Cancer Care Ontario
Action Cancer Ontario

Small Area Analysis of Cancer Incidence and Behavioural Risk Factors in the Erie-St. Clair Region

Dr. Laura Seliske[1] & Dr. Eric Holowaty[2]

1. Cancer Care Ontario

2. Dalla Lana School of Public Health, University of Toronto

Ontario
Cancer Care Ontario
Action Cancer Ontario

Conflict of Interest Declaration

We declare that there are no conflicts of interest to disclose regarding this presentation

Presentation Objectives

1. To understand the statistical approaches and caveats related to small area mapping and spatial analysis
2. To be familiar with public health spatial data sources and their strengths and limitations
3. To understand the multidisciplinary nature of public health spatial surveillance
4. To anticipate and adapt to a changing landscape in public health, information technology, legislation and policy formulation, and growing expectations for more granular spatial information

Project Team

- Dr. John McLaughlin, Samuel Lunenfeld Research Institute, Mount Sinai Hospital (PI)
- Dr. Eric Holowaty, Dalla Lana School of Public Health, University of Toronto (Co-PI)
- Todd Norwood, Cancer Care Ontario (Staff Scientist)
- Dr. Laura Seliske, Cancer Care Ontario (Research Associate)
- Crystal Palleschi, Lambton Public Health
- Susan Wang, Cancer Care Ontario (Biostatistician)

Background

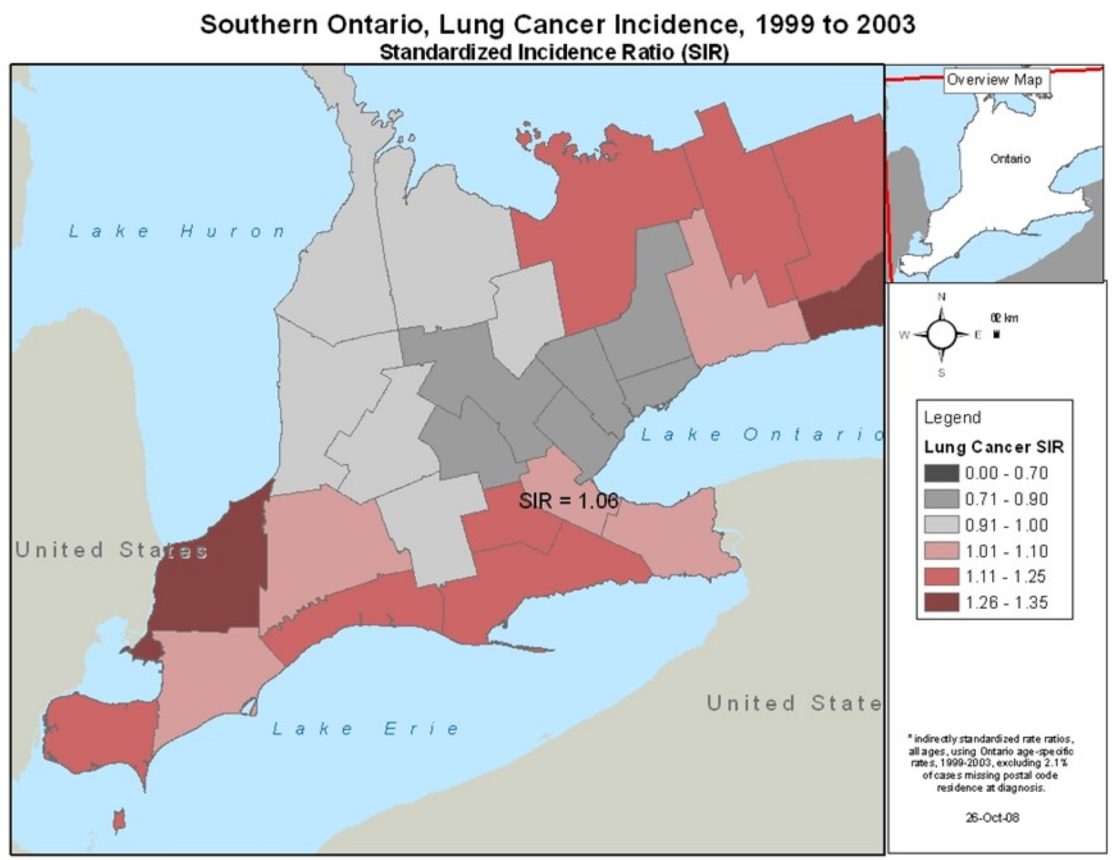
- Increased interest and use of GIS
- Computing power and software more accessible
- Georeferenced data readily available
- Rapid hazard appraisal and more granularity in community health profiling
- Advances in spatial analysis

Rationale for Erie St Clair LHIN

Erie-St. Clair region selected because:

- Widespread concerns about adverse effects of the environment
- A recognized need for further research to address these concerns
- Higher than expected rates for certain cancers across the region as a whole
- Existing collaborative relationships

Why small area/neighbourhood-level maps?



The Cancer Research Society (CRS) Project

“Advanced spatial analyses to characterize environmental impacts on cancer risk: Phase 1”

Objectives:

i. To determine neighbourhood-level incidence of cancer and behavioural risk factors (e.g. smoking, diet, physical activity, etc.) in the Erie St. Clair LHIN

ii. After adjustment for known risk factors, determine areas where cancer incidence remains elevated

iii. To provide an understanding of how these methods could be applied with occupational & environmental exposures

Data Sources

Health Event data

- Ontario Cancer Registry 2004-2008

Statistics Canada Census 2006

- Census geography
- Area classification
- Income quintiles

Link data

Postcode - EA/DA
CT, CMA/CA
CSD, CD boundaries
PCCF+

Tools
Methods

ArcGIS
R
WinBUGS

Canadian Community Health Survey

- Smoking prevalence
- Obesity
- Alcohol use
- Etc…

Environmental data

- National Pollutant Release Inventory
- Land Use Regression Models

Methods

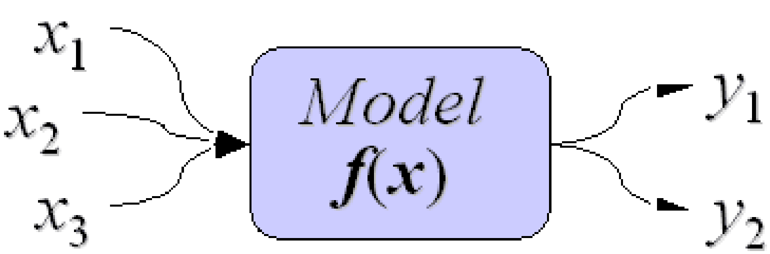
- Level of geography = Dissemination Area (DAs); population 400-700;
- Cancer incidence in Erie St. Clair LHIN compared to all of Ontario
- Bayesian methods used to create:
    - Smoothed standardized incidence ratios (SIRS) for the cancer outcomes
    - Prevalence of behavioural risk factors (e.g. smoking, alcohol use, obesity, etc.)
    - Incorporate behavioural risk factor estimates as covariates for cancer incidence models

Statistical Modeling Approach

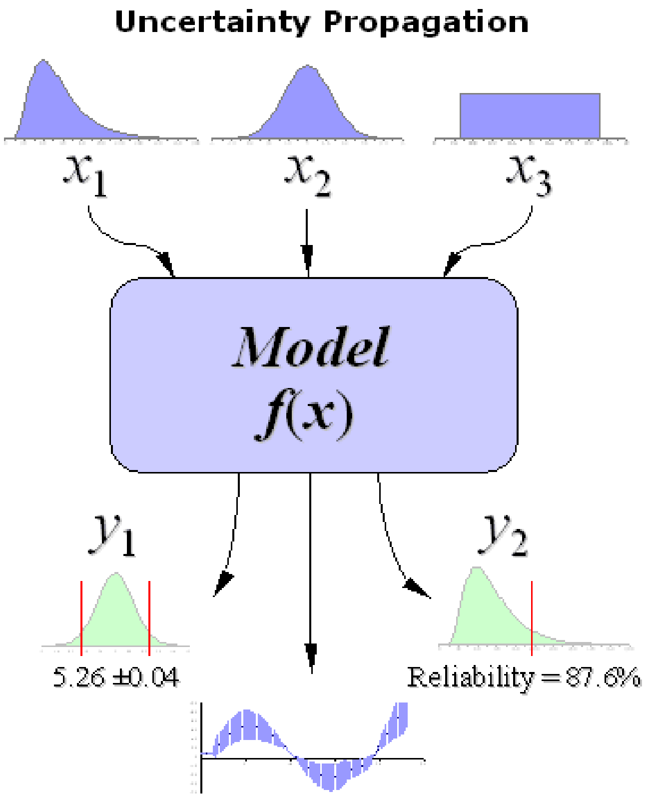
- **Problems:**
  - Small areas with a small number of cases (0,1,2…) produce variable/implausible SIRs
  - Spatial dependence
- **One solution:** Bayesian smoothing using hierarchical random effects models to detect areas at truly higher risk:
  - Allow for uncertainty due to low counts
  - Use spatial dependence to pool information from neighbouring areas
  - Can be used for many statistical distributions
  - Model diagnostics and goodness of fit statistics

Fixed vs. Random Effects

Deterministic model



Stochastic uncertainty



Challenges: Small Area Spatial Analysis

- Accuracy, granularity and completeness of exposure, health and population data, and boundary files
- Geocoding
- Modifiable Areal Unit Problem (MAUP)
- Current place of residence as a proxy for past exposure
- Problems adjusting for known confounders
- Necessity of using aggregated counts
- Data access and confidentiality restrictions

Cancer Types Included in the CRS Project

- 24 cancer types associated with environmental and occupational exposure
  - Examples: lung, colorectal, mesothelioma
- Some cancers provided unstable results, and were excluded from further analysis
  - Cervical cancer
  - Larynx cancer in females
  - Multiple myeloma in males
  - Stomach cancer in females
- Cancer types analyzed separately by sex where possible

Selected Cancer Incidence Results for the Erie St. Clair LHIN

- Maps for the LHIN by 2006 Census Dissemination Area (DA), adjusted for age and sex only
- Cancer incidence (SIRs):
  - DAs in red hues had incidence rates higher than Ontario
  - DAs in gray tones had incidence rates lower than Ontario
  - DAs in yellow had incidence rates similar to Ontario
- Posterior probabilities – statistical evidence that the SIRs were significantly elevated:
  - ≥95% credible limit
  - 90-94% credible limit
- SaT Scan results: Scanning statistic that identifies local clusters with elevated cancer incidence (observed vs. expected)

Cancer Incidence Maps

The cancer incidence maps are under review for publication. Once they are approved, the maps will be included in the slide presentation.

Additional Sites with Clustering

The following cancer sites also had statistically significant clustering (p-value ≤ 0.05):

- Bladder cancer in females
- Melanoma in males
- Melanoma in females
- Prostate in males
- Thyroid in females
- Testis in males

Behavioural Risk Factors

- There are many behavioural risk factors that may influence susceptibility to various cancers (e.g. smoking and lung cancer)

- Mapping DA-level estimates of risk factors has two key purposes:
  - Allows a greater understanding of neighbourhood-level differences in cancer risk
  - Can be incorporated into cancer incidence models

Behavioural Risk Factor Modeling: Smoking & Excess Body Weight Examples

- Smoking behaviour obtained from multiple cycles of the CCHS: 1.1, 2.1, 3.1, 2007/08 and 2009/10
- Ever smoking and excess body weight (aka overweight/obesity)
- Prevalence estimates account for age & CCHS cycle and DA-level income
- Maps have two components:
  - Prevalence estimates
  - posterior probabilities ($\geq$90% of Bayesian simulations exceed 80th percentile for the risk factor prevalence in Erie-St. Clair)

Behavioural Risk Factor Maps

The behavioural risk factor maps are under peer review for publication. Once they are approved, the maps will be included in the slide presentation.

Summary: Erie-St. Clair

- Small-area maps of cancer incidence and relevant behavioural risk factors feasible
- Behavioural risk factor estimates may help inform resource allocation
- Risk factor prevalence accounted for modest differences in unexplained variation

A few caveats…

- People are mobile
- Simple risk factor measures
- Other important information may be missing/unavailable
- Ecological fallacy

Shared Vision & Mission

Vision:

Establishment of a sustained, comprehensive, high quality, rapidly responsive system for spatial(-temporal) surveillance of public health problems/issues at the neighbourhood and community level.

Mission:

To ensure sufficient, skilled capacity, technical and scientific infrastructure, end-user support and ongoing methods and tools development and testing

Moving Forward: Key Issues

- Effective stakeholder engagement
- Important distinction between levels of complexity
  - Surveillance vs. research
- Significant “up front” work in data enhancement & harmonization
- Changes in policy, legislation, regulations may be required
- Communication: getting everyone on the same page

Thank you!!

Contact info: laura.seliske@cancercare.on.ca
eric.holowaty@sympatico.ca

Resources:

- **The GIS Primer**
  http://www.innovativegis.com/basis/primer/primer.html
- **Health and Environment Information Systems for Exposure and Disease Mapping and Risk Assessment**
  Jarup et al. Environmental Health Perspectives, June 2004. Vol. 112: 995-1045.
  Elliot et al. Environmental Health Perspectives, Aug 2008. Vol. 116: 1098-1130.
- **GIS and Public Health**
  Cromley EK and Lafferty SL. Guildford Press, 2002.
- **Developing the atlas of cancer in Queensland: methodological issues**
  Cramb et al. International Journal of Health Geographics, Jan 2011.
- **Feasibility and utility of mapping disease risk at the neighbourhood level within a Canadian public health unit: an ecological study**
  Holowaty et al. International Journal of Health Geographics, May 2010.
- **Estimating cancer risk in relation to tritium exposure from routine operation of a nuclear-generating station in Pickering, Ontario.**
  Wanigaratne et al. Chronic Diseases and Injuries in Canada. Sept 2013. Vol. 33: 278-289.

Resources:

- **Spatial Epidemiology: Methods and Applications**
  Elliot P, et al. Oxford University Press. 2000.
- **Applied Spatial Statistics for Public Health**
  Waller LA and Gotway CA. Wiley Interscience. 2004.
- **Geographic Information Systems and Public Health: Mapping the Future**
  Richards TB et al. Public Health Reports. July-Aug 1999. Vol. 112: 359-373.
- **Public Health and GIS**
  Rushton G et al. Annual Review of Public Health. May 2003. Vol. 24: 43-56.
- **Putting People on the Map: Protecting Confidentiality with Linked Socio-Spatial Data**
  Gutmann MP et al. National Research Council. 2007
  http://books.napedu/catalog/11865

Behavioural Risk Factor Modelling Equation (DA-level Portion)

$$logit\ (p_i) = \log\left(\frac{p_i}{1-p_i}\right) = \alpha_0 + b_1x_1 + b_2x_2 + u_i + v_i$$

Where:

p_i: percentage of smokers for the i^{th} DA

α_0: intercept

b_1: coefficient for the CCHS cycle (ref: CCHS 1.1)

b_2: coefficient for age group in 10 yr increments (ref: 50-59 yrs)

u_i: random effect for the i^{th} DA

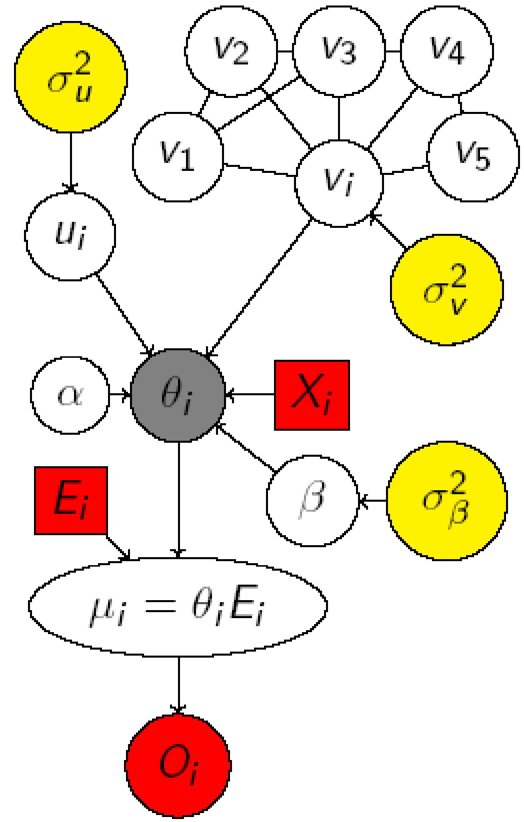
v_i: spatial effect for the i^{th} DA

The above equation was rearranged to obtain the percentage of smokers for each DA through the following equation:

$$p_i = \frac{e^{(\alpha_0 + b_1x_1 + b_2x_2 + u_i + v_i)}}{e^{(\alpha_0 + b_1x_1 + b_2x_2 + u_i + v_i)} + 1}$$

Bayesian Mapping Model

$$
\begin{aligned}
O_i &\sim Poisson(\mu_i) \\
\mu_i &= \theta_i E_i \\
\\
\log(\theta_i) &= \alpha + \beta X_i + u_i + v_i \\
u_i &\sim Normal(0, \sigma_u^2) \\
v_i | v_{-i} &\sim Normal(\textstyle\sum_{j \sim i} v_j / n_i, \sigma_v^2 / n_i) \\
\\
f(\alpha) &\propto 1 \\
f(\beta) &\propto 1 \\
\sigma_u^2 &\sim Gamma^{-1}(.001, .001) \\
\sigma_v^2 &\sim Gamma^{-1}(.001, .001)
\end{aligned}
$$



From Besag, York and Mollie, 1991.