Spatial Analysis in Public Health Practice and Research

Gerard Rushton
Department of Geography and Department of Health Management and Policy
The University of Iowa
Iowa City, Iowa, U.S.A.

Gerard-rushton@uiowa.edu
Health Geomatics Workshop 2008
Victoria, British Columbia
September 22, 2008

Spatial Analysis in Public Health

• I will discuss seven areas in public health where spatial analysis is used as a basis for leading mapping innovations;
• In each area, I will comment on the application area and the focus of the spatial analysis approach;
• In each case I will show that current approaches do not take advantage of recent developments in geospatial data availability and of new methods of spatial analysis that use such data in geographic information systems;
• I am critical of the current use of many spatial analysis methods that support mapping applications in public health.
1. **What method of age standardization should be used for making decisions about the allocations of health resources?**

- Rates mapped for targeting health interventions should be **INDIRECTLY AGE-STANDARDIZED** NOT Directly age-standardized as in several recent CDC disease atlases.
- Because interventions to improve the health of a community should be based on the demographic characteristics of the community not on the rates that would prevail if the community had the same demographic characteristics as the national population!
- Allocations of resources can be quite different if made on the basis of directly age-adjusted rates.
- An additional advantage of mapping indirectly age-standardized rates is that more geographic detail can be reliably shown.

Rushton, University of Iowa
“Men and Heart Disease indicates where those programs are most needed and can have the greatest benefit. It is my hope that Men and Heart Disease: An Atlas of Racial and Ethnic Disparities in Mortality will be used to guide the distribution of funds and resources to those communities of men experiencing excess mortality from heart disease and will promote the development of culturally sensitive prevention strategies.”

Age-standardization to the local pattern of ages?

• “The fact that populations vary by age must be incorporated into, not standardized out of, program planning and evaluation.” Kerner et al. 1988, p. 551.
Age-adjusted rates for cardiovascular disease mortality—all women in Iowa 1991-1995—rates spatially interpolated from County centroids)

DIRECT age-adjustment: NOT useful for targeting resources

INDIRECT age-adjustment: useful for targeting resources—not useful for directly comparing rates between areas

Source: Ishwari Sivagnanum, University of Iowa student.

Circles show the difference between the observed and the expected number of deaths (from Indirectly Age Adjusted Rates). Black is excess deaths—greater than expected; Brown is fewer than expected number of deaths.
### Conclusion on spatial analysis method

- For making resource allocation decisions:  
  Indirectly age-adjusted rates apply national rates to the age-sex populations of local areas to compute expected numbers of cases in a local area;  
  These numbers are reliable for small areas;  
  These small areas can be aggregated for overlapping spatial filter areas on a fine grid to give a spatially continuous map of observed / expected numbers of deaths or incidences.

### 2. What areas should be mapped?

- Current disease maps do not meet the needs of the public:  
  - They use political or administrative areas of different shapes and sizes and frequently are larger in population and area than are necessary to accurately portray rates;  
  - Technically, they show rates that differ in their reliability; the public is confused by this;  
  - When rates are “adjusted” to account for the different levels of support for the statistics on the map, they are further confused.
Infant Mortality Rates at Three Different Spatial Scales and Their Approximate Counterparts Using Available Census Administrative Areas

Des Moines, Iowa
1989 - 1992

Spatial Filters
1.2 miles

0.8 miles

0.4 miles

Zip Codes
Census tracts
Census block groups
Percent of simulated infant mortality rates at 4,000 grid locations on 1,000 maps that were less than the observed rate--0.8 mile filter

The null hypothesis for the Monte Carlo simulations is that each child has the same probability of dying as the region-wide rate.

3. What are the key characteristics of disease maps for public use?

- They must deal effectively with the problem of rate stability; otherwise the public focuses inappropriately on high rate areas that are usually the areas with the least reliable rates.
- The problem of rate stability arises because of the different spatial supports for the rates mapped—some areas have larger populations than others and larger populations give more stable and reliable rates.
- The problem is solved in this case by using an identical spatial basis of support—called a “spatially adaptive filter.”

See Talbot et al. 2000 and Tiwari and Rushton 2005
Spatial Analysis Method: Kernel density estimation

**Spatially Adaptive Filters**

**Fixed Filter (24 Miles)**

Legend:
- Filter Size (Miles):
  - Green: 10
  - Orange: 15
  - Blue: 20
  - White: 25
  - Red: 30

Fixed distance kernels
12 mile radius

Spatially adaptive kernels
100 breast cancer cases in each circle

Cedar Rapids
Coralville
Iowa City

Rushton, University of Iowa
Spatially Adaptive Filters

Iowa: Late stage colorectal cancer rates per thousand diagnosed cases 1993-1997

Spatially Adaptive Filter
Threshold > 100

24 Mile Filter

<table>
<thead>
<tr>
<th>Geographic Detail</th>
<th>Stability</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>High</td>
<td>Low</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Geographic Detail</th>
<th>Stability</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Low</td>
</tr>
</tbody>
</table>

Rushton, University of Iowa

Difference in Rates Determined using Fixed Filters and Spatially Adaptive Filters

Absolute Difference
None
1 - 25
26 - 50
51 - 75
Above 75
No Data

Rushton, University of Iowa
4. What are the three essential properties of disease maps designed for public use?

1. Cancer rates shown should be based, ideally, on approximately equal numbers of expected cancer cases and they should be indirectly age-sex adjusted;
2. Rates should be continuously distributed;
3. It should be clear to the user of a cancer map exactly what area has contributed data to any location examined and a familiar backdrop image should be viewed.
This map also has two of the three essential properties:

- it shows the rate as a continuous distribution (late-stage colorectal cancer rate);
- all rates shown are based on approximately the same number of expected late-stage cases (50 expected cases)

- it does not show the area that has contributed data to the rate;
- the user cannot “read-off” the rate as in the last map

Comparison of smooth and county maps of late-stage colorectal cancer in northwest Iowa, 1998-2003
This map has two of the three essential properties: it shows the rate as a continuous distribution (infant mortality); it shows the area that has contributed data to the rate in the middle of the red circle which the user has checked. The blue box provides detail of the data shown. The user can click anywhere on this map and immediately get the same information.

Here we drape the computed rate map over a conventional image.

Example of web-based spatial analysis using spatially adaptive filters for infant mortality rates in Des Moines, Iowa. The pop-up enquiry is shown on background of google earth and gives numbers of births, number of infant deaths, and infant mortality rate.

In this circle there were 1250 births and 29 deaths. Imr = 23.2

Source: Chetan Tiwari, University of Iowa
Technologies Used

• Opensource GIS software
  – PostgreSQL + PostGIS (spatial database)
  – Apache (web server)
  – Perl (programming language)
  – GDAL (geospatial data abstraction library)
  – Beta test site at: http://healthgis-01.iowa.uiowa.edu/webdmap

Source: Chetan Tiwari, The University of Iowa

Rushton, University of Iowa
Example of block group centroids within a 12 mile filter

Rushton, University of Iowa
Note poor level of geographic resolution caused by smoothing of 99 county centroids.

Colorectal Cancer Mortality Rate, 1999-2003

Note the better geographic resolution when mortality data is disaggregated to city and rest of county level
Need for Public Access to Disease Maps with choice of different views

Welcome to the Iowa Consortium for Comprehensive Cancer Control Cancer Maps Site

Prepared by
Kirsten Beyer, Zunqiu Chen, Veronica Escamilla and Gerard Rushton

Narrative Report on cancer maps A Guide to Interpretation

www.uiowa.edu/~gishlth/ICCCCMaps

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Incidence rates</td>
<td>Incidence rates</td>
<td>Incidence rates</td>
</tr>
<tr>
<td>County level</td>
<td>County level</td>
<td>County level</td>
</tr>
<tr>
<td>Smoothed map</td>
<td>Smoothed map</td>
<td>Smoothed map</td>
</tr>
<tr>
<td>Observed minus Expected Incidence</td>
<td>Observed minus Expected Incidence</td>
<td>Observed minus Expected Incidence</td>
</tr>
<tr>
<td>County level</td>
<td>County level</td>
<td>County level</td>
</tr>
<tr>
<td>Smoothed map</td>
<td>Smoothed map</td>
<td>Smoothed map</td>
</tr>
<tr>
<td>Late Stage Rates</td>
<td>Late Stage Rates</td>
<td>Late Stage Rates</td>
</tr>
<tr>
<td>County level</td>
<td>County level</td>
<td>County level</td>
</tr>
<tr>
<td>Smoothed map</td>
<td>Smoothed map</td>
<td>Smoothed map</td>
</tr>
<tr>
<td>Observed minus Expected late stage</td>
<td>Observed minus Expected late stage</td>
<td>Observed minus Expected late stage</td>
</tr>
<tr>
<td>incidences</td>
<td>incidences</td>
<td>incidences</td>
</tr>
<tr>
<td>County level</td>
<td>County level</td>
<td>County level</td>
</tr>
<tr>
<td>Smoothed map</td>
<td>Smoothed map</td>
<td>Smoothed map</td>
</tr>
<tr>
<td>Mortality Rates</td>
<td>Mortality Rates</td>
<td>Mortality Rates</td>
</tr>
<tr>
<td>County level</td>
<td>County level</td>
<td>County level</td>
</tr>
<tr>
<td>Smoothed map</td>
<td>Smoothed map</td>
<td>Smoothed map</td>
</tr>
</tbody>
</table>

Proportion of cases diagnosed in late stage is indirectly standardized by age.
Red areas indicate higher rates than expected and blue areas indicate lower rates than expected, given the statewide late stage cancer rate.

Data Sources: Cancer data from the Iowa Cancer Registry; County borders from the Natural Resources GIS Library hosted by the Iowa Department of Natural Resources and the US Geological Survey.

Created by: Zunqiu Chen, Dept. of Geoscience, University of Iowa June 2006
Rushton, University of Iowa 32
What is the relationship between rates of late-stage prostate cancer and rates of mortality from prostate cancer?
Spatial relationship between prostate cancer late-stage rate and prostate mortality

Prostate cancer mortality rate

Late-stage prostate cancer rate

$r = 0.047$

Relationship Between Rates of Male Prostate Cancer Late-Stage Diagnosis and Mortality in Iowa, 1998-2003

Note: Rates were calculated for 350 grids on a 5-mile uniform grid using adaptive spatial fitting with 50 expected late-stage cases and 50 expected deaths. Areas for which rates are calculated are not necessarily the same for late-stage and mortality rates. “High” is defined as greater than or equal to the expected number of cases, and “low” is defined as less than the expected number of cases. Data is from the Iowa Cancer Registry and the Iowa Department of Public Health.
What do we know about the spatial relationships between incidence rates of screening and rates of late stage cancer?

- Surprisingly little!
- There appears to be a universal assumption that higher rates of screening produce lower rates of late-stage cancers:
- “small-area analysis can be used to determine where to set up screening programs so that high-risk populations can be more easily recruited.” (Andrews et al. 1994, p.56.)
Late-stage cancer rate Vs. screening rate

Age-adjusted late-stage breast cancer rate in Iowa, 1998 – 2001
(state average = 0.1062)
Females, all ages

Age-adjusted mammography screening rate in Iowa, 1998 – 2001
(state average = 0.3191)
Iowa Medicare Population i.e. > 65 years

Rushton, University of Iowa


$r = -0.24$

Rushton, University of Iowa
5. What are the relationships between the availability of resources for treatment and the spatial choice of treatment?

• Example: Choice of lumpectomy with radiation and the location of radiation treatment facility.

Compare areas of high rates of lumpectomy/radiation with locations of radiation treatment facilities
Choices per 1,000 cases of localized breast cancer, 15 mile filter
6. What can be inferred by studying the spatial choices of patients for therapies?

- Visualizations of spatial choice
- Comparisons of chosen with potential places for interaction
- Imputing changes in spatial choices through time
- Finding distances to closest, second-closest, etc. places with given resources
This is a spider map.

A different view of the same cancer data: average distance to place of diagnosis for colorectal cancer cases in Iowa, 1993-1997.
7. Can methods for finding the optimal location of facilities ensure spatial access to essential health services?

- The models to find optimal locations for the locations of services are called “location-allocation” models.

“Siting a Women’s Health Facility: a location-allocation study of breast cancer screening services in Eastern Ontario” by N.A. Ross, M.W. Rosenberg and D.C. Pross

Distribution of women of screening age (50 - 69 years)

Location of diagnostic mammography facilities
Source: Ross et al., 1994, p. 156.
Result from one run of the location-allocation model: Scenario 5

<table>
<thead>
<tr>
<th>Location-allocation functions are not well-developed in most commercial GIS products</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Location-allocation heuristic models were a formal part of ArcInfo but are not yet formally available in ArcGIS—release is tentatively scheduled for 2009.</td>
</tr>
<tr>
<td>• Some steps in the heuristic models can be done in a database management model such as EXCEL. (see Rushton 1999).</td>
</tr>
<tr>
<td>• For example: “which service site if added from a list of feasible sites would most improve average distance of people to their nearest facility?”</td>
</tr>
</tbody>
</table>

---

<table>
<thead>
<tr>
<th>Limit</th>
<th>Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 km</td>
<td>50 km</td>
</tr>
<tr>
<td>Unserved demand</td>
<td>0</td>
</tr>
<tr>
<td>Demand with no alternate site</td>
<td>2220</td>
</tr>
<tr>
<td>Demand allocated</td>
<td>83138</td>
</tr>
<tr>
<td>Percent of total demand</td>
<td>100%</td>
</tr>
<tr>
<td>Distance travelled*</td>
<td>229761</td>
</tr>
<tr>
<td>Mean distance*</td>
<td>2.73</td>
</tr>
<tr>
<td>Maximum distance*</td>
<td>19</td>
</tr>
</tbody>
</table>

*1 unit = 5 km

## Conclusions: spatial analysis in public health practice and research

<table>
<thead>
<tr>
<th>Important questions</th>
<th>My response to question</th>
<th>Comment on spatial analysis method</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. What disease rate should be mapped when resource allocation decisions are in play?</td>
<td>Indirectly—not directly—age-adjusted rate maps should be used to support resource-allocation decisions.</td>
<td>Indirectly age-adjusted rates apply regional rates to local population characteristics to give expected numbers of disease incidences in any area. The observed number / expected number of incidences as well as the observed – the expected number, are most relevant to spatially allocate resources.</td>
</tr>
<tr>
<td>2. What areas should be mapped?</td>
<td>Areas of approximately equal population provide a consistent spatial basis of support.</td>
<td>Kernel-density estimation methods are the spatial analysis methods that best implement this principle. They operate on small-area, geocoded health data.</td>
</tr>
<tr>
<td>3. What kinds of maps should be available to the public?</td>
<td>The public should be given a choice of types of maps.</td>
<td>But there are important cartographic and statistical principles that should be followed.</td>
</tr>
<tr>
<td>4. What are the essential properties of disease maps?</td>
<td>1. Rates should be based on approximately equal numbers of persons at risk; 2. Rates should be continuously distributed spatially; 3. Visualization should show areas that support each rate observed.</td>
<td>Optimal location models are available in some GIS software but are often not available. There are technical problems in implementing these methods in GIS related to the spatial data structures that are optimal for solving these problems using available solution methods.</td>
</tr>
<tr>
<td>5. What relationships exist between the spatial pattern of health resources and consumer choices?</td>
<td>If undesirable relationships exist, consider changing spatial patterns of resources or otherwise dealing with problems of spatial access to essential resources.</td>
<td>Methods of revealed space preference show the sensitivity of consumer choice to spatial patterns of resources.</td>
</tr>
<tr>
<td>6. How do you measure the spatial relationships between two map patterns?</td>
<td>Correlate the values on each map over a spatial grid.</td>
<td>Be aware that assumptions that the correlated values are independent are not met.</td>
</tr>
<tr>
<td>7. How do you assess and improve the locations of services for known locations of need.</td>
<td>Apply principles of location-allocation analysis asking questions such as “best place to add a service;” “best place to drop a service;” “best places to add p service.”</td>
<td>Optimal location models are available in some GIS software but are often not available. There are technical problems in implementing these methods in GIS related to the spatial data structures that are optimal for solving these problems using available solution methods.</td>
</tr>
</tbody>
</table>
Some URLs related to this presentation

- To see application of spatially adaptive filtered cancer maps of Iowa, see:
  - [www.uiowa.edu/iowacancermaps/](http://www.uiowa.edu/iowacancermaps/)
  - [www.uiowa.edu/~gishlth/ICCCCMaps](http://www.uiowa.edu/~gishlth/ICCCCMaps)
- For software for producing spatially adaptive filtered maps see:
  - PC-based: [www.uiowa.edu/~gishlth/DMAP4](http://www.uiowa.edu/~gishlth/DMAP4)
  - Web-based (beta version):
    - [http://healthgis-01.iowa.uiowa.edu/webdmap](http://healthgis-01.iowa.uiowa.edu/webdmap)

Relationships between distance to mammography facilities and their use

References—targeting cancer services


References Continued

References on measuring spatial accessibility


References—location-allocation methods


Acknowledgements

• Support from the National Cancer Institute (NIH 1 R01 CA95961-01) and co-investigators, Michele West, Barry Greene, Jane Pendergast, Betsy Chrischilles, Marc Armstrong, Claire Pavlik, David Bennett.

• Research assistants Kirsten Beyer, Chiang Cai, David Haynes, Zunqiu Chen, Chetan Tiwari, Kara Wright

• The Iowa Cancer Registry, Chuck Lynch, Director

Thank you!