DEVELOPING AI/ML-READY AGING TRAJECTORY FILES

OLGA F. JARRÍN MONTANER & HAIQUN LIN

RUTGERS, THE STATE UNIVERSITY OF NEW JERSEY

PARENT GRANT: R33AG068931

PROGRAM OFFICER: PARTHA BHATTACHARYYA

RUTGERS
Community Health and Aging Outcomes Laboratory

Our Growing Research Team

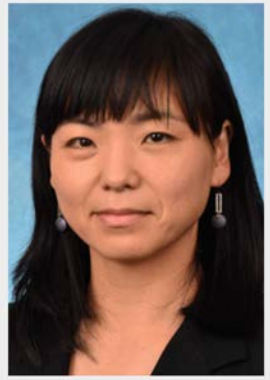
Olga F. Jarrín, PhD, RN
Principal Investigator

Haiqun Lin, MD, PhD
MPI of R33 & Co-I on R01

Zeeshan Ahmed, PhD
Co-Investigator on R33

Suzan Ahmad, PhD, RN
Co-Investigator on R33

Jason Roy, PhD
Co-Investigator on R33



Soko Setoguchi, MD, DrPH
Co-Investigator on R33

Marie Harvey, DrPH
Co-Investigator on R33 & R01

Dawn Kim, PhD
Research Assistant Professor

Anum Zafar, MPhil
Project Manager

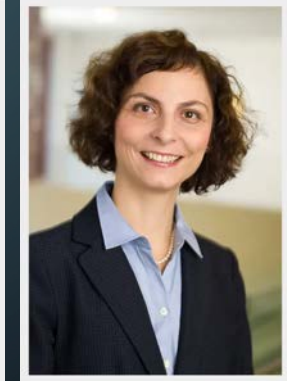
Maria Lopez, MA, MS
Research Coordinator

Weiyi Xia
PhD Student

Patrick Luib, MSN, APRN
PhD Student

Anaïs Mahone, MSW
Research Assistant

Charlotte Thomas-Hawkin...
Collaborator



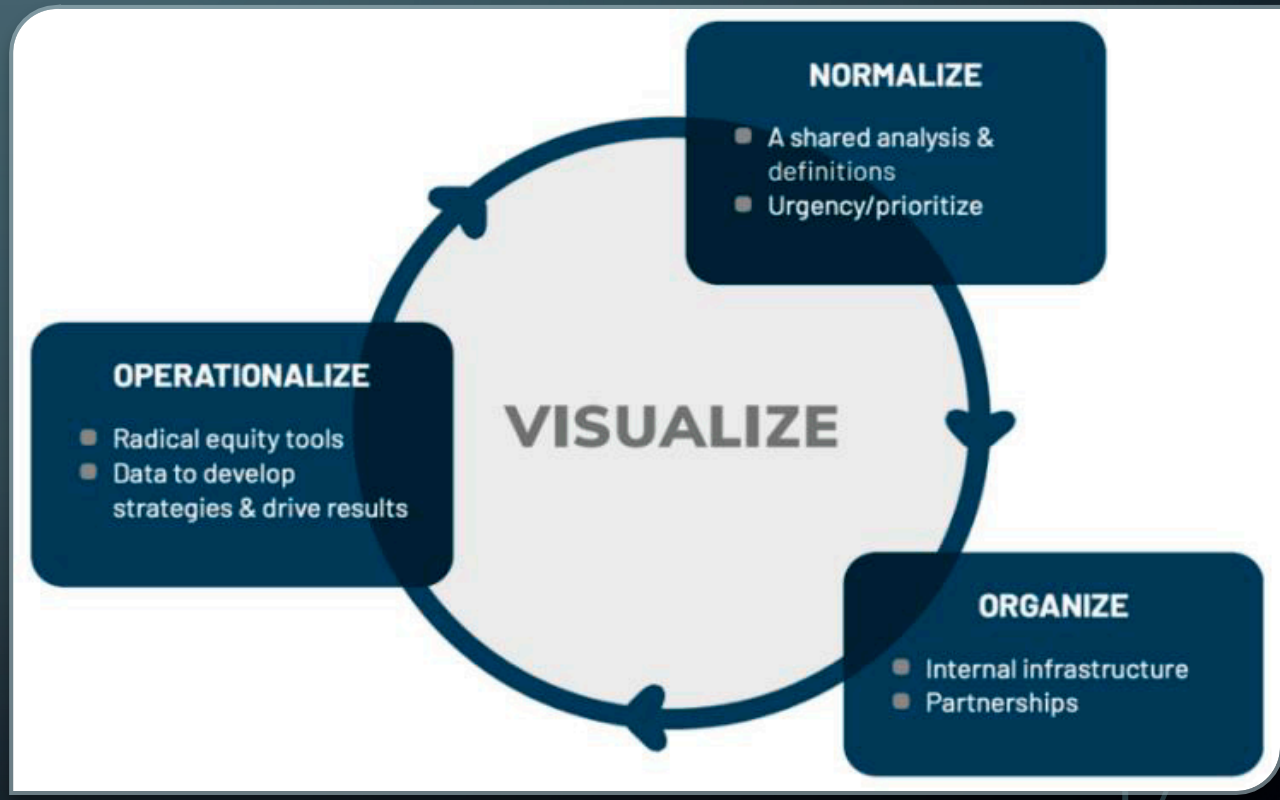
Irina Grafova, PhD
Collaborator

Ethan Halm, MD, MPH, MB...
Collaborator

RUTGERS
Community Health and
Aging Outcomes Laboratory

DEVELOPING AI/ML-READY DATASETS TO PREDICT LATE-LIFE AGING TRAJECTORIES OF MEDICARE BENEFICIARIES AND EMULATE TRIALS IN THE CMS VIRTUAL RESEARCH DATA CENTER.

We normalize, organize, and operationalize racial equity throughout data pre-processing and harmonization to address ethical issues, eliminate biases in datasets, algorithms, and applications; considering impacts and unintended consequences for disadvantaged or marginalized groups and health disparities.

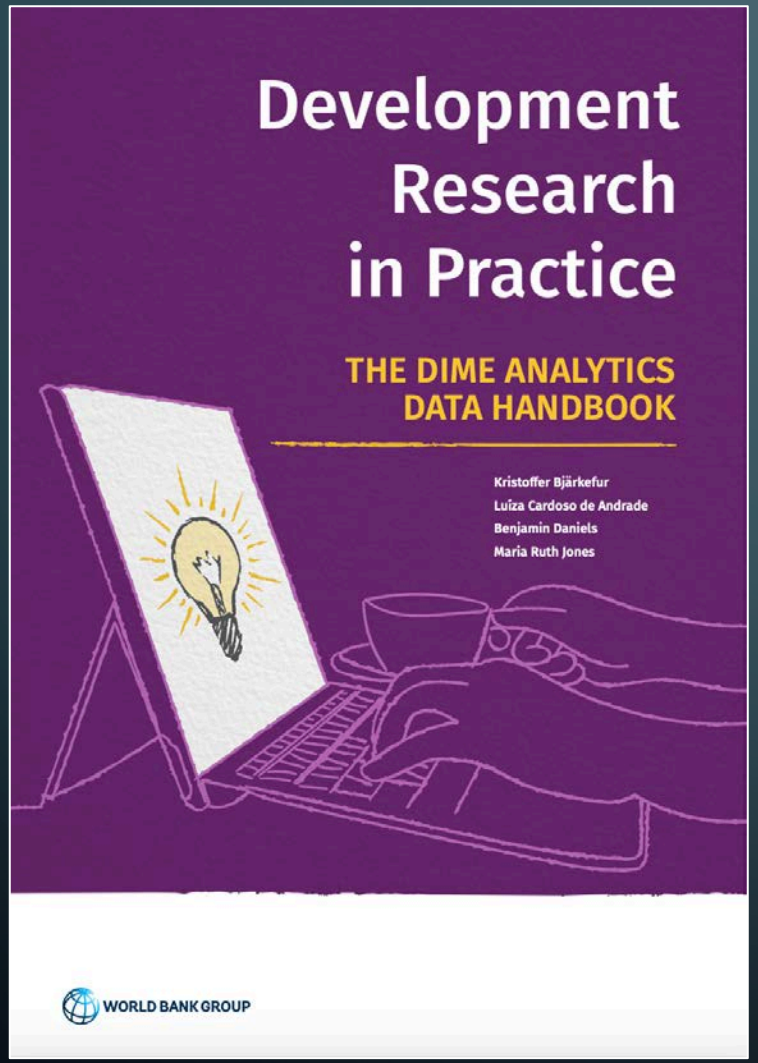


Government Alliance for Racial Equity (GARE) Framework

DATA PRE-PROCESSING & HARMONIZATION

- Standardization of documentation
- Irregularly spaced trajectories
- Augmented race/ethnicity variable
- Air quality exposure data ($PM_{2.5}$ & ozone)

STANDARDIZATION OF APPROACH



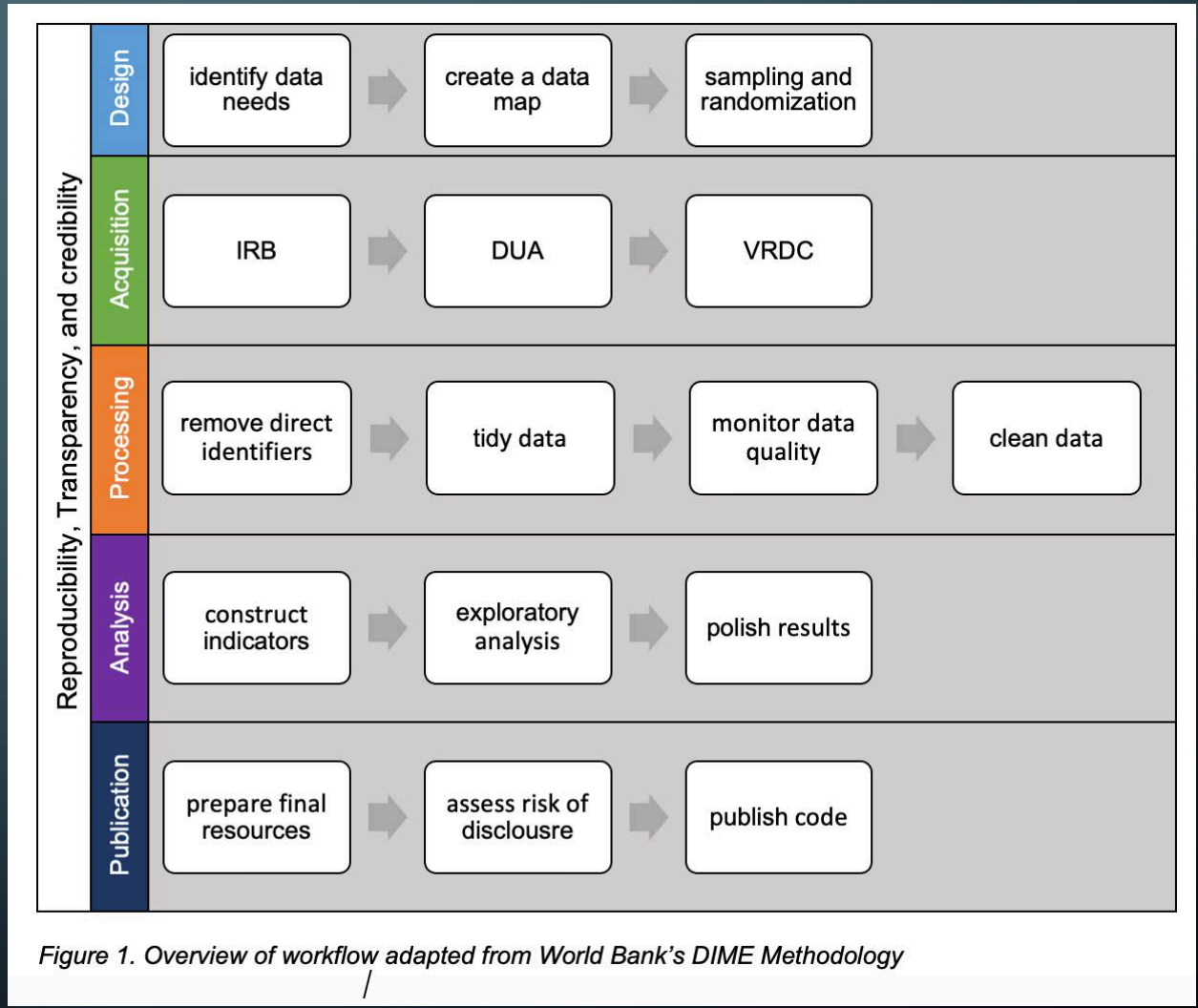
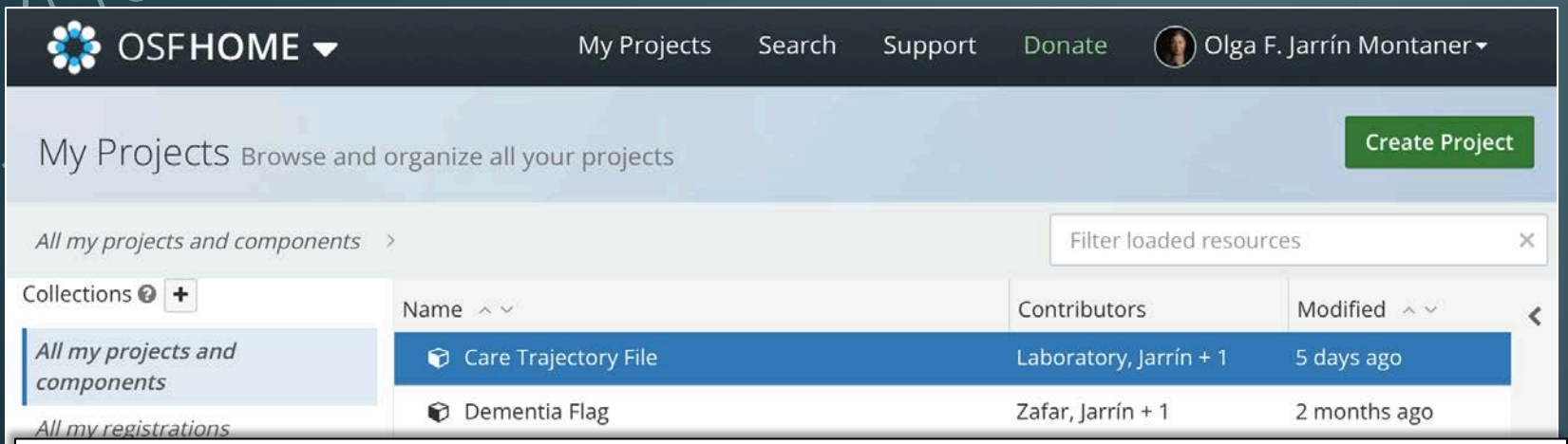


Figure 1. Overview of workflow adapted from World Bank's DIME Methodology

IRREGULARLY SPACED TRAJECTORIES



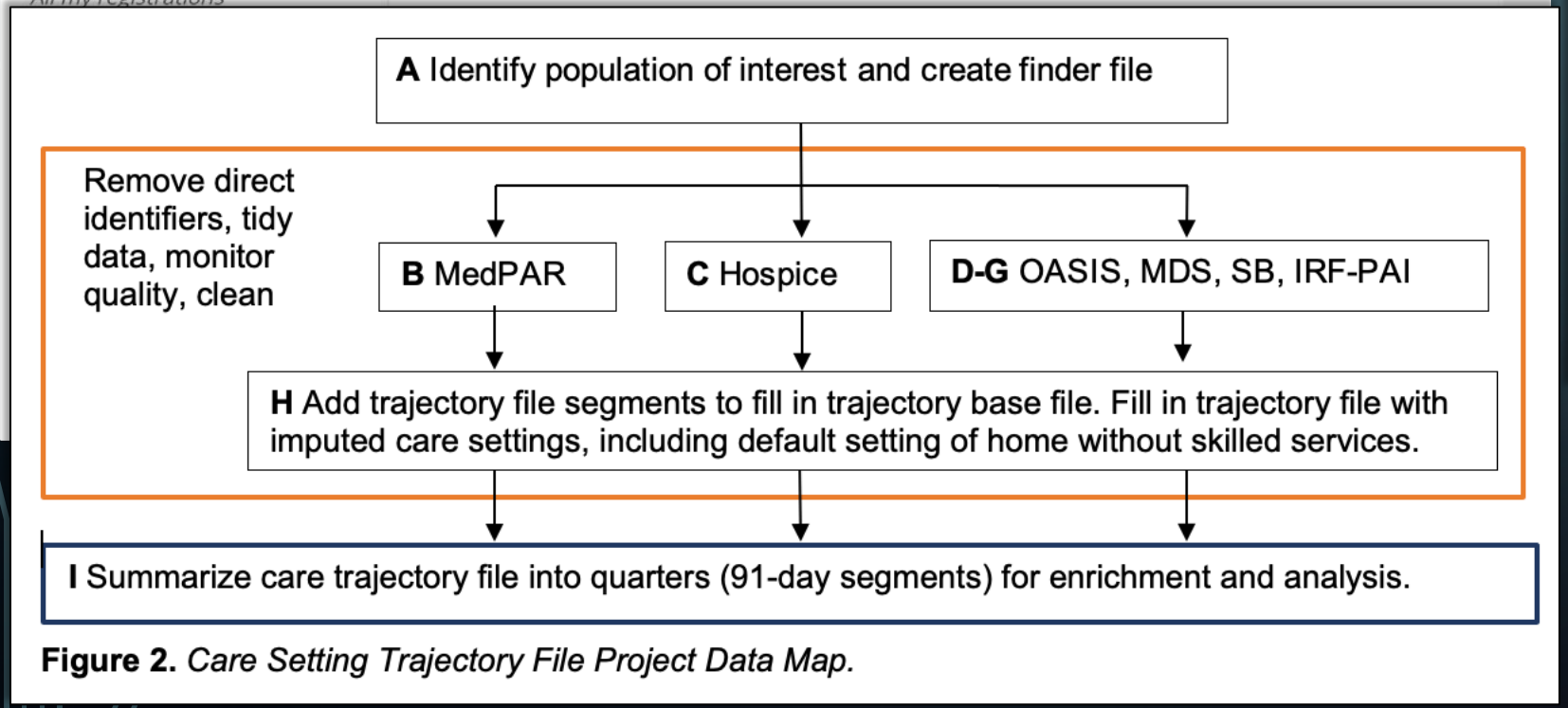
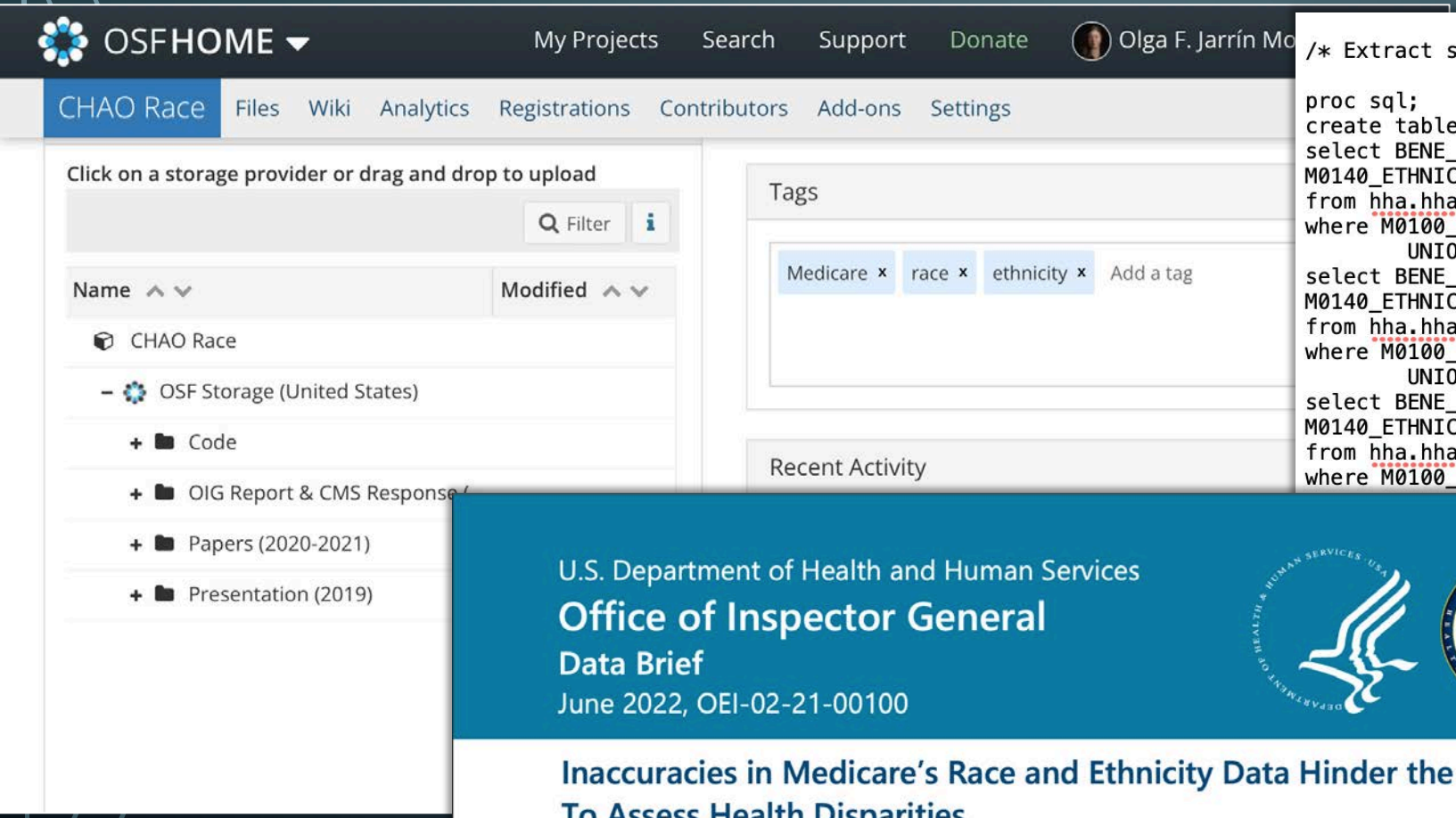


Figure 2. *Care Setting Trajectory File Project Data Map.*

```
/*****************************************************************
* STEP 2.1: MEDPAR DATA SECTION
*****************************************************************
            PART A: Use finder file of BENE_IDs to extract records a
*****************************************************************

PROC SQL;
CREATE TABLE MP_&YEAR. AS
SELECT A.*
FROM [DUA_SPECIFIC].MEDPAR&YEAR.[...](KEEP=BENE_ID SS_LS_SNF_IND_CD
INNER JOIN BASE_FILE_SRS AS B
ON A.BENE_ID=B.BENE_ID
ORDER BY BENE_ID;
QUIT;

/*****************************************************************
            PART B: Extract inpatient care dates from MedPAR
*****************************************************************

DATA MP_CST_&YEAR.;
SET MP_&YEAR.;
BY BENE_ID;
DSCHRG_DT_IMP=0;
DSCHRG_DT_IMP=DSCHRG_DT;
IF ADMSN_DT^=. AND DSCHRG_DT=. THEN DO;
DSCHRG_DT_IMP=ADMSN_DT+LOS_DAY_CNT;
DSCHRG_DT_IMP_FLAG=1;
END;

IF ADMSN_DT=DSCHRG_DT=. THEN ADMSN_FL=0;
ELSE ADMSN_FL=1;
CARE_SETTING=SS_LS_SNF_IND_CD;
EXP_SETTING_DT=DSCHRG_DT_IMP;
```

RUTGERS
Community Health and Aging Outcomes Laboratory

USING SELF-REPORTED RACE/ETHNICITY




```
/* Extract self-reported race variables from HHA (OASIS) */

proc sql;
create table HHA_assessment_race as
select BENE_ID, M0030_SOC_DT, M0140_ETHNIC_HISP, M0140_ETHNIC_ASIAN,
M0140_ETHNIC_BLACK, M0140_ETHNIC_NH_PI, M0140_ETHNIC_AI_AN, M0140_ETHNIC_WHI
from hha.hha_assessment_summary_2013
where M0100_ASSMT_REASON in ('01')
      UNION ALL
select BENE_ID, M0030_SOC_DT, M0140_ETHNIC_HISP, M0140_ETHNIC_ASIAN,
M0140_ETHNIC_BLACK, M0140_ETHNIC_NH_PI, M0140_ETHNIC_AI_AN, M0140_ETHNIC_WHI
from hha.hha_assessment_summary_2014
where M0100_ASSMT_REASON in ('01')
      UNION ALL
select BENE_ID, M0030_SOC_DT, M0140_ETHNIC_HISP, M0140_ETHNIC_ASIAN,
M0140_ETHNIC_BLACK, M0140_ETHNIC_NH_PI, M0140_ETHNIC_AI_AN, M0140_ETHNIC_WHI
from hha.hha_assessment_summary_2015
where M0100_ASSMT_REASON in ('01')

 M0140_ETHNIC_HISP, M0140_ETHNIC_ASIAN,
NIC_NH_PI, M0140_ETHNIC_AI_AN, M0140_ETHNIC_WHI
ry_2016
'01')

 M0140_ETHNIC_HISP, M0140_ETHNIC_ASIAN,
NIC_NH_PI, M0140_ETHNIC_AI_AN, M0140_ETHNIC_WHI
ry_2017
'01')

 M0140_ETHNIC_HISP, M0140_ETHNIC_ASIAN,
NIC_NH_PI, M0140_ETHNIC_AI_AN, M0140_ETHNIC_WHI
```

U.S. Department of Health and Human Services

Office of Inspector General

Data Brief

June 2022, OEI-02-21-00100

Inaccuracies in Medicare's Race and Ethnicity Data Hinder the Ability To Assess Health Disparities

Key Results

- Medicare's enrollment data on race and ethnicity have inaccuracies and other limitations, particularly for those identified as American Indian/Alaska Native, Asian/Pacific Islander, or Hispanic.

Why OIG Did This Review

The disparate impacts of the COVID-19 pandemic on various racial and ethnic groups have brought health disparities to the forefront. Health disparities are differences in health that adversely affect certain groups. People of color have been found to experience disparities in areas such as access to health care and quality of health care.[1] Such

RUTGERS

Community Health and Aging Outcomes Laboratory

FAIR AIR QUALITY (AND DATA!)

OSF HOME

My Projects Browse and organize all your projects

All my projects and components

Collections

- All my projects and components
- All my registrations
- All my preprints
- Bookmarks (0)

Contributors

Only contributors on loaded resources are filterable.

Community Health and Aging Outcomes Laboratory

Anum Zafar

Name

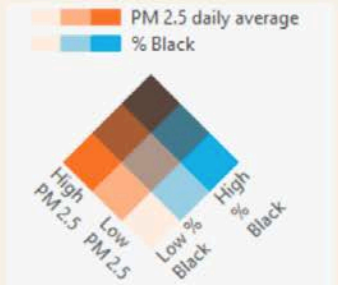
- Care Trajectory File
- Dementia Flag
- CHAO Race
- Annual Air Quality Data

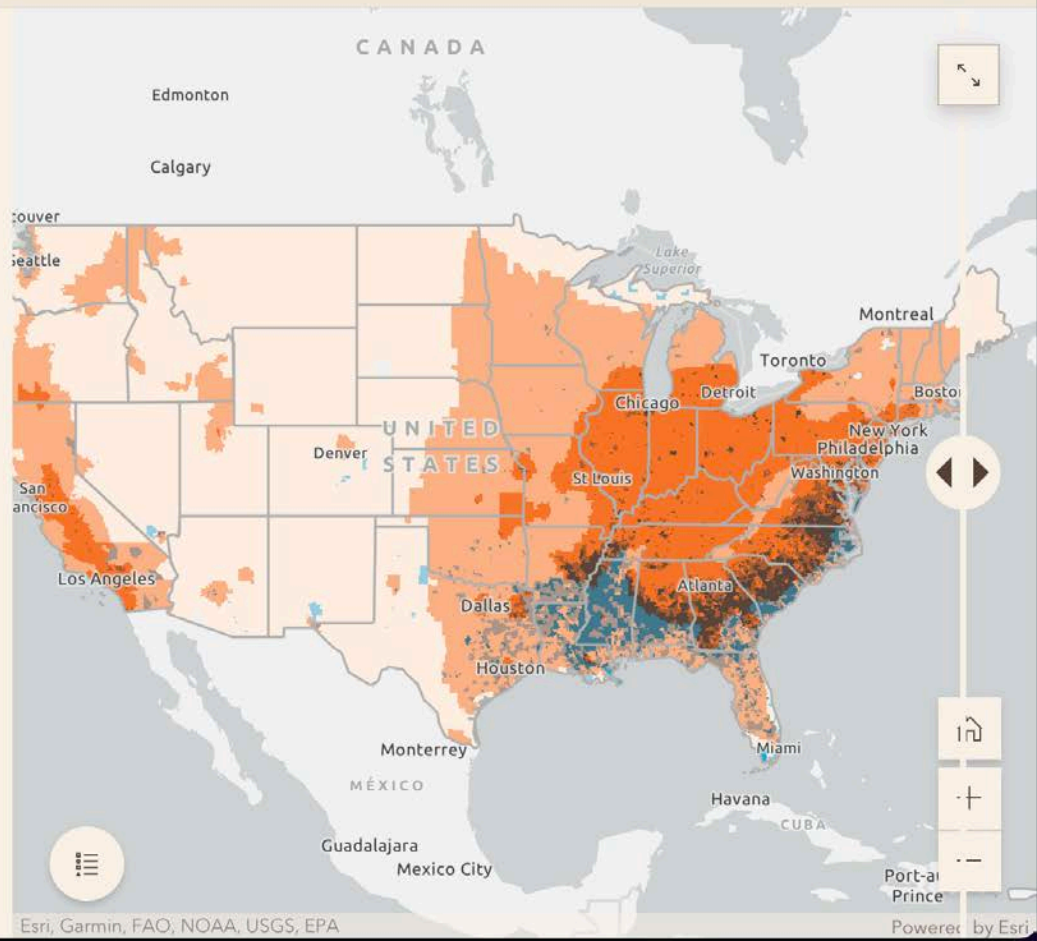
RUTGERS Air Quality and Health Equity

Air Pollution | Environmental Racism | About the Project

PM 2.5 & Black Communities

The maps show the relationship between PM 2.5 exposure and the percentage of the population in an area of Black race across the U.S. in 2002 (left) and 2018 (right). Black communities in Southern states are especially exposed to high level of PM 2.5.





RUTGERS
Community Health and Aging Outcomes Laboratory

CHALLENGES, FUTURE WORK

- Delays in accessing data in VRDC, access to Medicare files in July/August 2022, access to Medicaid files pending
- Awaiting access to SAS ML software in VRDC
- Seeking peers using VRDC Databricks environment for ML work (Python, R)
- Looking forward to the “fun” part ahead – testing and using pre-processing and harmonization work