ScholarWorks@GSU

Individual and Community Characteristics Associated with Recent Transmission of Tuberculosis in Rural Areas of the United States, 2011–2021

Authors	Stewart, Rebekah
Citation	Stewart, Rebekah. 2025. "Individual and Community Characteristics Associated with Recent Transmission of Tuberculosis in Rural Areas of the United States, 2011#2021." PhD diss., Georgia State University. https://doi.org/10.57709/37817936
DOI	https://doi.org/10.57709/37817936
Download date	2026-02-13 10:38:25
Link to Item	https://hdl.handle.net/20.500.14694/11587

GEORGIA STATE UNIVERSITY

SCHOOL OF NURSING

ACCEPTANCE

This dissertation, "Individual and Community Characteristics Associated with Recent Transmission of Tuberculosis in Rural Areas of the United States, 2011−2021," by Rebekah Stewart was prepared under the direction of the candidate's dissertation committee. It is accepted by the committee members in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Nursing in the Byrdine F. Lewis College of Nursing and Health Professions, Georgia State University.

____
Blake T. McGee, PhD, MPH, RN
Committee Chairperson

____
Maryam Haddad, PhD, MPH, RN, FNP-BC
Committee Member

____
Terri Pigott, PhD
Committee Member

____
Date

This dissertation meets the format and style requirements established by the College of Nursing and Health Professions. It is acceptable for binding, for placement in the University Library and Archives, and for reproduction and distribution in the scholarly and lay community by University Microfilms International.

____
Dawn M. Aycock, PhD, RN, ANP-BC, FAHA, FAAN, FPCNA
Director, PhD in Nursing Program
Byrdine F. Lewis College of Nursing and Health Professions

____
Kyungeh An, PhD, RN, FAAN, FAHA
Associate Dean, School of Nursing
Byrdine F. Lewis College of Nursing and Health Professions

AUTHOR'S STATEMENT

In presenting this dissertation as a partial fulfillment of the requirements for an advanced degree from Georgia State University, I agree that the Library of the University shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to quote from, to copy from, or to publish this dissertation may be granted by the author or, in his/her absence, by the professor under whose direction it was written, or in his/her absence, by the Director of the PhD Program in Nursing, Byrdine F. Lewis College of Nursing and Health Professions. Such quoting, copying, or publishing must be solely for scholarly purposes and will not involve potential financial gain. It is understood that any copying from or publication of this dissertation which involves potential financial gain will not be allowed without written permission from the author.

Rebekah Joy Stewart

NOTICE TO BORROWERS

All dissertation deposited in the Georgia State University Library must be used in accordance with the stipulations prescribed by the author in the preceding statement.

The author of this dissertation is:

Rebekah J. Stewart
Byrdine F. Lewis College of Nursing and Health Professions
Georgia State University
P.O. BOX 4019
Atlanta, GA 30302-4019

The director of this dissertation is:

Blake T. McGee, PhD, MPH, RN
Associate Professor
School of Nursing
Byrdine F. Lewis College of Nursing and Health Professions
Georgia State University
P.O. BOX 4019
Atlanta, GA 30302-4019

Users of this dissertation not regularly enrolled as students at Georgia State University are required to attest acceptance of the preceding stipulations by signing below. Libraries borrowing this dissertation for the use of their patrons are required to see that each user records here the information requested.

NAME OF USER	ADDRESS	DATE	TYPE OF USE (EXAMINATION ONLY OR COPYING)

CURRICULUM VITA

REBEKAH J. STEWART, PhD, MSN, MPH, NP-C, RN, FAANP

Education

PhD	2024	Georgia State University, Atlanta, Georgia
MSN	2014	Emory University, Atlanta, Georgia
MPH	2014	Emory University, Atlanta, Georgia
BSN	2004	Hope College, Holland, Michigan

Professional Employment

2024–Present	Division of TB Elimination, Centers for Disease Control and Prevention *Surveillance Team Lead*
2017–2023	Division of TB Elimination, Centers for Disease Control and Prevention *Nurse Epidemiologist*
2015–2017	Influenza Division, Centers for Disease Control and Prevention *Epidemic Intelligence Service Officer*
2015	Partners in Health, Maryland, Liberia *Ebola Control Officer*
2005–2012	Emergency Department, multiple U.S. hospitals *Registered Nurse/Rapid Response Nurse*

Professional Organizations

United States Public Health Service, *Commander*

American Association of Nurse Practitioners, *Fellow*

Sigma Theta Tau (STT) Internal Honor Society of Nursing, *Member*

Delta Omega Public Health Honor Society, *Member*

Honors

PhD Program Student Research Award, Georgia State University, 2022

Outstanding Service Medal, United States Public Health Service, 2021

Silver Bowl Award for Outstanding Graduating MSN Student, Emory University 2014

Robert W. Woodruff Full Tuition Scholarship, Emory University, 2012

Select Publications as First or Senior Author

Stewart RJ, Raz K, Burns S. Tuberculosis outbreaks in state prisons, United States, 2011–2019. *AJPH.* 2022. DOI: 10.2105/ajph.2022.306864

Chu T, Yousaf A, Chang K,…,**Stewart RJ.** Household Transmission of SARS-CoV-2 from Children and Adolescents. *NEJM*. 2021. DOI: 10.1056/NEJMc2031915

Labuda S, McDaniel CJ, Talwar A,…,**Stewart RJ.** Tuberculosis Outbreak Associated with Delayed Diagnosis and Long Infectious Periods in Rural Arkansas, 2010–2018. *Public Health Reports.* 2021. DOI: 10.1177/0033354921999167

Szablewski CS, Chang KT, McDaniel CJ,…,**Stewart RJ**, Lanzieri TM. SARS-CoV-2 Transmission Dynamics in a Sleep-Away Camp. *Pediatrics.* 2021. DOI: 10.1542/peds.2020-046524

Stewart R, Wortham J, Parvez F, et al. Tuberculosis Infection in Children. *The Journal for Nurse Practitioners.* 2020. DOI:10.1016/j.nurpra.2020.06.027

Stewart RJ, Rossow J, Eckel S, et al. Piloting and Evaluation of Text-Based Illness Monitoring for Detection of Novel Influenza A Virus Infections during an Influenza A(H3N2)v Outbreak in Michigan, 2016. *Journal of Public Health and Surveillance.*

Stewart RJ, Tsang CA, Pratt RH, et al. Tuberculosis – United States, 2017. MMWR Morb Mortal Wkly Rep 2018; 67:11. DOI: 10.15585/mmwr.mm6711a2

Stewart RJ, Rossow J, Thelen J, et al. Do Animal Exhibitors Support and Follow Recommendations to Prevent Transmission of Variant Influenza at Agricultural Fairs? A Survey of Animal Exhibitors after a Variant Influenza Outbreak in Michigan. *Zoonoses and Public Health.* 2017. DOI: 10.1111/zph.12425

Stewart RJ, Flannery B, Chung J, et al. Influenza Antiviral Prescribing for Outpatients with an Acute Respiratory Illness and at High-Risk for Influenza-Associated Complications during Five Influenza Seasons—United States, 2011-2016. *Clinical Infectious Diseases.* 2017. DOI: 10.1093/cid/cix922

ABSTRACT

INDIVIDUAL AND COMMUNITY CHARACTERISTICS ASSOCIATED WITH RECENT TRANSMISSION OF TUBERCULOSIS IN RURAL AREAS OF THE UNITED STATES, 2011–2021

BY

REBEKAH J. STEWART, PhD, MSN, MPH, NP-C, RN, FAANP

Problem: In the United States, social inequities are concentrated in rural areas, potentially increasing the risk for recent transmission (RT) of tuberculosis (TB), a precursor of TB outbreaks. Most TB cases occur among non-U.S.–born patients; however, little is known about TB in rural areas. This dissertation examined characteristics associated with RT in rural areas.

Methods: TB cases with genotyping results reported to the National Tuberculosis Surveillance System during 2011–2021 were included if they occurred in areas classified as non-metropolitan (rural) by the National Center for Health Statistics. Community characteristics included county-level poverty (American Community Survey) and number of primary care physicians (PCP) per 100,000 (Health Resources and Services Administration). Hierarchical generalized linear models were constructed to estimate the relative association of individual and county characteristics with RT.

Results: During 2011–2021, 7,607 rural U.S. residents were diagnosed with TB. Patients lived in counties with a mean of 20% (SD=7.0) of the population below poverty and 49 (SD=27) PCPs per 100,000. Most patients were male (66.0%), U.S.-born (63.8%), and identified as a racial or ethnic minority (74.1%). TB was attributed to RT for 847 (15.0%) of 5,633 cases with an RT designation. Most patients with TB attributed to RT lived in the U.S. Census South Region (65.6%), followed by the West (27.2%), Midwest

(6.4%), and Northeast (0.8%); nearly all were U.S.-born (90.3%). County and patient characteristics, except age, differed significantly by region. In hierarchical models, U.S. birth (odds ratio=5.05, 95% CI: 3.62−7.04), racial/ethnic minority identification (1.54, [1.13−2.11]), incarceration (0.46, [0.26−0.80]), age (0.97, [0.97−0.98]), and county PCP ratio (0.99, [0.98−1.00], p=.02) were associated with RT, whereas homelessness, substance misuse, comorbidities, and county-level poverty were not.

Conclusion: Unlike national epidemiology, most patients with TB in rural areas, and nearly all whose TB was attributed to RT, were U.S.-born. Healthcare providers in rural areas should consider TB as a potential diagnosis among symptomatic U.S.-born patients of racial or ethnic minority groups. More attention is needed to highlight the unique characteristics of patients with TB due to RT in rural areas.

INDIVIDUAL AND COMMUNITY CHARACTERISTICS ASSOCIATED WITH
RECENT TRANSMISSION OF TUBERCULOSIS IN
RURAL AREAS OF THE UNITED STATES, 2011–2021

by

REBEKAH J. STEWART, PhD, MSN, MPH, NP-C, RN, FAANP

A DISSERTATION

Presented in Partial Fulfillment of Requirements for the Degree of Doctor of Philosophy
in Nursing in the Byrdine F. Lewis
College of Nursing and Health Professions,
Georgia State University

Atlanta, GA
2024

COPYRIGHT

Copyright by

Rebekah J. Stewart

2024

ACKNOWLEDGEMENTS

This dissertation is dedicated to my son, Lewis Teddy Lennon, who entered my life early in my PhD journey and whose long weekend naps have allowed me to finish. Your infectious laugh and endless energy have been my motivation to keep going so I can spend my weekends riding bikes with you rather than doing dissertation research. May you live in a world where we finally eliminate tuberculosis and may you strive to reduce inequities wherever you find yourself.

I am so grateful to my committee members for the hours and dedication you have put into teaching me, mentoring me, and reviewing my work. Thank you, Dr. Blake McGee, for believing that public health is an important part of nursing and helping me pursue an epidemiologic study for my PhD in nursing. Your skillful mentorship and statistical acumen have been foundational for my success. Thank you, Dr. Maryam Haddad, for being a mentor, friend, and colleague for over 10 years and for agreeing to add dissertation committee member to your long list of responsibilities. Your critical review and deep understanding of TB literature and epidemiology has made everything I do professionally, including this dissertation, better. Thank you, Dr. Terri Pigott, for giving me personal crash courses in systematic reviews and multi-level modeling. I have learned so many important and useful methods from the two directed reading courses you so generously taught me; your calm demeanor and laugh made me feel so much more relaxed whenever I got overly anxious about the complex statistical methods I was learning.

I owe a mountain of thanks to Dr. Halley Riley of the Research Data Services at GSU who saved me when I couldn’t figure out how to format the county-level variables in my study and merge them in R. I am also grateful to Dr. Matt Hayat for lending me his statistical expertise on multiple occasions.

I am very grateful to all the local and state health department staff who take care of patients with TB disease and who provide the standardized reporting of every TB case in the United States. My gratitude also goes out to the Surveillance Team and Data Management Team in the Division of TB Elimination at CDC whose dedication and expertise in TB surveillance made this national analysis possible.

I also want to thank my partner, Patrick Lennon, who never doubted that I could work, become a mother, and do a PhD all at the same time, and who keeps saying “yes” to me no matter what I propose. Thank you for encouraging me to just take it “bird-by-bird” and for watching our little man during many long weekends of dissertating.

I would never have completed this dissertation without Dr. Laura Vonnahme, my friend, colleague, and cheerleader. She motivated me to pursue a PhD, helped me select the best classes to take, taught me R, took weekends away to dissertate with me, worked, mothered, and suffered with me, and encouraged me every step of the way. After these last four years, I don’t think there is anything the two of us cannot do.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACS	American Community Survey
AI/AN	American Indian/Alaska Native
CDC	Centers for Disease Control and Prevention
FIPS	Federal Information Processing Standard
HCV	Hepatitis C virus
HLM	Hierarchical Generalized Linear Model
HPSA	Health Professional Shortage Area
HRSA	Health Resources and Services Administration
ICC	Intra-Class Correlation
ICE	U.S. Immigration and Customs Enforcement
IGRA	Interferon-Gamma Release Assay
LTBI	Latent Tuberculosis Infection
M.tb	*Mycobacterium tuberculosis*
NCHS	National Center for Health Statistics
NH/PI	Native Hawaiian/Pacific Islander
NTSS	National Tuberculosis Surveillance System
OR	Odds Ratio
PCP	Primary Care Physician
RT	Recent Transmission
RVCT	Report of Verified Case of Tuberculosis
SES	Socioeconomic Status
SE	Standard Error
ST	Syndemic Theory
TB	Tuberculosis
TST	Tuberculin Skin Test
VIF	Variance Inflation Factor

CHAPTER 1

INTRODUCTION

Tuberculosis (TB) was called “the captain of all these men of death” in the 1600s and is as old as human history (Dubos & Dubos, 1952). Yet now in the 21st century, TB is still the leading infectious disease cause of death globally. Every year, 10 million people develop TB disease, and 1.5 million people die of TB-related causes (World Health Organization [WHO], 2022). This severe TB-related morbidity and mortality persists despite advances in science and medicine that make TB preventable and curable with antibiotics.

The global burden of TB is not shared equally. As a country’s socio-demographic index (a composite score based on income, education, and fertility rate) increases, the TB incidence, prevalence, and mortality rates due to TB decrease (Global Burden of Disease Tuberculosis Collaborators [GBD], 2017). As such, the United States has one of the lowest incidence and mortality rates of TB in the world. In 2022, there were 8,300 cases of TB reported, corresponding to an incidence rate of 2.5 cases per 100,000 people (Centers for Disease Control and Prevention [CDC]). Annually, approximately 10% of people with TB in the United States die with about 35% of those deaths attributable to TB (CDC, 2021).

However, even in the United States, the burden of TB is not distributed equally across the population. As incidence and mortality rates have declined, the disease has become concentrated in poorer, socially marginalized populations (CDC, 2021; Mancuso et al., 2016; Olson et al., 2012). While U.S.-born White people have an incidence rate of

0.3 cases per 100,000 people, the rate among non-U.S.–born Asian people is nearly 80 times higher at 23.8 cases per 100,000 people (CDC, 2021). In some U.S. communities experiencing ongoing transmission of TB, the rates of TB mirror those of the highest burden countries in the world (Asabor & Vermund, 2021).

This disparate nature of TB disease has been understood for decades with historical writings referring to the TB bacteria as a “necessary, but not sufficient” part of the cycle of disease that leaves those in the lowest rungs of society at the greatest risk (Dubos & Dubos, 1952; Knopf, 1914). In the United States, people living in rural areas are some of the most marginalized, with higher rates of poverty and chronic disease and lower access to healthcare as compared to their urban counterparts (Rural Health Information Hub, 2023). This dissertation will focus on exploring health disparities in TB transmission in rural America and understanding the complex interactions of rurality, TB risk factors, and clinical factors with TB transmission.

Tuberculosis Disease and Latent Tuberculosis Infection

Pathophysiology of Tuberculosis

TB is an airborne infectious disease caused by organisms of the *Mycobacterium tuberculosis* complex (Pai et al., 2016). *M. tuberculosis* most frequently infects the lungs; however, it can infect any part of the body. When *M. tuberculosis* enters the body through exposure to a person with infectious TB disease, the body can eliminate the pathogen through either innate immune responses or because of acquired T cell immunity (Pai et al., 2016). If the pathogen is not immediately cleared by the alveolar macrophages in the lungs, the dendritic cells or inflammatory monocytes transport *M. tuberculosis* to the pulmonary lymph nodes, where T cells and B cells form a granuloma around the bacteria. In most *M. tuberculosis* infections, the granuloma will contain the bacteria and the person infected will develop a latent TB infection (LTBI). However, if the bacteria overwhelm the granuloma, or if the granuloma breaks down due to immunocompromise, the bacteria can replicate in the lungs, causing a cavity that is teeming with bacteria, or can disseminate through the bloodstream to other parts of the body (Pai et al., 2016). This active process is termed TB disease.

Spectrum of TB Disease

TB is usually simplistically classified as one of two states – LTBI or TB disease. Using this strict categorization, a person with LTBI will usually test positive for TB infection with the tuberculin skin test (TST) and the interferon-gamma release assay (IGRA), but will not experience any symptoms of disease, is not infectious, and will not have microbiologic or imaging results consistent with TB disease (Pai et al., 2016). Conversely, a person with TB disease will usually also test positive with a TST and IGRA, but will have signs and symptoms of TB disease, will usually be infectious before treatment (in the case of pulmonary TB), and will usually have microbiologic and/or radiographic findings consistent with TB disease. Despite this simplistic categorization, it is increasingly understood that TB disease exists on a continuum with the possibility to move to and from infection to a latent state, subclinical TB disease, and active TB disease (Pai et al., 2016). Treatment for LTBI greatly reduces the risk of developing active TB disease (Njie et al., 2018).

Risk of Tuberculosis Infection

Risk of infection is tied to risk of exposure; this risk is generally higher in congregate settings such as correctional facilities (Lambert et al., 2016; Stewart et al., 2022) and shelters serving people experiencing homelessness (Bamrah et al., 2013; Self et al., 2021) as well as in communities with active TB transmission (Raz et al., 2022). In the United States, immigrants have the highest rates of infection (Miramontes et al., 2015); this is due to their much higher risk of exposure in their country of origin given the higher rates of TB disease outside the United States.

Epidemiology of Latent Tuberculosis Infection (LTBI)

Approximately 25% of the global population has LTBI (Cohen et al., 2019) however, there is significant geographic variability. High burden regions, such as those in sub-Saharan Africa and Southeast Asia, have prevalence estimates of 34.6% (95% CI: 24.4–42.9) and 36.0% (95% CI: 25.3–46.7), respectively, and low burden regions, such as those in The Americas and Europe, have prevalence estimates of 13.7% (11.0–16.3) and 12.2% (9.8–14.5), respectively (Cohen et al., 2019). In the United States, estimates of LTBI prevalence range from 3% (2.2%–5.2%), representing 8.9 (6.3–14.8) million people (Haddad et al., 2018), to 5% (4.2–5.8) (Miramontes et al., 2015), representing approximately 11 million people.

Risk of Progression to Tuberculosis Disease

While the immune system plays a poorly understood role in the risk of primary infection among those who are exposed, it is a key player in understanding risk among those who progress from infection to disease (Pai et al., 2016). Suppression of the immune system, especially cell-mediated immunity, increases the risk of progression.

HIV co-infection confers the greatest risk in progression to TB disease; HIV-infected individuals have a 30% risk of progressing to disease, compared with a 5%–10% risk among the general population (Markowitz et al., 1997; Moreno et al., 1993). Likewise, type 2 diabetes mellitus (Jeon & Murray, 2008) and excessive alcohol use (Rehm et al., 2009) approximately triple the risk of progression, and smoking doubles the risk of progression (Bates et al., 2007). Therefore, people with diabetes and those whose use excessive alcohol and who smoke are at higher risk of developing TB disease once infected.

Infectious Period of TB Disease

Due to the dynamic nature of TB infection and disease, a person with TB disease can be contagious for up to three months before developing symptoms and usually will remain contagious until treated with a TB regimen that greatly reduces the bacterial load in their body (CDC, 2005). During this time, they can transmit the *M. tuberculosis* bacteria to people with whom they share airspace. This period of time is called an infectious period. The average infectious period in the United States is unknown, but a prolonged infectious period is a well-known risk factor contributing to TB outbreaks (Labuda et al., 2021; Mindra et al., 2017). Because the duration of an infectious period is dependent on a person being diagnosed with and treated for TB disease, a person's health literacy and ability to access healthcare are key determinants in shortening an infectious period. The availability of a healthcare provider with the knowledge and tools to efficiently make the diagnosis and prescribe treatment is also an important determinant. Without these key components, a person's TB can remain undiagnosed and infectious for extended periods of time, an unfortunate situation that occurs frequently in the United States (Raz et al., 2022).

Recent Transmission of Tuberculosis and Tuberculosis Outbreaks

TB disease can arise from one of two situations: reactivation of an untreated or undertreated latent TB infection from the past or from an infection acquired recently (CDC, 2016), indicating recent transmission (RT). While all individuals diagnosed with TB should be promptly treated, those whose disease is attributed to RT require deliberate public health interventions to prevent or stop a TB outbreak. When RT is occurring in a community, swift public health action is warranted to identify where transmission is occurring, evaluate exposed people (i.e., contacts), and provide thorough evaluation and treatment to infected contacts. When RT is occurring in a community, healthcare providers in that community should also be made aware to aid in diagnosis and case management.

Recent transmission is a precursor for TB outbreaks, thus further emphasizing the importance of identifying and stopping TB transmission. Although there is not a nationally agreed upon number of TB cases related by RT that constitutes a TB outbreak, a TB outbreak consists of several cases (number varies by context and jurisdiction) that are related by recent transmission (Mindra et al., 2017; Powell et al., 2017; Raz et al., 2022). Therefore, preventing TB outbreaks starts with locating where recent transmission is occurring and identifying characteristics of people with TB disease attributed to RT to then direct interventions to stop transmission.

National-level analyses of TB outbreaks have found that outbreaks predominantly affect people of racial and ethnic minorities, those who are experiencing homelessness, and those who are incarcerated (Mindra et al., 2017; Raz et al., 2022). However, large outbreaks have also been documented in rural areas where few, if any, of the patients reported experiencing homelessness or incarceration (Bloss et al., 2011; Bock et al., 1998; Labuda et al., 2021). A recent study of large TB outbreaks in the United States found that of 24 outbreaks, five occurred in geographically remote or rural settings and 14 occurred primarily in counties considered to have moderate-to-high or high social vulnerability (CDC Social Vulnerability Index >0.5) (Raz et al., 2022). Despite these scientific advances in understanding outbreaks, no studies have been conducted to characterize recent transmission in rural areas in the United States — areas that have a preponderance of vulnerabilities that increase the risk for TB.

Tuberculosis and Health Disparities

A Disease of Social Inequity

TB has long been understood as a disease of social inequity (Ali, 2014; Knopf, 1914) with the highest rates of TB in lower income countries or in impoverished communities of higher income countries (Lönnroth et al., 2009). Over 95% of TB cases and deaths from TB are in low- and middle-income countries (WHO, 2022). Globally, TB incidence rates are most closely correlated with social and economic determinants such as the human development index, child mortality, and sanitation (Dye et al., 2009). As the human development index increases, childhood mortality decreases, and sanitation improves, rates of TB decrease. In a study conducted in 54 European big cities (defined as municipalities with more than 500,000 residents), even in countries with low annual incidence, TB rates concentrated in areas where homelessness, poverty, migrants, overcrowding, and substance misuse were more common (de Vries et al., 2014). In one city, TB rates differed more than fourfold between boroughs from fewer than 20 cases per 100,000 people to over 80 cases per 100,000.

Although the United States is a high-income country with one of the lowest rates of TB in the world (GBD, 2017), the correlation between social inequity and TB persists. A U.S.-based national study found that higher TB rates were associated with lower socioeconomic status (SES) across four measures of SES: education, crowding, income, unemployment (Olson et al., 2012). TB rates were highest in zip code quartiles with low SES; TB rates among U.S.-born people increased by five-fold or more in the two lowest SES quartiles as compared to the two highest SES quartiles. This phenomenon holds true for LTBI. Nationally, the prevalence of LTBI is higher among those living in poverty and among those with less than a high school education (Mancuso et al., 2016).

Health Disparities in Rural America

Poverty

Social inequities and health disparities are prevalent in rural America, providing optimal conditions for TB to thrive. A greater proportion of people in rural areas of the United States live below the poverty line as compared to those in urban areas (United States Census Bureau, 2020). Low educational attainment and unemployment are also higher in rural areas (Rural Health Information Hub, 2023).

Access to Healthcare

In addition to poverty, rural areas are often underserved in terms of healthcare access (Asabor & Vermund, 2021). People living in rural counties have a lower density of healthcare providers to serve them and must travel farther to reach healthcare facilities (Julia Foutz, 2017), which may further contribute to TB transmission through delayed diagnosis of TB disease (Labuda et al., 2021). The physician-to-population ratio in rural areas is 39.8 physicians per 100,000 people, compared with 53.3 physicians per 100,000 in urban areas (Hing & Chun-Ju, 2014), and the number of specialists per 100,000 people in rural areas is 30, in contrast to 263 in urban areas (Rural Health Information Hub, 2023). Rural residents are more frequently uninsured or covered by Medicaid insurance, leading to less access to care, and have greater transportation difficulties reaching healthcare providers (Rural Health Information Hub, 2023). All these factors introduce barriers to seeking and obtaining healthcare for TB.

Social-behavioral Risk Factors for TB

There is also a concentration of health-related risk factors for TB in rural areas. Smoking, alcohol consumption, and other substance use are all associated with higher TB risk (GBD, 2017; Mindra et al., 2017; Raz et al., 2022). Higher percentages of people smoke in rural areas (National Center for Chronic Disease Prevention and Health Promotion Office on Smoking and Health, 2014). Among racial and ethnic minority groups, the prevalence of binge drinking and alcohol use disorder is higher in non-metropolitan areas than in metropolitan areas (Dixon & Chartier, 2016). Adults living in rural areas have the highest rates of methamphetamine use (Substance Abuse and Mental Health Services Administration, 2023) and the odds of cannabis use (compared to non-use) among non-cigarette smokers in rural areas increased 219% during 2007–2017 (Coughlin et al., 2019). Because substance use is frequently associated with TB outbreaks (Mindra et al., 2017; Raz et al., 2022), the higher prevalence of these behaviors in rural areas increases the risk of outbreaks.

Higher Prevalence of Health Conditions in Rural Areas

Other health conditions are also more prevalent in rural areas. Some of these serve to emphasize the pervasive health disparities independent of TB risk. For example, the risk of stroke is higher in areas described as small rural and/or isolated compared with urban areas (Howard et al., 2017) as is the percentage of adults with diagnosed chronic obstructive pulmonary disease (Croft et al., 2018). Further highlighting the impact of health disparities, a recent study found that rural counties had higher percentages of potentially excess deaths from respiratory disease, and stroke (Garcia et al., 2019).

Another chronic disease more prevalent in rural areas, type 2 diabetes mellitus, is directly associated with increased risk for TB. Diabetes increases the risk of developing TB disease among people who are infected (Jeon & Murray, 2008), can increase health complications in people who have TB (GBD, 2017), and has accounted for a substantial amount of the TB burden in low to moderate TB incidence countries (Lee et al., 2018). Globally, the population attributable fraction of mortality due to diabetes among people with TB disease is 10.6% (GBD, 2017). Compared to urban populations, rural populations in the United States have a 16% higher prevalence of diabetes and a 20% higher diabetes-related hospital mortality (Dugani et al., 2021). Moreover, 62% of rural counties do not have a diabetes self-management education and support program (Rutledge et al., 2017).

Racial and Ethnic Disparities in TB in Rural Areas

Finally, the estimated proportion of people who identify as belonging to a racial and/or ethnic minority group in rural areas is between 18%–31% (Rural Health Information Hub, 2023). While this is a lower proportion than their urban counterparts at approximately 55%, it remains an important area of investigation. Inequities in TB for racial and ethnic minorities are stark. Rates of TB disease are more than six times greater among Black Americans than they are among White Americans (CDC, 2021). As mentioned previously, TB outbreaks are most prevalent among racial and ethnic minority populations (Mindra et al., 2017) and being of a racial or ethnic minority is associated with recent transmission of TB in the United States (Yuen et al., 2016). At the height of a TB outbreak in rural Alabama, the affected town had a TB rate of 253 cases per 100,000, rivaling the rates of TB infection in some of the world's highest burden countries (Asabor & Vermund, 2021). As stated succinctly by Apostolopoulos et al. (2018) "rural Black Americans live and work in particularly pathogenic environments that generate disproportionate and interacting chronic comorbidities compared to their White and/or urban counterparts."

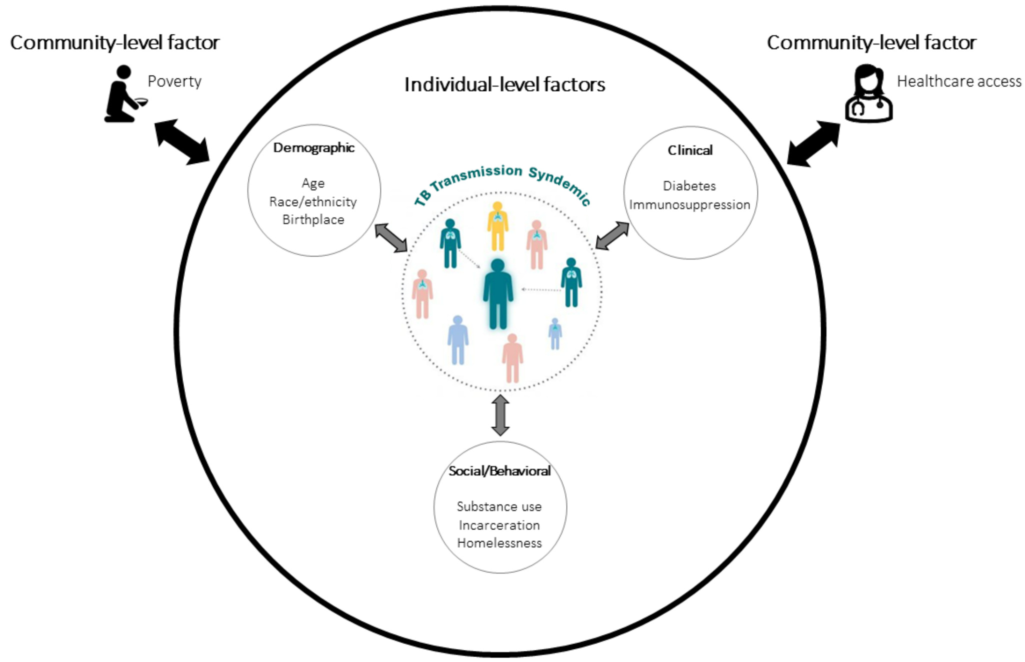
Theoretical Framework

Syndemic Theory

Syndemic Theory (ST) was developed by Merrill Singer in the mid-1990s to call attention to the synergistic nature of the health and social problems facing those living in poverty and with diminished access to healthcare (Singer et al., 2017). ST provides a framework to examine mutually reinforcing diseases and health issues under conditions of social inequality and inequity. In the initial conceptualization, three rules defined a syndemic: 1) two or more diseases cluster together in time or space; 2) these diseases interact in meaningful ways, whether social, psychological, or biological, and 3) harmful social conditions drive these interactions (Mendenhall et al., 2022; Singer, 1996). However, since the theory origination, disease interaction has not always been included and even when included, not well conceptualized and tested (Tsai & Burns, 2015). Much of the syndemic work since then has focused on social and structural drivers of disease and how these drivers interact with the disease in question (Mendenhall et al., 2022). Although ST was first used to study the clustering of HIV with other diseases and social problems (Mendenhall et al., 2022; Singer et al., 2017) and has been used most frequently in HIV research (Hossain et al., 2022), it is clear that TB should also be studied through a syndemic lens. As early as 1952, a microbiologist stated that TB "is a social disease…its understanding demands that the impact of social and economic factors on the individual be considered as much as the mechanisms by which tubercule bacilli cause damage to the human body" (Dubos & Dubos, 1952). Studying the "disease of social inequity" (Ali, 2014) with its stark disparities by race and ethnicity and socioeconomic status, even in one of the lowest incidence countries in the world, demands a syndemic approach. This study will examine how TB transmission interacts with the health disparities in rural America, including co-existing environmental, social-behavioral, and disease-related risk factors for TB that are more prevalent in rural areas (see Figure 1 for an illustration of the Syndemic Theoretical Framework applied to TB transmission in rural areas).

Figure 1

Syndemic Theoretical Framework as Applied to Tuberculosis Transmission in Rural Areas.



Significance to Nursing

Public health nurses are the backbone of public health in the United States with nearly 50,000 nurses comprising 16% of the total public health workforce (Beck et al., 2014). In local health departments that usually serve individual cities or counties, public health nurses comprise 20% of the workforce. In local health departments serving populations of less than 25,000 people, nurses are frequently the only staff member with clinical training; physicians are not usually included as staff until the population served includes at least 100,000 people (National Association of County & City Health Officials, 2019). Indeed, for counties of <25,000 population, most local health departments employ only one or two nurses who conduct all the clinical activities. Because of this, in smaller counties, nurses conduct most investigations of TB transmission and outbreaks and provide most of the case management for patients with TB. Therefore, understanding characteristics associated with TB transmission in rural areas is of critical importance to nurses' work and informs how they will direct limited resources to have the greatest impact.

The profession of nursing is rooted in the philosophy of modern nursing's founder, Florence Nightingale, who believed that individuals were complex and holistic beings whose health was greatly influenced by their environment (Selanders & Crane, 2012). Through her work and advocacy for patients, she laid the foundation of expectation that nurses will use empirical evidence to drive change that is beneficial to human health. This study will provide nurses, the primary healthcare providers in rural counties, the empirical evidence to understand the role of the rural environment on TB transmission and to advocate effectively for the health of the populations they serve.

Research Questions

The overarching research question is the following: What demographic, clinical, social-behavioral, and community-level characteristics are associated with TB transmission in rural areas in the United States, and how do those characteristics interact to drive TB transmission?

Research Question #1

Are there differences in demographic, clinical, and social-behavioral, and community-level characteristics between patients in rural areas whose TB disease is attributed to recent transmission compared with patients whose disease is not?

Research Question #2

Are there regional differences in demographic, clinical, social-behavioral, and community-level characteristics among patients with TB disease attributed to recent transmission in rural areas of the United States?

Research Question #3

What is the relative association of individual-level (e.g., age, race/ethnicity) and community-level characteristics (e.g., poverty, healthcare resources) with recent transmission among people with TB disease in rural areas in the United States?

CHAPTER 2

REVIEW OF THE LITERATURE

Recent transmission of TB can result in outbreaks (Raz et al., 2022; Yuen et al., 2016), which can be devastating for affected families and communities (Labuda et al., 2021; Mindra et al., 2017; Raz et al., 2022). Risk factors associated with TB infection and disease have been described extensively as well as risk factors associated with RT and outbreaks; however, little attention has been paid at a national perspective to TB risk factors specific to rural areas or examining rurality as a risk factor itself. This chapter provides an overview of the scholarly literature regarding TB in rural areas, estimates of recent transmission and associated characteristics, TB outbreaks in rural areas and associated characteristics, and the use of Syndemic Theory in examining co-existing diseases and social conditions that work synergistically to exacerbate health disparities in certain populations.

Tuberculosis in Rural Areas

Reports by CDC

Every year, CDC publishes a TB surveillance report, which includes TB incidence by state, by clinical and laboratory criteria, by residence in a congregate setting, and other stratifications for the previous year (CDC, 2021). Characteristics of TB cases are also presented by origin of birth and race and ethnicity and by attribution to RT. The number and incidence of TB cases are reported for metropolitan statistical areas (MSAs) with ≥500,000 population (i.e., urban MSAs) and RT estimates are displayed by county in a map. However, none of these annual estimates provides incidence rates in rural counties or delineates characteristics of TB cases by a rural/urban classification.

CDC's website has a section addressing TB in specific populations. This includes sections on racial and ethnic disparities, medical conditions, geographic disparities, children, pregnancy, correctional facilities, persons experiencing homelessness, and international travelers. Notably absent is any mention of TB in rural areas or the health disparities related to TB that may exist in rural areas.

Peer-Reviewed Publications

Two articles were identified that examined TB in geographically limited rural areas. In one article, published in 2010, the authors report the results of a population-based survey from 1997–2001 of persons with LTBI in a rural section of a county in central Florida (O'Donnell et al., 2010). They report that Black race was strongly associated with LTBI among U.S.-born persons (OR: 2.6, 95% CI: 1.3–5.5) and that risk factors for TB included HIV (OR: 27.4, 95% CI: 10.1–74.1), drug use (OR 4.6, 95% CI: 1.7–12.4), and Black race (OR: 3.4, 95% CI: 1.2–9.6). Another analysis compared TB incidence rates and trends over a 13 year period (1993–2005) in Appalachia by rural and urban classification (Wallace et al., 2008). They found the rates and trends were similar between urban and rural Appalachia and similar between the full Appalachia region and rural areas outside of Appalachia.

One national-level analysis was identified that reported on the association of rurality, poverty, and LTBI prevalence (Haddad et al., 2018). In the analysis, rural counties with higher-than-average LTBI prevalence were also more frequently counties with higher rates of poverty; this association was not observed in metropolitan counties. Further analysis of findings for rural areas or characteristics of patients in rural areas was not provided as this was not the focus of the paper.

Together, these articles and reports highlight the overall lack of attention that has been paid to TB in rural areas or to examining risk factors associated with TB disease among persons in rural areas.

Use of the Recent Transmission Algorithm

This study will use the recent transmission (RT) algorithm to determine whether cases of TB disease in rural areas are related by recent transmission. The RT algorithm uses molecular and surveillance data to examine relatedness of the TB bacteria among patients that are close geographically and in time. Since the development of the RT algorithm in 2015 (France et al., 2015), it has been used frequently by TB researchers to locate where TB RT is occurring and identify risk factors associated with RT. The first published use was in 2016 for a national RT estimate in the United States during 2011–2014 (Yuen et al., 2016). This paper also provided the first analysis of characteristics associated with RT. The authors found that 14% of TB cases nationally were attributable to RT with the burden of RT geographically heterogeneous and not explained by TB incidence; five of the eight states with very low TB incidence had counties in which ≥20% of genotyped cases were attributed to RT. In a multivariable analysis, characteristics associated with RT included younger age groups, being of a racial/ethnic minority group, experiencing homelessness, and substance use.

Starting in 2016, CDC began publishing national estimates of RT by county in their annual report of TB in the United States; in 2020–2021, 12.3% of TB cases nationally were attributed to RT (CDC, 2022b). Estimates of RT were higher among children aged 5–14 years (33.3%), among Non-Hispanic Native Hawaiian or Other

Pacific Islander persons (35.4%), among those residing in a correctional facility at the time of TB diagnosis (26.7%), among persons experiencing homelessness (27.1%), and among those who inject drugs (29.9%) (CDC, 2022b).

Since the first publication, the RT algorithm has been used to estimate RT in various subgroups. Kim et al (2022) estimated RT among birth cohorts of older age groups (RT estimate = 20% among US-born persons aged 50–90 years). Self et al (2021) used the RT algorithm to evaluate the role of RT as a driver of high TB incidence rates among persons experiencing homelessness (RT estimate = 37%; in 9 areas with higher than average TB incidence, RT estimate = 23%–82%). Others have used it to estimate the contribution of RT among persons living with HIV (RT estimate = 20%) (Schmit et al., 2020), the percent of multidrug-resistant TB that is attributed to RT (RT estimate = 6.7%) (Chen et al., 2020), and the percent of TB among healthcare personnel that is attributed to RT (RT estimate = 10%) (Mongkolrattanothai et al., 2019). The RT algorithm has also been used to find the inverse of RT — the proportion of TB due to reactivation of LTBI (Yelk Woodruff et al., 2021) — to estimate county-level LTBI prevalence in the United States (Haddad et al., 2018) — and most recently, to estimate the annual risk of LTBI reactivation into TB disease (Ekramnia, 2023).

Although the RT algorithm has been used frequently to estimate RT and characteristics associated with RT nationally and in select risk groups, estimates of RT specific to rural areas have not been explored, and characteristics associated with RT in rural areas remain unknown. A literature review of all articles indexed in PubMed with search terms of ["recent transmission" AND tuberculosis AND rural] returned 24 results; 20 were publications reporting on TB outside the United States. Three of the U.S.-based publications were reporting on declines in TB incidence in Arkansas (Berzkalns et al., 2014; Braden et al., 1997; France et al., 2007), the fourth was reporting on TB in Florida with one cluster in a rural part of Florida (Séraphin et al., 2016); none of the U.S.-based studies were recent enough to use the RT algorithm.

Tuberculosis Transmission and Outbreaks in Rural Areas

Because RT is a precursor for TB outbreaks (Raz et al., 2022; Yuen et al., 2016), reviewing the science on characteristics associated with TB outbreaks is important, especially considering there are no studies examining RT in rural areas.

National-Level Reports of Outbreaks

Outbreaks Investigated by CDC

National-level retrospective analyses of TB outbreaks investigated by CDC have found a high-proportion of persons involved in outbreaks had similar characteristics as those attributed to RT, including birth in the United States, being of non-Hispanic Black race/ethnicity, substance use, having markers of advanced disease and infectiousness (cavity on chest radiograph and acid-fast bacilli sputum smear positive), being incarcerated, and experiencing homelessness (Mindra et al., 2017; Mitruka et al., 2011). In the combined 48 outbreaks reported in these two publications, delayed diagnosis was a frequently cited factor contributing to the outbreak, including delays in seeking care and delays in diagnosis once under medical care. Incomplete contact investigations for pulmonary TB cases also played a role as did congregate settings such as overnight facilities serving persons experiencing homelessness and prisons. Transmission also occurred in households as well as in houses where people congregated to drink alcohol and in houses known as venues for buying and selling illicit drugs. Notably, neither study included rurality as a variable, therefore it is unknown which of these 48 outbreaks occurred in rural areas or whether characteristics associated with rurality contributed to any of the outbreaks.

Large Outbreaks of Tuberculosis in the United States

In 2022, Raz and colleagues published a study describing characteristics of large TB outbreaks in the United States using molecular surveillance methods (Raz et al., 2022). This publication was an important contribution to the literature on TB outbreaks nationally since the previous national-level reports (Mindra et al., 2017; Mitruka et al., 2011) only included outbreaks that were investigated by CDC, thus biasing the reporting to outbreaks and settings in which assistance was requested from federal investigators. In the 24 large outbreaks, defined as 10 or more cases related by RT, during 2014–2018, Raz et al identified many familiar themes including high proportions of U.S.-born people in racial and ethnic minority groups, people experiencing homelessness or incarcerated, and people using substances. Congregate settings were the primary site of transmission for eight outbreaks, eight of the outbreaks reported transmission primarily in non-familial social networks, and seven reported transmission primarily occurring within a family. Diagnostic delays were frequent and were postulated to reflect barriers to healthcare access. The authors also found that seven of the 24 outbreaks occurred in geographically remote or rural settings; of those, five outbreaks occurred primarily in counties considered to have moderate-to-high or high social vulnerability (CDC SVI >0.5). Although this study provided important insights not captured by others, it still leaves many holes for understanding RT in rural areas. First, this study only reported on large outbreaks of 10 or more cases, an outlier in terms of transmission; most TB clusters never grow beyond two cases (Yuen et al., 2016). Second, although this study identified outbreaks occurring in rural areas, no characteristics associated with the outbreaks were stratified by rural-urban designation or by geographic location, therefore comparisons of characteristics between rural and non-rural areas or in specific geographic regions could not be made. Finally, the designation of rural was undefined.

Individual Outbreak Reports in Rural Areas

To better encapsulate all the literature reported on TB outbreaks in rural areas in the United States, a review of the literature up to 2022 was conducted in PubMed, CINAHL, Web of Science, and Embase (see Appendix A for search terms). The review identified nine publications reporting on investigations of TB outbreaks in rural areas in the United States (Bloss et al., 2011; Bock et al., 1998; CDC, 2002; Dillaha et al., 2002; Funk, 2003; Labuda et al., 2021; Lowther et al., 2011; Phillips et al., 2004; Valway et al., 1998). Publication dates ranged from 1998 to 2021. Rural was defined inconsistently among the reports. In four publications, it was defined by the population size of the county where the outbreak was centered; in others the setting was either undefined or stated to be in a rural state or community. Table 1 provides the years and states involved as well as the definition provided for rurality. Below the table are additional key findings from the nine reports.

Table 1

Reports of Tuberculosis Outbreaks Occurring in Rural Areas

Author (year)	Outbreak year(s)	State	Definition of rural
Valway et al	1994–1996	Tennessee and Kentucky	Two counties with total population <14,000
Bock et al	1996–1997	Georgia	Population <10,000 in county ~ 400 square miles
CDC	2000–2001	Montana	"Rural Montana" – American Indian Reservation
Dillaha et al	1993–1999	Arkansas	"Rural county", population of 21,000
Funk	2000	Alaska	Not defined
Phillips et al	2001	Missouri	Not defined
Lowther et al (2011)	2008	Minnesota	Three adjacent rural counties with total population of 42,429
Bloss et al (2011)	2005–2007	Mississippi	Population <10,000
Labuda et al (2021)	2010–2018	Arkansas	"Rural county" and "Rural, impoverished community"

Transmission through Social Networks

A key finding from the review was that all nine publications report on outbreaks occurring in communities through close social networks. In an outbreak occurring in Georgia, initial cases were among people who participated in regular card playing and gambling that rotated locations among several homes and a work site. A local nightclub and county school were also named as potential sites of transmission (Bock et al., 1998). Similarly, an outbreak occurring among immigrants in Minnesota occurred due to a person with infectious TB playing in a band that frequented homes, community centers, church services, and other functions (Lowther et al., 2011). Other outbreaks were reported to occur among neighborhoods in close proximity (Bloss et al., 2011; Dillaha et al., 2002; Funk, 2003; Labuda et al., 2021) with a close social network that involved drinking and smoking inside one home reported as a primary site of transmission for one of the outbreaks (Bloss et al., 2011). Other social networks were closer still involving primarily family members and school contacts (Phillips et al., 2004), an Indian Reservation in rural Montana (CDC, 2002), and transmission involving a clothing factory (Valway et al., 1998). Notably, not one of these outbreaks in rural areas occurred among people residing in a congregate setting such as a prison or shelter, a difference noted from national outbreak reports.

Affected Populations

Of the outbreaks that report the race and/or ethnicity of people involved (n = 8), all except one of the outbreaks occurred among minority populations in the United States; in the one exception, the probable source case for many of the secondary cases is described as a U.S.-born White man, however race, ethnicity, and country of birth are not noted for most of the other cases (Valway et al., 1998). Other outbreaks occurred primarily among U.S.-born Black people (Bloss et al., 2011; Bock et al., 1998; Dillaha et al., 2002; Labuda et al., 2021), American Indians (CDC, 2002), Alaska Natives (Funk, 2003), and Guatemalan immigrants (Lowther et al., 2011). Race and ethnicity was not reported for cases and contacts in one of the publications (Phillips et al., 2004). Similar to observations for outbreaks nationally (Mindra et al., 2017; Mitruka et al., 2011), most outbreaks in rural areas occurred among U.S.-born people (Bloss et al., 2011; Bock et al., 1998; CDC, 2002; Dillaha et al., 2002; Funk, 2003; Labuda et al., 2021; Valway et al., 1998).

Risk Factors for Transmission

Delayed diagnosis of TB disease was reported as a factor facilitating transmission in all except two (CDC, 2002; Dillaha et al., 2002) of the nine outbreak reports. Substance use was reported as a contributing factor in three of the outbreaks (Bloss et al., 2011; Dillaha et al., 2002; Labuda et al., 2021); in another, alcohol use was reported as contributing to transmission (CDC, 2002). In two of the reports, inadequate ventilation in a school (Phillips et al., 2004) and a workplace setting (Valway et al., 1998) was theorized to exacerbate transmission. Three reports cited a hesitance among patients to name contacts (Bock et al., 1998; Dillaha et al., 2002; Labuda et al., 2021), who then went unevaluated for TB infection and subsequently developed TB disease. Cultural or language barriers between public health/clinical staff and patients were cited as a contributing factor in two reports (Funk, 2003; Lowther et al., 2011) as was distrust of the public health system (Dillaha et al., 2002; Labuda et al., 2021), which could also be the result of a cultural barrier, even if not expressly stated. Many of the communities in these rural areas were reported to be difficult to physically access, either due to their remote location or migratory patterns of community members, making evaluation and treatment efforts very challenging (Bloss et al., 2011; Bock et al., 1998; Funk, 2003; Lowther et al., 2011). And finally, insufficient public health resources to address the outbreak were reported as a facilitating factor for six of the outbreaks (Bloss et al., 2011; Bock et al., 1998; CDC, 2002; Funk, 2003; Labuda et al., 2021; Lowther et al., 2011). Frequently, local health departments responding to outbreaks had few clinical staff members with limited TB experience and expertise.

Summary of Individual Outbreak Reports in Rural Areas

Although these nine outbreak reports show some consistent themes, each is highly localized to the setting and time in which the outbreak occurred, some of which were almost 30 years ago. None of these reports provides a national perspective of outbreaks in rural areas or an understanding of RT in rural areas that does not result in an outbreak and a large investigation.

Theoretical Framework: Syndemic Theory in the Literature

This study was guided by the Syndemic Theory, a framework for studying diseases and social conditions that exist simultaneously and synergistically. The study of syndemics is a growing field; a 2022 “meta-knowledge analysis” identified 830 articles globally in the field of syndemics with the number of articles gradually increasing since 2003 (Hossain et al., 2022). Likely because Syndemic Theory was first used to study clustering of HIV with other diseases and social problems (Singer, 2006), the preponderance of articles focus on HIV and AIDS related syndemics (Hossain et al., 2022). Due to the synergistic and deadly combination of HIV infection and TB (Kwan & Ernst, 2011), it is not surprising that the inclusion of TB in HIV-related syndemic research occurred early, likely due at least in part to the call by the World Health Organization to create a collaborative approach to address the dual epidemic of TB and HIV in 2004 (WHO, 2004).

Syndemic of TB and Diabetes

While TB and HIV remain an important syndemic globally, a small minority (4.2% in 2021) of persons diagnosed with TB in the United States now have an HIV coinfection (CDC, 2022b). Instead the focus has shifted to understanding the syndemics of TB and other diseases (Trajman et al., 2022), including TB and diabetes (Houston et al., 2022; Magee et al., 2018; Mendenhall et al., 2017). The study of the convergence of TB and diabetes is important as diabetes has become the most frequently reported (23.9% in 2021) medical risk factor among persons with TB disease (CDC, 2022b), and diabetes is known to increase the risk of TB disease among persons infected with TB (Jeon & Murray, 2008).

Diabetes is a common medical condition in the United States; more than 133 million Americans are living with diabetes or prediabetes (CDC, 2024b). Those of a racial or ethnic minority group have a higher prevalence of diabetes than non-Hispanic White persons with the highest prevalence among American Indian or Alaska Native persons (14.5%) and non-Hispanic Black persons (12.1%) (CDC, 2024b). Among those with TB disease, the prevalence of diabetes among American Indian or Alaska Native persons and U.S.-born non-Hispanic Black persons is higher at 22.1% and 18.2%, respectively. This syndemic is further complicated by the geographic location of the diabetes belt, consisting of 644 counties in 15 mostly southern states (Barker et al., 2011), an area that has been subject to a number of documented TB outbreaks in rural areas (Bloss et al., 2011; Bock et al., 1998; Labuda et al., 2021; Valway et al., 1998).

In examining diabetes and TB from a syndemic perspective, Magee and colleagues (2018) published a review of the convergence of non-communicable diseases and tuberculosis, highlighting the pathophysiologic mechanisms and epigenetic factors in which the two diseases contribute to the other. Another review article takes the syndemic further and examines the interaction of TB, diabetes, and poverty (Mendenhall et al., 2017), highlighting that social factors associated with poverty, such as crowding, stress, and poor access to healthcare, increase the risk of acquiring both TB and diabetes, and complicate treatment and outcomes. The authors stress that research must recognize how social and medical problems cluster and interact within certain populations. Another article on this topic focuses on the syndemic caused by TB, diabetes, and detainment by U.S. Immigration and Customs Enforcement (ICE) (Houston et al., 2022). The authors construct a model arguing that ICE detention leads to malnutrition, medical neglect, chronic stress, and crowding — all of which lead to immune suppression, thus increasing the risk for TB and diabetes, and amplifying symptomology and poor outcomes. Similarly, authors in another study examine the interaction between chronic kidney disease and TB along with socioeconomic and demographic factors more prevalent in low- to middle-income countries, bringing attention to what they call "an emerging global syndemic" (Romanowski et al., 2016).

Syndemic Approach to TB Prevention and Control

Others have called for a syndemic approach to outbreak prevention and control citing that "social determinants of health are important contributors to the emergence and persistence of outbreaks; a number of infectious disease threats will not be optimally controlled unless social determinants of health are addressed; and using a syndemic lens can help inform novel approaches in clinical care" (Andermann, 2017). In a similar approach, another commentary focuses specifically on TB, emphasizing that current TB control policies ignore the impact of poverty on the susceptibility of immigrants to TB (Reitmanova & Gustafson, 2012) and that a syndemic approach to studying infectious diseases is necessary. Although the emphasis of this paper is on immigrants in Canada, the argument would also hold true for U.S.-born persons who are living in poverty.

As illustrated above, Syndemic Theory and use of a syndemic framework in research is gaining traction in the public health and clinical literature. Although this approach has been used predominantly to examine HIV-related syndemics, it has also been used to examine TB-related syndemics, most notably syndemics involving TB and diabetes. Interestingly, the theory analysis papers found in the literature have been narrative ones rather than ones testing the constructs of the theory using empirical evidence. Although commentaries from Canada focused on the importance of studying and addressing TB through the lens of social determinants of health, no studies were found that examined the potential syndemic caused by TB and rurality, recognizing that those in rural areas face several risk factors for TB, including higher poverty, decreased access to healthcare, and higher prevalence of diseases like diabetes.

Syndemic Theory Applied in Rural Areas

Syndemic Theory has been used to examine other diseases specifically in rural areas, most notably HIV and opioid-use disorders. In one paper, the syndemic of HIV, opioid overdose, and hepatitis C virus (HCV) infection are studied for county-level vulnerabilities using 29 county-level indicators in South Dakota (Wesner et al., 2020). The authors found that the most vulnerable counties were majority rural (62%) and part of a reservation (69%). Further, highlighting the role that community characteristics play in a syndemic, counties with higher rates of poverty, under-insurance, and lower coverage of primary care had high HCV infection rates. Another article published in 2018 argues for novel research approaches to understanding health disparities among minority populations in rural areas recognizing that syndemics of rural deprivation and chronic comorbidities contribute to poor health outcomes in Black Americans (Apostolopoulos et al., 2018).

Summary

TB disease is one of the oldest recorded human diseases, and general risk factors for TB infection and progression to TB disease are well known. As laboratory methods for characterizing relatedness of TB isolates have improved, epidemiologists have developed methods for estimating RT of TB that can be applied to national surveillance data, allowing for analyses of characteristics associated with RT nationally. Because RT leads to outbreaks, this is an important breakthrough for understanding TB epidemiology. However, these methods have never been applied to examine characteristics associated with RT in rural areas. Further, although analyses have been done of TB outbreaks nationally, none of these national-level analyses have described characteristics associated with outbreaks in rural areas. Of the outbreak reports in rural areas that are available, only one reports on an outbreak that occurred within the previous 10 years, and none provides a national perspective on TB outbreaks in rural areas. In the United States, rural areas are disproportionately impoverished and have higher rates of diseases such as diabetes. Together, these characteristics create a potential syndemic that can compound health disparities for people living in rural areas, however this potential syndemic has not been studied in the context of TB. This review demonstrates an important gap in the literature for understanding characteristics associated with recent transmission in rural areas and the role co-existing diseases and characteristics of the community have in propagating RT of TB.

CHAPTER 3

METHODOLOGY

This study used retrospective, cross-sectional, secondary data from the National Tuberculosis Surveillance System, the American Community Survey, and Area Health Resources Files from the Health Resources and Services Administration to conduct a descriptive, bivariate, and multi-level, multivariable analysis to explore factors associated with tuberculosis recent transmission in rural areas of the United States.

Data Sources

National Tuberculosis Surveillance System

Tuberculosis is a reportable disease in the United States; all 50 states and the District of Columbia require that every diagnosis of TB be reported to the local or state health department. State health departments then report every deidentified case of TB to CDC through the National Tuberculosis Surveillance System (NTSS) using the Report of Verified Case of Tuberculosis (RVCT). This report has been in effect since 1993 and includes individual patient data on demographic characteristics, geography, clinical characteristics, risk factors associated with TB, and laboratory results. This analysis used version 2 of the RVCT that was implemented in 2009 and was in effect during the study period. The full report in the version implemented in 2009 includes 49 original variables; of those, this analysis includes 29 variables, some of which are derived variables from the original 49 (Appendix B: Data Dictionary). Although some states require reporting of cases of latent TB infection (LTBI), these are not currently systematically reported to

CDC. Therefore, this analysis only included information on cases of TB disease, not LTBI.

Included Variables

The NTSS variables used for this study are detailed in the Appendix and can be grouped into three categories: 1) those used for inclusion criteria and dataset linking, 2) categorical demographic variables, and 3) categorical clinical variables. Variables used for inclusion criteria and matching included the date the patient's TB case was counted by the state, the state name, and the county name in which the case was counted. Most of the demographic variables that were included were ones previously shown to be associated with recent transmission or outbreaks in previous studies, including age group, origin of birth, race and/or ethnicity, alcohol and other drug misuse, experiencing homelessness, and incarceration at time of diagnosis (Labuda et al., 2021; Mindra et al., 2017; Raz et al., 2022; Yuen et al., 2016). Clinical variables included those that increase a person's risk of progression to TB disease such as HIV co-infection (Markowitz et al., 1997; Moreno et al., 1993), diabetes (Jeon & Murray, 2008), and other immunocompromising conditions (Pai et al., 2016).

American Community Survey

Using the individual's county of residence at time of TB diagnosis for linking, individual-level case data was linked to community-level data from the American Community Survey (ACS) 5-year estimates of the percent population below the federal poverty level (ACS code S1701) (United States Census Bureau, 2023a). ACS is an ongoing general household survey administered by the U.S. Census Bureau. Because only the 5-year estimates provide data for areas with populations of <20,000, these estimates were used instead of annual estimates. Five-year estimates from the 2011–2015 ACS release were linked to cases from 2011–2015 and 5-year estimates from the 2016–2020 ACS release were linked to 2016–2021 cases. Poverty estimates were included as continuous variables in all analyses.

Health Resources and Services Administration

Using the individual's county of residence for linking, individual-level case data were linked to county-level data from the Health Resources and Services Administration's (HRSA) Area Health Resource Files (AHRF) to determine the number of primary care physicians per 100,000 persons in each county. The county-level AHRF datasets include county totals of primary care physicians (MD and DO counts derived from the American Medical Association Masterfile) and county population estimates derived from the U.S. Census (Health Resources and Services Administration, 2024). It was initially proposed that this study would use county-level estimates of Health Professional Shortage Areas (HPSA) from HRSA. An HPSA is a geographic area, population group, or healthcare facility that HRSA has been designated as having a shortage of health professionals. Including HPSA estimates was not feasible for three reasons. First, HPSA estimates are derived from an index that includes the percent of the population below the federal poverty level. Because this measure was being included as a stand-alone county-level variable, including it as part of an index in the same study would lead to multicollinearity. Second, the HPSA dataset included duplicate FIPS codes for counties, leading to problems with linking, and third, some counties did not have estimates available. For all these reasons, the AHRF were identified as a better option for understanding county-level access to healthcare. Three county-level AHRF files were included to obtain estimates for each study year. The AHRF 2019‒2020 dataset included counts for study years 2011‒2018. The AHRF 2021‒2022 included counts for 2019 and 2020 and the AHRF 2022‒2023 dataset included counts for 2021. All three datasets were merged with NTSS and Census datasets. The year-specific number of primary care physicians in each county was divided by the population estimate for each county for that year and the corresponding ratio was multiplied by 100,000 to derive the ratio of primary care physicians per 100,000 persons for each county and year. These ratios were included as continuous variables in all analyses.

Recent Transmission Algorithm

The dependent variable in this study was attribution to recent transmission (RT) — that is, was the case of TB disease due to recent transmission of TB rather than reactivation of a latent infection acquired in the remote past? Once a person has TB disease, it is not possible to determine, based on clinical findings alone, whether that person's disease was attributable to RT or to reactivation of LTBI (Geng et al., 2005). Because knowing when RT is occurring guides critical public health action, many approaches have been used to determine RT (Kempf et al., 2005; McNabb et al., 2004; Moonan et al., 2012). In the United States, these approaches all take advantage of universal genotyping of culture-confirmed cases of TB (Ghosh et al., 2012). Genotyping allows for comparison of the molecular relatedness of *Mycobacterium tuberculosis* isolates; isolates that share the same genotype are more likely to be related by RT (McNabb et al., 2004). Using an RT algorithm to identify when recent transmission is occurring allows for estimations of RT in the absence of a field investigation, long considered the gold standard for understanding and validating TB transmission (France et al., 2015).

To find a universal approach for determining RT, the Division of TB Elimination at CDC compared four RT methods using gold-standard field investigation data from three states (France et al., 2015). The accuracy of the various methods ranged from a low of 88.5% for a state-based clustering method to a high of 94.4% using a novel plausible-source case method. A plausible-source case was defined as a case that 1) has a matching genotype with the secondary case, 2) involved a respiratory form of TB (pulmonary, laryngeal, pleural, or miliary) in a patient over 10 years of age, 3) was diagnosed within 2 years prior to the secondary case, and 4) resided within 10 miles of the secondary case. If a plausible source-case is identified, the secondary case is tagged as attributable to RT (France et al., 2015). Since 2015, the plausible-source case method has been CDC's standard for estimating RT in the United States.

The RT attribution variable is available for most culture-confirmed cases reported since 2011 (CDC, 2022b) and is usually referred to as the RT algorithm, which is how this dissertation will refer to it. In the NTSS dataset, applicable cases are tagged as attributable to RT or not attributable to RT using this validated algorithm. This variable was used as the dependent variable for analyses.

Multiple Imputation of Missing Recent Transmission Value

One key limitation of the RT algorithm is that it only includes cases with culture-confirmed disease, therefore missing approximately 20% of total cases (CDC, 2022b). Because children are frequently unable to produce a sputum specimen due to paucibacillary disease but have disease that is almost always the result of recent transmission (CDC, 2023b), imputing these missing values might create a more robust dataset. Because RT is likely missing at random conditional on demographic variables such as age and U.S.-born status, multiple imputation is a method that can be used to produce unbiased results while providing imputed values (Donders et al., 2006). This method makes use of the available observed data to create datasets with imputed RT outcomes and allows for a random selection of RT results from the distributions of the resulting datasets (Donders et al., 2006). Multiple imputation includes variance associated with the uncertainty inherent in imputed values, and therefore produces more accurate standard errors and confidence intervals than using single imputation or a missing indicator to accommodate missingness (Allison, 2002).

Although regression-based multiple imputation for RT had been planned for this dissertation, it was not implemented due to complications with the data analysis software. Instead, this analysis used a complete case analysis; missing data were noted and excluded from analysis similar to most published literature related to RT (Schmit et al., 2020; Self et al., 2021; Yuen et al., 2016).

Rural Area Classification Scheme

Using the county identifier, each case was classified as rural or urban based on the National Center for Health Statistics (NCHS) classification scheme (National Center for Health Statistics, 2014). Using this scheme, a metropolitan area contains a core urban area of ≥50,000 population. Nonmetropolitan areas are divided by micropolitan areas and noncore areas. Micropolitan areas contain an urban core of ≥10,000 population, but

<50,000. Noncore areas are those that are part of neither metropolitan nor micropolitan areas. This classification scheme is applied often in public health research (Ivey-Stephenson et al., 2017; James et al., 2018; Matthews et al., 2017; Moy et al., 2017). Some entities (e.g., American Hospital Association) and studies (Moy et al., 2017) consider any area that is not classified as metropolitan to be rural, while others only classify those in noncore areas as rural (James et al., 2018; Strosnider et al., 2017).

For this analysis, counties matching codes 1–4 (i.e., metropolitan counties) were considered urban. Counties matching code 5 or 6 (i.e., nonmetropolitan counties) were considered rural. Rural counties were further disaggregated into micropolitan counties (code 5) and noncore counties (code 6).

Inclusion and Exclusion Criteria

To build the final dataset, each county classified as nonmetropolitan was considered for inclusion. Second, each included county was analyzed to determine if there was at least one case of TB disease reported from the county during the study years; counties without a TB case were excluded from the analysis. The final dataset included all cases of TB disease during 2011–2021 that were reported from a county classified as nonmetropolitan; however, only cases with an RT designation were included for analyses pursuant to the research questions.

Data Management

The student principal investigator obtained approval for this study by the NTSS Analytic Steering Committee and was given access to the corresponding NTSS dataset. This dataset was stored on a CDC share drive that was only accessible using a CDC-issued laptop and was protected though a logical access card and personal pin. All analyses were conducted using the CDC-issued laptop and resulting datasets stored on the CDC share drive.

Variables from the ACS and AHRF were publicly available for download as csv and excel files. These files were also stored on the CDC share drive folder designated for this study. All analyses were conducted using RStudio version 4.3.2

Protection of Human Subjects

Before this analysis began on October 19, 2023, this dissertation was determined to be not human subjects research by the Institutional Review Board at Georgia State University (IRB Number: H24199, Reference Number: 377027). All data used in this study were from secondary data sources; no additional subjects were recruited, and no prospective data were collected. Although each patient represented in the dataset is assigned a unique state case number, the personally identifiable information linked to the state case number was retained by the state TB program and not included in the dataset and was not provided to the student principal investigator. NTSS data are collected for public health surveillance purposes and are protected by an Assurance of Confidentiality, a formal confidentiality protection authorized under Section 308(d) of the Public Health Service Act (CDC, 2017). All practitioners and researchers with access to NTSS data, including the student principal investigator, completed Assurance of Confidentiality training and signed agreements before obtaining access. ACS and AHRF variables are from publicly available datasets and are only reported in aggregate; individual person data is not reported or provided.

Analytic Approach

Research Question #1:

Are there differences in demographic, clinical, social-behavioral, and community-level characteristics between people in rural areas whose TB disease is attributed to recent transmission compared to people whose disease is not?

To answer this first research question, a descriptive analysis was performed to examine and compare frequencies and proportions of characteristics of people with TB disease in areas classified as rural according to the NCHS classification scheme detailed above. Cases were stratified into those attributed to recent transmission and those not attributed to recent transmission to allow for direct comparison. A bivariate analysis using Pearson's chi-square tests was used to determine if the proportions of the two groups were equal. The Mann-Whitney U test was used for comparisons between continuous variables. The resulting p-value was reported; a p-value of $<.05$ was considered statistically significantly different.

Sensitivity Analyses

Two sensitivity analyses were performed. First, cases from Alaska were excluded and the analysis re-run. Due to the remote nature of many Alaska villages, there is more molecular uniformity in some longstanding *M. tuberculosis* genotypes observed among Alaska Natives with TB disease. In addition, health officials conduct "TB Sweeps" during the summer months to identify to new cases of TB infection and disease (State of Alaska; Department of Health, 2021). Because the RT algorithm is dependent on both genotype matching and identification of a source case within the preceding 2 years, both of these factors (spurious genotype matches and artificial time clustering) could contribute to falsely elevated RT estimates among TB cases in rural Alaska.

The second sensitivity analysis excluded cases among people incarcerated in correctional facilities. Because people incarcerated in correctional facilities are sometimes reported from an administrative office address rather than the location of the facility (Stewart et al., 2022) and because factors associated with transmission might be different inside correctional facilities compared to the rest of the community, it was important to understand what impact, if any, that had on the results.

Research Question #2:

Are there regional differences in demographic, clinical, social-behavioral, and community-level characteristics of people with TB disease attributed to RT in rural areas of the United States?

To answer this second research question, a descriptive analysis was performed to examine and compare frequencies and proportions of characteristics of people with TB disease attributable to RT in areas classified as rural according to the NCHS classification scheme detailed above. Cases were further stratified by U.S. Census Regions: Northeast, Midwest, South, and West (United States Census Bureau, 2023b). Fisher's exact test was used to estimate significant differences in proportions among the four groups due to small cell sizes primarily in the Northeast Region. The Kruskal-Wallis test was used to compare differences among continuous variables in each of the regions. The resulting *p*-value were reported; a *p*-value of <0.05 was considered statistically significantly different.

As with Research Question #1, two sensitivity analyses were performed, first excluding cases reported from Alaska, and, secondly, excluding cases among those diagnosed while incarcerated.

For Research Questions #1 and #2, statistical testing was performed using Pearson's chi-square or Fisher's Exact test for categorical variables and the Mann-Whitney U or Kruskal Wallis test for continuous variables. The latter of each were used when the cell count or sample size were too small to meet the assumptions of the former. Chi-square and Fisher's Exact test assumptions state that 1) data should be in the form of frequencies or counts, 2) observations should be categorical on a nominal or ordinal scale, 3) each observation must be independent of the other observations, 4) the data should be representative of the population, and 5) row totals and column totals should be fixed. Additionally, chi-square tests should have expected frequency counts of at least 5. All the assumptions for chi-square tests were met; when the cell counts were too small, Fisher's Exact test was used. Assumptions of the Mann-Whitney U and Kruskal Wallis test include 1) independence of samples and 2) data is on an ordinal, interval, or ratio scale. Additionally, the Mann-Whitney U test states that the distribution of the two populations from which the samples are drawn should be similar in shape and of sufficient sample size. The assumptions of the Mann-Whitney U test were met for Research Question 1. For Research Question 2, the Kruskal Wallis test was used due to small and unequal numbers by regional group.

Research Question #3:

What is the relative association of individual-level (e.g., age, race/ethnicity) and community-level (i.e., poverty, primary care physician to population rate) characteristics *with recent transmission among people with TB disease in rural areas in the United States?*

To answer this third research question, a hierarchical generalized linear model (HGLM) using the logit function was constructed using individual-level independent variables from NTSS and county-level variables from ACS and HRSA. The outcome was a single binary variable indicating whether a patient's TB disease was attributed to recent transmission. To obtain sufficient power to conduct this study, a sample size of at least 400 TB cases was required (Green, 1991; Trickey, 2023). Ideally, the sample size would allow for five observations per highest level unit, which is the county for this study. To allow for this, the HGLM was run on a dataset that excluded counties with <5 cases. This dataset included 4,004 cases, well above the 400-case threshold. Individual-level characteristics either known to be associated with recent transmission generally or suspected to be contributing to recent transmission in rural areas were included in the model; county-level characteristics included the percent of the county living below the federal poverty line and the primary care physician to population rate. Table 2 provides previous estimates for individual-level variables' association with RT and the *a-priori* expectation of county-level variation for individual level-variables. Except for diabetes, the estimates in the table are derived from the first national analysis of characteristics associated with RT (CDC, 2022b; Yuen et al., 2016) and the corresponding estimates in CDC's RT report for 2021 (CDC, 2022b).

Although previous national RT estimates have reported state-level variation (Yuen et al., 2016), variation by county and by county-level variables have not been previously reported. Therefore, the expectation of county-level variation for individual-level variables is based on expert opinion from the student principal investigator and her colleagues at CDC rather than on empiric evidence.

Table 2

Rationale for Inclusion of Variables in the Hierarchical Generalized Linear Model

Level one variables	**Association with recent transmission (RT)**	**Expectation of county-level variation for individual-level variables**
Age	Younger age consistently associated with RT, especially pediatric cases ≤4 years (range: 41.0%–52.9%) and 5–14 years (range: 26.2%–33.3%) [1,2]	No county-level variation in RT by age or age group. Younger age groups have consistently been associated with RT in previous studies and it is unlikely that this would vary by poverty or access to healthcare.
Place of birth (U.S. vs outside the U.S.)	Birth in the United States consistently associated with RT (range: 24.8%–27.4%) as compared to cases among people born outside the United States (range: 7.5%–7.7%) [1,2]	Odds of RT will be higher among U.S.-born people vs. non-U.S.–born people regardless of county-level factors. Expect odds of RT among U.S.-born people to be even higher in counties with higher poverty and less access to healthcare.
Race/ethnicity	Higher proportions of RT observed among people identifying as American Indian/Alaska Native (39.8%), Native Hawaiian/Pacific Islander (30.2%), and Black (25.1%) [2] In the most recent update by CDC, the estimates of RT for these groups were the following: American Indian/Alaska Native: 41.6%, Native Hawaiian/Pacific Islander (35.4%), and Black (19.6%). The lowest RT was estimated in those identifying as Asian (6.9%) [1]	Expect racial/ethnic minority groups will have higher odds of RT regardless of county-level factors; however, also expect higher odds of RT among racial ethnic groups in counties with higher poverty and less access to healthcare. The differentiation of race/ethnicity in RT attribution may also be regional. For example, in most of the South, the expectation is that non-Hispanic Black people will have the highest proportion of those attributed to RT whereas in the West, American Indian/Alaska Natives will probably have the highest.

Incarcerated at diagnosis	Higher proportions of RT among people incarcerated at diagnosis (range: 17.3% – 26.7%), compared with those not incarcerated (range: 11.9%–14.3%) [1,2]	Expect the effect of incarceration to only vary at the county level in terms of county-level poverty and not by access to healthcare. This is because the poverty of the county could be related to conditions within the correctional facility (i.e., impoverished counties may not have the staff or health systems in place in the facility that would prevent and halt TB transmission quickly). However, issues like being in a HPSA should not factor into transmission that occurs within a facility.
Experiencing homelessness	Higher proportions of RT observed among people experiencing homelessness (range: 27.1%–37.5%), compared with those not experiencing homelessness (range 11.6%–13.0%) [1,2]	Experiencing homelessness is usually associated with higher odds of RT. Expect this may vary more by degree of rurality than by other county-level factors, however it stands to reason that there might be more RT associated with experiencing homelessness in counties with higher poverty and less access to healthcare.
Substance use	Higher proportions of RT among people with excess alcohol use (range: 22.4%–27.1%) and substance use (33.3%) [2] When stratified by injection and non-injection drug use, proportions of RT were similar among those using non-injection drugs (31.6%), compared with injection drugs (29.9%) [1]	In counties of higher poverty and less access to healthcare, those who use substances and have less access to resources to address substance use may have higher odds of RT.
Diabetes mellitus	RT estimates have not been previously reported among people with diabetes, however, given that diabetes	It is possible the effect of diabetes on RT could vary by access to healthcare. For example, if a person had

	is known to increase the risk of developing TB disease among those infected (Jeon & Murray, 2008) and higher proportions of people with diabetes in rural areas as compared to urban (Dugani et al., 2021), this is an important variable to include.	diabetes, and lived in an area with poor access to healthcare, their diabetes may be more uncontrolled, thus increasing their risk.
Immunosuppression (HIV and otherwise)	Higher proportions of RT among people with HIV co-infection (range: 15.5%–22.4%), compared with those not co-infected (12.4%–14.4%) [1,2]	The effect of immunosuppression on RT could vary by access to healthcare. As above with diabetes, if a person is immunosuppressed and lives in a county with poor access to healthcare, they may experience differential (and probably higher) odds of having TB attributable to RT than someone who is immunosuppressed but lives in a place with good access to healthcare.

Table references

1. (CDC, 2022b)
2. (Yuen et al., 2016)

Hierarchical Generalized Linear Model

The model to address research question #3 was a hierarchical generalized linear model using the logit link function (hereafter HGLM). HGLM is a complex form of multiple logistic regression that is used to analyze variance in the outcome variable when the predicting variables are at varying hierarchical levels (Snijders & Bosker, 2012). In this dissertation, the varying hierarchical levels are the level of the individual, which are grouped within the level of the county. HGLM should not be confused with block-wise logistic regression model building, which is sometimes referred to as hierarchical regression.

At level 1, the outcome (the probability of recent transmission) was modeled using individual-level predictors outlined in Table 2. The intercept-only model tests whether the mean log-odds of transmission varies across counties included in the analysis. Model 2 adds individual-level predictors of RT as discussed in Table 2. Models 3 and 4 examine whether the county-level conditional log-odds of RT (conditioned on individual-level factors) vary as a function of PCP-to-population rates or of county-level poverty.

<u>Intercept only model</u>

$$Ln\ (P\left(\frac{RT}{1-\ RT}\right))_{ij} =\ \beta_{0j}$$

$$\beta_{0j} =\ \gamma_{00} + u_{0j}$$

<u>Model 2: Individual-level variables</u>

$$Ln\ (P\left(\frac{RT}{1-RT}\right))_{ij} =\ \beta_{0j} +\ \beta_{1j}age + \beta_{2j}birthplace + \beta_{3j}incarceration + \beta_{4j}homelessness + \beta_{5j}substance\ use + \beta_{6j}clinical + \beta_{7j}race/ethnicity$$

<u>Model 3: Includes county PCP-to-population ratio</u>

$$Ln\ (P\left(\frac{RT}{1-RT}\right))_{ij} =\ \beta_{0j} +\ \beta_{1j}age + \beta_{2j}birthplace + \beta_{3j}incarceration + \beta_{4j}homelessness + \beta_{5j}substance\ use + \beta_{6j}clinical + \beta_{7j}race/ethnicity +$$

$$\beta_{0j} =\ \gamma_{00} + \gamma_{01}pcp\ to\ pop\ rate +\ u_{0j}$$

<u>Model 4: Includes percent of county population below poverty</u>

$$Ln\ (P\left(\frac{RT}{1-RT}\right))_{ij} =\ \beta_{0j} +\ \beta_{1j}age + \beta_{2j}birthplace + \beta_{3j}incarceration + \beta_{4j}homelessness + \beta_{5j}substance\ use + \beta_{6j}clinical + \beta_{7j}race/ethnicity +$$

$$\beta_{0j} =\ \gamma_{00} + \gamma_{01}poverty +\ u_{0j}$$

Model-Fit Procedures

The first model built was a random intercept model to predict the probability of RT and estimate the variation in the outcome across counties. This intercept-only model did not include any covariates. The parameters and standard errors were estimated using the maximum-likelihood estimator with a logit link. In this type of model, the coefficients and variance terms are jointly estimated by maximizing the likelihood of the predicted values given the data. The second model included level-1 (individual-level) variables. The third model included the county-level variable of primary-care physician to population rate as a fixed effect. The fourth model included level-1 (individual-level) variables and the county-level variable of county population below federal poverty line as a fixed effect. Both county-level variables were not included in the same model due to correlation noted between the two variables. For both the third and fourth models, the adjusted odds ratios with 95% confidence intervals were reported for fixed effects; random effects included the intra-class correlation and tau-squared for the county.

Assumptions

There are four major assumptions to address when fitting the HGLM with a logit link (Stoltzfus, 2011). Figures showing results of assumption testing can be found in Appendix C.

1) Error terms should be independent, indicating that independent sampling is used; logistic regression analysis should not be used for within subjects or matched-pair studies. This study did not employ a within-subjects or matched-pair design; however, this assumption is partially violated in an HGLM because the model applies the effect of nested cases within counties.

2) There was a linear relationship between continuous predictor variables and the log odds (or logit) of the dependent variable.
3) Lack of multicollinearity among categorical independent variables in the model was established by running a variance inflation factor (VIF) test and confirming there were no variables with a VIF>4. There was evidence of multicollinearity among the county-level variables; therefore, the HGLM was run separately with each county-level variable.
4) Lack of strongly influential outliers in the data was addressed by examining standardized residuals for cases in the dataset. No standardized residuals >3 were observed.

CHAPTER 4

RESULTS

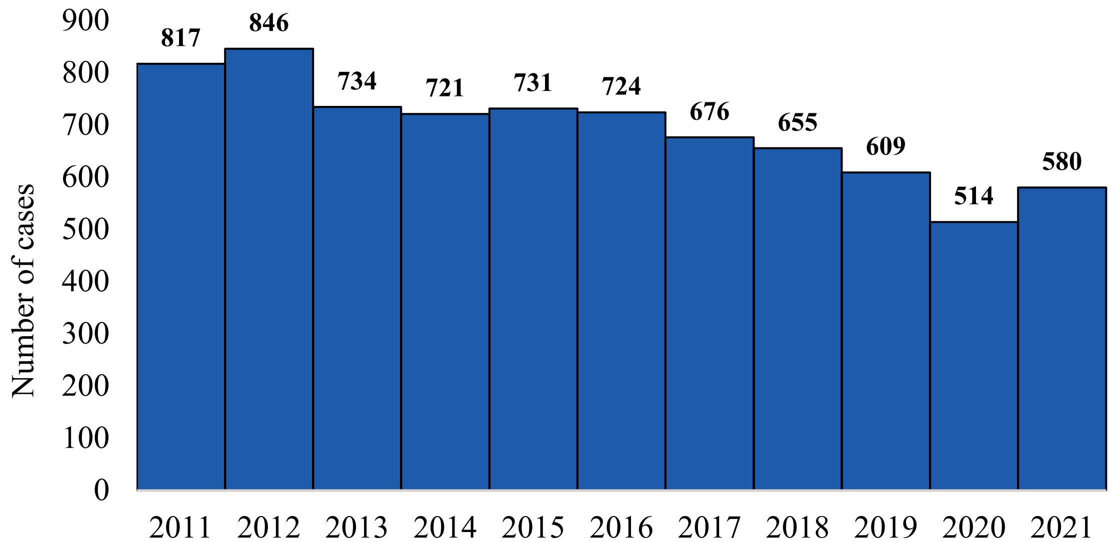
Chapter four discusses the results of three studies examining characteristics of patients diagnosed with TB disease in rural areas of the United States during 2011–2021 and the association of individual and county characteristics with the determination of whether the patient's TB disease was due to recent transmission (RT). Selected patient-level demographic characteristics included sex, age, origin of birth, and race/ethnicity. Selected patient-level social-behavioral characteristics included incarceration at diagnosis, housing status, and substance misuse. Selected patient-level clinical characteristics included diabetes, HIV co-infection, and non-HIV immunocompromising conditions, including the use of immunosuppressive therapies such as TNF-alpha inhibitors. Selected county-level characteristics included the U.S. Census Region, the percent of the county population living below the federal poverty level, and the number of primary care physicians per 100,000 population (PCP-to-population ratio).

Characteristics of patients with TB disease in rural areas

During 2011–2021, 7,607 patients diagnosed with TB disease in rural areas of the United States were reported to the National TB Surveillance System. The number of reported patients per year ranged from a low of 514 in 2020 to a high of 846 in 2012.

Figure 2.

Epidemic Curve of Patients Diagnosed with TB Disease in Rural Areas, United States, 2011–2021



Demographic Characteristics

Of all patients, 5,023 (66.0%) were male, 4,857 (63.8%) were born in the United States, and 5,616 (73.8%) identified as part of a racial/ethnic minority group (i.e., something other than White, non-Hispanic). The largest age category was 45–64 years, including 2,358 (31.0%) of the total reported patients with TB disease in rural areas (Table 3).

Table 3.

Demographic Characteristics of Patients Diagnosed with TB disease in Rural Areas of the United States, 2011–2021

Characteristic	**No**	**%**
Total	7,607	100.0
Sex		
Male	5,023	66.0
Female	2,582	33.9
Missing	2	0.0
Age (years)		

0–4	293	3.9
5–14	225	3.0
15–24	730	9.6
25–44	2,168	28.5
45–64	2,358	31.0
65+	1,832	24.1
Unknown/missing	1	0.0
Origin of birth		
U.S.-born	4,857	63.8
Non-U.S.–born	2,742	36.0
Unknown/missing	8	0.1
Race/ethnicity		
Hispanic	2,020	26.6
White, not of Hispanic ethnicity	1,964	25.8
Black, not of Hispanic ethnicity	1,755	23.1
Asian	806	10.6
American Indian/Alaska Native	778	10.2
Native Hawaiian/Pacific Islander	165	2.2
Multiple race, not of Hispanic ethnicity	92	1.2
Unknown/missing	27	0.4

Social-behavioral characteristics

Social risk factors for TB were present, but not ubiquitous. Misusing substances, including alcohol, non-injection drugs, and/or injection drugs, was reported among 1,467 (19.3%) patients; 820 (10.8%) were incarcerated at the time of diagnosis, and 303 (4.0%) were unhoused in the year before diagnosis (Table 4).

Table 4

Social-behavioral Characteristics of Patients Diagnosed with TB Disease in Rural Areas of the United States, 2011–2021

Characteristic	No	%
Total	7,607	100.0
Incarcerated at diagnosis		

Yes	820	10.8
No	6,746	88.7
Unknown/missing	41	0.5
Unhoused in previous year		
Yes	303	4.0
No	7,221	94.9
Unknown/missing	83	1.1
Substance misuse[1]		
Yes	1,467	19.3
No	6,000	78.9
Unknown/missing	140	1.8

1. Includes use of injection drugs, noninjection drugs, or heavy alcohol use in the past 12 months

Clinical Characteristics

Similarly, clinical factors that increase a person's risk of progressing from a latent TB infection (LTBI) to TB disease were present among some of the patients, but not a majority. Among the conditions examined, diabetes was the most frequent with 1,229 (16.2%) of all patients reporting diabetes (Table 5). Non-HIV immunocompromising conditions, including use of immunosuppressive therapies such as TNF-alpha inhibitors, were reported among 415 patients (5.5%); 251 (3.3%) of patients were co-infected with HIV.

Table 5

Clinical Characteristics of Patients Diagnosed with TB Disease in Rural Areas of the United States, 2011–2021

Characteristic	No	%
Total	7,607	100.0
Diabetes		
Yes	1,229	16.2
No	6,378	83.8
HIV co-infection		

Yes	251	3.3
No	6,397	84.1
Unknown/missing	959	12.6
Non-HIV immunocompromise[1]		
Yes	415	5.5
No	7,192	94.5
Tuberculosis disease attributed to recent transmission		
Yes	847	11.1
No	4,786	62.9
Unknown	1,967	25.9

1. Patient is immunocompromised because of either a medical condition that is not HIV (e.g., leukemia) or immunosuppressive therapy (e.g., TNF alpha inhibitor)

County Characteristics

Patients with TB disease in rural areas lived in counties that ranged from 5%–51% of the population living below the federal poverty line (median: 20) and in counties that had a mean of 49 primary care physicians per 100,000 persons (range: 0–419) (Table 6). The greatest number of patients lived in the South Region (4,840, 63.6%), followed by the Midwest (16.6%), the West (16.4%), and the Northeast (3.3%).

Table 6

Selected Characteristics of Counties Where Patients Diagnosed with TB Disease in Rural Areas of the United States Live, 2011–2021

Characteristic	**No**	**%**
Total	7,607	100.0
U.S. Census Regions		
Midwest	1,265	16.6
Northeast	253	3.3
South	4,840	63.6
West	1,249	16.4
NCHS[1] **rural designation**		
Micropolitan	4,380	57.6

Noncore	3,227	42.4
	Mean	**SD**
Percent county population living in poverty	20	7.0
PCP[2] per 100,000 persons in county	49	27.0

1. National Center for Health Statistics
2. Primary care physician

Recent Transmission

Overall, 847 patients had TB disease that was attributed to RT, accounting for 11.1% of the overall patient population; when excluding 1,967 patients for whom RT designation is unknown (25.9% of total patient population), 15.0% of the remaining patients' TB disease was attributed to RT (Table 5).

Research Question 1.

Are there differences in demographic, clinical, social-behavioral, and community-level characteristics between people in rural areas whose TB disease is attributed to recent transmission compared to people whose disease is not?

In rural areas, there were significant differences ($p < 0.05$) among most of the selected patient characteristics when comparing those whose TB disease was and was not attributed to RT; exceptions were differences in sex, the percent who were unhoused in the previous year, and the number of primary care physicians per 100,000 persons in the patient's county.

Significant Demographic and Social-Behavioral Differences

Most strikingly, 90.3% of patients whose TB was attributed to RT were born in the United States, in contrast to 60.4% of those whose TB was not attributed to RT. More patients whose TB was attributed to RT identified as being from racial/ethnic minority groups (i.e., not non-Hispanic White, [85.2%]), compared with those whose TB disease was not attributed to RT (71.0%). Patients whose TB was attributed to RT were younger, were more frequently incarcerated at diagnosis (17.4% vs. 9.9%) and reported more substance misuse (32.2% vs. 19.9%) (Table 7).

Table 7

Demographic and Social-Behavioral Characteristics of Patients Diagnosed with TB Disease in Rural Areas of the United States, by Attribution to Recent Transmission, 2011–2021

Characteristic	Attributed to recent transmission		Not attributed to recent transmission		*p*-value[1]
	No	**%**	**No**	**%**	
Total	847	15.0	4,786	85.0	
Sex					.120
Male	592	69.9	3,214	67.2	
Female	255	30.1	1,572	32.8	
Missing	0	0.0	0	0.0	
Age (years)					<.001
0–4	41	4.8	47	1.0	
5–14	33	3.9	36	0.8	
15–24	111	13.1	434	9.1	
25–44	307	36.2	1,323	27.6	
45–64	273	32.2	1,556	32.5	
65+	82	9.7	1,390	29.0	
Unknown/missing	0	0.0	0	0.0	
Origin of birth					<.001
U.S.-born	765	90.3	2,893	60.4	
Non-U.S.–born	82	9.7	1,889	39.5	
Unknown/missing	0	0.0	4	0.1	
Race/ethnicity					<.001
Hispanic	126	14.9	1,293	27.0	
White, not of Hispanic ethnicity	125	14.8	1,387	29.0	
Black, not of Hispanic ethnicity	273	32.2	1,041	21.8	
Asian	9	1.1	570	11.9	

American Indian/Alaska Native	268	31.6	361	7.5	
Native Hawaiian/Pacific Islander	22	2.6	68	1.4	
Multiple races, not of Hispanic ethnicity	24	2.8	46	1.0	
Unknown/missing	0	0.0	20	0.4	
Incarcerated at diagnosis					<.001
Yes	147	17.4	476	9.9	
No	696	82.2	4,286	89.6	
Unknown/missing	4	0.5	24	0.5	
Unhoused in previous year					.700
Yes	39	4.6	208	4.3	
No	800	94.5	4,526	94.6	
Unknown/missing	8	0.9	52	1.1	
Substance misuse[2]					<.001
Yes	273	32.2	951	19.9	
No	561	66.2	3,743	78.2	
Unknown/missing	13	1.5	92	1.9	

1. Statistical testing performed on observations with known results; unknown/missing results excluded
2. Includes use of injection drugs, noninjection drugs, or heavy alcohol use in the past 12 months

Significant Clinical Differences

All examined clinical characteristics were more frequently observed among patients whose TB disease was not attributed to RT (Table 8). Patients with diabetes comprised 19.0% of those whose TB disease was not attributed to RT, compared with 8.9% of those whose TB disease was attributed to RT. Less frequently observed overall was HIV co-infection (3.7% among non-RT vs. 2.5% among RT) and non-HIV immunocompromising conditions (6.2% among non-RT vs. 3.0% among RT).

Table 8

Clinical Characteristics of Patients Diagnosed with TB Disease in Rural Areas of the United States, By Attribution to Recent Transmission, 2011–2021

Characteristic	Attributed to recent transmission		Not attributed to recent transmission		*p*-value[1]
Total	847	15.0	4,786	85.0	
Diabetes					<.001
Yes	75	8.9	908	19.0	
No	772	91.1	3,878	81.0	
HIV co-infection					.040
Yes	21	2.5	179	3.7	
No	764	90.2	4,040	84.4	
Unknown/missing	62	7.3	567	11.8	
Non-HIV immunocompromise[2]					<.001
Yes	25	3.0	296	6.2	
No	822	97.0	4,490	93.8	

1. Statistical testing performed on observations with known results; unknown/missing results excluded
2. Patient is immunocompromised because of either a medical condition that is not HIV (e.g., leukemia) or immunosuppressive therapy (e.g., TNF alpha inhibitor)

Significant County-Level Differences

The vast majority (92.8%) of patients whose TB disease was attributed to RT lived in either the South or West region; among those who TB disease was not attributed to RT, 77.4% lived in the South or West region. The mean percent of the county population living below the federal poverty level was higher for patients whose TB disease was attributed to RT (24% vs. 20%) (Table 9) as was the median (24% vs. 19%) and IQR. As noted earlier, there were not significant differences in the mean number of primary care physicians at the county level (Table 9).

Table 9

Selected Characteristics of Counties Where Patients Diagnosed with TB Disease in Rural Areas of the United States Live, By Attribution to Recent Transmission, 2011–2021

Characteristic	Attributed to recent transmission		Not attributed to recent transmission		*p*-value
	No.	**%**	**No.**	**%**	
Total	847	15.0	4,786	85.0	
NCHS[1] rural designation					<.001
Micropolitan	439	51.8	2,849	59.5	
Noncore	408	48.2	1,937	40.5	
	Mean	**SD**	**Mean**	**SD**	
Percent county population living in poverty	24	7.0	20	7.0	<.001
PCP[2] per 100,000 persons in county	50	28.0	49	26.0	.092

1. National Center for Health Statistics
2. Primary care physician

Sensitivity Analyses

Sensitivity analysis 1. Excluding cases among patients in Alaska

Excluding cases reported from Alaska reduced the overall dataset by 459 observations, including 198 cases attributed to RT and 161 cases not attributed to RT. Among patients with an RT designation (n = 5,274), 12.3% were attributed to RT, an 18.0% reduction from the estimate in the full dataset of 15.0%. HIV co-infection was the only examined characteristic that was no longer statistically significant in this first sensitivity analysis. The number of patients with HIV co-infection among those attributed to and not attributed to RT remained nearly the same with the same number of patients (n = 21) with an HIV co-infection attributed to RT, and 177 (compared with 179) not attributed to RT. However, among those with negative tests for HIV co-infection, the percentages of those attributed and not attributed to RT nearly equalized when excluding Alaska (97% [n = 589] and 96% [n = 3,898], respectively), thus rendering the difference no longer statistically significant.

Other notable changes, although not changing the significance of the comparison, were the following. The percent of patients whose TB disease was attributed to RT and were incarcerated at diagnosis increased when Alaska was excluded, from 17.4% to 22.0%, suggesting that most RT associated with incarceration was occurring in other states. The percent of patients whose TB disease was attributed to RT in the West region decreased from 23.8% to 5.3%, suggesting that most RT occurring in the West region was occurring in Alaska. Finally, the mean and median number of primary care physicians per 100,000 population declined from 49 and 45, respectively to 41 and 38, suggesting that Alaska has more primary care physicians relative to their population than the mean rest of rural America (Appendix D, Table 1).

Sensitivity analysis 2. Excluding cases among patients incarcerated at diagnosis

Excluding the TB cases among persons incarcerated at diagnosis reduced the overall dataset by 820 observations, including 147 cases attributed to RT, 566 cases not attributed to RT, and 87 without an RT designation. Among patients with an RT designation (n = 5,010), 13.9% were attributed to RT, a difference of -7.1% from the full dataset. As with the first sensitivity analysis, HIV co-infection is the only characteristic that was no longer statistically different between those whose disease was and was not attributed to RT. The racial/ethnic makeup of patients whose disease was attributed to RT changed notably in the following ways: the percent of patients identifying as American Indian or Alaska Native increased from 31.6% to 37.6%, and the percent of patients identifying as Hispanic and non-Hispanic Black decreased from 14.9% to 11.4% and 32.2% to 29.3%, respectively. The county-level characteristics also held steady to those observed in the full dataset, except the mean and median of the primary care physicians for the population size among those whose TB disease was attributed to RT increased by 3 for both (Appendix D, Table 2).

Research Question #2:

Are there regional differences in demographic, clinical, social-behavioral, and community-level characteristics of people with TB disease attributed to RT in rural areas of the United States?

Attribution to Recent Transmission by U.S. Census Region

Most patients in rural areas whose TB disease was attributed to RT lived in the South Region (65.6%), followed by the West (27.2%), Midwest (6.4%), and Northeast (0.8%) regions. Although this pattern roughly aligns with the total volume of patients with TB disease in these regions (South: 4,840, Midwest: 1,265, West: 1,249, and Northeast: 253), the regional proportion attributed to RT among those with a known RT designation does not. The West Region accounted for 15.8% of the overall patient volume, but 27.2% of the patients whose TB disease was attributed to RT. In the West Region, 23.8% of patients' TB disease was attributed to RT (Table 10). The Midwest Region accounted for 16.6% of the overall patient volume, but only 6.4% of the overall patients whose TB disease was attributed to RT. Among those with an RT designation in the Midwest Region, 5.7% of the patients' TB disease was attributed to RT. Only seven patients in the Northeast Region had TB disease that was attributed to RT; this accounted for <1% of the overall number of patients with TB disease attributed to RT and 3.5% among patients with TB in the Northeast Region with a known RT designation. Attribution of TB disease to RT was most proportionate in the South Region, which accounted for 63.6% of patients overall and 65.6% of the total attributed to RT. The South Region RT estimate of 15.8% attributed to RT aligned most closely with national estimates of RT at approximately 12% (CDC, 2022b).

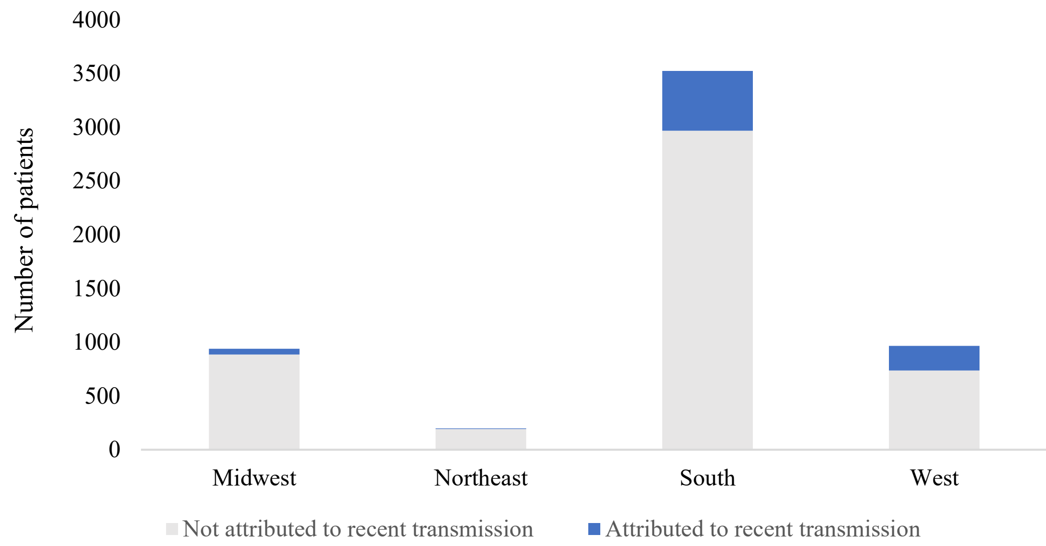
Table 10

Regional Counts and Proportions of Patients Whose TB Disease was Attributed to Recent Transmission in Rural Areas Among Patients with a Recent Transmission Designation, 2011–2021

Characteristic	Attributed to recent transmission		Not attributed to recent transmission		*p*-value
Total	847	15.0	4,786	85.0	
U.S. Census Regions					<.001
Midwest	54	6.4	887	18.5	
Northeast	7	0.8	193	4.0	
South	556	65.6	2,969	62.0	
West	230	27.2	737	15.4	

Figure 3

Attribution of TB Disease to Recent transmission, by U.S. Census Region, 2011–2021



Handling of Small Cell Counts

Due to small cell counts for variables among patients whose TB disease was attributed to RT, particularly in the Northeast Region (n = 7), two methods were explored for this analysis. First, the Midwest and Northeast regions were compared directly to determine if there existed significant differences in characteristics between the two regions. If no significant differences were identified, the two regions could be combined in the analysis. However, significant differences in several characteristics were identified for the two regions, including differences in origin of birth, incarceration at diagnosis, substance misuse, and the percent of population below the federal poverty level, so this method was not pursued further. In the second method, adopted for this analysis, age was reported as a continuous variable rather than as categories of age, and race/ethnicity identities were collapsed into two categories: 1) White, not of Hispanic ethnicity and 2) Other, being of a non-White minority group. Further, all clinical conditions examined in this analysis (i.e., diabetes, HIV co-infection, and non-HIV immunosuppression) were collapsed into one variable for "any clinical condition." Incarceration at diagnosis and reports of being unhoused in the previous year were removed from the regional analysis because there were no patients whose TB disease was attributed to RT in the Northeast Region who were either incarcerated at diagnosis or unhoused and only one patient for each of these risk factors in the Midwest Region.

Organization of results

First, the results will describe patient and county characteristics of patients whose disease is attributed to RT, stratified by U.S. Census Region. Second the results will describe significant differences between regions.

Midwest

As stated above, in the Midwest Region, among those with an RT designation, 5.7% of TB cases were attributed to RT. Rural patients whose TB disease was attributed to RT in the Midwest Region had a median age of 37 years (IQR: 22, 50). Nearly 80% were people born in the United States and approximately 33% identified as non-Hispanic White, the highest proportion in any of the regions. Approximately a third (37.0%) reported substance misuse. This region had the highest proportion of select co-existing clinical conditions at 24.1%. The median percent of the county population below the federal poverty line was 15 (IQR: 13, 20), and the median number of primary care physicians per 100,000 people was 47 (IQR: 35, 63), higher than the national median in this study.

Northeast

Findings from the Northeast Region should be interpreted with caution as only seven patients during the study years in rural areas the Northeast Region had their TB disease attributed to RT (3.5% of those with a known RT designation). Of those seven patients, more were male (57.1%), born in the United States (71.4%), and identified as part of a racial or ethnic minority group (71.4%). Only one reported misusing substances, and only one had a co-existing clinical condition examined in this study. The median age was 32 (IQR: 14, 46). The median percent of the county population below the federal poverty line was close to national estimates at 13 (IQR: 10, 13), and there was a median of 60 (IQR: 53, 99) primary care physicians per 100,000 people in the counties where these patients lived.

South

Of 3,525 with a known RT designation, 15.8% in the rural South were attributed to RT. Rural patients whose TB disease was attributed to RT in the South Region were mostly male (72.7%) with the highest median age of all the regions at 40 years (IQR: 28, 54). Nearly 90% were people born in the United States, and 82.0% identified as a non-White minority, primarily non-Hispanic Black. Substance misuse was reported by approximately a third (29.0%) of patients. The vast majority (87.1%) did not report any of the clinical conditions examined in this study. Patients lived in rural counties where a median of 22% (IQR: 10, 40) of the population lived below the federal poverty line. This region had the lowest primary care physicians to population ratio with a median of 36 (IQR: 29, 46).

West

Among 230 patients in rural areas of the West Region whose TB disease was attributed to RT, 66.5% were male with a median age of 38 years (IQR: 21, 53). Almost all (95.2%) were born in the United States, and almost all (97.8%) identified as part of a racial or ethnic minority group, of whom 90% was American Indian or Alaska Native. Substance misuse was reported more frequently (39.6%) in this region than other regions, and a small minority (5.7%) reported co-existing clinical conditions. Patients in the West Region lived in the most impoverished counties with a median of 27% (IQR: 24, 29) of the population living below the poverty line; however, the number of primary care physicians per 100,000 people was highest in the West Region at a median of 91 (IQR: 52, 104).

Significant Differences

When comparing characteristics among patients whose TB disease was attributed to recent transmission in the rural United States, there were significant differences ($p<0.05$) by region for all characteristics except for age.

Significant demographic and social-behavioral differences

Recent transmission in the South Region was predominantly among male patients (72.7%), compared with 57.4% in the Midwest, 57.1% in the Northeast, and 66.5% in the West (Table 11). These differences were statistically significant (p = .041). Origin of birth by region was also significantly different (p < .001). In the West Region, 95.2% of the TB disease attributed to RT occurred among U.S.-born persons, compared with 89.6% in the South, followed by 79.6% in the Midwest, and 71.4% in the Northeast (Table 11). Significant differences were also observed among patient's self-identified race and ethnicity. In the West Region, 97.8% of patients whose TB disease was attributed to RT were of a racial/ethnic minority group, compared with 82.0% in the South, 71.4% in the Northeast, and 66.7% in the Midwest. Substance misuse was also significantly different by region among patients whose TB disease was attributed to RT.

It was highest in the West Region at 39.6%, followed by the Midwest (37.0%), the South (29.0%), and the Northeast (14.3%).

Table 11

Demographic and Social-Behavioral Characteristics of Patients Whose TB Disease was Attributed to Recent Transmission, by U.S. Census Regions, 2011–2021

Characteristic	**Midwest**		**Northeast**		**South**		**West**		***p*-value**[1]
	No	**%**	**No**	**%**	**No**	**%**	**No**	**%**	
Total	54	6.4	7	0.8	556	65.6	230	27.2	
Sex									.041
Male	31	57.4	4	57.1	404	72.7	153	66.5	
Female	23	42.6	3	42.9	152	27.3	77	33.5	
Missing	0		0		0		0		
Origin of birth									<.001
U.S.-born	43	79.6	5	71.4	498	89.6	219	95.2	
Non-U.S.–born	11	20.4	2	28.6	58	10.4	11	4.8	
Unknown/missing	0		0		0		0		
Race/ethnicity									<.001
White, not of Hispanic ethnicity	18	33.3	2	28.6	100	18.0	5	2.2	
All racial/ethnic minority groups	36	66.7	5	71.4	456	82.0	225	97.8	
Unknown/missing	0		0		0		0		

Substance misuse[2]									.031
Yes	20	37.0	1	14.3	161	29.0	91	39.6	
No	33	61.1	5	71.4	384	69.1	139	60.4	
Unknown/missing	1	1.9	1	14.3	11	2.0	0		

Age (years)	Mean	SD	Mean	SD	Mean	SD	Mean	SD	.110
	36	21.0	33	28.0	41	17.0	38	21.0	

1. Statistical testing performed on observations with known results; unknown/missing results excluded
2. Includes use of injection drugs, noninjection drugs, or heavy alcohol use in the past 12 months

Significant clinical differences

Patients with TB disease attributed to RT in the Midwest had a significantly higher proportion of selected clinical conditions (24.1%), followed by the Northeast (14.3%), the South (12.9%), and the West (5.7%) (Table 12). The most frequently listed clinical conditions for all regions was diabetes, followed by non-HIV immunocompromising conditions for all regions except the South where HIV co-infection was more frequently reported. HIV co-infection among patients whose disease was attributed to RT was only observed in the South Region.

Table 12

Clinical Characteristics of Patients Whose TB Disease was Attributed to Recent Transmission, by U.S. Census Regions, 2011–2021

Characteristic	Midwest		Northeast		South		West		
	No	**%**	**No**	**%**	**No**	**%**	**No**	**%**	***p*-value**[1]
Total	54	6.4	7	0.8	556	65.6	230	27.2	
Any clinical condition[2]									<.001
Yes	13	24.1	1	14.3	72	12.9	13	5.7	
No	41	75.9	6	85.7	484	87.1	217	94.3	
Diabetes									N/A[3]
Yes	10	18.5	0		56	10.1	9	3.9	
No	44	81.5	7	100.0	500	89.9	221	96.1	
HIV co-infection									N/A[3]
Yes	0		0		21	3.8	0		
No	50	92.6	5	71.4	507	91.2	202	87.8	
Unknown/missing	4	7.4	2	28.6	28	5.0	28	12.2	
Non-HIV immunocompromise[4]									N/A[3]
Yes	3	5.6	1	14.3	17	3.1	4	1.7	
No	51	94.4	6	85.7	539	96.9	226	98.3	

1. Statistical testing performed on observations with known results; unknown/missing results excluded
2. Includes all clinical conditions listed below: diabetes, HIV co-infection, non-HIV immunocompromising conditions
3. Significance testing not done on sub-categories of clinical conditions due to cells containing zero

4. Patient is immunocompromised because of either a medical condition that is not HIV (e.g., leukemia) or immunosuppressive therapy (e.g., TNF alpha inhibitor)

Significant County-Level Differences

Overall, 57.6% of patients with TB disease in rural areas lived in a county that NCHS designated as micropolitan; the remaining 42.4% lived in a county designated as noncore (Table 13). These proportions align approximately with the proportions among those whose TB disease was attributed to RT: 51.8% in micropolitan counties and 48.2% in noncore counties. However, there were notable differences by region. In the West Region, 93.0% of patients whose TB was attributed to RT lived in noncore counties, compared with 42.6% in the Midwest, 30.6% in the South, and 14.3% in the Northeast.

The mean percent of the county population living below the federal poverty level was significantly higher for patients in the West Region (26%), followed by the South (24%), the Midwest (17%), and the Northeast (12%). Conversely, the mean number of primary care physicians per 100,000 persons was significantly higher for patients in the West (80 primary care physicians/100,000 persons), followed by 73 in the Northeast, 49 in the Midwest, and 39 in the South (Table 13).

Table 13

County Characteristics of Patients Whose TB Disease was Attributed to Recent Transmission, by U.S. Census Regions, 2011–2021

Characteristic	**Midwest**		**Northeast**		**South**		**West**		***p*-value**
	No	**%**	**No**	**%**	**No**	**%**	**No**	**%**	
Total	54	6.4	7	0.8	556	65.6	230	27.2	
NCHS[1] rural designation									<.001
Micropolitan	31	57.4	6	85.7	386	69.4	16	7.0	
Non-core	23	42.6	1	14.3	170	30.6	214	93.0	
	Mean	**SD[2]**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	
Percent county population living in poverty	17	7.0	12	3.0	24	6.0	26	8.0	<.001
PCP[3] per 100,000 persons in county	49	23.0	73	33.0	39	17.0	80	33.0	<.001

1. National Center for Health Statistics
2. Standard deviation
3. Primary care physician

Sensitivity Analyses

Sensitivity analysis 1. Excluding Cases from Alaska

When excluding cases reported from Alaska, substance misuse was no longer significantly different by U.S. Census Region; this finding should be interpreted cautiously due to small cell counts, including in the Northeast and West Regions (1 case and 7 cases, respectively). All other findings except age continued to be significantly different by U.S. Census Region; age was also not significantly different in the full dataset. Although the significance of the finding did not change, the percent of patients with any clinical condition in the West Region increased from 5.7% in the full dataset to 25% in the dataset excluding cases from Alaska. The percentages in the other regions remained similar (Appendix D, Table 3).

Sensitivity Analysis 2. Excluding Cases among Patients Incarcerated at Diagnosis

When excluding cases among patients who were incarcerated at diagnosis, two changes were observed: The sex of patients whose TB disease was attributed to RT was no longer statistically different between the U.S. Census Regions, and the percent of patients who misused any substances was no longer statistically different between the U.S. Census Regions. Other findings, including county-level findings, stayed relatively constant as compared to the full dataset (Appendix D, Table 4).

Research Question #3

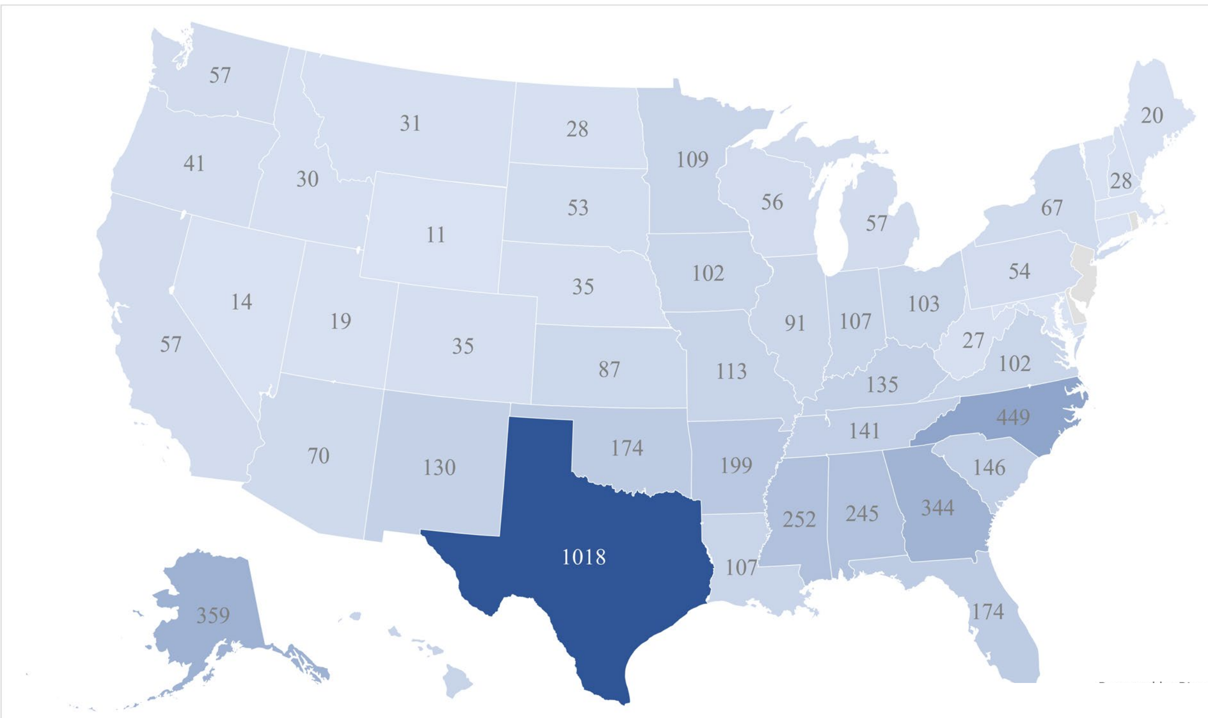
What is the relative association of individual-level and community-level characteristics with recent transmission among people with TB disease in rural areas in the United States?

Dataset Used for Analysis

Because the outcome of interest and dependent variable for this analysis was whether a patient's TB disease was attributed to RT, the parent dataset used for Research Question 3 excluded all cases with an unknown RT designation (n = 1,967). The resulting dataset included 5,633 cases, of which 847 (15.0%) were among patients whose TB disease was attributed to RT.

Figure 4

Distribution of TB Cases with RT Designation, by State, 2011–2021, n = 5633



First attempts to run the Hierarchical Generalized Linear Model (HGLM) on the resulting dataset failed to converge even when including only one level-1 (individual-level) predictor and zero level-2 (county-level) predictors other than the county designation (FIPS code). Examination of the dataset revealed that this was likely occurring because 401 of the counties contained only one case. Therefore, the dataset was restricted to counties with ≥5 cases. This restriction reduced the dataset to 4,004 cases (71% of original dataset) and reduced the number of counties from 1,199 to 344 (29% of original dataset). To determine if it might be possible to run the HGLM on a dataset further restricted to counties with ≥10 cases, a table and maps were derived to view the resulting datasets.

Table 14

Number of Observations, Counties, and States Included in Datasets Considered for Hierarchical Generalized Linear Model

Dataset	Total observations	Counties included	States included
Full dataset with RT designation	5,633	1,199	47
Dataset restricted to counties with ≥5 cases	4,004	344	42
Dataset restricted to counties with ≥10 cases	2,600	121	29

Figure 5

Distribution of TB Cases with Recent Transmission Designation and ≥5 Cases per County, by State, 2011–2021, n = 4,004

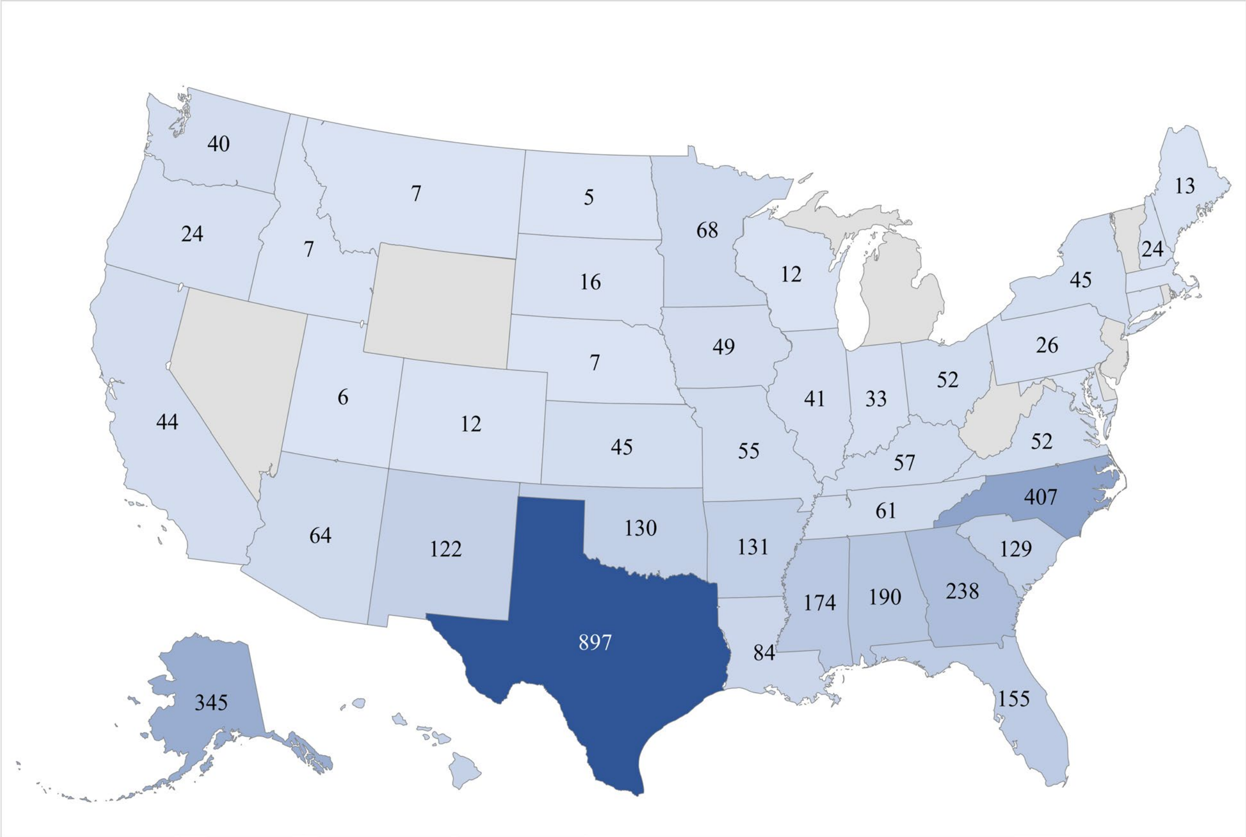
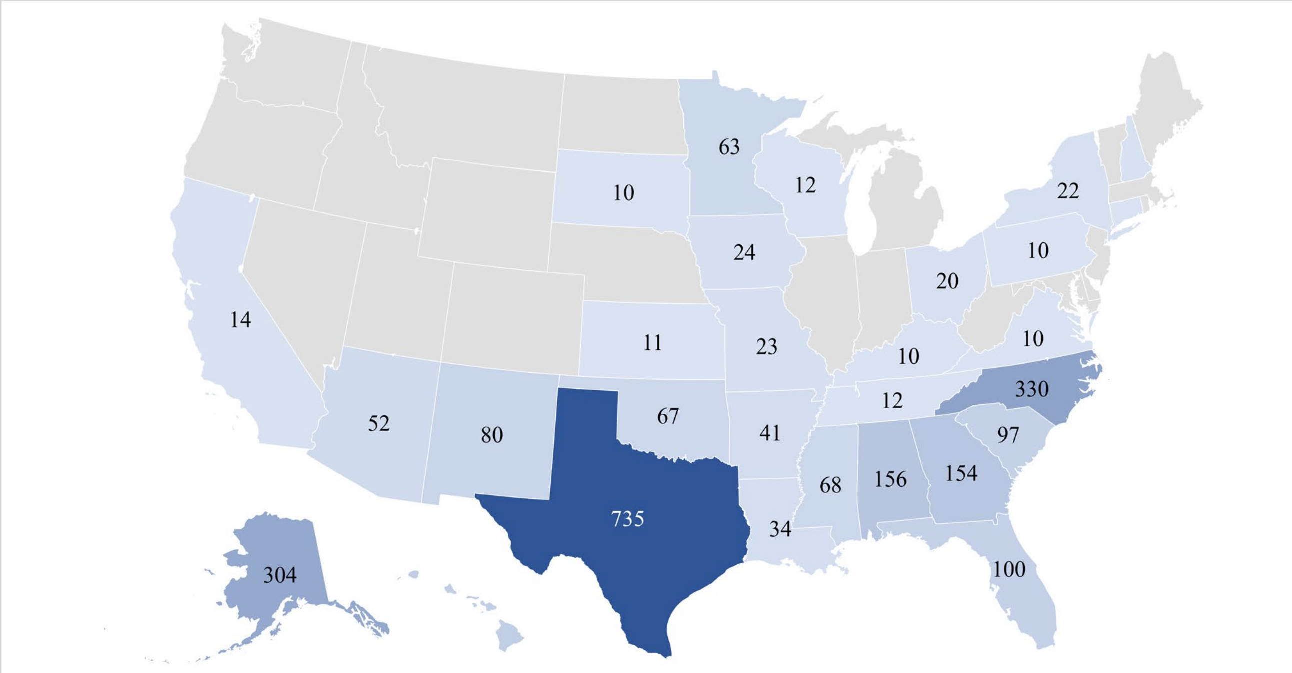


Figure 6

Distribution of TB Cases with Recent Transmission Designation and ≥10 Cases per County, by State, 2011–2021, n = 2,600



Because restricting the dataset to counties with ≥10 cases would exclude almost the entire West region and much of the Midwest and Northeast, thus resulting in primarily a regional analysis rather than a national one, the decision was made to run the HGLM on the dataset that includes counties with ≥5 cases.

Random Intercept Model

The first model was the random intercept model, which only included the county as the independent variable (using the FIPS codes as a categorical county identifier) with the RT outcome as the dependent variable. This model examined whether there exists variation across counties, indicating clustering and the need for a multi-level model. Clustering across counties was defined by an intra-class correlation (ICC) >0.01. With 4,004 observations across 344 counties, the variance of the random intercept across counties was 2.623 (SD: 1.62). The ICC is 0.44, indicating that 44% of the variation in the random intercept model was related to variation between counties. This clear indication of county-level clustering supports the use of a multi-level model. With no level-1 predictors in the model, the estimated odds ratio was 0.05 with a standard error (SE) of 0.01. The p-value was <0.01, indicating a significant effect. The model fit AIC for this model was 3084.3.

Table 15

Random Intercept Model of County FIPS[1] Code as Predictor of Recent Transmission, 2011–2021

Fixed Effects	**Odds Ratio**	**Lower CI[2]**	**Upper CI[2]**	***p*-value**
Intercept	0.05	0.04	0.07	<0.001
	Random Effects			
σ2	3.29			
τ00 FIPS[1] code	2.62			
ICC[3]	0.44			
N (FIPS code)	344			
Observations	4004			
AIC[4]	3084.3			

1. County identifier – Federal Information Processing Standard
2. 95% Confidence Interval
3. Intra-class correlation
4. Akaike Information Criterion

Inclusion of Level-1 Predictors in the Model

The inclusion of each pre-determined level-1 predictor in the model, starting with age, led to improvements in the model, as indicated by a decrease in the Akaike Information Criterion (AIC) value from 3,084 to 2,854. The number of observations declined slightly from 4,004 to 3,933.

Decision to Run Two Separate HGLM Models

It was determined that both county-level predictors of percent below poverty and the ratio of primary care physicians to population could not be included in the same model for two reasons. First, the two level-2 predictors were correlated (-0.166, p<0.001), indicating a significant negative correlation between counties. As the percent of the county population below poverty increased, the ratio of primary care physicians decreased. Second, when the HGLM was attempted with both predictors in the model, the model would not converge, producing an error that the model was nearly unidentifiable with a very large eigenvalue — that is, the model had difficulty estimating the parameters reliably, likely due to the correlation between the two variables.

Individual Level Predictors as Fixed, Rather Than Random Effects

None of the individual-level predictors were treated as random in the models. In other words, the assumption was made that the relationship between each individual-level predictor (e.g., age, race, incarceration) and RT operated similarly within in each county. This decision was made despite *a-priori* hypotheses that some of the individual-level predictors would operate differently with regards to RT on a county basis (Table 2). These hypotheses could not be tested due to the high number of counties that caused the models to not converge when including additional random effects beyond the random intercept. Both county-level variables were included as fixed effects rather than random effects in the models.

Model Including Rate of Primary Care Physicians per 100,000 Population

The ratio of primary care physicians per 100,000 population was added as a fixed effect to the model with all the level-1 predictors. The model fit improved from 2,854 to

2,767, and the number of observations declined by 57 to 3,876. Significant individual-level predictors included age (OR = 0.98, 95% CI: 0.97–0.98), being U.S.-born (OR = 5.05, 95% CI: 3.62–7.04), identifying as anything other than non-Hispanic White race/ethnicity (OR = 1.54, 95% CI: 1.13–2.11), and incarceration at diagnosis (OR = 0.46, 95% CI: 0.26–0.80) (Table 16). Individual-level predictors without a significant effect included being unhoused in the previous year, misuse of substances, and presence of one or more clinical conditions that increase the risk of progression to TB disease. With the inclusion of the level-1 predictors, the τ00 (the variability between counties) decreased from 2.62 to 1.73, and the ICC (the proportion of total variability in RT that is due to differences between counties) decreased from 0.44 to 0.34, indicating that the inclusion of the individual-level variables was explaining some of the across county variation. Although the odds ratio of the physician-to-population ratio was statistically significant ($p<0.05$), the odds ratio was 0.99 (95% CI: 0.98–1.00), suggesting the effect on the log odds of RT was minimal (Table 16).

Table 16

Hierarchical Generalized Linear Model of Individual-Level and County-Level Predictors of Recent Transmission with Inclusion of Number of Primary Care Physicians per 100,000 Persons in County, 2011–2021

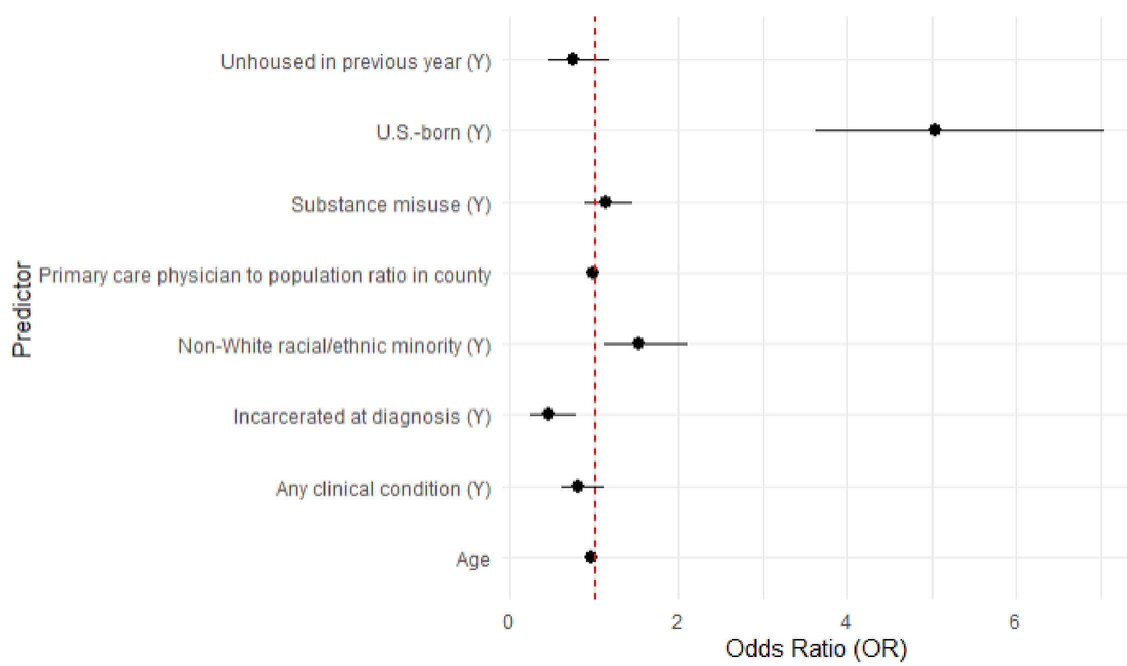
Fixed Effects	**Odds Ratio**	**Lower CI**[1]	**Upper CI**[1]	***p*-value**
Intercept	0.09	0.05	0.18	<.001
Age	0.97	0.97	0.98	<.001
U.S.-born (Y)	5.05	3.62	7.04	<.001
Non-White racial/ethnic minority (Y)	1.54	1.13	2.11	.007
Incarcerated at diagnosis (Y)	0.46	0.26	0.80	.006
Unhoused in previous year (Y)	0.75	0.47	1.19	.221
Misuse of substances (Y)[2]	1.15	0.90	1.46	.255
Any clinical condition (Y)[3]	0.82	0.61	1.12	.21
Primary care physician to population ratio in county	0.99	0.98	1.00	.022

Random Effects	
σ2	3.29
τ00 FIPS[4] code	1.73
ICC[5]	0.34
N (FIPS[4] code)	343
Observations	3,876
AIC[6]	2767.2

1. 95% Confidence Interval
2. Includes use of injection drugs, non-injection drugs, or heavy alcohol use in the past 12 months
3. Includes the following clinical conditions: diabetes, HIV co-infection, non-HIV immunocompromising conditions
4. County identifier – Federal Information Processing Standard
5. Intra-class correlation
6. Akaike Information Criterion

Figure 7

Forest Plot for Results of Hierarchical Generalized Linear Model of Individual-Level and County-Level Predictors of Recent Transmission with Inclusion of Number of Primary Care Physicians per 100,000 Persons in County, 2011–2021



Model Including Percent of County Population Below Federal Poverty Level

The PCP ratio was removed from the model and the percent of the county below the federal poverty level was added as a fixed effect. Similar to the model that included the county-level variable of provider to population ratio, the inclusion of this county-level variable exerted minimal change over the fixed effects of the individual-level predictors. The fixed effect of percent of county population below poverty (OR = 1.02, 95% CI: 0.99–1.05) was not statistically significant. The random effects in the model remained essentially the same. The model fit remained the same as the model without inclusion of this variable (2,854 in both) but was worse than the model that included the ratio of providers to 100,000 population (Table 17).

Table 17

Hierarchical Generalized Linear Model of Individual-Level and County-Level Predictors of Recent Transmission with Inclusion of Percent of County Population Below Poverty, 2011–2021

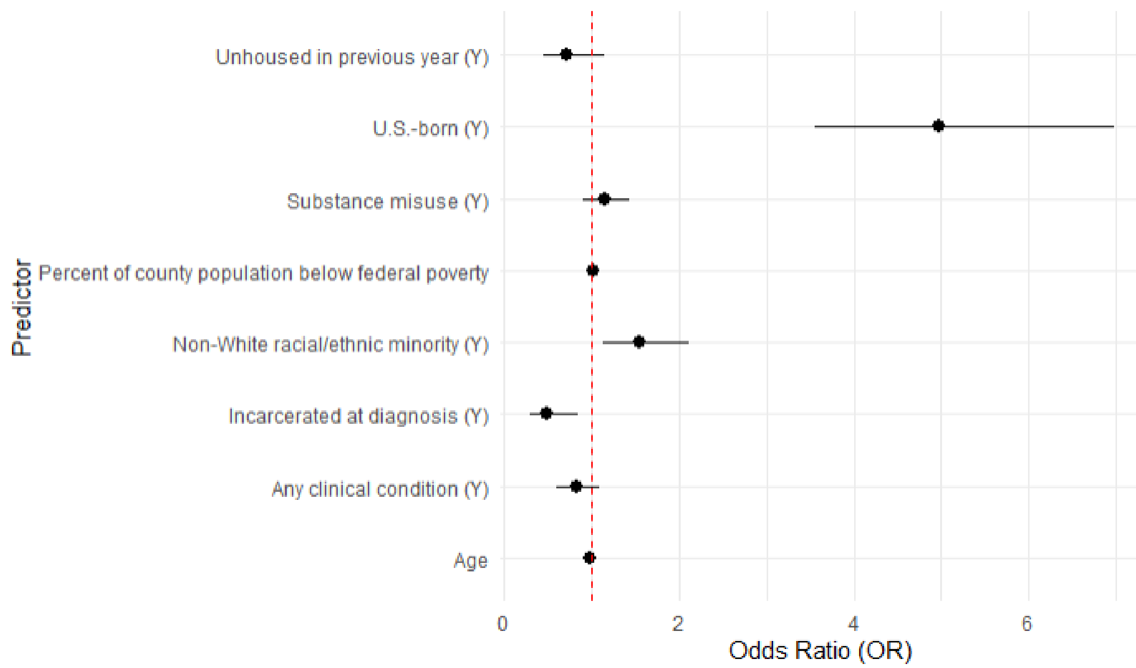
Fixed Effects	**Odds Ratio**	**Lower CI[1]**	**Upper CI[1]**	***p*-value**
Intercept	0.04	0.02	0.08	<.001
Age	0.97	0.97	0.98	<.001
U.S.-born (Y)	4.98	3.56	6.97	<.001
Non-White racial/ethnic minority (Y)	1.54	1.12	2.11	.007
Incarcerated at diagnosis (Y)	0.49	0.29	0.85	.011
Unhoused in previous year (Y)	0.71	0.45	1.14	.156
Misuse of substances (Y)[2]	1.14	0.90	1.44	.283
Any clinical condition (Y)[3]	0.82	0.60	1.10	.188
Percent of county population below federal poverty	1.02	0.99	1.05	.178

Random Effects		
σ^2	3.29	
τ_{00} FIPS[4] code	1.72	
ICC[5]	0.34	
N (FIPS[4] code)	344	
Observations	3933	
AIC[6]	2854.5	

1. 95% Confidence Interval
2. Includes use of injection drugs, non-injection drugs, or heavy alcohol use in the past 12 months
3. Includes the following clinical conditions: diabetes, HIV co-infection, non-HIV immunocompromising conditions
4. County identifier – Federal Information Processing Standard
5. Intra-class correlation
6. Akaike Information Criterion

Figure 8

Forest Plot for Results of Hierarchical Generalized Linear Model of Individual-Level and County-Level Predictors of Recent Transmission with Inclusion of Percent of County Population Below Poverty, 2011–2021



Overall, these findings show that across the 344 included counties, the log odds of RT varies, and individual-level predictors of age, origin of birth, race/ethnicity, and incarceration at diagnosis are significant predictors of the odds of RT, but county-level predictors are not associated with variation across counties in mean county-level odds of RT.

Two models were run without individual-level predictors, but with county-level predictors, one at a time. Although the county-level predictor of the percentage of the county below the federal poverty level did not have a significant effect on the odds of RT in the full model, it did have a small, but significant effect (OR = 1.04, 95% CI: 1.00–1.08) in the model that did not include individual-level predictors. This finding suggests that patients living in counties with higher poverty are more likely to have TB disease attributed to RT; however, when significant individual-level variables such as race and origin of birth are included, the effect of the county poverty level disappears.

Sensitivity Analyses

Sensitivity Analysis 1. Excluding Cases Reported in Alaska

Excluding cases from Alaska resulted in no meaningful changes to the results of either of the HGLMs (Appendix D, Tables 5 & 6).

Sensitivity Analysis 2. Excluding Cases among Patients Incarcerated at Diagnosis

When excluding cases incarcerated at diagnosis, misuse of substances changed from a non-significant effect (OR = 1.1, 95% CI: 0.90–1.44) to a significant effect (OR = 1.3, 95% CI: 1.02–1.66). Results presented are for HGLM that includes the percent below poverty; results are similar for the HGLM that includes the PCP-to-population ratio (Appendix D, Table 8). Although there were other small changes to effect sizes and confidence intervals of other variables in the model, those results were not meaningfully different from results in the full models.

Single-Level Logistic Regression Model with Inclusion of Both Individual-Level And County-Level Variables

To better understand what might be causing the changes in the fixed effects with the inclusion of the level-2 variables, a single-level logistic regression (i.e., not an HGLM) model was run, including all desired individual-level variables and both county-level variables. This post ad-hoc analysis had four noteworthy findings (Table 18):

1) The magnitude of the individual-level effects that were significant in the HGLM increased, mostly notably the effect of birth in the United States, with an odds ratio increasing from 4.98 (95% CI: 3.56–6.97) to 9.47 (95% CI: 7.37–12.32).
2) The direction of the effect for incarceration at diagnosis changed from negative in the HGLM (OR = 0.49, 95% CI: 0.29–0.85) to positive (OR = 1.73, 95% CI: 1.36–2.19) in the flat logistic regression model while continuing to be statistically significant.
3) Two individual-level variables (misuse of substances and presence of underlying clinical condition) that did not have significant effects in the HGLM had significant effects in the flat logistic regression model with the magnitude of the effect increasing in the same direction. Misuse of substances became significantly positively associated with RT whereas presence of a co-existing clinical condition was significantly negatively associated with RT.
4) Both county-level variables had significant effects; however, the odds ratios for both were 1.0 or almost 1.0.

Table 18

Single Level Logistic Regression Model Examining Characteristics Associated with Recent Transmission in Rural Areas, 2011–2021

Predictor	Odds Ratio	Lower CI[1]	Upper CI[1]	*p*-value
Intercept	0.06	0.02	0.04	<.001
Age	0.97	0.97	0.97	<.001
U.S.-born (Y)	9.47	7.37	12.32	<.001
Non-White racial/ethnic minority (Y)	2.47	1.98	3.12	<.001
Incarcerated at diagnosis (Y)	1.73	1.36	2.19	<.001
Unhoused in previous year (Y)	0.77	0.52	1.12	.181
Misuse of substances[2] (Y)	1.3	1.08	1.56	.006
Any clinical condition[3] (Y)	0.66	0.52	0.84	.001
Percent of county population below federal poverty	1.03	1.02	1.04	<.001
Primary care physician to population ratio in county	1.00	1.00	1.01	.025
Observations	5,474			
AIC[4]	3,717.1			

1. 95% Confidence Interval
2. Includes use of injection drugs, non-injection drugs, or heavy alcohol use in the past 12 months
3. Includes the following clinical conditions: diabetes, HIV co-infection, non-HIV immunocompromising conditions
4. Akaike Information Criterion

Overall, these findings suggest significant variability in RT between counties and that county-level effects not measured in this analysis are exerting effects on the individual-level characteristics associated with RT in rural areas.

CHAPTER 5

DISCUSSION

Chapter five provides a discussion of the study results and revisits the Syndemic Theory in light of the findings. Strengths and limitations of the study are presented, followed by a discussion of the implications for clinical and public health practice. Finally, recommendations for future research, policy, and implementation are provided.

Characteristics of Patients with TB Disease in Rural Areas

This study provided, for the first time, a national-level look at TB disease among people living in rural areas as well as a national-level examination of factors associated with recent transmission of TB disease in rural areas. During the study years of 2011 through 2021, at least 7,607 patients living in rural areas of the United States were diagnosed with TB disease. Of those, the majority (63.8%) were among people who were born in the United States. This finding is in direct contrast to the national epidemiology of TB, in which approximately 70% of TB cases are among people who were born in other countries outside the United States, primarily countries in which the incidence of TB is much higher than in the United States (CDC, 2021). This finding alone has important implications for clinical and public health practice. Most current recommendations for identifying TB instruct clinicians and public health practitioners to focus their efforts on people who were born in countries where TB is more common, such as CDC's Think. Test. Treat TB. campaign (CDC, 2023a). Indeed, the USPSTF recommendation for TB screening, most recently updated in May 2023, states that screening is recommended in populations at increased risk and states that populations at increased risk include "persons who were born in, or are former residents of, countries with high tuberculosis prevalence…" (Mangione et al., 2023). Although this recommendation is sound at a national level, reflecting the national epidemiology of TB, this recommendation would miss most TB cases in rural areas. It should be noted that these recommendations are focused on identifying people with asymptomatic LTBI, not symptomatic TB disease, and that this study does not provide incidence rates. It is possible, even likely, that incidence rates of TB disease in rural areas are higher among non-U.S.–born than U.S.-born persons given the much higher rates of TB disease in non-U.S.–born persons nationally and the lower density of non-U.S.–born persons in rural areas. However, this point still cannot be overstated. At a national level, most TB cases are among people born in other countries. In rural areas, clinicians and public health practitioners should be aware that most TB cases in their practice areas are going to occur among people who likely were born and raised in the United States.

Additionally, screening recommendations for TB focus on those with other clinical conditions that increase the risk of developing TB disease after infection (Mangione et al., 2023). HIV-infected individuals have a 30% risk of progressing to disease compared to a 5%–10% risk among the general population (Markowitz et al., 1997; Moreno et al., 1993) and diabetes approximately triples the risk of progression (Jeon & Murray, 2008). These clinically focused screening recommendations would also fall short in rural areas. A minority of patients in this study had HIV co-infection or other immunosuppressive conditions. While diabetes was reported among patients with TB at a higher prevalence than in the overall national population (16.2% compared with approximately 11% nationally), (CDC, 2024b) it was much lower than the prevalence observed among TB patients nationally at approximately 24% (CDC, 2021).

These findings beg the question: what is driving TB in rural areas? Tuberculosis has long been understood as a disease of social inequity (Knopf, 1914; Mancuso et al., 2016; Olson et al., 2012), as laid out in Chapter 1. Patients with TB disease in this study lived in counties where a mean of 20% of the population lived below the federal poverty line. Nationally, approximately 11% of the population lives below the federal poverty line (United States Census Bureau, 2023a), highlighting this disparity for patients in this study. From an access to care perspective, nationally there are approximately 80 primary care physicians per 100,000 people (HRSA, 2023). In this study, patients with TB lived in counties that had a mean of 49 primary care physicians per 100,000 people, strikingly less than the nationally mean. As expected, most patients with TB disease in rural areas identified as being of a racial or ethnic minority group. This aligns with national epidemiology of TB but is an even more stark finding in rural areas considering that the proportion of people who identify as part of a racial and/or ethnic minority group is lower in rural areas than in urban areas (Rural Health Information Hub, 2023). Infectious disease racial and ethnic disparities are not limited to TB; a recent commentary on the HIV epidemic concluded that it is “marked by geographical concentration in the US South and profound disparities between region and by race or ethnicity” (Sullivan et al., 2021). Likewise, findings from a review of COVID-19 pandemic literature found that Black and Hispanic Americans experienced higher rates of SARS-CoV-2 infection and COVID-19-related mortality (Mackey et al., 2021).

While the proportion of those who were unhoused was similar to national reports, the proportion of patients who were incarcerated at diagnosis was approximately twice what is seen nationally (CDC, 2021). This latter finding is likely due to the high density of state and federal prisons in rural areas (Huling, 2002) rather than a higher rate of incarceration among people living in rural areas.

Overall, the individual-level and county-level characteristics of patients with TB disease in rural areas support the theory that TB disease continues to occur among marginalized populations that experience social and health disparities that increase their risk for disease.

Although this study was focused on understanding characteristics associated with RT in rural areas, these findings related to TB disease in general in rural areas may be the most useful for driving public health action and policy. The rest of this chapter will focus on the three research questions that specifically address RT.

Discussion of Research Question 1. Characteristics Associated with Recent Transmission in Rural Areas

In rural areas of the United States, 15% of TB disease among patients with a known RT designation was attributed to RT. This proportion is higher than national estimates of approximately 12% (CDC, 2022b). Most individual-level characteristics examined by this study were significantly associated with RT, including age, origin of birth, race and/or ethnicity, incarceration status, and substance use. Interestingly, homelessness has been associated with RT in previous studies (CDC, 2022b; Self et al., 2021; Yuen et al., 2016), but not in this one.

Prior to this study, all information about TB transmission in rural areas was based on individual outbreak reports in localized settings (Bloss et al., 2011; Bock et al., 1998;

CDC, 2002; Dillaha et al., 2002; Funk, 2003; Labuda et al., 2021; Lowther et al., 2011; Phillips et al., 2004; Valway et al., 1998). This study suggests that many of the findings of those individual reports are consistent with the national picture. As with many of the individual reports, RT was more frequently reported among U.S.-born persons (Bloss et al., 2011; CDC, 2002; Dillaha et al., 2002; Funk, 2003; Labuda et al., 2021) and racial and ethnic minority groups (Bloss et al., 2011; CDC, 2002; Dillaha et al., 2002; Funk, 2003; Labuda et al., 2021); substance misuse was also significantly more common (Bloss et al., 2011; Dillaha et al., 2002; Labuda et al., 2021). Incarceration was significantly negatively associated with RT in this study which aligns with the individual outbreak reports that infrequently included known transmission in a congregate setting.

Previous outbreak reports in rural areas included clinical features of patients involved in outbreaks; however, without a comparator, it was unclear if these clinical features contributed to the outbreak. In this national study, the three co-existing clinical conditions of diabetes, HIV co-infection, and non-HIV immunosuppression were all more frequently observed among patients whose TB disease was not attributed to RT, although this finding was not statistically significant in the HGLM. In fact, the prevalence of diabetes among patients in rural areas whose TB disease was attributed to RT was lower than national estimates of diabetes generally and less than half than what it seen nationally among patients with TB disease.

In the bivariate analysis, patients whose TB disease was attributed to RT were significantly more likely to live in a county with higher rates of poverty than patients whose disease was not attributed to RT; however, ratios of primary care physicians to population were similar among the two groups. Interestingly, the reverse was true in the

HGLM – poverty was not significantly associated with RT and the PCP ratio was, however, with a very small effect size.

Overall, the findings from this first research question fill some important gaps in the literature for TB transmission in rural areas. Patients with TB disease in rural areas whose disease is attributed to RT are more frequently born in the United States, of a racial or ethnic minority group, are incarcerated at diagnosis, report substance misuse, and live in counties with higher rates of poverty. They are less likely to have certain co-existing medical conditions than patients whose TB disease is not attributed to RT.

Discussion of Research Question 2. Regional Differences in Characteristics Associated with Recent Transmission in Rural Areas

Findings at a national level improve understanding and can drive policy and funding. Interventions to prevent and stop infectious disease transmission, however, will always require an understanding at a more granular level. The second research question of this study aimed to identify and describe characteristics of patients whose TB disease was attributed to RT at a regional level. As described in the results, all characteristics except for age had significant differences by region. More patients with TB attributed to RT in the Midwest identified as White than in any other region. They were also the most likely to have an underlying health condition and more likely to misuse substances than patients in the Northeast or South Region. Those in the Northeast (n = 7) were the least likely to report substance misuse and the most likely to live in a region classified as micropolitan rather than non-core. After the Midwest, they were also more likely than the other two regions to identify as White, non-Hispanic. Patients in the South with TB attributed to RT were very likely to be male, U.S.-born, and identify as Black, non-

Hispanic. Almost all patients with TB attributed to RT in the West Region were U.S-born and almost all were among patients who identified as part of a racial or ethnic minority group, namely American Indian or Alaska Native. Substance misuse was reported most frequently in the West Region. Residents with TB in the West Region were the most likely to live in a rural county classified as non-core and lived in more impoverished counties than patients in other regions.

The West Region had the highest proportion of TB cases attributed to RT at nearly 24%. In the sensitivity analysis that excluded Alaska, only 32 remaining cases were attributed to RT (i.e., 5.3% of the total 917 remaining cases), changing the West Region from having the highest proportion of cases attributed to RT to the lowest. This finding could be interpreted in two ways. First, this could be due to systematic differences in the way TB cases are identified and reported in Alaska. As discussed in Chapter 3, longstanding molecular uniformity in *M. tuberculosis* genotypes and artificial time clustering due to case detection methods in rural Alaska could lead to falsely elevated estimates of RT. Second, Alaska has high rates of TB disease (7.9 per 100,000 in 2021) relative to the overall national TB incidence rate (2.4 per 100,000 in 2021) (CDC, 2021); rates of TB disease among indigenous people in the United States are on average at least 10 times higher than among patients with TB who identify as White (Springer et al., 2022). Although not recently documented in the literature, high rates of TB disease coupled with small, socially connected communities that spend much of the year indoors, could lead to TB transmission and outbreaks. Therefore, the finding that the West Region has the highest proportion of cases attributed to RT likely reflects both some degree of bias and some truth.

The findings from this second research question are particularly important given the dearth of previous literature examining TB in rural areas of the United States generally and the lack of national RT estimates and outbreak studies in rural areas. Previous outbreak reports from rural areas have described outbreaks primarily in the South Region (Bloss et al., 2011; Bock et al., 1998; Labuda et al., 2021; Valway et al., 1998) with two reports from the Midwest (Lowther et al., 2011; Phillips et al., 2004) and two reports from the West (CDC, 2002; Funk, 2003). This dissertation provides regional descriptions of patient characteristics whose TB disease is attributed to RT and highlights that both RT estimates and patient characteristics differ substantially across U.S. Census Regions, suggesting that regional studies may be more useful than national ones when seeking to identify drivers of TB transmission in rural areas and develop applicable interventions.

Discussion of Research Question 3. Relative Association of Characteristics Associated with Recent Transmission in Rural Areas

A HGLM was constructed to calculate the odds of RT for selected individual and county characteristics in rural areas of the United States. Although the decision to use a HGLM was determined *a-priori* due to the theoretical understanding that a patient's disease status and clinical outcomes are due to more than an individual's characteristics and lifestyle, this decision was reinforced by the results of the random intercept model, which had an intra-class correlation of 0.44, indicating that 44% of the variation was related to variation between counties. Although the ICC declined somewhat with inclusion of other explanatory variables, it stabilized at 0.34 in both final models, providing evidence that even with inclusion of all theoretically chosen explanatory variables available at both the individual and county level, there remained a significant amount of county-level variation in RT that was not explained by these factors. This finding has implications for future RT research. For data with a meaningful multilevel structure, making an assumption that the group structure is represented by only explanatory variables would be incorrect (Snijders & Bosker, 2012).

Importantly, the two county-level variables included in the models both had effect sizes of approximately 1, suggesting that the between county variation observed in the model was due to something other than county-level rates of poverty or access to healthcare (as defined by the ratio of primary care physicians to population). A manuscript currently undergoing peer-review (Mathilda Regan, 2024) used an area level measure called the Index of Concentration at the Extremes (ICE), a racialized measure of economic segregation in a mediation analysis. Of all the area-level measures in their study, which included poverty, unemployment, uninsurance, the ICE measure mediated the effects of the study outcomes more strongly than the other measures. Structuralized racism is an important driver of infectious disease incidence rates and outcomes (Noppert et al., 2023; Sullivan et al., 2021); considering the role of structuralized racism in regard to TB transmission is important. Two findings from the HGLM support that approach. In the adjusted model, the two characteristics with the strongest association with RT were being born in the United States and identifying as part of a racial or ethnic minority group. These findings were not surprising. Previous studies have consistently found that both RT and outbreaks are more common among U.S.-born persons and persons of racial or ethnic minority groups (Mindra et al., 2017; Mitruka et al., 2011; Raz et al., 2022; Yuen et al., 2016). Understanding the driving factors behind these findings and identifying paths to rectify these disparities are critical for achieving health equity and for continued work toward TB elimination in rural areas.

In contrast to previous national studies, being unhoused was not significantly associated with RT in either of the adjusted HGLMs, and incarceration at diagnosis was significantly negatively associated with RT. This finding aligns with individual outbreak reports from rural areas where congregate settings were rarely identified as a likely venue of transmission. Like those individual outbreak reports, it is likely that most of the RT in rural areas is occurring in community settings, including households and among close-knit social groups. Understanding and preventing RT in those communities will likely require understanding social structures within each individual's affected community. Of note, in the flat logistic regression model, the effect of incarceration at diagnosis was significantly positively associated with RT, a directional flip from the findings in the HGLM in which incarceration was significantly negatively associated with RT. This finding suggests that, at least in rural areas, the association of incarceration with RT depends largely on the county in which the patient is incarcerated; in some counties, incarceration increases the probability that the person acquired TB from RT, in others, the opposite is true. This finding should be further explored to understand which county level factors are influencing this association.

Unlike previous outbreak reports, substance misuse was also not significantly associated with RT when adjusting for other factors. This finding is surprising as nearly every outbreak report and RT study has reported on the misuse of substances, often consumed in group settings, to be an important factor in facilitating transmission. The findings from this study suggest that while substance misuse is more frequently observed among patients whose TB disease is attributed to RT, other factors are more important contributors to the variation in RT. Like incarceration, substance misuse was significantly associated with RT in the flat logistic regression model, again suggesting there exist county-level factors that vary and influence this association. Together these remarkable differences between a flat logistic regression model, used frequently in epidemiologic TB studies (Chen et al., 2020; Miramontes et al., 2015; Yuen et al., 2016), and a HGLM that accounts for county-level variation, provides evidence that HGLMs should be considered for national epidemiologic research, at least for RT. Furthermore, if the model parameters allow, it would be beneficial to allow individual-level predictors such as incarceration, substance misuse, and experiencing homelessness to be treated as both random and fixed effects rather than as fixed effects alone.

Revisiting the Syndemic Theory

This dissertation was guided by the framework provided by the Syndemic Theory, which posits that diseases and harmful social conditions exist simultaneously and synergistically (Mendenhall et al., 2022). These study findings do not support the theory that coexisting diseases are contributing to RT of TB. In the bivariate results of this study, all the coexisting medical conditions, including diabetes, HIV co-infection, and non-HIV immunosuppression were all significantly more frequent among patients whose disease was not attributed to RT. In the adjusted HGLM, the difference in co-existing conditions between the two groups was not significant. The reason for these findings is very likely the outcome selected in this analysis, namely RT, rather than TB disease itself. All patients included in this study had TB disease; there was not comparator without TB disease. In simple terms, RT can be understood as the inverse of reactivation

TB disease from a longstanding latent TB infection (Haddad et al., 2018). A key risk factor in reactivation TB is a compromised immune system, which could be caused by any of the coexisting medical conditions examined in this dissertation (e.g., HIV) (Markowitz et al., 1997; Moreno et al., 1993; Pai et al., 2016). Although an immunocompromised person exposed to TB bacteria would more likely develop TB disease than a latent infection, this is not possible to examine with NTSS data, which only includes patients with TB disease, not infection. Therefore, these coexisting clinical conditions could increase a person's risk of developing disease overall, rather than only development of TB disease from a recent infection. As such, the lack of association with RT observed in this study aligns with understandings of the disease process.

This study's findings provide some support for coexisting social conditions contributing to RT. In the unadjusted analysis, patients whose disease was attributed to RT were significantly more likely to live in counties with higher levels of poverty; however, in the HGLM, there was no detectable effect of poverty on RT. Although not statistically significant, patients whose disease was attributed to RT also lived in counties where the median number of primary care physicians to the population was lower. However, in the HGLM, the extremely small effect of this proxy for access to care was in the opposite direction, indicating that patients living in counties with lower ratios of primary care physicians were less likely to have disease that was attributed to RT. Again, the outcome of RT rather than TB disease likely explains this lack of strong association. TB is known to be a disease of social inequity and, indeed, patients included in this study were found to live in counties with higher levels of poverty and lower ratios of healthcare providers than the national averages. If we applied the Syndemic Theory to a study examining the risk of developing TB disease in rural areas, we would likely observe the harmful synergy posited by the theory. Another possibility for not seeing an effect in this study could be the selection of county-level factors. The HGLM shows county-level variation in RT, with differences at the county level influencing the association of individual factors with RT, suggesting that environmental factors do, in fact, play a role in RT. However, this study does not support the hypothesis that the county-level factors of poverty and primary healthcare access drive RT in rural areas.

Study Limitations

Limitations for this dissertation primarily derive from inherent characteristics of the primary data source. Although NTSS allows for national analyses that include all TB cases, the data are collected for surveillance purposes with the primary objective of tracking progress towards TB elimination and providing a summary of U.S. TB epidemiology. Not all variables that would be useful for understanding drivers of TB transmission are collected and some key information, such as infectious periods, diagnostic delays, social mixing behaviors, and sources of TB disease, were not included in this dataset. Findings are reported by public health officials based on their knowledge from interviews with patients and reviewing the medical record; certain medical conditions or perceived undesirable social behaviors may be subject to underreporting. Perhaps most importantly, NTSS only includes patients with TB disease. It does not include people without TB disease in the community nor does it include people with LTBI. Because of that, it is not possible to use this data source to understand the relative risk of certain characteristics with developing TB disease as compared to those without TB. This is an especially important limitation when considering the adjusted model's findings that poverty and presence of coexisting clinical conditions were not significantly associated with RT. These variables could likely be associated with developing TB disease; however, this hypothesis could not be tested with the available data.

A further limitation for the HGLM was the need to restrict the analysis to counties with at least five cases. Although the model would not run on the full dataset given the frequency of counties with only one or two cases, this adjustment excluded 1,629 patients and 855 counties from the analytic dataset. The counties and patients in the analytic dataset for the HGLM may not be representative of the counties and patients in the full dataset.

Finally, the Area Health Resource File, which was used to calculate the primary care physician to population ratio, did not have numbers of healthcare providers other than physicians available at the county level. Therefore, the rates used in this analysis only included primary care physicians and not nurse practitioners or physician assistants, despite the latter groups being more likely than physicians to work in rural areas (Agency for Healthcare Research and Quality, 2012).

Study Strengths

This dissertation also had many strengths. Most importantly, it is the first study to examine factors associated with RT of TB in rural areas of the United States. Beyond being the first, the findings in this study are nationally representative of all TB cases in rural areas, a leap forward from previously available literature that only provided localized outbreak reporting. This was made possible by using data from NTSS, a surveillance system that collects information of each newly reported cases of TB disease in the United States (CDC, 2022a). Unlike previous outbreak reports, this study defined rurality using a nationally recognized standard provided by the National Center for Health Statistics (National Center for Health Statistics, 2014), adding to the validity of findings and facilitating the reproducibility of results. Next this study used a validated algorithm (France et al., 2015) that combines molecular, geographic, and time data to determine RT. The RT algorithm has been used frequently in TB literature to provide national RT estimates and estimates among many subgroups, including older patients (Kim et al., 2022), persons experiencing homelessness (Self et al., 2021), and persons living with HIV (Schmit et al., 2020) among others. However, the RT algorithm has never been previously used to provide RT estimates specific to rural areas. Although NTSS provides many individual-level patient characteristics, including demographic, social-behavioral, clinical, and laboratory variables, it does not provide information for some of the key social determinants of health, including poverty and access to healthcare. Because TB is closely tied to social inequity and health disparities (Ali, 2014; Asabor & Vermund, 2021; Knopf, 1914; Olson et al., 2012), understanding the environment in which a patient with TB lives is important. This study was strengthened by including county-level measures of poverty and primary care physicians to population ratios; these measures, while not available for individual patients, provide an understanding of factors beyond an individual patient, that could contribute to infectious disease findings. Finally, this study was strengthened by the inclusion of a HGLM which accounted for between-county variation in RT (Snijders & Bosker, 2012), thus providing more accurate and reliable estimates of characteristics associated with RT.

Implications for Clinical Practice

This dissertation has important implications for clinical practice in rural areas. Although TB is a devastating disease that is still a leading cause of infectious disease mortality globally, in the United States, due to low incidence of TB disease, most U.S. healthcare providers have rarely, if ever, taken care of a patient with TB, and receive very little clinical training on TB. As a result, one of the key public health strategies for eliminating TB is striving to help healthcare providers consider TB on their differential diagnosis list when a patient has symptoms of TB. These efforts can be seen in the Think. Test. Treat. campaign developed by CDC (CDC, 2023a). However, part of these efforts to encourage providers to “think TB” is focused on helping providers recognize the groups at risk, which, in the United States, are primarily people who were born in or frequently traveled to a county with higher incidence of TB or who have immunocompromising conditions that increase their risk of progression to TB disease. This study shows that understanding of risk groups may not apply to rural areas. If healthcare providers in rural areas only consider TB in patients who were born in another country or are immunocompromised, they will miss most TB cases in their counties. Not only that, healthcare providers will also miss nearly all TB cases that are a part of active chains of transmission in their communities. In this way, this study not only has implications for office- and hospital-based clinicians, but also for the clinicians who work in the local health departments in rural areas, most of whom are nurses. Nurses in local health departments are the primary clinicians conducting infectious disease investigations and providing case management of patients. If a nurse working in a rural health department identifies a case of active TB disease, they should consider that RT may be occurring, especially if the patient is U.S.-born. Their investigation should focus on identifying other patients with TB disease as well as identifying TB infections among those exposed. The findings from this dissertation suggest that most patients involved in RT in rural areas will be U.S.-born and of a racial or ethnic minority group with differences in the minority group based on the U.S. Census Region. Efforts to stop transmission should focus on the risk groups within the community.

Recommendations for Future Research

As the first study examining TB disease in rural areas of the United States at a national level, there remain many avenues for future study. Perhaps most pressing is calculating incidence rates for TB disease in rural areas, stratified by key demographic groups. While case counts and proportions, as described in this study, allow for comparisons of risk factors and disease attribution, they do not facilitate understanding of population-specific risk. This study showed that most cases of TB disease in rural areas are among U.S.-born people and that nearly all RT was among U.S.-born people. However, this may primarily be a reflection of the underlying population. Rural areas have a much higher concentration of U.S.-born people than urban areas (Rural Health Information Hub, 2023). By calculating incidence rates, a researcher could show whether the findings in this study are simply reflecting the underlying population distribution or whether the risk of TB and RT in rural areas among U.S.-born persons is out of proportion with the population. Calculating rates would also help public health practitioners identify areas where the rates exceed the national average and therefore should be considered for focused interventions.

Although researchers are being encouraged to compare health outcomes among rural population subgroups and emphasize within-group analysis, as this study has done, over rural-urban comparisons (CDC, 2024a), an important step for understanding differences at a national level would be to stratify demographic, social, and clinical characteristics by urban/rural designations to allow for direct side-by-side comparison. In this study, each finding was compared to what has been reported nationally; however, national findings include findings from rural areas. This approach could be considered for a table(s) in CDC's annual TB surveillance report (CDC, 2021), which is published in the fall of each year. The methods used in this study could provide the framework for annual reporting by rural/urban stratifications, thus facilitating observations of trends over time.

As observed in the HGLM, RT varied significantly between counties; however, the county-level variables that were included in this study had very little effect, if any, on RT. Future research could explore which county-level factors do contribute to county-level differences in RT. A previous study in Michigan found that neighborhood density and neighborhood disadvantage were significantly associated with RT among U.S.-born persons (Noppert et al., 2017). This study found that most RT occurred among patients who identified as a race/ethnicity other than non-Hispanic White. Another recent study (Mathilda Regan, 2024) showed that a racialized measure of economic segregation mediated TB outcomes more than others, further supporting the need to consider area-level measures of structuralized racism in future analyses.

This study could not extend past 2021 due to changes in genotyping methods for TB; since 2022, all genotyping is done using whole-genome multi-locus sequencing (wgMLSTyping). Although wgMLSTyping is likely a more accurate method for genotyping as it examines approximately 70% of the *Mtb* genome rather than approximately 1% (CDC, 2024c), the RT algorithm has not yet been validated using this method. Once it is validated, it would be valuable to update this analysis using wgMLSTyping for RT determinations.

Finally, to aid research around access to healthcare in rural areas more broadly, the Health Resource Services Administration, which creates the Area Health Resource Files, should strive to find data sources that delineate the number of non-physician primary care providers, such as nurse practitioners, at the county level (Health Resources and Services Administration, 2024).

Recommendations for Public Health Policy and Practice

Due to the important demographic differences among patients with TB disease in this study versus what is understood and reported nationally, CDC should consider reporting characteristics of patients living in rural areas separately from those in urban areas to highlight those most at risk in different geographic localities. On a local level, TB resources in rural areas should be directed to U.S.-born persons in racial/ethnic minority groups.

Conclusion

In this first national study examining characteristics of patients with TB disease in rural areas of the United States, origin of birth, arguably the most important U.S. epidemiologic characteristic for TB, was nearly the inverse of national reports. Nationally, approximately 70% of all patients with TB were born in another country; this dissertation found that almost 64% of patients with TB in rural areas were born in the United States. Among patients whose disease was attributed to RT, more than 90% were born in the United States. In the bivariate analysis, several significant differences were found in characteristics of patients whose TB disease was attributed to RT. They were more frequently born in the United States, of a racial or ethnic minority group, incarcerated at diagnosis, reported substance misuse, and lived in counties with higher rates of poverty. They were less likely to have certain co-existing medical conditions than patients whose TB disease is not attributed to RT. In the adjusted HGLM, the odds of being U.S.-born, identifying as a racial or ethnic minority group, and being incarcerated remained significantly associated with RT. Although the log odds of RT significantly varied between counties, the county-level measures of poverty and rates of primary care physicians did not exert important effects. Estimates of RT differed by U.S. Census Region; among those with disease attributed to RT, several significant differences by region were identified, suggesting that regional studies may be more useful than national ones for understanding drivers of TB transmission in rural areas. Healthcare providers in rural areas should be educated that TB epidemiology in rural areas is different than TB epidemiology nationally and should consider TB as a potential diagnosis among symptomatic U.S.-born patients of racial and ethnic minority groups in addition to patients born outside the United States. Future research should focus on understanding population level risk in rural areas and county-level drivers of variation in RT. More attention is needed to highlight the unique characteristics of patients with TB disease broadly and attributed to RT in rural areas.

REFERENCES

Agency for Healthcare Research and Quality. (2012). *Primary care workforce facts and stats no. 3: Distribution of the U.S. primary care workforce*. https://www.ahrq.gov/sites/default/files/publications/files/pcwork3.pdf

Ali, M. (2014). Treating tuberculosis as a social disease. *Lancet*, *383*(9936), 2195. https://doi.org/10.1016/s0140-6736(14)61063-1

Allison, P. D. (2002). *Missing Data*. SAGE Publications, Inc. https://doi.org/https://doi.org/10.4135/9781412985079

Andermann, A. (2017). Outbreaks in the age of syndemics: New insights for improving Indigenous health. *Can Commun Dis Rep*, *43*(6), 125-132. https://doi.org/10.14745/ccdr.v43i06a02

Apostolopoulos, Y., Lemke, M. K., Hosseinichimeh, N., Harvey, I. S., Lich, K. H., & Brown, J. (2018). Embracing Causal Complexity in Health Disparities: Metabolic Syndemics and Structural Prevention in Rural Minority Communities. *Prev Sci*, *19*(8), 1019-1029. https://doi.org/10.1007/s11121-018-0924-3

Asabor, E. N., & Vermund, S. H. (2021). Confronting Structural Racism in the Prevention and Control of Tuberculosis in the United States. *Clin Infect Dis*, *73*(9), e3531-e3535. https://doi.org/10.1093/cid/ciaa1763

Bamrah, S., Yelk Woodruff, R. S., Powell, K., Ghosh, S., Kammerer, J. S., & Haddad, M. B. (2013). Tuberculosis among the homeless, United States, 1994-2010. *Int J Tuberc Lung Dis*, *17*(11), 1414-1419. https://doi.org/10.5588/ijtld.13.0270

Barker, L. E., Kirtland, K. A., Gregg, E. W., Geiss, L. S., & Thompson, T. J. (2011). Geographic distribution of diagnosed diabetes in the U.S.: a diabetes belt. *Am J Prev Med*, *40*(4), 434-439. https://doi.org/10.1016/j.amepre.2010.12.019

Bates, M. N., Khalakdina, A., Pai, M., Chang, L., Lessa, F., & Smith, K. R. (2007). Risk of tuberculosis from exposure to tobacco smoke: a systematic review and meta-analysis. *Arch Intern Med*, *167*(4), 335-342. https://doi.org/10.1001/archinte.167.4.335

Beck, A. J., Boulton, M. L., & Coronado, F. (2014). Enumeration of the governmental public health workforce, 2014. *Am J Prev Med*, *47*(5 Suppl 3), S306-313. https://doi.org/10.1016/j.amepre.2014.07.018

Berzkalns, A., Bates, J., Ye, W., Mukasa, L., France, A. M., Patil, N., & Yang, Z. (2014). The road to tuberculosis (Mycobacterium tuberculosis) elimination in Arkansas; a re-examination of risk groups. *PLoS One*, *9*(3), e90664. https://doi.org/10.1371/journal.pone.0090664

Bloss, E., Newbill, K., Peto, H., Rice, M. J., Ainsworth, G., Travnicek, R., Holcombe, M., Haddad, M. B., & Oeltmann, J. E. (2011). Challenges and opportunities in a tuberculosis outbreak investigation in southern Mississippi, 2005-2007. *South Med J*, *104*(11), 731-735. https://doi.org/10.1097/SMJ.0b0.3e318232679e

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. *Clin Infect Dis*, *27*(5), 1221-1226. https://doi.org/10.1086/514986

Braden, C. R., Templeton, G. L., Cave, M. D., Valway, S., Onorato, I. M., Castro, K. G., Moers, D., Yang, Z., Stead, W. W., & Bates, J. H. (1997). Interpretation of restriction fragment length polymorphism analysis of Mycobacterium tuberculosis isolates from a state with a large rural population. *J Infect Dis*, *175*(6), 1446-1452. https://doi.org/10.1086/516478

CDC. (2002). Tuberculosis outbreak on an American Indian reservation--Montana, 2000-2001. *MMWR Morb Mortal Wkly Rep*, *51*(11), 232-234. https://www.cdc.gov/mmwr/preview/mmwrhtml/mm5111a3.htm

CDC. (2005). Guidelines for the investigation of contacts of persons with infectious tuberculosis. Recommendations from the National Tuberculosis Controllers Association and CDC. *MMWR Recomm Rep*, *54*(Rr-15), 1-47. https://www.cdc.gov/mmwr/preview/mmwrhtml/rr5415a1.htm

CDC. (2016). *Basic TB facts*. Retrieved 03/28/2023 from https://www.cdc.gov/tb/topic/basics/default.htm

CDC. (2017, 10/12/2017). *Assuances of confidentiality*. Retrieved 06/30/2023 from https://www.cdc.gov/os/integrity/confidentiality/index.htm#:~:text=An%20Assurance%20of%20Confidentiality%20is,identifiable%20or%20potentially%20identifiable%20information.

CDC. (2021). *Reported Tuberculosis in the United States, 2021*. Retrieved 02/02/2023 from https://www.cdc.gov/tb/statistics/reports/2021/estimates.htm

CDC. (2022a, June 16, 2024). *Data hosting: National Tuberculosis Surveillance System*. https://www.cdc.gov/rdc/b1datatype/tuberculosis.htm

CDC. (2022b). *Reported tuberculosis in the United States, 2021. Estimates of recent transmission, 2020-2021*. https://www.cdc.gov/tb/statistics/reports/2021/ERT.html#:~:text=Among%20TB%20cases%20occurring%20in,3.9%25%20during%202018%E2%80%932019.

CDC. (2023a). *Think. Test. Treat TB*. Retrieved June 16, 2024 from https://www.cdc.gov/thinktesttreattb/index.html

CDC. (2023b). *Tuberculosis*. Retrieved 6/30/2023 from https://www.cdc.gov/tb/default.htm

CDC. (2024a, 4/18/24). *Call for papers | Rural health disparities: Contemporary solutions for persistent rural public health challenges*. Retrieved 6/16/24 from https://www.cdc.gov/pcd/call_for_papers/Rural_Health_Disparities.htm

CDC. (2024b). *National diabetes statistics report*. Retrieved June 16, 2024 from https://www.cdc.gov/diabetes/php/data-research/index.html

CDC. (2024c). *Tuberculosis whole-genome sequencing*. Retrieved 6/16/2024 from https://www.cdc.gov/tb/php/genotyping/whole-genome-sequencing.html

Chen, M. P., Miramontes, R., & Kammerer, J. S. (2020). Multidrug-resistant tuberculosis in the United States, 2011-2016: patient characteristics and risk factors. *Int J Tuberc Lung Dis*, *24*(1), 92-99. https://doi.org/10.5588/ijtld.19.0173

Cohen, A., Mathiasen, V. D., Schön, T., & Wejse, C. (2019). The global prevalence of latent tuberculosis: a systematic review and meta-analysis. *Eur Respir J*, *54*(3). https://doi.org/10.1183/13993003.00655-2019

Coughlin, L. N., Bonar, E. E., Bohnert, K. M., Jannausch, M., Walton, M. A., Blow, F. C., & Ilgen, M. A. (2019). Changes in urban and rural cigarette smoking and cannabis use from 2007 to 2017 in adults in the United States. *Drug Alcohol Depend*, *205*, 107699. https://doi.org/10.1016/j.drugalcdep.2019.107699

Croft, J. B., Wheaton, A. G., Liu, Y., Xu, F., Lu, H., Matthews, K. A., Cunningham, T. J., Wang, Y., & Holt, J. B. (2018). Urban-Rural County and State Differences in Chronic Obstructive Pulmonary Disease - United States, 2015. *MMWR Morb Mortal Wkly Rep*, *67*(7), 205-211. https://doi.org/10.15585/mmwr.mm6707a1

de Vries, G., Aldridge, R. W., Cayla, J. A., Haas, W. H., Sandgren, A., van Hest, N. A., & Abubakar, I. (2014). Epidemiology of tuberculosis in big cities of the European Union and European Economic Area countries. *Euro Surveill*, *19*(9). https://doi.org/10.2807/1560-7917.es2014.19.9.20726

Dillaha, J. A., Yang, Z., Ijaz, K., Eisenach, K. D., Cave, M. D., Wilson, F. J., Stead, W. W., & Bates, J. H. (2002). Transmission of Mycobacterium tuberculosis in a rural community, Arkansas, 1945-2000. *Emerg Infect Dis*, *8*(11), 1246-1248. https://doi.org/10.3201/eid0811.020299

Dixon, M. A., & Chartier, K. G. (2016). Alcohol Use Patterns Among Urban and Rural Residents: Demographic and Social Influences. *Alcohol Res*, *38*(1), 69-77.

Donders, A. R., van der Heijden, G. J., Stijnen, T., & Moons, K. G. (2006). Review: a gentle introduction to imputation of missing values. *J Clin Epidemiol*, *59*(10), 1087-1091. https://doi.org/10.1016/j.jclinepi.2006.01.014

Dubos, R., & Dubos, J. (1952). *The White Plague: Tuberculosis, Man, and Society* (3rd ed.). Rutgers University Press.

Dugani, S. B., Mielke, M. M., & Vella, A. (2021). Burden and management of type 2 diabetes in rural United States. *Diabetes Metab Res Rev*, *37*(5), e3410. https://doi.org/10.1002/dmrr.3410

Dye, C., Lönnroth, K., Jaramillo, E., Williams, B. G., & Raviglione, M. (2009). Trends in tuberculosis incidence and their determinants in 134 countries. *Bull World Health Organ*, *87*(9), 683-691. https://doi.org/10.2471/blt.08.058453

Ekramnia, M. (2023). Estimated rates of progression to TB disease for persons infected with Mycobacterium tuberculosis in the United States. *Epidemiology*, *Pending publication*.

France, A. M., Cave, M. D., Bates, J. H., Foxman, B., Chu, T., & Yang, Z. (2007). What's driving the decline in tuberculosis in Arkansas? A molecular epidemiologic analysis of tuberculosis trends in a rural, low-incidence population, 1997 2003. *Am J Epidemiol*, *166*(6), 662-671. https://doi.org/10.1093/aje/kwm135

France, A. M., Grant, J., Kammerer, J. S., & Navin, T. R. (2015). A field-validated approach using surveillance and genotyping data to estimate tuberculosis attributable to recent transmission in the United States. *Am J Epidemiol*, *182*(9), 799-807. https://doi.org/10.1093/aje/kwv121

Funk, E. A. (2003). Tuberculosis contact investigations in rural Alaska: a unique challenge. *Int J Tuberc Lung Dis*, *7*(12 Suppl 3), S349-352. https://www.ingentaconnect.com/contentone/iuatld/ijtld/2003/00000007/A00312s3/art00006?crawler=true

Garcia, M. C., Rossen, L. M., Bastian, B., Faul, M., Dowling, N. F., Thomas, C. C., Schieb, L., Hong, Y., Yoon, P. W., & Iademarco, M. F. (2019). Potentially Excess Deaths from the Five Leading Causes of Death in Metropolitan and Nonmetropolitan Counties - United States, 2010-2017. *MMWR Surveill Summ*, *68*(10), 1-11. https://doi.org/10.15585/mmwr.ss6810a1

GBD. (2017). The global burden of tuberculosis: results from the Global Burden of Disease Study 2015. *Lancet Infect Dis*. https://doi.org/10.1016/s1473-3099(17)30703-x

Geng, E., Kreiswirth, B., Burzynski, J., & Schluger, N. W. (2005). Clinical and radiographic correlates of primary and reactivation tuberculosis: a molecular epidemiology study. *JAMA*, *293*(22), 2740-2745. https://doi.org/10.1001/jama.293.22.2740

Ghosh, S., Moonan, P. K., Cowan, L., Grant, J., Kammerer, S., & Navin, T. R. (2012). Tuberculosis genotyping information management system: enhancing tuberculosis surveillance in the United States. *Infect Genet Evol*, *12*(4), 782-788. https://doi.org/10.1016/j.meegid.2011.10.013

Green, S. B. (1991). How many subjects does it take to do a regression analysis. *Multivariate Behavioral Research*, *26*(3), 499-510. https://doi.org/DOI: 10.1207/s15327906mbr2603_7

Haddad, M. B., Raz, K. M., Lash, T. L., Hill, A. N., Kammerer, J. S., Winston, C. A., Castro, K. G., Gandhi, N. R., & Navin, T. R. (2018). Simple Estimates for Local Prevalence of Latent Tuberculosis Infection, United States, 2011-2015. *Emerg Infect Dis*, *24*(10), 1930-1933. https://doi.org/10.3201/eid2410.180716

Health Resources and Services Administration. (2024). *Area Health Resources Files*. https://data.hrsa.gov/data/download?data=AHRF#AHRF

Hing, E., & Chun-Ju, H. (2014). State variability in supply of office-based primary case providers: United States, 2012. *NCHS Data Brief*, *151*. https://www.ruralhealth.us/NRHA/media/Emerge_NRHA/PDFs/db151.pdf

Hossain, M. M., Saha, N., Rodela, T. T., Tasnim, S., Nuzhath, T., Roy, T. J., Burdine, J. N., Ahmed, H. U., McKyer, E. L. J., Basu, B. K., & Ma, P. (2022). Global research on syndemics: a meta-knowledge analysis (2001-2020). *F1000Res*, *11*, 253. https://doi.org/10.12688/f1000research.74190.2

Houston, A. R., Lynch, K., Ostrach, B., Isaacs, Y. S., Nvé Díaz San Francisco, C., Lee, J. M., Emard, N., & Proctor, D. A. (2022). United States immigration detention amplifies disease interaction risk: A model for a transnational ICE-TB-DM2 syndemic. *Glob Public Health*, *17*(7), 1152-1171. https://doi.org/10.1080/17441692.2021.1919737

Howard, G., Kleindorfer, D. O., Cushman, M., Long, D. L., Jasne, A., Judd, S. E., Higginbotham, J. C., & Howard, V. J. (2017). Contributors to the Excess Stroke Mortality in Rural Areas in the United States. *Stroke*, *48*(7), 1773-1778. https://doi.org/10.1161/strokeaha.117.017089

HRSA. (2023). *State of the primary care workforce, 2023*. https://bhw.hrsa.gov/sites/default/files/bureau-health-workforce/data-research/state-of-primary-care-workforce-2023.pdf

Huling, T. (2002). Building a prison economy in rural America. https://www.prisonpolicy.org/scans/building.html

Ivey-Stephenson, A. Z., Crosby, A. E., Jack, S. P. D., Haileyesus, T., & Kresnow-Sedacca, M. J. (2017). Suicide Trends Among and Within Urbanization Levels by Sex, Race/Ethnicity, Age Group, and Mechanism of Death - United States, 2001-2015. *MMWR Surveill Summ*, *66*(18), 1-16. https://doi.org/10.15585/mmwr.ss6618a1

James, C. V., Moonesinghe, R., Wilson-Frederick, S. M., Hall, J. E., Penman-Aguilar, A., & Bouye, K. (2018). Racial/Ethnic Health Disparities Among Rural Adults-United States, 2012-2015: MMWR Surveillance Summaries / November 17, 2017 / 66(23);1-9. *J Health Care Poor Underserved*, *29*(1), 19-34. https://doi.org/10.1353/hpu.2018.0003

Jeon, C. Y., & Murray, M. B. (2008). Diabetes mellitus increases the risk of active tuberculosis: a systematic review of 13 observational studies. *PLoS Med*, *5*(7), e152. https://doi.org/10.1371/journal.pmed.0050152

Julia Foutz, S. A., Rachel Garfield. (2017). The role of Medicaid in rural America. *Kaiser Family Foundation*. https://www.kff.org/medicaid/issue-brief/the-role-of-medicaid-in-rural-america/

Kempf, M. C., Dunlap, N. E., Lok, K. H., Benjamin, W. H., Jr., Keenan, N. B., & Kimerling, M. E. (2005). Long-term molecular analysis of tuberculosis strains in alabama, a state characterized by a largely indigenous, low-risk population. *J Clin Microbiol*, *43*(2), 870-878. https://doi.org/10.1128/jcm.43.2.870-878.2005

Kim, S., Cohen, T., Horsburgh, C. R., Miller, J. W., Hill, A. N., Marks, S. M., Li, R., Kammerer, J. S., Salomon, J. A., & Menzies, N. A. (2022). Trends, Mechanisms, and Racial/Ethnic Differences of Tuberculosis Incidence in the US-Born Population Aged 50 Years or Older in the United States. *Clin Infect Dis*, *74*(9), 1594-1603. https://doi.org/10.1093/cid/ciab668

Knopf, S. A. (1914). Tuberculosis as a cause and result of poverty. *Journal of the American Medical Association*, *LXIII*(20), 1720-1725. https://doi.org/10.1001/jama.1914.02570200014004

Kwan, C. K., & Ernst, J. D. (2011). HIV and tuberculosis: a deadly human syndemic. *Clin Microbiol Rev*, *24*(2), 351-376. https://doi.org/10.1128/cmr.00042-10

Labuda, S. M., McDaniel, C., Talwar, A., Braumuller, A., Parker, S., McGaha, S., Blissett, C., Wortham, J., Mukasa, L., & Stewart, R. J. (2021). Tuberculosis Outbreak Associated With Delayed Diagnosis and Long Infectious Periods in Rural Arkansas, 2010-2018. *Public Health Rep*, 33354921999167. https://doi.org/10.1177/0033354921999167

Lambert, L. A., Armstrong, L. R., Lobato, M. N., Ho, C., France, A. M., & Haddad, M. B. (2016). Tuberculosis in Jails and Prisons: United States, 2002-2013. *Am J Public Health*, *106*(12), 2231-2237. https://doi.org/10.2105/ajph.2016.303423

Lee, P. H., Fu, H., Lee, M. R., Magee, M., & Lin, H. H. (2018). Tuberculosis and diabetes in low and moderate tuberculosis incidence countries. *Int J Tuberc Lung Dis*, *22*(1), 7-16. https://doi.org/10.5588/ijtld.17.0329

Lönnroth, K., Jaramillo, E., Williams, B. G., Dye, C., & Raviglione, M. (2009). Drivers of tuberculosis epidemics: the role of risk factors and social determinants. *Soc Sci Med*, *68*(12), 2240-2246. https://doi.org/10.1016/j.socscimed.2009.03.041

Lowther, S. A., Miramontes, R., Navara, B., Sabuwala, N., Brueshaber, M., Solarz, S., Haddad, M. B., Sodt, D., & Lynfield, R. (2011). Outbreak of tuberculosis among Guatemalan immigrants in rural Minnesota, 2008. *Public Health Rep*, *126*(5), 726-732. https://doi.org/10.1177/003335491112600515

Mackey, K., Ayers, C. K., Kondo, K. K., Saha, S., Advani, S. M., Young, S., Spencer, H., Rusek, M., Anderson, J., Veazie, S., Smith, M., & Kansagara, D. (2021). Racial and Ethnic Disparities in COVID-19-Related Infections, Hospitalizations, and Deaths : A Systematic Review. *Ann Intern Med*, *174*(3), 362-373. https://doi.org/10.7326/m20-6306

Magee, M. J., Salindri, A. D., Gujral, U. P., Auld, S. C., Bao, J., Haw, J. S., Lin, H. H., & Kornfeld, H. (2018). Convergence of non-communicable diseases and tuberculosis: a two-way street? *Int J Tuberc Lung Dis*, *22*(11), 1258-1268. https://doi.org/10.5588/ijtld.18.0045

Mancuso, J. D., Diffenderfer, J. M., Ghassemieh, B. J., Horne, D. J., & Kao, T. C. (2016). The Prevalence of Latent Tuberculosis Infection in the United States. *Am J Respir Crit Care Med*, *194*(4), 501-509. https://doi.org/10.1164/rccm.201508-1683OC

Mangione, C. M., Barry, M. J., Nicholson, W. K., Cabana, M., Chelmow, D., Coker, T. R., Davis, E. M., Donahue, K. E., Jaén, C. R., Li, L., Ogedegbe, G., Rao, G., Ruiz, J. M., Stevermer, J., Underwood, S. M., & Wong, J. B. (2023). Screening for Latent Tuberculosis Infection in Adults: US Preventive Services Task Force Recommendation Statement. *JAMA*, *329*(17), 1487-1494. https://doi.org/10.1001/jama.2023.4899

Markowitz, N., Hansen, N. I., Hopewell, P. C., Glassroth, J., Kvale, P. A., Mangura, B. T., Wilcosky, T. C., Wallace, J. M., Rosen, M. J., & Reichman, L. B. (1997). Incidence of tuberculosis in the United States among HIV-infected persons. The Pulmonary Complications of HIV Infection Study Group. *Ann Intern Med*, *126*(2), 123-132. https://doi.org/10.7326/0003-4819-126-2-199701150-00005

Mathilda Regan, T. B., Yunfei Li, Nicole Smartwood, Garrett R Beeler Asay, Ted Cohen, C Robert Horsburgh, Awal Khan, Suzanne M. Marks, Ranell L Myles, Joshua A Salomon, Julie L Self, Carla Winston, Nicolas A Menzies. (2024). Risk factors underlying racial/ethnic disparities in tuberculosis diagnosis and treatment outcomes, 2011-2019: a multiple mediation analysis of national surveillance data *Undergoing peer review*.

Matthews, K. A., Croft, J. B., Liu, Y., Lu, H., Kanny, D., Wheaton, A. G., Cunningham, T. J., Khan, L. K., Caraballo, R. S., Holt, J. B., Eke, P. I., & Giles, W. H. (2017). Health-Related Behaviors by Urban-Rural County Classification - United States, 2013. *MMWR Surveill Summ*, *66*(5), 1-8. https://doi.org/10.15585/mmwr.ss6605a1

McNabb, S. J., Kammerer, J. S., Hickey, A. C., Braden, C. R., Shang, N., Rosenblum, L. S., & Navin, T. R. (2004). Added epidemiologic value to tuberculosis prevention and control of the investigation of clustered genotypes of Mycobacterium tuberculosis isolates. *Am J Epidemiol*, *160*(6), 589-597. https://doi.org/10.1093/aje/kwh253

Mendenhall, E., Kohrt, B. A., Norris, S. A., Ndetei, D., & Prabhakaran, D. (2017). Non-communicable disease syndemics: poverty, depression, and diabetes among low-income populations. *Lancet*, *389*(10072), 951-963. https://doi.org/10.1016/s0140-6736(17)30402-6

Mendenhall, E., Newfield, T., & Tsai, A. C. (2022). Syndemic theory, methods, and data. *Soc Sci Med*, *295*, 114656. https://doi.org/10.1016/j.socscimed.2021.114656

Mindra, G., Wortham, J. M., Haddad, M. B., & Powell, K. M. (2017). Tuberculosis Outbreaks in the United States, 2009-2015. *Public Health Rep*, *132*(2), 157-163. https://doi.org/10.1177/0033354916688270

Miramontes, R., Hill, A. N., Yelk Woodruff, R. S., Lambert, L. A., Navin, T. R., Castro, K. G., & LoBue, P. A. (2015). Tuberculosis Infection in the United States: Prevalence Estimates from the National Health and Nutrition Examination Survey, 2011-2012. *PLoS One*, *10*(11), e0140881. https://doi.org/10.1371/journal.pone.0140881

Mitruka, K., Oeltmann, J. E., Ijaz, K., & Haddad, M. B. (2011). Tuberculosis outbreak investigations in the United States, 2002-2008. *Emerg Infect Dis*, *17*(3), 425-431. https://doi.org/10.3201/eid1703.101550

Mongkolrattanothai, T., Lambert, L. A., & Winston, C. A. (2019). Tuberculosis among healthcare personnel, United States, 2010-2016. *Infect Control Hosp Epidemiol*, *40*(6), 701-704. https://doi.org/10.1017/ice.2019.76

Moonan, P. K., Ghosh, S., Oeltmann, J. E., Kammerer, J. S., Cowan, L. S., & Navin, T. R. (2012). Using genotyping and geospatial scanning to estimate recent mycobacterium tuberculosis transmission, United States. *Emerg Infect Dis*, *18*(3), 458-465. https://doi.org/10.3201/eid1803.111107

Moreno, S., Baraia-Etxaburu, J., Bouza, E., Parras, F., Pérez-Tascón, M., Miralles, P., Vicente, T., Alberdi, J. C., Cosín, J., & López-Gay, D. (1993). Risk for developing tuberculosis among anergic patients infected with HIV. *Ann Intern Med*, *119*(3), 194-198. https://doi.org/10.7326/0003-4819-119-3-199308010-00003

Moy, E., Garcia, M. C., Bastian, B., Rossen, L. M., Ingram, D. D., Faul, M., Massetti, G. M., Thomas, C. C., Hong, Y., Yoon, P. W., & Iademarco, M. F. (2017). Leading Causes of

Death in Nonmetropolitan and Metropolitan Areas- United States, 1999-2014. *MMWR Surveill Summ*, *66*(1), 1-8. https://doi.org/10.15585/mmwr.ss6601a1

National Association of County & City Health Officials. (2019). *National profile of local health departments*. https://www.naccho.org/uploads/downloadable-resources/Programs/Public-Health-Infrastructure/NACCHO_2019_Profile_final.pdf

National Center for Chronic Disease Prevention and Health Promotion Office on Smoking and Health. (2014). *The health consequences of smoking - 50 years of progress: A report of the Surgeon General*. Centers for Disease Control and Prevention. https://www.ncbi.nlm.nih.gov/books/NBK179276/

National Center for Health Statistics. (2014). *2013 NCHS urban-rural classification scheme for counties*. https://www.cdc.gov/nchs/data/series/sr_02/sr02_166.pdf

Njie, G. J., Morris, S. B., Woodruff, R. Y., Moro, R. N., Vernon, A. A., & Borisov, A. S. (2018). Isoniazid-Rifapentine for Latent Tuberculosis Infection: A Systematic Review and Meta-analysis. *Am J Prev Med*. https://doi.org/10.1016/j.amepre.2018.04.030

Noppert, G. A., Hegde, S. T., & Kubale, J. T. (2023). Exposure, Susceptibility, and Recovery: A Framework for Examining the Intersection of the Social and Physical Environments and Infectious Disease Risk. *Am J Epidemiol*, *192*(3), 475-482. https://doi.org/10.1093/aje/kwac186

Noppert, G. A., Yang, Z., Clarke, P., Ye, W., Davidson, P., & Wilson, M. L. (2017). Individual- and neighborhood-level contextual factors are associated with Mycobacterium tuberculosis transmission: genotypic clustering of cases in Michigan, 2004-2012. *Ann Epidemiol*, *27*(6), 371-376.e375. https://doi.org/10.1016/j.annepidem.2017.05.009

O'Donnell, M. R., Chamblee, S., von Reyn, C. F., Ellerbrock, T. V., Johnson, J., Marsh, B. J., Moreland, J. D., Narita, M., Pedrosa, M., Johnson, L. S., & Horsburgh, C. R., Jr. (2010). Racial disparities in primary and reactivation tuberculosis in a rural community in the southeastern United States. *Int J Tuberc Lung Dis*, *14*(6), 733-740.

https://docserver.ingentaconnect.com/deliver/connect/iuatld/10273719/v14n6/s12.pdf?expires=1656262439&id=0000&titleid=3764&checksum=3C3838DBD20AB32DA827594BE74BA887&host=https://www.ingentaconnect.com

Olson, N. A., Davidow, A. L., Winston, C. A., Chen, M. P., Gazmararian, J. A., & Katz, D. J. (2012). A national study of socioeconomic status and tuberculosis rates by country of birth, United States, 1996-2005. *BMC Public Health*, *12*, 365. https://doi.org/10.1186/1471-2458-12-365

Pai, M., Behr, M. A., Dowdy, D., Dheda, K., Divangahi, M., Boehme, C. C., Ginsberg, A., Swaminathan, S., Spigelman, M., Getahun, H., Menzies, D., & Raviglione, M. (2016). Tuberculosis. *Nat Rev Dis Primers*, *2*, 16076. https://doi.org/10.1038/nrdp.2016.76

Phillips, L., Carlile, J., & Smith, D. (2004). Epidemiology of a tuberculosis outbreak in a rural Missouri high school. *Pediatrics*, *113*(6), e514-519. https://doi.org/10.1542/peds.113.6.e514

Powell, K. M., VanderEnde, D. S., Holland, D. P., Haddad, M. B., Yarn, B., Yamin, A. S., Mohamed, O., Sales, R. F., DiMiceli, L. E., Burns-Grant, G., Reaves, E. J., Gardner, T. J., & Ray, S. M. (2017). Outbreak of Drug-Resistant Mycobacterium tuberculosis Among Homeless People in Atlanta, Georgia, 2008-2015. *Public Health Rep*, *132*(2), 231-240. https://doi.org/10.1177/0033354917694008

Raz, K. M., Talarico, S., Althomsons, S. P., Kammerer, J. S., Cowan, L. S., Haddad, M. B., McDaniel, C. J., Wortham, J. M., France, A. M., Powell, K. M., Posey, J. E., & Silk, B. J. (2022). Molecular surveillance for large outbreaks of tuberculosis in the United States, 2014-2018. *Tuberculosis (Edinb)*, *136*, 102232. https://doi.org/10.1016/j.tube.2022.102232

Rehm, J., Samokhvalov, A. V., Neuman, M. G., Room, R., Parry, C., Lönnroth, K., Patra, J., Poznyak, V., & Popova, S. (2009). The association between alcohol use, alcohol use disorders and tuberculosis (TB). A systematic review. *BMC Public Health*, *9*, 450. https://doi.org/10.1186/1471-2458-9-450

Reitmanova, S., & Gustafson, D. L. (2012). Coloring the white plague: a syndemic approach to immigrant tuberculosis in Canada. *Ethn Health*, *17*(4), 403-418. https://doi.org/10.1080/13557858.2011.645156

Romanowski, K., Clark, E. G., Levin, A., Cook, V. J., & Johnston, J. C. (2016). Tuberculosis and chronic kidney disease: an emerging global syndemic. *Kidney Int*, *90*(1), 34-40. https://doi.org/10.1016/j.kint.2016.01.034

Rural Health Information Hub. (2023). *National rural health snapshop*. https://www.ruralhealth.us/about-nrha/about-rural-health-care#_ftn1

Rutledge, S. A., Masalovich, S., Blacher, R. J., & Saunders, M. M. (2017). Diabetes Self-Management Education Programs in Nonmetropolitan Counties - United States, 2016. *MMWR Surveill Summ*, *66*(10), 1-6. https://doi.org/10.15585/mmwr.ss6610a1

Schmit, K. M., Shah, N., Kammerer, S., Bamrah Morris, S., & Marks, S. M. (2020). Tuberculosis Transmission or Mortality Among Persons Living with HIV, USA, 2011-2016. *J Racial Ethn Health Disparities*, *7*(5), 865-873. https://doi.org/10.1007/s40615-020-00709-7

Selanders, L. C., & Crane, P. C. (2012). The voice of Florence Nightingale on Advocacy. *The Online Journal of Issues in Nursing*, *17*(1). https://doi.org/10.3912/OJIN.Vol17No01Man01

Self, J. L., McDaniel, C. J., Bamrah Morris, S., & Silk, B. J. (2021). Estimating and Evaluating Tuberculosis Incidence Rates Among People Experiencing Homelessness, United States, 2007-2016. *Med Care*, *59*(Suppl 2), S175-s181. https://doi.org/10.1097/mlr.0000000000001466

Séraphin, M. N., Lauzardo, M., Doggett, R. T., Zabala, J., Morris, J. G., Jr., & Blackburn, J. K. (2016). Spatiotemporal Clustering of Mycobacterium tuberculosis Complex Genotypes in

Florida: Genetic Diversity Segregated by Country of Birth. *PLoS One*, *11*(4), e0153575. https://doi.org/10.1371/journal.pone.0153575

Singer, M. (1996). A dose of drugs, a touch of violence, a case of AIDS: Conceptualizing the Sava Syndemic. *Free Inquiry in Creative Sociology*, *24*(2), 99-110. file:///C:/Users/yxp5/Downloads/6986-Article%20Text-13055-1-10-20160503.pdf

Singer, M. (2006). A dose of drugs, a touch of violence, a case of AIDS: Part 2. Further conceptualizing the sava syndemic. *Free Inquiry in Creative Sociology*, *34*(1), 39-53. file:///C:/Users/yxp5/Downloads/1568-Article%20Text-5749-1-10-20131008.pdf

Singer, M., Bulled, N., Ostrach, B., & Mendenhall, E. (2017). Syndemics and the biosocial conception of health. *Lancet*, *389*(10072), 941-950. https://doi.org/10.1016/s0140-6736(17)30003-x

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis, an introduction to basic and advanced multilevel modeling* (5 ed.). SAGE Publications. (1999)

Springer, Y. P., Kammerer, J. S., Silk, B. J., & Langer, A. J. (2022). Tuberculosis in Indigenous Persons - United States, 2009-2019. *J Racial Ethn Health Disparities*, *9*(5), 1750-1764. https://doi.org/10.1007/s40615-021-01112-6

State of Alaska; Department of Health. (2021). *Tuberculosis in Alaska - Summary Brief*. https://health.alaska.gov/dph/Epi/id/Documents/TB/2021-AK-TB-Summary-Brief.pdf

Stewart, R. J., Raz, K. M., Burns, S. P., Kammerer, J. S., Haddad, M. B., Silk, B. J., & Wortham, J. M. (2022). Tuberculosis Outbreaks in State Prisons, United States, 2011-2019. *Am J Public Health*, *112*(8), 1170-1179. https://doi.org/10.2105/ajph.2022.306864

Stoltzfus, J. C. (2011). Logistic regression: a brief primer. *Acad Emerg Med*, *18*(10), 1099-1104. https://doi.org/10.1111/j.1553-2712.2011.01185.x

Strosnider, H., Kennedy, C., Monti, M., & Yip, F. (2017). Rural and Urban Differences in Air Quality, 2008-2012, and Community Drinking Water Quality, 2010-2015 - United States. *MMWR Surveill Summ*, *66*(13), 1-10. https://doi.org/10.15585/mmwr.ss6613a1

Substance Abuse and Mental Health Services Administration. (2023). *2021 National Survey on Drug Use and Health* Retrieved 5/23/2023 from https://www.samhsa.gov/data/report/2021-nsduh-detailed-tables

Sullivan, P. S., Satcher Johnson, A., Pembleton, E. S., Stephenson, R., Justice, A. C., Althoff, K. N., Bradley, H., Castel, A. D., Oster, A. M., Rosenberg, E. S., Mayer, K. H., & Beyrer, C. (2021). Epidemiology of HIV in the USA: epidemic burden, inequities, contexts, and responses. *Lancet*, *397*(10279), 1095-1106. https://doi.org/10.1016/s0140-6736(21)00395-0

Trajman, A., Felker, I., Alves, L. C., Coutinho, I., Osman, M., Meehan, S. A., Singh, U. B., & Schwartz, Y. (2022). The COVID-19 and TB syndemic: the way forward. *Int J Tuberc Lung Dis*, *26*(8), 710-719. https://doi.org/10.5588/ijtld.22.0006

Trickey, A. W. (2023). *Estimating power and sample size*.

Tsai, A. C., & Burns, B. F. (2015). Syndemics of psychosocial problems and HIV risk: A systematic review of empirical tests of the disease interaction concept. *Soc Sci Med*, *139*, 26-35. https://doi.org/10.1016/j.socscimed.2015.06.024

United States Census Bureau. (2020). *Poverty*. Retrieved 06/18/2022 from https://www.census.gov/topics/income-poverty/poverty.html

United States Census Bureau. (2023a, 06/20/2023). *American Community Survey*. Retrieved 06/24/2023 from https://www.census.gov/programs-surveys/acs

United States Census Bureau. (2023b). *Census Regions and Divisions of the United States*. Retrieved 06/30/2023 from https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf

Valway, S. E., Sanchez, M. P., Shinnick, T. F., Orme, I., Agerton, T., Hoy, D., Jones, J. S., Westmoreland, H., & Onorato, I. M. (1998). An outbreak involving extensive transmission of a virulent strain of Mycobacterium tuberculosis. *N Engl J Med*, *338*(10), 633-639. https://doi.org/10.1056/nejm199803053381001

Wallace, R. M., Armstrong, L. R., Pratt, R. H., Kammerer, J. S., & Iademarco, M. F. (2008). Trends in tuberculosis reported from the Appalachian region: United States, 1993-2005. *J Rural Health*, *24*(3), 236-243. https://doi.org/10.1111/j.1748-0361.2008.00164.x

Wesner, C. A., Zhang, W., Melstad, S., Ruen, E., Deffenbaugh, C., Gu, W., & Clayton, J. L. (2020). Assessing County-Level Vulnerability for Opioid Overdose and Rapid Spread of Human Immunodeficiency Virus and Hepatitis C Infection in South Dakota. *J Infect Dis*, *222*(Suppl 5), S312-s321. https://doi.org/10.1093/infdis/jiaa231

WHO. (2004, 01/21/2004). *Pushing to rapidly scale-up measures to fight TB and HIV* https://apps.who.int/mediacentre/news/releases/2004/pr5/en/index.html

WHO. (2022). *Global Tuberculosis Report, 2022*. Retrieved 05/23/2023 from https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2022

Yelk Woodruff, R., Hill, A., Marks, S., Navin, T., & Miramontes, R. (2021). Estimated Latent Tuberculosis Infection Prevalence and Tuberculosis Reactivation Rates Among Non-U.S.-Born Residents in the United States, from the 2011-2012 National Health and Nutrition Examination Survey. *J Immigr Minor Health*, *23*(4), 806-812. https://doi.org/10.1007/s10903-020-01065-8

Yuen, C. M., Kammerer, J. S., Marks, K., Navin, T. R., & France, A. M. (2016). Recent Transmission of Tuberculosis - United States, 2011-2014. *PLoS One*, *11*(4), e0153728. https://doi.org/10.1371/journal.pone.0153728

APPENDICES

APPENDIX A

Database Search Terms for Literature Review of TB Outbreaks In Rural Areas

<u>PubMed Search Terms</u>

“Mycobacterium tuberculosis”[Mesh] OR “Tuberculosis”[Mesh] OR tuberculosis[tiab] OR tb[tiab]

AND

“contact tracing”[Mesh] OR “disease outbreaks”[Mesh] OR “disease transmission, infectious”[Mesh] OR contact*[tiab] OR transmission[tiab] OR outbreak*[tiab]

AND

“rural population”[Mesh] OR rural[tiab]

Alternatively

“rural population”[Mesh] OR rural[tiab] OR agricult*[tiab] OR south*[tiab] OR southeast*[tiab]

<u>CINAHL Search Terms</u>

((MH "Mycobacterium Tuberculosis") OR (MH "Tuberculosis+") OR TB OR Tuberculosis)

AND

((MH "Contact Tracing") OR (MH "Disease Outbreaks") OR (MH "Disease Transmission") OR contact* OR transmi* OR outbreak*)

AND

((MH "Rural Population") OR (MH "Rural Areas") OR (MH "Rural Health") OR rural)

<u>Web of Science Search Terms</u>

"mycobacterium tuberculosis" OR tuberculosis OR TB

AND

"contact trac*" OR "outbreak*" OR "transmi*"

AND

rural

<u>EMBASE Search Terms</u>

('tuberculosis'/exp OR 'mycobacterium tuberculosis'/exp OR tb:ab,ti OR tuberculosis:ab,ti)

AND

('contact examination'/de OR 'outbreak'/de OR 'disease transmission'/de OR outbreak:ab,ti OR transmi*:ab,ti)

AND

('rural area'/de OR 'rural population'/de OR rural:ab,ti)

<u>Full string:</u> ('tuberculosis'/exp OR 'mycobacterium tuberculosis'/exp OR tb:ab,ti OR tuberculosis:ab,ti) AND ('contact examination'/de OR 'outbreak'/de OR 'disease transmission'/de OR outbreak:ab,ti OR transmi*:ab,ti) AND ('rural area'/de OR 'rural population'/de OR rural:ab,ti)

APPENDIX B

National Tuberculosis Surveillance System Data Dictionary

Variable Name	Source	Code	Definition
CNTDATE	NTSS	MMDDYYYY	The month, day, and year that the health department responsible for counting TB cases verified the case as TB and included it in the official case count
STATE	NTSS	2-letter state code	Reporting jurisdiction of the TB case
COUNTY	NTSS	County Name	County name associated with the home address (permanent or temporary residence) for this patient
AGE	NTSS	3-digit number	Patient age at report date
AGE3	NTSS	0–4 5–14 15–24 25–44 45–64 65+ UNK	Patient age at report date
ORIGIN	NTSS	FBORN	Born outside the U.S. or in a U.S.-affiliated jurisdiction
		USBORN	Born in the U.S., born abroad of a U.S. citizen or born in a U.S.-affiliated jurisdiction
		UNK	Unknown origin of birth
RACEHISP	NTSS	AMIND	American Indian or Alaska Native, non Hispanic
		ASIAN	Asian, non-Hispanic
		BLACK	Black or African American, non-Hispanic
		HISP	Hispanic
		MULT	Multiple race, non-Hispanic
		NAHAW	Native Hawaiian or Other Pacific Islander, non Hispanic
		WHITE	White, non-Hispanic
		UNK	Unknown race or unknown ethnicity
SEX	NTSS	M	Patient’s sex at birth is male
		F	Patient’s sex at birth is female

		UNK	Patients sex at birth is unknown
HIVSTAT	NTSS	NEG	Laboratory HIV test result is negative
		POS	Laboratory HIV test result is positive
		IND	Laboratory HIV test result is indeterminate
		REFUSED	Refused to have HIV test done
		NOTOFFRD	Laboratory HIV test not offered
		TDUNK	Laboratory HIV test done, but result is unknown
		UNK	HIV status is unknown
HOMELESS	NTSS	N	This patient was NOT homeless at any time during that past 12 months prior to the TB diagnostic evaluation
		Y	This patient was homeless at any time during the past 12 months prior to the TB diagnostic evaluation
		UNK	Unknown whether this patient was unhoused at any time during the past 12 months before the TB diagnostic evaluation
IDU	NTSS	N	This patient has NOT injected illegal drugs within the past 12 months
		Y	This patient has injected illegal drugs within the past 12 months
		UNK	Unknown whether this patient has injected illegal drugs within the past 12 months
NONIDU	NTSS	N	This patient has NOT used noninjecting drugs within the past 12 months
		Y	This patient has used noninjecting drugs within the past 12 months
		UNK	Unknown whether this patient has used noninjecting drugs within the past 12 months
ALCOHOL	NTSS	N	This patient has NOT used alcohol to excess within the past 12 months
		Y	This patient has used alcohol to excess within the past 12 months
		UNK	Unknown whether this patient has used alcohol to excess within the past 12 months
CORRINST	NTSS	N	This patient was NOT an inmate of a correctional facility at the time of the TB diagnostic evaluation

		Y	This patient was an inmate of a correctional facility at the time of the TB diagnostic evaluation
		UNK	Unknown whether this patient was an inmate of a correctional facility at the time of the TB diagnostic evaluation
CORRTYPE	NTSS	FEDPRIS	This patient was a resident of a federal prison
		STAPRIS	This patient was a resident of a state prison
		LOCJAIL	This patient was a resident of a local jail
		CORRFAC	This patient was a resident of a juvenile correctional facility
		OTH	This patient was a resident of another correctional facility
		UNK	Unknown whether this patient was a resident of a correctional facility
STATUS	NTSS	ALIVE	Indicates this patient was alive at the time of diagnosis
		DEAD	Indicates this patient was dead at the time of diagnosis
		UNK	Unknown
		N	If CT scan result is abnormal, there is NO evidence of 1 or more lung cavities
		UNK	If CT scan result is abnormal, unknown whether there is evidence of 1 or more lung cavities
RISKDIAB	NTSS	Y	This patient has a diagnosis of diabetes mellitus either before or at the time of TB diagnosis
RISKIMMUNO	NTSS	Y	This patient had immunosuppression due to either a medical condition or medication, or immunosuppressive therapy
RISKTNF	NTSS	Y	This patient had recently received, or was receiving, TNF-alpha antagonist therapy at the time of TB diagnosis
RISKINFECT	NTSS	Y	This patient is a contact of an infectious TB patient within 2 years or less
RT	NTSS	Y	Indicates that this case of TB disease is due to recent transmission

APPENDIX C

Results of Assumption Testing for Hierarchical Generalized Linear Model

1) There should be a linear relationship between continuous predictor variables and the log odds (or logit) of the dependent variable.

Figure 1

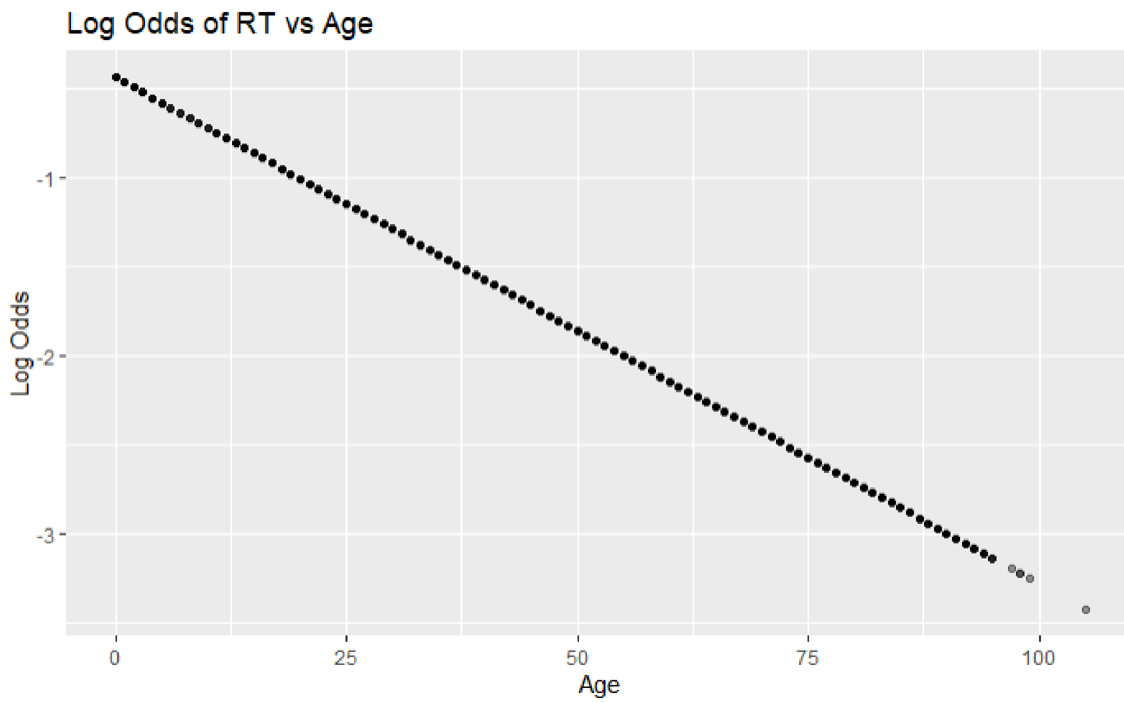


Figure 2

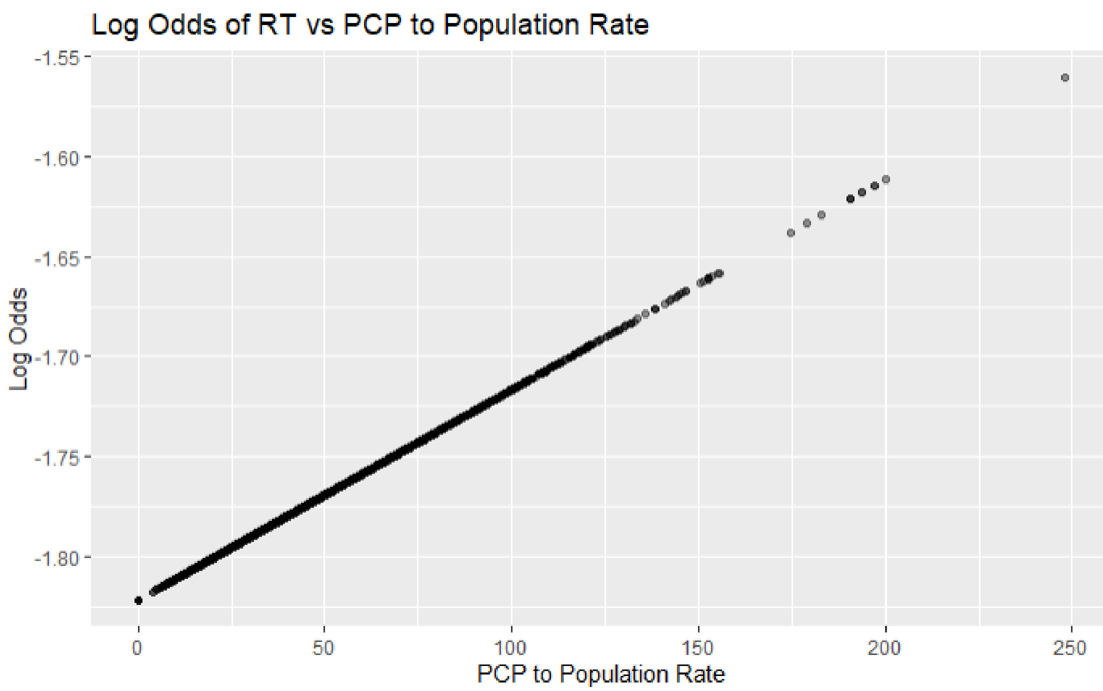
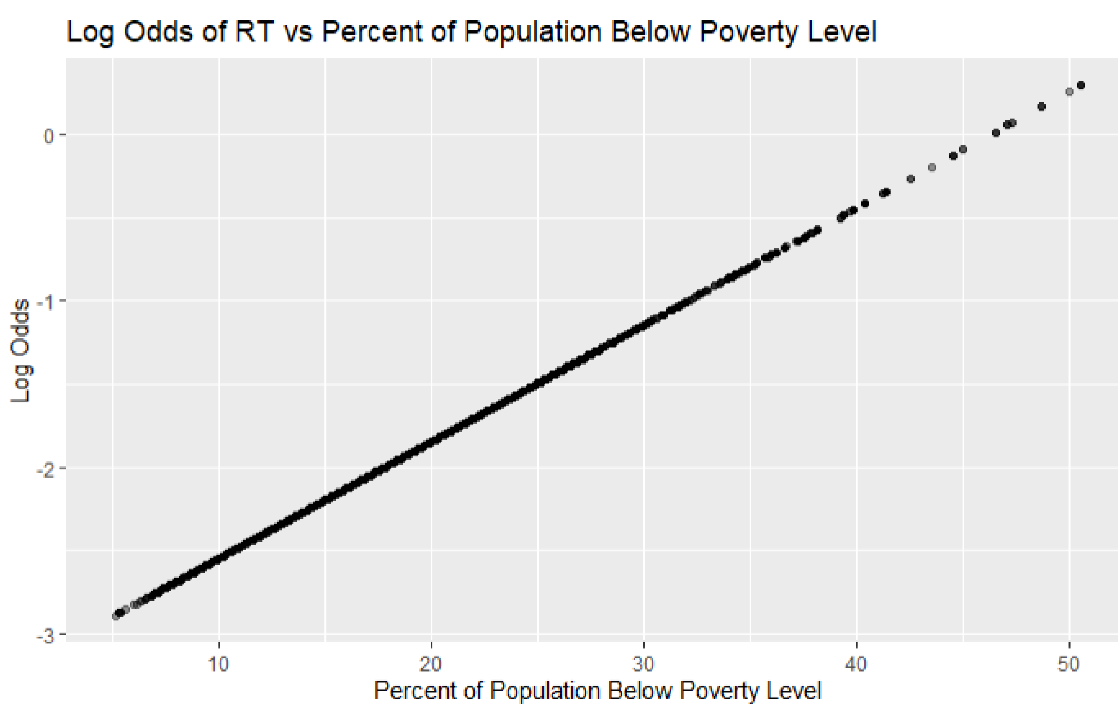


Figure 3



2) There should not be multicollinearity (or redundancy) among independent variables in the model. A variance inflation factor (VIF) test will be run to check for multicollinearity among independent variables. If a VIF>4 is observed, independent variables that are highly correlated will be evaluated to determine which ones can be removed, and the model will be re-run, observing differences in standard errors among the different models.

Table 1

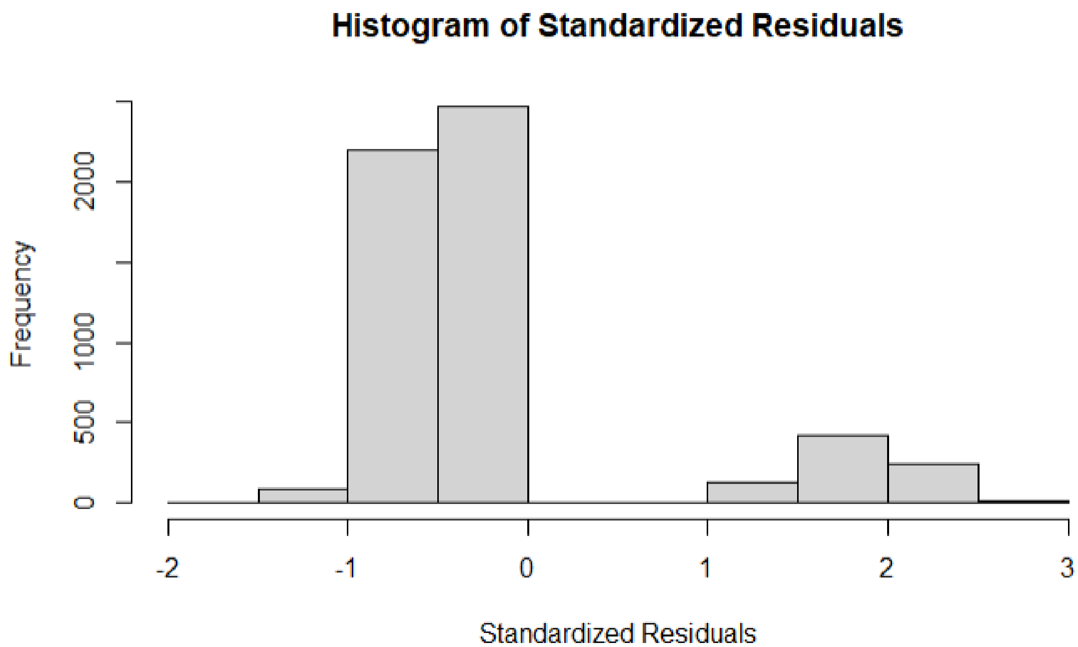
Results of Variance Inflation Factor Testing for Individual-Level Variables

Variable	**VIF**
Age	1.13
US-born	1.08
White race	1.07
Experiencing homelessness	1.04
Incarceration	1.04
Substance misuse	1.08
Presence of coexisting clinical condition	1.06

VIF testing revealed all individual-level variables had VIF <2.

There should not be strongly influential outliers in the data. This assumption only applies to continuous variables and will be addressed by examining standardized residuals for cases in the dataset. If standardized residuals >3 are observed, the model will be re-run after removing those outliers. Standardized residuals >3 were not observed in the histogram.

Figure 4


Histogram of Standardized Residuals
Frequency
0
500
1000
2000
-2
-1
0
1
2
3
Standardized Residuals

APPENDIX D

Results of Sensitivity Analyses

Table 1

Sensitivity Analysis Excluding Alaska for Estimates of RT and Characteristics Associated With RT, 2011–2021

Characteristic	Attributed to recent transmission		Not attributed to recent transmission		
	No	**%**	**No**	**%**	***p*-value**
Total	649	12.3	4625	87.7	
Sex					.090
Male	458	70.6	3109	67.2	
Female	191	29.4	1516	32.8	
Missing					
Age (years)					<.001
0–4	31	4.8	41	0.9	
5–14	10	1.5	32	0.7	
15–24	80	12.3	424	9.2	
25–44	252	38.8	1274	27.5	
45–64	213	32.8	1492	32.3	
65+	63	9.7	1362	29.4	
Unknown/missing	0		0		
Origin of birth					.001
U.S.-born	567	87.4	2761	59.7	
Non-U.S.–born	82	12.6	1860	40.2	
Unknown/missing	0		4	0.1	
Race/ethnicity					<.001
Hispanic	125	19.3	1288	27.8	
White, not of Hispanic ethnicity	125	19.3	1381	29.9	
Black, not of Hispanic ethnicity	273	42.1	1040	22.5	
Asian	9	1.4	549	11.9	
American Indian/Alaska Native	71	10.9	235	5.1	
Native Hawaiian/Pacific Islander	22	3.4	66	1.4	
Multiple race, not of Hispanic ethnicity	24	3.7	46	1.0	
Unknown/missing	0		20	0.4	
Incarcerated at diagnosis					<.001
Yes	143	22.0	471	10.2	

No	502	77.3	4130	89.3	
Unknown/missing	4	0.6	24	0.5	
Unhoused in previous year					.900
Yes	28	4.3	195	4.2	
No	613	94.5	4378	94.7	
Unknown/missing	8	1.2	52	1.1	
Substance misuse					<.001
Yes	189	29.1	884	19.1	
No	447	68.9	3651	78.9	
Unknown/missing	13	2.0	90	1.9	
Diabetes					<.001
Yes	74	11.4	899	19.4	
No	575	88.6	3726	80.6	
HIV co-infection					.300
Yes	21	3.2	177	3.8	
No	589	90.8	3898	84.3	
Unknown/missing	0		0		
Non-HIV immunocompromise					.002
Yes	21	3.2	295	6.4	
No	628	96.8	4330	93.6	
U.S. Census Regions					<.001
Midwest	54	8.3	887	19.2	
Northeast	7	1.1	193	4.2	
South	556	85.7	2969	64.2	
West	32	4.9	576	12.5	
NCHS[1] rural designation					.002
Micropolitan	439	67.6	2835	61.3	
Noncore	210	32.4	1790	38.7	
	Mean	**SD**	**Mean**	**SD**	
Percent of county population living in poverty	23	6	20	7	<.001
PCP[2] per 100,000 persons in county	41	18	48	25	<.001

1. National Center for Health Statistics
2. Primary care physician

Table 2

Sensitivity Analysis Excluding Cases Incarcerated at Diagnosis for Estimates of RT and Characteristics Associated with RT, 2011–2021

Characteristic	Attributed to recent transmission		Not attributed to recent transmission		
	No	%	No	%	*p*-value
Total	700	13.9	4310	86.0	
Sex					.800
Male	446	63.7	2767	64.2	
Female	254	36.3	1543	35.8	
Missing	0		0		
Age (years)					<.001
0–4	41	5.9	47	1.1	
5–14	33	4.7	36	0.8	
15–24	95	13.6	350	8.1	
25–44	220	31.4	1,070	24.8	
45–64	233	33.3	1,436	33.3	
65+	78	11.1	1,371	31.8	
Unknown/missing	0		0		
Origin of birth					<.001
Yes	621	88.7	2,703	62.7	
No	79	11.3	1,603	37.2	
Missing/Unknown	0		4	0.1	
Race/ethnicity					<.001
Hispanic	80	11.4	1,007	23.4	
White, not of Hispanic ethnicity	101	14.4	1,331	30.9	
Black, not of Hispanic ethnicity	205	29.3	935	21.7	
Asian	9	1.3	557	12.9	
American Indian/Alaska Native	263	37.6	353	8.2	
Native Hawaiian/Pacific Islander	22	3.1	65	1.5	
Multiple race, not of Hispanic ethnicity	20	2.9	43	1.0	
Unknown/missing	0		19	0.4	
Incarcerated at diagnosis					N/A
Yes	0		0		
No	696	99.4	4,286	99.4	

Missing/Unknown	4	0.6	24	0.6	
Unhoused in previous year					.400
Yes	34	4.9	179	4.2	
No	665	95.0	4,100	95.1	
Missing/Unknown	1	0.1	31	0.7	
Substance misuse					<.001
Yes	260	37.1	856	19.9	
No	433	61.9	3,385	78.5	
Missing/Unknown	7	1.0	69	1.6	
HIV co-infection					.100
NEG	624	89.1	3,638	84.4	
POS	16	2.3	144	3.3	
Missing/Unknown	60	8.6	528	12.3	
Diabetes mellitus					<.001
Yes	70	10.0	878	20.4	
No	630	90.0	3,432	79.6	
Missing/Unknown	0		0		
Non-HIV immunosuppression					.002
Yes	25	3.6	287	6.7	
No	675	96.4	4,023	93.3	
Missing/Unknown	0		0		
U.S. Census Regions					<.001
Midwest	53	7.6	864	20.0	
Northeast	7	1.0	179	4.2	
South	414	59.1	2,576	59.8	
West	226	32.3	691	16.0	
NCHS[1] rural designation					<.001
Micropolitan	306	43.7	2,590	60.1	
Non-Core	394	56.3	1,720	39.9	
	Mean	**SD**	**Mean**	**SD**	
Percent of county population living in poverty	27	7	19	7	<.001
PCP[2] per 100,000 persons in county	53	30	50	26	.300

1. National Center for Health Statistics
2. Primary care physician

Table 3

Characteristics of Cases in Rural Areas Whose TB Disease is Attributed to Recent Transmission, By U.S. Census Region, Excluding Alaska, 2011–2021

Characteristic	Midwest		Northeast		South		West		
	No	%	No.	%	No.	%	No.	%	*p*-value
Total	54		7		556		32		
Sex									.033
Male	31	57.4	4	57.1	404	72.7	19	59.4	
Female	23	42.6	3	42.9	152	27.3	13	40.6	
U.S.-born									<.001
Yes	43	79.6	5	71.4	498	89.6	21	65.6	
No	11	20.4	2	28.6	58	10.4	11	34.4	
Race/ethnicity									.042
Yes	18	33.3	2	28.6	100	18.0	5	15.6	
No	36	66.7	5	71.4	456	82.0	27	84.4	
Substance misuse									.4
Yes	20	37.0	1	14.3	161	29.0	7	21.9	
No	33	61.1	5	71.4	384	69.1	25	78.1	
Unknown/Missing	1	1.9	1	14.3	11	2.0	0		
Any clinical condition									.041
Yes	13	24.1	1	14.3	72	12.9	8	25.0	
No	41	75.9	6	85.7	484	87.1	24	75.0	

NCHS[1] rural designation									.031
Micropolitan	31	57.4	6	85.7	386	69.4	16	50.0	
Non-Core	23	42.6	1	14.3	170	30.6	16	50.0	
	Mean	**SD[2]**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	
Age (years)	36	21	33	28	41	17	45	23	.3
Percent county population living in poverty	17	7	12	3	24	6	23	10	<.001
PCP[3] per 100,000 persons in county	49	23	73	33	39	17	52	20	<.001

1. National Center for Health Statistics
2. Standard deviation
3. Primary care physician

Table 4

Characteristics of Cases in Rural Areas Whose TB Disease was Attributed to Recent Transmission, By U.S. Census Region, Excluding Cases Incarcerated at Diagnosis, 2011–2021

Characteristic	Midwest		Northeast		South		West		
	No.	%	No.	%	No.	%	No.	%	*p*-value
Total	53	10.4	7	1.4	414	81.2	226	4.4	
Sex									.7
Male	31	58.5	4	57.1	262	63.3	149	65.9	
Female	22	41.5	3	42.9	152	36.7	77	34.1	
Age (years)									.11
Mean (SD)	37 (21)		33 (28)		41 (19)		38 (21)		
Median (IQR)	37 (23, 50)		32 (14, 46)		42 (28, 56)		39 (21, 54)		
Range	0–75		0–80		0–89		0–88		
Origin of birth									<.001
U.S.-born	42	79.2	5	71.4	359	86.7	215	95.1	
Non-U.S.–born	11	20.8	2	28.6	55	13.3	11	4.9	
Race/ethnicity				0.0					<.001
White, not of Hispanic ethnicity	17	32.1	2	28.6	77	18.6	5	2.2	
All racial/ethnic minority groups	36	67.9	5	71.4	337	81.4	221	97.8	
Substance misuse									.8
Yes	19	35.8	1	14.3	152	36.7	88	38.9	
No	33	62.3	5	71.4	257	62.1	138	61.1	
Unknown/Missing	1	1.9	1	14.3	5	1.2	0		
Any clinical condition									<.001
Yes	13	24.5	1	14.3	67	16.2	13	5.8	
No	40	75.5	6	85.7	347	83.8	213	94.2	
NCHS[1] rural designation									<.001
Micropolitan	31	58.5	6	85.7	253	61.1	16	7.1	
Non-Core	22	41.5	1	14.3	161	38.9	210	92.9	
	Mean	**SD[2]**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	
Percent county population living in poverty	17	7	12	3	24	7	26	8	<.001

PCP[3] per 100,000 persons in county	48	24	73	33	42	18	80	33	<.001

1. National Center for Health Statistics
2. Standard deviation
3. Primary care physician

Table 5

Hierarchical Generalized Linear Model of Individual-Level and County-Level Predictors of Recent Transmission with Inclusion of Percent of County Population Below Poverty, Excluding Cases from Alaska, 2011–2021

Fixed Effects	Odds Ratio	Standard error	Lower CI	Upper CI	*p*-value
Intercept	0.05	0.37	0.02	0.10	<.001
Age	0.97	0.00	0.96	0.98	<.001
U.S.-born (Y)	4.56	0.17	3.25	6.39	<.001
White race (N)	1.44	0.16	1.05	1.98	.023
Incarcerated at diagnosis (Y)	0.48	0.30	0.27	0.86	.013
Unhoused in previous year (Y)	0.79	0.26	0.47	1.33	.371
Misuse of substances (Y)	1.19	0.13	0.92	1.54	.187
Any clinical condition (Y)	0.83	0.16	0.61	1.13	.246
Percent of county population below federal poverty	1.01	0.01	0.98	1.04	.296

Random Effects	
σ^2	3.29
τ_{00} FIPS code	1.54
ICC	0.32
N (FIPS code)	333
Observations	3588
Marginal R^2 / Conditional R^2	0.136/0.412
AIC	2441.1

Table 6

Hierarchical Generalized Linear Model of Individual-Level and County-Level Predictors of Recent Transmission with Inclusion of Number of Primary Care Physicians per 100,000 Persons in County, Excluding Cases from Alaska, 2011–2021

Fixed Effects	Odds Ratio	Standard error	Lower CI	Upper CI	*p*-value
Intercept	0.11	0.349408	0.06	0.22	.001
Age	0.97	0.003173	0.97	0.98	<.001
U.S.-born (Y)	4.60	0.169969	3.30	6.42	<.001
White race (N)	1.46	0.160073	1.07	2.00	.018
Incarcerated at diagnosis (Y)	0.46	0.296494	0.26	0.82	.009
Unhoused in previous year (Y)	0.80	0.265439	0.48	1.35	.413
Misuse of substances (Y)	1.20	0.132344	0.93	1.56	.162

Any clinical condition (Y)	0.84	0.157824	0.62	1.14	.259
Primary care physician to population ratio in county	0.99	0.004542	0.98	1.00	.018

Random Effects	
σ2	3.29
τ00 FIPS code	1.5
ICC	0.31
N (FIPS code)	333
Observations	3588
Marginal R2 / Conditional R2	0.144/0.412
AIC	2436.5

Table 7

Hierarchical Generalized Linear Model of Individual-Level and County-Level Predictors of Recent Transmission with Inclusion of Percent of County Population Below Poverty, Excluding Cases Incarcerated at Diagnosis, 2011–2021

Fixed Effects	**Odds Ratio**	**Standard error**	**Lower CI**	**Upper CI**	***p*-value**
Intercept	0.04	0.38	0.02	0.08	<0.001
Age	0.97	0.00	0.96	0.98	<0.001
U.S.-born (Y)	3.98	0.18	2.79	5.68	<0.001
White race (N)	1.6	0.18	1.13	2.27	0.009
Unhoused in previous year (Y)	0.67	0.25	0.41	1.10	0.113
Misuse of substances (Y)	1.3	0.12	1.02	1.66	0.037
Any clinical condition (Y)	0.81	0.16	0.59	1.11	0.172
Percent of county population below federal poverty	1.03	0.01	1.00	1.06	0.038

Random Effects	
σ^2	3.29
τ_{00} FIPS code	1.6
ICC	0.33
N (FIPS code)	343
Observations	3392
Marginal R^2 / Conditional R^2	0.154/0.430
AIC	2464.3

Table 8

Hierarchical Generalized Linear Model of Individual-Level and County-Level Predictors of Recent Transmission with Inclusion of Number of Primary Care Physicians per 100,000 Persons in County, Excluding Cases Incarcerated at Diagnosis, 2011–2021

Fixed Effects	**Odds Ratio**	**Standard error**	**Lower CI**	**Upper CI**	***p*-value**
Intercept	0.11	0.35	0.06	0.22	<.001
Age	0.97	0.00	0.97	0.98	<.001
U.S.-born (Y)	4.15	0.18	2.92	5.90	<.001
White race (N)	1.64	0.18	1.16	2.33	.006
Unhoused in previous year (Y)	0.70	0.25	0.43	1.15	.157
Misuse of substances (Y)	1.33	0.13	1.04	1.71	.026
Any clinical condition (Y)	0.81	0.16	0.59	1.11	.199
Primary care physician to population ratio in county	0.99	0.00	0.98	1.00	.018

Random Effects	
σ2	3.29
τ00 FIPS code	1.6
ICC	0.33
N (FIPS code)	342
Observations	3336
Marginal R2 / Conditional R2	0.141/0.423
AIC	2380.2