Identification of micro-regions with under-reported tuberculosis cases in Brazil, 2012-2014*

Gabriela Drummond Marques da Silva[1] – ⓘ orcid.org/0000-0002-1145-3940
Elisabeth Carmen Duarte[2] – ⓘ orcid.org/0000-0001-9148-5063
Oswaldo Gonçalves Cruz[3] – ⓘ orcid.org/0000-0002-3289-3195
Leila Posenato Garcia[4] – ⓘ orcid.org/0000-0003-1146-2641

[1]Universidade de Brasília, Programa de Pós-Graduação em Saúde Coletiva, Brasília, DF, Brazil
[2]Universidade de Brasília, Faculdade de Medicina, Brasília, DF, Brazil
[3]Fundação Oswaldo Cruz, Programa de Computação Científica, Rio de Janeiro, RJ, Brazil
[4]Instituto de Pesquisa Econômica Aplicada, Brasília, DF, Brazil

Abstract

Objective: to identify Brazilian micro-regions with under-reported tuberculosis cases, from 2012 to 2014. **Methods**: this was an ecological study using data from the Notifiable Health Conditions Information System (SINAN). The indirect estimator of the proportion of notified cases (EIPCN) was calculated as the mean between the proportion of cases that adhered to treatment and the proportion of those who underwent sputum smear microscopy. Negative Binomial Regression was used to investigate evidence of under-reporting in the micro-regions selected through EIPCN. **Results**: under-reporting was suspected in 89 (17.5%) micro-regions with EIPCN below 83%. The EIPCN rate ratio in the regression model was 0.996 (95%CI 0.988;1.003) considering all the data and equal to 0.987 (95%CI 0.974;0.999) excluding the 89 micro-regions with suspected under-reporting. **Conclusion**: evidence of tuberculosis case under-reporting was found in micro-regions where basic indicators of surveillance system quality had poor performance.

Keywords: Tuberculosis; Incidence; Epidemiological Monitoring; Ecological Studies; Quality Indicators, Health Care.

Correspondence:
Gabriela Drummond Marques da Silva – Universidade de Brasília, Programa de Pós-Graduação em Saúde Coletiva, Brasília, DF, Brazil. Postcode: 70910-900
E-mail: gabidrumm@gmail.com

Epidemiol. Serv. Saude, Brasília, 29(1):e2018485,2020

1

Introduction

Among communicable diseases, tuberculosis (TB) is the one that most kills worldwide. In 2017, 10 million new cases and 1.3 million deaths were recorded.[1] In Brazil, in the same year, 69,569 new TB cases were notified, this being equivalent to an incidence rate of 33.5 cases/100,000 inhabitants.[2] In 2016, there were 4,426 TB deaths in Brazil, resulting in a mortality rate equal to 2.1 deaths/100,000 inhab.[2]

According to the World Health Organization (WHO), globally some 3.6 million cases are not notified by the TB surveillance system.[1] Under-reporting occurs when part of the cases is not diagnosed by health centers and/or when part of diagnosed cases is not notified on the TB information system. In both situations, epidemiological surveillance of TB is jeopardized, as is the targeting of resources and planning of actions to address the disease.[3]

> *Increasing detection and, consequently, case notification, is one of the possible strategies for reducing the TB mortality rate in Brazil.*

In Brazil, the estimated proportion of notified cases was 87% in 2017.[1] High health professional turnover and delays in health service provision are some of the challenges to improved case notification in the country.[4] In addition, TB is little known by the population, including by health professionals working in primary care.[5] Lack of information about TB is known to be associated with case under-reporting.[6]

Under 15 year-olds comprise the age group with the lowest proportion of notified cases in Brazil.[1] Among the 239,000 deaths estimated for this age group, over 90% occurred in cases that did not have access to treatment.[7] As such, increasing detection and, consequently, case notification, is one of the possible strategies for reducing the TB mortality rate in Brazil.

Although WHO has estimates of the proportion of cases notified nationally in Brazil, including by age group, it does not measure this on the sub-national level. No studies having this objective for the entire country were found in the literature. The study closest to this theme evaluates the proportion of cases notified in the country, among those diagnosed by the TB surveillance system, using probabilistic database linkage.[8]

In view of this gap, the objective of this study is to identify Brazilian micro-regions with under-reported TB cases between 2012 and 2014.

Methods

Study design

This was a cross-sectional ecological study having as its units of analysis the 508 Brazilian micro-regions with a good or regular information system and at least 6 TB cases notified between 2012 and 2014. Micro-region information systems were considered to be good or regular when mean completeness was equal to or greater than 70% in relation to priority information system variables, selected by means of consulting specialists on the subject. The 508 areas account for 91.0% of Brazil's total micro-regions, and were used with the aim of minimizing bias and random fluctuation of the TB surveillance system quality indicators.[9]

New cases were considered to be those input onto the information system as "new case", "does not know" or not known. Cases closed due to change of diagnosis were excluded from the study. The TB incidence rate was calculated as the ratio between the number of resident cases diagnosed between 2012 and 2014, and the total population resident in the micro-region in the three years analyzed, multiplied by 100,000.

Data source

New cases and surveillance quality indicators were calculated using Notifiable Health Conditions Information System (SINAN) data for the period 2012-2014. In addition, the population estimates for 2012-2014 made by the Brazilian Institute of Geography and Statistics (IBGE) for the Federal Audit Court (TCU), were used as the source of data for the number of people resident in the micro-regions.[10]

The control variables were obtained from the 2010 IBGE Demographic Census and from the Justice Ministry's 2014 National Prison Information Survey (Infopen).[11,12]

Surveillance system quality indicators

We analyzed the indicators proposed by Silva et al. (2017).[9] Given that completeness is the proportion of cases with information provided for a given variable, the completeness of the following variables was measured: number of contacts examined, data treatment started, case closure status, sputum smear microscopy

at the second and sixth months of treatment. Mean completeness was also calculated for these five variables. Level of completeness was categorized as good (≥90%), regular (between 70%, inclusive, and <90%) and poor (<70%).

TB surveillance system timeliness was measured using the following indicators: notification timeliness (proportion of cases in which the difference between notification data and diagnosis date was less than or equal to 7 days); and treatment timeliness (proportion of cases starting treatment on the day of diagnosis). Timeliness was classified as good, regular or poor, following the same cut-off points used for completeness.

Moreover, indicators were considered to be "invalid" in micro-regions in which at least 15% of cases had incomplete or inconsistent data. Although there is no golden rule in the literature about this cut-off point, studies indicate that 5%, 10% and 20% incomplete data can result in biased inferences, depending on the generating mechanism in which information is missing.[13,14] In this study the probability of loss of data was taken to be non-random, and the 15% cut-off point suggested by Harrel (2001) was adopted in order to identify variables needing multiple imputation, i.e. requiring a more sophisticated imputation method for inferences to be made.[15]

Health professional acceptability in relation to the recommendations of the National Tuberculosis Control Program (PNCT) was measured using the following indicator: percentage of cases having sputum smear microscopy (proportion of new pulmonary cases having sputum smear microscopy); percentage of cases adhered to treatment (proportion of pulmonary cases not abandoning TB treatment, among cases not closed because of treatment failure, change of regimen and drug-resistant TB); and percentage of cases tested for HIV (proportion of cases tested for human immunodeficiency virus (HIV) infection, taking tests with results "in progress" as not having been performed).

The cut-off points for classifying acceptability as good, regular or poor were different for each indicator and were defined according to official document recommendations and recommendations made by TB surveillance specialists.[9] Acceptability was good in micro-regions in which at least 90% of cases had sputum smear microscopy, 95% of cases not lost to treatment and 85% tested for HIV. On the other hand, acceptability was poor in micro-regions where less than 80% of cases had sputum smear microscopy, 90% of cases not lost to treatment and 70% tested for HIV. The acceptability calculation was considered to be "invalid" in micro-regions in which 15% or more of cases had missing or unknown information in the variables needed to calculate acceptability.

Control variables

The control variables used in this study were: region (North, Northeast, Midwest, Southeast and South); logarithm of the population resident in the micro-region from 2012 to 2014; population deprived of liberty/10,000 inhab. in June 2014; urbanization rate (proportion of people resident in the urban area of the micro-region); unemployment rate (proportion of unemployed people, among those aged 16 years and over); and proportion of people living in households with more than two persons per bedroom in 2010.

Statistical analysis of the data

The cluster analysis of the acceptability and timeliness indicators was performed using a hybrid algorithm comprised of the k-means and hierarchical methods.[16,17] The number of groups was defined with the aid of a dendogram built using the Ward method, taking the Euclidian distance as the measurement of dissimilarity (available in the supplementary material to this article). The groups were compared according to the mean ($\bar{x}$) and standard deviation (s) of the timeliness and acceptability indicators.

The indirect estimator of the proportion of notified cases (EIPCN) was equal to the mean of the timeliness and/or acceptability indicators in which micro-regions with poor surveillance systems had low mean performance and great homogeneity, i.e. low standard deviation. The EIPCN cut-off point for defining micro-regions with suspected under-reporting was the point of the ROC curve with greatest sensitivity and specificity for representing micro-regions with poor TB surveillance.

Distribution of categorical indicators of surveillance quality in micro-regions with suspected under-reporting was compared with distribution in the other micro-regions. The difference between the groups was assessed using Person's chi-square test. In tables in which the expected number of cases was less than five, Fisher's exact test was used. In addition, a thematic EIPCN map was produced, in relation to the cut-off point found, with

the aim of visualizing the location of micro-regions with suspected under-reporting in the Brazilian territory.

A scatter plot was generated in relation to EIPCN and the TB incidence rate. The smoothing curve was adjusted using the generalized additive model of TB incidence per 100,000 inhab. in function of EIPCN, using the cubic spline as the smoothing function. This curve was represented on the scatter plot in order to assess its compatibility with the expected theoretical association. The assumption was made that empirical association deviations, in relation to expected theoretical association, would be the consequence of under-reported TB cases in the micro-regions.

Finally, the negative binomial regression model was adjusted, taking the number of TB cases diagnosed between 2012 and 2014 as the dependent variable, EIPCN as the independent variable and the logarithm of the population resident in the micro-region from 2012 to 2014 as the offset. We opted for this model owing to the inadequacy of the Poisson regression model, initially tested, in relation to the super-dispersion of the dependent variable data.[18]

The regression model, having EIPCN as the independent variable, was adjusted, both by excluding micro-regions with suspected under-reporting and also by considering these micro-regions. This model's rate ratio (RR) was compared with the rate ratio expected according to theoretical association between the variables.

A 5% significance level was used for all the statistical tests. The analyses were performed with the aid of R[19] version 3.5.1, using the Facto extra (https://CRAN.R-project.org/package=factoextra), mgcv (https://CRAN.R-project.org/package=mgcv) and MASS (https://CRAN.R-project.org/package=MASS) packages.

Ethical considerations

Ethical principles of research involving human beings were respected during this study, as standardized by National Health Council (CNS) Resolution No. 466/2012. This study was not submitted to a Research Ethics Committee as it used public domain secondary data, as per CNS Resolution No. 510/2016.

Results

The 508 micro-regions studied were divided into four groups. Group 1 (n=73) had the lowest mean proportion of cases having sputum smear microscopy ($\bar{x}$ =75.7%, s =9.5) and cases adhered to treatment ($\bar{x}$ =78.0%, s =7.9). The Group 2 micro-regions (187), in turn, had the highest mean proportion of cases having sputum smear microscopy ($\bar{x}$ =91.2%, s =5.3), adherence to treatment ($\bar{x}$ =91.7%, s =4.7) and tested for HIV($\bar{x}$ =85.9%, s =9.0) (Table 1).

Groups 3 (n=164) and 4 (n=84) had acceptability indicators with similar performance. In both groups, the mean percentage of cases having sputum smear microscopy was 88.2%, while for cases adhered to treatment it was 88.7%. Surveillance system timeliness, however, was lower in Group 4, both in relation to mean notification timeliness ($\bar{x}$ =57.6%, s =10.2), and also in relation to mean treatment timeliness ($\bar{x}$ =24.6%, s =10.8) (Table 1).

The Group 1 micro-regions were considered to have a poor surveillance system. Based on this, the conclusion was reached that a poor surveillance system can be characterized, in particular, by the low proportion of cases having sputum smear microscopy and adhering to TB treatment. Given that both of these correspond to acceptability indicators, their mean was called mean acceptability. This indicator was used as EIPCN due to its ability to distinguish between micro-regions with poor TB surveillance from the others.

The cut-off point for mean acceptability was 83%, as per the ROC curve. The sensitivity and specificity of this

Table 1 – Mean ($\bar{x}$) and standard deviation (s) of Brazilian micro-region tuberculosis surveillance system acceptability and timeliness indicators, 2012-2014

Indicator	Group							
	1 (n=73)		2 (n=187)		3 (n=164)		4 (n=84)	
	$\bar{x}$	s	$\bar{x}$	s	$\bar{x}$	s	$\bar{x}$	s
Notification timeliness	74.9	10.8	80.6	9.2	82.4	6.2	57.6	10.2
Treatment timeliness	51.0	19.2	63.8	17.5	42.1	16.5	24.6	10.8
Cases having sputum smear microscopy (%)	75.7	9.5	91.2	5.3	88.2	6.0	88.2	8.0
Cases adhered to treatment (%)	78.0	7.9	91.7	4.7	88.7	5.5	88.7	7.4
Cases tested for HIV (%)	60.0	19.3	85.9	9.0	54.6	14.5	58.1	18.4

Note: The most homogenous indicators (σ<10) are highlighted in gray.

point for identifying Group 1 micro-regions are equal to 98.6% and 97.0%, respectively. In all, 17.5% (n=89) of micro-regions had mean acceptability below 83%.

Distribution of surveillance quality indicators in the micro-regions with mean acceptability below 83.0%, in which under-reporting is suspected, was significantly different to distribution in micro-regions with mean acceptability great than or equal to this limit. Treatment timeliness distribution, however, was similar in both groups. This indicator had poor performance in around 80.5% of the micro-regions (Table 2).

Only 11.2% (n=10) of micro-regions with mean acceptability below 83% had good completeness for the number of contacts examined. Moreover, 62.9% (n=56) of them had good completeness for the case closure status variable, compared to 95.0% in the other micro-regions. Completeness of sputum smear microscopy in the second month was regular in the majority (n=51) of micro-regions with suspected case under-reporting, while completeness was good (n=246) in the majority of the other micro-regions. Almost a quarter (n=21) of micro-regions with mean acceptability below 83% had poor sputum smear microscopy completeness in the 6th month (Table 2).

Notification timeliness was regular or poor in 93.2% (n=83) of micro-regions with suspected TB case under-reporting. In addition, 74.2% (n=66) of these areas had a poor percentage of cases tested for HIV, while it was 48% (n=201) in the other micro-regions (Table 2).

Four (4.5%) micro-regions were identified as having mean acceptability below 83% in the Northern region, 33 (37.7%) in the Northeast, 21 (23.6%) in the Southeast, 17 (19.1%) in the South and 14 (15.7%) in the Midwest. These micro-regions are concentrated more in the states of Rio Grande do Sul (n=16, 18.0%), Minas Gerais (n=12, 13.5%), Bahia (n=11, 12.4%,), Mato Grosso (n=8, 9.0%), Rio de Janeiro (n=7, 7.9%) and Piauí (n=6, 6.7%) (Figure 1).

It is expected that the greater the mean acceptability of the TB surveillance system, the lower the mean TB incidence rate will be. In the micro-regions with mean acceptability greater than or equal to 83.0%, the empirical relationship between these indicators is linear and coherent with the theoretical relationship expected. In the other micro-regions, however, an increase can be seen in the mean TB incidence rate as a result of the increase in mean acceptability of the surveillance system (Figure 2).

Taking the data as a whole, the TB incidence rate is not significantly different in micro-regions with distinct mean acceptability, both in the crude negative binomial regression model (RR=0.996 - 95%CI 0.988;1.003), and in the adjusted model (RR=0.997- 95%CI 0.991;1.003). However, the rate ratio becomes less than one when the micro-regions with suspected under-reporting are excluded from the analysis (95%CI 0.974; 0.999 in the crude model and 95%CI 0.975; 0.996 in the adjusted model). In this case, when mean micro-region acceptability increases by 1 percentage point, on average the TB incidence rate decreases by 1.3% in the crude model and by 1.5% in the adjusted model) (Table 3).

Therefore, leaving the micro-regions with mean acceptability below 83% to one side, it can be concluded that the better the TB surveillance system is structured in terms of acceptability, the lower the expected incidence rate in the region will be. Moreover, the fact that this association was not significant for the complete database suggests the existence of TB case under-reporting in micro-regions with mean acceptability below 83%.

Discussion

Evidence of TB case under-reporting can be seen in micro-regions with a low proportion of new pulmonary cases having sputum smear microscopy and/or a low proportion of pulmonary cases that adhered to TB treatment. These micro-regions were mainly located in the country's Northeast and Southeast regions, while they were concentrated in the states of Rio Grande do Sul, Minas Gerais and Bahia.

Pelissari et al. identified that requesting sputum smear microscopy was positively associated with TB case detection in Brazilian municipalities with primary health care teams evaluated by the second cycle of the Program for Improving Primary Care Access and Quality (PMAQ-AB).[20] Moreover, a study conducted in Rio de Janeiro identified that death notification was higher among TB cases with bacteriological confirmation.[21] Both studies provide evidence that mean acceptability can be used as a proxy for TB case notification.

TB under-reporting in a given region can occur due to problems in three different stages of case care: access to the health system, TB diagnosis and notification on the information system.[22] The EIPCN proposed in this study has the potential to identify under-reporting happening in any of these stages. If, for example, this phenomenon
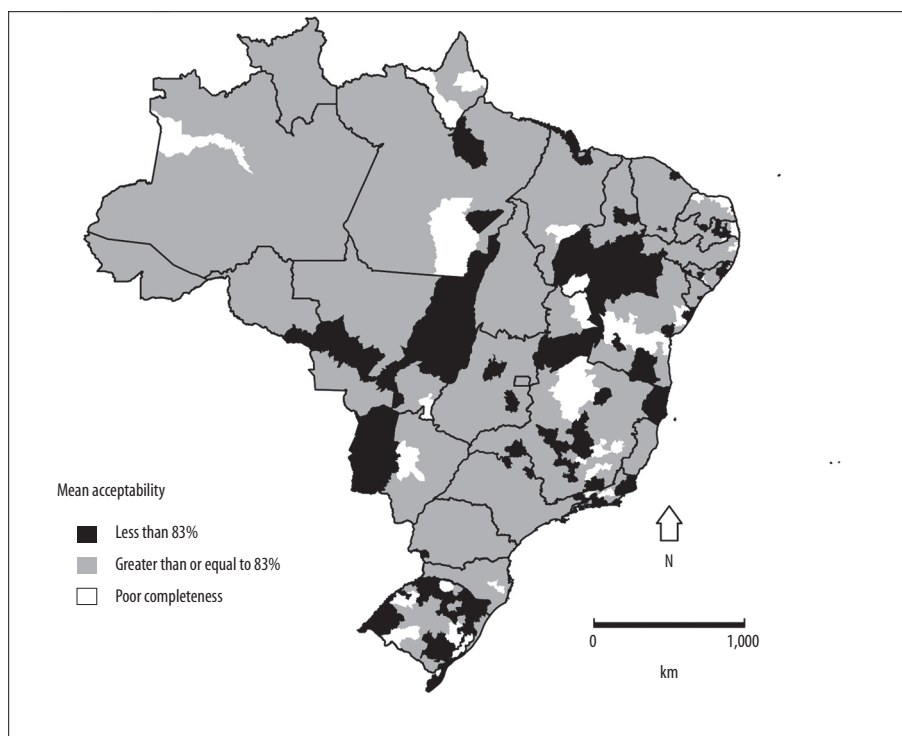
Note: Mean acceptability is equal to the meanian between proportion of tuberculosis cases having sputum smear microscopy and proportion of cases adhered to TB treatment, calculated only for micro-regions that have good or regular mean information system completeness (≥70%).

Figure 1 – Thematic map of tuberculosis surveillance system mean acceptability, Brazil, 2012 -2014

occurs owing to problems with sick people accessing health services, whether because of geographical, financial or cultural barriers, it is possible for this reality to be reflected in low adherence to TB treatment.[23,24]

If, on the other hand, under-reporting is caused by underdiagnosis, the proportion of cases having sputum smear microscopy may be useful for identifying this. Shortcomings with this indicator may also suggest low levels of sputum smear microscopy among symptomatic respiratory cases, given that both groups are submitted to the same surveillance system.[25]

Finally, Selig et al. (2004)[21] assume that there is association between bacteriological confirmation and TB case notification, given that in the case of more robust surveillance systems, health professionals are committed to fulfilling the main PNCT recommendations, and not just doing isolated activities. In this sense, mean acceptability is a good indirect estimator of the proportion of cases notified on the TB information system.

Oliveira et al. (2012)[8] used probabilistic linkage to measure the proportion of under-reported deaths on

the TB information system in Brazil in 2006. Given that under-reporting on the information system is one of the components of overall case under-reporting, the two measurements can be compared. They identified a high proportion of under-reporting on information systems in the Southeast and Northeast regions of Brazil, corroborating the results found in this study. On the other hand, the highest proportion of case under-reporting on the SINAN system was found in the Northern region, whereas our study did not find evidence of TB under-reporting for that region.

The divergence between the studies may be explained by the time difference in the data. Despite this, it is important to highlight that all areas with under-reporting were not necessarily identified by our study. Although evidence of this phenomenon exists in the areas selected, the possibility of its occurrence in other Brazilian micro-regions cannot be excluded.

The main assumption of this study is that TB case under-reporting exists in micro-regions with low quality surveillance systems and, for this reason, deviations

Table 2 – Description of tuberculosis surveillance system quality indicators, by cut-off point equal to 83% for mean acceptability[a] in the micro-regions, Brazil, 2012-2014

Indicator	Mean acceptability < 83%				Total		p-value
	No		Yes				
	n	%	n	%	n	%	
Completeness of number of contacts examined							
Good	188	44.9	10	11.2	198	39.0	<0.001[b]
Regular	170	40.6	42	47.2	212	41.7	
Poor	61	14.6	37	41.6	98	19.3	
Completeness data treatment started							
Good	400	95.5	80	89.9	480	94.5	0.081[c]
Regular	18	4.3	9	10.1	27	5.3	
Poor	1	0.2	-	-	1	0.2	
Completeness of case closure status							
Good	398	95.0	56	62.9	454	89.4	<0.001[c]
Regular	21	5.0	31	34.8	52	10.2	
Poor	-	-	2	2.2	2	0.4	
Completeness of sputum smear microscopy in the 2nd month							
Good	246	58.7	33	37.1	279	54.9	<0.001[b]
Regular	142	33.9	51	57.3	193	38.0	
Poor	31	7.4	5	5.6	36	7.1	
Completeness of sputum smear microscopy in the 6th month							
Good	164	39.1	22	24.7	186	36.6	0.003[b]
Regular	205	48.9	46	51.7	251	49.4	
Poor	50	11.9	21	23.6	71	14.0	
Mean Completeness							
Good	301	71.8	27	30.3	328	64.6	<0.001[b]
Regular	118	28.2	62	69.7	180	35.4	
Notification timeliness							
Invalid	4	1.0	1	1.1	5	1.0	0.047[c]
Good	48	11.5	5	5.6	53	10.4	
Regular	273	65.2	52	58.4	325	64.0	
Poor	94	22.4	31	34.8	125	24.6	
Treatment timeliness							
Invalid	8	1.9	4	4.5	12	2.4	0.395[c]
Good	9	2.1	1	1.1	10	2.0	
Regular	66	15.8	11	12.4	77	15.2	
Poor	336	80.2	73	82.0	409	80.5	
Cases that had sputum smear microscopy (%)							
Good	227	54.2	4	4.5	231	45.5	<0.001[b]
Regular	173	41.3	30	33.7	203	40.0	
Poor	19	4.5	55	61.8	74	14.6	
Cases that adhered to treatment (%)							
Invalid	4	1.0	20	22.5	24	4.7	<0.001[c]
Good	88	21.0	3	3.4	91	17.9	
Regular	134	32.0	2	2.2	136	26.8	
Poor	193	46.1	64	71.9	257	50.6	
Cases tested for HIV (%)							
Invalid	1	0.2	1	1.1	2	0.4	<0.001[c]
Good	121	28.9	8	9.0	129	25.4	
Regular	96	22.9	14	15.7	110	21.7	
Poor	201	48.0	66	74.2	267	52.6	
Total	**419**	**100.0**	**89**	**100.0**	**508**	**100.0**	

a) Mean acceptability is equal to the mean between proportion of tuberculosis cases having sputum smear microscopy and proportion of cases adhered to TB treatment.
b) Pearson's chi-square test.
c) Fisher's exact test.
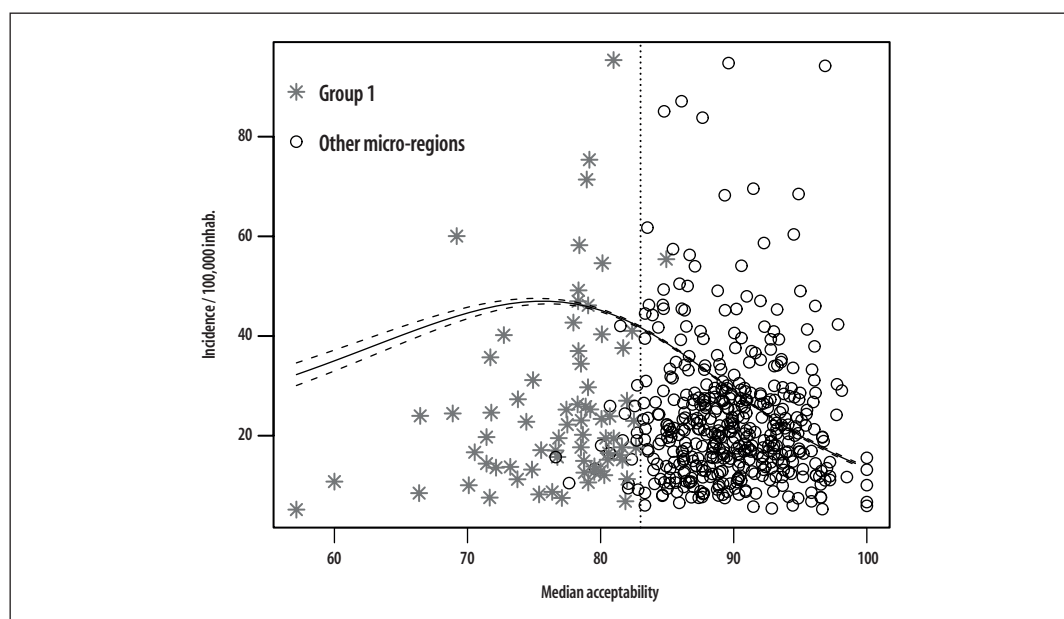
Note:
The vertical dotted line indicates mean acceptability equal to 83%.
Group 1: micro-regions with poor tuberculosis surveillance systems, selected by means of cluster analysis.
Other micro-regions: micro-regions with good or regular tuberculosis surveillance systems, according to cluster analysis.

Figure 2 – Smoothing curve (with 95% confidence interval) of tuberculosis surveillance system mean acceptability, according to tuberculosis incidence per 100,000 inhab. in Brazil's micro-regions, 2012-2014

Table 3 – Results of the regression model of tuberculosis incidence per 100,000 inhab. according to tuberculosis surveillance system mean acceptability, based on inclusion of micro-regions with mean acceptability below 83% in the database and use of control variables in the model, Brazil, 2012-2014

| Control variables | Micro-regions with acceptability <83% | | | |
| | Yes | | No | |
	Rate ratio	95%CI [a]	Rate ratio	95%CI [a]
Yes	0.997	0.991;1.003	0.985	0.975;0.996
No	0.996	0.988;1.003	0.987	0.974; 0.999

a) 95%CI : 95% confidence interval.

in empirical association between mean acceptability and TB incidence per 100,000 inhab., in relation to theoretical association, are generated due to this mechanism. However, it is also possible that the deviations found may have been caused by confounding factors, representing spurious associations. In order to minimize this problem, crude regression models and regression models adjusted by control variables were used.

A limitation of this study is the use of secondary data to evaluate TB case under-reporting. It was not possible to analyze 8.8% (n=49) of the Brazilian micro-regions because of low quality TB information systems. It is

recommended that these micro-regions be included in initiatives intended to enhance the TB surveillance system, so that they too are improved.

Another important limitation was the level of aggregation used for the analysis. The micro-regions are geographic units formed by a set of municipalities with independent Municipal TB Control Programs. As such, municipalities within the same micro-region may be heterogeneous with regard to surveillance system quality. Notwithstanding, it was considered that neighboring municipalities have greater likelihood of having similar surveillance systems than areas distant from each

other, given that they are subject to the same regional determinants. In addition, the level of aggregation chosen was useful for reducing random fluctuation of the quality indicators calculated at municipal level and caused as a result of a small number of notified cases.

Identification of areas with TB case notification problems on the sub-national level can be an expensive initiative if it is done by means of population surveys.[22]As such, this study proposed a simple method, based on secondary data, in order to identify such areas. The results found can be useful for targeting public policies aimed at improving case finding, diagnosis and notification. Moreover, the findings of this study can have an impact on reducing TB mortality, especially among children affected by the disease.

Authors' Contributions

Silva GDM took part in the conception of the article, data analysis and interpretation and wrote the manuscript. Duarte EC contributed to the conception of the article, data interpretation and critical review of the text. Cruz OG and Garcia LP planned the study design and took part in data interpretation and critical review of the manuscript. All the authors have approved the final version and are responsible for all aspects of the work, including the guarantee of its precision and integrity.

References

1. World Health Organization. Global tuberculosis report 2018 [Internet]. Geneva: World Health Organization; 2018 [cited 2018 Nov 10]. Available from: http://www.who.int/tb/publications/global_report/en/

2. Ministério da Saúde (BR). Brasil livre da tuberculose: evolução dos cenários epidemiológicos e operacionais da doença. Bol Epidemiol [Internet]. 2019 mar [citado 2020 jan 13];50(9):1-18. Disponível em: http://portalarquivos2.saude.gov.br/images/pdf/2019/marco/22/2019-009.pdf

3. Pinheiro RS, Andrade VL, Oliveira GP. Subnotificação da tuberculose no Sistema de Informação de Agravos de Notificação (SINAN): abandono primário de bacilíferos e captação de casos em outras fontes de informação usando linkage probabilístico. Cad Saúde Pública [Internet]. 2012 ago [citado 2020 jan 13];28(8):1559-68. Disponível em: http://www.scielo.br/pdf/csp/v28n8/14.pdf. doi: 10.1590/S0102-311X2012000800014

4. Reis SP, Harter J, Lima LM, Vieira DA, Palha PF, Gonzales RIC. Aspectos geográficos e organizacionais dos serviços de atenção primária à saúde na detecção de casos de tuberculose em Pelotas, Rio Grande do Sul, 2012. Epidemiol Serv Saúde [Internet]. 2017 jan-mar [citado 2020 jan 13];26(1):141-8. Disponível em: http://www.scielo.br/pdf/ress/v26n1/2237-9622-ress-26-01-00141.pdf. doi: 10.5123/s1679-49742017000100015

5. Ramdohr Sobrinho EC, Freitas KG, Figueiredo RM, Caliari JDS. A tuberculose na estratégia de saúde da família: o conhecimento dos agentes comunitários de saúde. Rev Eletr Enf [Internet]. 2013 abr-jun [citado 2020 jan 13];15(2):416-21. Disponível em: https://www.fen.ufg.br/revista/v15/n2/pdf/v15n2a14.pdf. doi: 10.5216/ree.v15i2.16982

6. Stop TB Partnership. Choosing an active case finding intervention. In: Improving tuberculosis case detection: a compendium of TB REACH case studies, lessons learnt and a monitoring and evaluation framework [Internet]. Geneva: Stop TB Partnership; 2014 [cited 2020 Jan 13]. Available from: http://www.stoptb.org/assets/documents/resources/publications/technical/TB_Case_Studies.pdf

7. Dodd PJ, Yuen CM, Sismanidis C, Seddon JA, Jenkins HE. The global burden of tuberculosis mortality in children: a mathematical modelling study. Lancet Glob Health [Internet]. 2017 Sep [cited 2020 Jan 13];5(9):e898-906. Available from: https://www.thelancet.com/journals/langlo/article/PIIS2214-109X(17)30289-9/fulltext. doi: 10.1016/S2214-109X(17)30289-9

8. Oliveira GP, Pinheiro RS, Coeli CM, Barreira D, Codenotti SB. Uso do sistema de informação sobre mortalidade para identificar subnotificação de casos de tuberculose no Brasil. Rev Bras Epidemiol [Internet]. 2012 set [citado 2020 jan

Epidemiol. Serv. Saude, Brasília, 29(1):e2018485,2020

9

13];15(3):468-77. Disponível em: http://www.scielo.br/pdf/rbepid/v15n3/03.pdf. doi: 10.1590/S1415-790X2012000300003

9. Silva GDM, Bartholomay P, Cruz OG, Garcia LP. Avaliação da qualidade dos dados, oportunidade e aceitabilidade da vigilância da tuberculose nas microrregiões do Brasil. Ciênc Saúde Coletiva [Internet]. 2017 out [citado 2020 jan 13];22(10):3307-19. Disponível em: http://www.scielo.br/pdf/csc/v22n10/1413-8123-csc-22-10-3307.pdf. doi: 10.1590/1413-812320172210.18032017

10. Ministério da Saúde (BR). DATASUS - Departamento de Informática do SUS. População residente - estimativas para o TCU [Internet]. Brasília: Ministério da Saúde: 2019 [citado 2019 maio 11]. Disponível em: http://tabnet.datasus.gov.br/cgi/deftohtm.exe?ibge/cnv/poptbr.def.

11. Instituto Brasileiro de Geografia e Estatística. Censo demográfico 2010 [Internet]. Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística; 201-. [citado 2019 maio 11]. Disponível em: https://sidra.ibge.gov.br/pesquisa/censo-demografico/demografico-2010/inicial

12. Ministério da Justiça (BR). Departamento Penitenciário Nacional. Infopen - levantamento nacional de informações penitenciárias [Internet]. Brasília: Ministério da Justiça; 201-. [citado 2019 maio 11]. Disponível em: http://dados.mj.gov.br/dataset/infopen-levantamento-nacional-de-informacoes-penitenciarias

13. Nunes LN, Klück MM, Fachel JMG. Uso da imputação múltipla de dados faltantes: uma simulação utilizando dados epidemiológicos. Cad Saúde Pública [Internet]. 2009 fev [citado 2020 jan 13];25(2):268-78. Disponível em: http://www.scielo.br/pdf/csp/v25n2/05.pdf. doi: 10.1590/S0102-311X2009000200005

14. Dong Y, Peng C-YJ. Principled missing data methods for researchers. Springer Plus [Internet]. 2013 May [cited 2019 May 4];2:222. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3701793/

15. Harrell FE. Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. New York: Springer New York; 2001. (Springer Series in Statistics).

16. Kassambara A. Hybrid hierarchical k-means clustering for optimizing clustering outputs [Internet]. Montpellier (FR): Data Novia; 2015 [cited 2020 Jan 13]. Available from: http://www.sthda.com/english/wiki/hybrid-hierarchical-k-means-clustering-for-optimizing-clustering-outputs

17. Johnson RA, Wichern DW. Applied multivariate statistical analysis. 6. ed. Upper Saddle River (NJ): Pearson; 2007. 808 p.

18. Gilberto AP. Modelos de regressão com apoio computacional [Internet]. São Paulo: Universidade de São Paulo; 2013 [citado 2020 jan 13]. 429 p. Disponível em: https://www.ime.usp.br/~giapaula/texto_2013.pdf

19. R Core Team. R: a language and environment for statistical computing [Internet]. Vienna: R Core Team; 2019 [cited 2020 Jan 13]. (R Foundation for Statistical Computing). Available from: http://www.R-project.org/

20. Pelissari DM, Bartholomay P, Jacobs MG, Arakaki-Sanchez D, Anjos DSO, Costa MLS, et al. Oferta de serviços pela atenção básica e detecção da incidência de tuberculose no Brasil. Rev Saúde Pública [Internet]. 2018 [citado 2018 nov 12];52:53. Disponível em: http://www.scielo.br/pdf/rsp/v52/pt_0034-8910-rsp-S1518-87872018052000131.pdf. doi: 10.11606/S1518-8787.2018052000131

21. Selig L, Belo M, Cunha AJLA, Teixeira EG, Brito R, Luna AL, et al. Óbitos atribuídos à tuberculose no Estado do Rio de Janeiro. J Bras Pneumol [Internet]. 2004 jul-ago [citado 2020 jan 13];30(4):335-42. Disponível em: http://www.scielo.br/pdf/jbpneu/v30n4/v30n4a06.pdf. doi: 10.1590/S1806-37132004000400006

22. Gibbons CL, Mangen M-JJ, Plass D, Havelaar AH, Brooke RJ, Kramarz P, et al. Measuring underreporting and under-ascertainment in infectious disease datasets: a comparison of methods. BMC Public Health [Internet]. 2014 Feb [cited 2020 Jan 13];14(1):147. Available from: https://bmcpublichealth.biomedcentral.com/articles/10.1186/1471-2458-14-147. doi: 10.1186/1471-2458-14-147

23. Hargreaves JR, Boccia D, Evans CA, Adato M, Petticrew M, Porter JDH. The social determinants of tuberculosis: from evidence to action. Am J Public Health [Internet]. 2011 Apr [cited 2020 Jan 13];101(4):654-62. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3052350/#!po=59.0909. doi: 10.2105/AJPH.2010.199505

24. Maciel EL, Reis-Santos B. Determinants of tuberculosis in Brazil: from conceptual framework to practical application. Rev Panam Salud Pública

[Internet]. 2015 Jul [cited 2020 Jan 13];38(1):28-34. Available from: https://www.ncbi.nlm.nih.gov/pubmed/26506318

25. Loureiro RB, Villa TCS, Ruffino-Netto A, Peres RL, Braga JU, Zandonade E, et al. Acesso ao diagnóstico da tuberculose em serviços de saúde do município de Vitória, ES, Brasil. Ciênc Saúde Coletiva [Internet]. 2014 abr [citado 2020 jan 13];19(4):1233-44. Disponível em: http://www.scielo.br/pdf/csc/v19n4/1413-8123-csc-19-04-01233.pdf. doi: 10.1590/1413-81232014194.01002013
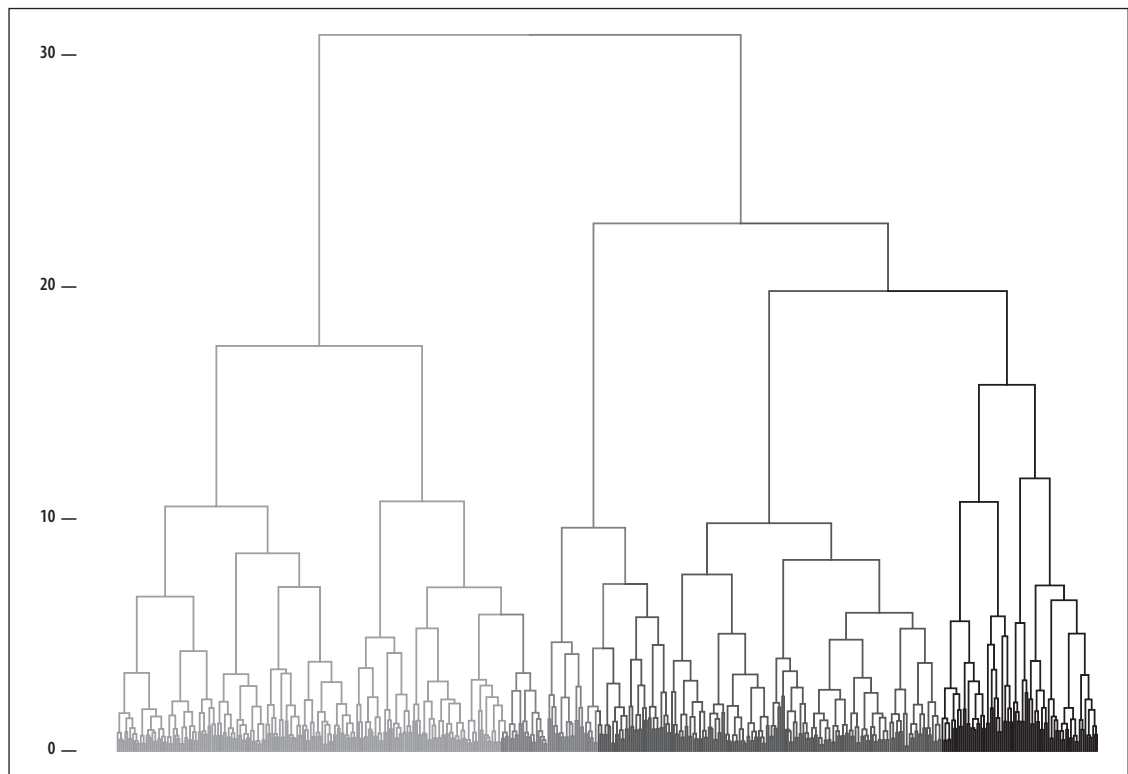
Suplementar material – Dendogram of cluster of cluster analysis of Brazil's micro-regions, 2012-2014, according to tuberculosis surveillance system acceptability and timeliness indicators