

Transmission of Tuberculosis in a South African Community With a High Prevalence of HIV Infection

Keren Middelkoop,^{1,2} Barun Mathema,^{6,7} Landon Myer,³ Elena Shashkina,⁷ Andrew Whitelaw,^{4,5} Gilla Kaplan,⁸ Barry Kreiswirth,⁷ Robin Wood,^{1,2} and Linda-Gail Bekker^{1,2}

¹Desmond Tutu HIV Centre, Institute of Infectious Disease and Molecular Medicine, ²Department of Medicine, and ³Division of Epidemiology and Biostatistics, School of Public Health and Family Medicine, University of Cape Town, ⁴Division of Medical Microbiology, University of Stellenbosch, and ⁵National Health Laboratory Service, Tygerberg Hospital, Cape Town, South Africa; ⁶Department of Epidemiology, Mailman School of Public Health, Columbia University, New York; and ⁷Public Health Research Institute Tuberculosis Center, and ⁸Laboratory of Mycobacterial Immunity and Pathogenesis, Public Health Research Institute, New Jersey Medical School–Rutgers, The State University of New Jersey, Newark

(See the editorial commentary by Borgdorff, Cain and DeCock on pages 8–9.)

Background. In settings of high tuberculosis transmission, little is known of the interaction between human immunodeficiency virus (HIV) positive and HIV-negative tuberculosis disease and of the impact of antiretroviral treatment (ART) programs on tuberculosis transmission dynamics.

Methods. *Mycobacterium tuberculosis* isolates were collected from patients with tuberculosis who resided in a South African township with a high burden of tuberculosis and HIV infection. Demographic and clinical data were extracted from clinic records. Isolates underwent IS6110-based restriction fragment length polymorphism analysis. Patients with unique (nonclustered) *M. tuberculosis* genotypes and cluster index cases (ie, the first tuberculosis case in a cluster) were defined as having tuberculosis due to reactivation of latent *M. tuberculosis* infection. Secondary cases in clusters were defined as having tuberculosis due to recent *M. tuberculosis* infection.

Results. Overall, 311 *M. tuberculosis* genotypes were identified among 718 isolates from 710 patients; 224 (31%) isolates were unique strains, and 478 (67%) occurred in 87 clusters. Cluster index cases were significantly more likely than other tuberculosis cases to be HIV negative. HIV-positive patients were more likely to be secondary cases ($P = .001$), including patients receiving ART ($P = .004$). Only 8% of cases of adult-adult transmission of tuberculosis occurred on shared residential plots.

Conclusions. Recent infection accounted for the majority of tuberculosis cases, particularly among HIV-positive patients, including patients receiving ART. HIV-negative patients may be disproportionately responsible for ongoing transmission.

Keywords. tuberculosis; transmission; HIV; molecular epidemiology; genotyping.

Despite recent gains in the global fight against tuberculosis, this disease remains a significant cause of morbidity and mortality in developing countries [1], and additional

strategies are required in combination with directly observed therapy, short-course (DOTS) to control endemic tuberculosis. These adjunctive strategies will need to be tailored to the specific nature of each epidemic. For example, where *Mycobacterium tuberculosis* infection rates are especially high and progression from recent infection is an important driver of the tuberculosis burden, an effective intervention to decrease transmission may have the greatest impact on tuberculosis control [2].

Molecular epidemiology has provided the tools to explore the diversity of infecting *M. tuberculosis* strains, the roles of reactivation and recent transmission in different settings, and the clinical and social factors

Received 25 March 2014; accepted 23 May 2014; electronically published 22 July 2014.

Presented in part: 6th International AIDS Society Conference on HIV Pathogenesis, Treatment, and Prevention, Rome, Italy, 17–20 July 2011. Abstract TUPE163.

Correspondence: Keren Middelkoop, PO Box 13801, Mowbray Observatory, Cape Town 7705, South Africa (keren.middelkoop@hiv-research.org.za).

The Journal of Infectious Diseases® 2015;211:53–61

© The Author 2014. Published by Oxford University Press on behalf of the Infectious Diseases Society of America. All rights reserved. For Permissions, please e-mail: journals.permissions@oup.com.

DOI: 10.1093/infdis/jiu403

associated with transmission. In the developed world, reactivation of previously acquired latent infection is responsible for a greater proportion of tuberculosis cases than recent transmission [3], whereas in sub-Saharan Africa, studies have reported that recent infection is responsible for much of tuberculosis disease [4–6].

Risk factors for *M. tuberculosis* transmission, such as young age [5], prior imprisonment [6], and visitation of social settings such as bars and churches [6], have been identified in various lower-middle income countries. Some studies from African countries with a high tuberculosis incidence have reported a lack of association between human immunodeficiency virus (HIV) infection and *M. tuberculosis* transmission [6–8], with one study reporting an increased proportion of tuberculosis cases due to recent infection among older HIV-positive patients [5]. However, there are few molecular epidemiological studies describing the interaction between HIV-positive and HIV-negative tuberculosis disease, and no studies have reported on the impact of a high coverage antiretroviral treatment (ART) program on *M. tuberculosis* transmission.

We assessed the interaction between HIV-positive and HIV-negative tuberculosis disease and the influence of both the HIV prevalence and a high-coverage ART program on *M. tuberculosis* transmission in a community where the burdens of tuberculosis disease and HIV infection are high.

METHODS

The study community has been described elsewhere [9–11]: it is a typical South African periurban township with clear community boundaries which houses a growing population of approximately 18 000 people. An estimated 10% of the community population migrates in per year, and emigration averages 4% [12]. The community consists of a formal sector with individually numbered serviced plots (approximately 78% of the community) and an informal sector of shacks sharing communal services. From 1 to 22 houses are built on a single plot (average density, 4 houses/plot). The prevalence of HIV infection among adults in the community is 23%–25% [10, 13, 14]. ART provision was initiated in 2005, and the program has followed national treatment guidelines: patients with CD4⁺ T-cell counts of <200 cells/μL or World Health Organization (WHO) stage 4 HIV disease eligible for ART [15], with the exception that from 2005 to 2008, patients with a CD4⁺ T-cell count of <350 cells/μL or WHO stage 3 or 4 disease were eligible for treatment [16]. The single primary care clinic provides tuberculosis care as per the South African National Tuberculosis Control Program guidelines [17], employing passive case detection with WHO-recommended DOTS. Despite the extensive ART program and tuberculosis treatment completion rates of nearly 80% [11, 18] high tuberculosis notification rates of approximately 2000 cases/100 000 persist

in this community [11]. Furthermore, the community *M. tuberculosis* transmission rate is high, with an annual risk of *M. tuberculosis* infection of 4.1% [19].

Patient Population

The study population comprised of all tuberculosis patients, resident in the community and notified from 2001 to 2010. Demographic and clinical data, were extracted from the tuberculosis register and clinic records. Social epidemiological data, including places of social congregation, were obtained by interviewer-administered questionnaires completed by patients with tuberculosis from 2006 to 2010. The study was approved by the Human Research Ethics Committee of the University of Cape Town and the Institutional Review Board of Rutgers, the State University of New Jersey. Patients provided written informed consent.

Laboratory Procedures

The sputum specimen collection and investigations have been previously described [9], but in short complied with the National Tuberculosis Program diagnostic protocol [17]. In addition, from 2006 all smear-negative specimens were also cultured on Lowenstein–Jensen (LJ) medium and isoniazid (INH) and rifampicin (RIF) susceptibility testing was extended to include all cultured specimens, regardless of clinical indication. All positive culture isolates were stored at –70°C at the research laboratory at the Institute of Infectious Disease and Molecular Medicine at the University of Cape Town. Culture stocks were shipped to the Public Health Research Institute (PHRI) Tuberculosis Center at Rutgers, the State University of New Jersey.

Genotyping Analysis

IS6110-based Restriction Fragment Length Polymorphism (RFLP) genotyping and analysis of *M. tuberculosis* strains performed at the PHRI Tuberculosis Center has been described in detail previously [9]. A strain family was defined as a group of strains that exhibited similar but not identical IS6110 hybridization profiles, suggesting relatedness by descent (eg, W-Beijing family) [20]. Strains with <6 copies of IS6110 are poorly differentiated by RFLP analysis and were excluded from analysis [20].

Definitions

An adult tuberculosis case was defined as a patient aged ≥15 years with tuberculosis. A unique strain was defined as an isolate with an RFLP pattern that occurred in only 1 patient within the study data set. Strain clusters were defined as >1 occurrence of a specific strain in different individuals during the study period [21]. The index case in each cluster was temporally defined as the first case diagnosed within the cluster, and subsequent patients in the clusters were defined as secondary cases. Unique cases and cluster index cases were presumed to be due to reactivation of latent *M. tuberculosis* infection or to recent infection acquired outside the study community (referred to as “reactivation cases” below), and secondary cases in clusters presumed to

be due to recent exogenous infection (“n-1 method” [21]). Retreatment tuberculosis cases, resulting from relapse of the same strain in the same individual, were excluded from clustering analyses; any case of retreatment tuberculosis that was due to a *M. tuberculosis* strain that differed from the strain causing the previous disease episode was included.

Data Analysis

Data were analyzed using Stata, version 13.0 (StataCorp, College Station, TX). Bivariate analyses used Wilcoxon rank sum and χ^2 tests, as appropriate. Multivariate logistic regression models were developed to examine factors associated with dominant strain families and with unique, index, and secondary cases. All statistical tests were 2-sided at an α level of 0.05. The analysis of the geographical distribution *M. tuberculosis* strains was restricted to residents in the formal sector, who use the assigned plot number as their address. The ArcMap 10 (Esri) geographic information system (GIS) was used to illustrate the spatial distribution of *M. tuberculosis* strains occurring in pairs during the course of the study.

RESULTS

From 2001 to 2010, there were 1353 smear- and/or culture-positive pulmonary tuberculosis cases diagnosed in the community. *M. tuberculosis* cultures were obtained from 1011 sputum-positive cases (75%), and 789 had RFLP patterns available (78% of collected specimens). Reasons for missing specimens or genotyping data are described elsewhere [9]. Patients with RFLP data did not differ from those without RFLP data with regard to age, sex, HIV status, or ART use (Table 1). We were less likely to obtain RFLP data for patients with tuberculosis who died ($P = .01$) and more likely to obtain RFLP data for patients with MDR tuberculosis ($P = .02$). There were no differences in other tuberculosis outcomes or the proportion of new cases and retreatment cases between patients with and those without RFLP data.

While genotyping data were available for 789 tuberculosis episodes, strains identified in 79 patients had <6 copies of IS6110. These patients did not differ from those with strains having ≥ 6 copies of IS6110 in terms of age, sex, HIV or ART status, new or retreatment tuberculosis cases, or tuberculosis treatment outcomes (data not shown).

The median age of the 710 patients with RFLP results was 32 years (interquartile range [IQR], 26–39 years), 14 patients (2%) were children, and 410 (58%) were male. A total of 24 patients (3%) had MDR tuberculosis, and no cases of extensively drug-resistant tuberculosis were identified. Treatment outcomes were as follows: 550 (77%) completed tuberculosis treatment, 59 (8%) were transferred out of the community before treatment completion, 60 (8%) interrupted treatment, 29 (4%) died, 10 (1%) did not respond to treatment, and 2 with MDR tuberculosis

Table 1. Comparison of Demographic and Tuberculosis Characteristics Among patients With and Those Without *Mycobacterium tuberculosis* Genotyping Data

Characteristic	Patients Without RFLP Data (n = 564)	Patients With RFLP Data (n = 789)	P Value
Age, y, median (IQR)	33 (27–41)	33 (27–40)	.31
Male sex	322 (57)	462 (59)	.59
Tuberculosis category			
New case	392 (70)	561 (71)	.53 ^a
Retreatment after previous cure	74 (13)	84 (11)	.16 ^a
Retreatment after previous completion	71 (13)	101 (13)	.91 ^a
Retreatment after treatment interruption	23 (4)	38 (5)	.52 ^a
Retreatment after treatment failure	4 (1)	5 (1)	.87 ^a
Confirmed MDR tuberculosis	7 (1)	24 (3)	.02
Tuberculosis outcome			
Cured/treatment completed	427 (76)	611 (78)	.44 ^a
Treatment interrupted	43 (8)	66 (8)	.62 ^a
Treatment failure	2 (0.4)	11 (1)	.05 ^a
Transferred out of community	48 (9)	66 (8)	.92 ^a
Died	41 (7)	31 (4)	.01 ^a
HIV status ^b			
Known	501 (89)	721 (91)	.12
Negative	189 (38)	259 (36)	
Positive	312 (62) ^c	462 (64)	.52
Receiving ART	49 (16)	105 (23)	.80 ^d

Data are median value (interquartile range) or no. (%) of specimens.

Abbreviations: ART, antiretroviral therapy; HIV, human immunodeficiency virus; MDR, multidrug resistant; RFLP, restriction fragment length polymorphism.

^a Compared with all other categories/outcomes combined. Outcomes for 7 specimens were unknown.

^b Data exclude those with an unknown HIV status.

^c $P = .001$, compared with all other strains.

^d Adjusted for year, to account for ART rollout.

had an unknown treatment outcome. Of 648 patients (91%) who were tested for HIV, 412 (64%) were HIV positive; of these 412, 96 (23%) were receiving ART at the time of tuberculosis diagnosis. Of the 710 patients with RFLP data, 9 were infected with 2 different *M. tuberculosis* strains; for 1 patient, the second strain was not fully genotyped (it was a variant of W-Beijing). Therefore, 718 isolates had RFLP patterns available for analysis.

Dominant Strain Families

Over the 10-year study period, 311 discrete *M. tuberculosis* strains were identified within 53 families. The dominant strain families were W-Beijing (32%), CC-related (30%), and BM (5%;

Table 2. Demographic and Tuberculosis Disease Characteristics of *Mycobacterium tuberculosis* Strain Families Identified in the Community

Characteristic	W-Beijing Family (n = 229)	CC-Related Family (n = 214)	BM Family (n = 38)	Other Strain Families (n = 237)
<i>M. tuberculosis</i> strains in family/families, no.	52	84	26	149
Age, y	33 (27–40)	31 (25–39)	32 (28–41)	33 (27–39)
Male sex	128 (56)	126 (59)	24 (63)	136 (57)
Tuberculosis category				
New case	158 (69)	153 (71)	27 (71)	171 (72)
Retreatment	71 (31)	61 (29)	11 (29)	66 (28)
Drug resistance				
Any	18 (8)	9 (4)	3 (8)	16 (7)
Confirmed MDR	12 (5)	2 (1)	1 (3)	9 (4)
HIV positive				
Overall ^a	147 (72)	120 (63)	15 (44)	132 (58)
Receiving ART	49 (16)	24 (20)	3 (20)	31 (23)
CD4 ⁺ T-cell count, cells/μL	151 (44–291)	114 (43–235)	105 (74–223)	118 (55–243)

Data are median value (interquartile range) or no. (%) of *M. tuberculosis* strains, unless otherwise indicated.

Abbreviations: ART, antiretroviral therapy; HIV, human immunodeficiency virus; MDR, multidrug resistant.

^a Data exclude those with an unknown HIV status.

Table 2). When compared to all other strains, the W-Beijing family was associated with HIV infection ($P = .001$), an association that persisted after adjustment for age and sex (odds ratio [OR], 1.8; 95% confidence interval [CI], 1.2–2.5). Among HIV-positive patients, the W-Beijing family was not associated with ART status ($P = .27$) or CD4⁺ T-cell count ($P = .34$). W-Beijing strains were not associated with monodrug-resistant and/or poly-drug-resistant tuberculosis ($P = .28$), although there was a borderline association with MDR tuberculosis ($P = .05$). The CC-related family was not associated with HIV status ($P = .98$) but was negatively associated with MDR tuberculosis, even after adjustment for age, sex, and HIV status (OR, 0.13; 95% CI, .02–.9). The BM family strain was associated with HIV-negative status, even after adjustment for age and sex (OR, 0.44; 95% CI, .2–.9), but was not associated with MDR tuberculosis ($P = .80$). There was no statistical association between these dominant *M. tuberculosis* families and age or sex.

Clustering

Of the 718 *M. tuberculosis* isolates, 224 (31%) were unique strains, 478 (67%) were clustered (Table 3), and 16 (2%) were reactivation of a previous strain in retreatment patients. There were 87 strain clusters in this community, ranging in size from 2 to 85 patients (mean cluster size, 5.5 patients). Of these 87 cluster index and 391 secondary cases, 38 (44%) and 245 (63%), respectively, were HIV positive. Use of the $n - 1$ method calculated that 54% of tuberculosis cases were due to recent exogenous infection. At least 61% of clusters were composed of both HIV-positive patients and HIV-negative patients, with a further 9% possibly including both HIV-positive patients

and HIV-negative patients (the HIV status unknown for some individuals in clusters). Overall, 23% and 7% of clusters comprised HIV-positive patients only and HIV-negative patients only, respectively. All HIV-negative clusters were couplets, whereas HIV-positive clusters ranged from 2 to 4 patients in size (median, 2 patients; IQR, 2–3 patients).

Figure 1 shows the proportions of unique, index and secondary *M. tuberculosis* isolates and the distribution of HIV in each of these groups. Among the 14 childhood tuberculosis cases, 9 (64%) were in clusters, and 2 were the index case in their cluster (ages, 6 and 13 years). In a multivariate model, clustering was not associated with study year, age, sex, HIV status, or ART use. However, index cases were significantly more likely than all other tuberculosis cases to be HIV negative (OR, 1.9; 95% CI, 1.2–3.1), but the 2 groups of cases did not differ in terms of age ($P = .11$) or sex ($P = .28$). HIV-negative index cases were more likely to be sputum-smear positive, compared with HIV-positive index cases (97% vs 71%; $P = .002$). The association between index cases and HIV-negative patients persisted in multivariate analysis that adjusted for age, sex, and sputum-smear-positive status (OR, 1.9; 95% CI, 1.2–3.0).

The cohort constituted 311 cases (44%) that were unique or cluster index cases (ie, reactivation cases), of which 51% were known to be HIV positive (Table 3). Compared with recent infection cases (ie, secondary cases), reactivation cases were more likely to occur in HIV-negative patients (Figure 1), and this association persisted after adjustment for age and sex (OR, 1.7; 95% CI, 1.2–2.4). Among HIV-positive patients, reactivation cases were less likely to be receiving ART, even after adjustment for age and sex (OR, 0.5; 95% CI, .3–.8), but they were not

Table 3. Demographic, Clinical, and Tuberculosis Characteristics of Clustered and Nonclustered *Mycobacterium tuberculosis* Strains Identified in the Community

Characteristic	Nonclustered Strains (n = 224)	Clustered Strains (n = 478)	Reactivation (ie, Index and Unique) (n = 311)	Secondary Cases (n = 391)	Index Cases (n = 87)	Unique and Secondary Cases (n = 615)
Age, y	33 (26–40)	32 (27–39)	32 (26–39)	33 (27–39)	31 (25–37)	33 (27–39)
Sex						
Male	131 (58)	276 (58)	179 (58)	228 (58)	48 (55)	359 (58)
Female	93 (42)	202 (42)	132 (42)	163 (42)	39 (45)	256 (42)
Tuberculosis category						
New case	170 (76)	339 (71)	236 (76)	273 (70)	66 (76)	443 (72)
Retreatment	54 (24)	139 (29)	75 (25)	118 (30)	21 (24)	172 (28)
Smear positive	178 (80)	398 (83)	252 (81)	324 (83)	74 (85)	502 (82)
Confirmed MDR tuberculosis	7 (3)	16 (3)	9 (3)	14 (4)	2 (2)	21 (3)
HIV status						
Known	206 (92)	432 (90)	282 (91)	356 (91)	76 (87)	562 (91)
Negative	84 (41)	122 (59)	122 (43)	111 (31)	38 (50)	195 (35)
Positive						
Overall	149 (34)	283 (66)	160 (57)	245 (69)	38 (50)	367 (65)
Receiving ART	22 (18)	73 (26)	26 (16)	69 (28)	4 (11)	91 (25)
CD4 ⁺ T-cell count, cells/ μ L	130 (71–218)	124 (47–245)	130 (72–219)	117 (44–252)	166 (84–219)	124 (55–243)

Data are median value (interquartile range) or no. (%) of *M. tuberculosis* strains.

Abbreviations: ART, antiretroviral therapy; HIV, human immunodeficiency virus; MDR, multidrug resistant.

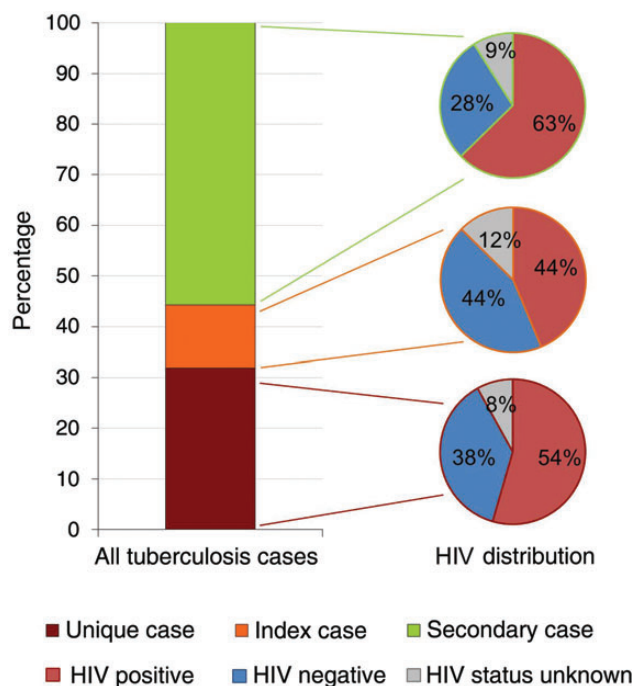


Figure 1. Distribution of unique, index and secondary *Mycobacterium tuberculosis* isolates, overall and by human immunodeficiency virus status.

associated with CD4⁺ T-cell count ($P = .84$). Conversely, secondary cases (composing 60% of the HIV-positive disease burden) were more likely to be HIV positive and receiving ART.

Among the 95 patients who were receiving ART at the time of tuberculosis diagnosis, 69 (73%) were secondary cases, 22 (23%) were unique cases, and 4 (4%) were index cases. The median time from starting ART to receiving the tuberculosis diagnosis was 46 months (IQR, 14–82 months). There was no difference in the median time from starting ART between secondary cases and either index cases ($P = .91$) or reactivation cases ($P = .18$).

Transmission in the Household and Other Locations

Of the 556 adult patients with tuberculosis living in the formal sector of the community, 249 (45%) shared a residential plot with 1 or more adults with tuberculosis. However, only 8% of adults shared a residential plot with another adult infected with the same *M. tuberculosis* strain. Figure 2 shows adult couplets distributed throughout the community. Of the 9 children included in clusters, 4 were on the same plot as an adult with tuberculosis who had the same *M. tuberculosis* strain, and 2 additional children who shared a residential plot were infected with the same strain.

Multivariate models that adjusted for age and sex showed that among the 514 patients who completed the social questionnaire,

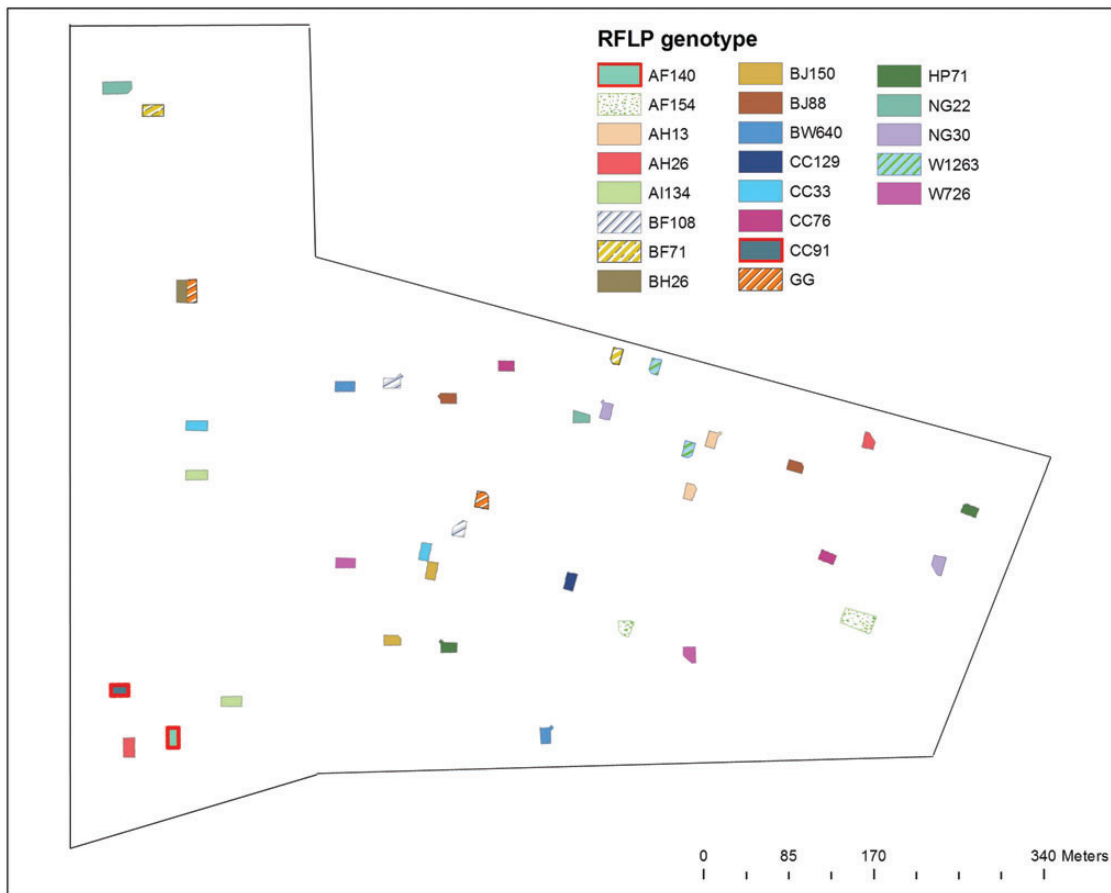


Figure 2. Couplet strains in the community. Thick red borders (AF140 and CC91) denote both patients with same strain on the same plot. Community orientation has been changed for purposes of confidentiality. Abbreviation: RFLP, restriction fragment length polymorphism.

clustering was not associated with shebeen (informal tavern), clinic/hospital, or church attendance in the previous 12 months ($P = .14$, $P = .71$, and $P = .78$, respectively). Clustering was also not associated with reported regular use of minibus taxis ($P = .43$).

DISCUSSION

This study community was geographically well defined, had a substantial tuberculosis and HIV burden, extensive diversity of *M. tuberculosis* strains and strain families, a high degree of tuberculosis treatment completion, and a high-coverage ART program. The study used GIS technology linked to genotyping and clinical information. Together, these factors enabled us to assess clustering and transmission in the community, both overall and by HIV and ART status. Recent transmission was responsible for as much as 54% of overall tuberculosis cases and for 60% of cases among HIV-positive patients. HIV-negative patients accounted for a disproportionately high number of cluster index cases.

It has been suggested that different *M. tuberculosis* lineages have adapted to specific human populations and that HIV infection may disrupt this geographic sympatric human host-*M. tuberculosis* relationship [22]. Here we report an association between W-Beijing strains and HIV-infected patients. Together with previous reports of W-Beijing as an emerging and diversifying *M. tuberculosis* family in Cape Town [23, 24], this finding may be an example of such a shift in host-pathogen relationships. Of interest was the lack of association between W-Beijing infection and ART treatment or $CD4^+$ T-cell count, which may indicate that this phenomenon is not driven by the extent of immune suppression, but rather is the result of an as yet not understood fundamental change in host protective immunity against *M. tuberculosis* W-Beijing infection following HIV infection.

Cluster analysis showed that HIV-positive tuberculosis disease and HIV-negative tuberculosis disease are not independent of one another, as indicated by the majority of clusters containing both HIV-positive and HIV-negative patients. Cluster index patients had nearly twice the odds of being HIV negative,

suggesting that HIV-negative patients may be disproportionately responsible for transmission in the community. Available data report that, following *M. tuberculosis* infection, HIV-positive patients progress more rapidly to tuberculosis, compared with their HIV-negative counterparts [25–27]. The more rapid progression to tuberculosis disease and possible clinical presentation of HIV-positive patients with tuberculosis, as well as characteristics such as paucibacillary disease, may account for the apparent lower transmission rate from these patients. However, this simple analysis does not take into account secondary transmission within clusters, and more-detailed genetic analysis, such as whole-genome sequencing of clusters, is required. Our study also reported that tuberculosis among HIV-positive patients was more commonly due to recent infection, and this trend was also evident in patients receiving ART. In Cape Town, >60% of patients have already had or currently have tuberculosis at ART initiation [28], and this is likely to be due to reactivation of latent infection. Subsequently, during ART, the tuberculosis risk for most patients is due to newly acquired *M. tuberculosis* infection. Ongoing transmission is therefore an important contributor to the persistent risk of tuberculosis among ART recipients. In addition to this time effect, the dynamic process of declining and recovering CD4⁺ T-cell counts, together with the higher CD4⁺ T-cell count threshold for initiating ART in this community, may also explain the lack of association with CD4⁺ T-cell count and reactivation of or reinfection tuberculosis disease.

The high rates of disease due to recent infection reported in this study are in keeping with the substantial tuberculosis transmission rates in this community [29]. Nonetheless, the proportion of reactivation versus newly infected cases is most likely overrepresented in this analysis. First, nearly 50% of children had unique strains or were index cases for a cluster. As tuberculosis in young children is predominantly the result of recent adult transmission [30], these data suggest that the source case for these children was missed in this study. Social interaction data from this community have previously shown that at least 50% of casual indoor contacts occurred outside of the study community [31], and this may partially explain unidentified source cases. Second, data censoring due to the data collection period and incomplete sampling (52% of sputum-positive pulmonary tuberculosis cases were included in the analysis) may have resulted in some strains being mistakenly classified as unique. Furthermore, transmission may result in sputum-negative pulmonary tuberculosis or extrapulmonary tuberculosis, and we were unable to assess *M. tuberculosis* genotypes in patients with these tuberculosis presentations. Therefore, recent transmission is likely responsible for a greater proportion of the tuberculosis cases than estimated in this study.

This study was unable to identify associations between specific community-based locations or public transport and clustering. Of note was the lack of association between ART use and

clustering and the wide variety of strains among patients receiving ART. This finding, together with the lack of association between clustering and clinic attendance, suggests that the clinic was not a significant location for tuberculosis transmission to vulnerable patients in this community. Also of interest was the low proportion of adult-to-adult transmissions occurring in households (<10%). This finding may, in part, explain the limited impact of household-based enhanced case-finding interventions [32].

We inferred recent transmission of tuberculosis from clustered strains. Despite the traditional interpretation of identical strains as epidemiologically linked, it is important to recognize that clustering is not synonymous with recent transmission. In areas of high incidence, the proportion of clustered cases is influenced by factors such as the background strains, the annual risk of infection, the age of the population, the geographic distribution [23, 33], the study duration [34, 35], the sampling strategies and percentages [36, 37], and the genotyping methods. The interpretation of clustering as evidence of transmission is enhanced in our study by the long study duration of 10 years [35, 38], the substantial tuberculosis transmission rates [19], and the high diversity of strains in the community. In addition, IS6110 fingerprinting is one of the most discriminatory typing techniques for isolates with >6 IS6110 bands, such as the CC and W-Beijing families [20].

A limitation of this study was that sputum samples and RFLP data were not collected for all patients with tuberculosis in this community over the study period. However, the incomplete sampling appears to be random: the only difference among patients with pulmonary tuberculosis for whom we did not obtain genotyping data was a higher mortality and lower MDR tuberculosis rate. Because of the incomplete sampling from the study population, the number of circulating strains, the number of clusters, and the size of clusters may have been underestimated [36, 37]. Moreover, our assessment of social factors associated with transmission used relatively crude measures for possible social influences, and in the future more-detailed studies of these factors will be required.

In summary, this study showed that the HIV-associated and HIV-unassociated tuberculosis epidemics were not independent. Furthermore HIV-positive patients may be responsible for proportionally less of the tuberculosis transmission in this community. Of importance is the finding that recent transmission is responsible for the majority of tuberculosis cases in a setting where the burdens of HIV infection and tuberculosis are high. This finding was particularly marked among HIV-positive patients, including patients receiving ART, suggesting that ART may not provide protection against progression to tuberculosis following recent infection. Finally, the household was not the key location for adult-to-adult transmission.

These findings have important implications for tuberculosis control. The impact of high ongoing transmission rates on INH

preventive therapy programs must be considered. Furthermore, our attention and interventions need to expand beyond HIV-positive patients and include reducing transmission from HIV-negative patients. Such efforts may be guided by the identification of the factors and locations associated with transmission in these settings.

Notes

Financial support. This work was supported by the National Institutes of Health (NIH; through Comprehensive Integrated Programme of Research on AIDS grant 1U19AI053217 to K. M., L. G. B., L. M., and R. W.; grants 1U19AI05321 and RO1 AI058736-02 to R. W.; grants RO1 AI66046 and RO1 AI 080737 to B. K.; grant RO1 AI 54361 to G. K.; and through the CIPRA-ZA Project 3 Extension 2007), the Royal Society Pfizer award (to L.-G. B.), the Hasso Plattner Foundation Mid-Career Development award (through the University of Cape Town to K. M.), and the Fogarty International Center, NIH (through Columbia University–Southern African Fogarty AIDS International Training and Research Program grant D43TW00231 to K. M.).

Disclaimer. The content of this publication does not necessarily reflect the views or policies of the National Institute of Allergy and Infectious Diseases, nor does mention of trade names, commercial projects, or organizations imply endorsement by the US government.

Potential conflicts of interest. All authors: No reported conflicts.

All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. World Health Organization. Global tuberculosis report, 2013. http://apps.who.int/iris/bitstream/10665/91355/1/9789241564656_eng.pdf. Accessed 11 March 2014.
2. Fine PE, Small PM. Exogenous reinfection in tuberculosis. *N Engl J Med* 1999; 341:1226–7.
3. Bauer J, Yang Z, Poulsen S, Andersen AB. Results from 5 years of nationwide DNA fingerprinting of *Mycobacterium tuberculosis* complex isolates in a country with a low incidence of *M. tuberculosis* infection. *J Clin Microbiol* 1998; 36:305–8.
4. Verver S, Warren RM, Munch Z, et al. Transmission of tuberculosis in a high incidence urban community in South Africa. *Int J Epidemiol* 2004; 33:351–7.
5. Glynn JR, Crampin AC, Yates MD, et al. The importance of recent infection with *Mycobacterium tuberculosis* in an area with high HIV prevalence: a long-term molecular epidemiological study in Northern Malawi. *J Infect Dis* 2005; 192:480–7.
6. Lockman S, Sheppard JD, Braden CR, et al. Molecular and conventional epidemiology of *Mycobacterium tuberculosis* in Botswana: a population-based prospective study of 301 pulmonary tuberculosis patients. *J Clin Microbiol* 2001; 39:1042–7.
7. Wilkinson D, Pillay M, Crump J, Lombard C, Davies GR, Sturm AW. Molecular epidemiology and transmission dynamics of *Mycobacterium tuberculosis* in rural Africa. *Trop Med Int Health* 1997; 2:747–53.
8. Asiimwe BB, Joloba ML, Ghebremichael S, et al. DNA restriction fragment length polymorphism analysis of *Mycobacterium tuberculosis* isolates from HIV-seropositive and HIV-seronegative patients in Kampala, Uganda. *BMC Infect Dis* 2009; 9:12.
9. Middelkoop K, Bekker LG, Mathema B, et al. Factors affecting tuberculosis strain success over 10 years in a high TB and HIV burdened community. *Int J Epidemiol* 2014; 43:1114–22.
10. Wood R, Middelkoop K, Myer L, et al. Undiagnosed tuberculosis in a community with high HIV prevalence: implications for tuberculosis control. *Am J Respir Crit Care Med* 2007; 175:87–93.

11. Middelkoop K, Bekker LG, Myer L, et al. Antiretroviral therapy and TB notification rates in a high HIV prevalence South African community. *J Acquir Immune Defic Syndr* 2011; 56:263–9.
12. Andrews JR, Wood R, Bekker L-G, Middelkoop K, Walensky RP. Projecting the benefits of antiretroviral therapy for HIV prevention: the impact of population mobility and linkage to care. *J Infect Dis* 2012; 206:543–51.
13. Middelkoop K, Bekker LG, Myer L, et al. Antiretroviral program associated with reduction in untreated prevalent tuberculosis in a South African township. *Am J Respir Crit Care Med* 2010; 182:1080–5.
14. Kranzer K, van Schaik N, Karmue U, et al. High prevalence of self-reported undiagnosed HIV despite high coverage of HIV testing: a cross-sectional population based sero-survey in South Africa. *PLoS One* 2011; 6:e25244.
15. South African Department of Health. National antiretroviral treatment guidelines. 2004. http://www.hst.org.za/uploads/files/sa_ART_Guidelines.pdf. Accessed 11 March 2014.
16. Sanne I, Orrell C, Fox MP, et al. Nurse versus doctor management of HIV-infected patients receiving antiretroviral therapy (CIPRA-SA): a randomised non-inferiority trial. *Lancet* 2010; 376:33–40.
17. Department of Health RSA. National tuberculosis management guidelines, 2009. http://familymedicine.ukzn.ac.za/Libraries/Guidelines_Protocols/TB_Guidelines_2009.sflb.ashx. Accessed 11 March 2014.
18. Cape Town TB Control. Progress report 1997–2003. Cape Town, South Africa: Health Systems Trust, 2004.
19. Middelkoop K, Bekker LG, Myer L, Dawson R, Wood R. Rates of tuberculosis transmission to children and adolescents in a community with a high prevalence of HIV infection among adults. *Clin Infect Dis* 2008; 47:349–55.
20. Mathema B, Kurepina NE, Bifani PJ, Kreiswirth BN. Molecular epidemiology of tuberculosis: current insights. *Clin Microbiol Rev* 2006; 19:658–85.
21. Murray M, Alland D. Methodological problems in the molecular epidemiology of tuberculosis. *Am J Epidemiol* 2002; 155:565–71.
22. Fenner L, Egger M, Bodmer T, et al. HIV infection disrupts the sympatric host–pathogen relationship in human tuberculosis. *PLoS Genet* 2013; 9:e1003318.
23. Middelkoop K, Bekker LG, Mathema B, et al. Molecular epidemiology of *Mycobacterium tuberculosis* in a South African community with high HIV prevalence. *J Infect Dis* 2009; 200:1207–11.
24. Cowley D, Govender D, February B, et al. Recent and rapid emergence of W-Beijing strains of *Mycobacterium tuberculosis* in Cape Town, South Africa. *Clin Infect Dis* 2008; 47:1252–9.
25. Daley CL, Small PM, Schecter GF, et al. An outbreak of tuberculosis with accelerated progression among persons infected with the human immunodeficiency virus. An analysis using restriction-fragment-length polymorphisms. *N Engl J Med* 1992; 326:231–5.
26. Selwyn PA, Hartel D, Lewis VA, et al. A prospective study of the risk of tuberculosis among intravenous drug users with human immunodeficiency virus infection. *N Engl J Med* 1989; 320:545–50.
27. Dooley SW, Villarino ME, Lawrence M, et al. Nosocomial transmission of tuberculosis in a hospital unit for HIV-infected patients. *JAMA* 1992; 267:2632–4.
28. Lawn SD, Myer L, Bekker LG, Wood R. Burden of tuberculosis in an antiretroviral treatment programme in sub-Saharan Africa: impact on treatment outcomes and implications for tuberculosis control. *AIDS* 2006; 20:1605–12.
29. Middelkoop K, Bekker LG, Liang H, et al. Force of tuberculosis infection among adolescents in a high HIV and TB prevalence community: a cross-sectional observation study. *BMC Infect Dis* 2011; 11:156.
30. Bloch AB, Snider DE Jr. How much tuberculosis in children must we accept? *Am J Public Health* 1986; 76:14–5.
31. Johnstone-Robertson SP, Mark D, Morrow C, et al. Social mixing patterns within a South African township community: implications for respiratory disease transmission and control. *Am J Epidemiol* 2011; 174:1246–55.

32. Ayles H, Muyoyeta M, Du TE, et al. Effect of household and community interventions on the burden of tuberculosis in southern Africa: the ZAMSTAR community-randomised trial. *Lancet* **2013**; 382:1183–94.
33. Braden CR, Templeton GL, Cave MD, et al. Interpretation of restriction fragment length polymorphism analysis of *Mycobacterium tuberculosis* isolates from a state with a large rural population. *J Infect Dis* **1997**; 175:1446–52.
34. van der Spuy GD, van Helden PD, Warren RM. Effect of study duration on the interpretation of tuberculosis molecular epidemiology investigations. *Tuberc Edinb* **2009**; 89:238–42.
35. Vynnycky E, Nagelkerke N, Borgdorff MW, van Soolingen D, van Embden JD, Fine PE. The effect of age and study duration on the relationship between “clustering” of DNA fingerprint patterns and the proportion of tuberculosis disease attributable to recent transmission. *Epidemiol Infect* **2001**; 126:43–62.
36. Murray M. Sampling bias in the molecular epidemiology of tuberculosis. *Emerg Infect Dis* **2002**; 8:363–9.
37. Glynn JR, Vynnycky E, Fine PE. Influence of sampling on estimates of clustering and recent transmission of *Mycobacterium tuberculosis* derived from DNA fingerprinting techniques. *Am J Epidemiol* **1999**; 149:366–71.
38. van der Spuy GD, Kremer K, Ndabambi SL, et al. Changing *Mycobacterium tuberculosis* population highlights clade-specific pathogenic characteristics. *Tuberc Edinb* **2009**; 89:120–5.