TITLE

Effect of clinical and socio-economic factors on the grouping in clusters of Mycobacterium *tuberculosis* clinical isolates in Elche (Spain).

AUTHORS

M. Ruiz(1), J.F. Navarro(2), J.C. Rodríguez(1), J.A. Larrosa(3), G. Royo(1).

(1) S. Microbiología. (2) S. Medicina Preventiva. Hospital General Universitario de Elche. Universidad Miguel Hernández. Elche (Alicante).(3)Universidad de Alicante.

ADDRESS

Montserrat Ruiz García

S. Microbiología. Hospital General Universitario de Elche

Camí de L'Almazara, 11

03203 Elche (Alicante). Spain

Fax: 34-966679108

e-mail: micro_elx@gva.es

RUNNING HEAD

RFLP-IS*6110* of *M. tuberculosis* in Elche.

KEY WORDS: *Mycobacterium tuberculosis*, RFLP, IS*6110*, epidemiology

SUMMARY

The study of tuberculosis by characterization of the patterns obtained using S*6110*-RFLP is a very useful tool when attempting to clarify the epidemiology of this disease. We studied the association that exists between the grouping in clusters of isolates and certain clinical, epidemiological and socio-economic characteristics in the Elche health district (Spain).

On multivariate analysis we found that the independent variables associated with grouping in clusters are: young patients (1-25 years of age), patients with high percentage of infection in the first circle of contacts (51%-100% of infection), patients who live in Elche or Santa Pola, and samples obtained using bronchoscopy.

This study enables us to identify the factors associated with the transmission of tuberculosis in our setting and provides data that contribute to a better understanding of the epidemiology of this disease and improved systems of control.

INTRODUCTION

We recently published a study in which we describe the molecular epidemiology of tuberculosis in Elche (Spain) applying the *IS6110*-RFLP technique to the clinical isolates in this region. An overall aggregation of 52.40% was obtained(1).

Our aim is to study the influence of the patients' clinical, epidemiological and socio-economic characteristics on the grouping in clusters detected using this methodology.

METHODS

Patients studied: 147 patients diagnosed between 1993 and 1999 in the Elche health district (Spain), who account for 59.34% of the patients diagnosed microbiologically as having tuberculosis.(1)

Patients' data: Name, age, sex, nationality, profession, attendance to a community centre, employment status (employed/unemployed), own telephone, place of residence, district/street, date disease was diagnosed, HIV serology, parenteral drug user (PDU), alcoholism, smoker, homeless, existence of other causes of immunodepression, pulmonary cavitation, previous history of tuberculosis, current treatment, **localization** of the disease, hospital admission, type of sample, quantification of the staining, quantification of the culture and susceptibility to first line anti-tuberculostatic drugs. All the data were obtained prospectively.

IS*6110*-RFLP technique: This was used following the standard protocol described by van Embden et al.(2)

Statistical analysis: We used SPSS. Univariate analysis of the relationship between grouping in clusters and the different co-variables was done calculating the raw Odds Ratio (OR) of prevalence with the corresponding 95% confidence intervals, and statistical significance was confirmed using Pearson's Chi-squared test or Fisher's exact test. In order to introduce the variables in the multivariate analysis, we used an unconditioned logistic regression procedure based on the likelihood ratio statistic and entered the variables manually. The categorical variables were factorized (dummy variables), and the missing values considered to be another category. Possible confounding variables (sex and age) or factors traditionally related to clusters in other studies (HIV, PDU) were forcibly included in the final model, even though they showed no statistical significance. Since the sample was small, we included in the final model predictive variables with a substantial OR whose statistical significance (p value) was ≥0.10.

RESULTS

1.- Patients' clinical-epidemiological characteristics

1.a.- Patients infected with HIV, PDU and the homeless: The 21 patients with HIV antibodies had a higher percentage of aggregation (71.43%) (p=0.017).(1) In patients who were PDU there was 72% aggregation and in the homeless 71.40%. 80% of HIV-infected patients were also PDU, and just under half were also homeless. 86% of homeless patients were HIV-infected and PDU.

1.b.- Immunodepressed patients: There was 66.7% aggregation in the 12 patients who were immunodepressed due to causes other than HIV. The causes considered were: treatment with corticoides, diabetes mellitus, tumors and chronic diseases.

1.c.- Patients who were alcoholic and smokers: The grouping in alcoholic patients is greater than in the general population (63.60%). 88.63% of the alcoholics were also smokers and the percentage of aggregation was 64%. The percentage of aggregation in smokers was 60%.

1.d.- Patients with cavitating tuberculosis: The percentage of aggregation in the 26 patients with cavitating tuberculosis was 57.70%; 88.5% of the patients with cavitating tuberculosis also had positive direct staining and the percentage of aggregation in this group was 56.50%.

1.e.- Patients with previous tuberculosis: The percentage of aggregation in the 14 patients with previous tuberculosis was 60%.

1.f.- Patients admitted to hospital: The percentage of aggregation in the 81 patients admitted to hospital on starting treatment was 59.3% as compared with 43.9 % in those who did not require admission.

1.g.- Type of sample studied: The percentage of aggregation was 80% in patients in whom the sample was obtained using fibrobronchoscopy (Table 3).

1.h.- Aggregation according to the contact study data: The data obtained are shown in table 1. We may point out that the percentage of aggregation is greater when the percentage of those infected in the first circle increases, when infection extends to the third circle studied, and when a greater number of families are infected.

If we consider the existence of prior contact with tuberculous patients, the percentage of aggregation is similar irrespective of whether there was prior contact or not, but it is slightly higher (58.3%) if we consider the most recent contacts (those since

1993).

2.- Socio-economic characteristics

2.a.- Professional activity: The percentage of aggregation is greater in persons who are employed (55.2%) than in those who are not (50.9%). Moreover, the percentage of aggregation is greater if we consider those whose professions brings them into contact with more people (66.7%).

2.b.- Residence in Elche: The percentage of aggregation in patients living in Elche (55.5%) is greater than in other patients who live elsewhere (inhabitants of nearby towns or tourists). Aggregation varies depending on the district in Elche and is greater in the poorest district (Los Palmerales). Aggregation is also greater in Santa Pola.

3.- Temporal distribution

The percentage of aggregation changes as a function of the period of time studied and is greater when a longer period of time is considered (Table 2).

4.- Results of the univariate analysis

On univariate analysis we find the following more frequently associated with clusters: younger patients, men, PDU, HIV+, the homeless, alcoholics, smokers, the immunodepressed, patients with pulmonary disease, those admitted to hospital, those diagnosed on bronchoscopy (Table 3), cases in which a greater proportion of people were infected in the first circle of contacts, cases which required the study of the greatest number of families and investigation up to the third circle (Table 4), professions related to the shoe industry and those which in theory involve greater contact with people, residents of the Palmerales district in Elche and residents of Santa Pola (Table 5).

5.- Results of the multivariate analysis

The following continue to be independent predictive factors of the association in clusters (Table 6) on multivariate analysis: An age under 26 years (RR=4.6, CI=1.3-15.9) and of 26-50 years (RR= 2.2, CI=0.77-6.3) compared with an age over 50 years; a percentage of people infected in the first circle of the contact study of 25%-50% (RR=2.94, CI=0.68-12.7) and above 50% (RR=3.83, CI=1,17-12.6) compared to the studies with no person infected in the first circle; residence in the town of Elche (RR=22.5, CI=1.8-282) compared with residence in other towns; residence in Santa Pola (RR=8.7, CI=0.75-101); and microbiological diagnosis made on bronchoscopy (RR=22.8, CI=2.9-176).

When the above variables are included, the model correctly predicts the appearance of 79% of the cases that are clusters and 69% of those that are not grouped according to RFLP.

DISCUSSION

Aggregation in clusters of cases of tuberculosis in a population is conditioned by many factors such as the thoroughness of the microbiological diagnosis, the time covered by the study, the contagiousness of the index case and of the secondary cases, the susceptibility of the contacts, the efficacy in the identification and prophylaxis of the contacts and certain socio-economic factors (overcrowded living conditions, attendance to a community centre, lack of health care, etc). Therefore, it is essential to know the characteristics of a population in order to correctly interpret the results obtained using RFLP.(3)

Elche is situated in the southeast of Spain, on the Mediterranean coast, and has a population of approximately 200,000. Around Elche there are many small towns, in particular Santa Pola with a population of 20,000. The main industry in the region is the footwear industry, and after 1960 this attracted a significant number of immigrant workers from other parts of Spain.(4). Nowadays, its demographic growth is stable, except in summer when thousands of tourists arrive each year.

The percentage of clusters in our study (52.4%) is relatively high compared with other studies(1), and this is mainly due to a longer follow-up time, since most secondary cases occur in the first 2-3 years' follow-up.

Although recent transmission of tuberculosis has been related to the association in *clusters* using IS*6110-RFLP*(5), this has recently begun to be questioned and transmission of the disease has been associated with certain characteristics of the population studied such as mobility, mixture and concentration (3). In addition, the existence of endemic strains that are preferentially transmitted in a population has been suggested (6). A study carried out in Arkansas(3) on a stable rural population with a low prevalence of HIV infection showed that there was no epidemiological relationship in 60% of the patients that appeared grouped in clusters using RFLP and, furthermore, there was evidence of previous tuberculosis in a third of these. This suggested that some of the clusters were due to simultaneous reactivations of strains that were endemic to the region.(3) In our study, we also observed a high percentage of association between patients with a previous history of tuberculosis and no known epidemiological relationship between them. However, it should be borne in mind that epidemiological studies using classical methods identify less than half of the groups of clusters found using RFLP (3, 5, 7), since it has been suggested that very brief contacts (8, 9) or contagion arising from contacts of which the patients were unaware (3) play an

important role in the transmission of tuberculosis.

There is a greater proportion of associated strains in under-developed countries with a high incidence of tuberculosis such as central Africa, for example, where there is more recent transmission(10). On the other hand, in countries with few cases of tuberculosis such as The Netherlands, there is more heterogeneity, that is, more reactivations.(11) In Norway, the country with the lowest incidence of tuberculosis in the world, only 19.7% of the patients were grouped in clusters(12) In developed countries with a high rate of incidence of the disease such as Japan, with 41.9 cases per 100,000 habitants in 1990, the isolates have a high degree of polymorphism. This may indicate that although the incidence is high, it could be mainly due to reactivations.(13) The percentage of aggregation obtained in our population is a reflection of a region in which tuberculosis is endemic, with a relatively high percentage of HIV+ patients and where there is also a significant circulation of strains. This percentage varies depending on the socio-economic characteristics of each area.

Various studies point out that there is greater aggregation in HIV+ patients (14,15,16); however, other studies indicate the opposite.(5,9) This discrepancy may be related to the existence of nosocomial transmission, the patients' socio-economic situation and the effect of how quickly the diagnosis is made and the efficacy of the treatment on the transmission and progression of the disease in these patients. In our study, we found a greater percentage of aggregation in patients who were infected with HIV, and this is statistically significant in the univariate analysis, although it does not appear as an independent risk factor in the multivariate analysis. It should be borne in mind that circulating strains may exist in the area, mainly transmitted between HIV+ patients during contacts involving risk practices (17).

In our study, the younger the age group, the greater the risk of clusters, which is

in agreement with the literature (5, 14, 16, 18), and this is associated with factors typical of young people such as going around in groups and frequenting bars (3, 7, 9). The greater susceptibility of such persons to new infections also plays a part (smaller percentage of reactors and of those vaccinated with BCG). (19,20)

The percentage of aggregation in alcoholic patients (63.60%) is greater than in the general population and this also coincides with the results obtained in other regions (7,18). In the Hamburg study, alcoholism is the variable that most strongly predicts forming part of a cluster (7). There is also more aggregation in patients who are smokers (60%), as occurs in other regions.(18)

Patients with cavitating tuberculosis are also more likely to be grouped in clusters, as some authors have published.(21)

The percentage of aggregation in patients with previous tuberculosis in our region is somewhat higher than in the general population(5,14,16), as opposed to the findings of other studies.(21) This could be interpreted as a case of simultaneous reactivations of endemic strains that are very prevalent in our community(3,6,22,23).

We found a clear relation between residence in the most urban area and aggregation in clusters(9,16,18). In addition, the greatest percentage of aggregation (66%) was found in a very poor district of Elche (Los Palmerales), which is the neighbourhood with the worst socio-economic indicators in the whole town (4), where drug dealing and consumption is most frequent and where there is the highest percentage of PDU, HIV and ex–prison inmates. There is also greater aggregation in Santa Pola, where there is a significant PDU and HIV+ population. The patients in Santa Pola are grouped with residents of Elche, probably due to the proximity of the two towns, which are only 15 km apart. Similar results were obtained in deprived areas in other cities such as New York(16) and San Francisco.(14)

To date, the strong association between samples obtained by fibrobronchoscopy has not been reported in the literature, except for contaminations associated with bronchoscopy (24). Our patients who underwent bronchoscopy form part of clusters with patients diagnosed by other procedures and in different years, and so transmission by bronchoscopy may be ruled out. We consider that the small number of patients who underwent this procedure, the existence of patients co-infected with HIV and the difficulty involved in diagnosing cases secondary to the original case may be some of the factors contributing to this phenomenon. These cases correspond to recent infections with a short clinical history, and so could be considered evidence of the improvement made in early diagnosis of this disease in Spain. Bearing in mind the small number of patients, further studies should be done on larger numbers of patients to confirm this association.

The aggregation in patients that required admission to hospital when newly diagnosed, is greater. This is associated with a high percentage of bacilli-bearing patients, since no nosocomial transmission was detected (25).

The association between grouping in clusters and greater proportion of patients infected in the contact study could be explained by the greater capacity and duration of contagion in some cases. This result might be due to the existence of supertransmissors in our area.(1,14)

Therefore, the data obtained using this technique, together with the patients' clinical and epidemiological data enable a better understanding of the epidemiology and transmission dynamics of tuberculosis to be obtained. Analysis of the results enables us to identify the factors associated with recent transmission in our setting, the possible existence of endemic strains in our community, and the true value of this methodology.

REFERENCES

(1) Ruiz M, Rodríguez JC, Navarro JF, Samper S, Martín C, Royo G. Molecular epidemiology of tuberculosis in Elche, Spain: a 7-year study. *J Med Microbiol* 2002; **51**: 273-277.

(2) van Embden JDA, Cave D, Cawford JT, et al. Strain Identification of *Mycobacterium tuberculosis* by DNA Fingerprinting: Recommendations for a standarized methodology. *J Clin Microbiol* 1993; 31; **2**:406-409.

(3) Braden CR, Templeton GL, Cave MD, et al. Interpretation of Restriction Fragment Length Polymorphism analysis of *Mycobacterium tuberculosis* isolates from a state with a large rural population. *J Infect Dis* 1997; **175**:1446-52.

(4) Larrosa JA. Atlas demográfico y social de la ciudad de Elche. Publicaciones de la Universidad de Alicante San Vicente del Raspeig. Alicante. 2000.

(5) van Soolingen D, Borgdorffff MW, de Haas PEW, Sebeck MMGG, Veen J, Dessens M, Krenmer, van Embden JDA. Molecular epidemiology of tuberculosis in the Netherlands: A nationwide study from 1993 through 1997. *J Infect Dis* 1999; **180**:726-36.

(6) Hermans PW, M, van Soolingen D, Dale JW, et al. Insertion element IS*986* from *Mycobacterium tuberculosis*: a useful tool for diagnosis and epidemiology of tuberculosis. *J Clin Microbiol* 1990; **28**:2051-2058.

(7) Diel R, Schneider S, Meywald-Walter K, Ruf CM, Rüsch-Gerdes S, Nieman S. Epidemiology of tuberculosis in Hamburg, Germany: long term population-based analysis applying classical and molecular epidemiological techniques. *J Clin Microbiol* 2002; **40**:532-539.

(8) Bauer J, Kok-Jensen A, Faurschou P, Thuesen J, Taudorf E, Andersen AB. A prospective evaluation of the clinical value of tuberculosis isolates in Denmark. *Int J*

Tuberc Lung Dis 2000; 4:295-299.

(9) Weis SE, Pogoda JM, Yang Z, et al. Transmission dynamics of tuberculosis in Tarrant County, Texas. *Am J Resp Crit Care Med* 2002; **166**:36-42.

(10) Kallenius G, Koivula T, Ghebremichael S, et al. Evolution and clonal traits of *Mycobacterium tuberculosis* in Guinea-Bissau. *J Clin Microbiol* 1999; **37**:3872-3878.

(11) van Soolingen D, Hermans PWM, de Haas PEW, Soll DR, van Embden JDA. The occurrence and stability of insertion sequences in *Mycobacterium tuberculosis* complex strains; evaluation of insertion sequence-dependent DNA polymorphism as a tool in the epidemiology of tuberculosis. *J Clin Microbiol* 1991; **29**:2578-2586.

(12) Dahle UR, Sandven P, Heldal E, Caugant DA. Molecular epidemiology of *Mycobacterium tuberculosis* in Norway. *J Clin Microbiol* 2001;**39**: 1802-1807.

(13) Takahashi M, Kazumi Y, Fukasawa Y, et al. restriction fragment length polymorphism analysis of epidemiologically related *Mycobacterium tuberculosis*. *Microbiol Immunol* 1993; **37**:289-294.

(14) Small P, Hopewell PC, Singh SP, et al. The epidemiology of tuberculosis in San Francisco: a population based study using conventional and molecular methods. *New Engl J Med* 1994; **330**:1703-1709.

(15) Genewein A, Telenti A, Bernasconi C, et al. Molecular approach to identifying route of transmission of tuberculosis in the community. *Lancet* 1993; **342**:841-844.

(16) Alland D, Kalkut GE, Moss AR, et al. Transmission of tuberculosis in New York city. An analysis by DNA fingerprinting and conventional epidemiological methods. *N Engl J Med* 1994; **330**:1710-1716.

(17) Camarena JJ, Artero A, Nogueira JM, Navarro JC, Olmos A, Blanquer R. Tuberculosis en pacientes con SIDA: aportación del análisis de los fragmentos de longitud polimórfica aislados de *Mycobacterium tuberculosis*. *Med Clin (Barc)* 1998,

111:721-724.

(18) van Deutekom H, Gerritsen JJJ, van Soolingen D, van Ameijden EJC, van Embden JDA, Coutinho RA. A molecular epidemiological approach to studying the transmission of tuberculosis in Amsterdam. *Clin Infect Dis* 1997; **25**:1071-1077.

(19) Vynnycky E, Fine PE. The natural history of tuberculosis: the implications of age-dependent risks of disease and the role of reinfection. *Epidemiol Infect* 1997; **119**:183-201.

(20) Vynnycky E, Nagelkerke N, Borgdorff MW, van Soolingen D, van Embden JD, Fine PE. The effect of age and study duration on the relationship between 'clustering' of DNA fingerprint patterns and the proportion of tuberculosis disease attributable to recent transmission. *Epidemiol Infect* 2001; **126**:43-62.

(21) Samper S, Iglesias MJ, Rabanaque MJ, et al. The molecular epidemiology of tuberculosis in Zaragoza, Spain: a retrospective epidemiological study in 1993. *Int J Tuberc Lung Dis* 1998; **2**:281-287.

(22) Kimerling ME, Benjamin WH, Lok KH, Curtis G, Dunlap NE. Restriction fragment length polymorphism screening of *Mycobacterium tuberculosis* isolates: population surveillance for targeting disease transmission in a community. *Int J Tuberc Lung Dis* 1998; **2**: 655-662.

(23) Dobbs KG, Lok KH, Bruce F, Mulcahy D, Benjamin WH, Dunlap NE. Value of *Mycobacterium tuberculosis* fingerprinting as a tool in a rural state surveillance program. *Chest* 2001; **120**:1877-1882.

(24) Gubler JG, Salfinger M, Gravenenitz A. Pseudoepidemic of nontuberculous mycobacteria due to contaminated bronchoscope cleaning machine. Report of an outbreak and review of the literature. *Chest* 1992; 1**01**:1245-1256.

(25) Roselle GA, Danko LH, Kralovic SM, Simbartl LA, Kizer KW. Tuberculosis in the

veterans healthcare system: a six-year review and evaluation of programme effectiveness. *Epidemiol Infect*. 2000; **125**:315-23.

Table 1: Percentage of aggregation according to contact study data

	% involvement 1st circle		**Circle involved**		**Nº families involved**		**Previous contact**	
	0%	25-50%	1º	3º	1	>5	No	Yes
Nº of patients	45	16	40	20	32	31	60	58
Percentage	31	11	27.4	13.7	22	21.2	41.1	39.7
% aggregation	44.4	65.9	50	65	43.8	67.7	48.3	49.2

Table 2: Percentage of aggregation as a function of the period of time studied

	1994-1995	**1994-1996**	**1994-1997**	**1994-1997**	**1994-1999**
% of isolates	15.8	29.6	56.6	85.5	100
% aggregation	25	26.6	30.2	36.9	52.40

Table 3: Relation between the probability that the cases belong to a cluster and health care variables

Variable/Category	Frequencies	% of clusters	Relative risk	Confidence intervals (95%)
HIV+				
No	51 (35.5%)	49	1	---
Yes	21 (15%)	71.4	2.6	0.87-7.76
Not known	72 (49.3%)	48.6	0.98	0.48-2.01
PDU				
No	114 (78.1%)	57	1	---
Yes	25 (17.1%)	72	2.57	0.94-6.63
Not known	7 (4.8%)	14.3	0.17	0.02-1.43
Alcoholism				
No	95 (65.1%)	49.5	1	---
Yes	44 (30.1%)	63.6	1.79	0.86-3.62
Not known	7 (4.8%)	14.3	0.17	0.02-1.44
Smoker				
No	73 (50%)	47.9	1	---
Yes	65 (44.5%)	60	1.63	0.83-3.20
Not known	8 (5.5%)	25	0.36	0.07-1.91
Immunodepression				
No	124 (84.9%)	52.4	1	---
Yes	12 (8.2%)	66.7	1.81	0.52-6.34
Not known	10 (6.8%)	30	0.39	0.10-1.57

Cavitation				
No	102 (69.7%)	52.9	1	---
Yes	26 (18%)	57.7	1.21	0.51-2.89
Noknown	18 (12.3%)	38.9	0.57	0.20-1.57
Previous TBC				
No	122 (83.6%)	52.5	1	---
Yes	14 (9.6%)	57.1	1.21	0.40-3.69
Not known	10 (6.8%)	40	0.60	0.16-2.25
Previous treatment				
No	128 (87.7%)	53.1	1	---
Yes	10 (6.8%)	60	1.32	0.36-4.61
Not known	8 (5.5%)	25	0.22	0.06-1.15
Present treatment				
Correctly administered	123 (84.2%)	52	---	---
Poorly administered	23 (15.8%)	52.2	1.01	0.41-2.45
Site				
Extrapulmonary	36 (24.7%)	41.7	1	---
Pulmonary	110 (75.3%)	55.5	1.75	0.81-3.7
Sample type				
Sputum	95 (65.1%)	53.7	1	---
Bronchoscope	10 (6.8%)	80	3.45	0.70-17.11
Others	41 (28.1%)	41.5	0.61	0.30-1.30

Hospital admission				
No	57 (39%)	43.9	1	---
Yes	81 (55.6%)	59.3	1.86	0.94-3.70
Not known	8 (5.4%)	37.5	0.77	0.17-3.52

Table 4: Relation between the probability that the cases belong to a cluster and variables obtained from the contact study. Univariate analysis

Variable/Category	**Frequencies**	**% of clusters**	**Relative risk**	**Confidence intervals (95%)**
% involved in 1st circle				
0%	45(31%)	44.4	1	---
1-25%	16(11%)	56.3	1.61	0.51-5.10
25-50%	16(11%)	65.9	1.61	0.51-5.10
51-100	41(28%)	39.3	2.41	1.00-5.77
Not known	28(19%)	52.1	0.81	0.31-2.11
Circle of contacts involved				
Up to 1st circle	40(27.4)	50	1	---
Up to 2nd circle	58(39.7)	55.2	1.23	0.55-2.76
Up to 3rd circle	20(13.7)	65	1.86	0.61-5.63
Not known	26(17.8)	38.5	0.65	0.24-1.72

Involvement up to the last circle				
No	69 (47.3%)	50.7	1	---
Yes	50 (34.2%)	60	1.46	0.70-3.04
Not known	27 (18.5%)	40.7	0.67	0.27-1.65
Nº of families involved				
1	32 (22%)	43.8	1	---
from 2 to 4	55 (37.7%)	54.5	1.54	0.64-3.71
> 5	31 (21.2%)	67.7	2.70	0.97-7.54
Not known	28 (19.1%)	39.3	0.82	0.30-2.33
Involvement outside the family				
No	97 (66.4%)	55.7	1	---
Yes	16 (10.9%)	56.3	1.02	0.35-2.97
Not known	33 (22.6%)	39.4	0.52	0.23-1.16
Contact with cases of TB				
No	60 (41.1%)	48.3	1	---
Yes after 1993	48 (32.9%)	58.3	1.48	0.70-3.22
Yes before 1993	10 (6.8%)	40	0.71	0.18-2.78
Not known	28 (19.2%)	53.6	1.23	0.50-3.03

Table 5: Relation between the probability that the cases belong to a cluster and socio-economic variables. Univariate analysis.

Variable/Category	Frequencies	% of clusters	Relative risk	Confidence intervals (95%)
Age				
≥51	50 (34.2%)	38	1	---
26-50	67 (45.9%)	56.7	2.14	1.01-4.52
1-25 years	27 (18.5%)	70.4	3.87	1.42-10.58
Not known	2 (1.4%)	0	0.03	0-∞
Sex				
Female	44 (30%)	47.7	1	---
Male	102 (70%)	53.9	1.28	0.63-2.63
Year of diagnosis				
1996-99	122 (83.6%)	51.6	1	---
1993-95	24 (16.4%)	54.2	1.11	0.46-2.63
Homeless				
No	126 (86.3%)	51.6	1	---
Yes	14 (9.6%)	71.4	2.35	0.70-7.88
Not known	6 (4.1%)	16.7	0.19	0.02-1.65
Attendance to a community centre				
No	112 (76.7%)	53.6	1	---
Yes	25 (17.1%)	52	0.94	0.40-2.24

Not known	9 (6.2%)	33.3	0.43	0.10-1.82
Profession				
No contact with people	65 (44.5%)	43.1	1	---
Contact	33 (22.6%)	66.7	2.64	1.10-6.33
Footwear industry	23 (15.7%)	65.2	2.47	0.92-6.65
Not known	25 (17.2%)	44	1.04	0.41-2.63
Employed				
No	53 (36.3%)	50.9	1	---
Yes	87 (59.6%)	55.2	1.19	0.60-2.35
Not known	6 (4.1%)	16.7	0.19	0.02-1.76
Residence				
Outside Elche	27 (18.5%)	37	1	---
Elche	119 (81.5%)	55.5	2.13	0.89-5.0
Area of residence in Elche				
Centre	28 (19.1%)	53.6	1	---
Altábix	14 (9.6%)	35.7	0.48	0.13-1.81
Palmerales	9 (6.2%)	66.7	1.77	0.36-8.35
Carrús-Toscar	34 (23.3%)	61.8	1.40	0.51-3.86
Pla-Sector V	27 (18.5%)	48.1	0.81	0.28-2.32
Other (Santa Pola and others)	4 (23.3%)	47.1	0.77	0.28-2.10

Table 6: Independent variables associated with belonging to a cluster. Multivariate analysis.

	Relative risk	Confidence interval	Significance (p)
Age			
≥51	1	---	0.05
26-50	2.20	0.77-6.23	0.14
1-25 years	4.56	1.31-15.90	0.01
% involvement in 1st circle			
0%			
1-25%	1	---	0.19
25-50%	1.90	0.44-8.18	0.39
51-100	2.94	0.68-12.73	0.14
	3.83	1.17-12.56	0.02
Residence			
Outside Elche	1	---	---
Elche	22.51	1.79-282.6	0.01
Area of residence in Elche			
Centre	1	---	0.16
Altábix	0.29	0.06-1.30	0.10
Palmerales	1.65	0.21-12.89	0.63
Carrús-Toscar	0.97	0.29-3.23	0.96
Pla-Sector V	0.59	0.162.08	0.41
Other (Santa Pola and others)	8.75	0.75-101.39	0.08

Sample type			
Sputum	1	---	0.11
Bronchoscope	22.78	2.93-176.95	0.003
Sex			
Female	1	---	---
Male	1.05	0.44-2.51	0.90
HIV+			
No	1	---	---
Yes	1.78	0.33-9.55	0.49
PDU			
No	1	---	0.21
Yes	2.12	0.42-10.75	0.36