Leveraging the OMOP Common Data Model to Support Distributed Health Equity Research

Sarah D. Gasman and William G. Adams[1]

[1] **Department of Pediatrics and Clinical and Translational Science Institute, Boston Medical Center and Boston University Chobanian & Avedisian School of Medicine, Boston, MA**

Background: To better support health equity research on a national scale, there is a pressing need to: 1) expand foundational data systems to include information on social and environmental factors that represent multilevel drivers of health; and 2) develop informatics tools to support visualization and statistical analyses in ways that are easy to use and share across research organizations.

Human health is impacted by individual and place-based social and environmental drivers of health (SEDoH).[1] Currently, patient-level SEDoH data represented in the OMOP-CDM is limited. This information is not often collected systematically or consistently by health systems, though efforts are being made to harmonize existing data to OMOP standards.[2] Additionally, the OHDSI community is working to transform publicly available place-based data to a standard representation[3] so that this information can be included as part of OHDSI analyses and research workflows.

We have developed a data platform and tool on top of existing OMOP infrastructure to enable health services researchers, clinicians, policy makers, and other stakeholders to explore relationships between health outcomes and place-based data, *themselves*, in a way that supports collaborative, inter-institutional research.

Objective: Develop a health equity research platform within a large urban safety-het health system to: 1) expand an existing de-identified Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) repository, the "Data for Equity (D4E) Platform", to include extensive census-level "place-based" social and environmental drivers of health (e.g. Child Opportunity Index (COI), Social Vulnerability Index (SVI), American Community Survey (ACS)); 2) develop a library of reusable analytic code to assess associations between a broad range of health outcomes and drivers of health; and 3) provide an easy to use data visualization tool, the "Health Equity Explorer (H2E)", to allow users to explore health equity themselves via graphs, maps, and statistical analyses.

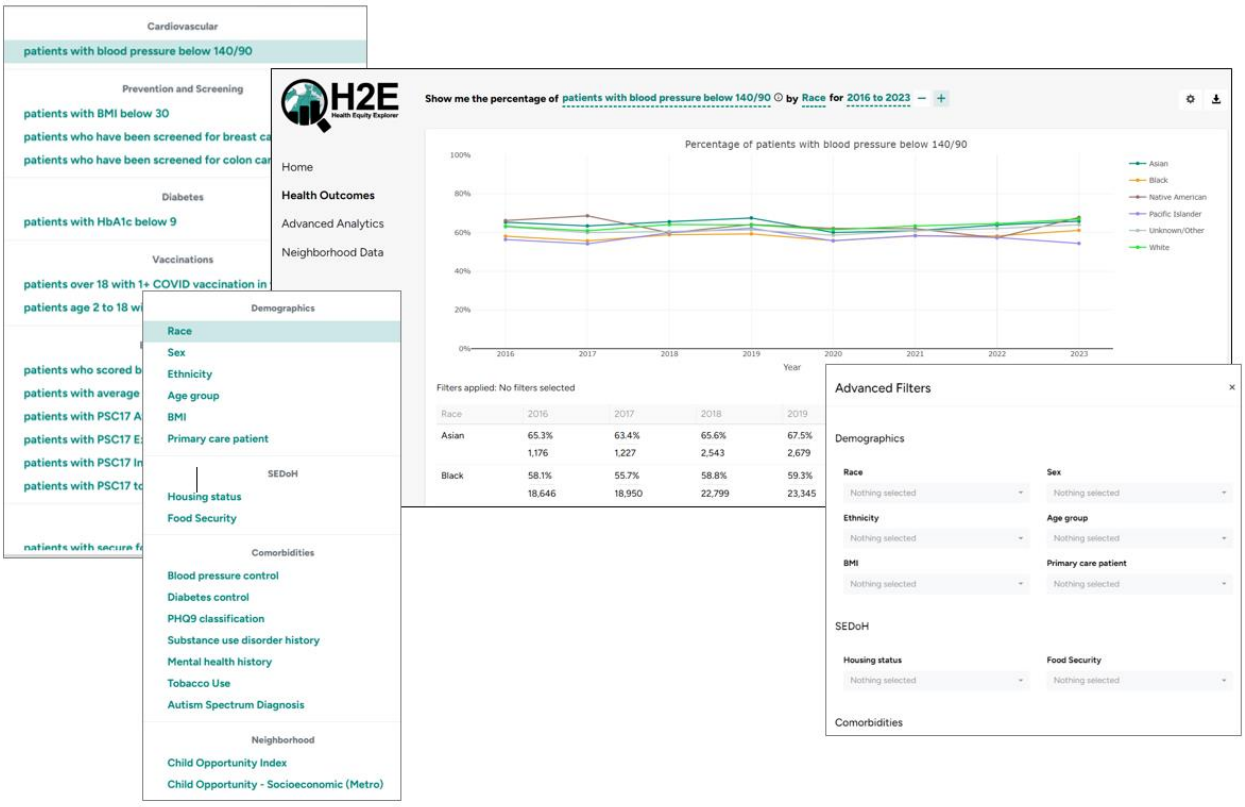
Methods: We used an interactive design process with input from multiple stakeholders including leaders of the Boston Medical Center “Health Equity Accelerator” and a community advisory board. We also leveraged materials from the Observational Health Data Sciences and Informatics (OHDSI) Community, open source statistical and application development tools, and standard measure specifications and value sets for target outcome measures. The application was initially developed in Tableau but was converted into an R-Shiny application between June and September 2023. H2E uses pre-computed health outcomes linked to a wide range of clinical and SEDoH variables to allow users to explore health equity and inequity via graphs, maps, and statistical analyses.

H2E uses data generated from the OMOP CDM that are also linked to a separate repository of geo-spatial data. Data are stored within the H2E data mart that is linked to, but separate from, an OMOP CDM. Health outcome data are generated by querying the OMOP-CDM using reusable SQL and R code. Patient current and historical addresses are geocoded and mapped to census tracts. H2E currently includes place-based data from the Social Vulnerability Index, Child Opportunity Index, and American Community Survey, but any zip code or FIPS code mapped data can be added to the data model.

Health outcomes are precomputed for eligible patients using standardized scripts that query a site's OMOP instance; sites can select which health outcomes to load to their data mart. Health outcomes are assessed once per patient per time period (year, quarter, month). Additionally, clinical, and place-based variables are assessed for eligible patients dynamically based on reference time period, not statically (for example, a patient will have a variable indicating a comorbidity only for time periods that occur after the diagnosis of that comorbidity). The computation of health outcomes and variables results in one database table per outcome in the results database. These outcome tables are then coalesced into a centralized data mart and joined with additional demographic and third-party place-based data. For this project, a Tableau-based prototype was developed and refined, followed by a 90-day software development sprint to convert the application into an R Shiny App in collaboration with a commercial software company, Appsilon through the "Data 4 Good" Program (https://data4good.appsilon.com/).

Results: We have developed an open-source, OMOP CDM-based, platform with clinical and place-based social and environmental data, re-usable code, and a visualization tool to support health and health equity exploration (visually and statistically). We have implemented a diverse and growing set of medical and behavioral health outcomes for children and adults in a way that allows dynamic exploration of the contributions of individual social need, and place-based social and environmental drivers of health. Summary data from analyses can be easily shared to support real-world distributed health equity research on a national scale.

Figure 1: Health Equity Explorer v2.0. Users can select from a list of precomputed health outcomes and drivers of health.



Scripts and code to generate outcomes and exposure variables are easily shareable. H2E also supports visualization of health outcomes by neighborhood maps (see Figure 2), and statistical modeling of the relationships between drivers and outcomes (including logistic regression), exporting of aggregate data and displays, and analyses of missingness (see Figure 3).

Figure 2: Display third-party GIS data and health outcomes by neighborhood.

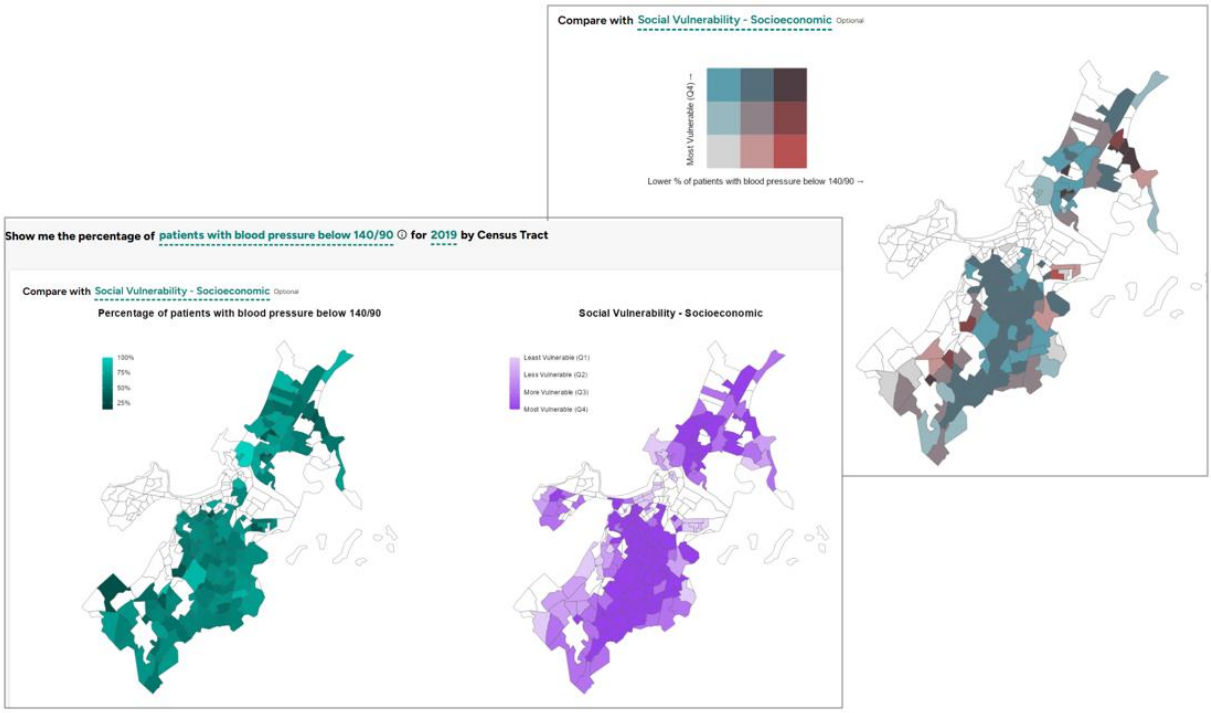
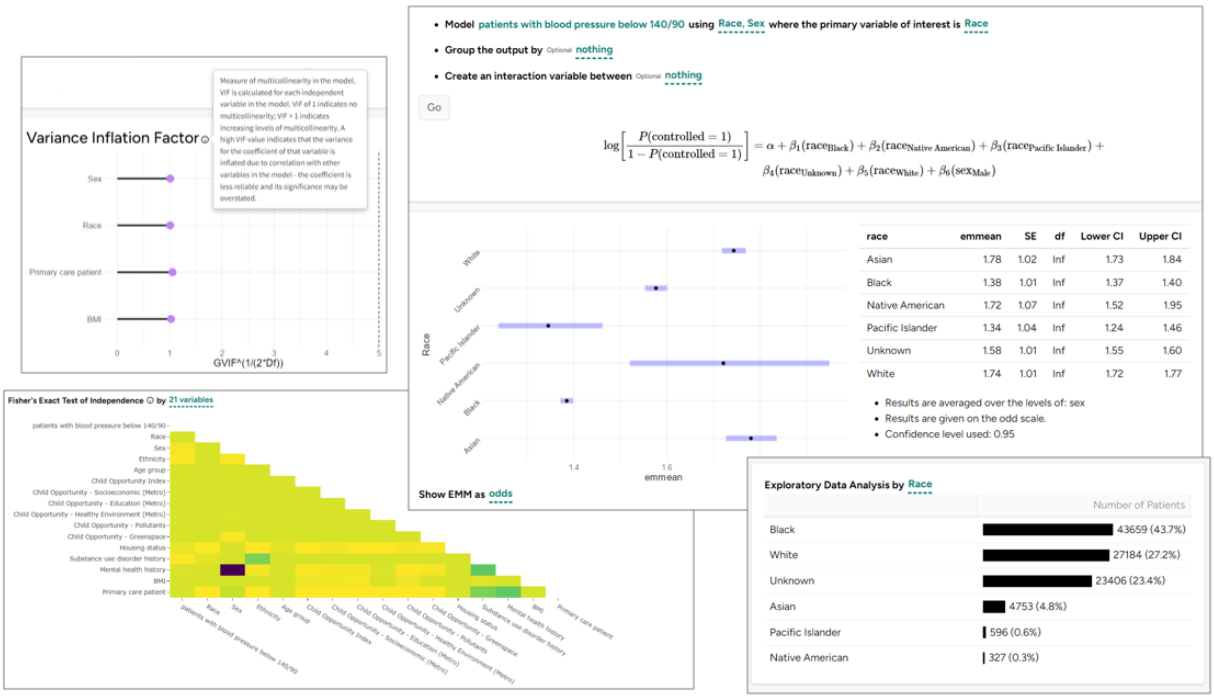


Figure 3: Advanced Analytics, including univariate and multivariate regression models and interaction terms. Further expanded in H2E v2.0 using R Shiny.



We have partnered with a commercial software development company to transform our Tableau-based H2E into an R Shiny App that can be shared broadly within the OHDSI and translational research communities. For this presentation, we will demonstrate the functionality of the application, provide technical specification, and explore potential new collaborations within the OHDSI Community.

Conclusion

The BMCHS Health Equity Explorer can be used to support dynamic and interactive explorations into the diverse drivers of health and health equity. With expanded use and partnerships, these tools have the potential to support distributed health equity research and intervention on a national scale.

References

1. How does place affect health? [Internet]. Centers for Disease Control and Prevention; 2023 [cited 2023 Jun 13]. Available from: https://www.atsdr.cdc.gov/placeandhealth/howdoesPlaceaffectHealth.html.
2. Phuong J, Zampino E, Dobbins N, Espinoza J, Meeker D, Spratt H, Madlock-Brown C, Weiskopf NG, Wilcox A. Extracting Patient-level Social Determinants of Health into the OMOP Common Data Model. AMIA Annu Symp Proc. 2022 Feb 21;2021:989-998. PMID: 35308947; PMCID: PMC8861735.
3. OHDSI GIS WG [internet]. [cited 13 June 2023]. Available from: https://ohdsi.github.io/GIS/index.html