Distribution Agreement

Signature:

______________________________ ______________

Claire Myers Date

Uneven Burden: A Multivariate Analysis of Differences in Sociodemographic, Comorbid, Risk, and Case Management Factors in Tuberculosis Cases in Rural vs Urban, U.S. from 2014-2019

By

Claire Myers

Master of Public Health

Global Epidemiology

Dr. Kenneth Castro, M.D.

Committee Chair

Uneven Burden: A Multivariate Analysis of Differences in Sociodemographic, Comorbid, Risk, and Case Management Factors in Tuberculosis Cases in Rural vs Urban, U.S. from 2014-2019

By

Claire Myers

B.A., Brown University, 2022

Dr. Kenneth Castro, M.D.

Committee Chair

An abstract of

A thesis submitted to the Faculty of the Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of Master of Public Health in Global Epidemiology

2025

Abstract

Uneven Burden: A Multivariate Analysis of Differences in Sociodemographic, Comorbid, Risk, and Case Management Factors in Tuberculosis Cases in Rural vs Urban, U.S. from 2014-2019

By Claire Myers

Background

Although tuberculosis (TB) occurs less frequently in the U.S. than globally, it remains a significant public health concern. Limited research has explored differences between individuals with TB living in rural versus urban areas of the U.S. This study addresses that gap by examining patient characteristics, risk factors, comorbidities, and case management practices among individuals reported with TB from rural and urban states, in 2014–2019.

Methods

We used publicly available data from the CDC Wonder Online Tuberculosis Information System (OTIS) to acquire data from 14 states, categorized as urban or rural, during 2014-2019. Multivariate logistic regression analyses were conducted to assess differences in demographics, TB disease site, HIV status, treatment completion within one year, and receipt of directly observed therapy (DOT).

Results

Compared to individuals in urban states, those with TB in rural states were more likely to be male (OR = 1.217, CI: 1.089–1.359); aged 25–44 (OR = 2.315, CI: 1.872–2.863), 45–64 (OR = 1.539, CI: 1.290–1.968), or 65+ (OR = 1.784, CI: 1.432–2.222); have pulmonary TB (OR = 1.388, CI: 1.235–1.561); complete therapy within a year (OR = 1.143, CI: 0.928–1.409); and receive DOT (OR = 1.301, CI: 1.163–1.454). They were less likely to be non-U.S.-born (OR = 0.335, CI: 0.293–0.382); Asian (OR = 0.207, CI: 0.171–0.252), Black or African American (OR = 0.990, CI: 0.874–1.121), or Hispanic/Latino (OR = 0.132, CI: 0.110–0.159); and less likely to be HIV-positive (OR = 0.177, CI: 0.125–0.252).

Discussion

These findings suggest that there are differences in individuals who live in rural areas who have TB compared those who live in urban areas. TB is a challenging and time-consuming disease to treat, especially for those in rural communities, and these differences allow us to better tailor treatment and resources to meet the needs of individuals battling TB most effectively in rural and urban communities.

Uneven Burden: A Multivariate Analysis of Differences in Sociodemographic, Comorbid, Risk, and Case Management Factors in Tuberculosis Cases in Rural vs Urban, U.S. from 2014-2019

By

Claire Myers

B.A., Brown University, 2022

Dr. Kenneth Castro, M.D.

Committee Chair

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of Master of Public Health in Global Epidemiology

2025

Introduction

Tuberculosis Background

In 2023, there were 9,615 cases of tuberculosis (TB) disease in the United States.[1] While the TB incidence in the U.S. is quite low, compared to the global incidence, TB disease continues to have a remarkable impact on the lives of affected individuals and the people around them.

TB disease is caused by *Mycobacterium tuberculosis*. These bacteria can result in asymptomatic infection or disease -- affecting many areas of the body and cause a plethora of debilitating symptoms. TB disease is most commonly found in the lungs and is known as pulmonary TB.[2] Untreated active pulmonary TB can spread from person-to-person through droplet nuclei containing bacteria and suspended in air. Transmission usually occurs in poorly ventilated and crowded spaces. While not spread through the air, extrapulmonary TB can occur almost anywhere in the body. Symptoms of TB are burdensome and debilitating, and include prolonged cough, fever, night sweats, weight loss, and immense pain and discomfort. People may lose the ability to work, run basic errands, and properly care for themself.[3] Untreated, this illness can result in a domino effect of job loss, income loss, homelessness, and the inability to attend to basic needs.

In addition to the challenges caused by symptoms, treatment of TB disease is intense and time consuming. Treatment of TB disease can range from 4 months to 9 months with multiple recommended effective drugs -- or longer depending on complications and patient adherence with their treatment regimen.[4] Comorbidities like HIV, diabetes, and chronic obstructive lung disease make disease treatment more complicated, as health professionals must consider interactions between diverse medications. Antimicrobial resistance is becoming an increasingly prevalent issue with all bacterial infections, especially TB. This may also affect the choice of treatment options. Lack of adherence to treatment can increase the risk of drug resistance in people with TB disease, potentially increasing the length of treatment.

TB in the U.S.

The geographic distribution of reported TB cases in the U.S. are concentrated in large urban areas of a few states, such as California, Florida, New York, and Texas.[5] Combined, these four states accounted for slightly over half (50.6%) of all reported TB cases in the U.S. in 2023. There are publications describing the risk factors and comorbidities of TB in these large U.S. urban areas. In contrast, there are few published reports describing details about the occurrence of TB in the rural U.S. Available reports include publications that examine TB outbreaks that occurred in rural communities. These publications are mostly case reports discussing the patient and contact investigations and the challenges and outcomes of the outbreak.[6,7]

Other published reports have assessed different risk factors and comorbidities of TB in the U.S. One of these describes the differences in the burden of TB among U.S. born and non-U.S. born individuals.[8] The authors found that the incidence of TB was consistently higher in non-U.S. born individuals. The authors discuss that incident TB cases among non-U.S. born individuals are likely the reactivation of infections acquired outside of the U.S. Incident cases among U.S.-born individuals are more likely to have been acquired through local disease outbreaks. Another publication assessed changes in epidemiologic trends in TB in the U.S. over time.[9] This manuscript found that between the years 2014 and 2017, non-U.S. born individuals were more likely to have diabetes than in previous years. Non-U.S. born persons were also less likely to be HIV positive than in previous years. One study assessed risk factors for TB multidrug resistance in the US.[10] This study did not find that HIV status, alcohol use, or homelessness were significantly associated with having multidrug resistant disease. However, they did find that previous TB disease was a strong risk factor for multidrug-resistant TB.

Most studies in the U.S. assess differences in TB in U.S.-born compared to non-U.S. born individuals. Other recent studies that explore associations between TB and risk factors or comorbidities mostly take place in low- and middle-income countries. The authors of a meta-analysis of 23 cohort studies from around the world found evidence of a two to four times increase in the risk of developing TB disease among individuals with diabetes compared to those without.[11] Another systematic review and meta-analysis found that HIV was associated with an increased risk of multidrug resistant TB.[12] It is important to continue to study TB in the U.S. to generate knowledge that can help us work towards elimination of the disease.

Study Objectives

While many studies examine intrapersonal risk factors and comorbidities associated with TB, little is known about how these relationships differ between rural and urban populations in the U.S. Prior to this study, we performed 23 key informant interviews with public health professionals working in rural health departments across the U.S. to discuss their experiences, challenges, and successes with active TB case management in their communities.[3] These local health departments may report, at most, 2 cases of active TB a year. In addition, these rural communities often have one or two staff members managing all communicable diseases and have few resources for active TB case management. From these interviews, we learned that some of the greatest barriers to management of TB in local, rural health departments include a lack of sufficient staff and resources, access to health facilities equipped to properly diagnose and treat

TB, lack of access to patient transportation to and from appointments, long commuting distances for public health professionals to provide directly observed therapy, lack of technology to provide resources for video directly observed therapy, and language barriers. Proper management of people with TB disease is labor-intensive, and requires resources often absent, or limited in local, rural health departments. Gaining a stronger understanding of the characteristics of people with TB disease living in rural areas could help these health departments tailor their resources to increase the efficiency and capability for the effective management of people with active TB. This study aims to explore the relationships between common TB risk factors, comorbidities, and case management related activities among people reported with TB who live in rural states compared to those who live in urban states of the U.S. between the years 2014-2019, prior to the disruptions associated with the COVID-19 pandemic. We explored relationships between sociodemographic variables, HIV status, TB disease site, completion of TB therapy within a year, and directly observed therapy among individuals with TB living in rural states compared to those living in urban states.

Methods

Data Source

Data for this analysis were obtained using the Centers for Disease Control and Prevention's (CDC) Online Tuberculosis Information System (OTIS).[13] OTIS is a CDC Wonder database containing information on TB cases that were verified and reported by cities or states to CDC's National TB Surveillance System (NTSS). This database includes information for 27 TB variables for cases reported from state health departments from all U.S. states, Washington D.C., and Puerto Rico from 1933-2023.

The OTIS CDC Wonder (Wide-ranging Online Data for Epidemiologic Research) database is a free, open-access database meant for use of data by public health and the public for research, decision making, education, and other public health activities.[14] The OTIS database provides data on TB case counts, grouped by a maximum of five variables. All data is deidentified, and any combination of grouped variables with a case count less than five is suppressed to protect patient information. For this analysis, data was extracted by exporting data for each component of the analysis individually by grouped variables and case counts. As this database is publicly accessible and does not contain personally identifiable information, no human subjects review was required.

Study Population

The study population for this analysis includes individuals living in the U.S. and Puerto Rico who have a verified case of tuberculosis reported to the NTSS and included in the OTIS database during 2014-2019. We excluded from our analysis the years after 2020 as the COVID-19 pandemic caused several disruptions to TB disease reporting, treatment, and management.[15] Case counts were extracted from Alabama, Arkansas, California, Florida, Kentucky, Maine, Mississippi, Montana, New York, North Dakota, South Dakota, Texas, Vermont, and West Virginia. States were selected based on the definition of the outcome variable described below.

Variables

The outcome measured in this analysis is whether the TB cases are from an urban designated state or rural designated state. Urban states were classified as those few U.S. states that collectively contributed at least 50% of the burden of TB in the country.[16] These states included California, Florida, New York, and Texas.[5] Rural states were classified as states where at least 40% of the total population resides in rural areas, as designated by the 2010 United States Census Bureau.[17] These states included Maine and Vermont, with 60-79.9% of the state's population living in a rural area, and Alabama, Arkansas, Mississippi, Montana, Kentucky, North Dakota, South Dakota, and West Virginia, with 40-59.9% of the state's population living in rural areas.

Sex, race/ethnicity, origin of birth, and age variables were analyzed to assess differences in the demographic composition of the individuals who have TB among those who live in urban states compared to those who live in rural states. Risk factors, comorbidities, and case management variables were also assessed to explore differences in these variables by state classification. The risk factors and comorbidities assessed include disease site and HIV status. The case management variables that were analyzed were completion of therapy within a year and type of directly observed therapy (DOT).

Data Analysis

Data cleaning and data analysis were completed using SAS version 9.4. Multivariate analyses compared TB cases reported from urban and rural states. Logistic regression models were run to obtain odds ratios and 95% confidence intervals (C.I.s) for each identified exposure using residence in a rural state as the desired outcome. We used chi-squared tests to assess the association between exposures of interest and the outcome. These chi-square tables are included in the appendix for reference. Except for completion of therapy, exposure variables were chosen based on the p-value of the chi-squared test. We chose p-values less than <0.05, as indicating a statistically significant association between the exposure and outcome of state rurality designation, to assess as a main exposure in the logistic regression models. Due to the constraints of the OTIS database, we could only consider a maximum of three covariates for each main exposure in the logistic regression models. We decided on covariates for each model based on previously known confounders, literature review, and data collected from the key informant interviews.

In total, we assessed five logistic regression models. Model 1 included all chosen demographic exposure variables (sex, race/ethnicity, origin of birth, and age). An odds ratio was obtained for each of these variables while controlling for all other variables in the model. Model 2 included HIV positive cases as the exposure of interest and controlled for race/ethnicity, injection drug use, and disease site. Model 3 included pulmonary disease site as the main exposure and controlled for HIV status, race/ethnicity, and injection drug use. Model 4 included completion of therapy within a year as the exposure of interest and controlled for HIV status, homelessness within the past year, and alcohol use. Model 5 included cases who received directly observed therapy as the exposure of interest and controlled for origin of birth, HIV status, and multidrug resistant TB. Collinearity was assessed in all models, and effect measure modification (EMM) and statistical interaction was assessed between the exposure of interest and all other covariates for models 2-5.

"Not applicable" and "not reported" data were excluded from these analyses and marked as missing. Additionally, due to the previously mentioned limitations of the OTIS database, case totals that were suppressed were also marked as missing.

Results

Study Population

The total number of TB cases reported during 2014 – 2019 in all 14 states included in this analysis was 30,297 (Table 1). The total number of cases in the rural states was 2,380 (7.85%) and 27,917 (92.14%) cases in the urban states. A total of 18,885 (62.33%) of these TB cases were male and 21,340 (70.57%) of reported cases were non-U.S. born individuals. There were 4,076 (13.45%) cases that were between 0-24 years old, 8,548 (28.21%) cases between 25-44 years old, 9,721 (32.09%) between 45-64 years old, and 7,952 (26.25%) cases that were 65 years old or older. There were 10,755 (35.66%) individuals who were Asian, non-Hispanic, 4,912 (16.29%) individuals who were Black, non-Hispanic, 10,787 (35.77%) were Hispanic or Latino, 3,279 (10.87%) were White, non-Hispanic, and 416 (1.38%) of individuals were of "Other" race/ethnicity. "Other" race/ethnicity was defined as someone who reported being American Indian or Alaska Native, non-Hispanic, Native Hawaiian or Other Pacific Islander, non-Hispanic, or Multiple Race, non-Hispanic. These race/ethnicity categories were combined as they each individually made up less than 5% of the total case count.

Table 1: Reported TB Cases by Sociodemographic Variables and Disease Characteristics, Urban and Rural States, U.S., 2014-2019

Table 1 (N = 30,297)*				
		Rural (n = 2,380)** (column %)	Urban (n = 27,917)** (column %)	Total (% of row n)
Sex (n = 30,297)	Male	1,542 (64.75%)	17,344 (62.13%)	18,886 (62.34%)
	Female	839 (35.35%)	10,573 (37.87%)	11,412 (37.67%)
Race and Ethnicity (n = 30,157)	Asian, Non-Hispanic	321 (13.49%)	10,436 (37.38%)	10,757 (35.67%)
	Black or African Amercian, non-Hispanic	894 (37.56%)	4,020 (14.40%)	4,914 (16.29%)
	Hispanic or Latino	291 (12.23%)	10,508 (37.64%)	10,799 (35.81%)
	White, Non-Hispanic	684 (28.74%)	2,595 (9.30%)	3,279 (10.87%)
	Other	179 (7.52%)	262 (0.94%)	441 (1.46%)
Origin of Birth (n = 30,239)	Non US Born	871 (36.60%)	20,469 (73.32%)	21,340 (70.57%)
	US Born	1,509 (63.40%)	7,390 (26.47%)	8,899 (29.43%)
Age (n = 30,297)	0-24 years old	362 (15.21%)	3,714 (13.30%)	4,076 (13.45%)
	25-44 years old	699 (29.37%)	7,849 (28.12%)	8,548 (28.21%)
	45-64 years old	760 (31.93%)	8,961 (32.10%)	9,721 (32.09%)
	65+ years old	559 (23.49%)	7,393 (26.48%)	7,952 (26.25%)
Disease Site (n = 30,287)	Pulmonary	1,812 (76.13%)	19,842 (71.07%)	21,654 (71.50%)
	Extrapulmonary or Both	567 (23.82%)	8,072 (28.91%)	8,639 (28.52%)
HIV Status (n = 26,858)	Positive	82 (3.45%)	1,385 (4.96%)	1,467 (5.46%)
	Negative or Indeterminate	2,152 (90.42%)	23,265 (83.34%)	25,417 (94.63%)
Completion of Therapy (n = 23,825)	Within 1 Year	1,728 (72.61%)	20,504 (73.45%)	22,232 (93.31%)
	> 1 year	143 (6.01%)	1,451 (5.20%)	1,594 (6.69%)
Directly Observed Therapy (n = 29,124)	Direct Only	1,452 (61.01%)	15,233 (54.57%)	16,685 (57.29%)
	Self Only or Both	812 (34.12%)	11,637 (41.68%)	12,449 (42.74%)

*Data for some variables missing because they are either not reported, not applicable, or suppressed.

**Due to missing, not reported, not applicable, or suppressed data, values may not add to column total, and column precents may not add to 100.

Table 2: Odds Ratios for Comparisons of Reported TB Cases from Rural vs Urban states, U.S., 2014-2019

Odds Ratio Table				
Exposures	**OR (null = 1)**	**Lower 95% C.I.**	**Upper 95% C.I.**	**Significant***
Male	1.217	1.089	1.359	Yes
Non-US Born	0.335	0.293	0.382	Yes
Asian, Non-Hispanic**	0.207	0.171	0.252	Yes
Black or African American, Non-Hispanic**	0.990	0.874	1.121	No
Hispanic or Latino**	0.132	0.110	0.159	Yes
25-44 years old***	2.315	1.872	2.863	Yes
45-64 years old***	1.593	1.290	1.968	Yes
65+ years old***	1.784	1.432	2.222	Yes
HIV Positive****	0.177	0.125	0.252	Yes
Pulmonary Disease	1.388	1.235	1.561	Yes
Completion of Therapy within a Year	1.143	0.928	1.409	No
DOT*****	1.301	1.163	1.454	Yes

*Significant odds ratios include 95% confidence intervals that do not cross the null value of 1.

**White, non-Hispanic is used as the reference group for these measures of association.

***0-24 years old is used as the reference group for these measures of association.

****HIV negative and HIV indeterminant is the reference group for is measure of association.

*****Self-administered therapy only and both DOT and self-administered is the reference group for this measure of association.

Model 1: Demographic Exposures

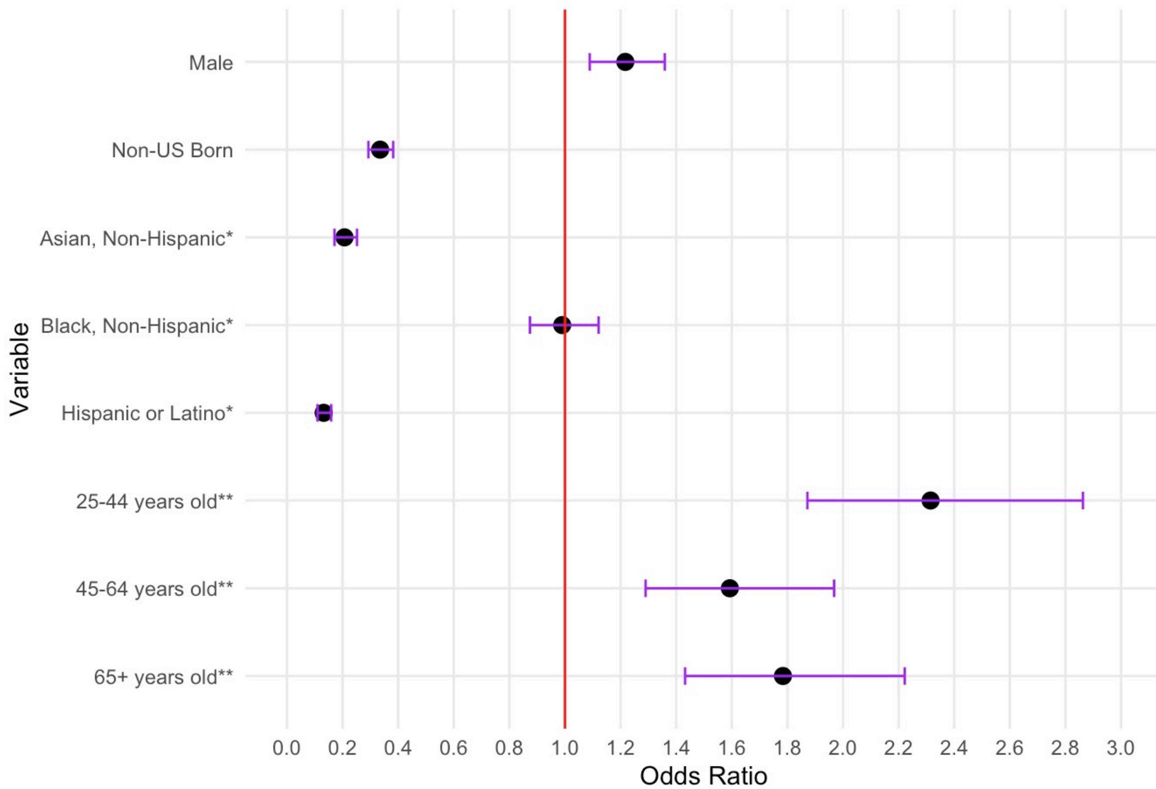
This first model included a logistic regression modeling the odds of certain demographic characteristics by rural vs urban states. The variables that were assessed in this regression were sex, age, race/ethnicity, and origin of birth. For the age variable, the age group 0-24 years old was considered as a reference group as it was the age group with the lowest frequency of TB cases. For race/ethnicity, TB cases who were White, non-Hispanic were considered as a reference group as it was the race/ethnicity with the lowest frequency of TB cases, and the "Other" race category was dropped as it made up less than 2% of cases and created collinearity violations, meaning all the race/ethnicity predictors were too closely related and the model would not run. Collinearity was assessed between all included exposures, and there were no collinearity violations. As EMM and statistical interaction were not of interest in this regression, we proceeded with the regression analysis.

After controlling for origin of birth, race/ethnicity and age, the odds that an individual with TB is male among reported cases living in a rural state from 2014-2019 is 1.217 (1.089, 1.359) times the odds that an individual with TB is male among those living in an urban state. After controlling all other covariates, the odds that an individual with TB is non-US. born among those living in a rural state is 0.335 (0.293, 0.382) times the odds that an individual with TB is non-US. born among those living in an urban state.

Compared to White, non-Hispanic individuals, after controlling for sex, origin of birth, age, and other race/ethnicity groups, the odds that an individual with TB is Asian, non-Hispanic among those living in a rural state is 0.207 (0.171, 0.252) times the odds that an individual with TB is Asian, non-Hispanic among those who lived in an urban state. The odds that an individual with TB is Black or African American, non-Hispanic among those living in a rural state is 0.990 (0.874, 1.121) times the odds that an individual with TB is Black or African American, non-Hispanic among those who lived in an urban state. The odds that an individual with TB is Hispanic or Latino among those living in a rural state is 0.132 (0.110, 0.159) times the odds that an individual with TB is Hispanic or Latino among those who lived in an urban state.

Compared to cases in age group 0-24 years old, after controlling for sex, origin of birth, race/ethnicity, and other age groups, the odds that an individual with TB is 25-44 years old among those living in a rural state is 2.315 (1.872, 2.863) times the odds that an individual with TB is 25-44 years old among those living in an urban state. The odds that an individual with TB is 45-64 years old among those living in a rural state is 1.593 (1.290, 1.968) times the odds that an individual with TB is 45-64 years old among those living in an urban state. The odds that an individual with TB is 65+ years old among those living in a rural state is 1.784 (1.432, 2.222) times the odds that an individual with TB is 65+ years old among those living in an urban state. There is evidence that the relationship between state rurality and reporting race/ethnicity of Black or African American, non-Hispanic among TB cases is not statistically significant given the 95% confidence interval crosses the null value of 1 (Table 2). However, there is evidence that the relationships between state rurality and the remainder of the demographic variables are statistically significant as none of the 95% confidence intervals cross the null value of 1.

Figure 1: Odds Ratio Estimates for Sociodemographic Variables, Comparing TB Cases in Rural vs Urban States



The odds ratios are comparing individuals with TB living in rural states to individuals with TB living in urban states.

*White, non-Hispanic is used as the reference group for these measures of association.

**0-24 years old is used as the reference group for these measures of association.

Model 2: HIV Status Exposure

The second logistic regression modeled the odds of being HIV positive comparing TB cases reported from rural vs urban states. After cleaning the data, there were a total of 26,858 cases of TB in this subset of states from 2014-2019. There were 3,439 cases that had missing, not reported, not applicable, or suppressed data for HIV status. In rural states, there were 2,211 TB cases that reported data on their HIV status, and 3.62% of TB cases were positive for HIV. In urban states, 24,647 TB cases reported data on their HIV status, and 5.62% of cases were HIV positive.

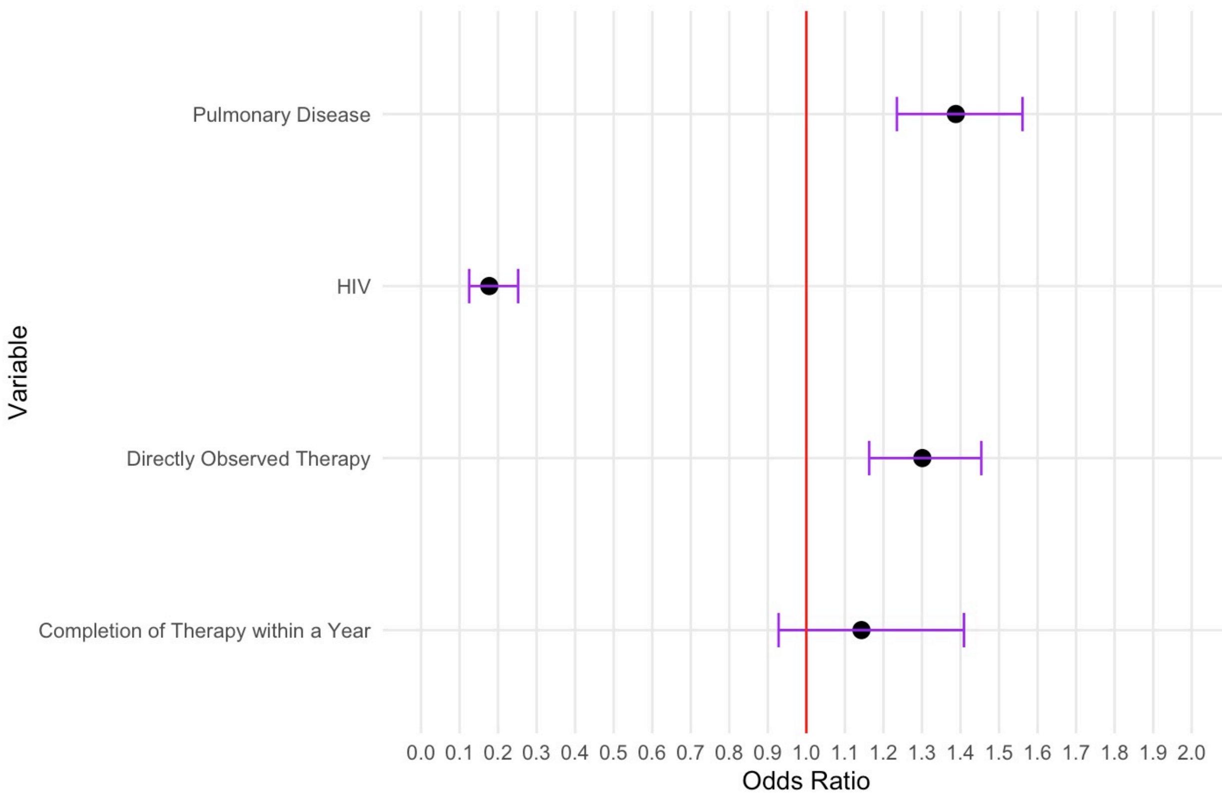
The initial model included HIV status, race/ethnicity, injection drug use, and disease site. EMM and statistical interaction were considered between HIV status and all other covariates. After running the collinearity assessment, the product terms for HIV status and race/ethnicity and disease site were collinear and dropped from the model. Injection drug use was assessed for EMM and statistical interaction, but no interaction was present, so this product term was also dropped from the model. After the assessment of confounding, race/ethnicity was found to be a confounder in this model and kept in the model. After further consideration, we decided to keep the injection drug use and disease site variables in the final model as well. After controlling for disease site, injection drug use, and race/ethnicity, the odds that an individual with TB is HIV positive among individuals who live in a rural state between 2014-2019 is 0.177 (0.125, 0.252) times the odds that an individual with TB has HIV among those who live in an urban state. There is evidence that there is a statistically significant relationship between HIV status and state rurality as the 95% confidence interval does not cross the null value of 1 (Table 2).

<u>Model 3: Disease Site Exposure</u>

The third logistic regression modeled the odds of having pulmonary TB disease by residence in a rural vs urban state. There was a total of 30,287 cases of TB with reported data on site of disease from 2014-2019, and 10 cases with missing, not reported, not applicable, or suppressed data for disease site. There were 2,070 cases from rural states had data reported for disease site, and 87.97% of those cases had pulmonary TB disease. In urban states, 27,914 cases had data reported for diseases site, and 71.08% of those cases had pulmonary disease.

The initial model included all the same variables as the previous HIV model, with disease site as the exposure of interest. The presence of interaction was considered in the initial model between each covariate and disease site. However, all interaction terms were dropped from the model due to a collinearity violation. After a confounding assessment of all the covariates, there was evidence of confounding by all the covariates together, and they were all kept in the final model. After controlling for presence of HIV, injection drug use, and race/ethnicity, the odds that an individual with TB had pulmonary TB between 2014-2019 among those living in a rural state is 1.388 (1.235, 1.561) times the odds that an individual with TB has pulmonary disease among those living in an urban state. As the 95% C.I. does not cross over the null value of 1, there is evidence that this measure of association is statistically significant (Table 2).

Figure 2: Odds Ratio Estimates for Pulmonary Disease, HIV Comorbidity, Directly Observed Therapy, and Completion of Therapy Within 1 Year, Comparing Reported TB Cases in Rural vs Urban States, U.S., 2014-2019



The odds ratios are comparing individuals with TB living in rural states to individuals with TB living in urban states.

<u>Model 4: Completion of Therapy within a Year Exposure</u>

The fourth logistic regression modeled the odds of completing TB therapy within a year by residence in rural vs. urban states. The total number of TB cases that had data reported on completion of therapy within a year between 2014-2019 was 23,825 with 6,472 missing, not reported, not applicable, or suppressed cases. There was a total of 1,870 cases included from rural states and 92.41% of those cases reported completing their TB therapy within one year. There was a total of 21,955 cases from urban states who reported information on when they completed their TB therapy, and 93.39% of those cases reported completing therapy within one year.

The initial model for this regression included the following variables: completion of therapy within a year, HIV status, homelessness within the past year, and alcohol use. Multidrug resistance was considered as a potential covariate but excluded from the model as data for multidrug resistant TB are "not applicable" for this variable.[18] The presence of interaction was considered in the initial model between each covariate and completion of therapy within a year. Interaction product terms between completion of therapy and alcohol use and completion of therapy and homelessness were dropped due to violations of collinearity. Interaction between completion of therapy within a year and HIV status was assessed. Ultimately, this product term was dropped due to a lack of evidence of interaction between these two variables. Additionally, there was no evidence of confounding between completion of therapy and any of the covariates, so all covariates were dropped from the model. The final model included a crude analysis of the association between completion of therapy within a year and state rurality. The odds that an individual with TB reported between 2014-2019 completed their TB therapy within a year among those living in a rural state is 1.143 (0.928, 1.409) times the odds that an individual with TB completed their treatment within a year among those living in an urban state. There is no evidence that this association is statistically significant as the 95% confidence interval crosses over the null value of 1 (Table 2).

<u>Model 5: Directly Observed Therapy Exposure</u>

The final logistic regression modeled the odds of having DOT by residence in rural vs urban states. There was a total of 29,124 TB cases between 2014-2019 that reported their type of therapy in the states included in this analysis. There were 1,173 cases that didn't report, had missing values, had not applicable data, or had suppressed data for their therapy type. Of TB cases with reported therapy type, 2,254 were from rural states, and 64.42% had DOT. Another 26,870 TB cases with reported therapy type were from urban states, and 56.69% had DOT.

The initial model of this regression included the variables: DOT, origin of birth, HIV status, and multidrug resistance. The presence of interaction was considered in the initial model between each covariate and DOT. The product term assessing interaction between DOT and multidrug resistance was dropped due to collinearity issues. After the evaluation of interaction between DOT and HIV status and between DOT and origin of birth, there was no evidence of interaction and both product terms were dropped from the model. Finally, a confounding assessment showed evidence of confounding by all covariates, so they were all kept in the final model. After controlling for HIV status, presence of multidrug resistant TB, and origin of birth, the odds of being on only DOT, rather than self-observed therapy or both self-observed therapy and DOT, among those living in a rural state is 1.301 (1.163, 1.454) times the odds of being on only DOT treatment among those living in an urban state. There was evidence that this association was statistically significant as the 95% confidence interval did not cross the null value of 1 (Table 2).

Discussion

Compared to individuals with TB who live in in urban states, individuals with TB living in rural states and reported between 2014-2019 were more likely to be male; older than 25 years old; specifically, between 25-44 years old; born in the US; and White, non-Hispanic over all other racial or ethnic group. In rural states, the data suggest that TB patients are less likely to be Asian, non-Hispanic or Hispanic or Latino. However, there is no evidence that there is a statistically significant difference in the odds that someone with TB in a rural state is Black or African American, non-Hispanic compared to those who are White, non-Hispanic.

Individuals with TB in rural states were more likely to have pulmonary TB but less likely to be HIV positive than those in urban states. Pulmonary TB disease is the most common site of TB disease, which aligns with seeing higher odds of having pulmonary TB disease in rural states compared to urban states. While we were unable to identify reasons for this observation through this analysis, we speculate that this association could reflect a lack of provider experience, or access to diagnostic procedures in rural states to properly diagnose extrapulmonary TB, also known as ascertainment bias. It would be interesting to further assess this relationship to explore if there is any other influence of outside factors. Risk factors of HIV and TB, such as injection drug use and risky sexual behavior, are more likely to be seen in urban areas than rural areas of the U.S., supporting the finding that it is more likely someone living with TB and is HIV positive is more likely to live in an urban state.

Both completion of TB therapy within a year and directly observed therapy were more likely to be seen in reported cases in rural states compared to in urban states from 2014-2019. While there wasn’t a significant association between completion of therapy within a year and state rurality, it would be intriguing to further explore this association as it would be interesting to analyze what factors are most influential in the timeliness of therapy completion. The completion of TB therapy may take longer in urban states due to a higher proportion of people with TB having risk factors that could delay treatment completion.

The public health professionals from the key informant interviews expressed challenges managing cases of TB due to significant communicable disease workloads and a lack of resources.[3] Because of this, it is notable that individuals with TB in rural states were more likely to be on directly observed therapy compared to self-administered therapy or a combination of both. Self-administered therapy may limit the resources and time needed to provide directly observed therapy to people living with TB in smaller health departments in rural jurisdictions. Nevertheless, these smaller health departments appear to have been able to prioritize the use of DOT as a treatment adherence assurance tool amid limited resources. These health departments may also lack the funding and resources to provide the technology that could make other, less labor-intensive methods for treatment monitoring, such as video observed therapy, possible. Lastly, directly observed therapy has long existed as the gold standard for TB treatment monitoring, and even less well-resourced health departments may have found cost-effective ways to provide directly observed therapy.

Strengths and Limitations

This study has substantial strengths. First, this analysis considers several different exposures. Many different demographic characteristics were assessed to compare rural vs urban states. These analyses highlight the importance of considering the demographic breakdown of the population of interest when doing public health work or designing interventions. This analysis shows that the demographic composition of individuals with TB is significantly different in urban states compared to rural states, so health and public health professionals should tailor case management activities and interventions to the affected populations. Also, some common risk factors, as well as TB case management factors, were analyzed. Combining analyses like these with qualitative data, such as the key informant interviews referenced in this paper, can provide further insight into the needs and experiences of TB patients in both rural and urban areas in the U.S.to create new ways to improve their care.

The OTIS database appeared sufficiently robust for these analyses. This database included substantial data collected for reported TB cases. Each analysis had a large sample size, with over 20,000 cases of TB. In addition, these data come from the systematic collection and reporting of national TB cases to CDC's National TB Surveillance System. These conform to standardized case definitions and variables.[19] Previous assessments of completeness of TB surveillance in the U.S. have documented that >95% of TB cases get reported to CDC.[20] Therefore, the database is likely robust and reliable. This database was free and open access, giving anyone the ability to research and learn more about TB cases in the U.S. and replicate or add on to this study.

This database came with limitations. The OTIS database required the user to choose which variables they wanted to group the data by and created an output with the grouped combinations and case counts. The more variables the user chooses to group by, the more cases become suppressed due to case counts less than 5, especially in rural states. While there was a possible total of 30,297 cases of TB reported by all the states of interest, there was data suppressed in all regressions, lowering the total case count, and creating potential bias associated with missing data. This database also only included state level data. It is challenging to accurately assign an entire state to rural or urban, as some rural states still have large urban areas and vice versa. We relied on categorization of rural states, using criteria previously defined by the 2010 United States Census Bureau. More recent categorizations would have resulted in the inclusion of New Hampshire, with 76 additional reported TB cases during the study period.[21] Ideally, we would have preferred to have access to county level data to explore these differences among rural and urban counties.

The diabetes variable was also of initial interest as an exposure at the beginning of this analysis. However, this variable was split into 3 categories: "Yes", "No", and "Not Reported." Data only existed if an individual with TB was diagnosed with diabetes or if they did not report their diabetes status. There were no cases who reported not having diabetes. Therefore, there was not a meaningful way to assess diabetes status by state of residence in this analysis, even though diabetes is a known comorbid condition for persons reported with TB disease.

Future Directions

This analysis took place with data from the years before the COVID-19 pandemic. The pandemic did affect TB in the US, and it would be interesting to consequently explore any changes among TB in rural compared to urban areas. Additionally, this analysis should be repeated with county level data to obtain a more accurate understanding of differences in individuals with TB in urban and rural counties. Demographics of individuals of TB in rural areas should be compared to the demographics of the general population of individuals living in rural areas to further examine differences of individuals with TB. More risk factors and exposures associated with TB should also be assessed based on the geographic residence of people with TB.

The key informant interviews referenced in this study provided useful background, insight, and inspiration for this analysis. Qualitative data are very powerful to capture the experiences of people living with disease and people managing or working in health care. It provides strong context to quantitative analyses and helps create a more complete picture of a given situation. These key informants expressed many challenges that their patients and themselves experience due to living in a rural area. Some of the biggest challenges included lack of access to health care and TB diagnoses due to inadequate provider knowledge of TB or minimal resources and funding in rural communities. It is important to further explore these challenges so that individuals with TB in rural areas of the U.S. can have the resources they need to receive a proper and timely diagnosis and treatment for their TB disease.

Acknowledgements

A special thank you to Dr. Kenneth Castro for all your guidance, mentorship, support, and encouragement throughout this process. Thank you to Lea Hamner, Dr. Rebekah Stewart, and Dr. Lauren Christiansen-Lindquist for sharing your expertise and advice. Your contributions were invaluable, and I’m deeply grateful for the role each of you played in shaping this work.

References

1. Williams, P. M. (2024). Tuberculosis—United States, 2023. *MMWR. Morbidity and Mortality Weekly Report*, *73*. https://doi.org/10.15585/mmwr.mm7312a4

2. CDC. (2024, October 4). *Clinical Overview of Tuberculosis Disease*. Tuberculosis (TB). https://www.cdc.gov/tb/hcp/clinical-overview/tuberculosis-disease.html

3. Myers, C., Hamner, L., & Castro, K. (n.d.). *Qualitative Assessment of Best Practices and Common Challenges to Active TB Case Management in Local Public Health Departments in Rural Settings of the U.S.*

4. CDC. (2024, June 14). *About Active Tuberculosis Disease*. Tuberculosis (TB). https://www.cdc.gov/tb/about/active-tuberculosis-disease.html

5. CDC. (2024, November 12). *Reporting Area Data*. Reported Tuberculosis in the United States, 2023. https://www.cdc.gov/tb-surveillance-report-2023/summary/reporting-area.html

6. Labuda, S. M., McDaniel, C. J., Talwar, A., Braumuller, A., Parker, S., McGaha, S., Blissett, C., Wortham, J. M., Mukasa, L., & Stewart, R. J. (2022). Tuberculosis Outbreak Associated With Delayed Diagnosis and Long Infectious Periods in Rural Arkansas, 2010-2018. *Public Health Reports (Washington, D.C. : 1974)*, *137*(1), 94–101. https://doi.org/10.1177/0033354921999167

7. Bock, N. N., Mallory, J. P., Mobley, N., DeVoe, B., & Taylor, B. B. (1998). Outbreak of tuberculosis associated with a floating card game in the rural south: Lessons for tuberculosis contact investigations. *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America*, *27*(5), 1221–1226. https://doi.org/10.1086/514986

8. Menzies, N. A., Hill, A. N., Cohen, T., & Salomon, J. A. (2018). The impact of migration on tuberculosis in the United States. *The International Journal of Tuberculosis and Lung Disease : The Official Journal of the International Union against Tuberculosis and Lung Disease*, *22*(12), 1392–1403. https://doi.org/10.5588/ijtld.17.0185

9. Armstrong, L. R., Winston, C. A., Stewart, B., Tsang, C. A., Langer, A. J., & Navin, T. R. (2019). Changes in tuberculosis epidemiology, United States, 1993–2017. *The International Journal of Tuberculosis and Lung Disease : The Official Journal of the International Union against Tuberculosis and Lung Disease*, *23*(7), 797–804. https://doi.org/10.5588/ijtld.18.0757

10. Chen, M. P., Miramontes, R., & Kammerer, J. S. (2020). Multidrug-resistant tuberculosis in the United States, 2011–2016: Patient characteristics and risk factors. *The International Journal of Tuberculosis and Lung Disease : The Official Journal of the International Union against Tuberculosis and Lung Disease*, *24*(1), 92–99. https://doi.org/10.5588/ijtld.19.0173

11. Al-Rifai, R. H., Pearson, F., Critchley, J. A., & Abu-Raddad, L. J. (2017). Association between diabetes mellitus and active tuberculosis: A systematic review and meta-analysis. *PLoS ONE*, *12*(11), e0187967. https://doi.org/10.1371/journal.pone.0187967

12. Sultana, Z. Z., Hoque, F. U., Beyene, J., Akhlak-Ul-Islam, Md., Khan, M. H. R., Ahmed, S., Hawlader, D. H., & Hossain, A. (2021). HIV infection and multidrug resistant tuberculosis: A systematic review and meta-analysis. *BMC Infectious Diseases*, *21*, 51. https://doi.org/10.1186/s12879-020-05749-2

13. *Online Tuberculosis Information System*. (n.d.). Retrieved March 30, 2025, from https://wonder.cdc.gov/wonder/help/tb.html

14. *General Help for CDC WONDER*. (n.d.). Retrieved March 30, 2025, from https://wonder.cdc.gov/wonder/help/main.html#WhatisWONDER

15. Schildknecht, K. R. (2023). Tuberculosis—United States, 2022. *MMWR. Morbidity and Mortality Weekly Report*, *72*. https://doi.org/10.15585/mmwr.mm7212a1

16. Division of Tuberculosis Elimination, N. C. for H., Viral Hepititis, STD, and TB Prevention. (2023). *Tuberculosis (TB) Disease in the United States 1993–2023** (p. 57). Centers for Disease Control and Prevention.

17. *Rural America*. (n.d.). Retrieved March 31, 2025, from https://mtgis-portal.geo.census.gov/arcgis/apps/storymaps/collections/189aa1dbd64c4c81b3b4a2b71124f6c6?item=1

18. *Online Tuberculosis Information System*. (n.d.). Retrieved April 15, 2025, from https://wonder.cdc.gov/wonder/help/tb.html#Completion%20of%20Therapy

19. CDC. (2025, February 19). *Technical Notes*. Reported Tuberculosis in the United States, 2023. https://www.cdc.gov/tb-surveillance-report-2023/commentary/notes.html

20. Curtis, A. B., McCray, E., McKenna, M., & Onorato, I. M. (2001). Completeness and timeliness of tuberculosis case reporting: A multistate study. *American Journal of Preventive Medicine*, *20*(2), 108–112. https://doi.org/10.1016/S0749-3797(00)00284-1

21. *Explore Rural Population in the United States | AHR*. (n.d.). Retrieved April 21, 2025, from https://www.americashealthrankings.org/explore/measures/pct_rural_b

Appendix

Study Population Breakdown

In rural states, 1,541 (64.75%) of the individuals reported with TB were male and 839 (35.25%) of individuals were female. 362 (15.21%) of individuals were between 0-24 years old, 699 (29.37%) were between 25-44 years old, 760 (31.93%) were between 45-64 years old, and 559 (23.49%) were 65 years or older. 321 (13.49%) of people with TB were Asian, non-Hispanic, 894 (37.56%) were Black or African American, non-Hispanic, 684 (28.74%) were White, non-Hispanic, 291 (12.23%) were Hispanic or Latino, and 179 (7.52%) of people were of "Other" race/ethnicity. 11 (0.46%) TB cases did not report their race or ethnicity. 1,509 (63.40%) of reported cases were individuals born in the US. 871 (36.60%) of cases were not born in the U.S.

In urban states, 17,344 (62.13%) of reported TB cases were male, and 10,573 (37.87%) of cases were female. 3,714 (13.30%) of cases were 0-24 years old, 7,849 (28.12%) were 25-44 years old, 8,961 (32.10%) were 45-64 years old, and 7,393 (26.48%) were 65 years or older. 10,436 (37.38%) of individuals with TB in these states were Asian, non-Hispanic. 4,020 (14.40%) of individuals were Black or African American, non-Hispanic. 2,595 (9.30%) of cases were White, non-Hispanic. 10,508 (37.64%) of individuals were Hispanic or Latino. 262 (0.94%) were "Other". 96 (0.34%) cases of TB did not report their race or ethnicity.

Chi-square Tests and Decisions

Sex, age, race/ethnicity, origin of birth, incident TB case, disease site, HIV status, diabetes diagnosis, multidrug resistant TB, completion of therapy within a year, and directly observed therapy were each individually compared to see if any of these variables had a significant association with rurality of the state in which they lived using a chi-square test. Sex, age, race/ethnicity, origin of birth, HIV status, disease site, diabetes, and directly observed therapy had significant associations with the rurality designation of the state with a p-value less than 0.05. All these variables were included as main exposures of tested logistic regressions in this analysis, except for diabetes. While diabetes, was significantly associated with the state type, there were no cases that reported not having diabetes (diabetes "No"). Cases were either classified as having diabetes (diabetes "Yes") or as a case that did not report their diabetes status. For this reason, we decided not to include diabetes in the analyses.

Completion of therapy within a year was not significantly associated with state rurality type individually within the chi-squared test, with a p-value of 0.1018. However, this variable was still considered as a main exposure in the model. Completion of therapy within a year was one of the only variables in the OTIS database that more directly related to TB case management and was identified as a common challenge in the key informant interviews. Many key informants mentioned challenges with lengthy treatment and having to alter the treatment regimen throughout the treatment period, sometimes further lengthening the treatment period. Therefore, completion of therapy within a year was considered as a main exposure in these analyses.

Summary of Chi-square Tables

Variable	*Chi square Test (p-value)*
Incident TB (vs Previous TB) x State	**0.4186**
Disease Site x State*	**<.0001**
DOT x State*	**<.0001**
Birth Origin x State*	**<.0001**
COT w/in 1 year x State	**0.1018**
HIV x State*	**<.0001**
Diabetes x State*	**<.0001**
Race/ethnicity x State*	**<.0001**
MDR x State	**0.2310**
Sex x State*	**0.0113**
Age x State*	**0.0021**

Chi-square Tables

Chi-Square Table Assessing the Association of Incident TB Cases vs State Classification (n=30,150) *

Variable	States with > 50% of TB Burden in the U.S.(urban) (col %)	States with > 40% individuals residing in rural areas (rural) (col %)	Total (% of total)
Incident TB	**26,405 (95.07%)**	**2,251 (94.70%)**	28,656 (95.04%)
Previous TB	**1,368 (4.93%)**	**126 (5.30%)**	1,494 (4.96%)
Total (% of total)	27,773 (92.12%)	2,377 (7.88%)	30,150

*147 missing values due to data suppression, not reported, or not applicable data

Ho: There is no association between state type and incident TB

Ha: There is an association between state type and incident TB

P-value: 0.4186

Chi-Square Table Assessing the Association of TB Disease Site x State Classification (n=30,287) *

Variable	States with > 50% of TB Burden in the U.S.(urban) (col %)	States with > 40% individuals residing in rural areas (rural) (col %)	Total (% of total)
Pulmonary TB	**19,842 (71.08%)**	**1,812 (76.17%)**	21,654 (71.48%)
Extrapulmonary TB or Both	**8,072 (28.92%)**	**567 (23.83%)**	8,639 (28.52%)
Total (% of total)	27,914 (92.15%)	2,379 (7.85%)	30,287

*10 missing values due to data suppression, not reported, or not applicable data

H0: No association between State Type and Disease Site

Ha: There is an association between State Type and Disease Site

P-value < 0.0001

Chi-Square Table Assessing the Association of DOT x State Classification (n=29,124) *

Variable	States with > 50% of TB Burden in the U.S.(urban) (col %)	States with > 40% individuals residing in rural areas (rural) (col %)	Total (% of total)
Direct	**15,233 (56.69%)**	**1,452 (64.42%)**	16,685 (57.29%)
Self or Both	**11,637 (43.31%)**	**802 (35.58%)**	12,439 (42.71%)
Total (% of total)	26,870 (92.26%)	2,254 (7.74%)	29,124

*1,173 missing values due to data suppression, not reported, or not applicable data

H0: No association between State Type and DOT

Ha: There is an association between State Type and DOT

P-value = <.0001

Chi-Square Table Assessing the Association of Birth Origin x State Classification (n=30,239) *

Variable	States with > 50% of TB Burden in the U.S.(urban) (col %)	States with > 40% individuals residing in rural areas (rural) (col %)	Total (% of total)
U.S.Born	**7,390 (26.53%)**	**1,509 (63.40%)**	8,899 (29.43%)
Non U.S.Born	**20,469 (73.47%)**	**871 (36.60%)**	21,340 (70.57%)
Total (% of total)	27,859 (92.13%)	2,380 (7.87%)	30,239

*58 missing values due to data suppression, not reported, or not applicable data

Ho: No association between birth origin and state type

Ha: There is an association between birth origin and state type

P-value: <.0001

Chi-Square Table Assessing the Association of Completion of Therapy within a Year vs State Classification (n=23,825) *

Variable	States with > 50% of TB Burden in the U.S.(urban) (col %)	States with > 40% individuals residing in rural areas (rural) (col %)	Total (% of total)
Within a Year	**20,504 (93.39%)**	**1,728 (92.41%)**	22,232 (93.31%)
Greater than a Year	**1,451 (6.61%)**	**142 (7.59%)**	1,593 (6.69%)
Total (% of total)	21,955 (92.15%)	1,870 (7.85%)	23,825

*6,472 missing values due to data suppression, not reported, or not applicable data

Ho: No association between completion of therapy and state type

Ha: There is an association between completion of therapy and state type

P-value: 0.1018

Chi-Square Table Assessing the Association of HIV Status vs State Classification (n=26,858) *

Variable	States with > 50% of TB Burden in the U.S.(urban) (col %)	States with > 40% individuals residing in rural areas (rural) (col %)	Total (% of total)
HIV +	**1,385 (5.62%)**	**80 (3.62%)**	1,465 (5.45%)
HIV -	**23,262 (94.38%)**	**2,131 (96.38%)**	25,393 (94.55%)
Total (% of total)	24,647 (91.77%)	2,211 (8.23%)	26,858

*3,439 missing values due to data suppression, not reported, or not applicable data

Ho: No association between HIV status and state type

Ha: There is an association between HIV status and state type

P-value: <.0001

Chi-Square Table Assessing the Association of Diabetes vs State Classification (n=30,297)

Variable	States with > 50% of TB Burden in the U.S.(urban) (col %)	States with > 40% individuals residing in rural areas (rural) (col %)	Total (% of total)
Diabetes +	**6,181 (22.14%)**	**341 (14.33%)**	6,522 (21.53%)
Not Reported	**21,736 (77.86%)**	**2,039 (85.67%)**	23,775 (78.47%)
Total (% of total)	27,917 (92.14%)	2,380 (7.86%)	30,297

Ho: There is no association between state type and diabetes

Ha: There is an association between state type and diabetes

P-Value: <.0001

Chi-Square Table Assessing the Association of Race/Ethnicity vs State Classification (n=30,157) *

Variable	States with > 50% of TB Burden in the U.S.(urban) (col %)	States with > 40% individuals residing in rural areas (rural) (col %)	Total (% of total)
Asian, Non-Hispanic	**10,436 (37.52%)**	**319 (13.60%)**	10,755 (35.66%)
Black or African American, Non-Hispanic	**4,020 (14.45%)**	**892 (38.02%)**	4,912 (16.29%)
White, Non-Hispanic	**2,595 (9.33%)**	**684 (29.16%)**	3,279 (10.86%)
Hispanic or Latino	**10,508 (37.78%)**	**279 (11.89%)**	10,787 (35.77%)
Other	**252 (0.91%)**	**172 (7.33%)**	424 (1.41%)
Total (% of total)	27,811 (92.22%)	2,346 (7.78%)	30,157

*140 missing values due to data suppression, not reported, or not applicable data

Ho: There is no association between race/ethnicity and state type.

Ha: There is an association between race/ethnicity and state type.

P-Value: <.0001

Chi-Square Table Assessing the Association of MDR vs State Classification (n = 23,464) *

Variable	States with > 50% of TB Burden in the U.S.(urban) (col %)	States with > 40% individuals residing in rural areas (rural) (col %)	Total (% of total)
MDR Yes	**298 (1.37%)**	**18 (1.03%)**	316 (1.35%)
MDR No	**21,417 (98.63%)**	**1,731 (98.97%)**	23,148 (98.65%)
Total (% of total)	21,715 (92.55%)	1,749 (7.45%)	23,464

*6,833 missing values due to data suppression, not reported, or not applicable data

Ho: There is no association between state type and multidrug resistance

Ha: There is an association between state type and multidrug resistance

P-value: 0.23

Chi-Square Table Assessing the Association of Sex vs State Classification (n = 30,297)

Variable	States with > 50% of TB Burden in the U.S.(urban) (col %)	States with > 40% individuals residing in rural areas (rural) (col %)	Total (% of total)
Female	**10,573 (37.87%)**	**839 (35.25%)**	11,412 (37.67%)
Male	**17,344 (62.13%)**	**1,541 (64.75%)**	18,885 (62.33%)
Total (% of total)	27,917 (92.14%)	2,380 (7.86%)	30,297

Ho: There is no association between state type and sex

Ha: There is an association between state type and sex

P-Value = 0.0113

Chi-Square Table Assessing the Association of Age vs State Classification (n = 30,297)

Variable	States with > 50% of TB Burden in the U.S.(urban) (col %)	States with > 40% individuals residing in rural areas (rural) (col %)	Total (% of total)
0-24 years	**3,714 (13.30%)**	**362 (15.21%)**	4,076 (13.45%)
25-44 years	**7,849 (28.12%)**	**699 (29.37%)**	8,548 (28.21%)
45-64 years	**8,961 (32.10%)**	**760 (31.93%)**	9,721 (32.09%)
65+ years	**7,393 (26.48%)**	**559 (23.49%)**	7,952 (26.25%)
Total (% of total)	27,917 (92.14%)	2,380 (7.86%)	30,297

Ho: There is no association between age and state type.

Ha: There is an association between age and state type.

P-value = 0.0021