

Title: Spatial and neighborhood data in the Collaborative Cohort of Cohorts for COVID-19 Research (C4R)

Author Names and Affiliations:

Jana A. Hirsch JAH474@drexel.edu (*Urban Health Collaborative and Department of Epidemiology and Biostatistics, Dornsife School of Public Health, Drexel University, Philadelphia, PA*)

Lilah M. Besser (*Department of Neurology, Miller School of Medicine, University of Miami, Miami, FL*)

Marcia Pescador Jimenez (*Department of Epidemiology, Boston University, Boston, MA*)

Stephen T. Dickinson (*Urban Health Collaborative, Dornsife School of Public Health, Drexel University, Philadelphia, PA*)

Talea Cornelius (*Center for Behavioral Cardiovascular Health, Columbia University Irving Medical Center, New York, NY*)

Stephen Francisco (*Urban Health Collaborative, Dornsife School of Public Health, Drexel University, Philadelphia, PA*)

Hoda S. Abdel Magid (*Department of Epidemiology and Population Health, Stanford University, Stanford, CA*)

Sandra Albrecht (*Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY*)

Norrina Bai Allen (*Institute for Public Health and Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL*)

Pallavi Balte (*Department of Medicine, Irving Medical Center, Columbia University, New York, NY*)

Sharrelle Barber (*The Ubuntu Center on Racism, Global Movements & Population Health Equity and Department of Epidemiology and Biostatistics, Dornsife School of Public Health, Drexel University, Philadelphia, PA*)

Lori Bateman (*Collaborative Studies Coordinating Center, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC*)

Jason P. Block (*Department of Population Medicine, Harvard Pilgrim Health Care Institute, Harvard Medical School, Boston, MA*)

Scott C. Brown (*Department of Public Health Sciences, Miller School of Medicine, University of Miami, Miami, FL*)

Russell Buhr (*David Geffen School of Medicine & Fielding School of Public Health at UCLA, Los Angeles, CA*)

Earle C Chambers (*Department of Family and Social Medicine and Department of Epidemiology & Population Health, Albert Einstein College of Medicine, New York, NY*)

Ann Chang (*Internal Medicine, University of California, San Francisco, CA*)

Ana V. Diez Roux (*Urban Health Collaborative and Department of Epidemiology and Biostatistics, Dornsife School of Public Health, Drexel University, Philadelphia, PA*)

Mitchell S. V. Elkind (*Department of Neurology, Vagelos College of Physicians and Surgeons, and Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY*)

Linda C Gallo (*Department of Psychology, San Diego State University, San Diego, CA*)

Penny Gordon-Larsen (*Department of Nutrition, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC*)

Jose Gutierrez, (*Department of Neurology, Vagelos College of Physicians and Surgeons, Columbia University, New York, NY*)

Carmen R Isasi (*Department of Epidemiology & Population Health and Department of Pediatrics, Albert Einstein College of Medicine, New York, NY*)

Peter James (*Department of Population Medicine, Harvard Medical School; Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, MA*)

Suzanne E. Judd (*Department of Biostatistics School of Public Health, University Alabama, Birmingham, AL*)

Alka Kanaya (*School of Medicine, University of California, San Francisco, CA*)

Namratha R. Kandula (*Department of General Internal Medicine and Epidemiology, Feinberg School of Medicine, Northwestern University, Chicago, IL*)

Joel D. Kaufman (*Department of Environmental and Occupational Health Sciences, University of Washington, Seattle, WA*)

Kiarri N. Kershaw (*Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL*)

Anna Kucharska-Newton (*Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC*)

Erin Kulick (*Department of Epidemiology and Biostatistics, College of Public Health, Temple University, Philadelphia, PA*)

Gina S. Lovasi (*Urban Health Collaborative and Department of Epidemiology and Biostatistics, Dornsife School of Public Health, Drexel University, Philadelphia, PA*)

Dave Mauger (*Department of Public Health Sciences, Pennsylvania State University, State College, PA*)

Leslie A. McClure (*College for Public Health and Social Justice, Saint Louis University, Saint Louis, MO*)

Steven Melly (*Urban Health Collaborative, Dornsife School of Public Health, Drexel University, Philadelphia, PA*)

Yvonne L. Michael (*Department of Epidemiology and Biostatistics, Dornsife School of Public Health, Drexel University, Philadelphia, PA*)

Gabriela R. Oates (*Heersink School of Medicine, University of Alabama at Birmingham, Birmingham, AL*)

Brenda Phillips (*Department of Public Health Sciences, Pennsylvania State University, State College, PA*)

Jessica A. Reese (*Department of Biostatistics and Epidemiology, University of Oklahoma Health Sciences Center, Oklahoma City, OK*)

Elizabeth A. Regan (*Division of Rheumatology, Department of Medicine, National Jewish Health, Denver, CO*)

Cameron Reimer (*Department of Epidemiology, Boston University, Boston, MA*)

Daniel A. Rodriguez (*Institute of Transportation Studies and Department of City and Regional Planning, University of California, Berkeley, CA*)

Tatjana Rundek (*Department of Neurology, Evelyn F. McKnight Brain Institute, Miller School of Medicine, University of Miami, Miami, FL*)

Mike Schembri (*School of Medicine, University of California, San Francisco, CA*)

Sharon Stein Merkin (*Geffen School of Medicine at University of California, Los Angeles, CA*)

Mario Sims (*Department of Social Medicine, Population and Public Health, University of California, Riverside School of Medicine, Riverside, CA*)

Carla Wilson (*National Jewish Health, Denver, CO*)

Yiyi Zhang (*Department of Medicine, Columbia University Irving Medical Center, New York, NY*)

Elizabeth Oelsner (*Department of Medicine, Columbia University Irving Medical Center, New York, NY*)

Target Journal: Health and Place

Counts:

- Title: 14 words; 96 characters with spaces
- Highlights: 5 bullets, each less than 85 characters including spaces
- Abstract: 112/100
- Words: 6626/ “4000-6000”
- Tables: 2
 - Table 1: Characteristics of C4R areas compared to non-C4R areas and the nation
 - Table 2: COVID context of C4R areas compared to the nation
- Figures: 2
 - Figure 1: Map of C4R coverage
 - Figure 2: Heatmap of C4R neighborhood measures
- Supplemental files:
 - Supplemental file 1: Basic cohort information
 - Supplemental file 2: Inventory form
 - Supplemental file 3: Known forthcoming data from cohorts

ABSTRACT

Neighborhood factors, encompassing social, built, and natural environments, may explain geographic differences in the impact of COVID-19 pandemic on populations. We catalogued spatial and environmental data in the Collaborative Cohort of Cohorts for COVID-19 Research (C4R), comprising 14 diverse US cohorts (>50,000 participants). C4R's calculated spatial coverage represented 28% of US land area and 52% of the total US population. C4R (vs. non C4R) areas were more urban, wealthy, with more foreign-born residents, and less car-dependent with lower proportion employed and green. Twelve cohorts collected neighborhood characteristics – most commonly social environment – based on participants' addresses. This work facilitates future harmonization of contextual data from C4R cohorts and neighborhood research on COVID-19 impact.

KEYWORDS

Neighborhood, Social Determinants of Health, Longitudinal, Harmonization, COVID-19, Geospatial, Measurement

INTRODUCTION

Social, natural, and built environments are powerful determinants of health behaviors, such as diet, and physical activity, and multiple outcomes, from cardiovascular and cerebrovascular diseases to brain health and cognitive decline (Besser, 2021, Besser et al., 2017, Smith et al., 2017, Malambo et al., 2016, Mayne et al., 2015, Mayne et al., 2021, Singh et al., 2019, Diez Roux et al., 2016, Michael et al., 2006, Won et al., 2016, Jimenez et al., 2021, Odden et al., 2023). By influencing available food resources, supportive locations for exercise or socialization, employment opportunities, access to nature, and more opportunities for healthful behaviors and activities, these environments impact individuals throughout their lives (Schulz and Northridge, 2004, Arcaya et al., 2016, Ellen et al., 2001). It is well-recognized that environmental and neighborhood supports are inequitably distributed; individuals who have lower income, have less formal education, or belong to marginalized groups live in neighborhoods with fewer resources and more exposure to environmental hazards (Conderino et al., 2021, Hirsch et al., 2016, Hirsch et al., 2023, Tsui et al., 2020, Liu et al., 2021b). As the coronavirus-19 (COVID) pandemic unfolded across the US, the burden of COVID infection, severity, poor recovery, and mortality were disproportionately high in low-income and lower resourced communities (Mude et al., 2021, McGowan and Bamba, 2022, Bryan et al., 2021). More evidence is needed on the intertwining links between disadvantage, environments, and COVID outcomes.

Extant research on how environmental factors within communities were associated with COVID outcomes (Weaver et al., 2022, Alidadi and Sharifi, 2022, Frumkin, 2021, Chen and Krieger, 2021) has demonstrated higher infection and mortality rates in larger metropolitan areas (Hamidi et al., 2020). Severity of COVID infection increased with an increase in air pollution (Bourdrel et al., 2021, Carballo et al., 2022, Marquès and Domingo, 2022), while reduced risk of COVID mortality was observed with increased greenspace exposure (Russette et al., 2021, Yang et al., 2022, Klompmaker et al., 2021). Evidence remains sparse and mixed for the impacts of density, land use mix, and other area-level factors such as food access or the social environment (Alidadi and Sharifi, 2022). Published studies have been limited in their geographic scope or participant diversity, which are important for understanding the associations of community and neighborhood characteristics with COVID outcomes. Furthermore, many prior studies are ecologic analyses which lack the individual longitudinal data necessary to disentangle existing health trajectories, health risks that affect COVID susceptibility and severity, and direct environmental impacts on COVID infection, outcomes, or long-term prognosis.

Several nationally representative, population-based longitudinal studies exist to understand complex risk factor development, progression of disease, and longitudinal, life-course predictors of health. Within the context of the COVID pandemic, these cohorts present a unique opportunity to examine how pre-existing health conditions and personal characteristics affected COVID risks and potential long-term changes in behaviors or health stemming from the pandemic period (e.g. decreased or increased physical activity from lockdown). Established early in the pandemic, the Collaborative Cohort of Cohorts for Covid-19 Research (C4R) brings together 14 cohort studies (**Supplemental File 1, list in abbreviations**) that have followed participants across the US for between 10 and 50 years. All cohorts are listed in abbreviations; details on each cohort and C4R published elsewhere (Oelsner et al., 2022). Briefly, C4R has collected information on COVID vaccination, infection, recovery, and mortality, as well as behavior and psychosocial outcomes from over 50,000 participants combined. Outside of the efforts of C4R, many of these cohorts have previously collected spatial and neighborhood contextual data about their participants, often over many exams and decades. These deeply phenotyped, well-characterized cohorts span diverse populations, life phases, and geographies, offering sufficient variability to better understand how environmental context impacts COVID and other health outcomes. Efforts to operationalize environmental data across different cohorts have been decentralized, often lead by independent investigators, resulting in variable data collection protocols and metrics (Lovasi et al., 2011, Brownson et al., 2009, Schaefer-McDaniel et al., 2010, Arcaya et al., 2016). Collaboration among cohorts and

harmonization of environmental data can help identify important factors that moderate COVID outcomes and disparities. Inventorying existing spatial, neighborhood data to map coverage is necessary for harmonization.

The goal of the current paper was to catalog available spatial data available from the C4R cohorts and describe the spatial extent of C4R participants. Our longer-term goal is to use findings from this paper to inform spatial/neighborhood measures harmonization and development in C4R. We expect this work will ultimately lead to well-powered research on the influence of spatial/neighborhood characteristics on COVID-19 outcomes and pandemic shifts across diverse geographies and participants. We expect these planned studies to have important implications for epidemiology, public health, planning, and policy.

METHODS

Cohort Geospatial Inventory

In 2021, three study team members (JAH, LMB, MPJ) drafted a geospatial inventory form to capture essential geospatial methods and measures within each C4R cohort (**Supplemental File 2**). The form documents each cohort's: (1) format and frequency of collecting and geocoding addresses, as well as the types of addresses collected (e.g., residential, work, school) and level of geocoding (e.g., latitude-longitude or Zip+4); (2) types of neighborhoods assessed, such as administrative data appended to geocodes (e.g., city, state, county) and the availability of buffers (e.g., Euclidian/straight line or network/street buffers); and (3) measures of the environment collected, created, or attached to participant geocodes. Specifically, the form assessed available neighborhood-level measures (i.e., already linked to participant locations) of the natural environment (e.g., greenspace, air quality), built environment (e.g., food stores, social and walking destinations, housing density), and social or economic environment (e.g., racial and ethnic composition, neighborhood residents' socioeconomic status). Within each of these three broad categories, the form asks about specific neighborhood metrics, the method of collection or derivation (e.g., survey, GIS, audit), the data source and year, the level of measurement (e.g., census tract or Euclidian buffer), and the time period(s) in which the measure was linked to locations (e.g., baseline, visit year). Lastly, the form collects COVID-specific neighborhood measures: number of cases, number of deaths, vaccination rates, policies, and disparity indices. For each COVID-specific neighborhood metrics, the form asks for the specific metrics available, the source of the data (e.g., survey, government agency), the level of measurement (e.g., city, zip), and the timing of measurement (e.g., by years, months). The draft inventory form was reviewed for clarity and completeness by three investigators and was piloted to assess usability in one of the C4R cohorts before finalization for use with the remaining 13 C4R cohorts.

Cohort Geospatial Data Acquisition Process and Ensuring Confidentiality

We (JAH, LMB, MPJ, STD, TC) contacted representatives for each of the 14 C4R cohorts requesting data manuals for each cohort's geospatial data, and offering a face-to-face meeting, if desired. Upon receiving data dictionaries/manuals from a cohort, information was extracted using the geospatial inventory form. We also looked for published manuscripts using measures of interest. This was done to minimize effort for each individual cohort. When each inventory was complete, we confirmed with the cohort that our assessment was accurate.

To provide data on the geographic coverage of C4R, we requested a list of geospatial units for participants' residential locations. We requested cohorts send the geographic administrative units to which participant addresses were geocoded (e.g., census tract, ZIP code, county, state) which were the smallest spatial resolution shareable without potential participant identification. While select cohorts have workplace locations, we did not include these. Geocoding data were provided to us in an anonymized

format (i.e., as a list of administrative geographies) and no individual participant data were linked to these locations.

Public Data Sources Linked to Geographic Coverage of C4R

Utilizing the ACS 2015-2019 5-Year Estimates and the geospatial data acquired for cohorts, we derived relevant population, social, and economic context information for a comparative analysis between areas where C4R participants have lived and those in which C4R participants have not lived. We used census tracts that match the 2015-2019 American Community Survey (ACS) boundaries, derived from the National Historical Geographic Information System (NHGIS) 2019 Census Tract Boundary File: United States and Puerto Rico (Manson et al., 2022). These measures linked included total population density (pop/km²), median household income, as well as percentages of older adults (≥ 65 years, ≥ 85 years), non-Hispanic Whites, foreign-born residents, home ownership (% rent, % own), residents below poverty, residents with formal educations (% with high school (HS) degree, % with college degree), residents employed, commuting by car, and crowding. Index of Concentration at the Extremes (ICE) was calculated from ACS data using standard methods (Krieger et al., 2016) for both the high-income White households versus low-income Black households and the high-income White non-Hispanic households versus low-income Hispanic households. ICE is scaled from -1 to +1 where -1 indicates that 100% of the population in the given area is concentrated in among the most deprived group, and +1 means that 100% of the population is concentrated in the most privileged group.

Data regarding urbanicity were derived from the McAlexander et al. community type classification (McAlexander et al., 2022). We operationalized urbanicity as the percent of total tracts that were classified as “high density urban”, “low density urban”, and “small town/suburban”, which make up all categories other than “rural” (and any undesignated tracts) as classified by McAlexander et al. To measure hospital access, we used staffed hospital bed counts per county from the Area Health Resources Files 2018-2019 to create a number of beds per 100,000 population density measure (Health Resources & Services Administration, 2022). Data on greenness was derived from the 2019 National Land Cover Database (NLCD) (Dewitz and U.S. Geological Survey, 2021) by including six categories of land cover previously used to generate a greenness measure at the census tract level (Yang et al., 2022). We reclassified the categories into one single “green” category and used zonal statistics to get the percent areal coverage per census tract.

We captured metrics of neighborhood COVID burden from two sources. Using the Johns Hopkins COVID-19 dashboard (cases per quarter), and the Centers for Disease Control (CDC) COVID Data Tracker (vaccination rates) (Dong et al., 2020, Centers for Disease Control and Prevention, 2022). We aggregated daily COVID-19 case counts per county from the Johns Hopkins COVID-19 Dashboard to quarterly counts then used population to create a quarterly cases per 100,000 population density measure. The CDC publishes data on COVID-19 vaccination rates per county, which we used to indicate completed primary series vaccination rates at the end of each quarter (Jan-Dec). We examined case rates from Q1 of 2020 until Q4 of 2021, and vaccination rates from Q1 of 2021 (when a vaccine was available to a majority of the population) to Q4 of 2021. Despite access to data for 2022, we did not examine these rates due to a lack of reporting that likely makes these data not representative of overall COVID levels.

Analyses

Inventory data on neighborhood measures within each cohort were grouped and summarized based on overall availability and construct. For example, since many cohorts did not obtain neighborhood housing type measures (i.e., age of housing, mix of multi-family vs. single family), but more collected home ownership (i.e., % within neighborhood who own or rent), we combined these into a single category (i.e.,

“Home ownership/housing conditions”). Similarly, many cohorts measured greenspace through park data or through satellite images, making it harder to differentiate between greenness measures and park access (which are related), so these were combined into “Greenspace/Parks.” Within each category, measures were counted and summarized based as 0 measures, 1 measure, 2-5 measures, 6 or more measures. For example, for “Access to Food Stores/Food Environment,” a cohort might have data on densities of fast-food establishments, restaurants, supermarkets, convenience stores, wholesale markets, and fruit and vegetable stands, as well as ratios of types of stores (e.g., healthy to unhealthy sources). This would count as 6+ measures. Similarly, data were summarized as being derived from GIS (e.g., the densities in the prior example) or derived from surveys to participants (e.g., self-reported access to fruit and vegetables within their neighborhood). We did not receive information from the cohorts indicating that they have audited any environmental characteristics using in-person audit tools. Finally, cohorts either had measures from only one time point (e.g., baseline or during an ancillary study) or multiple time points longitudinally (e.g., continuously, at each exam); therefore, we summarized data within each category using that dichotomy. All neighborhood measures were visualized in a heatmap that could illustrate data timing (color family), number of measures (intensity of color), and type of measures (pattern). Cleaning and summarization of inventory data was done in Excel, and the heatmap was created in R (ggplot2 package, RStudio 2023.06.1 with Base R version 4.3.0).

We received participant locations from each cohort in different aggregated geographic administrative levels and vintages. For example, one cohort provided counties (2015) in which participants had lived, while another provided tracts from four separate decennial censuses, ranging from 1980 to 2010. To harmonize these disparate boundaries, we merged and dissolved all files into one dataset and then overlaid it with 2019 census tracts from NHGIS. Tracts with at least 5% of their area overlapped by C4R participant geographies were included as “C4R areas” in the final tract-level datasets for comparison. All counties that these tracts fell in were used in final county-level datasets. Remaining census tracts and boundaries in the contiguous US were considered “non-C4R areas.” Only the contiguous US was considered for any aspect of this study, meaning that “national” measures refer to data for that subset only. All spatial processing was performed in ArcGIS Pro 2.9 (ESRI, Redlands, CA).

All public data sources on social, economic, and other contexts were joined to either the county- or tract-level geographies on matching Federal Information Processing Standards (FIPS) codes. These were then aggregated nationally, for C4R coverage area and for non-C4R coverage area. To calculate representativeness and coverage, total population and land area measures were summed in the aggregation. All other ACS 2015-2019 measures were aggregated by means. Therefore, final aggregated measures are means of tract-level measures, not percentages or indices of that level. For example, the HS Degree (%) measure is the mean of all tract-level percentages of persons with at least a high school diploma, rather than the total percentage of persons with a high school diploma. We used this method of aggregation for all additional measures. We described means and standard deviations (SD) for population, sociodemographic, economic, and environmental context data for C4R locations, non-C4R locations, and nationally. We used t-tests to calculate p-values comparing the means of C4R locations to means of non-C4R locations. All aggregation and analyses were done in Python 3.9 (Wilmington, DE).

RESULTS

Address and Geolocation Data

COPDGene was the only cohort that did not record addresses on inception of the cohort, instead recording addresses during a subsequent exam year. Most cohorts updated their participants’ address information each time an exam was performed, with some cohorts updating information more frequently.

For example, REGARDS recorded an initial address at the beginning of the study, then asked participants to update their addresses annually and confirmed addresses every six months during a follow-up call.

While collection and maintenance of address data was (and still is) done by all 14 cohorts, fewer cohorts (n=12) have processed these participants' addresses to identify their geographic location (i.e., geolocation or geocoding to a latitude-longitude for linkage to other geographic layers). Twelve of the cohorts (ARIC, CARDIA, COPDGene, FHS, HCHS/SOL, JHS, MASALA, MESA, NOMAS, REGARDS, SHS, SPIROMICS) have derived geospatial measures for administrative boundaries (e.g., tracts, zip codes, or counties), and six cohorts (ARIC, CARDIA, HCHS/SOL, JHS, MESA, REGARDS) have derived measures for Euclidian or network buffers (e.g., 1km around residence). Cohorts varied in use of Euclidean or Network buffers and size of buffers around residential addresses; some cohorts calculated buffer sizes using metric (1km, 3km, 5km) and others used imperial (0.5 mi, 1mi, 3mi, 5mi) (data not shown). Buffers used were also specific to certain GIS measure (e.g., some cohorts measured distance to nearest roadway, which is not a buffer, while others calculated densities of road network within a buffer around a participant's residence).

C4R Geospatial Coverage

C4R cohort participants covered 28% of the US land area, which includes 52% of the US total population (**Table 1**). Several locations within the US have participants from multiple C4R cohorts ("cohort overlap") (**Figure 1**). These areas include San Francisco, CA, Minneapolis, MN, Chicago, IL, Birmingham, AL, Winston-Salem, NC, Baltimore, MD, and New York, NY. These cities represent potential opportunities to harmonize and attach raw environmental GIS data available in one cohort to another cohort that may not have these data.

Neighborhood Measures

In our inventory of the C4R cohorts' geospatial data, twelve cohorts (ARIC, CARDIA, COPDGene, FHS, HCHS/SOL, JHS, MASALA, MESA, NOMAS, REGARDS, SHS, SPIROMICS) indicated that they have measured neighborhood geospatial data for their study participants (**Figure 2**). Two cohorts (FIP/PrePF, SARP) indicated they have not collected, processed, or linked any neighborhood geospatial data for their participants (blank columns, **Figure 2**).

Built Environment Data

Built environment data were collected by nine cohorts (ARIC, CARDIA, FHS, HCHS/SOL, JHS, MASALA, MESA, NOMAS, REGARDS). The most common data collected by cohorts was access to food stores or food environment. These cohorts measured neighborhood amenities for retail food stores, such as wholesale clubs, grocery and convenience stores, chain stores, co-ops, fast-food, specialty, meat and fish, and supermarkets. Three cohorts (JHS, MESA, REGARDS) utilized data from the National Establishment Time Series (NETS) dataset to calculate changes in the business environments of their cohorts over time (Hirsch et al., 2021). This dataset allows for cohorts to track the locations and types of businesses dating back to 1990, enabling longitudinal neighborhood change analysis. CARDIA, HCHS/SOL, JHS, MESA, and REGARDS all retained data on commercial buildings, land use, walkability, recreation/physical activity facilities, and alcohol/tobacco/marijuana outlets. These data were often derived from the NETS dataset, InfoUSA/ReferenceUSA, or similar dataset offerings from ESRI (i.e., ArcGIS Business Analyst). Traffic and road infrastructure, often primarily calculated using distance to arterial roads, were measured by seven cohorts (ARIC, CARDIA, HCHS/SOL, JHS, MESA, NOMAS, REGARDS); this was disparate from street connectivity, measured using both density of intersections, ratios of Network to Euclidean buffers, and other measures of connectivity.

Social Environment Data

Social environment data were the most plentiful data type connected to the C4R cohorts. Most cohorts have collected aggregated data on neighborhood socioeconomic status and data on urbanicity and population density, with all cohorts in possession of neighborhood-level data reporting some level of socioeconomic status. Most commonly, cohorts linked participant data to neighborhood employment data, median household income, education levels of residents, and poverty derived from the US decennial census and ACS. Cohorts that measured urbanicity of their participants' communities often relied on rural-urban commuting area (RUCA) codes, tract-based classification that uses standard census measures of population density, levels of urbanization, and journey-to-work commuting to characterize all US census tracts with respect to their rural/urban status (Economic Research Service, 2022).

Five cohorts reported measuring levels of neighborhood social engagement and cohesion (CARDIA, JHS, HCHS/SOL, MASALA, MESA). As expected for these types of constructs, these measures were assessed using survey tools administered to cohort participants. Uniquely, MESA measured these constructs simultaneously at the community level through administration of the survey to nearby, non-MESA participants and then created neighborhood estimates within a set buffer of MESA participant residences using Bayesian estimation techniques (Mujahid et al., 2007). Five cohorts incorporated levels of violence and crime (HCSC/SOL, JHS, MASALA, MESA, REGARDS). Sometimes perceptions of violence or crime were measured (e.g., in MESA and their community survey), while other times crime data were GIS-based from local municipalities or the CrimeRisk Index (provided through ESRI from Applied Geographic Solutions) (Applied Geographic Solutions, 2023).

Natural Environment Data

Natural environment data were the least available construct of geospatial data available among cohorts. Eight measured air quality (ARIC, CARDIA, HCHS/SOL, JHS, MESA, NOMAS, REGARDS, SPIROMICS), five measured greenspace/parks (CARDIA, HCHS/SOL, JHS, MASALA, MESA, REGARDS), three measured bluespace (CARDIA, HCHS/SOL, REGARDS), and only two measured hilliness/elevation (CARDIA, MESA) and temperature/weather (CARDIA, MESA). Cohorts that measured air quality kept data on multiple pollutants, especially nitrogen dioxide, ozone, particulate matter less than 10 μm or 2.5 μm (PM_{10} , $\text{PM}_{2.5}$), and nitrous oxide and measured in many different ways (e.g., modeling, averages during specific time periods). Some cohorts collected air quality data from EPA air pollution sensors (Environmental Protection Agency, 2023), while many used more complex modeling techniques to estimate fine-scale air pollution (Kirwa et al., 2021). Cohorts that measured greenspace/parks with park access generally quantified park density, size, and characteristics frequently sourced from local parks departments, city governments, or ParkServe (Trust for Public Land, 2023). When measuring natural vegetation, cohorts often used the NLCD from the USGS (Dewitz and U.S. Geological Survey, 2021) or the Normalized Difference Vegetation Index (NDVI), a measure of vegetation calculated using ArcGIS from Landsat or MODIS satellite imagery (Casey et al., 2017). Three cohorts measured some type of bluespace data (CARDIA, HCSC/SOL, REGARDS), measuring distances to the nearest shoreline and surface land use that is coded as water or ice. The two cohorts (CARDIA, MESA) that measured temperature (e.g., average variation, extreme temperature events) and weather (e.g., precipitation levels) tended to source data from the National Oceanic and Atmospheric Administration (NOAA) local climatological data archive (National Centers for Environmental Information, 2023). Almost no cohorts measured light pollution or noise, so these were not included in Figure 2.

COVID Pandemic Contextual Data

To date, none of the surveyed cohorts have linked neighborhood-level COVID data to participants' addresses. While the C4R project gathered self-reported COVID within participants' friend and family networks, a large gap remains in connecting participants to the COVID levels or pandemic responses (e.g., policies) they experienced around their residences.

Forthcoming Geospatial or Neighborhood Data

On or before November 2022, several cohorts reported geospatial measures being developed by investigators (through funded projects, ancillary studies, or approved manuscript proposals). These measures do not appear in this manuscript as they are not yet readily available for outside researchers through the coordinating centers. Examples include new segregation measures in ARIC, new disaster and climate measures in MESA, and new segregation, redlining, and neighborhood disinvestment measures in JHS (**Supplemental File 3**).

Sociodemographic, Economic, Environmental, and COVID Representativeness of C4R Geographies

Areas where C4R participants have lived or presently reside are more dense, urban, affluent, foreign-born, and formally educated and less car-dependent, with lower proportions employed and green than the non-C4R locations and the US (**Table 1**). Variability exists across cohorts (not shown). However, unsurprisingly, since many cohorts are based at institutions located in major cities, the geographic coverage of C4R is more dense (2613.08 population/km²) and urban (82.69) than the entire US (2087.58 and 75.61, respectively) and non-C4R locations (1575.32 and 68.72, respectively, $p < 0.001$). Relatedly, C4R areas have higher percentages of residents with a college degree and higher median household income, but also have a higher proportion of residents renting, below the poverty line, and unemployed. Tracking with their higher urbanicity, C4R locations have lower percentage of residents commuting by car and more racial and ethnic diversity as represented by both higher proportion foreign-born residents and lower non-Hispanic White residents. Interestingly, using ICE comparing only high-income White households to low-income Black households showed C4R locations have less concentrated privilege, while using ICE comparing high-income non-Hispanic White households to low-income Hispanic households showed C4R locations to have more concentrated privilege than non-C4R locations or the national average.

Estimates for the community COVID context of C4R participant locations showed that they generally experienced higher rates of infection and slightly higher rates of vaccination (**Table 2**). For example, in Q2 and Q3 of 2020, when the pandemic first peaked, the case rates for C4R locations were 536.74 cases per 100,000 and 1498.62 cases per 100,000 people, respectively, while the national rates were 489.68 cases per 100,000 people and 1435.37 cases per 100,000 people, respectively. We observed lower case rates in C4R locations in Q4 of 2020 and Q4 of 2021 than the case rates of the nation and non-C4R areas. Vaccination rates were always higher in the C4R locations than for the nation and non-C4R areas.

DISCUSSION

In the current study, we cataloged the available geospatial and neighborhood data and geospatial coverage of the 14 cohorts included in C4R. Twelve of the 14 cohorts had at least one neighborhood measure available at the time of this inventory. Social environmental data were most commonly included, with most cohorts having data on at least one neighborhood measure of socioeconomic status. The most common built environment measures were related to food access, followed by other destination-based

measures such as walkability. Natural environment data were available for the fewest cohorts, with emphasis on air quality or greenspace. The C4R sample is generally spatially and socially representative of the overall nation, covering 28% of US land area and 52% of the US population. Areas where C4R participants have lived or presently reside are denser and more urban, leaning more affluent, foreign-born, and formally educated and less car-dependent, employed, and proportion green than non-C4R locations or the US as a whole, although the magnitude of most of these differences was modest.

Our inventory of neighborhood measures provides an unprecedented overview of existing environmental cohort data and paves the way for future harmonization of measures across cohorts for understanding the COVID pandemic. Comprehensive characterizations of pre- and mid-pandemic neighborhood conditions with sufficient geospatial variability and participant diversity enables interdisciplinary research examining the contribution of environmental factors to COVID risk and disparities, including infection, morbidity, mortality, and persistent symptoms, as well as vaccination uptake. Place-based differences in COVID impact by urbanization or city characteristics have been demonstrated across countries (Naudé and Nagler, 2022, González-Val and Sanz-Gracia, 2022) and within the US (Bilal et al., 2022). Examinations of how neighborhoods impacted COVID outcomes have already informed some previous prevention efforts, such as opening up park space or closing streets to create protected pedestrian spaces (Firth et al., 2021, Douglas and Moore, 2022), and can inform the creation of environments that are resilient to future public health disasters (Bilal et al., 2022, Liu et al., 2021a, Kashem et al., 2021). The geographic representation and sociodemographic diversity of the C4R cohorts and their combined sample size permits research into how policies or neighborhood contexts contribute to COVID and related health disparities. Emergency measures in other domains implemented by some locales, such as eviction moratoriums, could be helpful to health and used as a basis for future policies to protect populations (Bilal et al., 2022, Leifheit et al., 2021, Sandoval-Olascoaga et al., 2021, Versey, 2021, Ali and Wehby, 2022).

The geographic and individual diversity of C4R across multiple jurisdictions could also be used to understand how neighborhood features and policies differentially impact COVID and other health outcomes depending on participant characteristics such as race and ethnicity, sex/gender, age, and socioeconomic status. This may help answer inquiries about the conditions and populations for which policies are most successful. Ultimately, while the COVID pandemic intensity wanes, lessons learned about neighborhoods during the COVID pandemic and recovery period can be applied to enhance future disaster preparedness (Bakerjian and Nguyen, 2022) as long as efforts are made to integrate individual-level risk information to avoid both ecological bias (Bakerjian and Nguyen, 2022) and confounding by prior conditions. Longitudinal data from C4R helps to overcome some of these limitations, allowing for better control of changing ecological conditions over time.

Harmonization of C4R cohorts' neighborhood environment and contextual data can additionally be performed and leveraged to understand critical shifts that occurred in built, social, and natural environments *post*-pandemic and their subsequent health impacts. These may be primary shifts that occurred early during the pandemic (e.g., policies around eviction, stay-home orders, pedestrianized streets) or enduring changes to retail environments, physical spaces, economic conditions, or population distributions. Evidence suggests that early pandemic prevention measures lowered pollution levels, limited access to public spaces, closed retail establishments, and produced lost income (Silva et al., 2021, Berman and Ebisu, 2020, Dang and Nguyen, 2021). A number of cities made intentional health-supportive changes to their environments (e.g., pedestrianized streets, parklets), with varying success retaining these as the pandemic has progressed (Rhoads et al., 2021, Rhoads et al., 2020, Rojas-Rueda and Morales-Zamora, 2021). Similar evidence has begun to emerge regarding shifts in health-related contexts and social environments, including crime (Abrams, 2021, Ashby, 2020, Boman and Gallupe, 2020), neighborhood trust or social cohesion (Jiang et al., 2022, Zangger, 2021), and movement patterns (Haslag and Weagley, 2022, González-Leonardo et al., 2022, Coven et al., 2023). As the country and world ease

into a “new normal,” it is essential to understand long-term impacts of the pandemic period on our lived experience, including neighborhood conditions. Connecting C4R neighborhood data to the extensive pre-pandemic (1975-2020) and planned (2020-present) deep phenotyping in the cohorts on physical function, cardiac measures, brain measures, sleep, biomarkers, and genotyping, among others (Oelsner et al., 2022), overcomes confounding by lack of knowledge regarding previous health behaviors and trajectories. Links with neighborhood data would provide a singular opportunity to understand how the COVID pandemic and neighborhood or community changes impacted other health conditions commonly assessed across the cohorts, such as cardiovascular risk, aging, and related diseases.

Despite the myriad opportunities ignited by available and potential C4R neighborhood data, obstacles exist for harmonization and future work. First, the pandemic disrupted data collection and infrastructure systems across the country and world. Specific to the US, there are known concerns about the quality and viability of Census 2020 and American Community Survey data that include the year 2020 (i.e., all five-year estimates from 2016-2025). This is similar for any other measure traditionally collected and processed in-person (e.g., building inspection measures), but is less likely to impact remote sensing measures (e.g., greenness). Furthermore, the validity and meaning of these measures may have changed due to pandemic restrictions (e.g., land uses designated for commercial purposes but were closed and their public functions suspended). Second, although many cohorts have measures within a domain, these measures use different metrics, were collected at different levels of resolution (e.g., counties v. census tracts), or were based on different underlying administrative and feature data. For example, some cohorts may source land use measures from parcel data collected in specific jurisdictions while others use national satellite imagery, or some may calculate intersection density for street networks while others use block length or a ratio of Network buffers to Euclidean buffers. Existing survey data are rich and comprehensive but very sparse, and without it, we may lack the ability to understand individuals’ perceptions of neighborhood amenities or culture pre-, mid-, and “post”-pandemic. Even within cohorts that collect survey data on perceived safety, social cohesion, and belonging, there are no signs of convergence toward a single standard set of tools. These inconsistencies make harmonization of existing data challenging. In addition, in several cohorts, neighborhood and spatial data were created, stored, and managed by ancillary studies rather than centrally within the coordinating center that currently interfaces with C4R. Harmonization is further hindered by the sensitive nature of spatial data; there are human subjects concerns when sharing geospatial data such as addresses, and the various cohorts and university IRBs may have different restrictions and concerns about the identifiability of participant latitudes-longitudes.

Despite privacy concerns, options for harmonization exist that circumvent some of these obstacles. One option is to create national-level metrics at a small administrative scale (i.e., block groups or tracts) that could be linked by each cohort to their participant locations. Alternatively, the creation of a national set of gridded metrics (i.e., as GIS rasters) would facilitate linkage to locations across a smooth, interpolated surface. These options create a balance between consistency across cohorts and ensuring data are backwards compatible, especially for measures where there is little to no ability to create historic metrics. While these workarounds may help with privacy concerns, some cohorts have limited internal resources, experience, and trained staff for managing geographic or spatial data to do the final linkage. Although harmonization is important to study neighborhood impact on COVID, pandemic impacts on everyday contexts, and their subsequent effects on health, these are only one arm of the many pressing scientific lines of inquiry possible for these cohorts. Ultimately, all cohorts are balancing different competing priorities and must remain accountable to their primary missions (e.g., cardiovascular, pulmonary), existing ancillary studies, and participants. Harmonization will require considerable financial resources across an extended timeline, and execution by a committed and experienced team of diverse researchers.

This paper provides a comprehensive description of the neighborhood and spatial data across cohorts, but is not without limitations. First, many cohorts were missing some of the details on software for geocoding

and, as in many transdisciplinary studies, we needed to establish common language around geography terminology and concepts (e.g., administrative boundary, uncertain geographic context problem, modifiable areal unit problem). Missing technical details may stem from structures in which neighborhood work is collected in ancillary studies and not distributed through the broader network of cohort researchers. When inventorying neighborhood data, the use of a survey and self-report by staff within cohort coordinating centers may introduce recall bias. We minimized these issues by collecting data documentation to confirm reported survey data and looking for published manuscripts using measures of interest. To facilitate harmonization, future work should prioritize documentation of software details and centralization of ancillary study datasets, code, and documentation at coordinating centers. In addition, as stated in our results, the inventory is a snapshot of what cohorts had in February-November 2022 and does not represent projects underway. For example, some cohorts had recently funded grants to incorporate new measures that were not yet calculated at the time of this inventory. We have tried our best to capture these forthcoming measures in the Supplemental files but may have missed upcoming research. Our calculation of spatial extent and national representation has limitations. Due to differences in human subject restrictions or levels of geocoding, cohorts sent address data at different spatial scales (e.g., cities, counties, census tracts) and across different spatial vintages (e.g., census 2000 tracts, census 2010 tracts). For anonymity, we only present data aggregated across all cohorts (i.e., counts of cohorts in a location, not the named cohort) and some cohorts additionally aggregated their data across time, providing all locations that a participant has lived across any exam. Cohorts with older adults (e.g., ARIC, MESA) have some participants who live in the southern US during the wintertime, and those wintertime locations are not necessarily captured in our presented data. Therefore, the data we presented on spatial representation combines time, spatial unit, and cohort, and the map of the geographic extent of C4R participants does not necessarily represent participants' locations during the COVID-19 pandemic. For research on specific exposure-outcome pairs, future work should gain approval and access to participant data on a smaller spatial and temporal resolution. The map also represents residential areas of C4R without weighting to account for differential counts of participants for given geographies (e.g., some census tracts may have one participant, some may have many). Similarly, our calculations of sociodemographic and COVID measures for the geographic extent required reaggregation of the spatial scale at which those data were provided (e.g., census tract for census, county for COVID data). This assumes an even distribution of each characteristic across the areal unit. Furthermore, we were limited to counties for national COVID data, and different reporting issues occurred at different points in the pandemic (lack of access to testing early in the pandemic, use of at-home tests later in the pandemic that may have not been reported to the city, etc.). Finally, to ease interpretation, all data presented are summaries and do not represent harmonized measures across the cohorts. More specific details on the availability of geospatial data for individual cohorts should be requested from the authors or directly from each cohort; access to these data should be requested through the approved processes of each cohort.

Despite these limitations, several strengths enhance the utility of this paper to guide future research on how environmental context shapes health outcomes and disparities. First, we used a centralized, systematic method for cataloging available geospatial data through a uniform data collection tool and process across all cohorts. Our graphical display of these data allows researchers to see potential synergies that would otherwise be difficult (or impossible) to recognize. In addition to sparking novel research questions and enabling well-powered analyses across diverse samples and regions, our clear summarization could be used as a model for other projects seeking to harmonize data across cohorts or as a guide for the creation of new measures (e.g., noise) that could be developed for one or more cohorts. Second, capturing key elements across these metrics, including differences in how data were collected (e.g., survey data vs. GIS data), allows researchers interested in utilizing these data to be informed about the inconsistencies across cohorts. This points researchers toward potential limitations of their analyses and provides information to interpret disparate study results across cohorts. Third, our systematic approach provides a roadmap for researchers who would like to collect their own geospatial data. For example, our collection process has already led to discussions across cohorts regarding best measures for

specific environmental features. We believe that this summarization will encourage selection and operationalization of measures and methods that increase the coherence of the evidence created. This would enable even more scientific rigor and lead to the creation of larger and more diverse datasets that can answer highly complex questions about the intersection of environment and human health. Finally, our diverse and interdisciplinary team of national leaders in neighborhood health demonstrates best practices for team science. Our process ensured that multiple viewpoints were represented, ultimately encouraging shared methods and promoting collaborations with a wide range of disciplines with great potential to improve public health.

CONCLUSION

In this study, we cataloged the geographic and spatial data available across a large geographically, socioeconomically, racially and ethnically diverse study of 14 pooled cohorts in the US. Our analyses described the spatial extent of the C4R population and examined representativeness compared to areas without C4R participants and compared to the nation overall, finding that C4R had relatively representative coverage of urban areas nationwide. We found that social environmental data were the most common types of contextual data collected, followed by built environment features. Our inventory also indicated that all cohorts were missing neighborhood COVID contextual data, several cohorts were lacking any geocoding or environmental data, and that even cohorts with contextual data had gaps in natural environment metrics, such as weather or heat, which may be relevant in an age of increasing climate-related events and vector-borne diseases (Caminade et al., 2019, Schiermeier, 2019, Strauss et al., 2021, Williams et al., 2019). Despite numerous advances in the assessment of neighborhood effects on health, our inventory of neighborhood measures highlights the need to develop a systematic methodology to facilitate harmonization. Agreement on common metrics, repositories of national environmental measures, and guidance documents for cohorts would encourage and enable valid and rigorous testing of the health effects of neighborhoods or their subsequent impacts on health disparities. Given the large quantity of data within the C4R project, methods will also need to be developed to integrate data across multiple spatial and temporal scales. Work remains to understand the implications of the COVID pandemic on our neighborhoods, the way we interact with them, and their subsequent impact on our health and well-being. During a time when the environmental context is rapidly changing due to human activities and substantial climate change shifts, lessons learned from the pandemic period may help us create more resilient, health-supportive, equitable communities. This project opens possibilities for research collaborations to understand the health effects of neighborhood changes experienced during the pandemic and work toward sustainable, thoughtful neighborhood interventions to support health and well-being as future disasters arise.

ACKNOWLEDGEMENTS AND FUNDING SUPPORT

C4R: C4R is supported by NHLBI-CONNECTS OT2HL156812

ARIC: The Atherosclerosis Risk in Communities study has been funded in whole or in part with Federal funds from the National Heart, Lung, and Blood Institute, National Institutes of Health, Department of Health and Human Services, under Contract nos. (HHSN268201700001I, HHSN268201700002I, HHSN268201700003I, HHSN268201700005I, HHSN268201700004I). Neurocognitive data are collected by U01 2U01HL096812, 2U01HL096814, 2U01HL096899, 2U01HL096902, 2U01HL096917 from the NHLBI, NINDS, NIA and NIDCD. Ancillary studies funded additional data elements. BP COG The BP COG Study is supported by the National Institute of Neurological Disorders and Stroke, National Institutes of Health, and US Department of Health and Human Services (grant R01 NS102715).

CARDIA: The Coronary Artery Risk Development in Young Adults Study (CARDIA) is conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with the University of Alabama at Birmingham (HHSN268201800005I, HHSN268201800007I), Northwestern University (HHSN268201800003I), University of Minnesota (HHSN268201800006I), and Kaiser Foundation Research Institute (HHSN268201800004I). This manuscript has been reviewed by CARDIA for scientific content. The neighborhood measures in CARDIA were developed with grants R01HL104580 and R01HL114091 from the NHLBI.

COPDGene: The project described was supported by Award Number U01 HL089897 and Award Number U01 HL089856 from the National Heart, Lung, and Blood Institute. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Heart, Lung, and Blood Institute or the National Institutes of Health. COPD Foundation Funding: COPDGene is also supported by the COPD Foundation through contributions made to an Industry Advisory Board comprised of AstraZeneca, Boehringer-Ingelheim, Genentech, GlaxoSmithKline, Novartis, Pfizer, Siemens, and Sunovion.

FHS: The Framingham Heart Study (FHS) acknowledges the support of Contracts NO1-HC-25195, HHSN268201500001I and 75N92019D00031 from the National Heart, Lung and Blood Institute. The use of geographic data in the FHS cohort was supported by R01 HL109262 (NHLBI) and P01 AG031093 (NIA). We also acknowledge the dedication of the FHS study participants without whom this research would not be possible.

HCHS/SOL: The Hispanic Community Health Study/Study of Latinos is a collaborative study supported by contracts from the National Heart, Lung, and Blood Institute (NHLBI) to the University of North Carolina (HHSN268201300001I, N01-HC-65233), University of Miami (HHSN268201300004I, N01-HC-65234), Albert Einstein College of Medicine (HHSN268201300002I, N01-HC-65235), University of Illinois at Chicago, HHSN268201300003I, N01-HC-65236 Northwestern Univ), and San Diego State University (HHSN268201300005I, N01-HC-65237). The following Institutes/Centers/Offices have contributed to the HCHS/SOL through a transfer of funds to the NHLBI: National Institute on Minority Health and Health Disparities, National Institute on Deafness and Other Communication Disorders, National Institute of Dental and Craniofacial Research, National Institute of Diabetes and Digestive and Kidney Diseases, National Institute of Neurological Disorders and Stroke, NIH Institution-Office of Dietary Supplements. Neighborhood measures in HCHS/SOL have been developed through HCHS/SOL ancillary study funding through NIH grant R01DK106209 (SOL-CASAS). Forthcoming neighborhood measures in HCHS/SOL are currently being developed through HCHS/SOL ancillary study funding through NIH grants R01HL148880 (SOL-Greenness), R01ES030994 (SOLAir), RC2HL101649 (SOL Sociocultural), R01HL148463, and (SOL San Diego CASAS Neighborhood Environments) R01DK106209.

JHS: The Jackson Heart Study (JHS) is supported and conducted in collaboration with Jackson State University (HHSN268201800013I), Tougaloo College (HHSN268201800014I), the Mississippi State Department of Health (HHSN268201800015I) and the University of Mississippi Medical Center (HHSN268201800010I, HHSN268201800011I and HHSN268201800012I) contracts from the National Heart, Lung, and Blood Institute (NHLBI) and the National Institute on Minority Health and Health Disparities (NIMHD). JHS neighborhood data were supported by the Center for Integrative Approaches to Health Disparities funded by National Institute of Minority Health and Health Disparities (NIMHD P60 MD002249), the Michigan Center for Integrative Approaches to Health Disparities funded by the NIMHD, by US Environmental Protection Agency and NIMHD: Environment Supplement (P60 MD002249 05S1), and from the NHLBI (R01HL148431).

MASALA: MASALA was supported by Grant Number R01HL093009 from the National Heart, Lung, And Blood Institute and the National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health, through UCSF-CTSI Grant Number UL1RR024131.

MESA: MESA and the MESA SHARe projects are conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with MESA investigators. Support for MESA is provided by contracts and grants 75N92020D00001, HHSN268201500003I, N01-HC-95159, 75N92020D00005, N01-HC-95160, 75N92020D00002, N01-HC-95161, 75N92020D00003, N01-HC-95162, 75N92020D00006, N01-HC-95163, 75N92020D00004, N01-HC-95164, 75N92020D00007, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169, R01-HL077612, R01-HL093081, R01-HL130506, R01-HL127028, R01-HL127659, R01-HL098433, R01-HL101250, and R01-HL135009 from the National Heart, Lung, and Blood Institute; R01-AG058969 from the National Institute on Aging; and by grants UL1-TR-000040, UL1-TR-001079, and UL1-TR-001420 from the National Center for Advancing Translational Sciences (NCATS). Funding for SHARe genotyping was provided by NHLBI Contract N02-HL-64278. This publication was developed under the Science to Achieve Results (STAR) research assistance agreements, No. RD831697 (MESA Air) and RD-83830001 (MESA Air Next Stage), awarded by the US Environmental Protection Agency. It has not been formally reviewed by the EPA. The views expressed in this document are solely those of the authors and the EPA does not endorse any products or commercial services mentioned in this publication. Whole genome sequencing (WGS) for the Trans-Omics in Precision Medicine (TOPMed) program was supported by the NHLBI. WGS for NHLBI TOPMed: Multi-Ethnic Study of Atherosclerosis (MESA) (phs001416.v1.p1) was performed at the Broad Institute of MIT and Harvard (3U54HG003067-13S1). Centralized read mapping and genotype calling, along with variant quality metrics and filtering were provided by the TOPMed Informatics Research Center (3R01HL-117626-02S1, contract HHSN268201800002I) (Broad RNA Seq, Proteomics HHSN268201600034I, UW RNA Seq HHSN268201600032I, USC DNA Methylation HHSN268201600034I, Broad Metabolomics HHSN268201600038I). Phenotype harmonization, data management, sample-identity QC, and general study coordination, were provided by the TOPMed Data Coordinating Center (3R01HL120393; U01HL-120393; contract HHSN268180001I). The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR001881, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center. The authors thank the other investigators, the staff, and the participants of the MESA study for their valuable contributions. A full list of participating MESA investigators and institutions can be found at <http://www.mesa-nhlbi.org>. NHLBI Pooled Cohorts Study: NIH/NHLBI R21HL153700, K23HL130627, R21HL129924, R21HL121457. MESA Neighborhoods is supported by the National Heart, Lung, and Blood Institute (R01HL071759), the National Institute on Aging (R01AG072634). Novel methods for neighborhoods research is supported by the National Heart, Lung, and Blood institute (R01HL131610).

NOMAS: Northern Manhattan Study is supported by the National Institute of Neurological Disorders and Stroke (NINDS R01NS29993) and the National Institute of Aging (NIA R01AG066162 and RF1AG074306). Forthcoming neighborhood measures in NOMAS are currently being developed through NOMAS ancillary study funding through RF1AG074306 (NOMAS-Greenness).

PrePF: Prevent Pulmonary Fibrosis cohort was established in 2000 and has been supported by: X01-ES101947, R01-HL095393, RC2-HL1011715, R21/33-HL120770, R01-HL097163, X01-HL134585, UH2/3-HL123442, P01-HL092870, UG3/UH3-HL151865, and DoD W81XWH-17-1-0597.

REGARDS: This research project is supported by cooperative agreement U01 NS041588 co-funded by the National Institute of Neurological Disorders and Stroke (NINDS) and the National Institute on Aging (NIA), National Institutes of Health, Department of Health and Human Service. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NINDS or the NIA. Representatives of the NINDS were involved in the review of the manuscript but were not directly involved in the collection, management, analysis or interpretation of the data. The authors thank the other investigators, the staff, and the participants of the REGARDS study for their valuable contributions. A full list of participating REGARDS investigators and institutions can be found at:

<https://www.uab.edu/soph/regardsstudy/> Neighborhood work was supported by the National Institute of Aging (grants R01AG049970, R01AG049970-S1, R01AG072634, R56AG049970), National Heart, Blood, and Lung Institute (grant R01HL14843), National Institute on Alcohol Abuse and Alcoholism (R01AA028552), Commonwealth Universal Research Enhancement (C.U.R.E) program funded by the Pennsylvania Department of Health – 2015 Formula award – SAP #4100072543, the Urban Health Collaborative at Drexel University, and the Built Environment and Health Research Group at Columbia University.

SARP: The research by the principal and co-principal investigators was funded by National Institutes of Health/National Heart Lung Blood Institute (NIH/NHLBI) Severe Asthma Research Program: U10 HL109164, U10 HL109257, U10 HL109146), U10 HL109172, U10 HL109250, U10 HL109168, U10 HL109152, U10 HL109086). Industry Partnerships also provided additional support: AstraZeneca, Boehringer-Ingelheim, Genetech, GlaxoSmithKline, MedImmune, Novartis, Regeneron, Sanofi and TEVA. Spirometers used in SARP III were provided by nSpire Health (Longmont, Colo).

SPIROMICS: SPIROMICS has been funded by contracts from the NIH/NHLBI (HHSN268200900013C, HHSN268200900014C, HHSN268200900015C, HHSN268200900016C, HHSN268200900017C, HHSN268200900018C, HHSN268200900019C, HHSN268200900020C), grants from the NIH/NHLBI (U01 HL137880 and U24 HL141762), and supplemented by contributions made through the Foundation for the NIH and the COPD Foundation from AstraZeneca/MedImmune; Bayer; Bellerophon Therapeutics; Boehringer-Ingelheim Pharmaceuticals, Inc.; Chiesi Farmaceutici S.p.A.; Forest Research Institute, Inc.; GlaxoSmithKline; Grifols Therapeutics, Inc.; Ikaria, Inc.; Novartis Pharmaceuticals Corporation; Nycomed GmbH; ProterixBio; Regeneron Pharmaceuticals, Inc.; Sanofi; Sunovion; Takeda Pharmaceutical Company; and Theravance Biopharma and Mylan. SPIROMICS AIR was supported by NIH NIEHS (R01ES023500).

SHS: The Strong Heart Study has been funded in whole or in part with federal funds from the National Heart, Lung, and Blood Institute, National Institute of Health, Department of Health and Human Services, under contract numbers 75N92019D00027, 75N92019D00028, 75N92019D00029, & 75N92019D00030. The study was previously supported by research grants: R01HL109315, R01HL109301, R01HL109284, R01HL109282, and R01HL109319 and by cooperative agreements: U01HL41642, U01HL41652, U01HL41654, U01HL65520, and U01HL65521.

The authors would like to thank Martha Daviglus, Priya Palta, and Neil Schneiderman for their contributions to these efforts.

ABBREVIATIONS

ARIC: Atherosclerosis Risk in Communities

C4R: Collaborative Cohort of Cohorts for COVID-19 Research

CARDIA: Coronary Artery Risk Development in Young Adults

CDC: Centers for Disease Control

COPDGene: Genetic Epidemiology of COPD

COVID: Coronavirus-19

FHS: Framingham Heart Study

GIS: Geographic Information Systems

HCHS/SOL: Hispanic Community Health Study/Study of Latinos

JHS: Jackson Heart Study

MASALA: Mediators of Atherosclerosis in South Asians Living in America

MESA: Multi-Ethnic Study of Atherosclerosis

NOMAS: Northern Manhattan Study

PrePF: Prevent Pulmonary Fibrosis

REGARDS: REasons for Geographic and Racial Differences in Stroke

SARP: Severe Asthma Research Program

SPIROMICS: Subpopulations and Intermediate Outcome Measures in COPD Study

SHS: Strong Heart Study

REFERENCES

- ABRAMS, D. S. 2021. COVID and crime: An early empirical look. *Journal of public economics*, 194, 104344.
- ALI, A. K. & WEHBY, G. L. 2022. State Eviction Moratoriums During The COVID-19 Pandemic Were Associated With Improved Mental Health Among People Who Rent: Study examines associations between the mental health of renters and state eviction moratoriums during COVID-19. *Health Affairs*, 41, 1583-1589.
- ALIDADI, M. & SHARIFI, A. 2022. Effects of the built environment and human factors on the spread of COVID-19: A systematic literature review. *Science of the Total Environment*, 850, 158056.
- APPLIED GEOGRAPHIC SOLUTIONS 2023. CrimeRisk. Thousand Oaks, CA.
- ARCAYA, M. C., TUCKER-SEELEY, R. D., KIM, R., SCHNAKE-MAHL, A., SO, M. & SUBRAMANIAN, S. 2016. Research on neighborhood effects on health in the United States: a systematic review of study characteristics. *Social Science & Medicine*, 168, 16-29.
- ASHBY, M. P. 2020. Initial evidence on the relationship between the coronavirus pandemic and crime in the United States. *Crime Science*, 9, 6.
- BAKERJIAN, D. & NGUYEN, A. 2022. Lessons Learned From Coronavirus Disease 2019 Recovery: Policy Implications for the Health and Well-Being of Older Adults. Oxford University Press US.
- BERMAN, J. D. & EBISU, K. 2020. Changes in US air pollution during the COVID-19 pandemic. *Science of the total environment*, 739, 139864.
- BESSER, L. 2021. Outdoor green space exposure and brain health measures related to Alzheimer's disease: a rapid review. *BMJ open*, 11, e043456.
- BESSER, L. M., MCDONALD, N. C., SONG, Y., KUKULL, W. A. & RODRIGUEZ, D. A. 2017. Neighborhood environment and cognition in older adults: a systematic review. *American journal of preventive medicine*, 53, 241-251.
- BILAL, U., MCCULLEY, E., LI, R., ROLLINS, H., SCHNAKE-MAHL, A., MULLACHERY, P. H., VAIDYA, V., KOH, C., DUREJA, K. & SHARAF, A. 2022. Tracking COVID-19 inequities across Jurisdictions represented in the Big cities health Coalition (Bchc): The COVID-19 health inequities in Bchc cities dashboard. *American Journal of Public Health*, 112, 904-912.
- BOMAN, J. H. & GALLUPE, O. 2020. Has COVID-19 changed crime? Crime rates in the United States during the pandemic. *American journal of criminal justice*, 45, 537-545.
- BOURDREL, T., ANNESI-MAESANO, I., ALAHMAD, B., MAESANO, C. N. & BIND, M.-A. 2021. The impact of outdoor air pollution on COVID-19: a review of evidence from in vitro, animal, and human studies. *European Respiratory Review*, 30.
- BROWNSON, R. C., HOEHNER, C. M., DAY, K., FORSYTH, A. & SALLIS, J. F. 2009. Measuring the built environment for physical activity: state of the science. *American journal of preventive medicine*, 36, S99-S123. e12.
- BRYAN, M. S., SUN, J., JAGAI, J., HORTON, D. E., MONTGOMERY, A., SARGIS, R. & ARGOS, M. 2021. Coronavirus disease 2019 (COVID-19) mortality and neighborhood characteristics in Chicago. *Annals of epidemiology*, 56, 47-54. e5.
- CAMINADE, C., MCINTYRE, K. M. & JONES, A. E. 2019. Impact of recent and future climate change on vector-borne diseases. *Annals of the New York Academy of Sciences*, 1436, 157-173.
- CARBALLO, I. H., BAKOLA, M. & STUCKLER, D. 2022. The impact of air pollution on COVID-19 incidence, severity, and mortality: A systematic review of studies in Europe and North America. *Environmental Research*, 114155.

- CASEY, J. A., JAMES, P., CUSHING, L., JESDALE, B. M. & MORELLO-FROSCH, R. 2017. Race, ethnicity, income concentration and 10-year change in urban greenness in the United States. *International journal of environmental research and public health*, 14, 1546.
- CENTERS FOR DISEASE CONTROL AND PREVENTION 2022. COVID Data Tracker. Atlanta, GA: Department of Health and Human Services.
- CHEN, J. T. & KRIEGER, N. 2021. Revealing the unequal burden of COVID-19 by income, race/ethnicity, and household crowding: US county versus zip code analyses. *Journal of Public Health Management and Practice*, 27, S43-S56.
- CONDERINO, S. E., FELDMAN, J. M., SPOER, B., GOUREVITCH, M. N. & THORPE, L. E. 2021. Social and Economic Differences in Neighborhood Walkability Across 500 U.S. Cities. *American Journal of Preventive Medicine*, 61, 394-401.
- COVEN, J., GUPTA, A. & YAO, I. 2023. JUE Insight: Urban flight seeded the COVID-19 pandemic across the United States. *Journal of urban economics*, 133, 103489.
- DANG, H.-A. H. & NGUYEN, C. V. 2021. Gender inequality during the COVID-19 pandemic: Income, expenditure, savings, and job loss. *World Development*, 140, 105296.
- DEWITZ, J. & U.S. GEOLOGICAL SURVEY 2021. National Land Cover Database (NLCD) 2019 Products (ver 2.0, June 2021): U.S. Geological Survey data release. Sioux Falls, SD: U.S. Geologic Survey.
- DIEZ ROUX, A. V., MUJAHID, M. S., HIRSCH, J. A., MOORE, K. & MOORE, L. V. 2016. The impact of neighborhoods on CV risk. *Global heart*, 11, 353-363.
- DONG, E., DU, H. & GARDNER, L. 2020. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet infectious diseases*, 20, 533-534.
- DOUGLAS, G. & MOORE, D. 2022. Analyzing the Use and Impacts of Oakland Slow Streets and Potential Scalability Beyond Covid-19.
- ECONOMIC RESEARCH SERVICE 2022. Rural-Urban Commuting Area Codes. Washington, DC: U.S. Department of Agriculture.
- ELLEN, I. G., MIJANOVICH, T. & DILLMAN, K.-N. 2001. Neighborhood effects on health: exploring the links and assessing the evidence. *Journal of urban affairs*, 23, 391-408.
- ENVIRONMENTAL PROTECTION AGENCY 2023. Air Data: Air Quality Data Collected at Outdoor Monifors Across the US. Research Triangle Park, NC.
- FIRTH, C. L., BAQUERO, B., BERNEY, R., HOERSTER, K. D., MOONEY, S. J. & WINTERS, M. 2021. Not quite a block party: COVID-19 street reallocation programs in Seattle, WA and Vancouver, BC. *SSM-population health*, 14, 100769.
- FRUMKIN, H. 2021. COVID-19, the built environment, and health. *Environmental health perspectives*, 129, 075001.
- GONZÁLEZ-LEONARDO, M., ROWE, F. & FRESOLONE-CAPARRÓS, A. 2022. Rural revival? The rise in internal migration to rural areas during the COVID-19 pandemic. Who moved and Where? *Journal of Rural Studies*, 96, 332-342.
- GONZÁLEZ-VAL, R. & SANZ-GRACIA, F. 2022. Urbanization and COVID-19 incidence: A cross-country investigation. *Papers in Regional Science*, 101, 399-415.
- HAMIDI, S., SABOURI, S. & EWING, R. 2020. Does density aggravate the COVID-19 pandemic? Early findings and lessons for planners. *Journal of the American Planning Association*, 86, 495-509.
- HASLAG, P. H. & WEAGLEY, D. 2022. From LA to Boise: How migration has changed during the COVID-19 pandemic. *Available at SSRN 3808326*.
- HEALTH RESOURCES & SERVICES ADMINISTRATION 2022. Area Health Resources Files. Rockville, MD.
- HIRSCH, J. A., GRENGS, J., SCHULZ, A., ADAR, S. D., RODRIGUEZ, D. A., BRINES, S. J. & DIEZ ROUX, A. V. 2016. How much are built environments changing, and where?:

- Patterns of change by neighborhood sociodemographic characteristics across seven U.S. metropolitan areas. *Social Science & Medicine*, 169, 97-105.
- HIRSCH, J. A., MOORE, K. A., CAHILL, J., QUINN, J., ZHAO, Y., BAYER, F. J., RUNDLE, A. & LOVASI, G. S. 2021. Business data categorization and refinement for application in longitudinal neighborhood health research: a methodology. *Journal of urban health*, 98, 271-284.
- HIRSCH, J. A., ZHAO, Y., MELLY, S., MOORE, K. A., BERGER, N., QUINN, J., RUNDLE, A. & LOVASI, G. S. 2023. National trends and disparities in retail food environments in the USA between 1990 and 2014. *Public Health Nutrition*, 26, 1052-1062.
- JIANG, N., WU, A. M. & CHENG, E. W. 2022. Social trust and stress symptoms among older adults during the COVID-19 pandemic: evidence from Asia. *BMC geriatrics*, 22, 330.
- JIMENEZ, M. P., DEVILLE, N. V., ELLIOTT, E. G., SCHIFF, J. E., WILT, G. E., HART, J. E. & JAMES, P. 2021. Associations between nature exposure and health: a review of the evidence. *International Journal of Environmental Research and Public Health*, 18, 4790.
- KASHEM, S. B., BAKER, D. M., GONZÁLEZ, S. R. & LEE, C. A. 2021. Exploring the nexus between social vulnerability, built environment, and the prevalence of COVID-19: A case study of Chicago. *Sustainable cities and society*, 75, 103261.
- KIRWA, K., SZPIRO, A. A., SHEPPARD, L., SAMPSON, P. D., WANG, M., KELLER, J. P., YOUNG, M. T., KIM, S.-Y., LARSON, T. V. & KAUFMAN, J. D. 2021. Fine-Scale air pollution models for epidemiologic research: insights from approaches developed in the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air). *Current environmental health reports*, 8, 113-126.
- KLOMPMAKER, J. O., HART, J. E., HOLLAND, I., SABATH, M. B., WU, X., LADEN, F., DOMINICI, F. & JAMES, P. 2021. County-level exposures to greenness and associations with COVID-19 incidence and mortality in the United States. *Environmental research*, 199, 111331.
- KRIEGER, N., WATERMAN, P. D., SPASOJEVIC, J., LI, W., MADURO, G. & VAN WYE, G. 2016. Public health monitoring of privilege and deprivation with the index of concentration at the extremes. *American journal of public health*, 106, 256-263.
- LEIFHEIT, K. M., LINTON, S. L., RAIFMAN, J., SCHWARTZ, G. L., BENFER, E. A., ZIMMERMAN, F. J. & POLLACK, C. E. 2021. Expiring eviction moratoriums and COVID-19 incidence and mortality. *American journal of epidemiology*, 190, 2503.
- LIU, C., LIU, Z. & GUAN, C. 2021a. The impacts of the built environment on the incidence rate of COVID-19: A case study of King County, Washington. *Sustainable cities and society*, 74, 103144.
- LIU, J., CLARK, L. P., BECHLE, M. J., HAJAT, A., KIM, S. Y., ROBINSON, A. L., SHEPPARD, L., SZPIRO, A. A. & MARSHALL, J. D. 2021b. Disparities in Air Pollution Exposure in the United States by Race/Ethnicity and Income, 1990-2010. *Environ Health Perspect*, 129, 127005.
- LOVASI, G. S., GRADY, S. & RUNDLE, A. 2011. Steps forward: review and recommendations for research on walkability, physical activity and cardiovascular health. *Public health reviews*, 33, 484-506.
- MALAMBO, P., KENGNE, A. P., DE VILLIERS, A., LAMBERT, E. V. & PUOANE, T. 2016. Built environment, selected risk factors and major cardiovascular disease outcomes: a systematic review. *PloS one*, 11, e0166846.
- MANSON, S., SCHROEDER, J., VAN RIPER, D., KUGLER, T. & RUGGLES, S. 2022. IPUMS National Historical Geographic Information System: Version 17.0 [dataset]. Minneapolis, MN: IPUMS University of Minnesota.
- MARQUÈS, M. & DOMINGO, J. L. 2022. Positive association between outdoor air pollution and the incidence and severity of COVID-19. A review of the recent scientific evidences. *Environmental Research*, 203, 111930.

- MAYNE, S. L., AUCHINCLOSS, A. H. & MICHAEL, Y. L. 2015. Impact of policy and built environment changes on obesity-related outcomes: a systematic review of naturally occurring experiments. *Obesity Reviews*, 16, 362-375.
- MAYNE, S. L., MITCHELL, J. A., VIRUDACHALAM, S., FIKS, A. G. & WILLIAMSON, A. A. 2021. Neighborhood environments and sleep among children and adolescents: a systematic review. *Sleep medicine reviews*, 57, 101465.
- MCALEXANDER, T. P., ALGUR, Y., SCHWARTZ, B. S., RUMMO, P. E., LEE, D. C., SIEGEL, K. R., RYAN, V., LEE, N. L., MALLA, G. & MCCLURE, L. A. 2022. Categorizing community type for epidemiologic evaluation of community factors and chronic disease across the United States. *Social sciences & humanities open*, 5, 100250.
- MCGOWAN, V. J. & BAMBRA, C. 2022. COVID-19 mortality and deprivation: pandemic, syndemic, and endemic health inequalities. *The Lancet Public Health*, 7, e966-e975.
- MICHAEL, Y. L., GREEN, M. K. & FARQUHAR, S. A. 2006. Neighborhood design and active aging. *Health & place*, 12, 734-740.
- MUDE, W., OGUOMA, V. M., NYANHANDA, T., MWANRI, L. & NJUE, C. 2021. Racial disparities in COVID-19 pandemic cases, hospitalisations, and deaths: A systematic review and meta-analysis. *Journal of global health*, 11.
- MUJAHID, M. S., DIEZ ROUX, A. V., MORENOFF, J. D. & RAGHUNATHAN, T. 2007. Assessing the measurement properties of neighborhood scales: from psychometrics to econometrics. *American journal of epidemiology*, 165, 858-867.
- NATIONAL CENTERS FOR ENVIRONMENTAL INFORMATION 2023. Local Climatological Data (LCD). Silver Spring, MD: National Oceanic and Atmospheric Administration.
- NAUDÉ, W. & NAGLER, P. 2022. COVID-19 and the city: Did urbanized countries suffer more fatalities? *Cities*, 131, 103909.
- ODDEN, M. C., LI, Y., THORPE JR, R. J., TAN, A., SIMS, K. D., RATCLIFF, J., MAGID, H. S. A. & SIMS, M. 2023. Neighborhood factors and survival to old age: the Jackson Heart Study. *Preventive medicine reports*, 35, 102360.
- OELSNER, E. C., KRISHNASWAMY, A., BALTE, P. P., ALLEN, N. B., ALI, T., ANUGU, P., ANDREWS, H. F., ARORA, K., ASARO, A. & BARR, R. G. 2022. Collaborative cohort of cohorts for COVID-19 research (C4R) study: study design. *American journal of epidemiology*, 191, 1153-1173.
- RHOADS, D., SOLÉ-RIBALTA, A., GONZÁLEZ, M. C. & BORGE-HOLTHOEFER, J. 2020. Planning for sustainable open streets in pandemic cities. *arXiv preprint arXiv:2009.12548*.
- RHOADS, D., SOLÉ-RIBALTA, A., GONZÁLEZ, M. C. & BORGE-HOLTHOEFER, J. 2021. A sustainable strategy for Open Streets in (post) pandemic cities. *Communications Physics*, 4, 183.
- ROJAS-RUEDA, D. & MORALES-ZAMORA, E. 2021. Built Environment, Transport, and COVID-19: a Review. *Current Environmental Health Reports*, 8, 138-145.
- RUSSETTE, H., GRAHAM, J., HOLDEN, Z., SEMMENS, E. O., WILLIAMS, E. & LANDGUTH, E. L. 2021. Greenspace exposure and COVID-19 mortality in the United States: January–July 2020. *Environmental research*, 198, 111195.
- SANDOVAL-OLASCOAGA, S., VENKATARAMANI, A. S. & ARCAYA, M. C. 2021. Eviction moratoria expiration and COVID-19 infection risk across strata of health and socioeconomic status in the United States. *JAMA network open*, 4, e2129041-e2129041.
- SCHAEFER-MCDANIEL, N., CAUGHY, M. O. B., O'CAMPO, P. & GEAREY, W. 2010. Examining methodological details of neighbourhood observations and the relationship to health: a literature review. *Social science & medicine*, 70, 277-292.
- SCHIERMEIER, Q. 2019. Europe's mega-heatwave boosted by climate change. *Nature*, 571, 155-155.

- SCHULZ, A. & NORTHRIDGE, M. E. 2004. Social determinants of health: implications for environmental health promotion. *Health education & behavior*, 31, 455-471.
- SILVA, A. L. P., PRATA, J. C., WALKER, T. R., DUARTE, A. C., OUYANG, W., BARCELÒ, D. & ROCHA-SANTOS, T. 2021. Increased plastic pollution due to COVID-19 pandemic: Challenges and recommendations. *Chemical Engineering Journal*, 405, 126683.
- SINGH, A., DANIEL, L., BAKER, E. & BENTLEY, R. 2019. Housing disadvantage and poor mental health: a systematic review. *American journal of preventive medicine*, 57, 262-272.
- SMITH, M., HOSKING, J., WOODWARD, A., WITTEN, K., MACMILLAN, A., FIELD, A., BAAS, P. & MACKIE, H. 2017. Systematic literature review of built environment effects on physical activity and active transport—an update and new findings on health equity. *International journal of behavioral nutrition and physical activity*, 14, 1-27.
- STRAUSS, B. H., ORTON, P. M., BITTERMANN, K., BUCHANAN, M. K., GILFORD, D. M., KOPP, R. E., KULP, S., MASSEY, C., MOEL, H. D. & VINOGRADOV, S. 2021. Economic damages from Hurricane Sandy attributable to sea level rise caused by anthropogenic climate change. *Nature communications*, 12, 2720.
- TRUST FOR PUBLIC LAND 2023. ParkServe. Los Angeles, CA.
- TSUI, J., HIRSCH, J. A., BAYER, F. J., QUINN, J. W., CAHILL, J., SISCOVICK, D. & LOVASI, G. S. 2020. Patterns in Geographic Access to Health Care Facilities Across Neighborhoods in the United States Based on Data From the National Establishment Time-Series Between 2000 and 2014. *JAMA Network Open*, 3, e205105-e205105.
- VERSEY, H. S. 2021. The impending eviction cliff: housing insecurity during COVID-19. American Public Health Association.
- WEAVER, A. K., HEAD, J. R., GOULD, C. F., CARLTON, E. J. & REMAIS, J. V. 2022. Environmental factors influencing COVID-19 incidence and severity. *Annual review of public health*, 43, 271-291.
- WILLIAMS, A. P., ABATZOGLOU, J. T., GERSHUNOV, A., GUZMAN-MORALES, J., BISHOP, D. A., BALCH, J. K. & LETTENMAIER, D. P. 2019. Observed impacts of anthropogenic climate change on wildfire in California. *Earth's Future*, 7, 892-910.
- WON, J., LEE, C., FORJUOH, S. N. & ORY, M. G. 2016. Neighborhood safety factors associated with older adults' health-related outcomes: a systematic literature review. *Social Science & Medicine*, 165, 177-186.
- YANG, Y., LU, Y. & JIANG, B. 2022. Population-weighted exposure to green spaces tied to lower COVID-19 mortality rates: A nationwide dose-response study in the USA. *Science of the Total Environment*, 851, 158333.
- ZANGGER, C. 2021. Help thy neighbor. Neighborhood relations, subjective well-being, and trust during the COVID-19 pandemic. *SocArXiv*. May, 19.

TABLES

Table 1: Population, sociodemographic, economic, and environmental characteristics for coverage of the C4R cohort geographies compared to national and areas not covered by C4R cohorts.				
Neighborhood characteristic	C4R Locations ¹	National	Non-C4R Locations	p-value ⁷
Land Area (km ²)	2,153,275	7,654,980	5,501,705	-----
Total Population ²	167,153,863	322,538,561	155,384,698	-----
Urbanicity ³ (%)	82.69	75.61	68.72	-----
	Mean (SD)	Mean (SD)	Mean (SD)	
Population density (pop/km ²) ²	2613.08 (5687.63)	2087.58 (4598.84)	1575.32 (3114.43)	<0.001
Age of residents: ≥65 (%) ²	16.33 (8.28)	16.40 (8.14)	16.47 (8.0)	0.025
Age of residents: ≥85 (%) ²	2.08 (2.01)	2.06 (1.95)	2.04 (1.88)	0.003
Non-Hispanic White Population (%) ²	57.04 (30.52)	61.48 (29.91)	65.82 (28.64)	<0.001
People of Color Population (%) ²	42.96 (30.52)	38.52 (29.91)	34.18 (28.64)	<0.001
Foreign-Born Residents (%) ²	13.27 (13.86)	12.65 (13.59)	12.04 (13.29)	<0.001
Median Household Income ²	68662.23 (35392.07)	66977.19 (33467.29)	65321.33 (31372.80)	<0.001
Rent (%) ²	38.53 (23.51)	36.56 (22.95)	34.63 (22.22)	<0.001
Own (%) ²	61.47 (23.51)	63.44 (22.95)	65.37 (22.22)	<0.001
Persons Below Poverty (%) ²	14.85 (12.00)	14.68 (11.60)	14.51 (11.20)	<0.001
HS Degree (%) ²	87.88 (9.89)	87.32 (10.36)	86.76 (10.77)	<0.001
College Degree (%) ²	33.18 (20.32)	30.7 (19.35)	28.27 (18.01)	<0.001
Employed (%) ²	93.99 (4.68)	94.19 (4.42)	94.38 (4.14)	<0.001
Commuting by Car (%) ²	83.03 (17.57)	84.81 (15.57)	86.56 (13.10)	<0.001
ICE (high-income White households vs. low-income Black households) ⁴	.161 (.258)	.181 (.227)	.201 (.188)	<0.001
ICE (high-income White non-Hispanic households vs. low-income Hispanic households) ⁴	.201 (.197)	.196 (.194)	.192 (.193)	<0.001
Crowding (occ housing units w/>1 person)	3.50 (5.07)	3.54 (5.15)	3.58 (5.22)	0.035

per room, %) ²				
Hospital Beds/100k Population ⁵	282.06 (396.79)	300.39 (493.41)	344.55 (669.18)	0.001
Greenness (%) ⁶	29.21 (26.71)	31.54 (27.85)	33.82 (28.74)	<0.001

¹Averages represent the areas where C4R cohort participants have recorded addresses. All geographies were aggregated across cohorts and available exam periods. Estimates represent the reaggregated population, sociodemographic, economic, and environmental characteristics for locations with participants and locations without participants allocated by geographic centroid.

²Data from ACS 2014-2018

³Urbanicity defined using McAlexander, T. P., Algur, Y., Schwartz, B. S., Rummo, P. E., Lee, D. C., Siegel, K. R., ... & McClure, L. A. (2022). Categorizing community type for epidemiologic evaluation of community factors and chronic disease across the United States. *Social sciences & humanities open*, 5(1), 100250.

⁴ICE=Index of Concentration at the Extremes from calculated using standard equations (Krieger et al., 2016).

⁵Data on Hospital Beds/100k population from the Area Health Resources Files 2018-2019 to create a number of beds per 100k population density measure (Health Resources & Services Administration, 2022)

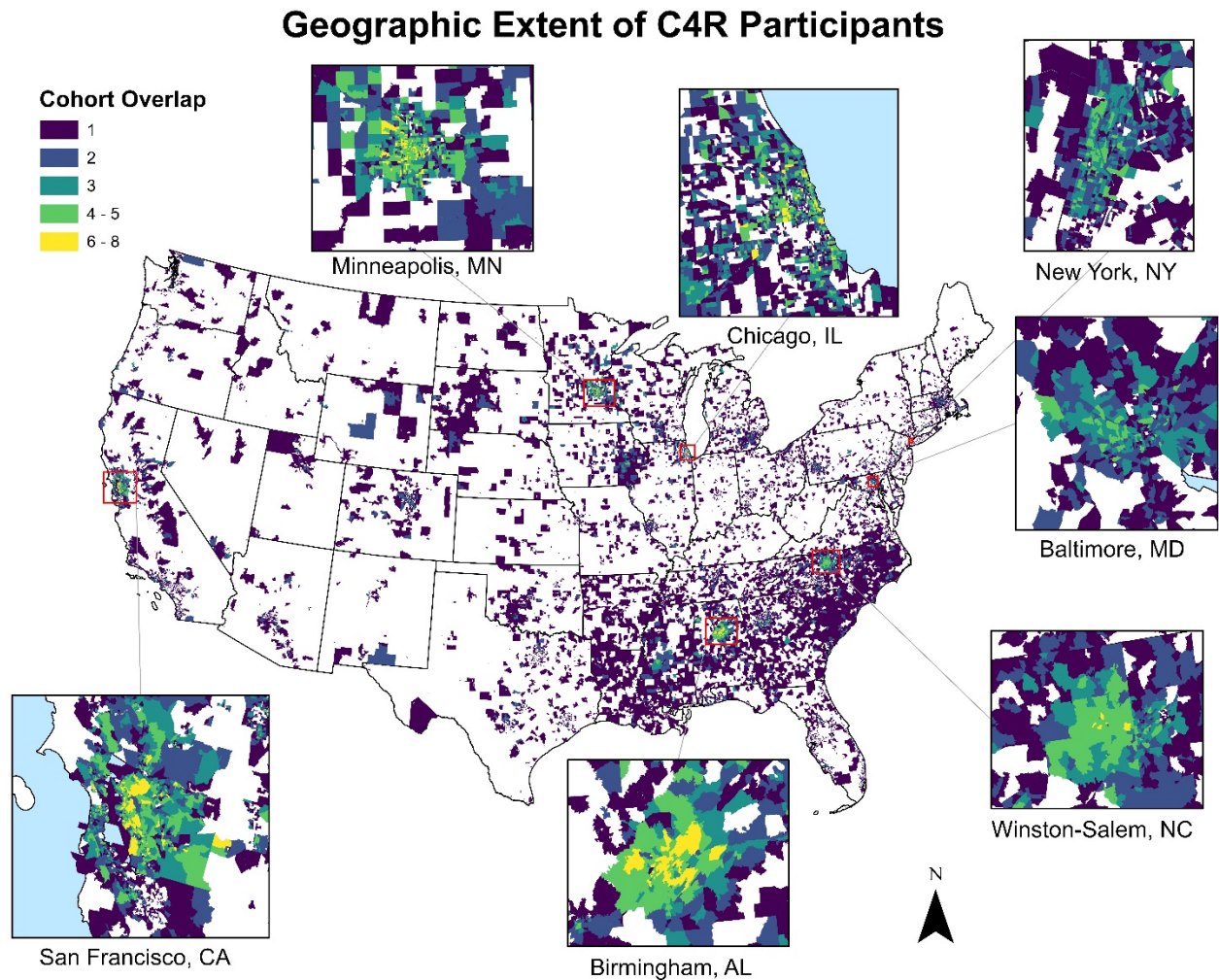
⁶Greenness data from 2019 NLCD (Dewitz and U.S. Geological Survey, 2021)

⁷P-value comparing means in C4R locations to means in non-C4R locations using t-test.

Table 2: COVID Experience for the C4R cohort geographies².				
	Case Rate (per 100,000) ¹		Vaccination Rate (%) ¹	
Variable	C4R Locations ²	National	C4R Locations ²	National
2020				
Q1	23.06	18.95	NA	NA
Q2	536.74	489.68	NA	NA
Q3	1498.62	1435.37	NA	NA
Q4	4501.01	4737.42	NA	NA
2021				
Q1	2887.01	2780.86	13.75	13.53
Q2	859.64	818.49	32.53	30.73
Q3	3600.22	3527.58	40.96	38.59
Q4	3432.20	3522.29	48.73	47.14
¹ COVID-related variables were derived from the Johns Hopkins COVID-19 dashboard (cases per quarter) and the CDC COVID Data Tracker (vaccination rates) (Dong et al., 2020, Centers for Disease Control and Prevention, 2022). ² Numbers represent the COVID context and experience of areas where participants are located. They do not represent case rates or vaccination rates for C4R participants themselves. Averages represent the areas where C4R cohort participants have recorded addresses. All geographies were aggregated across cohorts and available exam periods. Estimates represent the areal-weighted, reaggregated characteristics for locations with participants and nationally.				

FIGURES

Figure 1: Geographic Extent of C4R Participants.



Geographic coverage obtained from cohorts anonymized and aggregated by spatial unit (e.g., county, census tract) and time (e.g., baseline, most recent exam, all exams). Since participant locations were provided at varying time periods and geographies, the map does not necessarily represent locations during COVID-19 pandemic.

Figure 2: Neighborhood contextual data in the C4R cohorts by timeframe, measure type, and number of measures.

