The association of commuting patterns and the incidence of Tuberculosis in Lima, Peru: A network analysis

Anónimo

Abstract Background

Peru is one of the countries with the highest burden of Tuberculosis (TB) in the region of the Americas with 55 % of accumulated cases between Brazil, Peru and Mexico. In 2017, the TB incidence rate was the second highest after Haiti with an incidence of 135.2 cases per 100,000 inhabitants 1. Lima Metropolitan Area and Callao reported 64 % and 45 % of TB cases and deaths in the country, respectively 2. Social determinants are known to drive the spatial distribution of TB 3. Human mobility, for instance, is a key factor in infectious disease epidemiology and is often taken into account when studying diseases like malaria, dengue, and chikungunya 4–7. Commuting for various purposes, driven by population needs, contributes significantly to this mobility 8. In Lima, for instance, one-third of the population resides in peripheral areas and relies on public transportation for commuting. Unfortunately, these vehicles often operate at double their average capacity, leading to an increased risk of TB infection 9,10. The aim of this study is to further assess the association between commuting patterns and the incidence of TB at a district level in the Province of Lima.

Methodology

District-to-district commuting flow data was extracted from the 2017 Peru Census for commuter students, commuter workers, commuter students and/or workers. Each commuting flow dataset was represented as a graph and analyzed using network analysis techniques in R version 4.3.0. First, we did a subregional analysis to characterize the subregions according to their share of the total cumulative cases, their cumulative incidence rate (CIR), and the destination and origin of the people who commuted out and into them. Then, the entire networks at the district level were characterized using structural properties of graphs and several centrality measures were computed for the districts. We evaluated the association between the centrality measures and the CIR doing a descriptive correlation analysis. Finally, we built generalized linear and additive models to model the incidence rate using monetary and non-monetary poverty indices and pollutant concentration measurements as predictors, and we individually assessed whether the addition of one of the centrality measures that were most correlated with incidence improved the fit of the models.

Results

We found that the peripheral subregions of Lima (North, East and South) shared most of the cases of TB (83.5 %) and had the highest CIR, in comparison with Lima Center. Concurrently, these subregions had the highest percentage of people that commute to other subregions and within the same subregion to study or work or both. Analyzing the properties of the whole mobilization networks, we found that the network for commuting to work was denser (98 %) than the one to study (88 %). That means that commuting to work generates more mobilization between districts than commuting to study. Looking at the properties of the districts in the mobilization networks, with the commuting to study network we found a strong positive linear correlation between the intra-district flow rate and the CIR (Pearson corr. coeff. = 0.55), and a strong negative linear correlation between the out-strength centrality (out-district flow rate) and the CIR (Pearson corr. coeff. = -0.58). Using the commuting to work network, we found a strong positive monotonic correlation between out-closeness centrality and the CIR (Spearman corr. coeff. = 0.52). Using the commuting to study or work or both network, we found correlations similar to the ones found with the commuting to study network. In addition, we found that the generalized additive models fitted the data significantly better than the generalized lineal models, showing that the CIR of TB may be better modeled using non-linear terms of socio-demographic, pollution and mobilization predictors. The best base GLM included the percentage of people in monetary poverty and the concentration on nitrogen dioxide (NO_2). Using the centrality measures computed from the commuting to study network, we found that adding the in-degree centrality measure improved marginally the AIC of the base model by 0.2 %. On the other hand, the best base GAM included the percentage of people living with at least one unsatisfied basic need and the concentration of PM2.5, and that adding the in-strength centrality measure from the commuting to study network improved considerably the AIC of the base model by 18 %.

Conclusions

Palabras clave: Tuberculosis - Commuting patterns - Network analysis