January 2013

Neighborhood Level Socioeconomic Status And Rurality And Shiga Toxin-Producing Escherichia Coli Incidence: Connecticut, 2000-2011

Bridget Morrissey Whitney
Yale University, bridget.whitney@yale.edu

Neighborhood Level Socioeconomic Status and Rurality and Shiga Toxin-Producing *Escherichia coli* Incidence: Connecticut, 2000–2011

Bridget Whitney

Epidemiology of Microbial Diseases

Yale School of Public Health

2013

Acknowledgements

I would like to thank and share my upmost respect for Dr. James Hadler and Dr. Linda Niccolai for their guidance, suggestions, and evaluation of this analysis. Their encouragement and critiques were greatly appreciated and were paramount to this work.

Additionally, I would like to share my gratitude to Sharon Hurd, Paula Clogher, Danyel Olsen, James Meek, Kelly Bemis, and the Emerging Infections Program staff for their advice, support, and contributions to this project. Furthermore, I wish to acknowledge the Connecticut State Department of Public Health for their efforts and collaboration.

Finally, I wish to thank my fellow Yale School of Public Health students for their comradery and support throughout the thesis project.

Table of Contents

Abstract

Background: Shiga toxin *Escherichia coli* (STEC) O157 and other STEC strains are a well-known cause of enteric illness. National estimates are that STEC O157 causes approximately 96,534 illnesses every year in the United States, with another 168,698 illnesses caused by non-O157 STEC serotypes. Determining economic and sociodemographic factors associated with enteric disease incidence may provide new understandings of the transmission of these illnesses, particularly community transmission, and may prove useful in the prevention of disease.

Methods: A total of 764 incident STEC cases were reported in CT from 2000 to 2011. Incident cases were geocoded based on the case's address using ArcGIS. Incident cases were linked to neighborhood poverty level and neighborhood rurality level at the census tract level. Neighborhood poverty level was broken down into four categories for analysis: 0 – 4.99%, 5 – 9.99%, 10 – 19.99%, and greater than 20% of the population in the census tract living below the federal poverty line. Neighborhood rurality level was broken down into quartiles for analysis as well: 0 – 24.9%, 25 – 49.9%, 50 – 74.9%, and greater than 75% of housing units in the census tract considered rural. Twelve-year age-adjusted Shiga toxin *E. coli* incidence rates were calculated for each poverty category and each rurality category. Incidence rates were also determined by race/ethnicity.

Results: Of the 764 cases, 744 (97.4%) were able to be geocoded. Both neighborhood level poverty and neighborhood level rurality were found to be significantly associated with STEC incidence. Age-adjusted rates of all STEC infections revealed a trend of decreasing neighborhood poverty level and increasing STEC incidence ($p<0.001$); residents of the wealthiest census tracts were four times as likely to contract STEC compared to residents of the highest poverty census tracts. Age-adjusted rates of all STEC infections showed a trend of increasing neighborhood rurality and increasing incidence ($p<0.001$); residents of the most rural census tracts were 1.7 times as likely to contract STEC compared to residents of the most urban census tracts The same significant incidence associations were seen among O157 STEC cases and non-O157 STEC cases separately and were consistent across time periods, age, and race/ethnicity groups.

Conclusions: STEC incidence decreased as neighborhood poverty increased, showing a dose-response relationship with socioeconomic status, and increased as neighborhood rurality increased. These findings can be used to more effectively target education and interventions, especially in high-income neighborhoods, which include more rural neighborhoods in Connecticut. Area-based socioeconomic measures provide additional insights into the epidemiology of infectious diseases and can be used further to elucidate possible control and prevention measures. Future study implications include the need to better understand what risk exposures are driving the differences between higher and lower poverty areas, including among infants and children. What types of educational efforts are effective at reducing risk among those of higher SES also needs to be investigated. This analysis provides support that community-level risk factors play a larger role in the transmission of STEC.

Introduction

Escherichia coli (*E. coli*) is a principle enteric microorganism in the human gut.[1] As part of the normal flora, non-pathogenic *E. coli* is not harmful to its human host and actually provides benefits such as vitamin K synthesis.[1] A pathogenic *E. coli* strain, Shiga toxin-producing *Escherichia coli* (STEC), that expressed the O-antigen 157 and the H-antigen 7, was first noted in 1982 after two outbreaks of unusual severe bloody diarrhea and gastrointestinal illness were investigated by the Center for Disease Control and Prevention (CDC) for a link to foodborne illness.[2]

STEC O157:H7, and other STEC strains displaying different O-antigens are now well-known causes of enteric illness in the United States and in Connecticut (CT). More than 200 *E. coli* serotypes are known to produce shiga toxins with over 100 of these serotypes associated with human illness.[3,4] The initially recognized O157 serotype is the most common serotype isolated in North America, and in CT was confirmed in 41% of all STEC cases from 2000 to 2009.[4,5] Illness caused by STEC ranges from asymptomatic shedding, to mild diarrhea, bloody diarrhea, hemorrhagic colitis, and hemolytic-uremic syndrome (HUS).[4] HUS, a severe kidney disease, develops in approximately 5%-10% of people with STEC-associated diarrhea, roughly 10% of individuals who develop HUS die (3%-7%) or have permanent renal failure; up to 50% of people who develop HUS will have some level of renal damage.[3,6] There is significant risk of chronic and end-stage renal disease, persisting renal hypertension, diabetes mellitus, and neurological disorders in persons effected by HUS.[7] In the United States, development of HUS is most often associated with the O157 serotype.[8] HUS disproportionately affects children under the age of five years.[7,9]

It is estimated that each year 31 major pathogens acquired in the United States cause 9.4 million episodes of foodborne illness, 55,961 hospitalizations, and 1,351 deaths.[10] STEC infections account for a small proportion of these illnesses, but STEC can cause serious morbidity making it one of the most important emerging pathogens in food.[7] Estimates, after adjustment for underdiagnosis, show that STEC O157 causes approximately 96,534 (26,982–227,891) illnesses every year in the United States, with another 168,698 (17,163–428,522) illnesses caused by non-O157 STEC serotypes; STEC causes disease at rates of occurrence similar to other important enteric pathogens.[10] Both O157 and non-O157 serotypes of STEC are associated with sporadic and outbreak-linked disease, however surveillance data suggests that most STEC infections do not occur in an outbreak setting.[9] Both O157 and non-O157 serotypes of STEC also show summer seasonality in the United States, with a higher percentage of cases occurring in summer months.[4]

Livestock, in particular cows, and other ruminant animals are an important reservoir for STEC.[3,11] Important identified risk factors for the contraction of STEC include "foods of bovine origin," notably undercooked ground beef and unpasteurized milk.[6,9] Food and water sources associated with infection, such as apple cider, fresh vegetables and sprouts, and private well water sources, are thought to be contaminated by cattle feces.[6,9] Non-dietary risk factors, often known as recreational or environmental risk factors, include swimming in contaminated water, visits to petting zoos or farms (direct contact with infected animals), and human-to-human transmission in day care settings.[6,9] Almost all the current information on risk factors for Shiga toxin-producing *Escherichia coli* infections comes from surveillance information and outbreak investigations.[9]

Since 1996 the Foodborne Disease Active Surveillance Network (FoodNET), a CDC

surveillance program operating out of the Emerging Infections Program (EIP), has been tracking the incidence of O157 STEC infections in ten states including CT. CT started tracking the incidence of non-O157 STEC as well in 2000, taking advantage of the fact that some laboratories had switched from culture of O157 alone to testing for Shiga toxin. Shiga toxin positive specimens could be cultured to look for other serotypes of STEC in addition to O157. In 2011 the incidence of laboratory confirmed O157 STEC infections was 0.48 per 100,000 persons in CT and 0.99 per 100,000 persons in the entire FoodNET catchment area with contains approximately 15% of the population of the United States.[12] The "incidence" of laboratory confirmed non-O157 STEC infections in 2011, an underestimate as not all laboratories did the necessary Shiga toxin screening to initiate the process of detecting non-O157, was 0.62 per 100,000 persons in CT and 1.08 per 100,000 persons in the entire FoodNET catchment area.[12] There was a significant decrease in the reported incidence of O157 STEC infections in 2011 when compared to the data from 2006 to 2009; there has also been a sustained decline in the incidence of O157 STEC in the FoodNET catchment area since the initiation of FoodNET in 1996.[12] In 2011 the incidence goal of 1.00 cases per 100,000 persons for O157 STEC, set forth by Healthy People 2010, was significantly exceeded in CT.[12] The new Healthy People 2020 goal is to reduce O157 STEC incidence further, to 0.60 cases per 100,000 persons, and to reduce the incidence of post-diarrheal hemolytic uremic syndrome (HUS) from 1.8 cases per 100,000 children under the age of five per year (2005–2007) to 0.90 cases per 100,000 children.[13] Currently there is no Healthy People 2020 goal for non-O157 STEC. Reaching these goals will require further knowledge on the epidemiology of STEC, including understanding the socioeconomic factors that contribute to STEC infections in the United States.

Determining economic and sociodemographic elements associated with enteric disease

incidence may provide new understandings of the transmission of these illnesses, particularly community transmission, and may prove useful in the prevention of disease.[14] Previous studies have indicated that race/ethnicity, place of residence (e.g., urban versus rural), educational attainment, poverty, and age may affect the risk of infections with salmonellosis, shigellosis, and *E. coli* O157:H7.[14] There have been no published reports on the association between STEC and socioeconomic status in the United States. Routine public health surveillance interviews are attempted by FoodNET staff for all cases of O157 and non-O157 STEC reported to the state. These interviews, however, do not ascertain individual socioeconomic variables. For this analysis an alternative measure of socioeconomic status, neighborhood level socioeconomic status using census tract-level data, will be used. This analysis will provide a unique perspective on individual and community socioeconomic risk factors associated with Shiga toxin-producing *E. Coli* infection. The goal of this analysis was to use ArcGIS, geographic information system software, and neighborhood level census information to evaluate the association between Shiga toxin-producing *E. Coli* infections and area-based socioeconomic measures in CT.

Methods

Case Identification and Data Collection

Escherichia coli O157:H7 gastroenteritis and Shiga toxin-related disease (gastroenteritis) are both physician and laboratory reportable diseases in the state of Connecticut.[15] Laboratories are required to submit every *E. coli* O157 isolate as well as every positive shiga toxin broth to the CT State Public Health Laboratory (SPHL) for confirmation. If the SPHL is able to identify a shiga toxin-producing *E. coli* but unable to identify the O-antigen (the Connecticut SPHL is only able to identify the six most common O-antigens), the specimen is sent to CDC laboratory for further identification. Telephone interviews are attempted on all cases of STEC reported in CT residents using a standardized instrument: The *E. coli* O157 and other Shiga Toxin-Producing *E. coli* Questionnaire. Important variables extracted from the reportable disease forms and/or collected during interviews include address of residence, age, sex, and race/ethnicity. The non-demographic information collected via the questionnaire includes clinical information (symptomology, onset, hospitalization, treatment with antibiotics) and exposure history information (international travel, environmental exposures, food and water exposures).

Geocoding and Spatial analysis of incident cases

Geocoding is "the process of assigning a location, usually in the form of coordinate values (points), to an address by comparing the descriptive location elements in the address to those present in the reference material."[16] A total of 764 incident STEC cases were reported in CT from 2000 to 2011. An effort was made to geocode each incident case based on the case's address using Environmental Systems Research Institute's (ESRI) ArcGIS. For the majority of case addresses, the automatic settings for geocoding were employed. If automatic setting were unsuccessful, interactive geocoding was performed, which included looking up original case

report forms and checking and correcting number and spelling errors with Google Maps and USPS.com. The reference network used to geocode case addresses were 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.[17,18]

All shape files in this analysis were projected in ArcGIS using the North American Datum of 1983 (NAD 1983) projection, which is a state plane coordinate system. All geocoded data were then joined to their corresponding census tract using ArcGIS. Census tract socioeconomic data, percent of the population below the poverty line, as well as rurality, age, and population data were downloaded from the United State Census Bureau's American Community Survey (ACS) and linked to census tracts. All data was then imported into SAS for analysis.

Statistical analysis

Statistical analysis was limited to cases whose address could be geocoded in ArcGIS and successfully linked to a census tract. First, incident cases were linked to neighborhood poverty level and neighborhood rurality level at the census tract level. Neighborhood poverty level was measured as percent of the population in each census tract living below the federal poverty line at the time of the U.S. Census and neighborhood rurality level was measured as percent of housing units in each census tract that was considered rural (urban vs. rural) at the time of the U.S. Census. Data from the 2000 census was used for incident cases from 2000-2005 and data from the 2010 American Community Survey (ACS) was used for incident cases from 2006-2010. Next, all incident cases were aggregated into four race categories: Hispanic, non-Hispanic white, non-Hispanic black, and other. Age was then categorized into three groups, 0-4, 5-17, and over 18 years of age. Five age groups were originally used, 0-4, 5-17, 18-40, 40-64, and over 65 years

of age, but all adult age groups were found to have very similar incidence rates so they were consolidated. Neighborhood poverty level was broken down into four categories for analysis: 0 – 4.99%, 5 – 9.99%, 10 – 19.99%, and greater than 20%. These categories were chosen based on previous work done by the Public Health Disparities Geocoding Project.[17] Two census tracts from the 2000 census, totaling to 1008 people, could not be assigned to a poverty category due to missing data and were therefore excluded from the aggregated denominators. Two census tracts from the 2010 census, totaling to 3434 people, could not be assigned to a poverty category due to missing data and were therefore excluded from the aggregated denominators. No incident cases resided in these four census tracts. Neighborhood rurality level was broken down a priori into four groups for analysis as well: 0 – 24.9%, 25 – 49.9%, 50 – 74.9%, and greater than 75%. It was hypothesized that with an increasing percent living in rural areas, there might be a higher potential for contact with cattle and other farm animals.

Twelve-year crude and age-adjusted Shiga toxin *E. coli* incidence rates were calculated for each poverty category and each rurality category. Age-adjusted incidence rates were calculated using the direct method with weights taken from the average overall Connecticut populations of the 2000 and 2010 censuses. Crude incidence rates were compared between age groups, sexes, demographic groups, and two time periods. Comparisons between O157 and non-O157 cases were made along the same characteristics. Incidence rate ratios (IRRs) were calculated for age-adjusted rates using the 0-4.99% poverty group and the 0-24.9% rurality group as the references. Age-adjusted incidence rates for each poverty and rurality category for all STEC and for O157 and non-O157 separately were compared using IRRs. Associations between STEC incidence and poverty or rurality within major race/ethnic groups were examined. Chi-square tests for trend were performed to test the statistical significance of the gradients among

the four categories for poverty and rurality and 95% confidence intervals were calculated. A correlation between neighborhood level poverty and neighborhood level rurality was assessed, stratified analyses were performed, and Mantel Haenszel adjusted odds ratios (ORs) were calculated. All analyses were carried out in SAS 9.3 with the exception of chi-square tests for trend, which were carried out in EpiInfo 2000.

Results

Sample Characteristics

From 2000 to 2011, 764 incident cases of shiga toxin-positive specimens with culture confirmed *E. coli* were reported to the CT Department of Public Health and FoodNET. Of the 764 incident cases of STEC infections in Connecticut, 744 (97.4%) were matched on street address, ZIP/postal code, and city/state/province. All cases that did not automatically match in GIS were manually checked and matched to a census tract interactively in the GIS software; 10 cases were discarded because a P.O. Box was listed as an address and 10 were not matched because the address could not be found with the address locator. Cases that could not be matched in GIS did not differ significantly from matched cases (data not shown). The distribution of cases did not differ significantly by O-antigen with respect to age or sex, but did differ by race/ethnicity; non-Hispanic whites and non-Hispanic blacks were more prevalent among the O157 STEC compared to non-O157 STEC group ($p<0.001$) (Table 1).

Incidence and Neighborhood Level SES

Overall, the largest group defined by poverty in CT was the least impoverished group, 0-4.99% below the poverty line group (51.9%), followed by the 5-9.99% group (21.1%) and by the 10-19.99% group (15.5%) The smallest group was the poorest group (>20% below the federal poverty level) (11.5%). Neighborhood level poverty was found to be significantly associated with STEC incidence across all twelve years of data (Table 2, Figure 2), within first and second six years of data separately (Table 4, Figure 3), showing consistency across time, and among O157 and non-O157 cases (Table 2, Figure 3). Age-adjusted rates of all STEC infections revealed a trend of decreasing neighborhood poverty level associated with increasing STEC incidence ($p<0.001$) (Table 2, Figure 2). The lowest poverty group, 0-4.99% of the population

below the poverty line, had a 12-year age-adjusted incident rate of 2.46 (95% CI: 2.25, 2.67) cases per 100,000 person years, a rate 4.0-fold than the highest poverty group, >20% living below the federal poverty level. The other two groups had rates intermediate to these (Table 2, Figure 2). The same association between higher incidence and lower census-tract level poverty was seen among O157 STEC cases and non-O157 STEC cases separately ($p<0.001$) (Table 2) and among each age group (Figure 4). For 0-4 year olds, however, the strength of the association was the weakest (only 1.73 relative risk between the lowest and highest poverty groups) and only on the border of statistical significance ($p=0.055$).

Incidence and Neighborhood Level Rurality

Overall, the largest group defined by rurality in CT was the most urban group, 0-24.9% rural (83.4%), followed by 25-49.9% group (6.8%) and by the most rural group, >75% (5.9%) The smallest group was the 50-74.9% rurality group (3.9%). Neighborhood level rurality was also found to be significantly associated with STEC incidence across all twelve years of data (Table 3, Figure 6), within first and second six years of data separately (Table 4, Figure 7), and among O157 and non-O157 cases (Table 3). Age-adjusted rates of all STEC infections showed a trend of increasing census tract rurality with increasing incidence ($p<0.001$) (Table 3, Figure 6). The most rural census tracts in Connecticut, census tracts with >75.0% of housing units considered rural, had a 12-year age-adjusted incident rate of 2.97 (95% CI: 2.29, 3.65) cases per 100,000 person years, a rate 1.70-fold than the most urban census tracts, with 0 – 24.9% of housing units considered rural. The other two groups had rates intermediate to these (Table 3, Figure 6). The same association between higher incidence and higher census-tract level rurality was seen among O157 STEC cases and non-O157 STEC cases separately ($p<0.001$ and $p=0.014$, respectively) (Table 3) and among each age group (Figure 8). For 0-4 year olds, however, the

strength of the association was the weakest (2.28 relative risk between the lowest and highest rurality groups) and on the border of statistical significance ($p=0.077$).

A correlation between neighborhood level poverty and neighborhood level rurality was assessed and showed strong correlation between the two variables ($p<0.001$) (data not shown); more rural areas have less poverty. Stratified analyses were performed (data not shown) and Mantel Haenszel adjusted odds ratios (ORs) were calculated. After adjusting for census tract rurality, census tract level poverty was found to be an independent predictor of STEC incidence; the association did not change much and was still significant ($p<0.001$) (Table 5) (Figure 10). After adjusting for census tract poverty, census tract level rurality was also found to be an independent predictor of STEC incidence. This association was attenuated by adjustment for poverty but still significant ($p=0.005$) (Table 5) (Figure 11).

Race/Ethnicity

To test for trends among different racial and ethnic groups the race/ethnicity categories were condensed into two groups, non-white and non-Hispanic whites, due to a small number of cases and lack of statistical power in individual non-white groups (Hispanic, non-Hispanic black, and other races). Decreasing incidence of all STEC infections with increasing neighborhood poverty was seen in non-whites ($p=0.014$) and non-Hispanic white populations ($p<0.001$) in Connecticut (Figure 5). Increasing incidence of all STEC infections with increasing neighborhood rurality was seen in non-whites ($p=0.019$) and non-Hispanic white ($p=0.006$) populations (Figure 9).

Discussion

This analysis was carried out to assess socioeconomic and geographic neighborhood factors in relation to the incidence of Shiga toxin-producing *E. coli* (STEC) infection in the state of Connecticut. Both neighborhood level poverty and neighborhood level rurality were found to be significantly associated with STEC incidence. Age-adjusted rates of all STEC infections revealed a trend of decreasing neighborhood poverty level and increasing STEC incidence. Age-adjusted rates of all STEC infections showed a trend of increasing neighborhood rurality and increasing incidence. The same significant incidence associations were seen among O157 STEC cases and non-O157 STEC cases separately and were consistent across time periods, age, and race/ethnicity groups. The relationship between these factors and STEC incidence in the United States has not been previously demonstrated in the published literature and has implications for future research directions for prevention.

The magnitude of the relative risk between the highest and lowest incidence groups was large. Residents of the wealthiest census tracts, census tracts where less than 5% of the population live below the federal poverty line, were nearly four times or as likely to contract STEC as residents of the most impoverished census tracts, where more than 20% of the population live below the federal poverty line. The magnitude of the relative risk was similar for O157 STEC infections and for non-O157 STEC infections and the relationship persisted over time, across all age groups, and across race/ethnicity categories.

There are several possible explanations for higher incidence of STEC infections in wealthier neighborhoods. First, people of different socioeconomic backgrounds may have different prevalence of high risk exposures, including consumption of known risky foods and international travel. It is possible that persons of higher socioeconomic status eat more

undercooked meat of bovine origin, a known risk factor for STEC.[6,9] In addition, people of higher income categories may consume more raw milk, unpasteurized ciders, and raw produce, which have been linked to STEC outbreaks. [6,9] Wealthier residents may also have more means to travel internationally, a risk factor linked to non-O157 STEC incidence.[5] A study of campylobacter infections in Connecticut examined FoodNet population survey data to determine whether there was a relationship between income and exposure factors related to campylobacter.[19] It found that higher income people were more likely to travel internationally than low income.[19] It did not examine food consumption patterns specific to STEC, however. This is a possible area for future study to assess what may be driving the demonstrated socioeconomic gradient for STEC.

A second possible explanation for the higher incidence seen among higher income residents of Connecticut in this analysis is different health seeking behaviors among people of different socioeconomic backgrounds that could result in those in higher income neighborhoods being more likely to be diagnosed and counted than those in poor neighborhoods. From 2000-2003 FoodNET assessed factors associated with seeking medical care and submitting a stool specimen among persons with acute diarrheal illness and found that approximately 20% of people with acute diarrheal diseases sought medical care, 19% of whom submitted a stool specimen.[18] The analysis found that those with incomes at or below $25,000 were more likely to seek medical attention and submit a specimen for testing than those with incomes greater than $25,000.[18] This data suggests that income/socioeconomic status does not account for the differences observed here, although the numbers were small.

Age-adjusted STEC incidence rates showed an opposite dose-response trend of increasing neighborhood rurality level and increasing STEC incidence. Residents of the most

rural census tracts, census tracts where greater than 75% of the housing units are considered rural, were 70% more likely to contract STEC than residents of urban census tracts, where less than 25% of housing units are considered rural. This trend was present for both O157 and non-O157 STEC, over time, and within each age and race/ethnicity group examined.

It is not entirely clear what the association between STEC incidence and rurality means. The aim of this analysis was to try to assess farm animal exposure, in particular to cows and other ruminant animals, as they are an important reservoir for STEC.[3,11] We hypothesized that environmental exposures, including swimming in water contaminated with bovine feces and visiting petting zoos or farms (direct contact with infected animals), are more likely to occur among people living in rural settings.[6,9] Although we found an association of STEC incidence with rurality, we also found that rurality is associated with higher socioeconomic status in Connecticut and it is possible these variables are assessing the same exposures rather than different ones. Both rurality and poverty were independent predictors when adjusted for the other, so this is likely not the case. More research in Connecticut is needed to be able to separate out the interaction of these variables, including analyses of case exposure data and FoodNet population survey data to determine STEC risk factor differences by poverty level and rurality of residence.

There are several limitations in this analysis. First and foremost, the incidence of reported STEC is likely to be greatly underestimated. Only about 20% of persons with acute diarrhea seek medical attention, and not all laboratories test for all STEC types.[18,20] In addition, while all laboratories test for O157, only a limited number do Shiga-toxin testing, a prerequisite for identifying non-O157 STEC. Cases of STEC that were captured and reported through laboratory assessment may differ from the true burden of disease and may not be a representative sample

for the state of Connecticut, particularly non-O157 STEC. A second limitation of this analysis is the homogeneity of cases in regards to race. A large proportion of the Connecticut population is white, non-Hispanic and this data mirrored the population, with 79.1% of the cases identifying as non-Hispanic white. The number of Hispanic and non-Hispanic black and Asian cases were too small to look at each group separately. With a larger sample, and therefore more cases in other race/ethnicity categories, a more detailed analysis on socioeconomic status and race/ethnicity in regards to STEC incidence could be performed. A third limitation is the poverty measure used in this analysis. It is an area-based measure and does not directly measure individual income or SES. However, it does incorporate behaviors that may be influenced by the community one lives in, possibly including some risk behaviors related to STEC exposure. Another potential limitation of this analysis is what "rurality" means in Connecticut. Unfortunately, there is not a detailed definition from the American Community Survey as to what makes a housing unit "rural" for their assessment. This measure was used to try to assess frequent potential farm and farm animal exposures, but it is uncertain if this variable is really getting at this measure.

Although the causal factors behind the dose-response trend of decreasing neighborhood poverty level and increasing STEC incidence could not be determined, this study has some important implications, including control and research implications. Efforts to reduce STEC risk need to especially focus on higher SES groups and rural residents. More effective educational efforts are needed to discuss consequences of STEC infection, risk factors, and prevention among people of higher SES. Future study implications include the need to better understand what risk exposures are driving the differences between higher and lower poverty areas, including among infants and children. What types of educational efforts are effective at reducing

risk among those of higher SES also needs to be investigated. This analysis provides support that community-level risk factors play a larger role in the transmission of STEC.

In summary, Shiga toxin-producing *E. coli* show a strong dose-response trend of higher incidence among higher income and more rural census tracts in Connecticut. This is a new and exciting area of research with a need to better define why SES level makes a difference and to identify effective interventions to reduce the burden of disease, which has significant morbidity and mortality for those who are affected. The differences in the incidence of STEC by SES and rurality are probably multifactorial and complex. Improved understanding can help design and focus prevention messages. Future analyses should focus on understanding reasons for differences in STEC incidence rates.

Bibliography

1 Center for Food Safety and Applied Nutrition (U.S.). (U.S. Food and Drug Administration, Center for Food Safety and Applied Nutrition, College Park, Md.).

2 Armstrong, G. L., Hollingsworth, J. & Morris, J. G., Jr. Emerging foodborne pathogens: Escherichia coli O157:H7 as a model of entry of a new pathogen into the food supply of the developed world. *Epidemiologic reviews* **18**, 29-51 (1996).

3 Hedican, E. B. *et al*. Characteristics of O157 versus non-O157 Shiga toxin-producing Escherichia coli infections in Minnesota, 2000-2006. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **49**, 358-364, doi:10.1086/600302 (2009).

4 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).

5 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).

6 Thorpe, C. M. Shiga toxin-producing Escherichia coli infection. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **38**, 1298-1303, doi:10.1086/383473 (2004).

7 Karch, H., Bielaszewska, M., Bitzan, M. & Schmidt, H. Epidemiology and diagnosis of Shiga toxin-producing Escherichia coli infections. *Diagnostic microbiology and infectious disease* **34**, 229-243 (1999).

8 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).

9 Slutsker, L. *et al*. A nationwide case-control study of Escherichia coli O157:H7 infection in the United States. *The Journal of infectious diseases* **177**, 962-966 (1998).

10 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).

11 Karmali, M. A. Infection by Shiga toxin-producing Escherichia coli: an overview. *Molecular biotechnology* **26**, 117-122, doi:10.1385/MB:26:2:117 (2004).

12 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> (

14 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).

15 *Department of Public Health: Infectious Disease Reporting*, <http://www.ct.gov/dph/cwp/view.asp?a=3136&Q=453590&PM=1> (

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

17 *The Public Health Disparities Geocoding Project Monograph*, <http://www.hsph.harvard.edu/thegeocodingproject/webpage/monograph/methods.htm#aggabsm>

18 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).

19 Bemis, K. Neighborhood Level Socioeconomic Status and Campylobacter Incidence: Connecticut, 1999 – 2009. (2011).

20 Angulo, F. J. *et al*. Determining the burden of human illness from food borne diseases. CDC's emerging infectious disease program Food Borne Diseases Active Surveillance Network (FoodNet). *The Veterinary clinics of North America. Food animal practice* **14**, 165-172 (1998).

Tables and Figures

Table 1. Incidence of STEC by demographic features and comparison of cases by O antigen

	STEC (N=744)				STEC O Antigen				
					O157 (N=471)		non-O157 (N=273)		
Characteristic	N[a]	Crude IR	RR	P[b]	N[a]	%	N[a]	%	P[b]
Age (years)									0.20
0-4	124	4.88	4.52	<0.001	75	15.92	49	17.95	
5-17	296	4.00	3.70	<0.001	199	42.25	97	35.53	
≥18	324	1.08	ref		197	41.83	127	46.52	
Sex									0.53
Female	433	2.01	1.31	<0.001	270	57.32	163	59.71	
Male	311	1.53	ref		201	42.68	110	40.29	
Race/ethnicity									<0.001
Hispanic	50	1.04	0.55	<0.001	23	4.88	27	9.89	
Non-Hispanic white	589	1.89	ref		393	83.44	196	71.79	
Non-Hispanic black	20	0.53	0.28	<0.001	17	3.61	3	1.10	
Non-Hispanic Other	19	0.87	0.46	<0.001	10	2.12	9	3.30	
Unknown	66				28	5.94	38	13.92	
Time period				0.048					<0.001
2000-2005	390	1.86	ref		270	57.32	120	43.96	
2006-2011	354	1.69	0.91		201	42.68	153	56.04	

[a] Table values are N.
[b] P-value is for χ^2 test.

Table 2. Incidence Rate* and Ratios by Neighborhood Poverty Level, STEC, CT 2000-2011

	Neighborhood Poverty Level (% below poverty line)				
	0 - 4.99%	5.0 - 9.99%	10.0 - 19.99%	≥20.0%	p[a]
Total Population	1812235	735328	539848	402421	
All STEC					
Number of Cases	498	138	77	31	
Crude IR	2.29	1.56	1.19	0.64	
Age-Adjusted IR	2.46	1.73	1.23	0.63	<0.001
Age-Adjusted IR 95% CI	2.25, 2.67	1.46, 2.00	0.96, 1.50	0.40, 0.86	
Age-Adjusted IRR	1.00	0.71	0.50	0.25	
O157 STEC					
Number of Cases	310	99	44	18	
Crude IR	1.43	1.12	0.68	0.37	
Age-Adjusted IR	1.54	1.23	0.71	0.37	
Age-Adjusted IR 95% CI	1.38, 1.70	1.00, 1.46	0.50, 0.92	0.20, 0.54	
Age-Adjusted IRR	1.00	0.80	0.46	0.24	<0.001
Non-O157 STEC					
Number of Cases	188	39	33	13	
Crude IR	0.86	0.44	0.51	0.27	
Age-Adjusted IR	0.92	0.49	0.51	0.26	
Age-Adjusted IR 95% CI	0.79, 1.05	0.34, 0.64	0.34, 0.68	0.12, 0.40	
Age-Adjusted IRR	1.00	0.53	0.55	0.29	<0.001

* Per 100,000

[a] P-value is for χ^2 test for trend.

Table 3. Incidence Rate* and Ratios by Neighborhood Rurality, All STEC, CT 2000-2011

	Neighborhood Rurality Level (% of housing units considered rural)				
	0 – 24.9%	25.0 – 49.9%	50.0 – 74.9%	≥75.0%	P[a]
Total Population	2910246	237529	137045	205012	
All STEC					
Number of Cases	570	65	42	67	
Crude IR	1.63	2.28	2.55	2.72	
Age-Adjusted IR	1.75	2.42	2.61	2.97	
Age-Adjusted IR 95% CI	1.61, 1.89	1.85, 2.99	1.83, 3.39	2.29, 3.65	
Age-Adjusted IRR	1.00	1.38	1.49	1.70	<0.001
O157 STEC					
Number of Cases	355	45	26	45	
Crude IR	1.02	1.58	1.58	1.83	
Age-Adjusted IR	1.10	1.68	1.62	1.97	
Age-Adjusted IR 95% CI	0.99, 1.21	1.20, 2.16	1.01, 2.23	1.41, 2.53	
Age-Adjusted IRR	1.00	1.53	1.49	1.77	<0.001
Non-O157 STEC					
Number of Cases	215	20	16	22	
Crude IR	0.62	0.70	0.97	0.89	
Age-Adjusted IR	0.66	0.74	0.98	1.00	
Age-Adjusted IR 95% CI	0.57, 0.75	0.42, 1.06	0.50, 1.46	0.61, 1.39	
Age-Adjusted IRR	1.00	1.12	1.48	1.54	0.014

* Per 100,000

[a] P-value is for χ^2 test for trend.

Table 4. Incidence Rate* and Ratios by Time Periods, All STEC, CT

	Neighborhood Poverty Level (% below poverty line)				
	0 - 4.99%	5.0 - 9.99%	10.0 - 19.99%	≥20.0%	p[a]
2000-2005, N=390					
Crude IR	1.25	0.74	0.65	0.23	
Age-Adjusted IR	1.33	0.79	0.66	0.22	
Age-Adjusted IRR	1.00	0.59	0.49	0.16	<0.001
2006-2011, N=354					
Crude IR	1.04	0.82	0.55	0.39	
Age-Adjusted IR	1.12	0.94	0.58	0.40	
Age-Adjusted IRR	1.00	0.84	0.52	0.35	<0.001
	Neighborhood Rurality Level (% of housing units considered rural)				
	0 – 24.9%	25.0 - 49.9%	50.0 - 74.9%	≥75.0%	P[a]
2000-2005, N=390					
Crude IR	0.85	1.29	1.69	1.60	
Age-Adjusted IR	0.90	1.34	1.73	1.68	
Age-Adjusted IRR	1.00	1.48	1.95	1.86	<0.001
2006-2011, N=354					
Crude IR	0.79	1.00	0.94	1.11	
Age-Adjusted IR	0.86	1.10	0.95	1.26	
Age-Adjusted IRR	1.00	1.29	1.09	1.49	0.031

* Per 100,000

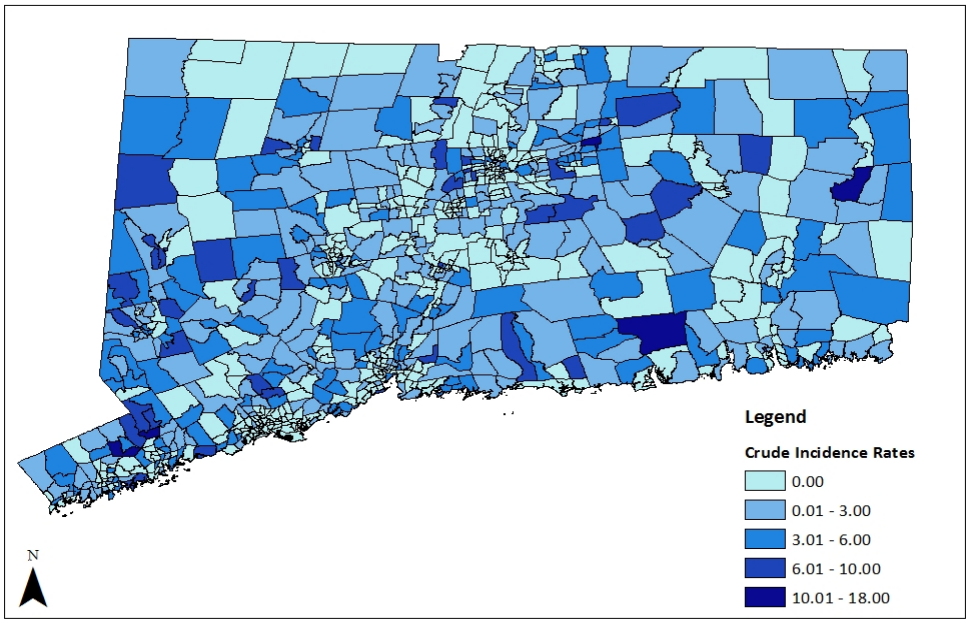
[a] P-value is for χ^2 test for trend.

Table 5. Mantel Haenszel Adjusted Odds Ratios, All STEC, CT

	Neighborhood Poverty Level (% below poverty line)				
	0 - 4.99%	5.0 - 9.99%	10.0 - 19.99%	≥20.0%	p[a]
Adjusted OR[b]	1.00	0.70	0.53	0.29	<0.001
	Neighborhood Rurality Level (% of housing units considered rural)				
	0 – 24.9%	25.0 - 49.9%	50.0 - 74.9%	≥75.0%	P[a]
Adjusted OR[c]	1.00	1.18	1.22	1.35	0.005

[a] P-value is for χ^2 test for trend.
[b] Adjusted for rurality.
[c] Adjusted for poverty.

Figure 1. Crude Incidence Rate of STEC per 100,000 person years by census tract, Connecticut, 2000 – 2011



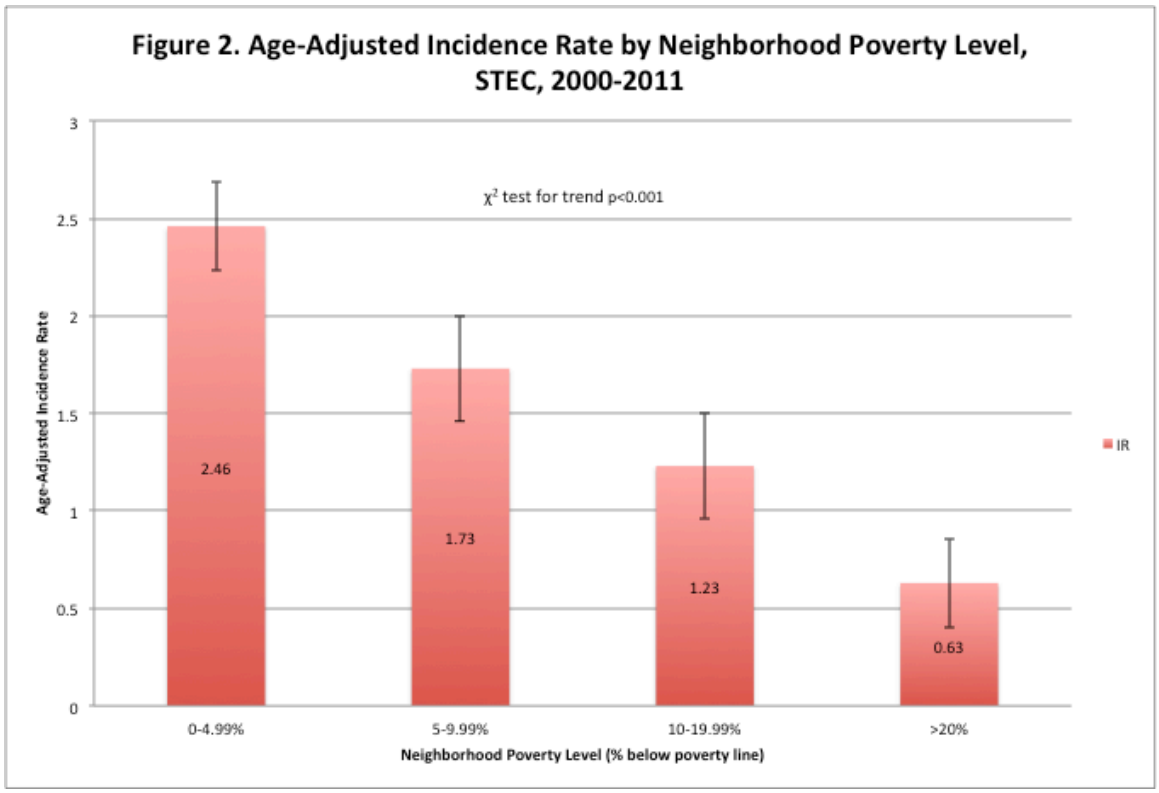


Figure 2. Age-Adjusted Incidence Rate by Neighborhood Poverty Level, STEC, 2000-2011
χ² test for trend p<0.001
Age-Adjusted Incidence Rate
3
2.5
2
1.5
1
0.5
0
2.46
1.73
1.23
0.63
0-4.99%
5-9.99%
10-19.99%
>20%
Neighborhood Poverty Level (% below poverty line)
IR

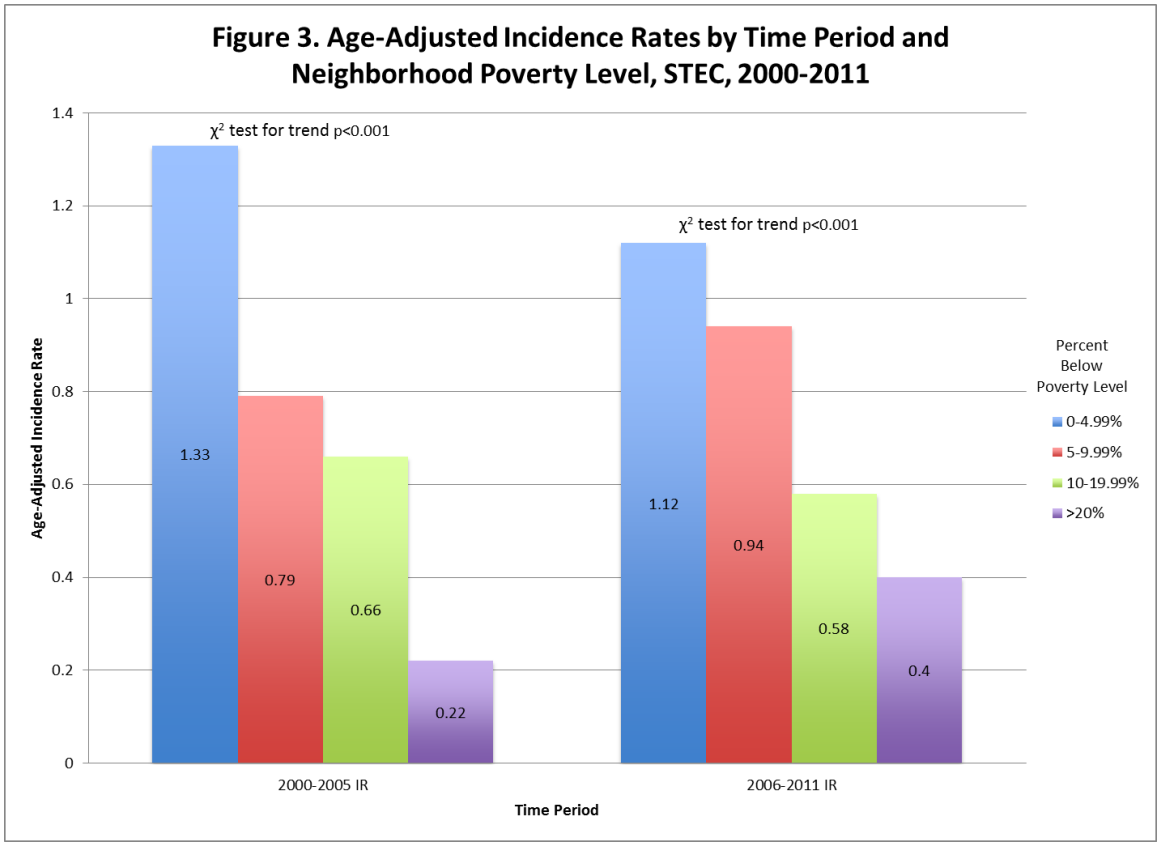


Figure 3. Age-Adjusted Incidence Rates by Time Period and Neighborhood Poverty Level, STEC, 2000-2011
χ² test for trend p<0.001
χ² test for trend p<0.001
Age-Adjusted Incidence Rate
1.4
1.2
1
0.8
0.6
0.4
0.2
0
1.33
0.79
0.66
0.22
1.12
0.94
0.58
0.4
2000-2005 IR
2006-2011 IR
Time Period
Percent Below Poverty Level
0-4.99%
5-9.99%
10-19.99%
>20%

Figure 4. Incidence Rate by Neighborhood Poverty Level and Age Group, STEC, 2000-2011

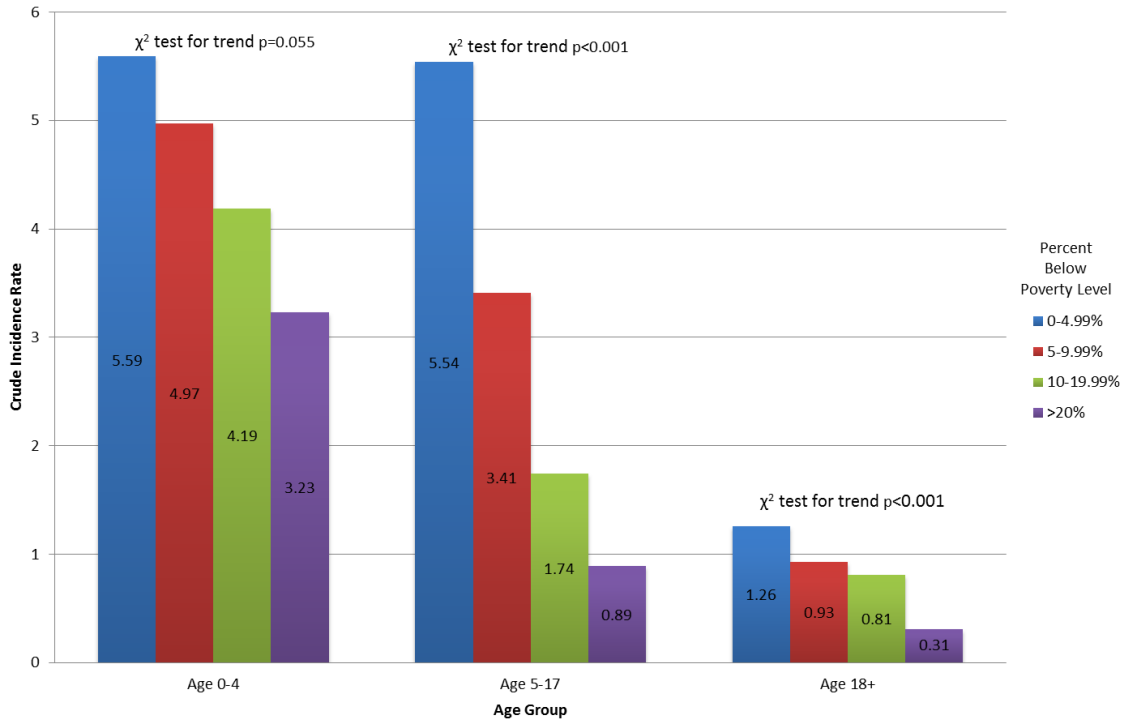
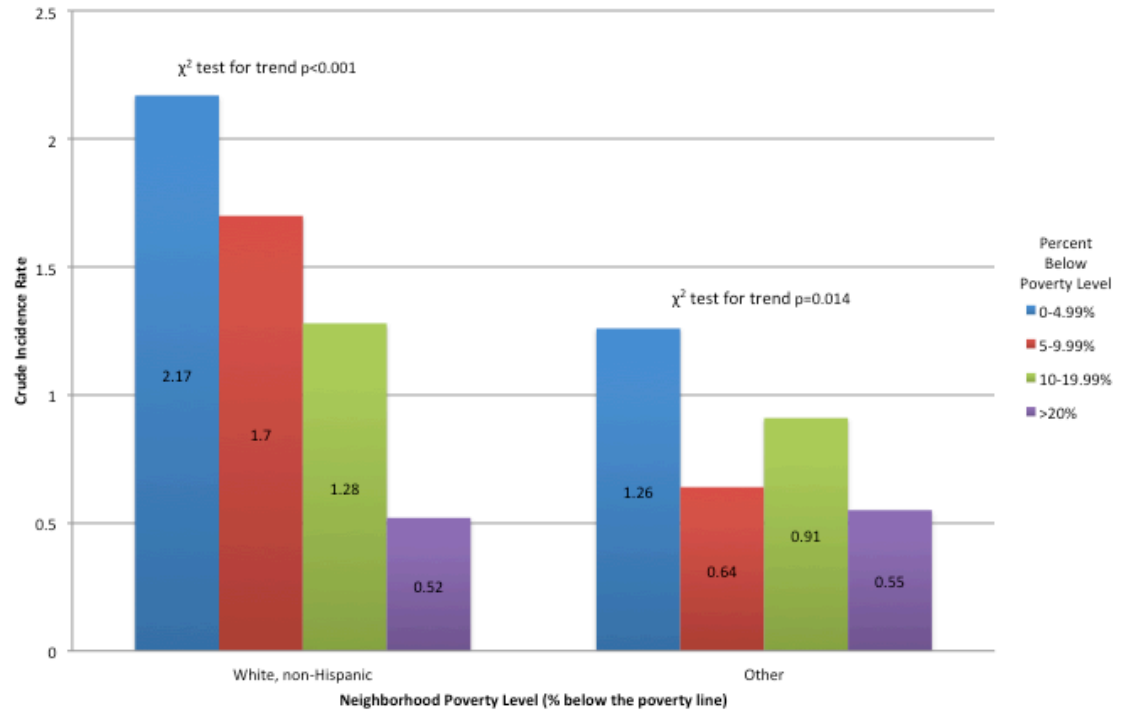


Figure 5. Crude Incidence Rate by Neighborhood Poverty Level and Race, STEC, 2000-2011



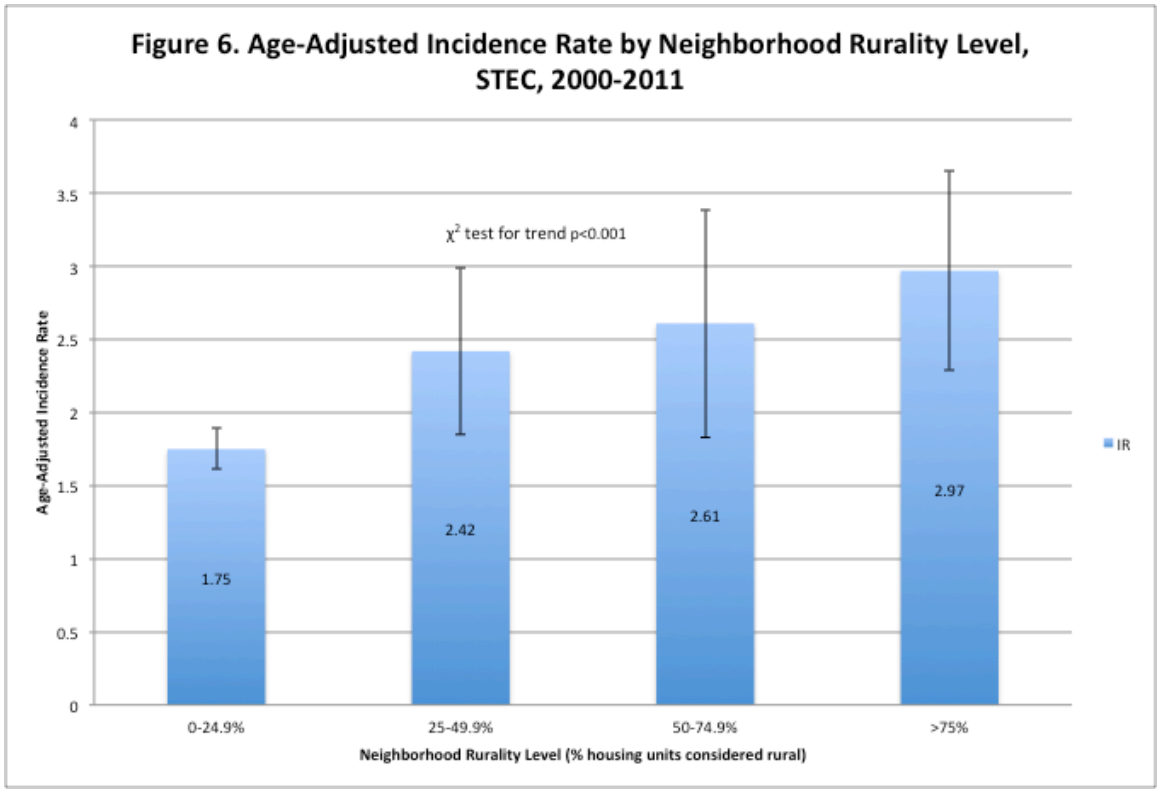


Figure 6. Age-Adjusted Incidence Rate by Neighborhood Rurality Level, STEC, 2000-2011
χ² test for trend p<0.001
Age-Adjusted Incidence Rate
4
3.5
3
2.5
2
1.5
1
0.5
0
1.75
2.42
2.61
2.97
IR
0-24.9%
25-49.9%
50-74.9%
>75%
Neighborhood Rurality Level (% housing units considered rural)

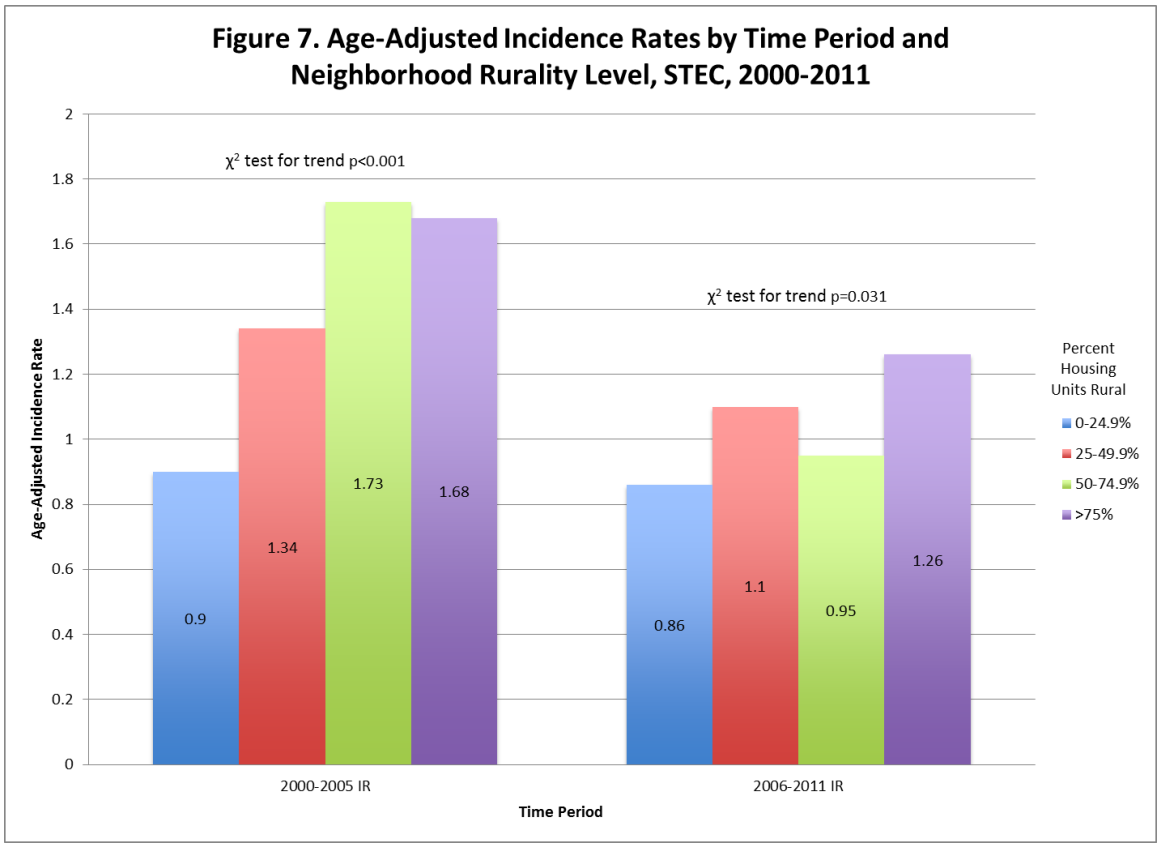


Figure 7. Age-Adjusted Incidence Rates by Time Period and Neighborhood Rurality Level, STEC, 2000-2011
χ² test for trend p<0.001
χ² test for trend p=0.031
Age-Adjusted Incidence Rate
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0.9
1.34
1.73
1.68
0.86
1.1
0.95
1.26
Percent Housing Units Rural
0-24.9%
25-49.9%
50-74.9%
>75%
2000-2005 IR
2006-2011 IR
Time Period

Figure 8. Incidence Rate by Neighborhood Rurality Level and Age Group, STEC, 2000-2011

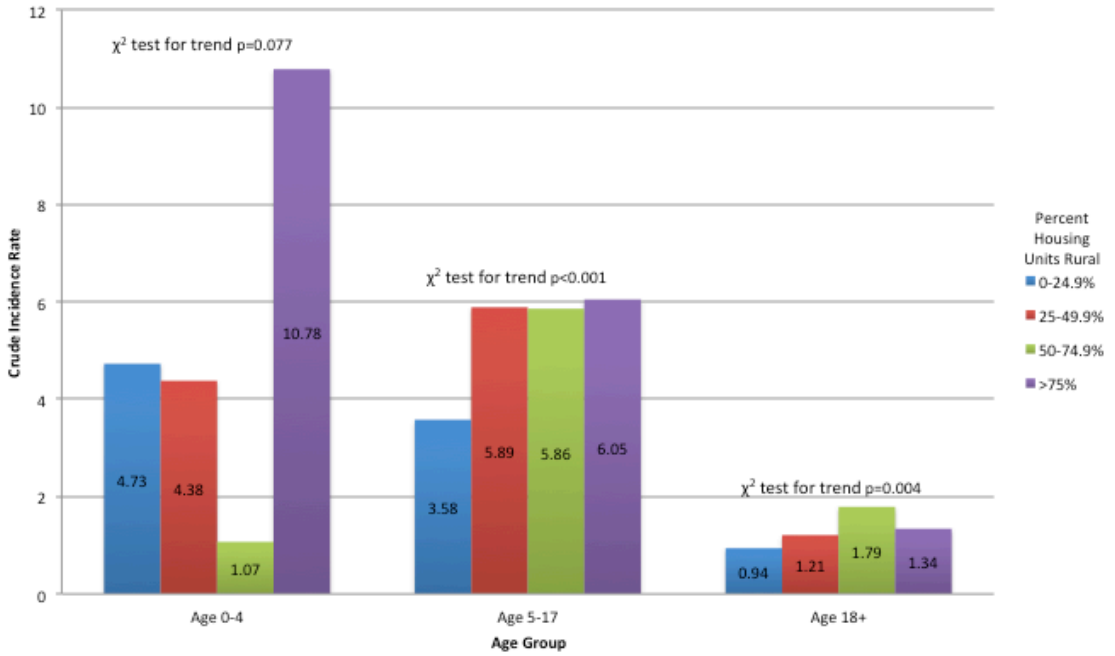
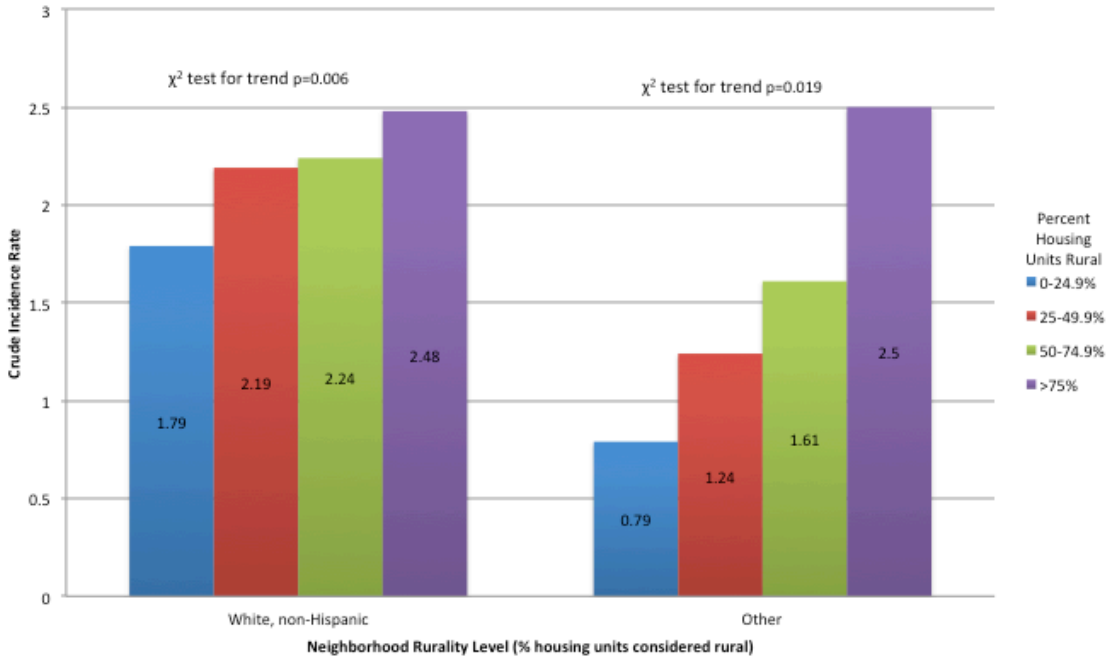


Figure 9. Crude Incidence Rate by Neighborhood Rurality Level and Race, STEC, 2000-2011



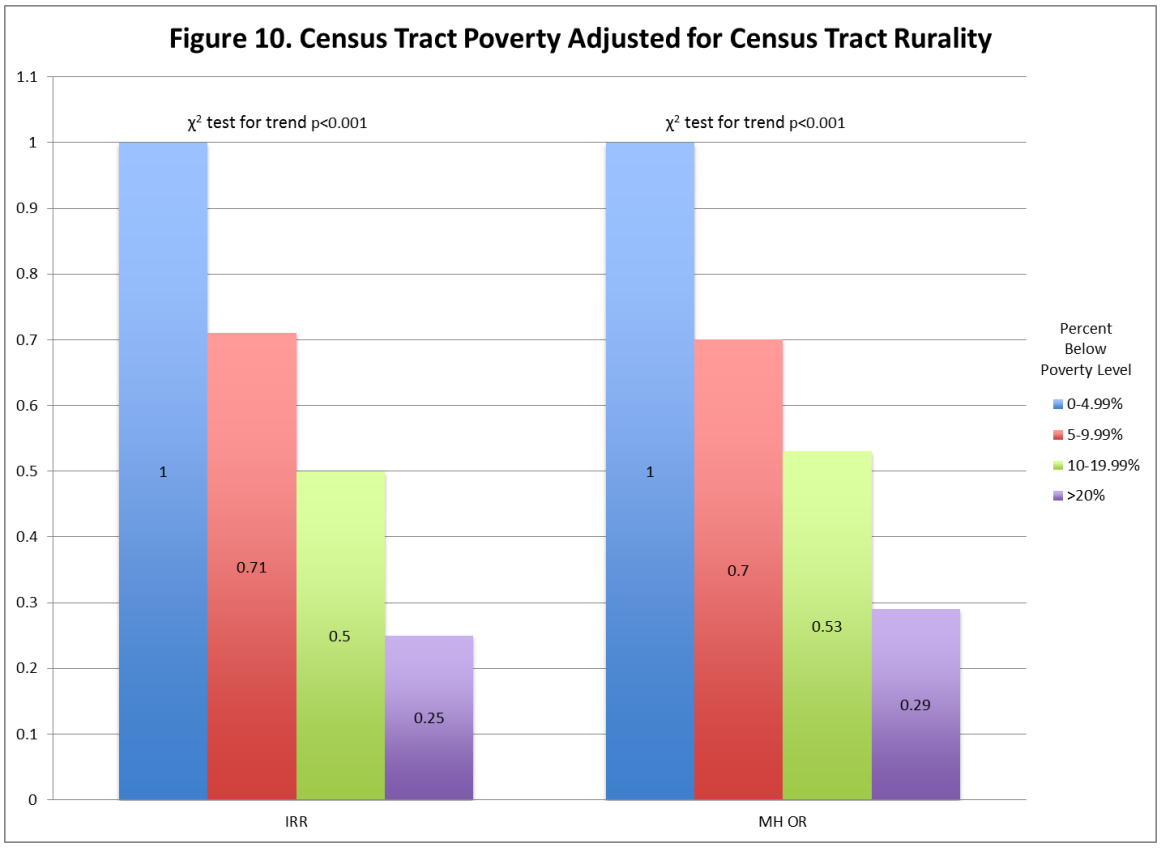


Figure 10. Census Tract Poverty Adjusted for Census Tract Rurality
χ² test for trend p<0.001
χ² test for trend p<0.001
1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1
0.71
0.5
0.25
1
0.7
0.53
0.29
IRR
MH OR
Percent Below Poverty Level
0-4.99%
5-9.99%
10-19.99%
>20%

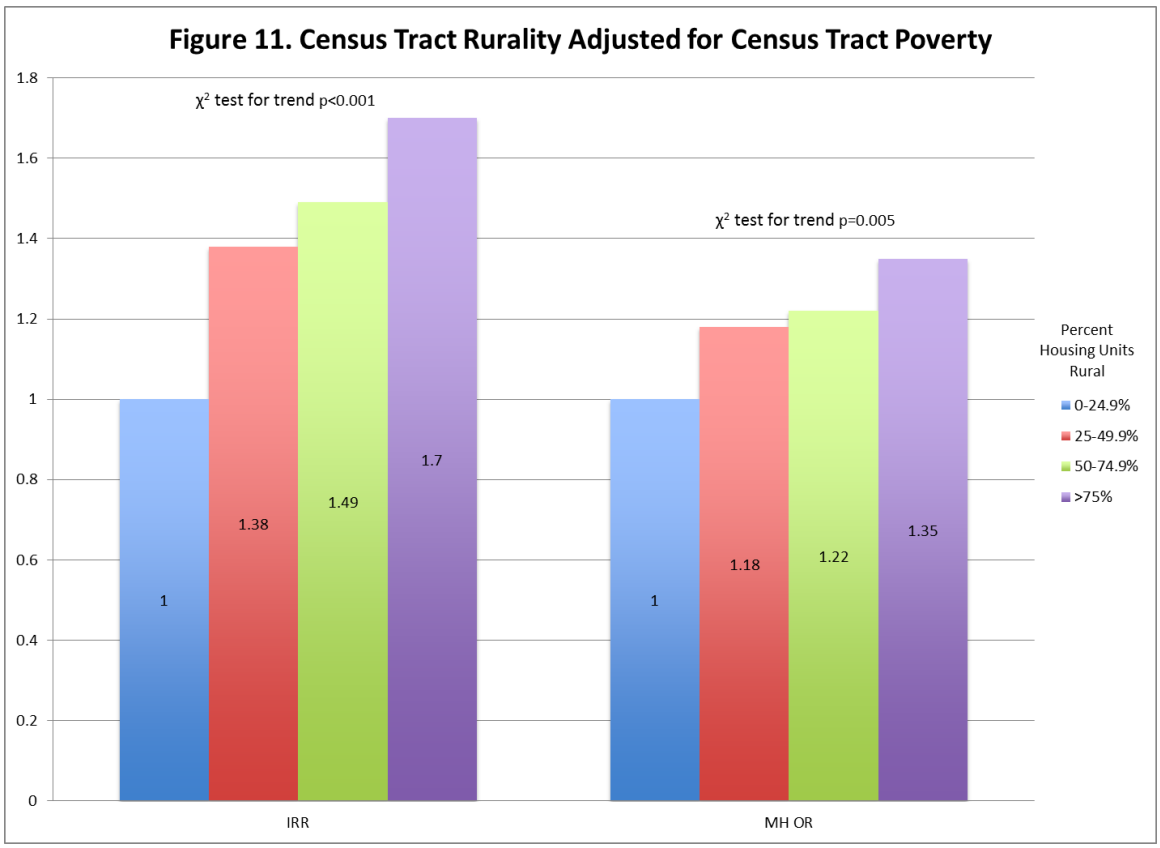


Figure 11. Census Tract Rurality Adjusted for Census Tract Poverty
χ² test for trend p<0.001
χ² test for trend p=0.005
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
1
1.38
1.49
1.7
1
1.18
1.22
1.35
IRR
MH OR
Percent Housing Units Rural
0-24.9%
25-49.9%
50-74.9%
>75%