# Cartogram Mapping and its Application to Cancer Data Visualization

*Stephanie Kobakian, Jessie Roberts and Dianne Cook*

## Contents

## Abstract

Choropleth maps are a traditional form of presenting spatial data. They are widely used in cancer atlases, leveraging the familiar geographical regions to communicate cancer statistics around the world. The presentation of this data has improved with the developments in the statistics used to measure the impact of cancer on a population. Web atlases have utilized developments in interaction and animation techonolgies to empower individual users to explore cancer statistics for their own purposes. There has also been progress and contributions to alternative displays that present spatial data with varying levels of focus on the spatial distribution of a statistic. This work highlights the impact of the varying sizes of areas created by the administrative boundaries that define communities and affects the way that spatial impact of cancer on Australian communities is understood. These displays are not often utilized in cancer atlases, this provides an opportunity to improve the communication of cancer statistics. Including alternative visualizations will provide users a new perspective on the spatial distribution of cancer statistics.

# 1 Introduction

Cancer statistics are usually delivered as an aggregated value for a geopolitical area. Presenting these statistics requires transforming individual observations into aggregations of communities as geographical units, in large part for privacy protection or political and policy purposes. The information could be as simple as counts per area (e.g. state, province, local government area, post/zip code). Counts alone are not sufficient to compare areas, because the populations of areas are all different. In this case, the counts data needs to be merged with population data to appropriately calibrate it to incidence, for example, rate per 100000 people. This type of data is collected on a routine basis for public health purposes, and may be made available to the general public as a service to the community. The task, then, is to examine what are the usual ways to communicate cancer statistics, to the public, are there alternative approaches, and what are the pros and cons of these choices.

A common approach to communicate cancer statistics, is to display statistics on a map. Using a choropleth map: the statistic is mapped to color and the geographic region is filled with this color. The viewer would then be able to examine the spatial distribution of the disease incidence, where there is a trend in longitude or latitude, or rural vs urban, or coastal vs inland, or even specific hot spots of the disease. Visualizing diseases on maps is often the first step in exploratory spatial data analysis and effectively helps in the formulation of hypotheses (Jahan et al. 2018). Disease maps help to present geographic patterns that may be overlooked in a table, obscuring the geospatially related statistics (Moore and Carpenter 1999). By providing a visual representation of cancer outcomes, geographic patterns of disease may be identified and effectively addressed with public health policy and actions. Exeter (2017) recognizes one of the key challenges with mapping spatial patterns of disease is the design of visualizations. This paper addresses the visualization techniques and their applications to cancer statistics. Highlighting the differences and historic use of these displays.

The paper is organized as follows. The next section describes the choropleth map which is the common approach to disease maps. Section 3 surveys atlases in use today. Section 4 describes an alternative display, the cartogram what may be useful for countries that have heterogeneously sized geographic units. The pros and cons of these approaches is discussed in Section 5. Disease maps are more useful when made interactive, and common options are described in Section 6, along with a discussion of benefits and disadvantages. The last section summarizes the paper and discusses future directions.

## 2 Map displays for disease data

A choropleth map is used to display differences in the geographical distribution of data by spatial unit by shading areas of a map. The geography is faithfully rendered, and the color rendering is designed to reveal spatial patterns among data values. A choropleth is constructed by drawing the geographic or political boundaries, and filling the shapes with colors to represent values of a measured variable (Tufte 1990). Figure 2 shows a choropleth of age-adjusted rate (per 100,000 people) of new cases of lung and bronchus in the USA, averaged over 2012 through 2016. The data was extracted from the official federal statistics by U.S. Cancer Statistics Working Group. (2018) on cancer incidence and deaths, produced by the Centers for Disease Control and Prevention (CDC) and the National Cancer Institute (NCI).

Early versions of choropleth maps used symbols or patterns instead of color. Bell et al. (2006) discuss the use of choropleths to visualize cancer data, and Walter (2001) gives an overview of the development of these maps for displaying disease data.

Utilizing the state boundaries can make a map familiar to read (Brewster and Subramanian 2010), and allows viewers to visually infer the spatial relationships in the data. The familiarity of the geography is a worthy consideration when presenting results of spatial analysis. Just as geographers are no longer the only creators of maps, Bell et al. (2006) suggests the audiences of spatial health data analysis have extended beyond researchers to the public, policymakers and the media.

Identifying and explaining spatial structures, patterns, and processes involves considering the individuals in communities and organizing communities into representable units (Moore and Carpenter 1999). In Figure 2, a west to east spatial trend of increasing rates, can be seen. There is also a spatial outlier – Utah has a noticeably lower rate that its neighbors. Also Kentucky has a noticeably high rate, and Maine also has a higher rate than its neighbors. There is something of a cluster of higher rates around the tobacco states.

While the areas are recognizable shapes, they are often politically driven boundaries with individual areas being of non-uniform size, containing different population densities and subject to change over time. The different population and geographical sizes of administrative areas can attract attention to the shades of the underpopulated but large areas (Tufte 1990), Skowronnek (2016) calls this an area-size bias. Choropleths can inhibit visual inference when presenting human related statistics as the display may draw attention away from the 'potentially more important results in the more populous communities' that are geographically smaller (Exeter 2017).

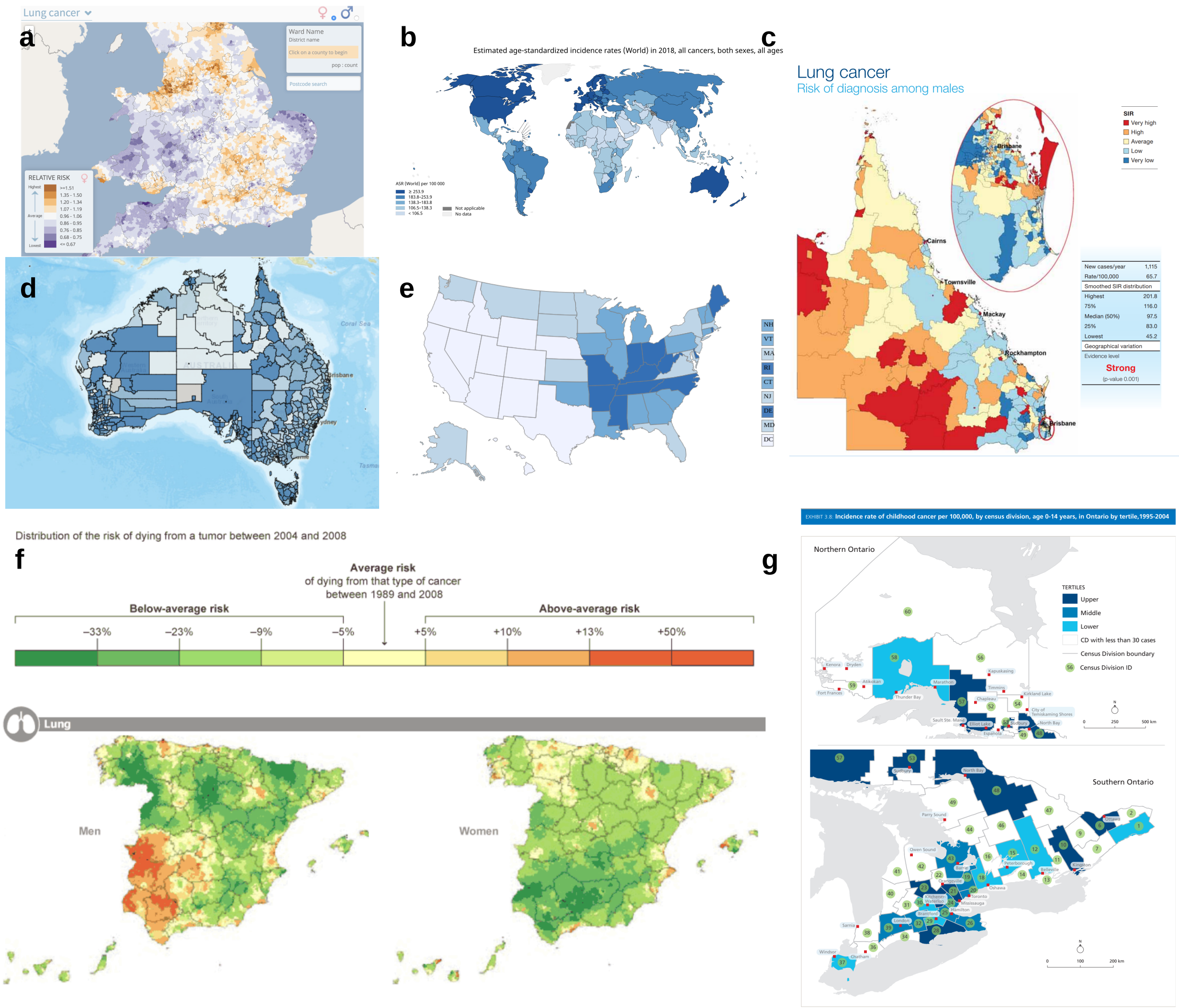


Figure 1: Publicly available choropleth cancer maps published between 2010 and 2015.

## 3 Cancer atlases

Choropleth maps can be useful devices for communicating information to public on a familiar map base. A cancer atlas is a choropleth map, or collection of maps, representing cancer incidence and mortality for a country, or group of countries. In epidemiology, choropleths are often used as a tool to study the spatial distribution of cancer incidence and mortality. The data collection methods of cancer mortality rates across regions, and the administrative control within regions lends itself to choropleth visualization. d'Onofrio et al. (2016) provides the definition of a cancer atlas, beginning with Haviland's maps in 1875, they attribute UK cancer atlases to Howe (1963), and early work in US cancer atlases can be attributed to Burbank (1971). The increasing development and use of disease maps can be attributed to the availability of geographic information system software (Exeter 2017). The choropleth maps presented levels via hatching or dots on a black and white scale. These atlases were key to developing hypotheses regarding areas with unusually high rates, geographic correlations, work related exposures, and high risk diets (d'Onofrio et al. 2016).

The presentation of cancer statistics has changed over time with greater access to computational power. Mortality rates are now often presented as relative rates of risk across the population, and age adjusted to correct for the higher prevalence of cancers in older populations. Howe (1989) describes Stock's development of the standardized mortality ratios through the 1930s. Table 1 summarizes the measures presented in published cancer atlases, and provides a definition of each measure.

Table 1: Measures used to report cancer statistics

| Measure | Details |
|---|---|
| 1. Count | Crude cancer counts |
| 2. Rate per 100,000 | Cancer incidence per 100,000 population |
| 3. New cancer cases per 100,000 | Specific methods could not be found |
| 4. IR (Incidence Ratio) | $(IR)_i = \frac{(Incidence\ Rate)_i}{Average\ Incidence\ Rate}$,<br>Cancer incidence rate in region $i$ over the average cancer incidence rate for the total region |
| 5. Age Adjusted Rate per 100,000 | Standardized by age structure or region |
| 6. Age Adjusted Relative Risk | RR standardized by age structure in each region $i$ |
| 7. SIR (Standardized Incidence Ratio) | IR standardized by age structure in each region $i$ |
| 8. Below or above Expected | Alternative expression of the SIR |
| 9. RER<br>(Relative Excess Risk) | $RER = \frac{(Cancer\ related\ mortality)_i}{Average\ cancer\ related\ mortality}$<br>Represents the estimate of cancer related mortality within five years of diagnosis<br>Also referred to as 'excess hazard ratio' |

Cancer maps are effective tools for communicating to wide range of audiences, including the general public and others not trained in statistical analyses. These visualizations enable non-expert audiences to interpret the outputs of sophisticated statistical analyses. Cruickshank (1947) as cited by Walter (2001), discusses using visuals as a 'formal statistical assessment of the spatial pattern'. Overwhelmingly, cancer maps utilized to communicate to the public and other non-expert audiences are choropleths. These atlases provided Incidence, Survival and Mortality rates to the public on areal map bases.

The Atlas of Cancer in Queensland (Cramb, Mengersen, and Baade 2011) focused on highlighting the difference in experience for those living in rural and disadvantaged areas, the Standardized Incidence Ratios (SIR) were modeled via Bayesian methods. The presentation of these rates considers not only the shapes of the areas, but also appropriate choices of color blind friendly color schemes, and categories of the values to highlight significantly different areas.

## 3.1 Overview of publicly available atlases

Roberts (2019) identified 33 publicly available cancer atlases, published between January 2010 and November 2015. All of these use choropleth maps. All but one of these were published by non-commercial organizations, including not-for-profits, government, research organizations, advocacy groups or government funded partnerships. The use of choropleths within the public domain mirrors the heavy use of choropleth maps within the research literature, discussed above. The cancer atlases identified by Roberts (2019) covered geographies from all around the world, most focused on single nations. Figure 1 displays a global (b), national (a, d, e, f) and state (c, g) choropleths. The sections below provide details on each of the maps within Figure 1.

### 3.1.1 The Environment and Health Atlas of England and Wales

Figure 1a contains an image from *The Environmental and Health Atlas of England and Wales* (Emperial College London - Small Area Health Statistics Unit 2010). This map shows the relative risk for women developing lung cancer in England and Wales in 2010. The cancer data used to generate this map came from Office for National Statistics (ONS) (England) and from the Welsh Cancer Intelligence and Surveillance Unit (WCISU).

### 3.1.2 Globocan 2012: Estimated Cancer Incidence, Mortality and Prevalence Worldwide

The map seen in Figure 1b is from the *Globocan 2012: Estimated Cancer Incidence, Mortality and Prevalence* (World Health Organization's International Agency for Research on Cancer 2018). This global map shows age standardized incidence rates (per 100,000) for all invasive cancers for both men and women, aggregated at a national level for 2018. This map is published by the World Health Organization's International Agency for Research on Cancer. Data was sourced from cancer registries in each country, contributing registries can be seen in the supplementary material on the cancer atlas website.

### 3.1.3 Atlas of Cancer in Queensland

Figure 1c shows an extract from *The Atlas of Cancer in QLD* (Queensland Cancer Registry 2011). This map was published by the Queensland Cancer Council and shows the relative incidence ratio of lung cancer in males in the state of QLD within Australia based on data from 1998 to 2007. Data to generate this map was sourced from the *Queensland Cancer Registry*.

#### 3.1.4 Bowel Cancer Australia Atlas

Figure 1d shows the *Bowel Cancer Australia Atlas* (Bowel Cancer Australia 2016). Published by *Bowel Cancer Australia* (Australia). This map shows the percentage of Australian males between 50 - 54 years of age that diagnosed with bowel cancer in 2016 in Australia. The source of the data is not provided.

#### 3.1.5 United States Cancer Statistics: An Interactive Cancer Statistics Website

The *United States Cancer Statistics: An Interactive Cancer Statistics Website* (U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute - Cancer Statistics Working Group 2019) can be seen in Figure 1e. This map contains the incidence rate per 100,000, of all cancer types for men and women in the United States in 2016, aggregated at the state level. The map was published by the *Centers for Disease Control and Prevention*. Incidence data seen in this map were compiled from cancer registries meeting U.S. Cancer Statistics data quality criteria covering 100% of the U.S. population.

#### 3.1.6 Map of Cancer Mortality Rates in Spain

The *Map of Cancer Mortality Rates in Spain* (El Pais 2014) can be seen in Figure 1f. The side by side maps show relative risk of lung cancer for men vs women based on data from 2004 to 2008. The source of the data and statistical methods are unknown.

#### 3.1.7 Atlas of Childhood Cancer in Ontario

Figure 1g shows an extract from the Atlas of Childhood Cancer in Ontario and specifically displays the *incidence rate of childhood cancers* (Pediatric Oncology Group of Ontario 2015) per 100,000 (by census division) for children aged 0-14, in Ontario from 1995 to 2004.

### 3.2 Common statistics displayed

Cancer maps are powerful visualizations that summarize complex statistical analyses, however the statistics represented in these maps cannot tell the entire story. Supplementary graphs and plots are often included to add more depth and information to the map. Bell et al. (2006) suggests additional materials such as tables, graphs, and text explanations support understanding and inference derived from maps, ensuring the message communicated will be consistent across a range of viewers. There are many visualizations used for displays of statistical summaries, these may be dot plots, bar plots, box plots, times series plots, cumulative distribution plots, scatter plots, Q-Q plots. These additional displays of the cancer distribution can provide alternative views of the cancer statistics, as well as the supporting statistics including error, confidence intervals, distributions, sample or population sizes, standard deviation and other measures. When presenting cancer maps, d'Onofrio et al. (2016) believes the intuition derived from maps must be 'validated by rigorous statistical analyses', the supplementary statistics help for this validation.

### 3.2.1 Geographic hierarchies

While atlases are often used to describe differences between areas, statistics may be displayed at different levels of aggregation. Global health statistics can be aggregated to administrative and arbitrarily defined regions, such as those used by the World Health organization and the United Nations (Ferlay J 2018). World atlases can allow for displays of data aggregated into continents, countries, states, provinces and congressional districts (Group 2019).

### 3.2.2 Population distribution

It is extremely likely that each population area will have a different number of people. The distribution of the population residing in all areas may also be communicated in a table or histogram display (Northern Ireland Cancer Registry 2011). Atlases can connect the population to the land available to them by communicating population density. Atlases can also connect the population to the land available to them by communicating population density.

### 3.2.3 Statistical uncertainty

Additional statistics that accompany an atlas often include a measure of the statistical uncertainty surrounding the statistics presented in a choropleth. In the review of atlases in the public domain, atlases were considered to report uncertainty to the non-expert user if they included a measure of statistical uncertainty either within or alongside the map (Roberts 2019). The maps considered used standard and well-known measures including credible intervals and standard deviation, statistical significance, box plots and distributions. Other methods involve providing adjacent maps or overlapping maps with symbols (Kronenfeld and Wong 2017). The maps employing uncertainty ranged from static documents or infographics, to interactive online visualizations. Communicating the statistical uncertainty associated with the estimates occurs using confidence intervals (CI), credible intervals (CrI), statistical significance levels, box plots, distribution plots, and reporting sample size and standard deviations. Close to half of the atlases identified (42%, n=14) included some measure of uncertainty. The most common measure used to represent uncertainty were credible or confidence intervals (CIs).

### 3.2.4 Demographics

Demographics include information regarding the age and sex distribution of the areas presented. Sex is an important cofactor for cancer atlases. As some cancers are sex specific, and others may be found in both males and females, atlases often specify the relevant sex as part of the visual output in the displays. Digital atlases allow for users to interact with the controls of the displays, they can select males, females or both depending on the type of cancers explored.

### 3.2.5 Socio-economic indicators

Socio-economic indicators can explain how the experience of cancer prevalence varies for various members of a society. These indicators include unemployment rates, poverty rates, remoteness, and education levels achieved though, only a few atlases also explored the impact of rurality on cancer rates. These rates may also be explored as percentages above or below the mean or median value for the set of spatial areas. The Human Development Index can be used to understand the socio-economic experience of a community, as well as Income levels which can be sourced from the World Bank list of economies (Ferlay J 2018). The areas are often ranked and allocated to quintiles; each quintile can be presented as categories describing the ranking.

One of the concerns of adding too much information to a map is the fear of cognitive overload (McGranaghan 1993) in which the user reaches an information threshold, beyond which will not be able to make sense of the information. These concerns are not unfounded, complexity and density of representation methods appear to overwhelm novice decision makers, while experts are able to use the detail more readily when making-decisions (Cliburn et al. 2002). Interactivity is a design feature within modern mapping methods that can be used to incorporate additional information and complexity without overloading the user.

Interactivity enables supplementary information to be incorporated into online atlases without cluttering the display. Interactive design features found in online cancer maps include tool tip features, drop down menus, data selection, zooming and panning allow users to explore the map as they want more information and allow flexibility in the display (Monmonier 2018). The use of these supports were found in various online cancer maps identified by Roberts (2019). The controls for basic interactive features are often placed outside of the plot space (Pedersen 2018), thus the map image is updated/replaced as the user interacts with the controls. For example, changing the population age or other demographic variables. Some more advanced interactions include direct interactions with the plot via the use of overlaid tool tip features, very few cancer atlases involve these more complex selection tools.

Additionally, interactivity allows the user to toggle between different variables, map views or multiple realizations of possible future scenarios (Goodchild, Buttenfield, and Wood 1994). Thus providing additional mechanisms for the users comprehension as well as the uncertainty of the available information (MacEachren (1992); Van der Wel, Hootsmans, and Ormeling (1994)).

These interactive features provide an opportunity for users to explore the additional information available. This helps users to understand and interpret the spatial distribution presented, as well as validate, explain or explore the presented statistics and their relationships to each other and/or their underlying spatial distribution. This allows relationships between spatial areas and diseases to be explored with sophistication in non-traditional but still 'cognitively accessible' ways (Carr, Wallin, and Carr 2000). The interactive features of the publicly available maps identified by Roberts (2019) allow the exploration of geographic hierarchies, population distribution, statistical uncertainty, demographics and socio-economic indicators. Carr, Wallin, and Carr (2000) suggested LM plots as a solution to linking cartography and statistical graphics.

# 4 Alternative to choropleths

## 4.1 Micromaps

Choropleths are uni-variate displays, they do not inherently support multi-dimensionality. This makes it difficult to pair demographic, environmental and other factors with spatial distributions.

## 4.2 Cartograms

Choropleths imply uniformity of data across the geographic space however population densities are extremely unlikely to be uniform (Skowronnek 2016). A cartogram displays abstracted geographic space, with the intention of improving the presentation of the statistic of interest. For a single variable of interest, each map area is changed to emphasize the distribution by representing the corresponding value, in comparison to the value of the other areas (Dougenik, Chrisman, and Niemeyer 1985). Changes in the map base are implemented by altering the boundaries, and therefore shapes, of individual areas.

Australia presents an extreme case of an urban rural divide. The land mass occupied by urban electoral districts is only 10% of Australia, yet 90% of the population live in these urban areas. To present election results on a choropleth map should be 'unthinkable', as it means diminishing the visual impact of majority of the electorates. A 1966 cartogram presented an alternative where boundary lines were largely straight line, and the result looked very little like the geographical shape of Australia. This issue is felt in any nation which experiences a spatially heterogeneous population distribution. As this feature of population distributions continues to intensify, the need for cartograms as an alternative to a choropleth map should only increase.

Choropleths may be considered true topological maps, however, if the land mass displayed covers enough of the globe, there must be a transformation or distortion to display the land in 2D (Monmonier 2018). The amount of distortion is related to the distance covered by the landmass displayed Tobler (1963). World map projections reflect the frequent perspectives used to view the earth. Choropleth maps will always be distorted if they cover enough of the globe, just like photographs of the globe from space. Choropleth creation requires choosing a map projection that shows a favorable distortion of the geography for presenting the set of spatial information. Selecting a display can prevent misinterpretation of global statistics, as global maps face the challenge of equitable displays of land mass on maps (Raisz 1963). If the statistic presented on the map base relies on physical distance and is influenced by the topology there is no transformation needed, beyond choosing a reasonable projection.

Event cartograms change the area of regions on a map depending on the amount of disease related events, but this does not consider the effects of land area and population (Kronenfeld and Wong 2017). The purposeful distortion of the map space, transformed according to population density, is beneficial when a uniform density of the map base is desired. Population then becomes a uniformly distributed background for the statistic presented (Berry, Morrill, and Tobler 1964). Dorling (2011) suggests 'population distribution is often extremely uneven in former British colonies', this makes the distortion necessary (Griffin 1980). When implementing a distortion of the geographical shape according to population, the resulting display is an area cartogram (Olson 1976), or population-by-area cartogram (Levison and Haddon Jr 1965).

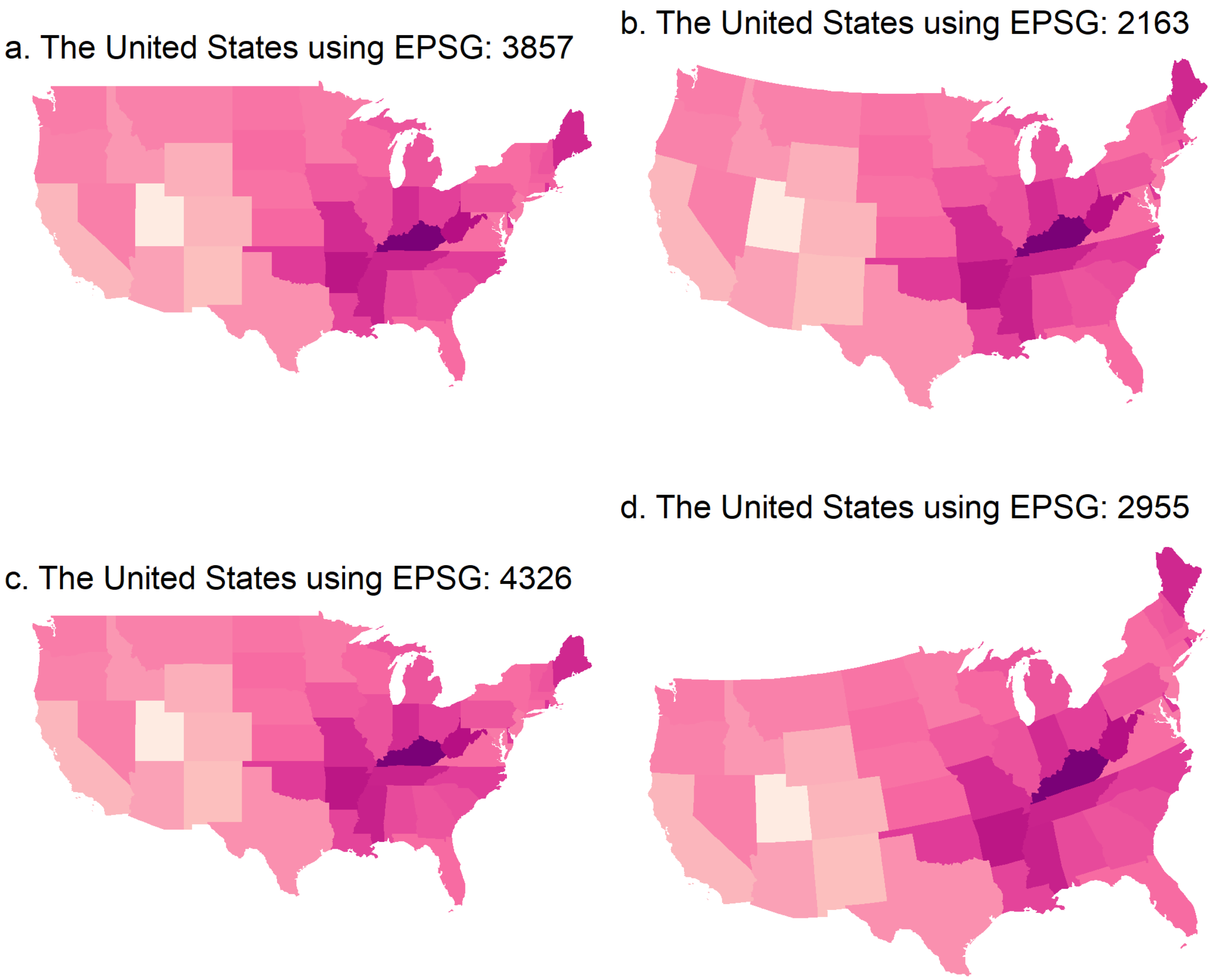


Figure 2: Four choropleth maps of the United States of America using various coordinate reference systems. Each state has been colored according to the average age-adjusted rate of incidence for lung and bronchus for females and males in the United States 2012-2016. The map projections alter the shapes and angles of the boundaries of each state. Maps a and b are similar in their straight edges, unlike maps c and d which curve on the northern United States border.

Cartograms provide an alternative visualization method for statistical and geographical information. The key difference between a choropleth and a cartogram is the desirable augmentation of the size, shape or distance of geographical areas (Dorling 2011). Monmonier (2018) suggests that white lies may be employed to create useful displays and map creators have the ability to draw lines that may distort the geometry and suppress features and it is easy for the average person to disregard the impact of transformations used to create cartograms. Cartograms may be seen as an extension of map transformations and projections. The favorable distortion is proportional to a value other other than actual earth size area (Olson 1976). A disadvantage of the conventional map is that sparsely populated rural areas may be emphasized, whereas the areas representing cities are very small, making interpretation of spatial patterns very difficult. The distortion of a cartogram accounts for the population density, preventing it from obscuring the spatial patterns (Levison and Haddon Jr 1965). The spatial transformation of map regions relative to the data emphasizes the data distribution instead of land size (Kocmoud and House 1998). When visualizing population statistics, Dorling (2011) considers this equitable representation design 'more socially just', or honest (Dent 1972), giving due attention to all members of the population and reducing the visual impact of large areas with small populations (Walter 2001). Howe (1989) suggests that 'cancer occurs in people, not in geographical areas' and Griffin (1980) believe that spatial socio-economic data, like cancer rates, are best presented on a cartogram for urban areas as the population map base avoids allocating 'undue prominence' to rural areas. Jahan et al. (2018) encourage the use of cartograms to highlight small areas and uncover local-level inequalities.

The creation of cartograms was historically in the hands of professional cartographers (Kraak 2017). Early approaches including John Hunter and Jonathan Young (1968) and Durham's wooden tile method, Skoda and Robertson's (1972) steel ball bearing approach and Tobler's (1973) computer programs (Dorling 2011). Geographical information systems allowed map users, and researchers to implement their own cartograms, but these systems are utilized depending on 'the effectiveness, efficiency, and satisfaction of the map products (Nielsen 1994), (Kraak 2017). Howe (1989) discusses the impact of electronic computer-assisted techniques.

There are many alternatives to consider, the intended audience of the map, and its purpose are key points in cartogram use and creation. Dorling (2011) reiterates: 'There is no "best" cartogram or method of creating cartograms just as there is no "best" map' (Monmonier and Schnell, 1988). Nusrat and Kobourov (2016) provided a framework to investigate implementations of the many algorithms presented, and the "statistical accuracy, geographical accuracy, and topological accuracy".

Table 2: Maps used to present statistics for the United States of America. Each state has been colored according to the average age-adjusted rate of incidence for lung and bronchus for females and males in the United States 2012-2016.

| Map display | Details |
|---|---|
| a. contiguous cartogram | Each stated shape has been distorted according to the population of the state in 2015. The state of California has become much larger due to it's large population density. This draws attention to the densely populated North East region, and detracts from the less populated Mid West. |

| Map display | Details |
|---|---|
| b. Non - Contiguous | The geographic shape of the states has been maintained, but the size has altered according to the population of the state in 2015. The state of California has remained closer to it's original size than it's surrounding states. The North East states have remained closer to their geographical size, in the case of Massachusetts and Connecticut. This draws attention to the densely populated North East region, and the sparse Mid West. |
| c. Dorling | The states have been represented by a circle, but the size was determined by the population of the state in 2015. The North East states remain closer to their neighbors, and may be displaced from their geographic location. The sparsity of the population in the Mid West is highlight by the distance between the circles, located at the geographic centroids. |
| d. Hexagon Tessellation | Each state is represented by a hexagon of equal size. The neighbouring states are easily contrasted, however the north east regions has been displaced from their geographic location. The sparsity of the population in the Mid West is highlight by the light yellow color, the age adjusted rate in Kentucky is the darkest and its neighbors are similar. |

### 4.2.1 Contiguous

A contiguous cartogram maintains connectivity of the map regions while areas are altered according to a statistic. This transformation often occurs at the expense of the shape of areas (Kocmoud and House 1998, @NAC, @TAAM). From a computer graphics perspective, Min Ouyang and Revesz (2000) explain the application of 'map deformation' to account for the value assigned to each area, they provide three methods for creating value-by-area cartograms. Examples include Tobler's Pseudo-Cartogram Method, Dorling's Cellular Automaton Method (2011), Radial Expansion Method of Selvin et al., Rubber Sheet Method of Dougenik et al., Gusein-Zade and Tikunov's Line Integral Method, Constraint-Based Method (Kocmoud and House) (1998).

An intentional goal when creating the 1966 Census population cartogram for Canada was to maintain contiguity, while attempting to keep the actual shape of places. The end result was a 'very accurate isodemo-graphic map of Canada'. This intentional design goal coincided with the rising interest in urban geography and presentation of social statistics. Figure 3 a shows a population contiguous cartogram of the United States. All states are visible and the shape of the United States overall is still recognizable. In contrast, Figure 4 a shows an Australian contiguous cartogram also based on population. The south east is completely distorted, areas with low population are still large very the map, as their initial area was used to determine the area allocated in the contiguous cartogram.

To be able to recognise the significant changes, a reader will usually have to know the initial geography to find the differences in the new cartogram layout (Olson 1976). Tobler's Conformal mapping means to preserve angles locally so that the shapes of small areas on a traditional map and a cartogram would be similar. Kocmoud and House (1998) presents this issue as conflicting tasks or aims, to adjust region sizes and retain region shapes.

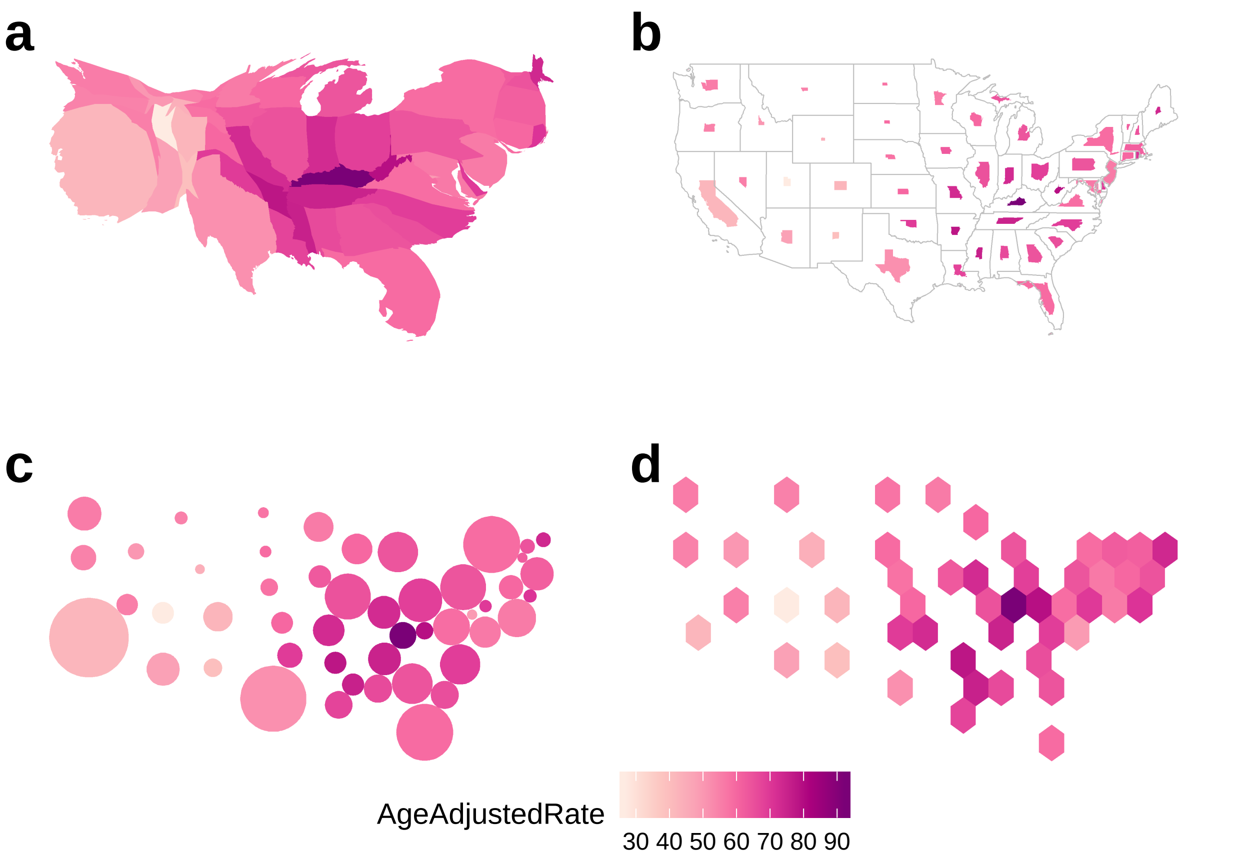


Figure 3: Common alternatives to maps, showing the same information as in Figure efchoroCRS for the United States of America: (a) contiguous cartogram, (b) non-contiguous, shape-preserved cartogram, (c) Dorling cartogram (non-contiguous), (d) hexmap (non-contiguous). In (a) - (c) the state has been resized, and reshaped, to match the 2015 population of the state. This provides a better sense of the extent of disease relative to the population in the country, and can help alleviate losing information about physically small but population dense states. In the hexmap (d) each state is given equal size, and thus equal emphasis.

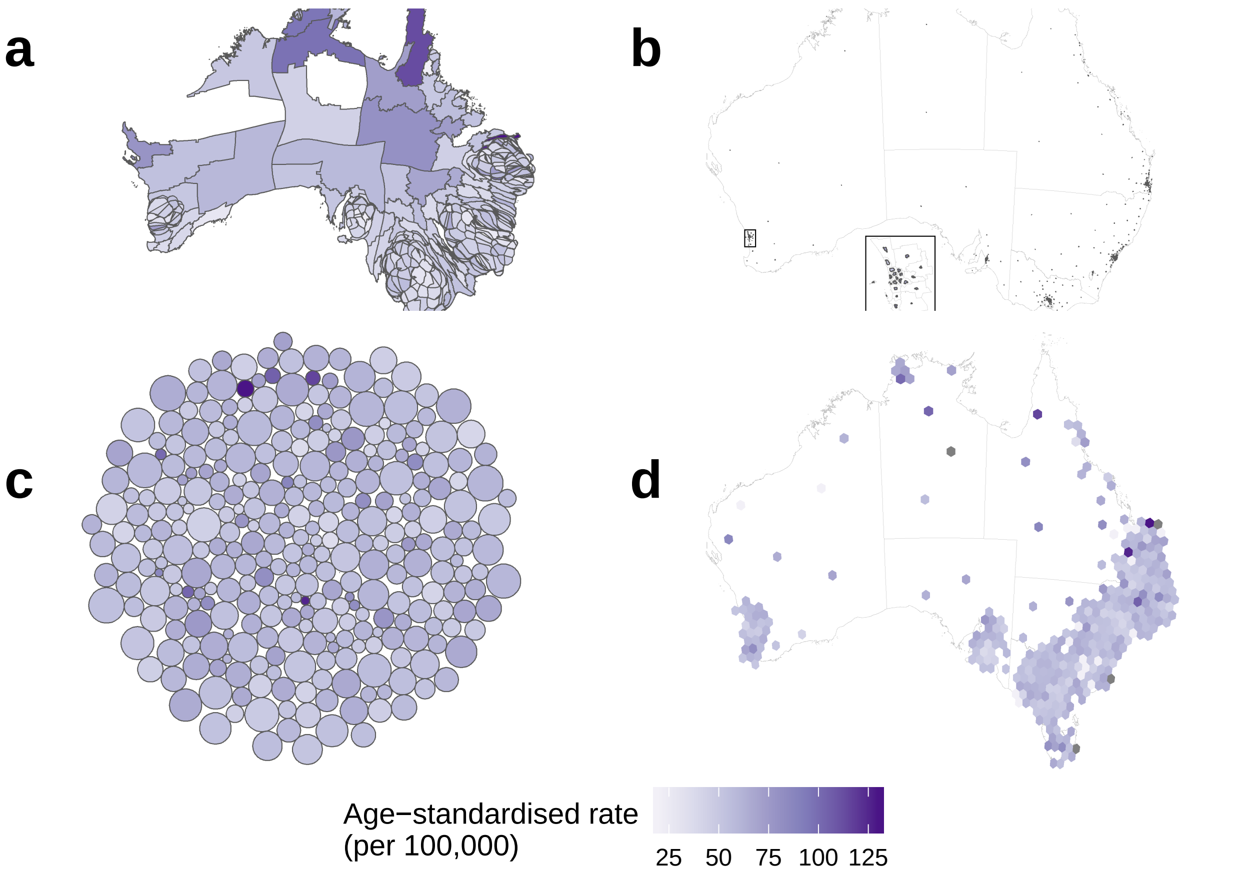


Figure 4: Alternative maps of the United States of America. Each state has been colored according to the average age-adjusted rate of incidence for lung and bronchus for females and males in the United States 2012-2016. Each stated shape has been distorted according to the population of the state in 2015. The state of California has become much larger due to it's large population density. This draws attention to the densely populated North East region, and detracts from the less populated Mid West.

### 4.2.2 Non-Contiguous

Non-Contiguous cartograms succeed in maintaining the shape of the areas presented. Each area stays in a similar position to their location on a choropleth map. The choropleth map base is often also plotted as a comparison point to highlight the change in area. The addition is the gap between areas, created as each individual area shrinks or grows according the associated value of the statistic. Olson (1976) discusses creation of these maps, the significance of the empty areas left between the geographic boundaries and the new shape, and the 'degree of difference from the original map that is the real message' of these displays.

As the trade-off regarding boundaries approaches simplicity, the distortion of region shapes on the contiguous cartogram presents an additional hurdle to visual recognition and this hurdle is not only eliminated on the non-contiguous cartogram but is replaced by the meaningful empty-space property (Keim et al. 2002, @NAC). The shapes are valuable for recognition and allows users to orient themselves on the display. Map creators can efficiently communicate with this kind of map by keeping the outlines or particular elements of the original in the new shape (Dent 1972). The scale of the areas does not impact on the shape recognition. However it may impact on the visibility of all areas if small areas expand beyond their boundaries. Figure 3 b shows small overlaps in the north east region of the United States. These overlaps are also shown in the non-contiguous cartogram of Australia in Figure 4 b.

### 4.2.3 Dorling

Daniel Dorling presented an alternative display engineers to highlight the spatial distribution and neighbourhood relationships without complex distortions of borders and boundaries. This approach opposes preserving the intricate shape details and is founded in the simple question put forward by Daniel Dorling (2011):

> "If, for instance, it is desirable that areas on a map have boundaries which are as simple as possible, why not draw the areas as simple shapes in the first place?"

He acknowledged the sophistication of contiguous cartograms but critiqued their 'very complex shapes,' he answers this with his implementation of maps created using 'the simplest of all shapes'. Circular cartograms use the same simple shapes for every region represented, and resizes the shapes according to the statistic represented or the population for a base map. This familiar shape may be more effective for understanding the spatial distribution than contiguous cartograms, as the 'nonsense' shapes used have 'no meaning' after distortions are applied Dent (1972). To produce a compelling map, a gravity model is applied to avoid overlaps, and keep spatial relationships with neighbouring areas over many iterations. This implementation can work for up to 'one hundred thousand' areas. Figure 3 c shows the displacement of the north east region of the United States in the Dorling cartogram, the rest of the country is still recognizable. This displacement is shown to an extreme in 4 c the Dorling cartogram of Australia.

The groundwork for this approach had been laid in the mid 1930's by Raisz, and rectangular cartograms provide dramatic comparisons and are especially useful for correcting misconceptions communicated by geographic maps. Tobler (2004) quotes the official definition of Value-Area Cartograms, the simplistic displays which represent each area as a single rectangle, sized according to the value of the statistic. This rectangular display also allows for tiling, where geographic neighbors placed in suitable relative positions also share borders, however contiguity may be sacrificed (Monmonier 2005). Rectangular cartograms allow for bi-variate displays, population can be effectively communicated by the size of each rectangular, and a second variable can be communicated using color (Kreveld and Speckmann 2007).

A similar method, where each geographic area is represented using a square, tessellated to create a square grid. This method has been used by FiveThirtyEight, Bloomberg Business, The Guardian, The Washington Post, The New York Times and NPR. Each area is represented by a square of the same dimensions. Figure 3 d shows a hexagon tile map representing each state using a single hexagon.

Recommended criteria to contrast mapping methods include average cartographic error, and maximum cartographic error, correct adjacency, maximum aspect ratio, and suitable relative positions (Kreveld and Speckmann 2007). However, this does not consider the issues with actually producing rectangular cartograms. Algorithms for the creation of rectangular cartograms

### 4.3 Tile map

A tile map provides a tessellated display of consistent shapes. Each tile is usually one unit of measurement, this could be geographic regions such as states, or population based where one tile is used for a consistent measure of population.

A simple tile map presents the areas in a tessellated grid display, where geographic neighbors are found next to each other, with some necessary displacement employed for regions with more than four neighbors. These tiles may be labelled or colored to represent a value. Tile maps may be difficult to create, they are best created manually, with additional time and care required as the amount of geographic areas to include increases. Cano et al. (2015) define the term 'mosaic cartograms' for hexagonal tile displays, where the number of tiles for each area can be used to communicate the statistic of regions. The complexity of the boundaries can be adjusted in the resulting display, as the size of the tiles used allows a trade-off be made between boundary complexity and simplicity.

### 4.4 Geofaceting

Tile maps can be extended to allow more than just simple coloring of tiles to present data. Hafen (2018) formalizes the term geofaceting to describe a grid display, the arrangement of tiles to create a grid that mimics the geographic topology of the set of areas. Like tile maps, the arrangement of the areas can be reused for other spatial distribution visualizations. Geofaceting have the functionality of facets, often used to replicate visualizations for each subset of the data, in this case the data subsets are geographic. The amount of information able to be communicated has increased from one value per region in a tile map to one visualization, this is a more flexible display. Virtually any plot display can be fit into the tile representing the areas, allowing displays of multiple variables or values per geographic entity. As the amount of areas increase care must be taken to ensure legibility of the displays. Xie, Hofmann, and Cheng (2014) acknowledges that this display suffers when the areas have very irregular shapes and large size disparities, as manual creation of the grid is still required.

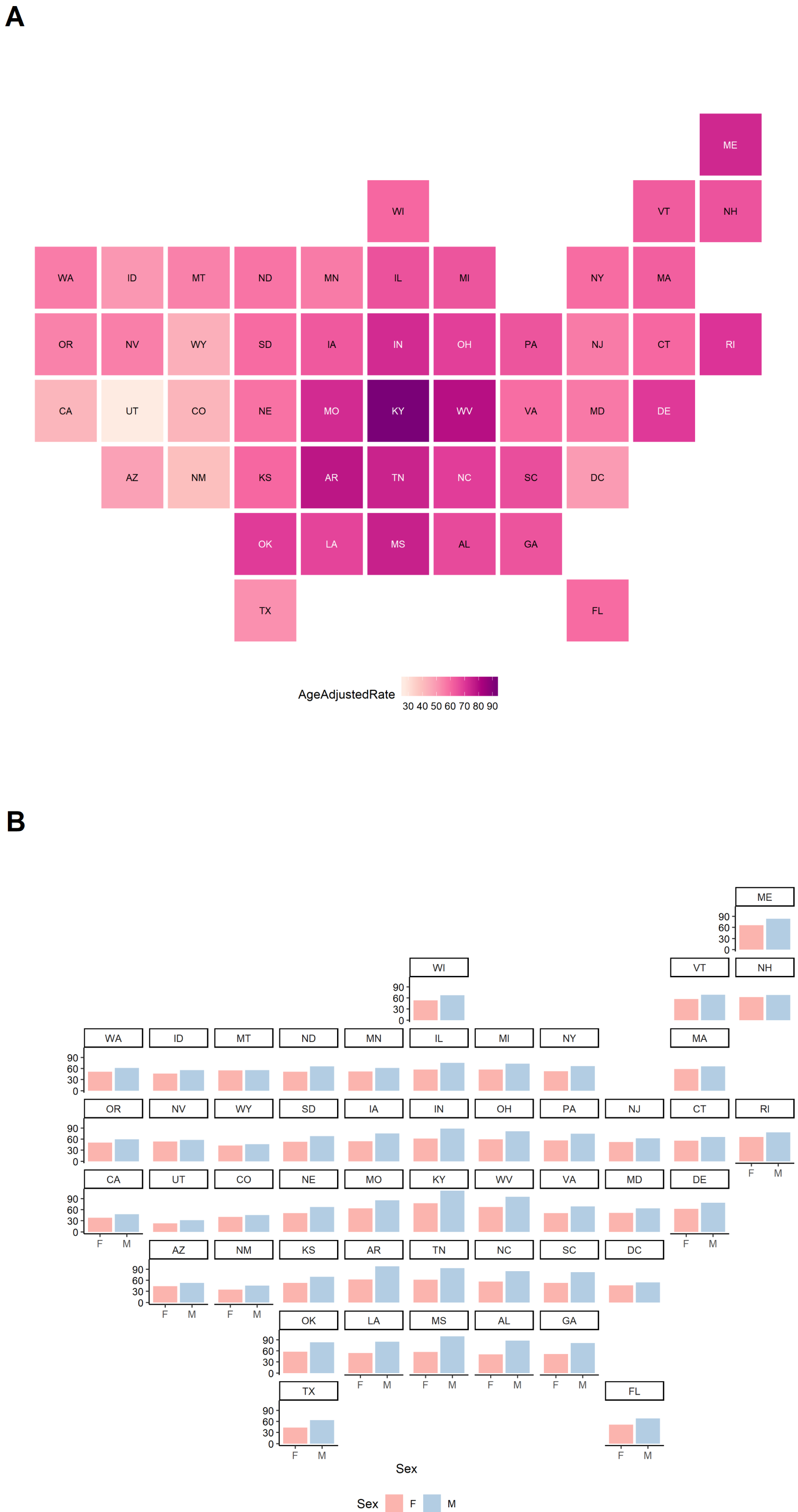


Figure 5: A tile map and geofaceted map of the Unites States of America. Each state in the tile map has been colored according to the average age-adjusted rate of incidence for lung and bronchus for females and males in the United States 2012-2016. Each state has been represented by a square, each square is the same size. The geofaceted map allows the rates for males and females to be

### 4.5 Multivariate displays

W., Carr, and Pearson (2015) present linked micromap plots to visually link geographic and statistical data, this serves as a solution to multi-dimensionality issues. This design is widely used in online atlas displays. Areas are grouped based on their value for one variable, and additional columns provide displays that contrast the areas in each group by other variables. The display juxtaposes of choropleth maps and statistical plots; it shows one map per group of the key separating variable, in a row with each additional statistical plot. Linked micromaps predominantly utilize the choropleth map for displays of spatial relationships, that are seen when spatial neighbours are alotted to the same group. It is one of several alternative displays that allow maps to become bivariate displays, commonly used to present both an estimate and the associated uncertainty.

Bivariate choropleth maps (Lucchesi and K. 2017) blend colour schemes to convey the intersection of categorised levels of an estimate and the associated uncertainty for each spatial area. Lucchesi and K. (2017) also suggests map pixelation, each region is broken into small pixels, the individual pixels are allocated values that reflect the uncertainty around the area's estimate by creating texture. This display can also be animated, with each frame produced by resampling the pixels. Areas with uncertain values will flicker more dramatically than areas with more certain values.

Solutions have been provided for bi-variate displays for movement data, containing Origin - Destination matrices and flow map representations (Yang et al. 2017). This method makes the maps and matrix connections clear, and easy to link geographic areas with link lines not crossing. Bar charts can also be added to these to show group totals.

## 5 Comparison of maps and cartograms

To choose an appropriate map display, the map creator must consider the intended user, and message the map will communicate. Choropleth displays utilize more traditional cartographic methods, they are usually true to the topography, displaying familiar boundaries of countries, states or administrative areas. Highlighting the geographic distribution, and differences between equal spatial units. When there is a relationship between the variable being mapped and geographical features, a choropleth map would pair this information visually. Choropleth methods require a decision to be made about the projection of the display, in this case aim to find a map projection that gives the least distorted representation of the geography (Tobler 1999). However a choropleth is limited to a uni-variate display, presenting a spatial distribution by filling each area with a color. These displays favor geographically large regions over small regions, though inset displays are often used to zoom in on a geographic area containing many small regions, this is seen in Figure 1 c of the state of Queensland, where the capital city Brisbane is highlighted by the inset map. Users can identify the areas relevant to them by the familiar boundaries, and can easily contrast the experience of geographic neighbors. The map is reusable and fit for many purposes.

Cartographers were historically the creators of geographic displays, and epidemiologists utilized the maps that had been prepared in advance. Cartograms with the population variable used to distortion the geographic map regions and boundaries are density equalising maps, that reduce the visual prominence of large, but low-population areas (Kronenfeld and Wong 2017). Cartogram displays incorporate the statistics and population into the design of the display, shifting and sometimes sacrificing familiar boundaries to draw attention to large outliers in the data space or population density. The unusual shapes are not longer familiar and may be difficult to compare. This can prevent effective comparison of the statistic. Cartograms also favor extreme statistics in the positive direction. Small extreme values are often lost due to the resulting small sized areas the do not draw user's attention. The change in the map based on one variable provides the opportunity for bi-variate displays, using the change in area for one variable, and color for another. The difference between the familiar map display helps highlight the impact of the disease on communities.

Alternative maps may replace the unexpected boundaries and shapes with familiar, simple shapes. This makes the spatial distribution the primary concept communicated by the display. Areas that are geographically close will maintain connectedness in some way, but the population and the statistic will dominate. Tile map displays may only allow one variable to be visualized if each region is given consistent map space, this occurs in geofaceting. Dorlings provide bi-variate displays as they use the size of the shape for one variable, and color for another. These alternatives are especially helpful for data aggregations where administrative boundaries break populations into groups. The experience of each community may be worth considering, as the experiences shared by the population within them may be similar due to the services and facilities they share. This display allows a more equitable view of each community, and does not minimise those that operate on a smaller geographic scale.

Faceting provides an opportunity to contrast the overall distribution of variables across a geographic space or collection of geographic areas. This display makes is difficult to compare specific regions.

Table 3: Mapping methods used to display cancer statistics

| | Choropleth | Contiguous | Non-contiguous | Dorling | Tile maps |
|---|---|---|---|---|---|
| Preserves shapes | Y | N | Y | N | N |
| Preserves neighbors | Y | Y | Y | S | S |
| Uniform use of shape | N | N | N | Y | Y |
| Bi-variate display[1] | N | Y | Y | Y | Y |

Creating maps of diseases now involves more decisions to be made by map makers, rather than cartographers. Technology has played an enormous role in increasing the opportunities for map makers. The computation and graphics power have made creation, alteration and interactivity possible. as these options have expanded and it is the objectives of the investigator that will drive the choices. Bell et al. (2006) and Moore and Carpenter (1999) have provided suggestions and comments to help map creators best communicate their health data and spatial analyses.

# 6 User interaction with maps

Interactivity and animation enhance online atlases by enabling additional information to be presented in the same display. Implementing movement and allowing for changes in the display invites the viewer to engage with the display, and the information presented.

Interactive methods involve direct manipulation of web displays, keeping a continuous view of the map as it changes due to the physical actions of users. The manipulation of the web display may

[1] A bi-variate display refers to a map that allows two variables to be plotted. One variable via the size of the area, and another by filling the area with a color.

occur by directly altering the display, clicking the controls or command-line prompts (Xie, Hofmann, and Cheng 2014). Pedersen (2018) explains that animations show the transitions between frames in a linear, fixed narrative to passive users, communicating a message by capturing and directing their attention. This may require direct user interaction on the display.

Web atlases have employed interactivity and animation to expand the communicative ability of a disease map. It is important to identify the primary questions that might be asked by the users of a map, and allow the display to change to best answer these questions. Bell et al. (2006) provide three questions that are generally asked when using a disease map:

- What is the mortality rate in a certain area?
- Are there geographic trends in the data, or regions of unusually high or low rates?
- Is the lung cancer mortality pattern similar to the pattern of smoking prevalence shown in a companion map?

A single map display may not effectively answer all three at the same time. This is especially true when some areas are small and difficult to identify clearly in a single display or resolution. However, with the ability to interact with a map, or animate the display, map creators can specify and answer these questions by directing the attention of a map user to the areas of interest.

## 6.1 Interaction and animation in publicly available atlases

Interactivity can be implemented within modern mapping methods to incorporate additional information and complexity without overloading the user. Adding too much information to a map may create cognitive overload (McGranaghan 1993) in which the user reaches an information threshold, beyond which will not be able to make sense of the information. Complexity and density of visualization methods can overwhelm novice decision makers, however experts are able to use the detail more readily when making-decisions (Cliburn et al. 2002). Supplementary information can be incorporated into online atlases without cluttering the display. Interactive design features found in online cancer maps include tool tip features, drop down menus, data selection, zooming and panning. Allowing users to explore the map for more information and providing flexibility in the display (Monmonier 2018). The use of these supports were found in various online cancer maps identified by Roberts (2019). Additionally, interactivity allows the user to toggle between different variables, map views or multiple realizations of possible future scenarios (Goodchild, Buttenfield, and Wood 1994). Thus providing additional mechanisms for the users comprehension as well as the uncertainty of the available information (MacEachren (1992); Van der Wel, Hootsmans, and Ormeling (1994)). For example, changing the population, age or other demographic variables. The controls for basic interactive features are often placed outside of the plot space (Pedersen 2018), the map image is updated/replaced as the user interacts with the controls. Some more advanced interactions include direct interactions with the plot via the use of overlaid tool tip features, very few cancer atlases involve these more complex selection interactive capabilities.

Relationships between spatial areas and diseases can be explored with sophistication in nontraditional but still 'cognitively accessible' ways (Carr, Wallin, and Carr 2000). This helps users to understand and interpret the spatial distribution presented, as well as validate, explain or explore the presented statistics and their relationships to each other and/or their underlying spatial distribution. The interactive features of the publicly available maps identified by Roberts (2019) allow the exploration of geographic hierarchies, population distribution, statistical uncertainty, demographics and socio-economic indicators. The affects of these actions are rapid, incremental, and reversible Perin (2014).

Where the needs of the audience is changeable and is the priority, the map creator can allow interactivity for map users to explore a data set through dynamic interactions that allow inspection of the data from many views (Dang, North, and Shneiderman 2001). Interactivity and animation provide an alternative to the traditional approach of a single map for a single use, and enable the map to become more exploratory. Interactivity allows users to drive their own exploration, controlling a display through variable controls, links, tool tips, and data selection tools (Pedersen 2018). Tool tips allow users to obtain additional information regarding specific geographic areas.

Development and adoption of this technology has allowed map users more control of the displayed message. Figure **??** presents forms of low level interactions with a user interface, these are the control menus for several online cancer atlases. The user interface for these online atlases allow combinations of cancer type, gender, and other factors to be shown on the map display and the user to move between and/or compare different views of these variables. These controls may also change the geographies shown, such as continents or countries, states or counties. These displays can show familiar shapes for users to orient themselves on the map, then zoom in to explore densely populated communities. Interactions with online atlases can also change the amount of detail shown.

## 6.2 Under-utilised techniques

It is uncommon to see linked displays utilized in online atlases, but there are several frameworks that support interactive displays of multivariate data. Carr, Wallin, and Carr (2000) suggested LM plots as a solution to linking cartography and statistical graphics. Linked displays allow for direct manipulation, Xie, Hofmann, and Cheng (2014) uses Swayne and Klinke's definition of high level interaction that occurs via data selection or brushing tools that highlight subsets of graphical elements such as polygons or points. Selections can be made statistically through interactions which allow subsets to be selected from statistical displays, such as tables of data. Monmonier implemented linking choropleths with statistical plots, adding a geographic component to scatter plots (Dykes and Unwin 1998), the use of link allows spatial dimensions to be reserved for geographic locations of areas. The statistical variation of additional variables can be presented in other appropriate displays. Low level interaction is more common, this occurs via the user interface or the map. Interacting with the map returns values or information used in the graphical display without referring to, of changing, the original data source of the map.

Animations communicate a story, showcasing a critical region by defining and progressing through a small set of keyframes, or waypoints (Esri Australia Pty. Ltd. 2019). Pedersen (2018) suggests the use of moving objects demands attention by taking advantage of the behaviour of the visual cognition system. They can be produced to not tolerate or require user input, creating useful, directed descriptions of processes, such as highlighting a particular region or time period (Monmonier 2018). Animations ensure all users follow the same interpretive path but creators cannot control the level of engagement.

Animations can be implemented to allow users varying degrees of control. Figure **??** a gives an example of how time can be included through animating displays. This is a basic form of interactivity with animation allowing users to play, pause and rewind films, where the film is intended to be watched in one direction but the position in the animation can be controlled. Bell et al. (2006) provide weather maps as a thoroughly developed example of animation of spatial displays to communicate information to the general public. The movement of a weather system will follow a forecasted path, all map users can follow the animated path of the weather system across the

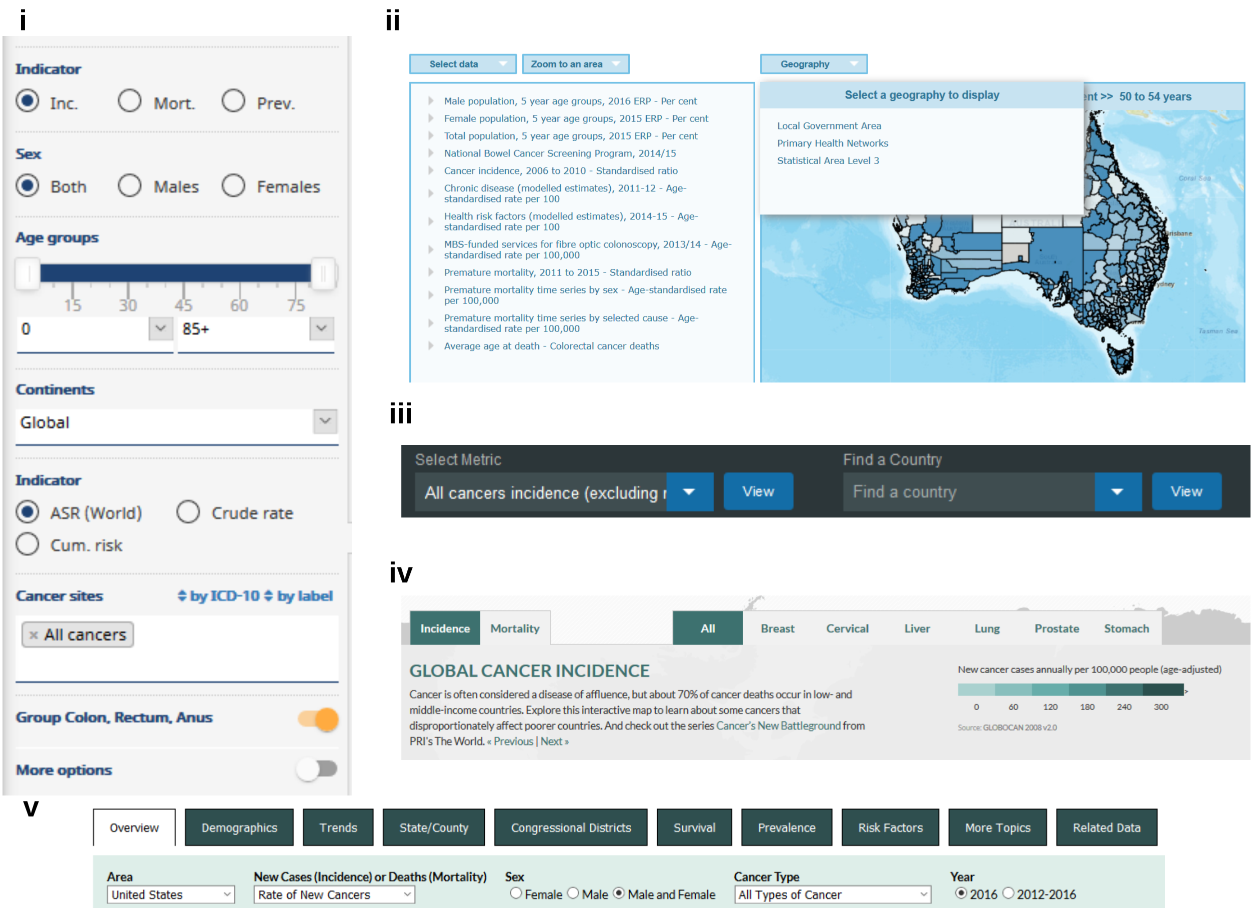


Figure 6: Interactive controls of displays in publicly available choropleth cancer maps. a. Globocan 2018: Cancer Today(Controls for indicator, sex, age groups, continents, and cancer types) , b. Bowel Cancer Australia Atlas (Controls for indicator, age, year, and geographical areas), c. The Cancer Atlas (Controls for predetermined combinations of cancer type or risk factors, mortality or incidence, and gender) d. Global Cancer Map (Controls for indicator, and cancer type), e. United States Cancer Statistics: Data Visualizations (Controls for demographics, trends, geography, indicators, risk factors, sex, cancer type, and year).

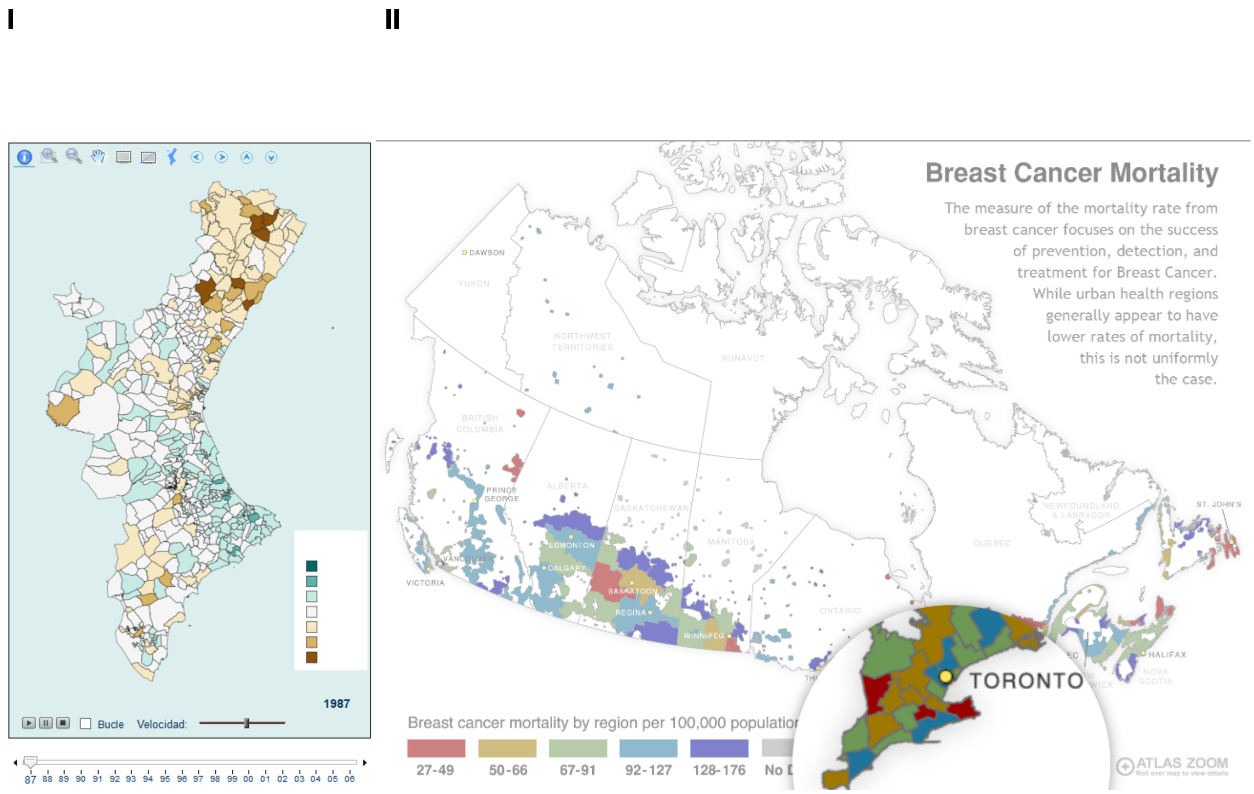


Figure 7: Interactivity and animations in publicly available choropleth cancer maps. a. Map of Cancer Mortality Rates in Spain(Controls for indicator, sex, age groups, continents, and cancer types) , b. Breast Cancer Mortality in Canada (Controls for indicator, age, year, and geographical areas)

geography over a specified period of time. Where the message is most important, static or animated graphics allow control over the display and interpretation. User interaction more allow clicks to progress between stages of more complex animations. The Australian Cancer Atlas (2018) provides tours which change the display to draw user's attention to areas on the map that are relevant to the story. This implementation of animation gives users tools to plan their own exploration.

# 7 Conclusions

This paper provides an overview of mapping practices as used for cancer atlases, and new approaches that could be adopted. The conventional approach is the choropleth map, and it is widely used. When there are small areas, as occurs in Australia where the population is concentrated on the coast, the information about cancer can be lost, and alternatives are needed. Making an inset can clarify congested regions but this breaks the viewers attention, because they need to shift focus from the map to the inset, and if there are many congested areas, many insets would be needed. The map alternatives, like cartograms and their variations, can be useful to allow the spatial distribution of cancer data to be digested.

There are many different statistics that are commonly used for display. The most basic is incidence rate, which is easy to understand. It is common to see relative rates in many maps which measure how far a region is above or below the average. The purpose of using a relative rate is, perhaps the desire to pinpoint the areas that need attention because they have higher than expected rates. However, we lose the incidence rate information and thus interpretability. A region might be much higher than average, but it may not be close to a health concern, because all regions have low incidence.

# 8 Acknowledgements

We would like to acknowledge Dr Earl Duncan (Research Associate at ARC Centre of Excellence for Mathematical & Statistical Frontiers, QUT). His time and effort given to edit and comment on this paper was invaluable.

We thank Dr Susanna Cramb (Spatial Modeller, Cancer Council Queensland) and Dr Peter Baade (Senior Research Fellow, Cancer Council Queensland) for providing the opportunity to cover this exploration of disease mapping methods in writing. This literature review would not be possible without the opportunity provided by Queensland University of Technology, and Cancer Council Queensland, and the roles of Professor Kerrie Mengersen (Professor of Statistics, Science and Engineering Faculty, QUT) and Dr Earl Duncan in contributing to the Australian Cancer Atlas and supervision of Stephanie's Research Masters.

It was in the development of this online cancer atlas that methods for disease map displays, and visual communication strategies were explored. We are thankful for the opportunity to write about these visualizations and the situations in which they are appropriate.

Several R (R Core Team 2019) packages were used to produce this paper. For data analysis the tidyverse package (Wickham 2017) provided many useful functions. ggplot2 package (Wickham 2016) was used to create maps, with RColorBrewer package (Neuwirth 2014) providing additional color palettes and ggthemes (Arnold 2019) package providing themes. The png package (Urbanek

2013) was used to access png images taken from online web atlases, cowplot package (Wilke 2019) was used to arrange these plots into grouped images.

The following packages were used to create transformations from the sf (Pebesma 2018) geographical shapes of the states of America (Bivand, Nowosad, and Lovelace 2019) and the Australian Statistical Areas (Level 3): The cartogram (Jeworutzki 2018) package creates contiguous, non-contiguous and Dorling cartograms. Hexagon tile map displays were created using the sugarbag package (Kobakian and Cook 2019).

# 9 References

Arnold, Jeffrey B. 2019. *Ggthemes: Extra Themes, Scales and Geoms for 'Ggplot2'*. https://CRAN.R-project.org/package=ggthemes.

Atlas, Australian Cancer. 2018. "Australian Cancer Atlas." Cancer Council Queensland, Queensland University of Technology, and Cooperative Research Centre for Spatial Information. 2018. https://atlas.cancer.org.au.

Bell, B Sue, Richard E Hoskins, Linda Williams Pickle, and Daniel Wartenberg. 2006. "Current Practices in Spatial Analysis of Cancer Data: Mapping Health Statistics to Inform Policymakers and the Public." *International Journal of Health Geographics* 5: 49. https://doi.org/10.1186/1476-072X-5-49.

Berry, Brian J. L., Richard L. Morrill, and Waldo R. Tobler. 1964. "GEOGRAPHIC Ordering of Information: NEW Opportunities" 16 (4): 39–44. https://doi.org/10.1111/j.0033-0124.1964.039_q.x.

Bivand, Roger, Jakub Nowosad, and Robin Lovelace. 2019. *SpData: Datasets for Spatial Analysis*. https://CRAN.R-project.org/package=spData.

Bowel Cancer Australia. 2016. "Bowel Cancer Australia Atlas." 2016. http://www.bowelcanceratlas.org/.

Brewster, Mark B., and S. V. Subramanian. 2010. "Cartographic Insights into the Burden of Mortality in the United Kingdom: A Review of 'The Grim Reaper's Road Map'." Journal Article. *International Journal of Epidemiology* 39 (4): 1120–2. https://doi.org/10.1093/ije/dyp395.

Cano, R. G., K. Buchin, T. Castermans, A. Pieterse, W. Sonke, and B. Speckmann. 2015. "Mosaic Drawings and Cartograms." *Computer Graphics Forum* 34 (3): 361–70.

Carr, D. B., J. F. Wallin, and D. A. Carr. 2000. "Two New Templates for Epidemiology Applications: Linkedmicromap Plots and Conditioned Choropleth Maps." *Statistics in Medicine* 19: 2521–38.

Cliburn, Daniel C, Johannes J Feddema, James R Miller, and Terry A Slocum. 2002. "Design and Evaluation of a Decision Support System in a Water Balance Application." *Computers & Graphics* 26 (6): 931–49.

Cramb, Susanna, Kerrie Mengersen, and Peter Baade. 2011. "Developing the Atlas of Cancer in Queensland: Methodological Issues." *International Journal of Health Geographics* 10 (January): 9. https://doi.org/10.1186/1476-072X-10-9.

Dang, G., C. North, and B. Shneiderman. 2001. “Dynamic Queries and Brushing on Choropleth Maps.” In *Proceedings Fifth International Conference on Information Visualisation*, 757–64. https://doi.org/10.1109/IV.2001.942141.

Dent, Borden D. 1972. “A Note on the Importance of Shape in Cartogram Communication.” *Journal of Geography* 71 (7): 393–401. https://doi.org/10.1080/00221347208981697.

d’Onofrio, Alberto, Chiara Mazzetta, Chris Robertson, Michel Smans, Peter Boyle, and Mathieu Boniol. 2016. “Maps and Atlases of Cancer Mortality: A Review of a Useful Tool to Trigger New Questions.” Journal Article. *Ecancermedicalscience* 10: 670–70. https://doi.org/10.3332/ecancer.2016.670.

Dorling, Daniel. 2011. “Area Cartograms: Their Use and Creation.” In *Concepts and Techniques in Modern Geography (CATMOG)*, 59:252–60. https://doi.org/10.1002/9780470979587.ch33.

Dougenik, James A., Nicholas R. Chrisman, and Duane R. Niemeyer. 1985. “AN Algorithm to Construct Continuous Area Cartograms.” *The Professional Geographer* 37 (1): 75–81. https://doi.org/10.1111/j.0033-0124.1985.00075.x.

Dykes, Jason, and David Unwin. 1998. “Maps of the Census: A Rough Guide.” *Case Studies of Visualization in the Social Sciences, Technical Report 43* 43 (January).

El Pais. 2014. “Map of Cancer Mortality Rates in Spain.” 2014. http://elpais.com/elpais/2014/10/06/media/1412612722_141933.html.

Emperial College London - Small Area Health Statistics Unit. 2010. “The Environmental and Health Atlas of England and Wales.” 2010. http://www.envhealthatlas.co.uk/eha/Breast/.

Esri Australia Pty. Ltd. 2019. “Animation Basics.” Environmental Systems Research Institute. 2019. https://pro.arcgis.com/en/pro-app/help/mapping/animation/overview-of-animation.htm.

Exeter, Daniel J. 2017. “Spatial Epidemiology.” Journal Article, 1–4. https://doi.org/10.1002/9781118786352.wbieg0283.

Ferlay J, Lam F, Ervik M. 2018. “Global Cancer Observatory: Cancer Today.” https://gco.iarc.fr/today.

Goodchild, M, B Buttenfield, and J Wood. 1994. “On Introduction to Visualizing Data Validity.” *Visualization in Geographical Information Systems*, 141–49.

Griffin, T. L.C. 1980. “Cartographic Transformation of the Thematic Map Base.” *Cartography* 11 (3): 163–74. https://doi.org/10.1080/00690805.1980.10438102.

Group, U. S. Cancer Statistics Working. 2019. “U.S. Cancer Statistics Data Visualizations Tool, Based on November 2018 Submission Data (1999–2016).” www.cdc.gov/cancer/dataviz.

Hafen, Ryan. 2018. “Introducing Geofacet.” *Ryanhafen*. https://ryanhafen.com/blog/geofacet/.

Howe, G. M. 1989. “Historical Evolution of Disease Mapping in General and Specifically of Cancer Mapping.” In *Cancer Mapping*, edited by Peter Boyle, Calum S. Muir, and Ekkehard Grundmann, 1–21. Berlin, Heidelberg: Springer Berlin Heidelberg.

Jahan, Farzana, Earl Duncan, Susanna Cramb, Peter Baade, and Kerrie Mengersen. 2018. “Making More of Spatial Maps: A Bayesian Meta-Analysis Approach.” In.

Jeworutzki, Sebastian. 2018. *Cartogram: Create Cartograms with R*. https://CRAN.R-project.org/package=cartogram.

Keim, D. A, S. C North, C Panse, and J Schneidewind. 2002. "Efficient Cartogram Generation: A Comparison." In *IEEE Symposium on Information Visualization, 2002. INFOVIS 2002*, 2002-:33–36. IEEE.

Kobakian, Stephanie, and Dianne Cook. 2019. *Sugarbag: Create Tessellated Hexagon Maps*. https://CRAN.R-project.org/package=sugarbag.

Kocmoud, Christopher J., and Donald H. House. 1998. "A Constraint-Based Approach to Constructing Continuous Cartograms." In.

Kraak, M. J. 2017. "Cartographic Design." In *The International Encyclopedia of Geography : People, the Earth, Environment, and Technology*, edited by D. Richardson, 1–16. United States: Wiley.

Kreveld, Marc van, and Bettina Speckmann. 2007. "On Rectangular Cartograms." *Computational Geometry* 37 (3): 175–87. https://doi.org/https://doi.org/10.1016/j.comgeo.2006.06.002.

Kronenfeld, Barry J., and David W. S. Wong. 2017. "Visualizing Statistical Significance of Disease Clusters Using Cartograms." Journal Article. *International Journal of Health Geographics* 16 (1): 19. https://doi.org/10.1186/s12942-017-0093-9.

Levison, M. E., and W. Haddon Jr. 1965. "THE Area Adjusted Map. AN Epidemiologic Device." Journal Article. *Public Health Reports* 80: 55–59.

Lucchesi, L. R., and Wikle C. K. 2017. "Visualizing Uncertainty in Areal Data with Bivariate Choropleth Maps, Map Pixelation and Glyph Rotation." *Stat*. https://doi.org/https://doi.org/10.1002/sta4.150.

MacEachren, Alan M. 1992. "Visualizing Uncertain Information." *Cartographic Perspectives*, no. 13: 10–19.

McGranaghan, Matthew. 1993. "A Cartographic View of Spatial Data Quality." *Cartographica: The International Journal for Geographic Information and Geovisualization* 30 (2-3): 8–19.

Min Ouyang, and P. Revesz. 2000. "Algorithms for Cartogram Animation." In *Proceedings 2000 International Database Engineering and Applications Symposium (Cat. No.PR00789)*, 231–35. https://doi.org/10.1109/IDEAS.2000.880581.

Monmonier, M. 2018. *How to Lie with Maps, Third Edition*. University of Chicago Press. https://books.google.com.au/books?id=MwdRDwAAQBAJ.

Monmonier, Mark. 2005. "Cartography: Distortions, World-Views and Creative Solutions." *Progress in Human Geography* 29 (2): 217–24. https://doi.org/10.1191/0309132505ph540pr.

Moore, Dale A., and Tim E. Carpenter. 1999. "Spatial Analytical Methods and Geographic Information Systems: Use in Health Research and Epidemiology." *Epidemiologic Reviews* 21 (2): 143–61. https://doi.org/10.1093/oxfordjournals.epirev.a017993.

Neuwirth, Erich. 2014. *RColorBrewer: ColorBrewer Palettes*. https://CRAN.R-project.org/package=RColorBrewer.

Northern Ireland Cancer Registry. 2011. "All-Ireland Cancer Atlas 1995-2007." 2011. http://www.ncri.ie/publications/cancer-atlases.

Nusrat, Sabrina, and Stephen G. Kobourov. 2016. "The State of the Art in Cartograms." *CoRR* abs/1605.08485. http://arxiv.org/abs/1605.08485.

Olson, Judy M. 1976. "NONCONTIGUOUS Area Cartograms." *The Professional Geographer* 28 (4): 371–80. https://doi.org/10.1111/j.0033-0124.1976.00371.x.

Pebesma, Edzer. 2018. "Simple Features for R: Standardized Support for Spatial Vector Data." *The R Journal* 10 (1): 439–46. https://doi.org/10.32614/RJ-2018-009.

Pedersen, Thomas Lin. 2018. "The Grammar of Animation." 2018. https://youtu.be/21ZWDrTukEs.

Pediatric Oncology Group of Ontario. 2015. "Incidence Rate of Childhood Cancers, Atlas of Childhood Cancer in Ontario (1985-2004)." 2015. https://www.pogo.ca/wp-content/uploads/2015/02/POGO_CC-Atlas-3-Incidence_Feb-2015.pdf.

Perin, Charles. 2014. "Direct Manipulation for Information Visualization." Theses, Université Paris Sud - Paris XI. https://hal.inria.fr/tel-01096366.

Queensland Cancer Registry. 2011. "The Atlas of Cancer in Qld, 1998 - 2007." 2011. https://cancerqld.org.au/research/queensland-cancer-statistics/queensland-cancer-atlas/.

Raisz, Erwin. 1963. "Rectangular Statistical Cartograms of the World." Journal Article. *Journal of Geography* 35 (1): 8–10. https://doi.org/10.1080/00221343608987880.

R Core Team. 2019. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Roberts, Jessie. 2019. "Communication of Statistical Uncertainty to Non-Expert Audiences." PhD thesis, Queensland University of Technology.

Skowronnek, Alsino. 2016. "Beyond Choropleth Maps – a Review of Techniques to Visualize Quantitative Areal Geodata." INFOVIS READING GROUP WS 2015/16. 2016. https://alsino.io/static/papers/BeyondChoropleths_AlsinoSkowronnek.pdf.

Tobler, W. 1999. *Unusual Map Projections*. Association of American Geographers.

Tobler, Waldo. 1963. "Geographic Area and Map Projections," January, 59–78. https://www.jstor.org/stable/212809.

———. 2004. "Thirty Five Years of Computer Cartograms." *Annals of the Association of American Geographers* 94 (1): 58–73.

Tufte, Edward R. 1990. *Envisioning Information*. Graphics Press.

Urbanek, Simon. 2013. *Png: Read and Write Png Images*. https://CRAN.R-project.org/package=png.

U.S. Cancer Statistics Working Group. 2018. "U.S. Cancer Statistics Data Visualizations Tool, based on November 2018 submission data (1999-2016)." www.cdc.gov/cancer/dataviz; U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute.

U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute - Cancer Statistics Working Group. 2019. "U.S. Cancer Statistics Data

Visualizations Tool, Based on November 2018 Submission Data (1999-2016).” 2019. www.cdc.gov/cancer/dataviz.

Van der Wel, Frans JM, Rob M Hootsmans, and Ferjan Ormeling. 1994. “Visualization of Data Quality.” In *Modern Cartography Series*, 2:313–31. Elsevier.

W., Pickle L., D. B. Carr, and J. B. Pearson. 2015. “MicromapST: Exploring and Communicating Geospatial Patterns in Us State Data.” *Journal of Statistical Software* 63 (3): 1–25. https://doi.org/10.18637/jss.v063.i03.

Walter, S. D. 2001. “Disease Mapping: A Historical Perspective.” Book.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

———. 2017. *Tidyverse: Easily Install and Load the 'Tidyverse'*. https://CRAN.R-project.org/package=tidyverse.

Wilke, Claus O. 2019. *Cowplot: Streamlined Plot Theme and Plot Annotations for 'Ggplot2'*. https://CRAN.R-project.org/package=cowplot.

World Health Organization's International Agency for Research on Cancer. 2018. “Globocan 2012: Estimated Cancer Incidence, Mortality and Prevalence.” 2018. http://globocan.iarc.fr/Pages/Map.aspx.

Xie, Yihui, Heike Hofmann, and Xiaoyue Cheng. 2014. “Reactive Programming for Interactive Graphics.” *Statistical Science* 29 (2): 201–13. http://www.jstor.org/stable/43288470.

Yang, Y., T. Dwyer, S. Goodwin, and K. Marriott. 2017. “Many-to-Many Geographically-Embedded Flow Visualisation: An Evaluation.” *IEEE Transactions on Visualization and Computer Graphics* 23 (1): 411–20. https://doi.org/10.1109/TVCG.2016.2598885.