Characterizing Socio-Demographic Tuberculosis Disparities and their Underlying Drivers across Subpopulations in both Low and High Burden Settings

by

Maheen Humayun

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Epidemiological Science)
in the University of Michigan
2024

Doctoral Committee:

Associate Professor Zhenhua Yang, Chair
Assistant Professor Gwenyth O.Lee
Research Scientist Wen Ye
Associate Professor Jon Zelner

Maheen Humayun

maheen@umich.edu

ORCID iD: 0000-0003-2172-3753

Dedication

I want to recognize the privileges that defined my opportunity set. A privilege as per its definition has little to do with one's ability but more with unequal and unjust social structures that benefit some demographics more than others. I have benefitted from these structures and I know many around me have had to overcome bigger obstacles to get exactly where I am. I am inspired by so many of you.

I also have the privilege of being surrounded by strong women from an early age. While grappling with many systemic barriers, they protested, resented, and fought in whatever capacity they could to teach me consciously or unconsciously that womanhood was resilience. My Nani often narrated how despite outperforming her male cousins, her family withdrew her education in grade 5. She lived on her dream of education by sending my mother to medical school despite limited financial resources. At a time when there was little to no conversation around gender equality, Ama taught me to question oppressive gender roles at a young age and was unwavering in her belief that education was the ticket to empowerment for her daughters. To my dad, I know that you raised your three girls fearlessly without letting loud misogynistic voices around you get to your head. In Khala, I have an unconditional support system that has always showed up, rain or shine. To my brilliant sisters, Feenu and Mariyam, you are the reason I have been able to chase my dreams, overcome mental health challenges, and navigate life away from home.

My 13-year-old self who was afraid of failing in middle school could have never imagined that I would move miles away from my home city, Rawalpindi, to Lahore for my

undergraduate education and then to the United States for my graduate education. I dared to dream because of the struggles and resilience of these many women who collectively kept raising the glass ceiling for me. I know that this continuous struggle is draining and requires many personal sacrifices; I hope to hold on tightly to their struggles and more so on days when it's harder to believe in myself.

Acknowledgements

This scholarly work builds on the great body of work done by Dr. Joseph H. Bates in Arkansas who served on my doctoral committee from 2021 to late 2023. He remained dedicated to this work until the very end of his life and provided his feedback even during illness. None of this work would be possible without committed collaborators at the Arkansas Health Department, City of Harare and University of Zimbabwe who shaped my understanding of global health justice. A special thank you to Dr. Leonard Mukasa and Dr. Joconiah Chirenda. This dissertation benefitted from the diverse academic backgrounds and nuanced perspectives of supportive committee members. I am so grateful for the opportunity to learn new statistical methods from Dr. Ye who always made herself available to answer my numerous questions while brainstorming creative ways to analyze the data presented here. Much of this work builds on the relentless efforts of many previous members of Yang Lab who were instrumental in data collection and analysis in Harare, Zimbabwe and Arkansas, USA.

Foremost, I want to thank my faculty advisor, Dr. Zhenhua Yang, who guided and mentored me through these years. You provided an encouraging environment where I was able to freely exchange ideas, experiment with new methods, grow as an independent researcher and express my authentic self. I am grateful for your genuine concern and support, which extended beyond our work together, allowing me to complete my PhD despite the ever-challenging COVID-19 pandemic. Science was more enjoyable with our fusion dishes, Michigan hikes and hallway conversations. I am honored to be your last PhD student!

In many of my friends and colleagues from the University of Michigan and my undergraduate lab at Lahore University of Management Sciences, I found a safe space to ask tough questions and a community that thrived on pulling each other up.

Preface

Despite being an ancient disease that has affected human life for centuries, tuberculosis (TB) continues to require public health attention as it is a leading infectious killer worldwide. In TB high burden settings, the risk of disease among the general population is widespread primarily due to ongoing recent transmission in the community[1]. The TB epidemic in many high burden countries is exacerbated by the overlapping human immunodeficiency virus (HIV) epidemic [1,2]. This contrasts with low incidence settings where the TB risk among the general population is markedly lower but with sustained pockets of high TB transmission among urban, houseless population in big cities and minority, indigenous population in rural areas[1,3,4]. Children are at an increased risk of TB in high burden settings while the older population seems to be at a higher risk in low incidence settings[1]. Reactivation of latent TB infection (LTBI), rather than recent transmission, is perceived to be the primary driver of TB in low incidence settings[1]. Given that reactivation of LTBI and recent transmission require specialized mitigation strategies, it is important to understand their relative role in driving the TB epidemic in a given population. The varied epidemiological profiles of TB epidemics across countries, regions and subpopulations necessitates the need for understanding the local context in which TB occurs before designing TB interventions.

Hence, the goal of my dissertation is to acknowledge and highlight the differences in the epidemiological factors that drive TB epidemics across diverse settings, including high and low burden settings. To achieve this goal, I have arranged my dissertation into 3 research chapters

(Chapter 2-4 of the Dissertation). In the first research chapter (Chapter 2), I characterize racial/ethnic disparities in TB incidence in Arkansas during 2010-2021 using detailed racial/ethnic categorizations, typically not available in routine state TB reports. By quantifying racial/ethnic disparities across various clinical presentations of TB, this study also provides insights related to race/ethnicity specific drivers of TB. This chapter has been published in *Emerging Infectious Diseases*[1].

In the second research chapter (Chapter 3), I integrate molecular typing data with traditional surveillance data to better understand the underlying drivers of racial/ethnic disparities observed in the first research chapter. As different interventions are required to interrupt on-going transmission compared to reactivation of LTBI, I examined the relative contribution of recent transmission versus reactivation of LTBI to the TB epidemic in Arkansas and across its racial/ethnic subpopulations. Hence, this chapter provides actionable knowledge for designing locally meaningful TB interventions.

The third research chapter (Chapter 4), published in *Open Forum Infectious Diseases*[2], is set in a high burden setting, Zimbabwe, where I characterize the gender epidemiology of TB. While the gender disparity in TB is widely reported, there is a dearth of studies that use granular categorizations of TB anatomic sites to understand if the effect of gender is consistent or varied across TB clinical presentations. Stratifying risk of TB-HIV coinfection and risk of various TB clinical types by gender provides a more holistic understanding of the role of gender in a TB-HIV high burden setting.

---

[1] Humayun, Maheen, et al. "Racial and Ethnic Disparities in Tuberculosis Incidence, Arkansas, USA, 2010–2021."Emerging Infectious Diseases 30.1 (2024): 116

[2] Humayun, Maheen, et al. "Effect of Gender on Clinical Presentation of Tuberculosis (TB) and Age-Specific Risk of TB, and TB-Human Immunodeficiency Virus Coinfection." Open Forum Infectious Diseases. Vol. 9. No. 10. US: Oxford University Press, 2022.

To summarize, this work provides evidence for why a one-for-all TB approach does not work given differences in the drivers of TB epidemics across diverse settings. The disaggregated analysis in this dissertation provides subpopulation specific understanding of TB drivers and unmasks the underlying disparities in TB incidence, which remain a major impediment in the progress towards TB elimination.

Table of Contents

List of Tables

List of Figures

Abstract

Tuberculosis (TB) continues to be one of the leading infectious killers, despite its long history as a human disease. In many high burden countries (HBCs), the TB epidemic is exacerbated by the overlapping HIV epidemic as HIV infection substantially increases the risk of primary TB infection, reinfection, and reactivation of LTBI. While the burden of TB is widespread among the general public in HBCs, in low incidence countries the TB epidemic often concentrates among socially and historically marginalized subpopulations.

In this dissertation, I characterized the racial/ethnic disparities in TB incidence in Arkansas, United States (U.S.), a low incidence state for TB. By applying an analytical framework that disaggregates TB surveillance data by detailed racial/ethnic categorizations, not widely used in previous state TB reports and studies, I found the risk of TB among Native Hawaiian/Pacific Islander (NHPI) persons to be highest (RR= 174, 95% CI: 140.6, 214.2), followed by Asian, Hispanic and non-Hispanic (NH) Black persons, compared to NH White persons after adjusting for age and sex. I also found that the risk of advanced TB disease at diagnosis was significantly higher for all racial/ethnic minorities, suggesting that the observed racial disparities may be driven by inequitable access to TB diagnosis.

The mere quantification of racial disparities is inconsequential without understanding the underlying mechanisms that drive racial disparities, as each mechanism requires a specialized TB mitigation strategy. By integrating genotyping of *Mycobacterium tuberculosis* clinical isolates of TB incident cases with traditional surveillance data, I quantified the relative

contribution of recent transmission and reactivation of LTBI across racial groups in Arkansas during 2010-2021 using time restricted genotypic clusters as a proxy for recent transmission. Approximately, 1/3rd of the TB cases in the state were clustered, and NHPI (RR=218.7, 95% CI: 150.6, 317.4) and NH Black (RR=7.4, 95% CI: 5.3, 10.2) persons were at a higher risk of clustered TB compared to NH White persons, suggesting a particular need to disrupt ongoing transmission among NHPI and NH Black communities. The risk of clustered TB among ≥65 year NH White persons was 6.5 times (95% CI: 1.52, 27.8) the risk among 15-24 year old NH White persons, demonstrating that previously recognized risk factors for clustered TB may have evolved over time and may not ubiquitously apply to all racial groups. Taken together, Arkansas needs a pro-equity and two-pronged approach for TB elimination, which considers the role of reactivation along with ongoing transmission despite its low overall burden.

This dissertation also characterized gender disparities in a TB-HIV syndemic setting, Zimbabwe. The large TB surveillance dataset (N=24 277) provided sufficient sample size to study the association of gender with TB across granular categorizations of TB anatomic sites, often unavailable in resource limited settings. Women of childbearing age were found to be at an increased risk of TB-HIV coinfection, which may result in increased antenatal transmission. When resolved by TB anatomic sites, women were found to have increased likelihood of severe forms of TB including abdominal TB (male/female odds ratio= 0.51, 95% CI: 0.39, 0.68) and TB of the bones/joints/spine (male/female odds ratio= 0.63, 95% CI: 0.45, 0.90).

To summarize, this dissertation generates new knowledge related to local epidemiological factors that differentially drive TB epidemics across subpopulations in low and high burden settings, highlighting why a one-for-all approach fails to address the local drivers of TB epidemics.

Chapter 1 Introduction

1.1 Etiology of tuberculosis

Tuberculosis (TB) in animals and humans is caused by a number of closely related species that collectively constitute the *Mycobacterium tuberculosis* complex (MTBC*)*. While TB in humans is primarily caused by *Mycobacterium tuberculosis (Mtb)*, other related species such as *M.africanum and M.bovis* have also been documented to cause TB in humans [5,6]. Within human adapted species, the accumulation of mutations over time has resulted in the emergence of several distinct genetic lineages that vary in geographic spread, virulence, transmissibility, immunogenicity and other functional/phenotypic components [6]. With the advent of sophisticated genomic tools, the evolving genetic diversity of *Mtb* strains has been further evaluated, resulting in identification of new lineages such as Lineages 8 and 9, and further classification of major lineages into "evolutionarily meaningful" sub lineages [7–9]

1.2 Transmission of tuberculosis

TB is an airborne infection that spreads from an infected person when they expel respiratory secretions containing infectious particles that travel as droplet nuclei, ranging between 1-5μm in size, before being inhaled by a susceptible person [5,10]. TB transmission from the infectious source to a susceptible host is dependent on the infectiousness of the source, proximity to the source, duration and frequency of contact with the infectious source, and characteristics of the pathogen[1,5]. The infectiousness of the source is determined by several factors including clinical manifestation of disease[1]. Acid fast bacilli (AFB) smear positive TB

cases are associated with an increased risk of infection among household contacts compared to smear negative TB cases due to a higher rate of infectious aerosol production[1,5,10]. Laryngeal and pulmonary TB (PTB) patients are able to efficiently expel infectious particles through activities such as singing, coughing, sneezing etc., into the environment [10]. Delay in diagnosis and treatment prolongs infectiousness of the source resulting in an increase in the number of contacts and duration of contact with the infectious source[11]. Once droplet nuclei are expelled into the environment, environmental factors such as limited UV light exposure, indoor setting, humidity and poor ventilation allow TB particles to remain viable until they can reach a susceptible host[1]. Conventionally, it was believed that TB transmission occurred among close contacts, however, molecular epidemiological studies demonstrated heterogeneity in where TB transmission occurs[1,12]. Household transmission explains only a small proportion of TB transmission as airborne transmission occurs with relatively brief exposure and in various social settings such as schools, public transport, bars, churches, hospitals, prisons, etc[1,12]

1.3 Pathogenesis of tuberculosis

In the host, infectious droplets pass through the nasal passage or mouth, bronchus and upper respiratory tract before they can settle in the alveoli [5,10]. A susceptible person who is exposed to *Mtb* might be able to clear the infection with the help of either innate or acquired T-cell immunity [13]. However, in some individuals the infection will not be eliminated and the bacteria will evade the first line of defense [13]. When alveolar macrophages are not able to completely destroy the bacteria, they begin to multiply intracellularly resulting in additional recruitment of immune cells which form a granuloma around the site of infection to limit further bacterial growth[11,14]. Granuloma can help control the infection but in some situations it will either fail to contain the infection resulting in early progression to disease, or sequester the

bacteria in a latent stage[14]. In the latter case, the granuloma can limit *Mtb* in a quiescent stage with no or slow *Mtb* growth, a stage referred to as 'latent TB infection' (LTBI), for sometimes as long as a few decades or even an entire lifetime[14,15]. LTBI is asymptomatic and noninfectious[5,14]. Based on a study that argues for reconceptualizing TB infection outcomes on a spectrum rather than a binary suggests that granulomas are heterogeneous, including in terms of bacterial growth, with possible slow or intermittent replication[16].

Bacteria may continue to multiply intracellularly and release from the granuloma causing primary TB with symptoms including cough, fever, weight loss, fatigue and lack of appetite [5,13]. Typically, if infection results in active disease within 18-24 months of infection, it is considered "primary disease" [11]. Such early progression to disease from recent infection is more common among young children aged 0-4 years and immunocompromised individuals [11]. Pediatric TB presents differently than adult TB, with over 90% of cases among children progressing to primary disease within 1 year of infection compared to a much longer interval for adults[17]. As a result, children have been recognized as a sentinel population for measuring recent transmission in a given population [1,17].

Only about 5% of those infected with *Mtb* will develop TB during the first two years of infection, while the remaining 5% develop disease at a later point[5]. LTBI can progress to active disease during one's lifetime, often triggered by changes in immunity related to old age, HIV infection, immunosuppressive treatment, etc[18–20]. Hence two distinct mechanisms, primary TB from recent infection and reactivation of a latent infection, can cause TB disease in a given population. Host related factors such as diabetes, malnutrition, HIV, occupational risk, health seeking behavior, and smoking increase the risk of progression to disease[1,10]. Hence, TB is brought about by an intricate play between host, pathogen and environmental factors.

While the TB disease most commonly affects the lungs, referred to as 'PTB', bacteria through the lymphatic system or bloodstream can also reach distant extrapulmonary sites including larynx, lymph nodes, pleura, bones, joints, brain and kidney resulting in 'extra pulmonary TB (EPTB)' [5,13]. The type of TB disease also has implications for transmissibility as PTB patients with symptoms such as vigorous coughing, cavitation and sputum smear positivity are more likely to contribute to transmission while EPTB patients do not generally contribute to transmission [1,5].

1.4 Diagnosis of tuberculosis

Based on Centers for Disease Control and Prevention (CDC) guidelines, medical history, physical examination, testing for *Mtb* infection, chest radiography and bacteriological examination of clinical specimens are the 5 major components of evaluating TB in suspected patients[5].

To detect *Mtb* infection, there are two widely used tests: the Interferon-gamma release assay (IGRA) and tuberculin skin test (TST)[5,13]. While these tests can help distinguish infected patients from uninfected individuals, they are not able to provide evidence of active TB disease[5,13]. Imaging techniques such as chest X-ray are commonly used to visualize radiological abnormalities as a result of PTB[5]. However, it is important to recognize that PTB may present differently in HIV and immunosuppressed patients who due to compromised immunity are less likely to present with lung cavitation which is believed to be a result of immune response to *Mtb*[5,13]. While radiological abnormalities help in diagnosing TB, it does not confirm TB diagnosis[5].

Bacteriological evidence is sought with the help of microscopy, culture examination and nucleic acid amplification tests[5]. For initial bacteriological evidence, sputum specimens are

evaluated for the presence of *Mtb* through AFB smear microscopy[5]. While AFB smear microscopy has a quick turnaround time, it fails to detect *Mtb* when bacterial load is relatively low resulting in positive culture despite negative smear results[5]. For children, obtaining sputum samples is particularly challenging as they typically do not produce sputum upon coughing and require more invasive procedures such as gastric aspiration[5,21]. Children also typically have a lower number of bacilli in their lung lesions, making smear microscopy suboptimal for pediatric TB diagnosis[5,17,21]. The sensitivity of sputum smear microscopy was documented to be below 1% in children aged 0-4 years and around 7% in those under 15 years[21]. HIV positive PTB patients also tend to have lower bacillary load in sputum samples compared to HIV negative PTB patients[22].

CDC recommends nucleic acid amplification (NAA) test for at least one respiratory specimen per patient[5]. Xpert MTB/RIF assay not only rapidly detects the presence of MTBC but also resistance to rifampicin by detecting mutations in the *RpoB* gene [5,23]. Culture examination remains the gold standard for confirming TB diagnosis and all diagnostic specimens should be culture examined regardless of AFB smear and NAA results[5]. In the absence of positive culture, TB may be diagnosed based on clinical symptoms alone[5]. These diagnostic challenges become increasingly important as they can potentially result in delayed treatment, worse outcomes and increased transmission.

1.5 Global status of the tuberculosis epidemic

Evidence of TB affecting human life dates back to 600 B.C.[24]. Despite affecting humans for centuries, TB remains a major global public health challenge. According to the World Health Organization's (WHO) Global Health Estimates, TB was the 13th leading cause of death in 2019 with an estimated incidence of 133 cases per 100,000 population in 2022 globally [25]. TB caused

1.3 million deaths in 2022 making it the second leading cause of death from a single infectious agent, superseded by only COVID-19[25]. Around 10.6 million people were estimated to have fallen ill with TB in 2022, reflecting an increase of 3.9% between 2020-2022 and reversing the downward trajectory of TB incidence that had been sustained over the last several years[25]. Post pandemic during 2020-2022, half a million excess deaths were attributed to pandemic related disruptions in TB care which would not have occurred had the pre-pandemic trend continued[25]. It is estimated that 25% of the world's population is infected with *Mtb* which translates into a 5-10% lifetime risk of developing active TB[5,25]. The risk of developing TB varies and is much higher in the presence of risk factors such as diabetes and HIV[5].

TB disproportionately affects the South East Asia and African regions accounting for 81% of the total TB related deaths[25]. The 30 high burden countries account for 87% of the total TB cases worldwide[25]. There is a remarkable disparity in TB incidence across regions and countries with TB incidence ranging from less than 10 cases per 100,000 population in countries mostly situated in the Americas or European regions to more than 500 cases per 100,000 population per year in countries mostly situated in the African region[25].

1.6 The dual epidemic of tuberculosis and human immunodeficiency virus infection

In 2022, 6.3% of people newly diagnosed with TB were living with HIV globally, however, this coinfection proportion exceeded 50% in parts of Southern Africa [25]. In TB infected individuals with untreated HIV infection and no LTBI treatment, the risk of progression to TB disease is 7-10 % per year compared to a 10% lifetime risk among TB infected people without any risk factors [5]. Among the 38 million people living with HIV, the risk of developing TB was 18 times higher compared to the rest of the world population in 2019 [26]. Additionally, the case fatality rate among HIV coinfected TB patients is approximately twice that of the HIV negative

TB patients [27]. Based on statistics from 2018, the treatment success rate for TB-HIV coinfected patients was 76% compared to an overall success rate of 85% among all new/relapse TB cases and TB-HIV coinfected individuals were three times more likely to die during TB treatment [26].

HIV infection substantially increases the risk of primary TB infection, reinfection and reactivation of latent TB [27]. This is because HIV decreases the immune system's ability to contain TB infection with the depletion of CD4 T lymphocytes [27]. The presentation of TB is often atypical among HIV coinfected individuals especially among those with advanced HIV and low CD4 cell count [27]. While PTB is the most common form of TB among HIV positive individuals, lower CD4 cell count is associated with lower likelihood of lung cavitation and higher likelihood of normal chest X-ray findings [27]. HIV positive individuals with increased immune suppression are more likely to present with EPTB, often concomitantly with PTB, unlike HIV negative individuals who typically present with EPTB alone[27]. While low bacillary load in sputum makes HIV coinfected patients less infectious, it also lowers the sensitivity of sputum microscopy for TB detection[22,27]. The atypical TB presentation among HIV coinfected individuals complicates TB prognosis, diagnosis and treatment.

1.7 Progress towards tuberculosis elimination goals

Based on the staggering TB statistics, WHO developed The End TB Strategy that aims to reduce the number of TB related deaths by 95%, TB incidence by 90% and TB affected households experiencing catastrophic costs, defined as exceeding 20% of the annual household income, to 0% by 2035 compared to the 2015 baseline [25]. The United Nations (UN) held its second high level meeting for TB in 2023 which endorsed WHO's TB elimination targets while also setting other important targets for 2027 such as increasing treatment coverage for TB

patients to 90%, providing preventive TB treatment to 90% of at risk individuals, and setting funding targets for research and treatment[25].

Recent global trends indicate that all WHO End TB 2020 milestones were missed by a significant margin and the progress towards 2025 milestones appears to be subpar[28]. TB deaths reduced by only 18% during 2015-2022 suggesting that the goal of 75% reduction by 2025 is unlikely to be met[25]. TB incidence reduced by only 8.7% between 2015-2022, far from the 2025 milestone of 50% reduction[25]. Around 50% of the TB affected households continue to experience catastrophic costs[25]. While global progress is abysmal, there are regional and country level differences in the progress towards TB elimination as 83 countries met the 2020 milestone of 20% reduction in TB incidence[25].

TB incidence has reduced substantially in low incidence countries, defined as having TB incidence below 10 cases per 100,000 population, but these countries struggle to make the final push towards pre-elimination (<1 case per 100,000 population) and elimination (<1 case per million population) despite being resourceful[29]. Most low incidence countries will have to bring TB incidence down by 2-10 times to meet pre-elimination targets[29]. In the United States, a significant decline of >50% reduction in TB incidence was recorded between 1993 and 2010, however, the annual rate of decline decreased from 6.1% in 2011 to 1.6% in 2019[29]. Based on an analysis of 33 low incidence countries, the current annual rate of decline will not be sufficient to meet elimination targets by 2050, except in one country[30].

Increased globalization has led to greater interdependence between countries and across regions as TB in one place is easily transported to another[30,31]. A substantial proportion of the TB epidemic in low incidence countries is driven by migrants from high burden countries[31]. While TB rates have declined and reached the pre-elimination target among U.S. born persons, the TB

rate among foreign-born persons continues to stay higher with slower annual decline in the U.S.[29]. Regardless of the country specific TB burden, a global concerted effort is required for sustained global TB elimination.

1.8 TB infection prevention and control (IPC)

In order to accelerate progress towards TB elimination, there is a need to strengthen current prevention and control efforts. The WHO has put forward recommendations for TB infection prevention and control (IPC) with three main components: administrative controls, environmental controls and respiratory protection[32]. The administrative controls are meant to reduce *Mtb* exposure in high risk environments through early detection, rapid initiation of effective treatment, respiratory isolation/separation during the infectious period, and promoting respiratory hygiene through health education and counseling[32]. Early case detection of active TB is crucial for timely initiation of treatment which reduces infectiousness of the source and duration of the infectious period resulting in reduced transmission[10,32].

Environmental controls are aimed at reducing infectious droplets in the air through improved ventilation or ultraviolet irradiation[32]. Respiratory protection with the use of medical masks and respirators is effective in reducing TB transmission if used properly[32]. These infection prevention and control measures are not only useful for healthcare settings but also for congregate settings such as correctional facilities, homeless shelters, dormitories, workplaces, refugee camps and hospices where crowding, limited ventilation, delayed diagnosis and limited access to care provides a conducive environment for TB transmission[32].

Immunization is an effective preventive strategy for disrupting the spread of communicable diseases, however, no new effective TB vaccines have been introduced for public use since the introduction of the first TB vaccine over a 100 years ago. BCG vaccination is the

only available vaccination against TB which is recommended for all healthy newborns in high burden settings[33]. Previous studies have reported that the efficacy of BCG against PTB among adults is highly variable, ranging between 0-80%[13]. For children, BCG tends to be efficacious against severe forms of TB such as TB meningitis and disseminated TB[5,13]. The variable efficacy of BCG across age groups and various forms of TB limits its utility for TB control. In the absence of a vaccine that can confer significant immunity against adult TB and reactivation of LTBI, it becomes even more important to strengthen epidemiological investigations aimed at addressing the shortcomings of current TB prevention efforts.

Under these circumstances, prompt identification of TB cases and provision of TB preventive treatment can be instrumental in curtailing the spread of TB. The vast majority of national TB programs are based on passive case finding which depends on a patient making contact with the healthcare system resulting in systematic exclusion of certain individuals due to various factors such as lack of health coverage, limited healthcare access, stigmatization, etc[34]. Hence, provider initiated screening programs can help overcome some of these factors resulting in early identification of missed cases and identification of individuals at an increased risk of progression to disease[34]. According to WHO, people living with HIV (PLHIV), close contacts of TB patients, people exposed to silica and people in prisons should always be screened for TB[34]. Those with an elevated risk of progression to disease such as PLHIV and those with increased exposure to TB disease such as household contacts of bacteriologically confirmed TB should be prioritized for TB preventive treatment to lower the risk of TB[35]. While these are broad recommendations, WHO recognizes the role of local epidemiology in making informed decisions about which groups need to be systematically screened for TB and offered preventive

treatment[34,35]. Epidemiological studies aimed at identifying locally relevant risk groups can help adapt TB preventive strategies at the community level.

1.9 Disparities in TB risk

Health disparities are systematic differences in health risks and outcomes, especially between persistently disadvantaged groups compared to advantaged groups[36]. Healthy People, which is an initiative that provides 10 year measurable objectives for improving health, recognizes that certain groups, based on their age, gender, socio-economic status (SES), disability status, race, ethnicity, religion, geographic location, sexual orientation etc., face greater social and economic barriers to health, often reflecting current and historical exclusion and discrimination[36,37]. While economic disadvantage is the lack of income, wealth and other material resources, social disadvantage is a result of how different groups are treated in a society based on their social identities[37]. Social and economic disadvantage is known to be associated with preventable death, disability and disease[37]. Such modifiable health differences are avoidable, unfair and unjust[36–38]. Health equity, embedded in the principles of social justice, is the absence of such avoidable health differences across various demographic, economic, geographic and social groups[36].

There is a vast body of work that establishes that the risk of TB is heterogeneous across subpopulations[39,40]. Lower SES has been found to be associated with an increased risk of TB infection, reinfection and progression to disease[40]. Lower socio-economic status is associated with proximate risk factors such as food insecurity, poor working conditions, crowded settings, poor ventilation, inaccessible health care and lack of awareness[41]. Data from the US shows that racial and ethnic minorities have an increased risk of LTBI, reactivation and extensive transmission[40,42]. Similarly, nativity is another important determinant of TB risk as foreign born

persons compared to U.S. born persons are at a higher risk of TB likely due to exposure to high TB prevalence in their country of origin and precarious socio-economic conditions after migration[40]. Race and nativity are associated with components of SES such as wealth, income and education, which determines the distribution of proximal factors such as nutrition, psychosocial stress, healthcare and physical environment[40]. Hence, SES mediates the relationship between race/nativity and risk of TB infection and progression through multiple pathways[40]. Gender and age disparities have also been observed for TB but these disparities play out differently across regions and types of TB disease[43–45].

TB control is unlikely to be achieved without understanding the social context in which TB epidemics occur. Intervening on upstream factors such as social, economic and environmental policies can reduce exposure to proximate risk factors[41]. Addressing the socio-economic drivers of TB, both upstream and proximate, at the individual, community, national and international levels will be instrumental in reducing TB incidence[41].

As TB incidence declined in low incidence settings, the remaining TB epidemic became concentrated in socially and historically disadvantaged groups resulting in an increase in TB disparities[40,46]. In high burden settings where resources are often limited, identification of groups at an increased risk can help deliver cost effective and impactful targeted interventions.

1.10 Molecular epidemiology of tuberculosis

Traditional epidemiology has been crucial in understanding TB transmission using contact tracing which allows determining epidemiological links between suspected cases based on patient interviews. However, the variable latency period of *Mtb* infection limits the ability of traditional epidemiology to fully understand TB transmission using the concentric circle or contact tracing approach alone[12]. By integrating molecular typing of *Mtb* clinical isolates with

traditional surveillance data, molecular epidemiology allows to track the movement of TB strains across the population. Molecular typing of the TB pathogen allows to correlate epidemiological characteristics of the disease with the biological properties of the pathogen[12].

Genotyping of *Mtb* isolates has helped resolve long standing knowledge gaps related to TB transmission. A genotypic cluster represents a set of cases with identical genotypes of *Mtb*, suggesting that these cases are a part of the same chain of transmission linked to a common strain[47]. Molecular studies revealed that estimates of recent transmission based on contact tracing were significantly lower when compared to rates of molecular clustering[12]. A study based in San Francisco showed that only 10% of the clustered patients were identified as epidemiologically linked based on contact investigations[12]. Population level molecular studies revealed that a substantial proportion of TB (20-50%) was driven by recent transmission at a time when it was believed that reactivation was the primary driver of TB in low incidence settings[12,47–49]. Differentiating incident TB cases due to reactivation of a historical infection from recent transmission with greater certainty is important because it prompts relevant authorities to employ correct mitigation plans[12].

Apart from identifying outbreaks or clusters unidentified by traditional epidemiological approaches, molecular studies provided new evidence of TB transmission through casual and limited contact in social settings or in the workplace[12,48,49]. In the context of recurrent TB, defined as another episode of TB after clinical resolution of the previous episode, molecular investigations challenged the long held belief that recurrent TB was largely a result of reactivation of the previous TB episode by demonstrating that a substantial proportion of recurrent TB resulted from new exogenous infections, especially in high burden settings[12,48].

Additionally, molecular studies are equipped to study mutations that lead to drug resistance and differences in the epidemiological characteristics of various lineages/strains[12,47].

While advanced molecular typing methods have revolutionized our understanding of TB, they are not a replacement for traditional epidemiologic methods and can best serve TB control efforts when integrated with conventional epidemiological investigations[12]. The precision offered by molecular studies in understanding TB transmission dynamics is dependent on the typing method used and how strain relatedness is defined[12]. Same genotypes do not always reflect recent transmission, especially in populations where strain diversity is limited and some strains happen to be more prevalent[47,49]. In these instances, patient interviews and contact investigations can help to reduce misclassification of cases as clustered[12,47,49].

1.11 Evolution of *Mtb* strain typing methods

The above-mentioned advancements in TB knowledge were made possible because of molecular typing tools that have gone through several generations of evolution over the last few decades. Before the advent of DNA typing methods, drug susceptibility and phage typing with limited discriminatory power were used to distinguish *Mtb* clinical isolates[12,48]. While otherwise highly conserved, *Mtb* genome has certain polymorphic regions that became the basis of typing methods[12]. The first generation of genotyping methods was based on DNA Restriction Fragment Length Polymorphism (RFLP) analysis using *Mtb*-specific insertion element IS*6110* as a marker. The IS*6110* RFLP has been used to estimate recent transmission, characterize the global spread of TB strains, and detect exogenous reinfection[48]. Polymorphic guanine-cytosine-rich repetitive sequence (PGRS) typing, another RFLP based method, was primarily introduced to tackle instances where discriminatory power of IS*6110* typing was questionable[48]. A major limitation of RFLP analysis is that it requires a large amount of DNA from a well-grown culture[50].

As a result, PCR-based methods, used in this dissertation as well, were introduced. These methods are faster because they require lower DNA yield[48,50]. Spacer oligonucleotide typing, a PCR-based method, uses the direct repeat (DR) locus which has 36 base pair long direct repeats interspersed by unique nonrepetitive sequences called spacers[12]. The variability in the presence of these spacers across strains is detected by the spoligotyping hybridization assay[12,48,50]. Another PCR-based method takes advantage of polymorphic loci where repetitive DNA sequences occur in tandem. These are called mycobacterial interspersed repetitive units (MIRU) or variable number of tandem repeats (VNTR) as the number of repeats vary across loci and strains, detectable through MIRU-VNTR typing[12,50].

The discriminatory power of the various *Mtb* strain typing methods is variable. Spoligotyping is less discriminatory than IS*6110* RFLP as indistinguishable spoligotypes were shown to have distinct IS*6110* patterns[12,48,50]. The discriminatory power of MIRU-VNTR typing depends on the number of MIRU loci typed[12]. The initial set of 12 discriminatory loci was expanded to 15 loci and subsequently to 24 loci. When more than 12 MIRU loci typing is combined with spoligotyping, it offers discriminatory power comparable to IS*6110* RFLP typing[12,50].

Despite the comparatively high discrimination power of 24-locus MIRU typing, it considers only 0.03% of the MTB genome when distinguishing between strains[50]. For greater precision in identifying clusters, molecular typing tools have now evolved to detect differences across the whole genome [50]. Whole genome sequencing (WGS) offers greater discriminatory power as it evaluates single nucleotide changes or SNPs over a considerably large genomic region. However, field validation of new molecular typing methods should precede real world epidemiologic application of these methods for meaningful interpretation of study findings.

Chapter 2 Racial and Ethnic Disparities in Tuberculosis Incidence, Arkansas, USA, 2010–2021

This chapter is published in *Emerging Infectious Diseases* (doi: https://doi.org/10.3201/eid3001.230778). Authors: Maheen Humayun, Leonard Mukasa, Wen Ye, Joseph H. Bates, Zhenhua Yang

2.1 Abstract

We conducted an epidemiologic assessment of disease distribution by race/ethnicity to identify subpopulation-specific drivers of tuberculosis (TB). We used detailed racial/ethnic categorizations for the 932 TB cases diagnosed in Arkansas, USA, during 2010–2021. After adjusting for age and sex, racial/ethnic disparities persisted; the Native Hawaiian/Pacific Islander (NHPI) group had the highest risk for TB (risk ratio 173.6, 95% CI 140.6–214.2) compared with the non-Hispanic White group, followed by Asian, Hispanic, and non-Hispanic Black. Notable racial/ethnic disparities existed across all age groups; NHPI persons 0–14 years of age were at a particularly increased risk for TB (risk ratio 888, 95% CI 403–1,962). The risks for sputum smear–positive pulmonary TB and extrapulmonary TB were both significantly higher for racial/ethnic minority groups. Our findings suggest that TB control in Arkansas can benefit from a targeted focus on subpopulations at increased risk for TB.

2.2 Introduction

Globally, tuberculosis (TB) is the 13th leading cause of death and a leading infectious killer, second to COVID-19 [51,52]. In 2021, TB affected 10.6 million persons and caused 1.6 million deaths worldwide [52]. Moreover, a quarter of the world's population is latently infected with *Mycobacterium tuberculosis*, which puts them at a 5%–10% lifetime risk of developing active disease [5,52,53]. Global trends indicate that the World Health Organization's (WHO) End TB Strategy milestones will likely be missed because TB incidence has reduced by only 10% and TB deaths by 5.9% during 2015–2021, compared with the desired goals of 20% reduction in incidence and 35% reduction in deaths [52,54]. The COVID-19 pandemic has further disrupted TB notification and treatment, reversing the progress toward global TB elimination while widening existing inequalities [52,55].

The TB epidemic is considered a major challenge in low-resource, high-burden countries; 30 high-burden countries represent 87% of the global burden [52]. However, the epidemiology of TB in low-incidence countries is characterized by concentration of the TB burden, sometimes as high as in high-burden countries, among socially and historically marginalized populations [30,31]. Before the impact of COVID-19 on TB notification, TB incidence continued to decline during 1993–2019 in the United States; however, the annual rate at which TB declined plateaued during the later years [56]. Similarly, a recent study from Arkansas reported no significant decline in TB incidence during 2009–2020 [57]. To move toward preelimination (1 case/100,000 population) and eventually elimination (<1 case/100,000 population), current TB interventions should be adapted to unique local challenges focusing on populations at increased risk for TB as suggested by WHO's action framework for TB elimination in low-incidence countries [30,31]. This framework calls for epidemiologic assessment of disease distribution in the local population by important

sociodemographic variables; however, such disaggregated analysis is typically not available in TB surveillance reports, limiting the ability of public health programs to develop pro-equity policies [30,31,58,59].

In the United States, race is a strong social determinant of health because it serves as a proxy for systemic and structural barriers to equitable opportunities for education, employment, earning, housing, and healthcare, which perpetuates racial discrimination and unjust distribution of resources that lead to adverse health outcomes [42,60]. In this study from Arkansas, USA, a state with TB incidence below the national average of 2.4 cases/100,000 population, we quantify the racial/ethnic disparities in TB incidence at the population level using detailed racial/ethnic categorizations that have not been widely used in previous TB studies in Arkansas [3,61,62]. This study will not only help map subpopulations at an increased risk for TB in a low-burden setting but also guide the development of targeted TB interventions in light of the underlying factors that differentially drive TB incidence across racial/ethnic groups.

This study used de-identified patient data that we retrospectively retrieved from the TB surveillance database maintained by the Arkansas Health Department. The University of Michigan Health Sciences and Behavioral Sciences Institutional Review Board (IRB-HSBS) determined that this study was not regulated.

2.3 Methods

2.3.1 Study Population

This study included all 932 TB cases diagnosed in Arkansas during 2010–2021. All of these cases met the Centers for Disease Control and Prevention (CDC)'s definition of a verified TB case. These verified cases met either laboratory or clinical case definition, including those verified by provider diagnosis as described in CDC's TB case reporting manual [63].

2.3.2 Data Collection and Data Sources

Demographic and clinical data available in this dataset were collected on verified TB cases using the standard CDC TB reporting form. We obtained US Census Bureau official population estimates for 2010–2019 from annual resident population estimates for 6 race groups (5 race-alone groups and 2 or more races) by age, sex and Hispanic origin for states and the District of Columbia from April 1, 2010–July 1, 2019; April 1, 2020; and July 1, 2020. We obtained population estimates for 2020 and 2021 from annual state resident population estimates for 6 race groups (5 race-alone groups and 2 or more races) by age, sex, and Hispanic origin from April 1, 2020–July 1, 2021 [64].

2.3.3 Data Analysis

This study had 4 objectives. The objectives were to characterize racial/ethnic disparities in TB risk; to determine if the observed racial/ethnic disparities were a result of underlying differences in sex and age distribution; to track age-specific incidence for racial/ethnic groups to draw inferences related to the underlying drivers of TB incidence; and to characterize racial/ethnic differences in advanced TB disease at diagnosis. From the TB surveillance dataset, we created a combined variable for race and ethnicity with 5 categories so that we did not consider racial/ethnic identities in isolation [65]. The Hispanic category included all racial subcategories. We categorized non-Hispanic persons into Asian, non-Hispanic Black, Native Hawaiian/Pacific Islander (NHPI), and non-Hispanic White categories. We did not include American Indian/Alaska Native (n = 4) and multirace (n = 1) categories in race/ethnicity-stratified results because of their small sample sizes. We categorized age as 0–14 years, 15–24 years, 25–44 years, 45–64 years, and ≥65 years. Consistent with the definition used by CDC,

US-born persons included those who were eligible for US citizenship at the time of birth regardless of place of birth.

To address the first objective, we calculated the overall TB incidence for the period 2010–2021, with corresponding 95% CI for the state and the 5 racial/ethnic groups mentioned. We calculated TB incidence per 100,000 population using the population estimates from the US Census Bureau [64]. We estimated TB incidence in the overall population and stratified by race/ethnicity and age group using Poisson regression with an offset term for the total population size. To further characterize TB-related disparities, we calculated risk ratio (RR) estimates for the race/ethnicity-combined variable, sex, and age group using Poisson regression. Because age and sex are important determinants of TB risk, we calculated race/ethnicity RRs that were adjusted for both age and sex concurrently using Poisson regression to achieve our second objective. In the absence of genotyping data, age-specific TB incidence can help draw inferences related to the underlying mechanisms that drive TB incidence. Previous studies demonstrated that TB among older adults strongly reflects reactivation of latent TB infection (LTBI) in low-burden settings, whereas TB is typically a result of recent transmission among young children [66,67]. Hence, the third objective determined how racial disparities tracked across age groups by reporting RRs for race/ethnicity from age-specific Poisson models adjusted for sex. In addition, to track age disparities within each racial/ethnic group, we reported RRs for age groups from race/ethnicity-specific Poisson models while adjusting for sex.

Pulmonary TB (PTB) often starts with minimum infiltrate and progresses to additional infiltrate; sputum smear positivity has been used previously as a proxy for delayed diagnosis [62]. The occurrence of extrapulmonary TB (EPTB) often reflects the spread of *M. tuberculosis* outside of the lungs due to the host's inability to contain the infection [68]. We

evaluated sputum smear–positive PTB and EPTB as important indicators for advanced disease. For those 2 outcomes, we calculated RRs and adjusted RRs for the race/ethnicity-combined variable, sex, and age group using Poisson regression.

We used the non-Hispanic White category as the reference group for race/ethnicity and female as the reference category for sex because those groups were at the lowest risk for TB compared with other variable categories. For age, no one category was at the lowest risk for TB across stratified results, so we used the ≥65 years age category as the reference group to follow how TB risk progressed with age. We determined model fit using a goodness-of-fit test. We conducted statistical analysis using the SAS OnDemand for Academics (SAS Institute Inc.,https://www.sas.com).

2.4 Results

2.4.1 Characteristics of Study Sample

Among the 932 TB cases in our study, 72% were bacteriologically confirmed through either nucleic acid amplification test or positive culture (**Table 2.6.1**). Most TB cases (76.5%) were exclusively diagnosed as PTB patients. Among all TB patients, 35.73% had a positive sputum smear result and 43.2% of the patients with PTB diagnosis (including those with both PTB and EPTB) had a positive sputum smear result. Of the total study sample, 86.6% of TB patients identified as non-Hispanic with diverse racial categorizations; 63.2% of the study patients were male. Most case-patients among Asian (91.7%), NHPI (67.4%), and Hispanic (82.4%) persons were not US born, and all of the non–US-born NHPI TB case-patients in this study were born in the Marshall Islands. For the non-Hispanic White group, only 2.3% of TB

cases were non–US born, and for the non-Hispanic Black group, only 6.1% of TB cases were non–US born.

2.4.2 Racial/Ethnic Disparity in TB Incidence

The overall TB incidence in Arkansas was 2.6 (95% CI 2.4–2.8) cases/100,000 population during 2010–2021 (**Table 2.6.2**). Upon stratifying by race/ethnicity, the NHPI persons (131.6 [95% CI 111.4–155.5] cases/100,000 population) had the highest incidence of TB followed by Asian (20.0 [95% CI 16.6–24.2] cases/100,000 population), Hispanic (4.8 [95% CI 4.0– 5.7] cases/100,000 population), non-Hispanic Black (4.4 [95% CI 3.9–5.0] cases/100,000 population), and non-Hispanic White persons (1.2 [95% CI 1.0–1.3] cases/100,000 population).

2.4.3 Racial/Ethnic Disparity after Adjusting for Sex and Age Differences

Based on the unadjusted model **(Table 2.6.3),** the risk for TB was many fold higher for all racial/ethnic groups when compared with the non-Hispanic White group. The risk for TB for NHPI persons was 113 (95% CI 92.1–137.7) times the risk for non-Hispanic White persons. Asian (RR 17.1, 95% CI 13.8–21.3) Hispanic (RR 4.0, 95% CI 3.3–5.0), and non-Hispanic Black (RR 3.8, 95% CI 3.2–4.5) persons all had higher risk for TB than non-Hispanic White persons. Male persons were at an 81% (RR 1.8, 95% CI 1.6–2.1) higher risk for TB than female persons. The risk for TB was 66% (RR 0.3, 95% CI 0.3–0.4) lower for the youngest group, 0–14 years of age, compared to the oldest age group. TB risk increased with age; the ≥65-year age group had the highest risk for TB.

After adjusting for age group and sex, racial/ethnic disparities continued to persist. NHPI persons were at the highest risk for TB compared with non-Hispanic White persons (RR 173.6,

95% CI 140.6– 214.2), followed by Asian (RR 21.6, 95% CI 17.3– 27.0), Hispanic (RR 5.9, 95% CI 4.8– 7.4), and non-Hispanic Black (RR 4.6, 95% CI 3.9–5.5) persons.

2.4.4 Age-Related Racial/Ethnic Disparities in TB Incidence

The statewide TB incidence was highest among the ≥65-year age group (3.8 [95% CI 3.3– 4.3] cases/100,000 population) whereas the youngest group, 0–14 years, had the lowest incidence (1.3 [95% CI 1.1–1.6] cases/100,000 population). Risk for TB increased with age in Arkansas (**Table 2.6.2**). The 0–14-year age group had the lowest risk for TB compared with the ≥65 year age group for non-Hispanic White (RR 0.05, 95% CI 0.02– 0.11), non-Hispanic Black (RR 0.18, 95% CI 0.11–0.30), Asian (RR 0.09, 95% CI 0.02–0.42), and Hispanic (RR 0.08, 95% CI 0.04–0.18) groups, (**Table 2.6.4**). We observed no significant age differences for the NHPI group ($p = 0.13$). NHPI persons 0–14 years of age had 888 (95% CI 403–1,962) times and NHPI persons ≥65 years of age had 55 (95% CI 20–148) times the risk of TB compared to similarly aged non-Hispanic White persons (**Table 2.6.5**).

2.4.5 Racial/Ethnic Disparity in Advanced Disease at Diagnosis

The risk for sputum smear-positive PTB was highest for NHPI persons (RR 138.8, 95% CI 94.7–203.7), followed by Asian (RR 14.4, 95% CI 9.5–22.0), Hispanic (RR 5.5, 95% CI 3.8–8.0), and non-Hispanic Black (RR 4.8, 95% CI 3.7–6.3) persons when compared with non-Hispanic White persons (**Table 2.6.6**). The risk for EPTB was highest for NHPI persons (RR 133.3, 95% CI 83.7–212.4), followed by Asian (RR 31.4, 95% CI 21.1–46.7), Hispanic (RR 5.3, 95% CI 3.3–8.3), and non-Hispanic Black (RR 4.5, 95% CI 3.1–6.3) persons compared to non-Hispanic White persons (**Table 2.6.6**).

2.5 Discussion

The TB incidence in the United States and Arkansas has been <10 cases/100,000 population for several years; however, neither the state nor the country has been able to make the final push toward preelimination and elimination targets. The underlying epidemiologic factors that drive the remaining TB epidemic in low-incidence countries differ from those in high-burden settings and also across subpopulations within low-incidence settings, which suggests the need for locally informed, context-specific TB interventions [69].

Our study provides an in-depth epidemiologic understanding of the concentrated TB epidemic in Arkansas that is not well captured by aggregated statewide estimates. We found remarkable disparities in TB incidence around the axes of race/ethnicity, sex, and age. Of particular importance were racial/ethnic disparities, which could not be explained by age and sex differences across racial/ethnic groups. Age-specific TB incidence and differences in clinical manifestation of TB at diagnosis across racial/ethnic categories hold important lessons for understanding the drivers of TB incidence and challenges related to health equity in Arkansas.

The racial/ethnic disparities that we observed in our study are consistent with previous studies conducted in Arkansas and the United States, which consistently reported racial disparities in risk for LTBI, recent transmission, and TB disease [3,42,46,61,70]. In 2021, a total of 88.1% of the TB cases reported in the United States were attributable to racial/ethnic minorities [56,71]. Such observed racial disparities can be explained as a consequence of structural racism that perpetuates health inequities primarily through 2 interlinked pathways: residential segregation and inadequate healthcare [40,60]. Persons from racial/ethnic minority groups are more likely to live in neighborhoods with high population density, limited healthcare access, poor housing conditions, and greater air pollution, which makes them more susceptible to acquiring TB

infection [1,40,60]. Moreover, those persons are more likely to experience conditions such as diabetes, HIV, and malnutrition that can contribute toward progression to TB disease [40,41]. In essence, race is associated with socioeconomic status (SES), including generational wealth, income, and education, which then mediates the relationship between race and susceptibility to infection and progression to disease through psychosocial stress, nutrition, physical environment, healthcare access, and immune function [39,40].

The NHPI persons in Arkansas had a strikingly high TB incidence of 131.6 cases/100,000 population, many times higher than their annual national-level TB incidence rate that has remained <20 cases/100,000 during 2003–2021 [56,72]. Most of the TB case-patients among NHPI in our study were born in the Marshall Islands, where TB incidence was 280.6 cases/100,000 population in 2020 and 343.2 cases/100,000 population in 2021 [56,72]. Decades of colonial rule and testing of nuclear weapons by the US government during 1946–1958 had socioeconomic repercussions for the health infrastructure of the Marshall Islands, where the prevalence of comorbidities that substantially increase TB risk is alarmingly high [73]. Under the Compact of Free Association, the people of the Marshall Islands can freely travel, live, and work in the United States, where they experience language, cultural, and economic barriers when accessing healthcare that can lead to infection, delayed diagnosis, and prolonged infectiousness with implications for community transmission [74,75]. Despite the high incidence of TB in the Marshall Islands, persons from that country are not required to undergo screening for LTBI or active disease upon arrival to the United States [4]. Screening those persons for TB may help in early diagnosis and treatment, thereby reducing the burden of TB in Arkansas and among NHPI persons.

TB is driven by various mechanisms, mainly by reactivation of LTBI or primary disease by a recent transmission, each of which requires specialized mitigation strategies [76]. TB incidence was 139.6 cases/100,000 population among 0–14-year-old NHPI children in our study; incidence was<2 cases/100,000 population among all other racial/ethnic groups of similar age. The risk among 0–14-year-old NHPI children was 888 times the risk among non-Hispanic White children of similar age. Because progression from infection to disease is rapid, TB among young children is a good marker for ongoing community transmission [17,77]. Given the high risk for sputum smear positive PTB and elevated TB risk among 0–14-year-olds, TB transmission appears to play a role in driving TB incidence among NHPI persons. Hence, to curtail the disproportionate TB burden for NHPI persons, mitigation strategies should focus on active case finding in addition to LTBI screening of adults to disrupt chains of transmission [4]. Curtailing the TB epidemic will also require ramping up contact tracing based on contact disclosure from TB patients, who often hesitate to name contacts because of stigma around TB in their communities [78]. Community-level advocacy and awareness using culturally appropriate tools can improve contact disclosure and equip the local community with the necessary information on TB diagnosis, treatment, and transmission [78].

For Hispanic, non-Hispanic White, and non-Hispanic Black persons, the risk for TB was highest among the oldest age group, indicating that TB is likely driven by reactivation of LTBI in those subpopulations [20]. Two previous studies conducted in Arkansas using genotyping data from 1997–2010 demonstrated that TB incidence among ≥65-year-olds was largely driven by nonclustered TB incidence, which is indicative of reactivation of LTBI [3,61]. In addition, foreign-born non-Hispanic White, non-Hispanic Black, Asian, and Hispanic persons had significantly higher risks for TB than did US-born persons (**Table 2.7.1**). The high risk among foreign-born

persons likely indicates reactivation, as suggested by a previous national-level study based on genotyping data [79]. Racial/ethnic groups with a high proportion of foreign-born TB cases can particularly benefit from TB control efforts focused on preventing reactivation of LTBI.

Sputum smear–positive PTB is highly infectious because its high bacillary load leads to an elevated risk for transmission [80,81]. In addition, the increased risk of smear-positive PTB at first diagnosis in persons from racial/ethnic minority groups likely points toward differential access to timely and adequate TB care, previously supported by a study from Arkansas, and such difference can result in severe disease with poor health outcomes [62,81]. Another study from Arkansas found that TB patients from rural Arkansas faced delays in receiving the correct diagnosis and were treated for other conditions for several months [78]. Our findings suggest the need to explore the barriers related to TB care that affect various subpopulations in Arkansas. Increasing awareness of TB among healthcare workers, especially in a time when TB incidence is low in the country can help equip them with the knowledge needed to make timely and accurate TB diagnosis [78].

The increased risk for EPTB, including concurrent PTB, among persons from racial/ethnic minority groups is indicative of elevated risk for advanced disease that is a diagnostic and a therapeutic challenge due to the dissemination of the disease [82]. An increased prevalence of comorbidities such as HIV might predispose racial/ethnic minority groups to EPTB, suggesting the need for improved management of risk factors that compromise host immunity [83]. The Arkansas Office of Minority Health and Health Disparities reported that among non-Hispanic Black persons, the mortality rate for HIV was 6 times higher and for diabetes was 2 times higher than the rates among non-Hispanic White persons during 2011–2015 [84]. TB preventive strategies should go beyond curtailing transmission and focus on improved

comanagement of noncommunicable conditions, which are often modifiable risk factors, by collaborating across health programs to provide more holistic patient-centered care [85].

This study relies on surveillance data that provides access to limited study variables. We used markers of advanced TB disease to make inferences related to access to timely and adequate care in Arkansas. To clarify the factors that limit access to timely care, future studies should collect qualitative data crucial for determining delays in healthcare and assessing if these delays are patient- or provider-related to bridge health inequities related to TB care. Distinguishing between ongoing community transmission and reactivation of remotely acquired TB infection is crucial when designing TB interventions, but lack of genotyping data in this study prevented reliable evaluation of the relative contribution of each of these 2 mechanisms across racial/ethnic groups [86]. Despite those limitations, this study provides incidence and RR estimates stratified by detailed racial/ethnic categories that had not been previously reported at the population level for Arkansas using the most recent surveillance data [3,58,61,62].

The state-level estimates of TB incidence in Arkansas are misleading because the progress toward TB elimination is unequally felt across racial/ethnic subpopulations. Our findings demonstrate that drivers of TB incidence vary across subpopulations, which necessitates designing context-specific TB interventions. Although our results may not be generalizable to other low-incidence settings, the racial/ethnic disparities we observed demonstrate the need for detailed disaggregated analysis of TB surveillance data by race/ethnicity while providing a framework for such an analysis in other US states.

Acknowledgments

We thank colleagues in the Tuberculosis Control Program at the Arkansas Health Department for their assistance in the collection of the TB surveillance data used in this study.

2.6 Tables

Table 2.6.1. Demographic and clinical characteristics of 932 TB patients diagnosed in Arkansas, USA, during 2010-2021

Characteristic	No. (%)
Diagnosis confirmation	
Bacteriologically confirmed	669 (71.78)
Clinically diagnosed	263 (28.22)
TB disease site	
Pulmonary	713 (76.50)
Extrapulmonary	164 (17.60)
Both	55 (5.90)
Sputum smear result	
Positive	333 (35.73)
Negative	473 (50.75)
Not done	126 (13.52)
Race/ethnicity	
Hispanic	125 (13.41)
Asian	109 (11.70)
Non-Hispanic Black/African American	245 (26.29)
Native Hawaiian/Pacific Islander	138 (14.81)
Non-Hispanic White	310 (33.26)
American Indian or Alaska Native	4 (0.43)
Multirace	1 (0.11)
Sex	
F	336 (36.05)
M	589 (63.20)
Unknown	7 (0.75)
Age group	
0–14 y	92 (9.87)
15–24 y	85 (9.12)
25–44 y	253 (27.15)
45–64 y	282 (30.26)
≥65 y	220 (23.61)
Origin of birth	
US born	613 (65.77)
Non-US born	319 (34.23)
Year	
2010	76 (8.15)
2011	85 (9.12)
2012	71 (7.62)
2013	72 (7.73)
2014	93 (9.98)
2015	90 (9.66)
2016	91 (9.76)
2017	85 (9.12)
2018	76 (8.15)
2019	65 (6.97)
2020	59 (6.33)
2021	69 (7.40)

*Percentages may not total 100 because of rounding. TB, tuberculosis.

Table 2.6.2. Average TB incidence by race/ethnicity, Arkansas, USA, 2010-2021

	Incidence (95% CI)					
Category	All ages	0–14 y	15–24 y	25–44 y	45–64 y	≥65 y
State	2.58 (2.42–2.76)	1.29 (1.05–1.59)	1.75 (1.42–2.17)	2.76 (2.44–3.13)	3.09 (2.75–3.48)	3.79 (3.32–4.33)
Hispanic	4.75 (3.99–5.67)	1.40 (0.80–2.47)	3.64 (2.26–5.86)	6.27 (4.76–8.28)	7.51 (5.25–10.75)	16.61 (10.02–27.56)
Asian	20.02 (16.60–24.16)	1.79 (0.45–7.14)	18.52 (11.17–30.72)	25.45 (19.29–33.59)	29.85 (21.33–41.77)	19.45 (9.73–38.90)
Non-Hispanic Black	4.42 (3.90–5.01)	1.65 (1.07–2.53)	2.98 (2.04–4.35)	4.45 (3.50–5.66)	6.34 (5.11–7.88)	8.54 (6.44–11.33)
NHPI	131.62 (111.39–155.51)	139.60 (105.51–184. 71)	108.08 (68.94–169.44)	109.27 (80.45–148.40)	206.17 (139.31–305.11)	158.67 (59.55–422.75)
Non-Hispanic White	1.17 (1.04–1.31)	0.16 (0.07–0.33)	0.22 (0.10–0.46)	0.66 (0.49–0.89)	1.50 (1.24–1.81)	2.88 (2.44–3.39)

*Incidence is no. cases/100,000 population. The table provides overall and age-stratified TB incidence for the state of Arkansas and for racial/ethnic groups as calculated by Poisson regression. NHPI, Native Hawaiian/Pacific Islander; TB, tuberculosis.

Table 2.6.3. Disparity in tuberculosis incidence by race, sex, and age, Arkansas, USA, 2010-2021

Covariate	Unadjusted risk ratio (95%CI)	Age- and sex-adjusted risk ratio (95%CI)
Race/ethnicity†		
Hispanic	4.07 (3.30–5.01)	5.94 (4.79–7.37)
Asian	17.14 (13.77–21.33)	21.61 (17.32–26.98)
Non-Hispanic Black	3.78 (3.20–4.48)	4.60 (3.88–5.45)
NHPI	112.64 (92.13–137.73)	173.58 (140.64–214.24)
Non-Hispanic White	Referent	Referent
Sex‡		
F	Referent	Referent
M	1.81 (1.59–2.07)	1.88 (1.64–2.15)
Age group§		
0–14 y	0.34 (0.27–0.44)	0.16 (0.12–0.20)
15–24 y	0.46 (0.36–0.59)	0.23 (0.18–0.30)
25–44 y	0.73 (0.61–0.87)	0.36 (0.30–0.43)
45–64 y	0.82 (0.68–0.97)	0.58 (0.49–0.70)
≥65 y	Referent	Referent

*RR calculated by Poisson regression. NHPI, Native Hawaiian/Pacific Islander; RR, risk ratio.

†Using unadjusted RR model 1.

‡Using unadjusted RR model 2.

§Using unadjusted RR model 3.

Table 2.6.4. Age disparities in tuberculosis incidence across racial/ethnic groups, Arkansas, USA, 2010-2021*

	Risk ratio (95% CI)				
Age group	Non-Hispanic White, no. cases = 305	Non-Hispanic Black, no. cases = 244	Asian, no. cases = 109	NHPI, no. cases = 138	Hispanic, no. cases = 124
0–14 y	0.05 (0.02–0.11)	0.18 (0.11–0.30)	0.09 (0.02–0.42)	0.87 (0.31–2.42)	0.08 (0.04–0.18)
15–24 y	0.07 (0.03–0.15)	0.33 (0.20–0.52)	0.92 (0.39–2.16)	0.68 (0.23–1.99)	0.21 (0.11–0.43)
25–44 y	0.22 (0.15–0.31)	0.50 (0.34–0.72)	1.27 (0.60–2.68)	0.68 (0.24–1.90)	0.37 (0.21–0.65)
45–64 y	0.50 (0.39–0.64)	0.72 (0.50–1.03)	1.51 (0.70–3.26)	1.29 (0.45–3.72)	0.44 (0.24–0.82)
≥65 y	Referent	Referent	Referent	Referent	Referent

*RR calculated by Poisson regression. NHPI, Native Hawaiian/Pacific Islander.

Table 2.6.5. Racial/ethnic disparities in tuberculosis incidence across age groups, Arkansas, USA, 2010-2021*

	Risk ratio (95% CI)				
Race/ethnicity	0–14 y, no. cases = 91	15–24 y, no. cases = 85	25–44 y, no. cases = 249	45–64 y, no. cases = 277	≥65 y, no. cases = 218
Hispanic	8.93 (3.52–22.68)	16.68 (6.92–40.22)	9.37 (6.22–14.13)	4.84 (3.23–7.26)	5.57 (3.27–9.48)
Asian	11.38 (2.36–54.78)	85.02 (34.67–208.53)	38.81 (25.75–58.50)	20.72 (14.08–30.50)	7.09 (3.48–14.46)
Non-Hispanic Black	10.50 (4.46–24.69)	13.77 (6.00–31.63)	6.83 (4.64–10.06)	4.36 (3.26–5.81)	3.06 (2.21–4.25)
NHPI	888.70 (402.55–1961.98)	496.40 (208.68–1180.84)	162.91 (105.93–250.55)	138.01 (89.26–213.38)	54.82 (20.30–148.08)
Non-Hispanic White	Referent	Referent	Referent	Referent	Referent

*RR calculated by Poisson regression. The table reports RR estimates with 95% CI while adjusting for sex. NH, non-Hispanic; NHPI, Native Hawaiian/Pacific Islander; RR, risk ratio.

Table 2.6.6. Disparities in advanced TB disease at diagnosis by race/ethnicity, sex, and age, Arkansas, USA, 2010-2021 *

	RR (95% CI) for sputum smear-positive PTB		RR (95% CI) for EPTB	
Covariate	Unadjusted	Age- and sex-adjusted	Unadjusted	Age- and sex-adjusted
Race/ethnicity†				
Hispanic	3.42 (2.39–4.90)	5.54 (3.84–8.00)	3.61 (2.31–5.66)	5.26 (3.32–8.34)
Asian	11.07 (7.28–16.82)	14.43 (9.46–22.02)	25.97 (17.59–38.35)	31.37 (21.09–46.66)
Non-Hispanic Black	3.80 (2.90–4.99)	4.80 (3.65–6.31)	3.68 (2.59–5.22)	4.46 (3.13–6.34)
NHPI	78.73 (54.40–113.94)	138.84 (94.65–203.67)	89.90 (57.41–140.78)	133.32 (83.68–212.41)
Non-Hispanic White	Referent	Referent	Referent	Referent
Sex‡				
M	2.27 (1.79–2.87)	2.40 (1.89–3.05)	1.43 (1.09–1.87)	1.49 (1.14–1.95)
F	Referent	Referent	Referent	Referent
Age group§				
0–14 y	NA	NA	0.20 (0.11–0.36)	0.10 (0.05–0.18)
15–24 y	0.42 (0.27–0.64)	0.22 (0.14–0.34)	0.37 (0.22–0.64)	0.20 (0.11–0.35)
25–44 y	0.73 (0.54–0.99)	0.37 (0.27–0.51)	0.93 (0.66–1.31)	0.48 (0.33–0.69)
45–64 y	1.04 (0.78–1.38)	0.75 (0.56–1.00)	0.66 (0.45–0.96)	0.48 (0.33–0.71)
≥65 y	Referent	Referent	Referent	Referent

*Estimates used Poisson regression, EPTB includes cases concurrently diagnosed with PTB. EPTB, extrapulmonary tuberculosis; NA, not available; NHPI, Native Hawaiian/Pacific Islander; PTB, pulmonary tuberculosis; RR, risk ratio; TB, tuberculosis.

†Using unadjusted RR model 1.

‡Using unadjusted RR model 2.

§Using unadjusted RR model 3.

¶Not available due to small sample size.

Supplemental Material for Chapter 2

To obtain U.S. born and foreign-born population estimates, the American Community Survey (ACS) 1-year Public Use Microdata Sample (PUMS) files were used for all years except 2020. Due to the COVID-19 impact, 1-year ACS Public Use Microdata Sample (PUMS) data file was not available for 2020 and was replaced with 1-year ACS Public Use Microdata Sample (PUMS) with experimental weights. For 2020, ACS 1-year PUMS with Experimental Weights was used. Poisson regression was used to calculate nativity RRs for NH White, NH Black, Asian and Hispanic groups. For NHPI group, negative binomial regression was used to calculate nativity RR as Poisson regression model did not provide a good fit.

Table 0.1. Nativity risk ratio with 95% CI across racial/ethnic groups, Arkansas, United States, 2010-2021 *

Nativity	NH White	NH Black	Asian	NHPI	Hispanic
U.S. born	Ref	Ref	Ref	Ref	Ref
Non–U.S. born	2.93 (1.38–6.20)	8.89 (5.27–14.98)	5.11 (2.58–10.10)	1.69 (0.85–3.36)	7.19 (4.54–11.40)

*NH, non-Hispanic; NHPI, Native Hawaiian/Pacific Islander; Ref, reference.

Chapter 3 What Drives Tuberculosis (TB) Disparities across Racial/Ethnic Groups? Lessons Learned from a Population-based Molecular Epidemiology Study in Arkansas, 2010-2021

3.1 Abstract

While significant racial TB disparities in Arkansas are well documented, what drives these disparities in the state is not well understood. Integrating molecular typing data of *Mtb* clinical isolates from incident TB cases with TB surveillance data can determine the relative contribution of recent transmission versus reactivation of latent TB infection (LTBI) across racial/ethnic groups. For the 627 TB cases included in this study, surveillance and genotyping data were obtained from the Arkansas Health Department (AHD) from 2010-2021. Cases with identical spoligotyping and 24-loci MIRU-VNTR typing patterns diagnosed within one year were considered as clustered while the rest were considered as non-clustered. Clustered cases were used as a proxy for measuring recent transmission. The proportion of clustered TB ranged between 16.7% in 2012 to 51.1% in 2020. The risk of clustered TB was significantly higher for NHPI persons (RR=218.7, 95% CI: 150.6, 317.4), and NH Black persons (RR=7.4, 95% CI: 5.3, 10.2) while the risk of non-clustered TB was higher for all racial/ethnic minorities, compared to NH Whites. Stratified analyses revealed that NH Black men and NH White men were at a significantly higher risk for clustered TB compared to their women counterparts. The risk of clustered TB among ≥65 year NH White persons was 6.5 times (95% CI: 1.52, 27.8) the risk among 15-24 year old NH White persons. This study provides new insights into what drives TB

disparities in Arkansas and calls for refocusing TB interventions based on race/ethnicity specific risk factors as traditionally recognized risk factors may not ubiquitously apply.

3.2 Introduction

Every subsequent year between 1993 and 2020 experienced a downward trend for tuberculosis (TB) rate, with the exception of 2015, in the United States (U.S.); possibly owing to increased federal funding and renewed attention for TB post the TB resurgence during 1986-1992[31,87]. The aggregated TB incidence in the U.S. was 2.5 cases per 100,000 population in 2022, however, the TB incidence was many fold higher among racial/ethnic minorities[87]. For example, the Native Hawaiian/Pacific Islander (NHPI) persons had a TB incidence of 24.4 cases per 100,000 population compared to the lowest incidence of 0.4 cases per 100,000 among Non-Hispanic (NH) White persons[87]. While TB incidence has substantially decreased at the national level, racial disparities continue to grow as shown by a national level study conducted during 1994-2016. For Asian persons, the risk of TB dramatically increased from 12.6 times in 1994 to around 30 times the risk among White persons in 2016 [70]. In settings where substantial reductions have been made in overall TB burden, the remaining TB epidemic often concentrates in segments of the population that have been historically and socially marginalized[30,88]. Given the contagious nature of TB, the growing inequality in the progress towards TB elimination is not only a social justice issue but also a potential threat to the gains made in the fight against TB.

In any given population, TB is mainly driven by two mechanisms: recent transmission where cases progress to active disease within a short duration of infection, or reactivation of a latent TB infection (LTBI) acquired in the past. After evading the first line of defense, if *Mtb)* cannot be contained within a granuloma formed by host immune cells, the initial infection progresses to active disease [13]. However, in some cases the tubercle bacilli can be contained

within the granuloma, with some evidence of slow or intermittent growth, during a stage referred to as LTBI[5,13,16]. The infection can be contained in this stage, sometimes for as long as a few decades, but with a 5-10% risk of progressing to active disease in one's lifetime[5,13]. While characterizing disparities is an important first step, identifying the mechanisms that drive these disparities is crucial, as each of the above-mentioned mechanisms require specialized mitigation strategies for TB control[10,89].

Traditional epidemiology, which relies on contact tracing, is not able to distinguish between these underlying mechanisms as it tends to assume that contacts of infectious individuals are infected by the source case without correlating epidemiological evidence with the biological properties of the pathogen[48]. Furthermore, identifying contacts is dependent on the patient's ability to provide reliable information and the investigator's skills[90]. These limitations of contact tracing have been addressed by molecular epidemiology studies, which integrate DNA fingerprinting of *Mtb* clinical isolates with traditional epidemiological data. These studies found evidence that TB transmission does not necessarily require prolonged exposure and can spread through casual and limited contact in social settings, which challenged our understanding of TB transmission based on contact investigations alone [12]. Molecular typing of TB isolates helps to distinguish between recent transmission and reactivation of LTBI as identical genotypes tend to represent transmission events related to the same strain if diagnosed within a shorter timeframe in a geographically defined population [3,48,50]. Molecular tools further enhanced the understanding of TB transmission dynamics by examining genetic diversity of *Mtb* strains, describing the extent of TB transmission by recent versus reactive disease and distinguishing exogenous reinfection from relapse[12].

Genotyping methods such as *IS 6110* Restriction Fragment Length Polymorphism (RFLP), followed by quicker PCR- based methods such as spoligotyping and mycobacterial interspersed repetitive unit variable number of tandem repeats (MIRU-VNTR) typing made it possible to examine strain diversity[12,48]. These methods evaluate a small portion of the *Mtb* genome and are currently being replaced by whole genome based approaches that distinguish strains based on single nucleotide changes analyzed over a much larger genomic region[50]. While molecular methods have been used widely used for examining risk factors for recent and reactive disease, their continuous application at the population level remains crucial for understanding the relative contribution of recent and reactive disease to the TB epidemic, allowing to optimally adapt TB interventions to the unique local epidemiological TB profile[31,91,92].

A recent study conducted in Arkansas reported remarkable racial/ethnic disparities in TB incidence during 2010-2021, however, the underlying mechanisms that drive TB incidence within each racial/ethnic subpopulation remain unclear[88]. By combining spoligotyping and 24-locus mycobacterial interspersed repetitive unit-variable number tandem repeat typing (MIRU-VNTR) with conventional TB surveillance data, we examine subpopulation specific drivers of TB in Arkansas. Integration of molecular tools with traditional epidemiology data sources generates knowledge that has greater utility for TB control given the complex nature of TB transmission[12].

3.3 Methods

3.3.1 Study Population

During the study period, 2010-2021, 933 new TB cases were diagnosed in Arkansas and 69% (n=644) of these were confirmed by bacterial culture. Of the 644 culture confirmed cases, 627 cases (97.4%) had complete genotyping data available. If more than one isolate was

available for a TB case, the earliest specimen with valid genotyping result was retained. Fisher's exact test was used to determine any systematic differences between culture-confirmed cases with and without genotyping data to examine if exclusion of those without genotyping data resulted in biasing the study sample.

3.3.2 Data Sources

TB surveillance and genotyping data were obtained from the Arkansas Health Department (AHD). Socio-demographic and clinical data on verified TB patients were collected by AHD using Centers for Disease Control and Prevention (CDC)'s Report of a Verified Case of Tuberculosis (RVCT) manual [63]. The TB Genotyping Information Management System (TB GIMS) database, maintained by the AHD, was used to access *Mtb* isolate genotyping information.

We used resident population estimates stratified by age, sex and race/ethnicity for incidence and risk ratio (RR) calculation. U.S. Census Bureau official population estimates for 2010-2019 were obtained from "Annual Resident Population Estimates for 6 race groups (5 Race Alone Groups and Two or More Races) by Age, Sex and Hispanic Origin for States and the District of Columbia: April 1, 2010 to July 1, 2019; April 1, 2020; and July 1, 2020" and population estimates for 2020 and 2021 were obtained from "Annual State Resident Population Estimates for 6 race groups (5 Race Alone Groups and Two or More Races) by Age, Sex and Hispanic Origin: April 1, 2020 to July 1, 2022"[64].

Given that this study used available deidentified patient data, it was determined as not regulated by the University of Michigan Health Sciences and Behavioral Sciences Institutional Review Board (IRB-HSBS).

3.3.3 Classification of clustered versus non-clustered cases

To distinguish between recent transmission and reactivation with improved specificity, we used time-restricted cluster definition that combines genotyping data with date of diagnosis[3,42,61]. National Tuberculosis (TB) Genotyping Service (NTGS) assigned *MTB* genotypes to TB cases based on unique combinations of spoligotype and 24-locus MIRU-VNTR genotyping patterns in the TB GIMS dataset[93]. If a genotype was shared by more than one case, it constituted a genotypic cluster. Within a genotypic cluster i.e. cases with identical genotypes, we applied time restriction such that each subsequently diagnosed case had its specimen collected within one year (365 days or less) of the specimen collection date of the previous case in the same cluster. Within a time restricted cluster, cases could be more than one year apart as long as each case had at least one other case diagnosed within the one-year time frame. Cases that did not meet the time restricted cluster definition were considered 'non-clustered' while those who met the time restricted definition were considered as 'clustered'. Clustered cases were used as a proxy for recent transmission while those that were non-clustered were used as a proxy for reactivation. Cases identified as Bovis or Bovis-BCG cases were excluded from (n=15) this analysis.

3.3.4 Determination of state-level proportion and incidence of clustered and non-clustered TB

The first aim of this study was to determine the relative contribution of the two underlying mechanisms, recent transmission and reactivation of remotely acquired TB infection, to the TB epidemic in Arkansas. To address this aim, we calculated the proportion of clustered cases among TB cases with available genotyping data for years between 2010 and 2021. We also calculated the incidence of clustered and non-clustered TB for the period, 2010-2021, in Arkansas with the corresponding 95% confidence interval (CI) using Poisson regression with an offset term for the total population size.

3.3.5 Incidence and risk ratio (RR) calculation for race/ethnicity, sex, and age groups

The second aim of the study was to characterize socio-demographic disparities in the risk of the two outcomes, recent transmission and reactivation, at the state level. To address this aim, Non-Hispanic (NH) persons were categorized into 4 racial categories including Asian, Black, Native Hawaiian/Pacific Islander (NHPI) and White while Hispanic persons regardless of race were treated as one category. Due to small sample size, the American Indian/Alaska Native (n=4) category was not included in the race/ethnicity-stratified results. Cases with missing values for sex were excluded (n=6). Age was categorized into the following categories: 0-14 years, 15-24 years, 25-44 years, 45-64 years, and ≥65 years. The two outcomes, clustered and non-clustered TB, were modeled using separate Poisson regression models. Incidence and risk ratios (RR) were calculated for race/ethnicity, sex, and age groups.

We used groups at the lowest risk of TB as the reference categories, i.e. females for sex and NH White for race/ethnicity. While the 0-14 year age group had the lowest risk across stratified analyses but this was likely due to an underestimation of TB disease in children due to diagnostic challenges around specimen collection and paucibacillary nature of disease[21]. As a result, we chose the 15-24 year age group as the reference category.

3.3.6 Race/ethnicity stratified risk ratio (RR) calculation

Third, we aimed to identify subgroups at an increased risk of recent transmission and reactivation within each racial/ethnic group. To address this aim, we reported race/ethnicity stratified RRs for age group and sex using Poisson regression models.

Model fit for all models was determined using a goodness of fit test. Statistical analysis was conducted using the SAS OnDemand for Academics (SAS Institute Inc., Cary, NC, USA).

3.4 Results

3.4.1 Characteristics of study population

In Arkansas during 2010-2021, 644 (69%) TB cases were culture confirmed and 97.4% of these culture positive cases had complete genotyping information available. Our study population that included culture confirmed cases with complete genotyping data was racially diverse with 36.8% NH White, 28.4% NH Black, 12.3% Hispanic, 11% Asian and 10.9% NHPI (**Table 3.6.1**). Males represented 64.3% and U.S. born persons represented 67.9% of the study sample. A small percentage (1.9%) of the study population was below the age of 15 years. Most cases (80.7%) presented with pulmonary TB alone. Through Fisher's exact test, we found that the distribution of gender, age group, place of birth and race/ethnicity was similar for culture confirmed cases with and without valid genotyping results (p-value>0.05). The difference in the distribution of clinical presentation was marginally significant with a p-value of 0.05.

3.4.2 Clustering of TB cases at the state level

After applying time restriction, cases clustered into 61 time-restricted clusters and the size of these clusters varied between 2 and 42 cases. Around 36.3% (n=222) of the culture confirmed cases with valid genotyping results (excluding Bovis and Bovis-BCG cases) were clustered based on the time-restricted definition. During the study period, the proportion of clustered TB cases among culture positive cases with valid genotyping data ranged between 16.7% in 2012 to 51.1% in 2020 (**Figure 3.6.1**). The proportion of non-clustered TB cases ranged between 80.5% in 2012 to 48.9% in 2020. The incidence of clustered TB during 2010-2021 was 0.61 cases per 100,000 population (95% CI: 0.53, 0.70) and the incidence of non-clustered TB was 1.09 cases per 100,000 population (95% CI: 0.98, 1.20) (**Table 3.6.2**).

3.4.3 Significant racial/ethnic disparities in clustered TB incidence

The risk of clustered TB was significantly higher for both NHPI (RR=218.7, 95% CI: 150.6, 317.4) and NH Black persons (RR=7.4, 95% CI: 5.3, 10.2) compared to NH White persons while the risk of clustered TB among Asian (RR= 1.62, 95% CI: 0.40, 6.65) and Hispanic persons (RR= 1.70, 95% CI: 0.87, 3.31) was comparable to NH White persons (**Table 3.6.2**).

Overall, we found that the risk of clustered TB among men was 1.93 times (95% CI: 1.46, 2.55) the risk among women. When stratified by race, we did not find gender disparity for NHPI (RR=0.91, 95% CI: 0.53, 1.57) and Hispanic (RR=3.61, 95% CI: 0.77, 16.99) persons (**Table 3.6.3**). However, the risk of clustered TB among NH White men and NH Black men was 2.8 times (95% CI: 1.56, 4.93) and 2.2 times (95% CI: 1.41, 3.34) the risk among NH White and NH Black women, respectively. Among Asians, there were only two clustered cases, both of which were diagnosed among men.

Statewide, the risk of clustered TB was significantly lower among 0-14 year olds (RR=0.17, 95% CI: 0.07, 0.38) compared to the 15-24 year age group (**Table 3.6.2**). The 25-44 year (RR=1.37, 95% CI: 0.89, 2.10), 45-64 year (RR= 1.50, 95% CI: 0.98, 2.30) and ≥65 year (RR= 0.79, 95% CI: 0.47, 1.33) age groups had a similar risk of clustered TB when compared to the 15-24 year age group. Upon stratification by race/ethnicity, the 0-14 year age group continued to be at a significantly lower risk for NH Black (RR= 0.20, 95% CI: 0.07, 0.62) and NHPI (RR=0.14, 95% CI: 0.04, 0.49) persons. Among Asian, NH White and Hispanic persons, there were no clustered cases among 0-14 year olds (**Table 3.6.3**). Among NH White persons the risk of clustered TB for 45-64 years and ≥65 years was 6.4 times (95% CI: 1.53, 26.9) and 6.5 times (95% CI: 1.52, 27.8) the risk of the 15-24 year reference category, respectively. For NH

Black, NHPI and Hispanic persons, the 25-44, 45-64 and ≥65 year age groups had a similar risk of clustered TB as the 15-24 year age group. Given TB case status, the probability of clustered TB for NHPI (probability=0.76, 95% CI= 0.65, 0.85) and NH Black (probability= 0.53, 95% CI: 0.45, 0.60) persons was over 50% (**Table 3.6.5**). The probability of clustered TB was 67% for the 0-14 year group (95 %CI: 0.38, 0.87) and 17% for the ≥65 year age group (95% CI: 0.12, 0.24).

3.4.4 Significant racial/ethnic disparities in non-clustered TB incidence

Compared to NH White persons, the risk of non-clustered TB was significantly higher for NHPI persons (RR=25.3, 95% CI: 15.1, 42.3), followed by Asian (RR=20.5, 95% CI: 15.4, 27.2), Hispanic (RR=4.08, 95% CI: 3.05, 5.45) and NH Black persons (RR=2.53, 95% CI: 1.94, 3.30) (**Table 3.6.2**).

Overall, the risk of non-clustered TB for men was 1.84 times (95% CI: 1.49, 2.26) the risk of non-clustered TB for women (**Table 3.6.2**). When stratified by race, we did not find gender disparity for the NHPI (RR=1.17, 95% CI: 0.44, 3.15), Hispanic (RR=1.61, 95% CI: 0.96, 2.68) and Asian (RR= 1.50, 95% CI: 0.92, 2.44) categories (**Table 3.6.4**). However, the risk of non-clustered TB for NH White men and NH Black men was 2.4 times (95% CI: 1.67, 3.30) and 1.6 times (95% CI: 1.05, 2.50) the risk for NH White and NH Black women, respectively.

Overall, the risk of non-clustered TB was significantly lower for 0-14 year olds, (RR=0.08, 95% CI: 0.03, 0.24) compared to that for the 15-24 year age group (**Table 3.6.2**). The 25-44 year (RR=1.54, 95% CI: 1.04, 2.30), 45-64 year (RR= 2.02, 95% CI: 1.38, 2.97) and ≥65 year (RR= 3.41, 95% CI: 2.33, 4.99) age groups had a higher risk of non-clustered TB when compared to the 15-24 year age group. Upon stratification by race/ethnicity, the 0-14 year age

group continued to have a significantly lower risk for NH Black (RR= 0.07, 95% CI: 0.01, 0.56) and Hispanic (RR=0.18, 95% CI: 0.04, 0.91) persons (**Table 3.6.4**). Among Asian and NH White persons, there were no non-clustered cases among 0-14 year olds. For NHPI persons, the risk of non-clustered TB among 0-14 year olds was comparable to that of the 15-24 year category. Among NH White persons, the risk of non-clustered TB for ≥65 year and 45-64 year age groups was 20 times (RR=19.7, 95% CI: 6.24, 62.3) and 7 times (RR=7.17, 95% CI: 2.23, 23.0) the risk for the 15-24 age group, respectively. For NH Black persons, the risk of non-clustered TB was highest for the ≥65 year age group (RR=4.05, 1.94, 8.43). For Hispanic persons, the risk of non-clustered TB for 25-44 years, 45-64 years and ≥65 years was 2.8 times (95% CI: 1.14, 6.64), 3.5 times (95% CI: 1.40, 8.89) and 8.7 times (5% CI: 3.15, 23.8) the risk of non-clustered TB for the 15-24 year age group, respectively. For NHPI and Asian persons, the 25-44, 45-64 and ≥65 year age groups had a similar risk of non-clustered TB as the 15-24 year age group.

3.5 Discussion

This study draws from important descriptive epidemiological studies, which reported significant racial/ethnic TB disparities in Arkansas and United States[46,57,70,88]. However, the mere quantification of racial disparities is inconsequential without catalyzing this knowledge to bridge existing health inequities. By delineating the underlying mechanisms that drive TB disparities in Arkansas, this study goes a step further by providing actionable knowledge crucial for tailoring TB interventions based on molecular epidemiological evidence of relative contribution of recent transmission and reactivation of LTBI to TB incidence in different subpopulations.

Our findings suggest that recent transmission continues to pose a major challenge to TB control in Arkansas. The drivers of TB disparity varied across racial/ethnic subpopulations as

NH Black and NHPI persons were at an increased risk of recent disease compared to NH White persons while all racial/ethnic minorities, especially Asian and NHPI persons, were at an increased risk of reactivation when compared to NH White persons. Apart from identifying racial disparities for reactivation and recent transmission, this study found that the effects of age and sex varied across racial/ethnic groups, calling for reevaluation of traditionally recognized risk factors for recent transmission and reactivation.

In Arkansas, a low incidence state, recently transmitted cases accounted for a substantial portion (36%) of the overall TB burden during the study period and the proportion of clustered cases was as high as 51% in 2020. The largely held belief that reactivation of remotely acquired TB infections mainly drives TB incidence in low incidence settings might be harmful as it overlooks the need for a two-pronged approach that disrupts local community transmission along with prevention of reactivation of LTBI [89,94]. It is important to recognize that estimates of clustering are dependent on the choice of the genotyping method and the definition of clustering used. Previous studies conducted in Arkansas and Michigan used a similarly strict definition for recent transmission by classifying cases with identical genotypes as recently transmitted cases only if they were diagnosed within a one-year timeframe allowing for improved specificity for identifying recent transmission[3,42,61].

NHPI and Black persons were at an increased risk of clustered TB suggesting significant ongoing community transmission within these subpopulations. Several previous studies have consistently reported higher risk of clustered TB among racial/ethnic minority groups [79,95]. However, previous TB studies conducted in Arkansas focused on NH Black and NH White categories resulting in limited understanding of what drives TB in other racial/ethnic groups [3,61]. The largest cluster of 42 cases identified in this study constituted of all NHPI persons and

spanned from 2013 to 2021. Given this evidence of sustained transmission among these subpopulations, there is a need for targeted active case finding and prompt initiation of effective treatment in order to reduce prolonged infectiousness and extensive transmission within the NH Black and NHPI communities[96]. To gain insights into TB transmission dynamics, we analyzed co-membership of 8 large clusters (>5 cases) in this study. We found that 4 of them were racially homogenous including a cluster of 42 NHPI, a cluster of 6 NH Black, a cluster of 6 NH Black and a cluster of 7 NH Black persons. Among the remaining 4 clusters, 3 of them had both NH White and NH Black persons while there was one cluster with both NH Black and Hispanic persons. While this analysis provides insights into TB transmission dynamics within and across subpopulations likely due to social mixing patterns, it is important to note that cluster size and identification of clustered cases can be impacted by factors such as differential access to TB diagnosis.

For Asian and Hispanic groups, the risk of clustered TB was comparable to NH White persons but the risk of non-clustered TB was many folds higher. Non-clustered TB is typically indicative of reactivation of TB infection acquired during a high incidence era or imported infection from endemic countries as patient genotypes do not match other circulating strains in the population at the time [92]. For groups at higher risk of reactivation, TB control efforts should focus more on management of risk factors that impair host immunity resulting in reactivation of LTBI. Homelessness, incarceration, immigration from high burden countries, diabetes, smoking, HIV and immune suppressive treatment are some of the recognized risk factors for reactivation[19,97]. Importation of infection can be minimized by screening immigrants pre and post migration by timely treating LTBI and active TB [98].

Similar to previous studies, older individuals in our study were at an increased risk of reactivation when compared to the younger age group[3,61,99]. This is likely explained by age related immune dysfunction and comorbidities that increase the risk of reactivation [18]. The older age group in low incidence settings is typically believed to represent earlier birth cohorts that were exposed to higher rates of TB resulting in reactivation during later years in life[3,46,61]. We found that overall the relative contribution of clustered TB versus non-clustered TB was lower for the oldest group (Table 3.6.5). For NH White and NH Black groups, the relative contribution of clustered TB versus non-clustered TB was only around 20% in the 65 years and over age groups.

Younger individuals are typically considered to be at a higher risk of recent transmission but in our study NH White persons aged above 45 years had a higher risk of recent transmission compared to younger counterparts (15-24 years) [91,95]. We suspected that this may be a result of outbreaks associated with congregate settings but we did not find these cases to be associated with long-term care or correctional facilities based on available surveillance data. It is also unlikely that these cases represent coincidental reactivation of a strain that was prevalent in the past among two or more cases as older NH Whites were represented among multiple different clusters in our study [91]. A prior study argued that matching genotypes in a rural setting where there is limited strain diversity do not always represent ongoing transmission[48,100]. While our study uses a time restricted cluster definition, which likely overcomes the limitation of this prior study, we believe further epidemiological investigations need to be conducted to elucidate the increased risk of recent transmission among older NH White persons. This finding suggests that traditionally recognized risk factors have evolved over time in low incidence settings, calling for

refocusing TB transmission control efforts to include newly identified groups at an increased risk of recent transmission.

Gender disparity was another theme observed from our findings; men were at increased risk of both reactivation and recent transmission compared to women. Previous studies have overwhelmingly reported higher risk for TB among men for a long time across diverse settings[46,95]. Researchers have attempted to explain these differences based on biological and non-biological factors such as genetic factors, sex hormones, health-seeking behavior, smoking pattern and workplace exposure[101,102]. Based on our stratified results, we observed that the gender disparity played out differently across racial groups. NH Black and NH White men were at a particularly increased risk for both recent transmission and reactivation, however, for other racial groups gender disparity was not observed. As studies continue to resolve the long-standing paradox of gender disparity in TB, they should also examine how gender and race interact together with TB risk using larger study samples that are powered to test these associations.

The integration of genotyping tools with surveillance data is particularly essential for understanding the complex transmission dynamics of TB due to its long latency period. While there is no clear gold standard for distinguishing recent transmission from reactivation, previous studies have reported greater discriminatory power, similar to that of IS*6110*-RFLP, for combined use of 24 loci MIRU-VNTR with spoligotyping over 12 loci MIRU-VNTR typing alone or when combined with spoligotyping [103,104]. Additionally, the 1-year time restricted clustering definition provides greater confidence in identifying recent transmission as previously described[3,61]. For a subset of study patients (n=200) diagnosed between 2018-2020, whole genome multilocus sequence typing (wgMLST) was available in the TB GIMS dataset[105]. Based

on wgMLST analysis, 46.3% (88/200) of cases were clustered compared to 40.5% (77/200) based on the time restricted definition used in our study.

While similar results were obtained from wgMLST analysis, it is important to recognize that the 24 loci MIRU-VNTR typing has lower molecular resolution compared to advanced methods such as WGS, which was not universally rolled out until 2022[106,107]. The 1-year cut-off can potentially misclassify patients as non-clustered especially those diagnosed at the very beginning or end of the study period. An inherent limitation of molecular TB studies is the underrepresentation of TB risk among young children due to diagnostic challenges around specimen collection resulting in possible underestimation of recent transmission[21]. Another intrinsic limitation of this and all other molecular epidemiological studies of TB is the exclusion of non-culture confirmed clinical cases given that *Mtb* genotyping data can only be obtained for culture-confirmed cases.

To summarize, this study highlights the need for an equity-based approach for TB elimination with a major focus on bridging existing racial disparities in Arkansas and similarly low incidence states in the U.S. Our work provides a useful framework that incorporates molecular and social dimensions of TB disease through an analysis disaggregated by age, sex and race for both clustered and non-clustered TB. Apart from evaluating the contribution of recent versus reactive disease, this study calls for refocusing TB interventions based on updated and race/ethnicity specific risk factors as traditionally recognized risk factors may have evolved over time and may not ubiquitously apply to all racial groups. A mechanistic understanding of what drives TB disparities at the local level provides an opportunity to adapt TB interventions to the evolving molecular TB epidemiology in Arkansas and its subpopulations.

3.6 Tables and Figures

Table 3.6.1. Comparison of culture confirmed cases diagnosed in Arkansas during 2010-2021 with and without valid genotyping data

		Cases with valid genotyping data (n=627)	Cases with incomplete genotyping data (n=17)	[a]P-value
Race				
	Hispanic	77 (12.3%)	1 (5.9%)	0.71
	American Indian /Alaska Native	4 (0.6%)	0 (0%)	
	Asian	69 (11%)	2 (11.8%)	
	Non-Hispanic Black	178 (28.4%)	3 (17.7%)	
	[c]NH/PI	68 (10.9%)	3 (17.7%)	
	Non-Hispanic White	231 (36.8%)	8 (47.1%)	
Gender				0.32
	Male	403 (64.3%)	14 (82.4%)	
	Female	218 (34.8%)	3 (17.7%)	
	Missing	6 (1%)	0 (0%)	
Age Group				0.92
	0-14 years	12 (1.9%)	0 (0%)	
	15-24 years	63 (10.1%)	1 (5.9%)	
	25-44 years	171 (27.3%)	6 (35.3%)	
	45-64 years	208 (33.2%)	5 (29.4%)	
	65 years and over	173 (27.6%)	5 (29.4%)	
U.S. Born				
	Yes	426 (67.9%)	11 (64.7%)	0.80
	No	201 (32.1%)	6 (35.3%)	
Clinical Presentation				0.05
	Pulmonary TB	506 (80.7%)	11 (64.7%)	
	Extra Pulmonary TB	84 (13.4%)	6 (35.3%)	
	[b]Both	37 (5.9%)	0 (0%)	

[a] P-value from Fisher's Exact test

[b] Diagnosed as both pulmonary TB and extra pulmonary TB

[c] Native Hawaiian/ Pacific Islander

* Due to rounding, percentages may not add to 100

Table 3.6.2. Incidence and risk ratio (RR) with 95% CIs for i) clustered TB and ii) non-clustered TB by race/ethnicity, age group and sex

	Clustered TB			Non-clustered TB		
	No	Incidence	Risk ratio (95% CI)	No	Incidence	Risk ratio (95% CI)
Overall	218	0.61 (0.53, 0.70)		389	1.09 (0.98, 1.20)	
Race/ethnicity						
NH Whites	59	0.23 (0.18, 0.29)	Ref	157	0.60 (0.51, 0.70)	Ref
NH Blacks	92	1.67 (1.36, 2.04)	7.37 (5.32, 10.22)	84	1.52 (1.23, 1.88)	2.53 (1.94, 3.30)
Asians	2	0.37 (0.09, 1.47)	1.62 (0.40, 6.65)	67	12.30 (9.68, 15.62)	20.45 (15.36, 27.22)
NH/PI	52	49.42 (37.66, 64.85)	218.65 (150.60, 317.43)	16	15.20 (9.32, 24.82)	25.28 (15.12, 42.29)
Hispanic	10	0.38 (0.21, 0.71)	1.70 (0.87, 3.31)	64	2.45 (1.92, 3.13)	4.08 (3.05, 5.45)
Age Group						
15-24 years	29	0.60 (0.41, 0.86)	Ref	33	0.68 (0.48, 0.95)	Ref
0-14 years	7	0.10 (0.05, 0.21)	0.17 (0.07, 0.38)	4	0.06 (0.02, 0.15)	0.08 (0.03, 0.24)
25-44 years	74	0.82 (0.65, 1.02)	1.37 (0.89, 2.10)	95	1.05 (0.86, 1.28)	1.54 (1.04, 2.30)
45-64 years	81	0.90 (0.72, 1.11)	1.50 (0.98, 2.30)	124	1.37 (1.15, 1.63)	2.02 (1.38, 2.97)
65 and over	27	0.47 (0.32, 0.68)	0.79 (0.47, 1.33)	133	2.31 (1.95, 2.74)	3.41 (2.33, 4.99)
Sex						
Female	76	0.42 (0.33, 0.52)	Ref	140	0.77 (0.65, 0.91)	Ref
Male	142	0.81 (0.68, 0.95)	1.93 (1.46, 2.55)	249	1.41 (1.25, 1.60)	1.84 (1.49, 2.26)

Table 3.6.3. Race/ethnicity stratified risk ratio (RR) with 95% CIs for clustered TB by age group and sex

		No	NH White	No	NH Black	No	Asian	No	NHPI	No	Hispanic
Sex											
	Female	16	Ref	31	Ref	0	Ref	26	Ref	2	Ref
	Male	43	2.78 (1.56, 4.93)	61	2.17 (1.41, 3.34)	2	N/A	26	0.91 (0.53, 1.57)	8	3.61 (0.77, 16.99)
Age group											
	15-24 years	2	Ref	14	Ref	0	Ref (type 3 p=0.663 3)	11	Ref	2	Ref
	0-14 years	0	N/A	4	0.20 (0.07, 0.62)	0		3	0.14 (0.04, 0.49)	0	N/A
	25-44 years	9	2.29 (0.49, 10.58)	33	1.45 (0.77, 2.70)	1		24	1.03 (0.50, 2.09)	5	1.47 (0.29, 7.59)
	45-64 years	28	6.41 (1.53, 26.90)	35	1.76 (0.95, 3.27)	1		13	1.72 (0.77, 3.85)	3	1.76 (0.29, 10.56)
	≥65 years	20	6.50 (1.52, 27.81)	6	0.69 (0.27, 1.81)	0		1	0.64 (0.08, 4.98)	0	N/A

Table 3.6.4. Race/ethnicity stratified risk ratio (RR) with 95% CIs for non-clustered TB by age group and sex

		No	NH White	No	NH Black	No	Asian	No	NH/PI	No	Hispanic
Sex											
	Female	48	Ref	34	Ref	28	Ref	7	Ref	23	Ref
	Male	109	2.35 (1.67, 3.30)	50	1.62 (1.05, 2.50)	39	1.50 (0.92, 2.44)	9	1.17 (0.44, 3.15)	41	1.61 (0.96, 2.68)
Age group											
	15-24 years	3	Ref	10	Ref	11	Ref	3	Ref	6	Ref
	0-14 years	0	N/A	1	0.07 (0.01, 0.56)	0	N/A	1	0.17 (0.02, 1.60)	2	0.18 (0.04, 0.91)
	25-44 years	16	2.71 (0.79, 9.30)	20	1.23 (0.57, 2.62)	26	0.98 (0.48, 1.98)	5	0.78 (0.19, 3.28)	28	2.75 (1.14, 6.64)
	45-64 years	47	7.17 (2.23, 23.04)	28	1.97 (0.96, 4.06)	24	1.56 (0.76, 3.18)	6	2.92 (0.73, 11.67)	18	3.53 (1.40, 8.89)
	≥65 years	91	19.72 (6.24, 62.28)	25	4.05 (1.94, 8.43)	6	1.08 (0.40, 2.92)	1	2.36 (0.25, 22.69)	10	8.66 (3.15, 23.82)

Figure 3.6.1. Proportions of clustered and non-clustered tuberculosis (TB) cases among all TB cases with a valid genotyping result diagnosed in Arkansas during 2010-2021. Solid black bar: clustered; dotted bar: non-clustered; and solid white bar: Bovis/Bovis-BCG. The numbers indicated within the bars are the yearly case numbers for each category.

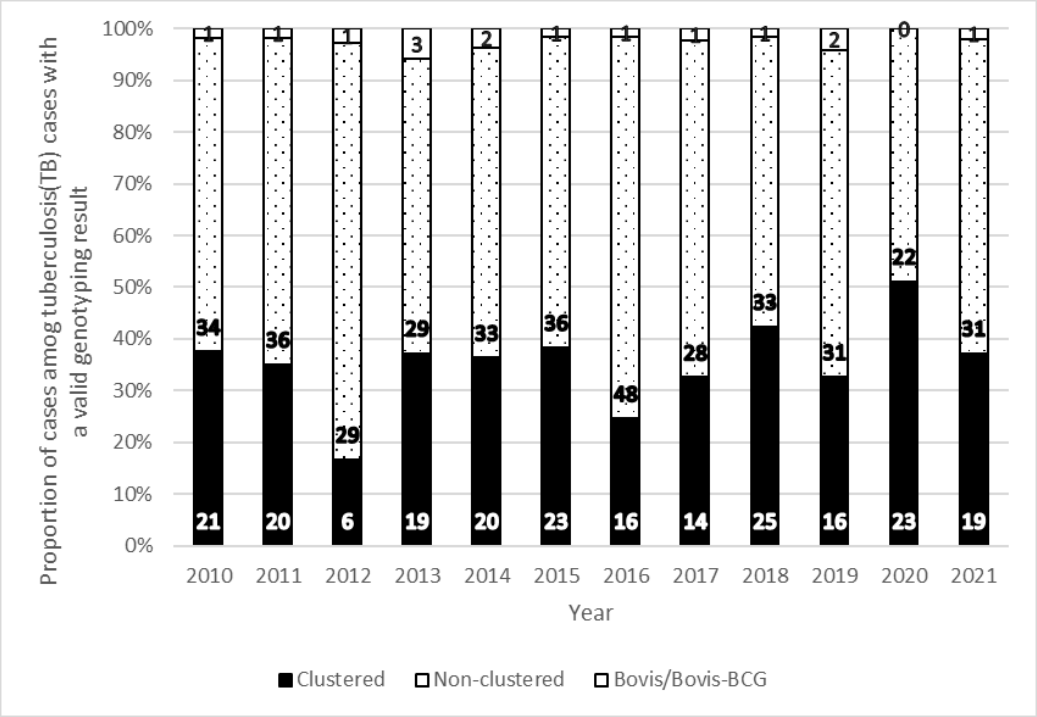


Table 3.6.5. Probability of clustered TB (95% CI) calculated using logistic regression

	All Races	NH White	NH Black	Asian	NH/PI	Hispanic
All age groups	0.36 (0.33, 0.40)	0.28 (0.23, 0.34)	0.53 (0.45, 0.60)	0.03 (0.01, 0.11)	0.76 (0.65, 0.85)	0.14 (0.07, 0.23)
Age group						
0-14 years	0.67 (0.38, 0.87)	[#]N/A	0.83 (0.37, 0.98)	[#]N/A	0.75 (0.24, 0.97)	[*]N/A
15-24 years	0.47 (0.35, 0.59)	0.40 (0.10, 0.80)	0.58 (0.38, 0.76)	[*]N/A	0.79 (0.51, 0.93)	0.25 (0.06, 0.62)
25-44 years	0.44 (0.36, 0.51)	0.35 (0.19, 0.54)	0.62 (0.49, 0.74)	0.04 (0.01, 0.22)	0.83 (0.65, 0.93)	0.15 (0.06, 0.32)
45-64 years	0.40 (0.34, 0.47)	0.39 (0.29, 0.50)	0.56 (0.43, 0.67)	0.04 (0.01, 0.24)	0.68 (0.45, 0.85)	0.14 (0.05, 0.36)
≥65 years	0.17 (0.12, 0.24)	0.19 (0.13, 0.27)	0.19 (0.09, 0.37)	[*]N/A	0.50 (0.06, 0.94)	[*]N/A

[#]No observations
[*]All cases in this cell were non-clustered

Chapter 4 Effect of Gender on Clinical Presentation of Tuberculosis (TB) and Age-Specific Risk of TB, and TB-Human Immunodeficiency Virus Coinfection

This chapter is published in *Open Forum Infectious Diseases* (doi: https://doi.org/10.1093/ofid/ofac512). Authors: Maheen Humayun, Joconiah Chirenda, Wen Ye, Innocent Mukerezdi, Hilda Angela Mujuru, Zhenhua Yang

4.1 Abstract

Background

Previous studies have shown gender differences in tuberculosis (TB) incidence; however, gender disparity has not been well documented across granular categorizations of anatomic sites affected by TB and in the presence of human immunodeficiency virus (HIV) coinfection, largely due to small sample size for less common TB clinical presentations and lack of detailed clinical data.

Methods

The study population included TB cases aged ≥15 years (n = 41, 266) diagnosed in Harare, Zimbabwe. This cross-sectional study estimated male-to-female ratio (M/F ratio) for (1) age-specific TB incidence, (2) age-specific HIV prevalence among incident TB cases, and (3) 9 types of TB defined by affected anatomic site.

Results

Males were at a 53% higher risk of TB compared to females (risk ratio [RR] = 1.53; 95% confidence interval [CI], 1.12–2.09). Based on adjusted odds ratios (aORs) from multinomial logistic regression model, the odds of abdominal TB (aOR = 0.51; 95% CI, 0.39–0.68), TB

bones/joints/spine (aOR = 0.63; 95% CI, .45–.90), and “other” extrapulmonary TB sites (aOR = 0.69; 95% CI = 0.59–0.81) versus pulmonary TB were lower among males compared to females. The risk of TB-HIV coinfection among males was 17% (RR = 0.83; 95% CI, 0.74–0.93) and 8% (RR = 0.92; 95% CI, 0.88–0.95) lower in the 15- to 24-year and 25- to 44-year age groups, respectively.

Conclusions

This study revealed a nuanced role of gender across finer categorizations of TB, indicating the need for future research to delineate underlying mechanisms driving gender disparities in TB. The finding that women had a greater likelihood of severe forms of TB and TB-HIV coinfection compared to men has important implications for women's health in TB-HIV high-burden settings.

4.2 Introduction

Tuberculosis (TB) has remained a major cause of global mortality and is the single leading infectious killer only second to coronavirus disease 2019 [28,51]. Although the TB epidemic affects all regions of the world, its burden is disproportionately shared by 30 countries that accounted for approximately 86% of the total estimated incident TB cases in 2020 [28]. In addition to global disparities across countries, the risk posed by TB varies across subpopulations defined by various sociodemographic characteristics including gender, age, race/ethnicity, immigration status, income level, and occupation in a given country [41,108]. Existing disparities serve as an impediment to the World Health Organization's (WHO) End TB Strategy goals and the progress towards global TB elimination. Without targeted interventions for high-risk groups, existing TB disparities will likely be magnified, and hence identification of high-risk groups is the first step towards a proequity policy [109].

Biological sex is an important determinant of health because sex differences in genetic, hormonal, and epigenetic regulation alter prevalence, manifestation, and treatment of disease [110]. However, biology alone does not explain the globally observed gender disparities because these differences also arise from the social construction of gender [110,111]. Prescribed gender roles, expectations, and norms may result in differences in access to material resources (prestige, power, nutrition), health-related behaviors (e.g., smoking and drinking), and exposure to psychosocial stressors (e.g., discrimination and violence) [110–113]. Current literature on gender epidemiology of TB demonstrates gender differences in prevalence/notification rate, disease manifestation, progression to disease, case-fatality rate, response to treatment, and along the continuum of TB pathogenesis and care [43,102,114,115].

Health inequalities defined across different strata of a single factor such as gender discount the actual context of people's lives where multiple identities/processes and their intersectionality together produce population health [116]. From previous studies, we know that gender interacts with sociodemographic factors such as age to modify the risk of TB [114,117]. These effects may be further modified by the presence of other risk factors such as human immunodeficiency virus (HIV), which is associated with an increased risk of new *Mycobacterium tuberculosis* infection, reactivation of latent TB infection, and reinfection [2]. The risk of TB infection is approximately 20 times higher among HIV-infected individuals, and the case-fatality rate is 2 times higher compared to HIV-negative individuals, worsening the progress towards TB elimination in high-burden settings [2,27].

Although many studies have already called for recognizing and targeting men as a high-risk group for TB interventions, the role of gender can be more nuanced across different contexts

and populations. In this Zimbabwe-based study, we resolved gender differences in TB incidence over time, by age, by TB clinical presentation defined by affected anatomic site and in the presence of HIV coinfection to better understand the effect of gender on TB in different sociodemographic and clinical contexts. As reported below, our multidimensional and intersectional approach to understanding gender disparity and mapping risk groups provides a more precise and pragmatic characterization of gender disparity, which can be useful for designing equity-driven policies, improved diagnosis, and targeted interventions.

4.3 Methods

4.3.1 Study Setting and Population

The setting of the current study, Zimbabwe, has been placed on the global TB watch list by the WHO and is among the top 30 high-burden countries for TB-HIV coinfection, requiring continued surveillance of the TB epidemic in the country [28]. The study sample included adult TB cases, 15 years and over, who were diagnosed (n = 41 266) in Harare, the capital city of Zimbabwe, during 2008 to 2017. The rate of sputum smear-positive cases among TB cases diagnosed in Harare in 2011 was 106.9 per 100 000, which was higher than the national average of 86.7 per 100 000 at the same time [118]. The cases diagnosed in Harare during that time included TB cases from both Harare city and its neighboring city, Chitungwiza.

4.3.2 Study Design and Data Collection

We conducted a time trend ecological analysis of (1) TB incidence during 2008–2017 and (2) HIV prevalence among incident TB cases during 2011–2017. A cross-sectional analysis was also conducted to examine gender disparity in TB clinical presentation defined by affected anatomic site.

Based on data availability, we retrospectively accessed 2 types of TB surveillance data for Harare city: (1) aggregated TB surveillance data for the period of 2008–2010 (n = 16 989) and (2) individual-level data on TB patients for the period 2011–2017 from the Harare City Health Department (n = 24 277). Both aggregated and individual TB surveillance data included information on sociodemographics and TB treatment history. Individual-level data additionally included information on HIV serostatus and other clinical variables. For population size estimates, population projections calculated by the Zimbabwe National Statistics Agency (ZimStat) based on 2012 population census data were used [119].

4.3.3 Data Analysis

The study estimated (1) overall and age stratified male to female ratio (M/F ratio) for TB incidence and HIV prevalence among incident TB cases, (2) trend of TB incidence and HIV prevalence among incident TB cases, and (3) M/F ratio for various TB clinical presentations.

4.3.3.1 Male to Female Ratio Calculation

We used negative binomial or Poisson regression models with gender as the explanatory variable to calculate the male to female risk ratio. We calculated overall and age-stratified M/F ratio for TB incidence using data between 2008 and 2017 for TB and M/F ratio for HIV prevalence among incident TB cases using data between 2011 and 2017. Age was categorized into 4 groups: 15–24 years, 25–44 years, 45–64 years, and ≥65 years. We used a goodness-of-fit test to choose between negative binomial and Poisson models. For some cases, we used Poisson regression because the negative binomial model failed to converge.

4.3.3.2 Trend of Tuberculosis (TB) Incidence and Human Immunodeficiency Virus Prevalence among Incident TB Cases

First, we examined the overall time trend for (1) TB incidence during 2008–2017 and (2) HIV prevalence among incident TB cases during 2011–2017 for Harare city. The TB incidence was calculated as incident cases of TB diagnosed in a given year per 100 000 population. Medium population projections based on 2012 Census data were used because they are considered to be most probable [119]. Because projections were only available for 2012 and onwards, we used the 2012 population estimate for years between 2008 and 2011. The HIV prevalence among incident TB cases was calculated as the percentage of HIV-positive TB patients among incident TB cases who had a documented HIV test result for a given year. The HIV data were not available for years between 2008 and 2010. To calculate annual percentage change (APC), we used negative binomial and Poisson regression models with year of diagnosis included in a linear form. In instances in which the negative binomial model failed to converge we used Poisson regression.

Second, we stratified overall time trends by gender to examine gender disparity in TB incidence and HIV prevalence among incident TB cases over time. Third, we studied gender-specific trends for each age group separately to observe the effect of age on observed gender disparity. To examine gender disparity, gender, year of diagnosis, and an interaction term between gender and year of diagnosis were included in the overall and age-stratified models.

4.3.3.3 Male to Female Ratio for Tuberculosis Clinical Presentations

To examine gender disparity for various TB clinical presentations, we used a multinomial logistic regression. The outcome variable, TB clinical presentation based on affected anatomic site, had 9 categories and was regressed on gender to obtain unadjusted male to female odds ratio (OR) (**Table 4.6.1**) Pulmonary tuberculosis (PTB) was used as the reference category for the outcome variable. For the adjusted model, year of diagnosis, healthcare setting, age group, treatment history, and HIV status were added as categorical variables to the model. Tuberculosis treatment history was dichotomized by collapsing cases with a treatment history of relapse, failed treatment, or loss to follow up as previously treated versus new cases with no treatment history.

4.3.4 Patient Consent Statement

This study used de-identified patient data and was approved by the University of Michigan's Health Sciences and Behavioral Sciences Institutional Review Board (IRB-HSBS). This study was determined unregulated by University of Michigan IRB-HSBS.

4.3.5 Data Availability

The data used in this article were accessed through paper records of Harare City Health Department for use at the University of Michigan. Upon reasonable request, these data can be shared with prior approval from the Harare City Health Department and relevant data providers.

4.4 Results

4.4.1 Characteristics of the Study Population

As described under Methods, we used data available from 2 different sources—aggregated and individual-level data—to conduct this study. Individual-level data were available for 24 277 TB cases diagnosed in Harare between 2011 and 2017 (**Table 4.6.1**) A majority of TB cases were male (58.5%) and between the productive age of 25 and 44 years (64.5%). New cases with no prior treatment history for TB accounted for approximately 91% of the study population. Among those with a prior treatment history for TB, 64.8% were men and 61.3% were in the 25- to 44-year age category. Although data on age, gender, treatment history, and diagnosis center were complete, approximately 7% of the study population did not have a documented HIV test result. The HIV seropositivity rate among those who had a documented test result for HIV was as high as 73.4%. Most of the TB patients presented with PTB (82.3%).

4.4.2 Overall and Age-Specific Gender Disparity in (1) Tuberculosis (TB) Incidence and (2) Human Immunodeficiency Virus Prevalence among Incident TB Cases

Overall, the risk of TB was 53% higher among males compared to females (risk ratio [RR] = 1.53; 95% confidence interval [CI], 1.12–2.09) (Table 4.6.2). The risk in younger age groups did not differ significantly, but the point estimate of M/F ratio consistently increased with age and the ≥65-year age group had the highest risk among males (RR = 1.92; 95% CI, 1.23–3.00). The overall risk of HIV prevalence among incident TB cases did not differ between males and females (RR = 0.95; 95% CI, .81–1.12). When stratified by age, males experienced 17% and 8% lower risk of HIV prevalence in 15- to 24-year and 25- to 44-year age groups, respectively (RR = 0.83, 95% CI = .74–.93 and RR = 0.92, 95% CI = .88–.95). With increase in age, the gender disparity in HIV prevalence reduced, putting males and females at comparable risk.

4.4.3 Trend of Age-Specific Gender Disparity in Tuberculosis Incidence Over Time

The TB incidence in Harare declined from 364 cases per 100 000 population in 2008 to 121 cases per 100 000 population in 2017 with an annual percentage decline of 14% (95% CI, −18.6% to −10.1%) (**Figure 4.6.1**) and a similar rate of decline (*P* value for interaction = .9) among males and females. In addition, TB incidence declined for both males and females across all age groups; however, females experienced a higher rate of decline for age groups between 15 and 64 years, although the difference was not statistically significant (*P* value for interaction > .05) (**Figure 4.6.2**).

4.4.4 Trend of Age-Specific Gender Disparity in Human Immunodeficiency Virus Prevalence among Incident Tuberculosis Cases over Time

Overall, the trend of HIV prevalence stayed stagnant among incident TB cases in Harare. It declined from 73% to 69% between 2011 and 2017 with an annual percentage decline of 0.5% (95% CI, −4.5% to 3.74%). The HIV prevalence increased by 0.08% per year among males (95% CI, −5.53 to 6.04) and declined by 0.95% per year among females (95% CI, −6.52 to 4.97) (**Figure 4.6.3**). In addition, HIV prevalence did not change significantly for either males or females across all age groups (*P* value for interaction > .05) (**Figure 4.6.4**).

4.4.5 Gender Disparity in Tuberculosis Clinical Presentation

The odds of presenting with TB at any of the 3 disease sites (abdominal TB, TB bones/joints/spine, and other extrapulmonary tuberculosis [EPTB]) compared to the reference category, PTB, were higher among females than males in the unadjusted model (**Table 4.6.3**) as shown by corresponding M/F ORs < 1. There was no significant gender difference for other TB subtypes.

The adjusted model yielded similar results. The odds of abdominal TB versus PTB were lower among males compared to females after adjusting for year of diagnosis, age group, HIV serological status, treatment history, and healthcare setting (aOR = 0.51; 95% CI, .39–.68). The adjusted odds of TB bones/joints/spine (aOR = 0.63; 95% CI, .45–.90) and other EPTB (aOR = 0.69 95% CI = .59–.81) were lower among males compared to females.

4.5 Discussion

A gap-focused approach to address health inequity not only reduces the absolute level of ill health but also aims to increase the rate of improvement among the disadvantaged group relative to the more advantaged group [112]. Given the communicable nature of TB, equitable improvements for all groups are key to reducing the overall burden of disease in a population. Without gender-stratified analysis, studies often leave out important information that can potentially accelerate our understanding of sex-specific pathogenesis, treatment, and prevention [120]. In this report, we characterize various dimensions of gender disparities related to the TB epidemic in a TB-HIV high-burden setting, Harare in Zimbabwe, during 2008 to 2017.

Our study found that males were at an increased risk of TB compared to females and the risk among males increased with age. Females were more likely to have EPTB at various sites including abdominal TB, TB bones/joints/spine, and other EPTB. The risk of TB-HIV coinfection was higher among females than among males in age groups: 15–24 years and 25–44 years. Although statistically insignificant, males and females experienced unequal rates of change in TB incidence and TB-HIV prevalence during the study period. Although men are widely known as a high-risk group for TB, our study suggests that it is important for TB prevention/control efforts to recognize the high burden of EPTB and TB-HIV coinfection among females.

The higher risk of TB among men in this study is corroborated by current literature that shows that generally men are twice as likely to be diagnosed with TB than women, which is possibly explained by behavioral and biological hypotheses [43,102,121]. Men tend to have a higher number of social contacts, work in high-risk occupations such as mining, and engage in high-risk behaviors such as smoking [121]. The biological hypothesis argues for increased genetic susceptibility to TB among men and immune responses modulated by sex hormones as possible explanations for observed gender differences [121].

There are only a few studies from high-burden or resource-limited settings that study gender differences in TB clinical presentation defined by granular categorizations of affected anatomic sites, often lacking sufficient sample size and relevant statistical analysis [44,122,123]. Many studies report M/F ratio among a sample of bacteriologically confirmed or smear-positive TB patients, which is likely to overrepresent PTB patients because it is the most prevalent and relatively easy to diagnose compared to EPTB [43]. We report M/F ratio specific to 9 different anatomical TB sites that showed that females were at a significantly higher risk of presenting with TB in bones/joints/spine, abdomen, and in sites captured by the "other EPTB" category compared to their risk of PTB relative to men. A recent study from Ghana, also a TB-HIV high-burden country, did not find significant differences between proportion of adult males and females affected by TB in spine, bones/joint, or abdomen, whereas a Pakistan-based study found all forms of EPTB, including abdomen, lymphatic, and osteoarticular spine TB, to be more common among females [124,125]. This study provides a reasonably large sample size to study gender differences across multiple sites of EPTB, but we remain cautious regarding misclassification bias resulting from under- or misdiagnosis of EPTB in this resource-deficient setting.

Although men are at an increased risk of TB, women in 15- to 24-year and 25- to 44-year age groups had a greater likelihood of presenting with TB-HIV coinfection. A study from sub-Saharan Africa found women in the 25- to 29-year age group to be twice as likely to be coinfected and 5 times more likely to be HIV infected compared to men [126]. Another previous study has also found that low $CD4^+$ cell count (<100) among EPTB-HIV-coinfected patients is associated with "severe" disease sites such as central nervous system (CNS), meningeal, or disseminated TB [127]. The higher likelihood of various forms of EPTB among females in our study may be attributable to higher HIV prevalence and low $CD4^+$ cell count. Based on our findings, targeted and earlier access to antiretroviral therapy for young women in Harare could be crucial for controlling the TB epidemic. Although the exact cause of increased likelihood of EPTB among women remains elusive, previous studies suggest that endocrinal factors in relation to menopause may explain the increased susceptibility to EPTB among women [45,128,129]. Apart from biological explanations, the increased likelihood of EPTB among women may be a result of under treatment or under diagnosis owing to structural and social factors that impede women's access to TB care, especially in male-dominated societies [130,131].

Based on traditional gender roles, women are generally primary caregivers and are involved in unpaid essential economic labor often involving subsistence farming in Africa [132]. The prevalence of TB-HIV coinfection among women of childbearing age poses risk of antenatal transmission and increase in child mortality trends. The disparate burden of TB-HIV coinfection among women is also a grim reflection of the prevalence of cultural, socioeconomic, behavioral, and structural risk factors that increase women's vulnerability to HIV in Africa [133]. In addition to gender inequality that limits women's access to resources such as healthcare, education, and wealth, patriarchal norms result in heightened social stigma and violent sexual practices against

women that are associated with increased risk of HIV. The effect of TB-HIV coinfection burden among productive age women extends beyond their own health and affects families, communities, and the broader development of the country.

In addition to providing important information on how disease patterns evolve in a population, epidemiological trends over time can be useful in evaluating the performance of existing disease control/prevention programs. The sudden drop in TB notification in 2014 likely suggests missed cases or data reporting error that can have serious implications for public health. After removing the 2014 data point in our sensitivity analysis, the APC for TB incidence did not significantly change (APC = −13.2%; 95% CI, −9.0% to 17.2%). This high-burden population provides a unique opportunity to investigate stratified associations due to the availability of a large sample. We are limited in our ability to make inferences related to TB and HIV temporality due to the cross-sectional study design. In addition, notification data from passive surveillance does not always truly represent disease incidence and is impacted by sample selection bias.

Conclusions

Men are well recognized as a high-risk group for TB, but disaggregated gender differences reveal that women are a high-risk group for various forms of EPTB and TB-HIV coinfection. Public health strategies for TB control/prevention must recognize that women are more likely to develop severe forms of TB and TB-HIV coinfection, which may result in worse outcomes for women. Future studies should attempt to delineate the underlying drivers of these observed gender differences. Although these gender differences may be partly explained by differences in biological susceptibility, they also reflect the role of social, structural, and historical factors that drive gender inequality in Africa.

4.6 Tables and Figures

Table 4.6.1. Characteristics of TB cases (n=24 277) diagnosed in Harare, Zimbabwe during 2011-2017

Patient characteristic	No (%[a])
Gender	
Male	14,206 (58.5)
Female	10,071 (41.5)
Age group	
15-24 years	2,797 (11.5)
25-44 years	15,659 (64.5)
45-64 years	4,772 (19.7)
≥ 65 years	1,049 (4.3)
Treatment history	
New case	21,984 (90.6)
Previously treated	2,293 (9.5)
Healthcare setting[b]	
BRIDH	15,160 (62.5)
WIDH	6,781 (27.9)
Other clinics	2,336 (9.6)
HIV serological status	
Positive	16,545 (68.2)
Negative	5,994 (24.7)
Unknown	1738 (7.2)
TB disease site	
PTB	19,968 (82.3)
Abdominal TB[c]	210 (0.9)
Lymphadenitis	453 (1.9)
Miliary TB	507 (2.1)
Other EPTB	671 (2.8)
Pericarditis	216 (0.9)

Pleural Effusion	1774 (7.3)
TB Bones/Joints/Spine	129 (0.5)
TB Meningitis	349 (1.4)

[a]Due to rounding, % may not add up to100
[b]BRIDH = Beatrice Road Infectious Disease Hospital; WIDH = Wilkins Infectious Diseases Hospital; Other clinics = other clinics throughout Harare
[c] Includes peritonitis and adenitis TB

Table 4.6.2. Age-stratified male to female ratio (M:F Risk Ratio) for TB incidence during 2008-2017 and prevalence of HIV among incident TB cases during 2011-2017 in Harare, Zimbabwe calculated using Poisson and negative binomial regression

Age Group	TB incidence M/F ratio (95% CI)	HIV prevalence M/F ratio (95% CI)
Overall	1.53 (1.12, 2.09)	0.95 (0.81,1.12)
15-24 years	1.14 (0.79, 1.65)	0.83 (0.74, 0.93)
25-44 years	1.25 (0.78, 2.01)	0.92 (0.88, 0.95)
45-64 years	1.61 (0.98, 2.62)	1.02 (0.94, 1.10)
≥65 years	1.92 (1.23, 3.00)	1.10 (0.86, 1.40)

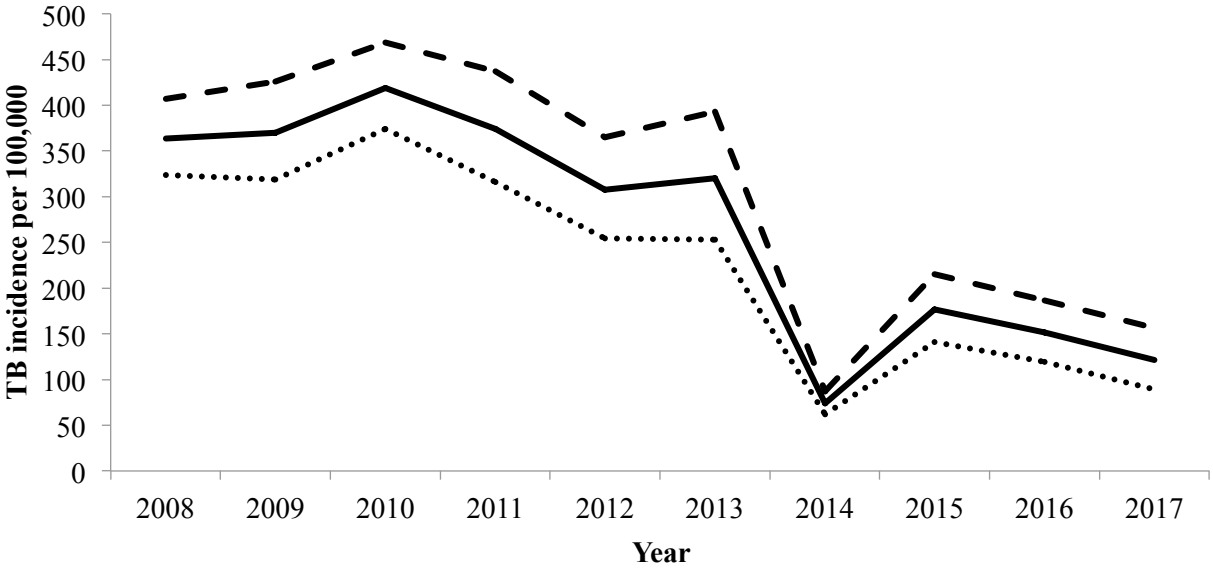
Abbreviations: CI, confidence interval; HIV, human immunodeficiency virus; TB, tuberculosis

Table 4.6.3. Gender-specific distribution of TB clinical presentation defined by affected anatomic site and male to female odds ratio (M: F risk ratio) for each anatomic site based on the analysis of 24 277 TB patients diagnosed in Harare, Zimbabwe during 2011-2017 using multinomial logistic regression

	Female (%)	Male (%)	Unadjusted M:F Odds Ratio (95% CI)	[a]Adjusted M:F Odds Ratio (95% CI)
PTB	8,189 (41.0)	11,779 (59.0)	Ref	Ref
Abdominal TB	121 (57.6)	89 (42.4)	0.51 (0.39, 0.67)	0.51 (0.39, 0.68)
Lymphadenitis	202 (44.6)	251 (55.4)	0.86 (0.72, 1.04)	0.88 (0.73, 1.07)
Miliary TB	212 (41.8)	295 (58.2)	0.97 (0.81, 1.16)	0.95 (0.79, 1.14)
Other EPTB	340 (50.7)	331 (49.3)	0.68 (0.58, 0.79)	0.69 (0.59, 0.81)
Pericarditis	98 (45.4)	118 (54.6)	0.84 (0.64, 1.10)	0.84 (0.64, 1.10)
Pleural Effusion	694 (39.1)	1080 (60.9)	1.08 (0.98, 1.20)	1.06 (0.96, 1.17)
TB Bones/Joints/Spine	65 (50.4)	64 (49.6)	0.69 (0.48, 0.97)	0.63 (0.45, 0.90)
TB Meningitis	150 (43.0)	199 (57.0)	0.92 (0.75, 1.14)	1.00 (0.80, 1.24)

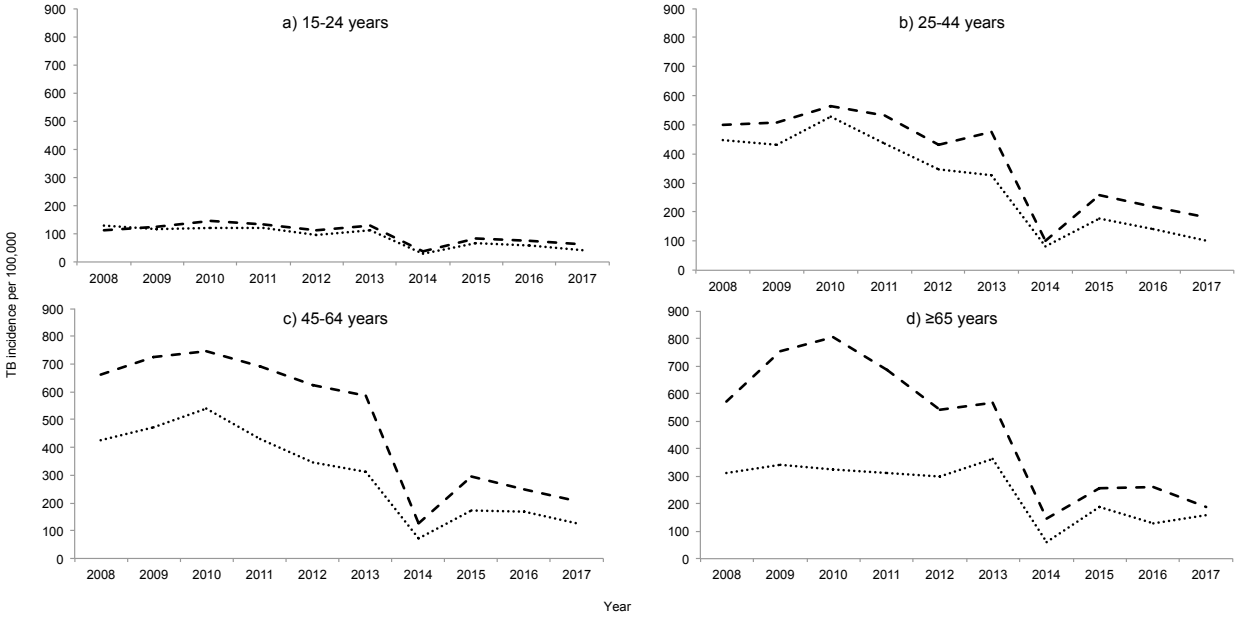
[a] Adjusted for year of diagnosis, HIV status, healthcare setting, age group and treatment history

Figure 4.6.1. Trend of TB incidence in Harare, Zimbabwe during 2008-2017. Solid line: the overall TB incidence; dashed line: TB incidence among males; dotted line: TB incidence among females. The table below shows the overall and gender-stratified annual percentage change (APC) for TB incidence with p-value for testing difference in the rate of decline between males and females



	Overall	Female	Male	P-value
APC (95% CI)	-14.43 (-18.55, -10.09)	-14.70 (-20.13, -8.90)	-14.26 (-19.75, -8.38)	0.91

Figure 4.6.2. Age-stratified gender disparity in TB incidence in Harare, Zimbabwe during 2008-2017. Panel a: among 15-24 years; Panel b: among 25-44 years; Panel c: among 45-64 years; Panel d: among >65 years. The dashed line and dotted lines show the trend of TB incidence among males and females, respectively. The table below shows the gender-stratified annual percentage change (APC) of TB incidence for each age group and p-values for testing difference in the rate of decline between females and males.

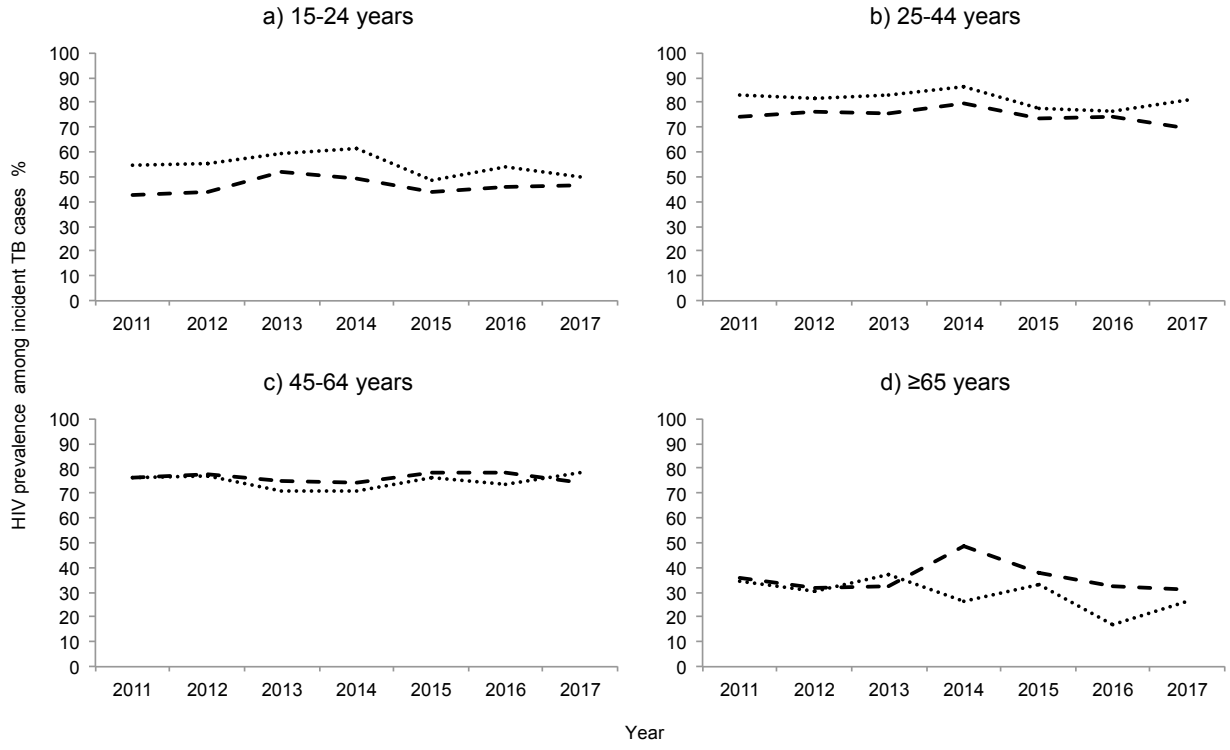


	15-24 years	25-44 years	45-64 years	≥65 years
Female APC (95% CI)	-12.29 (-17.76, -6.45)	-17.36 (-23.17, -11.11)	-16.07 (-22.03, -9.66)	-10.64 (-17.11, -3.67)
Male APC (95% CI)	-8.55 (-14.26, -2.46)	-13.00 (-19.04, -6.53)	-15.19 (-21.29, -8.62)	-15.49 (-21.59, -8.91)
P-value	0.37	0.33	0.85	0.30

Figure 4.6.3. Trends of HIV prevalence among incident TB cases in Harare, Zimbabwe during 2011-2017. Solid line: the overall HIV prevalence among incident TB cases; dashed line: HIV prevalence among incident male TB cases; dotted line: HIV prevalence among incident female TB cases. The table below shows the overall and gender stratified annual percentage change (APC) for HIV prevalence among incident TB cases with p-value for testing the difference in the change of prevalence over time between females and males.

	Overall	Female	Male	P-value
APC (95% CI)	-0.45 (-4.46, 3.74)	-0.95 (-6.52, 4.97)	0.08 (-5.53, 6.04)	0.81

Figure 4.6.4. Age-stratified gender disparity in HIV prevalence among incident TB cases in Harare, Zimbabwe during 2011-2017. Panel a: among 15-24 years; Panel b: among 25-44 years; Panel c: among 45-64 years; Panel d: among >65 years. The dashed and dotted lines show the trend of HIV prevalence among males and females, respectively. The table below shows the gender stratified annual percentage change (APC) of HIV prevalence for each age group and p-value for testing the difference in the change of prevalence over time between females and males.



	15-24 years	25-44 years	45-64 years	≥65 years
Female APC (95% CI)	-1.39(-5.20, 2.57)	-0.99 (-2.47, 0.52)	-0.08 (-2.98, 2.91)	-5.82 (-14.84, 4.14)
Male APC (95% CI)	0.89 (-3.35, 5.34)	-0.70 (-2.00, 0.61)	-0.07 (-2.30, 2.22)	-0.17 (-7.64, 7.92)
P-value	0.44	0.78	1.00	0.37

Chapter 5 Discussion

5.1 Summary of Main Findings

The goal of the work presented in this dissertation is to highlight the unique local epidemiological factors associated with TB epidemics across diverse settings. In the absence of an effective vaccine, existing diagnostic challenges and widening socio-demographic inequities, the need for epidemiological investigations to understand the local drivers of the TB epidemic is necessary to inform the development of locally relevant and pro equity TB interventions. I addressed this goal by conducting three studies and the main findings from each of them are summarized below.

In study 1, reported in Chapter 2, significant racial/ethnic disparities were found in TB incidence in Arkansas during the study period, 2010-2021, which persisted after adjusting for age and sex. The risk of TB among NHPI, Asian, Hispanic and NH Black persons was 113 (95% CI: 92.1–137.7), 17 (95% CI: 13.8, 21.3), 4 (95% CI: 3.3, 5.0) and 4 (95% CI: 3.2, 4.5) times the risk of TB among NH White persons in Arkansas, respectively. While racial disparities in TB are not particularly surprising given the vast body of previously published literature on the topic, this chapter quantified racial disparities for Arkansas at the population level using detailed racial/ethnic categorizations which are typically not available in state TB reports and underutilized in Arkansas based TB studies. These racial/ethnic disparities persisted across all age groups; the 0-14 year NHPI persons (RR=888, 95% CI: 403, 1962) were at a strikingly increased risk of TB when compared to similarly aged NH White persons. The risks of EPTB

and sputum smear PTB were significantly higher among racial/ethnic minorities, raising concerns related to inequitable delays in TB diagnosis in Arkansas.

In study 2, reported in Chapter 3, molecular typing of *Mtb* clinical isolates was integrated with TB surveillance data, which allowed us to quantify the relative contribution of recent transmission and reactivation of LTBI at the population level. At the state level, more than 1/3rd of TB cases were clustered during 2010-2021. NHPI (RR= 218.65, 95% CI: 150.60, 317.43) and NH Black (RR=7.37, 95% CI: 5.32, 10.22) persons were at a much higher risk of clustered TB while Asian (RR=1.62, 95% CI: 0.40, 6.65) and Hispanic (RR=1.70, 95% CI: 0.87, 3.31) persons experienced a similar risk of clustered TB when compared to NH White persons. As for non-clustered TB, all racial/ethnic minorities were at an increased risk compared to NH White persons, especially Asian (RR=20.45, 95% CI: 15.36, 27.22) and NHPI (RR= 25.28, 95% CI: 15.12, 42.29) persons. By quantifying the relative contribution of each of the two mechanisms by race/ethnicity, this study identified the utmost need to disrupt ongoing community transmission among NHPI and NH Black persons in Arkansas.

The risk of clustered TB was lower for the 0-14 year (RR=0.17, 95% CI: 0.07, 0.38) age group, when compared to the 15-24 year age group at the state level. The risk of clustered TB for 25-44 years (RR=1.37, 95% CI: 0.89, 2.10), 45-64 years (RR=1.50, 95% CI: 0.98, 2.30), and ≥65 years (RR=0.79, 95% CI: 0.47, 1.33) was comparable to the risk of clustered TB for 15-24 year age group. However, among NH White persons, the 45-64 years (RR=6.41, 95% CI: 1.53, 26.90) and ≥65 years (RR=6.50, 95% CI: 1.52, 27.81) had a higher risk of clustered TB compared to 15-24 years. The relative contribution of clustered TB versus non-clustered TB was higher among younger groups, even when stratified by race/ethnicity. These findings suggest that traditionally

recognized risk factors, such as young age, for clustered TB may have evolved over time and do not ubiquitously apply across racial/ethnic groups.

The risk of non-clustered TB at the state level increased with age, the 0-14 year (RR= 0.08, 95% CI: 0.03, 0.24) age group had the lowest risk while the ≥65 year (RR=3.41, 95% CI: 2.33, 4.99) age group had the highest risk. Among NH White persons, the risk of non-clustered TB was significantly higher for 45-64 years (RR=7.17, 95% CI: 2.23, 23.04) and ≥65 years (RR=19.72, 95% CI: 6.24, 62.28) when compared to the 15-24 years. Among NH Black persons, the risk of non-clustered TB was significantly higher for ≥65 years (RR=4.05, 95% CI: 1.94, 8.43). For Hispanic persons, the risk of non-clustered TB was significantly higher for 25-44 years (RR=2.75, 95% CI: 1.14, 6.64) 45-64 years (RR=3.53, 95% CI: 1.40, 8.89) and ≥65 years (RR=8.66, 95% CI: 3.15, 23.82). Elderly persons are more likely to develop TB as a result of reactivation of infection possibly acquired during times of historically high TB rates. Non-clustered TB can also suggest importation of infection from geographically distant areas such as TB endemic countries.

Gender disparity was consistently observed in Arkansas with men at 88% (95% CI: 1.64–2.15) higher risk of TB compared to women. The risk of both clustered (RR=1.93, 95% CI: 1.46, 2.55) and non-clustered TB (RR= 1.84, 95% CI: 1.49, 2.26) was significantly higher for men compared to women at the state level. Upon stratifying by race, NH Black men and NH White men were at a significantly higher risk of both clustered and non-clustered TB while gender disparity was not observed for other racial/ethnic groups. Men were also at a significantly higher risk of EPTB (RR=1.49, 95% CI: 1.14, 1.95) and sputum smear positive TB (RR= 2.40, 95% CI: 1.89, 3.05). While men are widely recognized as a high-risk group for TB, my dissertation

provides a more nuanced understanding of how gender and race interact with risk for various TB clinical phenotypes.

The effect of gender across diverse social and clinical contexts can vary and the characterization of these nuances may help uncover the mechanisms that drive gender disparities. In study 3 reported in Chapter 4, I characterized gender disparities in TB, TB-HIV coinfection and across granular categorizations of anatomic sites affected by TB. A large sample of around 24,000 TB patients and detailed clinical TB surveillance data from Harare, Zimbabwe, a high burden country for TB-HIV coinfection, allowed capturing statistical associations between gender and 9 different types of TB defined by the affected anatomic site. Similar to the low incidence setting, Arkansas, included in this dissertation, men (RR=1.53, 95% CI: 1.12, 2.09) were at an increased risk of TB compared to women in Harare. However, men aged between 15-24 years (RR=0.92, 95% CI: 0.74, 0.93) and 25-44 (RR=0.92, 95% CI: 0.88, 0.95) years were at a lower risk of coinfection when compared to women. The odds of abdominal TB (RR=0.51, 95% CI: 0.39, 0.68) 'other' EPTB (RR=0.69, 95% CI: 0.59, 0.81) and TB bones/joints/spine (RR=0.63, 95% CI=0.45, 0.90) compared to PTB were lower for men than women even after adjusting for HIV seropositivity.

To summarize, the epidemiological profile and drivers of TB vary across diverse settings and within subpopulations as demonstrated by the stratified analysis presented in this study. TB risk is concentrated among certain socio-demographic groups defined by race, ethnicity, age and gender, which potentially serve as an impediment to the progress towards global TB elimination. Furthermore, these disparities play out differently across diverse clinical, social and geographical contexts.

5.2 Public Health Implications

The significant racial/ethnic disparities identified through the first two studies of my dissertation have major implications for health equity in Arkansas, U.S. While there have been major reductions in TB incidence at the national and state level, the progress towards TB elimination has not been equitable with historically marginalized groups being left behind. For an infectious disease like TB, concentration of TB risk in certain groups translates into a threat for the entire community and the gains made in the fight against TB. This work has helped to map risk groups for targeted TB interventions and can serve as a guide for the development of pro-equity policies.

Through this work, it is evident that the state level estimates for TB are misleading as they mask the underlying disparities that may be fueling the remaining TB epidemic in low incidence states. The use of detailed racial/ethnic categorizations revealed that the TB incidence among NHPI persons in a low incidence state was comparable to the incidence of TB in some WHO high burden countries. Through an analysis stratified by important socio-demographics including age, gender, race/ethnicity and diverse clinical TB phenotypes, this dissertation provides a framework for an intersectional and disaggregated analysis of routinely collected TB surveillance data. While the findings of this study are relevant for Arkansas, this framework can be applied to other low incidence states in the U.S. to maximize the utility of TB surveillance data.

For public health practice, molecular epidemiologic investigations are essential as they help to decipher between the two mechanisms, recent transmission and reactivation of LTBI, each of which requires a specialized mitigation strategy. It is important to recognize that despite the low incidence of TB, ongoing community transmission plays a significant part in driving TB

incidence in Arkansas. Community outreach and improved contact tracing for NHPI and NH Black persons can help disrupt ongoing community transmission. Based on the race/ethnicity specific data presented, it is important to refocus TB mitigation efforts to include updated and non-traditional risk factors identified for clustered TB.

The findings from my study in Harare further the knowledge base of gender TB epidemiology and provide a nuanced understanding of the role of gender. Traditionally, men are widely recognized to be at an increased risk of TB but disaggregated analysis provided in this dissertation demonstrates that women are at an increased risk of more severe forms of TB and TB-HIV coinfection. These findings have implications for women's health due to poor prognosis and worse outcomes associated with HIV coinfection and disseminated TB. The increased risk of TB-HIV coinfection among women of childbearing age can result in increased antenatal transmission, suggesting that the implications of these findings go well beyond just women.

While the scale of the TB epidemic is global, the epidemiological factors that drive TB epidemics vary in diverse settings and across subpopulations. Consistent with the WHO's action framework for low incidence settings and the END TB Strategy, TB interventions need to be adapted at the local level as a one-for-all approach fails to consider population specific challenges and drivers of TB.

5.3 Limitations and Future Directions

In the end, I suggest future directions that can further the knowledge produced by this work while recognizing the limitations of the current study.

While we quantify the disparities in TB incidence and provide information on the drivers of TB incidence at the population level, the logistics of how this knowledge can be effectively translated into action, public health practice and policy requires further evaluation. Pilot

interventions may help to evaluate the feasibility and population level impact of targeted interventions based on this knowledge.

My dissertation demonstrated an increased risk of severe and advanced disease among racial/ethnic minorities in Arkansas. The mere quantification of these differences is not sufficient to equip public health programs to address them. A qualitative analysis of patient and provider related delays in TB care might help identify the areas along the care continuum that require immediate intervention. Similarly for gender disparities observed in Harare, we need to further evaluate why women are at an increased risk of severe TB. It is unclear if these gender disparities are a result of biological, social, behavioral or structural factors. Only by understanding the factors that drive these trends, we will be able to provide actionable knowledge required to intervene on these disparities.

An important aspect of understanding the gender disparity is to explore if access to ART is different across men and women as that may explain the observed HIV coinfection disparity. Additionally, future studies should also investigate gender disparity in HIV prevalence across age groups in Harare, as population wide HIV data was not available in this dissertation. Gender disparity persisted even after adjusting for HIV status but given the lack of data on CD4+ cell count may have resulted in residual confounding.

Since this dissertation utilized available TB surveillance data, it was limited in its ability to evaluate several key variables. The role of important behavioral factors such as smoking, and comorbidities such as obesity and diabetes in driving the TB epidemic in Arkansas, especially progression to TB disease, is an important area of investigation. The relative role of these comorbidities may also vary across racial/ethnic groups, thereby providing crucial knowledge for designing strategies particularly aimed at prevention of reactivation of LTBI.

As whole genome based methods are being rolled out for TB surveillance in the U.S. at the population level, future studies should attempt to use these advanced genotyping methods to determine the relative contribution of recent transmission and reactivation of LTBI with greater precision. Another important determinant of TB risk is nativity, which was partially explored in this dissertation (Chapter 2, Appendix). With the high risk of TB observed among non-U.S. born persons in Arkansas, it will be helpful to understand if this increased risk is driven by TB acquired in their country of origin or through exposure to precarious post immigration conditions in the U.S.

Bibliography

1 Mathema B, Andrews JR, Cohen T, *et al.* Drivers of Tuberculosis Transmission. *J Infect Dis* 2017; **216**. DOI:10.1093/infdis/jix354.

2 Bruchfeld J, Correia-Neves M, Kallenius G. Tuberculosis and HIV coinfection. *Cold Spring Harb Perspect Med* 2015; **5**. DOI:10.1101/cshperspect.a017871.

3 Berzkalns A, Bates J, Ye W, *et al.* The road to tuberculosis (*Mycobacterium tuberculosis*) elimination in Arkansas; a re-examination of risk groups. *PLoS One* 2014; **9**. DOI:10.1371/journal.pone.0090664.

4 Rothfeldt LL, Patil N, Haselow DT, Williams SH, Wheeler JG, Mukasa LN. Notes from the Field : Cluster of Tuberculosis Cases Among Marshallese Persons Residing in Arkansas — 2014–2015 . *MMWR Morb Mortal Wkly Rep* 2016; **65**. DOI:10.15585/mmwr.mm6533a7.

5 Centers for Disease Control and Prevention. Core Curriculum on Tuberculosis: What the Clinician Should Know. 2021; **SEVENTH ED**. https://www.cdc.gov/tb/education/corecurr/pdf/CoreCurriculumTB-508.pdf.

6 Coscolla M, Gagneux S. Consequences of genomic diversity in *mycobacterium tuberculosis*. Semin. Immunol. 2014; **26**. DOI:10.1016/j.smim.2014.09.012.

7 Ngabonziza JCS, Loiseau C, Marceau M, *et al.* A sister lineage of the *Mycobacterium tuberculosis* complex discovered in the African Great Lakes region. *Nat Commun* 2020; **11**. DOI:10.1038/s41467-020-16626-6.

8 Blouin Y, Hauck Y, Soler C, *et al.* Significance of the Identification in the Horn of Africa of an Exceptionally Deep Branching *Mycobacterium tuberculosis* Clade. *PLoS One* 2012; **7**. DOI:10.1371/journal.pone.0052841.

9 Freschi L, Vargas R, Husain A, *et al.* Population structure, biogeography and transmissibility of *Mycobacterium tuberculosis*. *Nat Commun* 2021; **12**. DOI:10.1038/s41467-021-26248-1.

10 Churchyard G, Kim P, Shah NS, *et al.* What We Know about Tuberculosis Transmission: An Overview. J. Infect. Dis. 2017; **216**. DOI:10.1093/infdis/jix362.

11 Long R, Divangahi M, Schwartzman K. Chapter 2: Transmission and pathogenesis of tuberculosis. *Can J Respir Crit Care, Sleep Med* 2022; **6**. DOI:10.1080/24745332.2022.2035540.

12 Mathema B, Kurepina NE, Bifani PJ, Kreiswirth BN. Molecular epidemiology of tuberculosis: Current insights. Clin. Microbiol. Rev. 2006; **19**. DOI:10.1128/CMR.00061-05.

13 Pai M, Behr MA, Dowdy D, *et al.* Tuberculosis. *Nat Rev Dis Prim* 2016; **2**: 16076.

14 Alsayed SSR, Gunosewoyo H. Tuberculosis: Pathogenesis, Current Treatment Regimens and New Drug Targets. Int. J. Mol. Sci. 2023; **24**. DOI:10.3390/ijms24065202.

15 Lillebaek T, Dirksen A, Baess I, Strunge B, Thomsen V, Andersen ÅB. Molecular evidence of endogenous reactivation of *Mycobacterium tuberculosis* after 33 years of latent infection. *J Infect Dis* 2002; **185**. DOI:10.1086/338342.

16 Lin PL, Flynn JL. The End of the Binary Era: Revisiting the Spectrum of Tuberculosis. *J Immunol* 2018; **201**. DOI:10.4049/jimmunol.1800993.

17 Ki HP, Shingadia D. Tuberculosis in children. Paediatr. Child Heal. (United Kingdom).

2017; **27**. DOI:10.1016/j.paed.2016.12.004.

18 Teo AKJ, Morishita F, Islam T, *et al.* Tuberculosis in older adults: challenges and best practices in the Western Pacific Region. Lancet Reg. Heal. - West. Pacific. 2023; **36**. DOI:10.1016/j.lanwpc.2023.100770.

19 Lin PL, Flynn JL. Understanding Latent Tuberculosis: A Moving Target. *J Immunol* 2010; **185**. DOI:10.4049/jimmunol.0903856.

20 Caraux-Paz P, Diamantis S, de Wazières B, Gallien S. Tuberculosis in the elderly. J. Clin. Med. 2021; **10**. DOI:10.3390/jcm10245888.

21 Basile FW, Nabeta P, Ruhwald M, Song R. Pediatric Tuberculosis Diagnostics: Present and Future. *J Pediatric Infect Dis Soc* 2022; **11**. DOI:10.1093/jpids/piac082.

22 Mugusi F, Villamor E, Urassa W, Saathoff E, Bosch RJ, Fawzi WW. HIV co-infection, CD4 cell counts and clinical correlates of bacillary density in pulmonary tuberculosis. *Int J Tuberc Lung Dis* 2006; **10**.

23 Lawn SD, Nicol MP. Xpert® MTB/RIF assay: Development, evaluation and implementation of a new rapid molecular diagnostic for tuberculosis and rifampicin resistance. Future Microbiol. 2011; **6**. DOI:10.2217/fmb.11.84.

24 Donoghue HD, Lee OYC, Minnikin DE, Besra GS, Taylor JH, Spigelman *M. Tuberculosis* in Dr Granville's mummy: A molecular re-examination of the earliest known Egyptian mummy to be scientifically examined and given a medical diagnosis. *Proc R Soc B Biol Sci* 2010; **277**. DOI:10.1098/rspb.2009.1484.

25 World Health Organization (WHO). Global tuberculosis report 2023. Geneva, 2023 DOI:Licence: CC BY-NC-SA 3.0 IGO.

26 World Health Organization. Global tuberculosis report 2020. 2020.

https://apps.who.int/iris/bitstream/handle/10665/336069/9789240013131-eng.pdf (accessed April 25, 2021).

27 Kwan C, Ernst JD. HIV and tuberculosis: A deadly human syndemic. Clin. Microbiol. Rev. 2011; **24**. DOI:10.1128/CMR.00042-10.

28 World Health Organization (WHO). Global tuberculosis report 2021. 2021.

29 Migliori GB, Dowdy D, Denholm JT, D'Ambrosio L, Centis R. The path to tuberculosis elimination: a renewed vision. *Eur Respir J* 2023; **61**. DOI:10.1183/13993003.00499-2023.

30 Lönnroth K, Migliori GB, Abubakar I, *et al.* Towards tuberculosis elimination: An action framework for low-incidence countries. Eur. Respir. J. 2015; **45**. DOI:10.1183/09031936.00214014.

31 World Health Organization (WHO). Towards tuberculosis elimination: an action framework for low-incidence countries. 2014 http://apps.who.int/iris/bitstream/handle/10665/132231/9789241507707_eng.pdf;jsessionid=8D9A66F6EB49FC0C6ACB771C7C114B74?sequence=1.

32 World Health Organization (WHO). WHO operational handbook on tuberculosis. Module 1: prevention - infection prevention and control. Geneva, 2023 https://iris.who.int/bitstream/handle/10665/372738/9789240078154-eng.pdf?sequence=1.

33 World Health Organization (WHO). WHO operational handbook on tuberculosis. Module 5: management of tuberculosis in children and adolescents. Geneva, 2022.

34 World Health Organization. WHO operational handbook on tuberculosis. Module 2: screening - systematic screening for tuberculosis disease. Geneva, 2021.

35 World Health Organization (WHO). WHO operational handbook on tuberculosis. Module

1: prevention - tuberculosis preventive treatment. Geneva, 2020.

36 U.S. Department of Health and Human Services. Health Equity and Health Disparities Environmental Scan. Rockville, MD, 2022.

37 Braveman P. What are health disparities and health equity? we need to be clear. *Public Health Rep* 2014; **129**. DOI:10.1177/00333549141291s203.

38 Lee H, Kim D, Lee SA, Fawcett J. The concepts of health inequality, disparities and equity in the era of population health. *Appl Nurs Res* 2020; **56**. DOI:10.1016/j.apnr.2020.151367.

39 Noppert GA, Clarke P, Hicken MT, Wilson ML. Understanding the intersection of race and place: The case of tuberculosis in Michigan. *BMC Public Health* 2019; **19**. DOI:10.1186/s12889-019-8036-y.

40 Noppert GA, Malosh RE, Moran EB, Ahuja SD, Zelner J. Contemporary Social Disparities in TB Infection and Disease in the USA: a Review. *Curr Epidemiol Reports* 2018; **5**. DOI:10.1007/s40471-018-0171-y.

41 Lönnroth K, Jaramillo E, Williams BG, Dye C, Raviglione M. Drivers of tuberculosis epidemics: The role of risk factors and social determinants. *Soc Sci Med* 2009; **68**. DOI:10.1016/j.socscimed.2009.03.041.

42 Noppert GA, Wilson ML, Clarke P, Ye W, Davidson P, Yang Z. Race and nativity are major determinants of tuberculosis in the U.S.: evidence of health disparities in tuberculosis incidence in Michigan, 2004-2012. *BMC Public Health* 2017; **17**. DOI:10.1186/s12889-017-4461-y.

43 Horton KC, MacPherson P, Houben RMGJ, White RG, Corbett EL. Sex Differences in Tuberculosis Burden and Notifications in Low- and Middle-Income Countries: A

Systematic Review and Meta-analysis. PLoS Med. 2016; **13**. DOI:10.1371/journal.pmed.1002119.

44 Ben Jmaa M, Ben Ayed H, Koubaa M, Hammami F, Damak J, Ben Jemaa M. Is there gender inequality in the epidemiological profile of tuberculosis? *Tunis Med* 2020; **98**.

45 Lin CY, Chen TC, Lu PL, *et al.* Effects of Gender and Age on Development of Concurrent Extrapulmonary Tuberculosis in Patients with Pulmonary Tuberculosis: A Population Based Study. *PLoS One* 2013; **8**. DOI:10.1371/journal.pone.0063936.

46 Kim S, Cohen T, Horsburgh CR, *et al.* 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* 2022; **74**. DOI:10.1093/cid/ciab668.

47 Murray M, Nardell E. Molecular epidemiology of tuberculosis: Achievements and challenges to current knowledge. Bull. World Health Organ. 2002; **80**. DOI:10.1590/S0042-96862002000600013.

48 Burgos M V, Pym AS. Molecular epidemiology of tuberculosis. *Eur Respir J* 2002; **20**: 54s LP-65s.

49 Barnes PF. MDC. Molecular Epidemiology of Tuberculosis. *N Engl J Med* 2003; **349**: 1149–56.

50 García de Viedma D, Pérez-Lago L. The Evolution of Genotyping Strategies to Detect, Analyze, and Control Transmission of Tuberculosis. In: Microbial Transmission. 2018. DOI:10.1128/microbiolspec.mtbp-0002-2016.

51 World Health Organization (WHO). WHO reveals leading causes of death and disability worldwide: 2000-2019. *World Heal Organ* 2020.

52 World Health Organization (WHO). Global Tuberculosis Report. Geneva, 2022

https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2022/tb-disease-burden/2-1-tb-incidence.

53 Houben RMGJ, Dodd PJ. The Global Burden of Latent Tuberculosis Infection: A Re-estimation Using Mathematical Modelling. *PLoS Med* 2016; **13**. DOI:10.1371/journal.pmed.1002152.

54 WHO. World Health Organization. The End TB Strategy. (2015). WHO, Geneva, Switzerland. http://www.who.int/tb/post2015_TBstrategy.pdf (Accessed 23rd April 2018). *World Heal Organ* 2015.

55 The Lancet Regional Health -Americas. Next steps in eliminating tuberculosis in the Americas. *Lancet Reg Heal -Americas* 2023; **21**. https://www.sciencedirect.com/science/article/pii/S2667193X23000868.

56 Filardo TD, Feng P-J, Pratt RH, Price SF, Self JL. Tuberculosis — United States, 2021. *MMWR Morb Mortal Wkly Rep* 2022; **71**. DOI:10.15585/mmwr.mm7112a1.

57 Renardy ME, Gillen C, Yang Z, *et al.* Disease phenotypic and geospatial features vary across genetic lineages for Tuberculosis within Arkansas, 2010–2020. *PLOS Glob Public Heal* 2023; **3**. DOI:10.1371/journal.pgph.0001580.

58 Arkansas Department of Health TCP. Tuberculosis Morbidity Annual Statistical Report 2016. 2016.

59 Heuvelings CC, de Vries SG, Grobusch MP. Tackling TB in low-incidence countries: improving diagnosis and management in vulnerable populations. Int. J. Infect. Dis. 2017; **56**. DOI:10.1016/j.ijid.2016.12.025.

60 Bailey ZD, Krieger N, Agénor M, Graves J, Linos N, Bassett MT. Structural racism and health inequities in the USA: evidence and interventions. Lancet. 2017; **389**.

DOI:10.1016/S0140-6736(17)30569-X.

61 France AM, Cave MD, Bates JH, Foxman B, Chu T, Yang Z. 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* 2007; **166**. DOI:10.1093/aje/kwm135.

62 Yang ZH, Gorden T, Liu DP, Mukasa L, Patil N, Bates JH. Increasing likelihood of advanced pulmonary tuberculosis at initial diagnosis in a low-incidence US state. *Int J Tuberc Lung Dis* 2018; **22**. DOI:10.5588/ijtld.17.0413.

63 Centers for Disease Control and Prevention. CDC Tuberculosis Surveillance Data Training: Report of Verified Case of Tuberculosis, Instructional Manual. https://www.cdc.gov/tb/programs/rvct/InstructionManual.pdf (accessed Jan 8, 2021).

64 Bureau USC. US Census Bureau. State population totals and components change: 2020-2022. 2022. https://www.census.gov/data/tables/time-series/demo/popest/2020s-state-total.html (accessed Feb 12, 2022).

65 Flanagin A, Frey T, Christiansen SL, Bauchner H. The Reporting of Race and Ethnicity in Medical and Science Journals: Comments Invited. JAMA - J. Am. Med. Assoc. 2021; **325**. DOI:10.1001/jama.2021.2104.

66 Wu IL, Chitnis AS, Jaganath D. A narrative review of tuberculosis in the United States among persons aged 65 years and older. J. Clin. Tuberc. Other Mycobact. Dis. 2022; **28**. DOI:10.1016/j.jctube.2022.100321.

67 Shingadia D, Novelli V. Diagnosis and treatment of tuberculosis in children. Lancet Infect. Dis. 2003; **3**. DOI:10.1016/S1473-3099(03)00771-0.

68 Baykan AH, Sayiner HS, Aydin E, Koc M, Inan I, Erturk SM. Extrapulmonary

tuberculosıs: an old but resurgent problem. Insights Imaging. 2022; **13**. DOI:10.1186/s13244-022-01172-0.

69 Houben RMGJ, Menzies NA, Sumner T, *et al.* Feasibility of achieving the 2025 WHO global tuberculosis targets in South Africa, China, and India: a combined analysis of 11 mathematical models. *Lancet Glob Heal* 2016; **4**. DOI:10.1016/S2214-109X(16)30199-1.

70 Khan A, Marks S, Katz D, *et al.* Changes in tuberculosis disparities at a time of decreasing tuberculosis incidence in the United States, 1994-2016. *Am J Public Health* 2018; **108**. DOI:10.2105/AJPH.2018.304606.

71 Division of Tuberculosis Elimination, National Center for HIV, Viral Hepatitis, STD, and TB Prevention C for DC and P. Health Disparities in TB. 2022 https://www.cdc.gov/tb/topic/populations/healthdisparities/default.htm.

72 Division of Tuberculosis Elimination, National Center for HIV, Viral Hepatitis, STD, and TB Prevention C for DC and P. Reported Tuberculosis in the United States, 2021. 2022 https://www.cdc.gov/tb/statistics/reports/2021/table2.htm.

73 Ritz, Sarah; Cash DH. Republic of the Marshall Islands Hybrid Survey: Final Report. 2018.

74 Riklon S, Alik W, Hixon A, Palafox NA. The ‘compact impact’ in Hawaii: focus on health care. *Hawaii Med J* 2010; **69**.

75 McElfish PA, Post J, Rowland B. A Social Ecological and Community-Engaged Perspective for Addressing Health Disparities Among Marshallese in Arkansas. *Int J Nurs Clin Pract* 2016; **3**. DOI:10.15344/2394-4978/2016/191.

76 Pareek M, Greenaway C, Noori T, Munoz J, Zenner D. The impact of migration on tuberculosis epidemiology and control in high-income countries: A review. BMC Med.

2016; **14**. DOI:10.1186/s12916-016-0595-5.

77 Swaminathan S, Rekha B. Pediatric tuberculosis: Global overview and challenges. Clin. Infect. Dis. 2010; **50**. DOI:10.1086/651490.

78 Labuda SM, McDaniel CJ, Talwar A, *et al.* Tuberculosis Outbreak Associated With Delayed Diagnosis and Long Infectious Periods in Rural Arkansas, 2010-2018. *Public Health Rep* 2022; **137**. DOI:10.1177/0033354921999167.

79 Yuen CM, Kammerer JS, Marks K, Navin TR, France AM. Recent transmission of tuberculosis - United States, 2011-2014. *PLoS One* 2016; **11**. DOI:10.1371/journal.pone.0153728.

80 Migliori GB, Nardell E, Yedilbayev A, *et al.* Reducing tuberculosis transmission: A consensus document from the World Health Organization Regional Office for Europe. *Eur Respir J* 2019; **53**. DOI:10.1183/13993003.00391-2019.

81 Tedla K, Medhin G, Berhe G, Mulugeta A, Berhe N. Delay in treatment initiation and its association with clinical severity and infectiousness among new adult pulmonary tuberculosis patients in Tigray, northern Ethiopia. *BMC Infect Dis* 2020; **20**. DOI:10.1186/s12879-020-05191-4.

82 Mbuh TP, Ane-Anyangwe I, Adeline W, Thumamo Pokam BD, Meriki HD, Mbacham WF. Bacteriologically confirmed extra pulmonary tuberculosis and treatment outcome of patients consulted and treated under program conditions in the littoral region of Cameroon 11 Medical and Health Sciences 1103 Clinical Sciences. *BMC Pulm Med* 2019; **19**. DOI:10.1186/s12890-018-0770-x.

83 Yang Z, Kong Y, Wilson F, *et al.* Identification of Risk Factors for Extrapulmonary Tuberculosis. *Clin Infect Dis* 2004; **38**. DOI:10.1086/380644.

84 Office of Minority Health and Health Disparities and Epidemiological Branch. Disparities in Diabetes Mellitus Mortality Among Blacks in Arkansas. .

85 World Health Organization. Framework for collaborative action on tuberculosis and comorbidities. Geneva, 2022.

86 Raz KM, Talarico S, Althomsons SP, *et al.* Molecular surveillance for large outbreaks of tuberculosis in the United States, 2014–2018. *Tuberculosis* 2022; **136**. DOI:10.1016/j.tube.2022.102232.

87 Reported Tuberculosis in the United States, 2022. 2022 https://www.cdc.gov/tb/statistics/reports/2022/national_data.htm#Data Tables.

88 Humayun M, Mukasa L, Ye W, Bates J, Yang Z. Racial and Ethnic Disparities in Tuberculosis Incidence, Arkansas, United States, 2010–2021. *Emerg Infect Dis J* 2024; **30**. DOI:10.3201/eid3001.230778.

89 Campbell JR, Dowdy D, Schwartzman K. Treatment of latent infection to achieve tuberculosis elimination in low-incidence countries. *PLoS Med* 2019; **16**. DOI:10.1371/journal.pmed.1002824.

90 Daley CL, Kawamura LM. The role of molecular epidemiology in contact investigations: A US perspective. Int. J. Tuberc. Lung Dis. 2003; **7**.

91 Heldal E, Dahle UR, Sandven P, *et al.* Risk factors for recent transmission of Mycobacterium tuberculosis. *Eur Respir J* 2003; **22**. DOI:10.1183/09031936.03.00019103.

92 Kamper-Jørgensen Z, Andersen AB, Kok-Jensen A, Bygbjerg IC, Thomsen VO, Lillebaek T. Characteristics of non-clustered tuberculosis in a low burden country. *Tuberculosis* 2012; **92**. DOI:10.1016/j.tube.2012.02.001.

93 Division of Tuberculosis Elimination, National Center for HIV, Viral Hepatitis, STD, and TB Prevention C for DC and P. GENType: New Genotyping Terminology to integrate 24-locus MIRU-VNTR. https://www.cdc.gov/tb/publications/factsheets/statistics/GenotypingTerminology.pdf.

94 Pareek M, Greenaway C. Programmatic screening for latent tuberculosis in migrants. Lancet Public Heal. 2022; **7**. DOI:10.1016/S2468-2667(22)00064-0.

95 Nava-Aguilera E, Andersson N, Harris E, *et al.* Risk factors associated with recent transmission of tuberculosis: Systematic review and meta-analysis. Int. J. Tuberc. Lung Dis. 2009; **13**.

96 Yuen CM, Amanullah F, Dharmadhikari A, *et al.* Turning off the tap: Stopping tuberculosis transmission through active case-finding and prompt effective treatment. Lancet. 2015; **386**. DOI:10.1016/S0140-6736(15)00322-0.

97 Ai JW, Ruan QL, Liu QH, Zhang WH. Updates on the risk factors for latent tuberculosis reactivation and their managements. Emerg. Microbes Infect. 2016; **5**. DOI:10.1038/emi.2016.10.

98 Dobler CC, Fox GJ, Douglas P, *et al.* Screening for tuberculosis in migrants and visitors from high-incidence settings: Present and future perspectives. Eur. Respir. J. 2018; **52**. DOI:10.1183/13993003.00591-2018.

99 Rajagopalan S. Tuberculosis and aging: A global health problem. *Clin Infect Dis* 2001; **33**. DOI:10.1086/322671.

100 Braden CR, Templeton GL, Donald Cave M, *et al.* Interpretation of restriction fragment length polymorphism analysis of *Mycobacterium tuberculosis* isolates from a state with a large rural population. *J Infect Dis* 1997; **175**. DOI:10.1086/516478.

101 Dabitao D, Bishai WR. Sex and Gender Differences in Tuberculosis Pathogenesis and Treatment Outcomes. In: Current Topics in Microbiology and Immunology. 2023. DOI:10.1007/978-3-031-35139-6_6.

102 Diwan VK, Thorson A. Sex, gender, and tuberculosis. Lancet. 1999; **353**. DOI:10.1016/S0140-6736(99)01318-5.

103 Christianson S, Wolfe J, Orr P, *et al.* Evaluation of 24 locus MIRU-VNTR genotyping of *Mycobacterium tuberculosis* isolates in Canada. *Tuberculosis* 2010; **90**. DOI:10.1016/j.tube.2009.12.003.

104 Teeter LD, Kammerer JS, Ghosh S, *et al.* Evaluation of 24-locus MIRU-VNTR genotyping in *Mycobacterium tuberculosis* cluster investigations in four jurisdictions in the United States, 2006–2010. *Tuberculosis* 2017; **106**. DOI:10.1016/j.tube.2017.05.003.

105 Division of Tuberculosis Elimination, National Center for HIV, Viral Hepatitis, STD, and TB Prevention C for DC and P. Transition to whole-genome multilocus sequence typing (wgMLST) for TB cluster detection. https://www.cdc.gov/tb/programs/genotyping/pdf/wgMLST-Training-Webinar.pdf.

106 Diel R, Kohl TA, Maurer FP, *et al.* Accuracy of whole-genome sequencing to determine recent tuberculosis transmission: An 11-year population-based study in Hamburg, Germany. Eur. Respir. J. 2019; **54**. DOI:10.1183/13993003.01154-2019.

107 Jajou R, de Neeling A, van Hunen R, *et al.* Correction: Epidemiological links between tuberculosis cases identified twice as efficiently by whole genome sequencing than conventional molecular typing: A population-based study(PLoS ONE (2018)13:4 (e0195413) DOI: 10.1371/journal.pone.0195413). PLoS One. 2018; **13**. DOI:10.1371/journal.pone.0197556.

108 Hargreaves JR, Boccia D, Evans CA, Adato M, Petticrew M, Porter JDH. The social determinants of tuberculosis: from evidence to action. *Am J Public Health* 2011; **101**. DOI:10.2105/AJPH.2010.199505.

109 Katrak S, Flood J. Latent tuberculosis and current health disparities in California: Making the invisible visible. Am. J. Public Health. 2018; **108**. DOI:10.2105/AJPH.2018.304529.

110 Mauvais-Jarvis F, Bairey Merz N, Barnes PJ, *et al.* Sex and gender: modifiers of health, disease, and medicine. Lancet. 2020; **396**. DOI:10.1016/S0140-6736(20)31561-0.

111 Phillips SP. Measuring the health effects of gender. *J Epidemiol Community Health* 2008; **62**. DOI:10.1136/jech.2007.062158.

112 Solar O, Irwin A. Social Determinants of Health Discussion Paper 2 (Policy and Practice). *World Heal Organ* 2010; : 1–79.

113 Gender in Tuberculosis Research. Geneva, 2004.

114 Holmes CB, Hausler H, Nunn P. A review of sex differences in the epidemiology of tuberculosis. Int. J. Tuberc. Lung Dis. 1998; **2**.

115 Thorson A, Diwan VK. Gender inequalities in tuberculosis: Aspects of infection, notification rates, and compliance. Curr. Opin. Pulm. Med. 2001; **7**. DOI:10.1097/00063198-200105000-00009.

116 Bauer GR. Incorporating intersectionality theory into population health research methodology: Challenges and the potential to advance health equity. *Soc Sci Med* 2014; **110**. DOI:10.1016/j.socscimed.2014.03.022.

117 Zhang X, Andersen AB, Lillebaek T, *et al.* Effect of sex, age, and race on the clinical presentation of tuberculosis: A 15-year population-based study. *Am J Trop Med Hyg* 2011; **85**. DOI:10.4269/ajtmh.2011.10-0630.

118 Noppert G, Yang Z, Sandy C, Chirenda J. Trends of sputum-smear positive tuberculosis in Zimbabwe: 2008-2011 Public Health. *BMC Res Notes* 2015; **8**. DOI:10.1186/s13104-015-1568-z.

119 Zimbabwe National Statistics Agency. Population Projections Thematic Report. 2015.

120 Scully EP. Sex, gender and infectious disease. Nat. Microbiol. 2022; **7**. DOI:10.1038/s41564-022-01064-5.

121 Nhamoyebonde S, Leslie A. Biological differences between the sexes and susceptibility to tuberculosis. *J Infect Dis* 2014; **209**. DOI:10.1093/infdis/jiu147.

122 Awasthi S, Saxena M, Ahmad F, Kumar A, Dutta S. Abdominal tuberculosis: A diagnostic dilemma. *J Clin Diagnostic Res* 2015; **9**. DOI:10.7860/JCDR/2015/13350.5887.

123 Sreeramareddy CT, Panduru K V., Verma SC, Joshi HS, Bates MN. Comparison of pulmonary and extrapulmonary tuberculosis in Nepal - A hospital-based retrospective study. *BMC Infect Dis* 2008; **8**. DOI:10.1186/1471-2334-8-8.

124 Ohene SA, Bakker MI, Ojo J, Toonstra A, Awudi D, Klatser P. Extra-pulmonary tuberculosis: A retrospective study of patients in Accra, Ghana. *PLoS One* 2019; **14**. DOI:10.1371/journal.pone.0209650.

125 Tahseen S, Khanzada FM, Baloch AQ, *et al.* Extrapulmonary tuberculosis in Pakistan- A nation-wide multicenter retrospective study. *PLoS One* 2020; **15**. DOI:10.1371/journal.pone.0232134.

126 Aliyu G, El-Kamary SS, Abimiku A, Blattner W, Charurat M. Demography and the dual epidemics of tuberculosis and HIV: Analysis of cross-sectional data from Sub-Saharan Africa. *PLoS One* 2019; **13**. DOI:10.1371/journal.pone.0191387.

127 Leeds IL, Magee MJ, Kurbatova E V., *et al.* Site of extrapulmonary tuberculosis is associated with HIV infection. *Clin Infect Dis* 2012; **55**. DOI:10.1093/cid/cis303.

128 Forssbohm M, Zwahlen M, Loddenkemper R, Rieder HL. Demographic characteristics of patients with extrapulmonary tuberculosis in Germany. *Eur Respir J* 2008; **31**. DOI:10.1183/09031936.00020607.

129 Peto HM, Pratt RH, Harrington TA, LoBue PA, Armstrong LR. Epidemiology of extrapulmonary tuberculosis in the United States, 1993-2006. *Clin Infect Dis* 2009; **49**. DOI:10.1086/605559.

130 Habib SS, Jamal WZ, Zaidi SMA, *et al.* Barriers to access of healthcare services for rural women— applying gender lens on tb in a rural district of sindh, pakistan. *Int J Environ Res Public Health* 2021; **18**. DOI:10.3390/ijerph181910102.

131 Perumal R, Naidoo K, Padayatchi N. TB epidemiology: Where are the young women? Know your tuberculosis epidemic, know your response. *BMC Public Health* 2018; **18**. DOI:10.1186/s12889-018-5362-4.

132 United Nations. Economic and Social Council; United Nations. Economic Commission for Africa. The socio-economic impact of HIV/AIDS among women and children UNICEF/Nairobi/Kenya. https://hdl.handle.net/10855/21643.

133 Ramjee G, Daniels B. Women and HIV in Sub-Saharan Africa. AIDS Res. Ther. 2013; **10**. DOI:10.1186/1742-6405-10-30.