Identification of rat hotspots and their effect on leptospirosis risk in a Brazilian slum community

Poppy Miller [1,4] Kate Hacker [2] Peter Diggle [1] Mike Begon [3] James Childs [2] Albert Ko [2] Federico Costa [2] Mitermayer Reis [2] Chris Jewell [1]

[1]**Lancaster University**

[2]**Yale University**

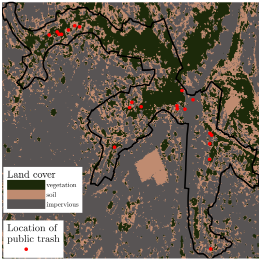
[3]**Liverpool University**

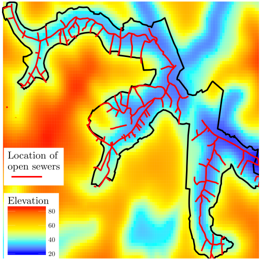
[4]**AgResearch**

November 26, 2018

Study area: Pau da Lima, Salvador, Brazil

- 88% squatters
- 66% did not finish primary school
- mean per capita daily household income: NZD$3.85
- open sewers and rubbish dumps





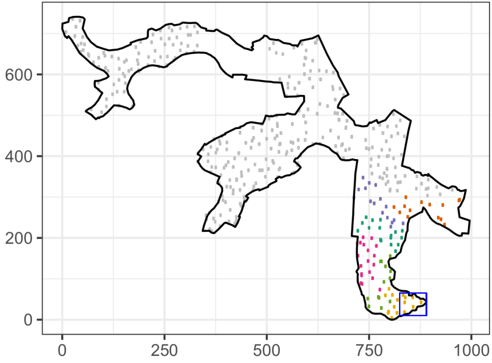
How you can get infected

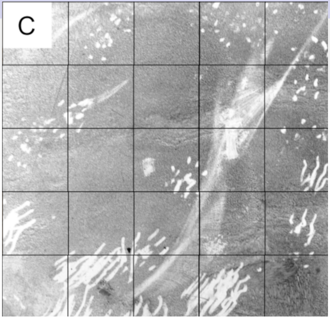
Leptospirosis

*Reservoir hosts are animals that harbors or nourishes a pathogen (a harmful organism) and serves as a source of infection.

Rat prevalence: study design

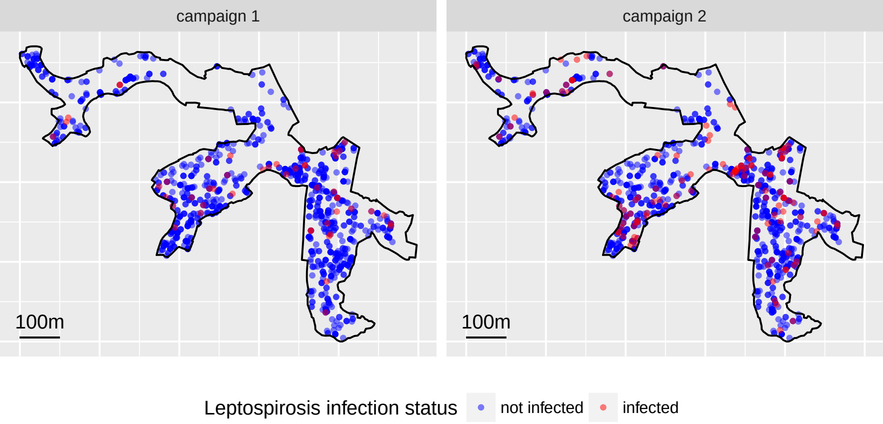
- **area**: 0.17 km^2
- **locations**: 340 + 100 close range
- **plates**: 5 per location over 2 days
- **campaigns**: 2 (dry and wet season)



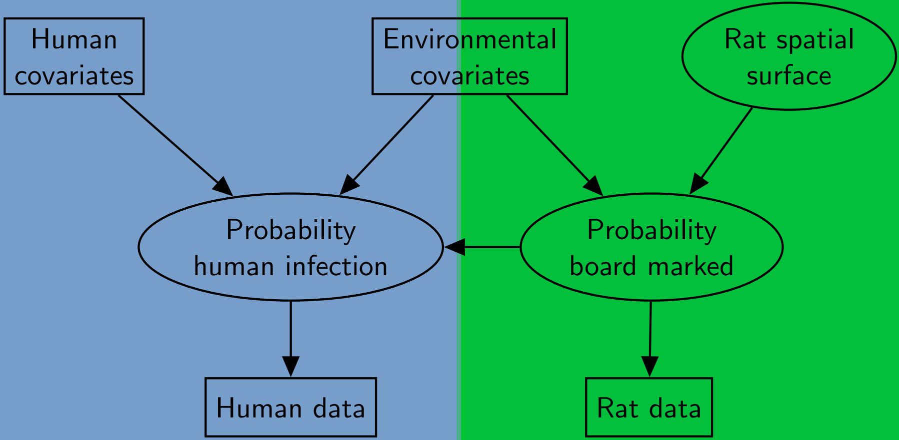


Human leptospirosis prevalence: study design

- 1110 residents in the study area
- titres measured before and after each rat tracking campaign
- pairs of titres determine residents infection status



Model overview



Rat model

$$m_{ic} \sim \text{Binomial}\left(p_{ic}, n_{ic}\right)$$
$$\text{cloglog}\left(p_{ic}\right) = x_{ic}^{T}\beta + S_{ic} + \log\left(\text{T}_{ic}\right)$$

$$S_c \sim \text{MVN}\left(0, \tau^2 + \Sigma_c\right)$$
$$\Sigma_{cws} = \sigma^2\left(1 + V_c\right)\exp\left(-V_c\right)$$
$$V_c = \left(\sqrt{3}b_{cws}\right)/\phi$$

- S_c : Matern 3/2 spatially correlated random effects (separate surface for each campaign)
- b_{cws} : distance between points w and s in campaign c (meters)

- m_{ic} boards positive for rat marks out of n_{ic} total for location i, and campaign c
- p_{ic} : probability of rat marks
- T_{ic} : offset (number of nights board exposed)
- X : rat covariate matrix
- Priors:
  $\beta_k \sim \text{Normal}\left(0, 100\right)$
  $\tau^2, \sigma^2 \sim \text{Gamma}\left(2, 0.5\right)$
  $\phi \sim \text{Gamma}\left(1.5, 0.05\right)$

Human Model

$$y_{jc} \sim \text{Bernoulli}\left(\pi_{jc}\right)$$
$$\text{logit}\left(\pi_{jc}\right) = z_{kjc}^{T}\gamma + \theta\left(x_j^T\beta + S_{jc}\right) + \delta_k$$

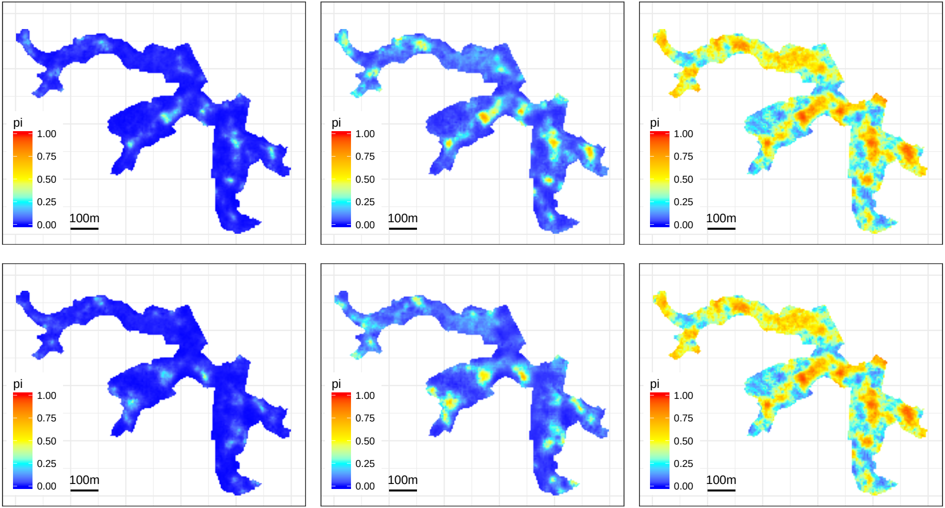
- y_{kjc} : human infection status for location j
- π_{kjc} : probability of human infection

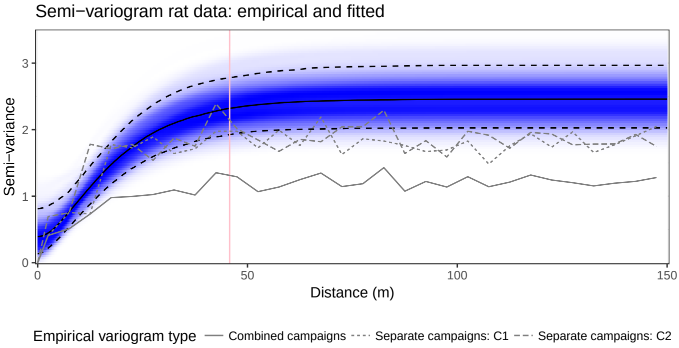
- Z : human covariate matrix
- X : rat covariate matrix at human locations
- S_{jc} : predicted spatial random effect at human locations
- γ, θ : coefficients
- δ_k : random effect for each individual k

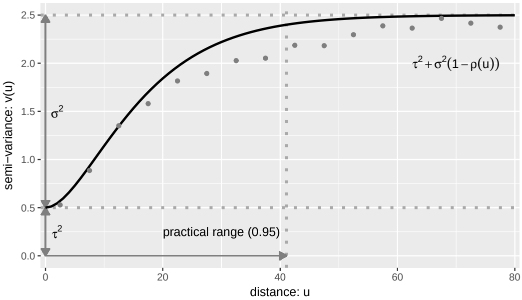
- Priors: $\gamma_k \sim \text{Normal}\left(0, 100\right)$, $\delta_i \sim \text{Normal}\left(0, \sigma_H\right)$, $\sigma_H \sim \text{Gamma}\left(2, 0.5\right)$

Predictive rat surfaces for campaigns 1 and 2

Lower CI (0.05) | **Median** | **Upper CI (0.95)**





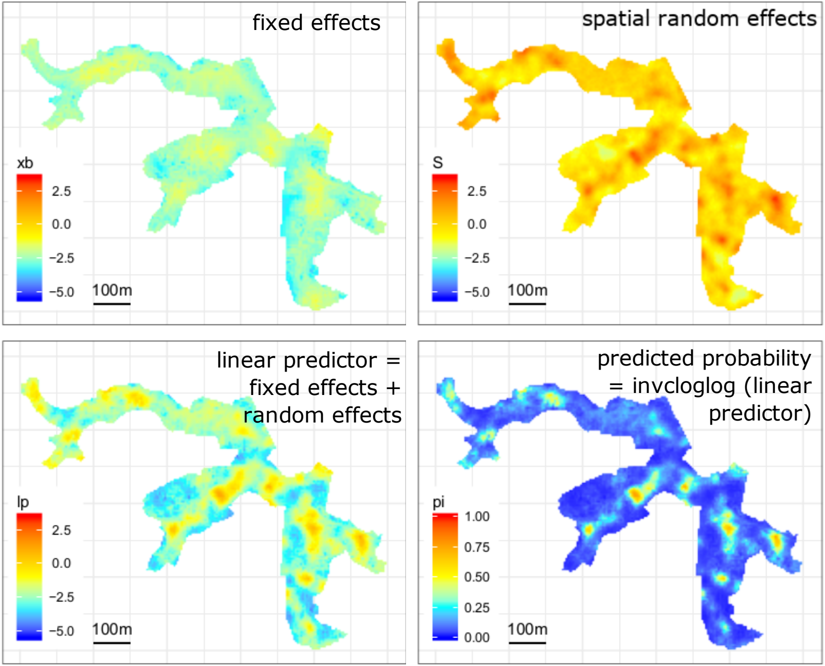


	Median (CI)
ϕ	17.31 (12.85, 22.86)
σ^2	2.03 (1.40, 2.71)
τ^2	0.39 (0.08, 0.85)

Rate ratios: rat model

Interpretation: The rate of rat mark deposition at the upper quartile value of a covariate is $RR_{U/L}$ multiplied by the rate at the lower quartile of the covariate.

	Rate ratio$_{U/L}$	**Data LQ**	**Data UQ**
Continuous variables			
Mean rainfall (mm)	1.46 (1.26, 1.66)	0.3	6.8
Distance 3d public dump (m)	0.60 (0.38, 0.82)	30.5	96.6
Distance 3d open sewer (m)	0.76 (0.60, 0.96)	9.6	17.6
Ground cover % soil	1.49 (0.95, 2.26)	6	41
Ground cover % vegetation	0.71 (0.43, 1.07)	0	57
Binary variables	**Rate ratio$_{1/0}$**		
Domestic / non-domestic	1.14 (0.82, 1.45)		


fixed effects
xb
2.5
0.0
−2.5
−5.0
100m
spatial random effects
S
2.5
0.0
−2.5
−5.0
100m
linear predictor = fixed effects + random effects
lp
2.5
0.0
−2.5
−5.0
100m
predicted probability = invcloglog (linear predictor)
pi
1.00
0.75
0.50
0.25
0.00
100m

Odds ratios: leptospirosis model (significant variables)

Interpretation: The odds of being infected with leptospirosis for a person with a covariate value at the upper quartile for that covariate are $OR_{U/L}$ times those at the lower quartile for that covariate.

	Odds ratio$_{U/L}$	**Data LQ**	**Data UQ**
Continuous variables			
Distance public dump (m)	0.44 (0.27, 0.63)	32.7	90.6
Log income (reias/month)	0.64 (0.30, 1.09)	log(1)	log(728)
Cumulative rainfall (m)	4.12 (2.45, 6.33)	0.56	1.70
Age (years)	13.28 (5.38, 27.01)	15	42
Rat linear predictor	1.03 (1.00, 1.07)	0.033	0.214
Binary variables	**Odds ratio$_{1/0}$**		
Male / Female	3.78 (1.96, 6.33)		

Practical implications

Target interventions to decrease leptospirosis risk:

- why does increasing rainfall increase risk?
- why are men and young people more at risk?
- increase incomes?
- remove or cover public dumps
- decrease rat numbers

Target interventions to reduce rat numbers:

- cover open sewers
- remove or cover public dumps
- rodenticide campaigns targeting rat hotspots

Current and future work

- Incorporate uncertainty in human infection status
- Extend model to more campaigns worth of data when available
  - Add campaign as a random effect
  - Add temporal correlations
- Formal model selection

CHICAS
combining health information, computation and statistics

Lancaster University

Yale SCHOOL OF PUBLIC HEALTH

Ministério da Saúde
FIOCRUZ
Fundação Oswaldo Cruz
Instituto Gonçalo Moniz

agresearch
āta mātai, mātai whetū

poppy.miller@agresearch.co.nz

Significant covariate effects: leptospirosis model

Parameter	Median	(CI)	Prob <0	Prob >0
Dist public dump (km)	-14.31	(-21.69, -7.51)		**1.000**
Total rainfall (m)	1.24	(0.85, 1.66)	**1.000**	
Age	0.18	(0.12, 0.23)	**1.000**	
Age >30 years	-0.18	(-0.25, -0.11)		**1.000**
Sex (male = 1)	1.33	(0.78, 1.90)	**1.000**	
Log income	0.25	(0.02, 0.49)	**0.982**	
Log income >40 reais/month	-0.73	(-1.24,-0.20)		**0.997**
Rat linear predictor	0.18	(-0.02, 0.38)	**0.966**	
σ (individual level random effect)	1.72	(1.13, 2.31)		

Odds ratios: leptospirosis model continuous variables

Interpretation: The odds of being infected with leptospirosis for a person with a covariate value at the upper quartile for that covariate are $OR_{U/L}$ times those at the lower quartile for that covariate.

	Odds ratio$_{U/L}$	**Data LQ**	**Data UQ**
Ground cover % soil	1.09 (0.70, 1.56)	3	37
Ground cover % vegetation	1.15 (0.95, 1.38)	0	17
Cumulative rainfall (m)	4.12 (2.45, 6.33)	0.56	1.70
Distance public dump (m)	0.44 (0.27, 0.63)	32.7	90.6
Distance open sewer (m)	1.12 (0.97, 1.29)	6.2	16.9
Age (years)	13.28 (5.38, 27.01)	15	42
Log income (reias/month)	0.64 (0.30, 1.09)	log(1)	log(728)
Rat linear predictor	1.03 (1.00, 1.07)	0.033	0.214

Odds ratios: leptospirosis model binary variables

	Odds ratio$_{1/0}$
Male / Female	3.78 (1.96, 6.33)
Ethnicity 2 / Ethnicity 1	1.57 (0.28, 4.62)
Ethnicity 3 / Ethnicity 1	1.64 (0.34, 4.87)
Ethnicity 4 and 7 / Ethnicity 1	3.77 (0.00, 54.66)
Literate / Illiterate	0.89 (0.45, 1.48)
Sewer exposed / not exposed	1.54 (0.81, 2.58)
Mud exposed / not exposed	1.24 (0.64, 2.07)
Flood exposed / not exposed	1.00 (0.51, 1.63)

Covariate effects: rat model

Parameter	Median (CI)		Prob <0	Prob >
Intercept	-3.06	(-3.52, -2.61)		**1.000**
Area soil	0.87	(-0.15, 2.00)	*0.943*	
Area soil squared	-3.76	(-6.89, -1.04)		**0.994**
Area veg 5m	-0.52	(-1.28, 0.17)		*0.934*
Area veg 5m squared	2.42	(0.04, 4.50)	**0.984**	
Mean rainfall	58.32	(38.14, 80.39)	**1.000**	
Dist dump	-16.11	(-26.50, -5.45)		**0.999**
Dist dump >70m	20.52	(6.14, 36.22)	**0.998**	
Domestic	0.13	(-0.17, 0.39)	0.795	
Dist open sewer	-19.87	(-38.14, -3.29)		**0.983**
Dist open sewer >40m	51.69	(12.29, 84.97)	**0.997**	
phi	17.31	(12.85, 22.86)		
sigmasq	2.03	(1.40, 2.71)		
tausq	0.39	(0.08, 0.85)		

Why the cloglog link?

Boards are marked at rate λ_i (Poisson process):

- p_i: probability that the number of rat marks is ≥ 1 in time period $[0, T_i]$
- $1 - p_i$: Prob (0 marks) in time period $[0, T_i]$

$$1 - p_i = \frac{\left\{\int_0^{T_i} \lambda_i dt\right\}^0 e^{-\int_0^{T_i} \lambda_i dt}}{0!} = e^{-\int_0^{T_i} \lambda_i dt} = e^{-\lambda_i T_i}$$

Rearrange to get cloglog link function:

$$\begin{aligned} \log\left(-\log\left(1 - p_i\right)\right) &= \log\left(\lambda_i T_i\right) \\ &= \log\lambda_i + \log T_i \\ &= d\left(x_i\right)' \beta + S\left(i\right) + \log T_i \\ &= \eta_i \end{aligned}$$