# Exploring the Spatial Relationship between Fetal Death Distribution and Toxics Release Inventory Sites in Georgia

Imam M. Xierali, Ph.D. and Gordon R. Freymann, M.P.H.
Georgia Department of Human Resources, Division of Public Health, Office of Health Information and Policy
2 Peachtree Street, Atlanta, GA 30303; ixierali@dhr.state.ga.us

## ABSTRACT

Environmental factors affect fetal development. This research uses exploratory data analysis tools in a Geographic Information System (GIS) as well as regression analysis to examine the spatial distribution of fetal death and live birth occurrence in the state of Georgia from 1996 to 2004 and the effect of potential exposures to toxic releases from Toxics Release Inventory (TRI) sites on birth outcomes in Georgia. Proximity to TRI sites is used as the proxy for exposure. Three conventional methods are employed. First, a traditional buffer analysis on statewide multi-year data is conducted and fetal death rates and proximity to suspect TRI sites correlated. Second, fetal death rates at census tract level are regressed on proximity to suspect TRI sites. Third, birth outcome (fetal death vs. live birth) is regressed on sex, mother's age, demographic cluster, and proximity to TRI sites. The result of this preliminary study reveals an inverse association between fetal death and proximity to suspect TRI sites in Georgia. The analysis may be helpful in understanding factors affecting birth outcomes in Georgia.

## INTRODUCTION

There is significant concern about exposure of the fetus to environmental pollutants during pregnancy. For instance, many studies have shown excess risk of birth anomalies in populations living near landfill sites (Dolk et al, 1998; Fielder et al, 2000; Vrijheid 2000). Choi et al (2006) observed increased risk for mothers living within one mile of a TRI site for having children diagnosed with brain cancer before five years of age, compared to living more than one mile away from a TRI facility.

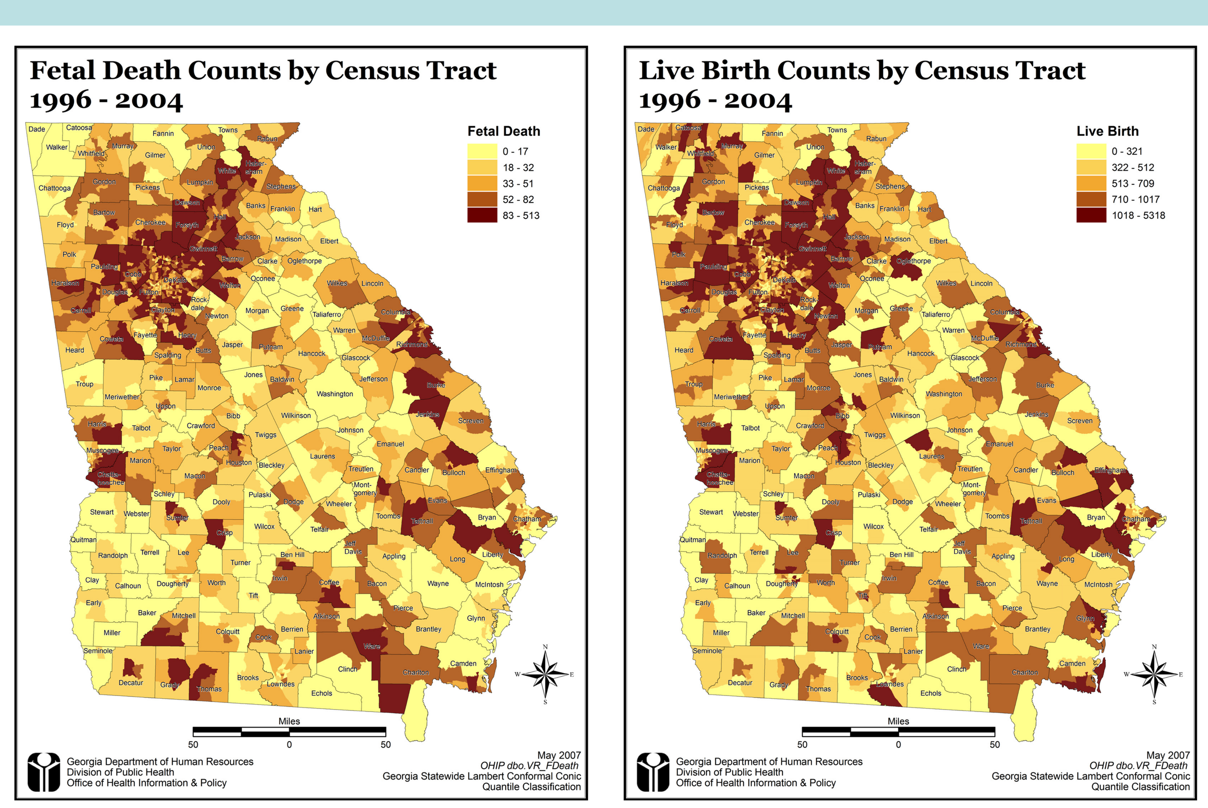
The health field concept proposes that "health status is influenced by human biology, the environment, the life styles of citizens, as well as the healthcare organization" (Georgia's New Health Outlook, 1976, p.8). Looking from an environmental perspective, this study examines whether living near a TRI site will increase the likelihood of fetal death occurrence in Georgia. Fetal death is defined as "Death prior to the complete expulsion or extraction from its mother of a product of human conception, irrespective of the duration of pregnancy; the death is indicated by the fact that after such expulsion or extraction the fetus does not breathe or show any other evidence of life such as beating of the heart, pulsation of the umbilical cord, or definite movement of voluntary muscles" (Georgia Division of Public Health, OASIS Web Query, http://oasis.state.ga.us/oasis/help/mch.html). In this study, fetal death rate is defined quantitatively as the ratio of fetal deaths over the sum of fetal deaths and live births multiplied by 1,000.

## DATA

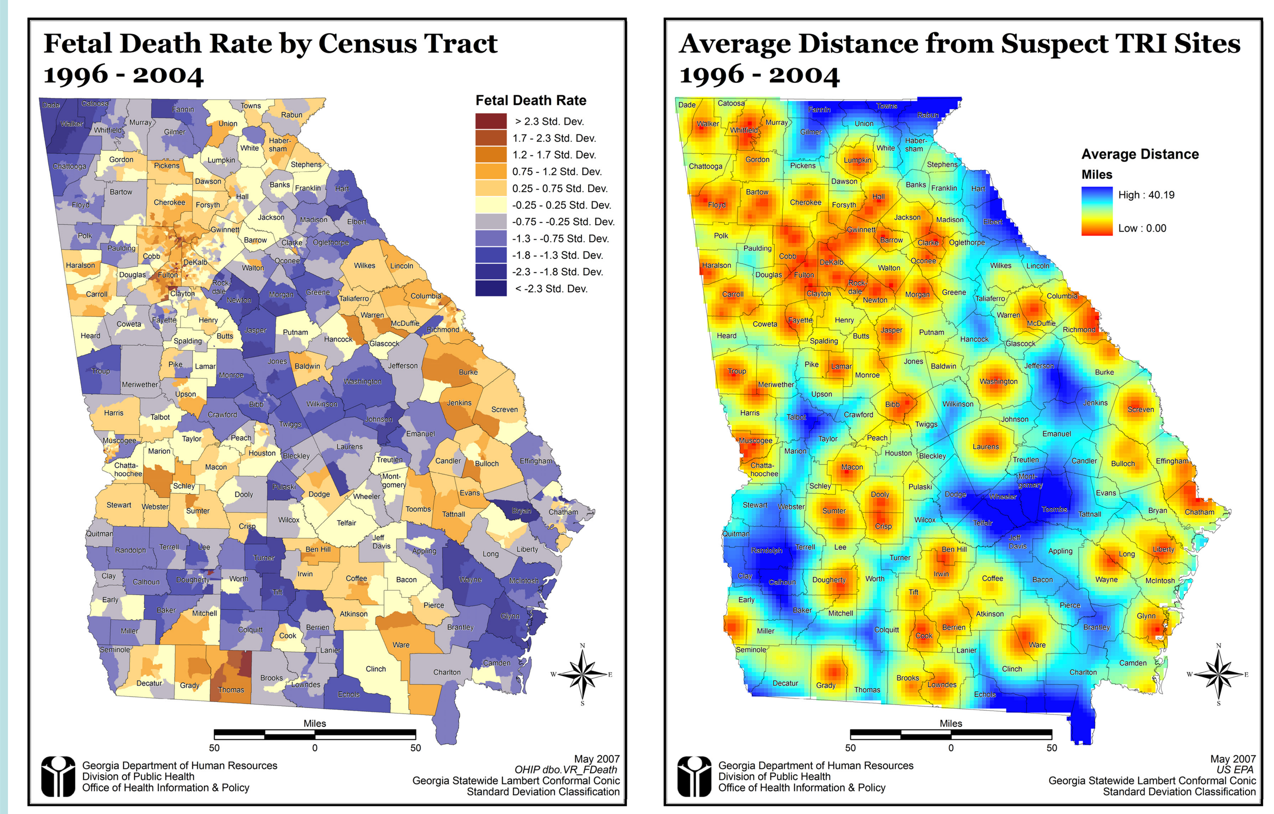
Georgia Division of Public Health Office of Health Information and Policy maintains large database of fetal death and birth records at individual level for Georgia. The data include not only demographic characteristics of the mother and father, but also the residential addresses of the mothers. These address data have been geo-referenced and linked to social economic data to aid geographic and statistical analysis. From 1996 to 2004 there have been 84,030 fetal deaths and 1,154,444 live births in Georgia. Only 16 cases or 0.02 percent out of these fetal deaths resulted from external causes such as accidents. Exposure to TRI pollutants may be a factor in many of the fetal deaths.

TRI data are from US Environmental Protection Agency (EPA). The data include pollutant chemical name, release amounts (estimated) to air, ground, and underground, as well as address information for each facility. The data are also geo-referenced by the facility address for the analysis. Distance from residential address to the nearest TRI site is used as the proxy for environmental risk exposure. TRI sites are selected so that only those sites that release chemicals harmful to fetal health are used in this round of analysis.

The harmful chemical lists are obtained from National Library of Medicine database. There are 161 chemicals linked to the fetal death not only in humans but also in animals. TRI sites that release any of these chemicals are included in the analysis; hereafter called the suspect TRI sites.

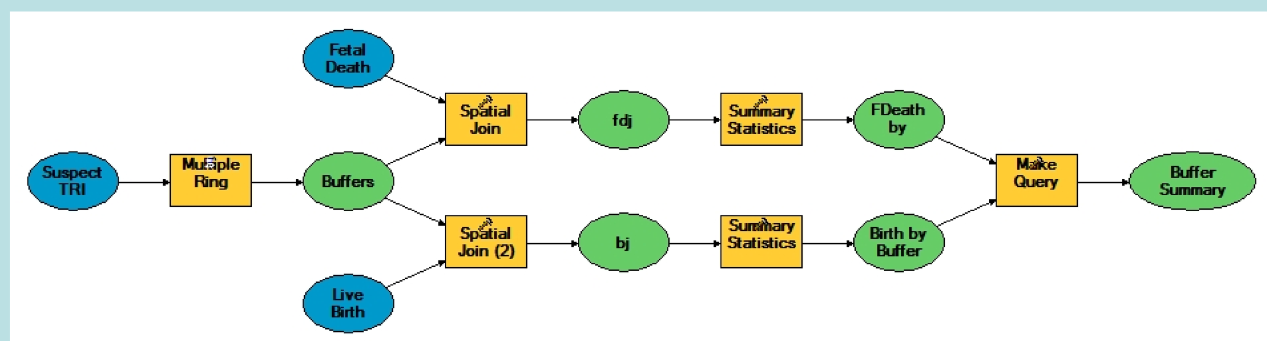


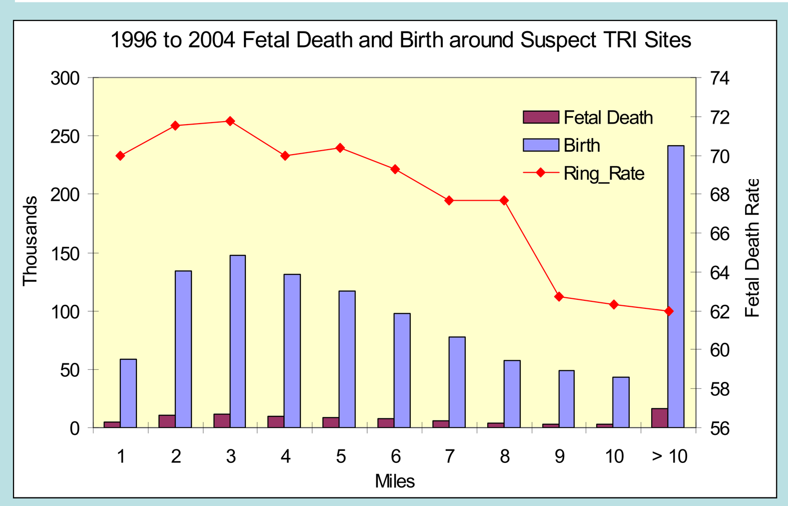

## KEY VARIABLES

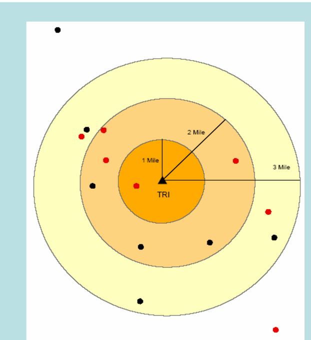



## METHOD 1: BUFFER ANALYSIS

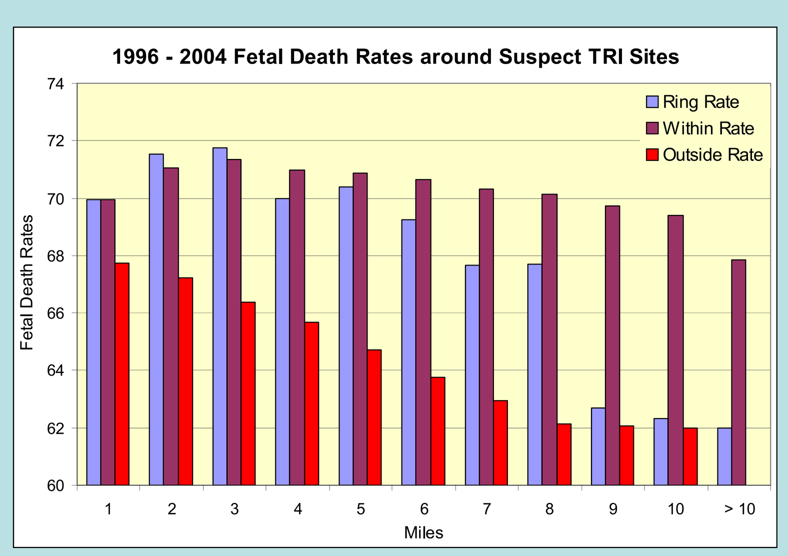
A set of buffer rings are placed around each of the suspect TRI sites. The resulting buffer rings are overlayed on the fetal death and live birth maps to count the number of deaths and births that fall within each ring. This is done for each year and the result tables are joined together for statistical analysis.



A three-mile 'ring rate', for example, is calculated from fetal deaths and live births that occurred between two and three miles away from a TRI site; a three-mile 'within rate' is calculated from all events that occurred within three miles from the site; a three-mile 'outside rate' is calculated from all events that occurred more than three miles away from the site.




The distance and fetal death rate for each year are stacked and correlated. Lower fetal death rates are observed as distance increases away from the suspect TRI sites.

**Correlations**

| | | | Distance | Ring Rate | Within Rate | Outside Rate |
|---|---|---|---|---|---|---|
| Spearman's rho | Distance | Correlation Coefficient | 1.000 | -.536** | -.114 | -.539** |
| | | Sig. (2-tailed) | . | .000 | .283 | .000 |
| | | N | 90 | 90 | 90 | 90 |
| | Ring Rate | Correlation Coefficient | -.536** | 1.000 | .761** | .728** |
| | | Sig. (2-tailed) | .000 | . | .000 | .000 |
| | | N | 90 | 90 | 90 | 90 |
| | Within Rate | Correlation Coefficient | -.114 | .761** | 1.000 | .655** |
| | | Sig. (2-tailed) | .283 | .000 | . | .000 |
| | | N | 90 | 90 | 90 | 90 |
| | Outside Rate | Correlation Coefficient | -.539** | .728** | .655** | 1.000 |
| | | Sig. (2-tailed) | .000 | .000 | .000 | . |
| | | N | 90 | 90 | 90 | 90 |

**. Correlation is significant at the 0.01 level (2-tailed).

## METHOD 2: SPATIAL REGRESSION

Fetal death rates at census tract level are regressed on proximity to TRI sites in a traditional, aspatial ordinary least square regression model. The residuals are checked for spatial autocorrelation using a first order spatial weight matrix (neighborhood table). Significant spatial autocorrelation is detected and a lag model is used instead. In the lag model, the average fetal death rates of neighboring tracts are also used to explain the fetal death rate of the tract under consideration.

## MODEL

$Y = X\beta + \varepsilon$ (OLS model)

$Y = \rho WY + X\beta + \varepsilon$ (Spatial Lag model)

$Y = \lambda WY + X\beta - \lambda W X\beta + \mu$ (Spatial Error Model)

Y is a n by one matrix of the dependent variable; X is a n by k matrix of the independent variables; *W* is a n by n spatial neighborhood matrix.

## ANALYSIS

Linear regression results show that the residuals present significant autocorrelation. The spatial lag model has the highest log-likelihood and the smallest AIC, indicating the best fit among the three models. The spatial lag model suggests that beyond distance, spatial dependence among fetal death rates of neighboring tracts is a significant factor for explaining the spatial distribution of fetal death rates in Georgia.

| *Model* | *R-Square* | *Log-likelihood* | *AIC* | *DF* | *N* |
|---|---|---|---|---|---|
| OLS | 0.0367 | -7683.81 | 15371.60 | 1616 | 1618 |
| Lag | 0.58065 * | -7128.36 | 14262.70 | 1615 | 1618 |
| Error | 0.58064 * | -7129.51 | 14263.00 | 1616 | 1618 |

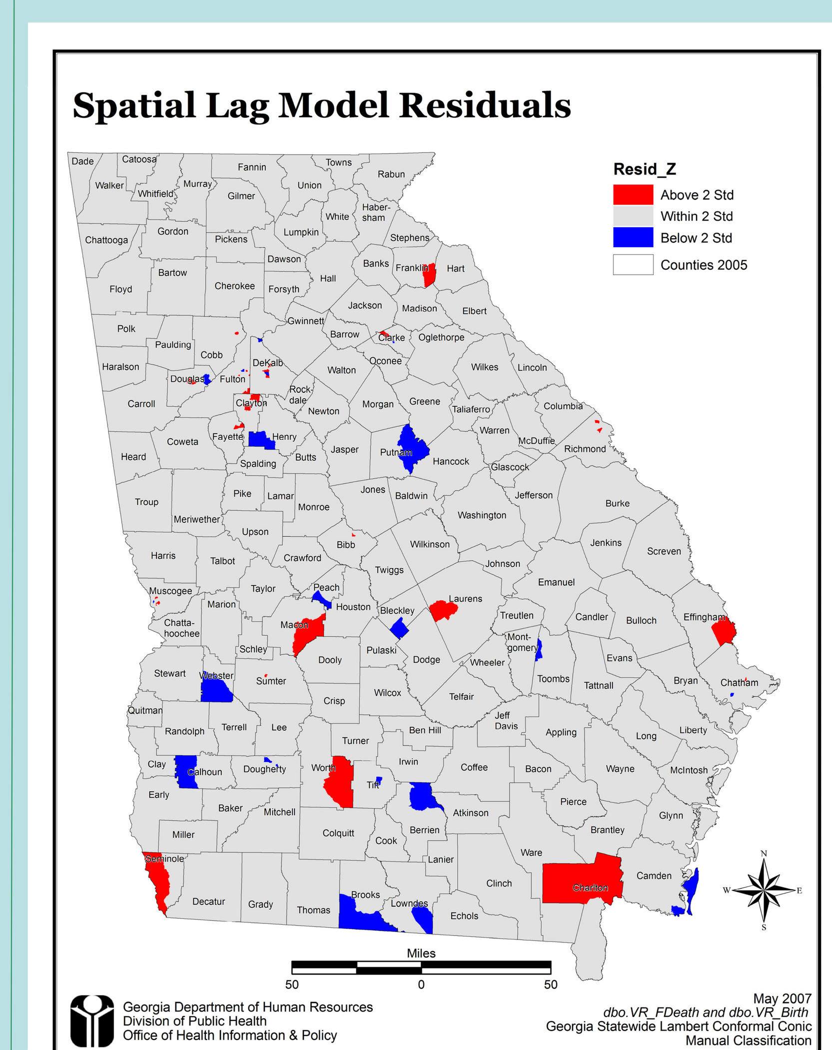
* Pseudo

| Model | Constant | Distance | Rho | Lambda |
|---|---|---|---|---|
| OLS | 72.692 ** | - 0.879 ** | ~ | ~ |
| Lag | 16.749 ** | - 0.221 ** | 0.773 ** | ~ |
| Error | 69.979 ** | - 0.468 ** | ~ | 0.776 ** |

** p < 0.01

| Spatial Dependence | Fetal Death Rate | Distance | Residual (OLS) | Residual (Lag) | Residual (Error) |
|---|---|---|---|---|---|
| Moran's *I* | 0.233 *** | 0.304 *** | 0.191 *** | -0.081 | -0.082 |
| Geary's *C* | 0.779 *** | 0.700 *** | 0.372 *** | 1.044 | 1.046 |

*** p < 0.001; ** p < 0.01




The spatial lag model has the best model fit. The lag model suggests significant neighborhood effects in the dependent variable. The residual map suggests 51 tracts have residuals beyond two standard deviations from the mean. Further analysis of all fetal deaths, live births, and suspect TRI sites in these tracts suggests that the accuracy of geocoding, especially of the suspect TRI sites in these tracts, is too coarse: only 35% of fetal deaths and 60% of TRI sites in these tracts had zip code level or worse accuracy.

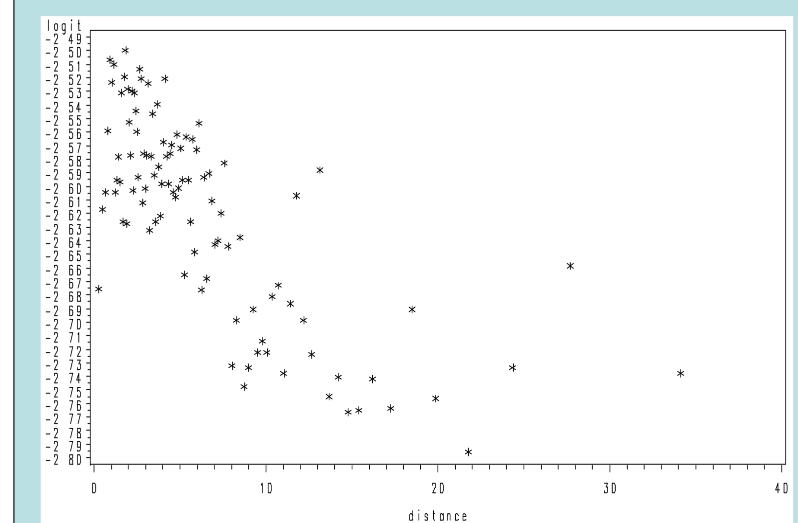
| Geocoding Accuracy (1996 - 2004) | | | | |
|---|---|---|---|---|
| Confidence | Overall | Fetal Death | Live Birth | TRI Sites |
| Street Level | 79.12% | 67.83% | 79.94% | 61.69% |
| Zipcode Level | 11.08% | 8.86% | 11.24% | 38.31% |
| Spatial Imputation | 9.80% | 23.31% | 8.82% | 0.00% |

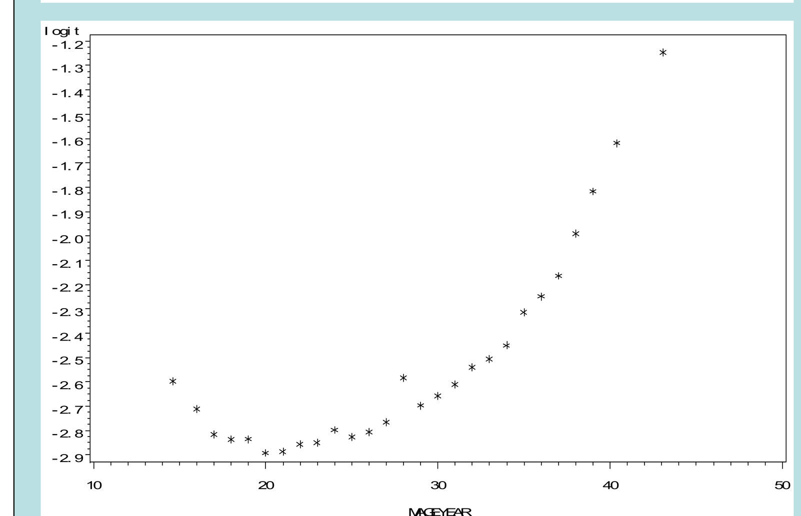
| Geocoding Accuracy at Outlier Tracts | | | | |
|---|---|---|---|---|
| Confidence | Overall | Fetal Death | Live Birth | TRI Sites |
| Street Level | 76.69% | 64.98% | 77.47% | 39.62% |
| Zipcode Level | 10.50% | 6.19% | 10.79% | 60.38% |
| Spatial Imputation | 12.80% | 28.84% | 11.75% | 0.00% |

## METHOD 3: LOGISTIC ANALYSIS

Fetal death and live birth data are combined and a dependent variable is defined as one for a fetal death incident and zero for a live birth. Distance from residential address of mothers to the nearest suspect TRI site is used as the proxy for environmental risk exposure. This distance and mother's age in years (MAGEYEAR) are used as explanatory variables.

Other variables explored include sex and demographic cluster major categories (Millard, unpublished work). The sex of the majority of fetal deaths is unknown, therefore, this variable was excluded from the model. Demographic profile cluster major categories could not predict the probability of a fetal death incident.




| Odds Ratio Estimate | | | |
|---|---|---|---|
| Effect | Point Estimate | 95% Wald Confi. Limit | |
| Distance | 0.995 | 0.993 | 0.996 |
| Mageyear | 1.045 | 1.044 | 1.046 |

| Analysis of Maximum Likelihood Estimate | | | | |
|---|---|---|---|---|
| Parameter | DF | Estimate | Wald Chi-Square | Pr > Chisq |
| Intercept | 1 | -3.7930 | 48997.2500 | <0.0001 |
| Distance | 1 | -0.0055 | 77.4054 | <0.0001 |
| Mageyear | 1 | 0.0442 | 6053.7700 | <0.0001 |

The outcome of the logistic regression suggests an inverse association between residential proximity to suspect TRI sites (in miles) and probability of fetal death. However, the overall impact of distance is quite small. Mother's age is a better predictor than the distance to suspect TRI sites, albeit with small impact as well.

## CONCLUSION

In this preliminary study, we examined the relationship between fetal deaths and TRI sites in Georgia. We found that using aggregate level analysis (Methods 1 and 2), fetal death rate is associated to the residential proximity to TRI sites. Buffer analysis suggests that fetal deaths and live births peak around areas three miles from the nearest suspect TRI sites. This finding questions the conventional approach of selecting a predefined buffer distance in similar environmental health studies. We also found significant neighborhood effects among fetal death rates at the census tract level. This means that a spatial regression approach is necessary for explaining the variations in fetal death rates at that level. Using the individual level data to predict the probability of a pregnancy ending with a fetal death, we found an inverse albeit weak association between the fetal death and the residential proximity to the nearest suspect TRI sites. At the individual level, the requirement for a high quality geocoding process is a necessity for an accurate prediction. When 38% of suspect TRI sites and 20% of birth outcomes are only accurate at the zip code level, the residential proximity measurement could pose serious hurdles on analysis implications. At the aggregate level, however, proximity to TRI sites is more stable. Overall, the three analyses suggest that fetal death events are associated with the residential proximity to the TRI sites under study.

More research is necessary to further understand the relationship between fetal deaths and environmental risks from hazardous sites. The future study may incorporate TRI data from neighboring states, and include other factors such as traffic emissions data and landfill sites.

***Acknowledgement:*** *Appreciation goes to Elaine J. Hallisey, Frank H. Millard, Colin K. Smith, David P. Austin, and Jeffery N. McMichael. This project would be impossible without their continuing encouragement and support.*

***Selected Reference:***

**Anselin, L. (1988)** Spatial Econometrics: Methods and Models. Dordrecht, Holland: Kluwer Academic Publishers.
**Choi, H.S., Y.K. Shim, W.E. Kaye, & P.B. Ryan (2006)** "Potential Residential Exposure to Toxics Release Inventory Chemicals during Pregnancy and Childhood Brain Cancer," *Environmental Health Perspectives,* 114(7): 1113-18.
**Cliff, A.D. & J.K. Ord (1972)** "Testing for Spatial Autocorrelation among Regression Residuals," *Geographical Analysis*, 4: 267-84.
**Dolk, H., M. Vrijheid, B. Armstrong, L. Abramsky et al (1998)** "Risk of Congenital Anomalies Near Hazardous-Waste Landfill Sites in Europe: the EUROHAZCON Study," *Lancet,* 352: 423-27.
**Fielder, H.M., C.M. Poon-King, S.R. Palmer, N. Moss, G. Coleman (2000)** "Assessment of Impact on Health of Residents Living Near the Nant-y-Gwyddon Landfill Site: Retrospective Analysis," *British Medical Journal*, 320: 19-22.
**Nuckols, J.A., M.H. Ward, & L. Jarup (2004)** "Using Geographic Information Systems for Exposure Assessment in Environmental Epidemiology Studies," *Environmental Health Perspectives,* 112(9): 1007-15.
**Vjirheid, M. (2000)** "Health Effects of Residence Near Hazardous Waste Landfill Sites: a Review of Epidemiologic Literature," *Environmental Health Perspective*, 108: 101-12.

Georgia DHR Publication Tracking Number: DPH087HW