OPEN ACCESS

ORIGINAL ARTICLE

Recent TB transmission, clustering and predictors of large clusters in London, 2010–2012: results from first 3 years of universal MIRU-VNTR strain typing

Esther L Hamblion,[1] Arnaud Le Menach,[1,2] Laura F Anderson,[3] Maeve K Lalor,[3] Tim Brown,[4] Ibrahim Abubakar,[3,5] Charlotte Anderson,[1] Helen Maguire,[1,2] Sarah R Anderson,[1] on behalf of the Public Health England Strain Typing Project Board

For numbered affiliations see end of article.

Correspondence to
Dr Esther L Hamblion, Field Epidemiology Services (Victoria), Public Health England, 151 Buckingham Palace Road, London SW1W 9SZ, UK; estherhamblion@hotmail.com

ELH and ALM contributed equally.

HM and SRA both supervised production of this work.

Linked

CrossMark

ABSTRACT

Background The incidence of TB has doubled in the last 20 years in London. A better understanding of risk groups for recent transmission is required to effectively target interventions. We investigated the molecular epidemiological characteristics of TB cases to estimate the proportion of cases due to recent transmission, and identify predictors for belonging to a cluster.

Methods The study population included all culture-positive TB cases in London residents, notified between January 2010 and December 2012, strain typed using 24-loci multiple interspersed repetitive units-variable number tandem repeats. Multivariable logistic regression analysis was performed to assess the risk factors for clustering using sociodemographic and clinical characteristics of cases and for cluster size based on the characteristics of the first two cases.

Results There were 10 147 cases of which 5728 (57%) were culture confirmed and 4790 isolates (84%) were typed. 2194 (46%) were clustered in 570 clusters, and the estimated proportion attributable to recent transmission was 34%. Clustered cases were more likely to be UK born, have pulmonary TB, a previous diagnosis, a history of substance abuse or alcohol abuse and imprisonment, be of white, Indian, black-African or Caribbean ethnicity. The time between notification of the first two cases was more likely to be <90 days in large clusters.

Conclusions Up to a third of TB cases in London may be due to recent transmission. Resources should be directed to the timely investigation of clusters involving cases with risk factors, particularly those with a short period between the first two cases, to interrupt onward transmission of TB.

Key messages

What is the key question?
► How much transmission of TB is occurring in London and what are the characteristics of cases that cluster and contribute to TB transmission?

What is the bottom line?
► Up to a third of TB cases in London can be attributed to recent transmission, more than double that found when previous population-based studies were undertaken over 15 years ago.

Why read on?
► We describe the characteristics of those in clusters attributable to recent transmission, explain the characteristics of those in large clusters compared to small clusters, providing an important update on TB transmission in London at a time when TB is being prioritised in England with the release of a national TB strategy.

INTRODUCTION

In the UK, TB incidence has been rising over the last 20 years[1,2] mainly in urban centres, especially London, one of the largest, most densely populated and ethnically diverse cities in Europe.[3] In 2012, nearly 3500 new TB cases were reported in London accounting for 39% of all UK cases,[2] and representing an incidence of 41 per 100 000 inhabitants higher than the national average of 14 per 100 000.[2,4] There was geographical variation across London with rates ranging from 117 per 100 000 in Newham local authority to 7 per 100 000 in Richmond-upon-Thames local authority.[2] The majority were young adults aged 20–39 years (55%) and non-UK born (83%). Nearly 1 in 10 had at least one social risk factor, with more than a third of these reporting more than one.[2]

Strain typing data, when combined with epidemiological data, enables the identification of patients with TB involved in the same chain of transmission. This assists the implementation of timely and appropriate control measures[5–7] such as enhanced contact tracing and active case finding.[8,9] Strain typing can distinguish between disease resulting from reactivation of an infection acquired previously and a new transmission event as it assumes that cases sharing indistinguishable strain types are recently infected.[10] The effectiveness of TB control activities in reducing or preventing infection can be evaluated by monitoring changes in recent TB transmission rates.[8,9]

Although reports on a number of outbreaks using multiple interspersed repetitive units-variable number tandem repeats (MIRU-VNTR) strain typing have been published,[11–14] few population level studies have been carried out in the UK based on strain typing to date. A London-wide strain typing study in 1995–1997 showed a low rate of active transmission of 14%,[15] and that the burden of disease in the city at that time was largely due to reactivation or importation of infection by recent migrants. The strain typing method used was restriction fragment length polymorphism (RFLP) typing, the gold standard at the time.[15]

Since 2010 universal prospective strain typing using 24-loci MIRU-VNTR has been implemented in England and Wales.[8] The objectives of our study were to estimate the proportion of TB cases in London attributable to recent transmission to identify the characteristics of those in clusters, and to compare the characteristics of the first two cases in small versus large clusters in order to predict risk factors for cluster growth.

METHODOLOGY
Study population
We included all culture-confirmed *Mycobacterium tuberculosis* (*M.tb*) cases, resident in London, with a notification date between 1 January 2010 and 31 December 2012, strain typed using 24-loci MIRU-VNTR[16 17] at the Mycobacteriology Reference Laboratories (MRLs) of Public Health England (PHE) (figure 1).

Definitions
A cluster needed to contain at least one case with a full 24-loci strain type. A clustered case was defined as any TB case from the study population with a strain with at least 23 typed loci, indistinguishable from that of at least one other case. A non-clustered case was defined as any TB case from the study population having a unique strain not shared by any other case. Clusters were defined as large if they contained five or more cases.

Geographical analysis was based upon the existing organisational jurisdictions of PHE in London (total population in 2012, 8.2 million). There are four health protection team (HPT) areas in London, each with a radius of approximately eight miles, and containing populations ranging from 1.4 to 2 million. Clusters that contained cases resident in more than one of the four HPT geographical footprints of London (South-West London, South-East London, North-West London, North-East and North-Central London), were classified as regional; those with all cases in the same HPT area were classified as local.

Notification interval (the number of days between notification date of the first and second case in a cluster) was estimated and dichotomised as short or long for each cluster according to whether it was more or less than 90 days. This was chosen based on previous work by Kik *et al*.[18]

Recent transmission of TB infection was presumed to have occurred when a case had an identical strain typing pattern to another case in the population during the 3-year period of the study.

Data collection
TB cases in London are notified by clinic staff to the PHE London TB Register, and data are then imported into the national Enhanced TB Surveillance (ETS) system. Isolate-related data were provided by the MRLs and matched to patient-related ETS data using patient identifiers as previously described.[4 19] We collected individual data including demographics (date of birth, gender, ethnicity, country of birth and year of entry to the UK), social risk factors (current or history of, alcohol misuse, drug misuse, imprisonment or homelessness), clinical features (site of disease, sputum smear status, dates of symptom onset, diagnosis and notification, BCG vaccination status, previous TB diagnosis) and microbiology (24-loci MIRU-VNTR strain type (the first isolate from each case), drug sensitivities and lineage).[19] The *M.tb* lineage was established as previously described.[20] We derived geographical categories (regional or local) from the patient's postcode of residence and size of cluster from the number of cases.

Data analysis
We calculated proportion of recent TB transmission by the 'n minus one' method according to the formula (number of clustered isolates−number of clusters)/number of isolates with a strain type.[21]

We described features of culture confirmed cases, clustered cases and of strain typed clusters. Characteristics of cases with a strain type were compared with those without using χ^2 tests for categorical variables (or Fisher exact test where necessary), and a t test for age.

We performed univariable logistic regression to identify individual risk factors for clustering using odd ratios (ORs), and multivariable logistic regression including sex and age, as potential confounders along with variables associated with clustering in univariable analysis (included if p ≤0.2). Likelihood ratio tests were used to assess possible interactions. Whether a case was born in the UK or not was excluded in the multivariable model due to colinearity with the composite variable UK/non-UK born and ethnicity.

We performed multivariable logistic regression with cluster size (large vs small) as an outcome. All explanatory variables relating to the characteristics of the first two notified cases in the cluster were included if associated with cluster size according to a univariable analysis (p ≤0.2) or potential confounders. An age group category of ≤30 years was chosen, as the highest rate and number of cases of TB in London is found in those aged 20–29 years. The variable indicating whether a case had any social risk factors was excluded from the model due to colinearity with the history of or currently being in prison variable which was included. Potential interactions were assessed using likelihood ratio tests.

RESULTS
Isolates and strain typed clusters
Between 1 January 2010 and 31 December 2012, 10 147 TB cases were notified in London residents (figure 1). Of those, 5728 (57%) were culture-confirmed as *M.tb*. Individuals with *M.tb* culture confirmation were younger than those without (37 vs 38 years old on average, p<0.01), more likely to be male (59% vs 41%, p<0.01), born outside the UK (57% vs 43%, p<0.01), have pulmonary disease (68% vs 32%, p<0.01), less likely to have extrapulmonary disease (46% vs 54%, p<0.01) and have at least one social risk factor (67% vs 33%, p<0.01). Of the 5728 culture-confirmed *M.tb* cases, 4790 isolates were typed with at least 23 loci (84%), and 2194 (46% (95% CI 44.4% to 47.2%)) were clustered in 570 clusters. The proportion of cases attributable to recent transmission in London was estimated at 34% (95% CI 32.6% to 35.3%).

The average cluster size was four individuals (range 2–55). Over half of clusters, 319 (56%) contained two individuals and just 35 (6%) had 10 or more individuals. There were 215 local (38%) and 355 (62%) regional clusters (figure 2). Of the 570
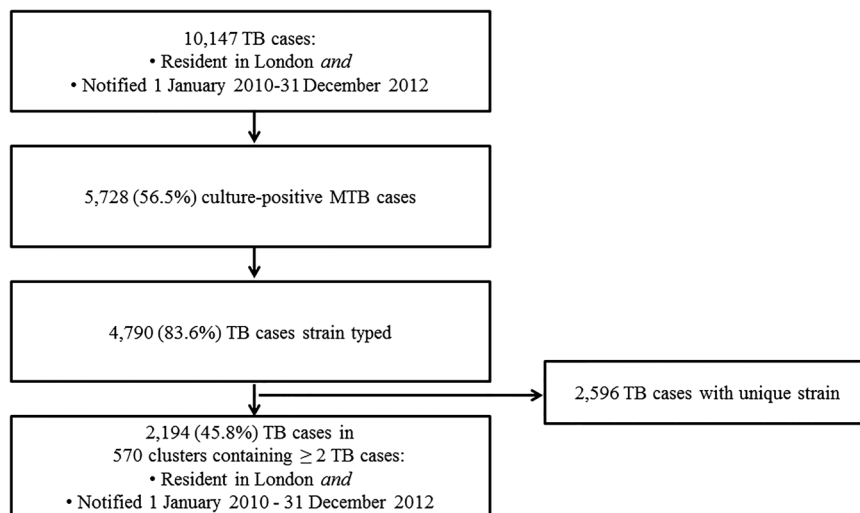
Figure 1 Number of reported cases of TB, including culture-positive cases, strain typed cases and clusters, London, 2010–2012. MTB, Mycobacterium tuberculosis.

clusters, 351 (62%) comprised only non-UK born individuals and 36 (6%) only UK born. A quarter of the clusters (n=139) had at least one individual with one or more social risk factors and in nine clusters (2%) all individuals had one or more social risk factor. The most common phylogenetic lineage was the Euro-American, reported for 230 (40%) of the clusters, 170 (30%) were of Central-Asian lineage, 73 (13%) of East-African-Indian lineage and 33 (6%) of Beijing lineage.

Characteristics of clustered cases

The proportion of clustered cases decreased with age from 79% in individuals below 10 years to 37% in those >70 years old (p<0.01). The proportion of clustering varied across ethnic groups according to the following distribution (p<0.01); 76% of black-Caribbean cases; 56% of black-African; 54% of white; 43% of Pakistani; 39% of Indian and 8% of Bangladeshi cases. Clustered cases were more likely to have been born in the UK than those not in clusters (67% vs 33%, p<0.01). If not born in the UK, clustered cases were more likely to be resident in the UK for a long time: 39% of recent migrants (entry to UK ≤4 years before notification) were clustered compared with 45% of those who entered the UK five or more years ago (p<0.01) (table 1). The proportion of clustering did not vary with

resistance pattern, whether any resistance (48% vs 46%, p=0.30), isoniazid mono-resistant (49% vs 46%, p=0.25) or multidrug resistance (53% vs 46%, p=0.19).

Factors associated with clustering

In the univariable analysis clustered cases were consistently more likely to have at least one social risk factor (57% vs 45%, p<0.001): 68% and 58% of those with a history of drug and alcohol misuse were clustered respectively, 54% of cases with a history of homelessness, and 68% of TB cases with a history of imprisonment were clustered (table 1).

In a stratified analysis currently having problem alcohol use modified the association between clustering and prison history (likelihood ratio test p=0.02). Clustered individuals who abused alcohol were 6.3 (2.2–18.0) times more likely to also have spent time in prison, while those without problem alcohol use were only 1.9 (1.3–2.8) times more likely to have spent time in prison (Mantel–Haenszel test of homogeneity χ^2 4.8, p value=0.03), therefore, an interaction term was added into the model. No other interactions were identified.

The final multivariable model adjusted for age, gender and all significant variables (p<0.05), demonstrated clustered cases were more likely to be male (aOR=1.2 (1.0–1.3)), born in the UK and from a black-Caribbean, black-African, white or Indian background (aOR=6.7 (3.7–12.2), 4.4 (2.6–7.6), 2.3 (1.7–3.2) and 1.9 (1.1–3.3), respectively) or born abroad and from a black-Caribbean (aOR=3.3 (1.8–5.8) or black-African (aOR=1.8 (1.5–2.2)) ethnic background, to have been previously diagnosed with TB (aOR=2.1 (1.5–3.0)), to have smear positive (aOR=1.5 (1.3–1.8)) or negative (aOR=1.4 (1.2–1.7)) pulmonary TB compared with extrapulmonary disease, to have a history of drug misuse (aOR=1.6 (1.0–2.4)), and to have a history of imprisonment and problem alcohol use (aOR=3.3 (1.2–9.3)) and were less likely to be born abroad and from a Bangladeshi ethnic background (aOR=0.5 (0.3–0.7)) (table 2).

Factors associated with small and large clusters

The analysis was carried out on 570 clusters in which the first case was notified between 1 January 2010 and 31 December 2012, 101 of which were classified as large (18%) and 469 (82%) as small (table 3).
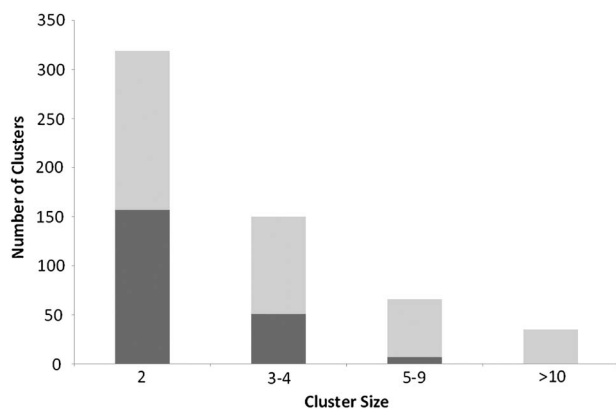

Figure 2 Number of clusters according to cluster size and geography (dark grey represents local cluster and light grey represents regional cluster), London, 2010–12 (N=570).

Tuberculosis

Table 1 Factors associated with TB clustering; a univariable logistic regression analysis, universal MIRU-VNTR, London 2010–2012

Variable	Total no of cases	No (%) clustered	OR (95% CI)	p Value*
Gender	4788			
Female	1876	830 (44.2)	Reference	
Male	2912	1363 (46.8)	1.1 (1.0 to 1.2)	0.08
Age	4790			
0–4	18	16 (88.9)	**9.3 (2.1 to 40.4)**	**<0.01**
5–9	15	10 (66.7)	2.3 (0.8 to 6.8)	0.13
10–14	76	44 (57.9)	1.6 (1.0 to 2.5)	0.05
15–19	266	162 (60.9)	**1.8 (1.4 to 2.3)**	**<0.01**
20–29	1628	755 (46.4)	Reference	
30–39	1169	509 (43.5)	0.9 (0.8 to 1.0)	0.14
40–49	679	298 (43.9)	0.9 (0.7 to 1.1)	0.27
50–59	421	187 (44.4)	0.9 (0.7 to 1.1)	0.47
60–69	252	114 (45.2)	1.0 (0.7 to 1.2)	0.74
>70	266	99 (37.2)	**0.7 (0.5 to 0.9)**	**0.01**
UK/non-UK born and ethnicity				
UK born				
Indian	68	39 (57.4)	**2.2 (1.4 to 3.6)**	**<0.01**
Pakistani	32	18 (56.3)	**2.1 (1.0 to 4.3)**	**0.04**
Bangladeshi	26	12 (46.2)	1.4 (0.6 to 3.1)	0.38
White	305	186 (61.0)	**2.6 (2.0 to 3.3)**	**<0.01**
Black-Caribbean	95	79 (83.2)	**8.1 (4.7 to 14.1)**	**<0.01**
Black-African	110	86 (78.2)	**5.9 (3.7 to 9.4)**	**<0.01**
Other†	74	54 (72.3)	**4.5 (2.6 to 7.5)**	**<0.01**
Non-UK born				
Indian	1364	515 (37.8)	Reference	
Pakistani	395	163 (41.3)	1.2 (0.9 to 1.5)	0.21
Bangladeshi	205	51 (24.9)	**0.5 (0.4 to 0.8)**	**<0.01**
White	186	79 (42.5)	1.2 (0.9 to 1.7)	0.22
Black-Caribbean	74	49 (66.2)	**3.2 (2.0 to 5.3)**	**<0.01**
Black-African	969	518 (53.5)	**1.9 (1.6 to 2.2)**	**<0.01**
Other†	777	297 (38.2)	1.0 (0.9 to 1.2)	0.83
Born in UK	4711			
UK	714	476 (66.7)	**2.8 (2.3 to 3.3)**	**<0.01‡**
Non-UK born	3997	1678 (42.0)	Reference	
If non-UK born, time since arrival	3406			
In UK <2 years	811	340 (41.9)	Reference	
In UK 2–4 years	835	310 (37.1)	0.8 (0.7 to 1.0)	0.05
In UK 5–9 years	703	309 (44.0)	1.1 (0.9 to 1.3)	0.43
In UK >9 years	1057	479 (45.3)	1.1 (1.0 to 1.4)	0.14
Country of birth	3290			
UK	714	476 (66.7)	Reference	
India	1224	457 (37.3)	**0.3 (0.2 to 0.4)**	**<0.01‡**
Pakistan	384	159 (41.4)	0.4 (0.3 to 0.5)	0.14
Somalia	431	269 (62.4)	**0.8 (0.6 to 1.1)**	**<0.01‡**
Bangladesh	208	50 (24.0)	**0.2 (0.1 to 0.2)**	**<0.01‡**
Nepal	189	102 (54.0)	**0.6 (0.4 to 0.8)**	**<0.01‡**
Nigeria	140	69 (49.3)	**0.5 (0.3 to 0.7)**	**<0.01‡**
London health protection team	4790			
North-East and central	1932	887 (45.9)	Reference	
North-West	1647	741 (45.0)	1.0 (0.8 to 1.1)	0.58
South-East	696	343 (49.3)	1.1 (1.0 to 1.4)	0.13
South-West	515	223 (43.3)	0.9(0.7 to 1.1)	0.29
Social risk factors	4670			
Any	493	279 (56.6)	**1.6 (1.3 to 1.9)**	**<0.01‡**
None	4177	1866 (44.7)	Reference	
Imprisonment	4557			
Yes	154	105 (68.2)	**2.6 (1.9 to 3.7)**	**<0.01‡**
No	4403	1983 (45.0)	Reference	

Table 1 Continued

Variable	Total no of cases	No (%) clustered	OR (95% CI)	p Value*
Homelessness	4624			
Yes	202	109 (54.0)	**1.4 (1.1 to 1.9)**	**0.02**
No	4422	2008 (45.4)	Reference	
Alcohol misuse	4273			
Yes	203	117 (57.6)	**1.7 (1.2 to 2.2)**	**<0.01‡**
No	4070	1837 (45.1)	Reference	
Drug misuse	4521			
Yes	174	119 (68.4)	**2.7 (1.9 to 3.7)**	**<0.01**
No	4357	1947 (44.8)	Reference	
Alcohol misuse and prison history	4214			
Yes	35	30 (85.7)	**7.3 (2.8 to 18.7)**	**<0.01**
No	4179	1891 (45.3)	Reference	
Clinical presentation	4790			
Extrapulmonary	760	760 (38.5)	Reference	
Pulmonary smear negative	1107	546 (49.3)	**1.6 (1.3 to 1.8)**	**<0.01**
Pulmonary smear positive	1247	667 (53.5)	**1.8 (1.6 to 2.1)**	**<0.01**
Pulmonary smear status unknown	464	221 (47.6)	**1.5 (1.2 to 1.8)**	**<0.01**
Previous TB	4567			
Yes	198	119 (60.1)	**1.8 (1.4 to 2.5)**	**<0.01**
No	4369	1968 (45.0)	Reference	
Resistance	4784			
Any	447	215 (48.1)	1.1 (0.9 to 1.3)	0.30
None	4337	1975 (45.5)	Reference	
Isoniazid mono-resistant	4784			
Yes	239	118 (49.4)	1.2 (0.9 to 1.5)	0.25
No	4545	2072 (45.6)	Reference	
MDR TB	4784			
Yes	81	43 (53.1)	1.3 (0.9 to 2.1)	0.19
No	4703	2147 (45.7)	Reference	
BCG vaccination	3741			
Yes	2727	1251 (45.9)	1.0 (0.9 to 1.2)	0.96
No	1014	466 (46.0)	Reference	
Year of notification	4790			
2010	1242	565 (45.5)	Reference	
2011	1716	758 (44.2)	0.9 (0.8 to 1.1)	0.48
2012	1832	871 (47.5)	1.1 (0.9 to 1.3)	0.26

*Significant results are in bold, those p≤0.2 included in multivariable model.
†'Other' ethnicity category includes all those who stated their ethnicity was Chinese, black-other or mixed/other.
‡Not included in final model due to colinearity with an included variable.
MDR, multi drug resistant; MIRU-VNTR, multiple interspersed repetitive units-variable number tandem repeats.

In the final multivariable model adjusted for age and gender, large clusters compared with small clusters were more likely to have <90 days (a short notification interval) between notification of the first two cases (aOR=2.9 (1.8–4.6)), cluster regionally rather than locally (aOR=10.3 (4.6–22.9)) and one of the first two notified cases to have a history of imprisonment (aOR=2.8 (1.1–7.3)) (table 3).

DISCUSSION

We estimated that a third of TB cases in London between 2010 and 2012 were due to recent transmission when applying the n−1 approach which denies the possibility of more than one index case. Our estimates of clustering (46%) and of recent transmission (34%) remain within observed ranges for low TB burden countries; estimates from England over the same time period show 52% clustering and 40% due to recent transmission;[4] a national study carried out in the Netherlands during 1993–1997 found estimates for clustering and recent transmission of 46% and 35%,

respectively[22]; a study of seven sentinel surveillance sites in the USA during 1996–2000 estimated clustering at 48%;[23] and a cross-sectional study during 2005–2009 estimated recent transmission as 23%,[9] although alternative laboratory techniques and a longer period of study were used in both the latter studies.

In London, previous estimates of recent transmission in 1995–1997 and in 1998 were lower at 14% and 7%, respectively,[15 24] indicating an increase in recent transmission over the last 15 years. The studies, however, are not directly comparable since an alternative laboratory technique was used, RFLP.

Individuals with both smear positive and smear negative pulmonary disease were more likely to cluster than those with extrapulmonary disease. Although it should be noted that those with extrapulmonary disease are less likely to be cultured. Patients with smear negative disease are less infectious but do transmit disease.[25 26] This could be a limitation of the laboratory methodology, or perhaps a reflection that there are more opportunities for infecting others due to the presumption that those with smear negative disease are less infectious, and

Table 2 Factors associated with TB clustering; a multivariable logistic regression analysis, universal MIRU-VNTR, London 2010–2012

Variable	Adjusted OR (95% CI)	p Value*
Gender		
Female	Reference	
Male	**1.2 (1.0 to 1.3)**	**0.02**
Age		
0–4	3.6 (0.8 to 17.5)	0.11
5–9	1.0 (0.3 to 3.7)	0.97
10–14	1.0 (0.6 to 1.8)	1.00
15–19	1.2 (0.9 to 1.7)	0.15
20–29	Reference	
30–39	0.9 (0.8 to 1.1)	0.31
40–49	**0.7 (0.6 to 0.9)**	**<0.01**
50–59	0.8 (0.6 to 1.0)	0.09
60–69	1.0 (0.8 to 1.4)	0.83
>70	**0.7 (0.5 to 1.0)**	**0.04**
UK/non-UK born and ethnicity		
UK born		
Indian	**1.9 (1.1 to 3.3)**	**0.02**
Pakistani	1.7 (0.8 to 3.5)	0.18
Bangladeshi	1.5 (0.7 to 3.4)	0.35
White	**2.3 (1.7 to 3.2)**	**<0.01**
Black-Caribbean	**6.7 (3.7 to 12.2)**	**<0.01**
Black-African	**4.4 (2.6 to 7.6)**	**<0.01**
Other	**4.0 (2.1 to 7.5)**	**<0.01**
Non-UK born		
Indian	Reference	
Pakistani	1.1 (0.9 to 1.4)	0.38
Bangladeshi	**0.5 (0.3 to 0.7)**	**<0.01**
White	1.0 (0.7 to 1.5)	0.85
Black-Caribbean	**3.3 (1.8 to 5.8)**	**<0.01**
Black-African	**1.8 (1.5 to 2.2)**	**<0.01**
Other	1.0 (0.8 to 1.2)	0.91
Social risk factors		
Homelessness		
Yes	0.8 (0.5 to 1.2)	0.30
No	Reference	
Drug misuse		
Yes	**1.6 (1.0 to 2.4)**	**0.03**
No	Reference	
Alcohol misuse and prison history		
Yes	**3.3 (1.2 to 9.3)**	**0.03**
No	Reference	
Clinical presentation		
Extrapulmonary	Reference	
Pulmonary smear negative	**1.4 (1.2 to 1.7)**	**<0.01**
Pulmonary smear positive	**1.5 (1.3 to 1.8)**	**<0.01**
Pulmonary smear status unknown	**1.3 (1.0 to 1.6)**	**0.04**
Previous TB		
Yes	**2.1 (1.5 to 3.0)**	**<0.01**
No	Reference	
MDR TB		
Yes	1.2 (0.7 to 2.0)	0.54
No	Reference	

*Significant results are in bold.
MDR, multi drug resistant; MIRU-VNTR, multiple interspersed repetitive units-variable number tandem repeats.

therefore, less precautions are taken in limiting onward transmission from these patients.[25] Patients with smear positive or negative disease do appear to be contributing to recent transmission of TB and should, therefore, be followed up appropriately during public health investigations.

Clustered cases were more likely to misuse drugs and currently be in or have a history of imprisonment along with problem alcohol use at diagnosis. In addition, one of the first two notified cases in large clusters was more likely to have a history of imprisonment than in small clusters. Cases who misuse drugs or alcohol may be less likely to promptly seek healthcare, and therefore, may have longer periods of infectiousness. Furthermore, they may be either reluctant to declare contacts or not be aware of the full names of contacts. It seems likely that contact tracing in these groups may not be as effective, leading to ongoing transmission.[27]

A possible explanation for the shorter interval between first cases in larger clusters may be a delay in diagnosis of the initial case resulting in a longer infectious period with greater potential to infect others or a breakdown in contact tracing for those early cases resulting in failure to offer prophylaxis to infected contacts. It could be that those in such clusters share a common exposure in the past or recently (in the UK or abroad) and manifest illness at a similar time. We have not examined individual epidemiological links between patients or yield of contact tracing for cases to explore if these are valid explanations but this will be the subject of further work.

Place and time of transmission is often difficult to identify, while we assume cases resident in London and sharing the same strain were recently infected in London, this may not be so. Recent migrants sharing the same strain may have acquired the common strain of TB circulating in their country of origin.[28]

Our study has some limitations. There was a high proportion (43%) of TB cases without a culture, and therefore, not strain typed. Cases without culture confirmation can contribute to transmission chains,[10] and therefore, we may have underestimated clustering. Isolates from cases reported at the beginning or at the end of the study period may have been misclassified as unique strains if they were in clusters occurring either before or at the end of the study period. This would also lead to underestimation of clustering. Conversely, strains with one missing loci appearing indistinguishable from another were included even though they may not be genetically related. This could have led to an increase in the estimate of clustering.

Some clusters may have been misclassified as smaller than they actually were should we have taken a wider geographical area than London into consideration. Furthermore, clusters classed as small near the end of the study period may actually be large given time to grow.[18] In addition, only characteristics of the first two notified cases were analysed, which do not necessarily represent the first cases in a chain of transmission, or cases with the earliest acquisition of infection. However, information about those cases may help promptly identify clusters more likely to grow. Although we studied cluster growth within a 3-year time period it is possible that studying the same population over a longer time period may yield somewhat different results. But we would not anticipate a major change in the findings or the direction of the results. We also recommend that this analysis is repeated using five years of data.

Our study provides important updated information about current TB transmission in London using new strain typing techniques more than 10 years after previous studies were published. One-third of cases were found to be due to recent transmission and we have identified population groups in whom that is more common. This provides useful information for TB services and policy makers to help identify where resources may be best deployed.

Table 3 Cluster size and association with demographics, social risk factors and clinical characteristics of the first two cases in a TB cluster; a univariable and multivariable logistic regression model, universal MIRU-VNTR, London 2010–2012 (N=570)

Variable	No (%) of large (≥5 cases) clusters	No (%) of small (2–4 cases) clusters	Univariable analysis		Multivariable analysis	
			OR (95% CI)	p Value*	Adjusted OR (95% CI)	p Value*
Total	101 (17.7)	469 (82.3)				
Gender						
≥One male	85 (84.2)	391 (83.4)	1.1 (0.6 to 1.9)	0.85	0.8 (0.4 to 1.6)	0.55
No male	16 (15.8)	78 (16.6)				
Age						
At least one case ≤30 years	77 (76.2)	329 (70.2)	1.4 (0.8 to 2.2)	0.22	1.5 (0.9 to 2.6)	0.14
No cases ≤30 years	24 (23.8)	140 (29.9)				
Country of birth						
≥One case UK born	29 (28.7)	137 (29.2)	1.0 (0.6 to 1.6)	0.92		
Non-UK born	72 (71.3)	332 (70.8)				
Social risk factors						
≥One case with any social risk factors	22 (21.8)	75 (16.0)	1.5 (0.9 to 2.5)	0.16†		
No social risk factors	79 (78.2)	394 (84.0)				
History of or currently in prison						
≥One case with prison history	10 (9.9)	12 (2.6)	**5.7 (1.9 to 17.5)**	**<0.01**	**2.8 (1.1 to 7.3)**	**0.04**
	91 (90.1)	457 (97.4)				
History of or currently homeless						
≥One case with history or currently homeless	13 (12.9)	31 (6.6)	1.8 (0.7 to 4.3)	0.21		
	88 (87.1)	438 (93.4)				
History of or current drug misuse						
≥One case with drug misuse	10 (9.9)	24 (5.1)	2.0 (0.7 to 5.8)	0.21		
	91 (90.1)	445 (94.9)				
History of or current alcohol misuse						
≥One case with alcohol misuse	10 (9.9)	37 (7.9)	1.7 (0.7 to 4.1)	0.26		
	91 (90.1)	432 (92.1)				
Clinical presentation						
First two cases extrapulmonary only	16 (15.8)	79 (16.8)	0.9 (0.5 to 1.7)	0.81		
	85 (84.2)	390 (83.2)				
First two cases pulmonary only	50 (49.5)	204 (43.5)	1.3 (0.8 to 2.0)	0.27		
	51 (50.5)	265 (56.5)				
≥One case with smear negative pulmonary TB	43 (42.6)	196 (41.8)	1.0 (0.7 to 1.6)	0.89		
	58 (57.4)	273 (58.2)				
≥One case with smear positive pulmonary TB	52 (51.5)	226 (48.2)	1.1 (0.7 to 1.8)	0.55		
	49 (48.5)	243 (51.8)				
Notification interval						
<90 days between first two cases	52 (51.5)	120 (25.6)	**3.1 (2.0 to 4.8)**	**<0.001**	**2.9 (1.8 to 4.6)**	**<0.01**
≥90 days between first two cases	49 (48.5)	349 (74.4)				
Drug resistance						
At least one case with any resistance	12 (11.9)	57 (12.2)	1.0 (0.5 to 1.9)	0.94		
No resistance	89 (88.1)	412 (87.9)				
Previous TB diagnosis						
≥One case with previous TB diagnosis	15 (14.9)	49 (10.5)	1.5 (0.8 to 2.8)	0.21		
No previous TB diagnosis	86 (85.2)	420 (89.6)				
Geography						
Clusters regionally	94 (93.1)	261 (55.7)	**10.7 (4.9 to 23.6)**	**<0.001**	**10.3 (4.6 to 22.9)**	**<0.01**
Clusters locally	7 (6.9)	208 (44.4)				

*Significant results are in bold, those p≤0.2 included in multivariable model.
†Not included in final model due to colinearity with prison history variable.
MIRU-VNTR, multiple interspersed repetitive units-variable number tandem repeats.

We would recommend that TB services target contact tracing efforts to those with social risk factors such as substance abuse and prison history,[29][30] and ensure all their contacts are identified, seen and screened to promote early case detection and prevent further transmission. Enhanced interventions are also needed such as the continuation of 'Find and Treat' as a service to actively case find among populations with social risk factors[31] and to ensure through outreach, that such patients complete treatment.[32] We recommend contact tracing should be thorough and prioritised in clusters where the time lag between the first two notified cases is less than 3 months. The advent of whole genome sequencing for TB is likely, when combined with the

epidemiological surveillance data, to enable further discrimination of the epidemiological links between individuals in clusters and will, therefore, help us better direct the resources required for improved TB control. Strategically, there have been changes that support the TB control effort; TB is being prioritised in the UK with the development of a national TB strategy[33] and within London the formation of a London TB Control Board. Our work contributes to the understanding of TB transmission in London and provides evidence for determining appropriate local TB control strategies.

Author affiliations
[1]Field Epidemiology Services (Victoria), Public Health England, London, UK
[2]European Programme for Intervention Epidemiology Training (EPIET), European Centre for Disease Control and Prevention, Stockholm, Sweden
[3]Tuberculosis Section, Respiratory Disease Department, Public Health England, London, UK
[4]National Mycobacterium Reference Laboratory, Institute of Cell and Molecular Science (ICMS), London, UK
[5]Centre for Infectious Disease Epidemiology, University College London, Mortimer Market Centre, London, UK

Contributors HM had the original idea and initiated the project. The project was overseen and led by SRA during September 2010 to September 2013 and thereafter by HM. ALM designed data collection tools and extracted data. ELH subsequently extracted further data and performed additional analyses. ALM wrote the initial statistical analysis plan, cleaned and analysed the original data and drafted the first version of the paper. ELH drafted subsequent versions after developing the statistical analysis plan and cleaning and analysing further data for 2012. All authors contributed to interpretation of findings, drafting of the paper, approval for publication and are accountable for the quality and integrity of the work. HM and SRA are guarantors and have provided final approval for submission.

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES
1 World Health Organisation. *Global tuberculosis report 2013*. 2013.
2 Public Health England. Tuberculosis in London: annual review (2012 data). 2013.
3 Anderson SR, Maguire H, Carless J. Tuberculosis in London: a decade and a half of no decline [corrected]. *Thorax* 2007;62:162–7.
4 Public Health England. *Tuberculosis in the UK—2013 report*. 2013.
5 de Vries G, van Hest RA, Burdo CC, et al. A *Mycobacterium tuberculosis* cluster demonstrating the use of genotyping in urban tuberculosis control. *BMC Infect Dis* 2009;9:151.
6 Gardy JL, Johnston JC, Ho Sui SJ, et al. Whole-genome sequencing and social-network analysis of a tuberculosis outbreak. *N Engl J Med* 2011;364:730–9.
7 Nava-Aguilera E, López-Vidal Y, Harris E, et al. Clustering of *Mycobacterium tuberculosis* cases in Acapulco: Spoligotyping and risk factors. *Clin Dev Immunol* 2011;2011:408375.
8 Okereke E. *TB Strain Typing Cluster Investigation Handbook for Health Protection Units*. London: Health Protection Agency, 2011. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/415232/TBStrain_Typing_Handbook_2nd_edition_September_2011.pdf

9 Moonan PK, Ghosh S, Oeltmann JE, et al. Using genotyping and geospatial scanning to estimate recent *Mycobacterium tuberculosis* transmission, USA. *Emerg Infect Dis* 2012;18:458–65.
10 de Vries G, Baars HW, Sebek MM, et al. Transmission classification model to determine place and time of infection of tuberculosis cases in an urban area. *J Clin Microbiol* 2008;46:3924–30.
11 Evans JT, Smith EG, Banerjee A, et al. Cluster of human tuberculosis caused by *Mycobacterium bovis*: evidence for person-to-person transmission in the UK. *Lancet* 2007;369:1270–6.
12 Maguire H, Brailsford S, Carless J, et al. Large outbreak of isoniazid-monoresistant tuberculosis in London, 1995 to 2006: case-control study and recommendations. *Euro Surveill* 2011;16. pii 19830.
13 Duthie S, Black C, Douglas G, et al. Tuberculosis outbreak associated with a mosque: challenges of large scale contact tracing. *Euro Surveill* 2008;13. pii: 19069.
14 Evans JT, Serafino Wani RL, Anderson L, et al. A geographically-restricted but prevalent *Mycobacterium tuberculosis* strain identified in the West Midlands Region of the UK between 1995 and 2008. *PLoS ONE* 2011;6:e17930.
15 Maguire H, Dale JW, McHugh TD, et al. Molecular epidemiology of tuberculosis in London 1995-7 showing low rate of active transmission. *Thorax* 2002;57:617–22.
16 Supply P, Allix C, Lesjean S, et al. Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of *Mycobacterium tuberculosis*. *J Clin Microbiol* 2006;44:4498–510.
17 Brown TJ, Nikolayevskyy VN, Drobniewski FA. Typing *Mycobacterium tuberculosis* using variable number tandem repeat analysis. *Methods Mol Biol* 2009;465:371–94.
18 Kik SV, Verver S, van Soolingen D, et al. Tuberculosis outbreaks predicted by characteristics of first patients in a DNA fingerprint cluster. *Am J Respir Crit Care Med* 2008;178:96–104.
19 Brown TJ, Evans JT, Sails A, et al. HPA *Mycobacterium tuberculosis* strain typing: a guide to data production and distribution. London: Health Protection Agency, 2012. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/433731/HPA_VNTR_280812_final.pdf
20 Brown T, Nikolayevskyy V, Velji P, et al. Associations between *Mycobacterium tuberculosis* strains and phenotypes. *Emerg Infect Dis* 2010;16:272–80.
21 Small PM, Hopewell PC, Singh SP, et al. The epidemiology of tuberculosis in San Francisco. A population-based study using conventional and molecular methods. *N Engl J Med* 1994;330:1703–9.
22 van Soolingen D, Borgdorff MW, de Haas PE, et al. Molecular epidemiology of tuberculosis in the Netherlands: a nationwide study from 1993 through 1997. *J Infect Dis* 1999;180:726–36.
23 Ellis BA, Crawford JT, Braden CR, et al. Molecular epidemiology of tuberculosis in a sentinel surveillance population. *Emerg Infect Dis* 2002;8:1197–209.
24 Love J, Sonnenberg P, Glynn JR, et al. Molecular epidemiology of tuberculosis in England, 1998. *Int J Tuberc Lung Dis* 2009;13:201–7.
25 Behr MA, Warren SA, Salamon H, et al. Transmission of *Mycobacterium tuberculosis* from patients smear-negative for acid-fast bacilli. *Lancet* 1999;353:444–9.
26 Neely F, Maguire H, Le Brun F, et al. High rate of transmission among contacts in large London outbreak of isoniazid mono-resistant tuberculosis. *J Public Health (Oxf)* 2010;32:44–51.
27 Lambregts-van Weezenbeek CS, Sebek MM, van Gerven PJ, et al. Tuberculosis contact investigation and DNA fingerprint surveillance in The Netherlands: 6 years' experience with nation-wide cluster feedback and cluster monitoring. *Int J Tuberc Lung Dis* 2003;7(Suppl 3):S463–70.
28 Dahle UR, Eldholm V, Winje BA, et al. Impact of immigration on the molecular epidemiology of *Mycobacterium tuberculosis* in a low-incidence country. *Am J Respir Crit Care Med* 2007;176:930–5.
29 Anderson C, Story A, Brown T, et al. Tuberculosis in UK prisoners: a challenge for control. *J Epidemiol Community Health* 2010;64:373–6.
30 Story A, Murad S, Roberts W, et al. Tuberculosis in London: the importance of homelessness, problem drug use and prison. *Thorax* 2007;62:667–71.
31 Jit M, Stagg HR, Aldridge RW, et al. Dedicated outreach service for hard to reach patients with tuberculosis in London: observational study and economic evaluation. *BMJ* 2011;343:d5376.
32 National Institute for Health and Clinical Excellence. *Identifying and managing tuberculosis among hard-to-reach groups*. NICE, 2012.
33 Public Health England. *Collaborative Tuberculosis Strategy for England—2015 to 2020*. Public Health England, 2015. https://www.gov.uk/government/publications/collaborative-tuberculosis-strategy-for-england

THORAX

Recent TB transmission, clustering and predictors of large clusters in London, 2010–2012: results from first 3 years of universal MIRU-VNTR strain typing

Esther L Hamblion, Arnaud Le Menach, Laura F Anderson, Maeve K Lalor, Tim Brown, Ibrahim Abubakar, Charlotte Anderson, Helen Maguire and Sarah R Anderson

Updated information and services can be found at:

http://thorax.bmj.com/content/71/8/749

These include:

References	This article cites 26 articles, 9 of which you can access for free at: **http://thorax.bmj.com/content/71/8/749#BIBL**
Open Access	This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/
Email alerting service	Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections	Articles on similar topics can be found in the following collections Open access (225) Health education (1221)

Notes