



BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<http://bmjopen.bmj.com>).

If you have any questions on BMJ Open's open peer review process please email info.bmjopen@bmj.com

BMJ Open

Identifying socioeconomic, epidemiological and operational scenarios for tuberculosis control in Brazil

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2017-018545
Article Type:	Research
Date Submitted by the Author:	06-Jul-2017
Complete List of Authors:	<p>Pelissari, Daniele; Ministerio da Saude, National Tuberculosis Program of Brazil</p> <p>Rocha, Marli; Ministerio da Saude, National Tuberculosis Program of Brazil</p> <p>Bartholomay, Patricia ; Ministerio da Saude, National Tuberculosis Program of Brazil</p> <p>Sanchez, Mauro; Federal University of Brasilia</p> <p>Duarte, Elisabeth; University of Brasília, Faculty of Medicine, Nucleo de Medicina Tropical</p> <p>Arakaki-Sanchez, Denise; Ministerio da Saude, National Tuberculosis Program of Brazil</p> <p>Dantas, Cíntia ; Ministerio da Saude, National Tuberculosis Program of Brazil</p> <p>Jacobs, Marina; Ministerio da Saude, National Tuberculosis Program of Brazil</p> <p>Andrade, Kleydson ; Ministerio da Saude, National Tuberculosis Program of Brazil</p> <p>Codenotti, Stefano ; Ministerio da Saude, National Tuberculosis Program of Brazil</p> <p>Andrade, Elaine Silva ; Ministerio da Saude, National Tuberculosis Program of Brazil</p> <p>de Araújo, Wildo ; University of Brasilia –UnB</p> <p>Costa, Fernanda ; Ministerio da Saude, National Tuberculosis Program of Brazil</p> <p>Ramalho, Walter; Universidade de Brasilia, Faculdade de Ceilândia;</p> <p>Universidade de Brasília, Núcleo de Medicina Tropical/Faculdade de Medicina</p> <p>Díaz-Quijano, Fredi; School of Public Health, University of São Paulo, Department of Epidemiology</p>
Primary Subject Heading:	Public health
Secondary Subject Heading:	Epidemiology, Health policy, Infectious diseases
Keywords:	Tuberculosis < INFECTIOUS DISEASES, EPIDEMIOLOGY, public health policy



For peer review only

Title: Identifying socioeconomic, epidemiological and operational scenarios for tuberculosis control in Brazil

*Daniele Maria Pelissari, National Tuberculosis Program, Ministry of Health. Address: SRTVN 701, Via W 5 Norte, Ed. PO 700 – 6º andar. Zip code: 70719-040, Brasília-DF, Brazil. Email: daniele.pelissari@saude.gov.br; daniele.pelissari@gmail.com. Telephone: (+55)61998431554

Marli Souza Rocha, National Tuberculosis Program, Ministry of Health. Brasília/DF, Brazil.

Patricia Bartholomay, National Tuberculosis Program, Ministry of Health. Brasília/DF, Brazil.

Mauro Niskier Sanchez, Department of Public Health, University of Brasilia – UnB. Brasília/DF, Brazil.

Elisabeth Carmen Duarte, Medical School, University of Brasilia –UnB. Brasília/DF, Brazil.

Denise Arakaki-Sanchez, National Tuberculosis Program, Ministry of Health. Brasília/DF, Brazil.

Cíntia Oliveira Dantas, National Tuberculosis Program, Ministry of Health. Brasília/DF, Brazil.

Marina Gasino Jacobs, National Tuberculosis Program, Ministry of Health. Brasília/DF, Brazil.

Kleydson Bonfim Andrade, National Tuberculosis Program, Ministry of Health. Brasília/DF, Brazil.

Stefano Barbosa Codenotti, National Tuberculosis Program, Ministry of Health. Brasília/DF, Brazil.

Elaine Silva Nascimento Andrade, National Tuberculosis Program, Ministry of Health. Brasília/DF, Brazil.

Wildo Navegantes de Araújo, University of Brasilia –UnB. Brasília/DF, Brazil.

Fernanda Dockhorn Costa, National Tuberculosis Program, Ministry of Health.

Walter Massa Ramalho, University of Brasilia –UnB. Brasília/DF, Brazil.

Fredi Alexander Diaz-Quijano, Department of Epidemiology, School of Public Health, University of São Paulo. São Paulo/SP, Brazil.

*Corresponding author

Role of the funding source: This study did not have funding.

Word count: 2,999

Number of figures and tables: 3 tables; 2 figures; 1 table as supplementary material

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

ABSTRACT

Objectives We aimed to identified socioeconomic (2010), epidemiological and health care operational indicators (2014/2015) associated with tuberculosis incidence in Brazil.

Design Ecological study

Settings The study was based on new tuberculosis cases (years 2001 to 2015) and epidemiological/operational indicators of the disease (2014 and 2015) from the Brazilian National Information System for Notifiable Diseases (SINAN) and the Mortality Information System (SIM). We also analyzed socioeconomic and demographic indicators obtained from the last population census (2010).

Participants The unit of analysis was the Brazilian municipalities, which in 2015 was 5570. We excluded five municipalities due to absence of socioeconomic information.

Primary and secondary outcome measures We analyzed socioeconomic, epidemiological and health care operational tuberculosis indicators. The adjusted incidence rate ratio (IRR) and 95% confidence intervals of independent variables associated with tuberculosis incidence rate were calculated by negative binomial regression, and municipalities were clustered by the k means method, considering the variables identified in the regression models.

Results Two socioeconomic clusters of municipalities were identified, according to unemployment rate and household crowding: a higher socioeconomic scenario (HSS) with 3482 municipalities and a mean tuberculosis incidence rate of 16.3/100,000 people; and a lower socioeconomic scenario (LSS) with 2083 municipalities and a mean TB incidence rate of 22.1/100,000 people. Then, in a second stage, we performed clustering for each group using epidemiological and operational indicators. This resulted in four subgroups defined by variables such as TB mortality rate and AIDS case detection rate. Moreover, some of the identified sub-scenarios were characterized by fragility in their information systems, while others by the proportion of vulnerable population among TB cases.

Conclusion The identified sub-scenarios highlights the country's high socioeconomic inequality and specific needs to implement the End TB Strategy. This classification can support evidence-based decision making, such as prioritization of focused actions.

Key word: tuberculosis, epidemiology, public health policy.

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

Strengths and limitations of this study
<ul style="list-style-type: none">• This study was based on national population data in a country of continental dimension.• Ours is the first study to consider factors associated with TB incidence rate in the clustering of municipalities to support the elaboration and implementation of a National Plan to End TB.• This methodology can be explored by other countries in order to consider their

respective contexts in the definition of strategies to guide their plans to end TB.

- The information quality could vary according to different sources and periods, which could be a limitation of this study.
- Inferences obtained are applicable to population groups, not to individuals. However, in order to support evidence-based decision making, ecological study can be a suitable design.

INTRODUCTION

In 2015, 10.4 million people had tuberculosis (TB) and 1.8 million died worldwide because of the disease.¹ In Brazil, similarly to other countries, TB incidence reduction (37.9/100,000 people in 2007 to 32.4/100,000 people in 2016)² seems to be associated with the improvement of the population living conditions³⁻⁵ and the performance of TB control programs.⁶ However, the burden of disease continues to be significant in the country, with 66,796 new cases registered in 2016.²

To move towards elimination, the World Health Organization (WHO) launched in 2014 the End TB Strategy, setting targets to be met by 2035, including a 90% reduction in TB incidence as compared to 2015.¹ The strategy is critical to energize the fight against the disease and mobilize resources, but needs to be adapted by the National TB Programs (NTP).

Some countries have already made progress developing their national plans. Among the strategies presented, we highlight the strengthening of existing TB services, the acceleration of the detection of cases in key populations, and the implementation of actions to reduce the gap in the cascade of TB care.⁷⁻⁹

The country has continental proportions, thus both socioeconomic indicators¹⁰ and those that reflect the performance of local TB programs² present a high degree of heterogeneity. Considering this context, and in order to support the "National Plan to End TB", we aimed to identify scenarios based on socioeconomic, epidemiological, and operational factors associated with TB incidence in Brazil.

METHODS

Type of study and data source

This is an ecological study, with the unit of analysis being the Brazilian municipalities, which in 2015 was 5570. We excluded five municipalities due to absence of socioeconomic information. Data on socioeconomic and demographic indicators were obtained from the last population census (2010).^{10,11} As for new TB cases (years 2001 to 2015) and epidemiological/operational indicators of the disease (2014 and 2015), we used data from the Brazilian National Information System for Notifiable Diseases (SINAN) and the Mortality Information System (SIM).¹⁰

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

Variables

The dependent variable was the TB incidence rate (/100,000) in 2015 and the independent variables were socioeconomic, epidemiological, and health care operational TB indicators.

We analyzed the following socioeconomic indicators: municipal human development index (HDI-M); average household income per capita; Gross Domestic Product (GDP) per capita; proportion of extremely poor, poor, and vulnerable to poverty; Gini coefficient; unemployment rate; illiteracy rate; proportion of the population living in households with more than two people per room represented by household crowding; infant mortality rate per 1,000 live births; and life expectancy at birth. Many of these indicators have already been identified in previous studies as TB determinants.¹²⁻¹⁴

Average household income per capita and GDP per capita were converted into US dollars using the average annual price of USD1.00 for 2010, which was 1.76 Brazilian currency (reais- R\$). We adopted the Brazilian definitions for the proportions of extremely poor, poor and vulnerable to poverty: proportion of individuals in the municipality with average household income per capita equal or less than USD40, USD80 and USD145, respectively.¹¹

The epidemiological indicators of TB were:

- AIDS case detection rate per 100,000 people;
- Proportion of new TB cases who were: HIV positive; prisoners; health professionals; indigenous; homeless; and, as a composite definition, proportion of TB cases from at least one of those vulnerable group. Those indicators were previously associated with an increased risk of TB in other studies;^{3-5,15}
- Proportion of TB retreatment;
- TB mortality rate per 100,000 people;

The health care operational indicators of TB considered in the analysis were:

- Proportions of new TB cases: in which contacts were examined; laboratory confirmed; tested for HIV; and treatment outcomes (cure, lost to follow-up, and no TB outcome registration);
- Proportion of sputum culture examination among retreatment cases;

Due to the availability of updated data at the time of analysis, data to calculate culture examination, treatment outcomes, and TB mortality rate refers to 2014, while the other indicators refers to 2015.

Statistical analysis

Statistical analysis was performed in two stages; each of them included a modelling to identify the factors associated with TB incidence. This was followed by a cluster analysis based in the

identified factors. First stage was focused in socioeconomic variables and the second in the epidemiological and operational indicators associated with TB incidence.

For the modelling, we included municipalities that presented mean annual variation of the triennial moving average of the incidence rate for the years 2001 to 2015 between -8 and 8%. By doing this, we intended to reduce possible biases due to the variability of values in small municipalities, and the intermittence in case notification.

Initially, bivariate models were estimated for each independent variable and those that presented an association with a p value <0.20 , were analyzed in a Spearman correlation matrix. Whenever correlation between independent variables ($r>0.50$) was identified, we selected the variable with the highest association with TB incidence rate. In the regression model, we used a stepwise forward selection method and variables with p value <0.05 were preserved in the model. We presented the association measures as the relative increment in the incidence rate.

All the models were adjusted by the population size of municipalities, which were classified as small (less than 20,000 inhabitants), medium (20,000 to 99,999 inhabitants) and large (100,000 inhabitants or larger).¹⁶

Socioeconomic indicators associated in a multiple model (primary model) were considered for a cluster analysis of all municipalities using the non-hierarchical k-means method. In this method, the algorithm seeks to reduce intra-group variance and maximize inter-group variance in relation to the Euclidean distance established by the indicators selected.¹⁷ For the definition of the number of clusters, we used the Elbow method which relates the number of clusters with the percentage of internal variation of the groups,¹⁷ adopting $> 60\%$ as the cut-off point and among these, the smallest number of possible clusters.

For the second stage, epidemiological/operational indicators were modelling in each of one socioeconomic scenarios, following a similar methodology described for socioeconomic factors. Factor associated with TB in these secondary models, as well as, TB mortality rate, were considered for a second cluster analysis, which subdivided the previous socioeconomic clusters into epidemiological/operational TB sub-scenarios. Because some operational indicators are only measured during care of TB cases, these second stage methods were applied only in municipalities with TB cases reported in 2014 and 2015.

Statistical analyses were performed with the Stata statistical package version 12.0, R version 3.3.1 and the cluster library.

Data used in this study do not include individual information from TB patients. Additionally, all data analyzed are publicly available in Brazil. Therefore, it was not necessary to submit the study to an institutional review board, which is in accordance with Resolution N° 510 of the National Health Council of Brazil.¹⁸

RESULTS

In 2015, 67,777 new TB cases were reported in Brazil, with an incidence rate of 33.1/100,000 people. Of the total number of municipalities, 3311 (59.5%) were eligible for the analysis for the

first model, due to their acceptable variability in reporting incident cases in the period 2001 to 2015, including 791 that did not present new TB cases in 2015.

With the exception of the unemployment rate, all remaining variables presenting association with the outcome and a p value smaller than 0.20 were strongly correlated (Spearman coefficient > 0.50). Considering that household crowding had a greater association with the TB incidence rate, the first model included this variable and the unemployment rate, adjusted by municipal population size (Table 1).

Table 1. Socioeconomic variables and association with tuberculosis incidence rate in Brazil (n= 3,311 municipalities^a)

Variable ^b	Mean (standard deviation-SD)	Median (IQ25%-IQ75%)	Relative increment of IR (95% CI) ^c	Relative increment of IR adjusted (95% CI) ^c
HDI-M	0.7 (0.1)	0.7 (0.6 - 0.7)	-4.8 (-36.9 to 43.7)	
Average household income per capita (USD)	280.3 (143.8)	257.9 (155.6 - 372.8)	-0.0 (-0.0 to -0.0) ^d	
GDP per capita (USD)	7510.4 (8630.3)	5555.2 (2909.4 - 9091.2)	0.0 (-0.0 to 0.0)	
Extremely poor (%)	11.4 (11.7)	6.5 (1.6 - 19.1)	-0.0 (-0.3 to 0.2)	
Poor (%)	23.4 (18.0)	18.5 (6.9 - 38.8)	0.1 (-0.0 to 0.3)	
Vulnerable to poverty (%)	44.1 (22.7)	42.6 (23.3 - 65.6)	0.2 (0.0 to 0.3) ^d	
Gini coefficient	0.5 (0.1)	0.5 (0.5 - 0.5)	116.0 (37.1 to 240.3) ^d	
Unemployment rate (%)	6.7 (3.7)	6.3 (4.2 - 8.6)	5.0 (4.2 to 5.8) ^d	3.9 (3.0 to 4.7) ^d
Illiteracy in population with ≥ 18 years (%)	17.2 (10.8)	13.9 (8.1 - 26.4)	-0.4 (-0.6 to -0.1) ^d	
Illiteracy in population with ≥ 15 years (%)	15.6 (9.8)	12.9 (7.2 - 23.8)	-0.4 (-0.7 to -0.1) ^d	
Household crowding (%)	26.4 (13.1)	24.7 (16.6 - 33.8)	1.2 (1.0 to 1.4) ^d	0.8 (0.6 to 1.1) ^d
Infant mortality rate (number of deaths in first year of life/1000 live births)	19.3 (7.2)	17.0 (13.7 - 24.1)	0.1 (-0.3 to 0.5)	
Life expectancy at birth (years)	73.1 (2.7)	73.4 (71.1 - 75.2)	-0.9 (-1.9 to 0.2)	

Abbreviations: HDI-M, municipal human development index; GDP, gross domestic product; SD, standard deviation; IQ, interquartile; IR, incidence rate; CI, confidence interval.

^a Municipalities with annual variation in TB incidence rate between -8% and 8% and at least one TB incidence case in 2015.

^b Variables measured in the last census (2010).

^c The association measure represents the relative increment in the incidence rate (IRR-1), adjusted by the population size of the municipality.

^d p value < 0.05.

Based in these two socioeconomic variables (Table 1), we identified a higher socioeconomic scenario (HSS) cluster, with 3482 municipalities, which presented better socioeconomic indicators when compared to the 2083 municipalities from the second cluster, the lower socioeconomic scenario (LSS). HSS cluster exhibited unemployment rates of up to 26.87%; and, household crowding values between 0.65% and 28.56%. On the other hand, LSS cluster exhibited unemployment rates of up to 39.15%; and, household crowding values between 26.65% and 88.64% (Fig. 1).

Figure 1. Distribution of Brazilian municipalities according to socioeconomic variables associated with tuberculosis incidence rate

The mean TB incidence rate in the LSS was 22.1/100,000 people (Table 2), which was significantly higher than that observed in the HSS, that was 16.3/100,000 people (IRR: 1.34; 95% CI: 1.26 - 1.41).

Table 2. Description of Socioeconomic Scenarios Associated With Tuberculosis Incidence Rate in Brazil (n= 5,565 Municipalities)

Variables ^a	Higher socioeconomic scenario n= 3482 municipalities		Lower socioeconomic scenario n= 2083 municipalities	
	Mean (SD)	Median (IQ25%-IQ75%)	Mean (SD)	Median (IQ25%-IQ75%)
TB incidence rate (cases/ 100,000 people)	16.3 (31.7)	10.2 (0 - 23.2)	22.1 (36.6)	16.9 (6.3 - 29.6)
HDI-M	0.7 (0.1)	0.7 (0.6 - 0.7)	0.6 (0.1)	0.6 (0.6 - 0.6)
Average household income per capita (USD)	330.5 (126.5)	324.1 (242.3 - 403.5)	181.3 (92.4)	150.8 (125.3 - 203.5)
GDP per capita (USD)	8661.1 (7707.6)	7259.7 (4758.8 - 10053.9)	4930.5 (8033.6)	2937.6 (2371.3 - 5004.5)
Extremely poor (%)	5.8 (7.3)	2.6 (1.1 - 7.4)	20.7 (11.9)	20.3 (12.0 - 28.7)
Poor (%)	14.2 (12.3)	9.7 (5.1 - 20.2)	38.3 (15.5)	40.3 (29.7 - 49.1)
Vulnerable to poverty (%)	33.0 (17.8)	29.3 (19.0 - 44.9)	62.4 (16.5)	67.0 (57.0 - 73.5)
Gini coefficient	0.5 (0.1)	0.5 (0.4 - 0.5)	0.5 (0.1)	0.5 (0.5 - 0.6)
Unemployment rate (%) ^b	5.1 (2.9)	4.9 (3.1 - 6.8)	8.3 (4.0)	7.6 (5.6 - 10.3)
Illiteracy in population with ≥ 18 years (%)	12.9 (8.0)	10.7 (7.2 - 16.0)	25.0 (10.3)	26.5 (17.1 - 32.9)
Illiteracy in population with ≥ 15 years (%)	11.7 (7.3)	9.8 (6.5 - 14.6)	22.7 (9.5)	23.9 (15.6 - 29.8)
Household crowding (%) ^b	17.1 (6.0)	17.3 (12.7 - 22.0)	38.5 (10.3)	35.7 (31.2 - 41.9)
Infant mortality rate (number of deaths in first year of life/1000 live births)	16.0 (5.0)	14.8 (12.8 - 17.5)	24.6 (6.9)	24.2 (19.4 - 29.0)
Life expectancy at birth (years)	74.3 (2.1)	74.5 (73.2 - 75.7)	71.1 (2.3)	71.1 (69.6 - 72.6)

Abbreviations: HDI-M, municipal human development index; GDP, gross domestic product; SD, standard deviation; IQ, interquartile.

^a With exception of TB incidence rate (2015), the other variables were measured in the last census (2010).

^b Variables identified during step 1 (model 1) used in step 2 with k-means method.

Among the 3482 HSS municipalities, 1125 had TB cases in 2014 and 2015, and were eligible (annual variation in TB incidence rate between -8% and 8%) for secondary modelling. In this model, the AIDS case detection rate and the proportion of new cases from at least one vulnerable group were positively associated with the TB incidence rate, while the proportion of contacts investigation among new TB cases presented an inverse association (Table 3).

Concerning the LSS, 1095 municipalities out of 2083 were eligible. The proportion of cases with no TB outcome registration was inversely associated with TB incidence rate, and the AIDS case detection rate and the proportion of cases from at least one vulnerable group were positively associated (Table 3).

Table 3. Epidemiological and Operational Tuberculosis Indicators Associated With Tuberculosis Incidence Rate Stratified by Socioeconomic Scenarios in Brazil

Variables ^a	Higher socioeconomic scenario n= 1125 municipalities ^b		Lower socioeconomic scenario n= 1095 municipalities ^b	
	Relative increment of IR (95% CI) ^c	Relative increment of IR adjusted (95% CI) ^c	Relative increment of IR (95% CI) ^c	Relative increment of IR adjusted (95% CI) ^c
Epidemiological indicators				
AIDS case detection rate (cases/100,000 people)	1.5 (1.2 to 1.7) ^e	1.4 (1.1 to 1.6) ^e	2.1 (1.7 to 2.5) ^e	2.0 (1.6 to 2.4) ^e
New cases from at least one vulnerable group (%) ^d	0.5(0.3 to 0.7) ^e	0.2 (0.1 to 0.4) ^e	0.7 (0.5 to 0.9) ^e	0.5 (0.3 to 0.7) ^e
TB-HIV confection among new cases (%)	0.3 (0.0 to 0.6) ^e		-0.2 (-0.5 to 0.2)	
New cases who were prisoners (%)	0.7 (0.4 to 0.9) ^e		1.2 (0.9 to 1.5) ^e	
New cases who were health professionals (%)	-0.5 (-1.3 to 0.2)		-0.2 (-1.0 to 0.6)	
New cases who were indigenous population (%)	1.1 (0.3 to 1.9) ^e		0.9 (0.5 to 1.2) ^e	
New cases who were homeless (%)	0.1 (-0.6 to 0.7)		0.1 (-0.7 to 1.0)	
Retreatment cases among total cases (%)	0.5 (0.2 to 0.8) ^e		0 (-0.3 to 0.3)	
Operational indicators (new TB cases)				
Contacts examination (%)	-0.3 (-0.4 to -0.1) ^e	-0.2 (-0.3 to -0.1) ^e	-0.0 (-0.2 to 0.1)	
Pulmonary cases with laboratory confirmation (%)	0.0 (-0.1 to 0.2)		-0.1 (-0.3 to 0.0)	
Tested for HIV (%)	0.1 (-0.0 to 0.2)		0.0 (-0.1 to 0.2)	
Cure (%)	-0.2 (-0.3 to -0.0) ^e		0.2 (0.0 to 0.3) ^e	
Lost to follow-up (%)	0.6 (0.3 to 0.9) ^e		0.3 (0.0 to 0.7) ^e	
No TB outcome registration (%)	0.3 (0.1 to 0.5) ^e		-0.3 (-0.5 to -0.1) ^e	-0.3 (-0.5 to -0.1) ^e
Culture examination (retreatment) (%)	0.1 (-0.1 to 0.2)		0.0 (-0.1 to 0.2)	

Abbreviations: IR, incidence rate; CI, confidence interval.

^a With exception of: cure, lost to follow-up, no TB outcome registration and culture examination (2014), the other variables were measured in 2015.

^b Municipalities with annual variation in TB incidence rate between -8% and 8% and at least one TB incidence case in 2014 and 2015.

^c The association measure represents the relative increment in the incidence rate (IRR-1), adjusted by the population size of the municipality.

^d HIV, health professional, prisoners, indigenous, and homeless.

^e p value < 0.05.

With the variables associated with the outcome in the previous models, and also considering the TB mortality rate, we defined three clusters for each socioeconomic scenario, totaling six sub-scenarios with TB cases in 2014 and 2015. For each scenario, a sub-scenario (1.0 and 2.0) was also defined including municipalities without TB notification in 2014 or 2015 (supplementary material).

Figure 2 shows municipalities scenarios distribution in Brazil. Regarding the sub-scenarios with TB cases in HSS, 1.1 showed the lowest mean rates of TB incidence, AIDS case detection and TB mortality. Sub-scenario 1.2, despite having relatively low mean rates of TB incidence, AIDS case detection, and TB mortality, had a high proportion of cases with no TB outcome registration. Sub-scenario 1.3 covers 27.8% of new TB cases reported in 2015 and presents the highest mean rates of TB incidence, AIDS case detection, and proportion of cases from at least one vulnerable group (22%) (supplementary material).

Concerning the LSS municipalities, the sub-scenario 2.1 had the highest mean proportion of cases with no TB outcome registration (81.8%) and the lowest mean for contact investigation

(36.5%) and HIV testing (52.3%). 2.2 shows a high TB incidence rate, the highest TB mortality and it has a high mean proportion of cases with no TB outcome registration (37%) and low HIV testing (53.5%). The sub-scenario 2.3 presents the highest mean proportion of new TB cases reported (56.3%), as a consequence of the inclusion of 14 capitals, among them São Paulo, the most populous city in the country. It has the highest mean AIDS case detection rate in the group of LSS and the second highest TB mortality rate among all sub-scenarios (supplementary material).

Figure 2- Brazilian's municipalities by TB incidence rate scenarios. Brazil, 2015.

DISCUSSION

The present study classified the 5565 Brazilian municipalities in two scenarios using socioeconomic variables associated with TB incidence rate in Brazil. After that, we performed a sub classification based on operational and epidemiological indicators associated with TB incidence rate.

In the socioeconomic scenarios definition, regarding the unemployment rate, studies in the United States,¹⁹ Spain,²⁰ and also in Brazil¹⁴ found an association of this variable and the risk of TB. At the individual level, unemployment has been associated with an increased risk of abusive use of alcohol and illicit drugs,²¹ and also with lost to follow-up during HIV treatment.²² These factors have already been associated with TB risk³⁻⁵ and could, at least partially, explain the association observed in our study.

In the United States and West Africa people living in crowding conditions had a higher risk of TB.^{23,24} In Brazil, this variable was already associated with TB incidence and analyzed as a potential mediator between socioeconomic determinants and TB incidence, because it may directly favors TB transmission by increasing the contact rate between *Mycobacterium tuberculosis* and susceptible people.¹⁴

The two socioeconomic scenarios divided the country according to its social disparities. The LSS, with municipalities predominantly in the North, Northeast, and Center-West regions, presented worse socioeconomic indicators than those of the HSS, with municipalities located predominantly in the South and Southeast regions. In addition, the mean TB incidence rate in the LSS was 34 higher than HSS.

Regarding the operational and epidemiological indicators, the AIDS case detection rate was positively associated with TB incidence in both of socioeconomic scenarios, which is similar to previous studies on the role of AIDS as a factor associated with TB risk at the contextual level.^{13,14,25}

The proportion of new cases from at least one vulnerable group was also another factor associated with TB incidence in both scenarios. Among vulnerable populations were included the prisoners. Specifically in Brazil, in 2014 there were approximately 607 thousand imprisoned people in 956 municipalities distributed in all regions of the country, with a prison occupation rate of 161%.²⁶

In the HSS, vulnerability was also correlated with the AIDS case detection rate which is higher in the South and Southeast regions (respectively 20.1% and 53.0% of the AIDS cases identified from 1980 to June 2016).²⁷ Regarding the LSS, vulnerability was correlated with indigenous populations, which are predominantly located in the North (37.4%), Northeast (25.5%), and Central-West regions (16.0%)²⁸ and present a higher risk of TB when compared to other races.⁴

In the HSS, we observed an inverse association between TB incidence rate and the percentage of contact investigation, which may represent the overall effect on transmission control, possibly through the identification and timely treatment. Finally, in the LSS the association with the proportion of cases with no TB outcome registration may represent failures in surveillance in collecting these data for the qualification of the information system.

Regarding the absence of TB cases in 2014 or 2015 in sub-scenarios 1.0 and 2.0, it is possible that there is an under-registration in these scenarios, mainly in the 2.0, where there are worse socioeconomic conditions, which were associated with a higher risk of TB. This suggests that activities related to TB detection should be strengthened especially in those groups of municipalities.

Regarding the sub-scenarios that reported cases in the two years of analysis from HSS cluster, group 1.1 has the lowest TB incidence rate, better socioeconomic indicators, and good TB epidemiological/operational indicators, suggesting an advanced stage in TB control.

The sub-scenario 1.2 (HSS-cluster), 2.1 and 2.2 (LSS-cluster) need improvement in the information system because the high proportion of cases with no TB outcome registration. This makes difficult to analyze the performance of TB control actions. Another challenge in these groups is the investigation of contacts which was particularly low in the 2.1 sub-scenario. On the other hand, the mean burden of proportion of cases from at least one vulnerable group was lower in group 2.2 (10.6%), revealing an endemic less concentrated in vulnerable populations. Although sub-scenario 2.1 has the highest percentage of HIV-TB co-infection in new cases (9.8%), it also has one of the lowest percentages of HIV testing (52.3%), suggesting the subdetection of HIV among people with TB.

The sub-scenario 1.3 presents the highest TB incidence rate, AIDS case detection rate and proportion of cases from at least one vulnerable group (22%), especially among prisoners (12.1%). In addition, this scenario is composed mainly by capitals, which could mean a more sensitive surveillance system. Despite being the group of municipalities with the highest TB risk,

the distribution of vulnerabilities reveals concentrated epidemic in some population groups, which requires distinct and focused strategies to control the disease.

Sub-scenario 2.3 has a reliable information system and good performance in operational activities (e.g. contact investigation and HIV testing), revealing that even with limited resources, it is possible to carry out effective disease control actions.

Finally, with the exception of the sub-scenario 2.0, all those in the LSS had a higher TB mortality rate than those in HSS. The sub-scenario 2.0, which did not have new cases, also had a higher mortality rate than the 1.0 group. Mortality is expected to be less underreported than the incidence, such as observed in other diseases.^{29,30} Thus, the use of this variable for the definition of clusters contributes to identify groups according to tuberculosis burden (represented in mortality) besides to the socioeconomic characteristics used for classification.

Limitation

As a common limitation of ecological studies, aggregate measures might differ from the individual ones. However, these studies provide an overview that contribute to direct decision-making in public policies. Underreporting of TB cases in Brazil is decreasing each year,¹ but may remain a potential limitation for the present study. The hypothesis is that the underreporting is not homogeneous among the municipalities, being higher in municipalities with worst socioeconomic indicators. Therefore, the magnitude of association measures may have been underestimated. The exclusion of municipalities that presented high variability in the incidence rate may have contributed to the reduction of this limitation.

Concerning data availability, the municipalities' clusterization was restricted to the variables available in different data sources, and socioeconomic indicators were only available for the year of the last census conducted in the country (2010).

Implications for public health and conclusion

The End TB Strategy proposes bold targets, and a prompt response from each country can be critical for their achievement. The grouping municipalities presented in this study was the initial step to assist in the elaboration of the Brazilian national plan.

Clustering enabled to define specific strategies for each sub-scenario, customized to tackle the disease. Efforts should be focused on strengthening information systems to provide a reliable picture of the epidemiological situation. In addition, specific actions to control the disease in vulnerable populations may contribute to the national reduction of TB risk. Municipalities in the LSS should receive additional support from local governments, given the limitations of resources, and, consequently, exacerbation of social vulnerabilities, which are reflected in the TB risk.

The heterogeneity of the socioeconomic and epidemiological situation of TB in Brazil, observed in this study, represents a great challenge for TB control in a country of continental proportions,

which may be also the reality of other countries. In this sense, data analysis approach proposed in this study identified sub-scenarios that will guide the NTP in Brazil in specifying appropriate actions to TB control, and could be considered by other countries. Additionally, the knowledge of local managers about their TB health services network and epidemiology, associated with robust data analysis methodologies should be used for decision-making.

CONTRIBUTORS

Daniele Maria Pelissari: conceived the study, contributed to the design of the study and results interpretation, coordinated and analyzed the data including epidemiological models and clusterization, wrote the first draft of the report, elaborated the illustrations, wrote the discussion of the sub-scenarios, revised the manuscript critically and approved the final version.

Marli Souza Rocha: conceived the study, contributed to the design of the study and results interpretation, analyzed the data including epidemiological models and clusterization, contributed to the discussion of the sub-scenarios, revised the manuscript critically and approved the final version.

Patricia Bartholomay: conceived the study, contributed to the design of the study and results interpretation, analyzed the data including epidemiological models and clusterization, contributed to the discussion of the sub-scenarios, revised the manuscript critically and approved the final version.

Mauro Niskier Sanchez: contributed to the design of the study and results interpretation, gave significant inputs in the first draft, contributed to the discussion of the sub-scenarios, revised the manuscript critically and approved the final version.

Elisabeth Carmen Duarte: contributed to the design of the study and results interpretation, gave significant inputs in the first draft, contributed to the discussion of the sub-scenarios, revised the manuscript critically and approved the final version.

Denise Arakaki-Sanchez: contributed to the design of the study and results interpretation, contributed to the discussion of the sub-scenarios, revised the manuscript critically and approved the final version.

Cíntia Oliveira Dantas: contributed to the design of the study and results interpretation, contributed to the discussion of the sub-scenarios, revised the manuscript critically and approved the final version.

Marina Gasino Jacobs: contributed to the design of the study and results interpretation, contributed to the discussion of the sub-scenarios, revised the manuscript critically and approved the final version.

Kleydson Bonfim Andrade: contributed to the design of the study and results interpretation, contributed to the discussion of the sub-scenarios, revised the manuscript critically and approved the final version.

Stefano Barbosa Codenotti: contributed to the design of the study and results interpretation, contributed to the discussion of the sub-scenarios, revised the manuscript critically and approved the final version.

Elaine Silva Nascimento Andrade: contributed to the design of the study and results, contributed to the discussion of the sub-scenarios, revised the manuscript critically and approved the final version interpretation.

Wildo Navegantes de Araújo: contributed to the design of the study and results interpretation, contributed to the discussion of the sub-scenarios, revised the manuscript critically and approved the final version.

Fernanda Dockhorn Costa: contributed to the design of the study and results interpretation, contributed to the discussion of the sub-scenarios, revised the manuscript critically and approved the final version.

Walter Massa Ramalho: contributed to the design of the study and results interpretation, contributed to the discussion of the sub-scenarios, revised the manuscript critically and approved the final version.

Fredi Alexander Diaz-Quijano: contributed to the design of the study and results interpretation, gave significant inputs in the first draft, elaborated the illustrations, contributed to the discussion of the sub-scenarios, revised the manuscript critically and approved the final version.

DATA SHARING STATEMENT

Contextual data are available from the Brazilian Health Ministry website (www.datasus.gov.br/tabnet/tabnet.htm); Brazilian Institute of Geography and Statistics website (<http://www.ibge.gov.br>); and The Human Development Atlas in Brazil website (<http://atlasbrasil.org.br/2013/>). Tuberculosis case data can be made available by Brazilian Health Ministry (<http://portalsaude.saude.gov.br>).

FUNDING

This study did not have funding.

COMPETING INTERESTS

None declared.

References

1. World Health Organization. *Global tuberculosis report 2016*. Geneva: World Health Organization, 2016. Available from: http://www.who.int/tb/publications/global_report/en/ [cited 2017 Jan 12].
2. Ministério da Saúde. Indicadores prioritários para o monitoramento do Plano Nacional pelo Fim da Tuberculose como Problema de Saúde Pública no Brasil. *Bol. Epidemiológico* 2017;48:1-6.
3. San Pedro A, Oliveira RM de. Tuberculose e indicadores socioeconômicos: revisão sistemática da literatura. *Rev Panam Salud Publica* 2013;33:294–301.
4. Lacerda SNB, Temoteo RC de A, Figueiredo TMRM de, Luna DT de L, de Sousa AN, Abreu C, et al. Individual and social vulnerabilities upon acquiring tuberculosis: a literature systematic review. *Int. Arch. Med.* 2014 Jul;7(35):1-8. doi: <http://dx.doi.org/10.1186/1755-7682-7-35>.
5. Millet J-P, Moreno A, Fina L, et al. Factors that influence current tuberculosis epidemiology. *Eur. Spine J.* 2013;22(4):539–48. doi: <http://dx.doi.org/10.1007/s00586-012-2334-8>.
6. Houben RMGJ, Menzies N, Sumner T, et al. Feasibility of achieving the 2025 WHO global tuberculosis targets in South Africa, China, and India: a combined analysis of 11

- mathematical models. *Lancet Glob. Health* 2016;4:e806–15. doi: [http://dx.doi.org/10.1016/S2214-109X\(16\)30199-1](http://dx.doi.org/10.1016/S2214-109X(16)30199-1).
7. Shin Y, Ki M, Sung N. A new strategy for tuberculosis control in North Korea. *Epidemiol. Health* 2015;37:e2015053. doi: <http://dx.doi.org/10.4178/epih/e2015053>.
 8. Sharma DC. New plan to end tuberculosis in south and southeast Asia. *The Lancet* 2017;25;389(10075):1183. doi: [http://dx.doi.org/10.1016/S0140-6736\(17\)30817-6](http://dx.doi.org/10.1016/S0140-6736(17)30817-6).
 9. Pai M, Correa N, Mistry N, et al. Reducing global tuberculosis deaths-time for India to step up. *The Lancet* 2017;25;389(10075):1174–1176. doi: [http://dx.doi.org/10.1016/S0140-6736\(17\)30790-0](http://dx.doi.org/10.1016/S0140-6736(17)30790-0).
 10. Ministério da Saúde. *Departamento de Informática do SUS: Informações de saúde (TABNET)*. Brasília: Ministério da Saúde, 2017. Available from: <http://www.datasus.gov.br/informacoes-de-saude/tabnet> [cited 2017 Mar 4].
 11. Programa das Nações Unidas para o Desenvolvimento – PNUD. *Atlas do Desenvolvimento Humano no Brasil*. Programa das Nações Unidas para o Desenvolvimento – PNUD, 2015. Available from: <http://www.atlasbrasil.org.br/2013/pt/consulta/> [cited 2017 Mar 4].
 12. Janssens J-P, Rieder HL. An ecological analysis of incidence of tuberculosis and per capita gross domestic product. *Eur. Respir. J.* 2008;32(5):1415–6. doi: <http://dx.doi.org/10.1183/09031936.00078708>.
 13. Dye C, Lönnroth K, Jaramillo E, et al. Trends in tuberculosis incidence and their determinants in 134 countries. *Bull. World Health Organ.* 2009;87:683–91.
 14. Pelissari DM, Diaz-Quijano FA. Household crowding as a potential mediator of socioeconomic determinants of tuberculosis incidence in Brazil. *PloS One* 2017;18;12(4):e0176116. doi: <http://dx.doi.org/10.1371/journal.pone.0176116>.
 15. Baussano I, Williams BG, Nunn P, et al. Tuberculosis incidence in prisons: a systematic review. *PLoS Med.* 2010;7(12):e1000381. doi: <http://dx.doi.org/10.1371/journal.pmed.1000381>.
 16. Ministério do Desenvolvimento Social e Combate à Fome. *Guia de Orientação Técnica – SUAS Nº 1 Proteção Social Básica de Assistência Social*. Brasília: Ministério do Desenvolvimento Social e Combate à Fome, 2005. Available from: http://www.pg.pr.gov.br/ctmas/wp-content/uploads/2011/07/guia_orientacao_tecnica.pdf [cited 2017 Mar 4].
 17. Trupti MK, Prashant RM. International Journal of Advance Research in Computer Science and Management Studies. *Rev. Determining Number Clust. K-Means Clust.* 2013;1:90–5.

18. Conselho Nacional de Saúde. *Resolução Nº 510, de 7 de abril de 2016*. Brasília: Conselho Nacional de Saúde, 2016. Available from: <http://conselho.saude.gov.br/resolucoes/2016/Reso510.pdf> [cited 2017 Jan 15].

19. Cantwell MF, McKenna MT, McCray E, et al. Tuberculosis and race/ethnicity in the United States: impact of socioeconomic status. *Am. J. Respir. Crit. Care Med.* 1998;157:1016–20.

20. Díaz de Quijano E, Brugal MT, Pasarín M, et al. Influencia de las desigualdades sociales, la conflictividad social y la pobreza extrema sobre la morbilidad por tuberculosis en la ciudad de Barcelona. *Rev. Esp. Salud Pública* 2001;75:517–28.

21. Compton WM, Gfroerer J, Conway KP, et al. Unemployment and substance outcomes in the United States 2002–2010. *Drug Alcohol Depend.* 2014;142:350-3. doi: <http://dx.doi.org/10.1016/j.drugalcdep.2014.06.012>.

22. Támara-Ramírez JR, Álvarez CA, Rodríguez J. Loss of follow-up and associated factors in patients enrolled in the HIV/AIDS program of the Hospital Universitario San Ignacio, Colombia, 2012-2013. *Biomed. Rev. Inst. Nac. Salud* 2016;3;36(2):265-75. doi: <http://dx.doi.org/10.1016/10.7705/biomedica.v36i2.2939>.

23. Myers WP, Westenhouse JL, Flood J, et al. An ecological study of tuberculosis transmission in California. *Am. J. Public Health* 2006;96:685.

24. Lienhardt C, Fielding K, Sillah JS, et al. Investigation of the risk factors for tuberculosis: a case-control study in three countries in West Africa. *Int. J. Epidemiol.* 2005;34:914–23.

25. Kwan CK & Ernst JD. HIV and tuberculosis: a deadly human syndemic. *Clin. Microbiol. Rev.* 2011;24(2):351-76. doi: <http://dx.doi.org/10.1128/CMR.00042-10>.

26. Ministério da Justiça. *Relatório do Levantamento Nacional de Informações Penitenciárias — Infopen*. Brasília: Ministério da Justiça, 2014. Available from: <http://www.justica.gov.br/noticias/mj-divulgara-novo-relatorio-do-infopen-nesta-terca-feira/relatorio-depen-versao-web.pdf> [cited 2017 Mar 21].

27. Ministério da Saúde. Departamento de DST, Aids e Hepatites Virais. Ano V - no 1 - 27a a 53a - semanas epidemiológicas - julho a dezembro de 2015 Ano V - no 1 - 01a a 26a - semanas epidemiológicas - janeiro a junho de 2016. *Bol. Epidemiológico - Aids E DST* (2016). Brasília: Ministério da Saúde, 2016.

28. Instituto Brasileiro de Geografia e Estatística. *Os indígenas no Censo Demográfico 2010: primeiras considerações com base no quesito cor ou raça, 2012*. Brasília: Instituto Brasileiro de

Geografia e Estatística, 2012. Available from:
http://www.ibge.gov.br/indigenas/indigena_censo2010.pdf [cited 2017 Mar 21].

29. Díaz-Quijano FA, Waldman EA. Factors associated with dengue mortality in Latin America and the Caribbean, 1995–2009: an ecological study. *Am. J. Trop. Med. Hyg.* 2012;86(2):328-34. doi: <http://dx.doi.org/10.4269/ajtmh.2012.11-0074>.

30. Gibbons CL, Mangen M-JJ, Plass D, et al. Measuring underreporting and under-ascertainment in infectious disease datasets: a comparison of methods. *BMC Public Health* 2014;11;14:147. doi: <http://dx.doi.org/10.1186/1471-2458-14-147>.

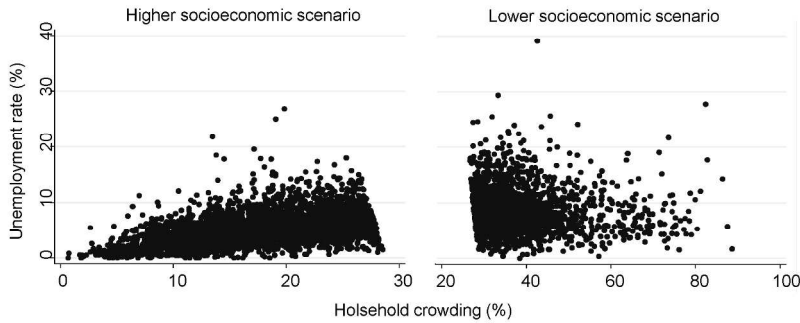


Figure 1. Distribution of Brazilian municipalities according to socioeconomic variables associated with tuberculosis incidence rate

250x190mm (300 x 300 DPI)

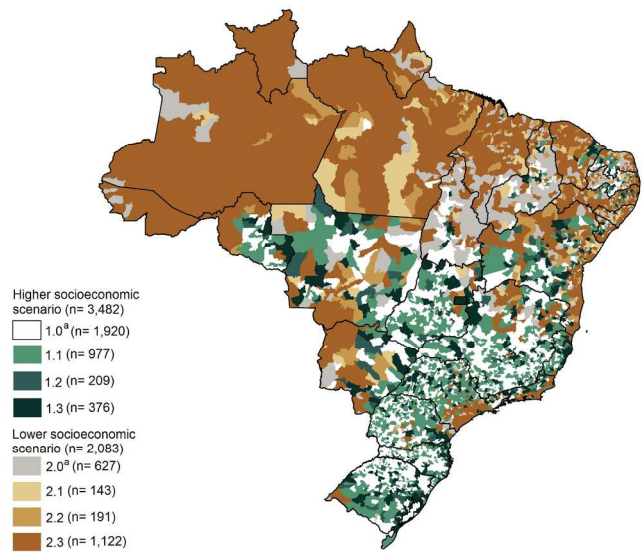


Figure 2 Brazilian's Municipalities by Tuberculosis Incidence Rate Scenarios. Brazil, 2015

^a Without cases in 2014 or 2015.

209x148mm (300 x 300 DPI)

Supplementary material- Socioeconomics, epidemiological and operational-tuberculosis indicators by tuberculosis scenarios in Brazil

Variables ^a	Higher socioeconomic n= 3482					Lower socioeconomic n= 2083		
	Subs. 1.0 ^b	Subs. 1.1	Subs. 1.2	Subs. 1.3	Subs. 2.0 ^b	Subs. 2.1	Subs. 2.2	Subs. 2.3
Socioeconomics indicators								
Number of municipalities (%)	1920 (34.5)	977 (17.6)	209 (3.8)	376 (6.8)	627 (11.3)	143 (6.6)	191 (3.4)	1122 (20.2)
Total population- No (%)	13465985 (6.6)	27729472 (13.6)	6030335 (2.9)	56866416 (27.8)	5365790 (2.6)	2962170 (1.4)	6373455 (3.1)	85643554 (41.9)
HDI-M	0.683	0.696	0.682	0.726	0.59	0.59	0.597	0.618
GDP per capita (USD)	14407	15619	13403	19565	6752	6653	7572	10203
Extremely poor (%)	6.1	5.5	7.6	3.9	23.3	23.3	21.3	18.7
Gini coefficient	0.47	0.49	0.49	0.51	0.53	0.53	0.54	0.54
Unemployment rate (%)	4.5	5.7	5.9	6.6	7.7	7.7	8.7	8.6
Household crowding (%)	16	18.1	18.6	19.8	36.6	38.5	38.5	39.5
Life expectancy at birth (years)	74.1	74.3	73.8	75	70.6	70.6	70.9	71.5
Infant mortality rate (number of deaths in first year of life/1000 live births)	16.1	16.1	17.3	14.8	25.9	26.1	25.2	23.6
Epidemiological indicators								
N of new cases (%)	978 (1.4)	5269 (7.8)	1353 (2)	18865 (27.8)	358 (0.5)	588 (2.9)	2205 (3.3)	38161 (56.3)
TB incidence rate (cases/ 100,000 people)	7.8	22.3	24.5	39.6	7	21.6	30.3	29.2
AIDS case detection rate (cases/100,000 people) ^{c,d}	6.4	11.1	11.9	21.6	5.2	6.4	8.2	10.1
TB mortality rate (deaths/ 100,000 people) ^{c,d}	0.7	1.3	1.6	1.6	1	1.1	2.3	2.1
New cases from at least one vulnerable group (%) ^{c,d,e}	9.9	14.3	19.2	22.6	12.1	15.9	10.6	14.3
TB-HIV confection among new cases (%)	5.9	8	7.7	8.8	6.9	9.9	3.6	5.6
New cases who were prisoners (%)	2.7	4.2	6.7	12.1	2.5	2.7	2.7	4.1
New cases who were health professionals (%)	0.7	1	1.5	1	1.1	0.7	1.3	0.9
New cases who were indigenous population (%)	0.4	0.7	1.6	0.6	1.5	2.7	2.4	3.5
New cases who were homeless (%)	0.7	1.3	1.9	1.4	0.7	1.1	1.1	1.1
Retreatment cases among total cases (%)	13.3	8.1	9.1	13.7	20.7	10.7	9.4	11.1
Operational indicators (new cases)								
Contacts examination (%) ^c	70.6	78.8	54.3	73.3	55.7	36.1	60.5	69.6
Pulmonary cases with laboratory confirmation (%)	65.4	70.4	64.0	71.0	66.5	66.3	66.2	71.4
Tested for HIV (%)	72.1	75.5	67.2	75.3	63.9	52.1	53.5	67.2
Cure (%)	66.8	84.8	32.2	73.1	65.4	14.2	51.1	79.7
Lost to follow-up (%)	5.1	5.8	2.8	8	3.8	1.9	5.2	8.2
No TB outcome registration (%) ^d	14.5	0	60.1	11	22.7	81.4	37	4.6
Culture examination (retreatment) (%)	35.5	39	25.1	44	25.5	26.1	21.5	27.8

^aSocioeconomic indicators: with the exception of the number of municipalities and total population (data from 2015), the other variables were measured in 2010. Epidemiological indicators: with the exception of TB mortality rate (data from 2014), the other variables were measured in 2015. Operational indicators: with the exception of cure, lost to follow-up, no TB outcome registration and culture examination (data from 2014), the other variables were measured in 2015. ^bWithout cases in 2014 or 2015; ^cVariables used in non-hierarchical clustering method for scenario 1; ^dVariables used in k-means method for scenario 2; ^eHIV, health professional, prisoners, indigenous and homeless. Municipal Human Development Index- HDI; Gross domestic product-GDP.

STROBE 2007 (v4) checklist of items to be included in reports of observational studies in epidemiology*

Checklist for cross-sectional studies (combined)

Section/Topic	Item #	Recommendation	Reported on page #
Title and abstract	1	(a) Indicate the study’s design with a commonly used term in the title or the abstract	2 and 3
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	3
Objectives	3	State specific objectives, including any pre-specified hypotheses	3
Methods			
Study design	4	Present key elements of study design early in the paper	3
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	3 and 4
Participants	6	(a) Cohort study—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up Case-control study—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls Cross-sectional study—Give the eligibility criteria, and the sources and methods of selection of participants	3 and 4
		(b) Cohort study—For matched studies, give matching criteria and number of exposed and unexposed Case-control study—For matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	4
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	3 and 4
Bias	9	Describe any efforts to address potential sources of bias	5
Study size	10	Explain how the study size was arrived at	3
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	4
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding	4 and 5
		(b) Describe any methods used to examine subgroups and interactions	4 and 5
		(c) Explain how missing data were addressed	5
		(d) Cross-sectional study—If applicable, describe analytical methods taking account of sampling strategy	-
		(e) Describe any sensitivity analyses	-

Results				
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed		5 and 6
		(b) Give reasons for non-participation at each stage		6 and 7
		(c) Consider use of a flow diagram		-
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders		6 and 7
		(b) Indicate number of participants with missing data for each variable of interest		3
Outcome data	15*	Cross-sectional study—Report numbers of outcome events or summary measures		6-8
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included		6 and 8
		(b) Report category boundaries when continuous variables were categorized		-
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period		-
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses		8 and 9
Discussion				
Key results	18	Summarise key results with reference to study objectives		9-11
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias		11
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence		9-11
Generalisability	21	Discuss the generalisability (external validity) of the study results		11-12
Other information				
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based		14

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at <http://www.plosmedicine.org/>, Annals of Internal Medicine at <http://www.annals.org/>, and Epidemiology at <http://www.epidem.com/>). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

Identifying socioeconomic, epidemiological and operational scenarios for tuberculosis control in Brazil: an ecological study

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2017-018545.R1
Article Type:	Research
Date Submitted by the Author:	27-Dec-2017
Complete List of Authors:	<p>Pelissari, Daniele; Ministerio da Saude, National Tuberculosis Program of Brazil</p> <p>Rocha, Marli; Ministerio da Saude, National Tuberculosis Program of Brazil</p> <p>Bartholomay, Patricia ; Ministerio da Saude, National Tuberculosis Program of Brazil</p> <p>Sanchez, Mauro; Federal University of Brasilia</p> <p>Duarte, Elisabeth; University of Brasília, Faculty of Medicine, Nucleo de Medicina Tropical</p> <p>Arakaki-Sanchez, Denise; Ministerio da Saude, National Tuberculosis Program of Brazil</p> <p>Dantas, Cíntia ; Ministerio da Saude, National Tuberculosis Program of Brazil</p> <p>Jacobs, Marina; Ministerio da Saude, National Tuberculosis Program of Brazil</p> <p>Andrade, Kleydson ; Ministerio da Saude, National Tuberculosis Program of Brazil</p> <p>Codenotti, Stefano ; Ministerio da Saude, National Tuberculosis Program of Brazil</p> <p>Andrade, Elaine Silva ; Ministerio da Saude, National Tuberculosis Program of Brazil</p> <p>de Araújo, Wildo ; University of Brasilia –UnB</p> <p>Costa, Fernanda ; Ministerio da Saude, National Tuberculosis Program of Brazil</p> <p>Ramalho, Walter; Universidade de Brasilia, Faculdade de Ceilândia;</p> <p>Universidade de Brasília, Núcleo de Medicina Tropical/Faculdade de Medicina</p> <p>Díaz-Quijano, Fredi; School of Public Health, University of São Paulo, Department of Epidemiology</p>
Primary Subject Heading:	Public health
Secondary Subject Heading:	Epidemiology, Health policy, Infectious diseases
Keywords:	Tuberculosis < INFECTIOUS DISEASES, EPIDEMIOLOGY, public health policy



For peer review only

Title: Identifying socioeconomic, epidemiological and operational scenarios for tuberculosis control in Brazil: an ecological study

*Daniele Maria Pelissari, National Tuberculosis Program, Ministry of Health. Address: SRTVN 701, Via W 5 Norte, Ed. PO 700 – 6º andar. Zip code: 70719-040, Brasília-DF, Brazil. Email: daniele.pelissari@saude.gov.br; daniele.pelissari@gmail.com. Telephone: (+55)61998431554

Marli Souza Rocha, National Tuberculosis Program, Ministry of Health. Brasília/DF, Brazil.

Patricia Bartholomay, National Tuberculosis Program, Ministry of Health. Brasília/DF, Brazil.

Mauro Niskier Sanchez, Department of Public Health, University of Brasilia – UnB. Brasília/DF, Brazil.

Elisabeth Carmen Duarte, Medical School, University of Brasilia –UnB. Brasília/DF, Brazil.

Denise Arakaki-Sanchez, National Tuberculosis Program, Ministry of Health. Brasília/DF, Brazil.

Cíntia Oliveira Dantas, National Tuberculosis Program, Ministry of Health. Brasília/DF, Brazil.

Marina Gasino Jacobs, National Tuberculosis Program, Ministry of Health. Brasília/DF, Brazil.

Kleydson Bonfim Andrade, National Tuberculosis Program, Ministry of Health. Brasília/DF, Brazil.

Stefano Barbosa Codenotti, National Tuberculosis Program, Ministry of Health. Brasília/DF, Brazil.

Elaine Silva Nascimento Andrade, National Tuberculosis Program, Ministry of Health. Brasília/DF, Brazil.

Wildo Navegantes de Araújo, Faculty of Ceilandia, University of Brasilia –UnB. Brasília/DF, Brazil.

Fernanda Dockhorn Costa, National Tuberculosis Program, Ministry of Health.

Walter Massa Ramalho, Faculty of Ceilandia, University of Brasilia –UnB. Brasília/DF, Brazil.

Fredi Alexander Diaz-Quijano, Department of Epidemiology, School of Public Health, University of São Paulo. São Paulo/SP, Brazil.

*Corresponding author

Role of the funding source: This study did not have specific funding.

Word count: 3,311

Number of figures and tables: 3 tables; 2 figures; supplementary material-1: 1 table; supplementary material-2: Access to data used in this study

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

ABSTRACT

Objectives We aimed to identify socioeconomic (2010), epidemiological and operational health care indicators (2014/2015) associated with tuberculosis incidence in Brazil.

Design Ecological study.

Settings The study was based on new tuberculosis cases and epidemiological/operational indicators of the disease from the Brazilian National Information System for Notifiable Diseases (SINAN) and the Mortality Information System (SIM). We also analysed socioeconomic and demographic indicators.

Participants The unit of analysis was the Brazilian municipalities, which in 2015 were 5570 and registered 67 777 new TB cases.

Primary and secondary outcome measures The tuberculosis incidence rate in 2015 was the primary outcome. We analysed as independent variables the socioeconomic indicators (2010), epidemiological and operational health care indicators of tuberculosis (2014 or 2015) using the negative binomial regression. Municipalities were clustered by the k means method considering the variables identified in multiple regression models.

Results We identified two clusters according to socioeconomic indicators associated with tuberculosis incidence rate (unemployment rate and household crowding): a higher socioeconomic scenario (n= 3482 municipalities) with a mean tuberculosis incidence rate of 16.3/100 000 population; and a lower socioeconomic scenario (2083 municipalities) with a mean tuberculosis incidence rate of 22.1/100 000 population. In a second-stage of clusterization we defined four subgroups in each of the socioeconomic scenarios using epidemiological and operational indicators such as tuberculosis mortality rate, AIDS case detection rate and proportion of vulnerable population among tuberculosis cases. Some of the sub-scenarios identified were characterized by fragility in their information systems, while others by the concentration of tuberculosis cases in key populations.

Conclusion Clustering in scenarios allowed the classification of municipalities according to the socioeconomic, epidemiological and operational indicators associated with tuberculosis risk. This classification can support targeted evidence-based decisions such as monitoring data quality for improving the information system, or to establish integrative social protective policies for key populations.

Key word: tuberculosis, epidemiology, public health policy.

Strengths and limitations of this study

- This study was based on national population data in a country of continental dimension (5,565 municipalities).
- The availability of indicators associated with tuberculosis, made it possible to consider both socioeconomic and epidemiological/operational approaches in the definition of municipality clusters for tuberculosis control.
- This methodology can be explored by other countries in order to guide their plans to end tuberculosis.
- Reporting and information quality may vary between sources and periods, which could affect estimate accuracy.
- Inferences obtained are applicable to population groups, not to individuals. However, ecological study can provide evidence to support public health decisions.

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

INTRODUCTION

In 2016, 10.4 million people had tuberculosis (TB) and 1.8 million died worldwide because of the disease.¹ In Brazil, similarly to other countries, TB incidence reduction (37.9/100 000 population in 2007 to 32.4/100 000 population in 2016)² seems to be associated with the improvement of the population living conditions^{3–5} and the performance of TB control programs.⁶ However, the disease burden continues to be significant in the country, with 66 796 new cases registered in 2016.²

In 2014, in a move towards elimination, the World Health Organization (WHO) launched the End TB Strategy, setting targets to be met by 2035, including a 90% reduction in TB incidence as compared to 2015.¹ The strategy is critical to energizing the fight against the disease and mobilizing resources, but needs to be adapted to the local context as any other health policy.⁷

Some countries have already made progress developing their national plans. Among the strategies presented, we highlight the strengthening of existing TB services, the acceleration of the detection of cases in key populations, and the implementation of actions to reduce the barriers of TB care.^{8–10}

Brazil is a country with continental dimensions, thus both socioeconomic indicators¹¹ and those that reflect the performance of local TB programs² present a high degree of heterogeneity. Considering this context, and in order to support the "National Plan to End TB",¹² we identified scenarios based on socioeconomic, epidemiological, and operational factors associated with the TB incidence rate.

METHODS

Type of study and data source

This is an ecological study, with the unit of analysis being the Brazilian municipalities, which were 5570 in 2015. We excluded five municipalities due to absence of socioeconomic information. Data on socioeconomic and demographic indicators were obtained from the last population census (2010).^{11,13} As for new TB cases (2015) and epidemiological/operational indicators of the disease (2014 and 2015), we used data from the Brazilian National Information System for Notifiable Diseases (SINAN) and the Mortality Information System (SIM).¹¹

Variables

The dependent variable was the TB incidence rate (/100 000) in 2015 and the independent variables were socioeconomic, epidemiological, and health care operational TB indicators. Many of these indicators have already been identified in previous studies as TB determinants.^{14–16}

The socioeconomic indicators analysed were:

- Municipal human development index (HDI-M);
- Average household income per capita;

- Gross Domestic Product (GDP) per capita;
- Proportion of the population that is extremely poor, poor, and vulnerable to poverty;
- Gini coefficient;
- Unemployment rate;
- Illiteracy rate;
- Proportion of the population living in households with more than two people per room represented by household crowding;
- Infant mortality rate per 1,000 live births;
- Life expectancy at birth.
- Population size of municipalities classified as small (less than 20 000 inhabitants), medium (20 000 to 99 999 inhabitants) and large (100 000 inhabitants or larger).¹⁷

Average household income per capita and GDP per capita were converted into US dollars (USD) using the average annual price in 2010 (1 USD \approx 1.76 Brazilian Reals- R\$). We adopted the Brazilian definitions for the proportions of the population that is extremely poor, poor and vulnerable to poverty: proportion of individuals in the municipality with average household income per capita equal or less than USD 40, USD 80 and USD 145, respectively.¹³

The epidemiological indicators of TB were:

- AIDS case detection rate per 100 000 population;
- Proportion of new TB cases who were: HIV positive; prisoners; health professionals; indigenous; homeless; and, as a composite definition, proportion of TB cases from at least one of those vulnerable groups. Those indicators were previously associated with an increased risk of TB in other studies,^{3-5,16,18}
- Proportion of TB retreatment;
- TB mortality rate per 100 000 population;

The operational health care indicators of TB considered in the analysis were:

- Proportions of new TB cases: in which contacts were examined; laboratory confirmed; tested for HIV; and treatment outcomes (cure, lost to follow-up, and no record of TB outcome);
- Proportion of sputum culture examination among retreatment cases;

Due to the availability of updated data at the time of analysis, data to calculate culture examination, treatment outcomes, and TB mortality rate refers to 2014, while the other indicators refers to 2015.

1

2

3

4

5 **Statistical analysis**

6

7 Statistical analysis was performed in two stages; each of them included a model to identify the

8 factors associated with TB incidence rate. This was followed by a cluster analysis based on the

9 factors identified. The first stage was focused on socioeconomic variables and the second on the

10 epidemiological and operational indicators associated with TB incidence rate.

11

12 Negative binomial regression was used to identify factors associated with the TB incidence rate

13 in 2015. For these regression analyses we only included municipalities that presented mean

14 annual variation of the triennial moving average of the incidence rate for the years 2001 to 2015

15 between -8 and 8%. By doing so, we intended to reduce possible biases due to the variability of

16 values in small municipalities, and possible intermittence in case reporting.

17

18

19 Independent variables associated with the TB incidence rate, with a p value <0.20 in simple

20 regression models, were analysed in a Spearman correlation matrix. Whenever correlation

21 between independent variables ($r>0.50$) was identified, we selected the variable with the highest

22 association with TB incidence rate. We used a stepwise forward selection method and preserved

23 the variables with a p value <0.05 in the multiple model. In addition, we adjusted the models by

24 the population size of municipalities. We presented the association measures as the relative

25 increment in the incidence rate.

26

27

28 Socioeconomic indicators associated in a multiple model (primary model) were considered for a

29 cluster analysis of all municipalities using the non-hierarchical k-means method. In this method,

30 the algorithm aims to reduce intra-group variance and maximize inter-group variance in relation

31 to the Euclidean distance established by the indicators selected.¹⁹ To define the number of

32 clusters, we used the Elbow method which relates the number of clusters with the percentage of

33 internal variation of the groups,¹⁹ adopting > 60% as the cut-off point and among these, the

34 smallest number of possible clusters.

35

36

37 For the second stage, epidemiological/operational indicators were modeling for each

38 socioeconomic scenarios, following a similar methodology described for socioeconomic

39 variables. Factors associated with TB in these secondary models, as well as, TB mortality rate,

40 were considered for a second cluster analysis, which subdivided the previous socioeconomic

41 clusters into epidemiological/operational TB sub-scenarios. Because some operational indicators

42 are only measured during care of TB cases, these second stage methods were applied only in

43 municipalities with TB cases reported in 2014 and 2015.

44

45

46 Statistical analyses were performed with the Stata statistical package version 12.0, R version

47 3.3.1 and the cluster library.

48

49 Data used in this study do not include personal identification of TB patients. Additionally, all

50 data analysed are publicly available in Brazil. According to local legislation (Resolution No. 510

51 of the National Health Council of Brazil),²⁰ research conducted exclusively with publicly data is

52 not evaluated by an institutional review board. However, this study was conducted according to

53 guidelines and standards for research involving human subjects.²¹

54

RESULTS

In 2015, 67 777 new TB cases were reported in Brazil, with an incidence rate of 33.1/100 000 population. The mean annual variation of the triennial moving average of the TB incidence rate in municipalities ranged from -22.6% to 41.8%. This interval was wider in small and medium municipalities (-22.7% to 41.9%) when compared to the larger ones (-7.3% to 14.6%). A total of 3311 (59.5%) municipalities presented a variation between -8% and 8% and were eligible for the analysis for the primary model, including 791 that did not present new TB cases in 2015.

With the exception of the unemployment rate, all remaining variables associated with the outcome ($p < 0.20$) were strongly correlated with each other (Spearman coefficient > 0.50). Considering that household crowding had a greater association with the TB incidence rate, the first model included this variable and the unemployment rate, adjusted by municipal population size (Table 1).

Table 1. Socioeconomic variables and association with tuberculosis incidence rate in Brazil (n= 3311 municipalities^a)

Variable ^b	Mean (standard deviation-SD)	Median (IQ25%-IQ75%)	Relative increment of IR (95% CI) ^c	Relative increment of IR adjusted (95% CI) ^c
HDI-M	0.7 (0.1)	0.7 (0.6 - 0.7)	-4.8 (-36.9 to 43.7)	
Average household income per capita (USD)	280.3 (143.8)	257.9 (155.6 - 372.8)	-0.0 (-0.0 to -0.0) ^d	
GDP per capita (USD)	7,510.4 (8630.3)	5,555.2 (2909.4 - 9091.2)	0.0 (-0.0 to 0.0)	
Extremely poor (%)	11.4 (11.7)	6.5 (1.6 - 19.1)	-0.0 (-0.3 to 0.2)	
Poor (%)	23.4 (18.0)	18.5 (6.9 - 38.8)	0.1 (-0.0 to 0.3)	
Vulnerable to poverty (%)	44.1 (22.7)	42.6 (23.3 - 65.6)	0.2 (0.0 to 0.3) ^d	
Gini coefficient	0.5 (0.1)	0.5 (0.5 - 0.5)	116.0 (37.1 to 240.3) ^d	
Unemployment rate (%)	6.7 (3.7)	6.3 (4.2 - 8.6)	5.0 (4.2 to 5.8) ^d	3.9 (3.0 to 4.7) ^d
Illiteracy in population with ≥ 18 years (%)	17.2 (10.8)	13.9 (8.1 - 26.4)	-0.4 (-0.6 to -0.1) ^d	
Illiteracy in population with ≥ 15 years (%)	15.6 (9.8)	12.9 (7.2 - 23.8)	-0.4 (-0.7 to -0.1) ^d	
Household crowding (%) ^e	26.4 (13.1)	24.7 (16.6 - 33.8)	1.2 (1.0 to 1.4) ^d	0.8 (0.6 to 1.1) ^d
Infant mortality rate (number of deaths in first year of life/1000 live births)	19.3 (7.2)	17.0 (13.7 - 24.1)	0.1 (-0.3 to 0.5)	
Life expectancy at birth (years)	73.1 (2.7)	73.4 (71.1 - 75.2)	-0.9 (-1.9 to 0.2)	

Abbreviations: HDI-M, municipal human development index; GDP, gross domestic product; SD, standard deviation; IQ, interquartile; IR, incidence rate; CI, confidence interval.

^aMunicipalities with annual variation in TB incidence rate between -8% and 8% and at least one TB incidence case in 2015.

^bVariables measured in the last census (2010).

^cThe association measure represents the relative increment in the incidence rate (IRR-1), adjusted by the population size of the municipality.

^dp value < 0.05 .

^eProportion of the population living in households with more than two people per room.

Based on these two socioeconomic variables, we identified a higher socioeconomic scenario (HSS) cluster, with 3482 municipalities, which presented better socioeconomic indicators when compared to the 2083 municipalities from the second cluster (Table 2), the lower socioeconomic

scenario (LSS). The HSS cluster exhibited unemployment rates of up to 26.87%; and, household crowding values between 0.65% and 28.56%. On the other hand, the LSS cluster exhibited unemployment rates of up to 39.15%; and, household crowding values between 26.65% and 88.64% (Fig. 1).

Figure 1. Distribution of Brazilian municipalities according to socioeconomic variables associated with the tuberculosis incidence rate

The mean TB incidence rate in the LSS was 22.1/100 000 population, which was significantly higher than that observed in the HSS, which was 16.3/100 000 population (IRR: 1.34; 95% CI: 1.26 - 1.41).

Table 2. Description of socioeconomic scenarios associated with tuberculosis incidence rate in Brazil (n= 5565 municipalities)^a

Variables ^b	Higher socioeconomic scenario n= 3482 municipalities		Lower socioeconomic scenario n= 2083 municipalities	
	Mean (SD)	Median (IQ25%-IQ75%)	Mean (SD)	Median (IQ25%-IQ75%)
HDI-M	0.7 (0.1)	0.7 (0.6 - 0.7)	0.6 (0.1)	0.6 (0.6 - 0.6)
Average household income per capita (USD)	330.5 (126.5)	324.1 (242.3 - 403.5)	181.3 (92.4)	150.8 (125.3 - 203.5)
GDP per capita (USD)	8661.1 (7707.6)	7259.7 (4758.8 – 10 053.9)	4930.5 (8033.6)	2937.6 (2371.3 - 5004.5)
Extremely poor (%)	5.8 (7.3)	2.6 (1.1 - 7.4)	20.7 (11.9)	20.3 (12.0 - 28.7)
Poor (%)	14.2 (12.3)	9.7 (5.1 - 20.2)	38.3 (15.5)	40.3 (29.7 - 49.1)
Vulnerable to poverty (%)	33.0 (17.8)	29.3 (19.0 - 44.9)	62.4 (16.5)	67.0 (57.0 - 73.5)
Gini coefficient	0.5 (0.1)	0.5 (0.4 - 0.5)	0.5 (0.1)	0.5 (0.5 - 0.6)
Unemployment rate (%) ^c	5.1 (2.9)	4.9 (3.1 - 6.8)	8.3 (4.0)	7.6 (5.6 - 10.3)
Illiteracy in population with ≥ 18 years (%)	12.9 (8.0)	10.7 (7.2 - 16.0)	25.0 (10.3)	26.5 (17.1 - 32.9)
Illiteracy in population with ≥ 15 years (%)	11.7 (7.3)	9.8 (6.5 - 14.6)	22.7 (9.5)	23.9 (15.6 - 29.8)
Household crowding (%) ^{c, d}	17.1 (6.0)	17.3 (12.7 - 22.0)	38.5 (10.3)	35.7 (31.2 - 41.9)
Infant mortality rate (number of deaths in first year of life/1000 live births)	16.0 (5.0)	14.8 (12.8 - 17.5)	24.6 (6.9)	24.2 (19.4 – 29.0)
Life expectancy at birth (years)	74.3 (2.1)	74.5 (73.2 - 75.7)	71.1 (2.3)	71.1 (69.6 - 72.6)

Abbreviations: HDI-M, municipal human development index; GDP, gross domestic product; SD, standard deviation; IQ, interquartile.

^aTotal of municipalities with socioeconomic data in Brazil were used in clusterization step.

^bWith exception of TB incidence rate (2015), the other variables were measured in the last census (2010).

^cVariables identified during step 1 (model 1) used in step 2 with k-means method.

^dProportion of the population living in households with more than two people per room.

Among the 3482 HSS municipalities, 1125 had TB cases in 2014 and 2015, and were eligible (annual variation in TB incidence rate between -8% and 8%) for a secondary modelling. In this analysis, the AIDS case detection rate and the proportion of new cases from at least one vulnerable group were positively associated with the TB incidence rate, while the proportion of contacts investigation among new TB cases presented an inverse association (Table 3).

As regards the LSS, 1095 municipalities out of 2083 were eligible. The AIDS case detection rate and the proportion of cases from at least one vulnerable group were positively associated with TB incidence rate. The proportion of cases with no record of TB outcome was inversely associated (Table 3).

Table 3. Epidemiological and operational tuberculosis indicators associated with tuberculosis incidence rate stratified by socioeconomic scenarios in Brazil

Variables ^a	Higher socioeconomic scenario n= 1125 municipalities ^b		Lower socioeconomic scenario n= 1095 municipalities ^b	
	Relative increment of IR (95% CI) ^c	Relative increment of IR adjusted (95% CI) ^c	Relative increment of IR (95% CI) ^c	Relative increment of IR adjusted (95% CI) ^c
Epidemiological indicators				
AIDS case detection rate (cases/100 000 population)	1.5 (1.2 to 1.7) ^e	1.4 (1.1 to 1.6) ^e	2.1 (1.7 to 2.5) ^e	2.0 (1.6 to 2.4) ^e
New cases from at least one vulnerable group (%) ^d	0.5 (0.3 to 0.7) ^e	0.2 (0.1 to 0.4) ^e	0.7 (0.5 to 0.9) ^e	0.5 (0.3 to 0.7) ^e
TB-HIV confection among new cases (%)	0.3 (0.0 to 0.6) ^e		-0.2 (-0.5 to 0.2)	
New cases who were prisoners (%)	0.7 (0.4 to 0.9) ^e		1.2 (0.9 to 1.5) ^e	
New cases who were health professionals (%)	-0.5 (-1.3 to 0.2)		-0.2 (-1.0 to 0.6)	
New cases who were indigenous population (%)	1.1 (0.3 to 1.9) ^e		0.9 (0.5 to 1.2) ^e	
New cases who were homeless (%)	0.1 (-0.6 to 0.7)		0.1 (-0.7 to 1.0)	
Retreatment cases among total cases (%)	0.5 (0.2 to 0.8) ^e		0 (-0.3 to 0.3)	
Operational indicators (new TB cases)				
Contacts examination (%)	-0.3 (-0.4 to -0.1) ^e	-0.2 (-0.3 to -0.1) ^e	-0.0 (-0.2 to 0.1)	
Pulmonary cases with laboratory confirmation (%)	0.0 (-0.1 to 0.2)		-0.1 (-0.3 to 0.0)	
Tested for HIV (%)	0.1 (-0.0 to 0.2)		0.0 (-0.1 to 0.2)	
Cure (%)	-0.2 (-0.3 to -0.0) ^e		0.2 (0.0 to 0.3) ^e	
Lost to follow-up (%)	0.6 (0.3 to 0.9) ^e		0.3 (0.0 to 0.7) ^e	
No TB outcome registration (%)	0.3 (0.1 to 0.5) ^e		-0.3 (-0.5 to -0.1) ^e	-0.3 (-0.5 to -0.1) ^e
Culture examination (retreatment) (%)	0.1 (-0.1 to 0.2)		0.0 (-0.1 to 0.2)	

Abbreviations: IR, incidence rate; CI, confidence interval.

^aWith exception of: cure, lost to follow-up, no record of TB outcome and culture examination (2014), the other variables were measured in 2015.

^bMunicipalities with annual variation in TB incidence rate between -8% and 8% and at least one TB incidence case in 2014 and 2015.

^cThe association measure represents the relative increment in the incidence rate (IRR-1), adjusted by the population size of the municipality.

^dHIV, health professional, prisoners, indigenous, and homeless.

^ep value < 0.05.

Using the variables associated with the outcome in the previous models, and also considering the TB mortality rate, we defined three clusters for each socioeconomic scenario, totaling six sub-scenarios with TB cases in 2014 and 2015. For each scenario, a sub-scenario (1.0 and 2.0) was also defined including municipalities without TB reporting in 2014 or 2015 (supplementary material-1).

Figure 2 shows the geographical distribution of municipalities according to the sub-scenarios. Regarding the sub-scenarios with TB cases in HSS, 1.1 showed the lowest mean rates of TB incidence, AIDS case detection and TB mortality. Sub-scenario 1.2, despite having relatively low mean rates of TB incidence, AIDS case detection, and TB mortality, had a high proportion of

cases with no record of TB outcome. Sub-scenario 1.3 covers 27.8% of new TB cases reported in 2015 and presents the highest mean rates of TB incidence, AIDS case detection, and proportion of cases from at least one vulnerable group (22%) (supplementary material-1).

Regarding the LSS municipalities, the sub-scenario 2.1 had the highest mean proportion of cases with no record of TB outcome (81.8%) and the lowest mean for contact investigation (36.5%) and HIV testing (52.3%). 2.2 shows a high TB incidence rate, the highest TB mortality and it has a high mean proportion of cases with no record of TB outcome (37%) and low HIV testing (53.5%). As a consequence of the inclusion of 14 capitals in the sub-scenario 2.3, it includes 56.3% of all new cases reported in 2015. It has the highest mean AIDS case detection rate in the group of LSS and the second highest TB mortality rate among all sub-scenarios (supplementary material-1).

Figure 2- Brazilian’s municipalities by the TB incidence rate scenarios. Brazil, 2015.

DISCUSSION

This study classified the 5565 Brazilian municipalities in two scenarios using socioeconomic variables associated with the TB incidence rate in Brazil. After that, we performed a sub classification based on operational and epidemiological indicators associated with the TB incidence rate.

Regarding socioeconomic indicators, the unemployment rate was associated with the risk of TB, as found in previous studies from the United States,²² Spain,²³ and also in Brazil¹⁶. At the individual level, unemployment has been associated with an increased risk of alcohol and illicit drugs abuse;²⁴ and also with lost to follow-up during HIV treatment.²⁵ These factors have already been associated with TB risk³⁻⁵ and could, at least partially, explain the association observed in our study.

Household crowding was also positively associated with the TB incidence rate. In the United States and West Africa people living in crowding conditions had a higher risk of TB.^{26,27} In Brazil, this variable was already associated with TB incidence and analysed as a potential mediator between socioeconomic determinants and TB incidence rate, because it may directly favors TB transmission by increasing the contact rate between infected and susceptible people.¹⁶

In our study, these indicators defined the two socioeconomic scenarios. The LSS, with municipalities predominantly in the North, Northeast, and Center-West regions, presented a higher incidence of TB than those of the HSS, with municipalities located predominantly in the South and Southeast regions. This suggested that classification of municipalities by socioeconomic variables could be highly functional to address the TB risk.

Regarding the operational and epidemiological indicators, the AIDS case detection rate was positively associated with the TB incidence rate in both of socioeconomic scenarios, which was consistent with previous studies in which AIDS has been a factor associated with TB risk at the contextual level.^{15,16,28}

The proportion of new cases from at least one vulnerable group was also another factor associated with TB incidence in both scenarios. One of the vulnerable populations included are the prisoners. Specifically in Brazil, in 2014 there were approximately 607 thousand imprisoned people in 956 municipalities distributed in all regions of the country, with a prison occupation rate of 161%.²⁹ This overcrowding may explain the high risk shown by this group in previous studies³⁰ and makes it a priority vulnerable group for TB control.

In the HSS, vulnerability was also correlated with the AIDS case detection rate, which is higher in the South and Southeast regions (respectively 20.1% and 53.0% of the AIDS cases identified from 1980 to June 2016).³¹ Regarding the LSS, vulnerability was correlated with indigenous populations, which are predominantly located in the North (37.4%), Northeast (25.5%), and Central-West regions (16.0%)¹¹ and present a higher risk of TB when compared to other races.⁴

We observed an inverse association between TB incidence rate and the percentage of contact investigation in the HSS, which may represent the overall effect on transmission control, possibly through the identification and timely treatment.³² Finally, in the LSS the association with the proportion of cases with no record of TB outcome may represent failures in surveillance in collecting these data for the qualification of the information system.

Regarding the absence of TB cases in 2014 or 2015 in sub-scenarios 1.0 and 2.0, it is possible that there is an under reporting in these scenarios, mainly in the 2.0, where there are worse socioeconomic conditions, which were associated with a higher risk of TB. This suggests that activities related to TB detection should be strengthened especially in those groups of municipalities.

Concerning the sub-scenarios that reported cases in the two years of analysis from the HSS cluster, group 1.1 has the lowest TB incidence rate, better socioeconomic indicators, and good TB epidemiological/operational indicators, suggesting an advanced stage in TB control. Sub-scenario 1.3 presents the highest TB incidence rate, AIDS case detection rate and proportion of cases from at least one vulnerable group (22%), especially among prisoners (12.1%). In addition, this scenario is composed mainly by capitals, which could mean a more sensitive surveillance system. Despite being the group of municipalities with the highest TB risk, the distribution of vulnerabilities reveals concentrated epidemic in some population groups, which requires distinct and focused strategies to control the disease. Sub-scenario 1.2 (HSS-cluster), 2.1 and 2.2 (LSS-cluster) need improvement in the information system because of the high proportion of cases with no record of TB outcome. This makes it difficult to analyze the performance of TB control

actions. Another challenge in these groups is the investigation of contacts which was particularly low in the sub-scenario 2.1. Although sub-scenario 2.1 has the highest percentage of TB-HIV coinfection in new cases (9.8%), it also has one of the lowest percentages of HIV testing (52.3%), suggesting the under detection of HIV among people with TB.

In the LSS, group 2.2 exhibited the highest incidence of TB but the lowest proportion of cases from at least one vulnerable group (10.6%), revealing an endemic situation that is less concentrated in vulnerable populations. Sub-scenario 2.3 has a reliable information system and good performance in operational activities (e.g. contact investigation and HIV testing), revealing that even with limited resources, it is possible to carry out effective disease control actions.

Finally, with the exception of sub-scenario 2.0, all those in the LSS had a higher TB mortality rate than those in the HSS. Sub-scenario 2.0, although it did not report new cases in 2014 or 2015, exhibited a higher mortality rate than the 1.0 group. Mortality is expected to be less underreported than incidence, as observed in other diseases.^{33,34} Thus, the use of this variable for defining clusters contributes to identifying groups according to TB burden besides to the other variables used for classification.

Limitation

As a common limitation of ecological studies, aggregate measures might differ from individual ones. However, these studies provide an overview that contributes to direct decision-making in public policies.

Underreporting of TB cases in Brazil is decreasing each year,¹ but may remain a potential limitation for this study. Since there is no information about TB case detection and latent TB infection in Brazilian municipalities, the overall burden cannot be estimated. Even so, we hypothesize that the underreporting is either homogeneous or higher in municipalities with worst socioeconomic indicators. Therefore, the magnitude of association between socioeconomic indicators and TB incidence may be higher than estimated in this study. The exclusion of municipalities that presented high variability in the incidence rate may have reduced the risk of information bias.

On other hand, although an important number of municipalities was excluded from the regression analysis, those localities were usually small, and the overall municipalities included made up 87.2% of the Brazilian population. In addition, only five municipalities (0.1% of the total) were excluded because of the absence of socioeconomic data. Therefore, we consider that the association identified in the multiple models can be widely extrapolated.

Concerning data availability, socioeconomic indicators were only available from the last census conducted in the country (2010). Therefore, recent socioeconomic trends and their impact on the current TB incidence rate could not be evaluated. However, we believe that the socioeconomic differences between municipalities have remained proportional in recent years, which allows their evaluation as a determinant of the TB incidence.

Implications for public health and conclusion

The End TB Strategy proposes bold targets, and a prompt response from each country can be critical for their achievement. We consider this work an innovative action for health public decisions, because we used secondary data available for most of the municipalities of the country with a robust data analysis, that recognizes the socioeconomic and operational diversity of a continental country. The grouping of municipalities presented in this study has already been applied in the National Plan to End¹² to support the implementation of efficient strategies.

Efforts should be focused on strengthening information systems to provide a reliable picture of the epidemiological situation, such as the implementation of monitoring strategies to ensure the quality of data collection.

The challenges of controlling TB in key populations require integrative collaboration with other government sectors. Establish social protective policies may contribute to the reduction of TB risk in especial groups such as prisoners and indigenous.

Municipalities in the LSS, besides additional resources, requires actions to reduce the exacerbation of social vulnerabilities, which are reflected in TB risk. That is why TB should be considered a priority in the public health agenda. In addition, municipalities from LSS scenarios that did not have record of TB case in 2015/2014 should focus on activities related to TB detection, especially active case find.

The heterogeneity of the socioeconomic and epidemiological situation in Brazil, observed in this study, represents a great challenge for TB control in a country of continental proportions, which may also be the reality of other countries. In this sense, our data analysis approach could be considered by other countries with available indicators in order to identify sub-scenarios to guide targeted actions to TB control.

CONTRIBUTORS

Daniele Maria Pelissari: conceived the study, contributed to the design of the study and the interpretation of results, coordinated and analysed the data including epidemiological models and clusterization, wrote the first draft of the report, prepared the illustrations, wrote the discussion, critically reviewed the manuscript and approved the final version.

Marli Souza Rocha: conceived the study, contributed to the design of the study and the interpretation of results, analysed the data including epidemiological models and clusterization,

contributed to the discussion, critically reviewed the manuscript and approved the final version.

Patricia Bartholomay: conceived the study, contributed to the design of the study and the interpretation of results, analysed the data including epidemiological models and clusterization, contributed to the discussion, critically reviewed the manuscript and approved the final version.

Mauro Niskier Sanchez: contributed to the design of the study and the interpretation of results, provided significant inputs in the first draft, contributed to the discussion, critically reviewed the manuscript and approved the final version.

Elisabeth Carmen Duarte: contributed to the design of the study and the interpretation of results, provided significant inputs in the first draft, contributed to the discussion, critically reviewed the manuscript and approved the final version.

Denise Arakaki-Sanchez: contributed to the design of the study and the interpretation of results, contributed to the discussion, critically reviewed the manuscript and approved the final version.

Cíntia Oliveira Dantas: contributed to the design of the study and the interpretation of results, contributed to the discussion, critically reviewed the manuscript and approved the final version.

Marina Gasino Jacobs: contributed to the design of the study and interpretation of results, contributed to the discussion, critically reviewed the manuscript and approved the final version.

Kleydson Bonfim Andrade: contributed to the design of the study and the interpretation of results, contributed to the discussion, critically reviewed the manuscript and approved the final version.

Stefano Barbosa Codenotti: contributed to the design of the study and the interpretation of results, contributed to the discussion, critically reviewed the manuscript and approved the final version.

Elaine Silva Nascimento Andrade: contributed to the design of the study and results, contributed to the discussion, critically reviewed the manuscript and approved the final version interpretation.

Wildo Navegantes de Araújo: contributed to the design of the study and the interpretation of results, contributed to the discussion, critically reviewed the manuscript and approved the final version.

Fernanda Dockhorn Costa: contributed to the design of the study and the interpretation of results, contributed to the discussion, critically reviewed the manuscript and approved the final version.

Walter Massa Ramalho: contributed to the design of the study and the interpretation of results, contributed to the discussion, critically reviewed the manuscript and approved the final version.

Fredi Alexander Diaz-Quijano: contributed to the design of the study and the interpretation of results, provided significant inputs in the drafts, prepared the illustrations, contributed to the discussion, critically reviewed the manuscript and approved the final version.

DATA SHARING STATEMENT

Contextual data are available from the Brazilian Health Ministry website (www.datasus.gov.br/tabnet/tabnet.htm); Brazilian Institute of Geography and Statistics website (<http://www.ibge.gov.br>); and the Human Development Atlas in Brazil website (<http://atlasbrasil.org.br/2013/>). Tuberculosis case data can be made available by Brazilian Health Ministry (<http://portalsaude.saude.gov.br>). More detailed information about how to access data is described at Supplementary Material-2.

FUNDING

This study did not have specific funding.

COMPETING INTERESTS

None declared.

References

1. World Health Organization. Global tuberculosis report 2017. Geneva: World Health Organization, 2017. Available from: <http://apps.who.int/iris/handle/10665/254762> [cited 2017 Dec 9].

2. Ministério da Saúde. Indicadores prioritários para o monitoramento do Plano Nacional pelo Fim da Tuberculose como Problema de Saúde Pública no Brasil. Bol. Epidemiológico 2017;48:1-6.

3. San Pedro A, Oliveira RM de. Tuberculose e indicadores socioeconômicos: revisão sistemática da literatura. Rev Panam Salud Publica 2013;33:294–301.

4. Lacerda SNB, Temoteo RC de A, Figueiredo TMRM de, Luna DT de L, de Sousa AN, Abreu C, et al. Individual and social vulnerabilities upon acquiring tuberculosis: a literature systematic review. Int. Arch. Med. 2014 Jul;7(35):1-8. doi: <http://dx.doi.org/10.1186/1755-7682-7-35>.

5. Millet J-P, Moreno A, Fina L, et al. Factors that influence current tuberculosis epidemiology. Eur. Spine J. 2013;22(4):539–48. doi: <http://dx.doi.org/10.1007/s00586-012-2334-8>.

6. Houben RMGJ, Menzies N, Sumner T, et al. Feasibility of achieving the 2025 WHO global tuberculosis targets in South Africa, China, and India: a combined analysis of 11 mathematical models. Lancet Glob. Health 2016;4:e806–15. doi: [http://dx.doi.org/10.1016/S2214-109X\(16\)30199-1](http://dx.doi.org/10.1016/S2214-109X(16)30199-1).

7. Baicker K & Chandra A. Evidence-Based Health Policy. N. Engl. J. Med. 2017;377:2413–5.

8. Shin Y, Ki M, Sung N. A new strategy for tuberculosis control in North Korea. Epidemiol. Health 2015;37:e2015053. doi: <http://dx.doi.org/10.4178/epih/e2015053>.

9. Sharma DC. New plan to end tuberculosis in south and southeast Asia. The Lancet 2017;25;389(10075):1183. doi: [http://dx.doi.org/10.1016/S0140-6736\(17\)30817-6](http://dx.doi.org/10.1016/S0140-6736(17)30817-6).

10. Pai M, Correa N, Mistry N, et al. Reducing global tuberculosis deaths-time for India to step up. The Lancet 2017;25;389(10075):1174-6. doi: [http://dx.doi.org/10.1016/S0140-6736\(17\)30790-0](http://dx.doi.org/10.1016/S0140-6736(17)30790-0).

11. Ministério da Saúde. Departamento de Informática do SUS: Informações de saúde (TABNET). Brasília: Ministério da Saúde, 2017. Available from: <http://www.datasus.gov.br/informacoes-de-saude/tabnet> [cited 2017 Mar 4].

12. Ministério da Saúde. Plano Nacional pelo Fim da Tuberculose como Problema de Saúde Pública. Brasília: Ministério da Saúde, 2017. Available from: <https://drive.google.com/file/d/0B0CE2wqdEaR-eVc5V3cyMVFPcTA/view> [cited 2017 Dec 9].

13. Programa das Nações Unidas para o Desenvolvimento – PNUD. Atlas do Desenvolvimento Humano no Brasil. Programa das Nações Unidas para o Desenvolvimento – PNUD, 2015. Available from: <http://www.atlasbrasil.org.br/2013/pt/consulta/> [cited 2017 Mar 4].
14. Janssens J-P, Rieder HL. An ecological analysis of incidence of tuberculosis and per capita gross domestic product. *Eur. Respir. J.* 2008;32(5):1415-6. doi: <http://dx.doi.org/10.1183/09031936.00078708>.
15. Dye C, Lönnroth K, Jaramillo E, et al. Trends in tuberculosis incidence and their determinants in 134 countries. *Bull. World Health Organ.* 2009;87:683–91.
16. Pelissari DM, Diaz-Quijano FA. Household crowding as a potential mediator of socioeconomic determinants of tuberculosis incidence in Brazil. *PloS One* 2017;18;12(4):e0176116. doi: <http://dx.doi.org/10.1371/journal.pone.0176116>.
17. Ministério do Desenvolvimento Social e Combate à Fome. Guia de Orientação Técnica – SUAS Nº 1 Proteção Social Básica de Assistência Social. Brasília: Ministério do Desenvolvimento Social e Combate à Fome, 2005. Available from: http://www.pg.pr.gov.br/cmas/wp-content/uploads/2011/07/guia_orientacao_tecnica.pdf [cited 2017 Mar 4].
18. Baussano I, Williams BG, Nunn P, et al. Tuberculosis incidence in prisons: a systematic review. *PLoS Med.* 2010;7(12):e1000381. doi: <http://dx.doi.org/10.1371/journal.pmed.1000381>.
19. Kodinariya, T. M. & Makwana, P. R. Review on determining number of Cluster in K-Means Clustering. *International Journal of Advance Research in Computer Science and Management Studies.* 2013;1:90–5.
20. Conselho Nacional de Saúde. Resolução Nº 510, de 7 de abril de 2016. Brasília: Conselho Nacional de Saúde, 2016. Available from: <http://conselho.saude.gov.br/resolucoes/2016/Reso510.pdf> [cited 2017 Jan 15].
21. Conselho Nacional de Saúde. Resolução No 466, de 12 de dezembro de 2012. Brasília: Conselho Nacional de Saúde, 2012. Available from: http://bvsms.saude.gov.br/bvs/saudelegis/cns/2013/res0466_12_12_2012.html [cited 2017 Dec 9].
22. Cantwell MF, McKenna MT, McCray E, et al. Tuberculosis and race/ethnicity in the United States: impact of socioeconomic status. *Am. J. Respir. Crit. Care Med.* 1998;157:1016–20.

23. Díaz de Quijano E, Brugal MT, Pasarín M, et al. Influencia de las desigualdades sociales, la conflictividad social y la pobreza extrema sobre la morbilidad por tuberculosis en la ciudad de Barcelona. *Rev. Esp. Salud Pública* 2001;75:517–28.

24. Compton WM, Gfroerer J, Conway KP, et al. Unemployment and substance outcomes in the United States 2002–2010. *Drug Alcohol Depend.* 2014;1;142:350-3. doi: <http://dx.doi.org/10.1016/j.drugalcdep.2014.06.012>.

25. Támara-Ramírez JR, Álvarez CA, Rodríguez J. Loss of follow-up and associated factors in patients enrolled in the HIV/AIDS program of the Hospital Universitario San Ignacio, Colombia, 2012-2013. *Biomed. Rev. Inst. Nac. Salud* 2016;3;36(2):265-75. doi: <http://dx.doi.org/10.1016/10.7705/biomedica.v36i2.2939>.

26. Myers WP, Westenhouse JL, Flood J, et al. An ecological study of tuberculosis transmission in California. *Am. J. Public Health* 2006;96:685.

27. Lienhardt C, Fielding K, Sillah JS, et al. Investigation of the risk factors for tuberculosis: a case-control study in three countries in West Africa. *Int. J. Epidemiol.* 2005;34:914–23.

28. Kwan CK & Ernst JD. HIV and tuberculosis: a deadly human syndemic. *Clin. Microbiol. Rev.* 2011;24(2):351-76. doi: <http://dx.doi.org/10.1128/CMR.00042-10>.

29. Ministério da Justiça. Relatório do Levantamento Nacional de Informações Penitenciárias — Infopen. Brasília: Ministério da Justiça, 2014. Available from: <http://www.justica.gov.br/noticias/mj-divulgara-novo-relatorio-do-infopen-nesta-terca-feira/relatorio-depen-versao-web.pdf> [cited 2017 Mar 21].

30. Bourdillon PM, Gonçalves CCM, Pelissari DM, et al. Increase in Tuberculosis Cases among Prisoners, Brazil, 2009–2014. *Emerg. Infect. Dis.* 2017;23:496–9.

31. Ministério da Saúde. Departamento de DST, Aids e Hepatites Virais. Ano V - no 1 - 27a a 53a - semanas epidemiológicas - julho a dezembro de 2015 Ano V - no 1 - 01a a 26a - semanas epidemiológicas - janeiro a junho de 2016. *Bol. Epidemiológico - Aids E DST* (2016). Brasília: Ministério da Saúde, 2016.

32. Pelissari DM, Bartholomay PO, Jacobs MG, et al. Health care basic services and detection of tuberculosis incidence in Brazil. *Rev. Saúde Pública* 2018 in press.

33. Díaz-Quijano FA, Waldman EA. Factors associated with dengue mortality in Latin America and the Caribbean, 1995–2009: an ecological study. *Am. J. Trop. Med. Hyg.* 2012;86(2):328-34. doi: <http://dx.doi.org/10.4269/ajtmh.2012.11-0074>.

34. Gibbons CL, Mangan M-JJ, Plass D, et al. Measuring underreporting and under-ascertainment in infectious disease datasets: a comparison of methods. BMC Public Health 2014;11;14:147. doi: <http://dx.doi.org/10.1186/1471-2458-14-147>.

For peer review only

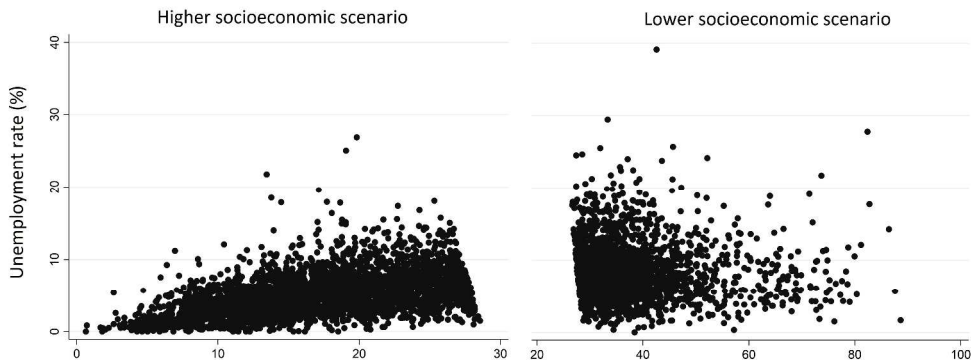


Figure 1. Distribution of Brazilian municipalities according to socioeconomic variables associated with the tuberculosis incidence rate

^aProportion of the population living in households with more than two people per room

Figure 1

695x346mm (300 x 300 DPI)

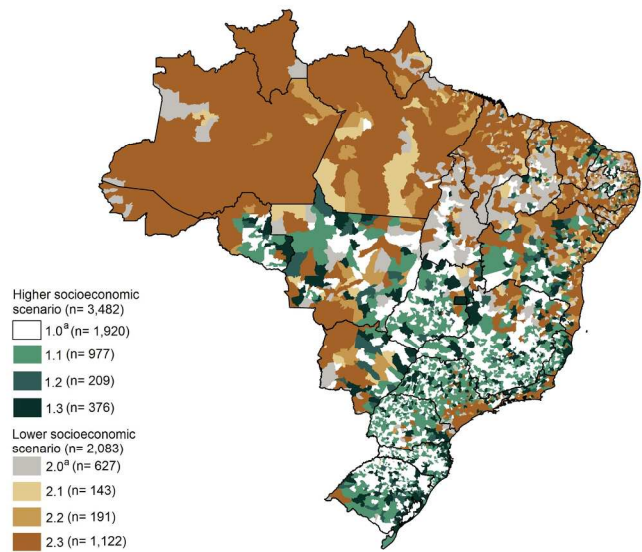


Figure 2 Brazilian's Municipalities by Tuberculosis Incidence Rate Scenarios. Brazil, 2015

^a Without cases in 2014 or 2015.

209x148mm (300 x 300 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

Supplementary material 1- Socioeconomic, epidemiological and operational-tuberculosis indicators by tuberculosis scenarios in Brazil

Variables ^a	Higher socioeconomic n= 3482					Lower socioeconomic n= 2083		
	Subs. 1.0 ^b	Subs. 1.1	Subs. 1.2	Subs. 1.3	Subs. 2.0 ^b	Subs. 2.1	Subs. 2.2	Subs. 2.3
Socioeconomics indicators								
Number of municipalities (%)	1920 (34.5)	977 (17.6)	209 (3.8)	376 (6.8)	627 (11.3)	143 (2.6)	191 (3.4)	1122 (20.2)
Total population- No (%)	13 465 985 (6.6)	27 729 472 (13.6)	6 030 335 (2.9)	56 866 416 (27.8)	5 365 790 (2.6)	2 962 179 (1.4)	6 373 455 (3.1)	85 643 554 (41.9)
HDI-M	0.683	0.696	0.682	0.726	0.59	0.589	0.597	0.618
GDP per capita (USD)	14 407	15 619	13 403	19 565	6752	6626	7572	10 203
Extremely poor (%)	6.1	5.5	7.6	3.9	23.3	23.4	21.3	18.7
Gini coefficient	0.47	0.49	0.49	0.51	0.53	0.54	0.54	0.54
Unemployment rate (%)	4.5	5.7	5.9	6.6	7.7	7.8	8.7	8.6
Household crowding (%)	16	18.1	18.6	19.8	36.6	38.8	38.5	39.5
Life expectancy at birth (years)	74.1	74.3	73.8	75	70.6	70.6	70.9	71.5
Infant mortality rate (number of deaths in first year of life/1000 live births)	16.1	16.1	17.3	14.8	25.9	26.1	25.2	23.6
Epidemiological indicators								
N of new cases (%)	978 (1.4)	5269 (7.8)	1353 (2)	18 865 (27.8)	358 (0.5)	588 (0.9)	2205 (3.3)	38 161 (56.3)
TB incidence rate (cases/ 100,000 people)	7.8	22.3	24.5	39.6	7	21.9	30.3	29.2
AIDS case detection rate (cases/100,000 people) ^{c,d}	6.4	11.1	11.9	21.6	5.2	6.6	8.2	10.1
TB mortality rate (deaths/ 100,000 people) ^{c,d}	0.7	1.3	1.6	1.6	1	1.8	2.3	2.1
New cases from at least one vulnerable group (%) ^{c,d,e}	9.9	14.3	19.2	22.6	12.1	15.6	10.6	14.3
TB-HIV confection among new cases (%)	5.9	8	7.7	8.8	6.9	9.8	3.6	5.6
New cases who were prisoners (%)	2.7	4.2	6.7	12.1	2.5	3	2.7	4.1
New cases who were health professionals (%)	0.7	1	1.5	1	1.1	0.5	1.3	0.9
New cases who were indigenous population (%)	0.4	0.7	1.6	0.6	1.5	2.4	2.4	3.5
New cases who were homeless (%)	0.7	1.3	1.9	1.4	0.7	1.3	1.1	1.1
Retreatment cases among total cases (%)	13.3	8.1	9.1	13.7	20.7	10.1	9.4	11.1
Operational indicators (new cases)								
Contacts examination (%) ^e	70.6	78.8	54.3	73.3	55.7	36.5	60.5	69.6
Pulmonary cases with laboratory confirmation (%)	65.4	70.4	64.0	71.0	66.5	66.8	66.2	71.4
Tested for HIV (%)	72.1	75.5	67.2	75.3	63.9	52.3	53.5	67.2
Cure (%)	66.8	84.8	32.2	73.1	65.4	14.6	51.1	79.7
Lost to follow-up (%)	5.1	5.8	2.8	8	3.8	1.5	5.2	8.2
No TB outcome registration (%) ^d	14.5	0	60.1	11	22.7	81.8	37	4.6
Culture examination (retreatment) (%)	35.5	39	25.1	44	25.5	20.4	21.5	27.8

^aSocioeconomic indicators: with the exception of the number of municipalities and total population (data from 2015), the other variables were measured in 2010. Epidemiological indicators: with the exception of TB mortality rate (data from 2014), the other variables were measured in 2015. Operational indicators: with the exception of cure, lost to follow-up, no record of TB outcome and culture examination (data from 2014), the other variables were measured in 2015. ^bWithout cases in 2014 or 2015; ^cVariables used in non-hierarchical clustering method for scenario 1; ^dVariables used in k-means method for scenario 2; ^eHIV, health professional, prisoners, indigenous and homeless. Municipal Human Development Index- HDI; Gross domestic product-GDP.

Supplemental material 2- Access to data used in this study

Municipalities in Brazil have a unique code (código IBGE) which facilitates merging data and is available in all data sources listed below.

The socioeconomic datasets from Brazil used in the current study are available at:

- “Departamento de Informática do SUS” repository:
 - Click the link: <http://datasus.saude.gov.br/>
 - Click in “Acesso à informação” (Information access) -> Informação de Saúde (TABNET) -> “Demográficas e Socioeconômicas” (Demographic and socioeconomic)
 - Select an indicator group, we used the following: População residente (resident population); Educação (education indicators); Trabalho e renda (labor and income indicators); Produto Interno Bruto (GDP per capita)
 - Inside each indicator group select an indicator
 - For each one select the “Abrangência Geográfica” (geographic dimension) field: “Brasil por município” (Brazil by municipality)
 - In the “Linha” field (line) select “Município” (municipality), in the “Coluna” field (column) select the “Ano” (year), and in the “Período Disponível” field (Available period), we used 2010
 - Click “Mostrar” (Show) and the indicator will be calculated for all municipalities
 - You now can export in different formats. Export options are at the bottom of the page
- Human Development Atlas repository:
 - Click the link: <http://www.atlasbrasil.org.br/2013/en/consulta/>
 - First, select a “Locality”: click “Municipalities” and then the checkbox: “All municipalities - Brazil”
 - Now, select the group of indicators and then the indicator, we used the following group indicators: MHDI; Demography; Income; Labor; Housing; Vulnerability
 - After loading data, download dataset by clicking at the icon in the right top of the page.

The datasets of tuberculosis indicators in Brazil is available by request at the Ministry of Health:

- Click the link: <https://esic.cgu.gov.br/sistema/site/index.aspx>
- Field the form to make a new registration
- Register your request of access data detailing with information and indicators you have interesting and send the request
- Ministry of Health has four weeks to answer the request

The datasets of aids notification by municipalities in Brazil is available in the “Departamento de Informática do SUS” repository:

- Click the link: <http://datasus.saude.gov.br/>
- Click in “Acesso à informação” (Information access) -> Informações de Saúde (TABNET) -> “Epidemiológica e Morbidade” (Epidemiologic and morbidity)
- Select: “Casos de Aids - Desde 1980 (SINAN)” (Aids cases- since 1980)
- In the “Abrangência Geográfica” field (geographic dimension) select “Brasil por Região, UF e município” (Brazil by region, state and municipality).
- In the “Linha” field (line) select “Município de residência” (resident municipality), and in the “Coluna” field (column) select the “Ano diagnóstico” (diagnose year). In the “Período Disponível” field (Available period) we used 2015
- Click in “Mostrar” (Show) and the number of aids cases will be calculated for all municipalities
- You now can export in different formats.

STROBE 2007 (v4) checklist of items to be included in reports of observational studies in epidemiology*

Checklist for cross-sectional studies (combined)

Section/Topic	Item #	Recommendation	Reported on page #
Title and abstract	1	(a) Indicate the study’s design with a commonly used term in the title or the abstract	2 and 3
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	3
Objectives	3	State specific objectives, including any pre-specified hypotheses	3
Methods			
Study design	4	Present key elements of study design early in the paper	3
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	3 and 4
Participants	6	(a) Cohort study—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up Case-control study—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls Cross-sectional study—Give the eligibility criteria, and the sources and methods of selection of participants	3 and 4
		(b) Cohort study—For matched studies, give matching criteria and number of exposed and unexposed Case-control study—For matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	4
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	3 and 4
Bias	9	Describe any efforts to address potential sources of bias	5
Study size	10	Explain how the study size was arrived at	3
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	4
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding	4 and 5
		(b) Describe any methods used to examine subgroups and interactions	4 and 5
		(c) Explain how missing data were addressed	5
		(d) Cross-sectional study—If applicable, describe analytical methods taking account of sampling strategy	-
		(e) Describe any sensitivity analyses	-

Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	5 and 6
		(b) Give reasons for non-participation at each stage	6 and 7
		(c) Consider use of a flow diagram	-
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	6 and 7
		(b) Indicate number of participants with missing data for each variable of interest	3
Outcome data	15*	<i>Cross-sectional study</i> —Report numbers of outcome events or summary measures	6-8
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	6 and 8
		(b) Report category boundaries when continuous variables were categorized	-
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	-
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	8 and 9
Discussion			
Key results	18	Summarise key results with reference to study objectives	9-11
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	11
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	9-11
Generalisability	21	Discuss the generalisability (external validity) of the study results	11-12
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	14

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at <http://www.plosmedicine.org/>, Annals of Internal Medicine at <http://www.annals.org/>, and Epidemiology at <http://www.epidem.com/>). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

Identifying socioeconomic, epidemiological and operational scenarios for tuberculosis control in Brazil: an ecological study

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2017-018545.R2
Article Type:	Research
Date Submitted by the Author:	07-Mar-2018
Complete List of Authors:	<p>Pelissari, Daniele; Ministry of Health, National Tuberculosis Program of Brazil</p> <p>Rocha, Marli; Ministry of Health, National Tuberculosis Program of Brazil</p> <p>Bartholomay, Patricia ; Ministry of Health, National Tuberculosis Program of Brazil</p> <p>Sanchez, Mauro; University of Brasilia</p> <p>Duarte, Elisabeth; University of Brasília, Faculty of Medicine, Nucleo de Medicina Tropical</p> <p>Arakaki-Sanchez, Denise; Ministry of Health, National Tuberculosis Program of Brazil</p> <p>Dantas, Cíntia ; Ministry of Health, National Tuberculosis Program of Brazil</p> <p>Jacobs, Marina; Ministry of Health, National Tuberculosis Program of Brazil</p> <p>Andrade, Kleydson ; Ministry of Health, National Tuberculosis Program of Brazil</p> <p>Codenotti, Stefano ; Ministry of Health, National Tuberculosis Program of Brazil</p> <p>Andrade, Elaine Silva ; Ministry of Health, National Tuberculosis Program of Brazil</p> <p>de Araújo, Wildo ; University of Brasília</p> <p>Costa, Fernanda ; Ministry of Health, National Tuberculosis Program of Brazil</p> <p>Ramalho, Walter; University of Brasília</p> <p>Díaz-Quijano, Fredi; School of Public Health, University of São Paulo, Department of Epidemiology</p>
Primary Subject Heading:	Public health
Secondary Subject Heading:	Epidemiology, Health policy, Infectious diseases
Keywords:	public health policy, Tuberculosis < INFECTIOUS DISEASES, Epidemiology < TROPICAL MEDICINE

SCHOLARONE™
Manuscripts

Title: Identifying socioeconomic, epidemiological and operational scenarios for tuberculosis control in Brazil: an ecological study

*Daniele Maria Pelissari, National Tuberculosis Program, Ministry of Health. Address: SRTVN 701, Via W 5 Norte, Ed. PO 700 – 6º andar. Zip code: 70719-040, Brasília-DF, Brazil. Email: daniele.pelissari@saude.gov.br; daniele.pelissari@gmail.com. Telephone: (+55)61998431554

Marli Souza Rocha, National Tuberculosis Program, Ministry of Health. Brasília/DF, Brazil.

Patricia Bartholomay, National Tuberculosis Program, Ministry of Health. Brasília/DF, Brazil.

Mauro Niskier Sanchez, Department of Public Health, University of Brasilia – UnB. Brasília/DF, Brazil.

Elisabeth Carmen Duarte, Medical School, University of Brasilia –UnB. Brasília/DF, Brazil.

Denise Arakaki-Sanchez, National Tuberculosis Program, Ministry of Health. Brasília/DF, Brazil.

Cíntia Oliveira Dantas, National Tuberculosis Program, Ministry of Health. Brasília/DF, Brazil.

Marina Gasino Jacobs, National Tuberculosis Program, Ministry of Health. Brasília/DF, Brazil.

Kleydson Bonfim Andrade, National Tuberculosis Program, Ministry of Health. Brasília/DF, Brazil.

Stefano Barbosa Codenotti, National Tuberculosis Program, Ministry of Health. Brasília/DF, Brazil.

Elaine Silva Nascimento Andrade, National Tuberculosis Program, Ministry of Health. Brasília/DF, Brazil.

Wildo Navegantes de Araújo, Faculty of Ceilandia, University of Brasilia –UnB. Brasília/DF, Brazil.

Fernanda Dockhorn Costa, National Tuberculosis Program, Ministry of Health.

Walter Massa Ramalho, Faculty of Ceilandia, University of Brasilia –UnB. Brasília/DF, Brazil.

Fredi Alexander Diaz-Quijano, Department of Epidemiology, School of Public Health, University of São Paulo. São Paulo/SP, Brazil.

*Corresponding author

Role of the funding source: This study did not have specific funding.

Word count: 3 543

Number of figures and tables: 3 tables; 2 figures; supplementary material-1: List of variables analysed in the study; supplementary material-2: Access to data used in this study; supplementary material-3: 1 table

ABSTRACT

Objectives To identify scenarios based on socioeconomic, epidemiological and operational health care factors associated with tuberculosis incidence in Brazil.

Design Ecological study.

Settings The study was based on new tuberculosis cases and epidemiological/operational variables of the disease from the Brazilian National Information System for Notifiable Diseases and the Mortality Information System. We also analysed socioeconomic and demographic variables.

Participants The unit of analysis was the Brazilian municipalities, which in 2015 numbered 5 570, but five were excluded due to absence of socioeconomic information.

Primary outcome Tuberculosis incidence rate in 2015.

Data analysis We evaluated as independent variables the socioeconomic (2010), epidemiological and operational health care indicators of tuberculosis (2014 or 2015) using negative binomial regression. Municipalities were clustered by the k means method considering the variables identified in multiple regression models.

Results We identified two clusters according to socioeconomic variables associated with the tuberculosis incidence rate (unemployment rate and household crowding): a higher socioeconomic scenario (n= 3 482 municipalities) with a mean tuberculosis incidence rate of 16.3/100 000 population and a lower socioeconomic scenario (2 083 municipalities) with a mean tuberculosis incidence rate of 22.1/100 000 population. In the second-stage of clusterization we defined four subgroups in each of the socioeconomic scenarios using epidemiological and operational variables such as tuberculosis mortality rate, AIDS case detection rate and proportion of vulnerable population among tuberculosis cases. Some of the sub-scenarios identified were characterized by fragility in their information systems, while others were characterized by the concentration of tuberculosis cases in key populations.

Conclusion Clustering municipalities in scenarios allowed us to classify them according to the socioeconomic, epidemiological and operational variables associated with tuberculosis risk. This classification can support targeted evidence-based decisions such as monitoring data quality for improving the information system or establishing integrative social protective policies for key populations.

Key word: tuberculosis, epidemiology, public health policy.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Strengths and limitations of this study
<ul style="list-style-type: none">• This study was based on national population data in a country of continental dimension (5 565 municipalities).• The availability of variables associated with tuberculosis made it possible to consider both socioeconomic and epidemiological/operational approaches in the definition of municipality clusters for tuberculosis control.• This methodology can be explored by other countries to guide their plans to end tuberculosis.• Reporting and information quality may vary between sources and periods, which could affect estimate accuracy.• Inferences obtained are applicable to population groups, not to individuals. However, ecological research can provide evidence to support public health decisions.

INTRODUCTION

In 2016, 10.4 million people had tuberculosis (TB) and 1.8 million died worldwide because of the disease.¹ In Brazil, similar to other countries, TB incidence reduction (37.9/100 000 population in 2007 to 32.4/100 000 population in 2016)² seems to be associated with the improvement of population living conditions^{3–5} and the performance of TB control programmes.⁶ However, the disease burden continues to be significant in the country, with 66 796 new cases registered in 2016.²

In 2014, in a move towards elimination, the World Health Organization (WHO) launched the End TB Strategy, setting targets to be met by 2035, including a 90% reduction in TB incidence compared to 2015.¹ The strategy is critical to energizing the fight against the disease and mobilizing resources, but needs to be adapted to the local context, as does any other health policy.⁷

Some countries have already made progress developing their national plans. Among the strategies presented, we highlight the strengthening of existing TB services, the acceleration of case detection in key populations, and the implementation of actions to reduce barriers to TB care.^{8–10}

Brazil is a country with continental dimensions, thus both socioeconomic indicators¹¹ and those that reflect the performance of local TB programmes² present a high degree of heterogeneity. Considering this context, and to support the "National Plan to End TB",¹² we identified scenarios based on socioeconomic, epidemiological, and operational factors associated with the TB incidence rate.

METHODS

Type of study and data source

This is an ecological study, with the unit of analysis being the Brazilian municipalities, which were 5 570 in 2015. We excluded five municipalities due to absence of socioeconomic information. Data on socioeconomic and demographic variables by municipality were only available from the last population census (2010).^{11,13} For new TB cases (2015) and epidemiological/operational variables of the disease (2014 and 2015), we used data from the Brazilian National Information System for Notifiable Diseases (SINAN) and the Mortality Information System (SIM).¹¹

Variables

The dependent variable was the TB incidence rate (new cases/100 000 population) in 2015, and the independent variables were socioeconomic, epidemiological, and health care operational TB variables. Many of these variables have already been identified in previous studies as TB determinants (supplementary material- 1).^{3,14–16}

The socioeconomic variables analysed were as follows:

- Municipal human development index (HDI-M);

- Average household income per capita;
- Gross domestic product (GDP) per capita;
- Proportion of the population that is extremely poor, poor, and vulnerable to poverty;
- Gini coefficient;
- Unemployment rate;
- Illiteracy rate;
- Proportion of population living in households with more than two people per room representing household crowding;
- Infant mortality rate per 1 000 live births;
- Life expectancy at birth.
- Population size of municipalities classified as small (less than 20 000 inhabitants), medium (20 000 to 99 999 inhabitants) and large (100 000 inhabitants or larger).¹⁷

Average household income per capita and GDP per capita were converted into US dollars (USD) using the average annual price in 2010 (1 USD ≈ 1.8 Brazilian Reals- R\$). We adopted the Brazilian definitions for the proportions of the population that are extremely poor, poor and vulnerable to poverty: proportion of individuals in the municipality with an average household income per capita equal or less than USD 40, USD 80 and USD 145, respectively.¹³

The epidemiological variables of TB were as follows:

- AIDS case detection rate per 100 000 population;
- Proportion of new TB cases who were: HIV positive, prisoners, health professionals, indigenous, homeless and, as a composite indicator, the proportion of TB cases from at least one of these vulnerable groups. Those vulnerabilities were previously associated with an increased risk of TB in other studies;^{3-5,18}
- Proportion of TB retreatment cases;
- TB mortality rate per 100 000 population;

The operational health care variables of TB considered in the analysis were as follows:

- Proportions of new TB cases: in which contacts were examined, laboratory confirmed, tested for HIV, and the treatment outcomes (cure, loss to follow-up, and no record of TB outcome);
- Proportion of sputum culture examination among retreatment cases;

Due to the availability of updated data at the time of analysis, the data to calculate culture examinations, treatment outcomes, and TB mortality rate refers to 2014, while the other epidemiological and operational variables refers to 2015.

Statistical analysis

Statistical analysis was performed in two stages; each of them included a model to identify the factors associated with TB incidence rate. This evaluation was followed by a cluster analysis based on the factors identified. The first stage was focused on socioeconomic variables and the second on the epidemiological and operational variables associated with TB incidence rate.

Negative binomial regression was used to identify factors associated with the TB incidence rate in 2015. For these regression analyses, we only included municipalities that presented a mean annual variation of the triennial moving average of the incidence rate for the years 2001 to 2015 between -8 and 8%. By doing so, we intended to reduce possible biases due to the variability of values in small municipalities and any possible intermittence in case reporting.

To obtain multiple regression models that were parsimonious and robust, we avoided including variables that were strongly correlated with each other or those that showed signs of multicollinearity (i.e., inversion of the correlation coefficient together with an increase in standard error). Whenever a strong correlation between independent variables (Spearman's $\rho > 0.6$) was identified, for the multiple model, we selected the variable with the highest association with TB incidence rate in univariate regression models. We used a stepwise forward selection method and preserved the variables with a p value < 0.05 in the multiple model. In addition, we adjusted the models by categories of the population size of municipalities (small: less than 20 000 inhabitants, medium: 20 000 to 99 999 inhabitants and large: 100 000 inhabitants or larger). We presented the association measures as the relative increment in the incidence rate (RIIR), which was obtained by subtracting one from the incidence rate ratio ($[IRR - 1] \times 100\%$).

Socioeconomic variables associated in a multiple model (primary model) were considered for a cluster analysis of all municipalities using the non-hierarchical k-means method. In this method, the algorithm aims to reduce intra-group variance and maximize inter-group variance in relation to the Euclidean distance established by the variables selected.¹⁹ To define the number of clusters, we used the Elbow method, which relates the number of clusters with the percentage of internal variation of the groups,¹⁹ adopting $> 60\%$ as the cut-off point and among these, the smallest number of possible clusters.

In a second stage, epidemiological/operational variables were modelled for each socioeconomic scenario, following a similar methodology described for socioeconomic variables. Factors associated with TB in these secondary models, as well as the TB mortality rate, were considered for a second cluster analysis, which subdivided the previous socioeconomic clusters into epidemiological/operational TB sub-scenarios. Because some operational variables were only measurable during the care of TB patients, these second stage methods were applied only in municipalities with TB cases reported in 2014 and 2015.

Statistical analyses were performed with the Stata statistical package version 12.0, R version 3.3.1 and the cluster library.

Patient and public involvement

Patient and public were not involved in this study because all variables studied correspond to data aggregated by municipalities. Therefore, researchers did not have access to any individual data or personal identification of tuberculosis patients. The results will be disseminated for Tuberculosis Control Programmes in municipalities and states and the grouping of municipalities presented in this study has already been incorporated into the National Plan to End TB.

Ethical aspects

All data analysed are publicly available in Brazil, and the procedure to access is described in supplementary material-2. According to Brazilian legislation (Resolution No. 510 of the National Health Council of Brazil),²⁰ studies conducted exclusively with publicly available data are not required to be evaluated by an institutional review board. This study was conducted according to the guidelines and standards for research involving human subjects.²¹

RESULTS

In 2015, 67 777 new TB cases were reported in Brazil, with an incidence rate of 33.1/100 000 population. The mean annual variation of the triennial moving average of the TB incidence rate in municipalities ranged from -22.6% to 41.8%. This interval was wider in small and medium municipalities (-22.7% to 41.9%) rather than in larger ones (-7.3% to 14.6%). A total of 3 311 (59.5%) municipalities presented a variation between -8% and 8% and were eligible for the analysis for the primary model, including 791 that did not present new TB cases in 2015.

Regarding socioeconomic variables, household crowding and unemployment rate exhibited the highest associations with TB incidence rate in both univariate and multiple analysis. The percentages of the poor and vulnerable to poverty population were not included in the multiple model because these factors were strongly correlated with household crowding. On the other hand, the Gini coefficient exhibited a moderate correlation with household crowding (Spearman rho=0.55) and a weak correlation with unemployment rate (Spearman rho: 0.31). However, the Gini coefficient was not preserved in the multiple model because of inversion of its regression coefficient and an increase in the standard error when adjusted. The other socioeconomic variables were not significantly associated with TB incidence rate in the multiple model (Table 1).

Table 1. Socioeconomic variables and association with tuberculosis incidence rate in Brazil (n=3 311 municipalities^a)

Variable ^b	Mean (standard deviation-SD)	Median (IQ25%-IQ75%)	RIIR (95% CI) ^c	(Adjusted) RIIR (95% CI) ^c
HDI-M	0.7 (0.1)	0.7 (0.6 - 0.7)	-4.8 (-36.9 to 43.7)	
Average household income per capita (USD)	280.3 (143.8)	257.9 (155.6 - 372.8)	-0.0 (-0.0 to -0.0) ^d	
GDP per capita (USD)	7 510.4 (8630.3)	5 555.2 (2 909.4 – 9	0.0 (-0.0 to 0.0)	

		0.91 (2)		
Extremely poor (%)	11.4 (11.7)	6.5 (1.6 - 19.1)	-0.0 (-0.3 to 0.2)	
Poor (%)	23.4 (18.0)	18.5 (6.9 - 38.8)	0.1 (-0.0 to 0.3)	
Vulnerable to poverty (%)	44.1 (22.7)	42.6 (23.3 - 65.6)	0.2 (0.0 to 0.3) ^d	
Gini coefficient (%)	51.0 (6.5)	51.0 (46.7 - 55.2)	0.8 (0.3 - 1.2) ^d	
Unemployment rate (%)	6.7 (3.7)	6.3 (4.2 - 8.6)	5.0 (4.2 to 5.8) ^d	3.9 (3.0 to 4.7) ^d
Illiteracy in the population with ≥ 18 years (%)	17.2 (10.8)	13.9 (8.1 - 26.4)	-0.4 (-0.6 to -0.1) ^d	
Illiteracy in the population with ≥ 15 years (%)	15.6 (9.8)	12.9 (7.2 - 23.8)	-0.4 (-0.7 to -0.1) ^d	
Household crowding (%) ^c	26.4 (13.1)	24.7 (16.6 - 33.8)	1.2 (1.0 to 1.4) ^d	0.8 (0.6 to 1.1) ^d
Infant mortality rate (number of deaths in the first year of life/1 000 live births)	19.3 (7.2)	17.0 (13.7 - 24.1)	0.1 (-0.3 to 0.5)	
Life expectancy at birth (years)	73.1 (2.7)	73.4 (71.1 - 75.2)	-0.9 (-1.9 to 0.2)	

Abbreviations: HDI-M, municipal human development index; GDP, gross domestic product; SD, standard deviation; IQ, interquartile; RIIR, relative increment in the incidence rate; CI, confidence interval.

^aMunicipalities with an annual variation in TB incidence rate between -8% and 8% and at least one new TB case in 2015.

^bVariables measured in the last census (2010).

^cThe association measure represents the relative increment in the incidence rate ($[(IRR-1) \times 100]$), adjusted for the population size of the municipality.

^dp value < 0.05.

^eProportion of the population living in households with more than two people per room.

Based on these two socioeconomic variables, we identified a higher socioeconomic scenario (HSS) cluster, with 3 482 municipalities, which presented better socioeconomic variables than the 2 083 municipalities from the second cluster, the lower socioeconomic scenario (LSS) (Table 2). The HSS cluster exhibited unemployment rates of up to 26.9% and household crowding values between 0.6% and 28.6%. On the other hand, the LSS cluster exhibited unemployment rates of up to 39.1%; and, household crowding values between 26.6% and 88.6% (Figure 1).

Figure 1. Distribution of Brazilian municipalities according to socioeconomic variables associated with the tuberculosis incidence rate

The mean TB incidence rate in the LSS was 22.1/100 000 population (Table 2), which was significantly higher than that observed in the HSS, which was 16.3/100 000 population (IRR: 1.3; 95% CI: 1.3 - 1.4).

Table 2. Description of socioeconomic scenarios associated with the tuberculosis incidence rate in Brazil (n=5 565 municipalities)^a

Variables ^b	Higher socioeconomic scenario n= 3482 municipalities		Lower socioeconomic scenario n= 2083 municipalities	
	Mean (SD)	Median (IQ25%-IQ75%)	Mean (SD)	Median (IQ25%-IQ75%)
TB incidence rate per 100 000 population ^c	16.3 (31.7)	10.2 (0 - 23.2)	22.1 (36.6)	16.9 (6.3 - 29.6)
HDI-M	0.7 (0.1)	0.7 (0.6 - 0.7)	0.6 (0.1)	0.6 (0.6 - 0.6)
Average household income per capita (USD)	330.5 (126.5)	324.1 (242.3 - 403.5)	181.3 (92.4)	150.8 (125.3 - 203.5)
GDP per capita (USD)	8 661.1 (7)	7 259.7 (4 758.8 - 10 053.9)	4 930.5 (8 033.6)	2 937.6 (2 371.3 - 5

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

	707.6)			004.5)
Extremely poor (%)	5.8 (7.3)	2.6 (1.1 - 7.4)	20.7 (11.9)	20.3 (12.0 - 28.7)
Poor (%)	14.2 (12.3)	9.7 (5.1 - 20.2)	38.3 (15.5)	40.3 (29.7 - 49.1)
Vulnerable to poverty (%)	33.0 (17.8)	29.3 (19.0 - 44.9)	62.4 (16.5)	67.0 (57.0 - 73.5)
Gini coefficient (%)	58.0 (6.0)	48.0 (43.9 - 52.0)	54.1 (5.8)	53.8 (50.2 - 57.5)
Unemployment rate (%) ^d	5.1 (2.9)	4.9 (3.1 - 6.8)	8.3 (4.0)	7.6 (5.6 - 10.3)
Illiteracy in the population with ≥ 18 years (%)	12.9 (8.0)	10.7 (7.2 - 16.0)	25.0 (10.3)	26.5 (17.1 - 32.9)
Illiteracy in the population with ≥ 15 years (%)	11.7 (7.3)	9.8 (6.5 - 14.6)	22.7 (9.5)	23.9 (15.6 - 29.8)
Household crowding (%) ^{d,e}	17.1 (6.0)	17.3 (12.7 - 22.0)	38.5 (10.3)	35.7 (31.2 - 41.9)
Infant mortality rate (number of deaths in the first year of life/1 000 live births)	16.0 (5.0)	14.8 (12.8 - 17.5)	24.6 (6.9)	24.2 (19.4 - 29.0)
Life expectancy at birth (years)	74.3 (2.1)	74.5 (73.2 - 75.7)	71.1 (2.3)	71.1 (69.6 - 72.6)

Abbreviations: TB, tuberculosis; HDI-M, municipal human development index; GDP, gross domestic product; SD, standard deviation; IQ, interquartile.

^aTotal of municipalities with socioeconomic data in Brazil that were used in the clusterization step.

^bWith the exception of the TB incidence rate (2015), the other variables were measured in the last census (2010).

^cIncidence rate ratio= 1.3; 95% CI= 1.3 - 1.4

^dVariables identified during step 1 (model 1) used in step 2 with the k-means method.

^eProportion of the population living in households with more than two people per room.

Among the 3 482 HSS municipalities, 1 125 had TB cases in 2014 and 2015 and were eligible (annual variation in TB incidence rate between -8% and 8%) for a secondary modelling. In this analysis, the AIDS case detection rate and the proportion of new cases from at least one vulnerable group were positively associated with the TB incidence rate, while the proportion of contacts investigation among new TB cases presented an inverse association (Table 3).

Regarding the LSS, 1 095 municipalities out of 2 083 were eligible. The AIDS case detection rate and the proportion of cases from at least one vulnerable group were positively associated with the TB incidence rate. In contrast, the proportion of cases with no record of TB outcome was inversely associated with TB incidence (Table 3).

Table 3. Epidemiological and operational tuberculosis variables associated with the tuberculosis incidence rate stratified by socioeconomic scenarios in Brazil (n=2 220 municipalities)

Variables ^a	Higher socioeconomic scenario n=1 125 municipalities ^b		Lower socioeconomic scenario n=1 095 municipalities ^b	
	RIIR (95% CI) ^c	(Adjusted) RIIR(95% CI) ^c	RIIR (95% CI) ^c	(Adjusted) RIIR(95% CI) ^c
Epidemiological				
AIDS case detection rate (cases/100 000 population)	1.5 (1.2 to 1.7) ^e	1.4 (1.1 to 1.6) ^e	2.1 (1.7 to 2.5) ^e	2.0 (1.6 to 2.4) ^e
New cases from at least one vulnerable group (%) ^d	0.5(0.3 to 0.7) ^e	0.2 (0.1 to 0.4) ^e	0.7 (0.5 to 0.9) ^e	0.5 (0.3 to 0.7) ^e
TB-HIV confection among new cases (%)	0.3 (0.0 to 0.6) ^e		-0.2 (-0.5 to 0.2)	
New cases who were prisoners (%)	0.7 (0.4 to 0.9) ^e		1.2 (0.9 to 1.5) ^e	
New cases who were health professionals (%)	-0.5 (-1.3 to 0.2)		-0.2 (-1.0 to 0.6)	
New cases who were from an indigenous population (%)	1.1 (0.3 to 1.9) ^e		0.9 (0.5 to 1.2) ^e	
New cases who were homeless (%)	0.1 (-0.6 to 0.7)		0.1 (-0.7 to 1.0)	
Retreatment cases among the total cases (%)	0.5 (0.2 to 0.8) ^e		0 (-0.3 to 0.3)	
Operational health care (new TB cases)				
Contact examination (%)	-0.3 (-0.4 to -0.1) ^e	-0.2 (-0.3 to -0.1) ^e	-0.0 (-0.2 to 0.1)	
Pulmonary cases with laboratory confirmation (%)	0.0 (-0.1 to 0.2)		-0.1 (-0.3 to 0.0)	

Tested for HIV (%)	0.1 (-0.0 to 0.2)	0.0 (-0.1 to 0.2)	
Cure (%)	-0.2 (-0.3 to -0.0) ^e	0.2 (0.0 to 0.3) ^e	
Lost to follow-up (%)	0.6 (0.3 to 0.9) ^e	0.3 (0.0 to 0.7) ^e	
No TB outcome registration (%)	0.3 (0.1 to 0.5) ^e	-0.3 (-0.5 to -0.1) ^e	-0.3 (-0.5 to -0.1) ^e
Culture examination (retreatment) (%)	0.1 (-0.1 to 0.2)	0.0 (-0.1 to 0.2)	

Abbreviations: RIIR, relative increment in the incidence rate; CI, confidence interval; TB, tuberculosis.

^aWith the exception of: cure, lost to follow-up, no record of TB outcome and culture examination (2014), the other variables were measured in 2015.

^bMunicipalities with an annual variation in TB incidence rate between -8% and 8% and at least one new TB case in 2014 and 2015.

^cThe association measure represents the relative increment in the incidence rate ($[(IRR-1) \times 100]$), adjusted for the population size of the municipality.

^dHIV, health professional, prisoners, indigenous, and homeless.

^ep value < 0.05.

Using the variables associated with the outcome in the previous models and considering the TB mortality rate, we defined three clusters for each socioeconomic scenario, making a total of six sub-scenarios with TB cases in 2014 and 2015. For each socioeconomic scenario, a sub-scenario (1.0 and 2.0) was also defined including municipalities without TB reporting in 2014 or 2015 (supplementary material- 3).

Figure 2 shows the geographical distribution of municipalities according to the sub-scenarios. Regarding the sub-scenarios with TB cases in HSS, sub-scenario 1.1 showed the lowest mean rates of TB incidence, AIDS case detection and TB mortality. Sub-scenario 1.2, despite having relatively low mean rates of TB incidence, AIDS case detection, and TB mortality, had a high proportion of cases with no record of TB outcome. Sub-scenario 1.3 covered 27.8% of new TB cases reported in 2015 and presented the highest mean rates of TB incidence, AIDS case detection, and proportion of cases from at least one vulnerable group (22.0%) (supplementary material-3).

Regarding the LSS municipalities, sub-scenario 2.1 had the lowest mean for contact investigation (36.5%) and HIV testing (52.3%), and the highest mean proportion of cases with no record of TB outcome (81.8%). Sub-scenario 2.2 showed a high TB incidence rate, the highest TB mortality and low HIV testing (53.5%) and a high mean proportion of cases with no record of TB outcome (37.0%). As a consequence of the inclusion of 14 capitals in sub-scenario 2.3, it includes 56.3% of all new cases reported in 2015. Furthermore, sub-scenario 2.3 has the highest mean AIDS case detection rate in the group of LSS and the second highest TB mortality rate among all sub-scenarios (supplementary material- 3).

Figure 2- Brazilian's municipalities by tuberculosis incidence rate scenario. Brazil, 2015.

DISCUSSION

This study classified the 5 565 Brazilian municipalities in two scenarios (LSS and HSS) defined by socioeconomic variables associated with the TB incidence rate in Brazil. Subsequently, we

performed a sub-classification based on operational and epidemiological variables associated with the TB incidence rate.

Regarding socioeconomic variables, the unemployment rate was associated with the risk of TB, as found in previous studies from the United States,²² Spain,²³ and Brazil.¹⁶ At the individual level, unemployment has been associated with an increased risk of alcohol and illicit drug abuse²⁴ and with loss to follow-up during HIV treatment.²⁵ These factors have already been associated with TB risk³⁻⁵ and could at least partially explain the association observed in our study.

Household crowding was also positively associated with the TB incidence rate. In several studies, including some developed in New Zealand,²⁶ Lima,²⁷ the United States and West Africa, people living in crowding conditions had a higher risk of TB.^{28,29} In Brazil, this variable was already associated with TB incidence and was considered a potential mediator between socioeconomic determinants and TB incidence rate because it may directly favour TB transmission by increasing the contact rate between infected and susceptible people.¹⁶

In our study, the LSS, with municipalities predominantly in the North, Northeast, and Centre-West regions, presented a higher incidence of TB than the HSS, with municipalities located predominantly in the South and Southeast regions. This suggested that classification of municipalities by socioeconomic variables could be highly functional to address TB risk.

Regarding the operational and epidemiological variables, the AIDS case detection rate was positively associated with the TB incidence rate in both socioeconomic scenarios, which was consistent with previous studies in which AIDS has been a factor associated with TB risk at the contextual level.^{15,16,30}

The proportion of new cases from at least one vulnerable group was also another factor associated with TB incidence in both scenarios. One of the vulnerable populations included is prisoners. Specifically, in Brazil, in 2014, there were approximately 607 thousand imprisoned people in 956 municipalities distributed in all regions of the country, with a prison occupation rate of 161%.³¹ This overcrowding may explain the high risk shown by this group in previous studies³² and makes it a priority vulnerable group for TB control.

In the HSS, vulnerability was also correlated with the AIDS case detection rate, which is higher in the South and Southeast regions (respectively, 20.1% and 53.0% of the AIDS cases identified from 1980 to June 2016).³³ Regarding the LSS, vulnerability was correlated with indigenous populations, which are predominantly located in the North (37.4%), Northeast (25.5%), and Central-West regions (16.0%).¹¹ These groups have presented a higher risk of TB than other populations.⁴

We observed an inverse association between the TB incidence rate and the percentage of contact investigation in the HSS, which may represent the overall effect on transmission control, possibly

through identification and timely treatment.³⁴ Finally, in the LSS, the association with the proportion of cases with no record of TB outcome may represent failures in surveillance in collecting these data for the qualification of the information system.

Regarding the absence of TB cases in 2014 or 2015 in sub-scenarios 1.0 and 2.0, it is possible that there is under-reporting in these scenarios, mainly in the sub-scenario 2.0, where there are worse socioeconomic conditions, which are associated with a higher risk of TB. This finding suggests that activities related to TB detection should be strengthened especially in those groups of municipalities.

Concerning the sub-scenarios that reported cases in the two years of analysis from the HSS cluster, group 1.1 has the lowest TB incidence rate, better socioeconomic indicators, and good TB epidemiological/operational indicators, suggesting an advanced stage in TB control. Sub-scenario 1.3 presents the highest TB incidence rate, AIDS case detection rate and proportion of cases from at least one vulnerable group (22.0%), especially among prisoners (12.1%). In addition, this scenario is composed mainly by capitals, which could mean a more sensitive surveillance system. Despite sub-scenario 1.3 corresponding to that with the highest TB risk, the distribution of vulnerabilities suggests a concentrated epidemic in some population groups, including HIV patients (8.8% of TB patients were co-infected) and prisoners (12.1% of new cases), which requires distinct and focused strategies such as screening and prompt treatment.

Sub-scenarios 1.2 (HSS-cluster), 2.1 and 2.2 (LSS-cluster) need improvement in the information system due to the high proportion of cases with no record of TB outcome. This makes it difficult to analyse the performance of TB control actions. Another challenge in these groups is the investigation of contacts, which was particularly low in sub-scenario 2.1. Although sub-scenario 2.1 has the highest percentage of TB-HIV coinfection in new cases (9.8%), it also has one of the lowest percentages of HIV testing (52.3%), suggesting the under-detection of HIV among people with TB.

In the LSS, group 2.2 exhibited the highest TB incidence but the lowest proportion of cases from at least one vulnerable group (10.6%), revealing an endemic situation that is less concentrated in vulnerable populations. Sub-scenario 2.3 has a reliable information system and good performance in operational activities (e.g., contact investigation and HIV testing), revealing that even with limited resources, it is possible to carry out effective disease control actions.

Finally, with the exception of sub-scenario 2.0, all those in the LSS had a higher TB mortality rate than those in the HSS. Sub-scenario 2.0, even though no new cases were reported in 2014 or 2015, exhibited a higher mortality rate than the 1.0 group. Mortality is expected to be less underreported than incidence, as observed in other diseases.^{35,36} Thus, the use of this variable for defining clusters contributes to characterizing groups according to TB burden besides the other variables used for classification.

Limitations

As a common limitation of ecological studies, aggregate measures might differ from individual ones. However, these studies provide an overview that contributes to direct decision-making in public policies.

Underreporting of TB cases in Brazil is decreasing each year,¹ but may remain a potential limitation for this study. Since there is no information about TB case detection and latent TB infection in Brazilian municipalities, the overall burden cannot be estimated. Even so, we hypothesize that the underreporting is either homogeneous or higher in municipalities with worst socioeconomic indicators. Therefore, the magnitude of association between socioeconomic indicators and TB incidence may be higher than estimated in this study. The exclusion of municipalities that presented high variability in the incidence rate may have reduced the risk of information bias.

On the other hand, although an important number of municipalities was excluded from the regression analysis, those localities were usually small, and the overall municipalities included made up 87.2% of the Brazilian population. In addition, only five municipalities (0.1% of the total) were excluded because of the absence of socioeconomic data. Therefore, we conclude that the association identified in the multiple models can be widely extrapolated.

Concerning data availability, socioeconomic variables by municipality were only available from the last census conducted in the country (2010). Therefore, recent socioeconomic trends and their impact on the current TB incidence rate could not be evaluated. However, we believe that the socioeconomic differences between municipalities have remained proportional in recent years, which allows their evaluation as a determinant of the TB incidence.

Implications for public health and conclusion

The End TB Strategy proposes bold targets, and a prompt response from each country may be critical for their achievement. We consider this work an innovative tool for public health decisions because we used secondary data available for most of the municipalities of the country with a robust data analysis that recognizes the socioeconomic and operational diversity of a continental country. The grouping of municipalities presented in this study has already been incorporated into the National Plan to End TB¹² to support the implementation of efficient strategies.

Efforts should be focused on strengthening information systems to provide a reliable picture of the epidemiological situation, such as the implementation of monitoring strategies to ensure the quality of data collection.

There is an inverse relationship between the amount spent with social protection and TB indicators (prevalence, incidence, and mortality).^{37,38} The challenges of controlling TB in key

populations (prisoners and indigenous) are probably related to their social marginalization and most likely require integrative collaboration with national social protection programmes run by other divisions of the government.

Municipalities in the LSS, besides additional resources, require actions to reduce the exacerbation of social vulnerabilities, which were reflected in TB risk. That is why TB should be considered a priority in the public health agenda. In addition, municipalities from LSS scenarios that did not have a record of TB cases in 2015/2014 should focus on activities related to TB detection, especially household-contact investigation as a strategy for active case finding, as this method has been shown to be more effective than standard passive case finding.^{39,40}

The heterogeneity of the socioeconomic and epidemiological situation in Brazil, observed in this study, represents a great challenge for TB control in a country of continental proportions, which may also be the reality in other countries. In this sense, our data analysis approach can be considered by other countries with available variables in order to identify sub-scenarios to guide targeted actions for TB control.

CONTRIBUTORS

Daniele Maria Pelissari: conceived the study, contributed to the design of the study and the interpretation of results, coordinated and analysed the data including epidemiological models and clusterization, wrote the first draft of the report, prepared the illustrations, wrote the discussion, critically reviewed the manuscript and approved the final version.

Marli Souza Rocha: conceived the study, contributed to the design of the study and the interpretation of results, analysed the data including epidemiological models and clusterization, contributed to the discussion, critically reviewed the manuscript and approved the final version.

Patricia Bartholomay: conceived the study, contributed to the design of the study and the interpretation of results, analysed the data including epidemiological models and clusterization, contributed to the discussion, critically reviewed the manuscript and approved the final version.

Mauro Niskier Sanchez: contributed to the design of the study and the interpretation of results, provided significant inputs in the first draft, contributed to the discussion, critically reviewed the manuscript and approved the final version.

Elisabeth Carmen Duarte: contributed to the design of the study and the interpretation of results, provided significant inputs in the first draft, contributed to the discussion, critically reviewed the manuscript and approved the final version.

Denise Arakaki-Sanchez: contributed to the design of the study and the interpretation of results, contributed to the discussion, critically reviewed the manuscript and approved the final version.

Cíntia Oliveira Dantas: contributed to the design of the study and the interpretation of results, contributed to the discussion, critically reviewed the manuscript and approved the final version.

Marina Gasino Jacobs: contributed to the design of the study and interpretation of results, contributed to the discussion, critically reviewed the manuscript and approved the final version.

Kleydson Bonfim Andrade: contributed to the design of the study and the interpretation of results, contributed to the discussion, critically reviewed the manuscript and approved the final version.

Stefano Barbosa Codenotti: contributed to the design of the study and the interpretation of results, contributed to the discussion, critically reviewed the manuscript and approved the final version.

Elaine Silva Nascimento Andrade: contributed to the design of the study and results, contributed to the discussion, critically reviewed the manuscript and approved the final version interpretation.

Wildo Navegantes de Araújo: contributed to the design of the study and the interpretation of results, contributed to the discussion, critically reviewed the manuscript and approved the final version.

Fernanda Dockhorn Costa: contributed to the design of the study and the interpretation of results, contributed to the discussion, critically reviewed the manuscript and approved the final version.

Walter Massa Ramalho: contributed to the design of the study and the interpretation of results, contributed to the discussion, critically reviewed the manuscript and approved the final version.

Fredi Alexander Diaz-Quijano: contributed to the design of the study and the interpretation of results, provided significant inputs in the drafts, prepared the illustrations, contributed to the discussion, critically reviewed the manuscript and approved the final version.

DATA SHARING STATEMENT

Contextual data are available from the Brazilian Health Ministry website (www.datasus.gov.br/tabnet/tabnet.htm); Brazilian Institute of Geography and Statistics website (<http://www.ibge.gov.br>); and the Human Development Atlas in Brazil website (<http://atlasbrasil.org.br/2013/>). Tuberculosis case data can be made available by Brazilian Health Ministry (<http://portalsaude.saude.gov.br>). More detailed information about how to access data is described at supplementary material-2.

FUNDING

This study did not have specific funding.

COMPETING INTERESTS

None declared.

References

1. World Health Organization. Global tuberculosis report 2017. Geneva: World Health Organization, 2017. Available from: <http://apps.who.int/iris/handle/10665/254762> [cited 2017 Dec 9].
2. Ministério da Saúde. Indicadores prioritários para o monitoramento do Plano Nacional pelo Fim da Tuberculose como Problema de Saúde Pública no Brasil. Bol. Epidemiológico 2017;48:1-6.
3. San Pedro A, Oliveira RM de. Tuberculose e indicadores socioeconômicos: revisão sistemática da literatura. Rev Panam Salud Publica 2013;33:294–301.
4. Lacerda SNB, Temoteo RC de A, Figueiredo TMRM de, Luna DT de L, de Sousa AN, Abreu C, et al. Individual and social vulnerabilities upon acquiring tuberculosis: a literature systematic review. Int. Arch. Med. 2014 Jul;7(35):1-8. doi: <http://dx.doi.org/10.1186/1755-7682-7-35>.

5. Millet J-P, Moreno A, Fina L, et al. Factors that influence current tuberculosis epidemiology. *Eur. Spine J.* 2013;22(4):539–48. doi: <http://dx.doi.org/10.1007/s00586-012-2334-8>.

6. Houben RMGJ, Menzies N, Sumner T, et al. Feasibility of achieving the 2025 WHO global tuberculosis targets in South Africa, China, and India: a combined analysis of 11 mathematical models. *Lancet Glob. Health* 2016;4:e806–15. doi: [http://dx.doi.org/10.1016/S2214-109X\(16\)30199-1](http://dx.doi.org/10.1016/S2214-109X(16)30199-1).

7. Baicker K & Chandra A. Evidence-Based Health Policy. *N. Engl. J. Med.* 2017;377:2413–5. doi: <http://dx.doi.org/10.1056/NEJMp1709816>

8. Shin Y, Ki M, Sung N. A new strategy for tuberculosis control in North Korea. *Epidemiol. Health* 2015;37:e2015053. doi: <http://dx.doi.org/10.4178/epih/e2015053>.

9. Sharma DC. New plan to end tuberculosis in south and southeast Asia. *The Lancet* 2017;25;389(10075):1183. doi: [http://dx.doi.org/10.1016/S0140-6736\(17\)30817-6](http://dx.doi.org/10.1016/S0140-6736(17)30817-6).

10. Pai M, Correa N, Mistry N, et al. Reducing global tuberculosis deaths-time for India to step up. *The Lancet* 2017;25;389(10075):1174-6. doi: [http://dx.doi.org/10.1016/S0140-6736\(17\)30790-0](http://dx.doi.org/10.1016/S0140-6736(17)30790-0).

11. Ministério da Saúde. Departamento de Informática do SUS: Informações de saúde (TABNET). Brasília: Ministério da Saúde, 2017. Available from: <http://www.datasus.gov.br/informacoes-de-saude/tabnet> [cited 2017 Mar 4].

12. Ministério da Saúde. Plano Nacional pelo Fim da Tuberculose como Problema de Saúde Pública. Brasília: Ministério da Saúde, 2017. Available from: <https://drive.google.com/file/d/0B0CE2wqdEaR-eVc5V3cyMVFPcTA/view> [cited 2017 Dec 9].

13. Programa das Nações Unidas para o Desenvolvimento – PNUD. Atlas do Desenvolvimento Humano no Brasil. Programa das Nações Unidas para o Desenvolvimento – PNUD, 2015. Available from: <http://www.atlasbrasil.org.br/2013/pt/consulta/> [cited 2017 Mar 4].

14. Janssens J-P, Rieder HL. An ecological analysis of incidence of tuberculosis and per capita gross domestic product. *Eur. Respir. J.* 2008;32(5):1415-6. doi: <http://dx.doi.org/10.1183/09031936.00078708>.

15. Dye C, Lönnroth K, Jaramillo E, et al. Trends in tuberculosis incidence and their determinants in 134 countries. *Bull. World Health Organ.* 2009;87:683–91.

16. Pelissari DM, Diaz-Quijano FA. Household crowding as a potential mediator of socioeconomic determinants of tuberculosis incidence in Brazil. *PloS One* 2017;18;12(4):e0176116. doi: <http://dx.doi.org/10.1371/journal.pone.0176116>.

17. Ministério do Desenvolvimento Social e Combate à Fome. Guia de Orientação Técnica – SUAS N° 1 Proteção Social Básica de Assistência Social. Brasília: Ministério do Desenvolvimento Social e Combate à Fome, 2005. Available from: http://www.pg.pr.gov.br/cmas/wp-content/uploads/2011/07/guia_orientacao_tecnica.pdf [cited 2017 Mar 4].
18. Baussano I, Williams BG, Nunn P, et al. Tuberculosis incidence in prisons: a systematic review. *PLoS Med*. 2010;7(12):e1000381. doi: <http://dx.doi.org/10.1371/journal.pmed.1000381>.
19. Kodinariya, T. M. & Makwana, P. R. Review on determining number of Cluster in K-Means Clustering. *International Journal of Advance Research in Computer Science and Management Studies*. 2013;1:90–5.
20. Conselho Nacional de Saúde. Resolução N° 510, de 7 de abril de 2016. Brasília: Conselho Nacional de Saúde, 2016. Available from: <http://conselho.saude.gov.br/resolucoes/2016/Reso510.pdf> [cited 2017 Jan 15].
21. Conselho Nacional de Saúde. Resolução No 466, de 12 de dezembro de 2012. Brasília: Conselho Nacional de Saúde, 2012. Available from: http://bvsms.saude.gov.br/bvs/saudelegis/cns/2013/res0466_12_12_2012.html [cited 2017 Dec 9].
22. Cantwell MF, McKenna MT, McCray E, et al. Tuberculosis and race/ethnicity in the United States: impact of socioeconomic status. *Am. J. Respir. Crit. Care Med*. 1998;157:1016–20.
23. Díaz de Quijano E, Brugal MT, Pasarín M, et al. Influencia de las desigualdades sociales, la conflictividad social y la pobreza extrema sobre la morbilidad por tuberculosis en la ciudad de Barcelona. *Rev. Esp. Salud Pública* 2001;75:517–28.
24. Compton WM, Gfroerer J, Conway KP, et al. Unemployment and substance outcomes in the United States 2002–2010. *Drug Alcohol Depend*. 2014;1;142:350-3. doi: <http://dx.doi.org/10.1016/j.drugalcdep.2014.06.012>.
25. Támara-Ramírez JR, Álvarez CA, Rodríguez J. Loss of follow-up and associated factors in patients enrolled in the HIV/AIDS program of the Hospital Universitario San Ignacio, Colombia, 2012-2013. *Biomed. Rev. Inst. Nac. Salud* 2016;3;36(2):265-75. doi: <http://dx.doi.org/10.1016/10.7705/biomedica.v36i2.2939>.
26. Baker M, Das D, Venugopal K, et al. Tuberculosis associated with household crowding in a developed country. *J. Epidemiol. Community Health* 2008;62:715–21. <http://dx.doi.org/10.1136/jech.2007.063610>

27. Wingfield T, Schumacher SG, Sandhu G, et al. The seasonality of tuberculosis, sunlight, vitamin d, and household crowding. *J. Infect. Dis.* 2014;210, 774–83. <http://dx.doi.org/10.1093/infdis/jiu121>

28. Myers WP, Westenhouse JL, Flood J, et al. An ecological study of tuberculosis transmission in California. *Am. J. Public Health* 2006;96:685.

29. Lienhardt C, Fielding K, Sillah JS, et al. Investigation of the risk factors for tuberculosis: a case-control study in three countries in West Africa. *Int. J. Epidemiol.* 2005;34:914–23.

30. Kwan CK & Ernst JD. HIV and tuberculosis: a deadly human syndemic. *Clin. Microbiol. Rev.* 2011;24(2):351-76. doi: <http://dx.doi.org/10.1128/CMR.00042-10>.

31. Ministério da Justiça. Relatório do Levantamento Nacional de Informações Penitenciárias — Infopen. Brasília: Ministério da Justiça, 2014. Available from: <http://www.justica.gov.br/noticias/mj-divulgara-novo-relatorio-do-infopen-nesta-terca-feira/relatorio-depen-versao-web.pdf> [cited 2017 Mar 21].

32. Bourdillon PM, Gonçalves CCM, Pelissari DM, et al. Increase in Tuberculosis Cases among Prisoners, Brazil, 2009–2014. *Emerg. Infect. Dis.* 2017;23:496–9.

33. Ministério da Saúde. Departamento de DST, Aids e Hepatites Virais. Ano V - no 1 - 27a a 53a - semanas epidemiológicas - julho a dezembro de 2015 Ano V - no 1 - 01a a 26a - semanas epidemiológicas - janeiro a junho de 2016. *Bol. Epidemiológico - Aids E DST* (2016). Brasília: Ministério da Saúde, 2016.

34. Pelissari DM, Bartholomay PO, Jacobs MG, et al. Health care basic services and detection of tuberculosis incidence in Brazil. *Rev. Saúde Pública* 2018 in press.

35. Díaz-Quijano FA, Waldman EA. Factors associated with dengue mortality in Latin America and the Caribbean, 1995–2009: an ecological study. *Am. J. Trop. Med. Hyg.* 2012;86(2):328-34. doi: <http://dx.doi.org/10.4269/ajtmh.2012.11-0074>.

36. Gibbons CL, Mangen M-JJ, Plass D, et al. Measuring underreporting and underascertainment in infectious disease datasets: a comparison of methods. *BMC Public Health* 2014;11;14:147. doi: <http://dx.doi.org/10.1186/1471-2458-14-147>.

37. Reeves A, Basu S, McKee M, et al. Social protection and tuberculosis control in 21 European countries, 1995-2012: a cross-national statistical modelling analysis. *Lancet Infect. Dis.* 2014;14, 1105–12. doi: [http://dx.doi.org/10.1016/S1473-3099\(14\)70927-2](http://dx.doi.org/10.1016/S1473-3099(14)70927-2)

38. Siroka A, Ponce NA, Lönnroth K. Association between spending on social protection and tuberculosis burden: a global analysis. *Lancet Infect. Dis.* 2016;16, 473–9. doi: [http://dx.doi.org/10.1016/S1473-3099\(15\)00401-6](http://dx.doi.org/10.1016/S1473-3099(15)00401-6)

39. Eang MT, Satha P, Yadav RP, et al. Early detection of tuberculosis through community-based active case finding in Cambodia. BMC Public Health 2012;12,469. doi: <http://dx.doi.org/10.1186/1471-2458-12-469>
40. Fox GJ, Nhung V, Sy DN, et al. Household-Contact Investigation for Detection of Tuberculosis in Vietnam. N. Engl. J. Med. 2018;378, 221–9. doi: <http://dx.doi.org/10.1056/NEJMoa1700209>

For peer review only

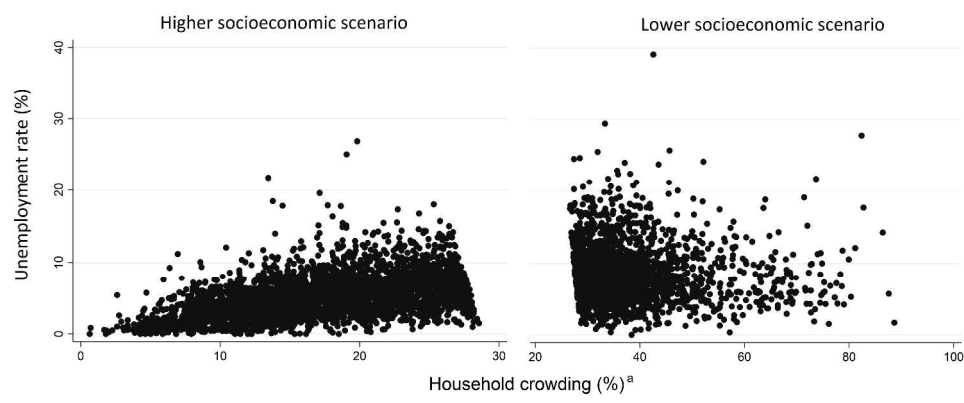


Figure 1. Distribution of Brazilian municipalities according to socioeconomic variables associated with the tuberculosis incidence rate

^aProportion of the population living in households with more than two people per room

704x394mm (300 x 300 DPI)

review only

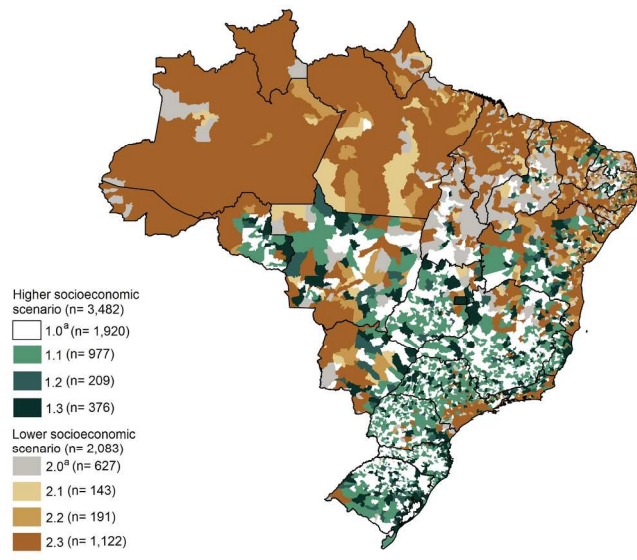


Figure 2 Brazilian's Municipalities by Tuberculosis Incidence Rate Scenarios. Brazil, 2015

^a Without cases in 2014 or 2015.

209x148mm (300 x 300 DPI)

Supplementary material 1- List of analysed variables and studies that found an association with tuberculosis incidence at the individual and contextual levels

Variables	Studies that found an association with tuberculosis incidence
Socioeconomic	
Municipal human development index	San Pedro A & Oliveira RM, 2013 ^{a,3} Dye C et al., 2009 ¹⁵ Pelissari DM & Diaz-Quijano FA, 2017 ¹⁶
Average household income per capita	San Pedro A & Oliveira RM, 2013 ^{3,a} Pelissari DM & Diaz-Quijano FA, 2017 ¹⁶
Gross domestic product (GDP) per capita	Janssens J-P & Rieder HL, 2008 ¹⁴ Dye C et al., 2009 ¹⁵ San Pedro A & Oliveira RM, 2013 ^{a,3} Pelissari DM & Diaz-Quijano FA, 2017 ¹⁶
Population that is extremely poor, poor, and vulnerable to poverty	Millet J-P et al., 2013 ^{a,5} Pelissari DM & Diaz-Quijano FA, 2017 ¹⁶
Gini coefficient	San Pedro A & Oliveira RM, 2013 ^{a,3} Pelissari DM & Diaz-Quijano FA, 2017 ¹⁶
Unemployment rate	San Pedro A & Oliveira RM, 2013 ^{a,3} Pelissari DM & Diaz-Quijano FA, 2017 ¹⁶
Illiteracy rate	San Pedro A & Oliveira RM, 2013 ^{a,3} Pelissari DM & Diaz-Quijano FA, 2017 ¹⁶
Household crowding	San Pedro A & Oliveira RM, 2013 ^{a,3} Lienhardt C et al., 2005 ²⁹ Myers WP et al., 2006 ²⁸ Baker M et al., 2008 ²⁶ Wingfield T et al., 2014 ²⁷ Pelissari DM & Diaz-Quijano FA, 2017 ¹⁶
Infant mortality rate	-
Life expectancy at birth	-
Population size of municipalities	-
Epidemiological	
AIDS case detection rate	San Pedro A & Oliveira RM, 2013 ^{a,3} Millet J-P et al., 2013 ^{a,5} Pelissari DM & Diaz-Quijano FA, 2017 ¹⁶
Proportion of new TB cases who were: HIV positive; prisoners; health professionals; indigenous; homeless	Baussano I et al., 2010 ^{a,18} San Pedro A & Oliveira RM, 2013 ^{a,3} Millet J-P et al., 2013 ^{a,5} Lacerda SNB et al., 2014 ^{a,4}
Proportion of TB retreatment	Millet J-P et al., 2013 ^{a,5}
TB mortality rate	-
Operational health care	
Proportions of new tuberculosis cases in which contacts were examined	San Pedro A & Oliveira RM, 2013 ^{a,3} Millet J-P et al., 2013 ^{a,5}
Proportions of new tuberculosis cases that were laboratory confirmed	-
Proportions of new tuberculosis cases tested for HIV	-
Treatment outcomes (cure, loss to follow-up, and no record of TB outcome) in the new tuberculosis cases	-
Proportion of sputum culture examination among retreatment cases	-

^a Systematic review

*The superscript number corresponds to the reference order

Supplemental material 2- Access to the data used in this study

Municipalities in Brazil have a unique code (código IBGE) which facilitates merging data and is available in all data sources listed below.

The socioeconomics datasets from Brazil used in the current study are available at the following:

- “Departamento de Informática do SUS” repository:
 - Click on the following link: <http://datasus.saude.gov.br/>
 - Click on “Acesso à informação” (Information access) -> Informação de Saúde (TABNET) -> “Demográficas e Socioeconômicas” (Demographic and socioeconomic).
 - Select an indicator group. We used the following: População residente (resident population); Educação (education indicators); Trabalho e renda (labour and income indicators); Produto Interno Bruto (GDP per capita).
 - Within each indicator group, select an indicator.
 - For each one, select the “Abrangência Geográfica” (geographic dimension) field: “Brasil por município” (Brazil by municipality).
 - In the “Linha” field (line), select “Município” (municipality), and in the “Coluna” field. (column), select the “Ano” (year); and in the “Período Disponível” field (Available period), we used 2010.
 - Click “Mostrar” (Show), and the indicator will be calculated for all municipalities.
 - You now can export in different formats. Export options are at the bottom of the page.
- Human Development Atlas repository:
 - Click on the link: <http://www.atlasbrasil.org.br/2013/en/consulta/>
 - First, select a “Locality”; click “Municipalities” and then the checkbox “All municipalities - Brazil”.
 - Now, select the group of indicators and then the indicator. We used the following group indicators: MHDI; Demography; Income; Labour; Housing; Vulnerability
 - After loading the data, download the dataset by clicking on the icon at the right top of the page.

The datasets of tuberculosis indicators in Brazil are available by request at the Ministry of Health, as follows:

- Click on the following link: <https://esic.cgu.gov.br/sistema/site/index.aspx>
- Complete the form to create a new registration.
- Register your request for access to data, detailing the information and indicators you have interest in, and send the request.
- The Ministry of Health has four weeks to answer the request.

The datasets of AIDS notification by municipalities in Brazil is available in the “Departamento de Informática do SUS” repository:

- Click on the following link: <http://datasus.saude.gov.br/>
- Click on “Acesso à informação” (Information access) -> Informações de Saúde (TABNET) -> “Epidemiológica e Morbidade” (Epidemiologic and morbidity).
- Select “Casos de Aids - Desde 1980 (SINAN)” (Aids cases- since 1980).
- In the “Abrangência Geográfica” field (geographic dimension), select “Brasil por Região, UF e município” (Brazil by region, state and municipality).
- In the “Linha” field (line), select “Município de residência” (resident municipality), and in the “Coluna” field (column), select the “Ano diagnóstico” (diagnose year). In the “Período Disponível” field (Available period), we used 2015.
- Click on “Mostrar” (Show), and the number of AIDS cases will be calculated for all municipalities.
- You now can export the data in different formats.

Supplementary material 3- Socioeconomic, epidemiological and operational-tuberculosis indicators by tuberculosis scenario in Brazil

Variables ^a	Higher socioeconomic n=3 482					Lower socioeconomic n=2 083		
	Subs. 1.0 ^b	Subs. 1.1	Subs. 1.2	Subs. 1.3	Subs. 2.0 ^b	Subs. 2.1	Subs. 2.2	Subs. 2.3
Socioeconomics indicators								
Number of municipalities (%)	1 920 (34.5)	977 (17.6)	209 (3.8)	376 (6.8)	627 (11.3)	145 (2.6)	191 (3.4)	1 122 (20.2)
Total population- No (%)	13 465 985 (6.6)	27 729 472 (13.6)	6 030 335 (2.9)	56 866 416 (27.8)	5 365 790 (2.6)	2 962 579 (1.4)	6 373 455 (3.1)	85 643 554 (41.9)
HDI-M	0.683	0.696	0.682	0.726	0.59	0.589	0.597	0.618
GDP per capita (USD)	14 407	15 619	13 403	19 565	6 752	9 226	7 572	10 203
Extremely poor (%)	6.1	5.5	7.6	3.9	23.3	9.4	21.3	18.7
Gini coefficient	0.47	0.49	0.49	0.51	0.53	0.54	0.54	0.54
Unemployment rate (%)	4.5	5.7	5.9	6.6	7.7	8.8	8.7	8.6
Household crowding (%)	16	18.1	18.6	19.8	36.6	38.8	38.5	39.5
Life expectancy at birth (years)	74.1	74.3	73.8	75	70.6	70.6	70.9	71.5
Infant mortality rate (number of deaths in the first year of life/1000 live births)	16.1	16.1	17.3	14.8	25.9	26.1	25.2	23.6
Epidemiological indicators								
N of new cases (% in relation to overall new cases)	978 (1.4)	5 269 (7.8)	1 353 (2)	18 865 (27.8)	358 (0.5)	586 (0.9)	2 205 (3.3)	38 161 (56.3)
TB incidence rate (cases/ 100 000 people)	7.8	22.3	24.5	39.6	7	9.9	30.3	29.2
AIDS case detection rate (cases/100 000 people) ^{c,d}	6.4	11.1	11.9	21.6	5.2	6.6	8.2	10.1
TB mortality rate (deaths/ 100 000 people) ^{c,d}	0.7	1.3	1.6	1.6	1	0.8	2.3	2.1
New cases from at least one vulnerable group (%) ^{c,d,e}	9.9	14.3	19.2	22.6	12.1	16.6	10.6	14.3
TB-HIV confection among new cases (%)	5.9	8	7.7	8.8	6.9	8.8	3.6	5.6
New cases who were prisoners (%)	2.7	4.2	6.7	12.1	2.5	3.8	2.7	4.1
New cases who were health professionals (%)	0.7	1	1.5	1	1.1	1.5	1.3	0.9
New cases who were from an indigenous population (%)	0.4	0.7	1.6	0.6	1.5	0.4	2.4	3.5
New cases who were homeless (%)	0.7	1.3	1.9	1.4	0.7	1.3	1.1	1.1
Retreatment cases among total cases (%)	13.3	8.1	9.1	13.7	20.7	11.1	9.4	11.1
Operational indicators (new cases)								
Contact examination (%) ^c	70.6	78.8	54.3	73.3	55.7	55.5	60.5	69.6
Pulmonary cases with laboratory confirmation (%)	65.4	70.4	64.0	71.0	66.5	65.8	66.2	71.4
Tested for HIV (%)	72.1	75.5	67.2	75.3	63.9	63.3	53.5	67.2
Cure (%)	66.8	84.8	32.2	73.1	65.4	65.6	51.1	79.7
Lost to follow-up (%)	5.1	5.8	2.8	8	3.8	5.5	5.2	8.2
No TB outcome registration (%) ^d	14.5	0	60.1	11	22.7	4.8	37	4.6
Culture examination (retreatment) (%)	35.5	39	25.1	44	25.5	45.4	21.5	27.8

^aSocioeconomic indicators: with the exception of the number of municipalities and total population (data from 2015), the other variables were measured in 2010. Epidemiological indicators: with the exception of TB mortality rate (data from 2014), the other variables were measured in 2015. Operational indicators: with the exception of cure, lost to follow-up, no TB outcome registration and culture examination (data from 2014), the other variables were measured in 2015. ^bWithout cases in 2014 or 2015; ^cVariables used in the non-hierarchical clustering method for scenario 1; ^dVariables used in the non-hierarchical clustering method for scenario 2; ^eHIV, health professional, prisoners, indigenous and homeless. Municipal Human Development Index- HDI; Gross domestic product-GDP.

STROBE 2007 (v4) checklist of items to be included in reports of observational studies in epidemiology*

Checklist for cross-sectional studies (combined)

Section/Topic	Item #	Recommendation	Reported on page #
Title and abstract	1	(a) Indicate the study’s design with a commonly used term in the title or the abstract	2 and 3
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	3
Objectives	3	State specific objectives, including any pre-specified hypotheses	3
Methods			
Study design	4	Present key elements of study design early in the paper	3
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	3 and 4
Participants	6	(a) Cohort study—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up Case-control study—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls Cross-sectional study—Give the eligibility criteria, and the sources and methods of selection of participants	3 and 4
		(b) Cohort study—For matched studies, give matching criteria and number of exposed and unexposed Case-control study—For matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	4
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	3 and 4
Bias	9	Describe any efforts to address potential sources of bias	5
Study size	10	Explain how the study size was arrived at	3
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	4
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding	4 and 5
		(b) Describe any methods used to examine subgroups and interactions	4 and 5
		(c) Explain how missing data were addressed	5
		(d) Cross-sectional study—If applicable, describe analytical methods taking account of sampling strategy	-
		(e) Describe any sensitivity analyses	-

Results				
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed		5 and 6
		(b) Give reasons for non-participation at each stage		6 and 7
		(c) Consider use of a flow diagram		-
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders		6 and 7
		(b) Indicate number of participants with missing data for each variable of interest		3
Outcome data	15*	Cross-sectional study—Report numbers of outcome events or summary measures		6-8
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included		6 and 8
		(b) Report category boundaries when continuous variables were categorized		-
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period		-
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses		8 and 9
Discussion				
Key results	18	Summarise key results with reference to study objectives		9-11
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias		11
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence		9-11
Generalisability	21	Discuss the generalisability (external validity) of the study results		11-12
Other information				
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based		14

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at <http://www.plosmedicine.org/>, Annals of Internal Medicine at <http://www.annals.org/>, and Epidemiology at <http://www.epidem.com/>). Information on the STROBE Initiative is available at www.strobe-statement.org.