CENSUS & DIFFERENTIAL PRIVACY:

PUBLIC HEALTH & HEALTH EQUITY

QUESTIONS

Nancy Krieger, PhD
Harvard T.H. Chan School of Public Health

QUESTIONS

Longstanding connections: census & public health

(1) Counts: denominators

(2) Place: metrics to characterize areas

(3) Time: temporal trends & discontinuities

(4) Counts for resources & representation: societal determinants of health

The US Census and the People's Health: Public Health Engagement From Enslavement and "Indians Not Taxed" to Census Tracts and Health Equity (1790–2018)

Nancy Krieger, PhD

Am J Public Health. 2019;109:1092–1100.

PATTERNING OF PARISIAN MORTALITY BY WEALTH: VILLERMÉ,1826

Arrondi-sement	N of residents	% untaxed rents	Mortality	
			Proportion	per 1000
2 (wealthiest)	65,623	7	1 in 62	16.1
3	44,932	11	1 in 60	16.7
1	52,421	11	1 in 58	17.2
4	46,624	15	1 in 58	17.2
11	51,766	19	1 in 51	19.6
6	72,682	21	1 in 54	18.5
5	56,871	22	1 in 53	18.9
7	56,245	22	1 in 52	19.2
10	81,133	23	1 in 50	20.0
9	42,932	31	1 in 44	22.7
8	62,758	32	1 in 43	23.3
12 (poorest)	80,079	38	1 in 43	23.3

Source: Villermé LR. Rapport fait par M. Villermé, et lu à l'Académie royale de Médicine, au nom de la Commission de statistique, sur une série de tableaux relatifs au mouvement de la population dans les douze arrondisements municipaux de la ville de Paris, pendant les cinq années 1817, 1817, 1819, 1820 et 1821. Archives Générales de Médicine 1826;10:216-247. The table is on p. 227.

REMINDER ABOUT RATES

Rate: $\dfrac{N\ of\ cases}{N\ of\ population\ at\ risk}$ per specified unit of time

- **Deflate denominator → inflated rate estimate**

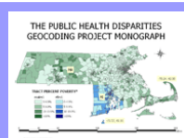- **Inflate denominator → deflated rate estimate**

AND: systematic error in denominator (e.g., for areas, for social groups) → systematic bias in rate estimation

What might be the impact of differential privacy on accurately measuring local and national distributions (by areas, by social groups) of:

-- rising death rates in US middle-age adults?

-- HIV/AIDS in US South?

QUESTION #1: COUNTS

The Public Health Disparities Geocoding Project Monograph

Geocoding and Monitoring US Socioeconomic Inequalities in Health: An introduction to using area-based socioeconomic measures

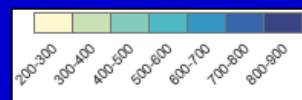| WHY?	READ MORE		HOW TO				TRY IT OUT!	TOOLS	
Executive Summary	Introduction	Publications	Geocoding	Generating ABSMs	Analytic Methods	Multi-level Modeling	Visual Display	Case Example	U.S. Census Tract Poverty Data	Glossary

Welcome to the Public Health Disparities Geocoding Project Monograph.

These pages present an introduction to geocoding and using area-based socioeconomic measures with public health surveillance data, based on the work of the Public Health Disparities Geocoding Project at the Harvard School of Public Health, Department of Society, Human Development, and Health.
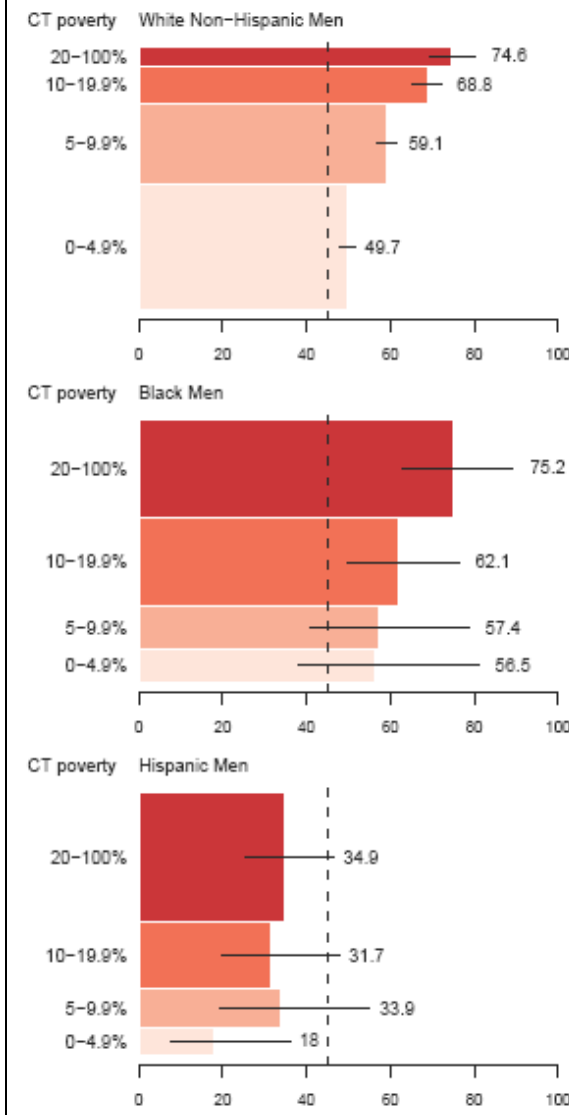
Health outcome (premature mortality rate) & *single* census tract variable (% below poverty)
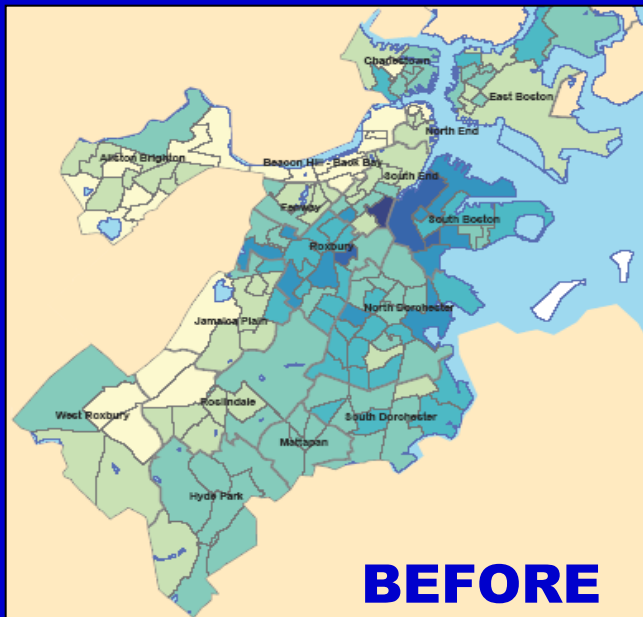
BEFORE

AFTER

CENSUS TRACT (CT) AGE-STANDARDIZED PREMATURE MORTALITY RATES (DEATH <75, PER 100,000) BEFORE AND AFTER ADJUSTING FOR CT POVERTY: BOSTON, 1999-2001

Legend: 200-300, 300-400, 400-500, 500-600, 600-700, 700-800, 800-900

LUNG CANCER MORTALITY: MA, 2000-2005

CT poverty — White Non-Hispanic Men
- 20-100%: 74.6
- 10-19.9%: 68.8
- 5-9.9%: 59.1
- 0-4.9%: 49.7

CT poverty — Black Men
- 20-100%: 75.2
- 10-19.9%: 62.1
- 5-9.9%: 57.4
- 0-4.9%: 56.5

CT poverty — Hispanic Men
- 20-100%: 34.9
- 10-19.9%: 31.7
- 5-9.9%: 33.9
- 0-4.9%: 18

Sources: Krieger et al in AJPH 2005 + 2003 + 2002 + 2001; AJE 2002; JECH 2003; PHR 2003; Chen et al J Urban Health 2006; NIH/NICHD R01 HD3685 + R25 HD047185.

Better Data for Better Health

Health data resources for analysis of the many factors which shape health in communities, and local efforts to achieve better health for all.

Health data are a powerful tool to help communities better understand the health of their neighborhoods and residents.

The health data sources in this collection range from public health and social service agencies to hospitals and insurers, and include multiple types of data at the state, county, city, and neighborhood levels. Across the nation, the growing variety of data sets now being aggregated and shared is providing an increasingly clear picture of health challenges that communities are experiencing, and driving residents, community leaders, policymakers, and advocates to come together to set common goals for improvement.

In connection with programs focused on creating healthier communities, the Robert Wood Johnson Foundation (RWJF) supports multiple efforts to leverage health data to advance efforts to achieve better health. With partners across the public, private and nonprofit sectors, we are working to expand the potential for data to illuminate health gaps in communities, and areas where action is needed. The more communities take advantage of the many forms of health data now available, the better they can target resources to assure everyone has a fair and just opportunity for health.

Robert Wood Johnson Foundation

HEALTH DATA RESOURCES

County Health Rankings

The County Health Rankings show that where we live matters to how long and how well we live. The Rankings rank nearly every county in the nation, based on 35 factors that impact health, such as high school graduation rates, housing, employment, income, and access to healthy foods.

500 Cities

A collaboration of RWJF, the Centers for Disease Control and Prevention, and the CDC Foundation, 500 Cities enables users to easily browse data about health in the nation's 500 largest cities, based on measures of health related to unhealthy behaviors, health outcomes, and prevention practices.

City Health Dashboard
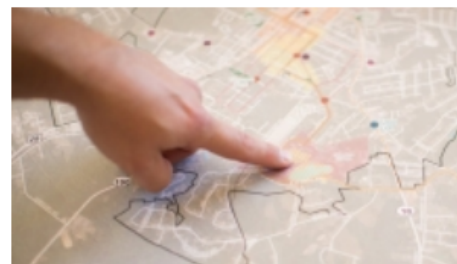
Developed with RWJF support, the City Health Dashboard allows users to see correlations between community-level factors that shape health in cities, such as housing affordability, unemployment, children in poverty, and access to nutritious foods.

Model-based estimates, by CT: Washington, DC

Asthma: 2016 (adults 18y+)
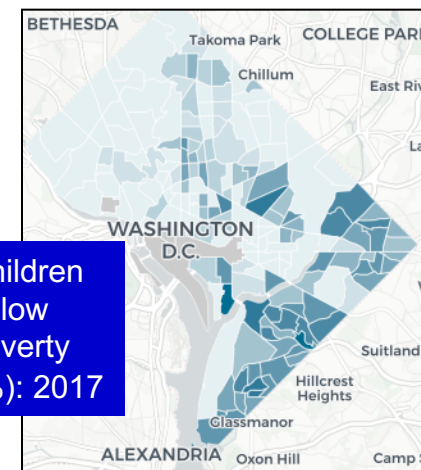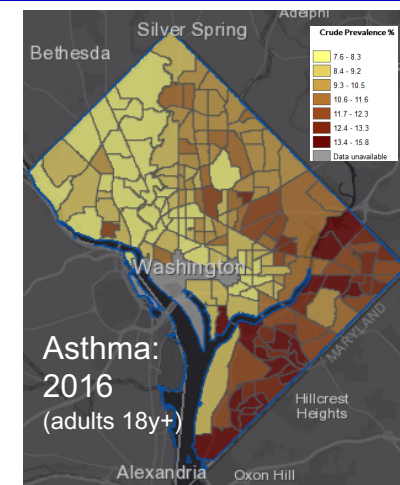
Children below poverty (%): 2017
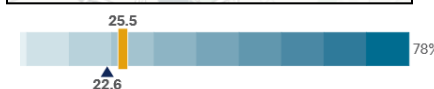
QUESTION #2: PLACE

Using the Index of Concentration at the Extremes at multiple geographical levels to monitor health inequities in an era of growing spatial social polarization: Massachusetts, USA (2010–14)

Nancy Krieger*, Rockli Kim, Justin Feldman and Pamela D Waterman
International Journal of Epidemiology, 2018, 788–819

Health outcome (child mortality rate) & area-based metrics for spatial social polarization, involving relations between social groups at each geographic level

Index of Concentration at the Extremes (per Massey, 2001)

$$ICE_i = (A_i - P_i)/T_i$$

where, say, for racialized economic segregation,

A_i = N of high income white non-Hispanic persons in neighborhood

P_i = N of low income black non-Hispanic persons in neighborhood

T_i = total N with known income in neighborhood

range: −1 (total deprivation) to 1 (total privilege)

Figure 1a: Child mortality incidence rate ratios* by ICE/poverty quintile (Q5: most privileged; referent) for the total population, Massachusetts, 2010-2014.
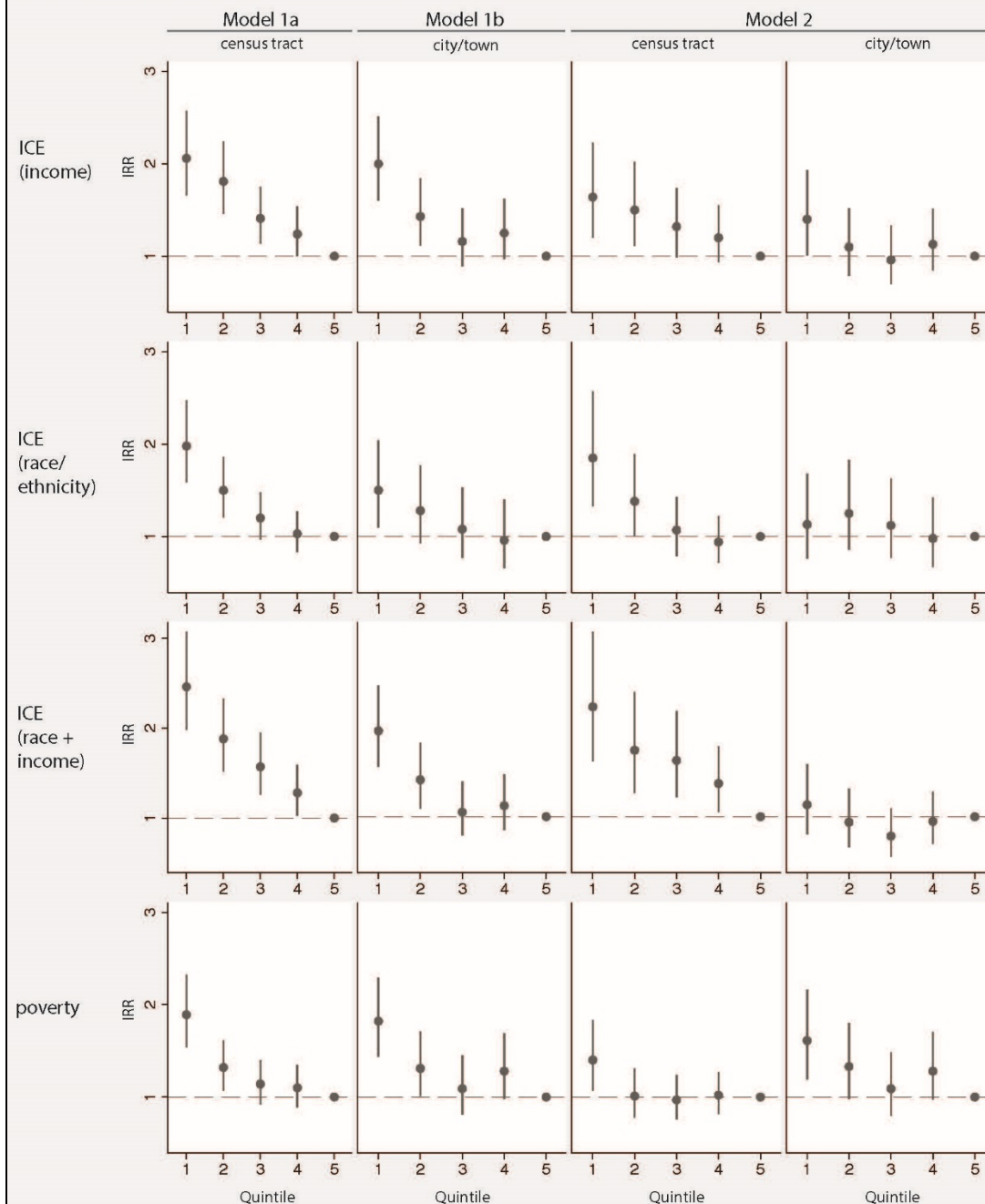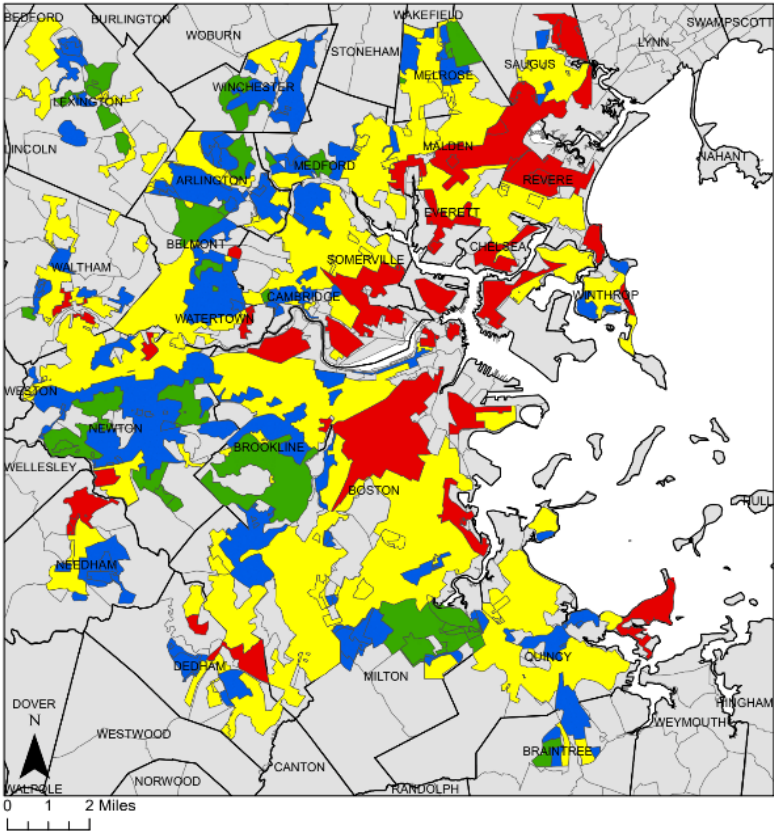
* Results from multilevel Poisson models for age-standardized mortality rates that adjusted for gender, race/ethnicity, and urbanicity

HISTORICAL REDLINING & CANCER STAGE AT DIAGNOSIS: BOSTON, 2011-2015

NOTE: CT HOLC status determined in relation to % of CT population in a given HOLC area, using *census block population counts*

Boston area HOLC map, 1938 (digitized)

Home Owners' Loan Corporation Grades

- 🟩 Best
- 🟦 Still Desirable
- 🟨 Definitely Declining
- 🟥 Hazardous

HOLC category (1938)	Census tracts (N = 151)	ICE: racialized economic segregation (2011-2015)		Poverty (2011-2015)	
	N	Score (mean)	Absolute difference (95% CI)	%	Absolute difference (95% CI)
Green + blue ("best" + "still desirable")	10	0.43	0	10.5	0
Yellow ("definitely declining")	79	0.06	-0.4 (-0.4, -0.3)	21.7	11.1 (8.6, 13.7)
Red ("hazardous)	62	0.14	-0.3 (-0.3, -0.2)	23.5	12.9 (10.2, 15.7)

QUESTION #3: TEMPORAL ISSUES

Why history matters for quantitative target setting:
Long-term trends in socioeconomic and racial/ethnic
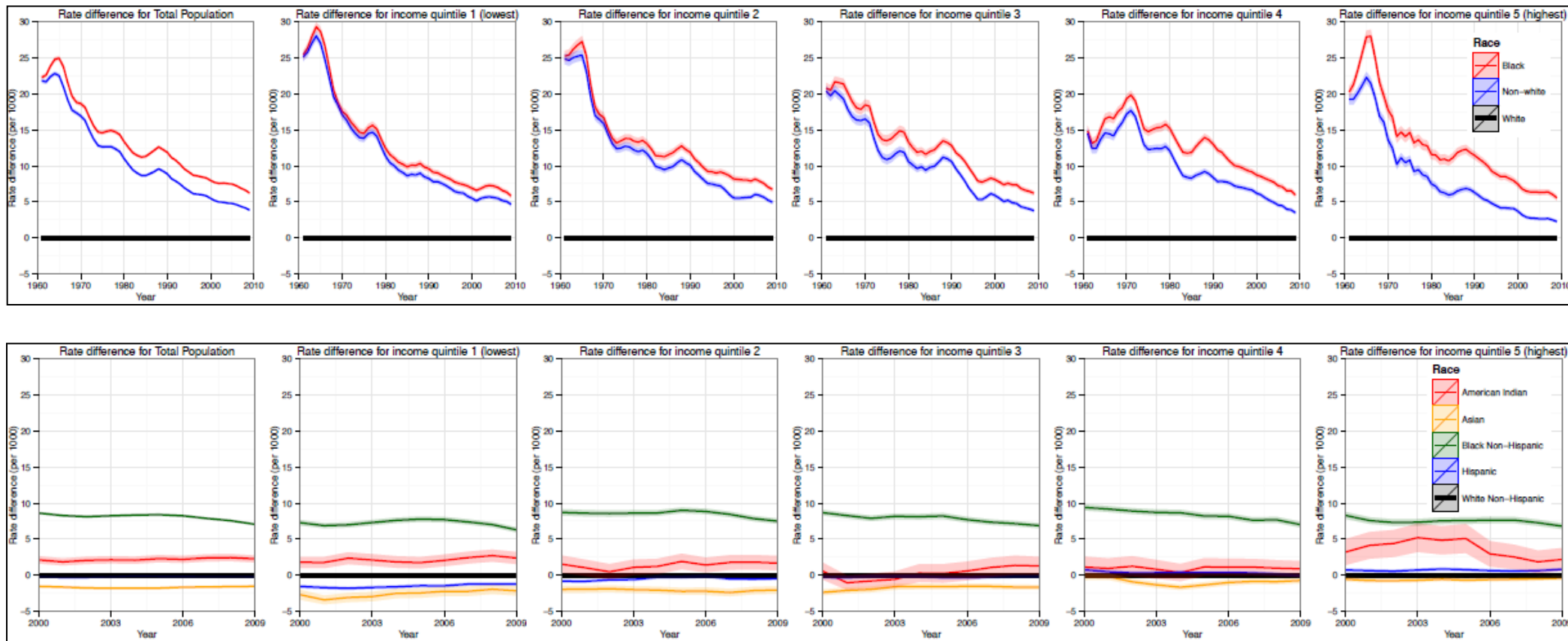inequities in US infant death rates (1960–2010)

Journal of Public Health Policy (2015) 36, 287–303

Nancy Krieger[a,*], Nakul Singh[b], Jarvis T. Chen[c], Brent A. Coull[b],
Jason Beckfield[d], Mathew V. Kiang[c], Pamela D. Waterman[c], and
Sofia Gruskin[e]

Discontinuities

1) 1997 OMB change: race + ethnicity

2) NCHS: change in age-standard

QUESTION #4: RESOURCES & REPRESENTATION

Federal programs relevant to health: funding formula use census data

Uses of Census Bureau Data in Federal Funds Distribution

A New Design for the 21st Century

Issued September 2017
Version 1.0
Prepared by Marisa Hotchkiss, Jessica Phelan

N =132 federal programs, $675 billion (FY 2015); among top 18, (each >$4 billion), 13 with direct health impacts (★)

Table 1: Federal Assistance Distributed Using Census Bureau Data in Fiscal Year 2015

	CFDA number[1]	Program Name	Federal Executive Department or Agency[2]	Fiscal Year 2015 Fund
★	93.778	Medical Assistance Program	HHS	$311,805,244,413
★	10.551	Supplemental Nutrition Assistance Program	USDA	$71,035,786,000
★	93.774 (part)	Medicare Part B Physicians Fee Schedule Services	HHS	$70,300,000,000
	20.205	Highway Planning and Construction	DOT	$38,479,013,855
	84.063	Federal Pell Grant Program	ED	$29,916,694,438
★	10.555	National School Lunch Program	USDA	$18,915,944,292
★	93.558	Temporary Assistance for Needy Families	HHS	$17,225,738,021
★	14.871	Section 8 Housing Choice Voucher	HUD	$15,761,488,440
	84.010	Title 1 Grants to Local Educational Agencies	ED	$14,253,154,251
	84.027	Special Education Grants to States	ED	$11,382,885,850
★	93.600	Head Start	HHS	$8,538,887,781
★	10.557	Special Supplemental Nutrition Program for Women, Infants, and Children	USDA	$6,062,899,861
	20.507	Federal Transit Formula Grants	DOT	$5,452,882,796
★	93.658	Foster Care Title IV-E	HHS	$5,409,221,818
★	93.596	Child Care Mandatory and Matching Funds of the Child Care and Development Fund	HHS	$5,314,103,129
★	14.195	Section 8 Housing Assistance Payments Program	HUD	$4,367,081,456
★	93.767	State Children's Health Insurance Program	HHS	$4,212,457,713
★	10.553	School Breakfast Program	USDA	$4,057,189,000

Political representation & redistricting: resource allocation and policies that are societal determinants of health

A Spatially Informed Analysis of Environmental Justice: Analyzing the Effects of Gerrymandering and the Proximity of Minority Populations to U.S. Superfund Sites

David E. Kramar, Aaron Anderson, Hayley Hilfer, Karen Branden, and John J. Gutrich

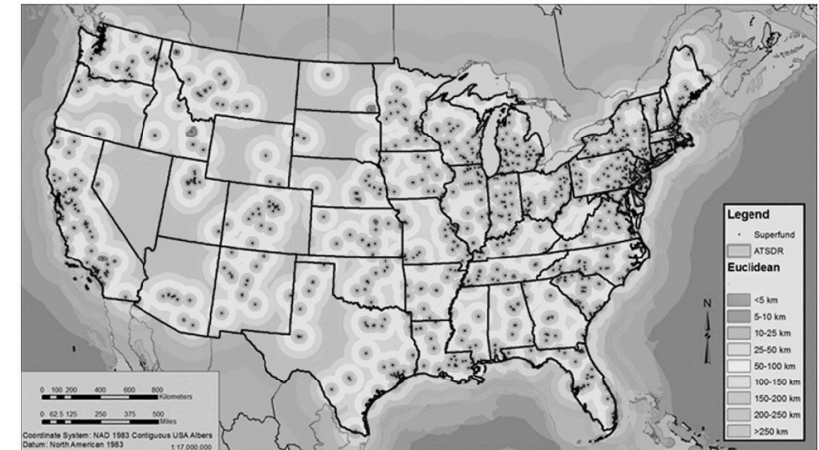ENVIRONMENTAL JUSTICE
Volume 11, Number 1, 2018

FIG. 1. Superfund sites located within the conterminous United States, and the Euclidean distance raster calculated from the superfund site locations

Key findings:
1) strong relationships between gerrymandering and (a) proximity to superfund sites, and (b) % white
2) "minority populations are effectively 'gerrymandered out' of the white and lower environmental hazards districts"

CORE QUESTIONS FOR PUBLIC HEALTH & HEALTH EQUITY

how will differential privacy affect:

COUNTS: for denominators and rates? for area-based metrics?

RELATIONSHIPS BETWEEN VARIABLES: within and across geographic levels?
  -- at the individual level (e.g., "race" x "ethnicity")?
  -- for area-based metrics?
  -- for analyses of health inequities? – especially for:
  -- "small" populations? (e.g., Indigenous populations; immigrant sub-groups)
  -- using census tract level data? block group data? or block data?

MONTORING TRENDS: in population health? in health inequities?

RESOURCES & REDISTRICTING: impact on these social determinants of health?

look forward to learning the answers – and sharing what I learn with my colleagues in public health & advocates for health equity