Yale University
EliScholar – A Digital Platform for Scholarly Publishing at Yale

Public Health Theses | School of Public Health

January 2016

Neighborhood-Level Socioeconomic Status And County Fairs And The Risk Of Shiga Toxin-Producing Escherichia Coli Incidence In Connecticut, 2012 To 2014

Gili Hrusa
Yale University, gili.hrusa@yale.edu

Follow this and additional works at: http://elischolar.library.yale.edu/ysphtdl

Recommended Citation
Hrusa, Gili, "Neighborhood-Level Socioeconomic Status And County Fairs And The Risk Of Shiga Toxin-Producing Escherichia Coli Incidence In Connecticut, 2012 To 2014" (2016). *Public Health Theses*. 1129.
http://elischolar.library.yale.edu/ysphtdl/1129

This Open Access Thesis is brought to you for free and open access by the School of Public Health at EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Public Health Theses by an authorized administrator of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact elischolar@yale.edu.

Neighborhood-Level Socioeconomic Status and County Fairs and the risk of Shiga Toxin-Producing *Escherichia coli* Incidence in Connecticut, 2012 to 2014

Gili Marie Hrusa
Epidemiology of Microbial Diseases
Yale School of Public Health
2016

Acknowledgements

"Knowledge in the end is based on acknowledgement"- Ludwig Wittegenstein

I would like to acknowledge and give thanks and Dr. James Hadler and Dr. Linda Niccolai for their agreeing to guide my project. Their suggestions have been instrumental in the shaping of my Master's thesis project into its final form. Apart from his technical input, I found Dr Hadler's patience and passion for public health truly inspiring. I would also like also to thank Dr. Jocelyn Mullins of the CT Department of Public Health for her continued help and thought-stimulating ideas for the project. I would be remiss in not acknowledging the staff at the Emerging Infections Program who collectively motivated and encouraged me during the process of writing this thesis. Specifically, I would also like to share my gratitude to Sharon Hurd, Paula Clogher, Danyel Olsen, Tamara Rissman, and Jim Meek who were always behind me with words of encouragement and support. Finally, I want to express a special appreciation to my colleagues and friends from the FoodCORE team.

Table of Contents

Abstract

Background: Shiga toxin-producing *E. coli* (STEC) is a reportable illness. It is associated with a spectrum of clinical manifestations that range from watery and/or bloody diarrhea to hemolytic uremic syndrome (HUS), a life threatening condition which causes acute renal failure in both children and adults. A study in Connecticut (CT) by Whitney et al. (2015) found a strong association between higher socioeconomic status (SES) and Shiga toxin-producing *E. coli* incidence. In the same study, there was a weaker association demonstrated between rurality and STEC incidence as well. The primary aim of this study is to investigate the relationship between SES disparities and the incidence of reported STEC-cases in CT between 2012 and 2014. By linking surveillance data with census tract information, populations can be identified that exhibit higher levels of disease burden. The value of geocoding and linking surveillance data to census data in describing the epidemiology of infectious diseases has been demonstrated in many instances. The study will also investigate the relationship of the geographical distribution and densities of cattle/ruminants and related environmental structures, specifically to farm animal-displaying county fairs, with human STEC infection in CT. County fairs have been suspected as an important point source for Shiga toxin-producing *E. coli* transmission to humans, following the work of Crump (2003) where agricultural fairs exhibiting livestock were implicated in STEC O157:H7 (STEC O157:H7) outbreaks. Data from Keen (2006) suggest that STEC O157 is common and highly transmissible among the livestock displayed at agricultural fairs and can persist in the environment well after the fair is concluded. Thus, fairs where livestock and other animals are displayed and kept for a time in very close contact conditions are worth exploring as sources of the STEC disease.

Methods: A total of 180 incident STEC cases were reported in CT from 2012-2014. For analysis, the cases were linked to neighborhood poverty level which was broken down into four categories based on percent of people living under the federal poverty line: 0–4.99%, 5–9.99%, 10–19.99%, and greater than 20%. County fairs in CT were compiled and geocoded with ArcGIS. Census Tracts were further classified as being in proximity to a county fair if they were located within the five mile buffer zone of a geocoded fair. Using national 2010 census information, three year age- and sex-standardized Shiga STEC incidence rates were calculated for each poverty category and census tract proximate and not proximate to a county fair. High risk cases with onset dates that fell within the ten day period of risk (3 day fair activity and 7 day STEC incubation period) were examined to see if people living in those census tracts at county fair time were at higher risk of STEC infection than they were when the fair was not in operation. County fair operation time was defined as a period of 94 days during summer and fall when significant fair activity occurs. Incidence rates were also analyzed by race/ethnicity and by year. Generalized linear models were employed to investigate the association of STEC risk with neighborhood poverty levels after adjusting for age, sex, race/ethnicity and county fair proximity. A similar model was constructed to see if there was an association of STEC and fair proximity while also adjusting for age, sex, race/ethnicity and neighborhood poverty level.

Results: One hundred and seventy-three of the 180 incident cases of STEC (96%) were geocoded based on the case's address using ArcGIS. Neighborhood poverty level was found to be significantly correlated with STEC risk. Overall, there was a progressive increase in IR with increasing age, females had 1.4-fold higher incidence rates than males, and non-Hispanic whites had both the most cases (83%) and an IR that was six-fold higher than non-Hispanic blacks. Age- and sex-standardized rates for all STEC infections (all serogroups) revealed an inverse relationship between poverty and STEC incidence. These results were consistent with the previous Whitney et al. (2015) study that found STEC risk highest in areas with less poverty. The STEC incidence in the second wealthiest SES category (5-9.99%) increased in 2012-2014 as compared to 2000-2011. The generalized linear model adjusted for age group, sex, race/ethnicity and proximity to fair, also showed a significant association between neighborhood poverty level and STEC risk. In terms of county fairs, the incidence rate was higher in the sum of census tracts with a fair in proximity (1.79 per 100 000 person years vs 1.42 per 100 000 person years, IRR 1.26), but the association was not statistically significant. The cases identified as being potentially exposed through a county fair had an incidence rate of 3.68 per 100,000 person years. This rate was found to be significantly higher in the 94 day fair activity period versus census tracts not in proximity to an operating county fair. However, again, the linear model found no significant association between STEC risk and operating fair proximity.

Conclusions: This study showed that in Connecticut STEC incidence decreases as neighborhood poverty increases. These results mirror the results of Whitney et al. (2015). These findings can be used to more effectively target

education and interventions, starting with higher-income neighborhoods. This data analysis failed to completely elucidate the relationship between the geographical location of county fairs and STEC risk. The generalized linear model showed no association between STEC risk and proximity to a county fair even though the univariate model suggested positive risk association. Looking to the future, studies designed to increase the understanding of the mechanism driving the differences in STEC risk between higher and lower income areas are warranted. Though the relationship between STEC risk and rurality and county fair proximity was not well elucidated, its further exploration appears justified as well. It is likely that the risk presented by county fairs is more complex than what can be represented by simple geographical location data.

Introduction

Food Borne Illnesses

Food safety is an area of high priority in public health in the United States and around the world. Food borne illnesses are caused by a wide variety of etiological agents including viruses, bacteria and a range of parasitic microorganisms. Despite technological advances in food safety, food borne illness persists as a significant public health burden. The Centers for Disease Control estimates that 1 in 6 Americans (48 million people) become ill annually from food borne disease. Of these, 128,000 are hospitalized and 3,000 die, either from one or more of 31 known pathogens or by unspecified etiological agents (Scallan 2011). There is a huge economic cost to such illnesses; *Salmonella* infections alone cause approximately 3.6 billion US dollars annually in direct medical costs and loss in productivity (Hoffman 2015). Despite mitigation efforts, factors such as climate change, pathogen evolution, antimicrobial resistance, and a sharp increase in the country's immune-compromised population (the elderly) contribute to an expanding threat where the severity of such illnesses is on the rise (Goulet 2012). A documented example of the latter is the augmented seriousness of health outcomes suffered by the elderly infected by *Salmonella* or *Listeria monocytogenes* as opposed to younger, more immune competent persons similarly infected (Goulet 2012).

Shiga toxin-producing *E. coli* (STEC)

Escherichia coli (E. coli) is a gram negative bacterium most commonly found as a commensal inhabitant of the lower intestine of mammals. Variants that are pathogenic to humans exist and present a significant health risk (Caprioli 2004). Disease due to Shiga toxin-

producing *E. coli* (STEC) is a reportable food borne illness associated with a spectrum of clinical manifestations that range from watery and/or bloody diarrhea to hemolytic uremic syndrome (HUS), a life threatening condition which causes acute renal failure, most often in children (Winfield 2003, Tarr 2005). HUS develops in approximately 5%-10% of people with STEC-associated diarrhea. Thorpe (2004) estimated that 3 to 7 percent of people suffering from HUS either die or have permanent renal impairment. The pathogenicity of STEC is due to a variety of virulence factors including the production of the Shiga toxin, which draws its name from its similarity to the toxin produced by *Shigella dysenteriae* (Melton-Cesa).

To date, a number of serotypes associated with STEC infections and illness have been identified. The most commonly isolate is serotype O157, which was first recognized as a potentially important human pathogen in 1982 following two outbreaks in the United States with cases characterized by severe bloody diarrhea and general gastrointestinal illness (Riley 1982). Non-O157 STEC strains have since been recognized to be important causes of diarrheal illness globally, with many associated with serious outbreaks as well as isolated cases (Johnson 2006).

The Foodborne Disease Active Surveillance Network (FoodNET)

Since 1996 the Food Borne Disease Active Surveillance Network (FoodNET), a CDC surveillance program operating out of the Yale School of Public Health's Emerging Infections Program (EIP), has been tracking the incidence of O157 and non-O157 STEC infections in ten states, one of which is Connecticut (Scallan 2007). In 2014 the incidence of laboratory confirmed O157 STEC infections in the United States was 0.92 per 100,000 persons. Despite improvements in food safety, there remains much to learn about how environmental and socioeconomic conditions contribute as STEC risk factors.

STEC Risk Factors

Human infection with STEC occurs from the inadvertent ingestion of mammalian fecal material, particularly from cattle infected with the bacterium. Many food items, from bovine products such as undercooked meat or unpasteurized milk, to feces-contaminated vegetables and apple cider are known to have acted as transmission vehicles (Armstrong 1996). Person to person transmission of STEC in family households and institutional settings has also been verified and has led to public health interventions to improve hygiene and avoid transmission in high risk settings such as food preparation and daycare facilities (Thorpe 2004).

Environmental Transmission and Risk factors for STEC

Environmental transmission is also important for STEC. Ruminants, in particular cattle, are natural reservoirs for STEC and typically appear asymptomatic while continually shedding the bacteria into the environment (Gansheroff and O'Brien 2000). *E. coli* O157 serotype is viable for extended periods outside of the host as demonstrated by Wang (1998), where survival for up to 70 days in bovine feces and water was verified. In addition, an extremely low infectious dose of 10 STEC bacteria cells in ingested materials can lead to a potentially disease-causing infection in humans (Feng 1995).

Studies have shown that both direct and indirect exposure to cattle play an important role in STEC infection (Wilson 1997, Michel 1999), however very few studies have attempted to investigate the spatial and geographical distribution of environmental risk factors associated with cattle and other ruminant exposure and its relationship to STEC incidence. Michel (1998) conducted a study in Ontario, Canada which evaluated the spatial association between livestock density and the incidence of verotoxigenic *E. coli* (VTEC) infection in human populations using a Geographical Information System (GIS) and spatial regression models. He found that the risk

for VTEC infection increased in areas where cattle density was higher. Similarly, a study in Sweden described and compared the spatial patterns of farmed cattle and VTEC infection in humans and found a strong correlation between variables of the agricultural structure and human VTEC infection incidence (Kistemann 2004). Studies of this nature have not, as yet, been conducted in Connecticut, nor has STEC been used as the target pathogen.

Study Design

This investigation is divided into two parts. The first focuses on socioeconomic status and its contribution to the STEC incidence observed in CT between 2012 and 2014. The second part of the study assesses the geographical distribution of a socio-environmental structure, farm animal-displaying county fairs, and its effect on with human STEC risk during the same time period.

Area-based Socioeconomic Status and STEC

In addition to the environmental component, this study extends previous work that explored variations in STEC incidence in CT due to neighborhood-level socioeconomic factors. Documenting and explaining the health impacts of social inequalities is important in all areas of public health, including for infectious diseases. A study by Whitney et al. (2015) identified demographic groups in CT at high risk for foodborne bacterial infections by relating STEC infection incidence and area-based socioeconomic status (SES) measures. She found that higher census tract–level SES was associated with a higher incidence of STEC. A more recent study in New York City found the same relationship (Greene 2015). These results are somewhat surprising given that infectious disease studies, including those on other food borne conditions, tend to demonstrate a consistently higher disease burden on populations of lower socioeconomic

status (Zappe-Pasturel 2013, Greene 2015). To test the findings of the Whitney et al. and Greene studies with STEC, this investigation includes an SES component to test its association with infection risk. Census tract–level poverty was used as a socioeconomic status variable by the Public Health Disparities Geocoding Project (Krieger 2005), and following their lead this project will use their coded data to relate SES disparities with incident cases of STEC reported in CT from the year 2012 to 2014, years not covered by the Whitney study.

Testing the Presence of County Fairs as a Risk Factor for STEC Infection

The exploration of county and agricultural fairs as a point source of risk for Shiga toxin-producing *E. coli* transmission to humans follows the work of Crump (2003) where fairs exhibiting livestock were strongly implicated in STEC O157:H7 (STEC O157:H7) outbreaks. Data from Keen (2006) suggest that STEC O157 is common and transmissible among displayed livestock and likely persists in the environment for some time after the fair is concluded. Connecticut, with these types of county fairs held throughout the state at certain times of the year (summer and fall), is a candidate for researching the role they play in STEC transmission risk to humans.

Methods

STEC Case Identification, Study Period and Data Sources

This analysis taps into data from the Foodborne Diseases Active Surveillance Network (FoodNet). Specifically, all incident reportable STEC infection cases in Connecticut in the time period 2012–2014. *Escherichia coli* O157:H7 gastroenteritis and any Shiga toxin–related disease are reportable by physicians and laboratories in CT (CT Department of Public Health). All isolated *E. coli* O157:H7 and non-O157 serotypes and broths from positive Shiga toxin test

results are sent to the Connecticut Department of Public Health (DPH) Laboratory for confirmation, isolation (Shiga toxin–positive broths), and serotyping. Demographic information is registered in all case report forms including street address, age, and sex of the case-patients. Apart from basic demographic information, epidemiologists attempt to interview all case-patients to obtain additional demographics, such as race, and identify and record possible exposures in the seven day period prior to the onset of clinical disease signs.

Geocoding and Spatial Analysis

A total of 180 incident STEC cases were reported in CT from 2012 to 2014. Of the 180 cases, 173 were geocoded as per the case's address using the Environmental Systems Research Institute's (ESRI) ArcGIS software. The automatic settings for geocoding were employed where possible, with manual geocoding being the fall-back option. The process included accessing the original case report forms and verifying/correcting number and spelling errors with Google Maps and USPS.com. The reference network used to geocode case addresses was TIGER (Topologically Integrated Geographic Encoding and Referencing) shape files from the United State Census Bureau and the North American Address Locator (ArcGIS 10 style) from ESRI.17. All shape files in this analysis were projected in ArcGIS using the North American Datum of 1983 (NAD 1983) projection. In this manner, ninety seven percent of all cases were successfully geocoded. Reasons for which three percent failed to be geocoded include the presence PO Box addresses as well as addresses as well as those that did not register in address locators. Additional problems were encountered with incomplete addresses, which were also impediments to geocoding. Once geocoding was complete, data were joined to their corresponding census tract using ArcGIS.

United States Census Data

In order to geocode and classify incident case patients to SES level, data from the US Census was accessed which contained household and population level information from census tracts. The poverty status of a census tract was recorded as the percentage of the tract's total population with family income below the federally defined poverty level. The 2010 American Community Survey (ACS) was used to determine census tract poverty level for the time period of interest, 2012-14 (uscensus.gov). Neighborhood poverty level data was operationalized by constructing four categories based on the percentage of residents in the census tract living below the poverty line: <5%, 5%–9.9%, 10%–19.9%, and >20% (Krieger et al 2005). To relate STEC disease incidence to SES, cases were coded according to the SES category and linked to their geocoded census tract.

Sources of Environmental Data: County Fairs in Connecticut

County and agricultural fairs that took place in Connecticut during the time period were identified by an online search. Each fair location was geocoded and linked to a census tract using procedures similar to those described above. In order to determine whether a census tract was in proximity to a county fair, a five mile buffer area for each fair was created. With a spatial join in ArcGIS, census tracts that were within the boundaries of the buffer zone were considered to be in proximity to the fair and coded as "yes" for data analysis (Fig 3). Five miles was chosen as the buffer zone based on an educated guess as to a reasonable distance an individual would travel to attend a county fair in their neighborhood. Cases within the fair buffer zone were further coded to determine whether the occurrence of clinical disease signs occurred within seven days of the county fair being operational.

Data and Statistical Analysis

Descriptive statistics were generated for the sample. Crude incidence rates were calculated for all cases (included non-geocoded) across age categories, sex, and race/ethnicity. Statistical analysis was limited to cases whose residential address was successfully geocoded in ArcGIS and linked to a specific census tract. Demographic features of geocoded cases were done by calculating counts and percentages. The probability that the counts stayed consistent throughout the study period was calculated with the Jonckheere-Terpesta test for trends which is similar to and agrees with the more familiar chi-square statistic for trends when there are only two comparison groups, but is the superior test when there are more than two categories. Incidence rates across the three years were also calculated and compared with probabilities from the main linear effect of a generalized model with a Poisson distribution.

For the portion of the study that involves investigating the trend and associations between SES level and risk of STEC infection, all tracts were assigned a percentage of the population living under poverty category (described above), and crude and age/sex standardized incident rates were calculated within four aggregated race categories (Hispanic, non-Hispanic white, non-Hispanic black and Other) and in tracts that are both in proximity and not in proximity to county fairs as defined by the five mile buffer area. The rates were also compared across O antigen with the probability of a chi square trend test to test for significance. Generalized linear models were constructed to see if there was an association between STEC risk with neighborhood poverty level after adjusting for age, sex, race/ethnicity and county fair proximity. A similar linear model was developed to test for an association of STEC risk and fair proximity while adjusting for age, sex, race/ethnicity and neighborhood poverty level.

Finally, STEC incidence rates were compared in census tracts with proximity to a fair and census tracts not in proximity to a fair. High risk cases with onset date that fell within the ten day period of risk (3 day fair activity and 7 day STEC incubation period) were isolated to see if people living in those census tracts during the time of a county fair were at higher risk of STEC disease than people in fair-proximate census tracts during times when the fair is shut down. Also compared were incidence rates tracts not proximate to fairs during and outside of the 94 day period during summer months when high fair activity exists in the state of CT.

Rates found in this study were routinely compared to those found in the earlier CT study by Whitney (2015), which covered the period 2000 to 2011. All analyses were done using SAS v9.4 (SAS Institute Inc., Cary, NC, USA) or EpiCalc. Statistical test probabilities of < 0.05 were taken as significant.

Results

Sample Characteristics and Crude Incidence Rates

From 2012 to 2014, 180 incident cases of Shiga toxin-positive specimens with culture-confirmed *E. coli* were reported to the CT Department of Public Health and FoodNET. Of these, 173 (96.1%) were matched on street address, ZIP/postal code, and city/state/province, linked to census tract and then mapped (Figure 1). All cases that did not automatically matched in GIS were manually checked and matched to a census tract interactively in the GIS software. Seven cases were discarded because of they were either listed as a P.O. Box or the address could not be found with any address locator.

Table 1 shows the demographic features of the 173 geocoded STEC cases in the three year study period. The number of cases did not vary significantly by year ($p=0.34$). The mean age of cases

ranged between 25.5 years and 31.9 years with no significant difference across the three years. The percentage of cases that were male and female did vary across the three years with the percentage females increasing from 2012 to 2014 (p =0.044). Race/ethnicity, distribution of cases in poverty level categories, and those in census tracts with proximity to a county fair also did not vary across the three-year period. Figure 4 shows graphically the crude incidence of STEC by neighborhood level poverty across socioeconomic status.

Crude incidence rates and incident rate ratios (IRRs) by age group, sex and race/ethnicity of all 180 incident cases are depicted in Table 2, which also compares the distribution of cases by these features for O157 versus non-O157 STEC. Overall, there was a progressive increase in IR with increasing age, females had 1.4-fold higher rates than males, and non-Hispanic whites had both the most cases (83%) and an IR that was six-fold higher than non-Hispanic blacks. The distribution of cases did not differ significantly by O-antigen with respect to age, sex or race/ethnicity. The overall STEC incidence rate per 100,000 by year was variable, with no specific trend: 1.35 for 2012, 1.94 for 2013 and 1.58 for 2014.

Incidence of STEC in CT and Neighborhood-level SES: 2012-2014

The largest group of STEC cases defined by poverty was the least impoverished category (0-4.99% living below the federal poverty level) accounting for 41.0% of the total. The second most frequent category among cases was the second least impoverished (5-9.99%), accounting for 35.3 % of the cases, followed by third least impoverished group (10-19.99%) at 16.2% of the cases. Graphically the distribution of the crude incidence by poverty level status is shown in Figure 4. The smallest group was the poorest (≥20%) accounting for only 7.5% (Table 2). Neighborhood poverty level was found to be significantly associated with STEC incidence during the researched time period (Table 4). For the combined STEC (both serogroups) the

highest incidence was found in the group with neighborhood level poverty of 5-9.99% (Crude IR= 2.12 per 100,000 person years) and the lowest for the most impoverish (crude OR=0.73). After standardizing for age and sex, the same trend was seen with a slightly larger difference between the highest and lowest incidence groups (2.18 per 100 000 person years versus 0.69). Notably, the difference in the rates between the 0-4.99% and 5-9.99% poverty categories was not statistically significant.

When examined by serogroup, there was a slight difference. For the O157 serotype, the highest crude and standardized incident rate was observed in the 0-4.99% poverty group (Standardized IR=0.80 per 100,000 person years). For non O157, similar to all STEC, the highest incidence occurred in the 5-9.99% poverty group with a standardized IR of 1.39 per 100 000 person years. For both serogroups, the most impoverished (≥20%) had the lowest rates.
Age and sex standardized rates for all STEC infections and each serogroup revealed a trend of decreasing poverty levels with higher STEC incidence ($p<0.001$ for each, Table 3, Figure 4). Using a generalized linear model adjusted for age group, sex, race/ethnicity and proximity to fair, this significant association between neighborhood poverty level and STEC risk was also found between the lowest and the highest SES group ($p=0.002$) (Table 5).

Trends in Race/Ethnicity, Socioeconomic Status and STEC

To probe whether there is a trend in STEC risk by SES status within race/ethnic groups, we tested the trend among white non-Hispanic and other, which was all other race ethnicities combined due to low counts in some of the race/ethnicity categories. Although the highest risk was in the 5-9.99% poverty level group in both whites and other races, there were statistically significant trends in each group for higher incidence with higher SES ($p=0.011$ and $p=0.045$, respectively, chi-square for trend). (Figure 5.)

Trends in Sex, Socioeconomic Status and STEC

To probe whether there was a trend in STEC risk by SES status by sex, we tested the trend among females and males. As seen previously, while the highest STEC risk was in the 5-9.99% poverty level group in both females and males, there tended to be trends in each group for higher incidence with higher SES ($p=0.015$ for females and $p=0.089$ for males, respectively, chi-square for trend) (Figure 6).

Trends in Age, Socioeconomic Status and STEC

To probe whether there was a trend in STEC risk by SES status among each age group, we tested three predefined age categories; 0-4 years, 5-17 years, and >17 years. For age categories 0-4 years and 5-17 years there was no significant association between STEC risk and SES ($p=0.87$ and $p= 0.19$ respectively). For those 18 and older, however, the highest STEC risk was in the 5-9.99% poverty level group, there was a significant trend was registered for higher incidence with higher SES ($p=0.002$, chi-square for trend). (Figure 7.)

Trends in STEC Incidence by Time Periods

To determine whether there was a difference in overall STEC incidence over the years, data from this study was compared to rates calculated by Whitney (2015) who covered the time periods of 2000-2006 and 2007-2011. Figure 8 shows the standardized rates by time period in CT across the four neighborhood level poverty categories. The trend of increasing SES and higher STEC incidence seems to hold true across the three time periods, however there is an increase in incidence in the 5-9.99% category as time progressed in the most recent period, 2012-2014 (Figure 8). The same comparison made across O antigens showed that for the O157 serotype the trend of lower STEC incidence as poverty levels increases was seen for both time

periods 2000-2011 and 2012-104 (Figure 9). Finally, for the non O157 serotypes there was an increase in incidence in both the second and third wealthiest census tracts over time (Figure 10).

County Fairs and STEC

To investigate whether there was a risk of STEC determined by being geographically located in an area of the state with proximity to a county fair (a presumably more agricultural area of the state), the year round incidence was calculated in census tracts that were close to a fair and those that were not. The incidence rate was higher in the sum of census tracts with a fair in proximity (1.79 per 100 000 person years vs 1.42 per 100 000 person years, IRR 1.26), but the difference was not statistically significant (p=0.17) (Table. 6).

Cases were then identified that may have been exposed through a county fair by determining if their date of onset coincided with the ten day fair activity time (three-day fair activity plus the seven day incubation period of STEC) of the fair whose buffer zone fell within overlapped with the case's census tract. Only six such cases were identified and the incidence was calculated for the ten day period at 3.68 per 100 000 person-years. To determine if this was higher than expected for the 94 day period in which all county fairs fell (July 10 – Oct 11 each year), this rate was compared to that of persons living in proximity to county fairs during the time when a fair was not in operation, to those living in census tracts with no proximity to county fairs, and to the sum of the two. The relative risk among these 6 qualifying cases was at least double of those of living in the same census tracts when a county fair was not happening, and more than triple of those living in non-county fair census tracts (Table 6). However, the former risk did not reach statistical significance (p=0.10) while the latter did (p=0.006).

The multivariate analysis of the 6 cases in the ten day activity period showed that those cases where 3.53 times more likely of STEC risk when compared to those cases in the 94 day period

with no fairs in proximity (p=0.008). Those six cases were also 2.62 times more likely to be at risk for STEC when compared to cases statewide (p= 0.027) (Table 7).

In order to see if there was an association between living close to a county fair and STEC risk, a general linear model was done while adjusting for the effects of sex, age, race/ethnicity and neighborhood poverty level. There was no association found between the risk of STEC and being in proximity to fair for all serotypes of STEC, O157 or non O157 (p=0.281, 0.801, 0.241 respectively) (Table 8).

Discussion

This analysis was undertaken to probe whether area-based socioeconomic status and geographic proximity to a potential point source for infection (farm animal displaying county fairs) amplified the risk of Shiga toxin producing *E. coli* (STEC) infection and disease in Connecticut. Neighborhood level poverty was found to be significantly associated with STEC incidence, but not in the way normally encountered in public health investigations. Age- and sex-standardized rates showed that there was a significant difference between the high and low poverty-level neighborhoods, but in this case, low risk for STEC infection was associated with higher poverty levels. During the time period addressed by this study (2012-2014), the incidence of diagnosed STEC was highest the SES category representing the second most affluent census tracts (5-9.99% of people living below the national poverty line). This was somewhat different from the results of Whitney et al. (2015), whose study encompassed two earlier time periods (2000-2006 and 2007-2011) and found highest STEC incidence in the census tracts classified in the wealthiest SES category (0-4.99% under the poverty line). Nonetheless, this study mirrors previous data in demonstrating that risk for STEC infection and disease is low where poverty is high and high where poverty is low. A positive correlation between infectious disease risk and

SES is not commonly seen in public health, where low SES typically contributes to a plethora of health risks, both for infectious and non-infectious diseases. Further research is necessary to completely understand why an affluent lifestyle leads to higher STEC risk, but it is clear that STEC risk factors are not uniformly distributed across the population in a geographic area. For example, the consumption of high STEC-risk foods such as undercooked meat, unpasteurized milk and ciders is more prevalent among persons of higher SES (Thorpe 2004). Conversely, the consumption of cheaper, processed food, which may not be preferable from a nutritional standpoint certainly provides very low-risk for STEC infection. International travel, also identified as an STEC risk factor, is much more common among higher SES residents (Hadler 2011). There are likely other factors that can be identified which contribute to concentrating risk in higher SES sectors and further investigations in this area appear warranted.

While suggestive, the relationship between STEC risk and proximity to county fairs is not well elucidated in this study. Although a statistically significant higher risk was registered in persons living in proximity to county fairs during the 3 days during and 7 days following fairs being held, there was no significant difference in STEC risk in census tracts that included a fair and those that didn't neither during the entire year nor for the 94 day period during which fairs were being held in one part of the state or another. The analysis suggests that there is a risk inherent to county fairs in that 94-day period of high fair activity, but the magnitude and socio-geographic distribution of the risk remains obscure. Because of a small sample size and (only 6 county fair high risk individuals were identified in the 3 year period), our statistical power is limited, but further examination of county fairs as a point source of STEC infection seems worthy of further investigation, perhaps as a part of multi-state study. General linear models

examining these relationships after adjustment for sex, age and poverty level were consistent with the univariate findings.

To test whether county fair presence is a point source for STEC infection in itself or simply another marker for rurality, we compared the rurality index of census tracts with proximity to a county fair to those without fair proximity. Rurality data was taken from American Community Survey (ACS) and is defined as the percentage of housing units considered rural in census tracts. It is worth mentioning that the definition of rural in CT is somewhat vague, as there is no detailed criteria identified by the ACS in classifying a housing unit as rural in their assessment. The mean rurality percentage of census tracts classified as "yes" to a county fair proximity was 23.34% and those without a county fair was 10.10% ($p=0.004$). What these data show is that, not surprisingly, county fairs tend to be held in more rural census tracts. It also provides evidence to suggest that county fairs are indeed a general marker for rurality in CT.

While our analysis failed to provide strong evidence implicating farm animal displaying county fairs as a significant point source for STEC infection, our interpretation may simply be a result of an analysis which does not accurately capture the movements of people in and around the popular events. Let us start with our questionable criteria behind classifying a census tract as proximate to a county fair. First, we somewhat arbitrarily chose the five mile radius around the fair to classify census tracts as proximate of not proximate. There is also an inherent assumption that individuals closer to the fair would have a higher probability of visiting the event and, perhaps, be more likely to make multiple visits. However, it is highly plausible that the motivation of individuals to visit and, more importantly, to spend extended periods at the fair goes beyond geographic proximity. Data on county fair seeking behavior in CT is absent, and it

may well be that our assumptions, based largely on previous personal experiences and not actual data, are off base. It should be noted that if a buffer larger than five miles was established, almost all of the census tracts in CT would have been classified as proximate, rendering the analysis infeasible. Also, when applying multiple selection criteria to an already small sample size of cases severely limits statistical power. This study, limited to a rather short time period, has a total sample size of 180 incident cases. Of these cases, only six of represented individuals that lived in proximity to fairs and had their infection occur when the fair was in operation. Thus the analysis is severely limited by the problem of small sample size.

Although STEC is reportable disease, it is still an infection that is almost certainly under detected (Scallan et al 2011). Therefore, estimations of risk and prevalence may not be entirely accurate if it is assumed that undiagnosed cases have similar onset patterns and demographic features as diagnosed cases. Only about 20% of persons with acute diarrhea seek medical attention, and not all laboratories routinely test for all STEC types (Scallan 2008). In addition, while all laboratories test for O157, only a limited number do Shiga-toxin testing, a prerequisite for identifying non-O157 STEC.

As shown by the data another limitation is the lack of heterogeneity of cases in regards to race and ethnicity. A large proportion of the Connecticut population is white, and the majority of STEC cases are also identified as white. Minority cases are invariably saddled with the problem of small sample size in statistical analyses. With a larger sample encompassing more than the three years of this study, with more cases in other race/ethnicity categories, a more accurate analysis on socioeconomic status, proximity to fair and race/ethnicity in regards to STEC incidence would be possible. Finally, poverty in census tracts is an area-based measure and does not directly measure individual income as a risk factor for STEC. Data taken from case

interviews remains critical in understanding the factors contributing to STEC incidence as it relates to socioeconomic status or to exposure to point source infection risk.

In conclusion, this study supports previous analyses of the relationship between SES and STEC incidence in CT; efforts to reduce STEC risk and disease burden need to focus more on higher and middle income SES groups. When examining whether county fairs pose a risk of exposure to STEC from a geographic perspective, the evidence suggested that the fair may indeed pose a risk but the evidence was far from definitive. Further analysis is warranted in this area which includes references to fair-seeking behavior in CT or in other, similar states.

Figures and Tables

Table 1. Demographic Features of Geocoded STEC Cases in CT (2012-2014)

Characteristic	Year 2012	2013	2014	p-value[a,b]	Total	CT (2010 Census)
Cases	48	69	56		173	3,554,876
Age in Years for All Cases, mean (SD)	28.8 (25.0)	31.9 (25.7)	25.5 (20.9)	0.340		
Age category years, n (%)				0.949		
1(0-4)	7 (14.6)	10 (14.5)	10 (17.9)		27 (15.6)	202,992
2(5-17)	15 (31.3)	16 (23.2)	14 (25.0)		45 (26.0)	623,033
3(18+)	26 (54.2)	43 (62.3)	32 (57.1)		101 (58.4)	2,728,851
Sex, n (%)				0.044		
Female	23 (47.9)	44 (63.8)	38 (67.9)		105 (60.7)	1,825,590
Male	25 (52.1)	25 (36.2)	18 (32.1)		68 (39.3)	1,729,286
Race/Ethnicity, n (%)				0.899		
White non-Hispanic	38 (79.2)	55 (79.7)	44 (78.6)		137 (79.2)	2,795,216
Black non-Hispanic	1 (2.1)	1 (1.4)	3 (5.4)		5 (2.9)	342,042
Hispanic	8 (16.7)	8 (11.6)	7 (12.5)		23 (13.3)	228,840
Other non-Hispanic	1 (2.1)	5 (7.2)	2 (3.6)		8 (4.6)	188,778
Percent Below Poverty, n (%)				0.444		
0-4.9%	20 (41.7)	30 (43.5)	21 (37.5)		71 (41.0)	1,327,918
5-9.9%	19 (39.6)	22 (31.9)	20 (35.7)		61 (35.3)	958,276
10-19.9%	5 (10.4)	14 (20.3)	9 (16.1)		28 (16.2)	673,403
>20 %	4 (8.3)	3 (4.3)	6 (10.7)		13 (7.5)	595,279
Proximity to Fair, n (%)				0.858		
No	19 (39.6)	26 (37.7)	23 (41.1)		68 (39.3)	1,598,239
Yes	29 (60.4)	43 (62.3)	33 (58.9)		105 (60.7)	1,956,637

[a]Probability for age in years is from a linear regression model.
[b]Probabilities for categorical variables are from Jonckheere-Terpstra tests for trends in proportions.

Table 2. Crude Incidence[a] of all STEC by Demographic Features and Comparison by O Antigen STEC (N=180*)

					O157 (N= 68)		Non-O157 (N=112)		
Characteristic	**N[a]**	**Crude IR**	**RR**	**p[b]**	**N[a]**	**%**	**N[a]**	**%**	**P[c]**
Age (years)									
0-4	27	4.45	3.43	$<10^{-7}$	9	13.2	18	16.1	
5 to 17	46	2.49	1.92	0.0001	24	35.3	22	19.6	0.07
>18	107	1.29	ref		35	51.5	72	64.3	
Sex									
Female	107	1.94	1.39	0.03	45	66.17	62	55.35	0.15
Male	73	1.40	ref		23	33.82	50	44.64	
Race/ethnicity									
Hispanic	23	1.60	0.85	0.47	5	7.35	18	16.07	
Non-Hispanic Black	5	0.49	0.26	0.002	1	1.47	4	3.57	0.30
Other	2	0.31	0.16	0.004	1	1.47	1	0.89	
Non-Hispanic White	144	1.88	ref		58	85.29	86	76.78	

a Incidence per 100 000 person years

b p value is for Chi-square

c p value is for Chi-square comparing distribution of O157 to non O157

Table 3. Crude Incidence Rates[a] of STEC by O Antigen by Year: Connecticut 2012-2014

Characteristic	Year 2012	2013	2014	p-value[b]
Population[c]	3554876	3554876	3554876	
All STEC				
Cases	48	69	56	
All STEC Crude IR (95% CI)	1.35 (1.02, 1.79)	1.94 (1.53, 2.46)	1.58 (1.21, 2.05)	<.001
O157				
Cases	18	32	17	
O157 Crude IR (95% CI)	0.51 (0.32, 0.80)	0.90 (0.64, 1.27)	0.48 (0.30, 0.77)	<.001
Non-O157				
Cases	30	37	39	
Non-O157 Crude IR (95% CI)	0.84 (0.59, 1.21)	1.04 (0.75, 1.44)	1.10 (0.80, 1.50)	0.730

[a]Rates and 95% confidence limits (CL) are in counts/100,000 person-years.
[b]Probabilities are for a main (linear) effect from a generalized linear model with a Poisson distribution.
[c]Population taken from 2010 US Census

Table 4. Incidence Rates[a] and Ratios by Neighborhood Status Poverty Level

	Neighborhood Poverty Level (% below poverty line)				
Characteristic	**0 - 4.99%**	**5.0 - 9.99%**	**10.0 - 19.99%**	**≥20%**	**p-value[b]**
Total Population	1327918	958276	673403	595279	
All STEC					
Number of Cases	71	61	28	13	
Crude IR	1.78	2.12	1.39	0.73	
Standardized IR (95% CI)	1.79 (1.37, 2.21)	2.18 (1.63, 2.73)	1.42 (0.89, 1.95)	0.69 (0.32, 1.07)	<.0001
Standardized IRR	1.00 ref	1.22	0.80	0.39	
O157 STEC					
Number of Cases	33	22	10	2	
Crude IR	0.83	0.77	0.49	0.11	
Standardized IR (95% CI)	0.80 (0.53, 1.08)	0.79 (0.46, 1.13)	0.52 (0.20, 0.84)	0.11 (0.00, 0.25)	<.0001
Standardized IRR	1.00 ref	0.99	0.65	0.13	
Non-O157 STEC					
Number of Cases	38	39	18	11	
Crude IR	0.95	1.36	0.89	0.62	
Standardized IR (95% CI)	0.99 (0.67, 1.30)	1.39 (0.95, 1.83)	0.91 (0.49, 1.32)	0.59 (0.24, 0.94)	<.0001
Standardized IRR	1.00 ref	1.41	0.92	0.60	

[a]Rates and 95% confidence limits (CL) are in counts/100,000 person-years. Standardized values used age group and sex values from the 2010 Census.
[b]Probabilities are for a Chi-square trend test for age-and sex-standardized rates.

Table 5. Model Adjusted STEC Incidence Rates[a] and Ratios by Neighborhood Poverty Level: CT 2012-2014

	Neighborhood Poverty Level (% below poverty line)			
Characteristic	**0 - 4.99%**	**5.0 - 9.99%**	**10.0 - 19.99%**	**≥20%**
Total Population	1327918	958276	673403	595279
All STEC				
Number of Cases	71	61	28	13
Model-adjusted IR (95% CI)	2.60 (1.99, 3.41)	3.13 (2.23, 4.39)	1.99 (1.31, 3.04)	0.98 (0.56, 1.71)
Adjusted IRR	1.00 (ref)	0.83	1.31	2.66
IRR Probability		0.308	0.242	0.002
O157 STEC				
Number of Cases	33	22	10	2
Model-adjusted IR (95% CI)	1.06 (0.69, 1.63)	1.02 (0.60, 1.72)	0.67 (0.28, 1.62)	0.15 (0.04, 0.61)
Adjusted IRR	1.00 (ref)	1.04	1.57	7.06
IRR Probability		0.890	0.258	0.009
Non-O157 STEC				
Number of Cases	38	39	18	11
Model-adjusted IR (95% CI)	1.44 (1.00, 2.09)	2.03 (1.39, 2.97)	1.27 (0.77, 2.10)	0.80 (0.43, 1.51)
Adjusted IRR	1.00 (ref)	0.71	1.14	1.80
IRR Probability		0.125	0.670	0.108

[a]Rates and 95% confidence limits (CL) are in counts/100,000 person-years. Rates (95% CI), rate ratios, and probabilities from generalized linear models with a Poisson distribution and adjusted for age group, sex, race-ethnicity, and proximity to fair (Yes/No).
[b]Probabilities are for post hoc contrasts with the reference group: (Poverty: 0 - 4.99%).

Table 6. Incidence Rates* in Census Tracts with County Fairs and Fair Activity Time Period*

Incidence Rates Year Round in Census Tract Categories Defined by Fair Proximity	Cases	Population (person-years)	Crude IR	RR (95% CI)	p[a] value
Census Tract with Fairs in Proximity	105	~5,942,275	1.77	1.24 (0.91,1.68)	0.17
Census Tract with No Fairs in Proximity	68	~4,755,245	1.43	ref	
Incidence Rates in Census Tracts with Fair Proximity or Exposure during 94 days of Fair Activity					
Census tracts with fairs in proximity – onset dates during the 10 days*after beginning of a fair	6	~162,921	3.68	**	
Census tracts with fairs in proximity – onset unrelated to fair timing	27	~1,508,380	1.79	ref gp1	0.10
Census tracts with no fairs in proximity	16	~1,221,374	1.30(0.79,2	ref gp2	0.006
Statewide risk for STEC during 94 days	43	~2,752,809	1.48	ref gp3	0.04

*per 100,000

* Time period of all Fair Activity in CT spans a 94 day period in July until October.

* Ten day risk period is the average 3 day period of a specific fair plus the seven day incubation period for STEC

[a]P value is for Chi Square

** RR (95%CI) for persons living in census tracts in proximity to a county fair and having onset during the 10 days after the beginning of a fair: Reference gp1 2.06 (0.85,4.98); Ref gp2 3.47 (1.36,8.87); Ref gp3 2.36 (1.004,5.54)

Table 7. STEC: Incidence Rates[a] in Census Tracts with County Fairs and Fair Activity Time, CT 2012-2014

Fair Activity	Cases	Adjusted IR (95% CI)	RR	p-value[b]
Incidence Rates in Census Tracts defined by Fair Proximity (Year-round)				
Census Tracts with Fairs in Proxmity	105	2.36 (1.85, 3.03)	1.11	0.495
Census Tracts with No Fairs in Proximity	68	2.12 (1.62, 2.77)		
Incidence Rates in Census Tracts with Fair Proximity or Exposure during 94 days of Fair Activity				
Onset date proximate to Fair Activity (10-day period)\	6	2.31 (1.00, 5.30)		
Census Tracts with Proximity to Fair (94-day period)	27	1.10 (0.71, 1.72)	2.09	0.103
Census Tracts with No Proximity to Fair (94-day period)	16	0.65 (0.38, 1.12)	3.53	0.008
Statewide Risk for STEC (94-day period)	43	0.88 (0.60, 1.28)	2.62	0.027

[a]Rates and 95% confidence limits (CL) are in counts/100,000 person-years.

[b]Probabilities are from a generalized linear model with a Poisson distribution and inverse link to place estimates on the original scale.

All models were adjusted for age group, sex, and percent poverty level

Table 8. Proximity to Fairs - Model[a] Adjusted IR and IRR for STEC: CT 2012-2014

Proximity to Fair	Adjusted IR (95% CI)	Adjusted IRR	P-Value[b]
STEC All			
Yes	1.97 (1.40, 2.77)	1.18	0.281
No	1.66 (1.17, 2.35)		
STEC O157			
Yes	0.50 (0.25, 1.01)	1.07	0.801
No	0.47 (0.24, 0.94)		
STEC Non-O157			
Yes	1.32 (0.87, 2.02)	1.27	0.241
No	1.04 (0.67, 1.62)		

[a]Rates and 95% confidence limits (CL) are in counts/100,000 person-years. Rates (95% CI), rate ratios, and probabilities from
generalized linear models with a Poisson distribution and adjusted for age group, sex, race-ethnicity, and percent poverty category
(in census tract).

[b]Probabilities are for Type 3 effects (adjusted for all the other effects in the model).

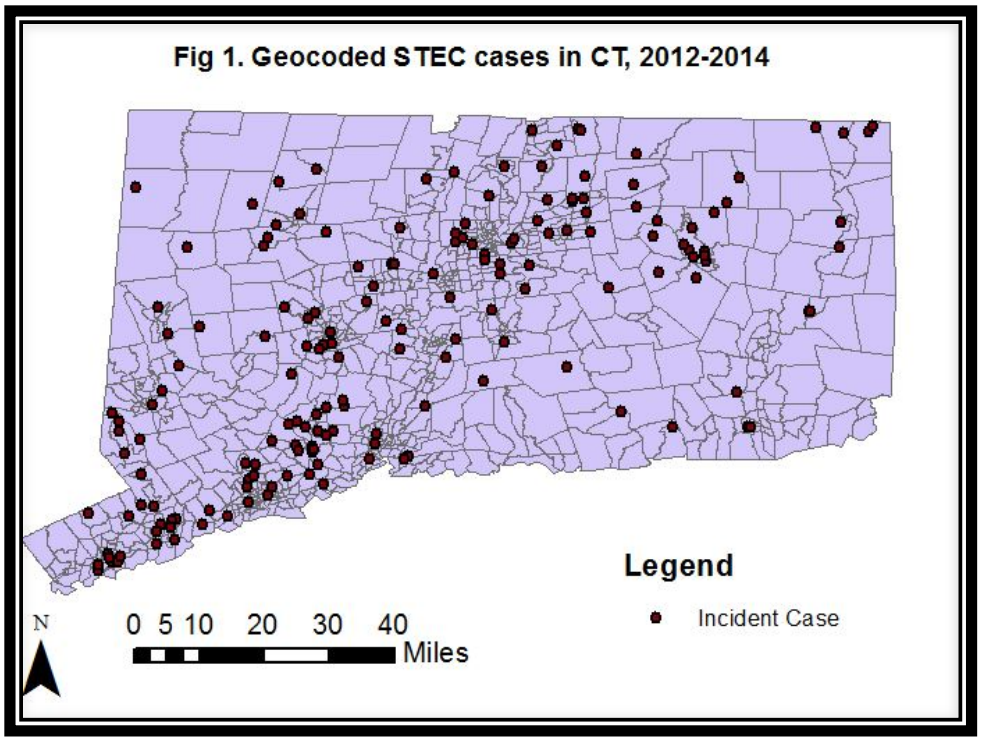


Fig 1. Geocoded STEC cases in CT, 2012-2014
Legend
Incident Case
N
0 5 10 20 30 40
Miles

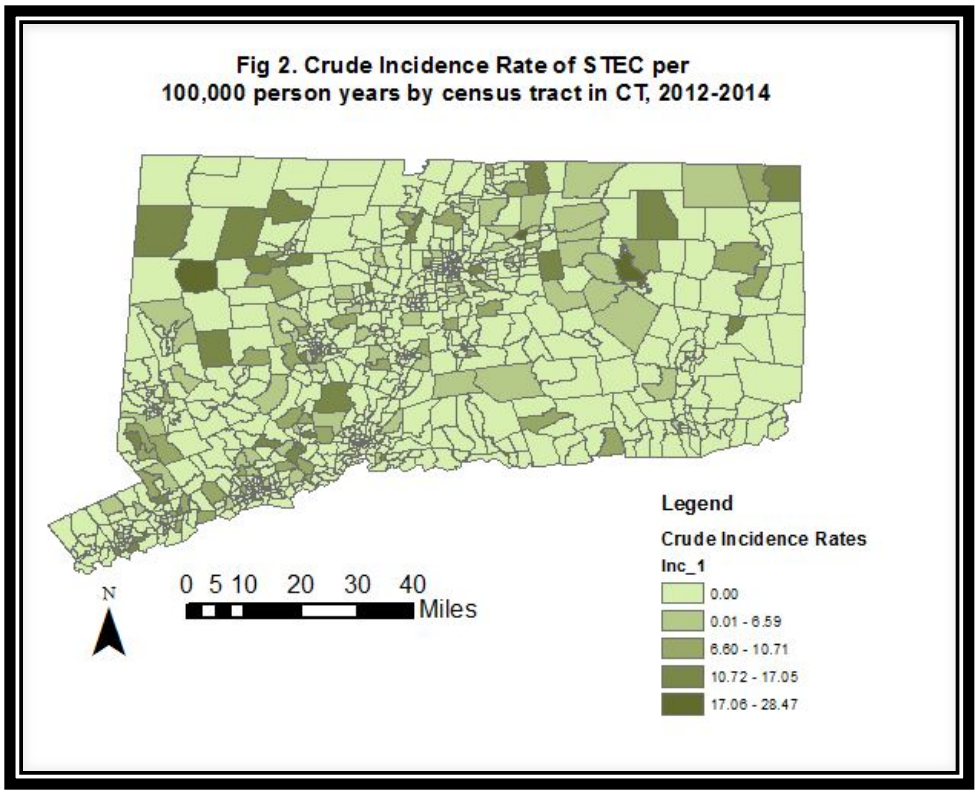


Fig 2. Crude Incidence Rate of STEC per
100,000 person years by census tract in CT, 2012-2014
Legend
Crude Incidence Rates
Inc_1
0.00
0.01 - 6.59
6.60 - 10.71
10.72 - 17.05
17.06 - 28.47
N
0 5 10 20 30 40
Miles

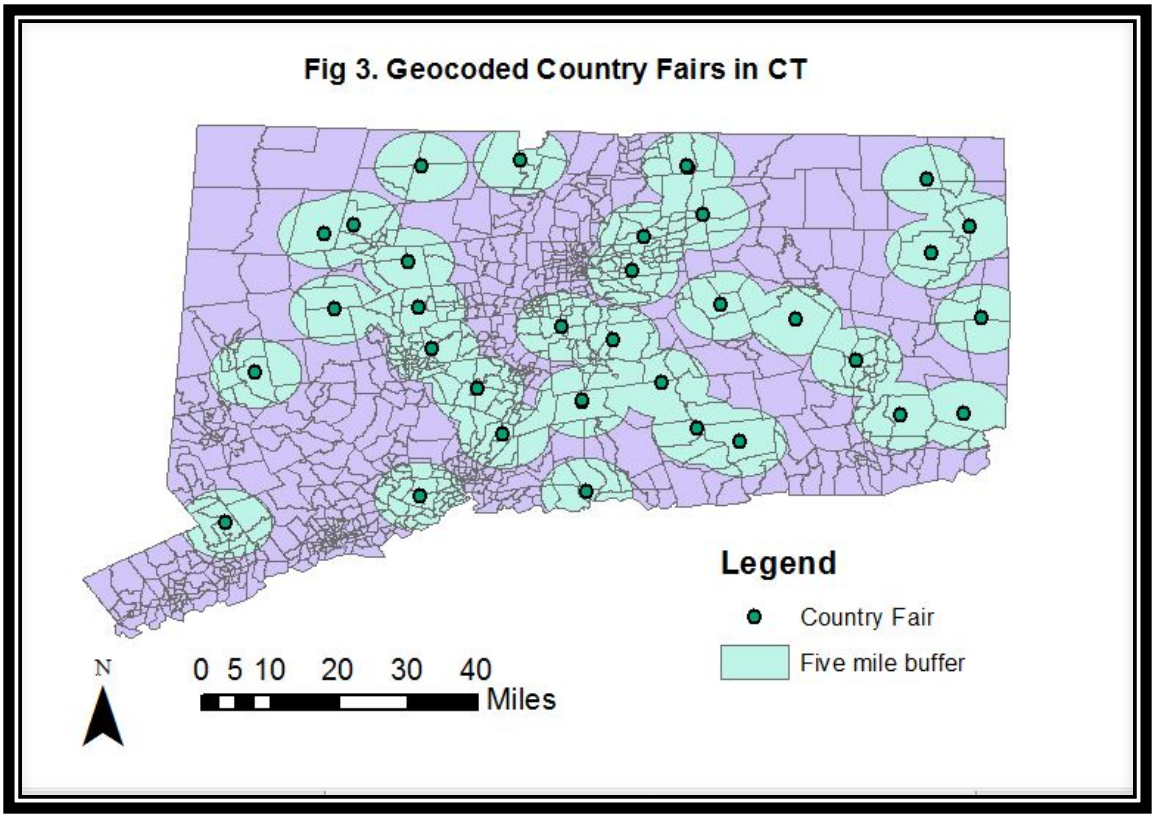


Fig 3. Geocoded Country Fairs in CT
Legend
Country Fair
Five mile buffer
N
0 5 10 20 30 40
Miles

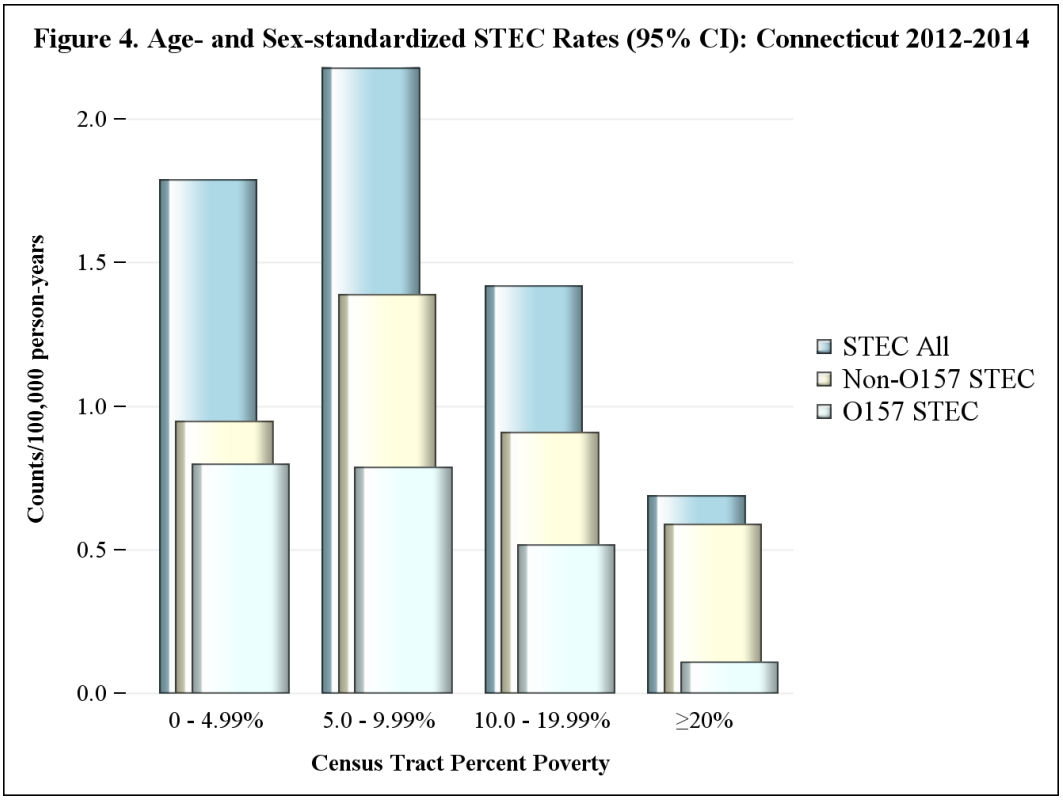


Figure 4. Age- and Sex-standardized STEC Rates (95% CI): Connecticut 2012-2014
Counts/100,000 person-years
2.0
1.5
1.0
0.5
0.0
0 - 4.99%
5.0 - 9.99%
10.0 - 19.99%
≥20%
Census Tract Percent Poverty
STEC All
Non-O157 STEC
O157 STEC

Figure 5. STEC Rates by Percent Poverty Level and Race: Connecticut 2012-2014

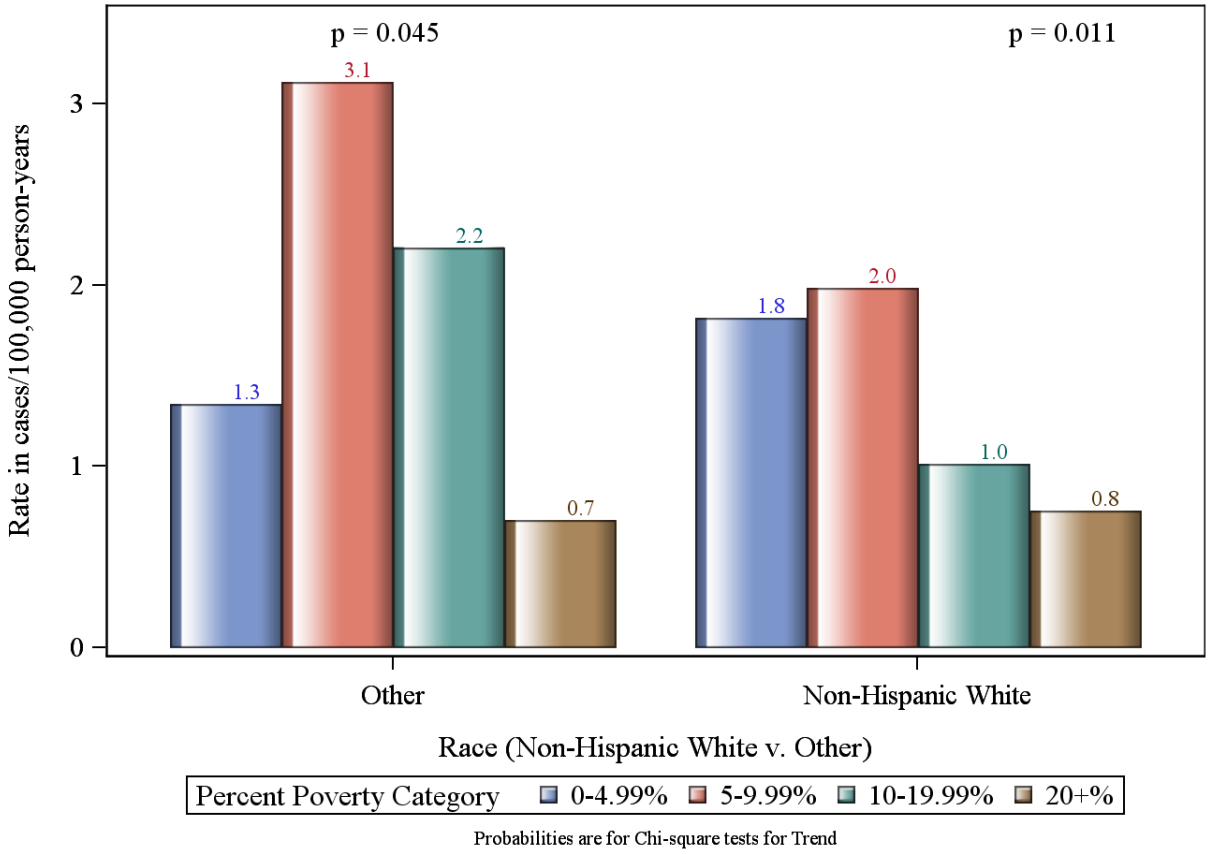


Figure 6. STEC Rates by Percent Poverty Level and Sex: Connecticut 2012-2014

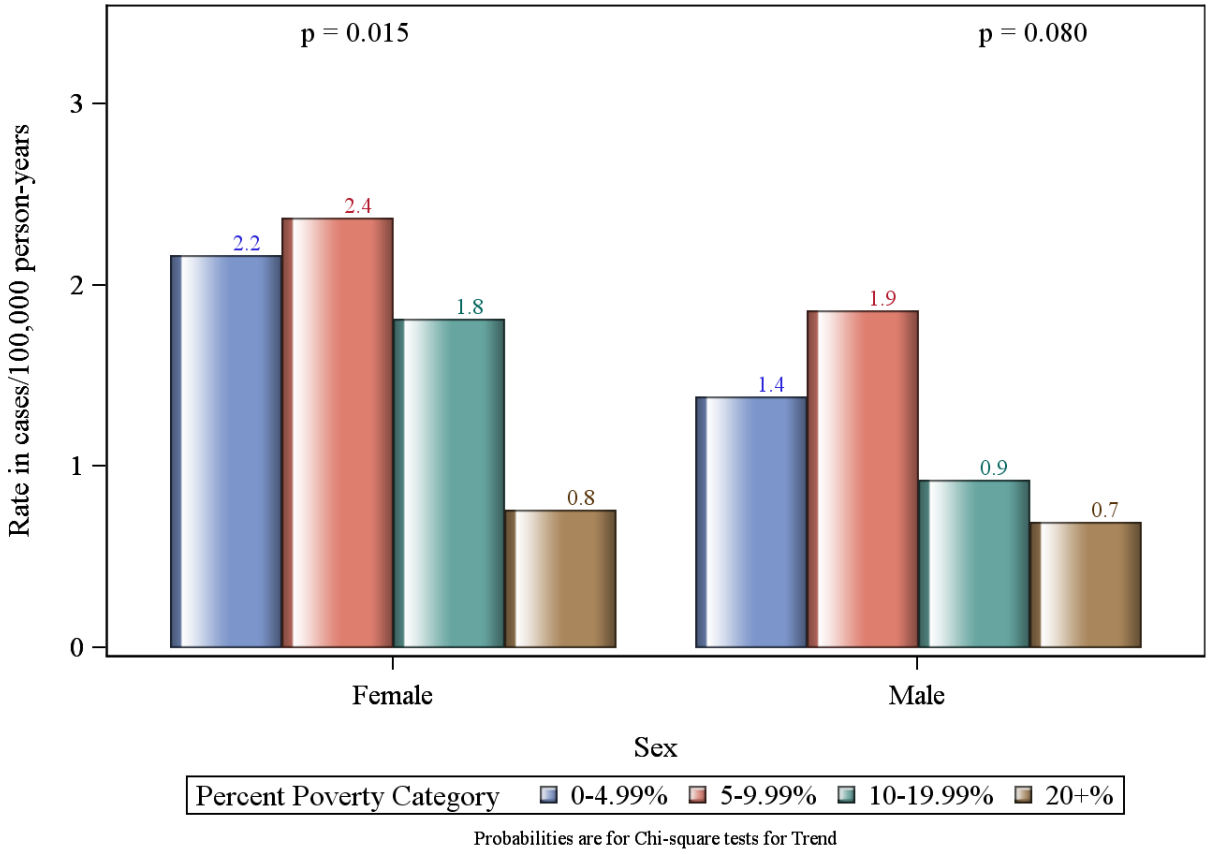
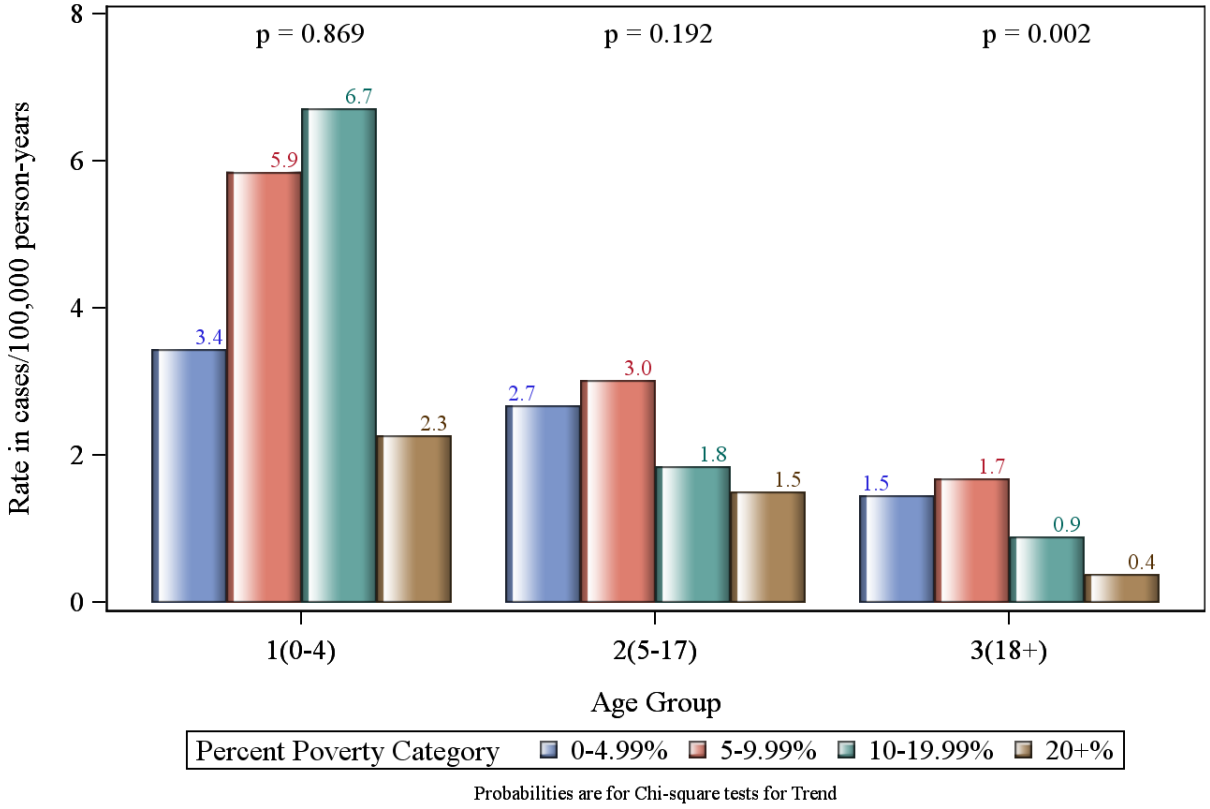


Figure 7. STEC Rates by Percent Poverty Level and Age Group: Connecticut 2012-2014



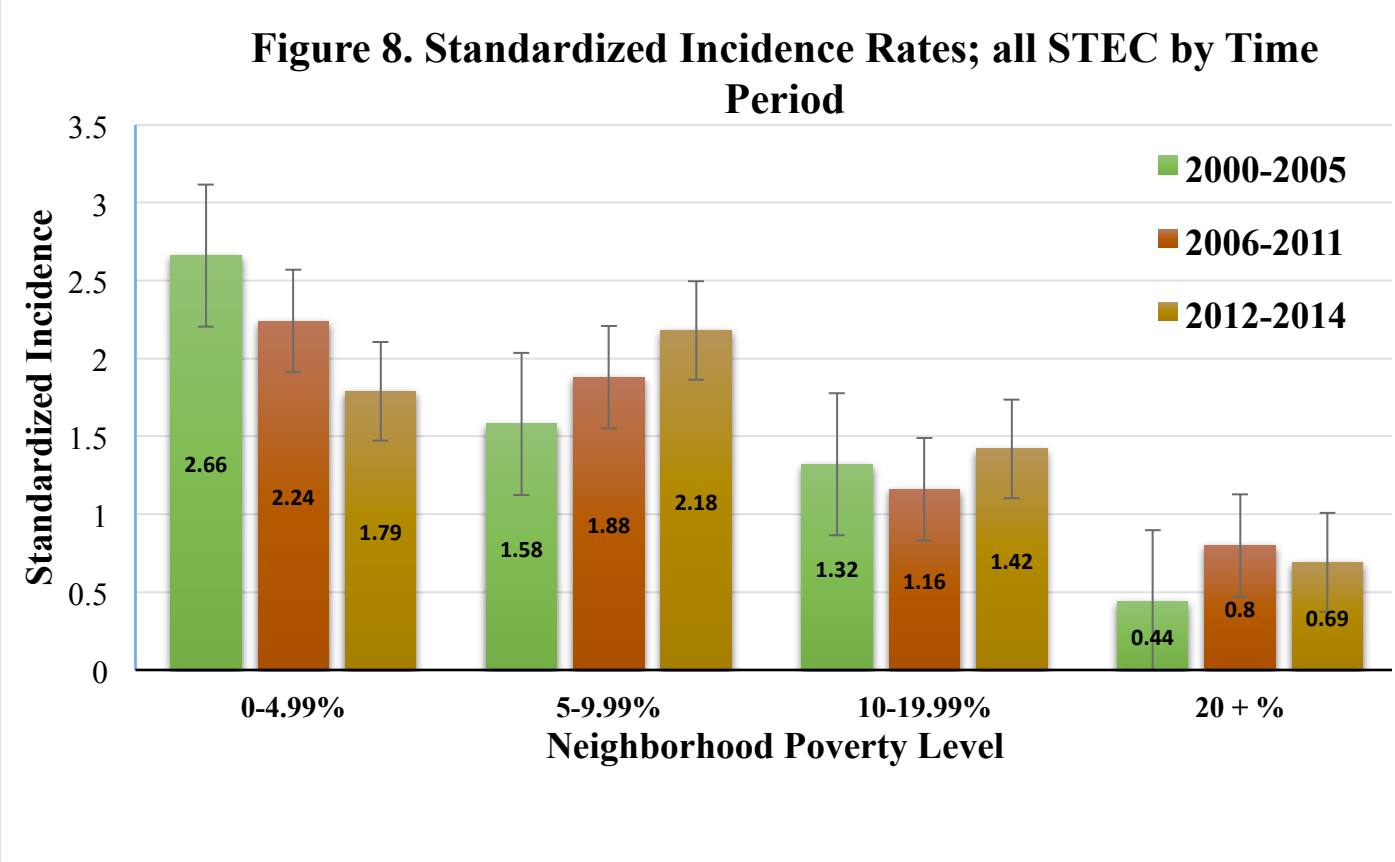


Figure 8. Standardized Incidence Rates; all STEC by Time Period
Standardized Incidence
3.5
3
2.5
2
1.5
1
0.5
0
2000-2005
2006-2011
2012-2014
2.66
2.24
1.79
1.58
1.88
2.18
1.32
1.16
1.42
0.44
0.8
0.69
0-4.99%
5-9.99%
10-19.99%
20 + %
Neighborhood Poverty Level

Figure 9. Standarized Rates O157 Incidence Rates Comparison by Time Period

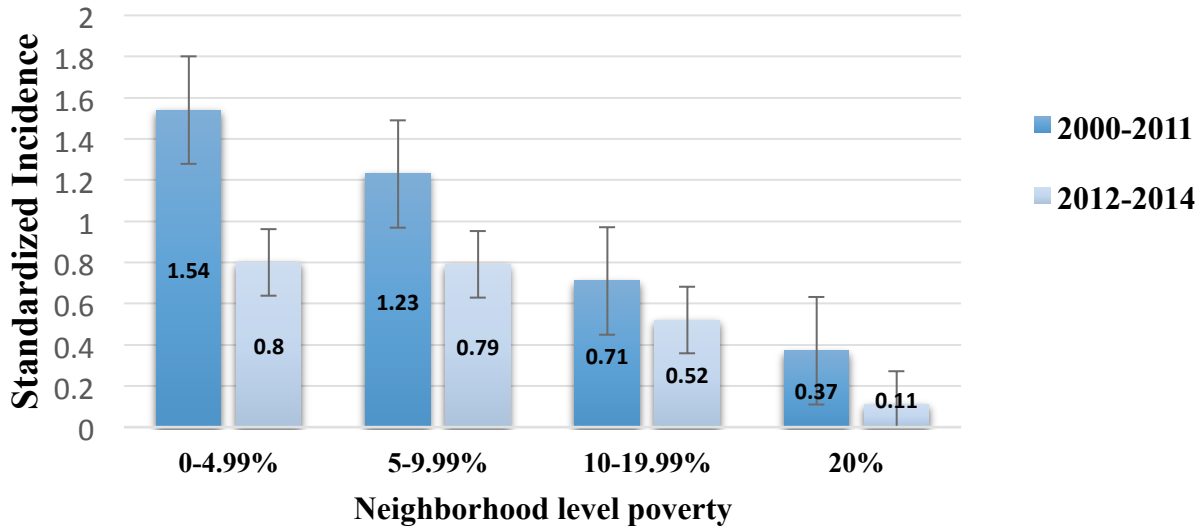
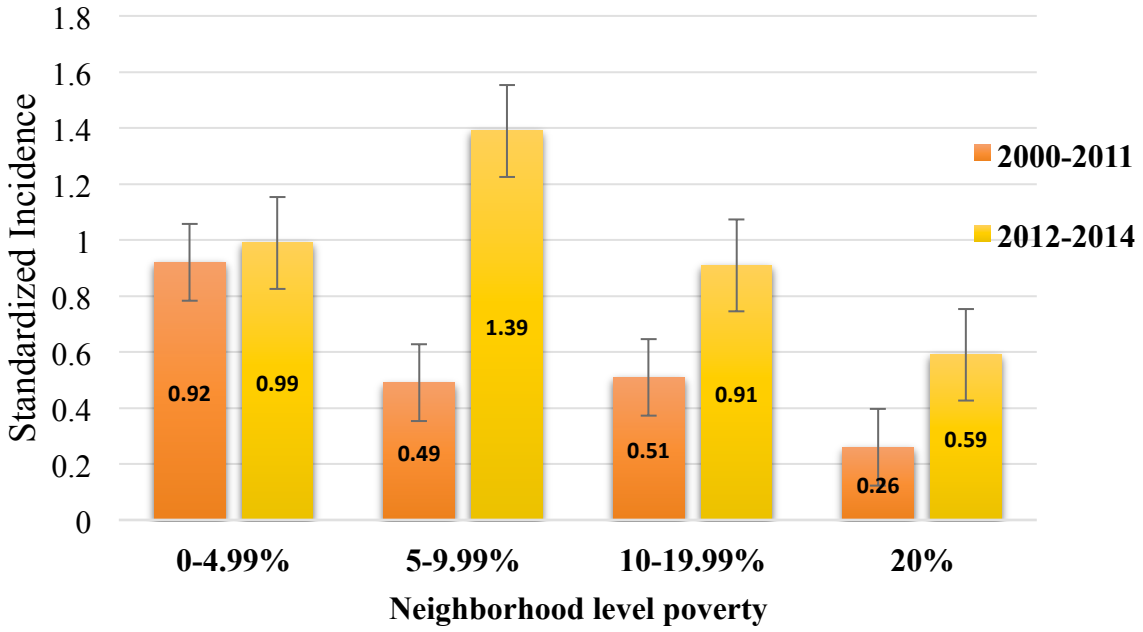


Figure 10. Standarized Incidence Rates non-O157 Comparison by Time Period



REFERENCES

Kistemann, T., Zimmer, S., Vågsholm, I., & Andersson, Y. (2004). GIS-supported investigation of human EHEC and cattle VTEC O157 infections in Sweden: geographical distribution, spatial variation and possible risk factors. *Epidemiology and infection*, *132*(03), 495-50

Greene, S. K., Levin-Rector, A., Hadler, J. L., & Fine, A. D. (2015). Disparities in Reportable Communicable Disease Incidence by Census Tract-Level Poverty, New York City, 2006–2013. *American journal of public health*, *105*(9), e27-e34.

Frank, C., Kapfhammer, S., Werber, D., Stark, K., & Held, L. (2008). Cattle density and Shiga toxin-producing Escherichia coli infection in Germany: increased risk for most but not all serogroups. *Vector-Borne and Zoonotic Diseases*, *8*(5), 635-644.

Friesema, I. H. M., Van De Kassteele, J., De Jager, C. M., Heuvelink, A. E., & Van Pelt, W. (2011). Geographical association between livestock density and human Shiga toxin-producing Escherichia coli O157 infections. *Epidemiology and infection*, *139*(07), 1081-1087.

Michel, P., Wilson, J. B., Martin, S. W., Clarke, R. C., McEwen, S. A., & Gyles, C. L. (1999). Temporal and geographical distributions of reported cases of Escherichia coli O157 [ratio] H7 infection in Ontario. *Epidemiology and Infection*, *122*(02), 193-200.

Innocent, G. T., Mellor, D. J., McEwen, S. A., Reilly, W. J., Smallwood, J., Locking, M. E., ... & Reid, S. W. J. (2005). Spatial and temporal epidemiology of sporadic human cases of Escherichia coli O157 in Scotland, 1996–1999. *Epidemiology and infection*, *133*(06), 1033-1041.

Zappe Pasturel, B., Cruz-Cano, R., Rosenberg Goldstein, R. E., Palmer, A., Blythe, D., Ryan, P., ... & Sapkota, A. R. (2013). Impact of Rurality, Broiler Operations, and Community Socioeconomic Factors on the Risk of Campylobacteriosis in Maryland. *American journal of public health*, *103*(12), 2267-2275.

Whitney, B. M., Mainero, C., Humes, E., Hurd, S., Niccolai, L., & Hadler, J. L. (2015). Socioeconomic Status and Foodborne Pathogens in Connecticut, USA, 2000–2011. *Emerging infectious diseases*, *21*(9), 1617.

Office of Disease Prevention and Health Promotion. Healthy People 2020 https://www.healthypeople.gov

Krieger N, Chen JT, Waterman PD, Rehkopf DH, Subramanian SV. Painting a truer picture of US socioeconomic and racial/ethnic health inequalities: The Public Health Disparities Geocoding Project. Am J Public Health. 2005;95(2):312---323.

Krieger, N., Williams, D. R., & Moss, N. E. (1997). Measuring social class in US public health research: concepts, methodologies, and guidelines. *Annual review of public health*, *18*(1), 341-378.

Simonsen, J., Frisch, M., & Ethelberg, S. (2008). Socioeconomic risk factors for bacterial gastrointestinal infections. *Epidemiology*, *19*(2), 282-290.

Chang, M., Groseclose, S. L., Zaidi, A. A. & Braden, C. R. An ecological analysis of sociodemographic factors associated with the incidence of salmonellosis, shigellosis, and E. coli O157:H7 infections in US counties. *Epidemiology and infection* **137**, 810-820, doi:10.1017/S0950268808001477 (2009).

Johnson KE, Thorpe CM, Sears CL. The emerging clinical importance of non-O157 Shiga toxin-producing Escherichia coli Clin Infect Dis 2006 ; 43:1587–95.CDC. *Foodborne Diseases Active Surveillance Network (FoodNet): Table and Figures-2011 Preliminary Data*, <http://www.cdc.gov/foodnet/data/trends/tables-2011.html> (2012).

Goulet V, Hebert M, Hedberg C, et al. Incidence of listeriosis and related mortality among groups atrisk of acquiring listeriosis. Clin Infect Dis **2012**; 54:652–660.

Melton-Celsa A.R., O'Brien A., Structure,biology, and relative toxicity of Shiga toxin family members for cells and animals, in:Kaper J.B., O'Brien A.D. (Eds.), *Escherichia coli* O157:H7 and other Shiga toxin producing*E. coli* strains, American Society

Hadler, J. L. *et al.* Ten-year trends and risk factors for non-O157 Shiga toxin-producing Escherichia coli found through Shiga toxin testing, Connecticut, 2000-2009. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **53**, 269-276, doi:10.1093/cid/cir377 (2011).

Brooks, J. T. *et al.* Non-O157 Shiga toxin-producing Escherichia coli infections in the United States, 1983-2002. *The Journal of infectious diseases* **192**, 1422-1429, doi:10.1086/466536 (2005).

Scallan, E. *et al.* Foodborne illness acquired in the United States--major pathogens. *Emerging infectious diseases* **17**, 7-15, doi:10.3201/eid1701.091101p1 (2011).

Scallan, E. *et al.* Factors associated with seeking medical care and submitting a stool

sample in estimating the burden of foodborne illness. *Foodborne pathogens and disease***3**, 432-438, doi:10.1089/fpd.2006.3.432 (2006).

Tarr, P. I., Gordon, C. A. & Chandler, W. L. Shiga-toxin-producing Escherichia coli and haemolytic uraemic syndrome. *Lancet* **365**, 1073-1086, doi:10.1016/S0140-6736(05)71144-2 (2005).

Riley L, Remis R, Helgerson S, et al. Hemorrhagic colitis associated with a rare Escherichia coli serotype. N Engl J Med 1983;308:681–5

Gould, L. H., Mody, R. K., Ong, K. L., Clogher, P., Cronquist, A. B., Garman, K. N., ... & Griffin, for the Emerging Infections Program FoodNet Working Group, P. M. (2013). Increased recognition of non-O157 Shiga toxin–producing Escherichia coli infections in the United States during 2000–2010: Epidemiologic features and comparison with E. coli O157 infections. *Foodborne pathogens and disease*, *10*(5), 453-460.

Armstrong GL, Hollingsworth J, Morris JG. Emergingfoodborne pathogens: Escherichia coli O157:H7 as amodel of entry of a new pathogen into the food supplyof the developed world. Epidemiol Rev 1996; 18: 29–53.

Wang G, Doyle MP. Survival of enterohemorrhagic Escherichia coli O157:H7 in water. J Food Protect 1998; 6: 662–667.

Tauxe RV. Emerging foodborne diseases : an evolving public health challenge. Emerg Infect Dis 1997; 3: 425–434.

Feng P. Escherichia coli serotype O157:H7: novel vehicles of infection and emergence of phenotypic variants. Emerg Infect Dis 1995; 1: 47–52.

An overview of geocoding, <http://webhelp.esri.com/arcgisdesktop/9.3/index.cfm?topicname=an_overview_of_geocoding> (2011).

CDC. *Foodborne Diseases Active Surveillance Network (FoodNet): Table and Figures-2011 Preliminary Data*, <http://www.cdc.gov/foodnet/data/trends/tables-2011.html> (2012).

13 *2020 Topics & Objectives: Food Safety* <http://healthypeople.gov/2020/topicsobjectives2020/objectiveslist.aspx?topicid=14> (

Keen, J. E., Wittum, T. E., Dunn, J. R., Bono, J. L., & Durso, L. M. (2006). Shiga-toxigenic Escherichia coli O157 in agricultural fair livestock, United States. *Emerg Infect Dis*, *12*(5), 780-786

Crump JA, Braden CR, Dey ME, Hoekstra RM, Rickelman-Apisa JM, Baldwin DA, et al. Outbreaks of Escherichia coli O157 infections at multiple county agricultural fairs: a hazard of mixing cattle, concession stands and children. Epidemiology Infect. 2003;131:1055–62

Winfield, M. D., & Groisman, E. A. (2003). Role of nonhost environments in the lifestyles of Salmonella and Escherichia coli. *Applied and environmental microbiology*, *69*(7), 3687-3694.