Survey Weighting Methods and Initial Health Outcome Findings from the New Jersey Health and Nutrition Examination Survey (NJHANES)
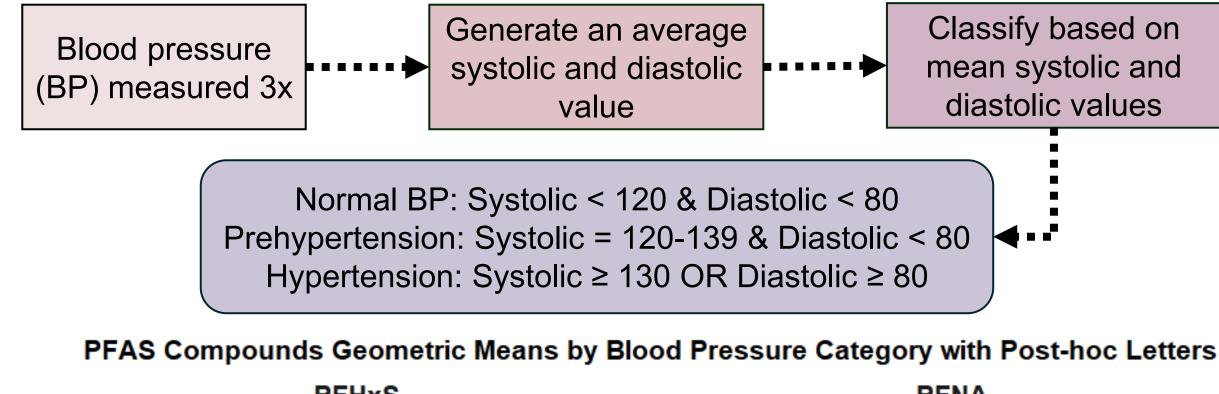
Chang Ho Yu[1], Colleen Donohoe[1], Meraya Sirimis[1,2], **Kayla Lam**[1,3], Elisabeth Cook[1], Eric Bind[1], and Zhihua (Tina) Fan[1]
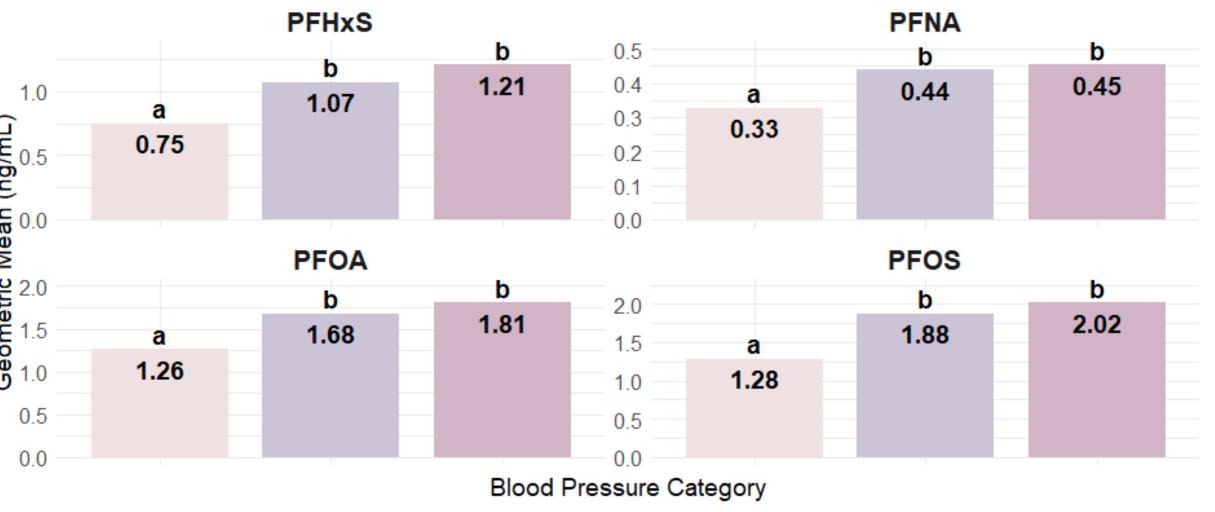
1. Environmental and Chemical Laboratory Services, Public Health & Environmental Laboratories, New Jersey Department of Health, Ewing, NJ 08628
2. The College of New Jersey, School of Public Health, Ewing, NJ 08618
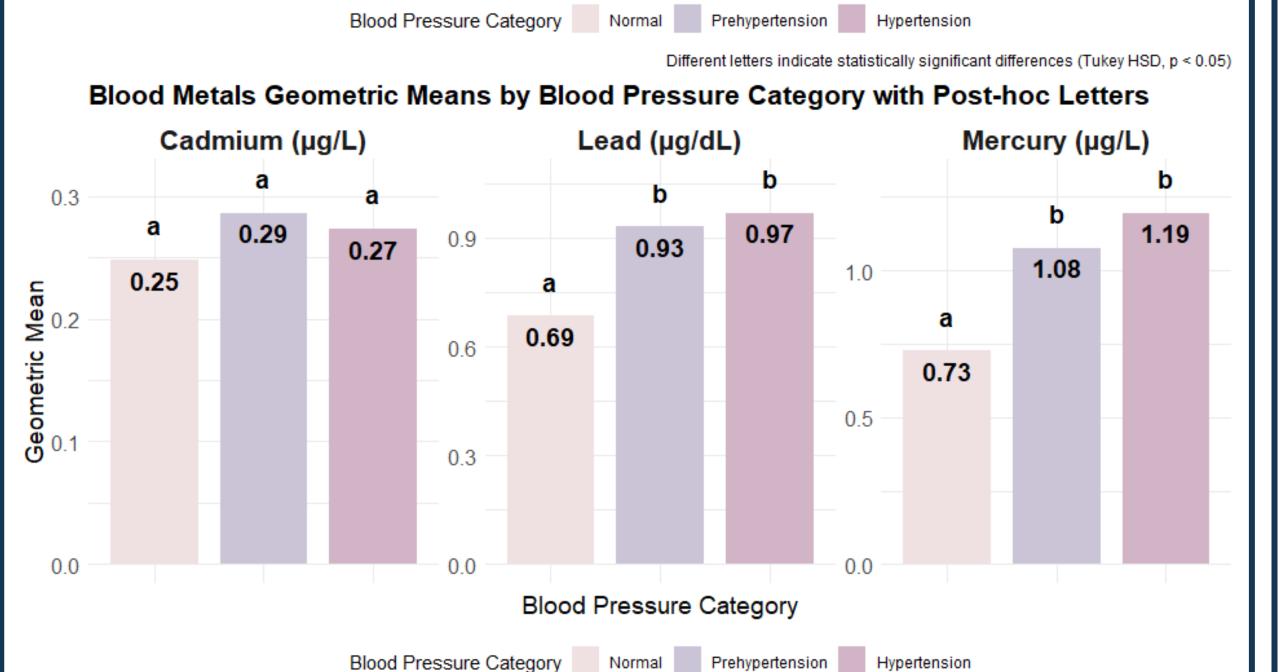3. APHL-CDC Fellow in Environmental Health, Bethesda, MD 20814

Introduction

- **New Jersey Health and Nutrition Examination Survey (NJHANES):** First probability-based population surveillance in New Jersey (NJ) that provides data on population-level exposures, health, and nutritional status to examine health outcomes and support public health planning/interventions

- **Study Population:** 503 sample persons (SPs) (ages 6 and older) have been enrolled since September 2021, with 446 completed assessments across NJ as of March 2025

- **Design:** Geographic area (GA) stratification and multi-stage clustered random sampling across NJ, modeled after the Centers for Disease Control and Prevention's (CDC) National Health and Nutrition Examination Survey (NHANES) (with modifications in stratification and clustering approaches)

- **Study Objectives:**
   a) To transfer development of individualized sample weights to provide design-consistent estimates for NJHANES; and
   b) To present initial health outcome findings from the ongoing NJHANES
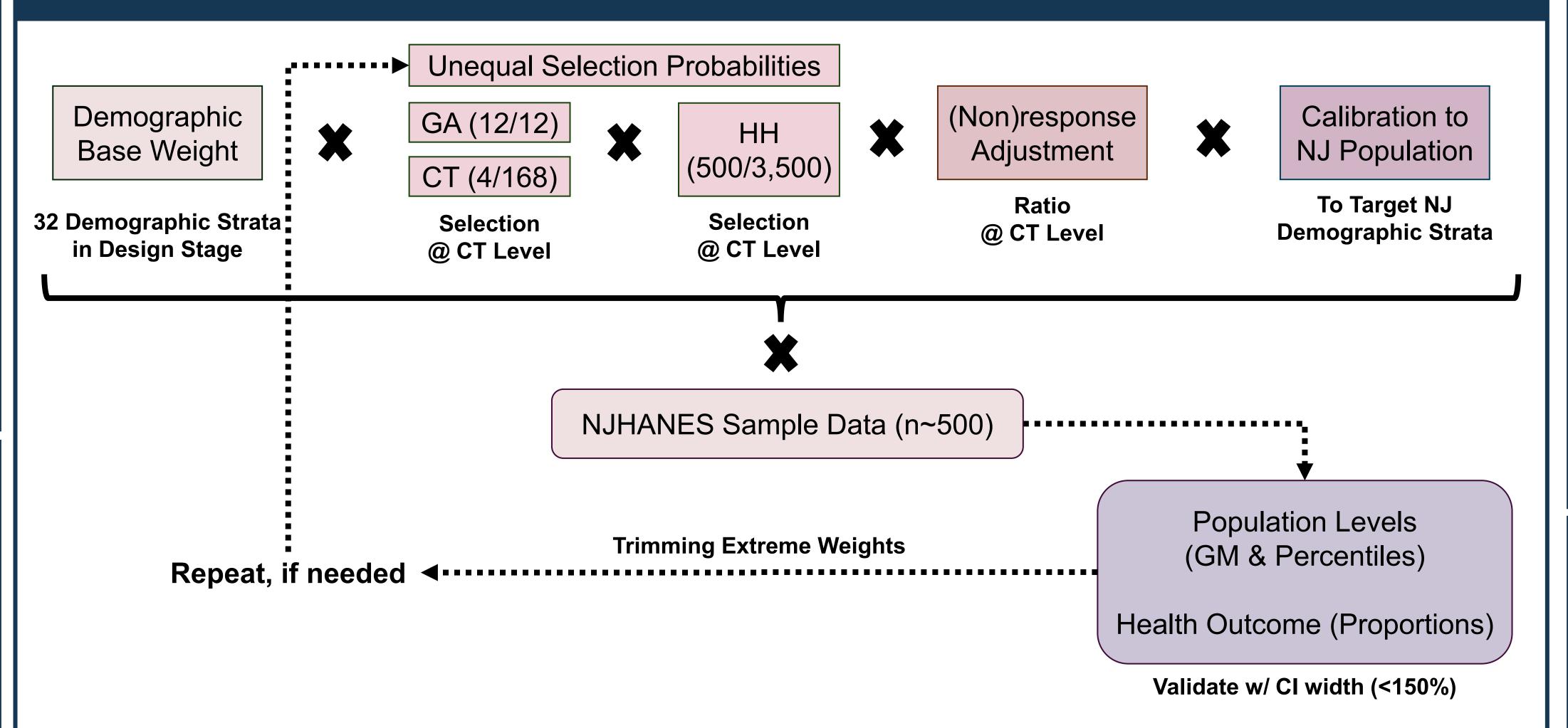
Results: Initial Health Findings

Blood pressure (BP) measured 3x → Generate an average systolic and diastolic value → Classify based on mean systolic and diastolic values

Normal BP: Systolic < 120 & Diastolic < 80
Prehypertension: Systolic = 120-139 & Diastolic < 80
Hypertension: Systolic ≥ 130 OR Diastolic ≥ 80

PFAS Compounds Geometric Means by Blood Pressure Category with Post-hoc Letters

PFHxS: a 0.75, b 1.07, b 1.21
PFNA: a 0.33, b 0.44, b 0.45
PFOA: a 1.26, b 1.68, b 1.81
PFOS: a 1.28, b 1.88, b 2.02

Blood Pressure Category: Normal, Prehypertension, Hypertension

Different letters indicate statistically significant differences (Tukey HSD, p < 0.05)

Blood Metals Geometric Means by Blood Pressure Category with Post-hoc Letters

Cadmium (µg/L): a 0.25, a 0.29, a 0.27
Lead (µg/dL): a 0.69, b 0.93, b 0.97
Mercury (µg/L): a 0.73, b 1.08, b 1.19

Blood Pressure Category: Normal, Prehypertension, Hypertension

Different letters indicate statistically significant differences (Tukey HSD, p < 0.05)

Key Associations Discovered:
- <u>PFAS</u>: NJHANES participants who were prehypertensive and hypertensive have significantly higher concentrations of the four selected PFAS (PFHxS, PFNA, PFOA, and PFOS)
- <u>Toxic Metals</u>: NJHANES participants who were prehypertensive and hypertensive have significantly higher concentrations of lead and mercury

Sample Weight Methodology

Demographic Base Weight (32 Demographic Strata in Design Stage) ✕ Unequal Selection Probabilities [GA (12/12), CT (4/168)] (Selection @ CT Level) ✕ HH (500/3,500) (Selection @ CT Level) ✕ (Non)response Adjustment (Ratio @ CT Level) ✕ Calibration to NJ Population (To Target NJ Demographic Strata)

✕

NJHANES Sample Data (n~500)

Population Levels (GM & Percentiles) / Health Outcome (Proportions) — Validate w/ CI width (<150%)

Trimming Extreme Weights — Repeat, if needed

Considerations and Adjustments Employed in NJHANES Sample Weighting

1. Base Weight Calculation
   - Incorporates selection probabilities at geographic area, census tract, and household levels established during NJHANES design
2. Selection Adjustment
   - Compensates for unequal household selection within each cluster and census tract
3. Response Adjustment
   - Accounts for census tract-level response versus non-response at second cluster stage

Population Pyramid of Sample Persons Targeted vs Recruited for NJHANES

Legend: Recruited, Targeted

Race/Ethnicity (B/W/H/O), Sex (M/F), Age — Count of Sample Persons

Group	Targeted	Recruited
O/F/60+	7	2
O/F/40-59	9	9
O/F/20-39	8	3
O/F/6-19	6	1
O/M/60+	7	4
O/M/40-59	9	3
O/M/20-39	8	3
O/M/6-19	5	4
W/F/60+	34	66
W/F/40-59	41	58
W/F/20-39	38	20
W/F/6-19	26	10
W/M/60+	33	60
W/M/40-59	39	28
W/M/20-39	37	18
W/M/6-19	25	11
H/F/60+	13	4
H/F/40-59	16	11
H/F/20-39	14	18
H/F/6-19	10	0
H/M/60+	12	3
H/M/40-59	15	3
H/M/20-39	14	5
H/M/6-19	9	3
B/F/60+	8	3
B/F/40-59	10	7
B/F/20-39	9	2
B/F/6-19	6	1
B/M/60+	8	4
B/M/40-59	8	4
B/M/20-39	9	3
B/M/6-19	6	3

4. Population Calibration
   - Aligns sample demographics with NJ 2020 census data
   - Ensures representative distribution across age, sex, and racial/ethnic groups
5. Weight Trimming
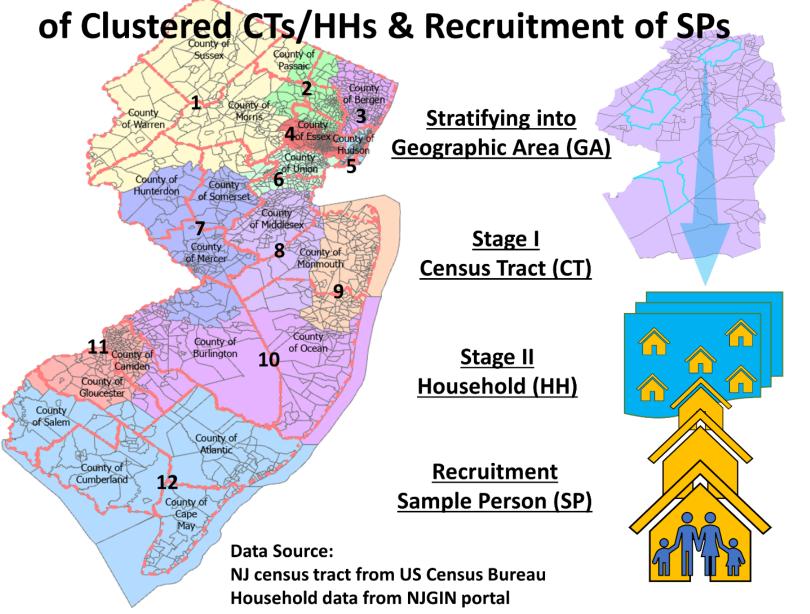   - Reduces extreme weights using statistical thresholds (>AVG + 3*SD) and ML techniques (K-means clustering, Isolation Forest, or XGBoost Regression)
   - Maintains statistical validity while controlling variance levels

NJHANES Design

NJHANES: Stratified by GA, Random Selection of Clustered CTs/HHs & Recruitment of SPs

Stratifying into Geographic Area (GA)

Stage I Census Tract (CT)

Stage II Household (HH)

Recruitment Sample Person (SP)

Data Source:
NJ census tract from US Census Bureau
Household data from NJGIN portal

Conclusions

- Data Validation and Integrity:
   - Issues with missing data due to unanswered survey questions that greatly hindered efforts on weight calculation and subsequent data analysis
      - Focus on ensuring key survey questions are answered and double-checked by study staff
      - Develop a back up plan to contact participants for missing information
   - Data integrity issues with survey responses and physical examination readings
      - Develop a thorough review protocol and implement a validation/correction process

- Study Implementation
   - Obtained and relied on inaccurate postal addresses
   - Fieldwork burden, partially due to the termination of mobile laboratory services in the study period
      - Considering onboarding a new mobile laboratory service that will collect and examine physical exams, leveraging current resources, and/or onboarding new personnel (such as APHL fellows)

Future Applications

- Complete field work and survey collection by Q2 2025
- Finalize sample weight methodology with appropriate adjustments, trimming, and calibration
- Validate findings through repeated analysis with finalized sample weights
- Conduct parallel analysis with NHANES data to expand NJHANES data utilization and impact on NJ health studies
- Sample size will be re-evaluated and determined with a balance of funding, capacity, and statistical validity for subdomain analyses

References

- Johnson CL, Dohrmann SM, Burt VL, Mohadjer LK. *National Health and Nutrition Examination Survey: Sample design, 2011–2014.* National Center for Health Statistics. Vital Health Stat 2(162). 2014.
- U.S. Census Bureau. 2020 Census Redistricting Data (PL 94-171). U.S. Census Bureau. 2021.

Acknowledgements