

CDC Sub-County Cancer Project

NAACCR June 25, 2020

Division of Cancer Prevention and Control

National Environmental Public Health Tracking Program



Sub-County Cancer Data Project

Partnership between

- National Environmental Public Health Tracking Program
- Division of Cancer Prevention and Control
- Ongoing collaboration – DCPC provides cancer incidence data at the state & county level to the Tracking Program. These data are mapped here:
<https://ephtracking.cdc.gov/DataExplorer/#/>
- Intended Outcome of *this* Project: Map cancer incidence data at a sub-county level



CDC Team

Division of Cancer Prevention and Control

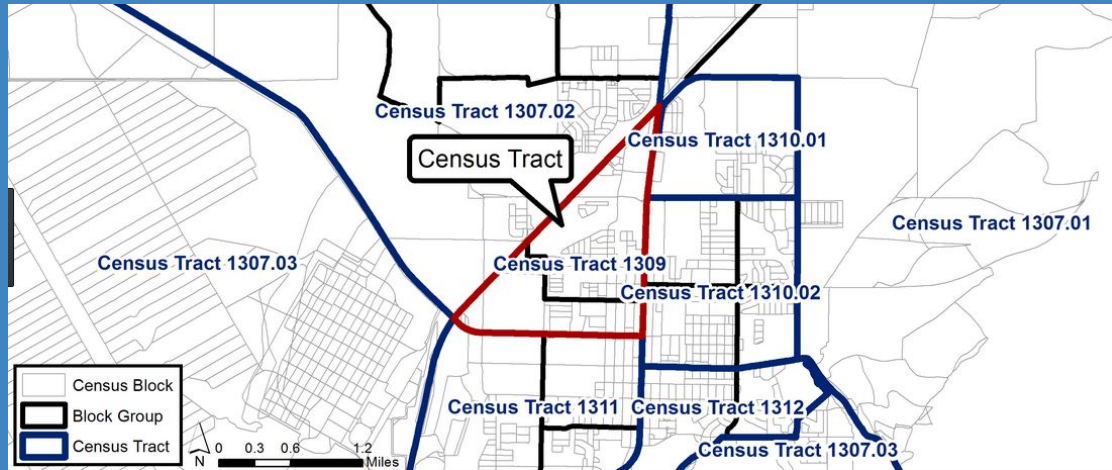
- Jane Henley
- Taylor Ellington
- Reda Wilson
- Simple Singh
- Manxia Wu
- Vicki Benard
- Lisa Richardson

Tracking Program

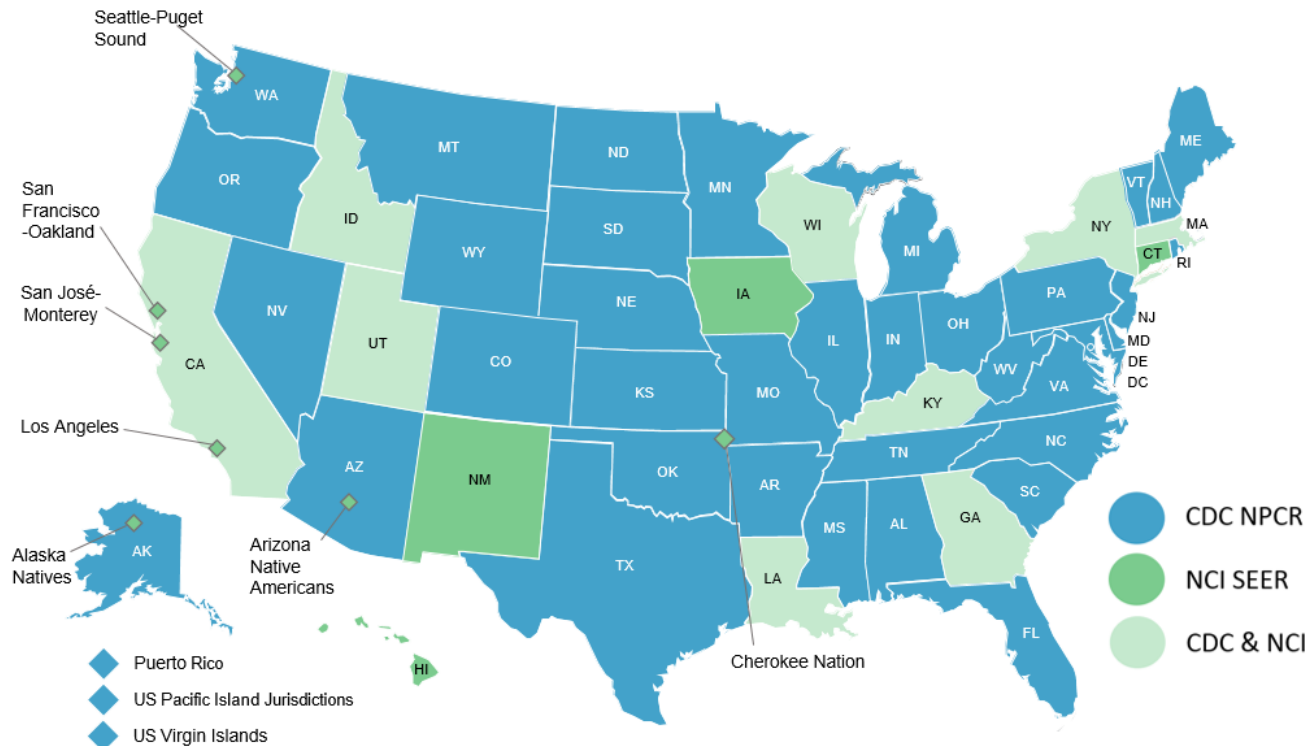
- Angie Werner
- Mackenzie Malone
- Aaron Vinson
- Heather Strosnider
- Fuyuen Yip

WHY USE SUB-COUNTY DATA?

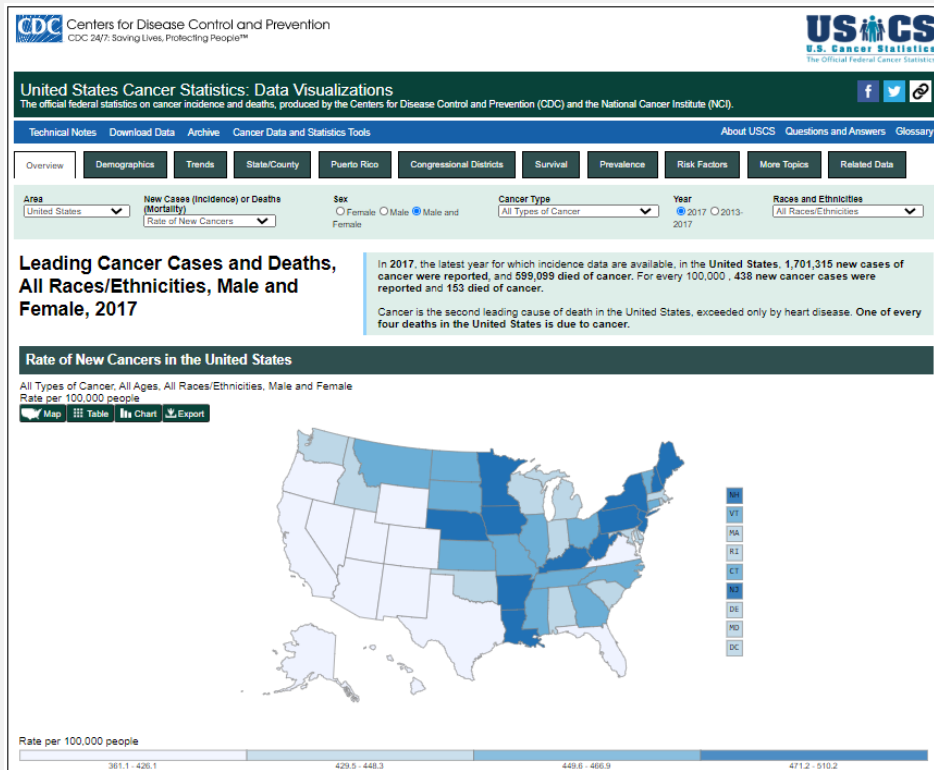
- Small area data can:
 - Highlight local variation
 - Allow for a better understanding of environmental health processes and impacts
 - Improve surveillance
 - Target interventions
- Small area data can also:
 - Create data reliability issues
 - Confidentiality issues



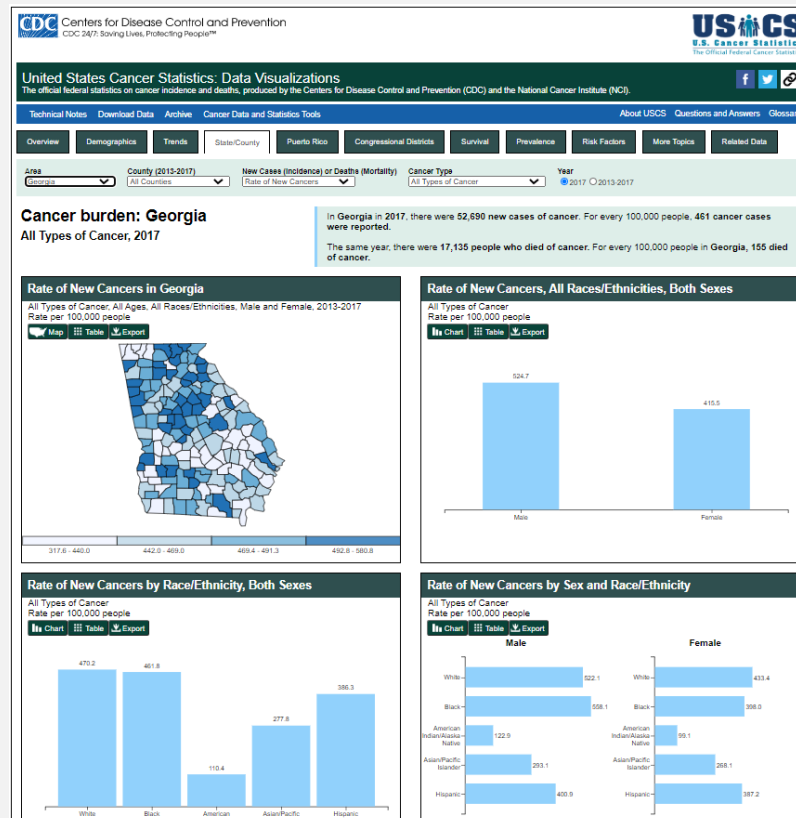
U.S. Cancer Surveillance



Local data are valuable. How local can we go?

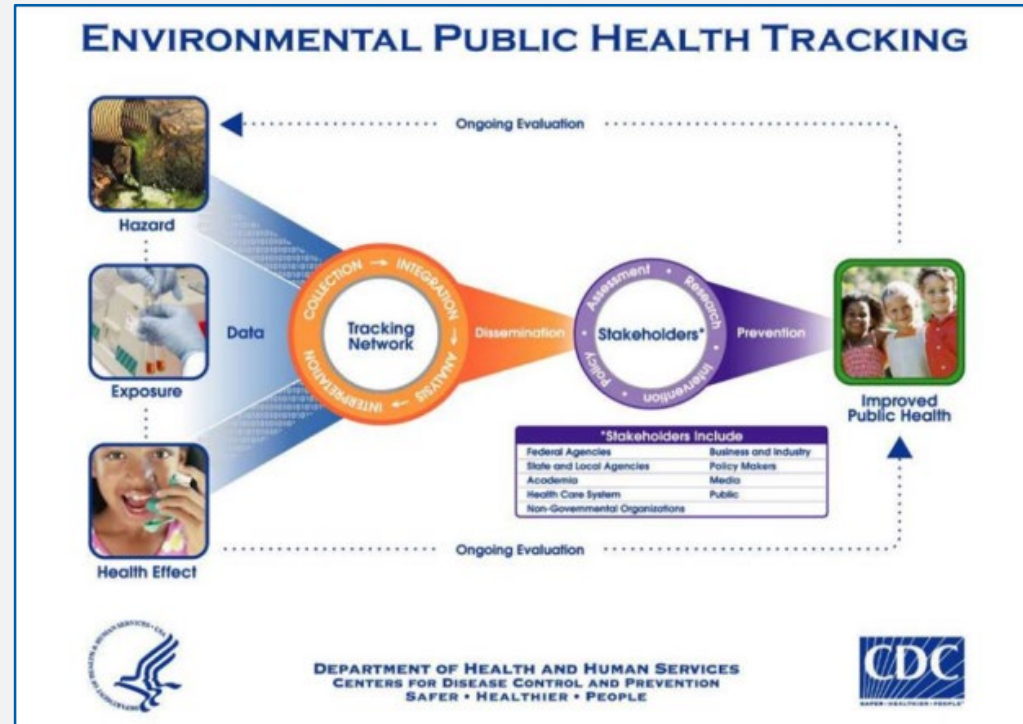


www.cdc.gov/cancer/dataviz



National Environmental Public Health Tracking Program

- **Vision:** Healthy informed communities
- **Mission:** To provide information from a nationwide network of integrated health and environmental data that drives actions to improve the health of communities



Funded Programs

26

25 states + 1 city

CDC-ASTHO Tracking Fellowships

43

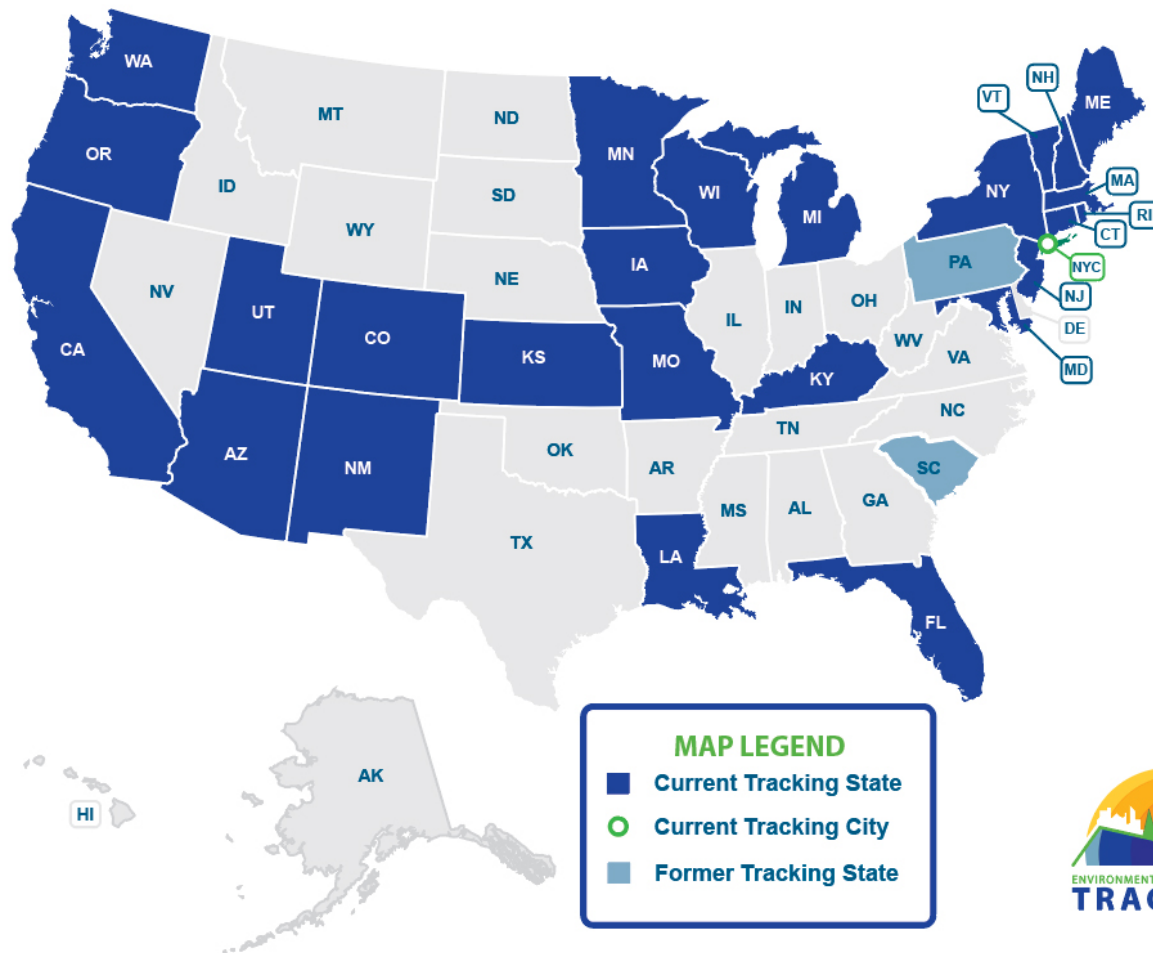
since 2009

Partnerships

CDC programs,
Federal agencies,
national organizations

Public Health Actions

500+



Air Quality

Asthma

Biomonitoring

Birth Defects

Cancer

CO Poisoning

Childhood Lead Poisoning

Chronic Obstructive Pulmonary Disease

Climate Change

Community Characteristics

Community Design

Community Drinking Water

Developmental Disabilities

Drought

Heart Disease

Heat Stress Illness

Lifestyle Risk Factors

Pesticide Exposures

Population and Vulnerabilities

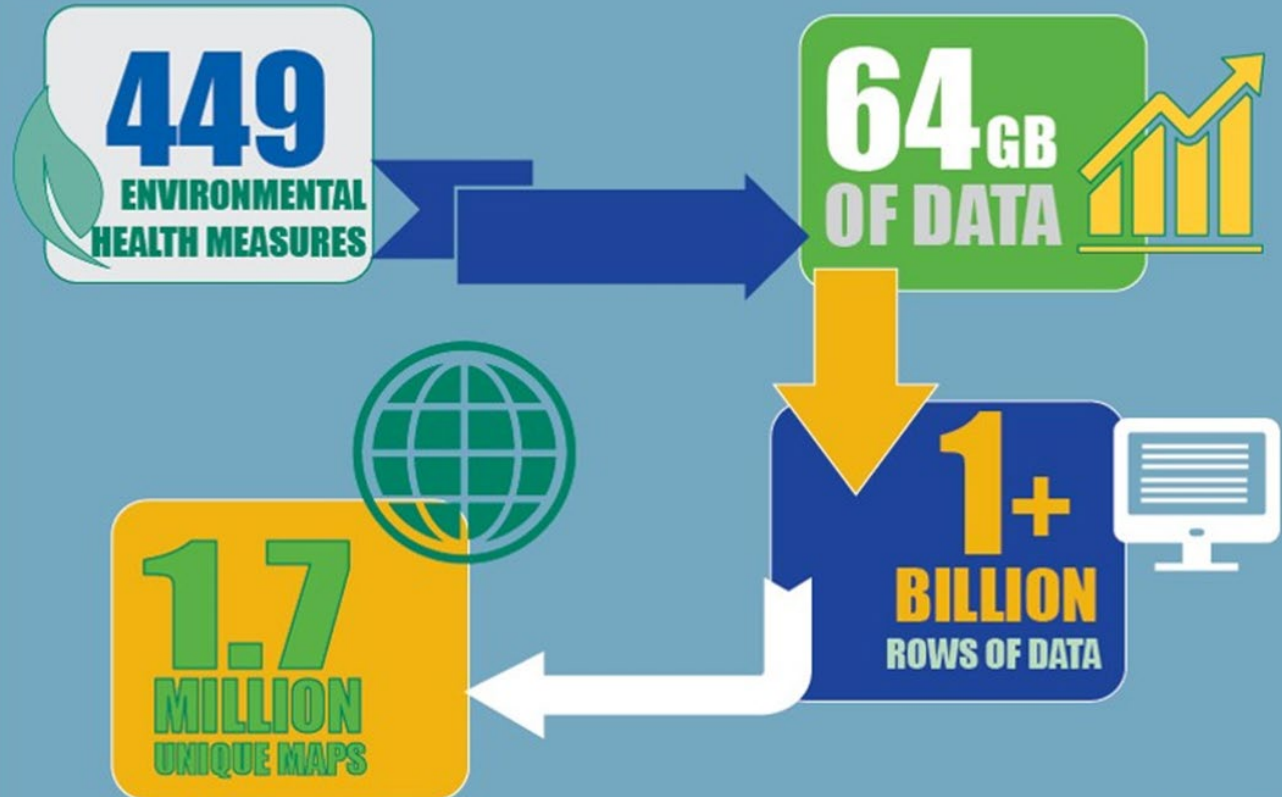
Reproductive & Birth Outcomes

Sunlight and UV

Toxic Substance Releases



ENVIRONMENTAL PUBLIC HEALTH TRACKING



SUB-COUNTY DATA PILOT

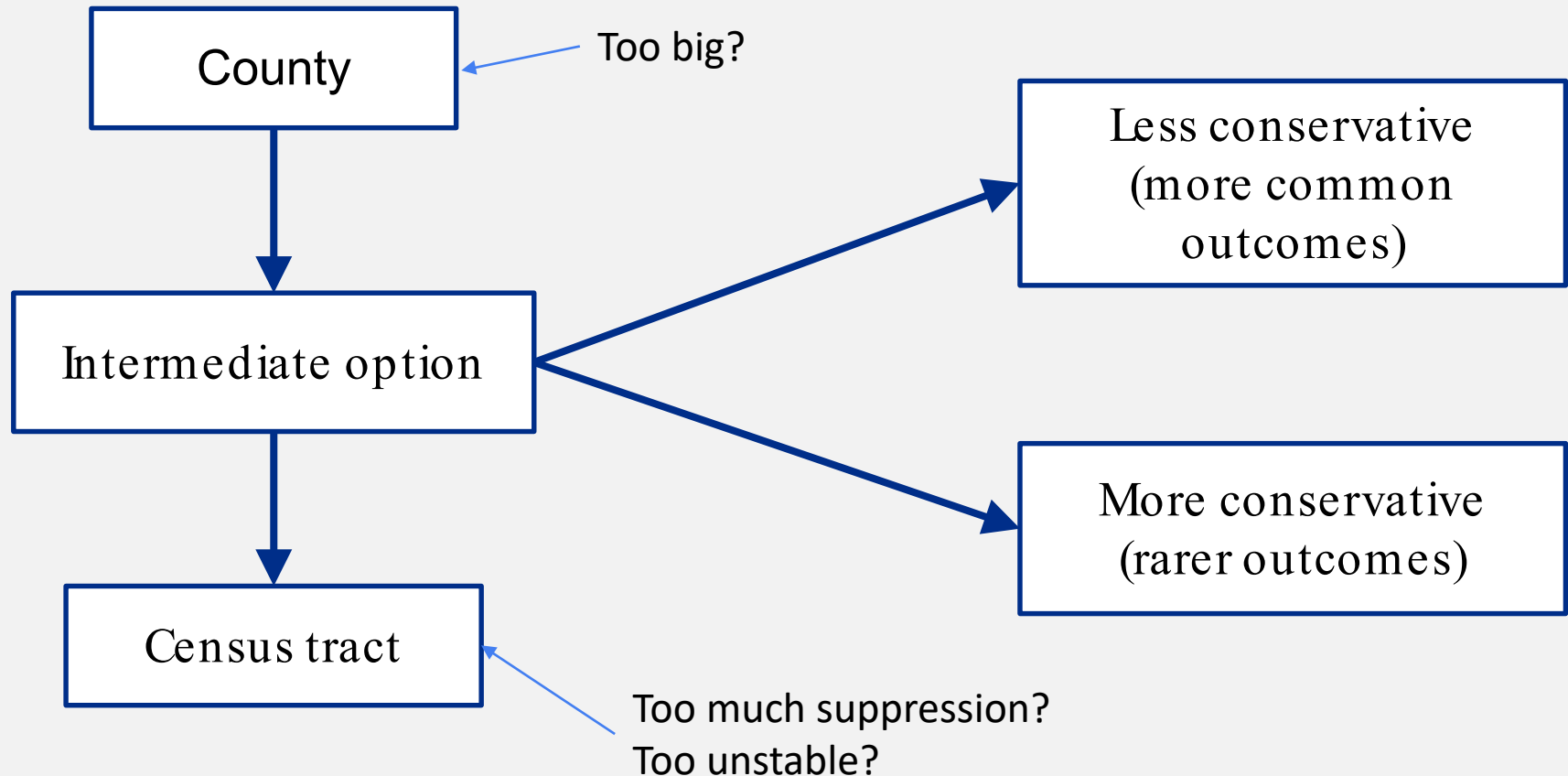
Acknowledgements

- **Tracking Grantees**
- **CDC Sub-County Content Workgroup (CWG)**
- **CDC Geospatial Standards and Network Development (SND) Team**
- **Ross Strategic**

2014 Tracking Program Sub-County Data Pilot – Tracking Grantees

- Developed **standardized sub-county geographies** to allow comparability across the Tracking Program's various datasets for environmental hazards, health exposures, and health outcomes.
- Created using the **Geographic Aggregation Tool** to merge based on the nearest population-weighted centroid until a specified threshold was reached.
- These standardized sub-county geographies
 - use Census tracts as the foundation,
 - have a hierarchical structure, and
 - nest within county boundaries.

Test aggregation schemes using census tracts as the foundation



Methods

- Calculated census tract-level expected case counts
- Shapefiles used as input for Geographic Aggregation Tool (GAT)
 - Created by NYS Department of Health's Environmental Health Surveillance Section to join neighboring geographic areas
- Ran GAT for various population thresholds
- Exported GAT data from ArcGIS, imported into SAS (prevalence, CI, RSE, suppression, case count distribution)

Considerations for Decision-Making

- Temporality (cross-sectional versus longitudinal)
- Compatibility between data and measures
- Ensuring protection of confidential data
- Feasibility of mapping cancer data at the sub-county level
- Communications and technical issues with display

STANDARDIZED SUB-COUNTY GEOGRAPHIES

Classification	Median case count range	Spatial aggregation level
Very common outcome	≥ 17.0 cases	Census tract
Common outcome	7.3 to 16.9 cases	Total population 5,000 persons
Rare outcome	1.9 to 7.2 cases	Total population 20,000 persons

Note: to ensure stability and protect confidentiality, spatial aggregation may also require temporal aggregation (eg, 3, 5, 7, or 10 year groups).

Cancer is considered a rare outcome.

Geographic aggregation

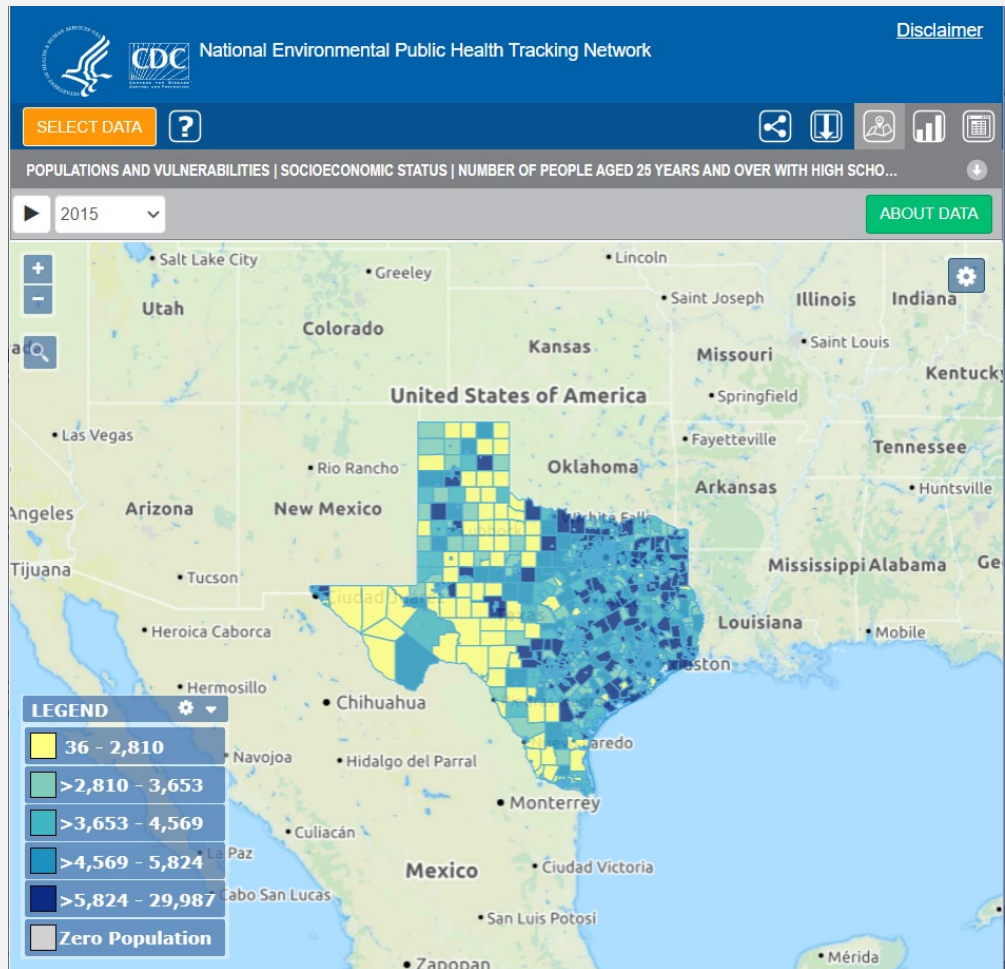
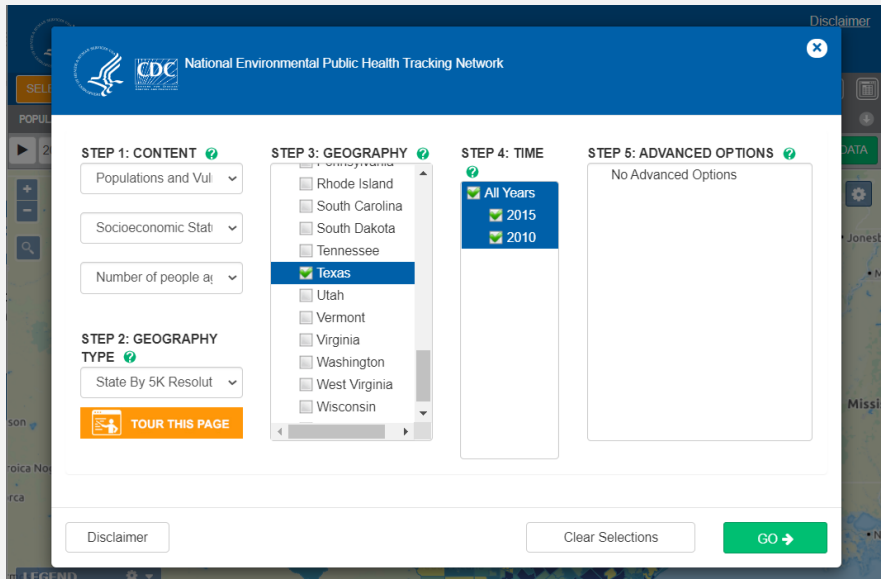
- **CDC will:**

- Use hierarchical structure aggregations with population-weighted centroid method with zero population tracts removed
- Exclude counties that do not meet population threshold from aggregation (TBD in how to treat those counties)
- Update geographies with new census years (e.g., 2010-2019 will use 2010 boundaries, 2020-2029 will use 2020 boundaries)
- Display census data on portal using these geographies
- Go through similar process as we look at different Nationally Consistent Data and Measures (NCDMs)

The Portal has Sub-County Maps for Populations and Vulnerabilities

Example - Texas (2015), number of people with at least high school education

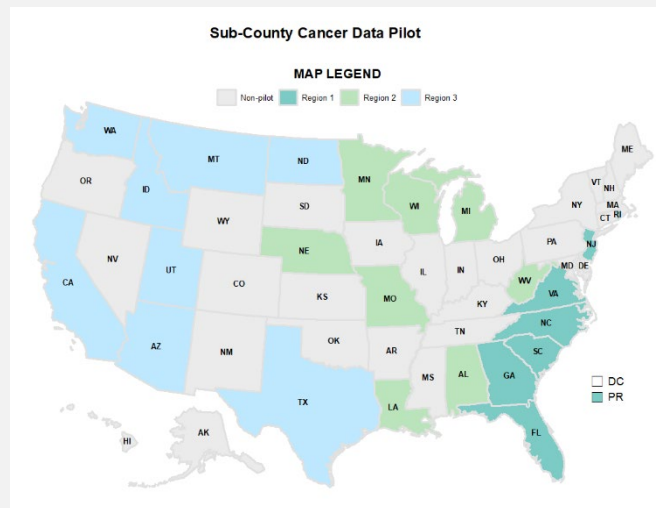
- Sub-county (5k)



SUB-COUNTY CANCER PILOT – PHASE 1

Sub-County Cancer Pilot Phase I

- June 2019 – February 2020
- 21 registries participated
- Participants provided summary tables and maps
- Tested spatiotemporal aggregation for cancer types at a sub-county level
- Discussed Nationally Consistent Data & Measures (NCDMs) recommendations to allow for multiple display options, including the overall spatiotemporal recommendation



Registry and Tracking Staff

Florida: Gary Levin, Chris DuClos, Jessi Joiner

Georgia: Rana Bayakly, Lyn Almon, Chrissy McNamara

New Jersey: Nan Stroup, Lisa Paddock, Barb Goun, Pam Agovino, Katharine McGreevy

North Carolina: Chandrika Rao, Christian Klaus, Dianne Enright

Puerto Rico: Diego Zavala

Rhode Island: Junhie Oh, Nancy LeBrun, Peter DiPippo, Jay Metzger, Lisa Garcia

South Carolina: Deborah Hurley, Susan Bolick, Stephanie Chiodini

Virginia: Leslie Hoglund, Sunney Wang, Taylor Guidry

Louisiana: Yong Yi, Lauren Maniscalco, Runa Bakshi, Kate Friedman

Michigan: Georgetta Alverson, Georgia Spivak, Jill Maras, Sydney Ogden, Thomas Largo, Jocelin Teachout, Amy Marquardt, Jeffrey Duncan

Minnesota: Paula Lindgren, Sally Bushhouse, Judy Punyko, Jessie Shmool, Blair Sevcik

Missouri: Jeanette Jackson-Thompson, Chester Schmaltz, Qiao Wang, Sherri Homan, Katie Long, Venkata Garikapaty, Elizabeth Semkiw, Jeff Wenzel, Scott Patterson, Jen Weaver, Hwei-Yiing Johnson

Nebraska: Lifeng Li, Christophe Irumva, Qianru Wu, Connie Ganz, Ming Qu, Julie Nielsen, Ge Lin, Janis Singleton, Han Liu

West Virginia: Steven Blankenship, Shawn Farley, Markie McCoy, Myra Fernatt

Wisconsin: Mary Foote, Jenna Staehler, Paul Cresswell, Jenny Camponeschi, Mireille Perzan

Arizona: Georgia Yee, Chris Newton, Matt Roach, Wes Korteum, Eric Thomas

California: Sandy Kwong, Vanessa Miguelino-Keasling, Eyobe Ako, Eric Roberts, Paul English, Alexa Wilkie

Idaho: Chris Johnson, Randi Rycroft, Bozena Morawski

Montana: Heather Zimmerman, Debbi Lemons

North Dakota: Lucy Zheng, S. Cristina Oancea

Texas: Paige Miller-Gianturco, Rebecca Sardell, Melanie Williams, Heidi Bojes, Emily Hall, Nusaybah Khan

Utah: Carol Sweeney, Kim Herget, Sam LeFevre, Greg Williams, Matt McCord

Washington: Johna Peterson, Mahesh Keitheri Cheteri, Buffi LaDue, Jennifer Sabel



■ Region 1

- Lung cancer
- Colorectal cancer
- Non-Hodgkin lymphoma
- Lung by strata (sex, age, ~~race, ethnicity~~)

■ Region 2

- Lung cancer
- Breast cancer (females only)
- Melanoma
- Lung by strata (sex, age, ~~race, ethnicity~~)

■ Region 3

- Lung cancer
- Prostate cancer (males only)
- Liver and IHB cancer
- Lung by strata (sex, age, ~~race, ethnicity~~)

Sub-county cancer pilot process

	Year 1	Year 2	Year 3	Year 4	Year 5	3-year period 1	3-year period 2	3-year period 3	5-year period
Census tract									
Median case count									
Census tracts with zero cases									
5k aggregation									
Median case count									
Geos with zero cases									
20k aggregation									
Median case count									
Geos with zero cases									

For each cancer, states discussed the number of census tracts included and the median case counts for each spatial aggregation.

Then the regional team decided on temporal aggregations to test for rate calculation table.

Table from step 1 to calculate median case counts across aggregations (temporal and spatial)

Sub-county cancer pilot process - Example

Table from step 1 to calculate median case counts across aggregations (temporal and spatial)
RHODE ISLAND – LUNG CANCER

Geo Level	No. Geos	Year 1	Year 2	Year 3	Year 4	Year 5	3-year	5-year
Census tract	241	4	3	4	4	4	11	18
5K	119	7	7	7	7	7	21	36
20K	30	30	31	32	32	33	88	156

Sub-county cancer pilot process - Example

Table from step 1 to calculate median case counts across aggregations (temporal and spatial)
MISSOURI – BREAST CANCER

Geo Level	No. Geos	Year 1	Year 2	Year 3	Year 4	Year 5	3-year	5-year	7-year	10-year
Census tract	1377	2	2	3	2	3	9	14	20	28
5K	696	5	5	5	5	6	18	29	40	56
20K	155	20	22	22	21	22	74	122	168	230

Sub-county cancer pilot process - Example

Table from step 1 to calculate median case counts across aggregations (temporal and spatial)
WASHINGTON – LIVER AND IHB CANCER

Geo Level	No. Geos	Year 1	Year 2	Year 3	Year 4	Year 5	3-year	5-year	7-year	10-year
Census tract	1442	0	0	0	0	0	1	2	3	4
5K	791	0	0	1	0	1	3	4	6	7
20K	189	2	2	3	2	3	12	19	25	32

Median Case Count Summary

Average median, minimum, and maximum case counts by aggregation level and cancer type

	Census Tract	5k Aggregation	20k Aggregation	
	Annual	5-year period	3-year period	5-year period
Lung Cancer				
Region 1	3 (0-19)	31 (1-116)	70 (16-238)	123 (29-322)
Region 2	3 (0-15)	30 (5-109)	64 (14-166)	97 (23-268)
Region 3	2 (0-19)	23 (0-119)	47 (8-164)	77 (14-268)
Colorectal Cancer	2 (0-12)	19 (2-77)	44 (11-144)	
non-Hodgkin Lymphoma	1 (0-7)	10 (1-37)		36 (12-94)
Melanoma of the Skin	1 (0-9)	16 (4-58)	32.5 (5-98)	
Female Breast	2 (0-15)	33 (4-106)	77 (24-202)	
Prostate Cancer	2 (0-22)	21 (2-104)	51 (13-157)	
Liver and IHB	1 (0-5)	5 (0-21)	9 (1-35)	

Threshold	Median lower 95% CI ^a	Median incidence ^a	Median upper 95% CI ^a	Min cases	Median cases	Max cases	Number of geos	Percent suppressed ^b	Percent unstable ^c
Census tract	x.x	x.x	x.x	x.x	x.x	x.x	xxx	x.x	x.x
Census tract w/ selected temporal aggregation	x.x	x							x
5k aggregation w/ selected temporal aggregation	x.x	x							x
20k aggregation w/selected temporal aggregation	x.x	x							x

For each cancer, states discussed the median case counts, the percent suppressed, and the percent unstable for each spatial aggregation.

Then the regional team came up with recommendations for display options.

^a 95% confidence intervals calculated using adjusted Gamma confidence intervals (edited SAS code from Owen Devine at end of document) and incidence (total new cases/total population) is calculated per 10,000 persons.

^b Use <16 cases for suppression flag. Each jurisdiction can also create an extra column to test their own suppression rule (if different).

^c RSE calculated using the following formula $(SE/rate)*100$, where $SE=\sqrt{(cases)/population}$ (SAS code at end of document). Flag any zero case geographic units as unstable so they are included in the percent unstable calculation.

Sub-county cancer pilot process - Example

Table from step 5 to calculate incidence (per 10,000) across aggregations (temporal and spatial)
RHODE ISLAND – LUNG CANCER

Geo Level	No. Years	Median LL	Median Rate	Median UL	Min cases	Median cases	Max cases	% sup-pressed	% un-stable
Census tract	1	2.07	8.47	19.53	0	4	13	92	100
Census tract	5	5.00	8.69	13.02	0	18	54	35	22
5K	3	5.60	8.95	13.14	0	21	74	29	11
5K	5	6.10	8.84	11.92	0	36	122	8	3
20K	3	7.54	9.38	11.23	26	97	145	0	0
20K	5	7.96	9.36	10.77	42	156	246	0	0

Sub-county cancer pilot process - Example

Table from step 5 to calculate incidence (per 10,000) across aggregations (temporal and spatial)
MISSOURI – BREAST CANCER

Geo Level	No. Years	Median LL	Median Rate	Median UL	Min cases	Median cases	Max cases	% sup-pressed	% un-stable
Census tract	1	2.44	13.45	35.35	0	3	15	89	99
Census tract	10	8.94	13.56	19.01	0	28	118	18	10
5K	5	9.58	14.23	19.80	0	29	97	10	3
5K	7	10.15	14.00	18.61	1	40	131	3	1
20K	3	12.08	15.38	19.19	21	74	198	0	0
20K	5	12.22	14.80	17.54	37	122	297	0	0

Missouri Cancer Registry (MCR) core activities are supported in part by a cooperative agreement between the Centers for Disease Control and Prevention (CDC) and the Missouri Department of Health and Senior Services (DHSS)(U58DP006299/02-03) and a Surveillance Contract between DHSS and the University of Missouri (MU)

Sub-county cancer pilot process - Example

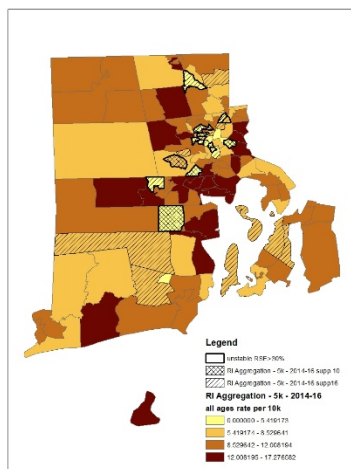
Table from step 5 to calculate incidence (per 10,000) across aggregations (temporal and spatial)
WASHINGTON – LIVER CANCER

Geo Level	No. Years	Median LL	Median Rate	Median UL	Min cases	Median cases	Max cases	% sup-pressed	% un-stable
Census tract	1	0.00	0.00	8.22	0	0	4	100	100
Census tract	10	0.25	0.92	2.08	0	4	35	95	97
5K	5	0.31	1.07	2.30	0	4	28	96	96
5K	7	0.36	1.02	2.00	0	6	41	95	88
20K	3	0.59	1.13	1.86	1	12	35	69	49
20K	5	0.65	1.11	1.66	0	19	57	36	20

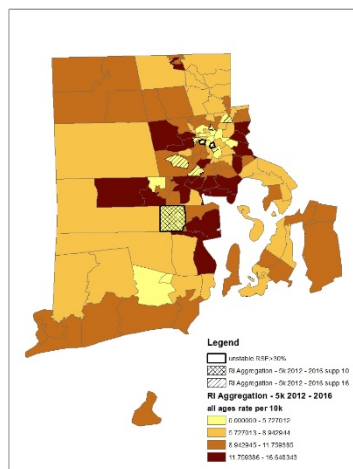
Sub-county cancer pilot process – Map Example

Rhode Island – Lung Cancer

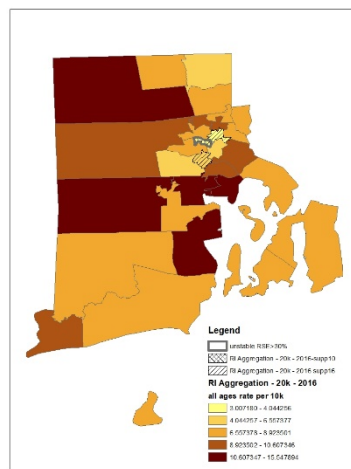
5K / 3 YR



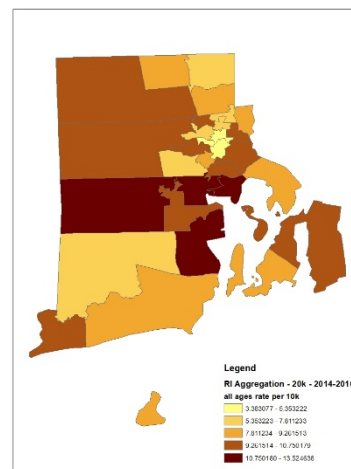
5K / 5 YR



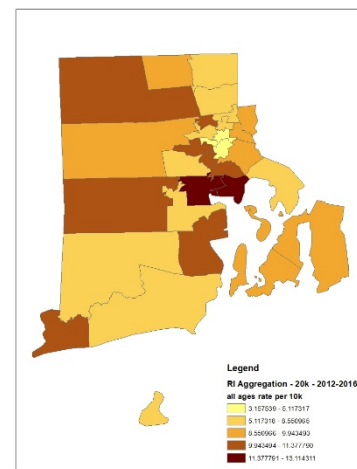
20K / 1 YR



20K / 3 YR



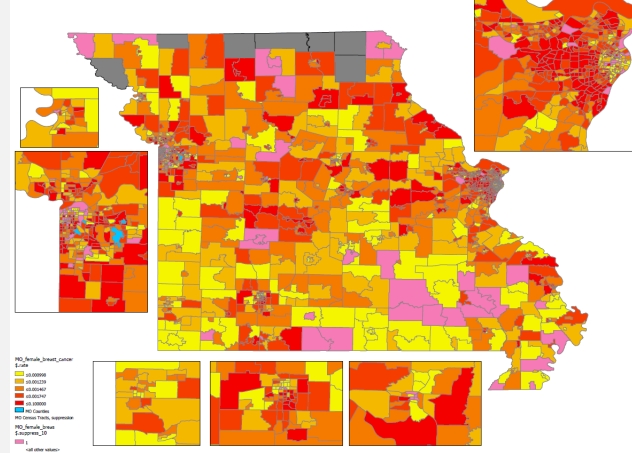
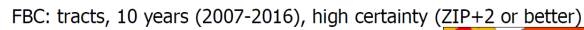
20K / 5 YR



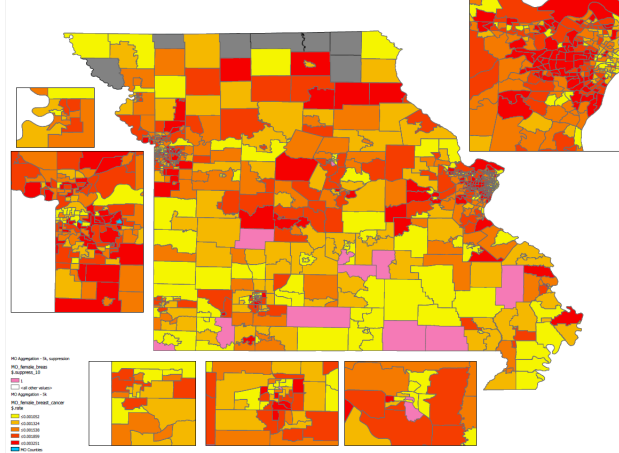
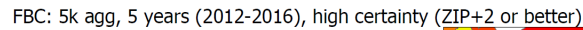
Sub-county cancer pilot process – Map Example

Missouri – Breast Cancer

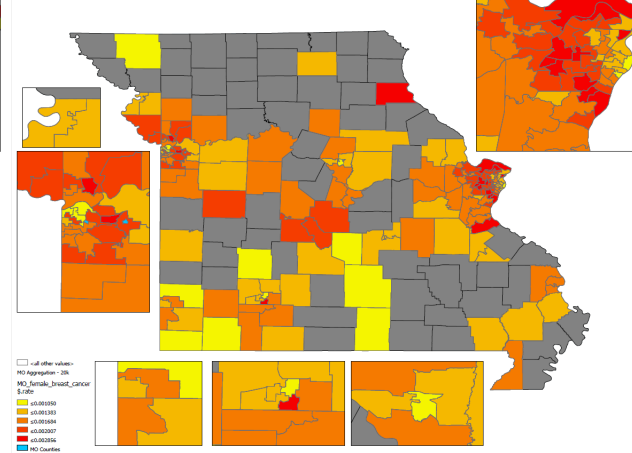
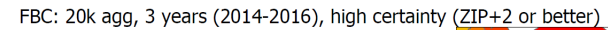
TRACT / 10 YR



5K / 5 YR



20K / 3 YR



Missouri Cancer Registry (MCR) core activities are supported in part by a cooperative agreement between the Centers for Disease Control and Prevention (CDC) and the Missouri Department of Health and Senior Services (DHSS)(U58DP006299/02-03) and a Surveillance Contract between DHSS and the University of Missouri (MU)

Cancer Type	Census Tract	5k aggregation	20k aggregation
Lung Cancer	Not feasible to display	5-year period	5-year period
Lung Cancer Female	Not feasible to display	5-year period	5-year period
Lung Cancer Male	Not feasible to display	5-year period	5-year period
Breast Cancer	10 -Year period	5-year period	3-year period
Prostate Cancer	10 -Year period	5-year period	3-year period
Colorectal Cancer	Not feasible to display	5-year period	3-year period
Melanoma	Not feasible to display	5-year period	3-year period
NHL	Not feasible to display	Not feasible to display	5-year period
Liver and IHB Cancer	Not feasible to display	Not feasible to display	Not feasible to display

Next Steps

- Obtain/develop annual census tract level denominators
- Visualize and test Tracking maps
- Release Tracking maps
- Evaluate variable to be included in NAACCR Data Dictionary
- Develop Census tract attribute file

Photo credit: By SA Mathieson, CC BY-SA 2.0,
<https://commons.wikimedia.org/w/index.php?curid=11287328>



SUB-COUNTY CANCER PILOT – PHASE 2

Sub-County Pilot Phase 2

- Evaluate additional cancer types
- Incorporate age-adjusted rates
- Evaluate alternate geos
(i.e., combine counties not meeting population thresholds, create third aggregation scheme with higher threshold)
- Develop messaging for portal
- Visualize and test Tracking maps
- Webinars



Photo credit: By Dave Croker, CC BY-SA 2.0,
<https://commons.wikimedia.org/w/index.php?curid=13230261>

Cancer Types on Tracking Network

- Acute myeloid leukemia
- Bladder cancer
- Brain and other nervous system cancer
- Breast cancer (females)
- Chronic lymphocytic leukemia
- Esophageal cancer
- Kidney cancer
- Larynx cancer
- Leukemia
- Liver cancer
- Lung and bronchus cancers
- Melanoma
- Mesothelioma
- Non-Hodgkin's lymphoma
- Oral cavity and pharynx cancer
- Pancreatic cancer
- Testicular cancer (males)
- Thyroid cancers
- Acute lymphocytic leukemia (childhood)
- Acute myeloid leukemia (childhood)
- Brain and central nervous system cancer (childhood)
- Leukemia (childhood)

Develop Messaging

- Participants expressed need for clear communication around cancer occurrence and environmental factors
- Will consider placement of messaging
- Will consider content of messaging
- Radon as example: <https://ibis.health.utah.gov/ibisph-view/indicator/view/Radon.HomeTest.html>
 - Includes definition of indicator and activities to reduce exposure

Visualize and Test Maps

- **New sub-county level maps will be developed and displayed on the Tracking Data Validation Portal**
 - **Secure portal**
 - **Opportunity to view maps, provide feedback, see what works/doesn't work**
- **Test display maps on test portal**
- **Release maps on public portal**

Phase 2 Timeline (Proposed)

- July – first call to discuss process; talk about messaging
- August – discuss Step 1 tables for selected cancer types; refine messaging
- September – discuss Step 5 tables and work on aggregations; refine messaging
- September – Webinar 1 – ALL NPCR/TRACKING WELCOME
- October – discuss maps for selected spatiotemporal aggregations; finalize messaging
- October – Midstream review webinar – ALL NPCR/TRACKING WELCOME
- November – discuss conclusions on multiple display options and overall recommendations
- December – test display maps on Tracking DVP
- December – Webinar 2 – ALL NPCR/TRACKING WELCOME
- January – Test display maps on test portal; wrap-up call
- February – Release maps on public portal
- February – Report-out webinar – ALL NPCR/TRACKING WELCOME

Tracking:

Angela K Werner, PhD
awerner@cdc.gov

Cancer Surveillance:

Jane Henley, MSPH
shenley@cdc.gov

Taylor Ellington, MPH

tellington@cdc.gov



For more information, contact NCEH

1-800-CDC-INFO (232-4636)

TTY: 1-888-232-6348 www.cdc.gov

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

