American Journal of Infection Control 47 (2019) 773–779

Contents lists available at ScienceDirect

American Journal of Infection Control

journal homepage: www.ajicjournal.org

Major Article

Disparities in the incidence of community-acquired *Clostridioides difficile* infection: An area-based assessment of the role of social determinants in Bernalillo County, New Mexico

William B. Hudspeth PhD, MPH, MS [a,*], Fares Qeadan PhD, MS [b], Erin C. Phipps DVM, MPH [c,d]

[a] *Infectious Disease Epidemiology Bureau, Epidemiology and Response Division, New Mexico Department of Health, Santa Fe, NM*
[b] *Department of Internal Medicine, Division of Epidemiology, Biostatistics, and Preventive Medicine, University of New Mexico Health Sciences Center, Albuquerque, NM*
[c] *New Mexico Emerging Infections Program, Albuquerque, NM*
[d] *Office for Community Health, University of New Mexico Health Sciences Center, Albuquerque, NM*

Key Words:
Socioeconomic status
Census tract
Health disparities
Geocoding
Community-associated

Background: Community-associated *Clostridioides difficile* infections (CA-CDIs) share many risk factors with health care–associated cases, although the role of socioeconomic factors is poorly understood. This study estimates the influence of several census tract–level measures of socioeconomic status on CA-CDI incidence rates.

Methods: CA-CDI case data from the New Mexico Emerging Infections Program were analyzed using quasi-Poisson regression modeling. Geocoded cases were assigned census tract–level socioeconomic measures to explore racial, ethnic and socioeconomic disparities in CA-CDI incidence.

Results: Regression modeling identified census tract–level socioeconomic measures as well as individual and medical measures that together accounted for 57% of the variance in CA-CDI rates. At the census tract level, socioeconomic factors associated with an increase in CA-CDI incidence included a high percentage of individuals lacking health insurance and a low percentage of individuals with low educational attainment. A sub-analysis that included racial and ethnic designation revealed that ethnicity had no significant effect, but compared to white race, other races were significantly more likely to acquire CA-CDI.

Conclusions: Although this work reveals the role of certain socioeconomic and race and ethnicity risk factors in the incidence of CA-CDI, it also underscores the complex relationships that exist between socioeconomic status and access to health care.

Clostridioides difficile is a gram-positive, anaerobic, spore-producing bacillus that is a known human pathogen, transmitted primarily from person to person through the fecal-oral route, with patient outcomes ranging in severity from mild diarrhea to fulminant colitis and death. Important risk factors for *C difficile* infection (CDI) are recent exposure to antibiotics and health care facilities and use of histamine 2 (H2) blockers and proton pump inhibitors (PPIs).[1,2] Surveillance data indicate that *C difficile* infects nearly half a million Americans each year, accounting for some 29,000 deaths in 2011.[1] Although the majority of cases are health care–associated (HA), of growing concern is that a substantial proportion of CDIs (32%) appear to be community-associated (CA). In the United States, incidence rates for CA-CDI of between 28.2 and 79.1 per 100,000 (compared with 45.7-155.9 per 100,000 for HA-CDI) have been reported through surveillance activities.[1]

Compared with patients with HA-CDI, CA infections appear to be primarily associated with younger people (median age of 50 compared with 70 for HA-CDI), women, and individuals with fewer comorbidities, less severe infections, lessened exposure to antibiotics, lower rates of mortality, and shortened hospitalizations.[3] Some individuals with CA-CDI, however, have poorer outcomes, with higher rates of colectomy compared to HA-CDI cases.[4] Furthermore, whereas CA-CDI shares many of the same individual and medical risk factors documented for HA-CDI, less well understood are the degree to which

* Address correspondence to William B. Hudspeth, PhD, MPH, MS, Infectious Disease Epidemiology Bureau, Epidemiology and Response Division, New Mexico Department of Health, 1190 St Francis Dr, Santa Fe, NM 87505.

E-mail address: william.hudspeth@state.nm.us (W.B. Hudspeth).

Funding/support: This research was made possible by the Centers for Disease Control and Prevention's New Mexico Emerging Infections Program, Mountain West Clinical Translational Research Infrastructure Network (grant 1U54GM104944), and the University of New Mexico Clinical and Translational Science Center (grant UL1TR001449), University of New Mexico Health Sciences Center, Albuquerque, New Mexico.

Conflicts of interest: None to report.

https://doi.org/10.1016/j.ajic.2018.12.014

socioeconomic and racial and ethnic factors shape potential differences in CA-CDI rates.

The Emerging Infections Program (EIP), established by the US Centers for Disease Control and Prevention in 1995, is responsible for surveillance in 10 states, including New Mexico (NM). The NMEIP, a partnership among the NM Department of Health, the University of New Mexico, and numerous regional hospitals, long-term care facilities, outpatient clinics, and laboratories, is tasked with surveillance of infectious diseases originating in food, water, health care facilities, and the community at large. The Healthcare-Associated Infections–Community Interface is one such surveillance program for which surveillance for CDI is a critical function. Since 1995, EIP has increasingly sought to address the health equity challenges posed by Healthy People 2020.[5] The program has evolved from focusing on monitoring social determinants through the collection of race and ethnicity data to conducting population-based surveillance using socioeconomic status (SES) measures to work toward health equity. Early work using EIP data to estimate racial and ethnic disparities in invasive *Streptococcus pneumoniae* infections and invasive group B *Streptococcus* disease began in the early to mid-1990s, with findings indicating higher rates among black populations than white populations.[6] Subsequent work, beginning in the late 1990s, used geocoded case address data to link to census tract socioeconomic data. Much of this work followed the groundbreaking methods used in the Public Health Disparities Geocoding Project to assess disparities linked to high-poverty census tracts.[7] EIP data have been used to document disparities linked to SES factors in invasive *S pneumoniae* infection incidence rates,[8] influenza,[9] and campylobacteriosis.[10] Thus far, no rigorous studies have sought to investigate the role of SES factors in variation in CDI rates. The goal throughout all of these studies has been to use SES measures both to document inequities and to develop unique intervention programs to explicitly address these inequities.[5] The present research addresses a high-priority EIP objective, namely, to use geocoded case data to determine whether census tract–level socioeconomic measures can explain possible racial and ethnic and SES disparities in CA-CDI incidence rates. As research objectives, the present study aims to (1) measure variability in CA-CDI rates according to sex, race and ethnicity, age, and area-level socioeconomics; (2) identify social determinants associated with CA-CDI; and (3) determine the degree to which socioeconomics explain observed racial and ethnic differences in incidence rates.

METHODS

Study population

The NMEIP CDI surveillance activity investigates all positive *C difficile* toxin or molecular assays originating from clinical, reference, and commercial laboratories that serve the Bernalillo County, New Mexico catchment area. Incident CA-CDI cases are defined as those individuals who meet the following criteria: (1) >12 months of age and living in the surveillance catchment area at the time of diagnosis, (2) provide a positive *C difficile* stool specimen collected in an outpatient setting or within 3 calendar days of hospital admission, (3) have no documented overnight stay in a health care facility during the 12 weeks prior to specimen collection, and (4) have not received a positive test result during the prior 8 weeks. All cases determined to be CA require the completion of a case report form, which includes data on treatment information, addresses, demographics, date and location of stool collection, exposures to health care settings, medical outcomes, clinical and radiographic findings, prior medications, pregnancy status, and underlying conditions. The case address information is used to determine whether the case resides within the surveillance catchment area (ie, Bernalillo County) and for geocoding to locate the case within a specific census tract. This analysis used NM CA-CDI data that were collected by the NMEIP for the 4-year time period 2011-2014.

Geocoding and linkage to census tract–level socioeconomic data

Geocoding of all CA-CDI cases for the 2011-2014 time period was performed using Quantum GIS version 2.8 (Open Source Geospatial Foundation, Chicago, Illinois), an open source desktop software package. These cases constituted the study sample used to analyze the associations of CA-CDI incidence rates with census tract–level demographic and SES risk factors. To link to and conduct the analysis of socioeconomic factors, several variables in the original case records were essential to the calculation of incidence rates and formed the basis for constructing analytical strata. These include census tract location, age, and sex (and ethnicity and race in the substudy). Socioeconomic data were collected primarily from the NM Department of Health Indicator-Based Information System.[11] Most of these data were derived from the US Census Bureau's 2010 American Community Survey data,[12] although several study datasets were collected directly from the US Census Bureau's American FactFinder.[13] The selection of appropriate socioeconomic data were based on the desire either to represent relative SES across census tracts or to model risk factors identified in other studies as associated with CA cases of *C difficile*. The 6 census tract–level variables used as proxies for SES were (1) the percentage of individuals living below the federal poverty level, (2) the percentage of individuals without health care insurance, (3) the median household income, (4) the percentage of individuals aged 16 years and older who are unemployed, (5) the percentage of individuals >25 years of age who do not possess a high school degree or its equivalent (low educational attainment), and (6) the percentage of occupied units with at least 1.5 persons per room (crowding).

For census tract–level SES proxies, each variable was stratified into 4 levels, for which each census tract was assigned a unique categorical value. For all but census tract median income, strata levels corresponded to 0%-4.9%, 5.0%-9.9%, 10.0%-19.9%, and 20.0% or more of the population in a census tract exposed to the respective SES measure. Thus, for categorized SES variables, high SES (the category corresponding to 0%-4.9%) was the reference level. For median income, continuous values were split into quintiles, creating a new categorical variable, with breaks at $26,500, $32,000, $37,130, and $44,700.

Age of individuals was assigned to 1 of 5 different categories (ie, 0-14 years, 15-24 years, 25-44 years, 45-64 years, and 65+ years). Race and ethnicity counts were aggregated into categories that corresponded to EIP case data reporting for these variables, which is based on the 2008 US Census Bureau NM population estimates by "race alone" and the 1997 Office of Management and Budget standard. Ethnicity was limited to non-Hispanic or Latino, Hispanic or Latino, or unknown. Race categories used in this study were white, Asian or Pacific Islander, Native American or Native Alaskan, African American, and unknown.

Individual and medical variables included in the EIP case report form were also assessed as potential factors affecting CA-CDI incidence and included medication use within the last 12 weeks (ie, antibiotics, PPIs, and H2 blockers), use of diagnostic nucleic acid tests, and history of peptic ulcer.

Case and population denominator counts were stratified by unique combinations of census tract, sex, and age category. The annual incidence rate of CA-CDI in each stratum was estimated by dividing the total number of cases for that stratum by the number of residents, aggregated for each stratum over the 4-year study time period. Thus, population counts, used as denominators for each of these analytical strata, provide a person-time denominator for the calculation of mean annual incidence rates. Age-adjusted rates were also calculated using the direct method for selected combinations of independent variables using the 2000 US standard million as the reference population.[14] A separate dataset was prepared that estimated annual incidence rates of CA-CDI across strata, defined not only by census tract, sex, and age category but also by race and ethnicity.

Statistical methods

Quasi-Poisson regression was used to assess associations between census tract–level socioeconomic indicators and variability in CA-CDI incidence rates while controlling for other relevant factors (ie, age, sex, medical risk factors). The use of quasi-likelihood was imposed because of violation of the assumptive equality of the mean and variance of the Poisson regression counts. By using quasi-Poisson regression, we could address overdispersion. Although we first attempted to use negative binomial modeling, the results proved to be very similar to those achieved using quasi-Poisson modeling. We chose the latter, however, as negative binomial models tend to upweight observations with small (relative to the mean) counts compared to the quasi-Poisson.[15] Quasi-Poisson regression was used to estimate the log of the number of cases, incidence rate ratios (IRRs), and 95% confidence intervals (CIs) across the analytical strata described earlier. For each unique combination of sex and age category, case counts of CA-CDI were the dependent, response variable, whereas the various SES factors, sex, race, age category, and medical variables were the independent variables.

Inclusion of individual and medical variables in the model required translation of information from the individual to area level as defined by the census tract (CT), sex, and age class area–based strata. This was accomplished by entering into the models the raw counts of the number of cases in each stratum who were exposed to the factor of interest. Again, the census tract strata population counts were used to normalize these case counts. This study does not assume that the percent exposures reported for the cases are reflective of exposures for the entire census tract. Rather, it assumes that these risk factors make a difference if the percent exposures reported for the cases are not uniform across census tracts.

To assess the effects of each individual SES variable on CA-CDI case counts, analysis began by running a bivariate analysis using SES variables as continuous, as well as categorical and indicator, variables. The use of both variable forms was employed to provide some transparency to the reader as to why 1 form was selected for final analysis. In addition to individual SES variables, sex (categorical) and age category (as a continuous variable) were simultaneously entered into the model. For individual and medical variables, each was entered as a continuous variable (ie, number of counts per stratum). All 2-way interactions among age group, sex, and census tract measure (for both continuous and categorical versions) were then examined as potential covariates in the model. Once significance for each was assessed, model building advanced to using multiple independent variables and examined outcomes when using both the continuous versions of the respective SES variables and their indicator variable analogues. In the case when a variable was significant in both settings (categorical and continuous), we chose the continuous form because it contained more information about the full variability extant in the covariate.[16] Finally, purposeful, stepwise, backward elimination was performed to determine a final model that could best explain variation in CA-CDI rates across strata.[17] In this case, full models included continuous and categorical expressions of the variables as well as interaction terms as described earlier. Covariate terms were removed stepwise if the *P* value was >.20. Significance was defined as a 2-tailed *P* <.05. All statistical analyses were conducted using Stata 14 software (StataCorp, College Station, TX). Quasi-Poisson regression models were run using the Stata generalized linear model command.

RESULTS

Characteristics of participants

The original source dataset consisted of 1,752 case records. After geocoding, a total of 1,672 incident CA-CDI cases from 2011-2014 could be located to rooftop level (remaining cases could not be geocoded and included PO Boxes and homeless persons). Characteristics of participants are found in Table 1. Race and ethnicity data were missing for roughly half the CA-CDI cases, and as such, statistical analyses involving race and ethnicity were based on a smaller sample size (n = 756 records, 45.2% of geocoded records).

Determinants of C difficile

Tables 2 and 3 show quasi-Poisson model results for individual SES variables. All models are adjusted for sex and age category. When entered as continuous variables (Table 2), none of the SES variables—including census tract poverty percentage, median income by census tract, percentage of individuals lacking health insurance, percentage of households with individuals ≥16 years of age who are unemployed, and percentage of households with crowding—were significantly associated with census tract–level incidence of CA-CDI at the 5% level of significance.

All SES measures more directly linked with income were not associated with census tract–level CA-CDI incidence rates. However, census tract–level percentage with low educational attainment was statistically significant at a higher level of significance (α = .10). The sister model, in which SES measures were entered as categorical and indicator variables (Table 3), revealed that census tract–level measures of the percentage of individuals lacking health insurance, percentage of individuals with low educational attainment, percentage

Table 1
Demographic and medical characteristics of 1,672 cases of geocoded community-associated *Clostridioides difficile* in Bernalillo County, New Mexico (2011-2014)

Category	Subcategory	Count	%	95% CI
Year	2011	303	18.1	16.35-20.04
	2012	465	27.8	25.71-30.01
	2013	443	26.5	24.43-28.67
	2014	461	27.6	25.48-29.77
Sex	Male	577	34.5	32.27-36.82
	Female	1,095	65.5	63.18-67.73
Age category	0-14	85	5.1	4.13-6.25
	15-24	84	5.0	4.07-6.18
	25-44	282	16.9	15.14-18.74
	45-64	556	33.3	31.03-35.55
	65+	665	39.8	37.45-42.14
Ethnicity (n = 997)	Non-Hispanic	583	58.5	55.38-61.50
	Hispanic	414	41.5	38.50-44.62
Race (n = 982)	White	908	92.5	90.63-93.96
	AI/AN	31	3.2	2.23-4.57
	Asian/Pacific	15	1.5	0.09-2.52
	Black	28	2.9	1.97-4.10
Antibiotic exposure (n = 1,526)	Yes	788	51.6	49.13-54.14
	No	738	48.4	45.86-50.87
PPI use	Yes	365	21.8	19.91-23.88
	No	1,307	78.2	76.12-89.09
H2 blocker use	Yes	131	7.8	6.64-9.23
	No	1,541	92.2	90.77-93.36
Peptic ulcer disease	Yes	11	0.1	0.04-1.18
	No	1,661	99.3	98.82-99.64

NOTE. Breakdown by individual-level variables and neighborhood-level variables. Sample size in parentheses for specific variables refers to the nonmissing sample size to calculate proportions.
AI/AN, American Indian/Alaska Native; *CI*, confidence interval; *H2*, histamine 2; *PPI*, proton pump inhibitor.

Table 2
Adjusted IRRs of community-associated *Clostridioides difficile* in Bernalillo County, New Mexico, for SES variables entered as continuous variables

Category	IRR (95% CI)	*P* value
Sex	1.283 (1.152-1.427)	<.001
Median age	1.021 (1.018-1.023)	<.001
% poverty level	1.003 (0.993-1.013)	.53
% no health insurance	1.004 (0.989-1.02)	.58
% low educational attainment (less than high school diploma or equivalent)	0.99 (0.979-1.001)	.07
% individuals 16 years or older who are unemployed	1.004 (0.987-1.022)	.65
% households with children under the age of 18	0.998 (0.99-1.005)	.56
% households with adults aged 65 or older	1.003 (0.994-1.013)	.46
% households with an average of ≥1.5 persons per room	1.009 (0.967-1.052)	.69
Median income	1.000 (1.000-1.000)	.52

NOTE. IRR values adjusted for both sex and age class (2011-2014).
CI, confidence interval; *IRR*, incidence rate ratio; *SES*, socioeconomic status.

of households with an average of ≥1.5 persons per room, and median income were all significant at the $\alpha = .05$ level. We note that census tract–level percentage of households with adults ≥65 years of age was significant, although at a higher level ($\alpha = .10$).

The full model (Table 4), containing some SES factors represented as continuous variables and some as categorical and indicator variables, explains approximately 59.55% of the variance in *C difficile* incidence rates across census tract–level strata defined by age and sex (as measured by R^2). Table 4 lists IRRs for the full model and provides 95% CIs and *P* values for continuous and categorical and indicator versions of the main covariates. Women exhibited incidence rates 28% higher than men (IRR = 1.282; 95% CI, 1.134-1.448); however, this rate depended significantly on the frequency of previous antibiotic use. In particular, for a 1-count increase in census tract–level previous antibiotic use, women were 16.49% more likely to experience CDI compared to men. For the main effect SES variables, census tract–level lack of health insurance and percentage of low educational attainment were significant predictors of CA-CDI rates. For those living in census tracts with high percentages (>20%) of individuals lacking health insurance, there was a significant, nearly dose-dependent increase in IRR values (IRR = 1.718; 95% CI, 1.40-2.108) relative to those living in census tracts with low percentages (<5%) of individuals lacking health insurance. Alternatively, there was a dose-dependent decrease in *C difficile* incidence as the percentage of individuals in a census tract with less than a high school education increased (IRR = 0.974; 95% CI, 0.959-0.989). However, this decrease depended significantly on the median age of individuals in the census tract. Specifically, for a 1% increase in low educational attainment, census tracts with a median age of 65+ were 2-fold more likely to experience *C difficile* incidence relative to those with a median age of 0-14 years. Finally, all 5 of the individual medical covariates were found to be risk factors for CA-CDI. Model fit was determined by assessing the distribution of the deviance residuals with normal plots of residuals, quantile plots of variables against a normal distribution, two-way scatter plots of residuals and expected values, and box plots of the residuals themselves. Plots and charts can be found in Appendix A.

Table 3
Adjusted IRRs of community-associated *Clostridioides difficile* in Bernalillo County, New Mexico, for SES variables entered as categorical variables

Covariate	Subcategory	IRR (95% CI)	*P* value
Sex	Male	Reference	
	Female	1.283 (1.156-1.422)	<.001
Age	Age	1.021 (1.018-1.023)	<.001
% poverty level	0-4.9	Reference	
	5-9.9	0.896 (0.738-1.087)	.263
	10-19.9	0.932 (0.757-1.148)	.507
	>20	1.084 (0.826-1.422)	.563
% no health insurance	0-4.9	Reference	
	5-9.9	1.438 (1.147-1.804)	.002
	10-19.9	1.302 (1.014-1.672)	.039
	>20	1.469 (1.038-2.078)	.030
% low educational attainment (less than high school diploma or equivalent)	0-4.9	Reference	
	5-9.9	0.993 (0.859-1.148)	.920
	10-19.9	0.816 (0.661-1.007)	.058
	>20	0.756 (0.58-0.984)	.037
% individuals 16 years or older who are unemployed	0-4.9	Reference	
	5-9.9	1.011 (0.864-1.183)	.890
	10-19.9	1.001 (0.823-1.218)	.992
% households with children under the age of 18	5-9.9	Reference	
	10-19.9	0.652 (0.354-1.199)	.169
	>20	0.670 (0.365-1.233)	.198
% households with adults aged 65 or older	0-4.9	Reference	
	5-9.9	0.265 (0.061-1.156)	.077
	10-19.9	0.253 (0.059-1.081)	.064
	>20	0.295 (0.069-1.253)	.098
% households with an average of ≥1.5 persons per room	0-4.9	Reference	
	5-9.9	1.651 (1.195-2.28)	.002
Median income	>$44,700	Reference	
	$37,130-$44,700	1.232 (1.028-1.477)	.024
	$32,000-$37,130	1.147 (0.909-1.448)	.248
	$26,500-$32,000	1.225 (0.942-1.594)	.130
	<$26,500	0.898 (0.649-1.242)	.514

NOTE. IRR values adjusted for both sex and age class (2011-2014).
CI, confidence interval; *IRR*, incidence rate ratio; *SES*, socioeconomic status.

Race and ethnicity subanalysis

Because more than half of CA-CDI cases lacked both race and ethnicity information, the racial and ethnic subanalysis dataset was considerably smaller and lacked the statistical power of the full dataset. A total of 756 cases were aggregated into strata defined by sex, age category, and ethnicity and race. In this case, a smaller subset of SES and individual covariates was significantly associated with variance in census tract–level CA-CDI incidence rates. Of interest is that self-reported ethnicity never approached statistical significance in any of the models constructed.

Table 5 presents the results for the final quasi-Poisson regression model and serves to illustrate the role of race in modeling IRRs for different levels of the covariates. As can be seen, none of the main effect SES variables reached statistical significance. There was, however, a significant interaction between sex and race. Specifically, at the census tract level, black women were 5.98 times more likely to experience CDI relative to white women (95% CI, 3.379-10.567). Again, at the census tract level, relative to white men, American Indian and Native Alaskan men exhibited an IRR that was more than 11 times as high (IRR = 11.657; 95% CI, 5.781-23.507), whereas the IRR for African American men was 27.681 (95% CI, 12.931-59.256) times higher than the IRR for white men. The full model was able to explain only 36.43% of the variance in *C difficile* incidence rates across analysis strata.

DISCUSSION

Although there have been a number of studies that have addressed racial and ethnic disparities in CDI rates,[18,19] with 1 study implicating disparities in access to health care,[20] the present research, to the best of the authors' knowledge, is the first to examine the relationship between census tract–level SES factors and incidence of CA-CDI. The incidence of CA-CDI at the census tract level was found to vary as a function of specific demographic, individual-level, and socioeconomic factors. These findings are relevant for the practice of public health as they relate to assessing interventions that can be addressed within the context of inequities in socioeconomic conditions and health care access in the United States.

Table 4
Final model for adjusted IRRs of community-associated *Clostridioides difficile* in Bernalillo County, New Mexico (2011-2014)

Covariate	Subcategory	IRR (95% CI)	*P* value
Sex	Male	Reference	
	Female	1.282 (1.134-1.448)	<.001
Age	0-14 years	Reference	
	15-24 years	1.472 (1.011-2.145)	.044
	25-44 years	0.801 (0.591-1.084)	.150
	45-64 years	0.968 (0.726-1.291)	.827
	≥65 years	1.952 (1.463-2.604)	<.001
% no health insurance	0-4.9%	Reference	
	5-9.9%	1.582 (1.352-1.852)	<.001
	10-19.9%	1.492 (1.282-1.737)	<.001
	>20%	1.718 (1.4-2.108)	<.001
% low educational attainment(less than high school diplomaor equivalent)		0.974 (0.959-0.989)	.001
% households with adults aged 65 or older		1.003 (0.998-1.008)	.268
Crowding (% households with morethan 1.5 persons per room)		1.135 (0.911-1.412)	.259
Previous antibiotic use		1.276 (1.193-1.364)	<.001
PPI		1.07 (1.028-1.113)	.001
H2 blocker		1.106 (1.023-1.195)	.011
Peptic ulcer comorbidity	No	Reference	
	Yes	1.572 (1.22-2.026)	<.001
Diagnostic NAT		1.087 (1.026-1.152)	.005
Interaction sex and antibiotic use	Female × antibiotic use Male	1.165 (1.071-1.267) per 1-count increase	<.001
Interaction age and low educational attainment	15-24 years × low educ 0-14 years	1.465 (1.024-2.096) per 1% increase in low educ	.0370
	≥65 years × low educ 0-14 years	1.994 (1.512-2.628) per 1% increase in low educ	<.001
	≥65 years × low educ 15-24 years	1.361 (1.034-1.790) per 1% increase in low educ	.0280
	≥65 years × low educ 25-44 years	2.481 (2.088-2.947) per 1% increase in low educ	<.001
	≥65 years × low educ 45-64 years	2.019 (1.767-2.307) per 1% increase in low educ	<.001

NOTE. Model R^2 = 59.55%.
CI, confidence interval; *educ*, education; *H2*, histamine 2; *IRR*, incidence rate ratio; *NAT*, nucleic acid test; *PPI*, proton pump inhibitor.

The link between poverty and high incidence rates of infectious diseases has been well supported for many decades, with the recognition that poverty is the single greatest risk factor for acquiring and succumbing to disease worldwide.[21] Health care workers have remarked on the poor's heightened vulnerability to communicable diseases and a general lack of medical care once they have been infected.[21] Low educational attainment is frequently cited as 1 of numerous components of SES tied to poverty.[22] It has been generally accepted that education is a critical ingredient of health because of the opportunities it creates to obtain better material living conditions.[22] In this regard, the results of the present analysis are somewhat anomalous. Although some of the more direct measures of poverty, including poverty percentage and median income, were not significantly associated with *C difficile* incidence rates at the census tract level in the quasi-Poisson modeling, the seemingly protective effects of living in neighborhoods with higher numbers of individuals with low educational attainment are not congruent with the widely observed association between poverty and burden of infectious disease. One may suggest that those with low education may not actively seek health care, thus actually lowering their potential exposures to CA-CDI and other infectious diseases typically associated with health care exposures. This has been attributed to the fact that low SES groups tend to have lower screening rates and may possibly resort to home remedies.[23] As such, case counts, as represented in the surveillance data, may underestimate the true incidence of CA-CDI. However, the fact that higher *C difficile* incidence rates are simultaneously and positively associated with the relative proportion of those who are medically uninsured falls within the existing normative paradigm. Herein lies the paradox, however, in that CDI, whether diagnosed as HA or CA, is commonly linked to health care exposure of 1 kind or another (ie, hospitals, outpatient clinics, long-term care facilities). Under this scenario, one might expect that access to health insurance and health facilities might constitute a risk factor for infection with *C difficile*. The association of higher incidence rates of CA-CDI among those who lack health insurance, as seen in this study, may indicate alternative health care–seeking behaviors in which individuals may be exposed to different kinds of health care facilities (ie, preference for treatment in emergency rooms where exposures may be higher) in which the risk of becoming infected with CA-CDI may be higher. The finding that both percentage of households with individuals ≥65 years of age and percentage of households with an average of ≥1.5 or more persons per room appear to be positively, although not significantly, associated with CDI rates is in keeping with expectations about risk factors for susceptibility to infectious disease.

Although the full quasi-Poisson regression model developed here accounts for a significant proportion of the variance in CA-CDI rates, the remaining, unexplained component in infection rate variance suggests that census tract–level analyses mask considerable within-tract variation in individual exposures. Future research questions should focus more closely on the complex, and interrelated factors that surround CA-CDI incidence rates. Earlier, it was discussed how poverty and lack of health insurance at the individual level may shape health-seeking behaviors and may revolve around whether individuals seek treatment at all, make use of specific categories of facilities where *C difficile* transmission is more likely, or are restricted from accessing health care facilities because of transportation limitations or other reasons related to geography. An examination of differences in health care facility use might provide important insights into geographic variation in CA-CDI incidence rates. The findings of this study also suggest complex relationships between race and ethnicity and SES factors—a topic of considerable recent research.[24] Efforts should be made to more directly assess how different racial and ethnic groups experience CA-CDI rates when they are embedded in variable socioeconomic contexts. Research should focus more explicitly on

Table 5
Adjusted IRRs of community-associated *Clostridioides difficile* in Bernalillo County, New Mexico, for SES and individual and medical covariates when including self-identified race

Covariate	Subcategory	IRR (95% CI)	*P* value
Sex	Men	Reference	
	Women	0.875 (0.765-1.002)	.053
Age		1.008 (1.005-1.012)	<.001
Race	White	Reference	
	AI/AN	11.657 (5.781-23.507)	<.001
	Asian/Pacific	46.859 (8.673-253.176)	<.001
	Black	27.681 (12.931-59.256)	<.001
% no health insurance		1.012 (0.999-1.024)	.071
% low educational attainment (less than high school diploma or equivalent)		0.992 (0.981-1.002)	.122
Previous antibiotic use		1.111 (1.032-1.197)	.005
PPI		1.123 (1.051-1.201)	.001
Diagnostic NAT		1.208 (1.1-1.326)	<.001
Interaction sex × race	Black women	5.976 (3.379-10.567)	<.001
	White women	Reference	
	Asian women	9.353 (5.444-16.067)	<.001
	White women	Reference	
	Black men	27.681 (12.931-59.256)	<.001
	White men	Reference	
	White men	1.142 (0.998-1.307)	.053
	White women	Reference	
	AI/AN men	11.657 (5.781-23.507)	<.001
	White men	Reference	
	AI/AN women	2.048 (0.999-4.200)	.050
	Black women	Reference	

NOTE. IRR values adjusted for census tract, sex, age class, and race (2011-2014). Model $R^2 = 36.43\%$.
AI/AN, American Indian/Alaska Native; *CI*, confidence interval; *IRR*, incidence rate ratio; *NAT*, nucleic acid test; *PPI*, proton pump inhibitor; *SES*, socioeconomic status.

differences in the way different racial and ethnic groups and individuals of different SES make decisions about health care use. Moreover, variation in health experience may also entail differences in treatment or diagnostic tests. For example, variation in the use of specific diagnostic tests may affect census tract–level rates of CA-CDI. The use of nucleic acid tests for CDI testing has been demonstrated to result in higher sensitivities than is the case for traditional toxin enzyme immunoassays used by many clinical laboratories, which are known to have a low sensitivity.[25] An approach that examines all of these factors provides an opportunity for community intercessions that seek to document, analyze, and develop interventions to address CA-CDI in communities at risk.

Several important limitations can be stipulated for this study. Census tract–level SES data collected from the American Community Survey, as well as analogous data collected from the NM Indicator-Based Information System, are ultimately derived from a sample of US—or NM state—populations, thus allowing for the possible misclassification of a given census tract with regard to SES level. The result of misclassification would be a trend toward the null hypothesis, with a consequent reduction in the power to discriminate SES differences between neighborhoods. SES data used in this analysis were derived from samples of individual census tracts, meaning that generalizations drawn about census tracts cannot be used to represent individual-level situations. At an interpretive level, this means that census tract SES data are used to assess the effects of neighborhoods on CDI risk, rather than the use of SES measures as proxies for individual-level SES. Even within a given census tract, neighborhoods of differing income and social disadvantage exist, which effectively acts to create additional misclassification at the census tract level.

Another important limitation of this study is the large number of case records that lacked race and ethnicity information. Although race was a highly significant predictor of variation in *C difficile* incidence rates at the census tract level, it is possible that underreporting may have introduced a certain degree of selection bias. The extraordinarily high IRR values seen for black individuals and Asian individuals are a case in point and are likely owing to the very small sample sizes observed for these racial groups.

The most important strength of this analysis is the fact that the NMEIP active surveillance of CDI has a high level of case ascertainment that is accomplished through participation of all clinical laboratories in and beyond the county to the state's reporting requirements. This limits selection bias and strengthens the internal validity of the study. Through active case finding, the case counts represented here are likely to be relatively complete. Finally, the relationships observed among various SES measures used in this study and CDI rates may potentially be generalizable for parts of the United States outside of Bernalillo County, NM.

Elucidation of the factors affecting infection with CA-CDI is essential, as this disease imposes an increasingly heavy burden on public health that extends outside of its putative associations with health care facility exposures. Recent studies have estimated the economic burden associated with CA-CDI and found an average attributable cost of $2,454-$29,000 per episode, with total annual costs in the United States surpassing $4.8 billion in excess medical costs.[26] Given that CA-CDI incidence rates are increasing both nationally and globally, it can be assumed that reported costs are likely underestimates and will likely increase going forward.[26] Although there is clearly a need for both primary and secondary prevention of CDI, including but not limited to eliminating the unnecessary use of antibiotics, curbing the prescription and use of PPIs and H2 blockers, and instituting better preventive precautions in health care facilities, more research is necessary to understand the linkages between CA-CDI rates and SES disparities as they relate to community-level risks. At present, there is only limited understanding of how possession of health insurance and access to health care may affect rates and opportunities for infection or to what degree SES may in some cases protect against CA-CDI. Future work in these areas should serve to reduce health disparities in CA-CDI.

Finally, given that previous studies have indicated that <25% of new, hospital-based cases can be attributed to transmission through potential ward-based inpatient sources, and that genetic studies indicate a diverse spectrum of alternative sources (potentially originating from both health care–related and environmental sources),[27] this work highlights the need for providers to consider interventions for CDIs that are transmitted within the community but have the potential to be introduced into health care facilities.

CONCLUSIONS

The risk of CA-CDI was found to be higher in those census tracts with higher percentages of individuals lacking health insurance coverage and those with low percentages of individuals with low educational attainment. Although not significant, percentage of households with individuals ≥65 years of age and percentage of households with an average of ≥1.5 or more people per room were also positive predictors of *C difficile* incidence rates. Thus, although some SES predictors of *C difficile* incidence rates did not reach significance, those that did appear to point in opposite directions with respect to commonly held assumptions about the relationship between poverty and incidence of infectious disease. In addition, as observed in earlier studies 1, a number of individual- and medical-level covariates were found to be positively associated with an increase in the incidence rate of CDI, including use of PPIs, H2 blockers, and antibiotics and presence of peptic ulcer.

Acknowledgments

We would particularly like to thank surveillance officers Emily Hancock and Nicole Kenslow from the New Mexico Emerging Infections Program for their data collection efforts.

SUPPLEMENTARY MATERIALS

Supplementary material associated with this article can be found in the online version at https://doi.org/10.1016/j.ajic.2018.12.014.

References

1. Lessa FC, Mu Y, Winston LG, Dumyati GK, Farley MM, Beldavs ZG, et al. Determinants of *Clostridium difficile* infection incidence across diverse United States geographic locations. Open Forum Infect Dis 2014;1:ofu048.
2. Carvour ML, Wilder SL, Ryan KL, Walraven C, Qeadan F, Brett M, et al. Predictors of *Clostridium difficile* infection and predictive impact of probiotic use in a diverse hospital-wide cohort. Am J Infect Control 2019;47:2-8.
3. Khanna S, Pardi DS, Aronson SL, Kammer PP, Orenstein R, St Sauver JL, et al. The epidemiology of community-acquired *Clostridium difficile* infection: a population-based study. Am J Gastroenterol 2012;107:89-95.
4. Kasper AM, Nyazee HA, Yokoe DS, Mayer J, Mangino JE, Khan YM, et al. A multicenter study of *Clostridium difficile* infection-related colectomy, 2000-2006. Infect Control Hosp Epidemiol 2012;33:470-6.
5. Hadler JL, Vugia DJ, Bennett NM, Moore MR. Emerging infections program efforts to address health equity. Emerg Infect Dis 2015;21:1589-94.
6. Robinson KA, Baughman W, Rothrock G, Barrett NL, Pass M, Lexau C, et al. Epidemiology of invasive *Streptococcus pneumoniae* infections in the United States, 1995-1998: opportunities for prevention in the conjugate vaccine era. JAMA 2001;285:1729-35.
7. 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: 312-23.
8. Burton DC, Flannery B, Bennett NM, Farley MM, Gershman K, Harrison LH, et al. Socioeconomic and racial/ethnic disparities in the incidence of bacteremic pneumonia among US adults. Am J Public Health 2010;100:1904-11.
9. Yousey-Hindes KM, Hadler JL. Neighborhood socioeconomic status and influenza hospitalizations among children: New Haven County, Connecticut, 2003-2010. Am J Public Health 2011;101:1785-9.
10. Bemis K, Marcus R, Hadler JL. Socioeconomic status and campylobacteriosis, Connecticut, USA, 1999-2009. Emerg Infect Dis 2014;20:1240-2.
11. NM Department of Health Indicator-Based Information System. Santa Fe (NM): New Mexico Department of Health - c2017. Available from: https://ibis.health.state.nm.us/. Accessed July 12, 2017.
12. census.gov. Washington, D.C: United States Census Bureau; c2017. Available from: https://www.census.gov/programs-surveys/acs/. Accessed February 22, 2017.
13. census.gov. Washington, D.C: United States Census Bureau; c2016. Available from: http://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml. Accessed December 2, 2016.
14. cancer.gov. Bethesda (MD): National Cancer Institute; c2016. Available from: http://seer.cancer.gov/stdpopulations/. Accessed November 15, 2016.
15. Ver Hoef JM, Boveng PL. Quasi–Poisson vs. negative binomial regression: how should we model overdispersed count data? Ecology 2007;88:2766-72.
16. Altman DG, Royston P. The cost of dichotomising continuous variables. BMJ 2006;332:1080.
17. Licht MH. Multiple regression and correlation. In: Grimm LG, Yarnold PR, editors. Reading and understanding multivariate statistics. Washington (DC): American Psychological Association; 1997. p. 53.
18. Yang S, Rider BB, Baeher A, Ducoffe AR, Hu DJ. Racial and ethnic disparities in health care–associated *Clostridium difficile* infections in the United States: state of the science. Am J Infect Control 2016;44:91-6.
19. Argamany JR, Delgado A, Reveles KR. *Clostridium difficile* infection health disparities by race among hospitalized adults in the United States, 2001 to 2010. BMC Infect Dis 2016;16:454.
20. Mao EJ, Colleen RK, Machan JT. Racial differences in *Clostridium difficile* infection rates are attributable to disparities in health care access. Antimicrob Agents Chemother 2015;59:6283-7.
21. Alsan MM, Westerhaus M, Herce M, Nakashima K, Farmer PE. Poverty, global health, and infectious disease: lessons from Haiti and Rwanda. Infect Dis Clin North Am 2011;25:611-22.
22. Regidor E, De Mateo S, Calle ME, Domínguez V. Educational level and mortality from infectious diseases. J Epidemiol Community Health 2002;56:682-3.
23. Fiscella K, Humiston S, Hendren S, Winters P, Jean-Pierre P, Idris A, et al. Eliminating disparities in cancer screening and follow-up of abnormal results: what will it take? J Health Care Poor Underserved 2011;22:83-100.
24. Williams DR, Priest N, Anderson NB. Understanding associations between race, socioeconomic status and health: patterns and prospects. Health Psychol 2016;35: 407-11.
25. Cohen J, Limbago B, Dumyati G, Holzbauer S, Johnston H, Perlmutter R, et al. Impact of changes in *Clostridium difficile* testing practices on stool rejection policies and *C difficile* positivity rates across multiple laboratories in the United States. J Clin Microbiol 2014;52:632-4.
26. Evans CT, Safdar N. Current trends in the epidemiology and outcomes of *Clostridium difficile* infection. Clin Infect Dis 2015;60(Suppl):66-71.
27. Eyre DW, Cule ML, Wilson DJ, Griffiths D, Vaughan A, O'Connor L, et al. Diverse sources of *C difficile* infection identified on whole-genome sequencing. N Engl J Med 2013;369:1195-205.