A multilevel regression analysis of incidence in tuberculosis in Bolivia

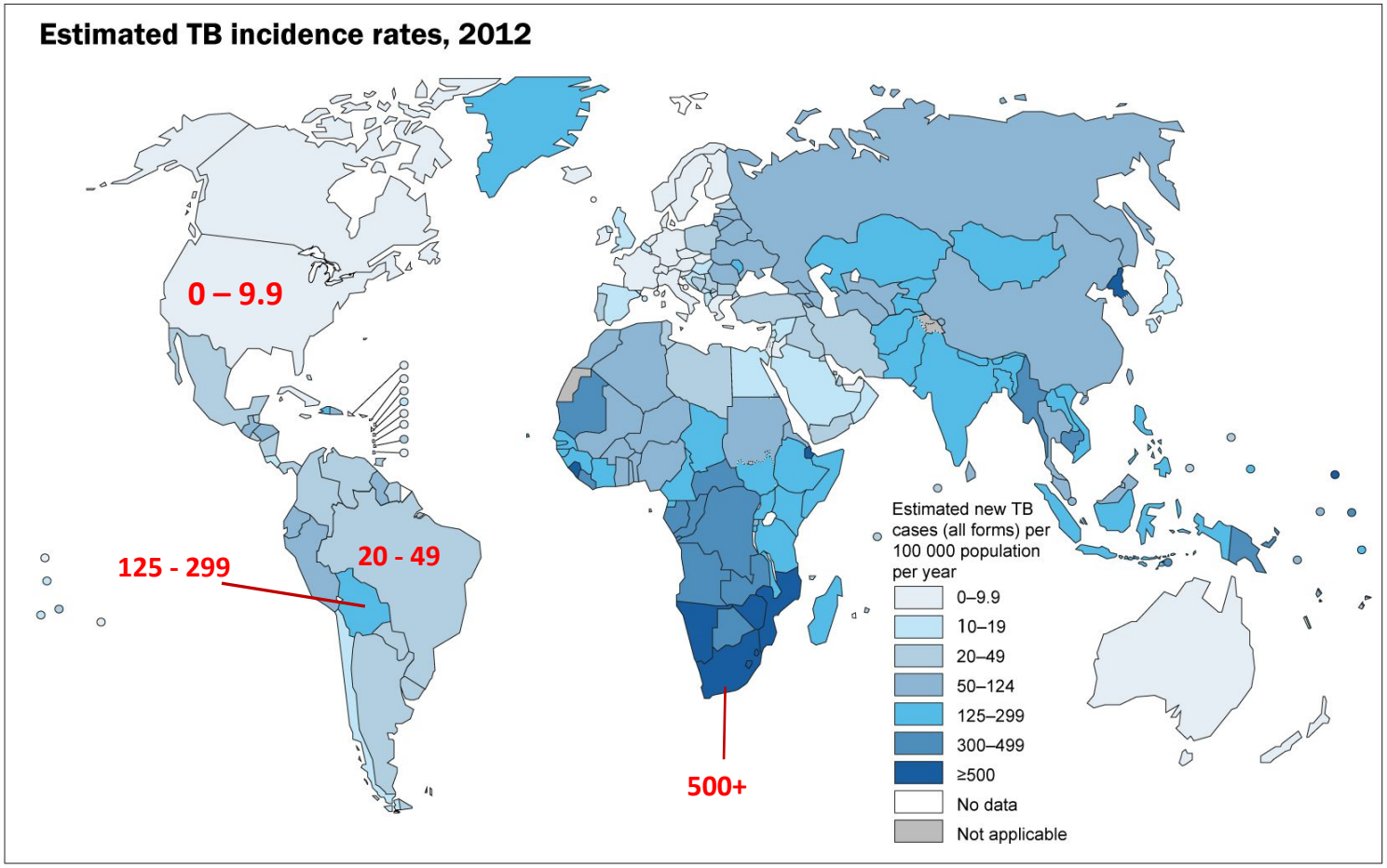
N. Saravia

April 2016

Presented at

Spatial Statistics and Disease Mapping Paper Session at

the AAG 2016 Annual Meeting.



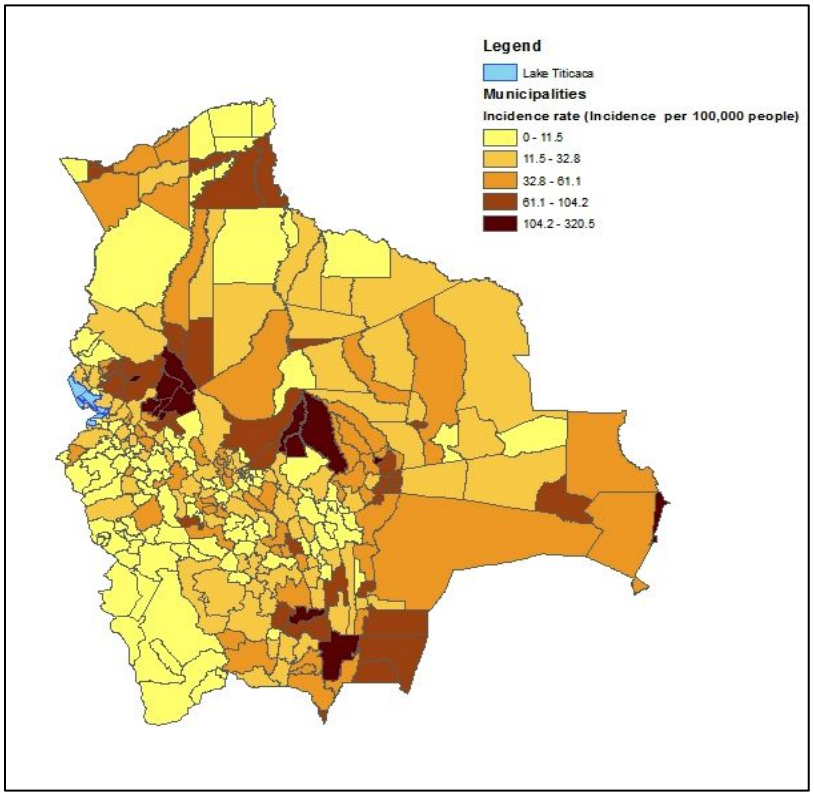
The boundaries and names shown and the designations used on this map do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted and dashed lines on maps represent approximate border lines for which there may not yet be full agreement.

Data Source: *Global Tuberculosis Report 2013*. WHO, 2013.

World Health Organization

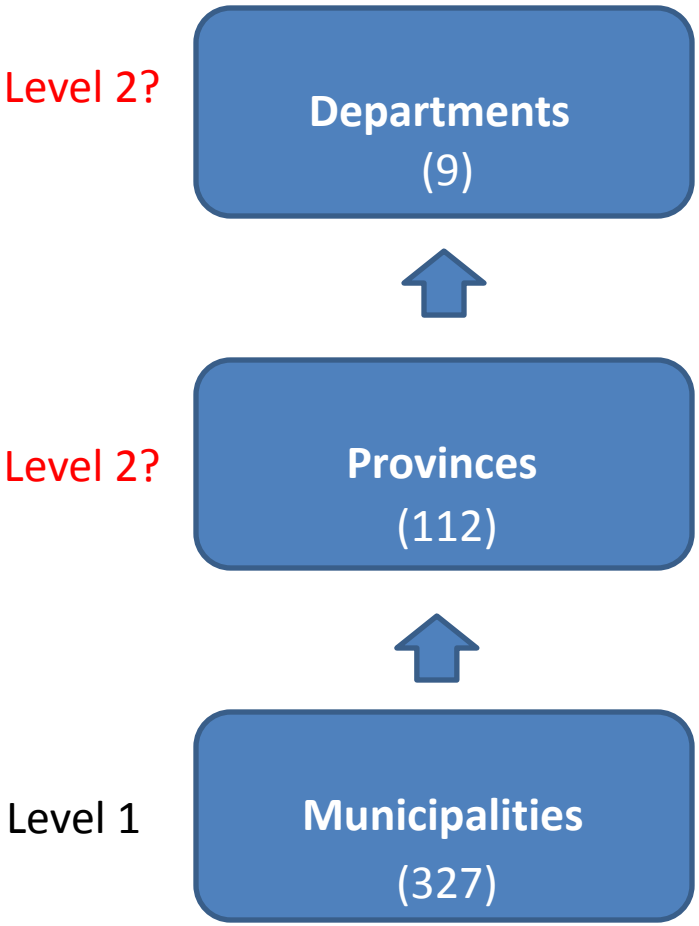
Geo-Administrative Departments

Incidence rate per Municipality



Nesting model based on geo-administrative structure

Geo-Administrative Classification

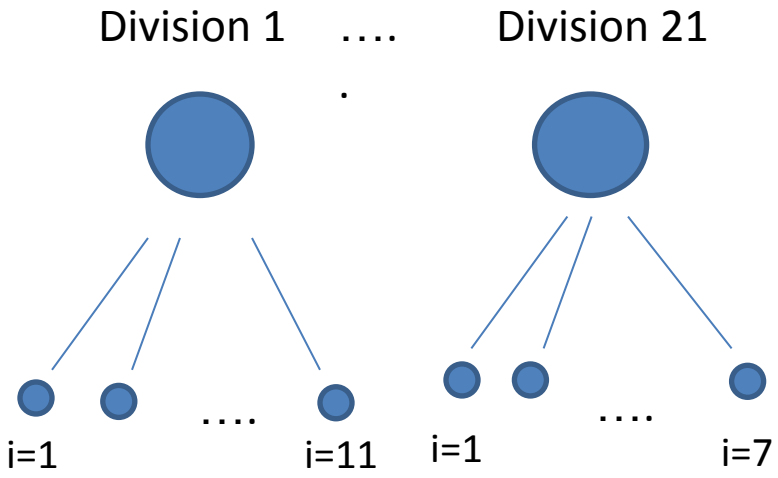


-In this structure, there are three to four municipalities per province. Too small to be representative units of level 1 within level 2 layer.

- Number of department units (9) is also too small to be considered as level 2.

Solution to improve the multilevel model:
A different nesting arrangement of the municipalities.

Nesting based on geo-divisions



-Municipalities grouped into country boundary-divisions and interior-divisions.

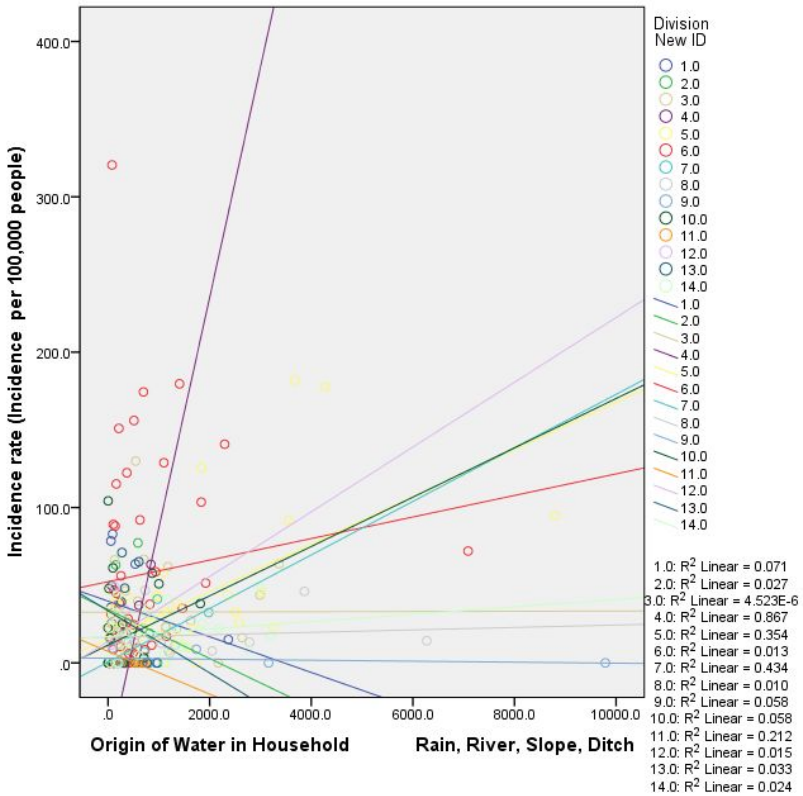
-Increases size of units per division without altering the data.

-21 divisions with an average of 17 municipalities per division.

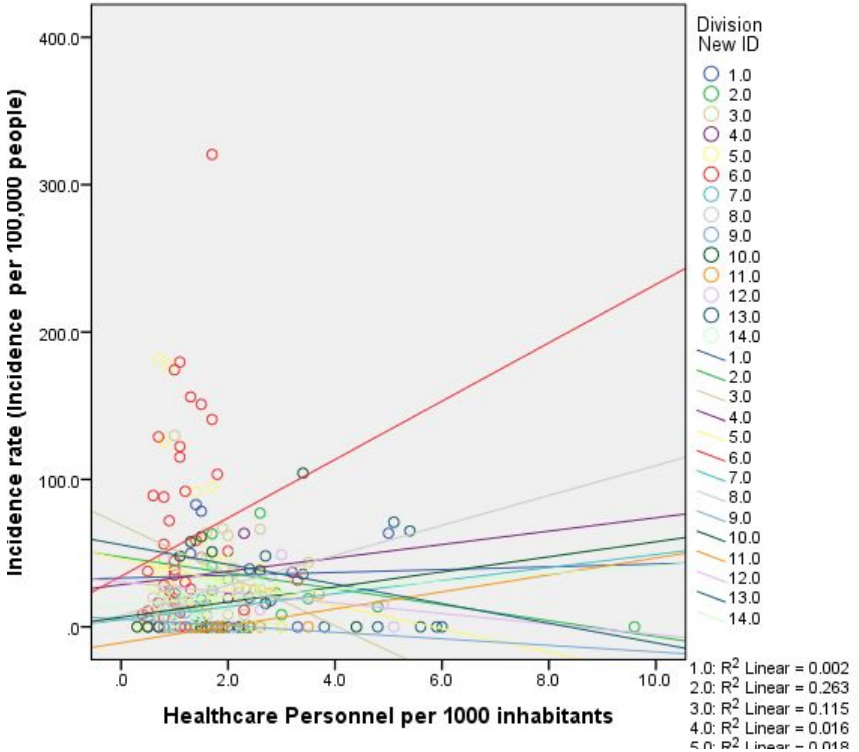
Divisions table

Department	Division ID	Number of Municipalities per Division
Beni	1	11
	2	8
Chuquisaca	3	25
	4	3
Cochabamba	5	45
La Paz	6	45
	7	8
	8	19
	9	8
Oruro	10	27
	11	8
Pando	12	6
	13	9
Potosi	14	28
	15	4
	16	6
Santa Cruz	17	42
	18	7
	19	7
Tarija	20	4
	21	7

Single-level regressions of Incidence vs. Predictors

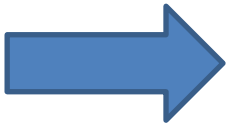


Incidence vs. Water origin (rain, river, etc.)



Incidence vs. healthcare personnel

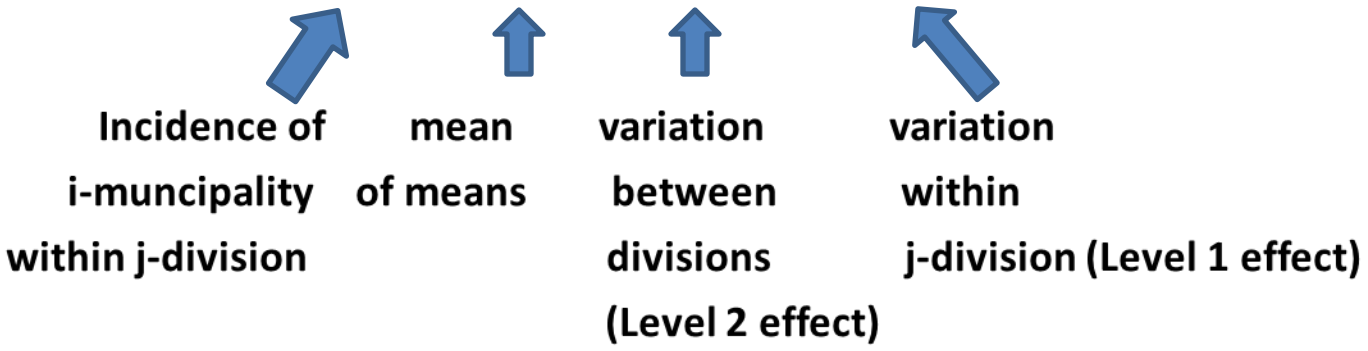
Great variance of slopes and intercepts for different divisions



Multi-level regression analysis

Two-level regression model for Incidence

- Null Model: $I_{ij} = \delta_{00} + \upsilon_{0j} + \varepsilon_{ij}$



- No predictors in this model

- Intraclass correlation $= \frac{\sigma_B^2}{(\sigma_B^2 + \sigma_W^2)}$ = 11.2% (>5%)

- σ_B^2 : variance between divisions
- σ_W^2 : variance within divisions

Level 1 Random Intercept Model

- $I_{ij} = \delta_{00} + u_{0j} + \varepsilon_{ij} + \alpha_{1j} X_{ij}$

Null Model

Variation due to predictor-X

Several runs of the SPSS Linear Mixed Model with different predictors were performed to obtain the correlation between the Incidence and predictors.

Multilevel Regression Models for Incidence versus several Predictors

Model #	Predictor (X_{ij})	Intercept			Slope			ICC (%)	BIC	BIC difference
		d_{00}	SE	p-value	d_{00}	SE	p-value			
1	Null Model	28.197	3.8	<0.001	---------			11.16	3309.686	0.00
2	Primary school coverage	29.694	3.1	<0.001	0.54	0.09	<0.001	6.50	3280.087	-29.599
3	Poverty	28.724	3.3	<0.001	- 0.59	0.11	<0.001	8.42	3285.256	-24.430
4	Heath personnel/1,000	29.176	3.9	<0.001	-3.16	1.96	<0.011	11.2	3309.90	0.214
5	Hospital rates (x100,000)	28.629	3.7	<0.001	-0.03	0.025	0.129	9.90	3313.569	3.883
6	Street discharge household rate	28.245	3.8	<0.001	$1.7x10^{-3}$	$3.8x10^{-4}$	0.65	10.94	3318.79	9.104
7	Income per capita	28.64	3.9	<0.001	-0.002	$3.1x10^{-3}$	0.448	11.51	3318.815	9.129
8	Sewage river discharge rate	28.266	3.8	<0.001	$3.1x10^{-4}$	$6.5x10^{-4}$	0.64	10.10	3322.30	12.614
9	Septic tank household rate	28.189	3.8	<0.001	$2.0x10^{-4}$	$4.6x10^{-4}$	0.651	10.60	3323.00	13.314
10	Water source from lake/pond rate/100,000	28.460	3.7	<0.001	$2.4x10^{-4}$	$2.0x10^{-4}$	0.215	10.58	3323.37	13.684
11	Water source from well rate/100,000	28.534	3.7	<0.001	$1.1x10^{-4}$	$6.5x10^{-5}$	0.093	11.30	3324.38	14.694
12	Public faucet water rate/100,000	28.406	3.7	<0.001	$1.2x10^{-4}$	$1.1x10^{-4}$	0.273	10.24	3324.9	15.214
13	Population density	28.512	3.7	<0.001	$3.6x10^{-5}$	$1.7x10^{-5}$	0.042	10.40	3325.621	15.935
14	Water source from rain/river rate/100,000	28.298	3.8	<0.001	$-1.8x10^{-4}$	$1.8x10^{-5}$	0.530	10.79	3329.3	19.614
15	Sewage house rate	28.189	3.8	<0.001	$-2.0x10^{-6}$	$1.0x10^{-5}$	0.84	11.20	3330.882	21.196
16	Tap water household rate/100,000	28.206	3.8	<0.001	$-1.0x10^{-6}$	$8.0x10^{-6}$	0.927	11.10	3331.279	21.593

Multilevel Regression Models for Incidence versus Multi – Predictors

Model #	Predictors (X_{ij})	Intercept			ICC (%)	BIC	Model comparison	BIC difference
		d_{00}	SE	p-value				
1	Null Model	28.197	3.8	<0.001	11.16	3309.686	--------	0.00
17	Poverty	28.724	3.3	<0.001	8.4	3285.256	17 vs 1	- 24.43
18	Model 17 + Prim. School coverage	29.638	3.05	<0.001	6.6	3271.286	18 vs 17	-13.97
19	Model 18 + Hospitals rate	29.725	3.04	<0.001	6.4	3276.726	19 vs 18	+5.44
20	Model 18 + Health personnel/1,000	30.857	2.99	<0.001	6.1	3261.486	20 vs 18	-9.80
21	Model 20 + Hospitals rate	30.837	3.01	<0.001	6.2	3266.990	21 vs 20	+5.50
22	Model 20 + Income per capita	30.62	2.99	<0.001	6.0	3270.157	22 vs 20	+8.67
23	Model 20 + Population density	30.852	2.99	<0.001	6.1	3269.964	23 vs 20	+8.48
24	Model 20 + Street Sew. discharge	30.852	3.0	<0.001	6.1	3270.913	24 vs 20	+9.43
25	Model 20 + Public water source	30.85	3.0	<0.001	6.2	3277.98	25 vs 20	+16.50

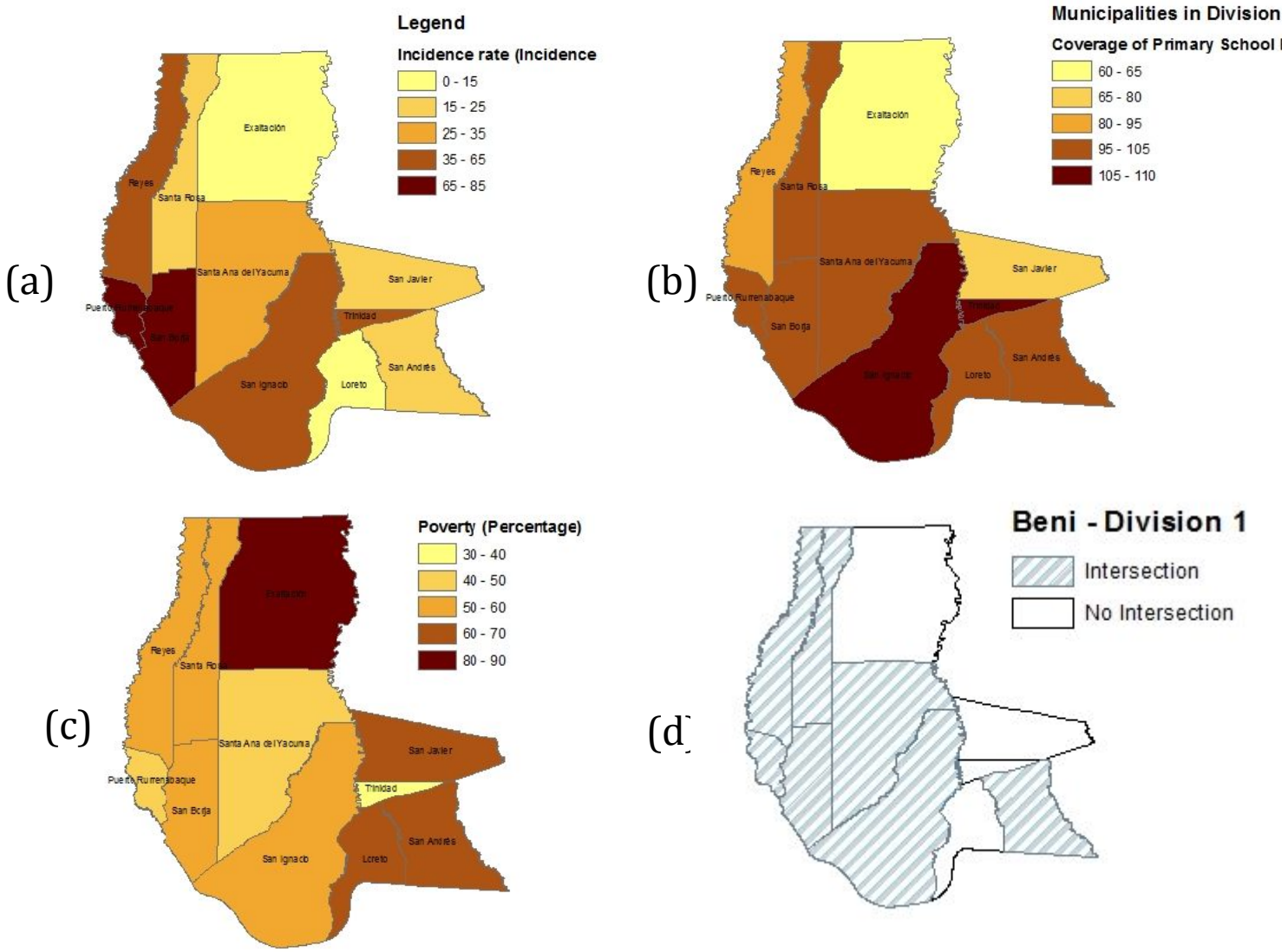


Figure 5. Maps of (a) Incidence, (b) Prim. School Cov., (c) Poverty. (d) Map pf Intersection , in Division 1, Beni department.

Data Sources:

Bolivian Department of Health Tuberculosis Database (2015)

WHO Global Tuberculosis Report (2013)