HHS Public Access

Author manuscript

J Registry Manag. Author manuscript; available in PMC 2022 April 04.

Published in final edited form as:

J Registry Manag. 2020 ; 47(2): 67–79.

Visualizing Cancer Incidence and Mortality Estimates by Congressional Districts, United States 2012–2016

Virginia Senkomago, PhD, MPH[a], **Trevor D. Thompson, BS**[a], **Lia C. Scott, PhD, MPH**[a], **Simple D. Singh, MD**[a], **Mary Elizabeth O'Neil, MPH**[a], **Reda Wilson, MPH**[a], **Jessica B. King, MPH**[a], **Melissa M. Jim, MPH**[a], **Hua Lu, MS**[b], **Manxia Wu, MD, MPH**[a], **Vicki B. Benard, PhD**[a], **Lisa C. Richardson, MD**[a]

[a]Division of Cancer Prevention and Control, US Centers for Disease Control and Prevention (CDC).

[b]Division of Population Health, US Centers for Disease Control and Prevention (CDC).

Abstract

Background: Cancer incidence and death rates in the United States are often published at the county or state-levels; examining cancer statistics at the congressional district (CD) level allows decision makers to better understand how cancer is impacting the specific populations they represent.

Methods: Cancer incidence data were obtained from the Centers for Disease Control and Prevention's National Program of Cancer Registries and the National Cancer Institute's Surveillance, Epidemiology, and End Results Program. Mortality data were obtained from the National Center for Health Statistics. CD rates were estimated by assigning the county-level age-adjusted rates to the census block and weighting those by the block population proportion of the CD. Those weighted rates were then aggregated over the blocks within the CD to estimate the district rate. Incidence rate estimates for 406 CDs and death rate estimates for 436 CDs were reported according to the boundaries for the 115th Congress of the United States. Maps showing rate estimates for all cancers combined, lung/bronchus, colorectal, female breast, cervical, and prostate cancer are presented by sex and race/ethnicity.

Results: The distribution of cancer incidence and death rates by CDs show similar patterns to those that have been observed at the county and state levels, with the highest cancer incidence and death rates observed in CDs in the South and Eastern regions.

Conclusion: This examination of cancer rates at the CD-level provides data that can be used to inform cancer control strategies at the local and national levels. Displaying the data with the Data Visualizations tool makes it easily accessible to the public and decision makers.

Keywords

cancer; cancer incidence; cancer mortality; congressional district; small-area estimates

Address correspondence to Dr. Virginia Senkomago, Division of Cancer Prevention and Control, CDC, 4770 Buford Hwy NE, S107-4, Atlanta, GA, 30341. Telephone: (404) 718-6369. vsenkomago@cdc.gov.

Introduction

The United States Cancer Statistics (USCS) provides statistics on cancer, including information on cancer incidence from national registries and mortality from the National Vital Statistics System, for the entire US population. These statistics include cancer incidence data from Centers for Disease Control and Prevention (CDC)'s National Program of Cancer Registries (NPCR) and the National Cancer Institute's (NCI) Surveillance, Epidemiology, and End Results (SEER) Program, as well as mortality data from CDC's National Center for Health Statistics (NCHS). These data are essential for public health professionals, researchers, the medical community, and decision makers to monitor cancer trends, support critical research, identify and evaluate programmatic efforts, and allocate resources where they are most needed. The data can be used to inform decision makers about the effectiveness of interventions to reduce or prevent cancer occurrence. The numbers of new cancer cases and deaths in the United States are often published at the county or state levels.[1,2] Examining cancer statistics at the congressional district (CD) level provides decision makers with additional data to better understand how cancer is impacting the specific populations they represent.

Two previous studies have reported on cancer death rates by CD; however, these studies reported death rates from older periods (1990–2001 and 2002–2011), and did not include cancer incidence data.[3,4] In this paper, we present estimates of cancer incidence and mortality rates by CD for the 5-year period of 2012–2016. We also describe the Data Visualizations tool that houses the official US federal cancer data, US Cancer Statistics, including the estimates for cancer incidence and mortality by CDs.

Methods

We obtained cancer incidence data from US Cancer Statistics that includes both data from CDC's NPCR and the NCI's SEER Program.[5] Mortality data were obtained from the NCHS. Block-level population data came from the US Census Bureau's 2010 Census Summary File 1.[6]

There are 436 CDs for 50 states and the District of Columbia in the United States. We estimated cancer death rates and counts for 2012–2016 for the 436 CDs according to the boundaries for the 115th Congress of the United States. Similarly, cancer incidence rates and counts were estimated for 424 CDs. Incidence data are not included for Kansas (4 CDs) and Minnesota (8 CDs) because county-level incidence data were not reported to CDC. Illinois opted not to present CD-specific estimated incidence case counts and rates (18 CDs). Therefore, we reported estimated incidence rates and counts for 406 CDs in this paper and in the online Data Visualizations tool.

CDs were defined by census blocks according to census geography hierarchy.[7] The District of Columbia and 7 states (Alaska, Delaware, Montana, North Dakota, South Dakota, Vermont, and Wyoming) have a single CD that follows state or federal district boundaries. For these states that have only 1 CD, the same state rates and counts were used. For states that have multiple CDs, rates were estimated by assigning the county-level age-adjusted

rates to the census block and weighting those by the block population proportion of the CD. Those weighted rates were then aggregated over the blocks within the CD to estimate the district rate. More specifically, the following steps were taken:

1. We obtained census block-level population counts by sex and race/ethnicity from the 2010 US Census Summary File 1.[6] We matched each census block to the assigned congressional district using the census Block Equivalency File downloaded from the census website.[8]
2. We calculated the age-adjusted county-level rates by sex and race/ethnicity using SEER*Stat software and merged with the block-level population counts by county.[9]
3. We calculated the cancer incidence and death rates for CDs as follows:

$$r_j = \sum_{i=1}^{n} \left(\frac{p_{ij}}{p_j}\right) r_i$$

Where r_i is the age-adjusted cancer incidence/death rate for block i (estimated as the corresponding county incidence/death rate), p_{ij} is the population counts of block i within CD_j p_j is the total population count for CD_j, and n is the number of blocks within CD_j. This method has been used in a previous study.[4] To calculate counts for CDs, we first estimated counts by CDs within each county based on the proportion of the county population for each CD that intersect with the county, then we aggregated the counts from different counties that intersect with the CD. The district count n_j was defined as

$$n_j = \sum_{k=1}^{m} \left(\frac{p_{kj}}{p_k}\right) n_k$$

Where n_k is the number of cases in county k, p_k is the population in county k, p_{kj} is the population of county k within district j, and m is the number of counties that intersect with CD_j. This gives the same estimates as weighting at the block level similar to the rate calculations above but is a more efficient calculation in terms of computing time. The district counts for both sexes combined were obtained by summing the male and female estimates.

Estimates were calculated by sex (both sexes, male, and female) and race/ethnicity (all races, non-Hispanic White, Black, and Hispanic). Block-level population data were not available by ethnicity for races other than White. As a result, the estimates for non-Hispanic Blacks could not be calculated; the Black race category includes both Hispanics and non-Hispanics.

Maps for incidence and mortality estimates for all cancers combined, and for 5 individual cancers (lung/bronchus, colorectal, female breast, cervical, and prostate) were constructed in ArcGIS (version 10.3). We chose to present data for these 5 individual cancers for which evidence-based prevention efforts (tobacco control and screening) are being implemented.[10,11] Estimates for cancer incidence and death rates by sex were displayed in quintiles on the maps to enable comparisons to previous studies.[3] Using quantile

classification for choropleth maps has also been found to convey information to readers with great accuracy in comparison to other data classification methods.[12] Data are suppressed for cells with fewer than 16 estimated cases. Incidence data for Hispanics in Delaware, Kentucky, and Massachusetts were suppressed at the states' request.

Results

Estimated Cancer Burden by Congressional Districts

Cancer incidence rates by congressional district.—Estimates for overall cancer incidence rates by CD ranged from about 320 to 610 new cancer cases per 100,000 persons per year (Figure 1). In general, cancer incidence rate estimates were highest in CDs in the Southern and Eastern regions of the country. When examined by race, CD incidence rate estimates were lower for Hispanics compared to other races. Black men had higher incidence rate estimates (rates in highest quintile) compared to men of other races across CDs in the Southern and Eastern regions. Among women, Non-Hispanic White women had higher incidence rates (rates in the third and fourth highest quintiles) compared to women of other races across CDs.

Incidence rate estimates for the 5 individual cancers by CD are shown in Figures 2 through 5. The distribution of lung cancer incidence rate estimates was similar to that of overall cancer incidence rate estimates. Among men, CDs in the Southern and Eastern part of the country had higher lung cancer incidence rate estimates, and Hispanic men had lower incidence estimates compared to men of other races. Among women, Non-Hispanic White women also appeared to have higher incidence rates compared to women of other races in CDs in the Southern and Eastern regions. The distribution of colorectal cancer incidence estimates was also similar to overall cancer incidence, with high incidence estimates seen in some Midwest states along with CDs in the South and East. The distribution of CDs with high incidence estimates for breast cancer appeared to differ by race; for non-Hispanic White women, CDs on the West and Northeast Coasts appeared to have higher incidence estimates. For Black women, CDs in the Appalachian region appeared to have higher incidence estimates for breast cancer. Cervical cancer incidence was higher in CDs for the South overall; among Black women, CDs in the Southeast had the highest rate estimates and among Hispanic women, CDs in the Southwest had the highest estimates. Prostate cancer incidence estimates showed a greater number of CDs with high incidence estimates for Black men in the Southern and Eastern regions compared to men of other races.

Cancer mortality rates by congressional district.—The distribution of the overall cancer death rates by CD was similar to that of incidence rates, ranging from about 100 to 290 cancer deaths per 100,000 persons per year (Figure 6). Overall cancer mortality rates were highest in CDs in the Southern and Eastern regions of the country and were lower for Hispanics compared to other races. Mortality estimates were highest for Blacks compared to other races; both Black men and women had a greater number of CDs with high mortality rate estimates (rates in highest quintile) compared to other races.

Mortality rate estimates for the 5 individual cancers by CD are shown in Figures 7 through 10. When examined by race, for all cancers, Black men and women had higher mortality

rates (rates in the highest quintile) in more CDs compared to men and women of other races. With the exception of prostate cancer, the distribution of cancer mortality rates by CD for the individual cancers was similar to that of overall cancer mortality with higher mortality rates seen in the Southern and Eastern regions of the country. Among non-Hispanic Whites, prostate cancer rates were highest in the CDs in the West, with lower rates in the Southeast.

Visualizing Cancer Burden Estimates by Congressional Districts

Cancer incidence and mortality estimates by congressional district are readily available to the public on the US Cancer Statistics Data Visualizations tool (www.cdc.gov/cancer/dataviz).[1] Estimates for overall cancer incidence and mortality, as well as those for the leading 20 cancer sites, are also easily accessible in the tool (Figure 11). Data graphics displayed in this tool include: (1) maps showing cancer incidence and mortality estimates for CDs in each state; (2) graphs ranking cancer incidence and mortality rates for leading cancer sites for each CD; (3) graphs ranking cancer incidence and mortality rates by sex and by race (non-Hispanic White, Black, and Hispanic) for each CD. The Data Visualizations tool also allows users to see the data used to generate the graphs and maps in table format by using the “Table” tab and to download data tables for cancer rates by CDs by using the “Export” tab.

Discussion

The distribution of cancer incidence and death rates by CD showed similar patterns to those that have been observed at the county and state levels, with the highest cancer incidence and death rates observed in CDs in the Southern and Eastern regions of the country.[1] The patterns of cancer death rate estimates by CD from this analysis are similar to those found in previous studies conducted for the 1990–2001 and 2002–2011 periods.[3,4]

Overall, Hispanics had lower incidence and death rates and Blacks appeared to have a greater number of CDs with higher incidence and death rates. The distributions of lung and colorectal cancer incidence and mortality rates by CD were similar to those observed for overall cancer rates. The examination of breast, cervical, and prostate cancer incidence and death rates by CD highlighted differences in rate distributions by race/ethnicity. Graphics of these cancer rate estimates by CD, and other national cancer data are readily accessible to the public on the USCS Data Visualizations tool (www.cdc.gov/cancer/dataviz).[1]

Key strengths of this analysis are that it uses national cancer incidence data and provides an examination of both cancer incidence and death rates, including estimates for 5 individual cancer rates. Limitations of this analysis include the assumptions made that the population proportions from the 2010 US census used for weighting still apply to the study period (2012–2016) and rates at the block-level were similar to those at the county-level. This may underestimate or overestimate rates at the CD-level. Of note, most CD estimates were derived from whole county rate estimates. Only a small portion of CDs (65 of the 436; 15%) were composed of 2 or more counties, with county boundaries traversing adjacent CDs. Geocoding cases at the CD-level could address these limitations in the future. Lastly, race/ethnicity categories examined were limited based on available county-level data; we were

unable to provide CD rate estimates for Asian-Pacific Islanders or American Indian/Alaskan Native people due to small case counts.

This examination of cancer rates at the CD-level provides data that can be used to inform cancer control strategies at the local and national levels. Displaying the data with the Data Visualizations tool makes it easily accessible to the public and decision makers.

Acknowledgments

We would like to thank the Geospatial Research, Analysis, and Services Program (GRASP) team at the Agency for Toxic Substances Disease Registry for their work in developing the US Cancer Statistics Data Visualizations tool, particularly Doug Day, Anil Kolli, and Lance Broeker.

References

1. US Cancer Statistics Working Group. US Cancer Statistics Data Visualizations Tool, based on November 2018 submission data (1999–2016). www.cdc.gov/cancer/dataviz.
2. Henley SJ, Singh SD, King J, Wilson RJ, O'Neil ME, Ryerson AB. Invasive cancer incidence and survival—United States, 2013. MMWR Morb Mortal Wkly Rep. 2017;66:69. [PubMed: 28125576]
3. Siegel RL, Sahar L, Portier KM, Ward EM, Jemal A. Cancer death rates in US congressional districts. CA Cancer J Clin. 2015;65:339–344. [PubMed: 26208318]
4. Hao Y, Ward EM, Jemal A, Pickle LW, Thun MJ. U.S. congressional district cancer death rates. Int J Health Geogr. 2006;5:28. [PubMed: 16796732]
5. Centers for Disease Control and Prevention (CDC). United States Cancer Statistics (USCS). https://www.cdc.gov/cancer/uscs/about/index.htm.
6. United States Census Bureau. 2010 Census Summary File 1—(name of state or United States) [machine-readable data files]. https://www.census.gov/data/datasets/2010/dec/summary-file-1.html.
7. United States Census Bureau Geography Program. https://www.census.gov/programs-surveys/geography.html.
8. United States Census Bureau. Geographies: 115th Congress block equivalency files. https://www.census.gov/geographies/mappingfiles/2017/dec/rdo/115-congressional-district-bef.html.
9. National Cancer Institute (NCI). Surveillance Epidemiology and End Results Program. SEER*Stat Software Version 8.3.5 https://seer.cancer.gov/seerstat/.
10. Hall IJ, Tangka FK, Sabatino SA, Thompson TD, Graubard BI, Breen N. Patterns and trends in cancer screening in the United States. Prev Chronic Dis. 2018;15:E97. [PubMed: 30048233]
11. Farrelly MC, Chaloupka FJ, Berg CJ, et al. Taking stock of tobacco control program and policy science and impact in the United States. J Addict Behav Ther. 2017;1(2):8. [PubMed: 30198028]
12. Brewer CA, Pickle L. Evaluation of methods for classifying epidemiological data on choropleth maps in series. Ann Assoc Am Geogr. 2002;92:662–681.

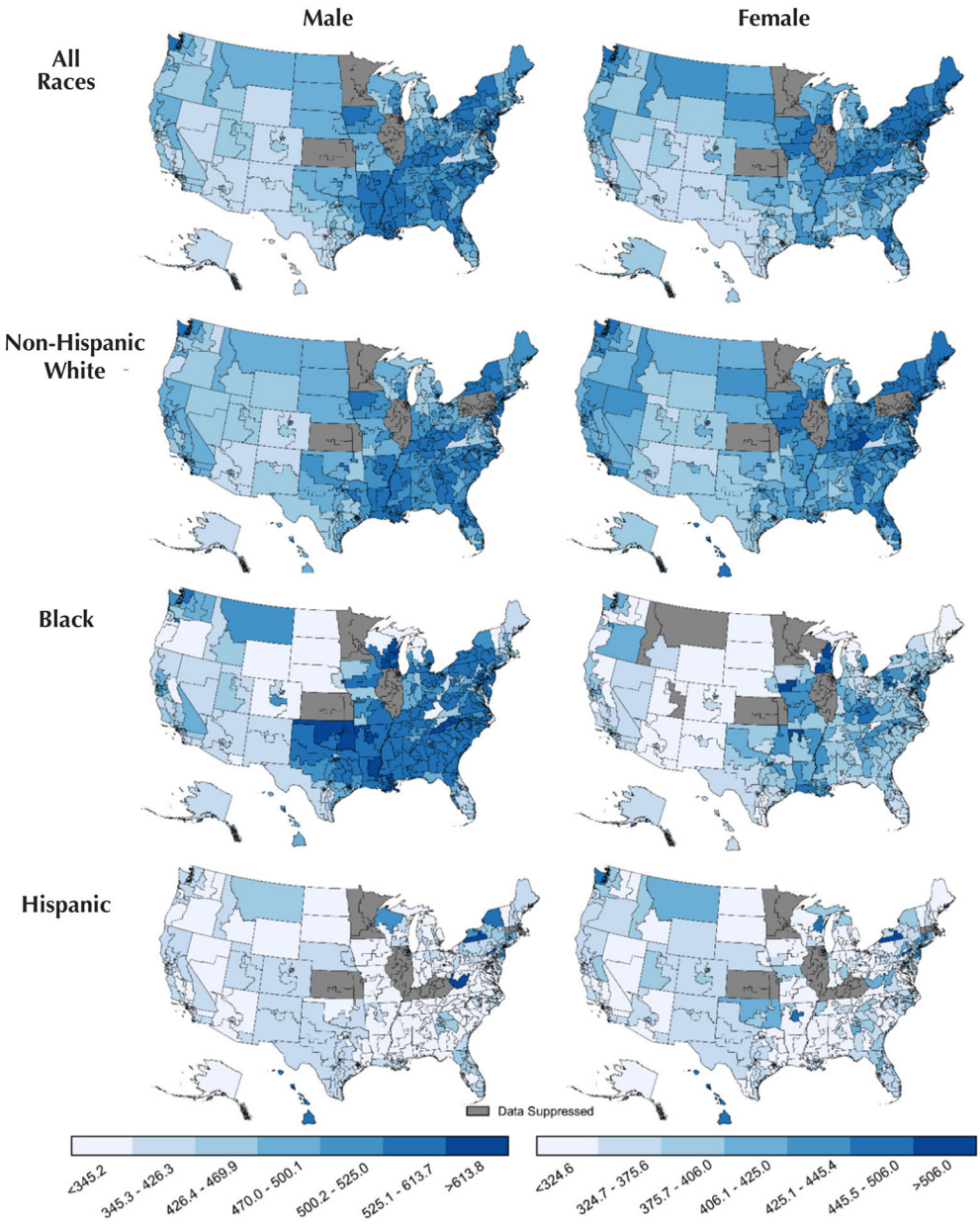

Figure 1. Overall Cancer Incidence Rate Estimates (per 100,000) by Sex and Race/Ethnicity in 406 Congressional Districts (CDs) from 2012–2016

Data are not included for Kansas (4 CDs), Minnesota (8 CDs), and Illinois (18 CDs). Data were suppressed for CDs with cancer case counts below 16 (these areas are depicted in gray).

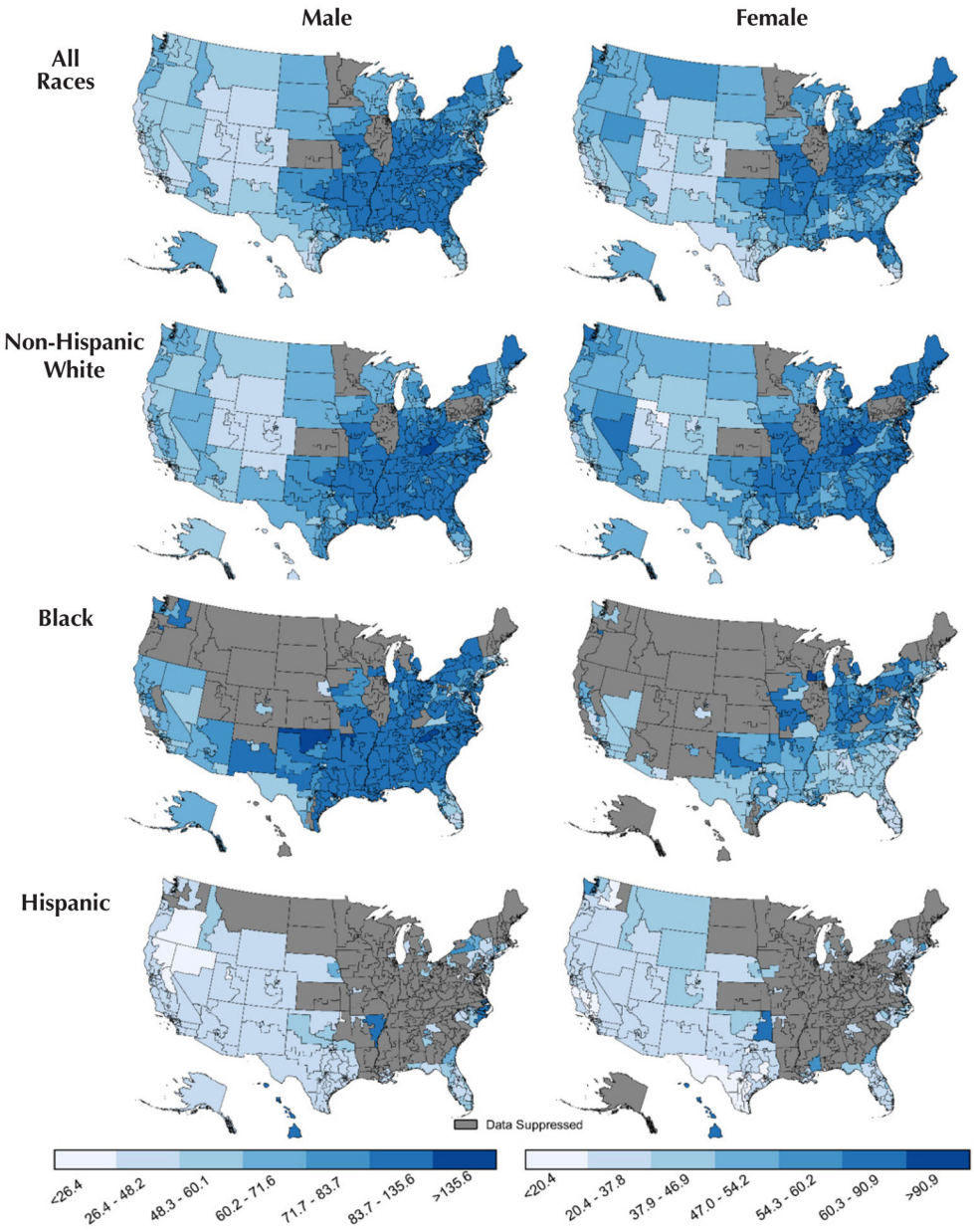

Figure 2. Lung Cancer Incidence Rate Estimates (per 100,000) by Sex and Race/Ethnicity in 406 Congressional Districts (CDs) from 2012–2016

Data are not included for Kansas (4 CDs), Minnesota (8 CDs), and Illinois (18 CDs). Data were suppressed for CDs with cancer case counts below 16 (these areas are depicted in gray).

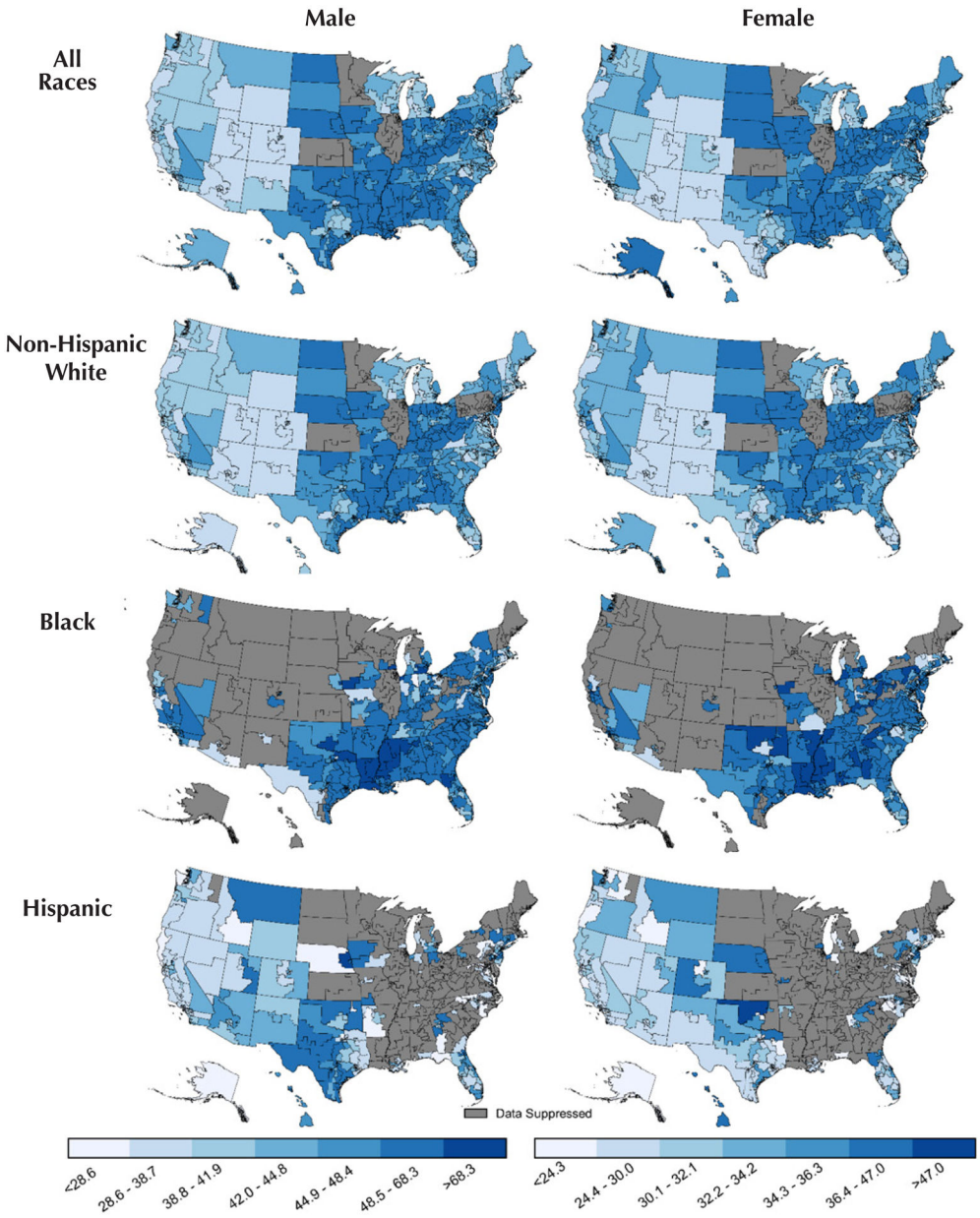

Figure 3. Colorectal Cancer Incidence Rate Estimates (per 100,000) by Sex and Race/Ethnicity in 406 Congressional Districts (CDs) from 2012–2016

Data are not included for Kansas (4 CDs), Minnesota (8 CDs), and Illinois (18 CDs). Data were suppressed for CDs with cancer case counts below 16 (these areas are depicted in gray).

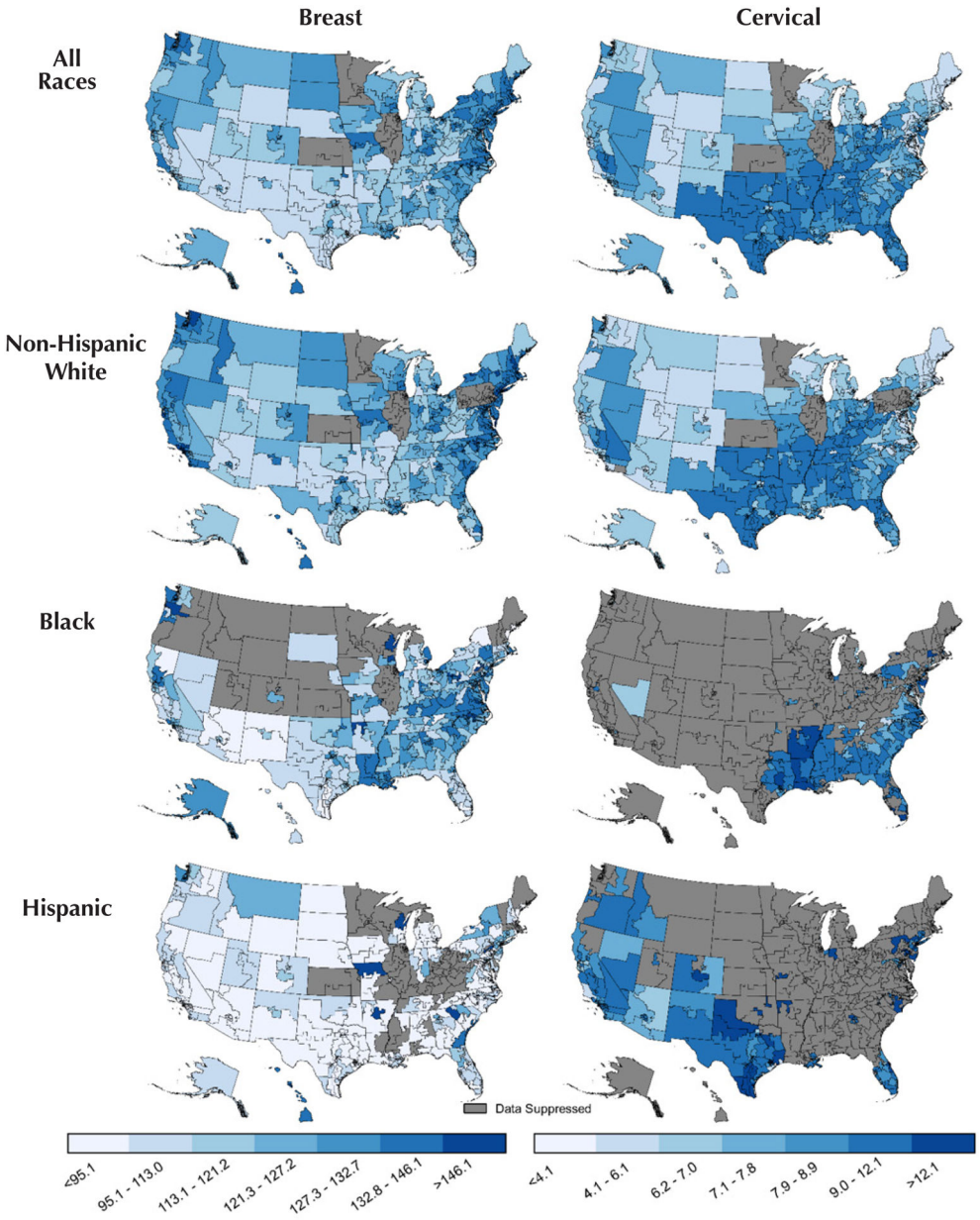

Figure 4. Breast Cancer and Cervical Cancer Incidence Rate Estimates (per 100,000) by Sex and Race/Ethnicity in 406 Congressional Districts (CDs) from 2012–2016

Data are not included for Kansas (4 CDs), Minnesota (8 CDs), and Illinois (18 CDs). Data were suppressed for CDs with cancer case counts below 16 (these areas are depicted in gray).

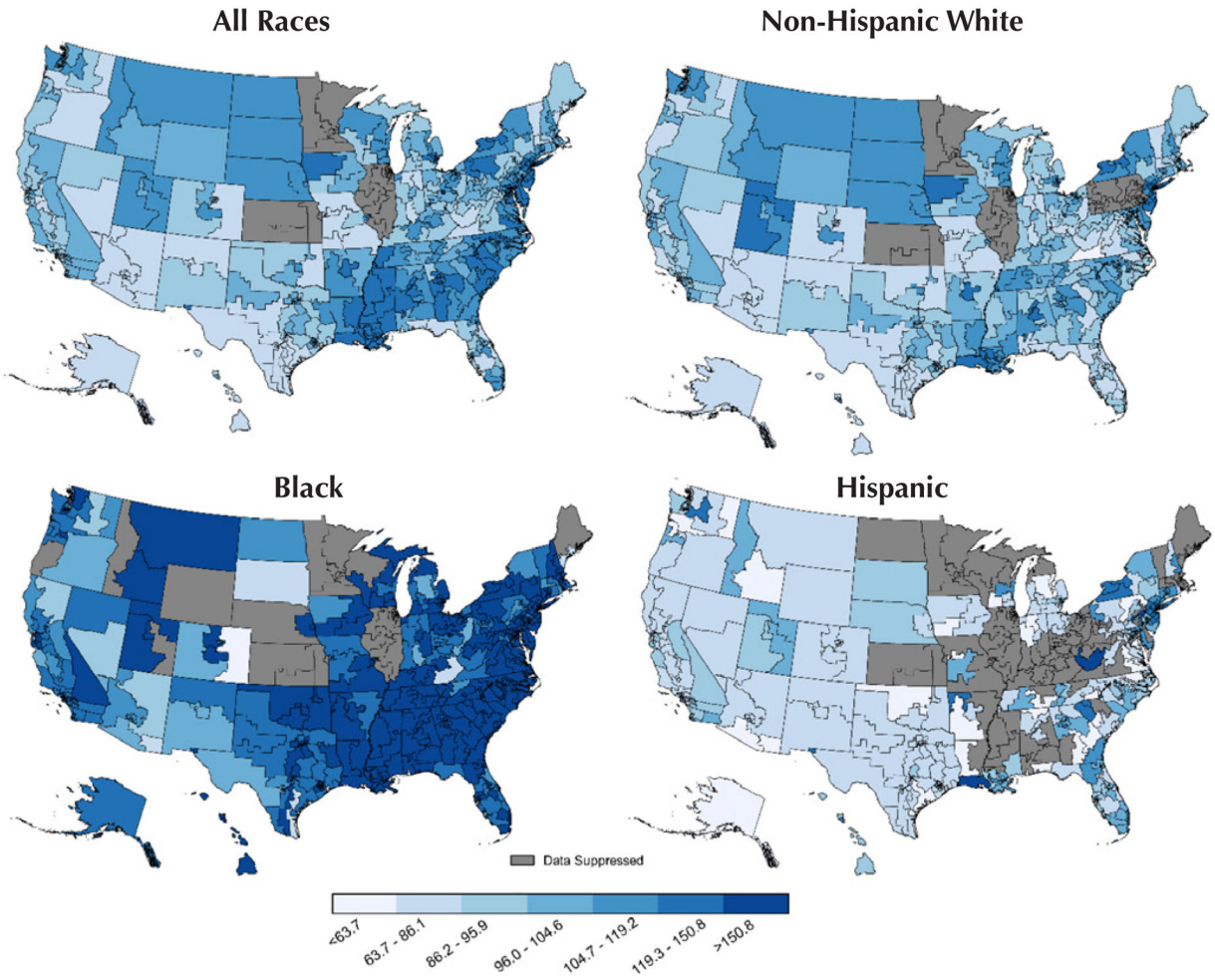

Figure 5. Prostate Cancer Incidence Rate Estimates (per 100,000) by Sex and Race/Ethnicity in 406 Congressional Districts (CDs) from 2012–2016

Data are not included for Kansas (4 CDs), Minnesota (8 CDs), and Illinois (18 CDs). Data were suppressed for CDs with cancer case counts below 16 (these areas are depicted in gray).

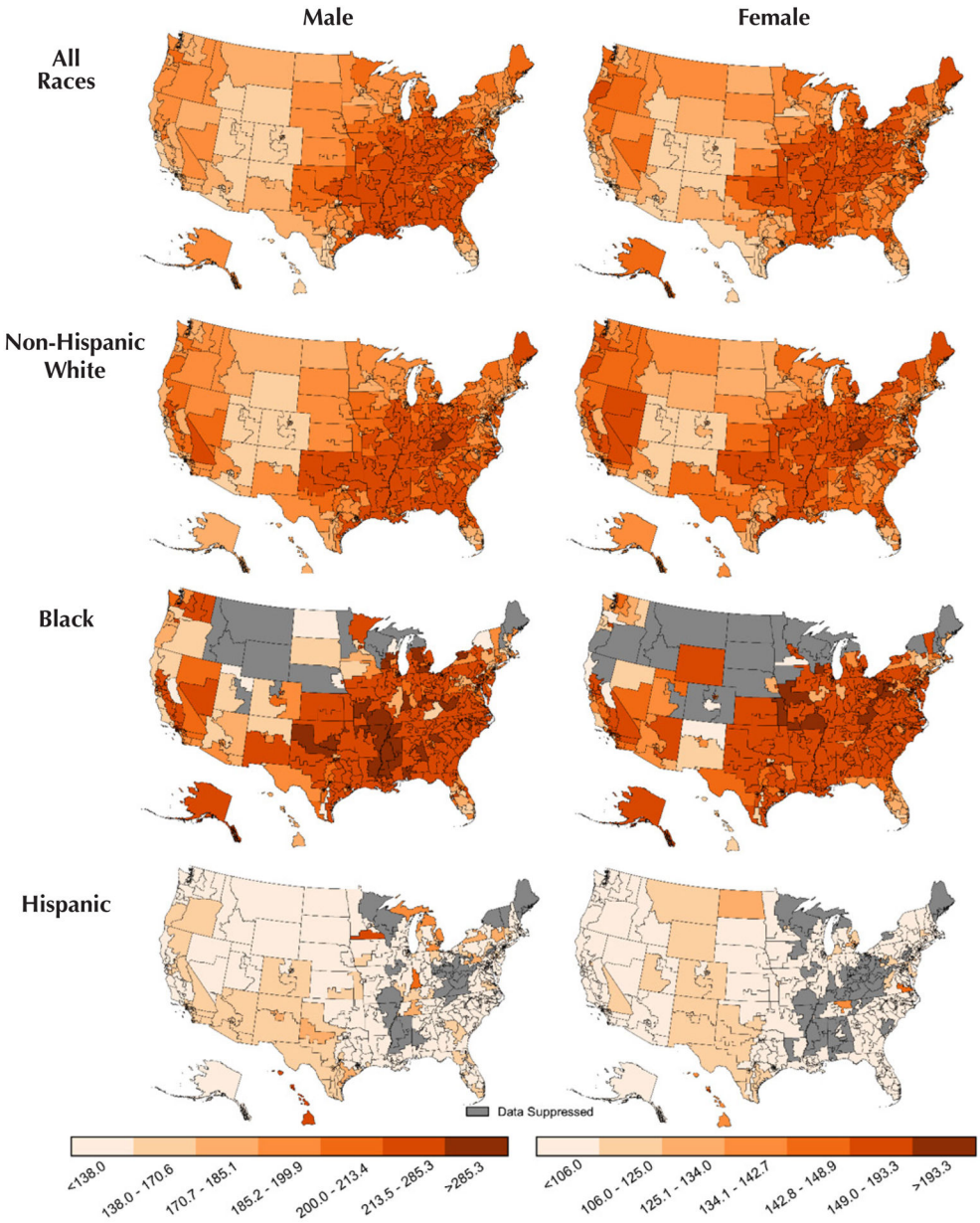

Figure 6. Overall Cancer Death Rate Estimates (per 100,000) by Sex and Race/Ethnicity in 436 Congressional Districts from 2012–2016

Data are suppressed for congressional districts with cancer case counts below 16 (these areas are depicted in gray).

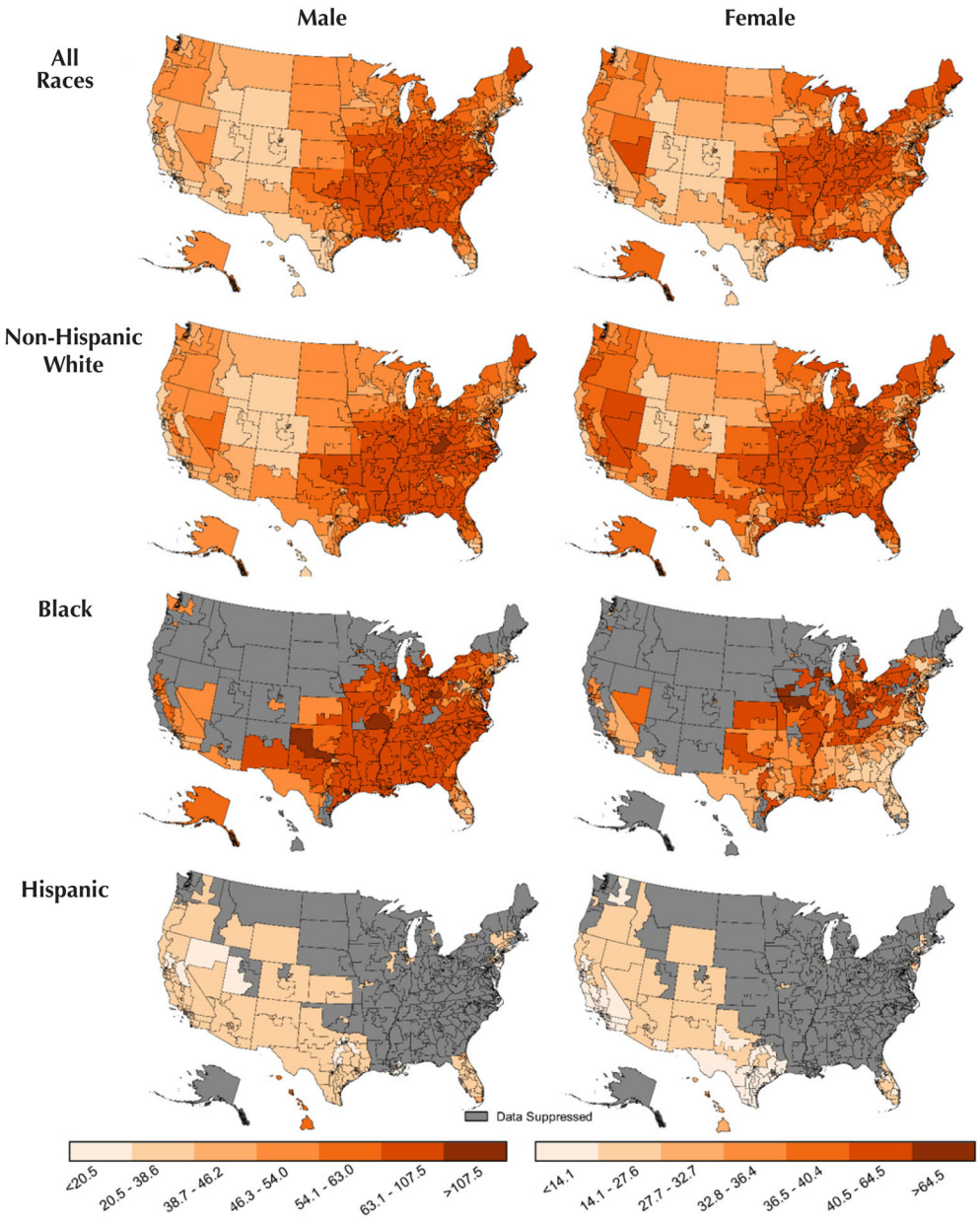

Figure 7. Lung Cancer Death Rate Estimates (per 100,000) by Sex and Race/Ethnicity in 436 Congressional Districts from 2012–2016

Data are suppressed for congressional districts with cancer case counts below 16 (these areas are depicted in gray).

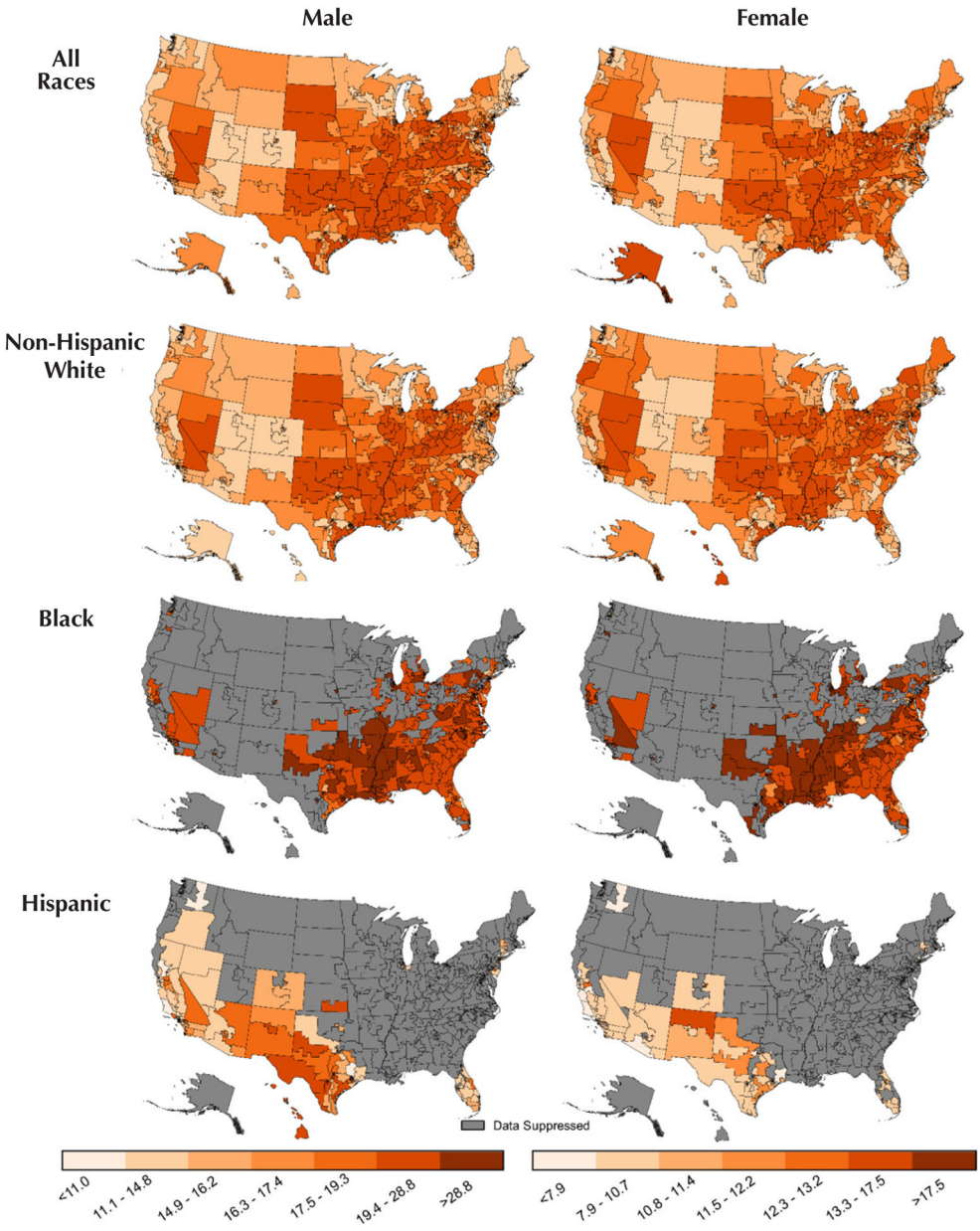

Figure 8. Colorectal Cancer Death Rate Estimates (per 100,000) by Sex and Race/Ethnicity in 436 Congressional Districts from 2012–2016

Data are suppressed for congressional districts with cancer case counts below 16 (these areas are depicted in gray).

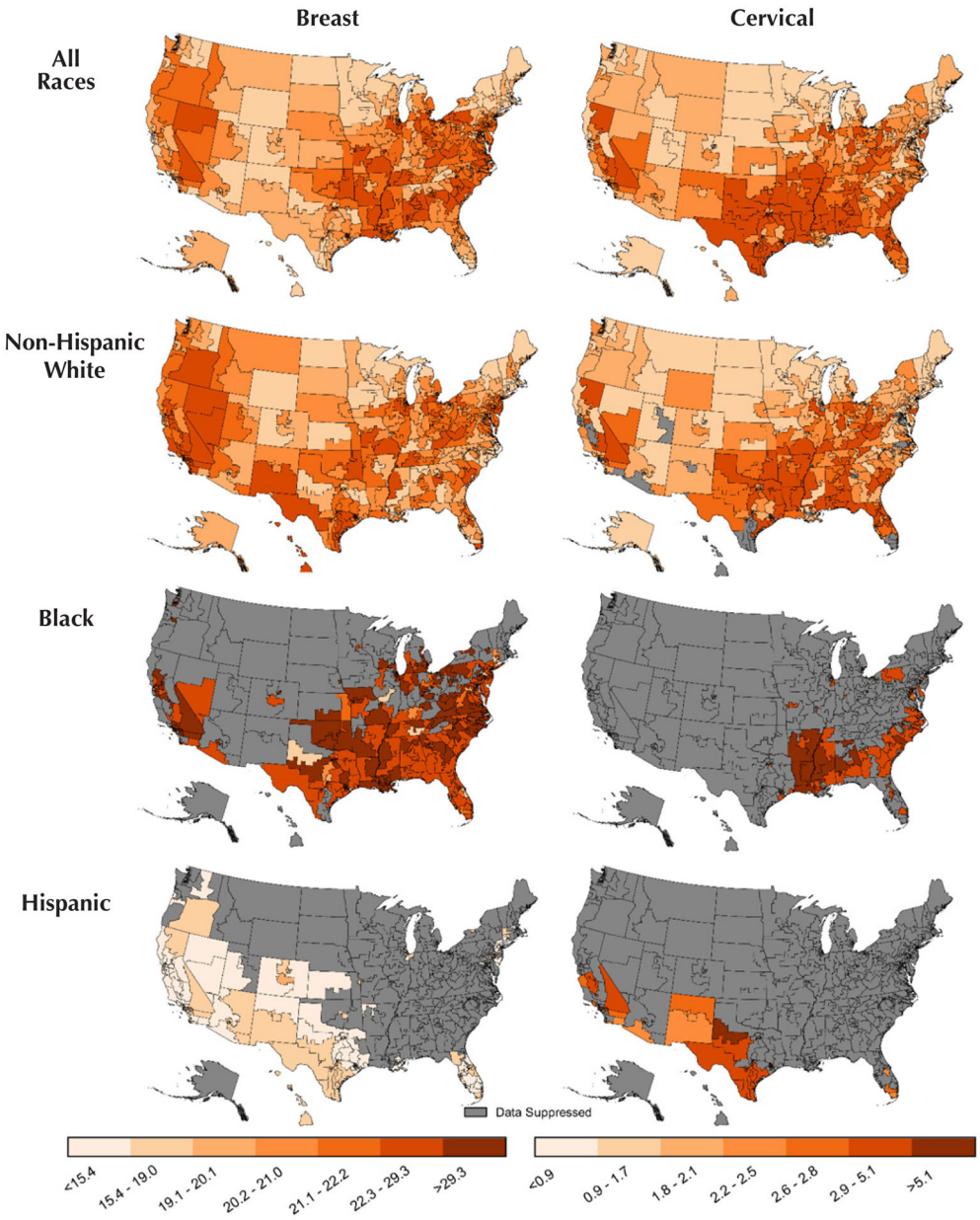

Figure 9. Breast Cancer and Cervical Cancer Death Rate Estimates (per 100,000) by Sex and Race/Ethnicity in 436 Congressional Districts from 2012–2016

Data are suppressed for congressional districts with cancer case counts below 16 (these areas are depicted in gray).

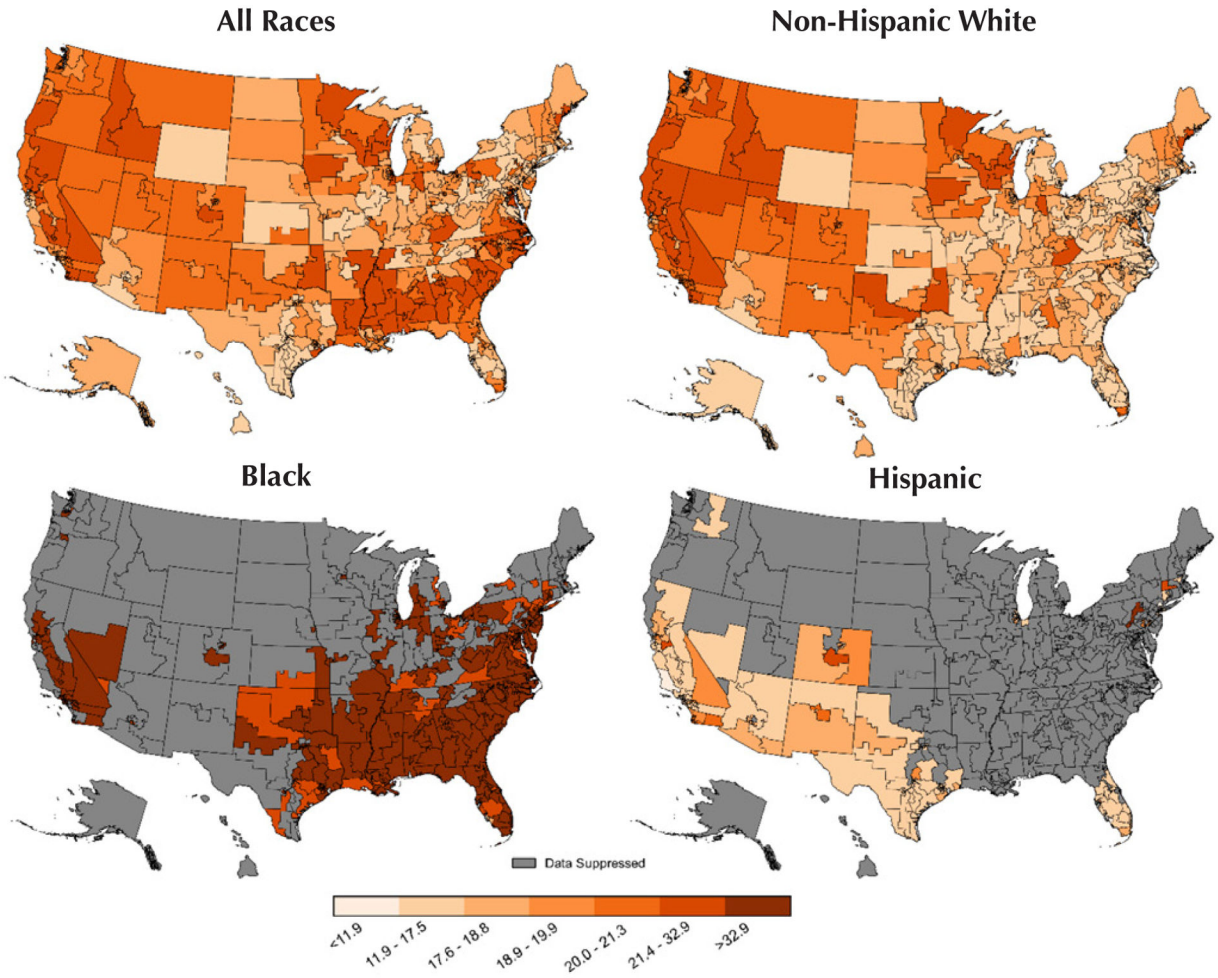

Figure 10. Prostate Cancer Death Rate Estimates (per 100,000) by Sex and Race/Ethnicity in 436 Congressional Districts from 2012–2016

Data are suppressed for congressional districts with cancer case counts below 16 (these areas are depicted in gray).

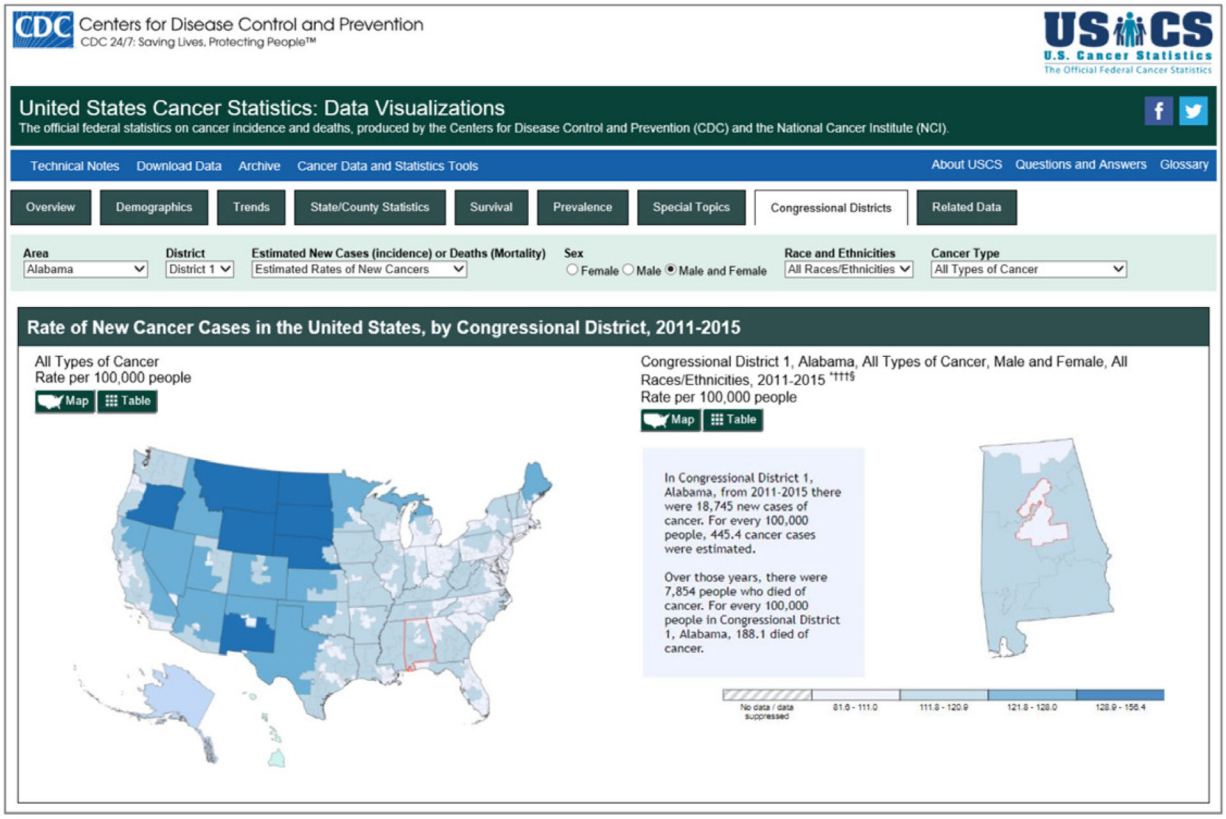

Figure 11.

Screenshot of US Cancer Statistics (USCS) Visualization Website that Enables the Public to View Cancer Incidence and Death Rate Estimates by State, County, and Congressional District