

Is it better to be rich in a poor area or poor in a rich area? A multilevel analysis of a case–control study of social determinants of tuberculosis

Ricardo Arraes de Alencar Ximenes,^{1,2*} Maria de Fátima Pessoa Militão de Albuquerque,^{1,3} Wayner V Souza,³ Ulisses R Montarroyos,¹ George T N Diniz,³ Carlos F Luna³ and Laura C Rodrigues⁴

Accepted 6 May 2009

Background Tuberculosis is known to have socio-economic determinants at individual and at area levels, but it is not known whether they are independent, whether they interact and their relative contributions to the burden of tuberculosis.

Methods A case–control study was conducted in Recife, Brazil, to investigate individual and area social determinants of tuberculosis, to explore the relationship between determinants at the two levels and to calculate their relative contribution to the burden of tuberculosis. It included 1452 cases of tuberculosis diagnosed by the tuberculosis services and 5808 controls selected at random from questionnaires completed for the demographic census. Exhaustive information on social factors was collected from cases, using the questionnaire used in the census. Socio-economic information for areas was downloaded from the census. Multilevel logistic regression investigated individual and area effects.

Results There was a marked and independent influence of social variables on the risk of tuberculosis, both at individual and area levels. At individual level, being aged ≥ 20 , being male, being illiterate, not working in the previous 7 days and possessing few goods, all increased the risk of tuberculosis. At area level, living in an area with many illiterate people and where few households own a computer also increased this risk; individual and area levels did not appear to interact. Twice as many cases were attributable to social variables at individual level than at area level.

¹ Departamento de Medicina Tropical, Universidade Federal de Pernambuco, Recife, Pernambuco, Brazil.

² Departamento de Clínica Médica, Universidade de Pernambuco, Recife, Pernambuco, Brazil.

³ Núcleo de Saúde Coletiva, Centro de Pesquisa Ageu Magalhães, FioCruz, Recife, Pernambuco, Brazil.

⁴ Department of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK.

* Corresponding author. Universidade Federal de Pernambuco-Medicina Tropical, Av. Professor Moraes Rego, s/n, Recife, Pernambuco 50670420, Brazil.

E-mail: ricardo.ximenes@pq.cnpq.br

Conclusions Although individual characteristics are the main contributor to the risk of tuberculosis, contextual characteristics make a substantial independent contribution.

Keywords Tuberculosis, social determinants, multilevel, Brazil, poverty

Introduction and background

The marked decrease in mortality from tuberculosis in England before the advent of drug treatment was an early indication of the effect of living conditions on tuberculosis.^{1–3} This effect has been shown often since and again, recently, after the re-emergence of tuberculosis in the late 1980s in both developing countries^{4,5} and developed countries.^{6–11} The association between poverty—at an individual level—and tuberculosis is plausible, and likely to be mediated by an increase in risk of infection in those living in crowded accommodation in areas of high incidence, and by increased risk of progression to disease in those with low resistance and reduced immunity because of under-nutrition or other socially determined factors.¹²

In a classic paper, Rose¹³ pointed out the importance of distinguishing between two kinds of etiological questions: what are the ‘cause of cases’ (factors that increase the individual risk) and the ‘cause of the incidence in populations’ (what variations in exposures between populations explain the variations in incidence between populations). For tuberculosis there is evidence that deprivation is both a cause of cases and a cause of incidence in populations. Evidence for the latter comes from traditional ecological analysis^{7,8,10} and from studies using geographical information systems.^{9,14–16} In developed countries, there is evidence that areas with higher incidence have worse housing conditions, crowding and higher population density,^{6,8,9,15–19} higher unemployment,^{6,7,9,15,19} higher proportion of immigrants from high-incidence areas,^{7–9,18,20} lower household income,^{9,21} higher inequality²² and also higher composite indexes of deprivation.^{6,8,10,21} Results from developing countries are similar, with particular emphasis on low income, education and unemployment.^{5,14,15,23}

Diez Roux²⁴ defended the importance of using a multilevel approach in the study of infectious diseases because it permits the investigation not only of the effects of individual factors (reflecting biology and life style) and area factors (reflecting socio-economic processes in the population), but also whether they act independently and whether they interact. Because this approach incorporates explicitly the social dimension it can more easily meet the objective of the ‘ecological epidemiology’ as proposed by Susser^{24–26} to identify how the context influences the health of individuals and groups. Finally, we share the values expressed by the World Health Organization (WHO) commission on social determinants: of health: a

commitment to the value of equity and the use of evidence as a basis for understanding and action.²⁷

The incidence of tuberculosis in Brazil (based on notified cases) has been relatively stable in the last decade, with roughly 50 per 100 000 inhabitants per year. In Recife, where this study was conducted, in 2005, the incidence was 136 per 100 000 inhabitants.²⁸ The National Program for Tuberculosis Control (PNCT) has standard procedures for investigation, and diagnosis in the whole country and the standardized treatment regimen is only delivered after the case is notified to the Surveillance System for Infectious Diseases (SINAN). Access to diagnosis, treatment and follow-up is free. As in many large cities, in Recife, poverty clusters in neighbourhoods, with the poorest areas having fewer basic services (water, sanitation, garbage collection) and the highest rates of population growth.²⁹

To further our understanding of social determinants of tuberculosis, and to inform policy decisions, it is helpful to know the relationship between the influence on tuberculosis of area and individual level factors, and also their relative contributions; in other words, what contributes more to tuberculosis: being poor or living in a poor neighbourhood. In the study reported here, we sought to answer these questions for the city of Recife, Brazil, using a case-control study with individual- and area-level information analyzed using a multilevel model.

Methods

The objectives of this study were to investigate the increase in risk of tuberculosis associated with social characteristics at the level of the individual and at the level of the area of residence; to explore whether these effects are independent, and whether they interact; and to estimate the relative contribution of each level to the burden of tuberculosis.

This is a case-control study with socio-economic information on individuals and on the area in which they live.

Study site and period was Recife, a city in the Northeast of Brazil with about 1.4 million inhabitants. Participants were recruited from May 2001 to July 2003.

Cases were subjects aged ≥ 7 years participating in a cohort study of newly diagnosed cases of pulmonary tuberculosis conducted to investigate predictors of successful tuberculosis treatment. Methods are

reported elsewhere,³⁰ but in short, patients diagnosed by the tuberculosis control programme were invited to participate in the study soon after diagnosis by the tuberculosis control programme personnel. About half of all eligible cases were invited to participate. Invitation was determined by availability of the health personnel in the unit of treatment (health centres of different levels of complexity) to interview patients and not by patient characteristics. This reflected internal organization of the service and was not related to the social characteristics of the area, as measured by quartiles of area social characteristics (data not shown). Those accepting, signed an informed consent form and were interviewed at the health unit by trained assistant nurses using a standard questionnaire, which included demographic information (like age and sex), address, and the socio-economic data used in this analysis. To investigate questions related to socio-economic factors, the questionnaire used a set of just under 50 questions from the census. These covered composition of the family, history of migration, years of schooling, type of work, whether at work, income, characteristics of the household and ownership of goods.

Controls were a random sample of subjects from the 2001 census database, aged ≥ 7 years and residents in the city of Recife. The selection was done allocating a number to all potential controls and using a random procedure in STATA. They were sampled from individuals who had been selected during the census to complete an in-depth form, which is applied to one in every 10 households (systematic sampling), using traditional census methodology.³¹ Controls were not interviewed by the study: we used their anonymous information on socio-economic factors, age, sex and census area of residence given to us in electronic format by the census bureaux.

The smallest unit from which census data is available is the census tract. Information on individuals linked to census tracts are not released by the Census Bureaux for confidentiality reasons. Census tracts are collated into census areas ('area ponderal'), defined by the Census Bureaux by selecting adjacent census tracts of similar population characteristics and infrastructure. Recife has 53 census areas, so a census area has, on average, approximately 26 000 inhabitants. Although we had detailed addresses for cases and could link them to census tracts, we were not given address or census tract data for controls to protect their anonymity, so the lowest census unit we could link cases and controls to was census area. Information about the census area of each case and each control was downloaded from the publicly available census database.³² All cases and controls therefore had the same information for the same variables at area and individual level, although, at the individual level, information was collected for cases by the study and for controls during the census.

Data were not available from the census on human immunodeficiency virus (HIV) infection. The prevalence of HIV in Recife is low ($< 1\%$ ³³), and among the cases, also likely to be low (8% of those tested³⁴).

Analysis

A database was constructed with individual- and area-level variables for cases and controls. We were aware that the study included only about half of all cases diagnosed with tuberculosis in Recife in residents during the study period, whereas controls were a random sample of all residents in the city. Because we had information on the number of cases notified to the Tuberculosis Control Program by area, we were able to calculate the degree of under-ascertainment of cases into the study by census area. We adjusted for the variation in completeness of ascertainment of cases into the study by giving weights to cases and controls in the analysis. All controls were given a weight of one and cases were given different weights according to their area of residence: for each area, the weight was the inverse of the ratio of notified cases to recruited cases in the area. The degree of under-ascertainment did not vary with wealth of the area.

The following individual-level variables were studied: sex, age, migration (always lived in Recife vs has ever lived out of Recife), illiteracy (not being able to read or write), whether the subject worked in the week preceding the interview, in-house access to piped water supply, number of the following goods possessed: radio, refrigerator, video, washing machine, microwave, computer, TV set, car and air conditioning. Age was grouped as 7–19, 20–34, 35–49, 50–64 and ≥ 65 years. The number of goods was grouped as follows. First, possession of each of the different goods was counted as 1, independent of the number of each good possessed (e.g. owning two cars counted as owning one good); then, the number of goods possessed was summed and grouped into four categories: 0 to 1, 2 to 3, 4 to 6, and 7 or more goods.

The characteristics of the area of residence were allocated to cases and to controls and categorized in quartiles according to the frequency in controls. The following area variables were studied: mean income of the head of the families in the area, mean number of schooling years of the head of the families in the area, mean number of inhabitants per household in the area, percent of literate individuals in the area, percent of individuals in the area who work, percent of households in the area with each of the following goods: refrigerator, video, washing machine, microwave oven, computer, air conditioning unit.

Bivariate analysis was performed for the variables at individual and area level, always controlling for age and sex. Variables that had shown a significant association ($P < 0.05$) in the bivariate model were introduced in the multivariable models. Multivariable analysis was performed initially separately for

Table 1 Composing the area-level variable computers and literacy

Computer and literacy	Proportion of literate residents	Proportion of households with a computer
High	Above median ^a	Top quartile ^b
Intermediate	Above median	Second and third quartile
Low	Below median	Lowest quartile

^aMedian value for literacy: 79% of the population is literate.

^bInter-quartile ranges: between 1st and 2nd quartile: 3.5% of households owns a computer; between 2nd and 3rd quartiles: 10.4% of households owns a computer; between 3rd and 4th quartiles: 19% owns a computer.

individual and area levels, weighting for the under-ascertainment of cases, and using a forward procedure. For the multivariable analysis of area variables, quartiles were grouped when appropriate [when their odds ratio (OR) of tuberculosis in adjacent quartiles was the same in the univariable analysis]. Because illiteracy and computer ownership were the best area predictors for tuberculosis in the multivariable analysis and because when the two variables were compared there were empty cells at the extreme values (no areas with a high proportion of computer owners had a high proportion of illiterate residents; and no areas with a high proportion of illiterate residents had a high proportion of computer owners), but there was some variation in the intermediate levels, we created a composite variable at area level combining the percent of the households in the area that had at least one computer and the percentage of residents in the area that were illiterate ('computing and literacy'). The final composite variable ('computer and literacy') had only three levels with roughly half the population in the lowest level, and a quarter each in the highest and intermediate levels. The composition of the variable is presented in Table 1.

After selecting the best model for individual variables and the best model for area variables, a final multilevel model was built with individuals as the first level and areas as the second level. This was done by backwards selection, starting with a model with all the significant variables in the area model and in the individual model and then withdrawing those that were no longer statistically significant. Interactions between the individual and the area-level variables were investigated. Population Attributable Fractions (PAFs)—the excess rate of disease in the total study population that is attributable to the exposure—for individual and area variables were calculated using the formula $PAF = [(\% \text{ of exposed cases}) \times (OR - 1)/OR]$ to estimate the contribution of each variable. PAFs were estimated for all individual levels and all area levels using an accepted procedure: it is worth noting that this is not the equivalent of adding the PAFs.³⁵ Gllamm was used to run the two-level random-intercept logit model using Stata 9.2. Ethical committee approval was granted to by the Universidade Federal de Pernambuco (UFPE) ethical committee.

Results

Included in the study were 1452 cases and 5808 controls, and 36% of cases and 54% of controls were female. The distribution of cases and controls according to the age groups defined in Methods were: 7–19 years: 12.4 and 28.0%; 20–34 years: 32.0 and 29.6%, 35–49 years: 32.7 and 22.6%; 50–64 years: 16.7 and 12.4%; and ≥ 65 years: 6.2 and 7.4%, respectively.

Bivariate analysis of individual- and area-level variables are presented in Tables 2 and 3. All individual variables tested were statistically significantly associated with tuberculosis at bivariate level. Age was a strong predictor of tuberculosis (Table 2). At multivariable level, history of migration was no longer significant. Access to piped water lost statistical significance when illiteracy was introduced in the model. Ownership of goods had the strongest association, and showed a clear reverse dose–response relationship to tuberculosis: using as baseline those with 7 goods or more, those with no goods or only 1 good had a 5-fold increase in risk; those with 2 to 3 goods, an increase of 3-fold and those with 4 to 6 goods, an increase of 1.7-fold. Illiteracy and not having worked in the previous week remained statistically significant (Table 4).

All area-level variables tested were statistically significantly associated with tuberculosis in the bivariable analysis (Table 3); however, in multivariable analyses, only the proportion of residents that are illiterate and the proportion of households that own a computer remained significant. As described in Methods, a composite variable 'computing and literacy' with three levels was created, and in the presence of this variable, no other area variable was statistically significantly associated with tuberculosis. Decreasing levels of 'computing and literacy' increase the risk of tuberculosis respectively by 1.5 times and doubled it (Table 4).

In the final model (Table 4), combining variables that were statistically significant at individual and area levels, all variables significant in the multivariable model at individual and area levels remained significant, with a small reduction on the strength of the association between tuberculosis and the extreme low levels of ownership of goods (at individual level) and between tuberculosis and computing and literacy at area level. This final model with individual and area levels was statistically significantly better than the model with

Table 2 Risk of tuberculosis associated with individual variables: bivariable, controlled for age and sex only

Variable	Percentage exposed cases	Percentage exposed controls	OR	95% CI	P
Individual level					
Age group ^a					
20–34 years	32.02	29.61	2.43	1.95–3.03	<0.001
35–49 years	32.71	22.57	3.29	2.69–4.02	<0.001
50–64 years	16.67	12.43	3.14	2.63–3.75	<0.001
≥65 years	6.20	7.40	1.99	1.61–2.45	<0.001
Sex^b					
Male	64.19	46.16	2.13	1.87–2.41	<0.001
Literacy^c					
Illiterate	18.32	10.42	1.88	1.56–2.28	<0.001
No computer ownership ^c	92.41	83.89	2.27	1.75–2.95	<0.001
Not working in the previous 7 days ^c	56.82	56.96	1.56	1.34–1.81	<0.001
Ownership of goods^c					
7+					
4–6	29.08	36.95	1.76	1.37–2.25	<0.001
2–3	53.12	39.44	3.03	2.34–3.91	<0.001
0–1	9.39	3.45	5.81	4.03–8.37	<0.001
Access to piped water ^c	19.23	12.40	1.69	1.27–2.25	<0.001
Migrant ^c	39.74	34.16	1.33	1.11–1.59	0.002

^aAdjusted for sex.^bAdjusted for age.^cAdjusted for age and sex.

CI = confidence interval.

Table 3 Risk of tuberculosis associated with area variables (bivariable)

Area level Inter-quartile ranges	Percentage exposed cases	Percentage exposed controls	OR	95% CI	P
Mean income of head of family in R\$^a					
2 between 366 and 618	28.86	26.08	1.00	0.73–1.37	0.995
3 between 619 and 873	25.55	24.78	0.83	0.59–1.17	0.283
4 >873	16.12	26.43	0.57	0.41–0.78	0.000
Mean years schooling of head of family					
2 between 5.1 and 5.6	27.82	22.81	0.86	0.63–1.17	0.337
3 between 5.7 and 6.3	19.70	24.76	0.63	0.48–0.82	0.001
4 >6.3	18.11	27.00	0.52	0.39–0.71	0.000
Mean number of habitants per household					
2 between 3.6 and 3.7	26.52	24.86	1.53	1.05–2.23	0.028
3 between 3.8 and 3.94	26.58	24.72	1.62	1.22–2.15	0.001
4 more than 3.9	31.54	25.50	1.83	1.32–2.53	0.000
Percentage of literate individuals					
2 between 78 and 79.3	31.75	23.23	0.94	0.69–1.30	0.732
3 between 79.4 and 83.5	19.08	25.64	0.66	0.49–0.88	0.005
4 >83.5	17.91	26.53	0.55	0.40–0.75	0.000

(continued)

Table 3 Continued

Area level Inter-quartile ranges	Percentage exposed cases	Percentage exposed controls	OR	95% CI	P
Percentage of individuals who work					
2 between 32.7 and 35.1	33.68	25.33	0.90	0.66–1.23	0.529
3 between 35.2 and 38.8	21.69	25.02	0.68	0.51–0.90	0.008
4 >38.8	23.31	25.62	0.47	0.34–0.65	0.000
Percentage of households with refrigerator					
2 between 86.9 and 89.5	30.99	24.02	0.85	0.62–1.15	0.294
3 between 89.6 and 92.6	18.60	24.40	0.59	0.46–0.75	0.000
4 >92.6	16.94	25.83	0.52	0.39–0.70	0.000
Percentage of households with video					
2 between 35.8 and 44.3	29.89	23.83	0.91	0.68–1.23	0.549
3 between 44.4 and 54.0	21.01	25.48	0.62	0.48–0.81	0.000
4 >54.0	14.88	23.74	0.52	0.38–0.72	0.000
Percentage of households with washing machine					
2 between 17.4 and 26.6	23.42	24.05	0.76	0.59–0.99	0.045
3 between 26.7 and 36.0	29.27	29.55	0.67	0.49–0.92	0.015
4 >36.0	13.57	21.81	0.49	0.35–0.69	0.000
Percentage of households with microwave					
2 between 6.9 and 13.0	27.48	23.71	0.95	0.70–1.30	0.748
3 between 13.1 and 20.9	26.31	26.22	0.79	0.57–1.10	0.158
4 >20.9	15.36	25.21	0.54	0.39–0.74	0.000
Percentage of households with computer					
2 between 3.7 and 10.3	27.69	25.62	0.89	0.65–1.22	0.469
3 between 10.4 and 19.5	30.92	26.93	0.83	0.62–1.11	0.210
4 >19.5	12.60	24.59	0.47	0.35–0.63	0.000
Average number of cars in the household					
2 between 0.18 and 0.30	30.30	27.88	0.76	0.57–1.00	0.057
3 between 0.31 and 0.48	21.83	24.00	0.58	0.46–0.75	0.000
4 more than 0.48	16.12	26.43	0.46	0.34–0.62	0.000
Average number of air-conditioning units in the household					
2 between 0.07 and 0.1	28.37	25.03	0.82	0.60–1.12	0.205
3 between 0.2 and 0.3	22.25	22.95	0.66	0.52–0.85	0.001
4 more than 0.3	16.87	28.19	0.47	0.35–0.65	0.000

^aUS \$1 roughly equivalent to R \$2.

individual data only ($P=0.011$). There were no statistically significant interactions between individual- and area-level variables.

Table 5 shows the proportion of cases attributable to each statistically significant individual- and area-level factor; the proportion attributable to all individual factors, to all area-level factors and to all factors identified. It is interesting to note that although extremes of poverty at individual level (having no goods or only 1 good) carry a very high increase in risk (OR=4.3) it only accounts for 7% of all cases, as very few controls (very few people in the population) are in

this category; and that possession of only 2 to 3 goods, although only increasing risk by 2.4-fold, accounts for 31% of all cases. The social variables in the model explained a very high proportion of all cases of tuberculosis, with individual-level variables explaining half the cases and area-level variables 29%. In Recife, 65% of all cases are explained by socio-economic variables, with one-and-a-half times more cases attributable to being poor than can be attributable to living in a poor neighbourhood. A proportion of cases remain unexplained, indicating that the study did not include all potential risks factors for tuberculosis.

Table 4 Multivariable analysis: risk of tuberculosis associated with individual variables, area variables and both

	OR (95% CI) individual factors	OR (95% CI) area factors	OR (95% CI) individual and area factors
Individual level			
Sex			
Male	2.20 (1.93–2.53)		2.21 (1.92–2.53)
Age group			
20–34 years	2.72 (2.07–3.59)		2.70 (2.06–3.55)
35–49 years	3.75 (2.90–4.85)		3.76 (2.91–4.86)
50–64 years	3.38 (2.73–4.20)		3.42 (2.74–4.25)
≥65 years	1.89 (1.46–2.45)		1.96 (1.52–2.54)
Illiterate	1.38 (1.15–1.66)		1.33 (1.11–1.61)
Not worked previous week	1.32 (1.13–1.53)		1.31 (1.13–1.52)
Possession of goods			
4–6	1.74 (1.36–2.23)		1.48 (1.16–1.90)
2–3	2.93 (2.24–3.84)		2.42 (1.86–3.15)
0–1	5.52 (3.57–7.64)		4.27 (2.88–6.34)
Area level			
Computers and literacy			
Intermediate		1.58 (1.25–2.00)	1.29 (1.00–1.67)
Low		2.12 (1.64–2.74)	1.59 (1.19–2.13)

Table 5 Proportion of cases attributable to individual- and area-level variables

	Percentage exposed cases	OR ^a	PAF	PAF
Individual level				
Ownership of goods				
4–6	0.29	1.5	0.10	
2–3	0.53	2.4	0.31	
0–1	0.09	4.3	0.07	
Literacy	0.18	1.3	0.04	
Worked in previous week	0.57	1.3	0.13	
All individual levels				0.50
Area level				
Computers and literacy				
Intermediate	0.24	1.3	0.06	
Low	0.63	1.6	0.24	
All area variables				0.29
All variables				0.65

^aORs used are adjusted by age and sex and all variables in the model.

Discussion

Our results show that there was a marked influence of social variables on risk of tuberculosis, both at individual and at area levels. We found a clear effect of

age on risk of tuberculosis similar to the literature. At individual level, being aged ≥20 years and possessing fewest goods increased markedly the risk of tuberculosis; male sex, being illiterate, not working in the previous 7 days also increased this risk. At area level, living in an area with many illiterate people and where few households own a computer, also increased the risk of tuberculosis; individual and area levels were independent and did not appear to interact. Many other area variables were significant by themselves (crowding, proportion working, low income) but lost significance when computer and literacy were included in the model. The statistically significant social variables in the model explained 65% of all cases of tuberculosis; with 50% of the cases being attributable to individual-level factors and 29% to area-level factors. The fact that the PAF for all identified factors is less than adding each PAF is not unexpected and results from the overlap of factors in the population.³⁵

The study does have some limitations. The area for which we had information was not very small (average 26 000 inhabitants), and it is possible that an area of this size still has some internal variation in wealth. This is not likely to be very marked as one of the criteria to aggregate census tracts into census areas is homogeneity (the other is contiguity). Any degree of internal heterogeneity would lead to an underestimation of the proportion of cases attributable to the area characteristics. Also, the study was potentially vulnerable to bias in relation to the area

variables as we did not include all diagnosed cases, with the proportion of diagnosed cases included in the study varying from area to area, whereas we did include as controls a representative random sample of all individuals in the areas. Fortunately we had data on notifications for each area, and therefore were able to control for this. We are confident that this removed the possibility of substantial bias. Social information was collected in controls as part of the census and in cases as part of the study, so that the interviewers and the setting were different for cases and controls (although the questionnaire used was the same). To reduce this bias, interviewers were rigorously trained and supervised. HIV infection was not measured but we believe that this has a limited impact given the small proportion of cases attributable to HIV in Recife. Finally, this is a study of diagnosed cases of tuberculosis, and all conclusions relate to those and not to undiagnosed cases in the population. Data from other studies provide indirect evidence that the fraction of cases which are never diagnosed is small.^{30,36} It is probably due to the fact that, in Brazil, access to treatment is easy as decentralization of the Tuberculosis Control Program is taking place [with the progressive transfer of activities from 'tuberculosis health units' to health units in the Family Health Program (FHP) and their teams]^{37,38} and investigation and treatment of patients are free and standardized for the whole country.³⁹

The study has many strengths. The controls were a random sample of the population of Recife, and control information was collected as part of the demographic census; detailed social information was collected at individual level using the questions developed for the census for both cases and controls. The sample size of approximately 1500 cases was large, giving the study substantial power, including power to detect interactions had an interaction been present. The multilevel analysis permitted the investigation in a single study of both individual and contextual variables (therefore exploring whether they are independent) and any interactions between them.

Many individual factors and all area factors were associated with tuberculosis in univariable analysis. This indicates how socially determined tuberculosis is, and that at least in this setting, deprived groups accumulate many aspects of deprivation. This may be more common in places like here where inequalities are marked. Because of this clustering we hesitate to propose biological pathways for the effect of these variables.

Of individual variables, illiteracy and ownership of goods were highly associated with tuberculosis and explained a good proportion of cases. These are simple measures of access to resources, and consistent with the literature. Not having worked in the previous week can reflect both increased risk associated with unemployment—consistent with much of the literature^{6,7,9,15,19}—but also can have resulted in part

from the individual not being well enough to work during the week before diagnosis. Unemployment is investigated using this question in the census in Brazil to address work both in the formal and informal sectors.

The composite measure we created, reflecting ownership of a computer and illiteracy, captured the effect of all other area variables, including proportion unemployed, income distribution and crowding. This again is an indication of clustering of aspects of deprivation in the same area; and of the centrality of education as it is related to all other aspects of poverty, at least in this setting.

This is the first study, to our knowledge, that investigates social determination of tuberculosis using a multilevel model (except for a study in South Africa of determination of reported levels of tuberculosis).⁴⁰ The independence of the effects of area-level factors and individual factors, shown for the first in this study, suggests a clear contextual risk: the risk of tuberculosis in an individual is increased by characteristics of the people around the individual; this is plausible as tuberculosis is an infectious disease, and infection must be acquired by direct contact with a case. It was surprising that the individual and the area factors did not interact: we expected individual wealth to reduce the effect of area of residence. Because of the very large sample size, the study would have had the power to find an interaction between individual and area factors if one was present, so this was not due to lack of power and must be real. Individual factors contributed to twice as many cases as area-level ones. We interpret this to reflect two things: first, that infection is acquired not only where people live, but also where they work or socialize; and secondly, that at least in this setting, poverty must be influencing the risk of progressing to disease more than the risk of acquiring infection. Since individual level variables explained almost twice as many cases of tuberculosis than area level variables, we conclude that based on the variables we had in the study, to avoid tuberculosis it is better to be rich in a poor area than to be poor in a rich area.

Finally, the fact that social determinants explained almost 65% of all cases is a clear confirmation of how essential poverty in its many aspects is to maintenance of tuberculosis, and reinforces the hope that social development and reduction in inequality may one day eliminate the disease. Although our study was in developing countries, we suggest this is true for developing and developed countries.

Conclusions

The study provides evidence of a substantial effect of both individual and area social characteristics on the risk of developing tuberculosis; these effects were independent and did not interact, and the

individual variables are responsible for roughly one-and-a-half as many cases.

Recommendation

Further research of determinants of tuberculosis should try to incorporate both contextual and individual factors. Targeting of control measures for tuberculosis—for example, active case finding—should consider not only individual characteristics but also characteristics of areas. Planning of public health interventions, for example, the decision of where to place a new health unit for tuberculosis diagnosis and treatment should be based not only on the incidence of diagnosed tuberculosis (since cases may be under diagnosed) but also in the area characteristics associated with risk of tuberculosis. Reducing poverty and inequality would have a major impact on tuberculosis.

Funding

(Brazilian) Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq; The British Council; REDE-TB do Brasil. RAAX was partially supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq (scholarship number 300917/2006-6).

Funding to pay the Open Access publication charges for this article was provided by Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq (scholarship number 300917/2006-6 to R.A.A.X.)

Acknowledgements

We would like to thank Katherine Fielding and Cibele Comini Cesar for advice.

Conflict of interest: None declared.

KEY MESSAGES

- The effect on tuberculosis of socio-economic characteristics at individual level and at area level are independent: the increase in risk of tuberculosis associated with living in a poor area is barely reduced when we take into account the individual characteristics; and the increase in risk of tuberculosis associated with being poor is barely reduced when we take into account effect the characteristics of the area of residence.
- The effect on tuberculosis of socio-economic characteristics at individual level and at area level do not interact: the increase in risk of tuberculosis associated with being poor is the same in people living in poor and rich areas; and the effect of living in a poor area is the same for rich and poor people.
- Individual level variables explained almost twice as many cases of tuberculosis than area level variables suggesting that, at least in this setting, to avoid tuberculosis it is better to be rich in a poor area than to be poor in a rich area.

References

- Berlinguer G. Medicina e Política. Hucitec, São Paulo: CEBES/HUCITEC 1978:52.
- Ruffino AN, Pereira JC. Mortalidade por tuberculose e condições de vida: O caso do Rio de Janeiro. *Saúde em Debate* 1981;**12**:27–34.
- Costa DC. Comentários sobre a tendência secular da tuberculose. *Cad Saúde Pública* 1988;**4**:398–406.
- Tupasi TE, Radhakrishna S, Quelapio MI *et al*. Tuberculosis in the urban poor settlements in the Philippines. *Int J Tuberc Lung Dis* 2000;**4**:4–11.
- Antunes JLF, Waldman EA. The impact of AIDS, immigration and housing overcrowding on tuberculosis deaths in São Paulo, Brazil, 1994-1998. *Soc Sci Med* 2001;**52**:1071–80.
- Spence DP, Hotchkiss J, Williams CS, Davies PD. Tuberculosis and poverty. *Br Med J* 1993;**307**:759–61.
- Mangtani P, Jolley DJ, Watson JM, Rodrigues LC. Socioeconomic deprivation and notification rates for tuberculosis in London during 1982–91. *Br Med J* 1995;**310**:963–66.
- Tocque K, Doherty MJ, Bellis MA, Spence DP, Williams CS, Davies PD. Tuberculosis notifications in England: the relative effects of deprivation and immigration. *Int J Tuberc Lung Dis* 1998;**2**:213–18.
- Kistemann T, Munzinger A, Dangendorf F. Spatial patterns of tuberculosis incidence in Cologne (Germany). *Soc Sci Med* 2002;**55**:7–19.
- Elender F, Bentham G, Langford I. Tuberculosis mortality in England and Wales during 1982-1992: its association with poverty, ethnicity and AIDS. *Soc Sci Med* 1998;**46**:673–81.
- Davies PDO. The world-wide increase in tuberculosis: how demographic changes, HIV infection and increasing numbers in poverty are increasing tuberculosis. *Ann Med* 2003;**35**:235–43.
- Waalder HT. Tuberculosis and poverty. *Int J Tuberc Lung Dis* 2002;**6**:745–46.
- Rose G. Sick individuals and sick populations. *Int J Epidemiol* 1985;**14**:32–38.
- Munch Z, Van Lill SW, Booyesen CN, Zietsman HL, Enarson DA, Beyers N. Tuberculosis transmission patterns in a high-incidence area: a spatial analysis. *Int J Tuberc Lung Dis* 2003;**7**:271–77.
- Souza WV, Carvalho MS, Albuquerque Mde F, Barcellos CC, Ximenes RA. Tuberculosis in intra-urban

- settings: a Bayesian approach. *Trop Med Int Health* 2007; **12**:323–30.
- 16 Souza WV, Ximenes R, Albuquerque MFM *et al.* The use of socioeconomic factors in mapping tuberculosis risk areas in a city of northeastern Brazil. *Rev Panam Salud Publica* 2000; **8**:403–10.
 - 17 Souza WV, Albuquerque MFM, Barcellos CC, Ximenes RAA, Carvalho MS. Tuberculose no Brasil: construção de um sistema de vigilância de base territorial. *Rev Saúde Pública* 2005; **39**:82–89.
 - 18 Leung CC, Yew WW, Tam CM *et al.* Socio economic factors and tuberculosis: a district based ecological analysis in Hong Kong. *Int J Tuberc Lung Dis* 2004; **8**:958–64.
 - 19 Ponticello A, Sturkenboom CJM, Simonetti A, Ortolani R, Mallerba M, Sanduzzi A. Deprivation, immigration and tuberculosis incidence in Naples, 1996–2000. *Eur J Infect Dis* 2005; **20**:729–32.
 - 20 Pavlovic M, Simic D, Krstic-Buric M *et al.* Wartime migration and the incidence of tuberculosis in the Zagreb region, Croatia. *Eur Respir J* 1998; **12**:1380–83.
 - 21 Krieger N, Waterman PD, Chen JT, Soobader MJ, Subramanian SV. Monitoring socioeconomic inequalities in sexually transmitted infections, tuberculosis, and violence: geocoding and choice of area-based socioeconomic measures – the public health disparities geocoding project (US). *Public Health Rep* 2003; **118**:240–60.
 - 22 Holtgrave DR, Crosby RA. Social Capital, Poverty and Income Inequality as predictor of Tuberculosis Case rates in the United States. *Abstract Book Natl HIV Prev Conf.* Atlanta: Emory University, Rollins School of Public Health, 2003; abstract no. MP 112.
 - 23 Muniyandi M, Ramachandran R, Gopi PG *et al.* The prevalence of tuberculosis in different economic strata: a community survey from South India. *Int J Tuberc Lung Dis* 2007; **11**:1042–45.
 - 24 Diez Roux AV. A glossary for multilevel analysis. *J Epidemiol Community Health* 2002; **56**:588–94.
 - 25 Susser M. The logic in ecological: I. The logic of analysis. *Am J Public Health* 1994; **84**:825–29.
 - 26 Susser M. The logic in ecological: II. The logic of design. *Am J Public Health* 1994; **84**:830–35.
 - 27 Kelly PM, Bonnefoy J. Measurement and knowledge network, *The Social Determinants of Health: Developing an Evidence Base for Political Action.* World Health Organization Commission on Social Determinants of Health, London: National Institute for Health and Clinical Excellence; 2007.
 - 28 SINAN/MS/SUS. <http://dtr2005.saude.gov.br/sinanweb/index.php?name=Tnet> (July 2007, date last accessed).
 - 29 SEPLAN. A dinâmica social do Recife. Diagnóstico. In: Recife PdCd, editor. *Recife.* <http://www.recife.pe.gov.br/secplanejamento/planodiretor/diagnostico-i.html> (7 Decemr 2007, date last accessed).
 - 30 Albuquerque MFPM, Ximenes RAA, Lucena-Silva N *et al.* Factors associated with treatment failure, drop out, and death in a cohort of tuberculosis patients in Recife, Pernambuco, Brazil. *Cad Saude Publica* 2007; **23**:1537–82.
 - 31 <http://www.sidra.ibge.gov.br/bda/tabela/listabl.asp?c=2106&z=1&o=23>.
 - 32 IBGE. *Base de informacao por setor censitario.* Instituto Brasileiro de Geografia e Estatística, Brazil, 2000 (in Spanish).
 - 33 Souza ES, Figueira FS, Silva GA *et al.* Soro-prevalência da infecção pelo HIV-1 em gestantes atendidas no IMIP, Recife- Brasil. *Revista do IMIP* 1995; **1**:925–29.
 - 34 Batista JAL, Albuquerque MFPM, Ximenes RAA *et al.* Smoking increases the risk of relapse after successful tuberculosis treatment. *Int J Epidemiol* 2008; **37**:841–51.
 - 35 Bruzzi P, Green SB, Byar DP, Brinton LA, Schairer C. Estimating the population attributable risk for multiple risk factors using case-control data. *Am J Epidemiol* 1985; **122**:904–14.
 - 36 Lindoso AABP, Waldman EA, Komatsu EK *et al.* Profile of tuberculosis patients progressing to death, city of São Paulo, Brazil, 2002. *RevSaúde Pública [online]* 2008; **42**:805–12.
 - 37 Paz AM, Siqueira MT. Avaliação da estratégia DOTS nas ações de controle na tuberculose em um centro de saúde do Recife. In: Livro de apresentação de pôsteres do 6º Congresso Brasileiro de Epidemiologia; 2004 Jun 19-23; Recife, Pernambuco. Rio de Janeiro: Abrasco; 2004. n. 3351 (in Spanish).
 - 38 Ruffino-Netto A, Villa TCS. *Tuberculosis Treatment. DOTS Implementation in Some Regions of Brazil. Background and Regional Features.* Millennium Institute-Rede TB. Brazilian Tuberculosis Research Network. CNPq. Conselho Nacional de Desenvolvimento Científico e Tecnológico, 2006.
 - 39 Ministério da Saúde. Secretaria de Políticas de Saúde. Manual técnico para o controle da tuberculose. Brasília: Ministério da Saúde; 2002 (in Spanish).
 - 40 Harling G, Ehrlich R, Myer L. The social epidemiology of tuberculosis in South Africa: a multilevel analysis. *Soc Sci Med* 2008; **66**:492–505.