Abstract

Setting: The situation of tuberculosis in the States and Federal District of Brazil.

Objective: To analyse the incidence of tuberculosis in the States, the Federal District and the five regions of Brazil.

Design: Transversal study, composed of diagnosed cases of tuberculosis in the period 2003-2012, in both sexes among residents of the States, recorded in the epidemiological report by the Ministry of Health. The results were obtained using the R-3.2.3 program.

Results: Amazonas and Rio de Janeiro presented the highest incidence of tuberculosis, in the country, and hence in the North and Southeast regions, respectively. The Federal District and Goiás were responsible for the low incidence in the Centre-West region. Healthcare is decentralised in the States and in practice tuberculosis remains inadequately controlled. The incidence decreased in most States, with minor variations, between 2004 and 2011, but there were significant differences among States and regions, which highlights the need to reassess policies for the control of tuberculosis in accordance with State and regional needs.

Conclusion: Detailed knowledge of the characteristics of tuberculosis patients and of levels of public health education is of vital importance to public health in Brazil, providing essential information for health policies to be directed toward effective control of the disease. Our study considers the context in the different States, taking into account the patients’ age and level of education, seeking to facilitate the design of programmes for the eradication of tuberculosis focused on the age groups most likely to suffer the onset of this disease.
Introduction

Tuberculosis (TB) has been a priority issue for healthcare authorities since 1930. The disease progresses slowly, but is ultimately fatal. Nevertheless, for a long time it was not subjected to serious examination, and to this day, it remains a serious public health problem in Brazil and elsewhere. The incidence of TB is associated with the social conditions and overall health status of the population. Brazil accounts for 80% of the nine million new cases identified each year in developing countries [1].

In 1990, the morbidity from TB in Brazil was 51.8 new cases per 100,000 inhabitants. By 2011, this rate had decreased to 38.4, a reduction of 25.9% in two decades [2]. TB is defined as a contagious disease caused by the Koch bacillus, which primarily affects the lungs. When an individual has poor health and inadequate nutrition, the bacteria have optimum conditions in which to provoke the disease.

According to the World Health Organization [3], TB is present worldwide, and the 2014 Global Report on TB presented data from 202 countries. According to this report, the total number of new TB cases and deaths from this cause was higher in 2013 than in previous years. In 2014, 9.6 million people were infected with TB and about 1.5 million died from it. Over 95% of these deaths took place in middle and low income countries. TB is also one of the five leading causes of death among women aged 15-44 years. An estimated one million children suffer from TB and 140,000 children will die because of the disease. According to the Brazilian Information System for Reported Diseases (SINAN), 254,230 cases of TB were reported nationwide in the period 2011-2013, hence the importance of studying the TB situation in the country. The aim of this study is to identify and analyse variations in the incidence of TB in Brazil, in the period 2011-2013. Specifically, we compare the variability between the mean incidences of TB (variability due to the factor) with the variability remaining within each level (residual variability). The probability $p$ is calculated to determine whether there exists a significant effect, and thus we analyse variations within and among groups of States, in all five regions of Brazil. We note that very few studies have been conducted specifically to determine associations regarding the incidence of TB in the States and the Federal District. The Ministry of Health publishes data on the annual incidence and prevalence of TB in Brazil, but does not conduct more detailed studies on this subject. The States provide the municipalities with technical assistance and human resources, and also evaluate healthcare actions and distribute information. The States’ responsibilities in terms of TB prevention, control, treatment and monitoring requires them to have epidemiological knowledge about its incidence, mortality and related socioeconomic factors. Accordingly, detailed knowledge of inter and intra-State variations in the incidence of TB will contribute to the effectiveness of actions to monitor, prevent and control the disease in each State. Furthermore, this knowledge will enable proper use to be made of public healthcare resources and policies.

Method

This paper presents a descriptive study, combined with a quantitative approach. We analyse the total numbers of new cases of all types of TB during the period 2003-2012, for each of the States, as reported in the 2014 Epidemiological Report of the Ministry of Health, based on data supplied by SINAN. The variables included are the incidence of TB (as an epidemiological indicator) and demographic data for the States and regions. The 26 Brazilian States, together with the Federal District are described as follows: Acre (AC), Alagoas (AL), Amapá (AP), Amazonas (AM), Bahia (BA), Ceará (CE), Espírito Santo (ES), Goiás (GO), Maranhão (MA), Mato Grosso (MG), Mato Grosso do Sul (MS), Minas Gerais (MG), Pará (PA), Paraíba (PB), Paraná (PR), Pernambuco (PE), Piauí (PI), Rio de Janeiro (RJ), Rio Grande do Norte.
(RN), Rio Grande do Sul (RS), Rondônia (RO), Roraima (RR), Santa Catarina (SC), Sao Paulo (SP), Sergipe (SE), Tocantins (TO) and Distrito Federal (DF).

Data on the incidence of TB in Brazil for the period 2001-2010, for the regions and States of residence, were obtained from the new cases notified, confirmed and reported by SINAN [4]. The latter information system can also be used at a more peripheral administrative level, i.e., at that of health units, following the decentralisation guidelines issued by the Unified Health System (SUS). If the health unit in a given municipality is not equipped with the necessary computers, the SINAN data can be updated at municipal or regional offices and/or at the State Health Secretariat [4].

The variance in a set of observations can be analysed by separating it into two or more parts associated with independent sources of variation, such as treatment, design grouping and error. The total variation is expressed as the sum of squares of deviations of observations from the overall mean (sometimes called the grand mean). The sums of squares associated with the specified sources of variation are divided by their degrees of freedom to derive the mean squares, which are estimates of the variance of the observations. The extent to which these estimates differ is subsequently used to assess whether sources such as treatments have significantly different effects [5].

Analysis of the variance is one of the methods most commonly used to test the homogeneity of variables [6]. For this purpose, we used R-3.2.3 statistical software, and specifically its granova.1w and granova.2w functions [7] to obtain a graphic representation of one and two factors, respectively. The graphic elements are spaced such that the average vector or estimates of the effects of the factor on the dependent variable form a straight line (shown in blue). The top of the chart shows the labels and the number of cases, and at the bottom, the deviation of the mean value from the overall mean (or the estimate of this effect). The mean value for each level is shown on the right. For each factor level, the residual variability is shown on the vertical axis. This representation of the data makes it evident that the distribution is not skewed, with outliers and that the variance is homogeneous (i.e., presenting a similar degree of dispersion throughout the chart), for the different factor levels in the States of the Northeast. The chart enables us to compare the factor variability (the variability between the average values of the levels), i.e., the variability between groups, represented by the area of the red square, with the residual variability, that which remains within each level, i.e., the variability between groups as represented by the area of the blue square.

The F statistic is the ratio of the variance among the sample means and the variance among individuals within the samples. It is used to test the null hypothesis that all the populations of the States have the same mean. Then, under the H0 hypothesis, the F statistic has an F distribution with G-1 degrees of freedom and N-G, where G is the number of States and N is the total number for incidence of TB in the Federative Units. The dependent variable is the number of cases of TB incidence and the factors, in each case, may be years, States or regions.

The residual is the difference between the observed and the predicted values. In ANOVA, it is designated as “residual means error variance or within-groups variance”. In the chart, the plot of residuals versus fitted values shows that the dispersion was similar throughout the fitted values. Although there was a slight deviation at the top of the Normal Q-Q curve, the chart presents no problem of heterogeneity.

In the “mvoutlier package, multivariate outlier detection based on robust methods, the “aq.plot” function detects outliers in the data and presents a list of possible points. In addition, the distribution function of chisq is plotted, together with two vertical lines corresponding to the chisq-quantile specified in the argument list (the default value is 0.975) and the fitted quantile. Three additional charts are
also created; the first shows the data, the second shows the outliers detected by the specified quantile of the chisq distribution, and the third shows the outliers detected by the fitted quantile [10].

Results and Discussion
Figure 1 also shows the incidence data, this time using a box-and-whisker plot to summarise the distribution of TB incidence with respect to years. In our analysis of the annual impact, the box-plot chart reveals a decreasing trend until 2007; then, from 2008 until 2012 the incidence of TB appears to stabilise. The top points outside the box represent the outliers for each year. In 2005, there were hardly any low outliers.

Figure 2 shows the dispersion of each Federative Unit, in the respective geographic regions. The incidence of TB in some communities in RJ is four to five times above or below the state average. In Rocinha in RJ, in 2010, there were just 386.7 cases per 100,000 inhabitants. This reinforces the principle that the control of TB depends on eliminating the inequalities arising from the social conditions in which individuals live and work. There exists a situation of extreme social vulnerability in the most affected communities, especially those where drug abuse is rife. In consequence, these areas do not receive adequate public health actions to prevent the disease [8].

Figure 2 also shows that AM, in the North region, presents a different pattern of TB incidence, that TO (with slight variations) has an incidence of below 20, and that RR has two outliers. In the Centre-West region, MS and MT have higher rates than GO and DF, while AM and RJ present higher mean rates than the others, with 67.0 and 74.5 respectively. In the Northeast, there was a reasonable uniformity among the States.

The slope of the fitting plane shows that the incidence of TB increased over time and that there were growing differences among the States. In each combination of factors, the residual variability is reasonably symmetric and balanced, which may indicate that the normality assumption required by the ANOVA method is being met. No values were found that distorted the analysis results.

Our analysis of the levels of the factor (region) for the study period (Figure 3) showed that the
Centre-West region (CO) had the lowest incidence of TB, followed by the South. The Northeast (NE) presented smaller variations among the States than elsewhere. The regions that presented the widest variations in incidence were the North and the South, with strong contributions from the States of AM and RJ, respectively.

The States RJ and DF have a higher incidence of TB, with a stronger influence of outliers, as is apparent in the Cook distance, i.e., residuals versus fitted values and scale-location (Figure 5). In the diagnostic charts, there are no problematic values.

Figure 4 shows the analysis of variance with one factor (the States). The dependent variable is the incidence of TB and the groups are the States. This chart shows the different levels of the factors, according to the mean values for TB cases (represented by red triangles), from left to right. For each factor level, there is little residual variability on the vertical axis, showing there is no asymmetric distribution or outliers, and that the variances are homogeneous.

The sequence of States at the top of Figure 4 is DF, TO, GO, MG, PR, SC, SE, PI, PB, RN, RO, RR, AP, MS, ES, MT, MA, AL, SP, BA, RS, AC, CE, PA, PE and AM. The States RO, RR, AP, MS, ES, MT, MA, AL and SP, at the top of the blue square, reveal a variation within the group (MS-within). In other words, the larger the red square, in comparison with the blue one, the stronger the effect of the factors (States and years) on the dependent variable (incidence of TB).

The total, between-treatments and error degrees of freedom are 269, 26 and 243, respectively. The variance ratio of 138.43 is determined from the F-distribution with 26 and 234 degrees of freedom, with a p-value of 0.000. The minimum values for significance at the 5% and 1% levels are 1.54 and 1.83, respectively.

The ratio $\text{MS. between}/\text{MS.within} = \frac{1931.15}{13.95} = 138.43 > 1.54$, and therefore there is no statistical evidence of equality of incidence of TB among the States. In graphic terms, the ratio between the mean values for the two squares is the $F$ value obtained, with $p < 0.001$, which leads us to conclude that there is a significant effect among the States. The larger the red square, in comparison with the blue one, the stronger the effect of the factor (States) on the dependent variable (incidence of TB).

The groups of States present large variations within each region, showing that some States invest more than others in combating TB. There are
large regional and State differences, especially in the North (AM) and the Southeast (RJ). RJ has a high concentration of the population living in economically-deprived neighbourhoods, which accounts for the high incidence of TB in this State. RO, RR, AP, MS, ES, MT, MA, AL and SP present large variations within each State and with respect to all others except RJ and AM.

**Figure 6** shows the outlier values shown with the quantiles that are summarised in **Figure 5**. It is clear that all the points above the 97.5% quantile and the fitted quantile correspond to the State of RJ. In other words, there is a wide variation in the incidence recorded in this State compared to the other without Brazil. According to [12], in RJ the limitations suffered by the healthcare services aggravate living conditions, impeding TB treatment, as is apparent in the differences observed between this State and the others.

**Conclusions**

The incidence of TB was highest in the years 2003, 2004 and 2011. The highest average numbers of TB cases are concentrated in the States of AM (North region) and RJ (Southeast region). These two States, thus, contribute significantly to the high incidence levels in their regions. The mean incidence over the 10-year study period in AM (66.9) was 20.1 points higher than in the region as a whole (46.8). Similarly, the incidence in RJ was 80.5% above the mean level for the Southeast region. From these findings, we conclude that actions to control TB should be carried out in a decentralised way, among the States. Within municipalities, a level of administration that is readily accessible to the population, efforts should be made to minimise the barriers between the patient and the continuity of treatment, in order to reduce levels of treatment abandonment. The levels of incidence recorded in the States also demonstrate, quite clearly, that in many cases treatment is not performed correctly.

In almost all the States, the incidence varied widely. AM and RJ presented the highest incidence of TB nationwide and in their respective regions (North and Southeast). The Federal District and Goiás were responsible for the low incidence of TB in the Centre-West.

There are large regional differences in the incidence of TB and in the mortality due to this disease, with higher levels in the States with the highest prevalence of HIV infection, such as RJ, and in those with poor access to health services, such as AM.

This study furthers our understanding of the stability or variability of the incidence of new cases of TB in Brazil, taking into account its geographic organisation, facilitating the design of policies to enhance individual and collective health, focusing especially on the population sectors most affected by TB in this country.

Problems can be identified in the form of specific characteristics of population groups, together with processes of social reproduction. Thus, communities or “particular socio-spatial groups” are taking shape. Strategies for primary healthcare attention need to be expanded and improved, to constitute a set of social actions aimed at communities in need, and thus promote the quality of life of the population [11].
References


