USING GOOGLE DISTANCE API TO ESTIMATE HEALTH CARE ACCESS
A Case Study from the New York State Congenital Heart Defects Surveillance Project
Tabassum Insaf (Tabassum.insaf@health.ny.gov)
New York State Department of Health/ Center for Environmental Health
University at Albany/ Department of Epidemiology and Biostatistics

Approaches to Evaluating Spatial Accessibility

Approach to calculating access to care in terms of transit

- Straight line distance
- Driving:
  - Driving Distance: Network Analysis
  - Driving Distance / Time (accounting for traffic): Google API (Application Programming Interface) or other routing databases
- Public Transit
  - Distance to Public Transit route or nodes: Mass transit databases
  - Time spent in Public Transit (including walking to and from the transit station): Google API (Application Programming Interface) or other routing databases
Comparing Straight Line, Driving and Public Transit Results

Google API

https://developers.google.com/maps/documentation/distance-matrix/

Google API details

- Use gmapsdistance package and a Google API key to access Google API – developed by Rodrigo Azuero Melo & David Zarruk
  
  https://cran.r-project.org/web/packages/gmapsdistance/gmapsdistance.pdf

- Other packages also available

- Google API issues
  - Free up to 2,500 requests per day.
  - $0.50 USD / 1,000 additional requests, up to 100,000 daily, if billing is enabled.
  - Purchase a commercial license from Google if you anticipate higher use.
  - Current and future traffic based on historical data but past traffic times not available
  - Issues with confidentiality. Only coordinates shared, can use geo-masking but will limit accuracy.
Transit Time
- Calculations are based on time the query was submitted, we assumed that the clinic visits will take place during normal office hours.

CASE STUDY

Rationale
- Differences in travel distance to appropriate care among individuals with birth defects can have a significant impact on healthcare utilization, and it has been suggested that geographic barriers may play a role in cessation of care among those with CHDs.
- Despite these findings, research characterizing geographic access to appropriate care among individuals with CHDs has been limited. One study calculated travel distances from zip codes to pediatric cardiology providers, but there has been no research using individual-level address data to measure distance to care for individuals with CHDs.
Study Details

- Study population to adolescents between the ages of 11 and 19 on January 1st, 2010 who met the surveillance system case definition.
- (ICD-9-CM) codes: 745.XX-747.XX, excluding
  - 746.86 (congenital heart block),
  - 747.1 (pulmonary arteriovenous malformation),
  - 747.6 (other anomalies of peripheral vascular system), and
  - 747.8 (other specified anomalies of the circulatory system) in one or more of the surveillance system data sources.
- We collected available sociodemographic, clinical, and residential address information for each case identified over the 2008-2010 time period.
- For our analyses, we used all residential addresses on record for each case. If a subject had multiple records, we included only those records which had unique addresses.
- There were a total of 10,834 records for 2,522 individuals. Of these 3058 were unique addresses with one address reported for 2,114 cases and two or more addresses reported for 408 cases. 95.4% of addresses were successfully geocoded to the rooftop. Rest (n=140) were geocoded to the zip code centroid.

Defining Facilities

- We validated this definition by assessing percent of cardiac procedures conducted at licensed and unlicensed facilities in the state. 97% of all adolescent cardiac surgical procedures were conducted at the 12 facilities in our analysis.
- Twelve facilities met this definition of surgical care, four of which were within our surveillance region.

Using gDistance package to calculate distance matrix and 3 closest facilities

<table>
<thead>
<tr>
<th>Case ID</th>
<th>Origin</th>
<th>Hospital ID</th>
<th>Destination</th>
</tr>
</thead>
<tbody>
<tr>
<td>10001</td>
<td>40.857421, -74.916066</td>
<td>AAAA</td>
<td>40.857421, -74.916066</td>
</tr>
<tr>
<td>10002</td>
<td>40.857421, -75.916066</td>
<td>BBBB</td>
<td>40.857421, -75.916066</td>
</tr>
<tr>
<td>10003</td>
<td>41.857421, -73.916066</td>
<td>CCC</td>
<td>41.857421, -73.916066</td>
</tr>
<tr>
<td>10004</td>
<td>40.857421, -79.916066</td>
<td>DDDD</td>
<td>40.857421, -79.916066</td>
</tr>
<tr>
<td>10005</td>
<td>46.857421, -73.916066</td>
<td>EEEE</td>
<td>46.857421, -73.916066</td>
</tr>
<tr>
<td>10006</td>
<td>44.857421, -73.916066</td>
<td>FFFF</td>
<td>44.857421, -73.916066</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Case ID</th>
<th>Origin</th>
<th>Hospital ID</th>
<th>Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>10001</td>
<td>40.857421, -74.916066</td>
<td>AAAA</td>
<td>27</td>
</tr>
<tr>
<td>10002</td>
<td>40.857421, -75.916066</td>
<td>BBBB</td>
<td>69</td>
</tr>
<tr>
<td>10003</td>
<td>41.857421, -73.916066</td>
<td>CCC</td>
<td>23</td>
</tr>
<tr>
<td>10004</td>
<td>40.857421, -79.916066</td>
<td>DDDD</td>
<td>72</td>
</tr>
<tr>
<td>10005</td>
<td>46.857421, -73.916066</td>
<td>EEEE</td>
<td>80</td>
</tr>
<tr>
<td>10006</td>
<td>44.857421, -73.916066</td>
<td>FFFF</td>
<td>80</td>
</tr>
</tbody>
</table>
Public Transit Time to Surgical Care

<table>
<thead>
<tr>
<th>Residential Address Characteristics</th>
<th>Predicted one-way drive time to specialty care (minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-poor census tracts</td>
<td></td>
</tr>
<tr>
<td>Southeastern New York</td>
<td>16.38 (15.51, 17.25)</td>
</tr>
<tr>
<td>Urban western New York</td>
<td>24.32 (23.30, 25.34)</td>
</tr>
<tr>
<td>Rural western New York</td>
<td>66.84 (63.50, 70.19)</td>
</tr>
<tr>
<td>Poor census tracts</td>
<td></td>
</tr>
<tr>
<td>Southeastern New York</td>
<td>14.32 (13.54, 15.09)</td>
</tr>
<tr>
<td>Urban western New York</td>
<td>16.72 (15.47, 17.97)</td>
</tr>
<tr>
<td>Rural western New York</td>
<td>68.70 (64.70, 72.70)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Region</th>
<th>Addresses where Public Transit data is available</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Addresses</td>
<td>2301 (75.2)</td>
</tr>
<tr>
<td>Drive time median (IQR), minutes</td>
<td>16.3 (13.3-21.0)</td>
</tr>
<tr>
<td>Public transit time median (IQR), minutes</td>
<td>53.1 (33.3-76.3)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>County</th>
<th>Addresses where Public Transit data is available</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allegany</td>
<td>33 (0)</td>
</tr>
<tr>
<td>Bronx</td>
<td>671 (99.9)</td>
</tr>
<tr>
<td>Cattaraugus</td>
<td>66 (0)</td>
</tr>
<tr>
<td>Chautauqua</td>
<td>96 (1.0)</td>
</tr>
<tr>
<td>Erie</td>
<td>729 (85.5)</td>
</tr>
<tr>
<td>Genesee</td>
<td>52 (81.0)</td>
</tr>
<tr>
<td>Kingman</td>
<td>155 (63.2)</td>
</tr>
<tr>
<td>Monroe</td>
<td>41 (9.8)</td>
</tr>
<tr>
<td>Orleans</td>
<td>41 (9.8)</td>
</tr>
<tr>
<td>New York</td>
<td>264 (93.1)</td>
</tr>
<tr>
<td>NYS</td>
<td>3058 (100.0)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>State</th>
<th>Addresses where Public Transit data is available</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>2301 (75.2)</td>
</tr>
<tr>
<td>Drive time median (IQR), minutes</td>
<td>16.3 (13.3-21.0)</td>
</tr>
<tr>
<td>Public transit time median (IQR), minutes</td>
<td>53.1 (33.3-76.3)</td>
</tr>
</tbody>
</table>
Next Steps

- Statewide hospitalization dataset
- Looking at 11-30 years old (transition to adulthood)
- Actual care received vs. nearest high level care

THANK YOU!!

“In many spheres of human endeavor, from science to business to education to economic policy, good decisions depend on good measurement.”
- Ben Bernanke

SPATIAL AND TEMPORAL ANALYSIS OF GASTROCHISIS IN MASSACHUSETTS AND TEXAS

Mahsa Yazdy, PhD, MPH
Director of the Center for Birth Defects Research and Prevention
Massachusetts Department of Public Health
Background on gastroschisis

- A birth defect where an infant’s bowels protrude from the abdomen wall
- More common in younger mothers
- Often occurs in the absence of other structural birth defects
- Recurrence risk is small
- Concordance in monozygotic twins is low
- Rarely associated with chromosomal anomalies/syndromes

Increasing prevalence of gastroschisis


A possible clue....

Gastroschisis has been observed to occur in clusters
Studying clusters of gastroschisis

- Many have been prompted by clinicians or concerned citizens
- Problems with assessing clustering this way
  - Population under study is often defined around the cluster
  - Case ascertainment can be questionable
  - Often small numbers
  - Limited data collection
- Systematic methods
  - Using birth defect surveillance data

Study aim

To use a rigorous and systematic method to identify if gastroschisis occurs in clusters of space or time and space using population-based data

Data source: surveillance data from MA and TX

- Cases: birth defects surveillance systems
- Controls: randomly selected from birth certificates
- Birth certificate address used as proxy for residence during pregnancy
- Covariates from birth certificates
  - Race/ethnicity, maternal age, education, smoking during pregnancy, and insurance status (in MA only)
Methods

- Generalized Additive Models (GAMs)
  - \( \text{Logit} \{p(x)\} = \alpha + \gamma'z + S(x_1, x_2) \)
  - \( \alpha \) is the intercept
  - \( z \) is a vector of the covariates
  - \( S(x_1, x_2) \) is a bivariate smooth function
  - Holding all the covariates constant and varying location gives us the relationship between location and the outcome adjusted for the covariates

- Space-time analysis: create series of maps over fixed time spans
Methods

• Determining risk factors to adjust for
  1) If risk factor changed the surface of the maps, kept in model
  2) If adding risk factor to model changed the degree of smoothing, kept in the model
• Both methods yielded similar risk factors
Methods

- Global test of significance
  - Testing if location was significant

- Local test of significance
  - Identified locations of statistically increased or decreased risks

Results: Massachusetts

- A total of 156 cases in Massachusetts
- Randomly selected 9,000 in-state births for controls
- Excluded 1 case and 81 controls because could not geocode address
- Total of 155 cases and 8,919 controls
- Births from January 2000 through December 2007

Results: Massachusetts

- Included in final models
  - Maternal age
  - Race/ethnicity

- The following were associated with gastroschisis but not included
  - Maternal education, smoking, and insurance status
Results: Massachusetts

Locations have been geographically altered to preserve confidentiality.

• Concern about excluded mothers – potential bias
• Imputed possible residential locations based on city and zip code
Results: Massachusetts

• Concern about excluded mothers – potential bias
• Imputed possible residential locations based on city and zip code
Results: Texas

- A total 1,756 cases in Texas
- Randomly selected 10,000 births for controls
- Excluded 69 case and 294 controls because could not geocode address
- Total of 1,687 cases and 9,706 controls
- Births from January 1999 through December 2008

Results: Texas

- Included in final models
  - Maternal age and race/ethnicity
- The following were associated with gastroschisis but not included
  - Maternal education and smoking
Imputed results: Texas

Conclusions

• Limitations
  • Residual confounding or spatial confounding
  • Edge effects
  • Using birth address as a proxy
  • Mobility during pregnancy ranges from 12-31%

• Strengths
  • Use of individual-level data so did not have to aggregate to artificial boundaries (e.g., census tract)
  • GAMs allowed for simultaneous adjustment of individual-level covariates
  • The use of smoothing term adapted to changes in population density to account for rural and urban areas

Conclusions

• Elevated odds ratios identified
  • Only Texas had statistically significant hot spots
  • Suggests that gastroschisis does cluster
  • Explore environmental or behavioral factors in areas where identified hot-spots
Acknowledgments

- Collaborators
  - Veronica Vieira, University of California, Irvine
  - Martha Werler, Boston University
  - Marlene Anderka, Massachusetts Department of Public Health
  - Peter Langlois, Texas Department of State Health Services
  - The staff at the Texas Department of Health and the Massachusetts Department of Public Health for their assistance in data preparation and collection


Resources and packages

R package (MapGAM) is available to map risk surfaces using individual-level data such as case-control study data or surveillance data, using generalized additive models (GAMs) to smooth over two-dimensional location while adjusting for any confounding variables.

MapGAM package: http://cran.r-project.org/web/packages/MapGAM/


Synthetic data and code used in the paper can be found at: http://www.bu.edu/sph/research/research-landing-page/superfund-research-program-at-boston-university/projects/project-2-analyzing-patterns-in-epidemiologic-and-toxicologic-data/project-2-resources-and-packages/