

NCI NAACCR Zone Design for Cancer Reporting Project and CDC Pilot Project to Visualize Cancer Data at the Sub-County Level

Welcome and Introduction

NAACCR 2020 Virtual Conference

Zaria Tatalovich, PhD
Surveillance Research Program, NCI

Logistics

- For technical issues please call: 217-698-0800 ext. 111.
- We will be taking questions at the very end
 - We ask that you type your question in the Q&A box as soon as you think of it
 - There will be a black bar at the bottom of your screen and you will see a box with Q&A - click on it then type in your question.

Presenters

Diane Ng, Westat

NCI/NAACCR Zone Design Project Overview, Challenges, Current Status, and Steps Forward

Debby Oh, Greater Bay Area Cancer Registry

CaliforniaHealthMaps.org: moving beyond county statistics to more meaningful geographies

Christopher Johnson, Cancer Data Registry of Idaho

The NCI/NAACCR Cancer Reporting Zone Project: Selecting the Preferred Zone Geography Alternative in Idaho

S. Jane Henley, Centers for Disease Control and Prevention
CDC Pilot Project to Visualize Cancer Data at Sub-County Level

Q&A Moderator: **David Stinchcomb**, Westat

Speaker bios: <https://www.naaccr2020.org/plenary-speakers>

NCI/NAACCR Zone Design for Cancer Reporting

Introduction

Background

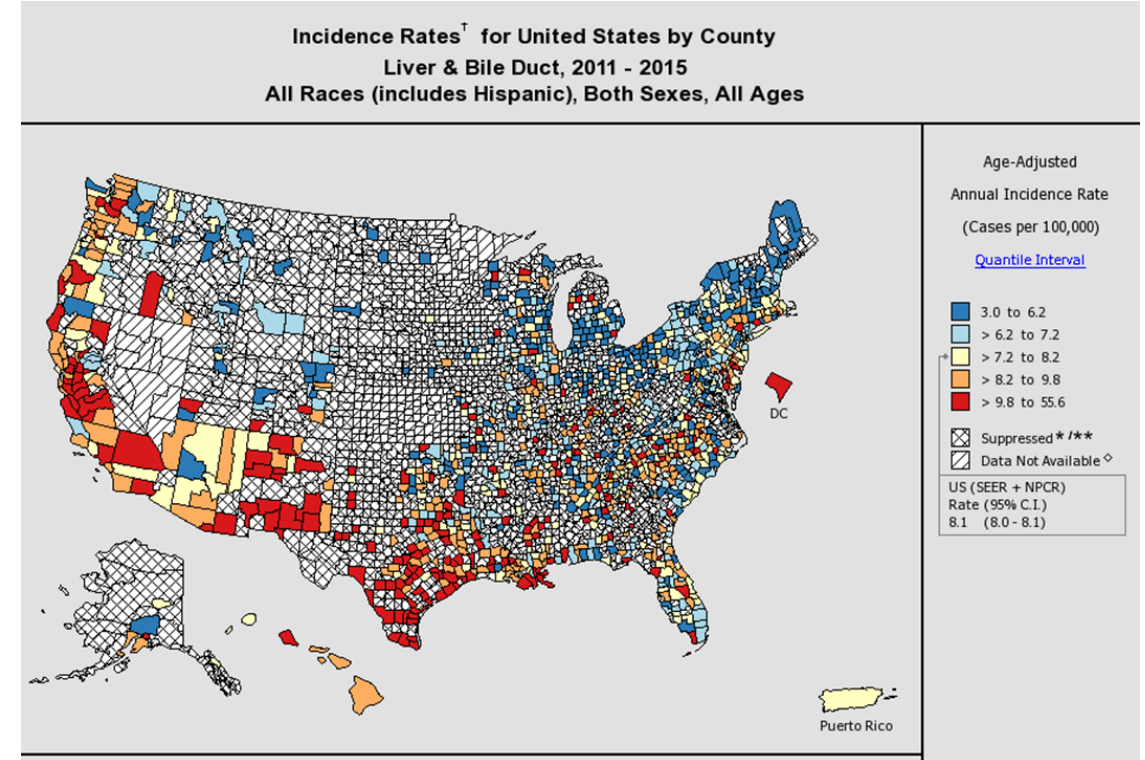
- The concept of zone design – optimizing geographic areas has a long history
 - Martin, D. (1998). Optimizing census geography: the separation of collection and output geographies. *International Journal of Geographical Information Science*, 12(7), 673–685.
 - Martin, D., Nolan, A., & Tranmer, M. (2001). The application of zone design methodology to the 2001 UK Census. *Environment and Planning A*, 33(11), 1949–1962.
- NCI funded projects that applied zone design in cancer control – (healthcare delivery, exposure assessment, neighborhood influences on cancer..)
- Couple of years ago our team decided to pursue the idea of applying the concept of zone design in cancer surveillance for the purpose of cancer reporting and analysis

Acknowledgements

- NCI
 - Mandi Yu
 - Denise Lewis
 - Li Zhu
 - Rocky Feuer
 - Westat
 - Dave Stinchcomb
 - Diane Ng
 - Matt Airola
 - NAACCR
 - Recinda Sherman
- Participating SEER and NPCR registries!!!

Motivation for NCI/NAACCR Zone Design Project

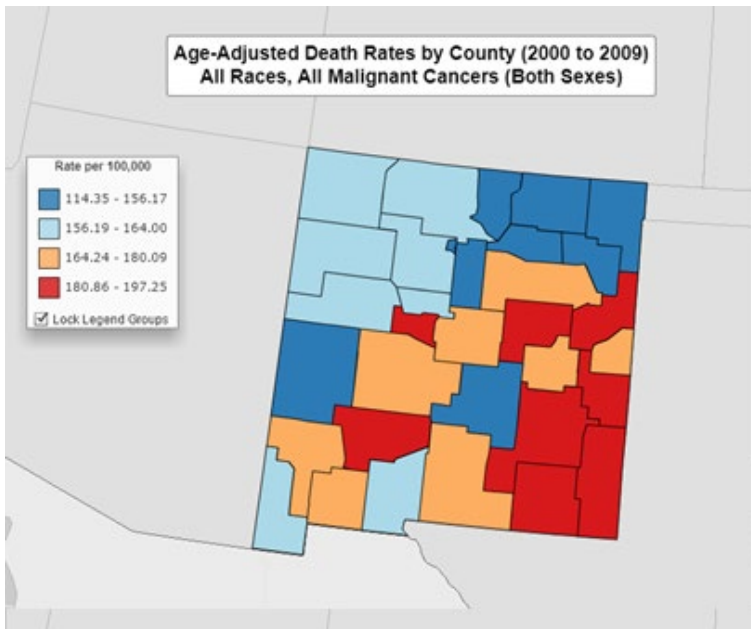
- Counties or Census Tracts are not very satisfactory geographic units for cancer reporting and analysis
- Issues:
 - Heterogeneity in population size – counties ~100 - 10 Mil
 - Unstable rates
 - Identifiability
 - Suppression



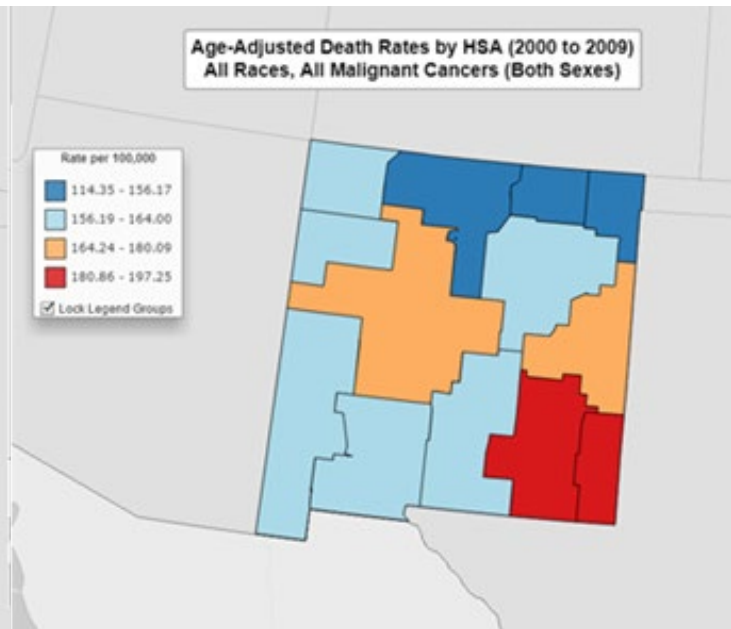
Dealing with Suppression

- From Counties to Health Service Areas (HSA), State Economic Areas (SEA)
- Aggregation of data over time
- Interpolation, extrapolation....

Counties



HSAs



- Variation in cancer rates lost
- Variation in population remains

Objectives

- Develop solution for cancer data release by geography that is more optimal than what we currently have
 - Relatively homogeneous populations
 - Large enough populations and case counts
 - Smaller proportion of areas with suppressed values
- Research data released with these zones should be easy to access (no special data use provisions)
- Cancer reporting zones should be meaningful to stakeholders

NCI / NAACCR Zone Design Project: Overview, Challenges, Current Status, and Steps Forward

Diane Ng and Dave Stinchcomb

Westat



Acknowledgements

› NCI

- Zaria Tatalovich
- Mandi Yu
- Denise Lewis
- Li Zhu
- Rocky Feuer

› NAACCR

- Recinda Sherman

› Participating SEER and NPCR registries

› Westat

- Dave Stinchcomb
- Matt Airola

Agenda

1. Background and project plan
2. Pilot study for process development
3. Methods and current status
4. Challenges and considerations
5. Steps forward

Agenda

1. Background and project plan
2. Pilot study for process development
3. Methods and current status
4. Challenges and considerations
5. Steps forward

Background

› County is not a very satisfactory geographic unit to use for cancer reporting

- Larger counties often have very heterogeneous populations
- Data for smaller counties often suppressed due to small numbers



Los Angeles County, CA
Pop: over 10 million



Loving County, TX
Pop: 134

› Census tracts (or collections of census tracts) are a much better unit for analysis, but are generally unavailable because of identifiability issues

- NCI and NAACCR have worked to make proxies for census tract available
 - Census tract poverty, SES, and urbanicity variables

Project plan

- Develop a set of cancer reporting zones that will:
 - Provide greater spatial resolution for large counties
 - Reduce suppression of data for small counties
 - Provide more meaningful data for communities & stakeholders
- Establish a common zone design method that can be applied to all states (with some flexibility)
- Work with individual states to apply the method

Agenda

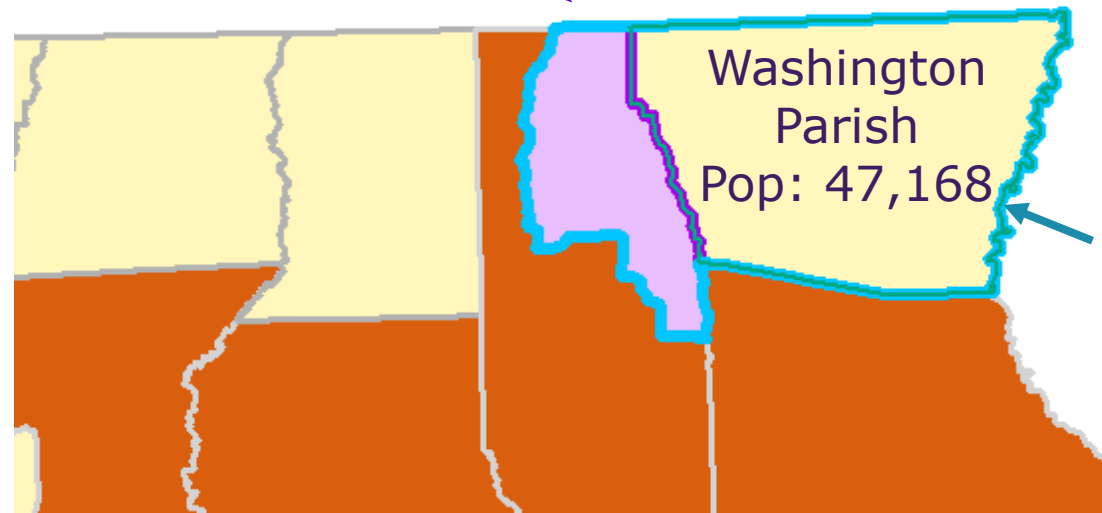
1. Background and project plan
- 2. Pilot study for process development**
3. Methods and current status
4. Challenges and considerations
5. Steps forward

Pilot study

- Initial activities
 - Tool evaluation
 - Initial zone construction tests
 - Reviewed results to determine an ideal target population for each zone
- Conducted pilot study with California and Louisiana
 - Addressed potential differencing issues
 - Differencing: a known problem in statistical disclosure control:
 - If tables are published for two sets of areas, users can compare the tables and produce new statistics for the areas formed by differencing, which may have populations below confidentiality thresholds. (Duke-Williams & Rees, 1998)

Differencing example – Washington Parish, Louisiana

Differencing Area
Pop: 10,143 (2 tracts)



Zone: all of Washington Parish
and part of Tangipahoa Parish
Pop: 57,311

Hypothetical* 5-year cancer incidence data:

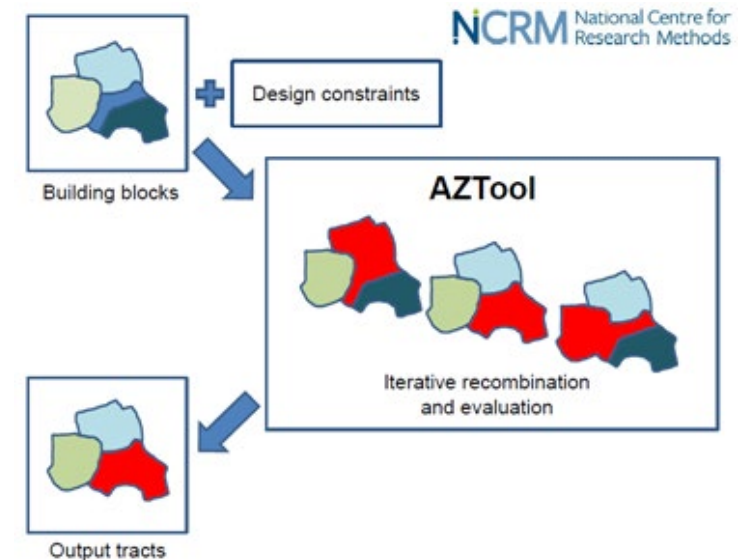
Area	Incidence Rate	Case Count	Population
Zone: Tangipahoa.Washington_1	69.8	20	57,311
Washington Parish	72.1	17	47,168
(differencing area)		3	10,143

* Populations are real but incidence rates and case counts are made up

Zone design process

➤ Based on findings, developed the following zone design process:

- Use AZTool to aggregate tracts and counties
- AZTool parameters:
 - Minimum and target population = 50,000
 - Homogeneity objectives
 - Urbanicity
 - Percent below poverty
 - Percent minority
 - Compactness objective
 - Even weights among objectives



The 2-step process

- Step 1: Aggregate census tracts in the large counties (populations over 100,000)
 - Zones cannot cross county boundaries
- Step 2: Aggregate:
 - The small and medium counties (populations less than 100,000)
 - With zones from Step 1 (with at least 50,000 people)

Agenda

1. Background and project plan
2. Pilot study for process development
- 3. Methods and current status**
4. Challenges and considerations
5. Steps forward

Methods and approach

- Invited registries to participate
 - SEER registries (September 2019)
 - All other U.S. registries (October 2019)
- Provided a questionnaire to obtain information on:
 - Geocoding practices
 - Existing geographic areas in the state/registry catchment area
 - Policies on release of cancer rates for geographic areas (state and sub-county reporting)
 - Cancer rate and count suppression rules

Methods and approach

- NCI and NAACCR reviewed the questionnaires to determine an order to begin work with each interested registry
- Considerations for ordering and selection:
 - Variability within a group of chosen registries
 - Size
 - Urbanicity
 - Readiness to participate

Current status

- 8 SEER registries responded to invitation to participate
 - 4 initially chosen to participate
- 13 additional registries responded to invitation through NAACCR to participate
 - 4 initially chosen to participate, but currently working with only 3

Current status

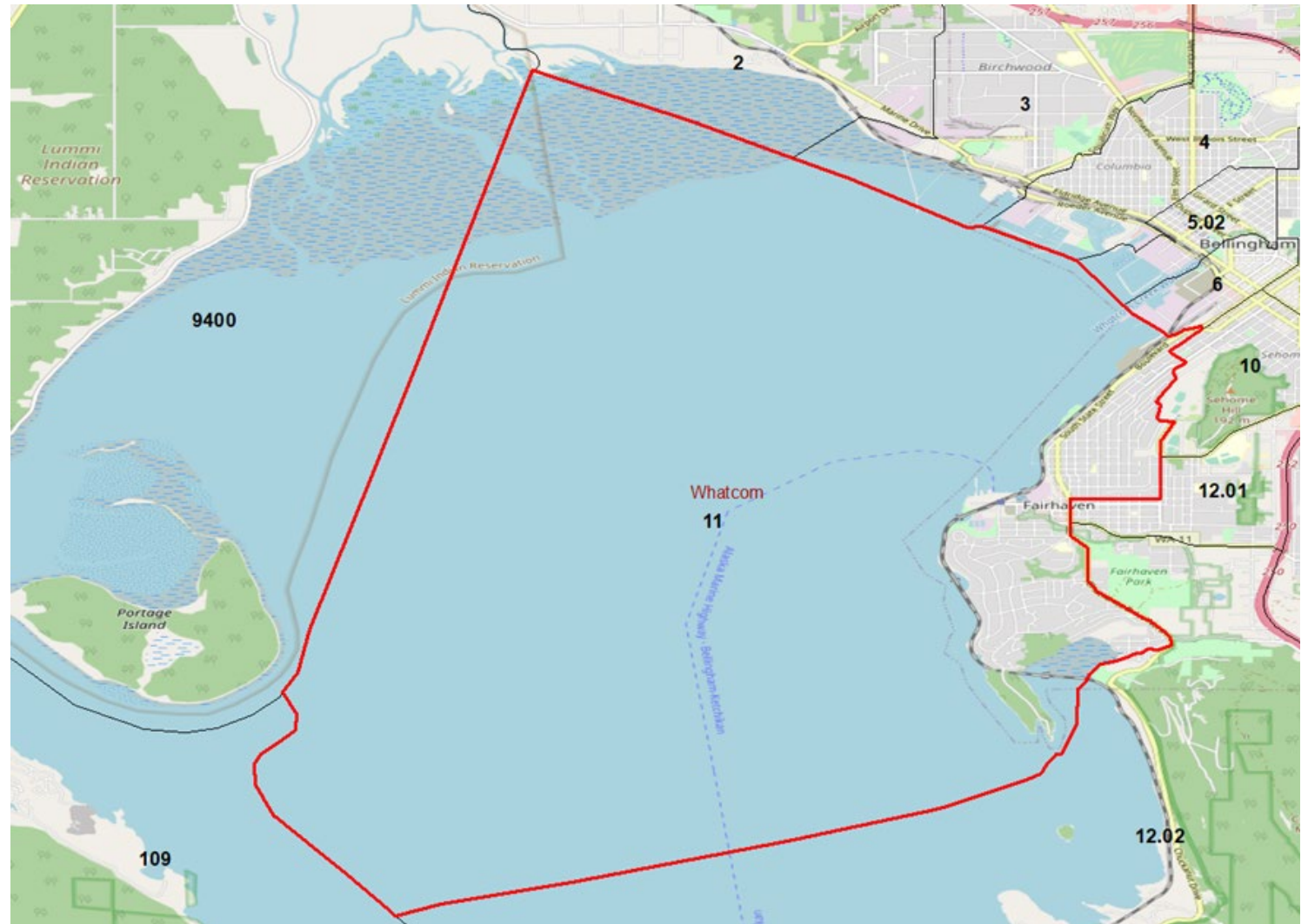
	SEER Registries	Non-SEER Registries
Total interested	8	11
Current participants	4	3
Step 1	3	1
Step 2	1	2

Agenda

1. Background and project plan
2. Pilot study for process development
3. Methods and current status
- 4. Challenges and considerations**
5. Steps forward

Challenges and considerations

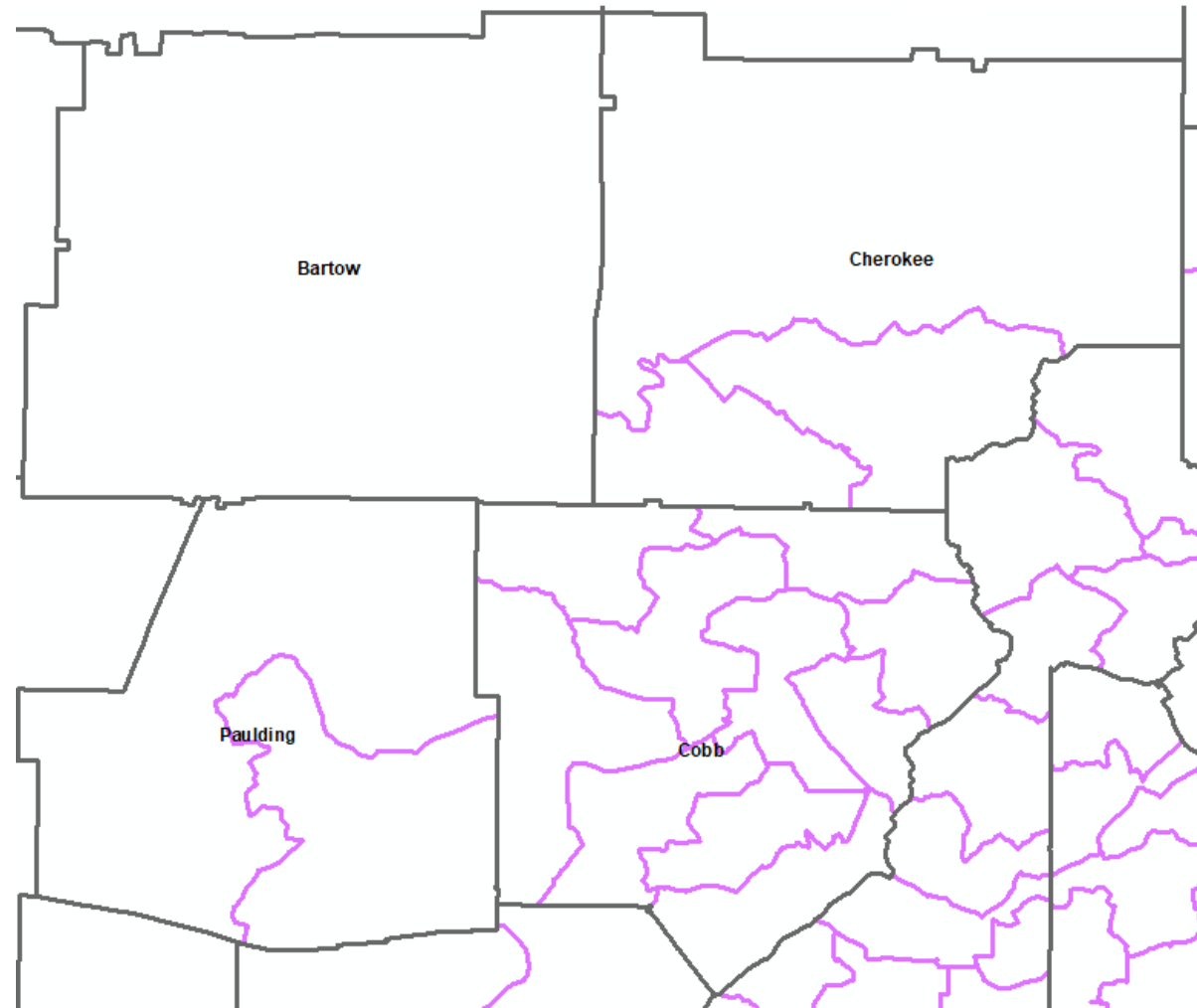
- Tract connectivity issues
 - Sometimes need to adjust adjacencies between tracts/counties for processing by either adding or deleting adjacencies



Example of tract where adjacencies needed to be deleted (Whatcom County, Seattle)

Challenges and considerations

- Not all large counties can be subdivided
 - Ex: Bartow County in Georgia
 - Population = 100,157
 - Could not combine census tracts to create zones with populations \geq 50,000



Some large counties in Georgia subdivided into zones

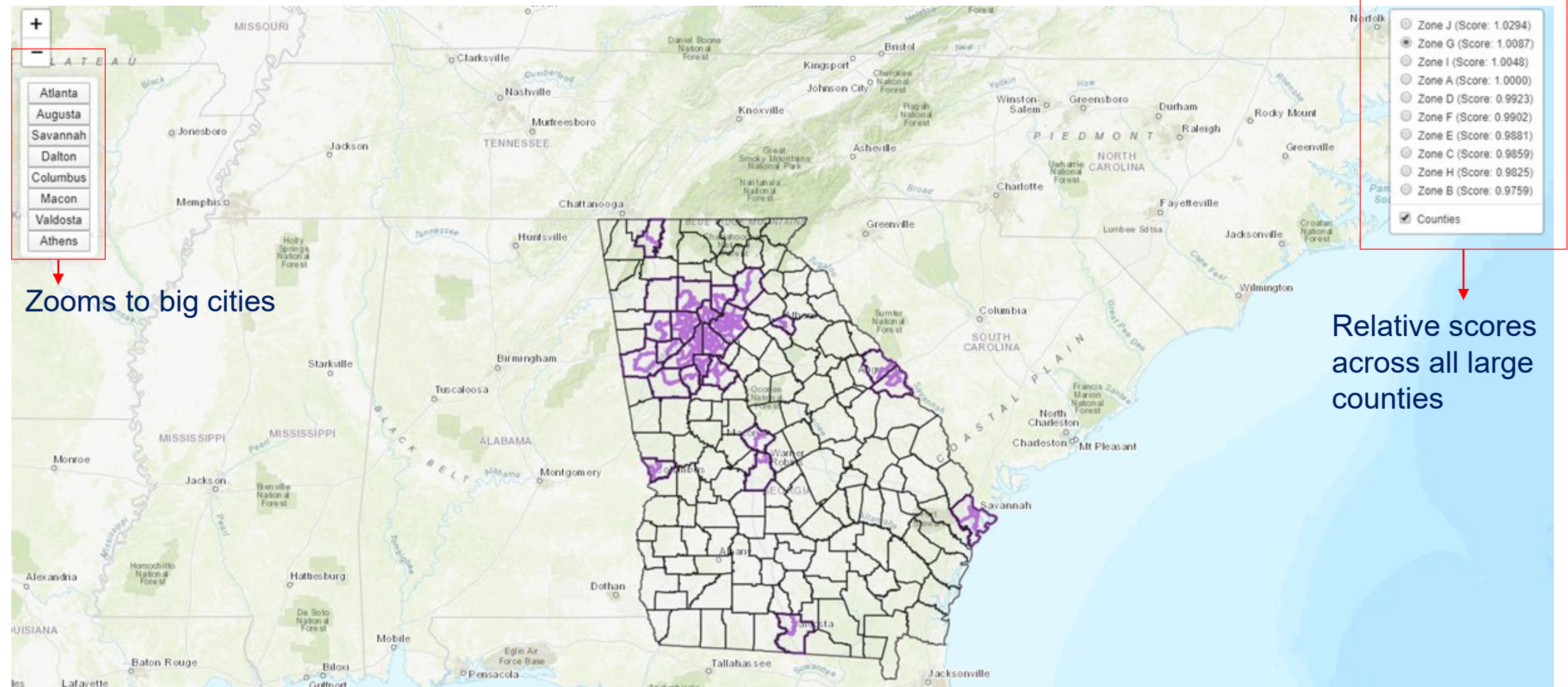
Challenges and considerations

➤ Additional modifications to the process

- Processing and review of zones based on geographic constraints (e.g. pre-defined public health districts, local health departments)
- Outputs and resources for review of zone alternatives
 - Statistics of each zone alternative by county
 - Web review tool

Web review tool

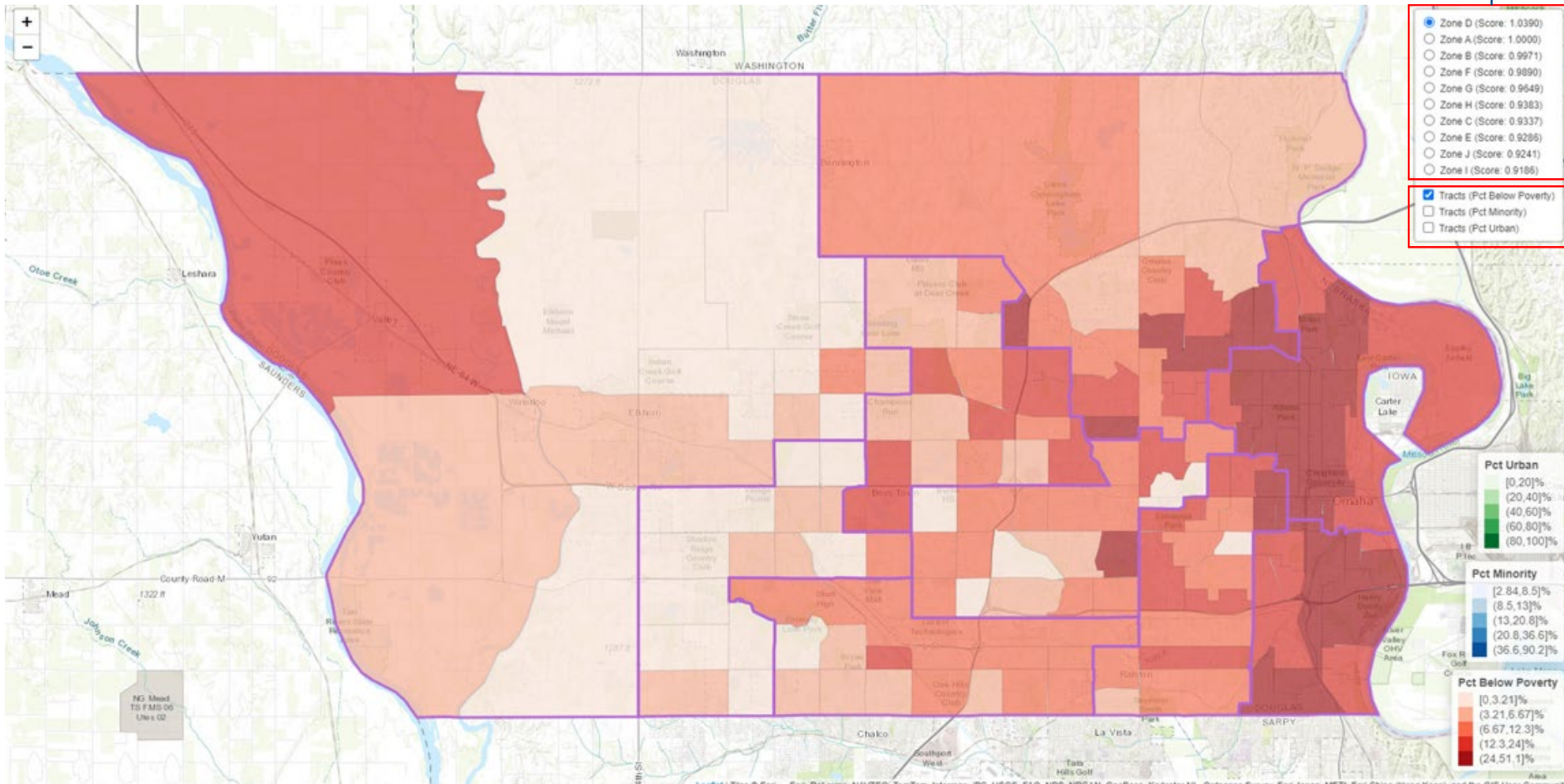
➤ Previous version



Web review tool

➤ Current version – individual tools by county

Relative scores for a specific county



Option for tract-level data layers

Agenda

1. Background and project plan
2. Pilot study for process development
3. Methods and current status
4. Challenges and considerations
- 5. Steps forward**

Steps forward

- Begin working with additional registries that have submitted questionnaires for participation
- Future invitations to other registries to participate in the project
 - If your registry is interested, contact us
- Zone-level reporting

Plans for zone-level reporting

› Websites with cancer rates by zone

- California example: CaliforniaHealthMaps.org
- Louisiana web tool in development:

Cancer Incidence ?

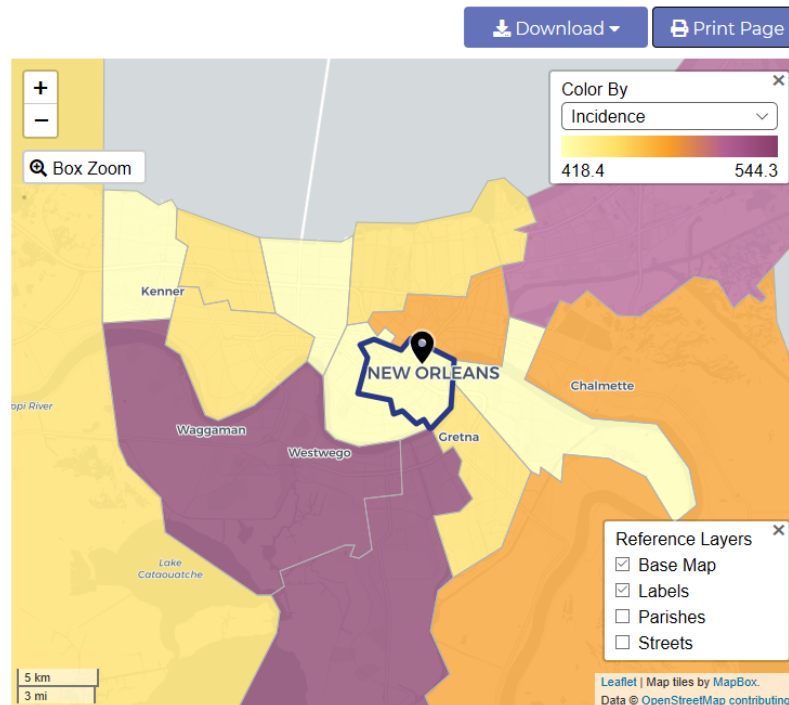
Location Search
New Orleans Q

Time Period
5-Year: 2011-2015 ▼

Cancer Site/Type
All Cancer Sites ▼

Sex ?
All Sexes ▼

Race/Ethnicity ?
All Ethnicities ▼



Orleans 5 (A9024) ×

Parishes: Orleans Parish
Places: New Orleans

5-Year: 2011-2015

All Cancer Sites

All Sexes

All Ethnicities

Cancer Statistics

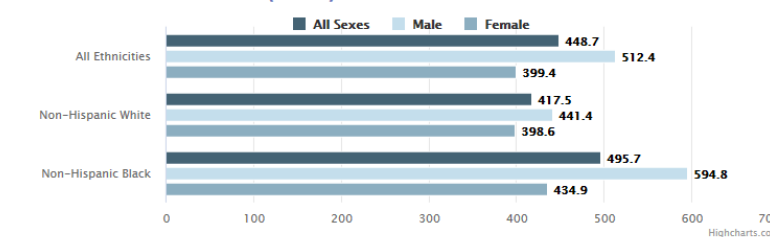
	Cases	Age Adjusted Incidence Rate (95% LCI, 95% UCI) ?
Zone	1,361	448.7 (424.1, 474.3)
Statewide	120,829	475.7 (473.0, 478.4)
Nationwide	8,136,881	454.5 (454.2, 454.8)

To ensure confidentiality and stable statistical rates, cancer incidence rates are not reported if based on fewer than 16 cancer cases. Distributions of selected population demographic factors are also provided for each zone. These data are from the decennial U.S. Census or the Census's American Community Survey.

Education

	Zone	Statewide	Nationwide
% With Bachelors Degree or Higher ?	40.0 %	22.6 %	30.3 %
% Did Not Finish High School ?	16.8 %	16.3 %	13.0 %

Age Adjusted Incidence Rate All Cancer Sites - Orleans 5 (A9024)



Plans for zone-level reporting

- › SEER*Stat database support
- › Report data by:
 - Site
 - Site and sex
 - Site, sex, and race/ethnicity
- › Range of reporting years can vary to meet suppression requirements
 - 1 year for common cancers
 - 5-10 years for less common cancers or more detailed breakdowns

Thank You

Contact us:

- Diane Ng: DianeNg@westat.com
- Zaria Tatalovich (NCI): tatalovichzp@mail.nih.gov

