

# Racial inequality in the annual risk of Tuberculosis infection in the United States, 1910-1933

J.L. Zelner<sup>1,2</sup>, C. Muller<sup>3</sup>, and J.J. Feigenbaum<sup>4</sup>

<sup>1</sup>Dept. of Epidemiology, University of Michigan School of Public Health

<sup>2</sup>Center for Social Epidemiology and Population Health, University of Michigan  
School of Public Health

<sup>3</sup>Dept. of Sociology, University of California, Berkeley

<sup>4</sup>Department of Economics, Princeton University

**Corresponding Author:**

Jonathan Zelner

Center for Social Epidemiology and Population Health

University of Michigan School of Public Health

Dept. of Epidemiology

1415 Washington Heights

SPH I, Suite 2675

Ann Arbor, MI USA 48109-2029

**Running Head:** Racial inequality in tuberculosis risk

# Abstract

Tuberculosis (TB) mortality rates in the U.S. fell rapidly from 1910-1933. However, during this period, racial disparities in TB mortality in the nation's rapidly expanding cities grew. Because of long delays between infection and disease, TB mortality a poor indicator of short-term changes in transmission. We estimated the annual risk of TB infection (ARTI) in 11 large U.S. cities to understand whether rising inequality in mortality reflected in rising inequality in ARTI using city-level TB mortality data compiled by the U.S. Dept. of Commerce from 1910 to 1933. We estimated ARTI for blacks and whites using pediatric extrapulmonary TB mortality data among African-Americans and whites our panel of cities. We also estimated age-adjusted pulmonary TB mortality rates for these cities. We find that the ratio of ARTI for blacks vs. whites increased from 2.1 (95% CI=1.7, 2.5) in 1910 to 4.2 (95% CI=3.6, 5.1) in 1933. This change mirrored increasing inequality in age-adjusted pulmonary TB mortality during this period. Our findings suggest that rising racial inequality in ARTI likely reflects migration and increasing population density and segregation in Northern cities during this period.

# 1 Introduction

During the first half of the 20th century in the United States, population-level Tuberculosis (TB) mortality rates fell rapidly from their heights in the 19th century<sup>1</sup>. These declines have been attributed to improved living conditions and medical care<sup>2,3</sup>. However, as TB mortality rates fell, racial disparities in TB mortality in the nation's rapidly expanding cities widened. In some cities, the TB mortality rate of African-Americans in 1930 was as high as four or five times that of whites<sup>4</sup>. One explanation for the durability of the disparity in TB mortality is that living and working conditions for African-Americans and whites diverged as U.S. cities grew: From 1916 to 1930, during the First Great Migration<sup>5</sup>, more than 1.5 million African-Americans migrated from the agrarian South into densely populated, segregated neighborhoods in the industrial North. This rapid influx of migrants, combined with discriminatory housing policies that artificially reduced the pool of available housing for African-Americans resulted in overcrowding that provided an ideal environment for TB transmission<sup>6</sup>.

The complicated natural history of TB infection makes it difficult to determine whether these divergent TB mortality rates reflected rising disparities in infection risk or in factors that differentially impacted case-fatality rates, such as advances in TB treatment that were unequally distributed by race. Because the epidemiology of TB is characterized by long and variable delays between infection and disease<sup>7</sup>, changes in transmission rates are not immediately reflected in temporal trends of pulmonary TB mortality at the population level<sup>8</sup>. In addition, the trajectory of TB mortality in the United States was dramatically altered by the 1918 influenza pandemic. Noymer<sup>9</sup> has demonstrated that people with TB were more likely to die during the flu pandemic than people without TB, leading to a rapid decline in TB mortality immediately after the pandemic. Noymer suggests that the flu hastened the decline of TB in the US by removing infectious people from the population. But the flu's effect on racial disparities in TB mortality in U.S. cities is unclear: increasing or stagnating transmission rates in increasingly segregated neighborhoods could have offset the effects of the reduction in the pool of infectious people.

To understand whether rising disparities in pulmonary TB mortality reflected changes in underlying transmission rates, we estimated the annual rate of TB infection (ARTI) for African-Americans and whites in large Northern cities from 1910 to 1933. We used the method of Vynnycky and Fine to indirectly estimate ARTI using pediatric TB meningitis (TBM) mortality data<sup>7</sup>. TB meningitis is a form of extra-pulmonary TB that causes inflammation of the membranes surrounding the brain, known as the meninges. Although TBM is an extra-pulmonary infection, people with TBM typically acquire their infection via exposure to someone with infectious pulmonary TB. All TB disease among young children is the result of recent infection of a susceptible individual, and so reflects short-term changes in risk to the population of susceptible people.

We focus on Northern cities in the U.S. to understand changing TB risks in cities that were growing rapidly in part as a result of the First Great Migration. In addition, historical demography has demonstrated that patterns of segregation were very different in Northern and Southern cities during this period, with Northern segregation characterized by the emergence of large, isolated ghetto areas that are particularly conducive to TB transmission<sup>6</sup>, and segregation in Southern cities characterized by lower residential density<sup>10</sup>.

## 2 Data

We used city-specific TB mortality data compiled by the U.S. Dept. of Commerce from 1910 to 1933<sup>11</sup> that we digitized and cleaned. These data include counts of pulmonary TB mortality in all years and specific causes of extra-pulmonary TB mortality from 1910-1920 and aggregated counts of extra-pulmonary TB from 1921-1933. We focus on the period from 1910 to 1933 both because of its historical importance and because after 1933 our mortality data switches from reporting deaths for whites and African-Americans to reporting deaths for whites and nonwhites.

To calculate per-capita mortality rates, we obtained age-specific census population estimates from these cities for the census years 1910, 1920, 1930 and 1940 using census extracts provided by IPUMS<sup>12</sup>. To obtain population estimates for our panel of cities during intercensal years, we ag-

gregated the population data for relevant age groups in all of the cities and linearly interpolated these values across census years.

In the Dept. of Commerce data, there are 16 cities (and counties in the case of boroughs within New York City), for which age-specific counts of pulmonary TB and extra-pulmonary TB deaths were available. Of these, 11 were in the northern U.S. : Baltimore, MD; Brooklyn, NY; Chicago, IL; Cincinnati, OH; Indianapolis, IN; Kansas City, MO; Manhattan, NY; Philadelphia, PA; Pittsburgh, PA; St. Louis, MO; Washington, DC, and the remaining 5 were in the south: Atlanta, GA; Birmingham, AL; Louisville, KY; New Orleans, LA; Richmond, VA. We denote cities in states that belonged to the former Confederacy as Southern cities, with the remaining states designated as northern cities.

## 3 Methods

### 3.1 Age-standardized pulmonary TB mortality rates

Estimates of racial disparity in pulmonary TB mortality should be adjusted for the age structure of the population because during the period we study millions of young African-Americans and whites moved from the rural South to cities in the North and South. Our estimates are age-adjusted to minimize the risk that they are confounded by the changing age composition of these cities from migration or changing birth and non-TB mortality rates. For more information on our approach to age-standardization, see the Supplementary Materials.

### 3.2 Estimating annual risk of TB infection

Vynnycky and Fine<sup>7</sup> used tuberculosis meningitis (TBM) mortality data for children less than 5 years old to estimate the annual risk of TB infection (ARTI) in England and Wales in the pre-chemotherapy era. They based their analysis on results from the Netherlands indicating that the

74 mortality rate for TB meningitis among children under 5 was roughly equivalent to 1% of the  
 75 population-level ARTI<sup>13</sup>, a pattern which was also found in Sweden for the period from 1925 to  
 76 1935<sup>14</sup>. We denote the ARTI in year  $t$  for group  $g$  as  $\lambda_t^g$ .

77 For the years 1910-1920, our data include counts of deaths from TB meningitis as well as other  
 78 categories of extra-pulmonary TB, including Pott's Disease, and "White swelling". Because total  
 79 TB mortality rates fell during the 1910s, extra-pulmonary causes of death in the raw data were  
 80 aggregated into a single category denoted "Other forms of Tuberculosis" from 1920 onward. To  
 81 estimate ARTI using these aggregated data, we used a Poisson model to predict the expected num-  
 82 ber of extra-pulmonary TB deaths in group  $g$  ( $D_{g,t}^{EP}$ ) as a function of the number of unobserved  
 83 TBM deaths ( $D_{g,t}^{TBM}$ ), where  $\alpha_g$  is the ratio of TBM deaths to extra-pulmonary deaths for group  $g$ :  
 84  $D_{g,t}^{EP} \sim \text{Poisson}(\alpha_g D_{g,t}^{TBM})$ . We used this information to estimate the ARTI ( $\hat{\lambda}_t^g$ ) for each racial  
 85 group,  $g$ , from the aggregate extra-pulmonary TB mortality data, where  $\hat{\gamma}_{g,t}$  is the expected num-  
 86 ber of extra-pulmonary TB deaths in group  $g$  during year  $t$ :

$$D_{g,t}^{EP} \sim \text{Poisson}(\hat{\gamma}_t^g) \quad (1)$$

$$D_{g,t}^{TBM} \sim \text{Poisson}(\alpha_g D_{g,t}^{EP}) \quad (2)$$

$$\lambda_t^g = 100 \times \hat{\gamma}_t^g / (\alpha_g N_{g,t}) \quad (3)$$

### 87 3.3 Per-case infection risk

88 We can use the ARTI estimates derived from Eq. 3 to approximate the number of new infections  
 89 generated by a single infectious pulmonary TB case in a given year among African-Americans  
 90 ( $\beta_t^B$ ) and whites ( $\beta_t^W$ ), respectively. We denote the total number of infectious TB cases in the  
 91 population during year  $t$  to be  $I_t$ , and the total population size to be  $N_t$ . The ARTI for a specific

group (i.e. black, white) in a given year can be decomposed into the product of 1) the population prevalence of smear-positive TB ( $I_t/N_t$ ) and 2) the group-specific per-case infection rate at time  $t$ ,  $\beta_t^g$ . Because  $I_t$  for this period cannot be observed directly, we approximate it from pulmonary TB mortality: the “Styblo rule” suggests that prior to the advent of effective TB antibiotics in the 1940s, there were approximately 4 prevalent TB cases for every one death from pulmonary TB<sup>15</sup>. A recent model-based analysis suggests that this approximation of pulmonary TB prevalence from mortality data is sensible for data from the pre-chemotherapy era<sup>16</sup>, which includes our entire study period. Consequently, we approximate the count of infectious pulmonary TB cases as  $I_t = 4 \times D_t$ , where  $D_t$  is the number of pulmonary TB deaths. We can then rearrange terms to estimate  $\beta_t$  for each group as follows, as in<sup>17</sup>:

$$\beta_t^g = \lambda_t^g \frac{N_t}{I_t} \quad (4)$$

We also modeled the relationship between city population density and  $\beta_t^g$  for each Northern U.S. city during the census years 1910, 1920, and 1930, when race-specific population estimates were most reliable. Our measure of population density is the average number of people living in a dwelling, which allows us to distinguish between cities characterized by low-density housing like houses and those characterized by high-density housing like apartments and tenements. We include dummy terms for each census year to adjust for trends not explained by changing city population density. For more information on this analysis, see the supplementary materials. We completed all analyses and data cleaning in R 3.3.1; we generated figures using ggplot2 0.9.3 and we fit models through MCMC using Stan 2.9<sup>18</sup>.

## 4 Results

### 4.1 Disparities in Pulmonary TB Mortality

Figure 1 shows age-adjusted pulmonary TB mortality rates for African-Americans and whites in our panel from 1910 to 1933. Panel A shows that age-adjusted mortality for African-Americans and whites fell, with a persistent gap between black and white pulmonary TB mortality rates. For whites, the age-adjusted pulmonary mortality rate fell from 158 cases/100K population (95 % CI = 155,160) in 1910 to 37/100K (95 % CI = 36,38) in 1933. For African-Americans, the age-adjusted TB mortality rate fell from 494/100K (95 % CI = 475,515) in 1910 to 196/100K (95 % CI = 189,203) in 1933. However, over this same period, the relative risk of death from TB rose for African-Americans as compared to whites, from 3.1 (95 % CI = 3.0,3.3) in 1910 to 5.3 (95 % CI = 5.0,5.5) in 1933. Panel B illustrates the growth in this ratio over time.

### 4.2 Annual Risk of TB Infection

Figure 2 depicts the annual risk of TB infection (ARTI) for African-Americans and whites from 1910 to 1930. The figure shows that for whites ARTI fell from 8.4% (95% CI = 7.8%, 8.9%) in 1910 to 1.7% (95% CI = 1.6%, 2.0%) in 1933. For African-Americans, the decline was less dramatic, falling from 17.3% (95% CI = 14.5%, 20.7%) in 1910 to 7.4% (95% CI = 6.2%, 8.9%) in 1933. The ratio of ARTI for blacks as compared to whites reached its lowest level in 1919 after the initial wave of the flu pandemic (RR = 2.1, 95% CI = 1.8, 2.5), but grew again as ARTI among African-Americans rebounded but continued to fall for whites, even as successive waves of pandemic flu hit some of the cities in our sample as late as 1920<sup>19</sup>. By 1933 the relative risk of TB infection for African-Americans as compared to whites was at its highest level (RR = 4.2, 95% CI = 3.4, 5.2) during our observation period.

For African-Americans, TBM deaths accounted for 68% of the total extra-pulmonary TB deaths during the period from 1910 to 1920 (95% CI = 56%, 83%). For whites, TBM accounted for a larger



share of extra-pulmonary deaths at about 82% (95% CI = 78%, 86%). These ratios and uncertainty in their estimation are reflected in the parameters  $\alpha_W$  and  $\alpha_B$  used to estimate  $\lambda_{g,t}$  for African-Americans and whites as in Eq. 3. Figure S1 shows the correspondence between the total number of extra-pulmonary TB deaths recorded in the dataset and the number of TB meningitis deaths in a given year. Additional models including an interaction term for race  $\times$  year indicated that the value of  $\alpha_g$  was stable over time for both groups.

### 4.3 Per-case infection risk

The ARTI for African-Americans and whites illustrated in Figure 2 is a product of the population prevalence of infectious (i.e. smear-positive) pulmonary TB and the group-specific rate of contact between susceptible individuals and infectious cases. Using the crude population prevalence of pulmonary TB in our cities (Figure S2) we can approximate the risk of TB infection posed by each prevalent case separately for African-Americans and whites, as described by Eq. 4.

Figure 3 shows that, for African-Americans, the number of new TB infections for every prevalent pulmonary TB case increased dramatically during the 1920s, whereas for whites this ratio continued to fall. On average, the risk of TB infection per pulmonary case fell by 2% per year for whites (95% CI = -2%, -1%), whereas for African-Americans, the number of new TB infections per prevalent case increased on average by 2% per year (95% CI = 1%, 3%).

Our analysis of the city-level per-case risk of infection during census years 1910, 1920 and 1930 shows that for each doubling of city population density, the per-case risk of infection for African-Americans and whites increased by 1.24 times (95% CI = 1.19, 1.29). This is equivalent to a ratio of 2.33 for the least vs. most dense Northern city in our data. After adjusting for city population density, there is still a residual decrease in per-case infection risk for whites in 1930 as compared to 1910 (RR = 0.88, 95% CI = 0.78, 1.00), and an increase for African-Americans (RR = 1.56, 95% CI = 1.12, 2.08). Across cities, increasing population density was associated with increasing per-case risk against a higher baseline risk for blacks as compared to whites.

City population size was strongly correlated with our measure of population density ( $r = 0.79$ ) and was omitted from the city-level analysis. As a result, it is difficult to determine whether the relationship between population size and ARTI reflects increasing population density alone or some other characteristic of city scaling. However, to understand the importance of city population density to the total burden of pediatric extrapulmonary TB in our data, we generated posterior simulations comparing the predicted number of pediatric extrapulmonary TB deaths in our dataset to a counterfactual scenario in which all cities had population density equal to the lowest value in the data. This allows us to account for the greater population size of the densely-populated, higher-risk cities in assessing the impact of this factor on total caseload. These simulations indicate that by 1930, city-level variation in population density accounted for approximately 40% (95% CI = 27%,53%) of black ARTI and 40% of white ARTI (95% CI = 33%,46%).

## 5 Discussion

Our findings confirm that the persistent and increasing racial disparities in pulmonary TB mortality discussed by Roberts<sup>4</sup> are reflected in patterns of age-adjusted mortality from a panel of large U.S. cities over a period of more than 20 years (Fig. 1). More critically, we can see there were similar disparities for African-Americans as compared to whites in the annual risk of TB infection, which more accurately measures short-term changes in risk than pulmonary TB mortality. By comparing our estimates of ARTI by race to the burden of pulmonary TB mortality in the population at large, we documented that the number of new TB infections for every pulmonary case increased among African-Americans at the same time that it decreased for whites (Fig. 3).

Although mortality from pulmonary TB fell for both groups, the per-pulmonary-case risk of TB infection increased for African-Americans and decreased for whites. This suggests that, for African-Americans, the rapid decline in pulmonary TB mortality immediately after the 1918 flu may have been offset by increasing or stagnating risks of infection from the dwindling pool of prevalent TB cases. For whites, in contrast, the per-case rate of infection consistently fell over the same period.

185 These diverging risks may reflect opposing trends in the quality of living conditions and medical  
186 care for African-Americans and whites in large cities during the First Great Migration. Our analy-  
187 sis of city-level variation in the per-case risk of infection during census years suggests that transmis-  
188 sion rates for both blacks and whites were greatest in the largest, most densely-populated cities,  
189 such as Chicago, New York, and Philadelphia, which absorbed the largest numbers of African-  
190 American migrants during this period.

191 Historical census data from this period<sup>20</sup> indicate that racial residential segregation - as measured  
192 by the isolation index - grew dramatically in a subset of the northern cities in our panel during  
193 the study period. Isolation is a measure of segregation that accounts for both the geographic sep-  
194 aration of African-Americans and whites as well as the black population share<sup>21</sup>. Scholars have  
195 proposed that it is a key explanation for the concentration of TB risk in black neighborhoods be-  
196 cause it indicates a high within-group intensity of contact<sup>6</sup>. This means that the results shown in  
197 Figure 3 cannot be interpreted straightforwardly as transmission rates, since this would imply uni-  
198 form mixing between black and white populations at a time when there is strong evidence that the  
199 rate of contact between African-Americans and whites was rapidly changing in large U.S. cities.  
200 Future dynamic modeling studies should use age-specific pulmonary TB and TBM mortality data  
201 as well as geographic information on case locations to understand the causal impact of segrega-  
202 tion on TB transmission in cities where such data are available. Our analysis represents a key step  
203 toward building these models.

204 A limitation of our findings is that we rely on aggregate extra-pulmonary TB mortality to estimate  
205 unobserved TB meningitis mortality for years from 1920 onwards. However, the stability of the  
206 ratio of TBM to aggregate extra-pulmonary TB deaths from 1910 to 1920 suggests that this approx-  
207 imation is reasonable for the period we observe. Including separate terms for the proportion of  
208 extra-pulmonary TB cases among African-Americans and whites ensures that our estimates of dis-  
209 parities are conservative: pooling black and white extra-pulmonary TB mortality would result in  
210 a larger value of  $\alpha_B$  and greater disparities in the ARTI between African-Americans and whites.  
211 Because our estimates of ARTI and crude TB prevalence are derived from mortality data, they are

necessarily approximate. However, this approximation of pulmonary TB mortality from mortality data has been shown to be sensible for data from the pre-chemotherapy era<sup>16</sup>. Another model-based analysis showed that the ARTI derived from TBM mortality data predicted age-specific patterns of pulmonary TB mortality<sup>17</sup>, suggesting that this is a good approximation of population-level risks. Further, our estimates of the relative annual risk of TB infection for African-Americans compared to whites are not affected by this approximation because ARTI for each group was calculated by multiplying estimates of the TBM mortality rate by the same constant. In addition, tuberculin skin test results from black and white U.S. Navy recruits in the late 1950s and early 1960s suggest that the disparities in infection risk that we documented here persisted into the chemotherapy era<sup>22</sup>.

We also rely on the relative risk of infection and death from TB as a measure of inequality in infection risk, although the absolute difference in age-adjusted TB mortality and ARTI declined during the study period. We believe that this is the relevant measure of health equity for the current analysis: across our panel of cities, ARTI and TB mortality among whites represents the best-case scenario for TB risk at any point in time and is the baseline against which gains in African-American infection risk should be judged. In addition, the rate of infection and death from TB among whites was still great enough during the study period that our results are not susceptible to inflated relative-risk estimates due to a denominator rapidly approaching zero. Finally, our analysis of per-case infection risk clearly indicates that both the absolute and relative per-case risk of infection increased during this period.

The finding that the number of new TB infections for every prevalent pulmonary TB case rose for African-Americans and declined for whites from 1910 to 1933 is consistent with the claim that increased transmission rates may have undercut reductions in TB mortality among African-Americans following the influenza epidemic. There are, however, other potential explanations for the disparities in pediatric TB meningitis mortality observed in our data that underlie these findings. The potential for differential reporting of TB-related mortality by race is an inherent limitation of any study of this type using administrative data. In addition, there have been sugges-

tions of differential susceptibility to extrapulmonary TB by race. For example, analyses of national data from the U.S. in the 1960s-1970s<sup>23</sup> and 1980s<sup>24</sup> showed that extrapulmonary TB accounted for a greater proportion of TB cases among African-Americans than whites, although these differences were relatively small, e.g. age & sex adjusted proportion of 18.9% vs 16.1% for African-Americans and whites, respectively<sup>24</sup>. It is also unclear whether these differences reflect genetic differences rather than disparities in SES and access to health care that impacted the risk of infection and disease. However, to the extent that racial differences in susceptibility to extrapulmonary TB drove the disparities observed in our data, they should not impact the changing patterns of disparity over time.

Another potential explanation for the increasing relative risk of TB meningitis mortality for African-Americans may be differing TB exposure profiles during the period of the Great Migration. In particular, African-American migrants from the agrarian south may have had greater exposure to *Mycobacterium bovis* than non-migrants. As a considerably larger proportion of *M. bovis* than *M. tuberculosis* infections result in extrapulmonary disease<sup>25</sup> it is possible that the trend of increasing disparity reflects, at least in part, these exposure differences. However, our reliance on pediatric TBM mortality should mitigate this risk somewhat, as there is no evidence of persistent human-to-human *M. bovis* transmission<sup>26</sup>.

Despite its historical focus, our study is part of a larger effort to develop a comprehensive framework for fighting TB and other infectious diseases in the present day: Migration into densely-populated, transitory urban neighborhoods is an important cause of TB and other infectious diseases in the crowded cities of the developing world<sup>27,28</sup>. Rural-to-urban migration following economic development can affect the physical environment in ways that make communities more susceptible to infectious diseases like TB and diarrheal disease<sup>29,30</sup>, and such migration is expected to increase dramatically in the coming decades<sup>31</sup>. Historical data on reductions and disparities in TB mortality, paired with an understanding of the social and biological forces driving these patterns, are essential tools for developing theories and methods to more effectively attack present-day global inequalities in infection risk.

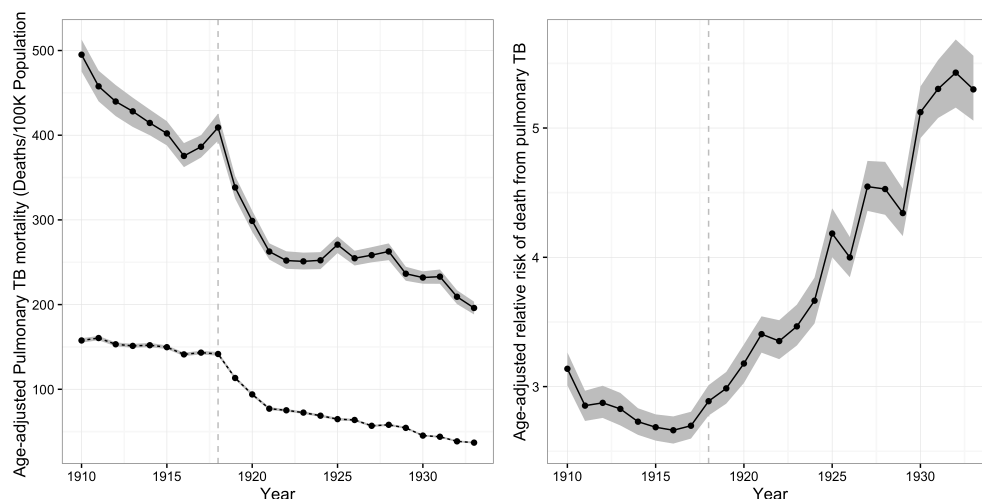


Figure 1: **Age-standardized TB mortality rates, by race.** The left-hand panel shows age-standardized TB mortality rates in the panel of cities from 1910 to 1933 in deaths per 100K population. Black mortality rates are indicated by the solid line and white TB mortality rates by the dashed line. The right-hand panel shows the adjusted relative risk (ARR) of death from pulmonary TB for African-Americans vs. whites from 1910 to 1933. The vertical dashed line indicates the timing of the 1918 influenza pandemic. Shaded regions in both panels indicated 95% posterior credible intervals.

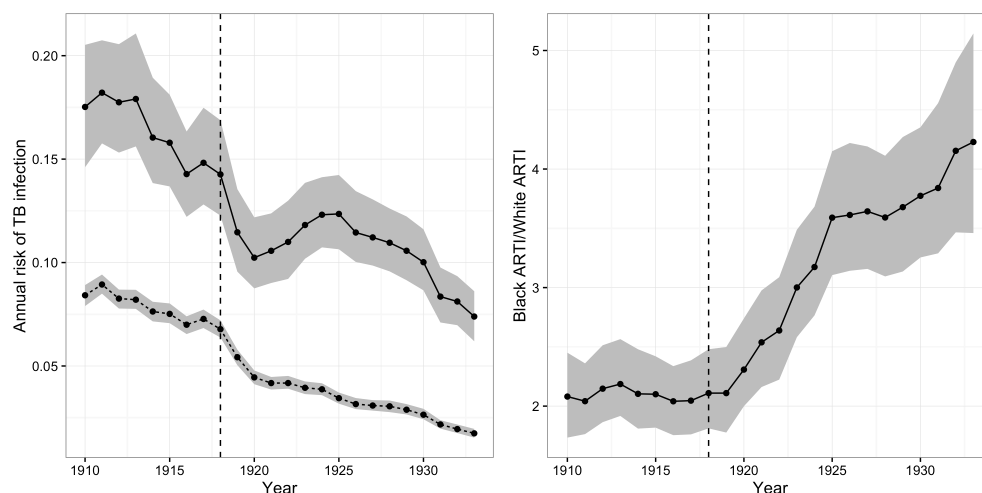


Figure 2: **Annual risk of TB infection (ARTI) by race, 1910-1933.** The left-hand panel illustrates the ARTI for African-Americans (solid line) and whites (dashed line) from 1910-1933. The right-hand panel illustrates the ratio of the ARTI for African-Americans vs. whites during this period. The gray shaded area in both panels illustrates the 95% posterior credible intervals (CIs) for these quantities.

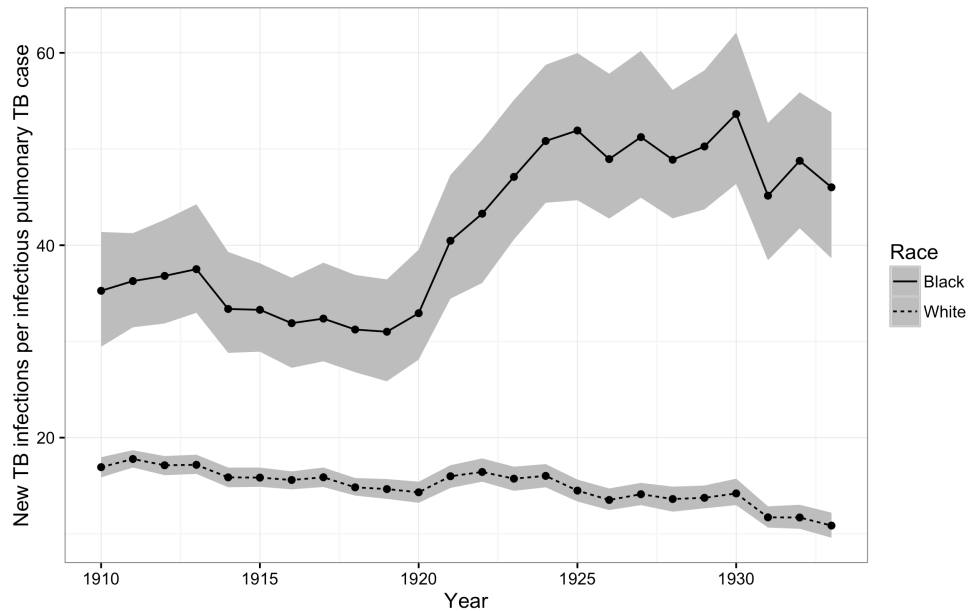


Figure 3: **New TB infections per prevalent pulmonary TB case, 1910-1933.** The figure illustrates trends in the number of new TB infections among African-Americans (solid line) and whites (dashed line) for every prevalent pulmonary TB case during the period from 1910-1933. Areas shaded in gray indicate 95% posterior credible intervals.

## References

1. Noymer A, Garene M. The 1918 influenza epidemic's effect on sex differentials in mortality in the United States. *Population and Development Review*. 2000;26(3):565-581.
2. McKeown T, Record R. Reasons for the decline of mortality in England and Wales during the nineteenth century. *Population Studies*. 1962;16:94-122.
3. D'Arcy Hart P, Wright PG. *Tuberculosis and social conditions in England with special reference to young adults*. National Association for the Prevention of Tuberculosis; 1939.
4. Roberts SK. *Infectious Fear: Politics, Disease, and the Health Effects of Segregation*. University of North Carolina Press; 2009.
5. Lee ES, Miller AR, Brainerd CP, Easterlin RA. *Population Redistribution and Economic Growth*,

- 277 *United States, 1870-1950*. Vol 1. Philadelphia: American Philosophical Society; 1957.
- 278 6. Acevedo-Garcia D. Residential segregation and the epidemiology of infectious diseases. *Social*  
279 *Science & Medicine*. 2000;51(8):1143-1161.
- 280 7. Vynnycky E, Fine P. The annual risk of infection with *Mycobacterium tuberculosis* in England  
281 and Wales since 1901. *International Journal of Tuberculosis and Lung Disease*. 1997;1(5):389-396.
- 282 8. Blower SM, McLean AR, Porco TC, et al. The intrinsic transmission dynamics of tuberculosis  
283 epidemics. *Nature Medicine*. 1995;1(8):815-821.
- 284 9. Noymer A. The 1918 influenza pandemic hastened the decline of tuberculosis in the United  
285 States: An age, period, cohort analysis. *Vaccine*. 2011;29:B38-B41.
- 286 10. Grigoryeva A, Ruef M. The Historical Demography of Racial Segregation. *American Sociological*  
287 *Review*. 2015;80(4):814-842.
- 288 11. U.S. Centers for Disease Control and Prevention. Vital statistics of the united states. [http:](http://www.cdc.gov/nchs/products/vsus.htm)  
289 [//www.cdc.gov/nchs/products/vsus.htm](http://www.cdc.gov/nchs/products/vsus.htm). Published 2016.
- 290 12. Ruggles S, Genadek K, Goeken R, Grover J, Sobek M. Integrated Public Use Microdata Series:  
291 Version 6.0. 2016. <https://usa.ipums.org/>.
- 292 13. Styblo K, Meijer J, Sutherland I. The transmission of tubercle bacilli: its trend in a human  
293 population. *Bulletin of the International Union of Tuberculosis and Lung Disease*. 1969;42:5-104.
- 294 14. Sjorgen I, Sutherland I. The risk of tuberculous infection in Sweden. *Tubercle*. 1975;56:97-112.
- 295 15. Styblo K. The relationship between the risk of tuberculous infection and the risk of developing  
296 infectious tuberculosis. *Bulletin of the International Union of Tuberculosis and Lung Disease*. 1985;60:117-  
297 119.
- 298 16. Begun M, Newall A, Marks G, Wood J. Revisiting Styblo's law: could mathematical models aid  
299 in estimating incidence from prevalence data? *Epidemiology and Infection*. 2015;143:1566-1565.
- 300 17. Vynnycky E, Fine P. Interpreting the decline in tuberculosis: the role of secular trends in effec-



301 tive contact. *International Journal of Epidemiology*. 1999;28:327-334.

302 18. RStan: the R interface to Stan, Version 2.8.0. 2015. <http://mc-stan.org/rstan.html>.

303 19. Olson DR, Simonsen L, Edelson PJ, Morse SS. Epidemiological evidence of an early wave of  
 304 the 1918 influenza pandemic in New York City. *Proceedings of the National Academy of Sciences*.  
 305 2005;102(31):11059-11063.

306 20. Logan JR, Zhang W, Turner R, Shertzer A. Creating the Black Ghetto: Black Residential Patterns  
 307 before and during the Great Migration. *The Annals of the American Academy of Political and Social  
 308 Science of the American Academy of Political and Social Science*. 2015;660(1):18-35.

309 21. U.S. Census Bureau. Measures of residential segregation. 2000.

310 22. Edwards P, Palmer C. Sensitivity to histoplasmin among Negro and white residents of different  
 311 communities in the USA. *Bulletin of the World Health Organization*. 1964;30:574-585.

312 23. Farer LS, Lowell AM, Meador MP. Extrapulmonary tuberculosis in the United States. *American  
 313 Journal of Epidemiology*. 1979;109(2):205-217.

314 24. Rieder H, Snider DE, Cauthen GM. Extrapulmonary tuberculosis in the United States. *American  
 315 Review of Respiratory Disease*. 1990;141(2):347-351.

316 25. Müller B, Dürr S, Hattendorf J, et al. Zoonotic *Mycobacterium bovis*-induced Tuberculosis in  
 317 humans. *Emerging Infectious Diseases*. 2013;19(6):899-908.

318 26. Grange J. *Mycobacterium bovis* infection in human beings. *Tuberculosis*. 2001;81(1/2):71-77.

319 27. Vries G de, Aldridge R, Cayla J, Sandgren A, Hest N van, Abubakar. Epidemiology of tubercu-  
 320 losis in big cities of the European Union and European Economic Area countries. *Euro Surveillance*.  
 321 2014;19(9):20726.

322 28. Alirol E, Getaz L, Chappuis F, Loutan L. Urbanisation and infectious diseases in a globalised  
 323 world. *The Lancet Infectious Diseases*. 2011;11:131-141.

324 29. Zelner JL, Trostle J, Goldstick JE, Cevallos W, House JS, Eisenberg JNS. Social connectedness  
 325 and disease transmission: social organization, cohesion, village context, and infection risk in rural

- 326 Ecuador. *American Journal of Public Health*. 2012;102(12):2233-2239.
- 327 30. Eisenberg JNS, Cevallos W, Ponce K, et al. Environmental change and infectious disease: How  
328 new roads affect the transmission of diarrheal pathogens in rural Ecuador. *Proceedings of the Na-*  
329 *tional Academy of Sciences*. 2006;103(51):19460-19465.
- 330 31. Goldstone JA. The new population bomb: The four megatrends that will change the world.  
331 *Foreign Affairs*. 2010;89(1):31-43.

# Supplementary Materials: Racial inequality in the annual risk of Tuberculosis infection, 1910-1933.

## 1 Supplementary Data

### 1.1 Panel Cities:

**North** (N = 11): Baltimore,MD; Brooklyn,NY; Chicago,IL; Cincinnati,OH; Indianapolis,IN; Kansas City,MO; Manhattan,NY; Philadelphia,PA; Pittsburgh,PA; St. Louis,MO; Washington,DC

**South** (N = 5): Atlanta,GA; Birmingham,AL; Louisville,KY; New Orleans,LA; Richmond,VA

## 2 Supplementary Methods

### 2.1 City-level variation in per-case risk

Because the population of small children in some cities in our panel is relatively small, particularly among African-Americans in the early years of the observation period, we cannot reliably interpolate population sizes for children under 5 between census years. We can, however, investigate the factors associated with variation in per-case risk at the city level during the 1910, 1920 and 1930 census years. We can model the per-case infection risk  $\beta$  as a function of city size, including yearly fixed effects by race. We can estimate the factors impacting city-level variation in  $\beta$  using a modified Poisson regression model, similar to the time series SIR model commonly applied to measles<sup>1</sup>. To do this, we estimate a Poisson regression model with the following underlying log-linear model, where  $X$  is a vector of city-specific covariates and  $\zeta$  is a vector of city/race regression coefficients:

$$\log(E(D_{ij}^{EP})) = \log(I/N) + \log(pop_{EP}) + X\zeta + \alpha_j \quad (1)$$

The inclusion of the offset term for pulmonary TB prevalence ( $\log(I/N)$ ) ensures that the coefficients in  $\zeta$  reflect the impact of covariates on per-pulmonary-case risk of infection rather than variation in ARTI. We do not include city-level random effects in this model to understand the relationship between city characteristics and variation in ARTI. This decision is justified by the relatively small number of cities in our sample. With more cities, we would be better able to differentiate between city characteristics and the impact of specific characteristics. However, our city-level results should obviously be interpreted in light of this small sample size.

In this analysis, we do not adjust for the differential reporting of TB meningitis mortality for blacks vs. whites and instead use the total count of extra-pulmonary TB cases, since we are primarily interested in factors impacting the relative risk of infection, such as city size, and these relationships would not be impacted by this adjustment.

### 3 Supplementary Results

#### 3.1 Age-adjusted pulmonary TB mortality

We standardized all pulmonary TB mortality rates using the distribution of ages in 1910 for people of all racial groups in the entire panel of cities. We linearly interpolated the population size at each age for the the entire panel of cities and calculated age-group specific pulmonary TB mortality rates for the age groups in the mortality data.

We used individual ages because, across years and cities, the bins used in the original data changed, from 5-year bins in some instances, to 10-year bins in others. For children under 5, mortality for exact ages were reported. Using the interpolated population data, we aggregated the population for the age groups included in each data bin (i.e. 5-9 years, 25-34 years, etc.) to estimate age-specific rates.

We estimated age-specific rates for each bin in the data using a Poisson model fit via MCMC. On each step of the MCMC estimation we calculated an age-adjusted pulmonary TB mortality rate by taking the weighted sum of these rates using population weights corresponding to the distribution of ages across all races in the 1910 census data for all cities in our data.

#### 3.2 City-level variation in ARTI

Table 1: Factors underlying change in city-level per-capita infection risk.

Variable	Median	Low CI	High CI
Intercept	0.04	0.04	0.05
Black	2.40	1.83	3.21
Year = 1920	0.82	0.75	0.91
Year = 1930	0.88	0.78	1.00
Black x 1920	1.15	0.84	1.59
Black x 1930	1.56	1.12	2.08
Log-population density	1.24	1.19	1.29
Black x Log-population density	1.28	1.17	1.41

Table 1 shows relative risks (RR) associated with city-level factors during the 1910, 1920 & 1930 census years. The model contains fixed-effects by race for each of the census years, to account for unmodeled change over time by race. It is important to note that because the estimates in this table reflect changes in the per-pulmonary-case risk of extrapulmonary TB in children under 5, the coefficient for race does not reflect the differential reporting of TB meningitis by race reported in the main text.

#### 3.3 TBM cases as a fraction of all extra-pulmonary TB cases: 1910-1920

Figure 1 illustrates the relationship between the count of extrapulmonary TB cases in our data and the number of TB meningitis cases in children under 5 among whites and African-Americans, respectively.

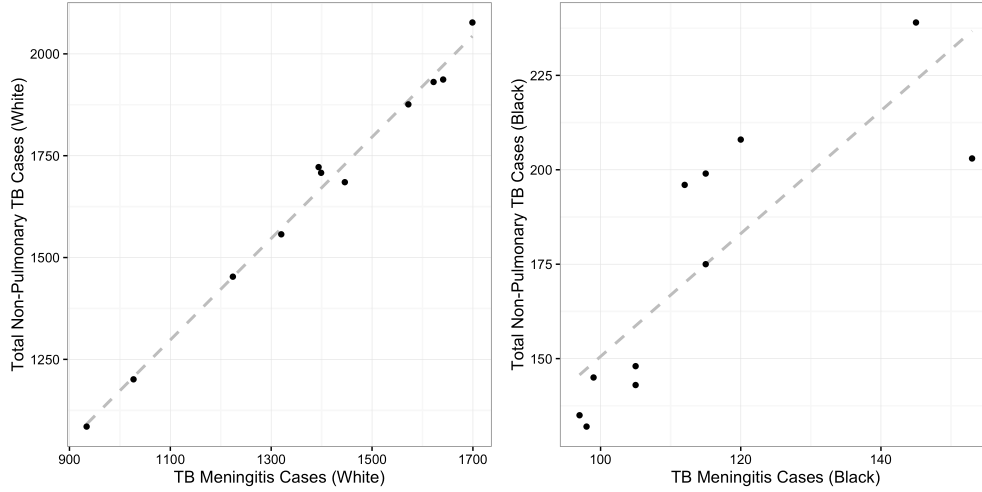


Figure 1: **TBM as a fraction of all non-pulmonary TB cases by race, 1910-1920.** The figure shows the ratio of aggregated non-pulmonary TB deaths to TB meningitis deaths for whites (left-hand panel) and blacks (right-hand panel). Each dot represents the total number of TBM and aggregated non-pulmonary TB deaths for all cities in the panel data.

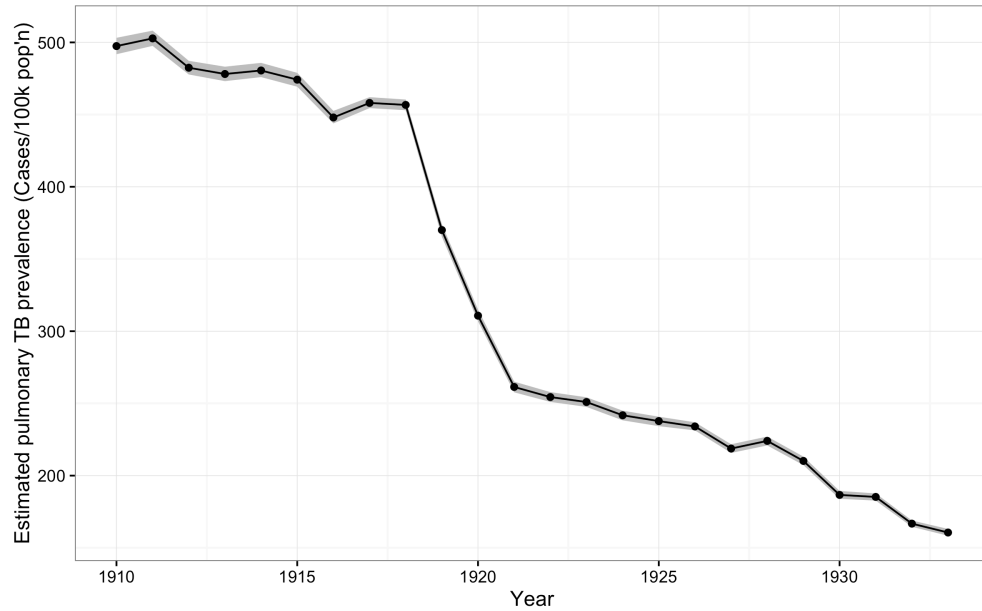


Figure 2: **Crude TB prevalence, 1910-1933.** Estimated population-level prevalence of infectious (i.e. smear-positive) TB estimated using aggregated crude TB mortality data across all Northern panel cities.

### 3.4 Combined Northern and Southern Results

In this section, we present estimates of age-adjusted TB mortality, the annual risk of TB infection by race, and the per-case risk of infection among individuals living in both the northern and southern cities in our dataset. These results are presented to demonstrate that our qualitative findings are not sensitive to the exclusion of these cities.

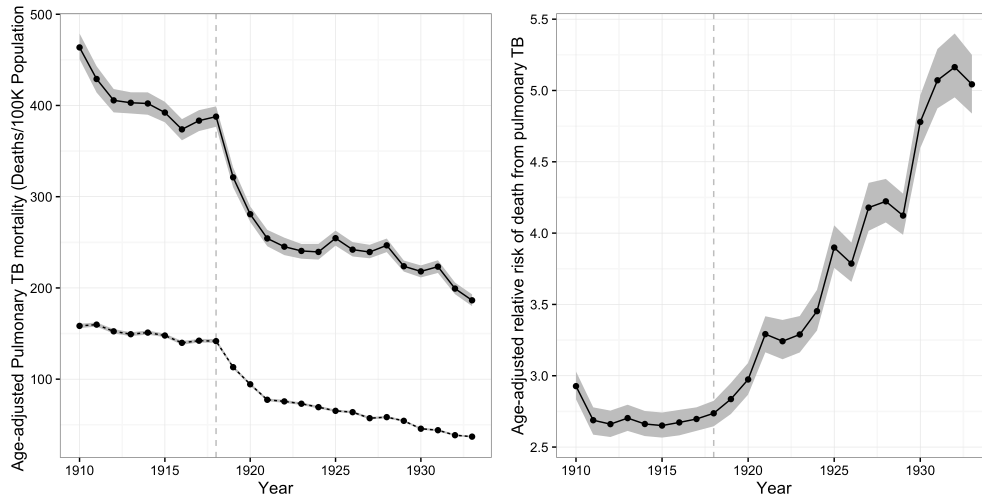


Figure 3: **Age-standardized TB mortality rates, by race.** The left-hand panel shows age-standardized TB mortality rates in the panel of cities from 1910 to 1933 in deaths per 100K population. Black mortality rates are indicated by the solid line and white TB mortality rates by the dashed line. The right-hand panel shows the adjusted relative risk (ARR) of death from pulmonary TB for African-Americans vs. whites from 1910 to 1933. The vertical dashed line indicates the timing of the 1918 influenza pandemic.

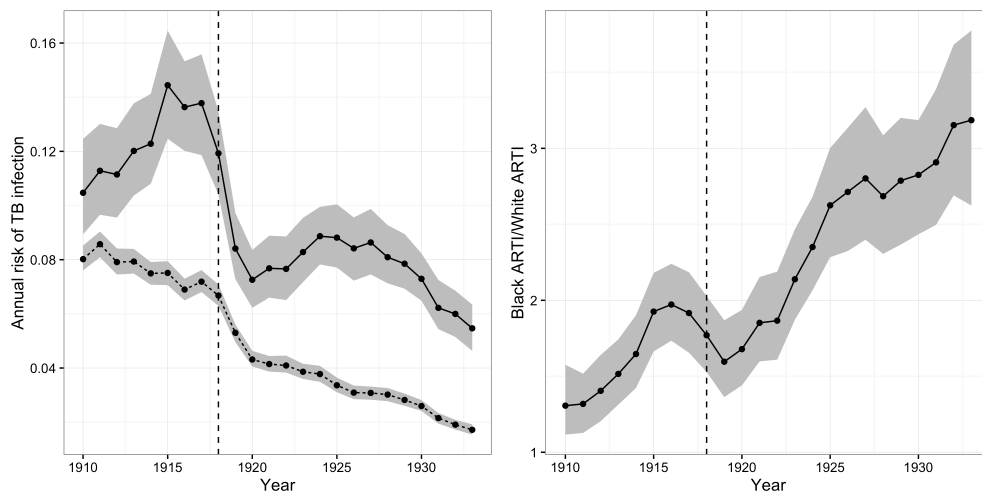


Figure 4: **Annual risk of TB infection (ARTI) by race, 1910-1933.** The left-hand panel illustrates the ARTI for African-Americans (solid line) and whites (dashed line) from 1910-1933. The right-hand panel illustrates the ratio of the ARTI for African-Americans vs. whites during this period. The gray shaded area in both panels illustrates the 95% posterior credible intervals (CIs) for these quantities.

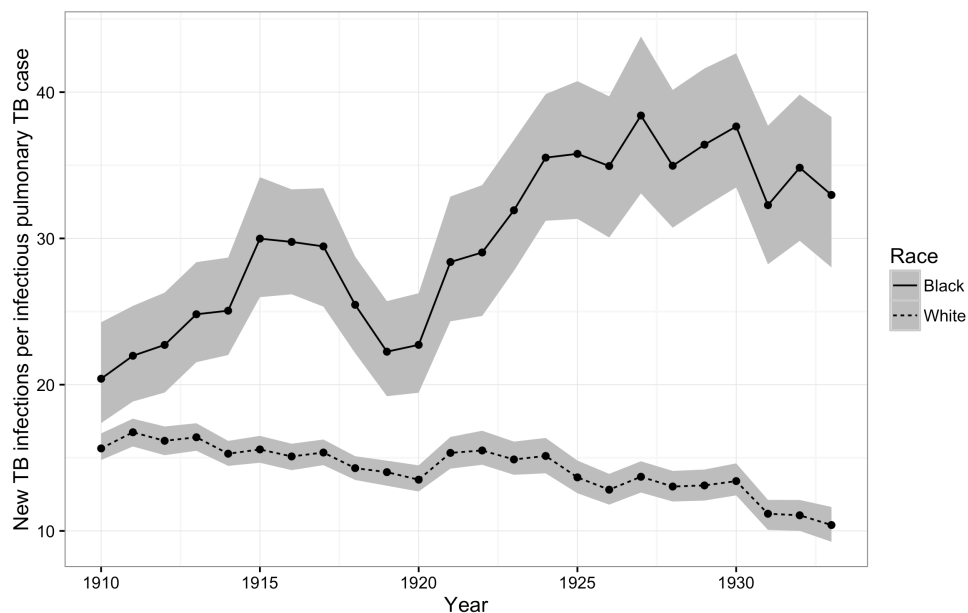


Figure 5: **New TB infections per prevalent pulmonary TB cases, 1910-1933.** The figure illustrates trends in the number of new TB infections among African-Americans (solid line) and whites (dashed line) for every prevalent pulmonary TB case during the period from 1910-1933.

### 3.5 Southern Results

In this section, we present estimates of age-adjusted TB mortality, the annual risk of TB infection by race, and the per-case risk of infection among individuals living in the southern cities in our dataset.

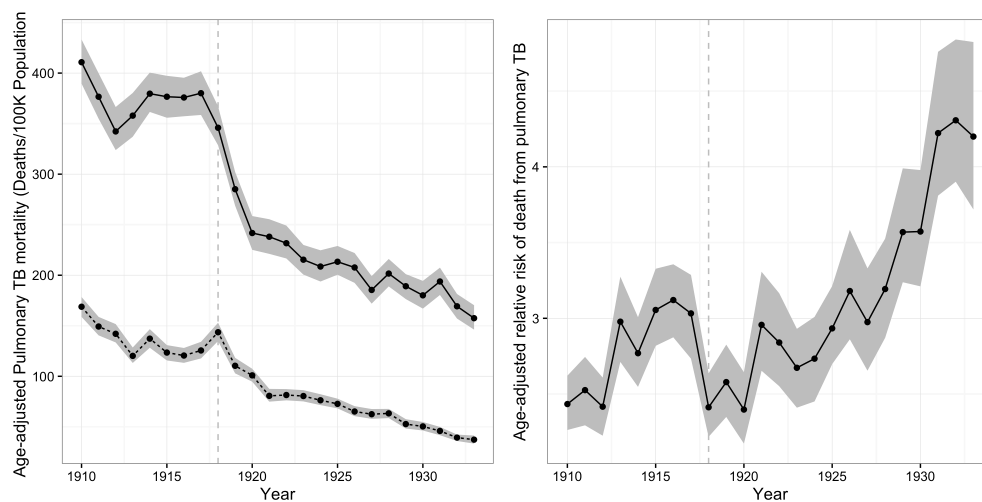


Figure 6: **Age-standardized TB mortality rates, by race.** The left-hand panel shows age-standardized TB mortality rates in the panel of cities from 1910 to 1933 in deaths per 100K population. Black mortality rates are indicated by the solid line and white TB mortality rates by the dashed line. The right-hand panel shows the adjusted relative risk (ARR) of death from pulmonary TB for African-Americans vs. whites from 1910 to 1933. The vertical dashed line indicates the timing of the 1918 influenza pandemic.

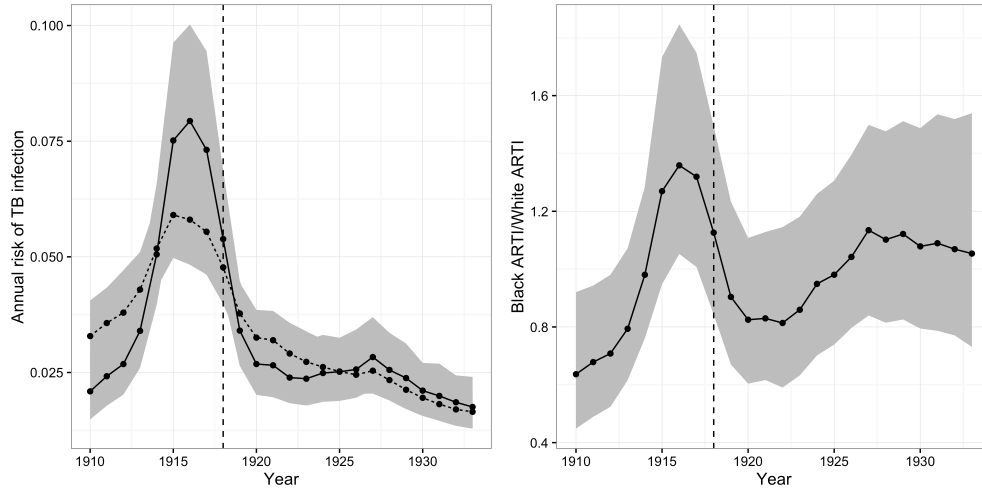


Figure 7: **Annual risk of TB infection (ARTI) by race, 1910-1933.** The left-hand panel illustrates the ARTI for African-Americans (solid line) and whites (dashed line) from 1910-1933. The right-hand panel illustrates the ratio of the ARTI for African-Americans vs. whites during this period. The gray shaded area in both panels illustrates the 95% posterior credible intervals (CIs) for these quantities.

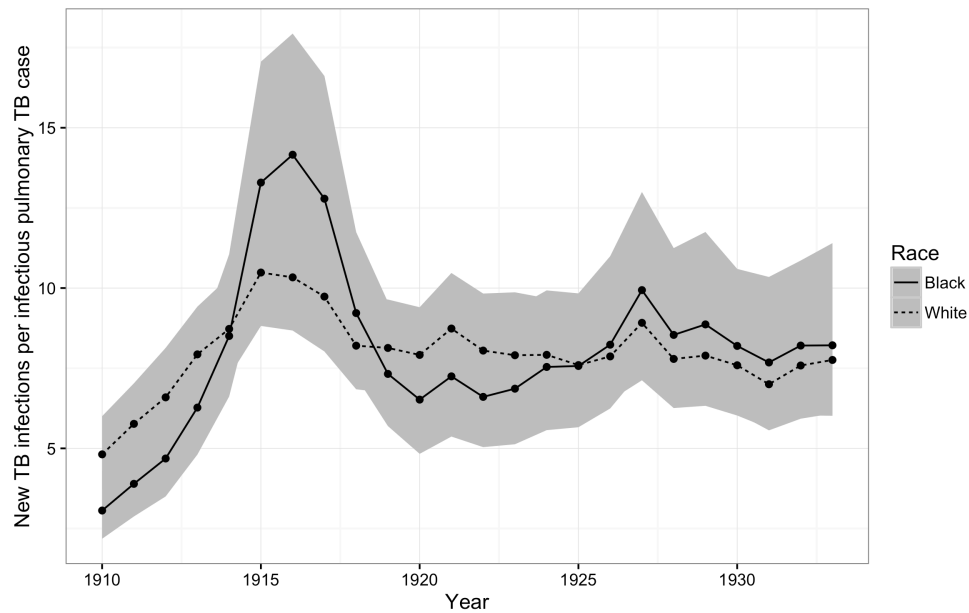


Figure 8: **New TB infections per prevalent pulmonary TB cases, 1910-1933.** The figure illustrates trends in the number of new TB infections among African-Americans (solid line) and whites (dashed line) for every prevalent pulmonary TB case during the period from 1910-1933.

## 4 Model code

```
data {
  int TBM_N; //Number of observations with disaggregated TBM and non_pulmonary_tb in data
  int<lower=TBM_N+1> N; //Total number of observations
```



```

int tbm_deaths[TBM_N]; //Count of TBM deaths by year by race
int non_pulmonary_deaths[N]; //Count of aggregate non-pulmonary deaths
int<lower=0,upper=1> black[N];
real<lower=0> pop[N];
int T; //Number of years in aggregated observations
int pulmonary_deaths[T]; //Aggregated pulmonary deaths
real<lower=0> total_pop[T]; //Total population

int year[N]; //Year indicator for indexing year terms
}

parameters {

  ## White and black rates of TBM:non-pulmonary-tb
  vector[2] tbm_beta;

  ## Non-pulmonary TB mortality rate for whites
  real np_tb_rate_w_intercept;
  vector[T-1] np_tb_rate_w_incr;

  ## Relative risk of non-pulmonary TB for blacks
  real np_tb_rate_rr_b_intercept;
  vector[T-1] np_tb_rate_rr_b_incr;

  vector[T] pulmonary_rate;
}

transformed parameters {
  vector[T] black_ari;
  vector[T] white_ari;

  vector[T] black_beta;
  vector[T] white_beta;

  vector[T] np_tb_rate_w;
  vector[T] np_tb_rate_rr_b;
  vector[T] prevalence;

  real white_beta_change;
  real black_beta_change;

  ## Add up increments to smooth term for white rate
  np_tb_rate_w[1] = np_tb_rate_w_intercept;
  for (t in 2:T) {
    np_tb_rate_w[t] = np_tb_rate_w[t-1] + np_tb_rate_w_incr[t-1];
  }

  np_tb_rate_rr_b[1] = np_tb_rate_rr_b_intercept;
  for (t in 2:T) {
    np_tb_rate_rr_b[t] = np_tb_rate_rr_b[t-1] + np_tb_rate_rr_b_incr[t-1];
  }
}

```

```

for (t in 1:T) {
  prevalence[t] = 4*exp(pulmonary_rate[t]);
}

for (t in 1:T) {
  white_ari[t] = 100*exp(tbm_beta[1] + np_tb_rate_w[t]);
  black_ari[t] = 100*exp(sum(tbm_beta) + np_tb_rate_w[t] + np_tb_rate_rr_b[t]);

  black_beta[t] = black_ari[t]/prevalence[t];
  white_beta[t] = white_ari[t]/prevalence[t];
}
}

model {

  tbm_beta ~ normal(0, 1);
  np_tb_rate_w_incr ~ normal(0, .1);
  np_tb_rate_rr_b_incr ~ normal(0, .1);
  pulmonary_rate ~ normal(0, 1);

  for (i in 1:TBM_N) {
    tbm_deaths[i] ~ poisson_log(log(non_pulmonary_deaths[i]) + tbm_beta[1] + tbm_beta[2]*black[i]);
  }

  for (i in 1:N) {
    int t;
    t = year[i];
    non_pulmonary_deaths[i] ~ poisson_log(log(pop[i]) + np_tb_rate_w[t] + np_tb_rate_rr_b[t]*black[i]);
  }

  for (i in 1:T) {
    pulmonary_deaths[i] ~ poisson_log(log(total_pop[i]) + pulmonary_rate[i]);
  }
}

```

## References

1. Bjørnstad ON, Finkenstädt B, Grenfell BT. Dynamics of measles epidemics: estimating scaling of transmission rates using a time series SIR model. *Ecological Monographs*. 2002;72:169-184.