Epidemiol. Infect. (2017), **145**, 1130–1134. © Cambridge University Press 2017
doi:10.1017/S0950268816003307

SHORT REPORT
Who are the patients that default tuberculosis treatment? – space matters!

C. NUNES[1,2]*, R. DUARTE[3], A. M. VEIGA[1,4] AND B. TAYLOR[5]

[1] *Escola Nacional de Saúde Pública, Universidade NOVA de Lisboa, Portugal*
[2] *Centro de Investigação em Saúde Pública, Universidade NOVA de Lisboa, Portugal*
[3] *Centro Diagnóstico Pneumológico de Vila Nova de Gaia, Faculdade de Medicina da Universidade do Porto, Portugal*
[4] *Lisbon School of Health Technology, Instituto Politécnico de Lisboa, Portugal*
[5] *Lancaster University, Faculty of Health and Medicine, Lancaster, UK*

Received 6 May 2016; Final revision 21 September 2016; Accepted 29 September 2016; first published online 16 January 2017

SUMMARY

The goals of this article are: (i) to understand how individual characteristics affect the likelihood of patients defaulting their pulmonary tuberculosis (PTB) treatment regimens; (ii) to quantify the predictive capacity of these risk factors; and (iii) to quantify and map spatial variation in the risk of defaulting. We used logistic regression models and generalized additive models with a spatial component to determine the odds of default across continental Portugal. We focused on new PTB cases, diagnosed between 2000 and 2013, and included some individual information (sex, age, residence area, alcohol abuse, intravenous drug use, homelessness, HIV, imprisonment status). We found that the global default rate was 4·88%, higher in individuals with well-known risk profiles (males, immigrants, HIV positive, homeless, prisoners, alcohol and drug users). Of specific epidemiological interest was that our geographical analysis found that Portugal's main urban areas (the two biggest cities) and one tourist region have higher default rates compared to the rest of the country, after adjusting for the previously mentioneded risk factors. The challenge of treatment defaulting, either due to other individual non-measured characteristics, healthcare system failure or patient recalcitrance requires further analysis in the spatio-temporal domain. Our findings suggest the presence of significant within-country variation in the risk of defaulting that cannot be explained by these classical individual risk factors alone. The methods we advocate are simple to implement and could easily be applied to other diseases.

Key words: Noncompliance, pulmonary tuberculosis, risk factors, spatial pattern, treatment.

Incomplete or failure of tuberculosis (TB) treatment increases the risk of antibiotic resistance, disease transmission, morbidity and mortality [1]. This study focuses on treatment default, which here we define as being off drugs for more than 8 weeks after completing at least 1 month of treatment. Definitions of possible outcomes and treatment regimens adopted in Portugal are fully aligned with WHO Guidelines [2]. Treatment default could be related to human behaviour at the patient level and/or to the provision of healthcare services. Since treatment default is modifiable it is crucial to try to understand the reasons these events occur. Transmission between individuals

* Author for correspondence: Professor C. Nunes, Escola Nacional de Saúde Pública, Universidade NOVA de Lisboa, Av. Padre Cruz, 1600-560 Lisboa, Portugal.
(Email: cnunes@ensp.unl.pt)

is mainly dependent on sputum-positive pulmonary cases, therefore, in a public health perspective, studies on this specific group (PTB) are necessary.

The main goals of this article are: (i) to understand how individual characteristics affect the likelihood of default to PTB treatment; (ii) to quantify the predictive capacity of these risk factors; and (iii) to evaluate if, having adjusted for the available individual factors, there was still significant residual spatial variation in default rates. We interpret the spatial effect as a possible warning for other individual factors not available (e.g. educational level), and for factors relating to the health service and thus identify regions in Portugal where defaulting rates are higher than can be explained by measured individual-level factors alone.

In this study we focus on new PTB cases, diagnosed between 2000 and 2013 in continental Portugal. We used data obtained from the national TB surveillance database (SVIG-TB) and included as risk factors individual information on sex (F/M), age, alcohol abuse (yes/no), (illicit) intravenous drug use (yes/no), homelessness (yes/no), HIV positivity (yes/no) and past/present imprisonment (yes/no). We modelled the spatial variation at the *freguesia* (small area) level – there are around 4000 such spatial units in Portugal. Our study only considers cases who have either defaulted from their treatment regimen or completed it; we do not consider the success of treatment here and all other possible outcomes were excluded (i.e. died, failed, not evaluated).

After producing descriptive statistics, we used logistic regression (LR) to model the impact of individual covariates on the odds of default. We quantify the predictive capacity of these risk factors using area under the receiver-operating characteristic (ROC) curve. We then used a logistic Generalized Additive Model (GAM) [3], which allowed us to incorporate individual covariates and a non-parametric spatial component to account for spatial autocorrelation [including latitude and longitude as a smoothed interaction term, $s(x,y)$].

The spatial component is used to model the variation in the risk of defaulting that cannot be captured through our measured individual-level covariates and in the present article, we attribute these effects to other individual risk factors or to contextual variables, namely differences in healthcare provision at the local level. For variable age we used four groups commonly used in the TB literature (0–14, 15–34, 35–64, $\geqslant$65 years). All analyses were conducted in R [4] using the *mgcv, gam* and *pROC* libraries, based on *glm, gam* and *roc* functions. Statistical significance was defined at the 5% level. Models were compared using Akaike's Information Criterion (AIC; lower values indicate better models) and we used backwards selection to choose the most parsimonious model. Ethics committee approval and informed consent were not required as data were derived from an Official National Surveillance System which had been previously anonymized.

Our study included 27 894 PTB cases, including 1361 patients who defaulted and 26 533 who completed their treatment regimens. Results from our regression models were based on 18 260 (65·46%) of cases due to incomplete covariate information.

The overall default rate in our study group was 4·88% and we found this to be higher in males (5·46% *vs.* 3·57%). The highest failure rate (6·12%) was in the 15–34 years age group, followed by the 35–64, 0–14, and $\geqslant$65 years age groups (4·65%, 4·61% and 2·44%, respectively). Individuals with known risk factors had higher default rates: homelessness (19·56% *vs.* 4·26%), HIV co-infection (14·62% *vs.* 2·78%), drug abuse (14·42% *vs.* 3·16%), imprisonment (9·52% *vs.* 4·44%), migration (8·85% *vs.* 4·37%) and alcohol abuse (8·39% *vs.* 3·67%). In single explanatory bivariate analyses, all results were statistically significant ($P < 0{\cdot}001$), partly due to the large size of our study sample.

The results from our statistical models are presented in Figure 1. Estimated crude odds ratios (OR) were obtained by univariable logistic binary models. The effect sizes were very high, with HIV, homelessness and intravenous drug use having ORs >5; the direction of estimated effects was found to be in agreement with the literature [5–9].

Following the backwards stepwise selection methods sex and imprisonment were not found to be statistically significant. The final multiple LR model yielded the following ORs: HIV positivity [OR 3·43, 95% confidence interval (CI) 2·85–4·13, yes *vs.* no], migrant status (OR 2·07, 95% CI 1·70–2·50, migrant *vs.* Portuguese), intravenous drug use (OR 2·76, 95% CI 2·27–3·36, yes *vs.* no). These are all high ORs and the direction of the effects are again in agreement with the literature [3–7]. Alcohol abusers and homeless individuals were also found to have high ORs close to 2 (OR 1·93, 95% CI 1·61–2·30 and OR 1·91, 95% CI 1·37–2·63, respectively). Our results suggest that older people have a greater chance of completing treatment (OR 1·03, 95% CI 0·44–2·07; OR 0·76, 95% CI 0·64–0·90, OR 0·71, 95% CI 0·50–0·99, respectively, for age groups 0–14, 35–64, $\geqslant$65 years,

https://doi.org/10.1017/S0950268816003307 Published online by Cambridge University Press

	(a) Odds ratio			(b) Spatial component – GAM
Variables (ref. class)	LR Crude	adjusted*	GAM *	
Sex (female)				
Male	1·559	n.s.	n.s.	
Age group (15-34) (ref.)				
0-14	0·742**	1·033**	1·011	
35-64	0·749	0·757	0·753	
≥65	0·385	0·714	0·719	
Alcohol (No)				
Yes	2·402	1·925	1·922	
Drugs (No)				
Yes	5·166	2·761	2·715	
Homeless (No)				
Yes	5·459	0·914	1·853	
HIV (No)				
Yes	5·999	3·430	3·291	
Migrant (PT)				
Migrant	2·134	2·072	1·893	
Imprisonment (No)				
Yes	2·264	n.s.	n.s.	

n.s., Not selected (stepwise method).
* Final model, using a stepwise selection method.
** Not statistically significant

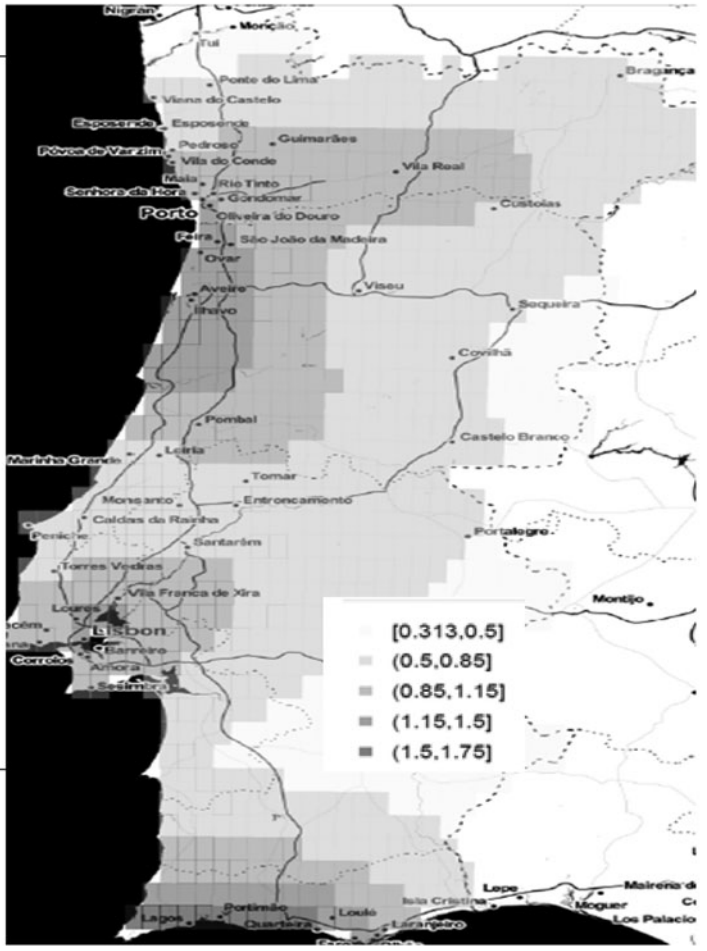

Fig. 1. Associations between individual risk characteristics and default treatment from both multiple regression models: Logistic regression (LR) and Generalized Additive Model (GAM). (*a*) Crude and adjusted odds ratios; (*b*) odds of failure over our study region, having adjusted for the other risk factors (spatial component of the GAM).

with the reference group being the 15–34 years age group).

Using the GAM (a multiple regression using a stepwise selection methods) the ORs were very similar but a smooth spatial term was also found to be statistically significant ($P = 0{\cdot}0472$). This term was found to improve the model – AIC from a value of 5308·4 for the LR model to 5296·622 for the GAM.

The GAM had a reasonable diagnostic accuracy in explaining the occurrence of treatment default (area under ROC 0·76, 95% CI 0·75–0·79).

Figure 1 shows a map of the spatial effect. This plot shows how the predicted odds of treatment default varied according to spatial location from our GAM.

A few studies have reported an association between treatment default and factors related to the social characteristics of patients [3–7]. In this study we found another component independent of these risk factors – this spatial component was very well defined. Three regions of Portugal were identified as having higher default rates – the two major cities (Lisbon and Oporto) and the south coast (Algarve). Porto and Lisbon are the biggest urban areas in Portugal, where more cases of TB are notified and where preventable risk factors (such as intravenous drug abuse, migration, homelessness) are more prevalent. These two cities are completely aligned with the epidemiology of TB in big cities in Europe [10] and are the only ones in Portugal with these characteristics. The third area identified (south coast of Portugal) is a tourist region receiving persons from other parts of the country or other countries with a

https://doi.org/10.1017/S0950268816003307 Published online by Cambridge University Press

very specific seasonal pattern (other tourist areas in Portugal do not have this seasonal characteristic) and with a small number of cases notified.

After having adjusted for individual risk factors, these areas were still clearly identifiable as having a higher default rate. There are other reasons, beside these individual risk factors, explaining the increased default rate, namely the default problem can occur unpredictably in unsuspected patients (with no classical non-compliant profiles) [11]. This observed spatial pattern could potentially be explained by the way our individual risk factors were measured or coded (i.e. what if alcohol consumption was not just 'no/yes'?), other unmeasured individual risk factors (such as social class, educational achievement, employment status, among others) or even organization issues [12], e.g. difficulty in providing treatment control or failure in the identification and follow-up of non-adherent patients. Additionally the status of Directly Observed Treatment-short course in TB treatment compliance management in Portugal could be relevant, but in our study we only know if it was recommended (and not if it was actually used).

Our study has some limitations. Although notification of TB in Portugal is mandatory, many subjects had at least some missing covariate data. We do not see this as a serious issue because: (1) comparisons of default rates, sex, age between those with and without missing data did not evidence different distributions; (2) between 2000 and 2013 there is a decreasing trend in the proportion of missing rates, suggesting global changes in practices of health services.

We wish to highlight that the use of the GAM methodological approach is very simple to implement and although arguably more principled modelling approaches are available, these often incur much greater computational expense. We would advocate the use of GAMs more widely at the exploratory stage due to the speed and simplicity with which models can be fitted.

In conclusion, we found that urban and tourist regions had higher default rates compared to the rest of the country, even after adjusting for all other identified individual risk factors (classical non-adherent patient profiles). The balance between the two different perspectives of default is well known: is it a problem caused by recalcitrant patients? or is healthcare system failing? [13]. The answer is not clear from our research and further analyses should be conducted that take into account both space and time as well as cultural factors [14]. Urban TB is well documented in Western Europe, but our study revealed another area with a completely different profile: a tourist area. Tourist areas have specific characteristics that mark them as different from other areas: they witness strong seasonal variations in population size and available health services, a generally deprived local population and also smaller numbers of cases. Our results are concordant with the literature, but they also clearly show that our scientific knowledge on treatment default is still incomplete. New research must be stimulated, with more comprehensive studies and new tailored strategies in order to increase compliance. The potential gain of using spatial or spatio-temporal methods (in general), in epidemiological and infection problems is illustrated here.

DECLARATION OF INTEREST

None.

ACKNOWLEDGEMENTS

The authors are very grateful to the coordinators of the National Program for Tuberculosis Control for access to data from the Tuberculosis Surveillance Database. This study was partially supported by FCT/MEC, Portugal (PTDC/SAUSAP/116950/2010 and SFRH/BSAB/1421/2014).

REFERENCES

1. **Grzybowski S, Enarson DA.** The fate of cases of pulmonary tuberculosis under various treatment programmes. *Bulletin of the International Union Against Tuberculosis* 1978; **53**: 70–75.
2. **World Health Organization.** Definitions and reporting framework for tuberculosis: 2013 revision. Geneva: World Health Organization, 2013.
3. **Wood SN.** *Generalized Additive Models: An Introduction with R.* Chapman and Hall/CRC, 2006.
4. **R Core Team.** R: A language and environment for statistical computing (https://www.R-project.org/). Accessed 7 April 2016.
5. **Faustini A, Hall AJ, Perucci CA.** Tuberculosis treatment outcomes in Europe: a systematic review. *European Respiratory Journal* 2005; **26**: 503–510.
6. **Vasankari T, *et al.*** Risk factors for poor tuberculosis treatment outcome in Finland: a cohort study. *BMC Public Health* 2007; **7**: 291.
7. **Dooley KE, *et al.*** Risk factors for tuberculosis treatment failure, default, or relapse and outcomes of retreatment in Morocco. *BMC Public Health* 2011; **11**: 140.

8. **Van Hest R, *et al.*** Tuberculosis treatment outcome monitoring in European Union countries: systematic review. *European Respiratory Journal* 2013; **41**: 635–643.
9. **Cayla JA, *et al.*** Current status of treatment completion and fatality among tuberculosis patients in Spain. *International Journal of Tuberculosis and Lung Disease* 2004; **8**: 458–464.
10. **De Vries G, *et al.*** Epidemiology of tuberculosis in big cities of the European Union and European Economic Area countries. *Eurosurveillance* 2014; **19**.
11. **Pablos-Méndez A, *et al.*** Nonadherence in tuberculosis treatment: predictors and consequences in New York City. *American Journal of Medicine* 1997; **102**: 164–170.
12. **Organisation for Economic Co-operation and Development.** *Geographic Variations in Health Care: What Do We Know and What Can Be Done to Improve Health System Performance?* Paris: OECD Publishing, 2014.
13. **Garner P, *et al.*** Promoting adherence to tuberculosis treatment. *Bulletin of the World Health Organization* 2007; **85**: 404–406.
14. **Verma G, *et al.*** Critical reflections on evidence, ethics and effectiveness in the management of tuberculosis: public health and global perspectives. *BMC Medical Ethics* 2004; **5**: 2.