To understand the epidemiology of tuberculosis (TB) and HIV co-infection in California, we cross-matched incident TB cases reported to state surveillance systems during 1993–2008 with cases in the state HIV/AIDS registry. Of 57,527 TB case-patients, 3,904 (7%) had known HIV infection. TB rates for persons with HIV declined from 437 to 126 cases/100,000 persons during 1993–2008; rates were highest for Hispanics (225/100,000) and Blacks (148/100,000). Patients co-infected with TB–HIV during 2001–2008 were significantly more likely than those infected before highly active antiretroviral therapy became available to be foreign-born, Hispanic, or Asian/Pacific Islander and to have pyrazinamide-monoresistant TB. Death rates decreased after highly active antiretroviral therapy became available but remained twice that for TB patients without HIV infection.
and higher for women. In California, HIV-associated TB has concentrated among persons from low- and middle-income countries who often acquire HIV infection in the peri-immigration period.

The modern resurgence of tuberculosis (TB) in conjunction with the HIV pandemic remains a major public health dilemma. In 2011, nine percent of all newly reported TB cases in the United States for which HIV status was known (1) and 13% (1.1 million cases) of cases reported worldwide (2) were associated with HIV co-infection. Despite compelling declines in TB incidence and associated deaths with use of highly active antiretroviral therapy (HAART), TB remains the leading cause of death among persons with HIV/AIDS (3,4).

California, the most populous state in the United States (38 million persons [12% of the US population]), reports the highest annual number of persons with TB (22.1% of total) and the second highest number of HIV-infected persons (103,073 [12.4%] cases) (5). Yet, because of prior restrictions on HIV reporting and limited systematic linking of state TB and HIV surveillance systems, California has not been included in key national surveillance reports of HIV incidence (6) or death in persons with TB–HIV (7).

Effective control of TB–HIV requires an understanding of the changing epidemiology of these diseases. To provide information for disease-reduction efforts and to improve survival among persons with TB–HIV, we retrospectively reviewed all incident TB–HIV cases in a 16-year period in California during which dynamic changes occurred in the HIV epidemic as a consequence of the introduction of HAART.

Methods

Study Population

We analyzed all TB cases reported to the California TB registry during January 1, 1993–December 31, 2008. California state law requires reporting of all verified cases of TB and HIV/AIDS (California Health and Safety Code Title 17§2505, and Section 121022 [2006]) to their respective programs. TB–HIV patients were identified through a statewide registry match with the California Office of AIDS by using Registry Plus Link Plus software (online Technical Appendix, wwwnc.cdc.gov/EID/article/19/3/12-1521-Techapp1.pdf). Annual state HIV prevalence was estimated through nonparametric back-calculation based on racial/ethnic group–specific counts of reported AIDS cases and reported AIDS-related deaths during 1981–2008 (online Technical Appendix Table 1) (8). Demographic, behavioral, and clinical information, including deaths, was abstracted from state surveillance forms (Report of a Verified Case of Tuberculosis and Adult HIV/AIDS Confidential Case Report). Surveillance data for both diseases have demonstrated high validity (9,10).

Definitions

HIV cases were classified in accordance with current Centers for Disease Control and Prevention (CDC) surveillance case definitions (11). For this analysis, patients not identified in the California HIV/AIDS registry were considered HIV-negative. TB diagnoses were grouped into 3 periods on the basis of HAART availability and approximately equal distribution of TB–HIV cases: pre-HAART (1993–1995), early HAART (1996–2000), and late HAART (2001–2008). Late diagnosis of HIV infection was defined as an AIDS diagnosis made <12 months after an initial diagnosis of HIV infection. Advanced immunosuppression was defined as a CD4+ T-lymphocyte count <50 cells/mm³; valid CD4+ T-lymphocyte counts were those collected within 6 months of TB diagnosis. Drug susceptibility testing was performed at local laboratories or at the California Microbial Diseases Laboratory (Richmond, CA, USA) by using the BACTEC 460TB System (Becton Dickinson Diagnostic Instruments, Sparks, MD, USA), BACTEC MGIT 960 Mycobacterial Detection System (Becton Dickinson), or the agar proportion method.

Statistical Analysis

Stratum-specific TB incidence per 100,000 population was calculated by dividing the number of incident cases by total (12) and HIV-infected population denominators. Clinical trends and demographic characteristics were described in 2 ways. First, we calculated prevalence ratios (PRs) and 95% CIs for comparison of characteristics associated with TB and TB–HIV cases (13). Second, we compared rates of annual percentage change (1993–2008) in prevalence of binary covariates by using logistic regression with robust SEs. This model included main effects for year (as a categorical variable) and HIV infection status, as well as the interaction between them.

Multivariate associations with death among TB–HIV patients were examined by using a generalized linear model with a log link and robust SEs to generate relative risk (RR) estimates (14); the model was a priori specified to include time period, age, sex, race/ethnicity, foreign birth, HIV risk factor, CD4+ T-lymphocyte count, sputum smear positivity, and interval between the diagnoses of HIV infection and TB. Death at diagnosis or during treatment was calculated from patients for whom outcome was known (3,754/3,904 [96.2%]); clinical outcomes during 2007 were excluded because of incomplete reporting. Multiple imputation was used to impute missing values for CD4+ T-lymphocyte count and viral load (15). The results obtained after multiple imputation were compared with those from an unimputed complete-case analysis (online Technical Appendix Table 2).
Proportions were compared by using χ² tests, and continuous variables were compared by using the Wilcoxon rank-sum test. All p values were 2-sided with α = 0.05 as the significance level. Data were analyzed by using Stata 12.1 (Stata Corporation, College Station, TX, USA) and R, version 2.13.2 (Foundation for Statistical Computing, Vienna, Austria).

Results

During 1993–2008, a total of 57,527 TB cases occurred in California, 3,904 (6.8%) of which were identified in a registry match with the California Office of AIDS. Overall, the proportion of TB–HIV cases decreased from 1,343 (9.2%) of 14,640 in the pre-HAART era to 1,254 (5.3%) of 23,812 in late HAART era. TB incidence among patients with HIV/AIDS declined from 437 cases per 100,000 persons to 126 per 100,000 (71% decrease); TB incidence rates for persons without known HIV co-infection declined from 16.5 cases per 100,000 persons to 7.0 per 100,000 (58% decrease) (Figure 1). Throughout the study period, incidence declined markedly in all strata of race/ethnicity but remained highest among Hispanics (225/100,000) and Blacks (148/100,000) with HIV/AIDS (Figure 2).

Demographic Trends

The median age of patients with known TB–HIV increased throughout the study period (Table 1, Appendix, wwwnc.cdc.gov/EID/article/19/3/12-1521-T1.htm); the number of persons >50 years of age with TB–HIV increased from 12% in the pre-HAART era to 21% in the late HAART era, out of proportion to the increase (from 40% to 46%) for TB cases alone (p<0.001 for difference in slopes). Likewise, female patients with TB–HIV increased from 11% to 17%, out of proportion to the increase (from 40% to 41%) for TB cases alone (p<0.001 for difference in slopes).

Foreign-born Persons

The proportion of TB–HIV patients who were foreign born increased from 37% in the pre-HAART era to 63% in the late HAART era, a greater increase than for TB patients without known HIV co-infection (from 67% to 77%; p<0.001 for difference in slopes). Seventy-six percent of foreign-born patients with TB–HIV (and 37% of all patients with TB–HIV) immigrated from Mexico or Central America. Among persons with HIV infection, TB was diagnosed a median of 11 years (interquartile range [IQR] 4–19 years) after entry into the United States, significantly longer than for persons without known HIV infection (8 years [IQR 2–18 years], p<0.001 by Wilcoxon rank-sum test). In contrast, TB–HIV patients were consistently younger (25 years [IQR 19–33 years]) than TB patients (31 years [IQR 20–49 years]; p<0.001 by Wilcoxon rank-sum test) at time of US entry. Immigrants from Mexico or Central America who had TB–HIV were younger (23 years [IQR 17–35 years]) than those from Southeast Asia or sub-Saharan Africa (36 years [IQR 24–55 years]; p<0.001 by Wilcoxon rank-sum test). In HIV-infected persons from sub-Saharan Africa, active TB developed sooner after immigration (2.6 years [IQR 0.2–5.7 years]) than in persons from all other regions (10.8 years [IQR 4.5–19.3 years]; p<0.001 by Wilcoxon rank-sum test). Compared with the pre-HAART era, patients with TB–HIV in the late HAART era were more likely to originate from Southeast Asia (PR 2.4; 95% CI 1.2–4.8) or sub-Saharan Africa (PR 4.6; 95% CI 2.1–10.1).

Race/Ethnicity

Hispanics accounted for 57% of all TB–HIV cases in the late HAART era (a 20% increase from the pre-HAART era); in contrast, the proportion of TB patients without known HIV co-infection who were Hispanic (38%) did not change. Asian/Pacific Islanders also represented an increasing proportion of TB–HIV patients in the late from the early HAART eras (8% vs. 4%; PR 2.2 [95% CI 1.6–3.0]). Among US-born persons, TB–HIV declined among Blacks (pre-HAART vs. late HAART; 31% vs. 17%) and Whites (23% vs. 12%) but not Hispanics (8% vs. 8%; p<0.001 for difference in slopes).

HIV Risk Group

From the pre-HAART era to the late HAART era among TB–HIV patients, the HIV transmission risk factors of men who have sex with men (MSM) (47% vs. 41%),
injection drug use (IDU) (22% vs. 15%), and MSM/IDU
(16% vs. 8%) decreased, whereas presumed heterosexual
transmission increased (6% vs. 16%; p=0.01 by χ²). MSM
was the most commonly reported HIV risk factor for all
racial/ethnic groups. Presumed heterosexual transmission
increased over time and was more common for TB–HIV
cases diagnosed in the late HAART era among Hispanics
(12%; 95% CI 10%–13%) and Blacks (11%; 95% CI 9%–
13%) than among Whites (4%; 95% CI 3%–5%). Similarly,
unknown or unreported HIV transmission risk factors were
more common among Hispanics (16%; 95% CI 15%–18%)
and Blacks (10%; 95% CI 9%–12%) than among Whites
(5%; 95% CI 4%–7%).

Clinical Trends

HIV Characteristics

Overall, AIDS developed within 12 months after HIV
diagnosis in 72% of patients, a percentage that did not sub-
stantially change throughout the study. Median CD4+ T-
lymphocyte count was 114 cells/mm³ (IQR 60–179 cells/
mm³) during the pre-HAART era and 100 cells/mm³ (IQR
55–150; p=0.01 by Wilcoxon rank-sum test) during the
late HAART era (Table 2, Appendix, wwwnc.cdc.gov/
EID/article/19/3/12-1521-T2.htm). Approximately 20% of
patients had advanced immunosuppression (CD4+ T-lym-
phocyte count <50 cells/mm³) during both the pre-HAART
and late HAART eras.

TB Characteristics

Patients with TB–HIV were more likely to be sputum
smear positive (PR 1.11; 95% CI 1.05–1.17) and less likely
to have culture-negative pulmonary TB (PR 0.58; 95% CI
0.49–0.68) than patients without known HIV co-infection
during the pre-HAART era. These differences diminished
in the late HAART era (PR 1.01; 95% CI 0.96–1.07 and PR
0.81; 95% CI 0.70–0.93, respectively).

Among initial isolates, pyrazinamide-mono-resistant
TB (PR 2.21; 95% CI 1.90–2.57) was more common, and
isoniazid-resistant (PR 0.69; 95% CI 0.60–0.79) and mul-
tidrug-resistant TB (PR 0.58; 95% CI 0.39–0.85) less com-
mon among TB–HIV patients than among patients without
known HIV co-infection. Pyrazinamide-monoresistant TB
among TB–HIV patients increased from 2% during the pre-
HAART era to 8% during the late HAART era, which was
out of proportion to the increase (from 1% to 3%) among
patients with TB alone (p<0.05 for difference in slopes).

Clinical Outcomes

Compared with the pre-HAART era, death during
treatment decreased in the late HAART era among TB–
HIV patients (30% vs. 14%) but not among TB patients
without known HIV co-infection (6.7% vs. 6.7%; p<0.001
for difference). In multivariate analysis, older age (RR 1.3
per 10 years; 95% CI 1.2–1.4), lower CD4+ T-lymphocyte
count (RR for reference value <50 cells/mm³ vs. >350–500
cells/mm³, 6.5; 95% CI 2.7–15.6), pre-HAART character-
istics (RR 2.2; 95% CI 1.9–2.6), sputum smear positivity
(RR 1.2; 95% CI 1.1–1.4), and female sex (RR 1.4; 95%
CI 1.1–1.7) were associated with increased risk for death,
whereas the heterosexual HIV risk group (relative to MSM,
RR 0.6; 95% CI 0.4–0.8) was protective (Table 3). Results
were similar in a sensitivity analysis excluding rather than
imputing missing CD4+ T-lymphocyte data (online Techni-
cal Appendix).

Discussion

In California, a dramatic decline in TB–HIV rates coin-
cided with the introduction of HAART and improvements
in TB control. The TB–HIV intersection has evolved from
one in which active TB and AIDS progressed in a margin-
alized, US-born population to an intersection increasingly
comprising persons from areas outside the United States
with elevated TB incidence who acquire HIV infection in
the peri-immigration period. TB–HIV-associated death has
decreased substantially but remains approximately twice
that associated with TB alone.

Population declines in TB–HIV after introduction of
HAART are well documented in low-income countries

Figure 2. Rates of tuberculosis (TB) for persons with HIV/AIDS,
California, USA, 1993–2008. Shaded areas represent 95%
bootstrap percentile CIs, by race. TB–HIV rates for Asians/Pacific
Islanders could not be calculated because of small numbers
of cases during some years. Annual state HIV prevalence was
estimated through nonparametric back-calculation on the basis
of racial/ethnic group–specific counts of reported AIDS cases
and reported AIDS-related deaths during 1981–2008 (online Technical
Appendix, wwwnc.cdc.gov/EID/article/19/3/12-1521-Technapp1.pdf).
Light gray, Whites; medium gray, Blacks; dark gray, Hispanics.
Despite the resurgence of TB in the United States during 1985–1992 (18), overall case rates unexpectedly decreased in some metropolitan areas before the availability of HAART because of improvements in TB control (19); in California, specific declines in TB–HIV began before the widespread use of HAART in 1996 (Figure 1). Improvements in programmatic TB control and declines in new annual HIV infections through the 1980s resulting from HIV awareness and prevention programs might have contributed to TB–HIV declines independent of HAART availability.

As in other high-income settings (20,21) and consistent with TB and TB–HIV (22) trends in the United States, immigrants from low-income countries increasingly represent the face of TB–HIV in California. Increased risk for TB–HIV among newcomers from sub-Saharan Africa reflects the hyperendemic nature of TB and HIV in this region. Yet, the demographic transformation of TB–HIV in California has been characterized largely by immigrants from Mexico or Central America and, to a lesser extent Southeast Asia, regions that have concentrated TB–HIV among Blacks. In the United States, new HIV infections (26) and TB–HIV (22,27) remain concentrated among Blacks. However, Hispanics have the highest TB–HIV rates in California, a finding corroborated by reports from municipalities along the US–Mexico border (28). Current rates of TB among HIV-infected Hispanics in California exceed rates in many World Health Organization–defined high-prevalence TB countries but are modest compared with the high rates in southern African countries (29) or marginalized HIV-infected populations in pre-HAART–era United States (30).

The dramatic decline in TB among HIV-infected persons in California has not been accompanied by a concurrent decrease in the proportion of patients who have advanced AIDS or late diagnosis of HIV infection, and TB rates continue to far exceed background TB incidence in the state. The distribution of CD4+ T-lymphocyte counts at TB diagnosis in California, even in the late HAART era, is not dissimilar to that found in studies from sub-Saharan Africa (31) or Southeast Asia (32). The opportunities missed are costly and underscore the need for early and innovative approaches to reach immigrants who are at particular risk for both infections. Implementation research that improves understanding of barriers to HAART as provided through key existing programs, such as the AIDS Drug Assistance Program and the Ryan White HIV/AIDS Program, is needed.

Although contrary to findings in much of the published literature from regions of high TB incidence, the higher prevalence of sputum smear positivity among HIV-infected patients is consistent with that found in prior studies from the United States (33). The prevalence of smear positivity is a function of multiple factors that affect the denominator (e.g., completeness of case reporting, reference standard testing) and numerator (e.g., quality of laboratory services, including staff workload and smear microscopy methods),

Table 3. Multivariate analysis of factors associated with deaths among HIV-infected TB patients, California, USA, 1993–2008*  

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Adjusted relative risk (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time period</td>
<td></td>
</tr>
<tr>
<td>2001–2008</td>
<td>Referent</td>
</tr>
<tr>
<td>1996–2000</td>
<td>1.18 (0.98–1.41)</td>
</tr>
<tr>
<td>1993–1995</td>
<td>2.21 (1.86–2.60)</td>
</tr>
<tr>
<td>Age†</td>
<td></td>
</tr>
<tr>
<td>Female Sex</td>
<td>1.29 (1.22–1.36)</td>
</tr>
<tr>
<td>1.36 (1.12–1.65)</td>
<td></td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td></td>
</tr>
<tr>
<td>White non-Hispanic</td>
<td>Referent</td>
</tr>
<tr>
<td>0.86 (0.73–1.01)</td>
<td></td>
</tr>
<tr>
<td>Black non-Hispanic</td>
<td>0.86 (0.67–1.09)</td>
</tr>
<tr>
<td>Asian/Pacific Islander</td>
<td>0.70 (0.11–4.52)</td>
</tr>
<tr>
<td>Foreign birth</td>
<td>0.65 (0.36–1.17)</td>
</tr>
<tr>
<td>HIV risk group‡</td>
<td></td>
</tr>
<tr>
<td>MSM</td>
<td>Referent</td>
</tr>
<tr>
<td>1.02 (0.88–1.19)</td>
<td></td>
</tr>
<tr>
<td>IDU</td>
<td>0.58 (0.43–0.78)</td>
</tr>
<tr>
<td>Unknown</td>
<td>1.24 (1.04–1.48)</td>
</tr>
<tr>
<td>Spumor smear positivity</td>
<td>1.23 (1.07–1.40)</td>
</tr>
<tr>
<td>CD4+ T-lymphocyte count, cells/mm³ §</td>
<td></td>
</tr>
<tr>
<td>&lt;50</td>
<td>6.45 (2.67–15.58)</td>
</tr>
<tr>
<td>50–99</td>
<td>5.57 (2.40–13.90)</td>
</tr>
<tr>
<td>100–199</td>
<td>3.09 (1.28–7.46)</td>
</tr>
<tr>
<td>200–349</td>
<td>1.47 (0.58–3.72)</td>
</tr>
<tr>
<td>350–499</td>
<td>1.99 (0.66–6.08)</td>
</tr>
<tr>
<td>TB as AIDS-defining diagnosis¶</td>
<td>1.22 (1.07–1.38)</td>
</tr>
</tbody>
</table>

*TB, tuberculosis; MSM, men who have sex with men; IDU, injection drug user.
†Per 10-year increase in age.
‡Categories are mutually exclusive; any IDU was included in the IDU category.
§CD4+ T-cell counts were imputed (online Technical Appendix, www.cdc.gov/EID/article/12-1521-Techapp1.pdf).
¶TB was considered the AIDS-defining event if TB and AIDS were reported within 6 months of each other.
and a combination of these factors probably contribute to differences in sputum smear positivity noted in areas of high versus low TB incidence.

Deaths associated with TB–HIV have declined remarkably in California since the pre-HAART era, consistent with national trends (7). Yet, although postmortem diagnoses of TB decreased significantly among patients with TB and with TB–HIV, patients with TB–HIV remain more than twice as likely to die during anti-TB treatment in the late HAART era as patients without known HIV co-infection. Women with HIV infection in the United States and other high-income areas have lagged behind men with respect to declines in mortality during the HAART era (34).

Our study has some potential limitations. First, TB–HIV cases may be underreported because TB patients without matches in the state HIV/AIDS registry were classified as HIV-negative. Persons who did not undergo HIV testing or whose providers did not adhere to the CDC AIDS case definition at TB diagnosis might not have been reported. However, assuming HIV awareness and testing have increased over time (35), misclassification of TB–HIV cases has declined in a time-dependent fashion. Moreover, during 1996–2006, the sensitivity of our case match procedure for capturing AIDS cases was ≈98% (95% CI 97.3%–98.7%) (36), and since confidential name-based HIV reporting began in 2006, <5% of TB–HIV cases were not also reported as AIDS cases (data not shown). Second, because antiretroviral therapy was unavailable in California Department of Public Health HIV/AIDS surveillance data, risk stratification according to HAART was not possible. Third, HIV prevalence estimates are a function of multiple parameters, some of which (race-stratified HAART-coverage, HIV incubation period, and migration patterns) carry considerable uncertainty. However, our estimates were subjected to multiple sensitivity analyses and are broadly consistent with extended back-calculation procedures undertaken by CDC (26). Applying national HIV rates to California would yield slightly higher HIV prevalence estimates. Last, trends in TB–HIV in California may not be generalizable to the United States as a whole.

In California, declines in TB–HIV incidence and death in the HAART era have been accompanied by a demographic shift toward foreign-born persons, particularly from Mexico and Central America. The opportunities for preventing TB and AIDS among foreign-born persons are underappreciated. Documentation of HIV status for TB patients in California (66% in 2008) remains below the national average (80%) (37). Screening and treatment completion rates for latent TB infection also are suboptimal despite the national standard and federal benefit (through the Health Resources and Services Administration) covering this practice (38). Improvements are hoped for with the recent availability of shorter course latent TB infection regimens (39). Progress toward comprehensive TB–HIV surveillance and recent lifting of legal barriers to HIV reporting at TB diagnosis are further cause for encouragement (40). TB and HIV programs must collaborate to monitor the confluence and changing epidemiology to ensure early detection of HIV and TB and to avert preventable deaths.

Acknowledgment

We thank the Office of AIDS and all local TB and HIV/AIDS staff in California’s health departments who reported the data used in this analysis. We also thank Eric Vittinghoff for his numerous insights and assistance with biostatistical methods.


This work was supported by the National Institutes of Health (K23 AI094251 to J.Z.M. and R01 AI076476 to P.C.H.) and the AMFDP/Robert Wood Johnson Foundation (J.Z.M.).

Dr Metcalfe is an assistant professor in the Division of Pulmonary and Critical Care Medicine at the University of California, San Francisco, and a doctoral candidate in the Division of Epidemiology at the University of California, Berkeley. His main research interests include TB and HIV co-morbidity in resource-limited settings, drug-resistant TB, TB diagnostics, and TB control in the United States.

References


Tuberculosis and HIV Co-infection, California, 1993–2008

Technical Appendix

I. Registry Cross-match

Prior to April 17, 2006, name-based reporting was restricted to confidential AIDS cases in California; after this date, confidential reporting of all verified cases of HIV infection was also required. Since the US Centers for Disease Control and Prevention (CDC) 1993 AIDS case definition, all patients with active TB and HIV infection are counted as AIDS cases (1). TB patients with HIV co-infection were identified through a statewide registry match with the California Office of AIDS using Registry Plus Link Plus software (2), a probabilistic record linkage program developed by CDC. Registry cross-match criteria included name, sex, race/ethnicity, and date and place of birth. Manual review of all matched cases was performed, with only those matches above a predetermined priority threshold considered to represent a co-infected case (3). Demographic, behavioral, and clinical information, including mortality, was abstracted from state surveillance forms (Report of a Verified Case of Tuberculosis and Adult HIV/AIDS Confidential Case Report). Surveillance data for both diseases have demonstrated high validity (4,5).

II. Nonparametric Back-calculation

Nonparametric back-calculation (6) calculates the HIV incidence rate based on AIDS diagnoses, using the incubation period distribution. During the highly active antiretroviral therapy (HAART) era, initiation of HAART prior to AIDS diagnosis by a fraction of patients leads to a slowing of HIV progression and delays AIDS diagnosis, effectively lengthening the average incubation period. Because the fraction of individuals initiating therapy is not well characterized by race over time in California, the realized distribution of the waiting time between HIV infection and AIDS diagnosis is not well known, and thus, while back-calculation may yield...
informative estimates of the HIV incidence, highly precise estimates of the incidence of HIV infection cannot be obtained.

We model the number of California diagnoses in ethnicity $k$ during year $i$ as follows. Let $I_{ik}$ denote the number of individuals infected during year $i$ for ethnicity $k$, and let $F_{ji}$ denote the probability of diagnosis in year $j$ given infection in year $i$. In general, observed diagnosis counts in any given year depend on the following: 1) changes in the diagnosis definition as occurred, for instance, in 1992; 2) delays in reporting; and 3) migration of HIV-positive case-patients into and out of California prior to AIDS diagnosis. Note also that, in general, the quantities $F_{ji}$ depend on the year $i$, due to treatment because individuals who have initiated HAART have a slowing rate of progression due to suppression of viral replication. In the absence of these features, we may use the method of nonparametric back-calculation to compute an estimate of the effective number infected during year $i$ as follows. The effective number infected is, in general, smaller than the true number of infected, because some of those truly infected initiated HAART and were therefore less likely to become diagnosed with AIDS than individuals who had not initiated HAART; the effective number infected provides an approximate equivalent untreated HIV population size, which may be useful in computing incidence rates, but is not useful for public health planning.

To compute the nonparametric back-calculation estimate, we assume independent Poisson counts for the diagnoses in each year (effectively ignoring diagnosis delay, a reasonable assumption given the time that has elapsed since the counts we are analyzing began). Then we have the following expected number of diagnoses in year $i$: $\lambda_i = \sum_{i \leq i} F_{ij} I_j$; for simplicity, we suppress the subscript $k$ for ethnicity in what follows. Let $i$ denote the vector of diagnoses from year $i=0$ to year $N$. Then, the log likelihood of a given sequence of counts $y_i$ is

$$\log L = \log \left( \prod_{i=0}^{N} e^{-\lambda_i} \frac{\lambda_i^{y_i}}{y_i!} \right) = \sum_{i=0}^{N} \left( -\lambda_i + y_i \log(\lambda_i) \right) + c \quad \text{(where $c$ is a term constant in the unknowns $I_{i,j}$)}.$$

Following Bacchetti et al. (1993), we add the following penalty term. Let $\theta_i = \log(I_i)$; the penalty is $q = \sum_{i=0}^{N-2} (\theta_i - 2\theta_{i+1} + \theta_{i+2})^2$. We then maximize $A(I) = \log(L) - kq$, where 1980 corresponds to year 0, and $N=29$ (for a total of 30 years of data, i.e., through 2009).
While the elegance of the nonparametric bootstrap derives in part from its relative freedom from
mechanistic assumptions, our application of this method to the post-HAART period requires
consideration of the change in the incubation period distribution at the population level caused
by a fraction of people starting therapy. Unfortunately, the rate of testing and the rate of starting
therapy at different times and for different CD4 progression stages are simply not known. We
therefore derived incubation period distributions by modifying the HIV staging model developed
by Longini et al (1989) (7) using the methods in Charlebois et al. (2011) (8); specifically, we
derive differential equations that we will use to derive the quantities $F_{ji}$. To derive them, we will
use a model of untreated HIV, but then assume that a certain fraction of individuals commence
therapy and that their progression is slowed. For untreated individuals, the 3 stages are 1)
infected but not seropositive 2), seropositive but asymptomatic, and 3) pre-AIDS symptoms and
abnormal hematology (see Longini et al. 1989 for details). We used empirical estimates for the
progression rates between these stages (7), yielding an untreated mean incubation period of
approximately 10 years (consistent with other findings, e.g. Cooley et al., 1996) (9). For treated
individuals, we assume the same stages, but with a smaller progression rate between them. We
also classify individuals by status ($k$) (1): person does not know he/she are infected (2), person
known HIV positive but untreated (3), person has initiated HAART but has not yet achieved full
virologic suppression, and (4) person on HAART and on treatment with the maximum virologic
suppression he/she will achieve. We let $Y_{jk}$ denote the number of HIV-infected people in stage $j$
and status $k$. We denote the testing rate in stage $j$ at time $t$ by $\theta_j(t)$, the rate of initiating
suppressive therapy by $\phi_j(t)$, the rate of achieving full virologic suppression by $\rho$, the rate of
discontinuation of therapy for individuals in virologic suppression by $\alpha$, the progression rate
among untreated individuals from stage $j$ to stage $j+1$ by $\gamma_j$ ($j=1, 2, 3$), and the relative
progression rate in status $k$ by $s_k$ ($s_1=1, s_2=1$). Denote the incidence rate by $I(t)$ (assumed
piecewise constant within each year). Mortality due to non-AIDS causes is assumed to be
negligible for simplicity. The equations are then

$$
\frac{dY_{jk}}{dt} = I(t)\delta_{j1}\delta_{k1} + Y_{j1}\theta_j(-\delta_{j1} + \delta_{j2}) + Y_{j2}\phi_j(-\delta_{j2} + \delta_{j3}) + \rho Y_{j3}(-\delta_{j3} + \delta_{j4}) +
Y_{j4}\alpha(-\delta_{j4} + \delta_{j2}) - Y_{jk}s_k\gamma_j + Y_{j-1} s_k Y_{j-1} (1 - \delta_{j1})
$$
where we separate subscripted indices with a comma when the index is an expression and not a single symbol, and $\delta_{ij}$ is a Kronecker delta. Numerical integration using the R package “odesolve” allows the matrix elements $F_{ji}$ to be computed (the initial conditions are $Y_{jk} = 0$ for all $j$ and $k$).

The numerical parameter estimates were as follows. Testing is assumed to be zero before time 5 (year 1985); treatment is zero before time 16 (year 1996). The progression parameters were derived from Longini et al (1989); $\gamma_1 = 5.49 \text{ yr}^{-1}$, $\gamma_2 = 0.228 \text{ yr}^{-1}$, and $\gamma_3 = 0.191 \text{ yr}^{-1}$. Unfortunately, population-based estimates for the testing and treatment rates by race over time are unavailable; we assumed that individuals have a testing rate of 0.3 yr$^{-1}$ prior to AIDS diagnosis (a rate which would underestimate the rate for urban men who have sex with men), a rate of cessation of antiretroviral therapy of 0.01 yr$^{-1}$, a rate of 4 yr$^{-1}$ of achieving full virologic suppression (i.e., a mean duration of 3 months before achieving full suppression), and that individuals on full suppression only progress at 10% of the rate of untreated individuals (which may in fact be an overestimate). In the base case scenario, we chose the rate of treatment in stage $j=2$ to be $\frac{1}{4}$ of that in stage $j=3$, and chose the rate in stage 3 so that by 2008, 50% of individuals with HIV but no AIDS diagnosis would be on treatment. Because many of these parameters are not well characterized, we conducted sensitivity analysis.

Of particular importance are the rates of treatment and testing by race. For the base case scenario, we assumed all rates were identical; the TB rates computed from these denominators thus contain no built-in assumption of differential rates of starting therapy.

We used parametric bootstrap (Efron & Tibshirani [10], p. 53–5) to compute confidence intervals for rates (with 100 replications). Specifically, for the denominator data, we assumed the estimated race-stratified incubation period over time and simulated AIDS incidence data given the random incubation period distribution for each year. We then estimated the HIV incidence from these simulated data, and from the HIV incidence, we computed the prevalence as indicated above; this yields a set of changing prevalences over time reflecting the expected variability given the sample sizes. For each of these, we simulated numerator (TB case) data using an independent Poisson distribution whose mean equaled the observed number of cases for the year in question, and computed the rate given these simulated numerators and denominators, yielding confidence intervals reflecting variability in the numerator and denominator.
Confidence intervals from parametric bootstrap do not reflect variability based on modeling assumptions. We therefore conducted the following sensitivity analyses (Technical Appendix Table 1) (1): assuming (unrealistically) no pre-AIDS treatment at all (2), assuming constant treatment rates sufficient to place 90% of HIV patients on HAART prior to diagnosis by 2009 (3), assuming treatment rates identical to the base case for 1996–2000, but assuming treatment rates 20% larger in 2001–2005, and a further 20% larger for 2006 to the present. We also conducted sensitivity analyses using different values for the smoothing parameter; values of 0 lead to wildly unstable estimates and are unsuitable (6). Finally, the 1992 (and earlier) AIDS case definition changes were not explicitly modeled; sensitivity analysis in which 10% of cases diagnosed in 1992 would have occurred in 1993 and 1994, however, did not substantially change our estimates.

Optimization was conducted by the Nelder-Mead method, using 25 randomly chosen starting values to reduce the chance of convergence to a merely local maximum. All calculations were conducted by using R v. 2.13 (http://cran.r-project.org) for MacIntosh (Apple, Cupertino, CA, USA).

III. Multiple Imputation

Multiple imputation allows asymptotically unbiased estimation of missing data under an assumption of missing at random conditional on measured variables (11). To recover missing values for CD4+ T-lymphocyte count (1,732 of 3,904 [44.4%]) and viral load (2,756 of 3,904 [70.6%]), we constructed 10 imputation data sets (12) per individual using Stata 12.2 (Stata Corporation, College Station, TX, USA) ‘mi’ suite of commands. CD4+ T-lymphocyte counts or HIV viral load not recorded or ascertained >6 months prior to TB diagnosis were considered “missing.” Demographic (age, sex, race/ethnicity, nationality, homeless status, correctional facility) and clinical variables (year of diagnosis, time from initial HIV diagnosis to TB diagnosis, time from initial HIV diagnosis to AIDS event, whether TB was the AIDS defining event, mode of HIV transmission, vital status at time of TB diagnosis, and tuberculin skin test reading) were used to impute missing CD4+ T-lymphocyte count and viral load values. Demographic, behavioral, and clinical information, including mortality, were abstracted from state surveillance forms (Report of a Verified Case of Tuberculosis, and Adult HIV/AIDS
Confidential Case Report). TB was considered the AIDS-defining event if TB and AIDS were reported within 6 months of each other. For each individual, regression analysis was done on each imputation data set, and results were combined to ascertain estimates by using standard multiple imputation techniques (13). CD4+ T-lymphocyte count was square root transformed and viral load was log-transformed prior to imputation. In order to avoid working with implausible imputations in the analysis, square root transformed CD4+ T-lymphocyte counts with negative values were truncated to 1 prior to back-transformation. Multivariate associations with mortality obtained following multiple imputation and from an unimputed “complete-case” analysis are demonstrated in the main text and below, respectively (Table 3; Technical Appendix Table 2). Overall goodness of fit for the complete-case analysis was assessed by using the Hosmer-Lemeshow test.

### Technical Appendix Table 1. Nonparametric back-calculation sensitivity analyses

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>TB rate, 1993</th>
<th>TB rate, 2008</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>White</td>
<td>Black</td>
</tr>
<tr>
<td>Baseline</td>
<td>0.00</td>
<td>202.30</td>
<td>768.6</td>
</tr>
<tr>
<td>Rate of testing</td>
<td>1.00</td>
<td>202.41</td>
<td>762.23</td>
</tr>
<tr>
<td>Treatment rate</td>
<td>0</td>
<td>205.95</td>
<td>775.75</td>
</tr>
<tr>
<td>Cessation rate</td>
<td>0.02</td>
<td>202.23</td>
<td>776.00</td>
</tr>
<tr>
<td>Cessation rate</td>
<td>0.01</td>
<td>201.25</td>
<td>772.20</td>
</tr>
<tr>
<td>Relative progression</td>
<td>0.05</td>
<td>202.22</td>
<td>767.02</td>
</tr>
<tr>
<td>Relative progression</td>
<td>0.20</td>
<td>203.01</td>
<td>740.13</td>
</tr>
<tr>
<td>Treatment fraction in 2008</td>
<td>0.90</td>
<td>191.78</td>
<td>743.86</td>
</tr>
<tr>
<td>Increase rate, 5 y</td>
<td>25%</td>
<td>198.50</td>
<td>776.40</td>
</tr>
<tr>
<td>Case definition change</td>
<td>0</td>
<td>202.30</td>
<td>769.96</td>
</tr>
</tbody>
</table>

### Technical Appendix Table 2. Multivariate analysis of factors associated with deaths among HIV-infected TB patients, California, 1993–2008

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Adjusted relative risk (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time period</td>
<td></td>
</tr>
<tr>
<td>2001–2008</td>
<td>Referent</td>
</tr>
<tr>
<td>1996–2000</td>
<td>1.07 (0.83–1.36)</td>
</tr>
<tr>
<td>Age†</td>
<td>1.29 (1.22–1.36)</td>
</tr>
<tr>
<td>Sex, F</td>
<td>1.33 (1.03–1.73)</td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td></td>
</tr>
<tr>
<td>White, non-Hispanic</td>
<td>Referent</td>
</tr>
<tr>
<td>Black, non-Hispanic</td>
<td>0.85 (0.66–1.09)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>1.03 (0.77–1.36)</td>
</tr>
<tr>
<td>Asian/Pacific Islander</td>
<td>0.93 (0.58–1.47)</td>
</tr>
<tr>
<td>Foreign birth</td>
<td>0.68 (0.53–0.88)</td>
</tr>
<tr>
<td>HIV risk group‡</td>
<td></td>
</tr>
<tr>
<td>MSM</td>
<td>Referent</td>
</tr>
<tr>
<td>IDU</td>
<td>1.08 (0.87–1.34)</td>
</tr>
<tr>
<td>Heterosexual contact</td>
<td>0.58 (0.39–0.86)</td>
</tr>
<tr>
<td>Unknown</td>
<td>1.35 (1.08–1.70)</td>
</tr>
<tr>
<td>Sputum smear positivity</td>
<td>1.23 (1.02–1.49)</td>
</tr>
<tr>
<td>CD4+ T-lymphocyte count, cells/mm³</td>
<td></td>
</tr>
<tr>
<td>&lt;50</td>
<td>5.21 (2.14–12.64)</td>
</tr>
<tr>
<td>50–99</td>
<td>3.69 (1.50–9.07)</td>
</tr>
<tr>
<td>100–199</td>
<td>2.92 (1.20–7.14)</td>
</tr>
<tr>
<td>200–349</td>
<td>2.25 (0.88–5.80)</td>
</tr>
<tr>
<td>350–499</td>
<td>Referent</td>
</tr>
<tr>
<td>&gt;500</td>
<td>1.59 (0.50–5.06)</td>
</tr>
<tr>
<td>Characteristic</td>
<td>Adjusted relative risk (95% CI)</td>
</tr>
<tr>
<td>---------------------------------------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>TB as AIDS-defining diagnosis§</td>
<td>0.94 (0.78–1.13)</td>
</tr>
</tbody>
</table>

*Complete-case analysis (CD4+ T-cell counts not imputed. TB, tuberculosis; MSM, men who have sex with men; IDU, injection drug use. †Per 10 year increase in age. ‡Categories are mutually exclusive; any IDU was included in the IDU category. §TB was considered the AIDS-defining event if TB and AIDS were reported within 6 months of each other.

References


