Evidence of localized clustering of gastroschisis births in North Carolina, 1999–2004

Elisabeth Dowling Root a,b,*, Robert E. Meyer c,d, Michael E. Emch a,b

a Department of Geography, University of North Carolina at Chapel Hill, Saunders Hall, Campus Box 3220, Chapel Hill, NC 27599, United States
b Carolina Population Center, University of North Carolina at Chapel Hill, Campus Box 8120, University Square, 123 W. Franklin Street, Chapel Hill, NC 27516, United States
c North Carolina Birth Defects Monitoring Program, State Center for Health Statistics, NC, United States
d Department of Maternal and Child Health, University of North Carolina at Chapel Hill, NC, United States

A R T I C L E   I N F O

Article history:
Available online 21 February 2009

Keywords:
Birth defects
Gastroschisis
Spatial scan statistic
Spatial cluster
Covariates
USA
SaTScan

A B S T R A C T

Examining the geographic distribution of birth defects can be useful in exploratory etiologic research. Identification of clusters of certain defects may uncover possible environmental or socio-economic risk factors and assist with the generation of hypotheses about underlying causes of these conditions. In North Carolina, the prevalence of gastroschisis, a serious abdominal wall defect, has increased over the past decade and anecdotal evidence from clinicians suggests the possibility of clustering of this condition. This study uses a spatial scan statistic to identify the location and extent of clusters of gastroschisis births in North Carolina between 1999 and 2004. Data on cases of gastroschisis were obtained from the North Carolina Birth Defect Monitoring Program (NCBDMP) and control births were chosen from all resident live births without birth defects contained in the North Carolina composite linked birth files. The clusters were controlled for five major risk factors (maternal age, race, parity, Medicaid status, maternal smoking) to ensure that the clusters were not artifacts of unequal population distribution. Results indicate a localized cluster of gastroschisis in the rural southern Piedmont of North Carolina which persists even after controlling for all major risk factors. Adjusting for these risk factors shifted the location of the cluster substantially, demonstrating the importance of adjusting for underlying population distribution. Results indicate a localized cluster of gastroschisis in the rural southern Piedmont of North Carolina which persists even after controlling for all major risk factors. Adjusting for these risk factors shifted the location of the cluster substantially, demonstrating the importance of adjusting for underlying population distribution. Since clusters persisted after adjusting for individual-level risk factors, environmental contaminants may explain the excess of gastroschisis cases. This study is among the first to assess spatial clustering of gastroschisis using GIS methods. This study also demonstrates the importance of controlling for covariates in spatial analysis and illustrates the usefulness of the spatial scan statistic in exploratory etiologic research.

© 2009 Elsevier Ltd. All rights reserved.

Introduction

Gastroschisis is a serious congenital abdominal wall defect in which the intestines and sometimes other internal organs develop outside the abdomen through an opening in the abdominal wall, resulting in the suspension of these structures in the amniotic sac. Gastroschisis is not usually associated with other birth defects and early diagnosis through ultrasound and modern surgical techniques have increased the survival rate to over 90% (Brantberg, Blaas, Salvesen, Haugen, & Eik-nes, 2004; Wilson & Johnson, 2004).

Estimates from state-based, active surveillance data compiled by the National Birth Defects Prevention Network suggest that the prevalence of gastroschisis in the United States is approximately 3.82 per 10,000 live births, although there is considerable geographic variation (Canfield, Honein, et al., 2006; NBDPN, 2007). In North Carolina, the birth prevalence of gastroschisis increased from 1.96 per 10,000 births in 1997 to 4.49 per 10,000 births in 2000 (Laughon et al., 2003) and has remained high with a rate of 4.26 per 10,000 live births in 2004. This apparent increase in gastroschisis is not unique to North Carolina; studies from around the world have reported an increase in prevalence over the past several decades (Calzolari et al., 1995; Forrester & Merz, 1999; Hougland, Hanna, Meyers, & Null, 2005; Martinez-Frias, Salvador, Prieto, & Zaplana, 1984; Penman, Fisher, Noblett, & Soothill, 1998; Penz, Menardi, & Brezinka, 1998; Rankin, Dillon, & Wright, 1999; Roeper, Harris, Lee, & Neutra, 1987; Williams, Kucik, Alverson, Olney, & Correa, 2005; Wilson & Johnson, 2004). These studies suggest a great deal of large-scale geographic variation in gastroschisis but very few have looked at small area variation of the defect. This is...
due, in part, to the relative rarity of gastroschisis births. The small number of cases born each year has traditionally made small-scale geographic analysis difficult. While the prevalence of gastroschisis appears to be increasing, the etiology remains uncertain. Young maternal age is one of the few risk factors consistently associated with gastroschisis. In general, studies have shown that women younger than 20 years of age have significantly higher risk of a gastroschisis birth compared to older women. Studies investigating the relationship between race/ethnicity and risk for gastroschisis show higher prevalence among Hispanic and white infants and lower prevalence among black infants (Canfield, Honein, et al., 2006; Lam & Torfs, 2006; Salihu et al., 2004). Since gastroschisis risk is significantly increased with young maternal age, a number of studies have investigated lifestyle and behavioral factors associated with younger women. Women who smoke or use alcohol during pregnancy may be more likely to have an infant with gastroschisis (Feldkamp, Alder, & Carey, 2008; Goldbaum, Daling, & Milham, 1990; Haddow, Palomaki, & Holman, 1993; Hougland et al., 2005; Torfs, Katz, Bateson, Lam, & Curry, 1996; Torfs, Velie, Oechslin, Bateson, & Curry, 1994; Werler, Mitchell, & Shapiro, 1992). Recently, Feldkamp, Reefhuis, Oechslin, & Holman, 1993; Hougland et al., 2005; Torfs, Katz, Bateson, Lam, & Curry, 1996; Torfs, Velie, Oechslin, Bateson, & Curry, 1994; Werler, Mitchell, & Shapiro, 1992) have reported an association between gastroschisis risk and genital urinary infections. Thus, it is possible that the geographic variation in gastroschisis rates may be due to unequal distribution of mothers with certain risk factors.

Prior epidemiological research indicates that the causes of gastroschisis, like many birth defects, are most likely complex and multifactorial and include not only maternal characteristics and behaviors but also environmental teratogens and genetic factors (Curry, McKinney, Thornton, & Stringer, 2000). With so much uncertainty周围_the etiology of this condition, most epidemiological studies are exploratory in nature, testing possible associations between these birth defects and socio-economic inequalities that are correlated with poor health outcomes or environmental contaminants that influence the development of other chronic conditions (e.g. cancer) or produce malformations in animal models (Brown, 1997). Understanding the geographic distribution of gastroschisis can be useful in exploratory etiologic research. Identification of disease clusters may uncover possible environmental or socio-economic risk factors and assist with the generation of hypotheses about the underlying socio-environmental causes of those clusters. Birth defects are particularly well suited to this type of geographic analyses because the lag time between exposure to environmental and socio-economic conditions and the development of the birth defect outcome is relatively short, at least for conditions that lack a substantial genetic component. This minimizes the potential bias introduced when study subjects move during the exposure period and allows for stronger hypotheses about the area-level factors that may cause the disease. Understanding the geographic distribution of diseases with a long latency period (e.g. cancer) may be less informative because study subjects are much more likely to move several times between exposure and diagnosis of the disease.

Researchers must also be careful in applying disease clustering techniques to ensure that identified clusters are not simply due to spatial variations in the density of the population being studied. Furthermore, if known covariates are not adjusted for, observed spatial patterns of birth defects may be due to the fact that individuals with similar risk factors live in the same geographic area, producing larger area-level patterns of disease. If, however, known individual-level risk factors are adjusted for and the cluster persists, environmental contaminants may be suspected as a possible cause of the birth defect. Given that such a wide variety of environmental, social and economic factors may influence the development of gastroschisis, it is important to understand how these factors interact and overlap in certain places to produce spatial patterns of disease.
high-quality proprietary street files (e.g. Dynamap, NAVERQ and TeleAtlas) and satellite imagery to locate addresses and are available at no cost. After each stage, the remaining unmatched addresses were geocoded using a different geocoding service. Since geocoding services use slightly different street files, this iterative process ensures that all possible addresses are geocoded.

Using this process, we matched 242 of the 264 cases (91.7%) and 11,651 of the 12,488 controls (93.3%). The 22 cases and 837 controls that could not be geocoded were excluded from the cluster analysis. Descriptive statistics were run on the unmatched versus matched records to see what, if any, differences existed between the two groups and we found some minor differences in race, parity and Medicaid status. We excluded multiples in this analysis since multiple births are not independent events (e.g. they share the same fetal environment) and we did not want to count these locations twice in the spatial cluster analysis.

Once data were geocoded, we assigned individual-level records to several census areas: blocks, block groups and tracts. The U.S. Census Bureau has developed a number of geopolitical areas to assist in the collection and reporting of census data. Census areas have a hierarchical structure. A census block is the smallest unit of geography bounded on all sides by visible features, such as streets, rivers or railroad tracks. Census block groups are clusters of contiguous blocks, typically containing from 600 to 3000 people and census tracts comprise groups of contiguous block groups and have a population ranging from 2500 to 8000. Initially, we used the individual-level point locations to scan for clusters but performed additional analyses with data aggregated to block, block group and tract in order to examine possible effects of geocoding errors. The results for all four analyses were nearly identical suggesting that any geocoding errors have a negligible effect on the size and location of identified clusters. We chose to use the census block groups for the remainder of the study because this geography yielded the highest p-values and is computationally less intensive and therefore faster to run on a desktop computer.

The data also contained potential covariates from the linked birth files including: mother’s and father’s age, race and ethnicity, marital status, number of prior births, month prenatal care began, mother’s smoking status, and whether or not Medicaid paid for the delivery.

**Statistical methodology**

We used the spatial scan statistic available in the SaTScan computer software package (Kulldorff, 1997; Kulldorff & Information Management Services, 2006; Kulldorff & Nagarwalla, 1995) to test for the presence of purely spatial clusters of gastrochisis and to identify their approximate location. We assumed the number of births in each census block group to be Poisson distributed. The method tests the null hypothesis that within any covariate group (age, race, parity, etc.) the risk of a gastrochisis birth is the same in all census block groups. This means that the expected covariate-adjusted rate of gastrochisis is constant throughout North Carolina. Despite the case-control design of our study, which would usually merit the use of a Bernoulli model, we chose to use the Poisson model for several reasons. First, prior research indicates several covariates that are related to gastrochisis including maternal age, race and parity. The Poisson model allows us to easily adjust for a large number of covariates, while the Bernoulli model does not. Second, in instances where there are few cases compared to controls (<10%) the Poisson model is a very good approximation to the Bernoulli model and produces slightly conservative p-values (Kulldorff, 1997, 2006). We examined study data using both the Bernoulli and Poisson models and results were identical though, as predicted, p-values for the Poisson model were slightly higher. Finally, the computing power necessary to run the Poisson model is significantly less than for the Bernoulli model, allowing the analysis to be run on a desktop computer.

The scan statistic detects clusters by gradually scanning an elliptical window across the entire study area, noting the number of observed and expected cases of gastrochisis inside the ellipse at each location (Kulldorff, 1997; Kulldorff, Athas, Feuer, Miller, & Key, 1998; Kulldorff & Nagarwalla, 1995). In this study, the center point of all census block groups in North Carolina served as the center for the ellipses. The radii of the ellipses vary continuously in size from zero to a user-defined maximum, which is a percentage of the total North Carolina population. The ability to vary the size of the ellipse is important because we usually do not know the size of the area covered by a cluster (Emch & Ali, 2002; Kulldorff, 1997). Thus, the location and size of the ellipse change creating an infinite number of distinct geographic areas. Each of these areas reflects a possible cluster. This method looks at varying spatial scales which is particularly appropriate for a birth defect with an unknown etiology because we do not know the scale at which the defect may exhibit spatial clustering.

The scan statistic uses a likelihood ratio test statistic, the methodology for which is described in detail elsewhere (Emch & Ali, 2002; Kulldorff, 1997; Kulldorff & Nagarwalla, 1995). For each ellipse, the likelihood of finding the observed number of gastrochisis births within the ellipse and outside the ellipse is calculated. The ellipse with the maximum likelihood is the most likely cluster, that is, the cluster least likely to be due to chance. In order to find the value of the test statistic, SaTScan uses a Monte Carlo simulation approach to find the maximum likelihood ratio over the entire range of ellipses. The same procedure (e.g. scanning the elliptical window of varying size across the study area) is repeated on a large number of random replications (we chose 9999). The maximum likelihoods of the study data and the Monte Carlo simulations are ranked in order to determine the distribution of the likelihood ratio and the corresponding p-value of the study data. SaTScan detects both primary and secondary clusters. The primary cluster is the window with the maximum likelihood ratio while secondary clusters are additional clusters that have high likelihood ratios but that do not overlap the primary cluster.

The maximum cluster size was initially set to include up to 50% of the population. However, repeated analyses showed that significant clusters included not more than 10% of the population, so we restricted the maximum cluster size to 25% of the population to minimize computing time. Both circular and elliptical windows of different shapes and angles were used to scan for clusters. We chose to use elliptical windows because prior research supports the possibility that an environmental contaminant induces gastrochisis (Dolk et al., 1998; Drongowski, Smith, Coran, & Klein, 1991; Fielder, Poon-King, Palmer, Moss, & Coleman, 2000; Torfs et al., 1996). The elliptical shape more accurately follows certain geographic features, such as watersheds and rivers, which we hypothesize could transport contaminants from their source. SaTScan does impose a penalty for using less compact shapes so that the cluster is not unnecessarily elongated in order to “cherry pick” cases over a larger area. We chose to focus on clusters with statistically significant p-values (<0.05), though we report one primary and one secondary cluster for each analysis.

After identifying statistically significant spatial clusters, the next step was to determine if these areas would change when the model was adjusted for known risk factors for gastrochisis. Since maternal age is the main risk factor consistently associated with gastrochisis, all analyses were age-adjusted using 3 categories: <20 years of age, 20–24 years, and 25 years or more. We also classified births by race (white, black and other), parity (no prior births vs. one or more prior births), Medicaid status (defined as the delivery paid for by Medicaid vs. other payer source) and maternal...
smoking (mother reported smoking during pregnancy vs. mother did not report smoking) and conducted separate analyses for each covariate. All covariates and their classifications were determined using univariate and logistic regression analysis in SAS Version 9.1 (data not shown). We chose to include covariates in the cluster analysis that had significant odds ratios in the regression analysis. Covariates were introduced into the spatial scan in an iterative manner and we controlled for not more than two covariates at a time. From a computational standpoint, we did not have a large enough sample of cases to partition the data into more than 2 or 3 covariate categories because p-values generated by the scan statistic become less reliable when locations have categories with no data (Kulldorff, 1997, 2006). In addition, we were not only interested in the geographic location of the cluster, but also how specific covariates would change that location. By adding in one covariate at a time and observing the change, we can see how the underlying geographic distribution of that covariate affects the distribution of cases in the State.

Results

Fig. 1 and Table 1 show the results of the unadjusted scan statistic. Two statistically significant clusters were identified, both located in the southern Piedmont region of North Carolina. The primary cluster \((p = 0.016)\) encompassed a larger area and included 50 cases of gastroschisis, approximately 2.42 times more cases than expected. The secondary cluster \((p = 0.046)\) was geographically smaller in size and included 12 cases, approximately 6 times more cases than expected.

Table 1 and Fig. 2 indicate how the results of the spatial cluster analysis changed when covariates were included in the model. When the analysis is adjusted for age, only one statistically significant cluster remains \((p = 0.043)\). This cluster is different in size and shape and includes fewer census block groups than the primary cluster found in the unadjusted model. It contains 26 cases, 3.3 times more cases than expected. The age-adjusted cluster is in the same general geographic region (e.g. the southern Piedmont), however, and encompasses portions of both clusters found in the initial unadjusted model. The log likelihood ratio (LLR) dropped from 12.93 to 12.23, indicating that age explains some of the excess in gastroschisis cases.

In two subsequent models adjusting for age plus race and age plus parity the LLR dropped, indicating a decrease in the strength of most likely cluster which signifies that race and parity explain some of the excess in gastroschisis cases. The location of the most likely cluster in both the age–race- and age/parity-adjusted models included the same census block groups. This cluster encompassed a larger area than the age-adjusted model and included a greater number of cases, 59, nearly 2.2 times more than expected. Although the p-values for both the age–race \((p = 0.051)\) and age–parity \((p = 0.053)\) adjusted analyses were of borderline significance at the \(p < 0.05\) level, the fact that the cluster persists in the same general geographic location across all covariate-adjusted models is compelling and merits further investigation.

The final two models adjusted for maternal age plus Medicaid status and age plus smoking status. For both analyses we found one significant cluster \((p = 0.014 \text{ for age/Medicaid and } p = 0.028 \text{ for age/smoking})\), both of which include the same census block groups and the same number of cases as the age-adjusted model. The LLR increased to 12.61 for the age/smoking-adjusted model and 13.45 for the age/Medicaid-adjusted model, indicating that these covariates do not explain the excess in gastroschisis cases. The relative risk within the cluster was 3.4 for the age/smoking model and 3.5 for the age/Medicaid model, the highest of any model. The fact that the size and shape of the age/Medicaid and age/smoking clusters are the same as the age-adjusted model suggests that Medicaid and smoking status do not explain any more of the excess of gastroschisis cases than age alone.

Discussion

The initial unadjusted model indicated two significant clusters of gastroschisis, the size and location of which changed dramatically when we adjusted the model for age, race, parity and smoking status, the four covariates with the strongest relationship to gastroschisis prevalence in this study population. The large cluster to the east of Charlotte disappears when age is adjusted for, which suggests a disproportionately large number of young mothers in...
the area is responsible for the large number of gastroschisis cases. While the clusters we detected using the covariate-adjusted models did not overlap perfectly, they did consistently include an area in the rural southern Piedmont just north of the cities of Gastonia and Charlotte. There appears to be a localized cluster of gastroschisis in North Carolina that persists through all analyses and merits further investigation.

This finding fills an important gap in the literature. Prior research on gastroschisis in North Carolina has shown a gradual increase in the birth defect over the past 10 years (Laughon et al., 2003; Wall & Meyer, 2006) and anecdotal evidence from health professionals has suggested a higher prevalence in certain geographic areas of the State. However, this is the first statistical analysis done to formally evaluate the possibility of spatial clusters and test whether the prevalence of gastroschisis is significantly higher within those clusters when compared to the rest of North Carolina. We used a spatial scan statistic because it does not require a priori knowledge of the geographic location, spatial scale or size of a cluster before conducting the analysis, thereby ameliorating the problem of preselection bias. The scan statistic also allows us to adjust for underlying population density and demographic characteristics so we can be more confident that observed clusters are not simply artifacts of unequal population distribution.

We believe this cluster of gastroschisis cannot be readily dismissed as a chance occurrence, and our future analyses will examine potential environmental exposures in this population. In this study, we adjusted for several risk factors: age, race, parity, smoking and Medicaid status (usually a proxy for low income or poverty). There are additional risk factors hypothesized in the literature for which we, unfortunately, do not have individual-level or population-level data. For example, recreational drug use (cocaine, amphetamine, marijuana, or LSD) has been linked with increased risk for gastroschisis (Forrester & Merz, 2006; Torfs et al., 1994) as have some over-the-counter medications such as pseudoephedrine and aspirin (Kozer et al., 2002; Torfs et al., 1996; Werler et al., 1992). Maternal nutritional deficits have also been

![Fig. 2. Close-up view of primary clusters of gastroschisis births detected using covariate-adjusted models, North Carolina, 1999–2004.](image)
linked to increased risk for gastroschisis (Lam & Torfs, 2006; Torfs, Lam, Schaffer, & Brand, 1998). Unfortunately we have no information on the local or regional variation of these behaviors, so we cannot tell if they partly explain the observed cluster.

The cluster we observed in this study encompasses a region of North Carolina that is both geologically and economically unique. The soil composition (mainly metamorphic rocks such as slate and gneiss) is unique to the western Piedmont and the observed cluster is sandwiched between the slopes of the Blue Ridge Mountains and an area of sandy soils referred to as the Sandhills. This combination of soil types, among other factors, influences groundwater recharge and discharge and surface water flow in the region. The cluster also covers one of the main textile producing areas of the state. Textile mills use considerable quantities of water for wet-processing activities such as washing, bleaching and dyeing and mill water is often laden with chemicals when it is discharged into surface and groundwater sources. While we certainly do not have enough information to suggest that textile mill practices are the cause of high gastroschisis rates in the rural southern Piedmont, the geographic pattern of the cases coupled with the density of textile operations and soil composition suggests a possible direction for future research. Since no data on occupation or industry association is available from the birth record, contextual data from the census on labor force participation could be used to examine this relationship.

The increase in birth prevalence of gastroschisis in different populations and different geographic locations over time also suggests the possibility of exposure to environmental contaminants. Studies examining the relationship between gastroschisis births and proximity to point source pollutants are rare and far from conclusive. The EUROHAZCON multicenter case–control study found an increased risk of gastroschisis within 3 km of a hazardous waste landfill but these results were only borderline significant (Dolk et al., 1998). Fielder et al. (2000) also found significantly higher rates of gastroschisis than expected in electoral wards within 3 km of a landfill site. However, a study by Morris et al. (2003) found no association between gastroschisis and omphalocele (another type of abdominal wall defect) and residence within 2 km of a landfill. Data on some environmental risk factors, such as landfill and hazardous waste sites, are publicly available and will be incorporated into future analyses of the present data in order to determine whether such environmental hazards may explain the excess of gastroschisis cases in our observed cluster.

It is important to put into perspective the magnitude of the excess risk observed within the cluster in this study. The cluster observed in the age-adjusted model contains 26 of the 240 gastroschisis cases, approximately 10 percent of all cases in an area with only 5 percent of the total population. The larger cluster observed in the age/race- and age/parity-adjusted models contained 59 of the 240 gastroschisis cases, nearly one quarter of all the cases that occurred in the state, but only about 9 percent of the population lives in this area. This translates to a more than two-fold greater odds of gastroschisis within both the age-adjusted and age/race- or age/parity-adjusted clusters (odds ratio of 2.6 and 2.2, respectively) when compared to the rest of North Carolina.

This study demonstrates the usefulness of spatial cluster analysis in exploratory etiological research of birth defects. The methods adjust for known risk factors for gastroschisis and illustrate the importance of adjusting spatial clusters for underlying population. If the purpose of cluster analysis is not only to identify the approximate location of clusters but also to target future research activities or public health initiatives, finding the location of the “true” cluster after adjusting for the underlying population distribution can prevent researchers from focusing such efforts in the wrong area. Furthermore, the spatial patterns observed in the data can be used to elicit etiological clues about birth defects such as gastroschisis, and generate hypotheses about the causal mechanisms responsible for the cluster. Comparing the socio-environmental characteristics of clustered versus non-clustered cases may reveal similarities or differences, which may, in turn, give clues to disease etiology (Draper, 1997; Williams, Alexander, Cartwright, & McNally, 2001).

It is important to keep in mind that the geographic boundaries of the clusters detected in this study are approximations of the “true” clusters. This means that while we know the general location of the cluster, we are uncertain as to the exact boundaries. As with any ecological analysis, we cannot say that the whole population living within the cluster area is at the same risk for giving birth to an infant with gastroschisis. Women have varying levels of risk, which depend on their individual characteristics, behaviors, and family histories. However, the presence of the cluster suggests that an added risk factor, perhaps environmental, may exist in that area. This geographic analysis uses residence at birth. Studies have shown that between 25 and 30 percent of women change residence between conception and birth (Fell, Dodds, & King, 2004; Khoury et al., 1988; Shaw & Malcoe, 1992). However, a majority of these moves appear to be local (e.g. within the same city or county) (Fell et al., 2004; Khoury et al., 1988) and the characteristics of women who do move are similar to those who do not (Canfield, Ramadhani, Langlois, & Waller, 2006).

Caution should be exercised when interpreting the results of geographic studies that use maternal residential address at delivery, especially if trying to ascribe the case of a cluster to some local environmental exposure.

In summary, we have identified a statistically significant excess of gastroschisis in the rural southern Piedmont of North Carolina which persists even after controlling for known risk factors. While gastroschisis has increased in North Carolina over the past decade and anecdotal evidence from clinicians in the State suggested the presence of one or more clusters of this birth defect, no spatial statistical analysis had been conducted until now. The spatial scan statistic enabled us to evaluate more reliably the location and strength of the clustering effect without the bias that could be introduced when researchers have some prior knowledge of the geographic location or size of a cluster. Future research will focus on possible environmental causes of the clustering.

References


