State Experiences with the Environmental Public Health Tracking Program

Steve Anderson
New Jersey Department of Environmental Protection

MARAMA
Data Analysis Workshop
College Park, Maryland
January 19, 2011
Overview

• Background on CDC Environmental Public Health Tracking Program

• Linkage Analysis of Benzene (NATA) and Leukemia

• Evaluation of EPA Hierarchical Bayesian predictions for PM2.5 and Ozone
What is EPHT?

• Funded by the Centers for Disease Control and Prevention

• Include both Health and Environmental Agencies

• “the ongoing collection, integration, analysis, and interpretation of data about environmental hazards, exposure to environmental hazards, and human health effects potentially related to exposure to environmental hazards”
  – Not a single study where you design a survey and develop new data to answer a specific question

• Rely on existing data tracked on a broader scale
  – Identify trends
  – Identify issues requiring additional research
  – Provide public information

• Goal is to ultimately build a national network with consistent health and environmental data
Environmental Public Health Tracking

Tracking Network

Collection → Integration → Analysis → Dissemination

Assessment → Research → Policy → Intervention → Prevention

Stakeholders:
- Federal Agencies
- State and Local Agencies
- Academia
- Health Care System
- Non-Governmental Organizations
- Business and Industry
- Policy Makers
- Media
- Public

Hazard → Exposure → Health Effect

Ongoing Evaluation

Improved Public Health

Department of Health and Human Services
Centers for Disease Control and Prevention
Safer • Healthier • People
State and Academic Partners
National Network

CDC Home
Centers for Disease Control and Prevention
Your Online Source for Credible Health Information

National Environmental Public Health Tracking Network

Home > Environments > Outdoor Air

Outdoor Air

Air pollution has been linked to many health problems such as asthma, heart disease, and adverse reproductive outcomes such as low birth weight.

Air Quality

National air quality has improved since the 1990’s, but many challenges remain in protecting public health and the environment from air quality problems.

Air and Health

Air Monitoring in the US

Air Contaminants

Tracking Hot Topics

- View our Network Tutorial Videos
- Learn how to protect your children from lead in toys
- Find out how CDC tracks lead exposure in children
- CDC Tracking Data
- Learn more about Carbon Monoxide and your health
- CDC Information
- Find out how CDC tracks Carbon Monoxide poisoning
- CDC Tracking Data
Schedule

- **Phase 1:**
  - 3 Pilot Demonstration Projects linking environmental data with health data
    - Cancer
    - Childhood blood lead
    - Birth defects

- **Phase 2:**
  - Focused on Network Implementation
  - NJ Network went “Live” December 2008 (at NJDHSS web site)

- Future Request for Proposals expected
Demonstration Projects

Goals: Develop and evaluate methods for linking health effects surveillance data with existing data for environmental hazards and exposures.

<table>
<thead>
<tr>
<th>Linkage Pilot</th>
<th>Health Outcome</th>
<th>Environmental Metric</th>
<th>Geographic Scale</th>
<th>Time Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cancer</td>
<td>Leukemia</td>
<td>Benzene NATA</td>
<td>Census Tract</td>
<td>1996 (estimate 1979 to 2002)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NJDEP Benzene Estimate</td>
<td>100 km grid</td>
<td>Avg Releases 1988 to 2003</td>
</tr>
<tr>
<td></td>
<td>Liver (angiosarcoma)</td>
<td>Vinyl Chloride NATA</td>
<td>Census Tract</td>
<td>1996 (estimate 1979 to 2002)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NJDEP Vinyl Chloride Estimate</td>
<td>100 km grid</td>
<td>Avg Releases 1988 to 2003</td>
</tr>
<tr>
<td></td>
<td>Bladder</td>
<td>THM Drinking Water</td>
<td>Water Purveyor Area</td>
<td>1979 - 1985</td>
</tr>
<tr>
<td>Lead</td>
<td>Childhood blood lead</td>
<td>Lead NATA</td>
<td>Census tract</td>
<td>1999</td>
</tr>
</tbody>
</table>
Outline of Linkage Project

- Exposure Assessment/Categories
- Cancer Incidence Data
- Quantify relationship between exposure and Leukemia
Some Challenges

• **Exposure Assessment**
  - **Latency**
    • period between cancer-initiating exposure and diagnosis of cancer that may be many years or decades long
  - **Migration**
    • people move frequently, especially over long periods of time
  - **Information on Confounders**
    • Most important factor that affects the rate of cancer is age
    • tobacco smoking history is not available
    • occupational history
    • diet

• **Health Outcome**
  - **Variability and Stability of Rates of Rare Diseases**
    • Most cancers considered rare diseases for statistical purposes
    • Need larger areas and longer time periods
  - **Geocoding of cancer incidence cases**
  - **Confidentiality**
    • Any information that could identify an individual cannot be released to anyone
    • Research (Institutional Review Board) vs. Surveillance
Distribution of NJ Benzene 1996 NATA

- All census tracts above the 10-6 cancer risk used as health benchmark

- Question: How to categorize census tracts into different exposure groups?
  - How many exposure categories?
  - Based on risk ranges? statistics?
Benzene Binary Cut Points

<table>
<thead>
<tr>
<th>Benzene Exposure Groups</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>.649</td>
<td>1.50</td>
</tr>
<tr>
<td>1.51</td>
<td>4.57</td>
</tr>
</tbody>
</table>

Legend
- NJ Counties
- Benzene (ug/m^3)
  - Yellow: 0.649 - 1.50
  - Red: 1.51 - 4.57
Benzene Three Risk Based Cut Points

<table>
<thead>
<tr>
<th>Benzene Exposure</th>
<th>Cancer Risk</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>4.99 to 14.99</td>
<td>Less than 1.5x10^-5</td>
</tr>
<tr>
<td>High</td>
<td>15.00 to 24.99</td>
<td>1.5x10^-5 - 2.4x10^-5</td>
</tr>
<tr>
<td></td>
<td>25.00 to 35.2</td>
<td>Greater than 2.5x10^-6</td>
</tr>
</tbody>
</table>

Legend
- NJ Counties
- Risk Ratio
  - 4.99 - 14.99
  - 15.00 - 24.99
  - 25.00 - 35.2

13
Exposure Categories

Benzene Concentration (ug/m³)
Cut points

Legend

Benzene ug/m³
- 0.649 - 1.400
- 1.400 - 2.500
- 2.500 - 4.570

Counties

- High exposure determined by statistical natural break (Jenks method in GIS)
- Medium and low cut point based on equal populations

Exposure Level (ug/m³) | Population
--- | ---
<1.4 | 3,785,122
1.4 - 2.5 | 3,781,901
2.5 - 4.57 | 845,494
Evaluation of Leukemia Data

- Data source: New Jersey State Cancer Registry (NJSCR)
- Time period for data: 1979 – 2002 (NATA is for 1996)
- Geographic scale: Census tract (NATA is census tract)
- Case Definition

Cases identified at autopsy or via death certificate were excluded because of lack of address information and uncertainty about the address of residence at diagnosis.

<table>
<thead>
<tr>
<th>Case Eligibility</th>
<th>Leukemia</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eligible</td>
<td>21,861</td>
<td>90.90%</td>
</tr>
<tr>
<td>Not Eligible*</td>
<td>2,201</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>24,062</td>
<td></td>
</tr>
</tbody>
</table>
Geocoding of Leukemia

- High Accuracy geocoding are cases that can be coded to X,Y points using the full address from the cancer registry, which is used to assign a census tract.

- Other cases can be assigned a census tract using the zip code centroid.
  - Many of these cases are in rural areas with PO Boxes instead of street addresses.
  - Disproportionately located in low exposure areas (selection bias).

<table>
<thead>
<tr>
<th>Geocode Success</th>
<th>Leukemia</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Accuracy</td>
<td>20,130</td>
<td>92.10%</td>
</tr>
<tr>
<td>Zip Code Centroid</td>
<td>1,579</td>
<td>7.20%</td>
</tr>
<tr>
<td>Uncodable</td>
<td>152</td>
<td>0.70%</td>
</tr>
<tr>
<td>Total Cases</td>
<td>21,861</td>
<td></td>
</tr>
</tbody>
</table>
Quantify relationship between exposure and Leukemia

• Census tract is the unit of analysis.
• Dataset contained
  – case counts
  – population estimates
  – Population by sex and age group.
  – Exposure metrics

• Cancer incidence rate ratios (RRs) were computed for each cancer type for levels of exposure metrics.
  – Rate Ratio = Incidence Rate / Incidence Rate in low exposure group

• Computed using the Poisson regression model.
  – a form of regression analysis used to model count data
  – Confidence intervals (95%) were generated for the RR estimates.

• RRs were adjusted for age group, sex, the percent of the population below the poverty level, and the percent of the population which is white.
Initial Linkage Results
(High Accuracy Geocoding only)

**Benzene Concentration (ug/m3) Cut points**

- 0.13 ug/m3 = 10^-6 health
- 1.4 ug/m3
- 2.5 ug/m3

**Legend**

- Yellow: 0.649 - 1.400
- Orange: 1.400 - 2.500
- Red: 2.500 - 4.570
- White: Counties

<table>
<thead>
<tr>
<th>Variable</th>
<th>Population</th>
<th>RR</th>
<th>95% CI</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benzene &lt; 1.40 μg/m³</td>
<td>3,785,122</td>
<td>1.00</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Benzene 1.40 - 2.49 μg/m³</td>
<td>3,781,901</td>
<td>1.09</td>
<td>1.06 - 1.12</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Benzene 2.50+ μg/m³</td>
<td>845,494</td>
<td>1.04</td>
<td>0.98 - 1.09</td>
<td>0.214</td>
</tr>
</tbody>
</table>
# Initial Linkage Results
**(High Accuracy Geocoding only)**

<table>
<thead>
<tr>
<th>Benzene Exposure</th>
<th>Average Annual Population</th>
<th>Cases</th>
<th>Adjusted RR</th>
<th>p-Value</th>
<th>95% Conf. Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2 Cut Points, Females only:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (≤1.39 μg/m³)</td>
<td>1,710,551</td>
<td>3,600</td>
<td>1.0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2 (&gt;1.39 ≤2.49 μg/m³)</td>
<td>1,923,059</td>
<td>4,332</td>
<td>1.082</td>
<td><strong>0.001</strong></td>
<td>1.033 1.133</td>
</tr>
<tr>
<td>3 (&gt;2.49 μg/m³)</td>
<td>427,663</td>
<td>897</td>
<td>1.114</td>
<td><strong>0.009</strong></td>
<td>1.028 1.208</td>
</tr>
<tr>
<td><strong>2 Cut Points, Males only:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (≤1.39 μg/m³)</td>
<td>1,624,819</td>
<td>4,755</td>
<td>1.0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2 (&gt;1.39 ≤2.49 μg/m³)</td>
<td>1,767,606</td>
<td>5,551</td>
<td>1.096</td>
<td><strong>0.000</strong></td>
<td>1.053 1.142</td>
</tr>
<tr>
<td>3 (&gt;2.49 μg/m³)</td>
<td>397,823</td>
<td>995</td>
<td>0.974</td>
<td>0.491</td>
<td>0.904 1.050</td>
</tr>
<tr>
<td><strong>Continuous, Females only:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Per μg/m³ increase</td>
<td>4,061,273</td>
<td>8,829</td>
<td>1.066</td>
<td><strong>0.001</strong></td>
<td>1.028 1.105</td>
</tr>
<tr>
<td><strong>Continuous, Males only:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Per μg/m³ increase</td>
<td>3,790,247</td>
<td>11,301</td>
<td>1.030</td>
<td>0.072</td>
<td>0.997 1.064</td>
</tr>
</tbody>
</table>
**FINAL Linkage Results**
(Add zip code centroid cases)

**Legend**

<table>
<thead>
<tr>
<th>Benzene ug/m3</th>
<th>RR</th>
<th>95% CI</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benzene &lt; 1.40 μg/m³</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Benzene 1.40 - 2.49 μg/m³</td>
<td>1.03</td>
<td>1.06</td>
<td>0.083</td>
</tr>
<tr>
<td>Benzene 2.50+ μg/m³</td>
<td>0.97</td>
<td>1.03</td>
<td>0.298</td>
</tr>
</tbody>
</table>

---

**Benzene Concentration (ug/m3)**

Cut points:
- 0.13 ug/m3 = 10-6 health
- 1.4 ug/m3
- 2.5 ug/m3

**95% CI P-value**
- Benzene < 1.40 μg/m³: 1 - -
- Benzene 1.40 - 2.49 μg/m³: 1.03 1.06 0.083
- Benzene 2.50+ μg/m³: 0.97 1.03 0.298
### FINAL Linkage Results
(Add zip code centroid cases)

<table>
<thead>
<tr>
<th>Benzene Exposure</th>
<th>Average Annual Population</th>
<th>Cases</th>
<th>Adjusted RR</th>
<th>p-Value</th>
<th>95% Conf. Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>2 Cut Points, Females only:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (≤1.39 μg/m³)</td>
<td>1,710,551</td>
<td>4,032</td>
<td>1.0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2 (&gt;1.39 ≤2.49 μg/m³)</td>
<td>1,923,059</td>
<td>4,586</td>
<td>1.024</td>
<td>0.229</td>
<td>0.985 1.065</td>
</tr>
<tr>
<td>3 (&gt;2.49 μg/m³)</td>
<td>427,663</td>
<td>937</td>
<td>0.913</td>
<td>0.015</td>
<td>0.849 0.982</td>
</tr>
<tr>
<td><strong>2 Cut Points, Males only:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (≤1.39 μg/m³)</td>
<td>1,624,819</td>
<td>5,334</td>
<td>1.0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2 (&gt;1.39 ≤2.49 μg/m³)</td>
<td>1,767,606</td>
<td>5,779</td>
<td>1.030</td>
<td>0.192</td>
<td>0.985 1.076</td>
</tr>
<tr>
<td>3 (&gt;2.49 μg/m³)</td>
<td>397,823</td>
<td>1,041</td>
<td>1.047</td>
<td>0.249</td>
<td>0.968 1.133</td>
</tr>
<tr>
<td><strong>Continuous, Females only:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Per μg/m³ increase</td>
<td>4,061,273</td>
<td>9,555</td>
<td>1.020</td>
<td>0.279</td>
<td>0.984 1.056</td>
</tr>
<tr>
<td><strong>Continuous, Males only:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Per μg/m³ increase</td>
<td>3,790,247</td>
<td>12,154</td>
<td>0.984</td>
<td>0.319</td>
<td>0.953 1.016</td>
</tr>
</tbody>
</table>
Summary of Cancer Linkage Pilot

- Collaboration between health and environmental agencies
- Broad data sets of environmental and health outcome data can be successfully evaluated
- Highlighted issues with differential loss of cases due to geocoding
  - rural areas in New Jersey have lower estimated benzene levels and higher rates of cases with addresses that cannot be exactly geocoded.
Evaluation of EPA Hierarchical Bayesian predictions for PM2.5 and Ozone

• Background
• County level Indicators
• Comparisons
  – With AQS FRM Monitors
  – With TEOM monitors
  – Variation within grids
• Health Impact Assessment
Background on HB data

• Public Health Air Surveillance Evaluation (PHASE)
  – Project with NY, Wisc., Maine
    • Multi-state (ME-NY-WI)
    • Multi-agency (CDC, EPA)
  – Air Quality characterization evaluation
    • How to apply regulatory data to public health
  – Methods development
    • How to link health and hazard data for analysis

• Evaluate different methods of characterizing air quality data that can be routinely linked with public health data
  – CMAQ alone
  – GIS interpolation of AQS data (kriging)
  – Combinations of data

• Hierarchical Bayesian approach selected as best
  – Combines AQS FRM monitoring data with CMAQ modeling
  – Spatial Resolution at 36km grid for US and 12km grid for Eastern US
  – Provides daily estimates
  – Results for 2001 to 2006
  – available on EPA website http://www.epa.gov/nerlesd1/land-sci/lcb/lcb_sfads.html
For more Information

Environmetrics
2010; 21: 48–65
Published online 18 August 2009 in Wiley InterScience (www.interscience.wiley.com)
DOI: 10.1002/env.984

- Nancy J. McMillan1*, y, David M. Holland2, Michele Morara1 and Jingyu Feng1
- 1 Statistics and Information Analysis, Battelle, 505 King Avenue, Columbus, OH 43201, U.S.A.
- 2U.S. Environmental Protection Agency, Office of Research and Development, Research Triangle Park, NC 27711-0111, U.S.A.
12 km Grids MARAMA Region

- GIS transformations needed to convert X/Y centroids to grids
• 3 different methods evaluated

Option 1: population weighted centroid for county

Option 2: Average of all grid centroids located in a county

Option 3: Weighted average of areas for all grids in the county

• Current indicators use Option 1
• HB data can fill in gaps of the monitoring system
• Indicators on CDC National Network
County Indicators—New Jersey

Title: Annual average ambient concentrations of PM 2.5 in micrograms per cubic meter, based on seasonal averages and daily measurement (monitor and modeled data)

Geographic Resolution: One State All Counties NJ

Time Resolution: 1 year : 2006

Map showing the distribution of PM 2.5 concentrations across New Jersey, with color codes indicating different concentration levels.

- Yellow: 10.1 - 10.4
- Green: >10.4 - 11.5
- Light Blue: >11.5 - 12.3
- Blue: >12.3 - 12.9
- Dark Blue: >12.9 - 14.2

Legend: Modeled data exclusively used to create measures.
New Jersey Grids and Monitors

Legend

- TEOM
- FRM Monitor

GRIDSnj
Counties
States
Comparison for New Jersey 2005 all FRM sites

2005 PM2.5 Comparison

$y = 0.8808x + 1.2626$

$R^2 = 0.9191$

AQS Monitor Value

HB Prediction

Under Prediction

Over Prediction
Comparison of Daily HB to Monitor Results

Daily Monitor-level Comparison

36-km and Monitor Comparison
Year: 2005

Daily NAAQS: 35 µg/m³

Over prediction

AQS PM<sub>2.5</sub> (µg/m³)

Under prediction

12-km and Monitor Comparison
Year: 2005

Daily NAAQS: 35 µg/m³

Over prediction

AQS PM<sub>2.5</sub> (µg/m³)

Under prediction
Comparison TEOM Site (Flemington)

2006 TEOM Comparison (Flemington)

\[ y = 0.787x + 3.329 \]

\[ R^2 = 0.6443 \]
Comparison TEOM Site (Millville)

\[ y = 0.6317x + 2.9966 \]
\[ R^2 = 0.6559 \]
Variation within a grid

Legend
- TEOM
- FRM Monitor
- GRIDSnjcounties
- Counties
- States

Staten Island

0.4 Miles

Rahway
Elizabeth Lab
Elizabeth-Mitchell Building
Variation within a grid

PM2.5 (ug/m³)

Day

PM25_PRED
FRM4eliz
FRMrahway
TEOMRahway
TEOMElizabeth
Health Impact Assessment

• PHASE products/outcomes:
  – EPA (& state) investigation of novel air quality estimating technique - EPA delivered monitor, CMAQ and fused AQ data
  – Study of “case-crossover” method with “how to” guide
  – Multi-state comparisons of AQ-health associations
  – Interactions with state public health agencies - Example of effective interagency collaboration

• Associations (preliminary) between AQ (8-hour O$_3$ and PM$_{2.5}$) and asthma, and AQ (PM$_{2.5}$) and myocardial infarction through case-crossover analyses
EPHTN Air Quality Health Impact Pilot Project

• Collaboration of CDC, EPA, Emory University (Coordinating Center), and four EPHTN grantees (Oregon, Florida, New York State and New York City)

• Development, testing, and application of consistent protocol for local HIA

• Will use EPA’s BenMAP
EPHTN Air Quality Health Impact Pilot Project

• CR function choice, sensitivity analyses
• Baseline and comparison pollution surfaces:
  – One common comparison scenario (e.g. policy-relevant background)
  – One or more scenarios of local interest
• Template for dissemination of results
• Two-year project launched this past summer
Near and Longer Term Goals for Pilot Projects

• Developing a consistent approach
• Routinely applying the most timely and locally relevant data to air quality health impact and benefits analyses
• Establishing or enhancing ties between with air quality managers and health agencies
• Improving public and stakeholder awareness of the public health dimensions of air pollution and the benefits of control measures.
Summary

• Collaboration between health and environmental agencies
• Environmental data used for health assessment
• Opportunities for more communication and involvement with air programs
THANK YOU

QUESTIONS???????