Airborne mammary carcinogens and breast cancer risk in the Sister Study

Nicole Niehoff, PhD MSPH
Postdoctoral fellow
Environment and Cancer Epidemiology Group
National Institute of Environmental Health Sciences

Hazardous Air Toxics

• 187 pollutants that are known or suspected to be carcinogenic or cause other serious health or environmental effects
• Distinct from criteria air pollutants (PM, O$_3$, CO, NO$_2$, Pb, SO$_2$)
• There are no nationwide ambient air quality standards for air toxics
• Numerous ambient sources:

EPA 2016, EPA 2017
Consideration of Multipollutant Exposures

• Exposure does not occur to single pollutants in isolation
  ➢ Joint effects of multiple pollutants may increase severity
  ➢ Exposures of interest may be correlated

• NIEHS (2011, 2015) and EPA (2016) have called interest to mixtures:
  EPA: “multi-pollutant control programs can save money and time, and achieve significant health, environmental and economic benefits, while reducing costs and burdens on sources of air pollution”

• There are a variety of methods available- it’s important to specify what question you are interested in evaluating
Biological Mechanisms: Air toxics and breast cancer
Carcinogenic Air Toxics

- Published review identified 216 chemicals associated with mammary gland tumors in at least one animal study
  - 29 are air toxics and available in the most complete nationwide data source of modeled concentrations, the National Air Toxics Assessment (NATA)

<table>
<thead>
<tr>
<th>Chemical</th>
<th>Carbon tetrachloride</th>
<th>Polycyclic organic matter</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,2-Dibromo-3-chloropropene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1,3-butadiene</td>
<td>Chloroprene</td>
<td>Propylene dichloride</td>
</tr>
<tr>
<td>1,4-dioxane</td>
<td>Ethylbenzene</td>
<td>Propylene oxide</td>
</tr>
<tr>
<td>2,4-dinitrotoluene</td>
<td>Ethylene dibromide</td>
<td>Styrene</td>
</tr>
<tr>
<td>2,4-toluene diisocyanate</td>
<td>Ethylene oxide</td>
<td>Toluene</td>
</tr>
<tr>
<td>2-chloroacetophenone</td>
<td>Ethylidene dichloride</td>
<td>Vinyl chloride</td>
</tr>
<tr>
<td>Acrylamide</td>
<td>Hydrazine</td>
<td>Vinylidene chloride</td>
</tr>
<tr>
<td>Acrylonitrile</td>
<td>Methylene chloride</td>
<td>Xylenes</td>
</tr>
<tr>
<td>Benzene</td>
<td>Nitrobenzene</td>
<td></td>
</tr>
<tr>
<td>Benzidine</td>
<td>o-Toluidine</td>
<td></td>
</tr>
</tbody>
</table>

Rudel 2007
The Sister Study

• Prospective observational cohort
  • 50,884 women, recruited from 2003-2009
  • Ages 35-74 at enrollment
  • Sister had been diagnosed with breast cancer, but no prior breast cancer diagnosis themselves at enrollment

• Excluded women without baseline address geocoded at census tract-level for linkage to exposure data and women with breast cancer diagnosis before enrollment was complete → n=49,718 included

• 2,975 breast cancer events (invasive or ductal carcinoma \textit{in situ}) through September 2016 (an average of 8.4 years after enrollment)
Certain Air Toxics were Associated with an Increased Risk of Breast Cancer

From a Cox proportional hazards model adjusted for age, race, residence type (urban/suburban/small town/rural), education, smoking
The Relationship Between Air Toxics and Breast Cancer was Stronger Among Overweight or Obese Individuals
10% of correlations >0.7
18% of correlations >0.5
Strongest: Ethylbenzene & xylenes (r=0.98)
Weakest: Ethylene dibromide & xylenes (0.001)
• Goal: Examine whether there are combinations of pollutants that may be more or less harmful for breast cancer than would be expected based on exposure to a single pollutant

• Classification and Regression Trees (CART)
  ➢ Classification trees: used for discrete outcomes (i.e. breast cancer)
  ➢ Regression trees: used for continuous outcomes
  ➢ A forward-selection, recursive partitioning approach
### Gini Index:
- Based on impurity functions
- Selects the variable resulting in binary groups that are most different with respect to the outcome

### Splitting Criteria
- Gini Index:
  - Based on impurity functions
  - Selects the variable resulting in binary groups that are most different with respect to the outcome

### Stopping Criteria
- Minimum # of cases in a node = 5
- Maximum number of levels on a branch = 5
- Total number of terminal nodes = 11

Lemon 2003, Loh 2011, Yohannes 1999
Conclusions

• Certain air toxics were associated with a higher risk of breast cancer
  - Methylene chloride, POM, propylene dichloride, and styrene
  - Biologically plausible: IARC group 1 or 2A; chromosomal instability, DNA damage, oxidative stress and inflammation, estrogenic

• These air toxics, with the exception of POM, were part of multipollutant groups that were identified in the classification tree
  - Methylene chloride was the highest on the tree

• Single pollutant analyses were stronger among those who were overweight or obese
  - BMI was used in the formation of branches with certain air toxics on the classification tree

Impact

• Ambient air toxic exposure is widespread
  ➢ Regulation of air toxics on a national scale is currently non-existent
  ➢ Estimation of air toxic concentrations has limitations

• Breast cancer is the most common cancer among women

• CART easily handles non-linear and non-additive associations
  ➢ Informed cut-points that may have been missed with traditional regression
  ➢ Identified high levels that may be important, but may impact a small number of women
  ➢ Investigator-driven parameters

• The findings from the classification tree may reflect harmful co-exposures for breast cancer of interest for future evaluation
Acknowledgements

• Co-authors
  Marilie Gammon, UNC
  Alexander Keil, UNC/NIEHS
  Hazel Nichols, UNC
  Lawrence Engel, UNC
  Dale Sandler, NIEHS
  Alexandra White, NIEHS

• Funding
  Intramural research program at NIEHS
  Cancer Control Education Program (T32CA057726)
  NIEHS Environmental Training Grant (T32ES007018)
Thank you!

Email: nicole.niehoff@nih.gov
Twitter: @nikkiniehoff
Exposure Assessment: National Air Toxics Assessment

• NATA is the only nationwide data source for air toxics

• 2005 version of the NATA was used in this dissertation
  ➢ In the middle of the enrollment period for the Sister Study
  ➢ Incorporates important assessment changes compared to previous years

• Source categories:
  ➢ Point (e.g. large factories, waste incinerators, airports)
  ➢ Non-point (e.g. prescribed burns, dry cleaners, small manufacturers)
  ➢ On-road mobile (e.g. cars, trucks, buses)
  ➢ Non-road mobile (e.g. airport ground support, trains, boats)
  ➢ Background and secondary formation
Data source inputs to create National Emissions Inventory
- state and local inventories
- existing databases from EPA regulatory programs
- emission factors and activity data
- revisions to source inventories from Risk and Technology Review
- EPA analyses supporting standard development

National Mobile Inventory Model (NMIM)
- consolidation of two models: Mobile Source Emission Factor Model (MOBILE) and NONROAD model
- vehicle, activity, and fuel data from states and federal agencies

Mobile Source National Emissions Inventory (NEI)

Dispersion model: HEM-3
Meteorology data
Release parameters
Background +
Point source ambient concentrations

Dispersion model: ASPEN
Meteorology data
Release parameters
Background +
Non-point source ambient concentrations

Dispersion model: HEM-3
Meteorology data
Release parameters
Background +
Mobile source ambient concentrations
CART Splitting Criteria

• Gini improvement measure
  1. Gini diversity index is calculated as $2p_{ij}(1-p_{ij})$ for the parent node and two child nodes
  2. Weighted diversity index of the two child nodes based on the proportion of the observations that end up in each node from the parent node
  3. Gini improvement measure = (parent node diversity index) – (weighted diversity index)

➢ All exposure variables are examined and the one (and its cut-point) that leads to the highest value of the Gini improvement measure is selected as the splitting point

Lemon 2003, Loh 2011, Yohannes 1999