Investigation of the Spatial and Temporal Distribution of Waterborne Disease in Southern Alberta

Stephanie Dixon & Daniel Gillis

May 12^{th}, 2006

Agenda

Introduction - Motivation

- 288 Canadian GI outbreaks (1974-2001)
  - 34% definitively waterborne
  - 21% likely waterborne
- Over 50% of outbreaks are preceded by extreme weather
  - Walkerton, E. coli, 2000
  - Milwaukee, Cryptosporidium, 1993
  - BC, ON, PQ
- High cattle density and GI illness risk are linked
  - Of 22 outbreaks, 3 linked water contaminated by cattle
  - Cranbrook cryptosporidium outbreak linked to cattle

Introduction - Motivation

- **Water safety is a major Public Health Concern**
- **Factors affecting risk need to be investigated**

In response, this study is aimed at determining the impact of agricultural practices and climate variables on the spatial-temporal distribution of waterborne disease in Alberta.

But first - we need some background

Introduction - Background

Zoonotic Background

- Campylobacter, Cryptosporidium, E. coli, Giardia & Shigella

- Over 64% of Canadian outbreaks are attributed to the above pathogens
- Infection causes diarrhea, bloody stool, vomiting, dehydration, etc.
- Affects of illness can last for a few days to years (depending on severity)

Data

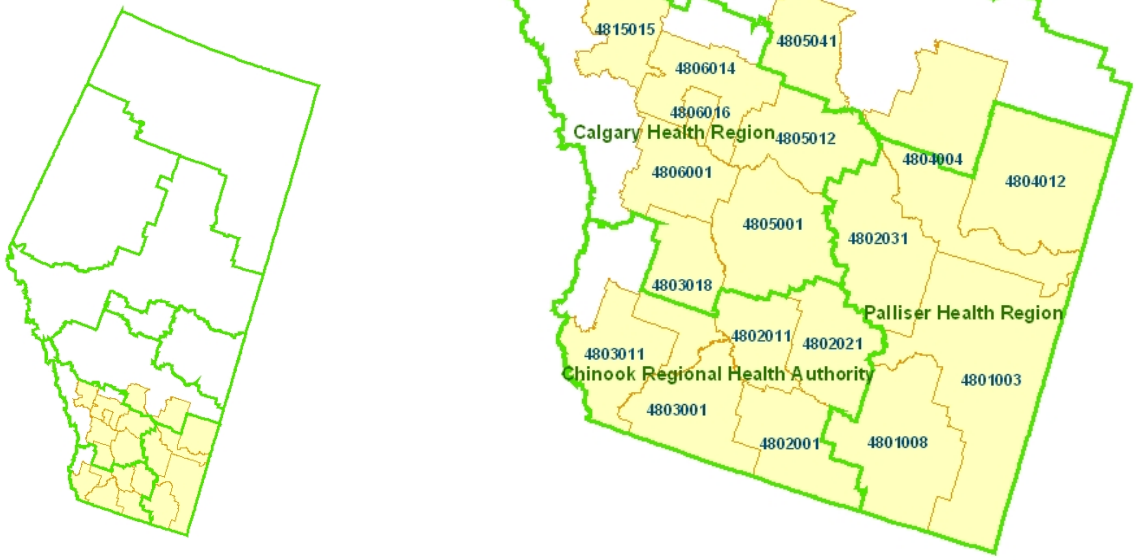
Data Sources Include

- CIHI (Canadian Institute for Health Information)
- Postal Code Reference File (Statistics Canada)
- GIS Data (University of Guelph)
- Agricultural Census Data (Government of Alberta)
- Climate Data (Environment Canada)

The data was aggregated by HU, CCS, Year, Season, Age & Sex

Data

- Consider the hierarchy



Data - Preliminaries

- Unadjusted rates suggest GI rates highest among
  - the elderly (65+)
  - females

- χ^2 significance tests indicate
  - season not significant
  - age, sex, year and CCS significant

Table 1: Contingency Table Results

variable	χ^2	**DF**	**p-value**
age	2374.85	2	0.000000
sex	14.70	1	0.000126
season	2.67	3	0.445527

Data - Preliminaries

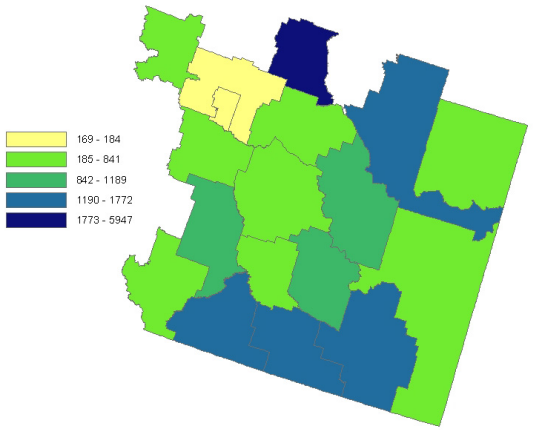
Tests of Spatial Relationship

- Used Moran's I and Geary's C with several contiguity matrices
  - nearest neighbour binary
  - row standardized
  - globally standardized
  - inverse geodesic distance weighted * binary

- Monte Carlo simulations (of 1 million samples)
- Geary's C = 0.5798 (p=0.0398)
- Moran's I = -0.0088 (p=0.2595)

Data - Preliminaries

Figure 1: Unadjusted Incidence Rates 1994



Modeling - SAS

- Proc GenMod with poisson distribution and log link
- Population at risk offset (α)
- Final model is

$$\ln\left(\frac{\mathbf{Count}_{hijklm}}{\alpha}\right) = \beta_0 + \beta_{\mathbf{HU}_h} + \beta_{\mathbf{CCS}_i} + \beta_{\mathbf{Year}_j} + \beta_{\mathbf{Season}_k} + \beta_{\mathbf{Age}_l} + \beta_{\mathbf{Sex}_m} + \beta_{(\mathbf{Age*Sex})_{lm}}$$

- Did not include climate or agricultural factors
- Did not consider spatial or temporal correlations

Modeling - WinBUGs

- Modeled with poisson distribution & log link, with offset α
- CAR hierarchical model used to account for spatial/temporal correlation
- Did not consider spatiotemporal interaction
- Final model is

$$\ln\left(\frac{\mathbf{Count}_{hijklm}}{\alpha}\right) = \beta_{\mathbf{Age}_l} + \beta_{\mathbf{nRD25}_{ijk}}$$

- Models compared using DIC
- Parameters assumed to be $\sim N(0, 0.001)$

Results - SAS

- Pearson estimate of scale parameter used to account for overdispersion

Table 2: SAS Modeling Results

Source	**Deviance**	**DF**	χ^2	$P(\chi^2 > X^2)$
Intercept	22.8814			
HU	13.9844	3	7077.18	< 0.0001
CCS	11.8767	14	1676.60	< 0.0001
Year	11.7427	6	106.54	< 0.0001
Season	11.7226	3	16.03	0.0011
Age	3.4421	2	6586.81	< 0.0001
Gender	3.3970	1	35.88	< 0.0001
Age*Gender	3.3432	2	42.79	< 0.0001

Results - WinBUGs

- Temporal considerations non-significant (Year, Season)
- Age & Sex models did not converge
- Only climate variable $nRD25$ was significant

Results - WinBUGs

- Estimate of climate effect -0.022
- Odds 95% CI (0.9611,0.9956)

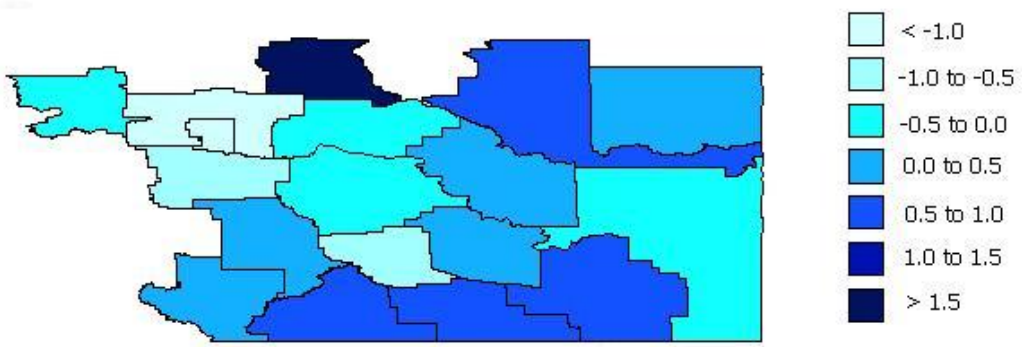
Table 3: CAR model parameter estimates

parameter	**mean**	**sd**	2.5%	**median**	97.5%
β_{AGE_1}	-7.548	0.016	-7.58	-7.55	-7.518
β_{AGE_2}	-8.216	0.012	-8.24	-8.22	-8.192
β_{AGE_3}	-6.475	0.013	-6.50	-6.48	-6.449
β_{nRD25}	-0.022	0.009	-0.04	-0.022	-0.002
σ	1.396	0.250	1.01	1.360	1.990

- The following plot illustrates the parameter estimates for the spatial components in each CCS

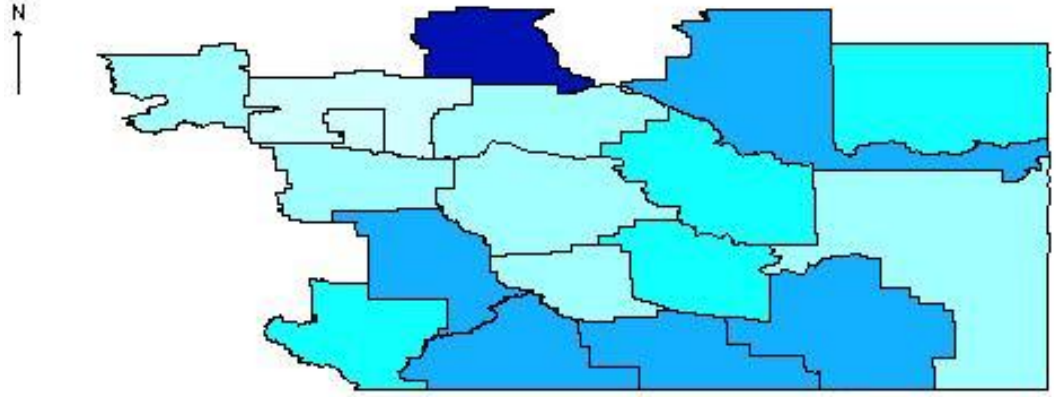
Results - WinBUGs

Figure 2: Spatial Parameter Estimates



Results - WinBUGs

Figure 3: WinBUGs IR Rates Spring 94



Conclusions

- Two methods explored
- Methods suggested different significant factors
  - Non spatial (GLM) excluded climate
  - Spatial (CAR) included climate
- Different results - Prefer CAR

Future Work

- Temporal considerations - different scale?
- Spatial Temporal interactions to be explored
- Other weight matrices
- Other variables - agriculture, watershed, etc.
- Convergence issues - reparameterize?
- Other spatial scales
- Non-nested modeling (i.e. non-hierarchical)
- Extend model to other provinces

Acknowledgements & Thanks

- Math & Stats Department, University of Guelph
- Public Health Agency of Canada
- Environment Canada
- ECCHO
  - Ecosystems Climate Change and Health Omnibus
  - further info can be found at www.eccho.ca
- NSERC & OGS