HHS Public Access

Author manuscript

Am J Epidemiol. Author manuscript; available in PMC 2017 August 08.

Published in final edited form as:

Am J Epidemiol. 2014 January 15; 179(2): 216–225. doi:10.1093/aje/kwt246.

Estimated Rate of Reactivation of Latent Tuberculosis Infection in the United States, Overall and by Population Subgroup

Kimberly M. Shea[*], **J. Steve Kammerer**, **Carla A. Winston**, **Thomas R. Navin**, and **C. Robert Horsburgh Jr.**

Department of Epidemiology, School of Public Health, Boston University, Boston, Massachusetts (Kimberly M. Shea, C. Robert Horsburgh); Northrup Grumman Information Systems, Atlanta, Georgia (J. Steve Kammerer); and Division of Tuberculosis Elimination, Centers for Disease Control and Prevention, Atlanta, Georgia (Carla A. Winston, Thomas R. Navin)

Abstract

We estimated the rate of reactivation tuberculosis (TB) in the United States, overall and by population subgroup, using data on TB cases and *Mycobacterium tuberculosis* isolate genotyping reported to the Centers for Disease Control and Prevention during 2006–2008. The rate of reactivation TB was defined as the number of non-genotypically clustered TB cases divided by the number of person-years at risk for reactivation due to prevalent latent TB infection (LTBI). LTBI was ascertained from tuberculin skin tests given during the 1999–2000 National Health and Nutrition Examination Survey. Clustering of TB cases was determined using TB genotyping data collected by the Centers for Disease Control and Prevention and analyzed via spatial scan statistic. Of the 39,920 TB cases reported during 2006–2008, 79.7% were attributed to reactivation. The overall rate of reactivation TB among persons with LTBI was estimated as 0.084 (95% confidence interval (CI): 0.083, 0.085) cases per 100 person-years. Rates among persons with and without human immunodeficiency virus coinfection were 1.82 (95% CI: 1.74, 1.89) and 0.073 (95% CI: 0.070, 0.075) cases per 100 person-years, respectively. The rate of reactivation TB among persons with LTBI was higher among foreign-born persons (0.098 cases/100 person-years; 95% CI: 0.096, 0.10) than among persons born in the United States (0.082 cases/100 person-years; 95% CI: 0.080, 0.083). Differences in rates of TB reactivation across subgroups support current recommendations for targeted testing and treatment of LTBI.

Keywords

latent tuberculosis infection; National Health and Nutrition Examination Survey; reactivation tuberculosis; tuberculosis

Despite substantial declines in the incidence of tuberculosis (TB) in the United States since 1993, more than 10,000 cases continue to occur each year (1). These cases may result from recent transmission of *Mycobacterium tuberculosis* or from reactivation of latent TB infection (LTBI), known as reactivation TB. Molecular epidemiologic studies carried out in

[*]Correspondence to Dr. Kimberly M. Shea, Department of Epidemiology, Boston University School of Public Health, 715 Albany Street, Room 322 East, Boston, MA 02118 (kimshea@bu.edu).

some areas of the United States suggest that the decline in TB incidence in these areas has been associated with an increased proportion of cases occurring from reactivation of LTBI (2–4), and an analysis of US TB cases identified during 2005–2009 indicated that nearly 77% of TB cases may be attributable to reactivation (5). Therefore, to continue to reduce TB incidence in the United States, it is important to understand which subgroups of the general population have the highest rates of progression from LTBI to active TB disease.

Some risk factors for reactivation TB have already been established. Human immunodeficiency virus (HIV)-infected persons progress from LTBI to active TB disease at 9 times the rate of HIV-uninfected persons (6). To a lesser extent, other medical risk factors, such as old healed TB disease, chronic renal failure, poorly controlled diabetes, or use of tumor necrosis factor-α-inhibitor therapy, also increase the risk of progression from LTBI to active TB (6). However, these conditions are relatively rare in the general population and are unlikely to account for the majority of reactivation TB cases.

Although some reports have suggested that foreign-born persons in the United States are more likely to have reactivation TB than US-born persons, it is unclear whether higher reactivation among foreign-born persons reflects a higher prevalence of LTBI in this population or an increased rate of progression from LTBI to TB disease, or both (5). In a southeastern US study that examined the rate of reactivation TB among persons with LTBI, Horsburgh et al. (7) concluded that increased progression from LTBI to active TB disease was associated with older age and foreign birth, but it is not clear whether this finding can be generalized to the entire US population. Since a substantial proportion of TB reported in the United States arises from older and foreign-born persons (1), it is important to determine whether the rate of progression from LTBI to TB disease is higher in these groups.

Conducting a cohort study is not practical for measuring the rate of reactivation TB among persons with LTBI in the United States, because TB occurs infrequently and because LTBI treatment, which is offered routinely, further decreases the rate of reactivation. However, we were able to estimate rates of reactivation TB by first using 2006–2008 genotyping and epidemiologic linkage information from the National Tuberculosis Genotyping Service to differentiate between cases of presumed primary TB and cases of reactivation TB. Next, we used population-based tuberculin skin testing (TST) results from the 1999–2000 National Health and Nutrition Examination Survey (NHANES) to estimate the number of US residents with LTBI, who were thus at risk for reactivation TB. Since prevention of reactivation TB by treatment of LTBI is a major goal of the US national strategy for TB elimination (6), identification of population subgroups with LTBI who are at increased risk of reactivation TB will help in targeting intervention strategies.

MATERIALS AND METHODS

Identification of clustered cases of TB

The Centers for Disease Control and Prevention (CDC) has collected information on newly reported cases of TB disease in the United States since 1953. Cases are currently reported to the National Tuberculosis Surveillance System using a standard case report form and a specified TB case definition for laboratory and clinical criteria (1). The National

Tuberculosis Genotyping Service was established in 2004 with the goal of using standard genotyping methods to genotype at least 1 *M. tuberculosis* isolate from each culture-confirmed TB case reported in the United States (8–10). Overall national genotyping surveillance coverage, defined as the proportion of reported culture-positive TB cases with a genotype result, was 66% in 2006 and approximately 80% in 2007 and 2008 (9).

TB cases reported to the National Tuberculosis Surveillance System during 2006–2008 were used in the current study and were linked to genotyping results from the National Tuberculosis Genotyping Service using a standard protocol (9). All TB cases that were not part of a genotype cluster, as defined below, were considered cases of reactivation TB. The underlying assumption for this approach is that TB cases connected by recent transmission should have indistinguishable genotypes and should occur within the same local jurisdiction as another genotypically related case. Conversely, cases due to reactivation of LTBI are expected to have unique genotypes unrelated to other cases identified within the past few years and within the same local jurisdiction (11). TB clusters were identified using methods previously described by Moonan et al. (11), with the exception that we used a single 3-year period in the current project. Briefly, TB genotypes were characterized using a standardized protocol for spacer oligonucleotide typing (spoligotyping) and 12-locus myco-bacterial interspersed repetitive unit–variable number of tandem repeats (MIRU-VNTR) typing. Genotype clusters were defined as at least 2 cases with indistinguishable TB genotypes reported within statistically significant geospatial zones as determined by a spatial scan statistic (SaTScan, Martin Kulldorff, Boston, Massachusetts; and Information Management Services, Inc., Silver Spring, Maryland) (12). In purely spatial analysis, we applied a discrete Poisson model assuming that the number of clustered TB cases was Poisson-distributed, with the underlying population at risk being all culture-positive TB cases identified during the project period. Recent transmission was evaluated for a 3-year period during 2006–2008 in an effort to utilize the highest-quality genotyping data (9) as close in time as possible to the NHANES data, while also increasing the probability of capturing sufficient subsequent TB cases for cluster identification (11). We aggregated cases by genotype using zip codes and assessed each genotype separately for evidence of clustering, using SaTScan software (version 8.0.2) to generate possible circular regions of different sizes and locations up to a maximum radius of 50 km. We calculated the likelihood ratio for each region in comparison with all possible regions, with the maximum likelihood ratio representing the circle most likely to identify clustering for the genotype. Monte Carlo simulation was used to determine the distribution of the likelihood ratio scan statistic under the null hypothesis of spatial randomness of TB cases and to generate P values for potential genotype clusters. We defined a significant cluster as one having a P value less than 0.05. Cases of TB that could not be genotyped because they were clinically diagnosed or whose genotypes were unavailable for other reasons were assigned cluster status in proportion to cases from the same risk group (e.g., US-born vs. foreign-born) that had isolates available for genotyping.

Identification of populations with LTBI

Because there are no data describing the prevalence of LTBI in the United States during 2006–2008, the prevalence of LTBI in the United States was estimated using TST results

collected in the 1999–2000 NHANES, a series of cross-sectional health examination surveys representative of the civilian noninstitutionalized US population. NHANES survey methodology has been previously described in detail (13).

NHANES survey data collected during 1999–2000 include demographic and TST results for 7,386 participants aged 1 year or older in whom TST was performed via the Mantoux technique using 0.1 mL of purified protein derivative S-1, the reference standard tuberculin used in the United States (14). The tests were administered by trained NHANES phlebotomists and read by trained NHANES TST readers 48–72 hours after placement (15).

NHANES participants with an induration of 10 mm or greater were classified as having LTBI, regardless of HIV infection status (because HIV status was unknown for NHANES participants). SAS software, version 9.2 (SAS Institute Inc., Cary, North Carolina), was then used to account for the NHANES sample design and weighting through application of methods described by Bennett et al. (16) to adjust for TST nonparticipation. The resulting prevalence of LTBI in the survey population overall and in subgroups of interest was assumed to approximate the prevalence of LTBI during 2006–2008 and was applied to 2006–2008 population estimates from the US Census to obtain estimates of the number of persons in the US population at risk for reactivation TB due to prevalent LTBI during 2006–2008.

Estimation of rates of reactivation TB

The rate of reactivation TB was calculated by dividing the number of nonclustered TB cases by the number of person-years at risk for reactivation TB. Rates were calculated overall and stratified by sex, age group, race/ethnicity, place of birth (US-born or foreign-born), and US Census region. Information about US Census region was not available in the public-use NHANES data set and was therefore accessed through the Research Data Center of the National Center for Health Statistics. Region 1 (Northeast) includes Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, New Jersey, New York, Pennsylvania, and Vermont; Region 2 (Midwest) includes Iowa, Indiana, Illinois, Kansas, Michigan, Minnesota, Missouri, Ohio, Nebraska, North Dakota, South Dakota, and Wisconsin; Region 3 (South) includes Alabama, Arkansas, Delaware, the District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, and West Virginia; and Region 4 (West) includes Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, New Mexico, Nevada, Oregon, Utah, Washington, and Wyoming. Consistent with US Census Bureau definitions, people were classified as foreign-born if they had not been born in the United States or an associated jurisdiction and neither parent was a US citizen. The number of person-years at risk for reactivation TB was calculated by multiplying the estimated number of persons with LTBI by 3 years, the duration of the study period. Confidence intervals for rate estimates were calculated using Wilson's method (17), and unadjusted rate ratios and 95% confidence intervals were calculated using SAS. Reactivation rates were calculated both by the "n" method, in which all nonclustered cases of TB were used in the numerator, and by the "$n - 1$" method, in which the first reported case in each cluster was assumed to be due to

reactivation and was therefore reclassified as nonclustered. Because the results from both methods were consistent, only results from the "*n*" method are presented.

Estimation of rates among HIV-infected persons

Because HIV-infected persons with LTBI have a much higher risk of developing active TB than HIV-uninfected persons, we estimated the rate of reactivation TB separately for HIV-infected and HIV-uninfected US residents. For reported TB cases aged 15–64 years with an unknown HIV infection status, we used Markov chain Monte Carlo multiple-imputation methods to assign HIV infection status based on an imputation model that used the characteristics of patients with known HIV status for each stratum. California was excluded from this analysis because the HIV status of TB cases reported in California was not reported to the CDC during 2006–2008. Next, because NHANES did not include information about HIV status and because the proportion of HIV-infected persons in the United States who have LTBI is unknown, we assumed that LTBI prevalence in HIV-infected persons was the same as the prevalence of LTBI in the US population (Appendix).

Ethical review

This project analyzed data collected for the purposes of routine disease surveillance and disease control and was not considered human subjects research requiring institutional review board approval. The protocol for the TB component of NHANES was reviewed and approved by the NHANES institutional review board.

RESULTS

Cases of reactivation TB

There were 39,920 TB cases reported to the CDC during 2006–2008, corresponding to an annual TB incidence rate of 4.4 per 100,000 US residents. Among 16,662 (42%) cases occurring in US-born persons, 12,710 (76.3%) were culture-positive, and among 23,171 (58%) cases occurring in foreign-born persons, 18,475 (79.7%) were culture-positive. (Place of birth was unknown for 87 cases.) Genotyping results were available for 9,133 (71.9%) culture-positive US-born cases and 13,724 (74.3%) culture-positive foreign-born cases (Table 1); 42 culture-positive cases were missing information on country of birth and were excluded. Comparison of genotyped and nongenotyped cases revealed that genotyped cases had different subpopulation distributions (Table 2); however, except for pediatric cases, in which the majority of TB diagnoses are based on a clinical case definition, the distribution of genotyped and nongenotyped cases generally differed by 5% or less.

Among the 22,857 TB cases with genotypes, 18,308 (80.1%) were not part of a genotype cluster and were attributed to LTBI reactivation. The percentage of TB cases attributed to reactivation differed for US-born (11,672 cases; 70.1%) and foreign-born (20,046 cases; 86.5%) persons. Applying risk-group-specific percentages of nonclustered cases to the number of risk-group-specific TB cases that could not be genotyped yielded an additional 5,274 and 8,173 cases of reactivation TB in US-born and foreign-born persons, respectively. The prevalence of LTBI based on NHANES data was found to vary by sex, age, race/ethnicity, US Census region, and country of birth (Table 3).

Rates of reactivation TB

The overall rate of reactivation TB among US residents with LTBI was estimated to be 0.084 cases per 100 person-years (95% confidence interval (CI): 0.083, 0.085). Table 3 presents the overall rate of reactivation TB stratified by sex, age group, race/ethnicity, and US Census region. Because the rate of reactivation TB among persons with LTBI was significantly higher among foreign-born persons (0.098 cases/ 100 person-years; 95% CI: 0.096, 0.10) than among US-born persons (0.082 cases/100 person-years; 95% CI: 0.080, 0.083), we further stratified rates by birthplace (Table 4).

Estimated rates in HIV-infected residents

Among patients aged 15–64 years not residing in California, there were 10,997 genotyped cases and an estimated 7,769 nongenotyped cases of reactivation TB during 2006– 2008. Among the 10,997 genotyped cases, HIV infection status was positive in 8.9%, negative in 77.1%, and unknown in 14.0%. Imputation indicated that 10.0% of genotyped patients with reactivation TB and 14.1% of nongenotyped patients with reactivation TB were HIV-infected, yielding an estimate of 2,198 HIV-infected cases of reactivation TB. Therefore, the US rate of reactivation TB among HIV-infected persons was estimated to be 1.82 cases per 100 person-years, assuming a 4.2% prevalence of LTBI in the HIV-infected population (Table 5).

DISCUSSION

We observed an overall rate of reactivation TB among persons with LTBI in the United States of approximately 0.084 cases per 100 person-years. This rate is substantially lower than the rates of 0.10–0.16 per 100 person-years observed in the 1950s (6) but higher than the rate of 0.040 observed in a recent population-based study in southern Florida (7). The decline in this rate in the United States over the past 5 decades probably represents elimination of a common high-risk subgroup from the population, namely persons with old, healed but untreated TB disease. These persons are known to have 5 times the reactivation rate of persons with LTBI without old, untreated disease, and such cases are much less common today than in the early chemotherapy era (7). The lower rates reported from the southern Florida study may reflect geographical variability that is partially explained by the high prevalence of infection with non-TB mycobacteria in that region, which can yield a positive test result even when *M. tuberculosis* is not present (18). Such cross-reactivity artificially increases the size of the population “at risk” (since persons with a cross-reactive TST test are “false-positive” for LTBI) and decreases the calculated reactivation rate.

The overall rate of reactivation was also substantially different between US-born (0.082 cases/100 person-years) and foreign-born (0.098 cases/100 person-years) persons. This difference is consistent with what has been observed in other studies (3, 5) and is probably the result of a higher prevalence of risk factors for reactivation, such as old, healed disease, poor nutrition or other immunosuppressive states, or more recent infection, among foreign-born persons. Unfortunately, we did not have individual-level information on these variables for the TST survey, so we could not adjust for these potential cofounders.

However, our results differ from those of a case-control analysis comparing 188 case-patients with 188 matched controls in Boston, Massachusetts, that did not find an increased rate of reactivation among foreign-born persons (19). That analysis differed from ours in terms of study design, population, and sample size. In addition, it is possible that some foreign-born persons with TB, particularly those who have recently arrived in the United States, may be misclassified as having unclustered (reactivation) TB because the related cases of that genotype are outside the geographical limit for clustering (20,21), or that the population of foreign-born persons living in the United States was underestimated in the US Census.

Among US-born persons with LTBI, we observed increased rates of reactivation at both ends of the age spectrum. Young children are known to be at increased risk of reactivation (22), a phenomenon thought to be explained by the immaturity of their immune systems. Older persons are known to have a higher prevalence of reactivation TB, but this has largely been attributed to the higher prevalence of LTBI in the elderly. Our results show that among US-born persons, rates of progression from LTBI to TB disease decline with age. This is consistent with the observations of a recent cohort study in Norway (23).

Among foreign-born persons with LTBI, we observed a low rate of reactivation among young children and increased rates of reactivation with increasing age. A possible reason for lower reactivation among young children could be over-estimation of LTBI in this population due to recent Bacillus Calmette-Guérin vaccination, whereas the increased rate in the oldest age group may be due to recent infections among some foreign-born elderly persons who had recently arrived in the United States.

There was also substantial variability of reactivation rates among racial/ethnic groups: Asians and Hispanics had elevated rates of reactivation compared with non-Hispanic whites among both US-born and foreign-born persons, and foreign-born blacks had elevated rates compared with foreign-born whites. This may be attributable to increased prevalence of old, healed disease, since these subgroups may have had decreased access to TB diagnosis and treatment. Additionally, some of these groups, particularly blacks and Hispanics, may have increased prevalence of HIV infection that was not accounted for in our study.

This analysis had several important limitations. First, there was potential for misclassification of cases of reactivation TB. We used lack of clustering as a proxy for reactivation TB, but there are circumstances in which recently transmitted TB cases may not cluster, such as when related cases are outside the geographical or temporal frame of the analysis. For example, we may have attributed some cases to reactivation because other cases with related genotypes were outside of our 3-year study period. In addition, partial sampling of genotypes may have led to overascertainment of reactivation cases if related cases were not detected because of missing genotyping results. Next, although TB clustering status was determined from all cases reported to the National Tuberculosis Genotyping Service, isolate submission is incomplete and imputation of results from cases without genotyping may have introduced bias. Genotype coverage was consistent for sex, age group, and race/ethnicity subpopulations but was lower in the Northeast in 2006 (45%) as compared with other years and other regions (60%–80%). This may have resulted in overestimation of

the rate of reactivation TB in the Northeast due to not having sufficient coverage to identify clusters. In addition, because we performed a large number of SaTScan runs for hundreds of different genotypes, it is possible that we overestimated the number of statistically significant clusters. This would mean that the reactivation rate we observed was an underestimate of the true rate.

Second, misclassification of LTBI also may have occurred. LTBI prevalence estimated during the 1999–2000 NHANES was probably slightly higher than LTBI prevalence during 2006–2008 (24), producing potential overestimation of rates in our study. However, we do not have reason to expect that our ratio measures comparing population subgroups would have been substantially affected. Misclassification of LTBI status also may have occurred because TST results were interpreted without regard to skin tests for non-TB mycobacteria or prior Bacillus Calmette-Guérin vaccination.

Third, although we were able to calculate the rate of reactivation TB by age group, waning TST reactivity in the elderly might mean that the true rate was even lower than what we observed.

Lastly, we did not have sufficient information in our data set to estimate the amount of reactivation separately for persons with and without diabetes, and our analysis of the amount of reactivation that can be attributed to HIV infection required substantial imputation. In our HIV imputation, we used a range of assumptions, all of which led to the same conclusion—namely that although the effect of HIV on reactivation was large, the effect of HIV on overall rates in the United States was small because of small numbers of HIV-infected persons in the population.

Targeted testing for and treatment of LTBI is a key strategy for elimination of TB in the United States, and the results of this analysis have important implications for the design and implementation of such programs, as well as for cost-effectiveness analyses. A recent study demonstrated that targeting specific populations for screening and treatment was highly cost-effective (25). However, that study assumed that the reactivation rate was constant among all populations. Our results indicate that certain groups, especially the foreign-born and racial and ethnic minorities, have increased rates of progression and will receive even greater benefit from testing and treatment. Conversely, persons in certain geographical areas may have lower rates of progression and may receive less benefit. Since these lower rates may be at least partially attributable to misclassification of LTBI status by skin testing, greater use of interferon-γ release assays for identification of LTBI may remove this source of misclassification. These results reinforce recent recommendations for targeted testing and treatment of LTBI as a cost-effective intervention, especially among foreign-born persons (25).

Acknowledgments

No specific funding was received for the study.

We gratefully acknowledge the staffs of the National Tuberculosis Genotyping Service contract laboratories, local and state public health laboratories, and local and state health departments, who collected the data included in these

analyses. We also acknowledge the staff of the Research Data Center at the National Center for Health Statistics, especially Stephanie Roberts, for assisting us in adding geographical information to our data set.

The findings and conclusions are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention, the Research Data Center, or the National Center for Health Statistics.

Some of the authors were employed by the Centers for Disease Control and Prevention during the conduct of this study.

Abbreviations

CDC	Centers for Disease Control and Prevention
HIV	human immunodeficiency virus
LTBI	latent tuberculosis infection
NHANES	National Health and Nutrition Examination Survey
TB	tuberculosis
TST	tuberculin skin testing

References

1. Centers for Disease Control and Prevention. Reported Tuberculosis in the United States, 2010. Atlanta, GA: Centers for Disease Control and Prevention; 2011.
2. Geng E, Kreiswirth B, Driver C, et al. Changes in the transmission of tuberculosis in New York City from 1990 to 1999. N Engl J Med. 2002; 346(19):1453–1458. [PubMed: 12000815]
3. Sharnprapai S, Miller AC, Suruki R, et al. Genotyping analyses of tuberculosis cases in U.S.- and foreign-born Massachusetts residents. Emerg Infect Dis. 2002; 8(11):1239–1245. [PubMed: 12453348]
4. Cattamanchi A, Hopewell PC, Gonzalez LC, et al. A 13-year molecular epidemiological analysis of tuberculosis in San Francisco. Int J Tuberc Lung Dis. 2006; 10(3):297–304. [PubMed: 16562710]
5. Ricks PM, Cain KP, Oeltmann JE, et al. Estimating the burden of tuberculosis among foreign-born persons acquired prior to entering the U.S., 2005–2009. PLoS One. 2011; 6(11):e27405. [PubMed: 22140439]
6. Horsburgh CR Jr. Priorities for the treatment of latent tuberculosis infection in the United States. N Engl J Med. 2004; 350(20):2060–2067. [PubMed: 15141044]
7. Horsburgh CR Jr, O'Donnell M, Chamblee S, et al. Revisiting rates of reactivation tuberculosis: a population-based approach. Am J Respir Crit Care Med. 2010; 182(3):420–425. [PubMed: 20395560]
8. Notice to readers. new CDC program for rapid genotyping of *Mycobacterium tuberculosis* isolates. MMWR Morb Mortal Wkly Rep. 2005; 54(2):47.
9. Ghosh S, Moonan PK, Cowan L, et al. Tuberculosis genotyping information management system: enhancing tuberculosis surveillance in the United States. Infect Genet Evol. 2012; 12(4):782–788. [PubMed: 22044522]
10. Cowan LS, Diem L, Monson T, et al. Evaluation of a two-step approach for large-scale, prospective genotyping of *Mycobacterium tuberculosis* isolates in the United States. J Clin Microbiol. 2005; 43(2):668–695.
11. Moonan PK, Ghosh S, Oeltmann JE, et al. Using genotyping and geospatial scanning to estimate recent *Mycobacterium tuberculosis* transmission, United States. Emerg Infect Dis. 2012; 18(3): 458–465. [PubMed: 22377473]
12. Kulldorff M. A spatial scan statistic. Commun Stat Theory Methods. 1997; 26(6):1481–1496.

13. National Center for Health Statistics. National Health and Nutrition Examination Survey 1999–2000 General Data Release Documentation. Hyattsville, MD: National Center for Health Statistics; 2012. http://www.cdc.gov/nchs/data/nhanes/nhanes_99_00/general_data_release_doc.pdf [Accessed March 31, 2012]
14. Lee E, Holzman RS. Evolution and current use of the tuberculin test. Clin Infect Dis. 2002; 34(3): 365–370. [PubMed: 11774084]
15. National Center for Health Statistics. National Health and Nutrition Examination Survey Analytic Guidelines. Hyattsville, MD: National Center for Health Statistics; 2010. http://www.cdc.gov/nchs/data/nhanes/nhanes_03_04/nhanes_analytic_guidelines_dec_2005.pdf [Accessed December 7, 2010]
16. Bennett DE, Courval JM, Onorato I, et al. Prevalence of tuberculosis infection in the United States population: the National Health and Nutrition Examination Survey, 1999–2000. Am J Respir Crit Care Med. 2008; 177(3):348–355. [PubMed: 17989346]
17. Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. Stat Med. 1998; 17:857–872. [PubMed: 9595616]
18. Larson EM, O'Donnell M, Chamblee S, et al. Dual skin tests with *M. avium* sensitin and PPD to detect misdiagnosis of latent tuberculosis infection. Int J Tuberc Lung Dis. 2011; 15(11):1504–1509. [PubMed: 22008764]
19. Patel S, Parsyan AE, Gunn J, et al. Risk of progression to active tuberculosis among foreign-born persons with latent tuberculosis. Chest. 2007; 131(6):1811–1816. [PubMed: 17413054]
20. Murray M. Sampling bias in the molecular epidemiology of tuberculosis. Emerg Infect Dis. 2002; 8(4):363–369. [PubMed: 11971768]
21. Glynn JR, Bauer J, de Boer AS, et al. Interpreting DNA fingerprint clusters of *Mycobacterium tuberculosis*. Int J Tuberc Lung Dis. 1999; 3(12):1055–1060. [PubMed: 10599007]
22. Horsburgh, CR., Moore, M., Castro, KG. Epidemiology of tuberculosis in the United States. In: Rom, WN., Garay, SM., editors. Tuberculosis. 2. Philadelphia, PA: Lippincott Williams & Wilkins; 2004. p. 31-45.
23. Wiker HG, Mustafa T, Bjune GA, et al. Evidence for waning of latency in a cohort study of tuberculosis. BMC Infect Dis. 2010; 10:37. [PubMed: 20178619]
24. Winston CA, Navin TR. Birth cohort effect on latent tuberculosis infection prevalence, United States. BMC Infect Dis. 2010; 10:206. [PubMed: 20626871]
25. Linas BP, Freedberg KA, Horsburgh CR. Priorities for screening and treatment of latent tuberculosis in the United States. Am J Respir Crit Care Med. 2011; 184(5):590–601. [PubMed: 21562129]
26. Schafer, JL. Analysis of Incomplete Multivariate Data. London, United Kingdom: Chapman & Hall Ltd; 1997.

APPENDIX

HIV Imputation

The multiple imputation procedure in SAS, version 9.2, was used to impute missing data for all variables that were included in the imputation model (26) (e.g., not limited to human immunodeficiency virus (HIV) infection status) using the Markov chain Monte Carlo (MCMC) method. We established that the pattern of missing data was arbitrary (e.g., nonmonotonic) and assumed that missing values were missing at random. MCMC further assumes that the data are multivariate normal, and it uses Bayesian inference to draw from the posterior distribution of the missing data that is modeled on the observed data (26). A benefit of the MCMC method is that it may be used with arbitrary patterns of missing data. We used 5 imputations to derive our results and used separate models for genotyped and nongenotyped cases of tuberculosis (TB). HIV infection rates were computed for clustered,

nonclustered, and nongenotyped TB case groups, and the mean value of the proportion of TB case-patients who were HIV-infected was estimated from the 5 imputations.

The variables used in the imputation model included age, male sex, Hispanic ethnicity, black race, race other than black or white, foreign birth, cluster status (genotyped only), HIV positivity, homelessness, excess alcohol use, injection drug use, noninjection drug use, and residence in a correctional facility.

Each of the variables listed above was categorized as present or absent (0 or 1), with the exception of age, which was continuous. Homelessness, excess alcohol use, injection drug use, and noninjection drug use were categorized as present if they had been present at any time during the 12 months prior to TB diagnostic evaluation; residence in a correctional facility was based on having been an inmate in a correctional facility at the time of TB diagnostic evaluation.

The estimated number of patients with reactivation TB who were HIV-infected was calculated from the count of cases that were nonclustered (or estimated nonclustered from the nongenotyped data), and the proportion of patients who were HIV-infected was obtained from the imputed data (10.02% for genotyped cases and 14.11% for nongenotyped cases):

$$\text{Genotyped TB cases: } 10,997 \text{ cases} \times 0.1002 = 1,102 \text{ HIV- infected}$$

$$\text{Nongenotyped TB cases: } 7,769 \text{ cases} \times 0.1411 = 1,096 \text{ HIV- infected}$$

Table 1

Characteristics of Tuberculosis Cases Reported to the Centers for Disease Control and Prevention, Overall and by Place of Birth,[a] United States, 2006–2008

	Overall		US-Born		Foreign-Born	
	No.	%	No.	%	No.	%
TB cases reported to the CDC	39,920		16,662		23,171	
Culture-positive cases	31,259	78.3	12,710	76.3	18,475	79.7
Genotyping results available	22,857	73.1	9,133	71.9	13,724	74.3
No. of nonclustered cases among those with genotyping results	18,308	80.0	6,398	70.1	11,873	86.5
Estimated no. and % of reactivation TB cases	31,841	79.7	11,672	70.1	20,046	86.5

Abbreviations: CDC, Centers for Disease Control and Prevention; TB, tuberculosis.

[a]Among 39,920 total TB cases, 87 had an unknown place of birth; these cases were excluded from the US-born and foreign-born columns of the table.

Table 2

Characteristics of Genotyped and Nongenotyped Tuberculosis Cases Reported to the Centers for Disease Control and Prevention, United States, 2006–2008[a,b]

	Genotyped TB Cases (Culture-positive)		Nongenotyped TB Cases			
			Culture-positive		Clinical[c]	
	No.	%	No.	%	No.	%
Total	22,857		8,328		8,648	
Sex						
Male	14,410	63.0	5,080	61.0	4,973	57.5
Female	8,444	36.9	3,194	38.4	3,669	42.4
Unknown	3	0.0	54	0.6	6	0.1
Age group, years						
0–14	469	2.1	189	2.3	1,701	19.7
15–24	2,829	12.4	839	10.1	881	10.2
25–44	7,868	34.4	2,760	33.1	2,592	30.0
45–64	7,017	30.7	2,543	30.5	2,429	28.1
≥65	4,674	20.4	1,997	24.0	1,044	12.1
Unknown	0	0.0	0	0.0	1	0.0
Race/ethnicity						
Non-Hispanic white	3,845	16.8	1,491	17.9	1,394	16.1
Non-Hispanic black	5,779	25.3	2,317	27.8	2,389	27.6
Hispanic	6,487	28.4	2,385	28.6	2,813	32.5
Other	6,709	29.4	2,122	25.5	2,041	23.6
Unknown	37	0.2	13	0.2	11	0.1
US Census region						
Northeast	3,789	16.6	1,697	20.4	1,780	20.6
Midwest	3,012	13.2	858	10.3	1,064	12.3
South	8,779	38.4	3,836	46.1	3,480	40.2
West	7,277	31.8	1,937	23.3	2,324	26.9
Birthplace						

	Genotyped TB Cases (Culture-positive)		Nongenotyped TB Cases			
			Culture-positive		Clinical[c]	
	No.	%	No.	%	No.	%
US-born	9,133	40.0	3,577	43.0	3,952	45.7
Foreign-born	13,724	60.0	4,751	57.0	4,696	54.3

Abbreviation: TB, tuberculosis.

[a] Among 39,920 total TB cases, 87 had unknown place of birth (42 genotyped cases, 32 culture-positive nongenotyped cases, and 13 clinical nongenotyped cases); these cases are not shown in the table.

[b] Comparison of genotyped cases with nongenotyped cases revealed that genotyped cases were more likely to be male, to reside in the western region of the United States, and to have been born outside of the United States; genotyped and nongenotyped cases also differed by age group and race/ethnicity ($P < 0.001$ for all χ^2 comparisons).

[c] Clinical cases are TB cases diagnosed using laboratory or provider verification in the absence of a culture result.

Table 3

Rates of Reactivation Tuberculosis per 100 Person-Years, by Population Subgroup, United States, 2006–2008

	No. of Cases[a]	% of US Population With Latent TB Infection	US Population Estimate[b]	No. of PY at Risk for Reactivation TB[b]	Rate of Reactivation TB per 100 PY	95% CI[c]	Rate Ratio	95% CI
Total	31,841	4.2	301,580	37,815	0.084	0.083, 0.085		
Sex								
Male	19,293	5.1	147,861	22,837	0.084	0.083, 0.086		Reference
Female	12,568	3.2	153,719	14,900	0.084	0.083, 0.086	1.0	0.97, 1.0
Age group, years								
1–14	1,504	1.1	62,904	2,097	0.072	0.068, 0.075	0.62	0.59, 0.66
15–24	3,513	2.4	42,533	3,055	0.115	0.111, 0.119	1	Reference
25–44	10,483	5.0	85,616	12,803	0.082	0.080, 0.083	0.71	0.69, 0.74
45–64	9,211	6.5	74,052	14,484	0.064	0.062, 0.065	0.55	0.53, 0.58
≥65	6,862	5.5	36,475	6,066	0.113	0.110, 0.116	0.98	0.95, 1.02
Race/ethnicity								
Non-Hispanic white	5,551	1.9	213,057	12,408	0.045	0.044, 0.046	1	Reference
Non-Hispanic black	7,191	7.0	37,896	7,965	0.090	0.088, 0.092	2.0	1.9, 2.1
Hispanic	9,531	9.9	36,094	10,758	0.089	0.087, 0.091	2.0	1.9, 2.1
Asian/other	9,470	11.4	14,533	4,990	0.190	0.186, 0.194	4.2	4.1, 4.4
US Census region								
Northeast	5,995	3.2	55,946	5,317	0.113	0.110, 0.116	1	Reference
Midwest	4,082	3.2	67,186	6,527	0.063	0.061, 0.065	0.56	0.55, 0.59
South	12,291	4.6	108,828	15,003	0.082	0.080, 0.084	0.74	0.71, 0.77
West	9,432	5.4	69,620	11,313	0.083	0.082, 0.085	0.44	0.43, 0.46
Birthplace								
US-born	11,672	1.8	265,164	14,319	0.082	0.080, 0.083	1	Reference
Foreign-born	20,046	18.7	36,416	20,429	0.098	0.096, 0.10	1.2	1.2, 1.3

Abbreviations: CI, confidence interval; PY, person-years; TB, tuberculosis.

[a] Numbers of cases in population subgroups may not exactly sum to total numbers of cases because of estimation and rounding.

[b] Per 1,000 population.

[c] Wilson's method.

Am J Epidemiol. Author manuscript; available in PMC 2017 August 08.

Table 4

Rates of Reactivation Tuberculosis per 100 Person-Years, by Population Subgroup, According to Birthplace (US Birth vs. Foreign Birth), United States, 2006–2008

	No. of Cases[a]	% of US Population With Latent TB Infection	US Population Estimate[b]	No. of PY at Risk for Reactivation TB[b]	Rate of Reactivation TB per 100 PY	95% CI[c]	Rate Ratio	95% CI
				US-born				
Total	11,672	1.8	265,164	14,319	0.082	0.080, 0.083		
Sex								
Male	7,436	2.1	129,441	8,155	0.091	0.089, 0.093	1	Reference
Female	4,243	1.4	135,723	5,700	0.074	0.072, 0.077	0.82	0.78, 0.86
Age group, years								
1–14	981	0.3	60,539	545	0.180	0.165, 0.192	1.6	1.5, 1.8
15–24	749	0.6	37,949	683	0.110	0.102, 0.118	1	Reference
25–44	2,393	1.2	69,201	2,491	0.096	0.092, 0.100	0.88	0.81, 0.95
45–64	4,095	3.3	64,886	6,424	0.064	0.062, 0.066	0.58	0.54, 0.63
≥65	3,303	4.7	32,589	4,595	0.072	0.069, 0.074	0.66	0.61, 0.71
Race/ethnicity								
Non-Hispanic white	4,455	1.1	201,365	6,645	0.067	0.065, 0.069	1	Reference
Non-Hispanic black	4,426	5.6	35,237	5,920	0.075	0.073, 0.077	1.1	1.1, 1.2
Hispanic	1,958	1.7	21,542	1,099	0.178	0.171, 0.186	2.7	2.5, 2.9
Asian/other	800	2.7	7,020	569	0.141	0.131, 0.151	2.1	1.9, 2.3
US Census region								
Northeast	1,548	0.9	48,145	1,290	0.120	0.114, 0.126	1	Reference
Midwest	1,617	1.8	63,310	3,430	0.047	0.045, 0.050	0.40	0.39, 0.40
South	6,165	1.9	97,866	5,575	0.111	0.108, 0.113	0.93	0.88, 0.98
West	2,332	2.3	55,843	3,832	0.061	0.058, 0.064	0.34	0.32, 0.36
				Foreign-born				
Total	20,046	18.7	36,416	20,429	0.098	0.096–0.100		
Sex								
Male	11,806	22.7	18,420	12,544	0.094	0.092, 0.096	1	Reference
Female	8,246	14.4	17,996	7,774	0.106	0.104, 0.108	1.1	1.1, 1.2

	No. of Cases[a]	% of US Population With Latent TB Infection	US Population Estimate[b]	No. of PY at Risk for Reactivation TB[b]	Rate of Reactivation TB per 100 PY	95% CI[c]	Rate Ratio	95% CI
Age group, years								
1–14	525	11.9	2,367	845	0.062	0.057, 0.068	0.40	0.36, 0.44
15–24	2,749	12.8	4,584	1,760	0.156	0.150, 0.162	1	Reference
25–44	8,098	20.6	16,414	10,144	0.080	0.078, 0.082	0.51	0.49, 0.53
45–64	5,114	25.3	9,165	6,956	0.074	0.072, 0.076	0.47	0.45, 0.49
≥65	3,557	11.9	3,886	1,387	0.256	0.248, 0.265	1.6	1.6, 1.7
Race/ethnicity								
Non-Hispanic white	1,101	17.9	9,040	4,854	0.023	0.021, 0.024	1	Reference
Non-Hispanic black	2,762	20.0	2,297	1,378	0.200	0.193, 0.208	8.8	8.1,9.7
Hispanic	7,510	17.7	16,478	8,750	0.086	0.084, 0.088	3.8	3.5,4.1
Asian/other	8,667	24.4	8,601	6,296	0.138	0.135, 0.141	6.1	5.6, 6.6
US Census region								
Northeast	4,412	16.9	7,800	3,949	0.112	0.108, 0.115	1	Reference
Midwest	2,461	22.4	3,885	2,612	0.094	0.091, 0.098	0.84	0.80, 0.89
South	6,114	18.7	10,965	6,149	0.099	0.097, 0.102	0.89	0.86, 0.93
West	7,079	16.8	13,766	6,944	0.102	0.100, 0.104	0.44	0.42, 0.45

Abbreviations: CI, confidence interval; PY, person-years; TB, tuberculosis.

[a] Numbers of cases in population subgroups may not exactly sum to total numbers of cases because of estimation and rounding.

[b] Per 1,000 population.

[c] Wilson's method.

Table 5

Estimated Rate of Reactivation Tuberculosis Among HIV-infected and HIV-uninfected Tuberculosis Patients Aged 15–64 Years Not Residing in California, United States, 2006–2008

	Estimated No. of Reactivation TB Cases	Estimated % of US Population With Latent TB Infection	Estimated US Population	Estimated No. of PY at Risk for Reactivation TB	Estimated Rate of Reactivation TB per 100 PY	95% Confidence Interval
HIV-infected	2,198	4.2	961,000	121,100	1.82	1.74, 1.89
HIV-uninfected	16,568	4.2	182,243,000	22,850,000	0.073	0.070, 0.075

Abbreviations: HIV, human immunodeficiency virus; PY, person-years; TB, tuberculosis.