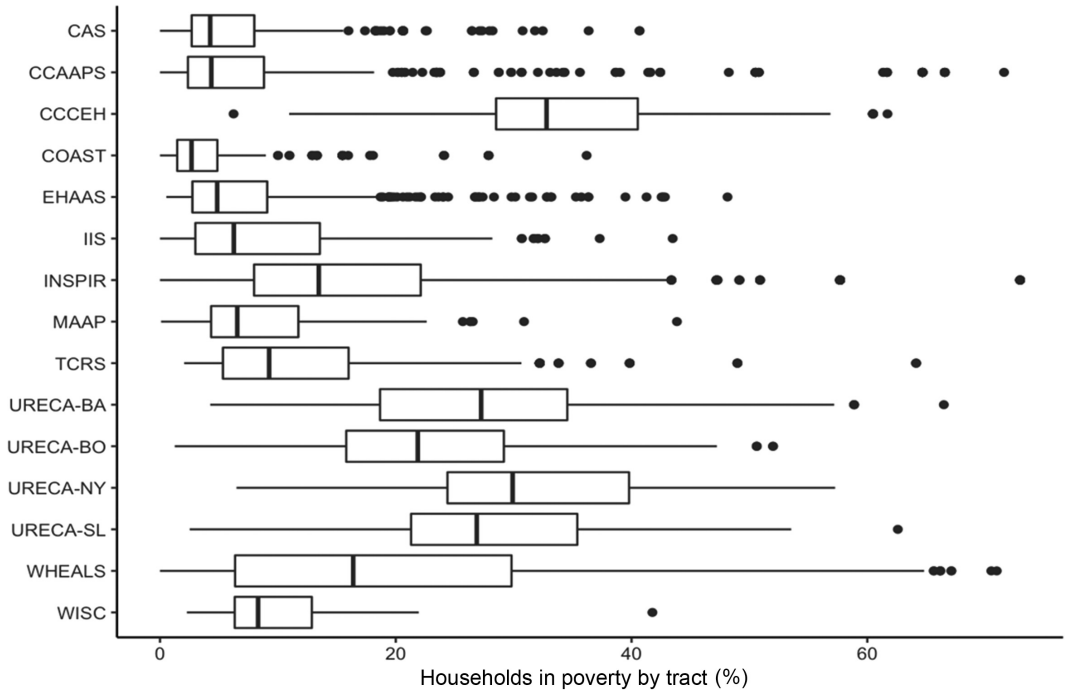


Fig. 2. Percentage* of household poverty in census tacts where CREW participants resided at birth. CAS, Childhood Allergy and Asthma Study; COAST, Childhood Origins of Asthma Study; CCAAPS, Cincinnati Childhood Allergy and Air Pollution Study; CCCEH, Columbia Center for Children's Environmental Health; EHAAS, Epidemiology of Home Allergens and Asthma Study; INSPIRE, Infant Susceptibility to Pulmonary Infections and Asthma Following RSV Exposure; IIS, Infant Immune Study; MAAP, Microbes, Allergy, Asthma, and Pets; TCRS, Tucson Children's Respiratory Study; URECA, Urban Environment and Childhood Asthma (BA, Baltimore; BO, Boston; NY, New York; SL, St. Louis), WHEALS, Wayne County Health, Environment, Allergy, and Asthma Longitudinal Study; WISC, Wisconsin Infant Study Cohort.

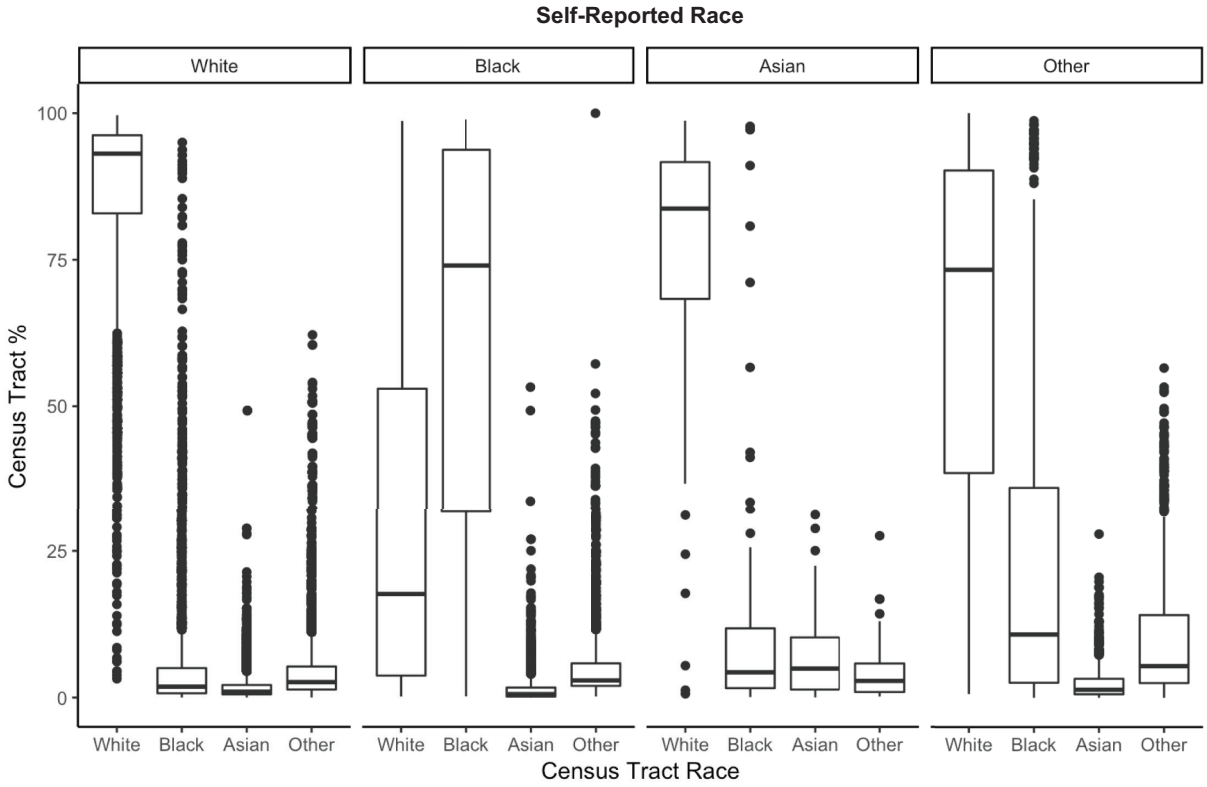


Fig. 3. Comparison of census tract race to participants' self-reported race.

https://doi.org/10.1017/cts.2021.7 Published online by Cambridge University Press

Participants in CREW cohorts were born over nearly four decades, and, as detailed in the accompanying supplement, linking geocoded addresses with the appropriate census tract at the time of participant birth was critical to our approach. By including census tract boundaries and accompanying data from 1980, 1990, 2000, and 2010 in our DeGAUSS software, we were able to link study participants to the appropriate census tract and data for their date of birth. Of note, the DeGAUSS platform is flexible and amenable to additional geographic data and analyses, including derivation of nearby greenspace, distance to roadways and other locations, estimating gridded air pollution exposures, area-based indices of neighborhood deprivation, and others. Specific details and example uses are available on the DeGAUSS website [12,13].

As part of the NIH Environmental influences on Child Health Outcomes (ECHO) program, the objective of CREW is to understand the etiology of childhood asthma and determine environmental and genetic contributors. As such, CREW offers a unique opportunity to examine the significantly higher rates of asthma prevalence, hospitalization, and morbidity among children residing in households and neighborhoods with lower SES, as compared to White children residing in higher socioeconomic status (SES) households and communities [18–20]. Multiple influences contribute to these disparities including disproportionately higher exposure to air pollution, poor housing, limited access to care, and other adverse physical and toxicant contributors, and inherited factors may increase individuals' sensitivity to these [21–25]. However, environmental exposures and inherited factors alone do not fully explain the observed health disparities. For that reason, hardships linked to poverty, including discrimination, stress, family turmoil, violence, instability, and others have been posited to play an important role in the social patterning of disease [26]. These social determinants of health may be separate from access to medical care and are important drivers not only in asthma morbidity, but also in a wide range of both adult and pediatric health outcomes, including mortality [27,28]. The importance of socioeconomic factors is highlighted by observations that disparities in health outcomes across strata of SES remain present within racial/ethnic groups [28].

Our methods and results should be considered, however, in the context of some limitations. Census tract boundaries alone may not accurately describe individuals' neighborhood experience and the use of decennial census data rather than ACS data resulted in reduced temporal precision but allowed us to increase our historic reach to 1980. There are also limitations to the US Census data including a lack of specificity in certain variables such as ethnicity. For example, Hispanic participants in the TCRS and IIS in Tucson, AZ, likely differ in ancestry from Hispanic participants from cohorts in New York City, NY. Finally, our approach requires expertise at individual institutions to obtain patient or electronic health records.

In conclusion, we demonstrated the use of a distributed approach to conduct geospatial analyses for the 12 CREW cohorts that is also applicable to other multisite studies. Future applications of our method will include additional gridded data including land cover, air pollution models, and meteorological information. Future research to understand the etiology of childhood asthma will incorporate longitudinal residential locations throughout childhood and multilevel analyses of individual and neighborhood environments.

Supplementary material. To view supplementary material for this article, please visit https://doi.org/10.1017/cts.2021.7.

Acknowledgments. 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). The authors would particularly like to thank Michael Evans, Andrew Moseby, Zhouwen Liu, Brent Olson, Suzanne Havstad, Soma Datta, Howard Andrews, and Agustin Calatroni for running the DeGAUSS software to link census and address data at their cohort sites.

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 Riley, and Jacqueline 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, and 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, and Lynn Taussig.

Infant Immune Study (IIS): Anne L. Wright*, Fernando D. Martinez*, Wayne Morgan,Debra A. Stern, Dean Billheimer, Brian Hallmark, Paloma Beamer, Nathan Lothrop, Lynn Taussig, 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, and 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, and 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, and 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, and 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, and 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, and Jaqueline Jezioro; Mount Sinai School of Medicine, New York, NY: H Sampson, R Sperling, and 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, and 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, and 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, and B. Stoffel; National Institute of Allergy and Infectious Diseases, Bethesda, MD: P Gergen, A Togias, E Smartt, and K Thompson.

Cincinnati Childhood Allergy and Air Pollution Study (CCAAPS): Gurjit K. Khurana Hershey*, Patrick H. Ryan*, Jocelyn M. Biagini Myers*, Grace K. LeMasters*, Kristi Curtsinger, Liza Murrison*, Jeffrey W. Burkle, Christopher Wolfe, Zachary Flege, David Morgan, Kristina Keidel, Krista Tensing, and Taylor Groeschen.

The Infant Susceptibility to Pulmonary Infections and Asthma Following RSV Exposure (INSPIRE): Vanderbilt University Medical Center, Nashville, TN: Tina V. Hartert MD, MPH**, 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 Gebretsadik MPH*, Kayla Goodman, 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, Emory University, Atlanta, GA: Larry J. Anderson MD, Tatiana Chirkova 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, and G Stouffer.

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

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

Biomedical Informatics and Biostatistical Core (BBC), University of Wisconsin-Madison: Umberto Tachinardi*, Mark Craven*, Eneida Mendonca, Lisa Gress, Adam Nunez, and Laura Ladick; University of Manchester: Philip Couch, Camille Johnson, John Ainsworth, Victoria Turner; Imperial College, London, UK, Adnan Custovic.

Genetics Core, University of Chicago. 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, and Jaime Hart.

Microbiome Core, University of California-San Francisco: Susan Lynch* and 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.

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

Columbia University: P01 ES09600, R01 ES008977, P30 ES09089, R01 ES013163, R827027.

Tucson Children's Respiratory Study (TCRS): R01HL132523, R01HL56177, K25HL103970.

Infant Immune Study (IIS): R01AI42268, K25HL103970.

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, NCRR/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.

Disclosures. The authors have no conflicts of interest to declare.

References

1. **Ajayakumar J, Curtis AJ, Curtis J**. Addressing the data guardian and geospatial scientist collaborator dilemma: how to share health records for spatial analysis while maintaining patient confidentiality. *International Journal of Health Geographics* 2019; **18**: 30.
2. **Boulos MN, Curtis AJ, Abdelmalik P**. Musings on privacy issues in health research involving disaggregate geographic data about individuals. *International Journal of Health Geographics* 2009; **8**: 46.
3. **Brownstein JS, Cassa CA, Kohane IS, Mandl KD**. Reverse geocoding: concerns about patient confidentiality in the display of geospatial health data. *AMIA Annual Symposium Proceedings* 2005; **2005**: 905.
4. **Zandbergen PA**. Ensuring confidentiality of geocoded health data: assessing geographic masking strategies for individual-level data. *Advances in Medicine* 2014; **2014**: 567049.
5. **Brokamp C, Wolfe C, Lingren T, Harley J, Ryan P**. Decentralized and reproducible geocoding and characterization of community and environmental exposures for multisite studies. *Journal of the American Medical Informatics Association* 2018; **25**: 309–314.
6. **Gern JE, Jackson DJ, Lemanske RF, *et al.*** The Children's Respiratory and Environmental Workgroup (CREW) birth cohort consortium: design, methods, and study population. *Respiratory Research* 2019; **20**(1): 115.
7. **Beck AF, Moncrief T, Huang B, *et al.*** Inequalities in neighborhood child asthma admission rates and underlying community characteristics in one US county. *The Journal of Pediatrics* 2013; **163**: 574–580.
8. **Akinbami LJ, Moorman JE, Liu X**. Asthma prevalence, health care use, and mortality: United States, 2005-2009. *National Health Statistics Reports* 2011; **2011**: 1–14.
9. **Akinbami LJ, Moorman JE, Garbe PL, Sondik EJ**. Status of childhood asthma in the United States, 1980-2007. *Pediatrics* 2009; **123**(Suppl 3): S131–S145.
10. **Akinbami LJ, Moorman JE, Simon AE, Schoendorf KC**. Trends in racial disparities for asthma outcomes among children 0 to 17 years, 2001-2010. *The Journal of Allergy and Clinical Immunology* 2014; **134**: 547–553 e5.
11. **Brokamp C**. DeGAUSS: Decentralized Geomarker Assessment for Multi-Site Studies. *Journal of Open Source Software* 2018; **3**(30): 812. doi: 10.21105/joss.00812
12. **Decentralized Geomarker Assessment for Multi-Site Studies**. [Internet] [cited Dec 17, 2020]. (https://degauss.org/)
13. **https://github.com/degauss-org/crew_census** [Internet] [cited Dec 17, 2020].
14. **Burbank AJ, Sood AK, Kesic MJ, Peden DB, Hernandez ML**. Environmental determinants of allergy and asthma in early life. *Journal of Allergy and Clinical Immunology* 2017; **140**: 1–12.
15. **Chen E, Schreier HM, Strunk RC, Brauer M**. Chronic traffic-related air pollution and stress interact to predict biologic and clinical outcomes in asthma. *Environmental Health Perspectives* 2008; **116**: 970–975.
16. **Clougherty JE, Levy JI, Kubzansky LD, *et al.*** Synergistic effects of traffic-related air pollution and exposure to violence on urban asthma etiology. *Environmental Health Perspectives* 2007; **115**: 1140–1146.

17. **Gupta RS, Zhang X, Springston EE**, ***et al.*** The association between community crime and childhood asthma prevalence in Chicago. *Annals of Allergy, Asthma & Immunology* 2010; **104**: 299–306.
18. **Gold DR, Wright R**. Population disparities in asthma. *Annual Review of Public Health* 2005; **26**: 89–113.
19. **Keet CA, Matsui EC, McCormack MC, Peng RD**. Urban residence, neighborhood poverty, race/ethnicity, and asthma morbidity among children on Medicaid. *The Journal of Allergy and Clinical Immunology* 2017; **140**: 822–827.
20. **Sullivan PW, Ghushchyan V, Kavati A, Navaratnam P, Friedman HS, Ortiz B**. Health disparities among children with asthma in the United States by place of residence. *The Journal of Allergy and Clinical Immunology* 2019; **7**: 148–155.
21. **Beck AF, Huang B, Chundur R, Kahn RS**. Housing code violation density associated with emergency department and hospital use by children with asthma. *Health Affairs (Project Hope)* 2014; **33**: 1993–2002.
22. **Ji H, Khurana Hershey GK**. Genetic and epigenetic influence on the response to environmental particulate matter. *The Journal of Allergy and Clinical Immunology* 2012; **129**: 33–41.
23. **Loss GJ, Depner M, Hose AJ**, ***et al.*** The early development of wheeze. Environmental determinants and genetic susceptibility at 17q21. *American Journal of Respiratory and Critical Care Medicine* 2016; **193**: 889–897.
24 **Martenies SE, Milando CW, Williams GO, Batterman SA**. Disease and health inequalities attributable to air pollutant exposure in detroit, Michigan. *International Journal of Environmental Research and Public Health* 2017; **14**(10): 1243.
25. **Rauh VA, Landrigan PJ, Claudio L**. Housing and health: intersection of poverty and environmental exposures. *Annals of the New York Academy of Sciences* 2008; **1136**: 276–288.
26. **Williams DR, Sternthal M, Wright RJ**. Social determinants: taking the social context of asthma seriously. *Pediatrics* 2009; **123**(Suppl 3): S174–S184.
27. **Wold Health Organization CoSDoH**. *Closing the gap in a generation: health equity through actionon the social determinants of health*. CSDH final report 2008. Geneva: WHO.
28. **Braveman P, Gottlieb L**. The social determinants of health: it's time to consider the causes of the causes. *Public Health Reports* 2014; **129**(Suppl 2): 19–31.

https://doi.org/10.1017/cts.2021.7 Published online by Cambridge University Press