Explaining the Geographic Distribution of Colorectal Cancer Survival: an Iowa example

North American Association of Central Cancer Registries Annual Conference, Louisville KY
June 21, 2011

Kevin Matthews
Kevin-matthews@uiowa.edu
PhD Student, Health Geography
University of Iowa, Department of Geography
Iowa Cancer Maps
http://www.uiowa.edu/iowacancermaps/

Colorectal Cancer, 2000 – 2005

Standardized Incidence Rate (SIR)

Standardized Mortality Rate (SMR)

Late-Stage Incidence Rate
Measures of Disease Burden

• Common descriptive measures in traditional epidemiology
  – Incidence
  – Mortality
  – Late-Stage Rates
  – Survival

• With a geocode, we can identify the geographic patterns of each measure

• Mapping incidence, mortality and late-stage rates are established in health-related geographic research

• If a diagnosis date and a date of death is available, we can map SURVIVAL as well
Current Activity in **Small Area Survival Mapping**


Survival in Cancer Atlases

• One Example: Cancer Atlas of the United Kingdom and Ireland 1991–2000

• Measures of burden of disease
  – “…a compendium of trends in cancer incidence, mortality and survival…”
  – “…data on incidence and mortality and analyses of survival trends…”
  – Five other similar statements about incidence, mortality and survival

• Number of Maps per Measure
  – 21 maps of Cancer Incidence (1 for each type of cancer)
  – 21 maps of cancer mortality
  – 0 maps of cancer survival

Source: Cancer Atlas of the United Kingdom and Ireland 1991–2000
SURVIVAL: Summary Measures

1. Five Year Survival Rates (Kaplan-Meier)

2. Standardized Survival Rate (SSR) (Cox Proportional Hazards Model)

3. Survival Clusters
Data and Methods
Population Characteristics

• Study Population
  – All persons in Iowa aged 50 and older
  – Newly diagnosed with Colorectal Cancer
  – between January 1, 1997 and December 31, 2008

• 18,260 Colorectal Cancer Cases

• 4,867 Colorectal Cancer deaths

• 13,393 Censored Cases
Data Overview

For each Colorectal Cancer Case:

• **Attribute Data (from Iowa Health Registry)**
  – Age at diagnosis
  – Stage at Diagnosis
  – Gender

• **Geographic**
  – Address-level geocodes for residence of case at time of diagnosis
  – Source: Iowa Cancer Registry

• **Temporal**
  – Study: January 1, 1997 to December 31, 2008
  – Resolution: Monthly (144 months)
  – Date of diagnosis: Month and Year
  – Date of death: Month and Year
  – SURVIVAL TIME: Computed by subtracting the number of months from diagnosis to the number of months since death
Data Overview

• Exclusion Criteria
  – SEER Stage 9
  – Persons less than 50
  – A recurrent diagnosis for the same individual

• Censoring Criteria
  – Diagnosees alive at the end of the study timeframe (December 2008)
  – Death from causes other than Colorectal cancer
    • Defined by ICD-10 SEER Cause-specific Death
    • Unknown cause
Data Format

- **○** = Colorectal Cancer Diagnosis  (n = 18,260)
- **×** = Colorectal Cancer Death  (n = 4,787)
- **●** = Censored  (n = 13,393)

Person A
Survival Time: 63 months

Person B
Survival Time: Unknown
(measured as 26 months)

Person C
Survival Time: Unknown
(measured as 130 months)

Person D
Survival Time = 10 months

Time (t):  
- t0 = Jan 1, 1997  
- t1 = Jan 31, 1997  
- t13 = 1-Year Survival  
- t61 = 5-Year Survival  
- t144 = Dec 2008
The data to be summarized and mapped are these Survival Times.
Example Survival Time Data

<table>
<thead>
<tr>
<th>Case</th>
<th>Date of Diagnosis</th>
<th>Date of Death</th>
<th>Survival Time</th>
<th>Cumulative Time at Risk</th>
<th>Distance to Grid Point</th>
<th>X-coord</th>
<th>Y-coord</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Jun 1998</td>
<td>Feb 2001</td>
<td>63 months</td>
<td>63</td>
<td>2 miles</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>Oct 2005</td>
<td>--</td>
<td>38 months</td>
<td>101</td>
<td>4 miles</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>Jan 1997</td>
<td>--</td>
<td>143 months</td>
<td>244</td>
<td>5 miles</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>Jan 1997</td>
<td>Nov 1997</td>
<td>10 months</td>
<td>254</td>
<td>10 miles</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- If this example were the entire study population, the **Total Time-At-Risk** = 254 months

- If we set our threshold to be 250 months, case D would be excluded from the spatial partition
Methods Overview

• Spatially Adaptive Filtering
  – to create a set of spatial partitions and identify the number of diagnosees within each partition
  – ArcGIS 9.3

• Kaplan-Meier Survival Analysis
  – to calculate Survival Probability for all persons within each spatial partition
  – STATA 10

• Cox Proportional Hazards Model
  – to compare survival probability for all persons within each spatial partition to the entire state
  – STATA 10

• Inverse Distance Weighting
  – Interpolation method between centroids of the spatially adaptive filters to generate a “surface” of survival probabilities
  – ArcGIS 9.3 – Geostatistical Analyst
Spatially Disaggregating the Kaplan-Meier Survival Estimate and Cox Proportional Hazard Model
The Kaplan-Meier Survival Estimator

Estimates the probability of survival beyond Time (t) for a population based on an aggregation of individual survival times.

\[ s(t) = \prod_{j \mid t_j \leq t} \frac{(n_j - d_j)}{n_j} \]
Colorectal Survival Rates in Iowa (1997 to 2008)

<table>
<thead>
<tr>
<th>Time (months)</th>
<th>Survival Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.00</td>
</tr>
<tr>
<td>1</td>
<td>0.93</td>
</tr>
<tr>
<td>2</td>
<td>0.86</td>
</tr>
<tr>
<td>3</td>
<td>0.80</td>
</tr>
<tr>
<td>4</td>
<td>0.75</td>
</tr>
<tr>
<td>5</td>
<td>0.70</td>
</tr>
<tr>
<td>6</td>
<td>0.66</td>
</tr>
<tr>
<td>7</td>
<td>0.62</td>
</tr>
<tr>
<td>8</td>
<td>0.58</td>
</tr>
<tr>
<td>9</td>
<td>0.54</td>
</tr>
<tr>
<td>10</td>
<td>0.50</td>
</tr>
</tbody>
</table>

1-Year Survival * = 77.2%
5-Year Survival* = 66.7%
10-Year Survival* = 63.2%
SPATIALLY DISAGGREGATE the Kaplan-Meier Survival Estimate by County in Iowa
Five-Year Survival Rate

Counties

Five-Year Survival Rate

County

- .79 to .75
- .74 to .70
- .69 to .65
- 64 to .60
- 60 to .53
Standardized Survival Rate* Counties

*Age-Adjusted
Standardized Survival Rate

**SIGNIFICANT** Counties

*Age-Adjusted*
How to Create Smoothed Maps

• Spatially Adaptive Filtering
  – Overlapping spatial partitions (usually circles)
  – Partitions share observations
  – Radius increases or decreases based on a user-specified threshold
    • 3600 Months

• For the cases within each spatial partition, we compute:
  – Kaplan-Meier 5-Year Survival Estimates
  – Cox Proportional Hazards Model
    • Statistically validate the geographic distribution of survival
    • Adjust for Age and other individual level covariates
Grid Points and Spatial Partitions

4,000 Grid Points

4,000 Spatial Partitions
Example: 100 points and partitions subset from the 4,000 sample points
Example: SIMULATED Case Distribution
Example: 1 point and partition subset from the 4,000 sample points
Within this partition:

- Five-Year Survival Estimates
- Standardized Survival Rates

Using the cases within each partition, 4,000 independent experiments are conducted to obtain a survival estimate for each of the 4,000 partitions.
Introducing the Standardized Survival Ratio (SSR)

• Mapping the Kaplan-Meier has two major deficiencies
  – Does not assess the statistical significance of the local survival estimates
  – Does not provide a means for age-adjustment

• The Cox Model provides a method to compare two or more Kaplan-Meier Survival estimates

• Comparing the local survival experience with the statewide experience provides a baseline to assess the statistical significance

• Interpretation: Is the survival experience greater or lesser than the statewide experience

• The SSR is calculated using Cox Proportional Hazards Model
Standardized Survival Rate (SSR)
Statistically Significant* Areas of High and Low Survival

* p-value < 0.05
Validation
Using cases within each type of survival area
Survival Rate

Time (months): 0 = Jan 1997 and 144 = December 2008

1-Year Survival * = 77.2%
5-Year Survival* = 66.7%
10-Year Survival* = 63.2%

*Statewide Rate

Low Survival (p < 0.01)
Low Survival (p < 0.05)
Low Survival (p > 0.05)
High Survival (p > 0.05)
High Survival (p < 0.05)
High Survival (p < 0.01)
State-wide Survival Rate
```
. stcox High_p01 High_NS High_p05 Low_p01 Low_NS Low_p05
    failure _d:  crc_death == 1
    analysis time _t:  survival

Iteration 0:  log likelihood = -45877.782
Iteration 1:  log likelihood = -45732.441
Iteration 2:  log likelihood = -45729.295
Iteration 3:  log likelihood = -45729.272
Iteration 4:  log likelihood = -45729.272
Refining estimates:
Iteration 0:  log likelihood = -45729.272

Cox regression -- Breslow method for ties

No. of subjects = 18260  Number of obs    =  18260
No. of failures =   4867
Time at risk    =  856616
Log likelihood  =  -45729.272  LR chi2(6)     =  297.02
                        Prob > chi2    =   0.0000

        _t | Haz. Ratio  Std. Err.     z  P>|z|  [95% Conf. Interval]
-------------+--------------------------------------------------
   High_p01  |   .4170219  .0824634   -4.42 0.000     .2830344    .6144387
   High_NS  |   .8635561  .0393526   -3.22 0.001     .7897706    .9442352
   High_p05  |   .7585186  .0407568   -5.14 0.000     .6826992    .8427584
  Low_p01   |   1.591178  .0772424    9.57 0.000     1.446764    1.750006
  Low_NS   |   1.271585  .0634765    4.81 0.000     1.153066    1.402286
  Low_p05  |   1.402668  .0565693    8.39 0.000     1.296063    1.518042
```

Research Directions

• Goals of this Research
  – To see future atlases of cancer with maps of survival
  – To establish the Standardized Survival Rate (SSR) as the standard map design in future atlases of cancer

• Application of the SSR other areas and/or other types of cancer
  – Iowa’s population tends to be somewhat older, racially homogenous and reside in low population density area

• Multilevel Cox Model to provide potential explanations for spatial patterns of survival
  – Socioeconomic Status and Demographic Characteristics
  – Access to Primary Care or Oncological Specialists
Acknowledgments

• Iowa Health Registry
  – Charles Lynch
  – Daniel Olson

• Iowa Department of Public Health
  – Donna Johnson

• Iowa Cancer Maps Team
  – Gerard Rushton (mentor)
  – Qiang Cai
  – Chetan Tiwari
  – Kirsten Beyer
QUESTIONS?

Kevin Matthews
PhD Student, Health Geography
University of Iowa, Department of Geography
(kevin-matthews@uiowa.edu)
Explaining the Geographic Distribution of Colorectal Cancer Survival: an Iowa example

North American Association of Central Cancer Registries Annual Conference, Louisville KY
June 21, 2011

Kevin Matthews
Kevin-matthews@uiowa.edu
PhD Student, Health Geography
University of Iowa, Department of Geography
Additional Slides
Sample Points
Buffered Sample Points
All 4000 Spatial Partitions
Statistically Significant Sample Points
Buffered Points to reveal Areas of High and Low Survival
Statistically Significant Areas of High and Low Survival
Standardized Survival Rate (SSR)