500 Cities
Local Data for Better Health

James B. Holt, MPA, PhD

December 7, 2016

Outline

- Project overview
- Cities
- Measures
- Methods
- 500 Cities Open Data
- Examples of 500 Cities maps
- Expected benefits for public health
- Questions

Project Overview

- Robert Wood Johnson Foundation and CDC Foundation
- In partnership with the Centers for Disease Control and Prevention
- Project launched in October 2015
- First-of-its kind data analysis for the 500 largest American cities, and the census tracts within these cities, to estimate and report data for a select number of chronic disease measures, using data from the Behavioral Risk Factor Surveillance System
- Data estimates will be made available through:
  - CDC Chronic Data Portal (Socrata Open Data) – December 2016
  - Interactive website that will allow users to view, explore and download city and tract-level data – scheduled launch in early 2017

Purpose and Need

❑ **Purpose:**

- Provide high quality, small-area estimates for behavioral risk factors that influence health status, health outcomes, and use of clinical preventive services
- Identify emerging health problems
- Inform development and implementation of effective and targeted public health prevention activities *in America's cities*
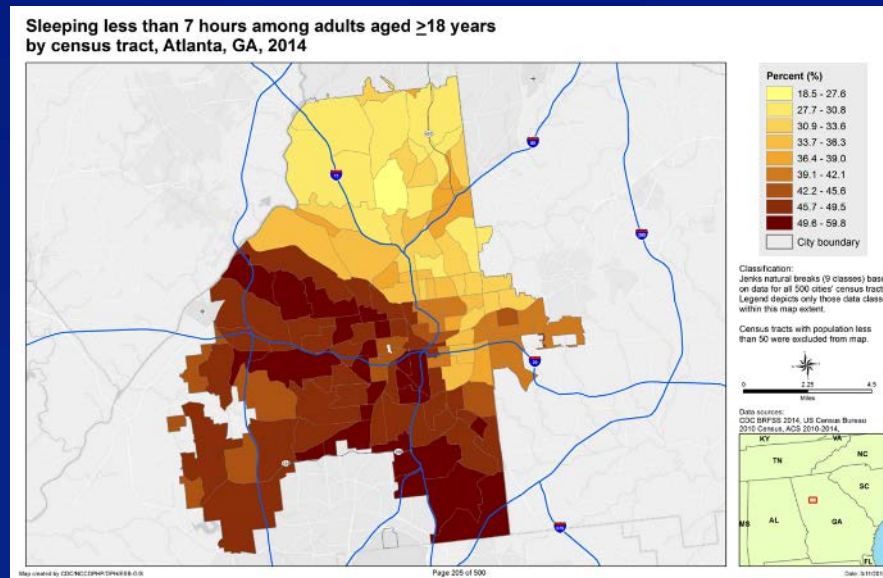
❑ **Need:**

- Limited data currently available at county and metropolitan levels
- No data available on a large scale for small areas within counties or cities
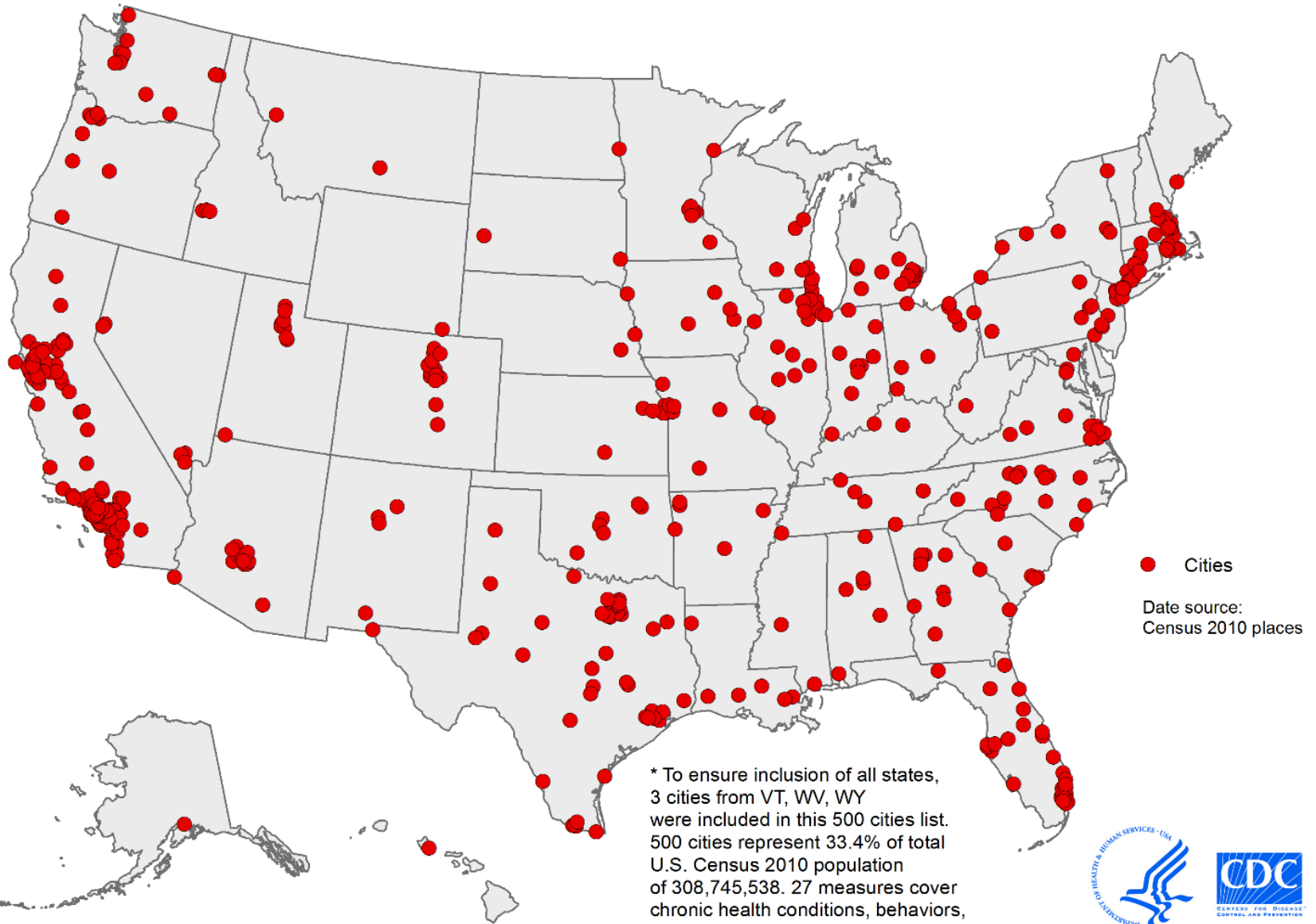
Our Vision

To go from this:
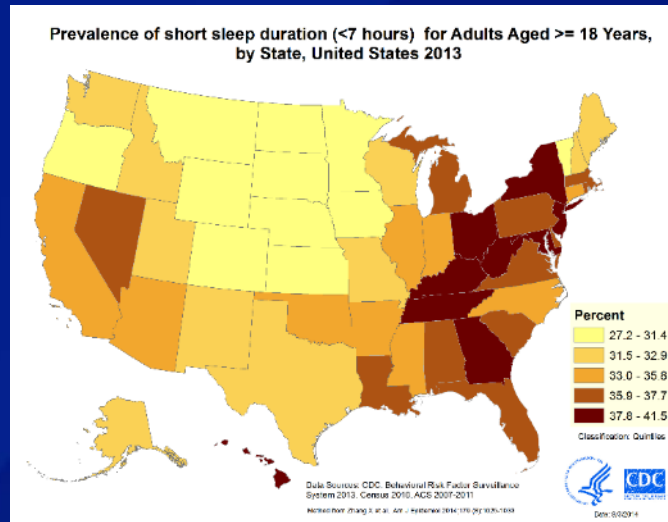
To also
have
this:

How do we define "City"?

- US Census Bureau, 2010 Population Counts
- Incorporated Places
    - Legally established to provide governmental functions
    - Always within a state, but may cross county boundaries
    - Includes: cities, towns, villages, boroughs
    - Excludes: boroughs in AK (treated as counties); towns in New England states, New York, and Wisconsin (treated as Minor Civil Divisions); boroughs in New York (treated as Minor Civil Divisions)
- Exception – Honolulu
    - No Incorporated Places in Hawaii
    - Honolulu is a major urban center
    - Consultations with Hawaii Department of Health
    - City and County of Honolulu

Largest 500 US Cities*, by 2010 Population

Cities

Date source:
Census 2010 places

* To ensure inclusion of all states,
3 cities from VT, WV, WY
were included in this 500 cities list.
500 cities represent 33.4% of total
U.S. Census 2010 population
of 308,745,538. 27 measures cover
chronic health conditions, behaviors,
and risk factors.

DEPARTMENT OF HEALTH & HUMAN SERVICES · USA

CDC
CENTERS FOR DISEASE
CONTROL AND PREVENTION

Date: 2/24/2016

Why Census Tracts?
Census Geography Hierarchy

Standard Hierarchy of Census Geographic Entities
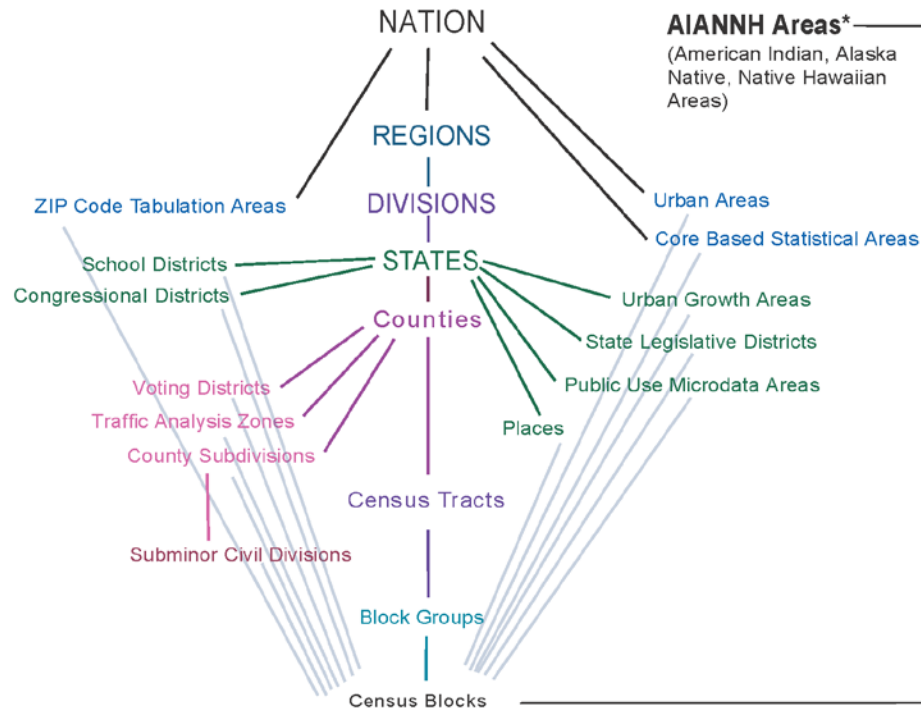
Facts about the 500 Cities

- ❑ **Each state represented**
  - ▪ From 1 per state to 121 per state (CA)
- ❑ **Population range (1 thru 497):**
  - ▪ 66,135 (Schenectady, NY) to 8,175,133 (New York City, NY)
  - ▪ 3 additional cities
    - • Cheyenne, WY (59,466)
    - • Charleston, WV (51,400)
    - • Burlington, VT (42,417)
- ❑ **Total population = 103,745,538 (<ins>33.4% of total US</ins>)**
- ❑ **Number of census tracts ~ 28,000**
  - ▪ From 8 per city to 2,140 per city
  - ▪ From less than 1 square mile to over 642 square miles in size
  - ▪ Population ranges from less than 50 to 28,960 per tract

Measure Selection Criteria

- Measures are amenable to public health intervention
- Reflect public health priorities to address leading causes of morbidity and mortality
- Preventive services are consistent with US Preventive Services Task Force recommendations
- Exhibit substantial, meaningful variation at the city and census tract level
- Can be estimated for small area levels from existing, regularly-collected surveillance data – Behavioral Risk Factor Surveillance Survey (BRFSS)

27 Chronic Disease Measures

- <u>Fills a niche</u> for health data at the city and census tract level, which are not presently available, while not duplicating health-related data that are available elsewhere
- Compliments similar state-level measures that are available elsewhere
- 5 measures related to unhealthy behaviors
- 13 measures related to health outcomes
- 9 measures related to prevention

Unhealthy Behavior Measures

- ❑ Binge drinking among adults aged ≥18 years
- ❑ Current smoking among adults aged ≥18 years
- ❑ No leisure-time physical activity among adults aged ≥18 years
- ❑ Obesity among adults aged ≥18 years
- ❑ Sleeping less than 7 hours among adults aged ≥18 years

Health Outcome Measures

- Arthritis among adults aged ≥18 years
- Current asthma among adults aged ≥18 years
- High blood pressure among adults aged ≥18 years
- Cancer (excluding skin cancer) among adults aged ≥18 years
- High cholesterol among adults aged ≥18 years who have been screened in the past 5 years
- Chronic kidney disease among adults aged ≥18 years
- Chronic obstructive pulmonary disease among adults aged ≥18 years
- Coronary heart disease among adults aged ≥18 years
- Diagnosed diabetes among adults aged ≥18 years
- Mental health not good for ≥14 days among adults aged ≥18 years
- Physical health not good for ≥14 days among adults aged ≥18 years
- All teeth lost among adults aged ≥65 years
- Stroke among adults aged ≥18 years

Prevention Measures

❑ Current lack of health insurance among adults aged 18–64 years

❑ Visits to doctor for routine checkup within the past year among adults aged ≥18 years

❑ Visits to dentist or dental clinic among adults aged ≥18 years

❑ Taking medicine for high blood pressure control among adults aged ≥18 years with high blood pressure

❑ Cholesterol screening among adults aged ≥18 years

❑ Mammography use among women aged 50–74 years

❑ Papanicolaou smear use among adult women aged 21–65 years

❑ Fecal occult blood test, sigmoidoscopy, or colonoscopy among adults aged 50–75 years

❑ Older adults aged ≥65 years who are up to date on a core set of clinical preventive services

- Men: Flu shot past year, PPV shot ever, Colorectal cancer screening
- Women: Same as above, and Mammogram past 2 years

Measure Definitions

- Each measure has a detailed data definition, along with background statistics, description of indicator significance, limitations, and related Healthy People 2020 objectives

- Most definitions are direct replications from recent Chronic Disease Indicators (CDI) updates (2015)
  - Developed in partnership with the Council of State and Territorial Epidemiologists and the National Association of Chronic Disease Directors
  - Subjected to rigorous scientific review and clearance
  - Published in 2015 MMWR R&R and CDI website (http://www.cdc.gov/cdi)

Holt JB, Huston SL, Heidari K, Schwartz R, Gollmar C, Tran A, Bryan L, Liu Y, Croft JB, 2015. Indicators for Chronic Disease Surveillance. MMWR Recomm Rep 2015;64(No. RR-1):1-250.

Measure Definitions - Example

Binge drinking among adults aged ≥18 years

Demographic group	Adults aged ≥18 years.
Numerator	Adults aged ≥18 years who report having five or more drinks (men) or four or more drinks (women) on an occasion in the past 30 days.
Denominator	Adults aged ≥18 years who report having a specific number, including zero, of drinks on an occasion in the past 30 days (excluding those who refused to answer, had a missing answer, or answered "don't know/not sure").
Measures of frequency	Annual prevalence: crude and age adjusted (standardized by the direct method to the year 2000 standard U.S. population, distribution 9 [*1*]) with 95% confidence intervals and by demographic characteristics when feasible.
Time period of case definition	Past 30 days.
Background	In 2010, a total of 17.1% of adults reported binge drinking on an occasion in the past 30 days (*2*). Binge drinking prevalence is higher among men, persons aged 18–34 years, whites, and those with household incomes ≥$75,000 (*2*).
Significance	Excessive alcohol use accounted for an estimated average of 88,000 deaths and 2.5 million years of potential life lost (YPLL) in the United States each year during 2006–2010 (*3*), and an estimated $223.5 billion in economic costs in 2006 (*4*). Binge drinking accounted for more than half of those deaths, two thirds of the YPLL (*5*), and three fourths of the economic costs (*4*). Binge drinking also is a risk factor for many health and social problems, including motor-vehicle crashes, violence, suicide, hypertension, acute myocardial infarction, sexually transmitted diseases, unintended pregnancy, fetal alcohol spectrum disorders, and sudden infant death syndrome (*6,7*). In the United States, binge drinking accounts for more than half of the alcohol consumed by adults (*8*). However, most binge drinkers are not alcohol dependent (*9,10*).
Limitations of indicator	The indicator does not convey the frequency of binge drinking or the specific amount of alcohol consumed.
Data resources	Behavioral Risk Factor Surveillance System (BRFSS).
Limitations of data resources	As with all self-reported sample surveys, BRFSS data might be subject to systematic error resulting from noncoverage, nonresponse, or measurement bias. In an effort to address noncoverage issues related to phone use, BRFSS began including cell phone interviews in the 2011 data collection. Due to changes in sampling and weighting methodology, 2011 is a new baseline for BRFSS, and comparisons with prior year data are inappropriate. A recent study using BRFSS data found that self-reports identify only 22%–32% of presumed alcohol consumption in states, based on alcohol sales (*11*).
Related recommendations	• *Healthy People 2020* objective SA-14.3: Reduce the proportion of persons engaging in binge drinking in the past 30 days—adults aged ≥18 years. • CDC Prevention Status Report: Excessive alcohol use (*12*).

1. Klein RJ, Schoenborn CA. Age adjustment using the 2000 projected U.S. population. Healthy people 2010 statistical notes, no. 20. Hyattsville, MD: US Department of Health and Human Services, CDC, National Center for Health Statistics; 2001. Available at http://www.cdc.gov/nchs/data/statnt/statnt20.pdf.
2. CDC. Vital signs: binge drinking prevalence, frequency, and intensity among adults—United States, 2010. MMWR 2012;61:14–9.
3. CDC. Alcohol-related disease impact (ARDI) application. Atlanta, GA: US Department of Health and Human Services, CDC; 2013. Available at http://apps.nccd.cdc.gov/DACH_ARDI/Default.aspx.
4. Bouchery EE, Harwood HJ, Sacks JJ, Simon CJ, Brewer RD. Economic costs of excessive alcohol consumption in the U.S., 2006. Am J Prev Med 2011;41:516–24.
5. Stahre M, Roeber J, Kanny D, Brewer RD, Zhang X. Contribution of excessive alcohol consumption to deaths and years of potential life lost in the United States. Prev Chronic Dis 2014;11:130293.
6. National Institute of Alcohol Abuse and Alcoholism. Tenth special report to the U.S. Congress on alcohol and health. Bethesda, MD: US Department of Health and Human Services, National Institutes of Health; 2000.
7. Warren KR, Hewitt BG, Thomas JD. Fetal alcohol spectrum disorders. Alcohol Res Health 34;2011:4–14.
8. Office of Juvenile Justice and Delinquency Prevention. Drinking in America: myths, realities, and prevention policy. Washington, DC: US Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention; 2005. Available at http://www.udetc.org/documents/Drinking_in_America.pdf.
9. Dawson DA, Grant BF. LI T-K. Quantifying the risks associated with exceeding recommended drinking limits. Alcohol Clin Exp Res 2005;29:902–8.
10. Woerle S, Roeber J, Landen MG. Prevalence of alcohol dependence among excessive drinkers in New Mexico. Alcohol Clin Exp Res 2007;31:293–8.
11. Nelson DE, Naimi TS, Brewer RD, Roeber J. U.S. state alcohol sales compared to survey data, 1993–2006. Addiction 2010;105:1589–96.

Methodology: Small Area Estimation

- ❏ Problem: insufficient (or no) sample sizes to create direct survey estimates at the city level
- ❏ Small area estimation (SAE) enables the prediction of prevalence for "small" areas (geographically or statistically) where there are small or no samples
- ❏ Numerous approaches to SAE
- ❏ Our models predict the statistically expected risk of health behaviors or conditions for:
  - ▪ 208 demographic groups (defined by age, gender, race/ethnicity)
  - ▪ Further adjusted by where they live (state, county, neighborhood)
- ❏ Acknowledges that health is the product of a person's gene code and behavior and their ZIP Code

Small Area Estimation Method: Multilevel Regression and Post-Stratification (MRP)

❑ **Construct a multilevel regression model to predict health outcomes using individual characteristics and spatial contexts:**

- Individual-level fixed effects (age, sex, race/ethnicity, education)
- Area-level fixed effects (county-level poverty)
- Area-level random effects (county and state)

Zhang X, Holt JB, Yun S, Lu H, Greenlund KJ, Croft JB, 2015. Validation of multilevel regression and poststratification methodology for small area estimation of health indicators from the Behavioral Risk Factor Surveillance System (BRFSS). American Journal of Epidemiology 182(2):127-137.

Zhang X, Holt JB, Lu H, Wheaton AG, Ford ES, Greenlund K, Croft JB, 2014. Multilevel regression and poststratification for small area estimation of population health outcomes: a case study of chronic obstructive pulmonary disease prevalence using BRFSS. American Journal of Epidemiology 179:1025-1033.

MRP (cont.)

- Apply the parameters from the multilevel models to the census population to obtain probability of health risk or outcome at the individual level

- Apply this probability to any target population to compute the estimated prevalence – in 500 Cities this is the census block level

- Block-level estimates are aggregated to produce census tract estimates and city estimates

MRP Validation

- Validation studies confirm the strong consistency between MRP model-based SAEs and direct survey estimates at state, county, and sub-county levels

  - Validated internally against BRFSS direct estimates
  - Validated externally against MO County-level Study direct estimates
  - Validated externally against American Community Survey and against CMS Medicare claims data
  - External validation currently underway at sub-county level with Boston city BRFSS and Southeast PA Household Survey. and (soon) Mecklenburg County (NC), Hennepin County (MN), and DataHaven (CT)

Advantages of Using MRP for Creating Small Area Estimates

- Reliable estimates for areas with small (or no) samples
- High precision
- Flexible: combines individual- and area-specific information relevant to small area estimation of outcomes of interest
- Modeling nationally and predicting locally.
- Can be used for any geography needed.