Spatio-temporal co-occurrence of hotspots of tuberculosis, poverty and air pollution in Lima, Peru

Gabriel Carrasco-Escobar1,2, Alvaro Schwalb3, Kelly Tello-Lizarraga4, Percy Vega-Guerovich5 and Cesar Ugarte-Gil3,6,7*

Abstract
Growing evidence suggests pollution and other environmental factors have a role in the development of tuberculosis (TB), however, such studies have never been conducted in Peru. Considering the association between air pollution and specific geographic areas, our objective was to determine the spatial distribution and clustering of TB incident cases in Lima and their co-occurrence with clusters of fine particulate matter (PM2.5) and poverty. We found co-occurrences of clusters of elevated concentrations of air pollutants such as PM2.5, high poverty indexes, and high TB incidence in Lima. These findings suggest an interplay of socio-economic and environmental in driving TB incidence.

Keywords: Tuberculosis, Hotspots, air pollution

Background
In 2017, 10 million new cases of tuberculosis (TB) occurred worldwide [1] which constitute a major health burden that strains middle- and low-income countries. Many socio-economic factors within these countries are frequently associated with higher TB incidence such as poverty, unemployment, low income, overcrowding, and population density [2]. It is well known that TB is prone to spatial aggregation often in poor areas of cities and can even be associated with a higher risk of infection, as observed in Southern Ethiopia where the risk is 4.16 times higher inside a cluster [3]. The use of geographical surveillance in public health allows for the detection of areas with a high prevalence or incidence of a particular disease in order to identify socio-economic factors associated with the phenomenon [4]. These methods have been applied to TB transmission [5]. Spatial information contributes to appropriate decision-making with a more efficient budget and human resources allocation and has been used previously in infectious diseases to detect hotspots and epidemics [6].

Environmental factors such as pollution and suspended particles are considered to play important roles in the development of TB [7]. This is explained by secretion clearance impairment by epithelial cells of the respiratory tract which is the primary defense mechanism against Mycobacterium tuberculosis [8].

Studies on spatio-temporal distribution of TB cases have principally addressed their association with demographic and geographical predictors for multidrug-resistant TB in Peru [9–11], nonetheless, their possible association with environmental factors such as air pollution have yet to be of interest. This descriptive study sought to determine the spatial distribution and clustering of TB cases in Lima, Peru and determine co-occurrence with clusters of PM2.5 and economic index.

Methods

Study design
Ecological analysis using Peruvian Ministry of Health (MoH) data of TB incidence cases from 2015 to
2017 and high-resolution fine particulate matter with aerodynamic diameter of 2.5 μm or smaller (PM_{2.5}) from the National Aeronautics and Space Administration (NASA) [12].

**Study area and population**

This study was conducted in Lima, Peru. Lima has a population of almost 9 million in an area of 2672 km². It is divided in 43 districts (Fig. 1a), some of which are the most densely populated districts in Peru. Overall, Lima is the most developed province in the country with the largest internal migration rate.

**Data sources**

**Tuberculosis cases**

All new TB cases are reported in the 342 health centers of the MoH in Lima; new cases diagnosed in hospitals are reported from their corresponding health center. We obtained incidence data from 2015 to 2017. TB diagnosis in the health centers is based on clinical suspicion (cough for more than 2 weeks, fever, night sweats) and/or microbiological confirmation. All reported new pulmonary TB cases (smear positive or negative) were used for this study. The populations of each district were provided by the National Institute of Statistics and Informatics (INEI) via the RETrieve of DATa for small Areas by Microcomputer (REDATAM) platform; incidence rates were computed at district-level per 100 000 habitants. Data handling and formatting were performed using R software.

**Particulate matter 2.5 μm (PM_{2.5})**

Air quality data from the Socioeconomic Data and Applications Center (SEDAC) by NASA was used [13]. SEDAC provides annual global surface concentrations (micrograms per cubic meter, μg/m³) of mineral dust and sea-salt filtered atmospheric PM_{2.5} [12, 14]. From 2015 to 2016, PM_{2.5} gridded datasets were provided at a spatial resolution of 0.01 degrees (~ 1.11 km at the equator) [14]. This product is the computation of aerosol optical depth (AOD) from multiple satellite instruments including the NASA Moderate Resolution Imaging Spectroradiometer (MODIS), Multi-angle Imaging Spectro Radiometer (MISR), and the Sea-Viewing Wide Field-of-View Sensor (SeaWiFS). SEDAC used a GEOS-Chem chemical transport model to relate the total column measure of aerosol to near-surface PM_{2.5} concentration and a geographically weighted regression (GWR) with global ground-based measurements to predict and adjust for the residual PM_{2.5} bias per grid cell in the initial satellite-derived values, previously validated [15]. We processed the raster data in Google Earth Engine (GEE) and summarized as the median PM_{2.5} value per year and district, scaled by 100.

**Poverty level**

Poverty level at the individual, household and neighborhood level was provided by INEI [16]. The downscaling of poverty indicators [17] were conducted using data from the National Census and the National Survey of Households (ENAHO), a specialized survey that includes detailed information about incomes and expenditures. Information on level of poverty was only provided for 2016, however, no significant relative changes were observed in previous estimates from 2013 [18]. Dalenius-Hodges method [19] was used to compute five poverty-stratum at individual, household and neighborhood levels. A dimension reduction was conducted using a principal component analysis (PCA) in order to assign a single poverty level value per district. Final poverty index was computed as the additive inverse of all principal components with an eigenvalue greater than 1, higher

---

**Fig. 1** Study area in Lima Province, Peru. **a** Districts in Lima, each color represents a district. **b** Tuberculosis incidence rate (cases/100 000 pop) in Lima, 2015–2017
values reflecting an impoverished population and lower values a wealthy population.

Spatial analysis
Spatial autocorrelation of TB cases, poverty, and PM$_{2.5}$ were assessed using global Moran’s $I$ statistics to describe the overall spatial dependence in the entire study area. In addition, local Getis-Ord $Gi^*$ statistic (a type of Local Indicator of Spatial Association - LISA) was used to identify local patterns and high-risk areas. A first-order queen contiguity-based weighted neighborhoods (districts with contiguous patterns and high-risk areas) was used to identify local variation in TB (GLMM) was created to assess the importance of poverty and PM$_{2.5}$ as drivers of spatial variation in TB. The Gini Index was calculated to assess the disproportionate distribution of cases in Lima districts. The Gini Index was calculated as follows:

$$
\log(\rho_{st}) = \alpha + \gamma_{t(a)} + \phi_s + \upsilon_s
$$

Where the TB rate for each district and year ($\log(\rho_{st})$) is modeled by 1) an intercept ($\alpha$), 2) an exchangeable random effects for each year ($\gamma_{t(a)}$) to account for inter-annual variation in TB over time (yearly random effect); and 3) spatially unstructured ($\phi_s$), and structured ($\upsilon_s$) random effects using a convolution prior that combines area-specific overdispersion and a neighborhood dependency structure [24]. Covariates (poverty level and PM$_{2.5}$) were added to this model and model parameters were estimated within a Bayesian framework using Integrated Nested Laplace Approximation (INLA) [25], an alternative to Markov Chain Monte Carlo (MCMC) methods. Models were fitted using the ‘INLA’ package in R software.

Results
Baseline characteristics and TB incidence
A total of 28,381 new pulmonary TB cases were reported during the study period (2015–2017) with stable yearly rates. The incidence rate at district-level ranged between 18.8–214, 20.2–216, and 0–199 cases per 100,000 inhabitants in 2015, 2016, and 2017, respectively (Fig. 1b). The average PM$_{2.5}$ concentration at district-level was slightly higher in 2016 (29.5 $\mu g/m^3$) than in 2015 (26.2 $\mu g/m^3$). The average PM$_{2.5}$ was highly heterogeneous across districts in Lima, ranging between 13.8–40.4 $\mu g/m^3$ in 2015, and 16.5–44.8 $\mu g/m^3$ in 2016.

Spatial clustering and co-occurrence
An overall strong spatial autocorrelation was observed during the 2015–2017 period for TB cases (Moran’s $I$ range: 0.24–0.36, $P < 0.01$), PM$_{2.5}$ (Moran’s $I$ range: 0.55–0.56, $P < 0.001$), and poverty level (Moran’s $I = 0.353$, $P = 0.003$) (Supplementary Fig. 1). Stable high-risk clusters of TB cases were observed in the central-east part of Lima and low-risk clusters in the south-west (Fig. 2a). High-risk clusters of PM$_{2.5}$ were detected in the south and central-east of Lima, with low-risk clusters located in the south. Furthermore, cold spots of poverty level (cluster of wealthy districts) were located in the central-west. Statistically significant co-occurrence of clusters (seven categories - cold spot 99% confidence, cold spot 95% confidence, cold spot 90% confidence, not significant, hot spot 99% confidence, hot spot 95% confidence, hot spot 90% confidence) of TB cases and PM$_{2.5}$ (Kendall’s $W = 0.596$; $P = 0.046$), TB cases and poverty level (Kendall’s $W = 0.4714$; $P = 0.003$), and the three variables combined (Kendall’s $W = 0.4606$; $P = 0.001$) were observed.

Mixed-effects models and inequity indexes
A moderate concentration of TB cases among Lima districts were observed in the study period (Gini Index range = 0.27–0.29) (Fig. 2b). The general trends of TB

Statistical analysis
The Gini Index was calculated to assess the disproportionate distribution of cases in Lima districts. The Gini coefficient is a common measure of the inequality among values of a frequency distribution (TB cases). It is defined as a ratio with values between 0 (perfect equality) and 1 (perfect inequality). The Gini coefficient was computed using the ‘ineq’ package in R software.

A negative binomial generalized linear mixed model (GLMM) was created to assess the importance of poverty level and PM$_{2.5}$ as drivers of spatial variation in TB incidence rate across Lima. A baseline model was formulated as follows:
cases with PM$_{2.5}$ and poverty level are present in Fig. 2c. The slight decrease in TB incidence, in spite of the increasing poverty index, can be explained by the efforts of the MoH and National TB Program towards TB control among low-resource districts. The spatio-temporal Bayesian mixed-effects negative binomial multivariate regression shows that PM$_{2.5}$ (adjusted relative risk [ARR] = 1.31; 95% credible interval [CI]: 1.17–1.50) and poverty level (ARR = 1.14; 95% CI: 1.11–1.17) were associated with TB incidence rate in Lima, accounting for the spatio-temporal structure of the districts (Table 1); spatial and temporal random effects are shown in Fig. 3.

**Discussion**

This study describes the skewed spatial distribution of TB cases in Lima. A strong spatio-temporal clustering of poverty, environmental fine particulate matter PM$_{2.5}$ and incident TB cases was observed during the study period. Importantly, clusters of elevated PM$_{2.5}$ concentrations, poverty, and high TB incidence significantly co-occur during the 2015–2016 period, suggesting that socioeconomic determinants and environmental conditions interplay as important components of transmission and clustering of TB in this setting.

Lima is the capital with one of highest ambient air pollution levels in Latin America [26]; however, its burden on health remains largely unknown. Previous studies in Peru have explored the effect of air pollution as a determinant of asthma [27]. The most important effects of air pollution manifest in the upper respiratory tract where it alters immune response [8], contributing as a susceptibility factor for various respiratory diseases [28]. One study showed impaired the expression of CD69, IFN-γ and TNF-α when human peripheral blood mononuclear cells are exposed to *Mycobacterium tuberculosis* [29]. Likewise, exposure to air pollution has been linked to many substantial adverse effects on human health; in particular people with chronic respiratory diseases such as asthma and chronic obstructive pulmonary disease.

These findings are consistent with previous studies supporting the hypothesis of air pollution as an environmental determinant of TB [30]. One study in Jiangsu (China) showed association between TB and long-term pollution exposure (using PM$_{2.5}$, PM$_{10}$, SO$_2$ and NO$_2$ measurements); similar findings were found in North Carolina for

**Table 1** Estimates of Spatio-temporal Bayesian mixed-effects negative binomial regression

| Variables       | Unadjusted |                  |                  | Adjusted         |                  |                  |
|-----------------|------------|------------------|------------------|------------------|------------------|------------------|
|                 | RR         | Std. Dev.        | 95% credible interval | RR         | Std. Dev.        | 95% credible interval |
| PM$_{2.5}$      | 1.306      | 0.063            | 1.154–1.484      | 1.318            | 0.064            | 1.166–1.504      |
| Poverty level   | 1.142      | 0.013            | 1.112–1.173      | 1.142            | 0.013            | 1.113–1.172      |

RR Relative Risk, PM$_{2.5}$ fine particulate matter
long-term pollution exposure (PM$_{10}$ and PM$_{2.5}$) [7]. Short-term exposure to outdoor pollution also seems to be a factor for active TB [31] and can also increase the risk for TB infection [32]. Poverty condition also seems to be a key variable for unhealthy concentrations of PM$_{2.5}$ and high-risk of TB alike. Although poverty has been proven to be a driver of the TB epidemic, it is also an important determinant of outdoor and indoor air pollution. The concentrations of the pollutants stated above are often highest largely in the urban areas of low- and middle-income countries [33]. Rapid rural-urban migration, as experienced in Lima, has created overcrowded districts characterized by poverty and increased air pollution concentrations [34]. The inequalities of TB are seldom looked at from an environmental perspective; however, these findings suggest the complex interaction between socio-economic factors and environment pollution in the transmission of TB in Lima.

Some limitations are acknowledged in this study. First, the high spatial resolution of the NASA-SEDAC PM$_{2.5}$ estimates could not be fully harnessed due to the coarse spatial resolution of the TB data at the district level. Some districts present a high heterogeneity in PM$_{2.5}$ between sub areas that were not included in this study. Additionally, NASA-SEDAC estimates lack a high temporal resolution; only yearly estimates were provided and seasonal variations in PM$_{2.5}$ have been reported [35]. Also, the effect of long-term exposure to PM$_{2.5}$ and other markers such as PM$_{10}$, SO$_2$ and NO$_2$ were not evaluated due missing information. Finally, TB reporting system (SIGTB) was established in 2015 and NASA-SEDAC only provided data until 2016, resulting in a short study period to observe the relation between PM$_{2.5}$ and TB incidence.

Conclusions
This study describes the co-occurrence of clusters of elevated concentrations of air pollution (measured by PM$_{2.5}$), poverty and high-risk areas of TB in Lima. These findings support previous studies and address the interplay of socio-economic and environmental drivers of TB incidence that will help to tailor TB control interventions. Further studies should be done to confirm these results at the individual level.

Abbreviations
TB: Tuberculosis; PM$_{10}$: Fine particulate matter; MoH: Ministry of Health; NASA: National Aeronautics and Space Administration; INEI: National Institute of Statistics and Informatics; REDATAM: Data Recuperation for Small Areas by Microcomputers; SEDAC: Socioeconomic Data and Applications Center; AOD: Aerosol Optical Depth; MODIS: Moderate Resolution Imaging Spectroradiometer; MISR: Multi-angle Imaging SpectroRadiometer; SeaWiFS: Sea-Viewing Wide Field-of-View Sensor; GWR: Geographically Weighted Regression; GEE: Google Earth Engine; ENAHO: National Survey of Households; PCA: Principal Component Analysis; LISA: Local Indicator of Spatial Association; GLMM: Generalized linear mixed model; INLA: Integrated Nested Laplace Approximation; MCMC: Markov Chain Monte Carlo; SO$_2$: Sulfur dioxide; NO$_2$: Nitrogen dioxide

Acknowledgments
We thank the Ministry of Health and the Directorate of Prevention and Control of TB (DPCTB) for providing such useful data to researchers.

Authors’ contributions
All authors contributed to study design. GC-E gathered the data and GC-E, KT-L and PV-G analyzed the data. GC-E, AS and CU-G contributed to drafting of the manuscript. All authors read and approved the final manuscript.

Additional file 1: Figure S1. Global Moran’s I of tuberculosis incidence, poverty, and PM$_{2.5}$ in Lima 2015–017.
Funding
Gabriel Carrasco-Escobar was supported by NHL/Fogarty International Center Global Infectious Diseases Training Program (D43 TW007120). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Availability of data and materials
All datasets are presented in the main paper and the supplementary materials. Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
1Health Innovation Lab, Instituto de Medicina Tropical Alexander von Humboldt, Universidad Peruana Cayetano Heredia, Lima, Peru. 2Division of Infectious Diseases, Department of Medicine, University of California San Diego, La Jolla, CA, USA. 3Instituto de Medicina Tropical Alexander von Humboldt, Universidad Peruana Cayetano Heredia, Lima, Peru. 4School of Public Health and Administration, Universidad Peruana Cayetano Heredia, Lima, Peru. 5School of Economics and Finance, Universidad del Pacifico, Lima, Peru. 6School of Medicine, Universidad Peruana Cayetano Heredia, Lima, Peru. 7IB Centre, London School of Hygiene and Tropical Medicine, London, UK.

Received: 11 October 2019 Accepted: 5 March 2020

Published online: 24 March 2020

References
1. World Health Organization. World Tuberculosis Report 2018. 2018.
2. Trauer JM, Dodd PJ, Gomes MGM, Gomez GB, Houben RMGJ, McBryde ES, et al. The importance of heterogeneity to the epidemiology of tuberculosis. Clin Infect Dis Off Publ Infect Dis Soc Am. 2019;69:159–66.
3. Tadesse S, Enguasselassie F, Hagos S. Spatial and space-time clustering of tuberculosis in Gurage zone, Southern Ethiopia. PLoS One. 2018;13:e0198353.
4. Dominkovic P, Granell C, Pérez-Navarro A, Casals M, Ocaña A, Cañay JA. Development of spatial density maps based on geoprocessing web services: application to tuberculosis incidence in Barcelona, Spain. Int J Health Geogr. 2011;10:62.
5. Shaweno D, Kamahar M, Alene KA, Ragonnet R, Clements AC, Trauer JM, et al. Methods used in the spatial analysis of tuberculosis epidemiology: a systematic review. BMC Med. 2018;16:193.
6. Wang X, Yin S, Li Y, Wang W, Du M, Guo W, et al. Spatiotemporal epidemiology of, and factors associated with, the tuberculosis prevalence in northern China, 2010-2014. BMC Infect Dis. 2019;19:365.
7. Smith G, Schoenbach VL, Richardson DB, Gammon MD. Particulate air pollution and susceptibility to the development of pulmonary tuberculosis disease in Northern China: an ecological study. Int J Environ Health Res. 2014;24:103–12.
8. Huff RD, Carlsten C, Hirota JA. An update on immunologic mechanisms in the respiratory mucosa in response to air pollutants. J Allergy Clin Immunol. 2019;143:1989–2001.
9. Shah L, Choi HW, Berrang-Ford L, Henostroza G, Krippa H, Zamudio C, et al. Geographic predictors of primary multidrug-resistant tuberculosis cases in an endemic area of Lima. Peru. 2014;19:1307–14.
10. Lin HH, Shin SS, Contreras C, Aserincas L, Piacore CJ, Cohen T. Use of spatial information to predict multidrug resistance in tuberculosis patients, Peru. Emerg Infect Dis. 2012;18:811–3.
11. Zelen JA, Murray MB, Becerra MC, Galea J, Lecca L, Calderon R, et al. Identifying hotspots of multidrug-resistant tuberculosis transmission using spatial and molecular genetic data. J Infect Dis. 2016;213:287–94.
12. van Donkelaar A, Martin RV, Brauer M, Hsu NC, Kahn RA, Levy RC, et al. Global estimates of fine particulate matter using a combined geophysical-statistical method with information from satellites. Environ Sci Technol. 2016;50:3762–3.
13. SEDAC. Global Annual PM2.5 Grids from MODIS, MISR and SeaWIFS Aerosol Optical Depth (AOD) with GWR, v1: Satellite-Derived Environmental Indicators. Available from: https://sedac.ciesin.columbia.edu/data/set/4984994-annual-gwr-pm25-modis-misr-seawifs-aod. Cited 2019 Dec 19.
14. van Donkelaar A, Martin RV, Brauer M, Hsu NC, Kahn RA, Levy RC, et al. Global Annual PM2.5 Grids from MODIS, MISR and SeaWIFS Aerosol Optical Depth (AOD) with GWR, 1998–2016. Palisades NY: NASA Socioeconomic Data and Applications Center (SEDAC). 2018.
15. van Donkelaar A, Martin RV, Brauer M, Hsu NC, Kahn RA, Levy RC, et al. Global estimates of fine particulate matter using a combined geophysical-statistical method with information from satellites, models, and monitors. Environ Sci Technol. 2016;50:3762–72.
16. INEI. Perú: Planos Estratificados de Lima Metropolitana a Nivel de Manzanas según Ingreso Per Cápita del hogar 2016. Available from: https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib1403/Index.html. Cited 2019 Dec 19.
17. Elbers C, Lanjouw JO, Lanjouw P. Micro-level estimation of poverty and inequality. Econometrica. 2003;71:355–64.
18. INEI Instituto Nacional de Estadistica e Informatica. Mapa de pobreza provincial y distrital 2013. 2015. Available from: http://repositorio.minedu.gob.pe/handle/123456789/3638. Cited 2019 Dec 19.
19. Dalenius T, Hodges JL Jr. Minimum variance stratification. J Am Stat Assoc. 1955;50:88–101.
20. Caldas de Castro M, Singer BH. Controlling the false discovery rate: a new application to account for multiple and dependent tests in local statistics of spatial association. Geogr Anal. 2006;38:180–208.
21. Griffith DA. The boundary value problem in spatial statistical analysis. J Reg Sci. 1983;23:377–87.
22. Biscarri O, Dzul-Manzanilla F, Gomez-Dantés H, Pavia-Ruz N, Hadidz TJ, Lenhart A, et al. Spatio-temporal coherence of dengue, chikungunya and Zika outbreaks in Merida, Mexico. PLoS Negl Trop Dis. 2018;12:e0006298.
23. Legendre P. Species associations: the Kendall coefficient of concordance revisited. J Agric Biol Environ Stat. 2005;10:226.
24. Besag J, Green P, Higdon D, Mengeser K. Bayesian computation and stochastic systems. Stat Sci. 1995;10:35–62.
25. Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. J R Stat Soc Ser B Methodol. 2009;71:319–92.
26. Bell ML, Cifuentes LA, Davis DK, Cushing E, Telles AG, Gouveia N. Environmental health indicators and a case study of air pollution in Latin American cities. Environ Res. 2011;111:57–67.
27. Orellano P, Quaranza N, Reynoso J, Balbi B, Vasquez J. Association of outdoor air pollution with the prevalence of asthma in children of Latin America and the Caribbean: a systematic review and meta-analysis. J Asthma Off J Assoc Care Asthma. 2018;55:S174–86.
28. Schraufnagel DE, Balmes JR, Cowl CT, De Matteis S, Jung S-H, Mortimer K, et al. Air pollution and noncommunicable diseases: a review by the forum of international respiratory societies’ environmental committee, part 1: The damaging effects of air pollution. Chest. 2019;155:409–16.
29. Iribonque O, Carranza C, Sarkar S, Torres M, Choi HT, Nwoko J, et al. Urban air pollution particulates suppress human T-cell responses to mycobacterium tuberculosis. Int J Environ Res Public Health. 2019;16:21.
30. Popovic I, Soares Magalhaes RJ, Ge E, Marks GB, Dong G-H, Wei X, et al. A systematic literature review and critical appraisal of epidemiological studies on outdoor air pollution and tuberculosis outcomes. Environ Res. 2019;170:33–45.
31. Zhu S, Xia L, Wu J, Chen S, Chen F, Zeng F, et al. Ambient air pollutants are associated with newly diagnosed tuberculosis: a time-series study in Chengdu, China. Sci Total Environ. 2018;631–632:457–55.
32. Lai TC, Chiang CY, Wu CF, Yang SL, Liu DP, Chan CC, et al. Ambient air pollution particulates suppress human T-cell responses to mycobacterium tuberculosis. Int J Environ Res Public Health. 2019;16:21.
33. Cesari D, De Benedetto GE, Bonasoni P, Busetto M, Dinoi A, Merico E, et al. Seasonal variability of PM2.5 and PM10 composition and sources in an urban background site in southern Italy. Sci Total Environ. 2018;612:202–13.