Tech Talk

Build Health Equity with Data Science

MARCH 1, 2022
Lisa M. Lines, PhD, MPH
Jamie L. Humphrey, PhD, MPH
Matt Brown, MS
Cindy D'Annunzio, MBA
Determinants of Health

Clinical Factors

Behavioral Factors

Social Factors

RTI Rarity™

an AI solution of RTI International
“What are the social determinants of…”

- Life Expectancy
- Infant Mortality
- Overdose Mortality
- Cancer Mortality
There is a **23 year difference** in life expectancy between these two neighborhoods, less than 2 miles apart.
Area-Based Analyses: Selected Geographic Levels

<table>
<thead>
<tr>
<th>CENSUS BLOCK GROUP – 9-DIGIT ZIP CODES</th>
<th>CENSUS TRACT</th>
<th>ZIP CODE TABULATION AREA – 5-DIGIT ZIP CODES</th>
<th>CITY</th>
<th>COUNTY/PARISH</th>
<th>MSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Block group 080050807002</td>
<td>n=1,442</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tract 080700</td>
<td>n=5,801</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZIP Code 80012</td>
<td>n=52,588</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aurora</td>
<td>n=369,011</td>
<td></td>
<td></td>
<td>Arapahoe County</td>
<td></td>
</tr>
<tr>
<td>Denver Area</td>
<td>n=644,560</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

OTHERS: VOTING DISTRICTS, SCHOOL DISTRICTS, POLICE PRECINCTS, HOSPITAL REFERRAL REGIONS, ETC.

STATE | CENSUS DIVISION | CENSUS REGION
---|-----------------|-----------------
How We Created Local Social Inequity Scores

Conceptual model → Data curation → Random forests → Ranked percentile scores → Validation
Simplified Illustration of Random Forest Algorithm

Splitting on randomly selected variables

Original data

Test sets yield error rates & variable importance

Training set (approx. 2/3)

Test set (approx. 1/3)

Training set

Test set

Training set

Test set

Training set

Test set

Cumulative estimate

Cumulative estimate

Cumulative estimate

Cumulative estimate

Tree 1

Tree 2

...
# Social Determinants of Health (SDoH) Data at RTI

<table>
<thead>
<tr>
<th>Main Federal Data Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CDC:</strong></td>
</tr>
<tr>
<td>- USALEEP project</td>
</tr>
<tr>
<td>- PLACES (21 BRFSS measures for chronic conditions and healthy behaviors)</td>
</tr>
<tr>
<td>- Environmental Public Health Tracking Network</td>
</tr>
<tr>
<td>- Compressed Mortality File</td>
</tr>
<tr>
<td><strong>TidyCensus:</strong> American Community Survey demographic data</td>
</tr>
<tr>
<td><strong>US DOT:</strong> Transportation measures</td>
</tr>
<tr>
<td><strong>USDA:</strong> Food and nutrition data</td>
</tr>
<tr>
<td><strong>EPA:</strong> Walkability index, air pollution</td>
</tr>
<tr>
<td><strong>CMS:</strong> HCRIS data</td>
</tr>
<tr>
<td><strong>FBI:</strong> Uniform Crime Reporting data</td>
</tr>
<tr>
<td><strong>HUD:</strong> Subsidized housing data</td>
</tr>
<tr>
<td><strong>HIFLD:</strong> Infrastructure data</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Other Federal and Private Data Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Opportunity Atlas</strong></td>
</tr>
<tr>
<td><strong>Social Capital Project</strong></td>
</tr>
<tr>
<td><strong>Baseline Resilience Indicators for Communities</strong></td>
</tr>
<tr>
<td><strong>United States Drought Monitor</strong></td>
</tr>
<tr>
<td><strong>RWJF County Health Rankings</strong></td>
</tr>
<tr>
<td><strong>Brandeis Diversity Data Kids – Child Opportunity Index</strong></td>
</tr>
<tr>
<td><strong>Brown University's Longitudinal Tract Data Base</strong></td>
</tr>
<tr>
<td><strong>Mapping the Meal Gap</strong></td>
</tr>
<tr>
<td><strong>MIT Election Data and Science Lab</strong></td>
</tr>
</tbody>
</table>
Visualizing Complex Data
Life expectancy by local social inequity: 5 states
Life expectancy and LSI across the US
Maps of Life Expectancy and Social Inequity: Florida

Life Expectancy Estimates

Local Social Inequity Scores
Maps of Life Expectancy and Social Inequity: United States

Predicted Life Expectancy

Local Social Inequity Scores
Unpacking the LSI scores – Rochester, NY

<table>
<thead>
<tr>
<th>High LSI Census Tract</th>
<th>Low LSI Census Tract</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local Social Inequity Score: 0.98</td>
<td>Local Social Inequity Score: 0.04</td>
</tr>
<tr>
<td>Predicted Life Expectancy: 75.30</td>
<td>Predicted Life Expectancy: 83.10</td>
</tr>
<tr>
<td>Poverty: 73.28%</td>
<td>Poverty: 8.02%</td>
</tr>
<tr>
<td>Disability: 17.45%</td>
<td>Disability: 8.57%</td>
</tr>
<tr>
<td>Average Home Value: $94k</td>
<td>Average Home Value: $260k</td>
</tr>
<tr>
<td>Smoking Prevalence: 36.80%</td>
<td>Smoking Prevalence: 9.80%</td>
</tr>
<tr>
<td>Uninsured: 8.61%</td>
<td>Uninsured: 5.16%</td>
</tr>
<tr>
<td>Unemployment: 20.85%</td>
<td>Unemployment: 4.43%</td>
</tr>
</tbody>
</table>
Where are the 99$^{th}$ percentile tracts?
Census Tract with Highest LSI Score in US

Macon, GA
Bibb County

**Highest LSI Tract**

- Non-White: 96%
- Single Parents: 98%
- Median Household in 1990: $24k
- Median income in 2012-2016: $13k
- Unemployment: 30%
- At or below poverty line: 57%
LSI scores by ZIP code across SIM2 states

Connecticut

Idaho

Ohio
Bivariate Analysis of LSI vs SIM2 & LSI vs VBP - Ohio
## Top Predictors of Life Expectancy Across the US

<table>
<thead>
<tr>
<th>VI rank</th>
<th>Indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Nursing home residents, %, 2010</td>
</tr>
<tr>
<td>2</td>
<td>Mentally unhealthy &gt;14 days per month, % of adults, 2018</td>
</tr>
<tr>
<td>3</td>
<td>Tooth loss prevalence rate, % of adults 65+, 2018</td>
</tr>
<tr>
<td>4</td>
<td>Smoking prevalence rate, %, 2018</td>
</tr>
<tr>
<td>5</td>
<td>Child Opportunity Index, 2010-2015</td>
</tr>
<tr>
<td>6</td>
<td>Probability of reaching the top 20% of income (among children born the same year), 1978-2015</td>
</tr>
<tr>
<td>7</td>
<td>Rural/urban commuting area code</td>
</tr>
<tr>
<td>8</td>
<td>Foreign-born, %, 2010-14</td>
</tr>
<tr>
<td>9</td>
<td>Physically unhealthy &gt; 14 days per month, adults, %, 2018</td>
</tr>
<tr>
<td>10</td>
<td>Teen births per 1k females aged &lt; 20 years, 2010-14</td>
</tr>
</tbody>
</table>
What is driving the algorithm?
Comparison with Other SDoH Composite Measures

Percent variance explained

- Social Vulnerability Index (SVI)
- Area Deprivation Index (ADI)
- Social Deprivation Index (SDI)
- Local Social Inequity (LSI)
State Innovation Models, Round 2 (SIM2): Bivariate associations in Ohio – More hospital care, less primary care in highest-risk ZIP codes

Figure 1. Any inpatient admission by quartile and year

Figure 2. Number of PCP visits by quartile and year

Figure 3. Number of ED visits by quartile and year

n=12.1 million person-years
Validation: Linking LSI Scores with Medicaid Data in Multivariable Triple-Difference Models

Figure 1. Total Spending

Figure 2. Inpatient Admissions
Modeling Other Outcomes
Observed overdose deaths vs. model-predicted overdose deaths

N=4,269
ZIP codes in 4 states
Predictors: 30
RMSE: 3.1
Adj R² = 0.9672

Age-adjusted drug overdose mortality per 100k
Observed cancer deaths vs. model-predicted cancer deaths

Age-adjusted cancer mortality per 100k

N=72,770
US Census tracts
Predictors: 140
RMSE: 7.8
Adj R² = 0.9487

Cancer mortality rate per 100k, 2015-19
Fitted values
Validation vs Other SDoH Composites – Drug Overdose Mortality and Cancer Mortality

- Drug overdoses - 4 states
  - n=4,269 ZCTAs
  - SVI explained 5%
- Cancer mortality - cross-US
  - n=72,100 tracts
  - LSI explained 31%
Risk Adjustment for Social Risk Factors

- Does not remove the requirement to address them
- Does not imply lower standards of care are acceptable
- Does not unfairly penalize providers for factors outside of their control
- Does not unfairly reward providers simply for being in better-off/less at-risk communities
Use Cases for RTI Rarity Data and Scores

- Control for social factors in analyses of access to care, quality, and cost
- Use in evaluations to identify patterns of higher-risk communities
- Target interventions to higher-risk populations
- Inform policies to address root causes
The Value

Value

AI

RA

Equity

Place
TechTalk

Questions and Discussion

Contact cdannunzio@rti.org for more on how RTI Rarity can help you meet your project goals

rti.org/TechTalk