Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Mitchell Dvorak Date

Finding TB transmission: An assessment of CDC's prioritization of TB genotype clusters in the United States, 2016-2019

By

Mitchell Dvorak
Master of Public Health

Epidemiology

_____

Clinton McDaniel, MPH
Committee Chair

_____

Kala Marks Raz, MPH
Committee Member

_____

Kristin Nelson, PhD, MPH
Committee Member

Finding TB transmission: An assessment of CDC's prioritization of TB genotype clusters in the United States, 2016-2019

By

Mitchell Dvorak

B.S. in Biochemistry
University of Wisconsin – Madison
2015

Thesis Committee Chair: Kristin Nelson, PhD, MPH

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology
2023

Abstract

Finding TB transmission: An assessment of CDC's prioritization of TB genotype clusters in the United States, 2016-2019

By Mitchell Dvorak

The Division of Tuberculosis Elimination (DTBE) at the Centers for Disease Control and Prevention conducts routine surveillance of tuberculosis (TB) clusters in the United States and collaborates with local and state partners to control spread. The DTBE cluster priority assignment, on a scale of Priority 1 to 4, is a tool that allows for DTBE to systematically review and respond to TB clusters based on perceived risk of recent or ongoing transmission. While this process, in its current state, has occurred since 2016, no study has evaluated genotype-matched TB cluster growth following DTBE priority assignment or identified case characteristics associated with cluster growth with the context of DTBE priority assignment. Therefore, we aimed to 1) identify difference in cluster growth by DTBE cluster priority assignment level, 2) identify patient characteristics associated with DTBE cluster priority assignment, and 3) identify patient characteristics that were associated with clusters attaining at least one additional case during the two-year follow up. This analysis utilized National TB Surveillance System and TB Genotyping Information Management System data from 2013-2021 and 2016-2019, respectively. We identified a statistically significant difference (p-value<0.0001) in both cluster growth and number of cases added in a two-year follow up by DTBE cluster priority assignment. Among clusters assigned Priority 1 by DTBE, 53% gained at least one additional case in the two years following a cluster alert. Alternatively, among clusters assigned Priority 4 by DTBE, only 20% gained at least one additional case in the two-year follow up. Analysis of case characteristics showed HIV coinfection as a significant predictor for cluster growth in a multivariate model [aOR=1.7; 95% CI= (1.1, 2.5)] but was not significant in a model for DTBE priority assignment. This analysis validates the effectiveness of DTBE cluster priority assignment in identifying TB clusters likely to gain additional cases but indicates that HIV status may be given additional weight when assigning cluster priority.

Finding TB transmission: An assessment of CDC's prioritization of TB genotype clusters in the United States, 2016-2019

By

Mitchell Dvorak

B.S. in Biochemistry
University of Wisconsin – Madison
2015

Thesis Committee Chair: Kristin Nelson, PhD, MPH

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology
2023

ACKNOWLEDGEMENTS

I wish to express my deepest gratitude to my joint field advisors, Clinton McDaniel, MPH, and Kala Marks Raz, MPH, for their incredible guidance and mentorship, and Kristin Nelson, PhD, MPH, my faculty advisor, for her steadfast encouragement. I also wish to thank my friends, family, and my partner for their unwavering love and support.

Table of Contents

Introduction

Decreases in tuberculosis (TB) incidence in the United States have been observed since the early 1990s.[1] Despite declines in TB incidence in the U.S. over the previous decades until 2021, the U.S. has not yet achieved TB elimination (defined as one case per 1,000,000 persons per year) and the future spread of TB in the United States remains uncertain.[1,2]

TB disease can be caused by recent *Mycobacterium tuberculosis* infection or by reactivation of latent TB infection (LTBI) acquired in the remote past.[3-5] Reactivation of LTBI remains a concern as it may result in new transmissions and seed or expand TB clusters.[2,6] TB clusters are comprised of ≥2 patients with matching *M. tuberculosis* genotypes in the same geographic area.[3,7] Without public health intervention and given certain transmission dynamics, clusters may expand and become TB outbreaks. Previous studies define TB outbreaks as ≥6 cases with the same genotype within the same geographic area.[3,7] TB outbreak response can be expensive and resource demanding.[7] Therefore, identifying and interrupting recent transmission remains an important strategy for TB control.

Previous studies have shown the success of cluster surveillance and the importance of identifying clusters that attain additional cases in rapid succession in preventing TB outbreaks.[7,8] Prompt identification of clusters that are likely to grow remains an important activity for decreasing TB incidence and is necessary to prevent future outbreaks.[4,9]

Public health professionals utilize contact tracing of patients with confirmed TB and genotyping of tuberculosis isolates to identify transmission networks. Genotyping is an important component of TB cluster investigation and is based on the principle that cases involved in the same chain of transmission will share the same genotype.[9-14] Since 2004, universal genotyping

has been performed for at least one *M. tuberculosis* isolate from each culture-confirmed TB case in the United States.[3] In 2009, GENType genotyping, which characterizes genotypes using a combination of spacer oligonucleotide typing (spoligotyping) and 24-locus mycobacterial interspersed repetitive unit–variable number tandem repeat (24-locus MIRU-VNTR) typing, was implemented to provide increased molecular resolution compared to previous methods.[15,16] However, these conventional genotyping methods only investigate <1% of the *M. tuberculosis* genome, so the molecular resolution can be insufficient for differentiating isolates presumably involved in recent transmission.[8,13,15-18] Therefore, in 2018, the National TB Molecular Surveillance Center began performing universal whole-genome sequencing (WGS) for all culture-positive TB isolates in the United States as part of a planned transition to replace conventional genotyping methods with whole-genome multi-locus sequencing typing (wgMLSType). Whole-genome multi-locus sequencing typing is a genotyping scheme that relies on WGS data and offers increased molecular resolution by expanding coverage of the genome to ~70%. All *M. tuberculosis* isolates from 2018 forward are assigned a wgMLSType.[8,14]

In 2011, the Division of TB Elimination (DTBE) implemented a national TB cluster alert system to facilitate early and systematic detection of TB clusters. The cluster alerts utilize the log-likelihood ratio (LLR) statistic to identify spatiotemporal concentrations of genotype-matched TB cases that may represent recent transmission during the previous three years.[3,7,11] Each week, medium and high level alerts are generated if the LLR statistic for a particular genotype and county crosses a threshold of ≥5 and ≥10, respectively. Since 2016, DTBE staff have routinely assessed all TB cluster alerts using demographic, clinical, genotypic, epidemiologic, and social risk factor data reported to the National TB Surveillance System. Each cluster alert is prioritized by consensus on a scale from 1 to 4 based on the likelihood for recent

transmission and the potential for continued cluster growth.[14] Clusters that are assessed as a high priority (Priority 1 and 2) indicate a high level of concern for recent transmission and continued cluster growth, whereas clusters that are assessed as a low priority (Priority 3 and 4) indicate a low level of concern for recent transmission and continued cluster growth. The priority assessment serves as a guide for determining the degree to which DTBE staff will engage with state TB programs on follow-up action for TB clusters.[14]

The cluster assessment and prioritization process involves consideration of numerous factors. Characteristics about the cluster – including the rarity and distribution of the genotype, presence of a sudden increase in the number of cases in the past two years, diagnosis of children <5 years of age, or association with a known outbreak – can be indicators that cases in the cluster are related by recent transmission.[7,14,19] Patient characteristics are also an important consideration when evaluating the likelihood of recent transmission and the potential for future transmission. Social risk factors such as a history of homelessness, substance use, or diagnosis in a congregate setting, such as a correctional or healthcare facility, can increase concern for broad, ongoing exposures. Additionally, demographic, clinical, and social risk factor data have been found to be useful for predicting clusters that are at risk for additional growth.[14] U.S.-born patients represent a large portion of cases in clusters with ≥3 cases, even though most reported TB cases in the U.S. are among non-U.S.-born patients.[8] Patients in clusters that belong to marginalized communities, including persons experiencing homelessness, persons with a history of substance use, and incarcerated persons, have been shown to indicate risk for cluster growth.[7,14,20-25] Additionally, clusters containing persons with HIV coinfection may indicate recent transmission, due to increased susceptibility and rapid progression of TB disease among people living with HIV.[26]

DTBE has prioritized clusters and collaborated with state and local public health agencies to respond to concerning TB clusters and outbreaks since 2016 and has adapted the cluster prioritization protocol to incorporate current genotyping practices, consider relevant clinical and social risk factors of cases, and track cluster growth patterns. Prioritizing tuberculosis clusters likely to represent recent transmission or that are at risk for growth help focus limited resources where interventions will have the greatest public health impact.[7,14] While previous studies have examined risk factors for outbreak development and have been utilized to inform DTBE cluster prioritization and response, no study has been conducted to analyze cluster growth following DTBE response process and characteristics of cases associated with general cluster growth.[7,8] While TB outbreaks are resource demanding and must be controlled, most routine surveillance and programmatic work from public health professionals at the local, state, and federal level involve single case or small clusters (<6 cases).[7,27] Understanding and increasing the accuracy with which DTBE identifies clusters likely to grow can improve the collaborative public health response among federal, state and local partners and lead to more targeted interventions for interrupting ongoing TB transmission.[9]

Therefore, we analyzed TB cluster alerts in the United States to assess cluster growth following DTBE prioritization and response to 1) determine the presence of statistically significant difference in cluster growth, measured by number of additional cases within two years following a cluster alert, by DTBE prioritization level, 2) identify patient demographic, clinical, or social risk factors among clusters that are significantly different by DTBE cluster priority level assignment, and 3) identify patient demographic, clinical, or social risk factor characteristics among clusters that are significantly different from those of clusters that added additional cases. Given the robust prioritization protocol, which considers numerous factors

involved in cluster transmission dynamics, we hypothesize that differences in cluster growth by initial cluster prioritization level will be significant, with high priority clusters (Priority 1 and 2) being more greatly associated with cluster growth following DTBE cluster prioritization and response.

Methods

Data Source and Study Population

To evaluate DTBE's cluster prioritization process, we used data on cluster alerts from the TB Genotyping Information Management System (TB GIMS), including date of alert, alert level, and DTBE prioritization assignment, in conjunction with patient surveillance data from the National Tuberculosis Surveillance System (NTSS) containing data on patient demographic, clinical, and social risk characteristics and any known epidemiologic linkages.

We utilized NTSS data on *M. tuberculosis* cases reported during 2013–2019 that were identified in a GENType-based cluster alert in the 50 U.S. states and the District of Columbia during 2016–2019. GENType clusters can alert multiple times within the same county, which might result in the same case being included in more than one cluster alert. Therefore, we excluded re-alerts of GENType clusters from this analysis.

Analyzing Cluster Growth

Cluster growth was defined as the addition of one or more GENType-matched cases reported in the alerted county during the two years following alert date. This threshold for cluster growth was selected based on most (60.9%) of GENType-based clusters not gaining additional

cases and a large proportion (46.6%) of GENType-based clusters that gained additional cases adding only one additional case in the subsequent two years.

We used the Kruskal-Wallis test to assess differences in cluster growth, measured by median number of cases added to a cluster in two years following a cluster alert, by DTBE priority assignment (Priority 1-4). We used the Wilcoxon rank-sum test to determine whether there are statistical differences in the median number of cases added in the two years following the alert date stratified by DTBE priority level (high vs. low priority).

Social and Clinical Characteristics of Cases

We employed logistic regression to identify characteristics of cases in cluster alerts that were associated with cluster growth and with DTBE assigned priority level (high vs. low priority).

We evaluated NTSS variables that are commonly assessed by DTBE staff when prioritizing cluster alerts: race/ethnicity, sex, defined age-groups, country of birth, HIV status, history of homelessness in the year prior to diagnosis, residence in correctional facility or long-term care facility at the time of diagnosis, primary occupation in the year prior to diagnosis, death at diagnosis or during treatment, treatment completion, drug susceptibility results, diabetes, renal disease, and the presence of known epidemiologic links among cases. To evaluate any substance use as a social risk factor, we created a variable to identify cases with reported excess alcohol use or use of illicit drugs in the year prior to diagnosis. We evaluated the impact of patient infectiousness on cluster growth using site of disease, sputum smear results, and presence or absence of cavitary lesions on chest radiography. We considered three levels of infectiousness, including most infectious (pulmonary TB, sputum smear positive and/or cavitary

lesions identified through chest radiography), somewhat infectious (pulmonary TB but not sputum smear positive and no cavitary lesions), and not infectious (extrapulmonary TB). Additionally, we evaluated drug resistance as the presence of isoniazid and/or rifampin resistance.

We performed bivariate analyses for each outcome of interest – cluster growth and assigned priority level – to obtain unadjusted odds ratios. Furthermore, we identified statistically significant predictors to include in the multivariate models by performing backwards elimination. We controlled for confounding by DTBE priority assignment in the cluster growth multivariate model. Statistical analyses were evaluated an alpha level of 0.05.

Ethics Statement

This analysis was performed on data that are used for public health practice and are collected as part of routine TB surveillance. CDC deemed this analysis not to be human subjects research, and therefore, does not require IRB approval. Analyses were performed using SAS version 9.4 (SAS Institute, Inc.). Figures were developed using R statistical software (*version 4.2.1*).

Results

During 2013–2019, a total of 63,920 verified cases of *M. tuberculosis* were reported to the NTSS by the 50 U.S. states and the District of Columbia. Of these, 48,191 (75.4%) had genotyping results available and were eligible to be included in a cluster alert. During 2016–2019, 816 GENType-based clusters alerted in TB GIMS (Figure 1). Of these, 169 (20.7%) were

re-alerts that were excluded from the analysis. The remaining 647 unique cluster alerts contained 1,970 cases. The median number of cases per alert was 3 cases with a range of 2–14 cases. Among the 647 cluster alerts, DTBE assigned 115 (17.8%) as Priority 1, 155 (24.0%) as Priority 2, 265 (40.1%) as Priority 3, and 112 (17.3%) as Priority 4.

Cluster growth by DTBE priority assignment

Of the 270 clusters assigned a high priority (Priority 1 and Priority 2), 131 (48.5%) gained at least one additional case in the two years following a cluster alert. Of the 377 clusters assigned a low priority (Priority 3 and Priority 4), 122 (32.4%) gained at least one additional case in the subsequent two-year period. Sixty-one (53.0%) Priority 1 clusters, 70 (45.2%) Priority 2 clusters, 100 (39.5%) Priority 3 clusters, and 22 (19.6%) Priority 4 clusters added ≥1 case within the subsequent two-year period of the alert (Figure 2a). The difference in the proportion of clusters that grew was significantly different by grouped (high vs low, p-value<0.0001) and individual (Priority 1-4, p-value<0.0001) DTBE priority assignment.

Among clusters that grew, the median number of additional cases was 2 in Priority 1 clusters (interquartile range [IQR]: 1–4 cases), 2 in Priority 2 clusters (IQR: 1–3), 1.5 in Priority 3 clusters (IQR: 1–3) and 1 in Priority 4 clusters (IQR: 1–2) (Figure 2b). The median number of cases added to a cluster significantly differed by DTBE priority assignment (p-value < 0.0001). Additionally, the difference in the median number of additional cases between cluster Priorities 1 and 3, and all priority assignments vs Priority 4 assigned clusters were significant (p-value<0.0001).

Characteristics of TB cases in clusters by DTBE prioritization assignment and cluster growth

Cases in clusters that DTBE assigned as high priority were more likely to be U.S.-born (71.1%), non-Hispanic Black (41.4%), and have HIV coinfection (7.2%) as compared with cases in clusters that DTBE assigned as low priority (58.4%, 29.6%, and 4.2%, respectively) (Table 1). Social risk factors, including substance use in the year before diagnosis (37.3%), homelessness within the year before diagnosis (14.7%), and incarceration at diagnosis (7.8%) were more commonly reported among cases in clusters that DTBE assigned as high priority than those in clusters that DTBE assigned as low priority (27.1%, 7.4%, and 2.2%, respectively). Additionally, cases in clusters that DTBE assigned as high priority were more likely to have any isoniazid or rifampin resistance (8.6%) as compared with cases in clusters that DTBE assigned as low priority (4.2%).

Cases in clusters that experienced growth were more likely to be male (70.3%), U.S.-born (67.2%), non-Hispanic Black (39.3%), Hispanic (33.9%), and non-Hispanic Native Hawaiian/Other Pacific Islander (3.4%) as compared with cases in clusters that did not gain additional cases (63.2%, 61.9%, 32.4%, 26.5%, and 2.0%, respectively). Social risk factors, including any substance use (37.7%), homelessness within the year before diagnosis (13.6%), and incarceration at diagnosis (7.0%) were more prevalent among cases in clusters that grew as compared with cases in clusters that did not gain an additional case (26.6%, 8.4%, and 3.0%, respectively).

Bivariate statistical analyses

Patient characteristics associated with high DTBE priority assignment and cluster growth include U.S. birth [unadjusted odds ratio [OR] (95% CI) =1.7 (1.4, 2.1); 1.3 (1.1, 1.5)], non-Hispanic Black race [1.8 (1.4, 2.4); 1.8 (1.4, 2.3)], homelessness within the year before diagnosis

[2.2 (1.6, 2.9); 1.7 (1.3, 2.3)], incarceration at diagnosis [3.6 (2.2, 5.7); 2.4 (1.6, 3.7)], reported substance use [1.6 (1.3, 2.0); 1.7 (1.4, 2.1)], and HIV coinfection [1.7 (1.2, 2.6); 2.0 (1.3, 3.0)] (Table 1).

Any isoniazid or rifampin resistance was significantly associated with high priority assignment [2.2 (1.5, 3.2)] but not cluster growth [0.8 (0.6, 1.2)]. Alternatively, male sex [1.4 (1.1, 1.6)], Hispanic ethnicity [1.8 (1.4, 2.4)], and Native Hawaiian or Pacific Islander racial/ethnic background [2.7 (1.4, 5.2)] are characteristics significantly associated with cluster growth but not high priority assignment [1.1 (1.0, 1.4); 1.1 (0.9, 1.5); and 0.6 (0.3, 1.3), respectively].

Multivariate statistical analyses

Results from these multivariate analyses (Table 2) showed U.S. origin of birth [adjusted OR (95% CI) =1.6 (1.2, 2.0) and 1.3 (1.0, 1.7)], non-Hispanic Black race [1.9 (1.4, 2.5) and 1.7 (1.2, 2.3)], and any substance use [1.3 (1.1, 1.7) and 1.5 (1.2, 1.9)] to be significantly associated with high priority assignment and cluster growth, respectively (Table 2).

Hispanic ethnicity [2.4 (1.7, 3.3)] and non-Hispanic Native Hawaiian/Other Pacific Islander [4.6 (2.4, 8.8)] were significantly associated with cluster growth but were not significantly associated with DTBE priority assignment.

Additionally, people living with HIV had higher odds of being in a cluster with growth compared to HIV-negative individuals [1.72 (1.13, 2.63)]. While this factor was a significant predictor in cluster priority assignment, it was not a significant predictor in a multivariate model for DTBE priority level assignment.

Discussion

This retrospective cohort study is the first study to analyze cluster growth following DTBE cluster prioritization. DTBE prioritizes alerted TB clusters on a weekly basis and the assigned cluster priority informs interaction of DTBE staff with local partners to control *M. tuberculosis* transmission.[14] Therefore, DTBE cluster priority assignment is a defining process in the control of TB clusters nationally.

Additionally, while previous studies have assessed patient demographic, clinical, and social risk factors associated with resulting TB outbreaks, this study identified patient demographic, clinical, and social risk factors associated with clusters attaining at least one additional case in the two years following a cluster alert in conjunction with cluster priority assessment.[8,22] While preventing and controlling TB outbreaks must remain a key concern for DTBE staff, outbreaks are rare events and most public health work is related to individual case management and small clusters.[21,27] Notably, most clusters do not gain additional cases following a cluster alert. Understanding patient risk factors associated with clusters attaining at least one additional case in the two years following a cluster alert that are under-captured by DTBE priority assignment can help inform prioritization and lead to better collaboration with local partners to control TB clusters.

Overall, DTBE cluster priority assignment appears to be successful, since a greater portion of clusters assigned high priority grew by at least one case in the two years following alert compared to clusters assigned low priority. The proportion of clusters that gained at least one additional case and the number of cases added to a cluster in the two-year follow-up were significantly different by DTBE priority assignment (p-value<0.0001). Notably, DTBE excelled

at identifying clusters with low likelihood of attaining an additional case, as <20% of clusters assigned Priority 4 grew in the following two years. In comparison, more than 50% of clusters assigned Priority 1 attained additional cases in the two-year follow up. Furthermore, clusters assigned low priority by DTBE that did result in cluster growth gained fewer median number of cases in the two years following a cluster alert compared to cluster assigned high priority. These results highlight the overall success of the established program for TB cluster evaluation in the U.S. However, not all clusters that gained at least one additional case in the following two years were assigned a priority level aligning with resulting cluster growth outcome, indicating that there remain opportunities for improvement in the cluster prioritization process. Some low priority clusters gained many additional cases following a cluster alert, notably one cluster assigned Priority 4 that gained 15 additional cases in a two-year follow-up. While this appears concerning, some GENTypes are common and not all cases added to a cluster alert are attributable to direct transmission between cases within a cluster. Higher resolution genotyping (i.e., wgMLSType) provides advanced resolution to discern such clusters.

Bivariate associations of patient characteristics with DTBE prioritization level and cluster growth provide unadjusted effect measures that can be compared across models to assess statistical significance of unadjusted associations. While certain characteristics had discordant statistical significance between the two models in the bivariate analysis, none had statistically significant, directionally opposite associations, signaling general agreement between characteristics associated with high DTBE prioritization level and TB clusters that gained at least one additional case during the two years following a cluster alert.

While most characteristics associated with the multivariate cluster growth model are captured in the DTBE priority assignment model, HIV coinfection was a characteristic that was

not significant in the DTBE priority assignment model but was a significant predictor for cluster growth. HIV coinfection is a major risk factor for progression to TB disease and increases risk for poor health outcomes.[26] Therefore, DTBE typically considers newly diagnosed TB cases in people living with HIV to be more likely due to recent TB infection.[14,26] While DTBE staff do consider patient HIV status during cluster priority assignment, additional weight may need to be given to this factor when assessing for likelihood of cluster growth.

Despite significant associations between cluster growth and both Hispanic and Native Hawaiian/other Pacific Islander race/ethnicity, these characteristics were not significantly associated with high DTBE prioritization. This discrepancy is likely attributable to common GENType strains of TB and low TB strain variation among these populations. Conventional genotyping methods using MIRU-VNTR have been shown to have varying performance based on *M. tuberculosis* strain lineage.[15] While attainment of additional cases among clusters with *M. tuberculosis* strain lineage one (L1) could indicate recent transmission, further analysis with more to discriminatory genotyping methods (i.e., wgMLSType) could provide different results. Further, since universal WGS became available in 2018, DTBE staff have utilized such higher resolution genotyping to assess likelihood of recent transmission between cases, which may result in lower DTBE priority level assignment, particularly among clusters with L1 lineage.

Any isoniazid or rifampin resistance was positively associated with high DTBE priority assignment but was not associated with cluster growth in a multivariate model. Drug resistance, particularly to the drug rifampin, is a major concern due to increased difficulty of successful treatment.[4,28] Therefore, local programs may initiate a thorough responses to such clusters, regardless of DTBE intervention, leading to reduced transmission of drug-resistant strains. Although any isoniazid or rifampin drug resistance was not associated with cluster growth, drug-

resistant TB remains a concern due to public health implications and should continue to be prioritized by DTBE staff.

Homelessness and incarceration at diagnosis have previously been identified as risk factors for TB outbreaks.[7,8,21,22,27] Persons experiencing homelessness or incarceration spend more time in congregate settings, which are often poorly ventilated and crowded with people with elevated risk of TB, and may yield unsuccessful contact evaluations.[21,29] Such conditions are ideal for the spread of an airborne pathogen. While both homelessness and incarceration at diagnosis were significant predictors for DTBE priority assignment level, neither homelessness nor incarceration at diagnosis were significantly associated with cluster growth in a multivariate model. Previous studies have identified homelessness and incarceration at diagnosis as risk factors in the context of large outbreaks (defined as ≥ six cases).[7,8] These risk factors were likely not significant in our multivariate model because large outbreaks are rare events and our analysis focused more broadly on clusters attaining at least one additional case in the two years following a cluster alert.

This analysis has multiple limitations. First, this analysis excludes non-genotyped cases of TB that may be involved in recent transmission. TB genotypes require cultured *M. tuberculosis* isolates from patient samples. Clinically diagnosed cases, often including small children and patients with extrapulmonary TB, typically do not provide samples and are not genotyped. Furthermore, patients with pulmonary TB who refuse to provide a sputum sample remain not genotyped. These cases may not be counted for initial alerts or enumerated in cases added to a cluster in the two-year follow up. Additionally, conventional genotyping methods may overestimate clustering due to lower molecular resolution of *M. tuberculosis* genomes compared to whole-genome sequence-based methods. Because wgMLSType has greater

molecular resolution than conventional genotyping methods for assessing the likelihood of recent transmission among cases, we hope to perform a similar analysis with wgMLSType for clusters starting in 2018 and onward once additional data becomes available.

Furthermore, the bivariate and multivariate analyses were performed at the individual case level and do not account for clustering of patient characteristics within TB clusters. Using methods that do adjust for clustering of patient characteristics, we would expect point estimates to mirror values obtained in this analysis but that confidence intervals would be larger, which would result in a further reduced multivariate model due to loss of statistical significance. Future analyses will account for patient characteristic clustering.

Conclusion

We identified bivariate associations of patient characteristics and DTBE cluster priority assignment and cluster growth. HIV coinfection was significantly associated with cluster growth in a multivariate model but was not a statistically significant predictor in a multivariate model for DTBE cluster priority assignment. Future analyses that adjust for clustering of patient characteristics and utilize whole-genome sequencing data will be necessary to validate this finding. Overall, the DTBE cluster prioritization process should consider giving additional weight to HIV coinfection when assigning TB cluster priority.

References

1. Filardo TD, Feng PJ, Pratt RH, Price SF, Self JL. Tuberculosis - United States, 2021. MMWR Morb Mortal Wkly Rep 2022;71(12):441-446. DOI: 10.15585/mmwr.mm7112a1.
2. Dowdle WR, Centers for Disease C. A strategic plan for the elimination of tuberculosis in the United States. MMWR Suppl 1989;38(3):1-25. (https://www.ncbi.nlm.nih.gov/pubmed/2495432).
3. Kammerer JS, Shang N, Althomsons SP, Haddad MB, Grant J, Navin TR. Using statistical methods and genotyping to detect tuberculosis outbreaks. Int J Health Geogr 2013;12:15. DOI: 10.1186/1476-072X-12-15.
4. Frieden TR, Sterling TR, Munsiff SS, Watt CJ, Dye C. Tuberculosis. Lancet 2003;362(9387):887-99. DOI: 10.1016/S0140-6736(03)14333-4.
5. Ghosh S, Moonan PK, Cowan L, Grant J, Kammerer S, Navin TR. Tuberculosis genotyping information management system: enhancing tuberculosis surveillance in the United States. Infect Genet Evol 2012;12(4):782-8. DOI: 10.1016/j.meegid.2011.10.013.
6. Walter ND, Painter J, Parker M, et al. Persistent latent tuberculosis reactivation risk in United States immigrants. Am J Respir Crit Care Med 2014;189(1):88-95. DOI: 10.1164/rccm.201308-1480OC.
7. Althomsons SP, Kammerer JS, Shang N, Navin TR. Using routinely reported tuberculosis genotyping and surveillance data to predict tuberculosis outbreaks. PLoS One 2012;7(11):e48754. DOI: 10.1371/journal.pone.0048754.
8. Wortham JM, Li R, Althomsons SP, Kammerer S, Haddad MB, Powell KM. Tuberculosis Genotype Clusters and Transmission in the U.S., 2009-2018. Am J Prev Med 2021;61(2):201-208. DOI: 10.1016/j.amepre.2021.02.006.
9. DTBE strategic plan. Centers for Disease Control and Prevention. . Centers for Disease Control and Prevention. .
10. National Tuberculosis Controllers A, Centers for Disease C, Prevention. Guidelines for the investigation of contacts of persons with infectious tuberculosis. Recommendations from the National Tuberculosis Controllers Association and CDC. MMWR Recomm Rep 2005;54(RR-15):1-47. (https://www.ncbi.nlm.nih.gov/pubmed/16357823).
11. Lindquist S, Allen S, Field K, et al. Prioritizing tuberculosis clusters by genotype for public health action, Washington, USA. Emerg Infect Dis 2013;19(3):493-6. DOI: 10.3201/eid1903.121453.
12. Teeter LD, Vempaty P, Nguyen DT, et al. Validation of genotype cluster investigations for Mycobacterium tuberculosis: application results for 44 clusters from four heterogeneous United States jurisdictions. BMC Infect Dis 2016;16(1):594. DOI: 10.1186/s12879-016-1937-9.
13. France AM, Grant J, Kammerer JS, Navin TR. A field-validated approach using surveillance and genotyping data to estimate tuberculosis attributable to recent transmission in the United States. Am J Epidemiol 2015;182(9):799-807. DOI: 10.1093/aje/kwv121.
14. Prevention CfDCa. Prioritizing tuberculosis genotype clusters for further investigation & public health action. August 2017

(https://www.cdc.gov/tb/programs/genotyping/Prioritizing_Tuberculosis_Genotype_Clusters_August2017.pdf).
15. 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 (Edinb) 2017;106:9-15. DOI: 10.1016/j.tube.2017.05.003.
16. Oelemann MC, Diel R, Vatin V, et al. Assessment of an optimized mycobacterial interspersed repetitive- unit-variable-number tandem-repeat typing system combined with spoligotyping for population-based molecular epidemiology studies of tuberculosis. J Clin Microbiol 2007;45(3):691-7. DOI: 10.1128/JCM.01393-06.
17. Moonan PK, Ghosh S, Oeltmann JE, Kammerer JS, Cowan LS, Navin TR. Using genotyping and geospatial scanning to estimate recent mycobacterium tuberculosis transmission, United States. Emerg Infect Dis 2012;18(3):458-65. DOI: 10.3201/eid1803.111107.
18. Roetzer A, Diel R, Kohl TA, et al. Whole genome sequencing versus traditional genotyping for investigation of a Mycobacterium tuberculosis outbreak: a longitudinal molecular epidemiological study. PLoS Med 2013;10(2):e1001387. DOI: 10.1371/journal.pmed.1001387.
19. Gunasekera KS, Zelner J, Becerra MC, et al. Children as sentinels of tuberculosis transmission: disease mapping of programmatic data. BMC Med 2020;18(1):234. DOI: 10.1186/s12916-020-01702-x.
20. Raz KM, Talarico S, Althomsons SP, et al. Molecular surveillance for large outbreaks of tuberculosis in the United States, 2014-2018. Tuberculosis (Edinb) 2022;136:102232. DOI: 10.1016/j.tube.2022.102232.
21. Mindra G, Wortham JM, Haddad MB, Powell KM. Tuberculosis Outbreaks in the United States, 2009-2015. Public Health Rep 2017;132(2):157-163. DOI: 10.1177/0033354916688270.
22. Haddad MB, Mitruka K, Oeltmann JE, Johns EB, Navin TR. Characteristics of tuberculosis cases that started outbreaks in the United States, 2002-2011. Emerg Infect Dis 2015;21(3):508-10. DOI: 10.3201/eid2103.141475.
23. Oeltmann JE, Kammerer JS, Pevzner ES, Moonan PK. Tuberculosis and substance abuse in the United States, 1997-2006. Arch Intern Med 2009;169(2):189-97. DOI: 10.1001/archinternmed.2008.535.
24. Stewart RJ, Raz KM, Burns SP, et al. Tuberculosis Outbreaks in State Prisons, United States, 2011-2019. Am J Public Health 2022;112(8):1170-1179. DOI: 10.2105/AJPH.2022.306864.
25. Lai CC, Hsieh YC, Yeh YP, et al. A pulmonary tuberculosis outbreak in a long-term care facility. Epidemiol Infect 2016;144(7):1455-62. DOI: 10.1017/S0950268815002265.
26. Aaron L, Saadoun D, Calatroni I, et al. Tuberculosis in HIV-infected patients: a comprehensive review. Clin Microbiol Infect 2004;10(5):388-98. DOI: 10.1111/j.1469-0691.2004.00758.x.
27. Mitruka K, Oeltmann JE, Ijaz K, Haddad MB. Tuberculosis outbreak investigations in the United States, 2002-2008. Emerg Infect Dis 2011;17(3):425-31. DOI: 10.3201/eid1703.101550.
28. WHO. Global tuberculosis report 2022. Geneva, Switzerland: World Health Organization, 2022.

29. Connors WJ, Hussen SA, Holland DP, Mohamed O, Andes KL, Goswami ND. Homeless shelter context and tuberculosis illness experiences during a large outbreak in Atlanta, Georgia. Public Health Action 2017;7(3):224-230. DOI: 10.5588/pha.17.0040.

Tables and Figures

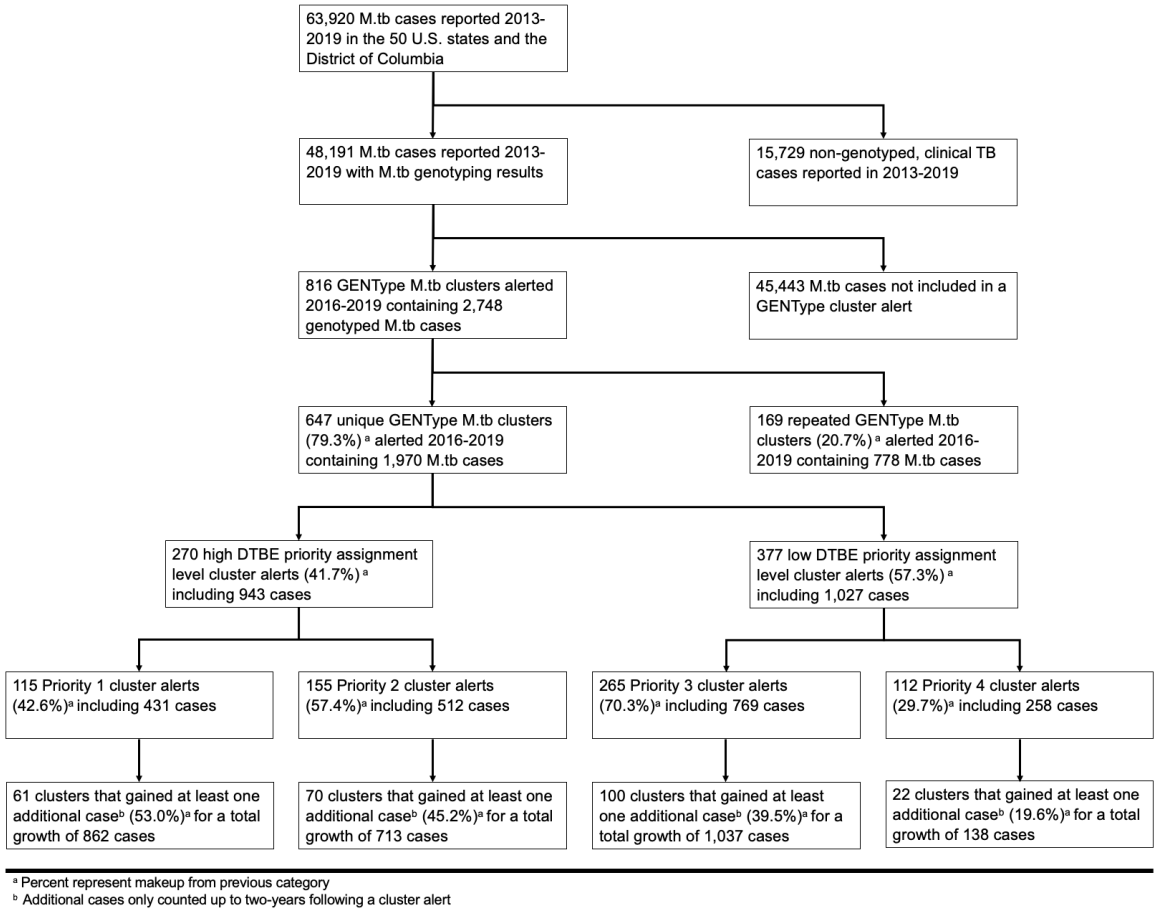

Figure 1: Flow diagram of cases and cluster alerts by DTBE priority assessment and cluster growth and inclusion in analysis

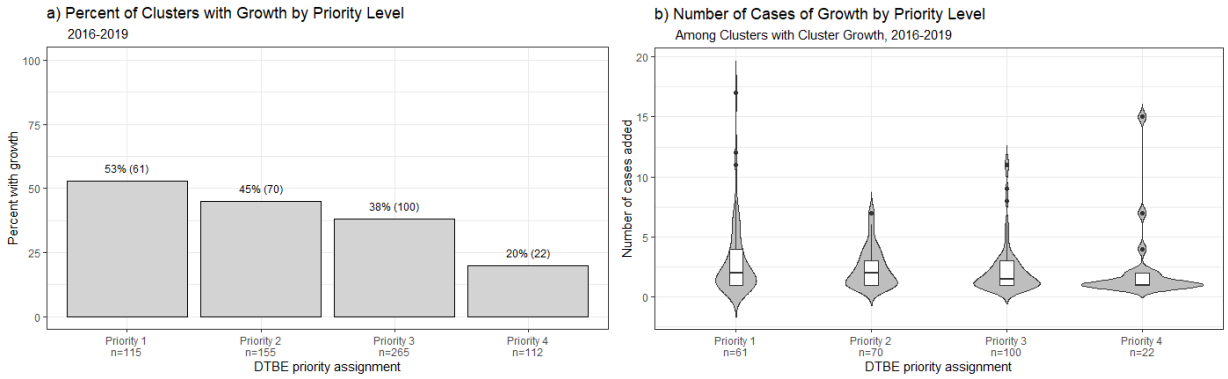

Figure 2: Differences in cluster growth by DTBE priority assignment. a) Number and proportion of clusters that experienced growth of ≥1 additional case in the two years following an alert, stratified by DTBE priority assignment. b) Distribution in the number of cases added to each cluster in the two years following an alert among clusters that experienced growth, stratified by DTBE priority assignment.

Table 1: Frequency of patient characteristics among TB cases in cluster alerts, stratified by DTBE priority assignment (high vs. low priority) and whether the cluster experienced growth of ≥1 case in the two years following the alert (cluster growth vs. no cluster growth). Bivariate associations and 95% confidence intervals for patient characteristics and DTBE priority assignment and whether a cluster experienced growth are presented.

	DTBE Priority Assignment[A]			**Cluster Growth[B]**		
Patient Characteristics	**High Priority** N=943 (47.9%)	**Low Priority** N=1027 (52.1%)	**OR (95% CI)**	**Cluster Growth** N=955 (48.5%)	**No Cluster Growth** N=1015 (51.5%)	**OR (95% CI)**
Demographic Characteristics						
U.S.-born	670 (71.1)	600 (58.4)	**1.7 (1.4, 2.1)**	642 (67.2)	628 (61.9)	**1.3 (1.1, 1.5)**
Male	642 (60.1)	670 (65.2)	1.1 (1.0, 1.4)	671 (70.3)	641 (63.2)	**1.4 (1.1, 1.6)**
Age, years						
<5	38 (4.0)	27 (2.6)	1.2 (0.7, 2.1)	25 (2.6)	40 (3.9)	**0.6 (0.3, 1.0)**
5-14	16 (1.7)	21 (2.0)	0.7 (0.4, 1.4)	12 (1.3)	25 (2.5)	**0.5 (0.2, 0.9)**
15-24	117 (12.4)	132 (12.9)	0.8 (0.60, 1.1)	104 (10.9)	145 (14.3)	**0.7 (0.5, 0.9)**
25-44	359 (38.1)	316 (30.8)	Ref (1)	353 (37.0)	322 (31.7)	Ref (1)
45-64	319 (33.8)	388 (37.8)	**0.7 (0.6, 0.9)**	363 (38.0)	344 (33.9)	1.0 (0.8, 1.2)
≥65	94 (10.0)	143 (13.9)	**0.6 (0.4, 0.8)**	98 (10.3)	139 (13.7)	**0.6 (0.5, 0.9)**
Race/Ethnicity						
American Indian/Alaska Native	48 (5.1)	43 (4.2)	1.6 (1.0, 2.5)	32 (3.4)	59 (5.8)	0.8 (0.5, 1.3)
Asian	86 (9.1)	129 (12.6)	0.9 (0.7, 1.3)	64 (6.7)	151 (14.9)	**0.6 (0.4, 0.9)**
Black	390 (41.4)	304 (29.6)	**1.8 (1.4, 2.4)**	375 (39.3)	319 (31.4)	**1.8 (1.4, 2.3)**
Hispanic	263 (27.9)	330 (32.1)	1.1 (0.9, 1.5)	324 (33.9)	269 (26.5)	**1.8 (1.4, 2.4)**
Multiple Races	8 (0.9)	10 (1.0)	1.1 (0.4, 3.0)	8 (0.8)	10 (1.0)	1.2 (0.5, 3.1)
Native Hawaiian/Other Pacific Islander	20 (2.1)	32 (3.1)	0.6 (0.3, 1.3)	32 (3.4)	20 (2.0)	**2.7 (1.4, 5.2)**
White	126 (13.4)	175 (17.0)	Ref (1)	120 (12.6)	181 (17.8)	Ref (1)
Social Risk Factors						
Any substance use						
Yes	352 (37.3)	278 (27.1)	**1.6 (1.3, 2.0)**	360 (37.7)	270 (26.6)	**1.7 (1.4, 2.1)**
Unknown	11 (1.2)	9 (0.9)	1.6 (0.6, 3.8)	16 (1.7)	4 (0.4)	**5.1 (1.7, 15.4)**
Homelessness within year before diagnosis						
Yes	139 (14.7)	76 (7.4)	**2.2 (1.6, 2.9)**	130 (13.6)	85 (8.4)	**1.7 (1.3, 2.3)**
Unknown	6 (0.6)	3 (0.29)	2.4 (0.6, 9.5)	5 (0.5)	4 (0.4)	1.4 (0.4, 5.3)
Incarcerated at diagnosis	74 (7.8)	23 (2.2)	**3.6 (2.2, 5.7)**	67 (7.0)	30 (3.0)	**2.4 (1.6, 3.7)**
Resident of long-term care facility at diagnosis	16 (1.7)	12 (1.2)	1.5 (0.7, 3.1)	13 (1.4)	15 (1.5)	0.9 (0.4, 1.9)

Primary occupation at diagnosis						
Healthcare worker	25 (2.7)	18 (1.8)	1.5 (0.8, 2.8)	20 (2.1)	23 (2.3)	0.9 (0.5, 1.7)
Migrant worker	6 (0.6)	27 (2.6)	**0.2 (0.1, 0.6)**	11 (1.2)	22 (2.2)	0.5 (0.3, 1.1)
Other	894 (94.8)	962 (93.7)	Ref (1)	909 (95.2)	947 (93.3)	Ref (1)
Unknown	18 (1.9)	20 (2.0)	1.0 (0.5, 1.8)	15 (1.6)	23 (2.3)	0.7 (0.4, 1.3)
Clinical Characteristics						
Clinical markers of infectiousness [C]						
Pulmonary, sputum-smear positive and/or cavitary TB disease	352 (36.9)	336 (33.1)	1.1 (0.8, 1.5)	339 (36.0)	349 (34.0)	0.9 (0.6, 1.3)
Pulmonary, sputum-smear negative, non-cavitary TB disease	522 (54.7)	610 (60.1)	1.0 (0.7, 1.4)	533 (56.5)	599 (58.3)	0.7 (0.5, 1.0)
Extrapulmonary TB disease only						
History of previous TB diagnosis	37 (3.9)	57 (5.6)	0.7 (0.4, 1.0)	52 (5.5)	42 (4.1)	1.3 (0.9, 2.0)
HIV coinfection						
Yes	68 (7.2)	43 (4.2)	**1.7 (1.2, 2.6)**	72 (7.5)	39 (3.8)	**2.0 (1.3, 3.0)**
Unknown	56 (5.9)	81 (7.9)	0.8 (0.5, 1.1)	53 (5.6)	84 (8.3)	**0.7 (0.5, 1.0)**
Other immunocompromising condition	41 (4.4)	44 (4.3)	1.0 (0.7, 1.6)	40 (4.2)	45 (4.4)	0.9 (0.6, 1.5)
Organ transplant recipient	4 (0.4)	7 (0.7)	0.6 (0.2, 2.1)	5 (0.5)	6 (0.6)	0.9 (0.3, 2.9)
Diabetic	143 (15.2)	180 (17.5)	0.8 (0.7, 1.1)	165 (17.3)	158 (15.6)	1.1 (0.9, 1.4)
End-stage renal disease	11 (1.2)	22 (2.1)	0.5 (0.3, 1.1)	14 (1.5)	19 (1.9)	0.8 (0.4, 1.6)
Deceased at TB diagnosis or during treatment	93 (9.9)	90 (8.8)	1.1 (0.8, 1.5)	90 (9.4)	93 (9.2)	1.0 (0.8, 1.4)
Any Isoniazid or Rifampin resistance	81 (8.6)	43 (4.2)	**2.2 (1.5, 3.2)**	54 (5.7)	70 (6.9)	0.8 (0.6, 1.2)
Epidemiologic Risk Factors						
Contact with an infectious TB patient	209 (22.2)	245 (23.9)	0.9 (0.7, 1.1)	173 (18.1)	281 (27.7)	**0.6 (0.5, 0.7)**
Missed contact of identified TB patient	38 (4.0)	28 (2.7)	1.5 (0.9, 2.5)	37 (3.9)	29 (2.9)	1.4 (0.8, 2.3)
Incomplete treatment of latent TB infection	39 (4.1)	34 (3.3)	1.3 (0.8, 2.0)	32 (3.4)	41 (4.0)	0.8 (0.5, 1.3)
Epidemiologically linked to another TB case	232 (24.6)	225 (21.9)	1.2 (0.9, 1.4)	194 (20.3)	263 (25.9)	**0.7 (0.6, 0.9)**
Completion of treatment						
Treatment completed	810 (85.9)	893 (87.0)	Ref (1)	815 (85.3)	888 (87.5)	Ref (1)
Treatment incomplete	36 (3.8)	35 (3.4)	0.9 (0.6, 1.4)	44 (4.6)	27 (2.7)	**1.8 (1.1, 2.9)**
Incomplete due to death	69 (7.3)	64 (6.2)	0.8 (0.6, 1.2)	62 (6.5)	70 (7.0)	1.0 (0.7, 1.4)
Unknown	28 (3.0)	35 (3.4)	1.1 (0.7, 1.9)	34 (3.6)	29 (2.9)	1.3 (0.8, 2.1)

* Counts and percentages provided are by column (i.e., DTBE Priority Assignment level, Cluster Growth outcome).

* Bolded odds ratios and associated 95% confidence intervals indicate a statistically significant bivariate association between the variable level and outcome (DTBE Priority Assignment or Cluster Growth).

[A] DTBE Priority Assignment levels indicated are High Priority, which contains DTBE priority assignments of Priority 1 and Priority 2, and Low Priority, which contains DTBE priority assignments of Priority 3 and Priority 4.

[B] Cluster Growth is defined as the addition of ≥1 cases in the two years following a cluster alert.

[C] Clinical Markers of Infectiousness is a variable that considers a combination of the three variables used to assess infectiousness of a TB case: pulmonary TB, sputum smear, and cavitary disease. Highly Infectious included pulmonary disease, sputum-smear positive and/or cavitary TB. Moderately Infectious included pulmonary disease, sputum-smear negative and non-cavitary TB. Non-Infectious included cases with extrapulmonary TB

Table 2: Adjusted associations between patient characteristics and DTBE priority assignment (high vs. low priority) and whether the cluster experienced growth of ≥1 case in the two years following the alert (cluster growth vs. no cluster growth). Adjusted odds ratios and 95% confidence intervals are presented.

	DTBE Priority Assignment[A] (High vs Low)	Cluster Growth[B] (Yes vs No)
Patient Characteristics	**Adjusted OR (95% CI)**	**Adjusted OR (95% CI)**
Demographic Characteristics		
U.S.-born	**1.6 (1.2, 2.0)**	**1.3 (1.0, 1.7)**
Age, years		
<5	1.3 (0.7, 2.2)	-
5-14	0.8 (0.4, 1.6)	-
15-24	0.8 (0.6, 1.1)	-
25-44	Ref (1.0)	-
45-64	**0.7 (0.5, 0.8)**	-
≥65	0.6 (0.5, 0.9)	-
Race/Ethnicity		
American Indian/Alaska Native	**1.8 (1.1, 2.9)**	0.9 (0.5, 1.5)
Asian	**1.6 (1.0, 2.4)**	1.0 (0.6, 1.5)
Black	**1.9 (1.4, 2.5)**	**1.7 (1.3, 2.3)**
Hispanic	1.4 (1.0, 1.9)	**2.4 (1.7, 3.3)**
Multiple Races	1.3 (0.5, 3.6)	1.5 (0.6, 4.1)
Native Hawaiian/Other Pacific Islander	1.2 (0.6, 2.3)	**4.6 (2.4, 8.8)**
White	Ref (1.0)	Ref (1.0)
Social Risk Factors		
Any substance use		
Yes	**1.3 (1.1, 1.7)**	**1.5 (1.2, 1.9)**
Unknown	1.0 (0.4, 2.9)	**6.3 (1.8, 22.2)**
Homelessness within year before diagnosis		
Yes	**1.9 (1.4, 2.7)**	-
Unknown	2.8 (0.5, 16.5)	-
Incarcerated at diagnosis	**2.8 (1.7, 4.6)**	-
Clinical Characteristics		
History of previous TB diagnosis	**0.6 (0.4, 0.9)**	-
HIV coinfection		
Yes	-	**1.7 (1.1, 2.5)**
Unknown	-	0.8 (0.5, 1.1)
Any Isoniazid or Rifampin resistance		
Resistance present	**2.5 (1.7, 3.8)**	-
Resistance profile unknown	1.2 (0.6, 2.3)	-
Epidemiologic Risk Factors		
High DTBE priority assignment	*n/a*	**1.8 (1.5, 2.1)**

***** Bolded odds ratios and associated 95% confidence intervals indicate a statistically significant association between the variable level and outcome (DTBE Priority Assignment or Cluster Growth).

[A] DTBE Priority Assignment levels indicated are High Priority, which contains DTBE priority assignments of Priority 1 and Priority 2, and Low Priority, which contains DTBE priority assignments of Priority 3 and Priority 4.

[B] Cluster Growth is defined as the addition of ≥1 cases in the two years following a cluster alert.