The following full text is a publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/152882

Please be advised that this information was generated on 2019-09-06 and may be subject to change.
Transmission and Progression to Disease of Mycobacterium tuberculosis Phylogenetic Lineages in The Netherlands

Hanna Nebenzahl-Guimaraes,a,b,c Lilly M. Verhagen,d,e Martien W. Borgdorff,f,g Dick van Soolingen,a,h

National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands; Life and Health Sciences Research Institute (ICVS), School of Health Sciences, University of Minho, Braga, Portugal; ICVS/3Bs, PT Government Associate Laboratory, Braga/Guimarães, Portugal; Wilhelmina Children’s Hospital Utrecht, Utrecht, The Netherlands; Laboratory of Pediatric Infectious Diseases, Radboud University Medical Centre, Nijmegen, The Netherlands; Public Health Service, Amsterdam, The Netherlands; Department of Clinical Epidemiology, Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands; Department of Medical Microbiology, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands.

The aim of this study was to determine if mycobacterial lineages affect infection risk, clustering, and disease progression among Mycobacterium tuberculosis cases in The Netherlands. Multivariate negative binomial regression models adjusted for patient-related factors and stratified by patient ethnicity were used to determine the association between phylogenetic lineages and infectivity (mean number of positive contacts around each patient) and clustering (as defined by number of secondary cases within 2 years after diagnosis of an index case sharing the same fingerprint) indices. An estimate of progression to disease by each risk factor was calculated as a bootstrapped risk ratio of the clustering index by the infectivity index. Compared to the Euro-American reference, Mycobacterium africanum showed significantly lower infectivity and clustering indices in the foreign-born population, while Mycobacterium bovis showed significantly lower infectivity and clustering indices in the native population. Significantly lower infectivity was also observed for the East African Indian lineage in the foreign-born population. Smear positivity was a significant risk factor for increased infectivity and increased clustering. Estimates of progression to disease were significantly associated with age, sputum-smear status, and behavioral risk factors, such as alcohol and intravenous drug abuse, but not with phylogenetic lineages. In conclusion, we found evidence of a bacteriological factor influencing indicators of a strain’s infectivity, and we will refer to these two properties that affect the strain’s transmissibility, namely, a decreased ability to infect and a lower clustering index in ancient phylogenetic lineages compared to their modern counterparts. Confirmation of these findings via follow-up studies using tuberculin skin test conversion data should have important implications on M. tuberculosis control efforts.

Curbing tuberculosis (TB) transmission is a challenge in high-burden countries. However, even in low-prevalence settings, controlling TB is an important requirement due to human migration from higher-incidence areas to Western countries (1). In Western countries, studies on transmission are more feasible, as all cases undergo extended diagnostic algorithms and all clinical and demographic data are recorded. Current molecular typing methods, such as variable number of tandem repeat (VNTR) typing and restriction fragment length polymorphism (RFLP) typing, allow identification of clusters of Mycobacterium tuberculosis isolates with identical genotypes that, in population-based studies, reveal recent transmission (2, 3). Spoligotyping and VNTR typing can identify the genotype family of the isolate, revealing bacterial variation via the identification of phylogenetic lineages (4, 5).

While many studies have elucidated the variation in the disease’s spread and outcome attributable to host and environmental factors, there is also evidence that bacterial factors may affect the spread of tuberculosis (6). In The Netherlands, for example, one study showed that the number of positive contacts around a case increases with growing cluster size (7). In a subsequent study in the same setting, cluster size growth was not different between phylogenetic lineages after controlling for host risk factors (8). However, this study could not distinguish between transmission rates and progression to disease. There are, however, indications that progression to disease is partly dependent on bacterial variation. It has, e.g., been postulated that some Mycobacterium africanum strains might transmit equally well as other M. tuberculosis complex strains but might be less associated with progression to disease (9). We will refer to these two properties that affect the degree of clustering as infectivity (the bacterium’s ability to establish an initial infection in the human host) and progression to disease (the bacterium’s capacity to produce disease) (10).

In the low-incidence context of The Netherlands, with a globally representative cohort of patients, we aim to determine differences in indices of infectivity, clustering, and estimated progression to disease of different mycobacterial lineages using fingerprinting data and contact investigation. This will provide insights into the role of bacteriological factors in TB transmission, which itself may affect future TB control measures.

MATERIALS AND METHODS

Data collection and DNA fingerprinting. The National Institute for Public Health and the Environment (RIVM) is a reference laboratory for secondary laboratory diagnosis of all TB cases in The Netherlands, offer-
ing identification, drug susceptibility testing, and molecular typing for each TB case. DNA fingerprints of all nationwide *M. tuberculosis* complex isolates and their cluster statuses have been stored in an RFLP/VNTR database since 1993. The registration committee of The Netherlands Tuberculosis Register (NTR) approved this retrospective study and provided anonymized demographic and clinical information for patients. Because these data are deidentified by name, DNA fingerprinting results were matched by sex, date of birth, year of diagnosis, and postal code. All notified culture-positive cases of *M. tuberculosis* between 1993 and 2011 were included in the study. For patients with multiple isolates sharing identical fingerprints, only the isolate with the earliest diagnosis date was included. Contaminating isolates were excluded.

Isolates recovered from patients between 1993 and 2009 underwent IS6110 typing and polymorphic GC-rich sequence (PGRS) RFLP typing (n = 15,073), and those from 2004 onward were subjected to VNTR typing (n = 5,870) (11, 12). In the period of 2004 to 2008, both RFLP and 24-locus VNTR typing were performed to obtain a smooth transition in typing methods and to evaluate VNTR typing performance (3). In addition, 4,433 randomly selected isolates were spoligotyped (n = 4,433). We defined a cluster as a group of patients who shared *M. tuberculosis* isolates with identical RFLP or VNTR patterns or, if strains had fewer than five IS6110 copies, identical PGRS RFLP patterns.

**Conventional contact investigation.** Systematic contact investigation by TB Public Health Services in The Netherlands is conducted per the stone-in-the-pond principle, in which the decision to extend conventional contact investigation to the next ring of contacts is based on the prevalence of infection in the investigated ring (13). Contacts are defined by the frequency and intimacy of their contacts with the TB index case. The tuberculin skin test (TST) is used to investigate presumably exposed contacts. If the number of TST-positive contacts in the first ring suggests a high spread of tuberculosis, a larger ring of contacts is investigated. We have defined positive contacts as contacts with a TST induration ≥10 mm and/or contacts who received a diagnosis of TB disease. If contact investigations become very large, identified TB infections and secondary cases are less likely to be related to the index case. To minimize the probability that positive contacts in our research were unrelated to the defined index case, we only included contacts in the first ring around the index patient. First-ring contacts are defined as contacts that are physically close to the index patient, considering environmental factors, such as room size, ventilation, air purification, and air circulation. In addition, the patient and the contact must be able to indicate where they met and must have a long-standing relationship to qualify as a first-ring contact. Examples of first-ring contacts are household members, close work colleagues, and close friends.

**Classification into phylogenetic lineages.** The phylogenetic lineages of isolates were determined using a combination of spoligotyping, the MIRU (mycobacterial interspersed repetitive unit) best match analysis offered by the MIRU-VNTRplus online tool, and RFLP similarity, as described in a previous study using the same data set (8, 14). Three species (*M. africanaum*, *Mycobacterium bovis*, and *M. tuberculosis*) and four major phylogenetic lineages of *M. tuberculosis* were identified: the Euro-Ameri- can, Central Asian strain (CAS), East African Indian (EAI), and Beijing genotypes. Strains not assigned a phylogenetic lineage or assigned to multiple major phylogenetic families per cluster were not analyzed. Strains classified as either T or U (unknown) also were excluded due to the ambiguity of these classifications.

**Definitions.** For our infectivity index, we took the mean number of positive contacts around each patient who underwent contact investigation. We excluded patients with missing data on contact investigation or those who had zero contacts investigated, as well as those for whom we lacked ethnicity information. Because TB transmission almost exclusively results from patients with pulmonary TB, we also excluded patients with extrapulmonary TB, leaving us with a total of 2,809 cases (Fig. 1).

For our clustering index, we used the number of secondary cases occurring within 2 years of the index case diagnosis. The 2-year cutoff has been shown to best reflect recent transmission as opposed to disease reactivation (1, 15). We defined index cases as patients who had strains with RFLP or VNTR patterns not seen in other patients in the previous 2 years. We searched for index cases based on RFLP-typing data from 1995 to 2007 and for index cases based on VNTR typing from 2007 to 2009. We excluded RFLP-defined index cases from 1993 and 1994 and VNTR-defined index cases from 2005 and 2006 (n = 2,684), because we could not determine whether the strains of these index cases were unobserved in the previous 2 years. Similarly, we excluded RFLP-defined index cases occurring after 2007 and VNTR-defined index cases occurring after 2009 (n = 950), because we could not follow these index cases for a full 2 years. Secondary cases from these index cases (included in the counts) were also excluded. Finally, we excluded cases between 1995 and 2007 occurring <2 years after a previous patient with the same RFLP fingerprint yet diagnosed >2 years after a cluster’s start (n = 722) and cases occurring between 2007 and 2009 that occurred <2 years after a previous patient with the same VNTR fingerprint yet >2 years after a cluster’s start (n = 40). After excluding extrapulmonary cases, 4,432 patients remained: 2,881 nonclustered index patients, 607 index patients who were the first patient of a cluster, and 944 secondary cases within 2 years of a cluster’s start (Fig. 1).

Finally, estimates of progression to disease were calculated as risk ratios (RR) of the population risk of disease given exposure to a risk factor by the population risk of infection given exposure to the same risk factor (dividing the clustering odds ratios [ORs] by the infectivity ORs).

**Statistical analysis.** We used a multivariate negative binomial regression model to determine the association between phylogenetic lineages and the infectivity and clustering indices. Since TST is poorly specific among *Mycobacterium bovis* BCG contacts and positive TSTs may represent old infections, we divided our data sets into native and foreign-born (FB) cohorts in order to address important differences between the two: FB patients are often BCG vaccinated (in contrast to native Dutch patients, who are not), while the prevalence of infection is higher among FB patients. Second-generation patients (born to FB patients) were included in the native cohorts, given that, like native patients, they are not BCG vaccinated and they have already been born in a setting of lower prevalence of infection. Studies carried out in The Netherlands have also previously demonstrated that contact investigation practices vary by demographic characteristics of the index patient (16). As such, in both analyses, we adjusted for index patient-related factors, including demographic, behavioral, and sputum smear status. In addition, the logarithm of the number of investigated contacts around a source case was used as an offset in the multivariate model assessing the association between phylogenetic lineages and the spread index, since the greater number of contacts around a source are investigated, the likelier it is to detect TST positive contacts. Variables with P values of ≤0.20 were entered into the multivariate model. Crude and adjusted ORs are presented with 95% confidence intervals (CIs). Estimates of TB progression were calculated for any risk factor that was significant in either multivariate regression model. To calculate the variance for the estimate of TB progression, we performed a bootstrapping procedure, running our multivariate negative binomial regression models 10,000 times on bootstrapped data sets. The median of the resulting 10,000 RR values was used as the estimate of TB progression, while the 2.5th and 97.5th percentiles were used as the 95% cutoffs for the estimate CI. All analyses were conducted using SAS (Windows version 9.3), SPSS program for Windows version 20.0 (SPSS Inc., Chicago, IL, USA) and R (version 3.1.2 for Windows).

**RESULTS**

Between January 1993 and December 2011, 18,294 isolates were collected from the same number of notified TB cases in The Netherlands, and their clustering statuses were ascertained, of which 15,601 (85%) were successfully matched with the NTR data. Of these, 15,224 (98%) were noncontaminated *M. tuberculosis* cul-
tures with completely ascertained information on host risk factors. After phylogenetic lineage assignment, there were 10,738 isolates that were *M. bovis*, *M. africanum*, or *M. tuberculosis* of the Euro-American, Beijing, CAS, or EAI lineages (Fig. 1). The mean age of the patients carrying these strains was 41 years (standard deviation, 20 years); 6,394 (60%) were male; and 7,762 (72%) were foreign born.

**Mycobacterial genotypes.** The Euro-American lineage was predominant in both the infectivity (78% in native cohort; 56% in FB cohort) and clustering (79% in native cohort; 64% in FB cohort) data sets. In contrast, both *M. africanum* and *M. bovis* represented less than 1% of all cases in the infectivity data set. In the clustering data set, both *M. africanum* and *M. bovis* represented only 2% of all cases (Tables 1 and 2).

**Infectivity by mycobacterial lineage.** The proportion of cases in which a contact investigation was performed in The Netherlands was approximately equal between lineages, though slightly lower in the FB cohort for Beijing and EAI compared to the native counterpart (Fig. 2). The average number of TST-positive contacts declined significantly in the >65 years age category in the native cohort and in the <20 years age category in the FB cohort. Smear positivity was associated with an increased average number of TST-positive contacts in both native and FB cohorts. There were no significant differences in infectivity by gender, homelessness, and alcohol use in the two cohorts, although use of intravenous drugs in the native populations and rural residence in the FB population were associated with a decreased average number of TST-positive contacts. The mean number of TST-positive contacts around an index case was significantly lower for *M. bovis* than for the Euro-American reference lineage in the native population in multivariable analysis. In the FB population, *M. africanum* and EAI presented a significantly lower number of TST-positive contacts (Table 1).

**Clustering by mycobacterial lineage.** The number of secondary cases declined significantly with increasing age (>65 years) in both the native and FB cohorts. Smear positivity was also associ-
ated with an increased number of secondary cases in both cohorts, and female gender was associated with an increased number of secondary cases only among the FB. Rural residence was associated with a decreased number of secondary cases only in the FB cohort. Relative to the Euro-American reference in the multivariable analysis, the number of secondary cases was significantly lower for *M. bovis* in the native-born population and for *M. africanum* in the FB population (Table 2).

Estimates of progression to disease by mycobacterial lineage. Estimates of progression to disease were significantly lower in the >65 years age category in both ethnic cohorts and significantly higher in the 0- to 19-years age category in the FB cohort. Additionally, in the FB-born population, estimates of progression to disease were significantly lower in smear-negative patients. Both alcohol and drug abuse were significantly associated with higher estimates in the native population. No

---

**TABLE 1** Risk factors among native and foreign-born index cases for infectivity (number of TST-positive contacts per index case)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Native cohort</th>
<th>Mean no. of TST-positive contacts/ index case</th>
<th>Univariate analysis</th>
<th>Multivariate analysis, relative no. (95% CI)</th>
<th>Foreign-born cohort</th>
<th>Mean no. of TST-positive contacts/ index case</th>
<th>Univariate analysis</th>
<th>Multivariate analysis, relative no. (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, yr</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–19</td>
<td>99</td>
<td>1.02</td>
<td>0.96 (0.65–1.4)</td>
<td></td>
<td></td>
<td>200</td>
<td>0.83</td>
<td>1.11 (0.83–1.5)</td>
</tr>
<tr>
<td>20–39</td>
<td>363</td>
<td>0.99</td>
<td>1 (Ref)</td>
<td></td>
<td></td>
<td>1024</td>
<td>0.75</td>
<td>1 (Ref)</td>
</tr>
<tr>
<td>40–64</td>
<td>339</td>
<td>0.82</td>
<td>0.91 (0.69–1.2)</td>
<td></td>
<td></td>
<td>459</td>
<td>0.60</td>
<td>0.81 (0.65–1.0)</td>
</tr>
<tr>
<td>≥65</td>
<td>207</td>
<td>0.51</td>
<td>0.50 (0.36–0.70)</td>
<td></td>
<td></td>
<td>118</td>
<td>0.49</td>
<td>0.66 (0.45–0.97)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>635</td>
<td>0.78</td>
<td>1.12 (0.93–1.5)</td>
<td></td>
<td></td>
<td>1,125</td>
<td>0.73</td>
<td>1 (Ref)</td>
</tr>
<tr>
<td>Female</td>
<td>373</td>
<td>0.93</td>
<td>1.0 (0.82–1.3)</td>
<td></td>
<td></td>
<td>676</td>
<td>0.66</td>
<td>0.73 (0.65–0.82)</td>
</tr>
<tr>
<td>Smear positivity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>395</td>
<td>0.52</td>
<td>0.50 (0.39–0.65)</td>
<td>0.48 (0.38–0.62)</td>
<td></td>
<td>754</td>
<td>0.47</td>
<td>0.54 (0.45–0.65)</td>
</tr>
<tr>
<td>Positive</td>
<td>613</td>
<td>1.04</td>
<td>1 (Ref)</td>
<td>1 (Ref)</td>
<td></td>
<td>1,047</td>
<td>0.87</td>
<td>1 (Ref)</td>
</tr>
<tr>
<td>Lineage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Euro-American</td>
<td>786</td>
<td>0.91</td>
<td>1 (Ref)</td>
<td>1 (Ref)</td>
<td></td>
<td>1,157</td>
<td>0.75</td>
<td>1 (Ref)</td>
</tr>
<tr>
<td>Beijing</td>
<td>112</td>
<td>0.67</td>
<td>0.73 (0.50–1.1)</td>
<td>1.1 (0.77–1.6)</td>
<td></td>
<td>182</td>
<td>0.55</td>
<td>1.7 (0.54–1.0)</td>
</tr>
<tr>
<td>CAS</td>
<td>30</td>
<td>0.23</td>
<td>0.26 (0.11–0.61)</td>
<td>0.64 (0.27–1.5)</td>
<td></td>
<td>198</td>
<td>0.86</td>
<td>1.2 (0.86–1.5)</td>
</tr>
<tr>
<td>EAI</td>
<td>62</td>
<td>0.56</td>
<td>0.62 (0.37–1.04)</td>
<td>0.78 (0.47–1.3)</td>
<td></td>
<td>237</td>
<td>0.53</td>
<td>0.71 (0.54–0.94)</td>
</tr>
<tr>
<td><em>M. africanum</em></td>
<td>5</td>
<td>0.40</td>
<td>0.44 (0.067–2.9)</td>
<td>0.44 (0.055–3.5)</td>
<td></td>
<td>15</td>
<td>0.27</td>
<td>0.36 (0.11–1.1)</td>
</tr>
<tr>
<td><em>M. bovis</em></td>
<td>13</td>
<td>0.23</td>
<td>0.25 (0.068–0.94)</td>
<td>0.23 (0.059–0.94)</td>
<td></td>
<td>12</td>
<td>0.17</td>
<td>0.22 (0.052–0.96)</td>
</tr>
<tr>
<td>Residency</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>690</td>
<td>0.90</td>
<td>1 (Ref)</td>
<td>1 (Ref)</td>
<td></td>
<td>1,067</td>
<td>0.79</td>
<td>1 (Ref)</td>
</tr>
<tr>
<td>Rural</td>
<td>318</td>
<td>0.70</td>
<td>0.78 (0.60–1.0)</td>
<td>0.78 (0.60–1.0)</td>
<td></td>
<td>734</td>
<td>0.58</td>
<td>0.74 (0.62–0.89)</td>
</tr>
<tr>
<td>Alcohol abuse</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>969</td>
<td>0.85</td>
<td>1 (Ref)</td>
<td>1 (Ref)</td>
<td></td>
<td>1,779</td>
<td>0.70</td>
<td>1 (Ref)</td>
</tr>
<tr>
<td>Yes</td>
<td>39</td>
<td>0.54</td>
<td>0.64 (0.33–1.2)</td>
<td>0.59 (0.30–1.2)</td>
<td></td>
<td>22</td>
<td>0.59</td>
<td>0.84 (0.36–1.9)</td>
</tr>
<tr>
<td>Drug abuse</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>953</td>
<td>0.86</td>
<td>1 (Ref)</td>
<td>1 (Ref)</td>
<td></td>
<td>1,722</td>
<td>0.70</td>
<td>1 (Ref)</td>
</tr>
<tr>
<td>Yes</td>
<td>55</td>
<td>0.40</td>
<td>0.47 (0.26–0.83)</td>
<td>0.43 (0.24–0.78)</td>
<td></td>
<td>79</td>
<td>0.71</td>
<td>0.84 (0.36–1.9)</td>
</tr>
<tr>
<td>Traveler to country of endemicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>973</td>
<td>0.85</td>
<td>1 (Ref)</td>
<td>1 (Ref)</td>
<td></td>
<td>1,759</td>
<td>0.69</td>
<td>1 (Ref)</td>
</tr>
<tr>
<td>Yes</td>
<td>35</td>
<td>0.51</td>
<td>0.61 (0.30–1.2)</td>
<td>0.53 (0.25–1.1)</td>
<td></td>
<td>42</td>
<td>1.02</td>
<td>1.5 (0.83–2.6)</td>
</tr>
<tr>
<td>Homeless</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>978</td>
<td>0.84</td>
<td>1 (Ref)</td>
<td>1 (Ref)</td>
<td></td>
<td>1,753</td>
<td>0.70</td>
<td>1 (Ref)</td>
</tr>
<tr>
<td>Yes</td>
<td>30</td>
<td>0.73</td>
<td>0.88 (0.43–1.8)</td>
<td>0.48 (0.24–0.78)</td>
<td></td>
<td>48</td>
<td>0.69</td>
<td>0.98 (0.56–1.7)</td>
</tr>
<tr>
<td>Site of disease</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pulmonary</td>
<td>891</td>
<td>0.86</td>
<td>1 (Ref)</td>
<td>1 (Ref)</td>
<td></td>
<td>1,430</td>
<td>0.73</td>
<td>1 (Ref)</td>
</tr>
<tr>
<td>Pulmonary + extrapulmonary</td>
<td>117</td>
<td>0.67</td>
<td>0.78 (0.53–1.1)</td>
<td>1.0 (0.70–1.5)</td>
<td></td>
<td>371</td>
<td>0.60</td>
<td>0.83 (0.66–1.0)</td>
</tr>
</tbody>
</table>

*Ref, reference.*

October 2015 Volume 53 Number 10 Journal of Clinical Microbiology jcm.asm.org 3267
significant differences were found across phylogenetic lineages (Table 3).

DISCUSSION

In this study, we observed variations in the infectivity and clustering indices of different phylogenetic subgroups of *M. tuberculosis*, *M. bovis*, and *M. africanum* after controlling for clinical and demographic index host factors. *M. africanum* and *M. bovis* showed both significantly lower infectivity and clustering indices in the FB and native populations, respectively. A significantly lower infectivity was also observed for the EAI lineage in the larger FB population.

Our findings around *M. africanum* are consistent with previous experiments characterizing its reduced ESAT-6 (early secretory antigenic target-6) immunogenicity and candidate genes behind its attenuated phenotype (17). However, they are only

### TABLE 2 Risk factors among native and foreign-born index cases for clustering (number of secondary cases within 2 years of an index case)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Native cohort</th>
<th>Mean no. of second cases per index case</th>
<th>Univariate analysis</th>
<th>Multivariate analysis, relative no. (95% CI)</th>
<th>P</th>
<th>Multivariate analysis, relative no. (95% CI)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of index cases</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age, yr</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–19</td>
<td>59</td>
<td>0.46</td>
<td>1.02 (0.53–1.90)</td>
<td>0.001</td>
<td>1.04 (0.56–1.93)</td>
<td>1.04 (0.56–1.93)</td>
<td>0.001</td>
</tr>
<tr>
<td>20–39</td>
<td>216</td>
<td>0.45</td>
<td>1 (Ref)</td>
<td>1 (Ref)</td>
<td>1,411</td>
<td>0.27</td>
<td>1 (Ref)</td>
</tr>
<tr>
<td>40–64</td>
<td>267</td>
<td>0.34</td>
<td>0.65 (0.41–0.98)</td>
<td>0.68 (0.46–1.02)</td>
<td>569</td>
<td>0.31</td>
<td>1.14 (0.86–1.51)</td>
</tr>
<tr>
<td>≥65</td>
<td>495</td>
<td>0.1</td>
<td>0.22 (0.12–0.33)</td>
<td>0.21 (0.13–0.32)</td>
<td>195</td>
<td>0.08</td>
<td>0.30 (0.16–0.55)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>387</td>
<td>0.23</td>
<td>0.87 (0.61–1.23)</td>
<td>0.4255</td>
<td>0.027</td>
<td>1 (Ref)</td>
<td>1 (Ref)</td>
</tr>
<tr>
<td>Male</td>
<td>650</td>
<td>0.27</td>
<td>1 (Ref)</td>
<td>1,476</td>
<td>0.31</td>
<td>1 (Ref)</td>
<td>1 (Ref)</td>
</tr>
<tr>
<td>Smear positivity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>393</td>
<td>0.18</td>
<td>0.58 (0.40–0.83)</td>
<td>0.0027</td>
<td>0.027</td>
<td>6.2 (0.44–0.88)</td>
<td>0.48 (0.38–0.61)</td>
</tr>
<tr>
<td>Positive</td>
<td>644</td>
<td>0.3</td>
<td>1 (Ref)</td>
<td>1,375</td>
<td>0.36</td>
<td>1 (Ref)</td>
<td>1 (Ref)</td>
</tr>
<tr>
<td>Lineage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Euro-American</td>
<td>756</td>
<td>0.25</td>
<td>1 (Ref)</td>
<td>1,385</td>
<td>0.3</td>
<td>1 (Ref)</td>
<td>0.081</td>
</tr>
<tr>
<td>Beijing</td>
<td>73</td>
<td>0.36</td>
<td>1.44 (0.79–2.62)</td>
<td>0.1971</td>
<td>1.20</td>
<td>0.68–2.11</td>
<td>0.76 (0.66–1.32)</td>
</tr>
<tr>
<td>CAS</td>
<td>28</td>
<td>0.29</td>
<td>1.15 (0.41–3.09)</td>
<td>0.79 (0.30–2.10)</td>
<td>290</td>
<td>0.28</td>
<td>0.93 (0.64–1.36)</td>
</tr>
<tr>
<td>EAI</td>
<td>118</td>
<td>0.33</td>
<td>1.34 (0.82–2.19)</td>
<td>0.90 (0.54–1.50)</td>
<td>335</td>
<td>0.28</td>
<td>0.74 (0.51–1.07)</td>
</tr>
<tr>
<td>M. africanum</td>
<td>8</td>
<td>0.13</td>
<td>0.50 (0.04–5.22)</td>
<td>0.26 (0.03–2.53)</td>
<td>65</td>
<td>0.15</td>
<td>0.52 (0.22–1.21)</td>
</tr>
<tr>
<td>M. bovis</td>
<td>54</td>
<td>0.06</td>
<td>0.22 (0.06–0.78)</td>
<td>0.11 (0.006–0.56)</td>
<td>22</td>
<td>0.09</td>
<td>0.30 (0.06–1.67)</td>
</tr>
<tr>
<td>Residency</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>785</td>
<td>0.35</td>
<td>1 (Ref)</td>
<td>1,547</td>
<td>0.34</td>
<td>1 (Ref)</td>
<td>0.004</td>
</tr>
<tr>
<td>Rural</td>
<td>252</td>
<td>0.23</td>
<td>0.65 (0.45–0.94)</td>
<td>0.99 (0.69–1.43)</td>
<td>904</td>
<td>0.24</td>
<td>0.70 (0.55–0.89)</td>
</tr>
<tr>
<td>Alcohol abuse</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1,009</td>
<td>0.24</td>
<td>1 (Ref)</td>
<td>2,430</td>
<td>0.28</td>
<td>1 (Ref)</td>
<td>0.964</td>
</tr>
<tr>
<td>Yes</td>
<td>28</td>
<td>0.68</td>
<td>2.8 (1.22–6.42)</td>
<td>1.88 (0.89–3.99)</td>
<td>21</td>
<td>0.29</td>
<td>1.93 (0.28–3.74)</td>
</tr>
<tr>
<td>Drug abuse</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1,017</td>
<td>0.25</td>
<td>1 (Ref)</td>
<td>2,398</td>
<td>0.28</td>
<td>1 (Ref)</td>
<td>0.905</td>
</tr>
<tr>
<td>Yes</td>
<td>20</td>
<td>0.65</td>
<td>2.64 (0.98–7.03)</td>
<td>1.42 (0.58–3.49)</td>
<td>53</td>
<td>0.26</td>
<td>1.53 (0.42–2.18)</td>
</tr>
<tr>
<td>Traveler to country of endemicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>988</td>
<td>0.26</td>
<td>1 (Ref)</td>
<td>2,391</td>
<td>0.28</td>
<td>1 (Ref)</td>
<td>0.167</td>
</tr>
<tr>
<td>Yes</td>
<td>49</td>
<td>0.10</td>
<td>0.39 (0.14–1.09)</td>
<td>0.26 (0.09–0.70)</td>
<td>60</td>
<td>0.15</td>
<td>0.53 (0.22–1.30)</td>
</tr>
<tr>
<td>Homeless</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1,028</td>
<td>0.25</td>
<td>1 (Ref)</td>
<td>2,401</td>
<td>0.27</td>
<td>1 (Ref)</td>
<td>0.233</td>
</tr>
<tr>
<td>Yes</td>
<td>9</td>
<td>0.33</td>
<td>1.31 (0.25–6.99)</td>
<td>1.46 (0.58–3.49)</td>
<td>50</td>
<td>0.44</td>
<td>1.61 (0.74–3.49)</td>
</tr>
<tr>
<td>Site of disease</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pulmonary</td>
<td>883</td>
<td>0.26</td>
<td>1 (Ref)</td>
<td>1,949</td>
<td>0.29</td>
<td>1 (Ref)</td>
<td>0.129</td>
</tr>
<tr>
<td>Pulmonary + extrapulmonary</td>
<td>154</td>
<td>0.19</td>
<td>0.94 (0.58–1.53)</td>
<td>0.92 (0.67–0.99)</td>
<td>0.58–1.07</td>
<td>0.92 (0.67–0.99)</td>
<td></td>
</tr>
</tbody>
</table>

a Ref, reference.
partially consistent with those from a study conducted in the Gambia, where *M. africanum* was shown to transmit equally well to household contacts but less likely than *M. tuberculosis* to progress to disease (9). While numbers in our native population were too low to detect any associations in both indices, in the larger FB cohort, our findings suggest that lower infectivity might also be a component of the overall lower transmissibility of *M. africanum*. Perhaps because of this lower infectivity, we did not observe the previously reported lower estimate of progression to disease in *M. africanum*. Possible explanations for this disparity may lie in the slightly different definition of infectivity used in the Gambia, where they used the incidence of TST conversion (using a follow-up period of 3 months) specifically within households as the outcome. In addition, we may not be comparing exactly the same genotype; in our FB cohort, only 3 of 183 (1.7%) *M. africanum* strains with a known birth country came from the Gambia.

In a cohort of native and FB TB cases in Montreal, the EAI lineage was also significantly associated with lower number of TST-positive contacts around index cases and with less clustering (lower proportion of patients clustering, as defined by identical RFLP or spoligotypes) in multivariable analysis (18). It is interesting to observe this trend in our study, which includes only pulmonary cases of EAI, given the association this lineage has with the extrapulmonary site of disease (16). In a secondary cohort of only FB cases in the Montreal study, the EAI lineage was significantly associated with less TST positivity but not with less clustering (18). This again agrees with our study, where we observed a significant association of EAI with lower infectivity but not with lower clustering. These findings on the EAI genotype are hard to explain using the molecular epidemiological data from Vietnam, where approximately 40% of cases are caused by EAI strains and another 40%, by the Beijing genotype strains (19). If EAI strains are less successful at infecting, one would expect them to disappear in a few generations and be replaced by other, more fit, strains. This shift is perhaps occurring at the very moment, as Beijing genotype isolates have been associated with a lower age of patients and, hence, with active transmission.

Although *M. bovis* was spread significantly less in the native

---

**TABLE 3** Estimates of progression to disease by risk factor

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Median of bootstrapped progression-to-disease RR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, yr</td>
<td>Native cohort</td>
</tr>
<tr>
<td>0–19</td>
<td>1.09 (0.54–2.39)</td>
</tr>
<tr>
<td>20–39</td>
<td>1 (Ref)</td>
</tr>
<tr>
<td>40–64</td>
<td>0.83 (0.56–1.25)</td>
</tr>
<tr>
<td>≥65</td>
<td>0.76 (0.58–0.94)</td>
</tr>
<tr>
<td>Smear positivity</td>
<td>Negative</td>
</tr>
<tr>
<td>Negative</td>
<td>0.62 (0.27–1.52)</td>
</tr>
<tr>
<td>Positive</td>
<td>1 (Ref)</td>
</tr>
<tr>
<td>Lineage</td>
<td>Euro-American</td>
</tr>
<tr>
<td></td>
<td>1 (Ref)</td>
</tr>
<tr>
<td></td>
<td>CAS</td>
</tr>
<tr>
<td></td>
<td>EAI</td>
</tr>
<tr>
<td></td>
<td><em>M. africanum</em></td>
</tr>
<tr>
<td></td>
<td><em>M. bovis</em></td>
</tr>
<tr>
<td>Residency</td>
<td>Urban</td>
</tr>
<tr>
<td></td>
<td>Rural</td>
</tr>
<tr>
<td>Alcohol abuse</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Drug abuse</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
</tr>
</tbody>
</table>

*Ref, reference.*
population and although the estimates of average number of secondary cases were lower than other lineages, the fact that there were three documented secondary cases (from three different index cases with unique fingerprints) does not rule out the possible occurrence of transmission of M. bovis in The Netherlands, where pasteurization practices have been in place for decades. Ingestion of unpasteurized dairy products has been suggested as the likely route of infection in extrapulmonary cases in second-generation immigrants in The Netherlands who may have traveled back to their country of origin (20). Yet, all three M. bovis index cases with secondary cases in our clustering cohort also had pulmonary manifestations; two of these index cases were FB but had no indication of recent travel to a country of M. bovis endemcity. Indeed, instances of human-to-human transmission of M. bovis have been documented in other settings (21, 22). Together these observations suggest that, from a public health perspective, contact investigation and treatment of pulmonary M. bovis patients should not altogether differ from those of M. tuberculosis patients.

Unlike studies conducted in other populations, where the Beijing strain was associated with greater virulence and transmissibility, we did not find that the Beijing strain had higher indices of infectivity, clustering, or progression to disease in The Netherlands (23, 24). This is concordant with other recent studies conducted in similarly low-incidence, immigrant-receiving settings, such as the United States and Canada, which concluded that Beijing strains are no more of a public health threat than non-Beijing strains (25, 26). The observed higher success rate of Beijing strains may therefore result from circumstances characteristic of high-prevalence settings, such as mass use of BCG vaccination, development of resistance, crowding of the human population, and other unknown factors.

Other clinical and demographic factors positively associated with either infectivity or clustering indices, such as smear positivity, a lower age, and residing in an urban area, have been similarly described in previous studies (27–29). The significantly lower estimate of progression to disease given an elderly source likely reflects a lower dose of infection (due to a less close contact) and propensity for older patients to have older contacts themselves, as well as the higher proportion of long, latent infections (possibly associated with lower virulence) in this age category (30). Likewise, the significant association between alcohol and drug abuse with higher estimates of progression to disease can be linked to the direct effects of both substances on immunity, the indirect effects of substance-related disorders (i.e., malnutrition), and other potential confounding factors, such as homelessness (31, 32). There are two possible reasons behind the less-expected association between use of intravenous drugs and the lower average number of TST-positive contacts in the native cohort. Contacts of drug abusers are often intravenous drug users themselves, a scenario in which the accurate definition of a first-ring contact is prone to misclassification (contacts could be misclassified as first-ring contacts while they actually do not have much contact with an index case and, therefore, do not become TST positive). It has also been described that drug use can comprise cellular immunity (even in the absence of HIV infection) so that TST positivity in drug users is lower (33, 34).

The low prevalence setting of this study means that the investigation of the role of the M. tuberculosis genotype on transmission is less likely to be confounded by a high background infection pressure, where a TST result is more likely to fail at distinguishing recent from past infection. Furthermore, in The Netherlands there is no routine BCG vaccination program that could affect the interpretation of TST results, making TST a suitable tool for the detection of recent M. tuberculosis infection in contact investigations. This advantage applies solely to the native cohort, however, as patients in the FB cohort are far more likely to have been BCG vaccinated than native patients (40% versus 8%, respectively) and have had higher exposure to TB in their country of origin; both of these factors might lead to an overestimation of infectivity. It is encouraging, however, to observe the same trend of lower infectivity in EAI result in another study which did adjust for the probability of previous latent TB (18). On the other hand, the facts that FB patients often have FB contacts and that contact tracing in this group is less efficient imply that we might have also underestimated infectivity (and, by implication, biased the progression to disease index upward) in this group. The same reasoning applies to cases of addiction to alcohol and drugs, where an increased likelihood of homelessness means infected contacts are less likely to be found.

It is important to remember the potential shortcomings from the molecular epidemiology data underpinning these findings. A lack of clinical follow-up data of infected contacts meant that we were unable to link infected contacts to secondary cases and, thus, to estimate the proportion of secondary cases infected by a specific index case. In this low-burden country, however, there is likely a large overlap in the number of infected contacts around an index case and the number of secondary cases occurring within 2 years of that index case. It nevertheless meant that we could not control for risk factors across the transmission chain, such as rates of latent TB treatment and existing medical risk factors in secondary cases, which could influence the likelihood of progression to disease or the susceptibility to infection of the host, respectively. Studies using a prospective cohort approach (i.e., with access to household contacts and TST conversion data) that can bypass some of these issues are warranted to confirm these findings.

In sum, the lower infectivity or overall transmissibility observed in this study for M. bovis, M. africanum, and EAI—all, ancient lineages—matches the hypothesis that modern strains, as a consequence of their access to rapidly increasing numbers of susceptible hosts, have been selected for more rapid disease progression and transmission (35). Validation of this scenario via future experimental studies could have important implications on how TB control efforts may be determined not only by index case characteristics, but also by a bacterial signature, such as phylogenetic lineage.

ACKNOWLEDGMENTS

This study was supported by the Portuguese Foundation for Science and Technology (FCT) (reference SFRH/BD/33902/2009 to H.N.-G).

We thank Rogier Donders, Megan Murray, and Maha Farhat for their statistical support and discussion of the methodology used. We thank the staff of the RIVM mycobacteriological laboratory for their work on the RFLP and VNTR typing of M. tuberculosis isolates, and we thank the Municipal Health Services for their voluntary collaboration in the nationwide tuberculosis surveillance.

REFERENCES
