Tuberculosis inequalities and socio-economic deprivation in Portugal

D. Apolinaário,* A. I. Ribeiro,†‡ E. Krainski,§ P. Sousa,† M. Abranches,# R. Duarte†‡**††

*Serviço de Pneumologia, Centro Hospitalar de Trás-os-Montes e Alto Douro, Vila Real, †EpiUnit, Instituto de Saúde Pública, Universidade do Porto, Porto, ‡Faculdade de Medicina, Universidade do Porto, Porto, Portugal; §Department of Statistics, Federal University of Paraná, Curitiba, Paraná, Brazil; #Norwegian University for Science and Technology, Trondheim, Norway; †Faculty of Engineering of the University of Porto, Porto, **Serviço de Pneumologia, Centro Hospitalar de Vila Nova de Gaia/Espinho, ††Chest Disease Centre, Vila Nova de Gaia, Portugal

SUMMARY

OBJECTIVE: To analyse the geographical distribution of tuberculosis (TB) in Portugal and estimate the association between TB and socio-economic deprivation.

METHODS: An ecological study at the municipality level using TB notifications for 2010–2014 was conducted. Spatial Bayesian models were used to calculate smoothed standardised notification rates, identify high- and low-risk areas and estimate the association between TB notification and the European Deprivation Index (EDI) for Portugal and its component variables.

RESULTS: Standardised notification rates ranged from 4.41 to 76.44 notifications per 100 000 population. Forty-one high-risk and 156 low-risk municipalities were identified. There was no statistically significant association between TB notification rate and the EDI, but some of its variables, such as the proportion of manual workers and the percentage unemployed, were significantly and directly associated with TB notification, whereas the variable ‘proportion of residents with low education level’ showed an inverse relationship.

CONCLUSION: Wide inequalities in TB notification rates were observed, and some areas continued to exhibit high TB notification rates. We found significant associations between TB and some socio-economic factors of the EDI.

KEY WORDS: TB; socio-economic factors; spatial analysis

DUE TO ITS WIDESPREAD geographical distribution, along with the emergence of multidrug resistance and co-infection with the human immunodeficiency virus (HIV), tuberculosis (TB) remains a severe public health problem.1 Portugal is among the 10 poorest countries of the European Union, and recently underwent a deep economic recession due to the global economic crisis, which led many countries to reduce their support for TB control programmes.2 Portugal, a country with a medium to low incidence of TB, reached an incidence of 20 per 100 000 population for the first time in 2014, and has one of the highest TB burdens in Europe.3,4

The scientific and public health communities recognise the fact that to attain the TB targets, including its eradication, it is necessary to expand the scope of TB control beyond the usual efforts by addressing one of the drivers of the epidemic: the socio-economic determinants of health.5 Precarious social and economic conditions, such as living in crowded and poorly ventilated spaces or delayed diagnoses due to barriers to accessing health services, lead to increased TB transmission and exposure to other TB risk factors, such as malnutrition, alcoholism or other comorbidities.5,6 There has been increasing interest in the study of socio-economic factors as determinants of TB, and several associations have been identified.6 In Portugal, it was found that those areas with the highest TB incidence were also those with high incidence of HIV infection, incarceration, overcrowding, unemployment and immigrants.7 Research into social determinants is thus essential to understand the geographic inequalities in TB incidence and to define targeted intervention strategies.5

Deprivation is a multidimensional concept that includes a combination of factors, such as low education levels, low social status, crowding and unemployment, among others.8 Due to the multidimensional nature of the concept of deprivation and for practical reasons (public health professionals prefer to use a single index to identify high-risk areas for targeted community and place-based initiatives rather than an array of individual variables), socio-
economic deprivation is often measured using composite indicator indices. Not only are these indices used in public health practice, they have also been increasingly employed in epidemiological research. Several studies have reported significant associations between socio-economic deprivation indices and the occurrence of TB.8–10

Until 2015, apart from the creation of the European Deprivation Index (EDI), no other validated or robust multivariate index of deprivation was available for Portugal. The EDI is a transnational index of socioeconomic deprivation that was developed for five European countries.11 It is formed by a weighted combination of variables obtained from the national census that are most highly associated with the country’s specific individual deprivation indicator. The variables included and the corresponding weights for Portugal are shown in Table 1.12 As the choice of this set of variables was based on the Townsend conceptualisation of deprivation and on robust statistical criteria, they represent the best set of variables available for capturing deprivation in Portugal available at that time and at the municipality level. Furthermore, there is evidence to suggest that all the variables comprising the EDI (education, occupation, crowding, housing conditions, age and unemployment) were individually associated with TB. We therefore hypothesised that by using a multivariate index we would be able to better understand the relationship between TB and deprivation in Portugal.

In addition to examining the link between TB incidence and socio-economic determinants from a public health perspective, it is crucial to identify priority areas with a higher risk of disease. Spatial analysis is an essential tool that includes the mapping of diseases and their determinants, and the study of potential geographical correlations.18 Spatial analysis is crucial for planning health activities for the monitoring and prevention of diseases to reduce inequalities, and spatial analysis tools are now commonly used for various diseases, including TB.7,15,19,20

In addition to what has already been studied in Portugal on the subject,7,19,21 it is important to recognise the spatial distribution of TB in recent years and identify the current high- and low-risk areas. The economic recession in the 2000s may have widened the geographic asymmetries in TB incidence, as several studies have reported that economic crises escalate social and economic polarisation of health and disease.22 It is therefore crucial to investigate the role of socio-economic factors on these geographic differentials and to assess the usefulness of multivariate indices, such as the EDI, in the understanding of the socio-economic determinants of TB.

The aim of the present study was to analyse the spatial distribution of TB notification across municipalities in Portugal and to evaluate the association between TB notification and the EDI and the various socio-economic factors included in its calculus.

### MATERIALS AND METHODS

#### Design and study area

An ecological spatial study was conducted in Portugal (continental and archipelagos) using municipalities as units of analysis. In 2014, there were 308 municipalities in Portugal, with an average population of 33 684 (range 459–509 312).23

#### Study population and data sources

We analysed all TB cases notified in Portugal from 2010 to 2014 based on data obtained from the national surveillance system for TB, Sistema de Vigilância da TB (SVIG-TB), disaggregated by municipality, sex and quinquennial age groups. The source of population data was the National Statistics Institute, Lisbon, Portugal (https://www.ine.pt/), and the estimated population by municipality for the period studied was recorded and stratified by sex and quinquennial age groups.

#### Tuberculosis notification rates

TB cases notified per 100 000 residents were calculated at the municipality level. The mean TB notification rate for the 5 study years was used to reduce random fluctuations in notification rates.

#### Covariates: socio-economic deprivation and its components

The EDI for Portugal (Table 1) was used to classify small areas according to their level of socio-economic deprivation,12 as well as the eight census variables that composed the EDI, which were individually included as covariates. The EDI and the individual variables were standardised and classified into five quintiles (from q1, the least deprived, to q5, the most deprived) because, since its inception11 and in its applications,24,25 the EDI has been discretised as quintiles to facilitate the identification of vulnerable communities, interpret results and compare geographic patterns of deprivation and health events.

### Table 1  Formula for the calculation of the EDI score, Portugal

<table>
<thead>
<tr>
<th>Score calculus</th>
<th>EDI = % non-owned households (Z score) × 1.193</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+ % households without indoor flush lavatory (Z score) × 1.456</td>
</tr>
<tr>
<td></td>
<td>+ % residents with low education level (≤6th grade) (Z score) × 1.292</td>
</tr>
<tr>
<td></td>
<td>+ % household with ≤5 rooms (Z score) × 0.404</td>
</tr>
<tr>
<td></td>
<td>+ % unemployed looking for a job (Z score) × 0.376</td>
</tr>
<tr>
<td></td>
<td>+ % female residents aged ≥65 years (Z score) × 0.255</td>
</tr>
<tr>
<td></td>
<td>+ % households with no bath/shower (Z score) × 0.060</td>
</tr>
<tr>
<td></td>
<td>+ % residents employed in manual occupations (Z score) × 0.013</td>
</tr>
</tbody>
</table>

EDI = European Deprivation Index.
Moreover, we found that the EDI, its component variables and TB followed a generally linear relationship, which did not imply any other specific cut-off values.

**Statistical analysis**

To take into account the spatial autocorrelation and large variance of small areas, we used a hierarchical Bayesian spatial model. To guarantee that the geographic patterns and associations assessed were not driven by the different age structures of the Portuguese municipalities, notification rates were age-standardised.

We assumed that the response variable, TB cases \(O_i\), in each \(i\)th area follows a Poisson distribution, where \(E_i\) is the expected number of cases and \(\theta_i\) the relative risk (RR) or standardised notification rate (Equation 1). We used the Portuguese TB notification rates by sex and age group (5-year age groups) as reference to compute the expected number of cases:

\[
O_i \sim \text{Poisson}(E_i, \theta_i) \\
\log(\theta_i) = \alpha + \beta x_i + s_i
\]  

(Equation 1)

where \(\alpha\) is an intercept quantifying the average number of TB cases in the 308 areas, and \(\beta x_i\) is the effect of each covariate. Each covariate was discretised into five quintiles and included in the model under the usual regression framework: the first class was used as the reference level; a coefficient was then fitted for every other class, which was then used to compute RRs. To ensure that the relationship between TB and the covariates remained linear, we also ran a flexible regression model to inspect the shape of the associations and look for inflection points; no inflection points were detected.

The area-specific effect \(s_i\) was modelled based on a Besag, York and Mollié (BYM) model, with the parameterisation suggested by Dean et al. (Equation 2.2): \(28\)

\[
s_i = \tau(\sqrt{\phi}u_i + \sqrt{1 - \phi}v_i)
\]  

(Equation 2.2)

where \(u_i\) is the structured effect and \(v_i\) the unstructured effect. The \(u_i\) effect was scaled to make the model more intuitive and interpretable, so that \(\phi\) expresses the proportion of the spatial effect due to the structured part and \(1/\tau\) is the marginal variance of \(s_i\).

First, we fitted a univariable model to evaluate the individual association between TB and the EDI. We then examined the associations between TB and each of the socio-economic factors comprising the EDI (Table 1). Finally, we fitted a multivariable model (final model), which was chosen based on the different goodness-of-fit measures used (deviance information criteria, Watanabe-Akaike information criterion and conditional predictive ordinate) and where we retained the socio-economic factors significantly associated with TB.

Associations were expressed in RRs, which express the ratio between the risk of TB of the covariate class and the risk of the reference class (the least deprived were used as the reference). For example, an RR of 1.10 in a certain covariate class meant that the risk of TB for that class was 10% higher than the risk in the reference class. An RR would be considered significantly higher or lower if its 95% credible intervals (95%CrIs) did not include the value 1. RRs and 95%CrIs were derived from their posterior means and quintiles. Posterior distributions were obtained using the Integrated Nested Laplace Approximation (INLA), which was implemented in the R-INLA library. \(30\)

The geographic distribution of TB notification was obtained from the final model. We also used the function ‘excursions’ to identify high- and low-risk areas. \(31,32\) This method uses the posterior joint distribution computed from INLA and takes into account the dependence structure, allowing us to accurately identify areas where the notification rate is greater or smaller than a certain threshold. In this study, we defined 20 cases per 100 000 as the threshold to determine high-/low-risk municipalities. A significance level of 0.05 was set up. TB-standardised notification rates and high- and low-risk areas were mapped using ArcMap 10.2.1 (Environmental Systems Research Institute, Redlands, CA, USA).

As patient data were fully anonymised, data protection permission and ethical approval were not required.

**RESULTS**

From 2010 to 2014, 12 592 cases of TB were notified; 71.7% were pulmonary, and 65.5% of the cases occurred in men. The mean age of the patients was 47.4 ± 18.8 years. The mean crude TB notification rate between 2010 and 2014 was 23.9/100 000 (95%CrI 23.5–24.3).

The spatial distribution of the standardised notification rates is shown in Figure A. High- and low-risk municipalities are shown in Figure B. Standardised notification rates ranged from 4.41 to 76.44 notifications/100 000. The 41 highest-risk municipalities were mainly observed in the Porto metropolitan area, particularly the suburban municipalities, followed by the Lisbon metropolitan area, the North coast, the Alentejo coast and Algarve. The 136 lowest-risk municipalities were essentially in the Centre region and inner Alentejo.

We did not find a statistically significant association between TB notifications and the EDI (Table 2). We did, however, find statistically significant associations with some of the variables comprising the EDI (proportion of households with ≤5 rooms, proportion of non-owned households, percentage of residents with low education level and percentage unemployed) in the univariate analysis (Table 2). In
the multivariate analysis, the covariates ‘percentage of manual workers’ and ‘percentage unemployed’ were significantly and positively associated with a higher risk of TB in most socio-economically deprived municipalities. On the other hand, the variable ‘percentage of residents with a low education level’ had an inverse relationship with TB (Table 2).

DISCUSSION

We observed strong inequalities in the spatial distribution of TB in Portugal, with high-risk municipalities corresponding mainly to large urban centres. In the analysis of the association between TB and socio-economic deprivation, the EDI was not associated with TB, and only some of its variables associated with occupation, unemployment and education were related.

During the study period, Portugal was still classified as an intermediate incidence country based on global TB notification rates and, at the municipality level, large asymmetries were confirmed, with standardised notification rates ranging from 4.41 to 76.44 notifications/100 000. Some municipalities

<table>
<thead>
<tr>
<th>Variable</th>
<th>EDI RR (95%CrI)</th>
<th>Household with ≤ 5 rooms RR (95%CrI)</th>
<th>No bath or shower RR (95%CrI)</th>
<th>No indoor flush lavatory RR (95%CrI)</th>
<th>Non-owner households RR (95%CrI)</th>
<th>Women aged ≥ 65 years RR (95%CrI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: least deprived</td>
<td>1.00 (reference)</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>2</td>
<td>1.08 (0.90–1.30)</td>
<td>1.01 (0.83–1.24)</td>
<td>0.90 (0.75–1.09)</td>
<td>1.12 (0.93–1.34)</td>
<td>1.02 (0.83–1.25)</td>
<td>0.94 (0.77–1.14)</td>
</tr>
<tr>
<td>3</td>
<td>1.14 (0.94–1.38)</td>
<td>1.10 (1.09–1.35)</td>
<td>1.03 (0.87–1.27)</td>
<td>0.95 (0.78–1.17)</td>
<td>1.06 (0.87–1.30)</td>
<td>0.94 (0.76–1.17)</td>
</tr>
<tr>
<td>4</td>
<td>1.19 (0.96–1.46)</td>
<td>1.33 (1.05–1.68)</td>
<td>0.94 (0.75–1.17)</td>
<td>0.97 (0.79–1.20)</td>
<td>1.24 (1.01–1.53)</td>
<td>0.97 (0.77–1.23)</td>
</tr>
<tr>
<td>5: most deprived</td>
<td>1.12 (0.90–1.38)</td>
<td>1.68 (1.29–2.20)</td>
<td>0.90 (0.71–1.14)</td>
<td>0.93 (0.75–1.17)</td>
<td>1.55 (1.24–1.94)</td>
<td>0.86 (0.65–1.12)</td>
</tr>
</tbody>
</table>

* Only covariables with significant associations.
EDI = European Deprivation Index, RR = relative risk, CrI = credible interval.
had notification rates higher than 50/100,000, far above the threshold for high incidence. The spatial distribution of TB has previously been studied in Portugal at the municipality level, although only for pulmonary cases in 2010, and these studies also reported a high degree of spatial heterogeneity.\textsuperscript{7,19,21} High-risk municipalities were found mainly in the principal urban centres (Porto and Lisbon metropolitan areas). This is consistent with previous studies that identified these critical spatiotemporal TB foci. Some municipalities on the Alentejo coast and Algarve also constituted high-risk areas, probably due to the large migrant population in this region during the summer.

Although the association between poverty and TB has been recognised for many years,\textsuperscript{5,6} the composite socio-economic deprivation index was not significantly associated with the TB notification rate in our study; however, some of its component variables were related. This result was also confirmed in another study looking at the impact of a multivariate index and its components.\textsuperscript{10} As it is a multivariable indicator, it theoretically should better reflect socio-economic deprivation than individual indicators. This might be related to the fact that not all socio-economic variables that comprise the EDI have the same epidemiological relevance to the risk of TB. The use of a composite indicator may therefore not be ideal in such cases. However, it is also well-known that TB incidence depends on other factors in addition to socio-economic ones, such as comorbidities and certain risk behaviours, and a proven spatial association with the incidence of pulmonary TB in Portugal,\textsuperscript{7} and these were not analysed in the present study. Another limitation is that, despite the use of an in-depth geographic-based analysis, there were substantial disparities in socio-economic level in some municipalities (especially in urban settings), and thus aggregated indicators were given as average values, which lacked discriminative power.

When analysed independently, two EDI variables—the percentage of manual workers and the proportion unemployed—showed a significant direct association with TB notification, i.e., TB rates were higher in municipalities with higher proportions of manual workers and those with higher unemployment rates. This observation is consistent with previous studies that associated high TB incidence with low household income\textsuperscript{33,34} and unemployment,\textsuperscript{7,17} which may be related to the higher poverty levels in these populations. In the case of manual workers, suboptimal working conditions and occupational exposure to inhalatory risks may contribute to this association.\textsuperscript{13,14} However, the EDI covariate ‘proportion of residents with low education level’ showed a significant inverse correlation with the number of notified TB cases, with the risk being higher for groups with a higher education level. This observation is in contrast to findings from other studies,\textsuperscript{9,34} and may be explained by lower diagnostic suspicion and a delayed diagnosis reported among persons with university degrees,\textsuperscript{15} with a consequent increase in the transmission period.

A possible study limitation is the geographic heterogeneity in TB notifications, which may have resulted in notification bias. However, as TB notification is compulsory in Portugal, this is unlikely and, as a mean TB incidence rate of 5 years was used, the effect of the potential bias in results was probably mitigated.

Our study had several strengths. First, we used robust statistical methods to characterise geographic patterns, which allowed us to identify the priority areas for intervention. Adequate statistical methods were also used to estimate the association between socio-economic factors and TB. The BYM model included two random effects: spatially structured and spatially unstructured. The former accounted for the spatial structure of the observations, and the latter dealt with non-spatial variability. These effects act as latent confounders, which we could not account for.

In conclusion, we found wide inequalities in TB notification rates and were able to identify areas with a particularly high risk; these should be subjected to further detailed geographic analyses to identify the

### Table 2 (continued)

<table>
<thead>
<tr>
<th>Univariate analysis</th>
<th>Multivariate analysis*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low education level ( (&lt;6^{th} \text{ grade}) )</td>
<td>( \text{RR (95%CrI)} )</td>
</tr>
<tr>
<td>Manual occupations</td>
<td>( 0.90 ) (0.75–1.07)</td>
</tr>
<tr>
<td>Unemployed looking for a job</td>
<td>( 0.79 ) (0.66–0.96)</td>
</tr>
<tr>
<td>Low education level ( (&lt;6^{th} \text{ grade}) )</td>
<td>( 0.68 ) (0.55–0.84)</td>
</tr>
<tr>
<td>Manual occupations</td>
<td>( 0.75 ) (0.60–0.92)</td>
</tr>
<tr>
<td>Unemployed looking for a job</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
complex mechanisms underlying TB, which would help develop specific intervention strategies. We found significant associations between some (but not all) socio-economic factors and TB, which suggest that the use of multivariate indices might not be useful for understanding the determinants of some diseases, such as TB.

Acknowledgements
This work was supported by contributions from Iceland, Liechtenstein and Norway through the European Economic Area Grants, under the Public Health Initiatives Programme (PT 06, grant number 138071).

The authors are grateful to C Rodrigues (Centro Hospitalar de Trás-os-Montes e Alto Douro, Portugal) for help in the elaboration of the database for the statistical analysis.

Conflicts of interest: none declared.

References
OBJECTIF: Analyser la distribution géographique de la TB au Portugal et estimer l’association entre TB et le dénuement socioéconomique.

MÉTHODES: Une étude écologique a été réalisée au niveau d’une municipalité, basée sur les notifications de la TB entre 2010 et 2014. Des modèles spatiaux bayésiens ont été utilisés pour calculer les taux de notification standardisés lissés, délimiter les zones de risque élevé/faible et estimer l’association entre la notification de la TB et l’index européen de défavorisation (IED) pour le Portugal et les variables qui le composent.

RESULTATS: Les taux standardisés de notification sont allés de 4,41 à 76,44 notifications par 100 000 habitants. Quarante et une municipalités à haut risque et 156 à faible risque ont été identifiées. Il n’y a pas eu d’association statistiquement significative entre le taux de notification de la TB et l’IED, mais certaines de ses variables, comme le « pourcentage de travailleurs manuels » et le « pourcentage de chômeurs », ont été significativement et directement associées avec la notification de la TB, tandis que la variable « pourcentage de résidents avec un faible niveau d’instruction » a montré une relation inverse.

CONCLUSION: De grandes inégalités des taux de notification de la TB ont été observées et quelques zones ont encore eu des taux de notification élevés de TB. Nous avons découvert une association significative entre la TB et certains facteurs socioéconomiques de l’IED.

RESUMEN

OBJETIVO: Analizar la distribución geográfica de la tuberculosis (TB) en Portugal y evaluar la asociación entre la TB y la penuria socioeconómica.

MÉTODOS: Se llevó a cabo un estudio ecológico a escala municipal, a partir de las notificaciones de TB registradas durante el periodo del 2010 al 2014. Mediante modelos bayesianos espaciales se calcularon las tasas de notificación normalizadas suavizadas, se delimitaron las zonas de alto y bajo riesgo y se evaluó la asociación entre la notificación de la TB y el índice europeo de privación (IED, por European Deprivation Index) para Portugal, con las variables que lo componen.

RESULTADOS: Las tasas de notificación normalizadas oscilaron entre 4,41 y 76,44 por 100 000 habitantes. Se encontraron 41 municipios donde el riesgo de contraer TB era alto y 156 donde era bajo. No se observó una asociación con significación estadística entre la tasa de notificación de TB y el IED; sin embargo, algunas de sus variables como ‘el porcentaje de trabajadores manuales’ y ‘el porcentaje de desempleados’ se correlacionaron de manera directa y significativa con la notificación de TB y la variable ‘porcentaje de residentes con bajo grado de instrucción’ reveló una correlación inversa.

CONCLUSIÓN: Se observaron desigualdades notables en las tasas de notificación de la TB en las diferentes zonas, que siguen siendo altas en unas pocas regiones. Se encontraron asociaciones significativas entre la TB y algunos factores socioeconómicos del EDI.