

Community Health Map: A Geospatial and Multivariate Data Visualization Tool for Public Health Datasets

Awalin Sopan, Angela Song-Ie Noh

Human-Computer Interaction Lab, Department of Computer Science, University of Maryland

Ginnah Lee, Paul Rosenfeld

Department of Electrical and Computer Engineering, University of Maryland

Sohit Karol

Department of Kinesiology, University of Maryland

Ben Shneiderman

Human-Computer Interaction Lab, Department of Computer Science, University of Maryland

Abstract

Trillions of dollars are spent each year on health care. The U.S. Department of Health and Human Services keeps track of a variety of health care indicators across the country, resulting in a large geospatially multivariate data set. Current visualization tools for such data sets make it difficult to make multivariate comparisons and show the geographic distribution of the selected variables at the same time. Community Health Map is a web application that enables users to visualize health care data in multivariate space as well as geospatially. It is designed to aid exploration of this huge data repository and deliver deep insights for policy makers, journalists, consumer groups and academic researchers. Users can visualize the geospatial distribution of a given variable on an interactive map, and compare two or more variables using charts and tables. By employing dynamic query filters, visualizations can be narrowed down to specific ranges and regions. Our presentation to policy makers and pilot usability evaluation suggest that the Community Health Map provides a comprehensible and powerful interface for policy makers to visualize health care quality, public health outcomes, and access to care in an effort to help them to make informed decisions about improving health care.

Keywords: Healthcare data, Geospatial multivariate visualization, Graphical user interface, Information visualization, Informed decision-making

Video demonstration of Community Health Map in 3 minutes: <http://www.youtube.com/watch?v=gLMjms5VzXQ>

1. Introduction

As the health care debate raged on in the Senate, there were numerous opinions about what is wrong with health care in the U.S. While everyone agrees that the cost of health care is exploding for both individuals and businesses, the reason why costs rise but quality doesn't necessarily follow suit appears to be a much more open question. This is further complicated by the fact that costs of health care are rising much quicker in some parts of the country than in others. Portions of the health care system are plagued by waste, inefficiency, and in some cases, fraud (DeMott [1997], Himmelstein et al. [2009], Meier and Homann [2009]). With such a large amount of money at stake, commentators

Email addresses: awalin@cs.umd.edu (Awalin Sopan), angela@cs.umd.edu (Angela Song-Ie Noh), ginnah@umd.edu (Ginnah Lee), prosenf1@umd.edu (Paul Rosenfeld), skaro1@umd.edu (Sohit Karol), ben@cs.umd.edu (Ben Shneiderman)

on all sides want to ensure that funds are used efficiently and effectively. Although the U.S. Department of Health and Human Services (HHS) keeps track of the health care indicator variables across the country, only a limited number of tools are available for exploration and visualization of this huge data repository (Fisher et al. [2009]) to aid the policy makers. Community Health Map is a web-based tool designed to visualize health care data. During development of this tool, HHS staff provided continuing feedback, along with the raw data used to create the visualizations. The Community Health Map is designed for users with experience in the health care industry and it does not require programming experience. Users should be able to achieve these goals:

1. Visualize health performance, access and quality indicators for their county.
2. Filter the regions by demographic characteristics (such as age, income, poverty or education). The filters also allow users to compare their zone of interest to the national average.
3. Produce visual results for presentations that would spur policy makers to take action to improve the health and quality of health care in their regions.

The contribution of the Community Health Map is that it enables a wider range of non-technical users, such as mayors, local health officials, journalists, and consumer groups, to dynamically explore the geospatial distribution of health data. The map uses easy to understand color coding and provides detailed information in a multivariate table and chart simultaneously. Users can compare map regions for an overview and then go to the table and chart for more detail.

The Open Government Directive has produced many initiatives to provide public access to many government datasets, which has triggered a dramatically increased interest in visualization tools to explore these datasets. A prime requirement has been that these visualization tools be powerful enough to enable significant discoveries, while easy enough to use by non-programmers with only moderate technical skills, such as policy makers, journalists, consumer groups, and diverse academic researchers.

This paper reviews existing tools to visualize health care data, along with their capabilities and limitations. Next, it briefly describes the structure of raw data and design inspiration. This is followed by the interface description, and by sample insights obtained with the Community Health Map. Finally, it details the feedback obtained from a usability evaluation conducted with the domain experts, both from the Department of HHS as well as three professionals in the health care industry. The paper concludes with the future directions for this tool.

2. Related Work

Data visualization has been used as an important tool to gain insights into health care data sets, which are typically multivariate, discrete and at different granularity levels. C. Tominski [2008] draws attention to the importance of multivariate data representation with respect to time and space. Showing geographically distributed health related data in a choropleth map using sequential color scheme has been proven very useful for data driven knowledge discovery (Maceachren et al. [1998], EL and MJ [2004]). While designing visual analytics tools for healthcare data, it is crucial to maintain the simplicity as the targeted users need not have technical expertise. Gill-garcia and Pardo Gil-Garca and Pardo [2005] identified usability and ease of use as two notable technical challenges for successful e-government. Usability has become a stronger design consideration in government website design producing critical discussions (Fedorowicz and Dias [2010]) and extensive guidelines (Donker-Kuijter et al. [2010], National Cancer Institute [2010]). Previous works have employed different desktop applications designed for multivariate analysis to study health care datasets (Keahey [1998], Madigan and Curet [2006]). However, most of desktop tools for multivariate data visualization typically require users to spend a significant amount of time learning the tool. To have better impact of the available public data, access, usage and impact of the service are important factors for the end-users (Verdegem and Verleye [2009]) and for these purposes a webbased tool is a better fit. Unfortunately, there are a few web applications for health care data visualizations. Among these, most of such tools using geospatial visualizations only have the capability to compare a couple of variables at a given time (Fisher et al. [2009]).

One of the most popular tools in this category is the Dartmouth Atlas (The Dartmouth Institute [2010], Cooper [1996]) which is shown in Figure 2. The atlas features an interactive map of the United States and divides up the regions based on the states and the Hospital Referral Regions (HRRs). It is designed to visualize only two variables, which can be selected by a drop down menu - either the total reimbursements for the year 2007, or the annual growth

rate for the selected region from 2001-2007. The map is then colored in five different gradients, depicting the geographical variation for the selected variables. Selecting a particular state or HRR using the cursor also shows the average for the selected region and the corresponding national average in a new layer.

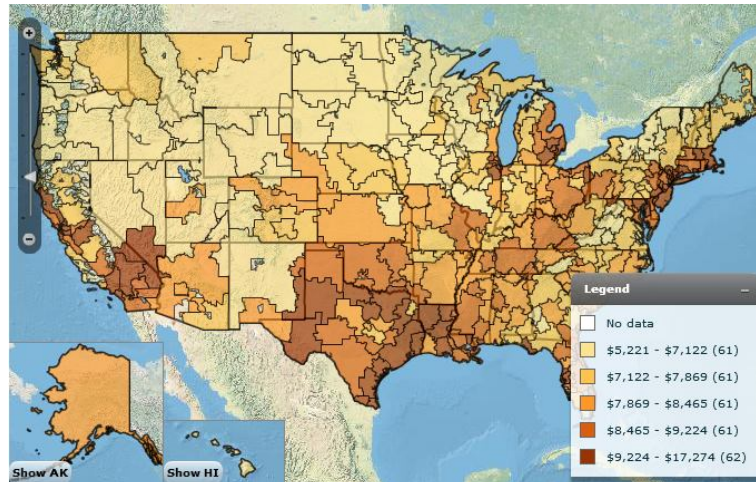


Figure 1: Dartmouth Atlas showing Medicare reimbursement per enrollee distributed over the Hospital Referral Regions

A modified and slightly improved version of the Dartmouth Atlas is available on the New York Times website (New York Times [2009]). Although the interactive map available on this tool only features the HRRs and not the states, variables from three broad categories - reimbursement, enrollees, and surgery rates are visualized. Like the original Dartmouth Atlas, only one variable can be visualized at a given time, and mouse over on a certain region shows a comparison of selected variable value of that region to the state average and national average using bar charts. The Health Statistics Map (Plaisant [1993]) uses dynamic query to visualize epidemiological data provided by National Center for Health Statistics (NCHS) on the US map. When the Interactive Cancer Atlas (InCA) (Interactive Cancer Atlas [2010]) shows only one variable in the table for all the US states, the Health Profiles Interactive (Association of Public Health Observatories [2010]) shows multiple variables in the table but only for a selected region along with the map view.

One of the web-based tools that allows comparisons of different variables and has been used in several studies is the Hospital Compare (Department of Health and Human Services [2009], Werner and Bradlow [2006]). The tool features an interactive wizard which gives users multiple options to select a geographical region of interest, a single medical condition and returns the most appropriate hospitals based on the input criteria. The tool also allows side by side comparison of up to three hospitals on three selected variables. However, Hospital Compare lacks the ability to compare different geographical regions across the state or the country for different variables, and is restricted only to the comparison of three hospitals in the chosen location at the county level.

The ability to visualize multiple variables on the map and compare them using a table and charts at the same time can provide valuable insights which might not be possible to obtain from current tools. Although a conceptual framework and the potential capabilities of such a tool have been discussed in the past, there are no similar applications existing to the best of our knowledge (Lu [2005]). While Geographic Information Systems (GIS) have been found to be a useful method to visualize the health care data (Denise Castronovo [2008], Villalon [1999]), there also have been concerns that using interactive maps without any other form of data visualization might be subject to misinterpretation (Goodman and Wennberg [1999]). The HealthVisPCP (Edsall et al. [2001]) uses a parallel coordinate plot and scatterplot linked to a GIS. Pennsylvania Cancer Atlas (Bhowmick et al. [2008]) deals with cancer data of Pennsylvania state and includes coordinated table, scatter plot and map view; however the table can show data of only one type of cancer at a time. However, as different geographical regions can have different demographical distribution, only the choropleth map may not suffice to provide a deeper understanding for hypothesis generation. It is also very important

to show if how the regions with similar demographics are doing in different areas of healthcare. (Maciejewski et al. [2010]) show a technique to show similar demographic region with same color in a choropleth map, though the distribution of health data in similar regions cannot be shown in the same map. With our dynamic query filtering, we provide a way for the users to select similar regions according to their demographics and visualize the healthcare data for those regions.

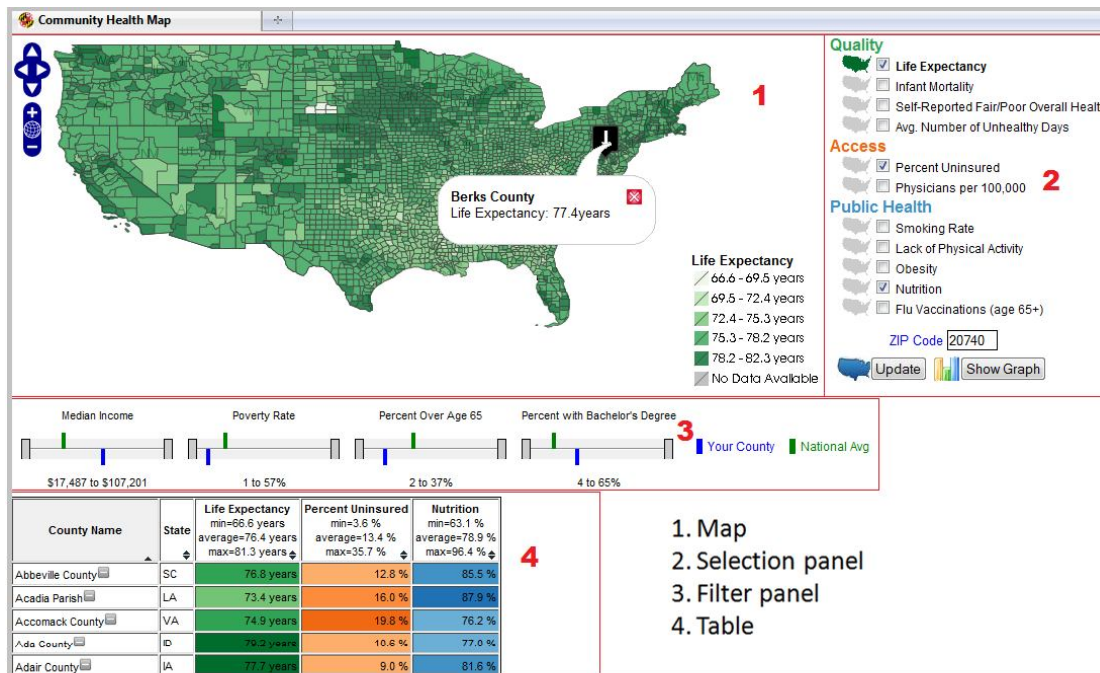


Figure 2: The Community Health Map interface consists of a map (top), hinted double-sided sliders for map and table filtering (middle), a color coded table (bottom), and the selection panel (right). The selection panel allows users to click the link to render a layer on the map, or check boxes to display a table with the selected variables.

Our Community Health Map gains its inspiration both from the Dartmouth Atlas as well as Hospital Compare. Community Health Map aims to provide the combined features of a simple GIS application having the capabilities of comparing multiple variables at any given time on a map and showing the insights using conventional graphs and tables as well enabling exploration of data such as zoom, pan, filtering, and access to raw data.

3. Data and Design

Community Health Map manipulates two types of datasets: healthcare data and demographic data. These data sets are combined with geospatial data sets which define the boundaries of regions such as counties in order to generate the map. The initial data set came from several sources and in different formats. The geospatial data was distributed as shape files; county level shape files came from the U.S. Census TIGER/Line website ¹. The demographic data used for the filter variables came from U.S. Census datasets distributed in Excel format compiled at the county level. The data filtering mechanism allows users to select counties based on specific demographic variables. The health care data are also broken down to county level and they were provided by Patrick Conway at the Department of Health and Human Services in excel format. These eleven health care variables are categorized into three functional

¹<http://www.census.gov/geo/www/tiger/>

areas: Quality, Accessibility and Public Health. This data is a modified and sanitized version of the data provided by the Community Health Status Indicators project ². Before any of this data could be used, all of it was loaded into the PostGIS relational database and the health data and demographic data was cross referenced to the geographic boundaries. The tool can be extended to accommodate any number of health care variables along with demographic variables for filtering purpose.

One of the design goals of Community Health Map was to simplify usability. Since the tool broadly targets *the people in charge*, we do not make any assumptions about technical prowess of these individuals. For this reason we expect that a web application would be the best solution since our target users are not computer experts. Any user with a modern web browser can view the data with different representations such as a heat-map, tables, and charts, which are generated without any additional plugin, binaries to download, or datasets to import. Another important design goal of Community Health Map is to make it an *app* in the true sense of the word. Community Health Map does far more than just simply push a few tables and map the relevant data for users to look at. Community Health Map can be qualified more as a *Rich Internet Application* in that all interactions happen within a page itself. Users are free to zoom and pan around the map, sort table columns, or pick a new variables with charts to explore all without ever leaving the page.

4. Community Health Map User Interface

Community Health Map embodies three coordinated visualizations: a map display, data table, and chart view. This allows users to see the geospatial as well as multivariate features of the data. Users inputs are provided to the system in three ways: the ZIP code entry box, selection panel, and dynamic query filter panel. The visualization dynamically changes according to these inputs. Figure 2 shows the Graphical User Interface (GUI) of Community Health Map. Essentially, the interface can broadly be divided in four distinct parts:

- The top left panel consists of a pannable and zoomable interactive map of United States.
- The middle left panel contains hinted double-sided filters that allow the users to filter the map and data tables simultaneously.
- The bottom left panel consists of data tables. The raw data is displayed in a sortable table with cells colored based on the quintiles of the data.
- The right side vertical selection panel contains a list of variables, classified into three categories, which can be selected using check boxes. Doing a mouse-over on any variable name shows its description in a small bubble. The ZIP code entry box lets users provide a ZIP code that will be selected on the map.

4.1. Map View

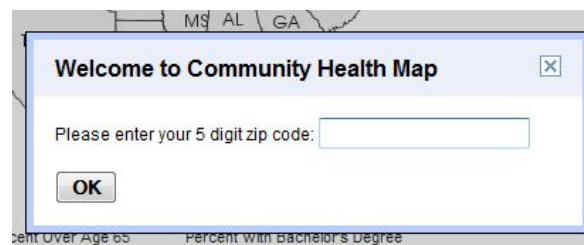


Figure 3: The welcome screen asks users to put a ZIP code representing their region of interest. Then, the tool will find out matched county for that ZIP code.

²<http://www.communityhealth.hhs.gov/HomePage.aspx>



Figure 4: Another method for comparing multiple variables with their national average is through the graphs. The modal dialog floats above the page so that users do not lose the context of what they were doing previously. Users can save the graphs to record their insights and send to others.

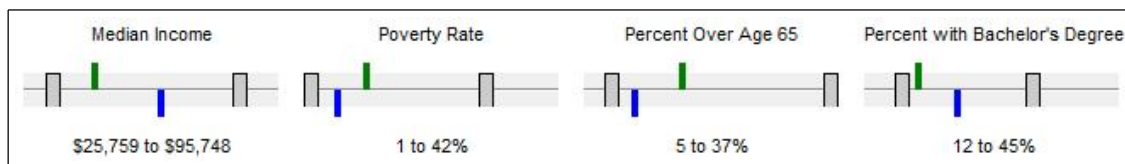


Figure 5: Double-sided slider for dynamic query filtering, annotated with the national average (green marker) of a demographic variable and the average for the **county** (blue marker) currently selected. The current range of the selection is shown below each filter; for example, in this figure, Poverty Rate is selected from 1 to 42% so counties with poverty rate more than 42% will not be shown in the map or the table.

When users start the Community Health Map, the ZIP code entry box will be shown. Users will be asked to put an area zip code which they are interested in (see Figure 3). Then, the arrow shielded black box will come up indicating the selected area in the map. To visualize the data on the map, users begin by clicking the map-icon beside a variable to draw the heat map overlay. Once a variable is selected, the label of the variable becomes bold and the geographical distribution is shown by heat maps varying in 5 different shades. Users can only select one variable at a time to be shown on the map. Users can zoom (using the map controls or mouse wheel), pan, and obtain the variable value of a region by clicking on the map. Clicking on any region on the map shows the name of the selected county, along with its value for the chosen variable in a small bubble.

4.2. Table and Chart view

To view multiple variables, the Community Health Map provides two additional views apart from the map: a chart view and a table view. Users can use the checkboxes in the selection panel to check the variables of interest and click the **Update** button (see Figure 2). The table shows the values of selected variables within the filter range so that the map and table always stay consistent. Cells are also color coded for easy comparisons with other variables and geographical regions. More intense colors in a cell indicate higher value for a variable. The table cells always mimic the map colors for consistency. The table columns are sortable that provides easy comparison of values. The county to which the area with the given ZIP code belongs is highlighted in the table with a yellow background (See Figure 9). When the selected county is filtered out by dynamic query filter, however, the highlighted color gets to become gray from yellow. Users can also create a bar chart for the geographic area of interest. The **show graph** button triggers a modal dialog containing a chart to hover above the page while dimming the interface underneath (see Figure 4). Each chart consists of a blue bar showing the value for the selected variable on that county and a green bar showing its corresponding national average as a benchmark. The charts are generated dynamically, so users can create charts

showing various variables for the selected region and their respective national average. The charts are particularly convenient for saving a snapshot of an insight or just having one more avenue to perform comparisons.

4.3. *Dynamic Query Filters*

One of the key features of Community Health Map is its ability to employ dynamic query filters. Four filtering options represented by double-sided sliders are present right below the map. The filters represent the Median Income, Poverty Rate, Percent of People Over Age 65, and Percent of People with Bachelors Degree. When the users change the range of the filter variables, both the map and data table update simultaneously by removing entries that are out of the selected range. The double-sided sliders are hinted with benchmark values such as the national average with green and the currently selected region's value with blue (see Figure 5). By selecting ranges that surround their position on the filter, the users can get a sense of which regions are similar to their own and immediately see how their colors compare to the colors that are remaining on the map and in the table. This is an important feature since users might try to use socioeconomic factors as excuses to rationalize poor performance or high cost. The ability to filter out data based on such demographic variables can be helpful in removing the credibility of such excuses. Initially the filters ranges are adjusted to national minimum to national maximum values for each demographic variable. When users move the sliders, the filters show the values for the current selected range below the filters. Each time users pick a new ZIP code, the **your region** marker shifts to the corresponding position on all filter sliders.

4.4. *Color Coding*

Each category of variables are assigned a particular hue, Green for Quality, Orange for Access and Blue for Public Health. When users are looking at the table they can easily understand which category a column belongs to by the color. Each variable value range is divided into five equal length portions and given a shade of the assigned hue so that difference in intensity corresponds to difference in the values. The sequential color schemes are derived from ColorBrewer Brewer [2004].

5. **Application Example**

This section presents insights we were able to achieve using Community Health Map. Having a geographical heat map is effective for showing variables that appear to be clustered around a particular region.

Figure 6 shows an overview of the entire US showing the Average Number of Unhealthy Days in the Last Month. There is a very clear dark spot that develops around eastern Kentucky into West Virginia (circled in red). The counties in this area all fall into the 4th and 5th quintiles and no other place on the map has such a dark area. A professor at University of Maryland pointed out that Kentucky and West Virginia are mining centers in the US. West Virginia and Kentucky have several underground coal mines, which is a plausible explanation of poor health. Furthermore, after zooming into this area and switching the map to life expectancy (Figure 6), it becomes very clear that the area of poor health corresponds to low life expectancy. Since the map re-renders completely in place, users can keep their eye on the area of interest to be able to identify correlations with ease. Users can then choose any other variable to draw on the map and look for other correlations. While it is somewhat intuitive that coal miners might have poorer health and lower life expectancy, we are encouraged that the tool allows users to gain these types of insights in a straight forward way.

The map filtering mechanism is also extremely effective in comparing and contrasting different segments of the population, such as wealthier vs. poorer areas. Wealthier areas were those with mean median income above national average and poverty rate below national average, while poorer were below the median. These groups were selected by adjusting the double-sided filters for Median Income and National Average. Figure 7 shows how wealthier and poorer counties report their overall health status respectively. These two maps show the same region of the United States, but we can see that the colors shown on the wealthier map are much lighter than the colors shown on the poorer map. This implies that wealthier counties tend to have better health than poorer ones. Again, while this insight may seem to be obvious, it demonstrates the ability of users to easily compare the quality of health care across socioeconomic classes, which is more of a challenge with other tools. The filtering mechanism is also useful for trying to pinpoint whether specific demographic variables determine poor quality or high costs. These four variables should allow users to dissect the United States in interesting ways to look for interesting or unexpected patterns.

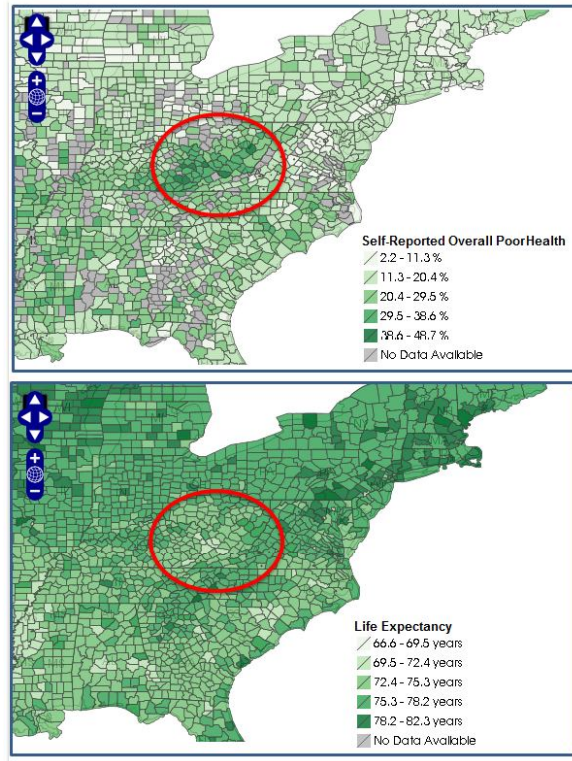


Figure 6: Top: The circled region having poor health around eastern Kentucky and West Virginia. The dark region circled in red shows the high number of Self-reported Overall Poor Health. Bottom: Same area of the map showing Life Expectancy. Lighter color indicates lower value of life expectancy.

Figure 8 shows the distribution of life expectancy over south central states. Life expectancy is relatively very high in Nebraska, Minnesota, Iowa and Colorado. However, four counties in the middle point of the map show distinct white spots, which represents very low life expectancy. In the detail view, we zoomed in and click to see the county names in South Dakota: Shannon, Bennett, Jackson, and Todd.

Then to analyze the correlation with other health indicator variables, we selected on infant mortality, percent uninsured and smoking rate variables. To compare Life Expectancy, we sorted the table by this variable. From Figure 9, Shannon County in South Dakota, shows the the standard we used for comparison. We can easily notice that top four counties with low expectancy are all from South Dakota and those are the counties with white spot from the map view. Except Bennett county, the three other counties also show high rate of infant mortality. We assume some poor environmental health conditions could influence this outcome. We find out from the table view, that those four counties have relatively high percent of uninsured population and high smoking rate compared to other counties.

We assume some health care access factors could influence these results. Therefore, for a more detailed comparison we selected two more variables, Percent uninsured and Physicians per 100,000, and generated the chart view. From Figure 10 we find out that Shannon County has more uninsured people and fewer physicians per 100,000 people. This comparison chart supported our hypothesis. Shannon County has twice as high percentage of uninsured population as the national average. Also it has significantly lower rate of physicians per 100,000 compared to the national average. These results show that there should be more support on health access to these counties to overcome poor environmental health conditions.

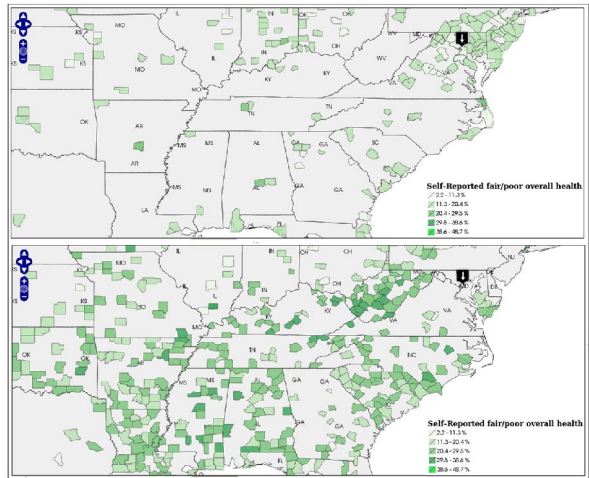


Figure 7: Top: Self-reported Overall Poor Health for wealthy counties of the US (low poverty rate, high median income). Bottom: Same variable for the same region, but this time only considering poorer counties. Darker gradient indicates greater percentage of self reported poor overall health.

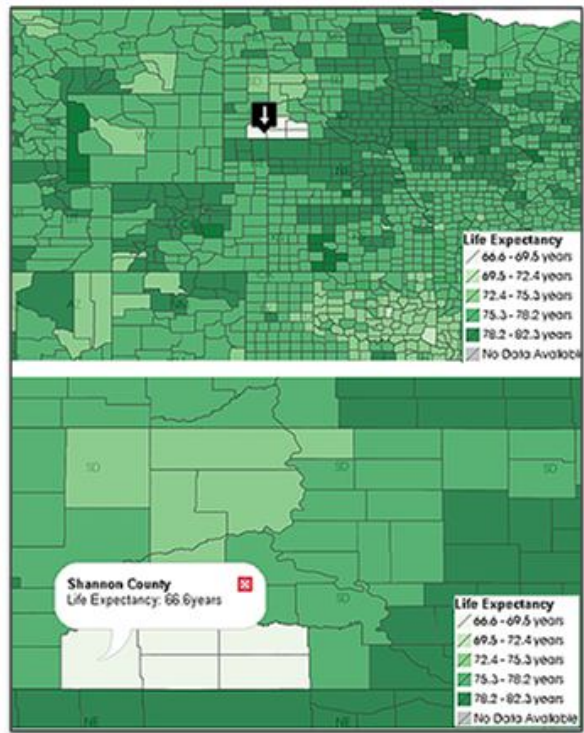


Figure 8: Top: Overview of Life Expectancy map for north central states. Bottom: Zoomed version of the map across South Dakota

6. Usability Evaluation

During the three months of development of Community Health Map we worked closely with Todd Park (the Chief Technical Officer of HHS), Karen Milgate (Director of Centers for Medicare and Medicaid Services), and Patrick

County Name	State	Life Expectancy	Infant Mortality	Percent Uninsured	Smoking Rate
		min=66.6 years average=76.4 years max=81.3 years	min=0.0 deaths average=7.3 deaths max=23.7 deaths	min=3.6 % average=13.4 % max=35.7 %	min=3.6 % average=23.4 % max=46.2 %
Bennett County	SD	66.6 years	3.3 deaths	24.1 %	28.8 %
Jackson County	SD	66.6 years	13.8 deaths	27.2 %	20.2 %
Shannon County	SD	66.6 years	15.7 deaths	29.3 %	42.4 %
Todd County	SD	66.6 years	14.9 deaths	30.8 %	39.9 %
Baltimore city	MD	69.6 years	11.1 deaths	17.4 %	22.6 %
Marlboro County	SC	69.6 years	10.6 deaths	19.2 %	35.7 %
Petersburg city	VA	69.6 years	12.6 deaths	18.1 %	26.0 %
Phillips County	AR	69.6 years	10.4 deaths	26.0 %	20.9 %

Figure 9: Color coded table shows the correlation among four variables from quality (colored by green), access (colored by orange) and public health indicator (colored by blue). Table is sorted by Life Expectancy. Also, the selected county, Shannon County is highlighted by yellow background, which makes users find out easily where their region of interest lies.

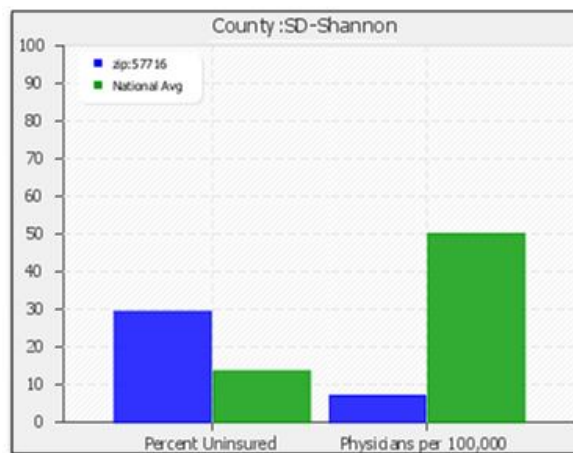


Figure 10: Several health indicator values of Shannon County are compared to the national average. Blue ones show the values of Shannon county and green ones show the national average.

Conway (Chief Medical Officer of HHS) and implemented their suggestions. Their assessment of primary users and use cases guided our development. After completion of each milestone we sought their feedback and made changes accordingly so that the Community Health Map would be useful as well as simple to use. In our final demo to Bill Corr (Deputy Secretary of HHS), Todd Park and his team, we received positive feedback on the usability and flexibility of the tool. In spite of their brief use of Community Health Map, they found it to be superior to the many interactive maps they had studied. During our final demo Todd Park remarked:

This is incredibly cool, I think it is spare, elegant, and intuitive. The White House and HHS see this as a powerful seed for changing health care in the United States.

An initial usability evaluation with a user interface expert and some of graduate students specializing in Health Informatics at the University of Maryland, College Park also revealed interesting results. Subjects were given a four minutes video demonstration of the Community Health Map and allowed to use it for twenty minutes. They were asked to perform five tasks incorporating the use of filter, interactive map, table and chart:

- Find the areas with similar income to their own ZIP Code.
- Identify which one of these areas is richest.
- Find which health care variable in their selected ZIP Code area needs the most improvement as it is far below the national average.

- Find which county is performing best for some health care indicator variable and then to find the top 10 other counties that are next to that county for that variable.
- Find which county has most '% of Uninsured People' among all counties that have poverty rate above national average.

We designed the tasks especially to evaluate whether the users are comfortable enough using different aspects of the interface like clicking on the map to see the value of selected variable, using double ended slider for filtering, sorting table and using the chart for comparison. We asked them to think aloud so that we could keep record of their reactions. Finally they were asked to explore the tool on their own and work with variables of their choice, based on their expertise and research interests. They also provided their general feedback at the end of the evaluation.

While one of the subjects had a few minor suggestions about the usability of the interface, they agreed that a small demonstration before the sessions was necessary for them to use the tool effectively. None of the subjects reported any major usability challenges and were able to perform all the tasks easily. Especially, they all liked that this tool is a web-based program which means no additional effort for using it such as difficult installation process. One of the subjects, a Health Informatics consultant in the Washington D.C. area, said:

This tool is simple, innovative and powerful. I haven't come across anything like this before and hope professionals in my field will have access to it sometime soon.

Another subject, who is also a PhD candidate in Public Health wanted to verify the hypothesis about the correlation between lack of physical activity and obesity. The subject was easily able to do his research using the tables which shows both of variables at the same time for not only all areas but also only the areas which he was more interested in by filtering, and comparing those numbers with the national average. Further, this subject also appreciated the simplicity of the interface, along with the ability to filter the results using the dynamic query filters.

After getting suggestion from our user interface expert, we modified the interface to make it more intuitive to use. We made the color of selected county's bar in the chart same as the color of hinting bar for selected county's demographic indicator in the query filter sliders (it was different before). We also added the feature that mousing over a variable name in selection view will show a tooltip of the variable's description and data source.

7. Conclusion and Future Directions

The Community Health Map was designed as a simple web application to enable a broad range of users to explore health care data for the geographical regions of their interest. Some of the key features of Community Health Map are its ability to:

- analyze data at the level of counties
- visualize the data simultaneously on a geographical map, charts and tables
- use dynamic query filters to selectively visualize healthcare data of regions with similar demographic features and identify who is doing better in this aspect(a region can improve its health care condition by following the strategy of other regions with similar demographic pattern but better health care condition)
- dynamically change the visualization of the data by selecting a specific range of income, poverty levels, age and education of the chosen geographical regions and
- compare multiple geographical regions as well as variables side by side.

The features, along with the simplicity make Community Health Map a powerful tool to analyze the health care data in United States and assist in discovering important trends and valuable insights. Our clients, the Department of Health and Human Services envision this as a tool that can spur action by decision makers and streamline health care spending.

This prototype should act as an inspiration to more complete implementations what would include more demographic, outcome, and cost variables, as well as richer geographic region selection features. An expanded implementation would support linking to other datasets, collaboration among users, and history-keeping to enable backtracking during exploration.

Further testing would be needed to understand user problems and user requests for additional features. While the current design emphasizes visualization which presents some difficulties for users with visual disabilities, an improved version could support spatial discovery with sonification and tabular exploration with text-to-speech tools.

Although Community Health Map was designed for the health care data, it can easily be used to analyze any multivariate data set with a geographic spread, and could be of particular interest for users in the education or finance domains.

8. Acknowledgements

We would like to thank Patrick Conway and Karen Milgate for taking the time to supply us with clean datasets that were easy to use and Todd Park who provided the initial inspiration for this project and his continued enthusiasm over the tool made it that much more satisfying to develop. Finally we thank Catherine Plaisant and other members of the Human-Computer Interaction Lab of University of Maryland for their valuable suggestions on the design of our user interface and usability test.

References

- New York Times (2009). Regional differences in costs and care. Online. <http://www.nytimes.com/ref/business/20070611GAPGRAPHIC.htm>;accessed 04-Dec-2009.
- The Dartmouth Institute (2010). Dartmouth atlas of health care. <http://www.dartmouthatlas.org/>;accessed 11-Aug-2010.
- Association of Public Health Observatories (2010). Health profiles interactive. Online. <http://www.apho.org.uk/default.aspx?QN=HPINTERACTIVE>.
- Bhowmick, T., Robinson, A., Gruver, A., MacEachren, A., and Lengerich, E. (2008). Distributed usability evaluation of the pennsylvania cancer atlas. *International Journal of Health Geographics*, 7(1):36.
- Brewer, C. A. (2004). Color research applications in mapping and visualization. In *Color Imaging Conference*, pages 1–3. IS&T - The Society for Imaging Science and Technology.
- C. Tominski, P. Schulze-Wollgast, H. S. (2008). Visual methods for analyzing human health data. *Encyclopedia of Healthcare Information Systems*, pages 1357–1364. Information Science Reference, 2008, ISBN 978-1-59904-889-5.
- Cooper, M. M. (1996). The dartmouth atlas of health care: what is it telling us? *Health Syst Rev*, 29(3):44–5, 47.
- DeMott, K. (1997). Healthcare practices vary widely from town to town: regional dartmouth atlas. *Health Syst Lead*, 4(1):2–3.
- Denise Castronovo, E. N. (2008). A conceptual framework for visualizing seasonality in spatiotemporal health data using dynamic maps. In *American Public Health Association 136th Annual Meeting*.
- Department of Health and Human Services (2009). Hospital compare. <http://www.hospitalcompare.hhs.gov>.
- Donker-Kuijjer, M. W., de Jong, M., and Lentz, L. (2010). Usable guidelines for usable websites? an analysis of five e-government heuristics. *Government Information Quarterly*, 27(3):254 – 263.
- Edsall, R. M., MacEachren, A. M., and Pickle, L. (2001). Case study: design and assessment of an enhanced geographic information system for exploration of multivariate health statistics. pages 159–162.
- EL, K. and MJ, K. (2004). Geovisualization to support the exploration of large health and demographic survey data. *International Journal of Health Geogr*, 3(1):12.
- Fedorowicz, J. and Dias, M. A. (2010). A decade of design in digital government research. *Government Information Quarterly*, 27(1):1 – 8.
- Fisher, E. S., Bynum, J. P., and Skinner, J. S. (2009). Slowing the growth of health care costs – lessons from regional variation. *N Engl J Med*, 360(9):849–852.
- Gil-Garca, J. R. and Pardo, T. A. (2005). E-government success factors: Mapping practical tools to theoretical foundations. *Government Information Quarterly*, 22(2):187 – 216.
- Goodman, D. C. and Wennberg, J. E. (1999). Maps and health: The challenges of interpretation. *Journal of Public Health Management and Practice*, 5(4):1.
- Himmelstein, D. U., Wright, A., and Woolhandler, S. (2009). Hospital computing and the costs and quality of care: A national study. *Am J Med*.
- Interactive Cancer Atlas (2010). Interactive cancer atlas. Online. <http://apps.nccd.cdc.gov/DCPCINCA>;accessed 11-Aug-2010.
- Keahey, T. A. (1998). Visualization of high-dimensional clusters using nonlinear magnification. pages 228–235.
- Lu, X. (2005). A framework of web gis based unified public health information visualization platform. In *Computational Science and Its Applications ICCSA 2005*, pages 265–268. Springer Berlin.
- Maceachren, A. M., Brewer, C. A., and Pickle, L. W. (1998). Visualizing georeferenced data: Representing reliability of health statistics. *Environment and Planning A*, pages 1547–1561.
- Maciejewski, R., Drake, T., Rudolph, S., Malik, A., and Ebert, D. S. (2010). Data aggregation and analysis for cancer statistics - A visual analytics approach. pages 1–5.
- Madigan, E. and Curet, O. (2006). A data mining approach in home healthcare: outcomes and service use. *BMC Health Services Research*, 6(1):18.

- Meier, B. D. and Homann, D. (2009). Fraud in the health-care system from the perspective of the public health insurance companies. *Gesundheitswesen*.
- National Cancer Institute (2010). Nci research-based web design and usability guidelines. Online. <http://www.usability.gov/guidelines/index.html>;accessed 04-Dec-2009.
- Plaisant, C. (1993). Facilitating data exploration: Dynamic queries on A health statistics map.
- Verdegem, P. and Verleye, G. (2009). User-centered e-government in practice: A comprehensive model for measuring user satisfaction. *Government Information Quarterly*, 26(3):487 – 497.
- Villalon, M. (1999). Gis and the internet: tools that add value to your health plan. *Health Manag Technol*, 20(9):16–18.
- Werner, R. M. and Bradlow, E. T. (2006). Relationship Between Medicare's Hospital Compare Performance Measures and Mortality Rates. *JAMA*, 296(22):2694–2702.