Using Watershed Boundaries to Map Adverse Health Outcomes: Examples From Nebraska, USA

Brittany Corley
University of Nebraska Medical Center

Shannon L. Bartelt-Hunt
University of Nebraska-Lincoln, sbartelt2@unl.edu

Eleanor Rogan
University of Nebraska Medical Center

Donald Coulter
University of Nebraska Medical Center

John Sparks
University of Nebraska Medical Center

See next page for additional authors

Follow this and additional works at: https://digitalcommons.unl.edu/civilengfacpub

Corley, Brittany; Bartelt-Hunt, Shannon L.; Rogan, Eleanor; Coulter, Donald; Sparks, John; Baccaglini, Lorena; Howell, Madeline; Liaquat, Sidra; Commack, Rex; and Kolok, Alan S., "Using Watershed Boundaries to Map Adverse Health Outcomes: Examples From Nebraska, USA" (2018). Civil Engineering Faculty Publications. 121.
https://digitalcommons.unl.edu/civilengfacpub/121

This Article is brought to you for free and open access by the Civil Engineering at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Civil Engineering Faculty Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
Authors
Brittany Corley, Shannon L. Bartelt-Hunt, Eleanor Rogan, Donald Coulter, John Sparks, Lorena Baccaglini, Madeline Howell, Sidra Liaquat, Rex Commack, and Alan S. Kolok
Using Watershed Boundaries to Map Adverse Health Outcomes: Examples From Nebraska, USA

Brittany Corley1, Shannon Bartelt-Hunt2, Eleanor Rogan3, Donald Coulter4, John Sparks4, Lorena Baccaglini1, Madeline Howell3, Sidra Liaquat1, Rex Connmack5 and Alan S Kolok3,6

1Department of Epidemiology, University of Nebraska Medical Center, Omaha, NE, USA.
2Department of Civil Engineering, University of Nebraska–Lincoln, Omaha, NE, USA.
3Department of Environmental, Agricultural & Occupational Health, University of Nebraska Medical Center, Omaha, NE, USA. 4Division of Hematology/Oncology, Department of Pediatrics, University of Nebraska Medical Center, Omaha, NE, USA. 5Department of Geography & Geology, University of Nebraska Omaha, Omaha, NE, USA. 6Idaho Water Resources Research Institute, University of Idaho, Moscow, ID, USA.

ABSTRACT: In 2009, a paper was published suggesting that watersheds provide a geospatial platform for establishing linkages between aquatic contaminant, the health of the environment, and human health. This article is a follow-up to that original article. From an environmental perspective, watersheds segregate landscapes into geospatial units that may be relevant to human health outcomes. From an epidemiologic perspective, the watershed concept places anthropogenic health data into a geospatial framework that has environmental relevance. Research discussed in this article includes information gathered from the literature, as well as recent data collected and analyzed by this research group. It is our contention that the use of watersheds to stratify geospatial information may be both environmentally and epidemiologically valuable.

KEYWORDS: Watershed, agrichemicals, environmental health, epidemiology, agricultural runoff

Introduction

Contaminants in water can cause adverse health effects. A poignant example of this is the contamination of Flint, Michigan, drinking water with lead and the subsequent elevated blood levels in children. In that community, lead contamination is a posttreatment issue, as the corrosive water from the Pontiac River solubilized lead from the distribution system that delivers water to individual households.1 The geospatial distribution of lead exposure in Flint is relatively easy to map as the source of the exposure (the drinking water) is easy to identify, and the latency period between the initial exposure and elevated blood lead levels is short.1

Unlike the relationship between lead in drinking water and elevated blood lead levels, there are other examples where the water contamination is diffuse and the latency period for adverse health effects may be years or decades. One such example would be the development of cancers or birth defects when individuals are exposed to water contaminated with agrichemicals.2 The agrichemicals can include pharmaceuticals, such as steroids or antibiotics used on livestock, nutrients, such as nitrates and phosphates, as well as herbicides and insecticides. Agrichemical residues have been found in food, water, and juices.3–5 Although the concentrations found are within set safe limits, the true health risk at these levels is not well understood and could be subjected to synergistic effects.6 Pesticides in particular have been detected in human breast milk, which has led to concerns about prenatal exposure and various health effects in children.7 In this case, the chemical source can be agricultural fields, and efforts to map the geospatial organization of the exposure must invariably involve mapping large areas upstream from the communities whose water is affected.

Efforts to map the geospatial distribution of diseases are often conducted by first compartmentalizing the relevant geography into established geographical census units, such as census blocks and block groups, census tracts, zip codes, counties, or states. Although this may be appropriate for some environmental exposures, it may not be at all appropriate for waterborne agrichemicals, steroids, and antibiotics because the pathways by which these contaminants travel do not respect anthropogenic geospatial boundaries.8 Rather, these contaminants become mobile when rainstorms induce surface runoff that transports the chemicals from land and deposits them into local waterways. These waters ultimately flow downstream within well-defined watersheds.

A watershed is a topographic area within which surface and shallow groundwater drain to a specific point.9–10 States or other geographic regions can easily be divided into specific watersheds, as everyone live within one watershed or another. Furthermore, when it comes to waterborne contaminant exposure, two individuals who live miles apart, but within the same watershed, may experience similar exposures, whereas two...
individuals who live close to each other, but in different watersheds, may experience very different exposure profiles.

The central hypothesis of this article is that there is an advantage in conducting geospatial analysis relative to adverse health outcomes using watersheds, rather than anthropogenic census tracts, particularly with respect to agrichemical runoff. We contend that the relationship between watershed geography and contaminant distribution is critical for certain classes of chemical contaminants, and this article illustrates a methodology for investigating that relationship.

Relationship Between Watershed Boundaries and Population Geography

The watershed

From an epidemiologic perspective, exposure assessment is more complicated when dealing with environmental health studies than it is in occupational health studies. Exposure assessment is the process of measuring the magnitude, frequency, and duration of exposure to a chemical, and occupational health studies typically have excellent assessments of exposure due to the defined geospatial boundaries (ie, the workplace) and the use of predefined exposure definitions by job title.

In contrast to occupational exposure assessment, two of the largest problems encountered when attempting to define a chemical exposure through natural waters are the lack of defined boundaries and the spatial heterogeneity of exposure. Poorly defined boundaries (also known as fuzzy objects) occur when there is no clear boundary of an object in geographical information systems (GIS), an eventuality that is common when dealing with highly variable metrics within a geography, such as soil type. Spatial heterogeneity of exposure occurs when there is an uneven distribution of various concentrations of chemicals and exposures within a given spatial area, and it has been a problem in studies of exposures to airborne contaminants. When the population was organized by population geography, spatial heterogeneity of exposure has caused problems in studies looking at agricultural exposures and birth defects, environmental movement of contaminants and neural tube defects, urban environment exposures and cancer incidence, and iodine exposure and thyroid cancer.

We propose that environmental assessment of contaminant exposure via natural waters can best be dealt with when the watershed is used as the defining geospatial boundary. A watershed is defined as an “area of land where drainage of streams and rainfall meet at a common outlet, such as the outflow of a reservoir, mouth of a bay, or any point along a stream channel.” A watershed is also an area of connectivity where any activity that affects the water quality, quantity, or rate of movement at one location can change the characteristics of a watershed downstream, providing a common level of exposure between contaminants.

The US Geological Society (USGS) has subdivided the United States into successively smaller hydrologic units (HUs) which can be classified as follows: regions (HU 2), subregions (HU 4), basins (HU 6), subbasins (HU 8), watershed HU (10), and subwatershed (HU 12), respectively. There are 21 major regions in the United States, and these are either major river drainage systems or drainage systems for several rivers, such as the Missouri Region and Texas Gulf Region, respectively. Next, there are 222 subregions in the United States; these include drainage areas for a river system. There are 370 basins in the United States, which are nested within the subregions. The cataloging unit, which is what the term watershed most frequently is used to represent, has 2264 units across the United States.

The population

Although the geospatial distribution of waterways can be represented through HUs, the most common way that geospatial distribution of humans is represented in the United States is through census entities. These entities include the nation, regions, divisions, states, counties, census tracts, census block groups, and census blocks (Table 1).

Relative to human health studies, the most commonly used population geographies are city and county. Studies that have looked at water exposure and human health outcomes often have limitations of potential classification error. Classification error is a type of information bias in which study participants are assigned to an incorrect classification group. For example, if a person is assigned to a county, but the county contains multiple watersheds, this can cause misclassification, as the assumed exposure may be very different from the actual one. For studies focusing on exposures to natural surface water, we recommend using the watershed as the primary geography.

Overlap

For watersheds to be used in environmental exposure assessment, it is necessary to overlap the human census data with the HU data. Unfortunately, watersheds and population geographies were developed by different groups of professionals for very different reasons, and, consequently, there is very little overlap between the two. The watershed geography focuses on natural water flow, whereas the population geography focuses on governmental delineations and how the population is organized. For this reason, very few points of overlap are seen within the two (Figure 1). Unfortunately, the lack of overlap between watershed HU and human census tracts is not improved when different levels of organization are used. Due to this geographic mismatch, it is not prudent to assume that exposures to contaminated water are consistent across census groups, such as counties or state boundaries.
Choosing the appropriate HU

To map the incidence of adverse health outcomes by watershed, one of the HUs needs to be selected above the others. In Nebraska, the selection of HU was driven by two different descriptors: the exposure profile (land use) found within the state and the underlying population density.

In Nebraska, soil and precipitation patterns change from East to West, as well as agricultural land use. The term “land use” is defined as the land’s purpose relative to human activity and is usually, but not always, related to land cover. For lands to be useful for agriculture, certain environmental factors are required, including soil conditions and climate (e.g., soil texture, mineralogy, precipitation patterns) which determine the suitability for crop production (e.g., type of fertilizer/pesticide application, type of crop). Furthermore, as the environmental conditions of Nebraska change geospatially, it is likely that agrichemical exposure also changes accordingly.

Population density is defined as “the number of people living per unit of area (e.g., per square mile); the number of people relative to the space occupied by them.” Parts of Nebraska have a very sparse population density. The variance in population density is high not only within states but also between states. Some states, such as California (Figure 2A), are primarily urban and therefore are more densely populated. Even within traditional rural or farming states, such as Kentucky and Nebraska, the population density can vary greatly. For example, Kentucky (Figure 2B) is for the most part even populated with a scattering of urban centers. Nebraska (Figure 2C), however, is very sparsely populated with very few densely populated urban centers.

Nebraska can be divided using 6 different HU codes (Figure 3). If the HU delineation were too large, then watersheds with vastly different agricultural practices and therefore exposures would be combined, thereby reducing the effectiveness of the analysis. For example, HU code 2 encompasses the entire Missouri River Valley and the entire state of Nebraska. Obviously, this does not allow any discrimination of exposures within the state and includes so much terrain that there is a vast amount of geographic variability within it. The HU codes 4 and 6 were also too large and had the same problems as HU code 2.

If the HU code selected is too small, then the population within each watershed designation would be too small to allow any meaningful analysis, as many, perhaps the majority, of these watersheds would not cover a geography where an adverse health impact had occurred.
A geospatial analysis favoring a watershed approach lends itself much more readily to some adverse health outcomes rather than others. For example, waterborne contaminants have been linked to birth defects, pediatric cancers, and thyroid cancer. Although all 3 diseases differ in incidence, cause, and outcomes, the analysis of potential risk factors may benefit from the use of a watershed approach. For this reason, they will be used as case studies for this article.

**Birth defects**

The leading cause of infant mortality in the United States is birth defects or congenital abnormalities. The cost of birth defect–related hospitalizations for all age groups represents 5.2% of total costs for all hospital discharges. Not only are birth defects costly but they also affect 1 in every 33 live births in the United States. In 2011, Nebraska had a higher burden of birth defect–related death in relation to the nation, with rates of 1.94 per 100,000 and 1.27 per 100,000, respectively.

The risk factors for birth defects are mostly unknown and vary depending on the type. The most notable risk factors include alcohol, illicit drug use in pregnancy, smoking, obesity, diabetes mellitus, phenylketonuria, multiple gestations, advanced maternal age, advanced paternal age, family history, folic acid deficiency, medication exposures, and radiation exposure.

Several studies support the hypothesis that agrichemicals play a role in the cause of certain birth defects. Pesticides are known to be both reproductive and neurotoxic agents and have been shown to be teratogenic in animal studies. Pesticides, including atrazine, alachlor, and chlorpyrifos, are classified...
as endocrine disruptors, whereas bifenthrin and diuron are developmental toxicants.\textsuperscript{15} Studies have shown a potential association between pesticide exposure before or during pregnancy and various types of birth defects.\textsuperscript{60–53}

Applications of spatial assessments for birth defects in relation to agricultural land use have provided further insight regarding the cause of birth defects. For instance, spatial attributes such as elevation, soil types, lithology, watersheds, fertilizer use, and neighborhood characteristics are associated with specific neurological birth defects.\textsuperscript{17,54}

**Pediatric cancers**

Cancer is the second most common cause of death among children in the United States.\textsuperscript{55} A child born in the United States has 0.35% chance of developing cancer before 20 years of age; this is equivalent to an average of 1 in 285 children being diagnosed with cancer before 20 years of age.\textsuperscript{56} The Nebraska rates of pediatric cancer were reported to be above the national average in 2010 to 2012; however, the trend has regressed back to the national average in recent years.\textsuperscript{57} The cause of childhood cancer is mostly unknown, but some cases can be linked to genetic causes.\textsuperscript{58,59} Despite the unknown cause for most childhood cancers, recent research has linked certain cancers to environmental factors, such as hematologic malignancies to oil and gas production,\textsuperscript{60} renal cancers to industrial and pesticide pollution exposure,\textsuperscript{61} retinoblastomas to pesticides,\textsuperscript{62} and leukemia, neuroblastoma, and hepatic tumors to crop production proximity.\textsuperscript{63}

**Thyroid cancer**

In recent years, there have been stable diagnostic rates for thyroid cancer from 2010 to 2012,\textsuperscript{64} with rates for men as 1 in 169 and for women as 1 in 58.\textsuperscript{64} Whereas in 2017, rates for thyroid cancer were reported as 1 in 163 for men and 1 in 57 for women.\textsuperscript{65} In Nebraska, the rates for thyroid cancer overall and for women are higher than the national average: 19.4 per 100,000 (women, Nebraska), 12.7 per 100,000 (overall, Nebraska), and 17.3 per 100,000 (women, national), and 11.7 per 100,000 (overall, national).\textsuperscript{66}

There are 4 main types of primary thyroid carcinoma: papillary, follicular, anaplastic, and medullary. These groups typically share risk factors but not always.\textsuperscript{67} Other risk factors for thyroid cancer include genetic predisposition, radiation exposure, iodine intake, preexisting thyroid disease, age, sex, hormonal and reproductive factors, geographic factors, ethnic factors, diet, and drug exposure.\textsuperscript{68} Environmental factors other than radiation exposure have been examined in recent studies, including heavy metals\textsuperscript{69} and pesticide exposure.\textsuperscript{29}

**Population at risk**

A population at risk is defined as the population that has a chance of developing a disease or condition of interest. This article features 3 different adverse health outcomes: birth defects, pediatric cancer, and thyroid cancer, and each of the 3 has a different population at risk.

**Defining the population at risk**

Finding a population at risk for an adverse health outcome implies that the researcher understands who in the population is at risk. One way to define this population is to use your case definition, ie, how the cases are determined to truly be a case. For this article, the following definitions were used, and the cases were gathered from the Nebraska Birth Defects Registry, and the Nebraska Cancer Registry at the Nebraska Department of Health and Human Services.

**Birth defects**

The definition used for birth defects was any congenital anomaly from a baby resulting from a live birth that was recorded in the Nebraska Birth Defects Registry from 1995 to 2014.\textsuperscript{30} Based on this definition, the population at risk was infants born alive in Nebraska from 1995 to 2014.

**Pediatric cancer**

The definition used for pediatric cancer was any malignancy occurring in someone aged 19 years and below, which was recorded in the Nebraska Cancer Registry from 1987 to 2014.\textsuperscript{70} Based on this definition, the population at risk was any child living in Nebraska aged 19 years and below from 1987 to 2014.

**Thyroid cancer**

The definition used for thyroid cancer was any case of thyroid cancer occurring at any age that was recorded within the Nebraska Cancer Registry from 1987 to 2014.\textsuperscript{70} Based on this definition, the population at risk was any person living in Nebraska from 1987 to 2014.

**Classifying population data based on watershed delineations**

Converting the geography of populations into the geography of watersheds may result in misplacing individual cases in a geographically incorrect watershed. Clearly, the smaller the population geography unit, the lower is the probability of misclassification error. An example of this is shown in Figure 4. The initial watershed map (Figure 4A) shows Nebraska with the HU 8 watersheds overlaid on the state. As mentioned above, counties (Figure 4B) do not overlap well with watersheds. The mismatch is exacerbated by the method used to assign counties to watersheds in GIS, which is by county center. Zip codes (Figure 4C) and census blocks (Figure 4D) overlap better with watersheds. For this article, census blocks were chosen to categorize the population at risk.
According to the US Cancer Statistics Working Group (USCS), relative incidence rates containing less than 16 cases are unstable and prone to error. Based on this proposition, the following minimum populations per watersheds are required. For example, for birth defects, a population of 600, on average, is necessary; to obtain more than 16 cases. For this reason, HU codes 10 and 12 were too small due to very low population numbers, particularly in the panhandle (western) section of the state. Based on these observations, HU 8 was used for preliminary mapping.

Incidence Rate Calculations
To determine which watersheds to include when mapping statewide adverse health impacts, it was first necessary to determine watershed incidence rate. The true incidence rate \( \left( I_T \right) \) is the number of cases that occur over a given time divided by the current population at risk and is typically reported per 100,000 (equation (1)). General incidence rate for each watershed is calculated as follows:

\[
\text{Incidence rate} (I_T) = \frac{\text{New Cases}_{\text{Watershed}}}{\text{Population at Risk}_{\text{Watershed}} \times \text{Time of analysis}} \times 100,000
\]  

Due to the different time periods examined and the different sources of at-risk populations, there were 3 different incidence rates used in this analysis that were based on the general crude incidence rate calculation for each watershed. For example, thyroid cancer cases were aggregated more than 28 years and featured a population at risk that was equivalent to the entire population (all ages); therefore, Census 2000 population counts were multiplied by 28 years to calculate the total population at risk. In contrast, birth defects were aggregated more than 19 years and featured a population at risk that was equivalent to all live births from 1995 to 2014. Regardless of the slight differences in how incidence rates were calculated, the final data sets that were developed could then be mapped by watershed.

Although the 3 adverse health impacts that this article reports can all be considered rare, pediatric and thyroid cancers are rarer than birth defects. The probability of a birth defect is approximately 3%, and the incidence rate is likely be mostly stable in Nebraska. For pediatric and thyroid cancers, the probability is much lower (0.3% for pediatric cancers and thyroid cancers between 0.6% for men and 1.7% for women). What this means functionally is that that maps developed for pediatric and thyroid cancers are, by necessity, more variable than those developed for birth defects.

Determining Which Watersheds to Include
The main way of excluding watersheds from analysis is based on a percent error calculation. In this approach, one calculates the difference adding one additional case per watershed would introduce to the incidence calculation. The first step was to complete the percent error calculation (equation (3)), which includes two parts: the true incidence (equation (1), \( I_T \)) and the “error” incidence (equation (2), \( I_E \)), which is the incidence if one additional case was present in the watershed. Generalized “error” incidence rate and percent error for each watershed is calculated as follows:

\[
\text{Percent Error} = \left( \frac{I_E - I_T}{I_T} \right) \times 100
\]  

For birth defects, thyroid cancer and pediatric cancer, the percent error cutoff chosen was 20. The USCS reported a high rate of error to be 25% and that is the cutoff they have used; however, due to the small population sizes in rural Nebraska, this was relaxed to 20% to limit the number of excluded watersheds. Based on this cutoff for birth defects, 22 of 72 watersheds were removed from analysis, whereas for pediatric cancer, 35 of 72 watersheds were excluded from analysis, and for thyroid cancer, 29 of 72 watersheds were excluded from analysis.
For birth defects and thyroid cancer, watershed exclusion based on percent error was conducted maps were created (Figure 5). For pediatric cancers, unfortunately, over 40% of the available HU 8 watersheds were excluded. Due to the overall low population values and high percent error, the decision was made to recreate the pediatric cancer graph using HU 6, to allow to more data to be usable within the state. Thus, pediatric cancers are reported using both HU 6 and 8 (Figure 6).

**Results**

In Figures 5 and 6, the unadjusted incidence rates for Nebraska, for birth defects, pediatric cancer and thyroid cancer are mapped. For birth defects (Figure 5A), the incidence ranges from 0 to 7692 per 100000; for pediatric cancers (Figure 6B), the incidence ranges from 0 to 177 per 100000; and finally for thyroid cancers, the incidence ranges from 3.25 to 16.96 per 100000.

When incidence rates across the 3 different adverse health outcomes were compared with each other, there were no significant intercorrelations (Table 2). Intercorrelations might have occurred if one or more of the watersheds were contaminated with a key aquatic compound that is known to be associated with one of more adverse health outcomes. When the watersheds were viewed in composite, some light was shed on this lack of intercorrelation. For birth defects, the watersheds of interest include the Loup River (28), the north fork of the Elkhorn River (30), the lower Platte River (50), the lower Elkhorn (52), the upper Republican (58), and the south fork of the big Nemaha River (68). For pediatric cancers, the watersheds of interest include the Turkey River (3), the Cedar River

![Figure 5. Nebraska HU 8 profile: (A) birth defect incidence (1995-2014) and (B) thyroid cancer incidence (1987-2014). HU indicates hydrologic unit.](image)

![Figure 6. Nebraska pediatric cancer (1987-2014): (A) HU 6 and (B) HU 8. HU indicates hydrologic unit.](image)

<table>
<thead>
<tr>
<th></th>
<th>BIRTH DEFECT INCIDENCE</th>
<th>PEDIATRIC CANCER INCIDENCE</th>
<th>THYROID CANCER INCIDENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Birth defect incidence</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pediatric cancer incidence</td>
<td>-0.07</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Thyroid cancer incidence</td>
<td>0.12</td>
<td>0.07</td>
<td>1</td>
</tr>
</tbody>
</table>

Abbreviation: HU, hydrologic unit.
For example, thyroid cancer has been linked to higher average levels of nitrate in water supplies (exceeding 5 mg/L). There was also a suggestion for the potential of a link between volcanic elements in water and papillary thyroid cancer, although this potential link needs to be investigated further. Although, overall, pediatric cancer has been linked to pesticides in water, liver cancer has been specifically linked to arsenic in water supplies. Similarly, specific birth defects have been linked with various waterborne exposures, including central nervous system defects linked with trihalomethanes, carbon tetrachloride, trichloroethylene, and dichloroethylenes. Oral cleft defects were linked with trihalomethanes, carbon tetrachloride, trichloroethylene, tetrachloroethylene, and dichloroethylenes. Major cardiac defects were linked with trihalomethanes, benzene, and 1,2-dichloroethane. In addition, congenital cardiac disease was linked with trichloroethylene, dichloroethylene, and chromium in ground water. Neural tube defects were linked with carbon tetrachloride, trichloroethylene, and benzene.

Conclusions
The central hypothesis of this article was that there is an advantage in conducting geospatial analysis relative to adverse health outcomes using watersheds, rather than by anthropogenic census tracts, particularly with respect to agrichemical runoff. We contend that the relationship between watershed geography and contaminant distribution is critical for certain classes of chemical contaminants, and this article illustrates a methodology for investigating that relationship.

This use of HUs as geographic spatial polygons for systems that are not necessarily strictly hydrologic in nature has been documented previously. The HUs have previously been used in ecological modeling, which is commonly applied to human health behavior research. It has recently been noted that watersheds seldom circumscribe regions of similarity that influence water quality. Omernik et al correctly point out that HUs are not only composed of watersheds but they are also parts of watersheds. Consequently, from a strict hydrologic point of view HUs may not represent watersheds. Nevertheless, from an epidemiologic point of view, HU delineation brings a natural, rather than an anthropogenic, focus to the process of geospatial mapping of adverse health impacts. Although the delineation of the 3 adverse health impacts featured in this article did not result in strong intercorrelations, we still think that the use of HUs is a novel and dramatic improvement.

Due to the preliminary nature of this methodology, two important factors were not previously discussed which include the inclusion of sociodemographic information and the potential effects of upstream processes on water quality within the watershed. Due to the nature of the case studies, used social demographic data were not included. This was done to avoid the potential of introducing an ecological fallacy within the data. However, if one were to conduct a cohort study, a case-control study, or a cross-sectional study, then correction for sociodemographic data would easily be included and is necessary to have corrected and usable incidence rates.

There is potential for upstream human pressure or agricultural activities that influence water quality downstream taking into account water routing, evapotranspiration, precipitation, and other climate variables. This was not discussed previously in this article due to the preliminary nature of the study. However, in future refinements of this methodology, this will be included.

Future Work
In the future, we plan to refine this methodology and incorporate water quality data into the approach. Environmental data sets on water quality will include the water quality data from the Environmental Protection Agency STORage and RETrieve (EPA STORET) data set, as it is comprehensive and includes data from several sources. This process may prove to be complicated as was suggested by Omernik et al for some HU delineations may experience contaminant input from multiple sources. A method to average water quality over the spatial HU scale used in the analysis will need to be developed. A first step of this may be to compare the HUs used above with the land use maps for Nebraska to quantify the variation within each HU area. Future work also includes applying this process to other states within the Midwest to observe whether they show similar profiles.

Overall, the methodology demonstrated in this article is a way to identify areas of interest with respect to watersheds and human health. As demonstrated by this study, there appears to be link between specific watersheds and the incidence of birth defects, pediatric cancer, and thyroid cancers in the state of Nebraska.

Acknowledgements
The authors would like to thank the Nebraska Birth Defects Registry (NBDR) and the Nebraska Cancer Registry for the data used in this article.

Author Contributions
JS, DC, and ASK conceived and designed the experiments. BC, MH, and SL analyzed the data. BC wrote the first draft of the manuscript. ASK contributed to the writing of the manuscript. BC, ASK, LB, SB-H, and ER agree with manuscript results and conclusions and jointly developed the structure and arguments for the paper. BC, ASK, LB, and SB-H made critical revisions and approved final version. All authors reviewed and approved the final manuscript.