



Using tuberculosis patient characteristics to predict future cases with matching genotype results

J. E. Oeltmann, E. S. Click, P. K. Moonan

<http://dx.doi.org/10.5588/pha.13.0098>

Setting: United States.

Background: It is unknown whether tuberculosis (TB) case or patient characteristics can predict the likelihood of future related TB cases.

Objective: To estimate the likelihood for future related cases, i.e., cases with matching TB genotypes within the same county diagnosed within the 2 years following the year of reporting of each included case.

Design: We considered all TB cases with genotyping results reported in the United States during 2004–2010. Predictive scores were calculated based on patient characteristics by dividing the number of patients who were not the last case in a county-level TB genotype cluster by the total number of patients.

Results: Overall, there was a 30.8% chance that a future related case would be detected during the 2 years following the report year of any given case. Future related cases were detected in 34.7% of instances following the diagnosis of smear-positive cases, 51.9% of instances following the diagnosis of a homeless patient and 45.2% of instances following the diagnosis of a patient who reported substance abuse. Predictive scores ranged by race (White 13.9%, Native Hawaiian 43.8%) and age group (≥ 65 years 13.1%, 0–4 years 43%), and were higher for US-born patients.

Conclusions: Behavioral and sociodemographic factors can help predict the likelihood of future related cases and can be used to prioritize contact investigations.

Tuberculosis (TB) genotyping can help identify TB cases that may be attributable to recent transmission. When *Mycobacterium tuberculosis* isolates cluster by genotype (i.e., ≥ 2 patients with *M. tuberculosis* isolates with matching genotype results), they often represent recent transmission, particularly when the patients reside in similar communities or regions.¹ Combining routinely collected surveillance and genotyping data can identify ongoing transmission within a community.² Using a national genotyping surveillance database, we characterized TB patients who represented the first or index case diagnosed within a localized TB genotype cluster of cases.

The objective of this analysis was to identify patient characteristics associated with index case status, and to assess whether the patient characteristics of known cases can be used to predict the likelihood of future genotype-clustered cases in the same county, hereafter called future related cases. This information could help those involved in local TB control to prioritize and enhance contact investigations of cases most

likely to be the first of one or more future related cases.

METHODS

Data for this analysis came from incident TB cases reported to the National Tuberculosis Surveillance System (NTSS) of the Centers for Disease Control and Prevention (CDC) from 50 states and the District of Columbia. The NTSS includes demographic, behavioral and clinical information for each reported patient.³ We obtained genotyping data from the CDC's National Tuberculosis Genotyping Service (NTGS), consisting of spacer oligonucleotide typing and 12-locus mycobacterial interspersed repetitive unit-variable number of tandem repeats (MIRU-VNTR) typing data.⁴ NTSS and NTGS are linked at the patient level. We considered all TB cases reported between 1 January 2004 and 31 December 2010 with genotyping results and linked NTSS data. From this data set, two subsets of data were created: one to assess factors associated with index-case status, and another to calculate predictive scores based on reported case characteristics. Further details regarding data restrictions for each of the two subsets of data are provided below.

Factors associated with index patients

Genotype clusters were defined as ≥ 2 cases with matching genotyping (spoligotype and 12-locus MIRU-VNTR) results reported within the same county. Index cases were the first case reported in a cluster, based on treatment start date. Secondary cases were the remaining clustered cases that occurred after the treatment start date of the index case. Index cases were included only if they began treatment on or after 1 January 2006, and no cases with a matching genotype had been observed in the county in the previous 2 years. This restriction was created to reduce misclassification of secondary cases as index cases. Unique cases were cases that were not clustered. Unique cases began treatment between 1 January 2006 and 31 December 2008 and were excluded if one or more cases with matching genotype results were observed during the 2 years before (2004–2005) or after (2009–2010) the 3-year inclusion period. These exclusion criteria were created to reduce misclassification of secondary or index cases as unique cases.

We calculated the distribution of demographic, clinical and behavioral characteristics among index cases and unique cases and assessed differences using Pearson's χ^2 test. Specific characteristics (independent variables) examined were sex, age, race/ethnicity, country of origin, substance abuse, history of home-

AFFILIATIONS

Division of Tuberculosis Elimination, US Centers for Disease Control and Prevention, National Centers for HIV, Viral Hepatitis, STD, and Tuberculosis Prevention, Atlanta, Georgia, USA

CORRESPONDENCE

John E Oeltmann
Division of Tuberculosis Elimination
Centers for Disease Control and Prevention
1600 Clifton Rd, MS E-10
Atlanta, GA 30333, USA
Tel: (+1) 404 639 5334
e-mail: jeo3@cdc.gov

ACKNOWLEDGEMENTS

The authors are grateful to National TB Genotyping Service scientists at the California Department of Health Services, Sacramento, CA, and the Michigan Department of Community Health, Lansing, MI, USA, the local and state TB program and laboratory personnel who participate in surveillance and genotyping activities, and the Surveillance, Epidemiology, and Outbreak Investigations Branch, Division of TB Elimination, The National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention, Centers for Disease Control and Prevention, Atlanta, GA, USA. The authors thank T Navin and J Grant for their helpful review.

KEY WORDS

tuberculosis; genotype; transmission

Received 7 November 2013
Accepted 1 February 2014

PHA 2014; 4(1): 47–52
© 2014 The Union

lessness, incarceration at the time of diagnosis, sputum smear status, site of disease and human immunodeficiency virus (HIV) status. In the NTSS, excess alcohol use, injection drug use, and non-injection drug use are defined as positive if reported within the year before TB diagnosis. We defined substance abuse as having any one of these three behaviors reported.

We used multivariate logistic regression to assess the association between being an index case, relative to unique cases (dependent variable), and the abovementioned independent factors. A significance level of 0.01 was used for all statistical tests.

Predictive scores of patient characteristics

To estimate the ability with which individual patient characteristics can predict one or more future related cases, we calculated predictive scores for patient and clinical characteristics for both US-born and foreign-born patients. For any given case, a future related case was a case with matching genotyping results diagnosed in the same county during the 2 years following the report year of each included case. All cases in a cluster were categorized as either the last case diagnosed in a cluster or not the last case diagnosed in a cluster. Cases reported in 2006, 2007 and 2008 were included in this analysis. All cases reported in 2006 were defined as either the last case in a cluster or not, based on data available from 2006, 2007 and 2008. Cases reported in 2007 were similarly defined based on data from 2007, 2008 and 2009, and cases reported in 2008 were defined based on data from 2008, 2009 and 2010. Therefore, for each case included, there were at least 2 years of follow-up time during which another case with matching genotype results could have been diagnosed within the same county. This length of time is referred to as the follow-up period. A predictive score was calculated for each sociodemographic, behavioral and clinical characteristic. Predictive scores for all cases and for cases with individual characteristics were estimated by dividing the number of cases that were not the last case in a county-based cluster by the total number of cases (unique cases + clustered cases). This proportion represents the percentage of instances at which, given a specific characteristic, a future related case will be diagnosed.

Approval by an institutional review board was not required because data were collected and analyzed for this project as part of routine TB surveillance; the project was therefore not considered research involving human subjects.

RESULTS

Factors associated with index patients

Of 51 527 reported culture-positive cases with genotyping data, 2918 were eligible index cases and 13 612 were eligible unique cases; 16 485 were considered secondary cases and were not included in this portion of the analysis. A total of 18 512 were excluded from this analysis to avoid potential misclassification as a unique ($n = 16 432$) or index case ($n = 2080$) (Table 1).

Distributions of all patient and disease characteristics differed significantly for index and unique cases (Table 2). A larger proportion of index cases were male, born in the United States, were substance users, were homeless and had a history of incarceration. With the exception of Asians and Whites, the proportions of all other racial/ethnic groups were larger for index cases.

Crude and adjusted odds ratios (ORs) and 99% confidence levels (CIs) for the assessment of the relation between patient and disease characteristics and being an index case relative to a unique case are presented in Table 3. After controlling for disease characteristics known to be related to TB transmissibility (sputum smear

TABLE 1 Number of cases eligible and ineligible for analysis of factors associated with index case status by year

	2004	2005	2006	2007	2008	2009	2010
Index	2080*				2918		
Unique	7654*			13 612		8778*	

*Ineligible cases.

status, site of disease), the odds of being an index case were significantly increased for all age groups relative to the oldest patients, race/ethnic minority patients, patients born in the United States and those who were substance users.

Predictive scores of patient characteristics

During the period from 1 January 2006 to 31 December 2008, 8364 (30.8%) of 27 146 total eligible cases were not the last clustered patients reported in the same county during the follow-up period. This suggests that there is roughly a 30% chance a related case will be diagnosed in the same county following the diagnosis of any given case. Predictive scores for specific patient characteristics are provided in Table 4. In general, predictive scores were higher among US-born patients, males and minorities. Among US-born patients, predictive scores for age groups exhibited the widest range of values (from 49.8% for the youngest patients to 18.1% for the oldest). Following the diagnosis of TB among the youngest patient (age 0–4 years), a future related TB case was thus identified 49.8% of the time, while following the diagnosis of TB among the oldest (≥ 65 years) patients, a future related case was only identified 18.1% of the time. Among foreign-born patients, diagnosis of TB among the youngest patients was the least predictive of a future related case (predictive score = 11.8%).

Future related cases were detected 51.9% of the time following the diagnosis of a homeless patient, 45.2% of the time following the diagnosis of a patient who reported substance use, and 38.8% of the time following diagnosis of patients with TB within a correctional facility (Table 3). As expected, sputum smear-positive disease had a higher predictive score than sputum smear-negative disease (34.7 vs. 28.9%) and pulmonary disease had a higher predictive score than extra-pulmonary disease (33.0% vs. 22.1%).

DISCUSSION

The odds of being an index case relative to a unique case were influenced by patient characteristics. As suspected, characteristics associated with index cases also had the highest predictive scores for diagnosis of future related cases. We believe that characteristics associated with index cases, when noted, can serve as a predictor of future related cases. Prioritizing patients with these characteristics for contact investigations could result in a more efficient use of resources. These future related cases, in theory, should be identified during contact investigations around the reported case. Ideally, these individuals would be identified early enough to prevent progression to active TB disease. This information can therefore be used to prioritize and enhance contact investigations around those patients most likely to be followed by later diagnosed cases with matching genotypes and within the same county. Three characteristics stood out as being the most predictive of future related cases: age ≤ 4 years, homelessness and substance abuse. Following the diagnosis of a patient with any one of these characteristics, there was roughly a 50% chance that a future related case would be diagnosed.

An index case is not necessarily the source of a chain of TB

TABLE 2 Demographic, social and clinical characteristics of index and unique TB cases ($n = 16\,530$)

Characteristic	Index patients* ($n = 2918$) n (%)	Unique patients† ($n = 13\,612$) n (%)	Pearson's χ^2 P
Sex			
Male	1924 (65.9)	8349 (61.3)	<0.001
Female	992 (34.0)	5262 (38.7)	
Unknown‡	2 (0.1)	1 (0.0)	
Age, years			
0–4	64 (2.2)	85 (0.6)	<0.001
5–14	27 (0.9)	103 (0.8)	
15–24	396 (13.6)	1652 (12.1)	
25–44	1023 (35.1)	4645 (34.1)	
45–64	889 (30.8)	3762 (27.6)	
≥65	509 (17.4)	3365 (24.7)	
Race/ethnicity			
Asian	644 (22.1)	3940 (28.9)	<0.001
Black	735 (25.2)	2942 (21.6)	
Hispanic	1016 (34.8)	3836 (28.2)	
Native American	37 (1.3)	142 (1.0)	
Native Hawaiian	42 (1.4)	54 (0.4)	
White	427 (14.6)	2644 (19.4)	
Other/unknown	17 (0.6)	54 (0.4)	
Country of birth			
US-born	1189 (40.8)	4478 (32.9)	<0.001
Foreign-born	1718 (58.9)	9100 (66.9)	
Unknown	11 (0.4)	34 (0.3)	
Substance abuse			
Yes	621 (21.3)	1799 (13.2)	<0.001
No	2205 (75.6)	11 543 (84.8)	
Unknown	92 (3.2)	270 (2.0)	
Homelessness			
Yes	209 (7.2)	550 (4.1)	<0.001
No	2666 (91.4)	12 979 (95.4)	
Unknown	43 (1.5)	83 (0.6)	
TB diagnosis at correctional facility			
Yes	147 (5.0)	442 (3.3)	<0.001
No	2760 (94.6)	13 162 (96.7)	
Unknown	11 (0.4)	442 (3.3)	
HIV status			
Positive	198 (6.8)	742 (5.5)	<0.001
Negative	1587 (54.4)	8159 (59.9)	
Unknown	1133 (38.8)	4711 (34.6)	
Sputum smear result			
Positive	1654 (56.7)	6210 (45.6)	<0.001
Negative	850 (29.1)	5190 (38.1)	
Unknown	414 (14.2)	2212 (16.3)	
Disease site			
Pulmonary only	2303 (78.9)	9609 (70.6)	<0.001
Extra-pulmonary only	342 (11.7)	2593 (19.1)	
Both	269 (9.2)	1394 (10.2)	
Unknown	4 (0.2)	16 (0.1)	

*The first patient reported within a county-based genotype cluster.

†Patients who were not part of a county-based genotype cluster.

‡Not included in χ^2 calculations.

TB = tuberculosis; HIV = human immunodeficiency virus.

TABLE 3 Crude and adjusted associations between demographic, social and clinical characteristics and index case status

Characteristic	Crude OR	99%CI	Adjusted OR	99%CI
Sex*				
Female	Referent		Referent	
Male	1.2	1.1–1.4	1.1	0.9–1.2
Age, years				
0–4	5.0	3.2–7.8	3.8	2.4–6.2
5–15	1.7	1.0–3.1	1.7	0.9–3.1
15–24	1.6	1.3–1.9	1.6	1.3–2.0
25–44	1.5	1.3–1.7	1.5	1.2–1.7
45–64	1.6	1.4–1.8	1.4	1.2–1.7
≥65	Referent		Referent	
Race/ethnicity				
Asian	1.0	0.9–1.2	1.6	1.3–2.0
Black	1.6	1.3–1.8	1.7	1.4–2.0
Hispanic	1.6	1.4–1.9	2.0	1.7–2.5
Native American	1.6	1.0–2.6	1.4	0.9–2.4
Native Hawaiian	4.8	2.8–8.3	6.7	3.8–11.9
White	Referent		Referent	
Unknown/multiple	1.9	0.9–4.0	2.6	1.2–5.6
Country of birth				
Foreign-born	Referent		Referent	
US-born	1.7	0.7–4.2	1.7	1.4–1.9
Unknown	1.4	1.3–1.6	1.1	0.4–3.0
Substance use				
No	Referent		Referent	
Yes	1.8	1.6–2.1	1.4	1.2–1.6
Unknown	1.8	1.3–2.5	1.2	0.8–1.7
Homelessness				
No	Referent		Referent	
Yes	1.8	1.5–2.3	1.3	0.9–1.6
Unknown	2.5	1.5–4.1	1.8	1.0–3.2
TB diagnosis at a correctional facility*				
No	Referent		Referent	
Yes	1.6	1.2–2.0	1.1	0.8–1.4
HIV status				
Negative	Referent		Referent	
Positive	1.4	1.1–1.7	1.2	0.9–1.6
Unknown	1.2	1.1–1.4	1.4	1.3–1.6
Sputum smear result				
Negative	Referent		Referent	
Positive	1.6	1.4–1.8	1.4	1.2–1.6
Unknown	1.1	0.9–1.4	1.2	0.9–1.5
Disease site*				
Extra-pulmonary only	Referent		Referent	
Pulmonary only	1.8	1.6–2.1	1.6	1.3–1.9
Both	1.5	1.2–1.8	1.3	0.9–1.6

*Not included are cases with unknown sex ($n = 3$), TB diagnosed at a correctional facility ($n = 19$) and disease site ($n = 20$).

OR = odds ratio; CI = confidence interval; TB = tuberculosis; HIV = human immunodeficiency virus.

TABLE 4 Predictive scores of demographic, social and clinical characteristics to predict future related TB cases

Characteristic	Total (n = 27 146)		US-born (n = 13 081)		Foreign-born (n = 14 065)	
	Predictive score* %	99%CI	Predictive score* %	99%CI	Predictive score* %	99%CI
Overall	30.8	30.1–31.5	34.9	33.8–36.0	27.0	26.1–28.0
Sex						
Male	32.4	31.5–33.3	36.6	35.2–37.9	28.3	27.0–29.5
Female	28.2	27.0–29.3	31.7	30.0–33.5	25.2	23.7–26.7
Age, years						
0–4	47.4	39.5–55.4	49.8	41.6–58.0	11.8	0.6–44.1
5–14	32.1	24.5–40.4	41.7	30.7–53.2	20.4	11.4–32.0
15–24	31.0	29.0–33.0	38.3	35.0–41.7	26.2	23.8–28.7
25–44	30.9	29.7–32.2	38.3	36.3–40.3	25.7	24.2–27.2
45–64	35.9	34.5–37.3	40.0	38.1–41.9	30.9	28.9–32.9
≥65	21.7	20.2–23.2	18.1	16.2–20.1	25.7	23.5–28.0
Race/ethnicity						
Asian	29.1	27.8–30.4	28.0	25.4–30.6	29.4	27.9–31.0
Black	37.4	35.8–38.9	44.4	42.5–46.2	19.6	17.2–22.1
Hispanic	30.5	29.2–31.8	35.2	32.9–37.7	28.1	26.5–29.7
Native American	43.5	36.2–51.0	44.0	36.6–51.6	16.7	0.1–74.6
White	22.8	21.1–24.6	24.4	22.5–26.4	14.9	11.6–18.7
Other	41.8	33.7–50.2	46.2	32.7–60.1	39.2	29.2–49.9
Substance use						
Yes	45.2	43.3–47.1	48.3	46.0–50.6	37.7	34.3–41.2
No	27.8	27.0–28.6	30.4	29.2–31.7	25.7	24.7–26.7
Homelessness						
Yes	51.9	48.7–55.1	57.1	53.3–60.7	36.7	30.6–43.1
No	29.5	28.8–30.2	32.6	31.5–33.8	26.8	25.8–27.7
TB diagnosis at a correctional facility						
Yes	38.8	34.6–43.0	46.7	41.1–52.5	27.8	22.0–34.1
No	30.5	29.8–31.3	34.4	33.3–35.5	27.0	26.0–28.0
HIV status						
Positive	34.8	31.7–37.9	44.6	40.5–48.7	19.7	15.8–24.1
Negative	28.3	27.3–29.2	33.9	32.6–35.3	22.4	21.2–23.7
Unknown	34.3	33.1–35.6	34.4	32.5–36.4	34.2	32.6–35.9
Sputum smear result						
Positive	34.7	33.6–35.8	38.8	37.2–40.3	30.5	29.1–32.0
Negative	28.9	27.7–30.1	32.8	31.1–34.7	25.7	24.2–27.2
Unknown	21.3	19.5–23.1	24.8	22.2–27.6	17.8	15.5–20.3
Disease site						
Pulmonary only	33.0	32.1–33.8	37.0	35.7–38.2	29.0	27.8–30.1
Extra-pulmonary only	22.1	20.5–23.7	24.1	21.6–26.8	20.6	18.6–22.7
Both	29.6	27.3–31.8	34.3	31.0–37.8	25.4	22.5–28.4
Unknown	10.5	0.6–40.4	11.1	0.1–58.5	10.0	0.1–54.4

*Number of cases (with a given characteristic) who were not the last case in a county-based cluster divided by the total number of cases (with a given characteristic).
TB = tuberculosis; CI = confidence interval; HIV = human immunodeficiency virus.

transmission. Using TB surveillance data, which does not collect information about interpersonal connections between patients, we were not able to determine the true source patient who infected the later diagnosed patients. Index patients are simply the first reported among a series of cases, defined here as having matching genotype and county of residence. For example, the odds of being an index case was highest for children aged 0–4 years, as young children with TB tend to be sentinels for recent transmission rather than the actual source.⁵ However, when a childhood case (aged <4 years) was diagnosed in a foreign-born patient, there was only an 11.8% chance of detecting a future related case. We believe this is because foreign-born children with TB who are living in the United States most likely acquired their

infection before immigration to the United States, and if the source of their infection is still overseas it would not be captured in the NTSS.

While index cases may not always be source cases, our findings are similar to those from studies that assessed risk factors for genotype clustering,⁶ recent transmission⁷ and involvement in outbreaks,⁸ reinforcing the notion that index cases characterized by these risk factors may often be the source case for future cases. We found only one study that assessed characteristics associated with the generation of secondary cases using epidemiologic but not genotyping data. Rodrigo et al. reported similar results associated with substance use, suggesting that substance users are more likely to be associated with secondary cases.⁹

Previous studies have described the characteristics of the first two patients within a genotype cluster and reported that the infectiousness of the first two patients was associated with greater cluster growth in New York City,¹⁰ and that <3 months between diagnosis of the first two patients, one or both patients being young (age <35 years), both patients residing in an urban setting and both patients being from sub-Saharan Africa were associated with greater cluster growth in The Netherlands.¹¹ While our results share some common findings, they are difficult to compare to these previous studies due to the differing case definitions and methodology. We choose to focus on the first case, rather than the first two cases, within a cluster because we aimed to provide those involved in TB control with a simple, rapid method to evaluate the likelihood of secondary cases. We are of the opinion that at the time of diagnosis of a patient characterized by any of the significant risk factors identified from this analysis, TB control authorities can expect an increased likelihood of future related cases and can immediately plan to increase case-finding efforts.

The reasons why some TB strains circulate locally, while others do not, are not completely understood. However, some if not all reasons might be explained by the characteristics of the persons infected rather than the strain itself. A noteworthy finding was that having a history of substance use or homelessness had a significantly higher predictive score than did sputum smear-positive or pulmonary disease, the two primary indicators used to initiate a contact investigation.⁵ We therefore believe that some social and behavioral factors are in fact more important than disease characteristics (smear and chest radiograph results) for determining the likelihood of the spread of TB disease, and should be considered when prioritizing cases for contact investigation.

Lack of access to routine health care, social gathering behaviors and inadequate TB control measures among substance users may explain the increased likelihood of secondary cases following the diagnosis of TB among substance users or the homeless.^{12,13} Similarly, lack of access to routine care may also explain the increased likelihood of secondary cases following the diagnosis of TB among minorities,¹⁴ as the longer a TB case remains undiagnosed, the greater the chances of continued transmission. It is also possible that some cases of TB among those without routine access to care were never diagnosed and reported. If so, our estimates of the association between substance use, homelessness, minority race and clustering would be biased toward the null and the true association between these factors and clustering may therefore be even stronger.

Some additional limitations should be considered. First, as submission of isolates for genotyping is not mandatory, all reported culture-positive cases are not represented. Second, county-based genotype clustering serves only as a proxy for recent TB transmission in the absence of epidemiologic links between patients. Third, cluster size is time-dependent. Had we chosen to use a longer follow-up period to define clusters, cluster size could only get larger, but we do not believe this would significantly al-

ter our main results regarding factors related to index cases or the relative differences in predictive scores associated with specific characteristics. Furthermore, predictive scores would increase if applied to a community with a higher prevalence of TB transmission. Finally, as behavioral data were self-reported, the validity of these data is not known.

CONCLUSION

We were able to use demographic, clinical, and behavioral data to identify cases that were associated with an increased risk of future related cases within a county. After controlling for HIV, sputum smear status and site of disease, we found increased odds of related secondary cases associated with young age, minority race/ethnicity, birth in the United States, homelessness and substance use. An enhanced contact investigation following the TB diagnosis in patients with any of these characteristics may help to prevent future cases.

References

- Barnes P F, Cave M D. Molecular epidemiology of tuberculosis. *N Engl J Med* 2003; 349: 1149–1156.
- Moonan P K, Ghosh S, Oeltmann J E, Kammerer J S, Cowan L S, Navin T R. Estimating recent transmission of *M. tuberculosis* in the United States based on genotyping and geospatial scanning. *Emerg Infect Dis* 2012; 18: 458–465.
- Centers for Disease Control and Prevention. Reported tuberculosis in the United States, 2012. Atlanta, GA, USA: US Department of Health and Human Services, CDC, 2013.
- Centers for Disease Control and Prevention. Notice to readers: new CDC program for rapid genotyping of *Mycobacterium tuberculosis* isolates. *MMWR Morb Mortal Wkly Rep* 2005; 54: 47.
- National Tuberculosis Controllers Association; Centers for Disease Control and Prevention (CDC). Guidelines for the investigation of contacts of persons with infectious tuberculosis. Recommendations from the National Tuberculosis Controllers Association and CDC. *MMWR Recomm Rep* 2005; 54(RR-15): 1–47.
- Fok A, Numata Y, Schulzer M, FitzGerald M J. Risk factors for clustering of tuberculosis cases: a systematic review of population-based molecular epidemiology studies. *Int J Tuberc Lung Dis* 2008; 12: 480–492.
- Nava-Aguilera E, Andersson N, Harris E, et al. Risk factors associated with recent transmission of tuberculosis: systematic review and meta-analysis. *Int J Tuberc Lung Dis* 2009; 13: 17–26.
- Mitruka K, Oeltmann J E, Ijaz K, Haddad M B. Tuberculosis outbreak investigations in the United States, 2002–2008. *Emerg Infect Dis* 2011; 17: 425–31.
- Rodrigo T, Caylà J A, García de Olalla P, et al. Characteristics of tuberculosis patients who generate secondary cases. *Int J Tuberc Lung Dis* 1997; 1: 352–357.
- Driver C R, Macaraig M, McElroy P D, et al. Which patients' factors predict the rate of growth of *Mycobacterium tuberculosis* clusters in an urban community? *Am J Epidemiol* 2006; 164: 21–31.
- Kik S V, 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.
- Lofy K H, McElroy P D, Lake L, et al. Outbreak of tuberculosis in a homeless population involving multiple sites of transmission. *Int J Tuberc Lung Dis* 2006; 10: 68.
- Oeltmann J E, Kammerer J S, Pevzner E S, Moonan P K. Tuberculosis and substance abuse in the United States, 1997–2006. *Arch Intern Med* 2009; 169: 189–197.
- Smedley B D, Stith A Y, Nelson A R, eds. Unequal treatment: confronting racial and ethnic disparities in health care. Washington, DC, USA: National Academies Press, 2003.

Contexte : Etats-Unis.

Cadre : On ne sait pas si les caractéristiques d'un cas de tuberculose (TB) ou du patient permettent de prévoir la probabilité de contamination dans l'avenir.

Objectif : Estimer la probabilité de cas présentant un génotype similaire dans le même conté et dans une période de 2 ans suivant le cas index.

Schéma : Nous avons étudié tous les cas de TB avec génotypage déclarés aux Etats-Unis entre 2004 et 2010. Les scores prédictifs ont été calculés en fonction des caractéristiques du patient en divisant le nombre de patients qui n'étaient pas le dernier cas d'un groupement de génotypes au niveau d'un conté par le nombre total de patients.

Résultats : Le risque global de nouveau cas lié à un autre cas était de 30,8% pendant les 2 années suivant l'année de déclaration de tout nouveau cas. Ces contaminations ont été détectées dans 34,7% des circonstances après diagnostic d'un cas à frottis positif, 51,9%, après diagnostic d'un patient sans domicile fixe et 45,2%, après diagnostic d'un patient toxicomane. Les scores prédictifs variaient en fonction de l'ethnie (Blancs 13,9% ; Amérindiens/Hawaiens 43,8%), l'âge (>65 ans 13,1% ; 0-4 ans 43%) et étaient plus élevé chez les patients nés aux Etats-Unis.

Conclusion : Les facteurs comportementaux et socio-démographiques peuvent contribuer à prévoir la probabilité d'infection de cas dans le futur et peuvent servir à prioriser les recherches de sujets contacts.

Marco de referencia: En los Estados Unidos de América se desconoce si las características clínicas de un caso de tuberculosis (TB) o las características del paciente permiten pronosticar la probabilidad de aparición de futuros casos de TB relacionados.

Objetivo: Evaluar la probabilidad de aparición en el futuro de casos relacionados — es decir, casos con genotipos equivalentes, diagnosticados en el mismo condado, durante los 2 primeros años después del año de notificación de cada caso incluido.

Métodos: Se consideraron en el estudio todos los casos de TB notificados en los Estados Unidos que contaban con resultados de genotipificación entre el 2004 y el 2010. Se calcularon las puntuaciones pronósticas en función de las características del paciente, al dividir el número de pacientes que no fueron el último caso de un conglomerado genotípico de TB a escala del condado, por el número total de pacientes.

Resultados: En general, se observó una probabilidad de 30,8% de

aparición de un caso relacionado, durante los 2 años que siguieron al año de notificación de cualquier caso dado. Se observaron casos relacionados en el 34,7% de las veces después del diagnóstico de casos con baciloscopia positiva; el 51,9% de las veces después del diagnóstico de una persona sin domicilio; y en el 45,2% de las veces tras el diagnóstico de un paciente que refería consumo de drogas. Las puntuaciones pronósticas oscilaron, con respecto a la etnia, entre 13,9% en la etnia blanca y 43,8% en los nativos de Hawái); con respecto al grupo de edad, entre 13,1% a partir de los 65 años y 43% en el grupo entre 0 años y 4 años; y el índice pronóstico fue más alto en los pacientes nacidos en los Estados Unidos.

Conclusión: Los factores comportamentales y sociodemográficos contribuyen a predecir la probabilidad de aparición de casos futuros relacionados con un caso de TB y se pueden utilizar con el propósito de priorizar las investigaciones de contactos.