

1 Linking rattiness, geography and environmental degradation to
2 spillover *Leptospira* infections in marginalised urban settings: an
3 eco-epidemiological community-based cohort study in Brazil

4 Max T. Eyre^{1,2}, Fábio N. Souza³, Ticiana S. A. Carvalho-Pereira³, Nivison Nery Jr.³,
5 Daiana de Oliveira³, Jaqueline S Cruz³, Gielson A. Sacramento³, Hussein Khalil^{3,4}, Elsie A.
6 Wunder Jr^{5,6}, Kathryn P. Hacker⁷, James E. Childs^{5,6}, Mitermayer G. Reis^{3,5}, Mike Begon⁸,
7 Peter J. Diggle¹, Albert I. Ko^{5,6}, Emanuele Giorgi^{1,*}, and Federico Costa^{1,3,5,6,*}

8 ¹Centre for Health Informatics, Computing, and Statistics, Lancaster University Medical School, UK

9 ²Liverpool School of Tropical Medicine, Liverpool, UK

10 ³Institute of Collective Health, Federal University of Bahia, Salvador, Bahia, Brazil

11 ⁴Swedish University of Agricultural Sciences, Umeå, Sweden

12 ⁵Oswaldo Cruz Foundation, Brazilian Ministry of Health, Salvador, Bahia, Brazil

13 ⁶Department of Epidemiology of Microbial Diseases, Yale School of Public Health, New Haven, USA

14 ⁷University of Pennsylvania, Philadelphia, Pennsylvania, USA

15 ⁸Department of Evolution, Ecology and Behaviour, University of Liverpool, Liverpool, UK

16 *These authors contributed equally.

17 September 19, 2021

18 **Abstract**

19 **Background:** Zoonotic spillover from animal reservoirs is responsible for a significant global public
20 health burden, but the processes that promote spillover events are poorly understood in complex urban
21 settings. Endemic transmission of *Leptospira*, the agent of leptospirosis, in marginalised urban communities
22 occurs through human exposure to an environment contaminated by bacteria shed in the urine of the rat
23 reservoir. However, it is unclear to what extent transmission is driven by variation in the distribution of
24 rats or by the dispersal of bacteria in rainwater runoff and overflow from open sewer systems.

25 **Methods:** We conducted an eco-epidemiological study in a high-risk community in Salvador, Brazil, by
26 prospectively following a cohort of 1,401 residents to ascertain serological evidence for leptospiral infections.
27 A concurrent rat ecology study was used to collect information on the fine-scale spatial distribution of
28 ‘rattiness’, our proxy for rat abundance and exposure of interest. We developed and applied a novel
29 geostatistical framework for joint spatial modelling of multiple indices of disease reservoir abundance and
30 human infection risk.

31 **Results:** The estimated infection rate was 51.4 (95%CI 40.4, 64.2) infections per 1,000 follow-up
32 events. Infection risk increased with age until 30 years of age and was associated with male gender.
33 Rattiness was positively associated with infection risk for residents across the entire study area, but this

34 effect was stronger in higher elevation areas (OR 3.27 95%CI 1.68, 19.07) than in lower elevation areas
35 (OR 1.14 95%CI 1.05, 1.53).

36 **Conclusions:** These findings suggest that, while frequent flooding events may disperse bacteria in
37 regions of low elevation, environmental risk in higher elevation areas is more localised and directly driven by
38 the distribution of local rat populations. The modelling framework developed may have broad applications
39 in delineating complex animal-environment-human interactions during zoonotic spillover and identifying
40 opportunities for public health intervention.

41 **Funding:** This work was supported by the Oswaldo Cruz Foundation and Secretariat of Health
42 Surveillance, Brazilian Ministry of Health, the National Institutes of Health of the United States (grant
43 numbers F31 AI114245, R01 AI052473, U01 AI088752, R01 TW009504 and R25 TW009338); the Well-
44 come Trust (102330/Z/13/Z), and by the Fundação de Amparo à Pesquisa do Estado da Bahia (FAPES-
45 B/JCB0020/2016). M.T.E is supported by an MRC doctorate studentship. F.N.S. participated in this
46 study under a FAPESB doctorate scholarship.

47
48 **Keywords:** zoonotic spillover; leptospirosis epidemiology; transmission pathways; rat reservoir; rodent-
49 borne zoonoses; multivariate model-based geostatistics; joint spatial modelling.

50 1 Introduction

51 Zoonotic spillover, the transmission of pathogens from infected vertebrate animals to humans, is responsible for
52 a significant public health burden globally. Understanding the processes that promote spillover transmission
53 is essential for improving our ability to predict and prevent spillover events, but for many zoonoses, such
54 as *Leptospira interrogans*, *Escherichia coli* O157 and *Giardia* spp., they are poorly understood [1]. This is
55 due to the complex nature of the spillover system, in which the probability of transmission is governed by
56 dynamic interactions in space and time between ecological, epidemiological, behavioural and immunological
57 factors that determine pathogen pressure, exposure and host susceptibility. Zoonotic spillover research must
58 explore interactions between the environment, disease reservoirs and local epidemiology, presenting two central
59 challenges: i) the need for transdisciplinary studies at the animal-human disease interface (a One Health
60 approach) that accurately collect data on multiple components of the spillover process at common temporal
61 and spatial scales at which these events take place; ii) the development of integrative approaches to jointly
62 analysing these diverse datasets within a spatially and temporally explicit framework [1–3].

63 Leptospirosis, a neglected zoonotic disease caused by pathogenic bacteria from the genus *Leptospira*, is an
64 important example of zoonotic spillover. Globally, it is estimated to cause more than one million cases and

65 over 58,000 deaths each year [4], with an annual global burden of 2.9 million DALYs [5]. This burden falls
66 heavily on marginalised urban populations in low- and middle-income countries who live in areas characterised
67 by high population density, poor quality housing and inadequate provision of healthcare, sanitation, and
68 waste management services. In these settings, leptospiral infection occurs through contact with water or soil
69 contaminated with leptospires shed in the urine of the principal reservoir, the Norway rat (*Rattus norvegicus*)
70 [6]. These areas produce the socio-ecological conditions that allow rodent populations to proliferate and
71 leptospires to persist for long periods in the environment [7]. Residents consequently have frequent, intense
72 and largely unavoidable exposure to the contaminated environment, often exacerbated by their geographical
73 vulnerability to flooding events [8]. In response, the World Health Organisation (WHO) has convened the
74 Leptospirosis Burden Epidemiology Reference Group (LERG) which has recommended “Targeted intervention
75 based on the improved knowledge of disease ecology” [9], highlighting the current knowledge gap for *Leptospira*
76 transmission mechanisms and target points for effective intervention.

77 Multiple studies have helped to elucidate key aspects of the *Leptospira* transmission cycle in urban set-
78 tings, identifying socioeconomic vulnerability, household environment and behavioural exposures as important
79 determinants of infection risk [7, 10–19]. However, these variables have been unable to explain fine-scale spa-
80 tial variation in risk [10, 12]. This is likely to be driven by the high spatial and temporal heterogeneity in
81 environmental risk, observed in recent studies of *Leptospira* in soil, and surface and sewage waters [6, 20, 21].
82 These findings lead to two key questions: i) to what extent does environmental contamination by localised
83 rat shedding drive infection risk, rather than exposure to leptospires that have been dispersed by rainwater
84 runoff and overflowing sewer systems; and ii) how does this change across the geography of a community, for
85 example at different elevation levels?

86 Establishing a dynamic link between rats, the environment and *Leptospira* transmission is complicated
87 by the difficulty of measuring and modelling the rat contamination process. However, urban Norway rats
88 have been found to have high *Leptospira* prevalence and shedding rates worldwide [22–28]. This suggests
89 that rat abundance may be predictive of environmental risk, and could be used as a proxy for this shedding
90 process. While several studies have identified associations between infection risk and household rat sightings
91 and infestation [10, 12, 22, 29–31], their ability to explore fine-scale spatial variation in risk was limited by a
92 reliance on household infestation surveys or aggregation of incidence and abundance indices to a common coarse
93 spatial scale. All modelled abundance as a regression covariate, thereby not accounting for uncertainty in its
94 measurement. The absence of methods to formally integrate abundance into analyses of spillover mechanisms
95 is an issue for rodent-borne zoonoses more widely [3, 32].

96 There is no gold-standard index of abundance and field teams use a range of imperfect indices, such as
97 traps, infestation surveys and track plates. In our previous work, we developed a multivariate generalized
98 linear geostatistical model for joint spatial modelling of multiple imperfect abundance indices [33]. We use the
99 term ‘abundance’ loosely here to denote all ecological processes that are associated with animal abundance,
100 for example animal presence and activity, and that can be used to quantify exposure to a disease of interest.
101 This methodology was then used to model the spatial distribution of ‘rattiness’, our proxy for rat abundance,
102 at a fine scale within a community in Salvador, Brazil [33]. The spatial distribution of rattiness was highly
103 heterogeneous, suggesting that it could be a driver of micro-heterogeneity in infection risk.

104 To analyse reservoir host abundance and infection data at fine spatial scales, we propose that a framework
105 should i) account for spatial correlation in human and reservoir host data; ii) jointly model multiple imperfect
106 indices of abundance while accounting for the appropriate sampling distribution of each index; iii) account for
107 uncertainty in abundance indices, iv) allow for the prediction of abundance and infection risk at all locations
108 within the study area, and v) quantify the uncertainty associated with those predictions. Several studies have
109 attempted to model spatial associations between disease reservoir or vector abundance and human infection
110 for leptospirosis [34–36], tularemia [37]; Lyme disease [38]; West Nile Virus [39]; dengue fever [40]; and Lassa
111 fever [41]. However, none of the approaches used satisfy all five of the above conditions. The development of
112 new tools for the joint spatial analysis of abundance and human infection may consequently be beneficial for
113 the study of other zoonoses and vector-borne diseases [42].

114 The aim of this study was to develop a flexible modelling framework for zoonotic spillover to explore
115 whether rattiness, acting as a proxy for local leptospiral contamination by Norway rats, can explain spatial
116 heterogeneity in leptospiral transmission in a high-risk urban community in Brazil. We extend the rattiness
117 framework of Eyre et al. [33] to include human infection risk. We describe findings from a transdisciplinary
118 eco-epidemiological study which comprises a prospective community-based cohort study with two serosurveys
119 and a fine-scale rat ecology study. The ecology study was used to collect information on the spatial distribution
120 of rat abundance, our exposure of interest, in the period between the two surveys using multiple abundance
121 indices. Then, we explore associations between infection risk, rattiness and a range of measured environmental
122 and individual risk factors.

123 2 Materials & methods

124 2.1 Study design

125 2.1.1 Study area

126 The study was conducted in Pau da Lima community (13°32'53.47" S; 38°43'51.10" W), a marginalised
127 informal settlement located in the city of Salvador, Northeast Brazil. The study site has an area of 0.25km²
128 and is characterised by three connected valleys with large elevation gradients, high population density and a
129 heterogeneous environment of vegetation, paved surfaces and exposed soil (Figure 1). There are significant
130 gradients in socioeconomic status and infrastructure quality over small elevation increases - with the most
131 marginalised members of the community living at lower elevations. The community suffers from low quality
132 housing, poor provision of waste management services and inadequate drainage and sanitation systems [12,
133 43]. Residents are consequently often unable to avoid intense exposure with mud and floodwater. These factors
134 result in abundant rat populations [33] and a high estimated annual *Leptospira* infection rate of 35.4 (95%
135 CI, 30.7, 40.6) infections per 1,000 annual follow-up events [12]. For this reason, Pau da Lima has become an
136 exemplar for investigating urban *Leptospira* transmission in Brazil over the last 15 years.

137 2.1.2 Serosurveys

138 We conducted a prospective community cohort study with two serosurveys carried out in August-October
139 2014 and January-April 2015. After an initial census of the study site, all ground floor households were visited
140 and inhabitants who met the eligibility criteria of ≥ 5 years of age who had slept ≥ 3 nights in the previous
141 week in a study household were invited to join the study.

142 During each survey trained phlebotomists collected blood samples from participants and administered a
143 modified version of the standardised questionnaire used previously [12, 29]. Information was collected on
144 demographic and socioeconomic indicators, household environmental characteristics and exposures to poten-
145 tial sources of environmental contamination in the previous six months (the average time between the two
146 serosurveys). Study data were collected and managed using REDCap electronic data capture tools [44] and
147 all individual data were anonymised. The locations of sampled households are shown in Figure 1A. If an indi-
148 vidual was not found during a sample collection visit their house was revisited at least five times on different
149 days of the week.

150 The microscopic agglutination test (MAT) was used to determine titers of agglutinating antibodies against

It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

151 pathogenic *Leptospira* in sera obtained from the blood samples collected in each serosurvey. Serological samples
152 were reacted with a panel of two *Leptospira* reference strains that are dominant in Pau da Lima: *Leptospira*
153 *interrogans* serovars Copenhageni (COPL1) and Cynopteri 3522C (C3522C). These two strains have been
154 shown to have the same performance in identifying MAT seroconversion in our prospective studies as the
155 WHO recommended battery of 19 reference serovars. When agglutination was observed at a dilution of 1:50,
156 the sample was titrated in serial two-fold dilutions to determine the highest agglutination titer. The study
157 outcome of leptospiral infection was defined as seroconversion, an MAT titer increase from negative to $\geq 1:50$,
158 or a four-fold increase in titer for either serovar between paired samples from cohort subjects. All laboratory
159 analyses were performed in the Laboratory Pathology and Molecular Biology at Fiocruz, Salvador. As part
160 of quality control procedures two independent evaluations were conducted by Yale University for all infected
161 subjects and 8% of all samples, with high concordance between results.

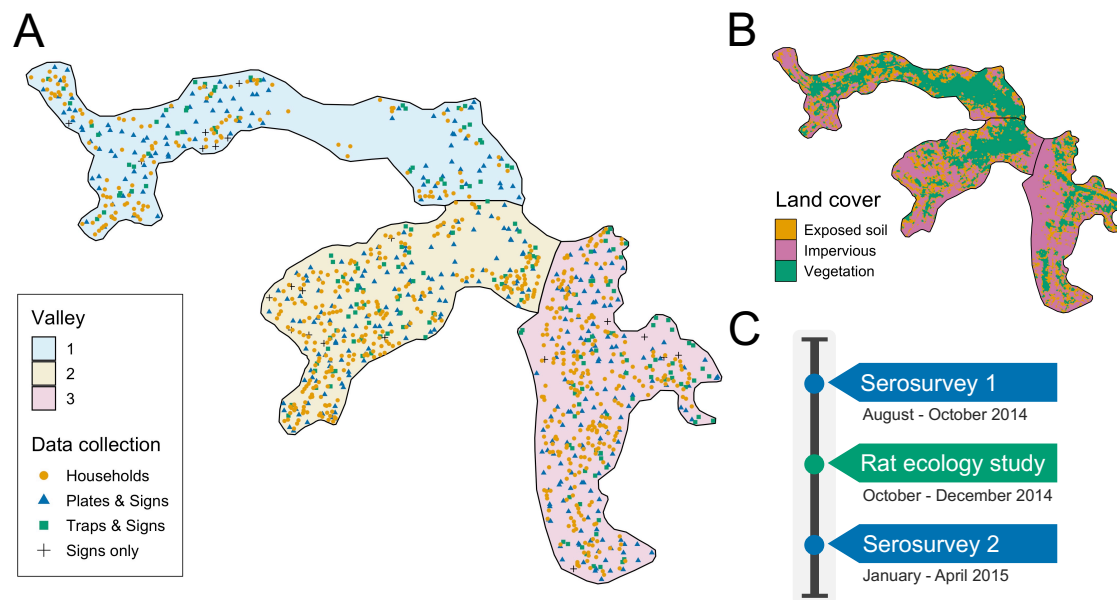


Figure 1: A) Map of the three valleys within the study site in Pau da Lima, with household locations for the serosurveys marked as orange circles. Locations sampled in the the rat ecology study are shown for each of the rat abundance indices as follows: Plates & Signs (track plates, burrows, faeces and trails), Traps & Signs (traps, burrows, faeces and trails) and Signs only (burrows, faeces and trails); B) Land cover classification map (impervious cover is defined as man-made structures e.g. pavement and buildings); C) Study timeline for the two community serosurveys and rat ecology study.

162 **2.1.3 Rat ecology study**

163 To estimate exposure risk due to local rat contamination between the two serosurveys, a cross-sectional rat
164 ecology study was conducted from October to December 2014. As has been described previously [33], the aim
165 of this study was to collect data on the fine-scale spatial variation in rat reservoir population abundance. Data
166 were collected for five indices of rat abundance: live trapping, track plates, number of active burrows present,
167 presence of faecal droppings and presence of trails. Rat trapping was carried out at 189 locations, randomly
168 distributed across the study area (see Panti-May et al. [45]). Two traps were deployed for 4 consecutive 24-hour
169 trapping periods at each location. Trapping success and trap closure without a rat, a common malfunction,
170 were recorded after each 24-hour period. Track plates were placed at 415 locations for two consecutive 24-hour
171 periods following the standardised protocol for placement and survey developed and validated previously [46],
172 with five plates placed at each location in the shape of a ‘five’ on a die. After each 24-hour period plates were
173 repainted and any lost plates were recorded and replaced. On the first day of trapping or plate placement,
174 a survey for signs of rat infestation, adapted from the Centers for Disease Control and Prevention [47] and
175 validated in the study area [29], was conducted within an area of 10m radius around each trapping or plate
176 location to record the number of active burrows and the presence of faecal droppings and trails. In total,
177 595 independent locations were sampled for traps, track plates and the three survey indices for signs of rat
178 infestation. The spatial distribution of these locations is shown in Figure 1A. At 21 locations, theft and local
179 gang violence meant that data for track plates and traps was not collected and only the three survey indices
180 for signs of rat infestation were used.

181 **2.1.4 Environmental data**

182 In addition to the environmental survey conducted at each household location, we also collected information
183 for three spatially continuous environmental variables: elevation relative to the bottom of each valley, distance
184 to large public refuse piles and the proportion of land cover classified as impervious (man-made structures)
185 within a 30-metre radius. The land cover variable was created from Digital Globe’s WorldView-2 satellite
186 imagery (8 bands) taken on February 17, 2013 which was classified using a maximum likelihood supervised
187 algorithm and validated with ground truthed data collected from 20 randomly selected sites of size 5m by 5m.
188 The classification map is shown in Figure 1B.

189 2.1.5 Ethics

190 Participants were enrolled according to written informed consent procedures approved by the Institutional
191 Review Boards of the Oswaldo Cruz Foundation and Brazilian National Commission for Ethics in Research,
192 Brazilian Ministry of Health (CAAE: 01877912.8.0000.0040) and Yale University School of Public Health (HIC
193 1006006956).

194 For the rat ecology study, the ethics committee for the use of animals from the Oswaldo Cruz Foundation,
195 Salvador, Brazil, approved the protocols used (protocol number 003/2012), which adhered to the guidelines
196 of the American Society of Mammalogists for the use of wild mammals in research [48] and the guidelines
197 of the American Veterinary Medical Association for the euthanasia of animals [49]. These protocols were
198 also approved by the Yale University’s Institutional Animal Care and Use Committee (IACUC), New Haven,
199 Connecticut (protocol number 2012–11498).

200 2.2 Joint modelling rat abundance and human infection: the rattiness-infection 201 framework

202 The developed geostatistical modelling framework jointly models multiple rat abundance indices as measure-
203 ments of a common latent process, called rattiness. Rattiness at each household location contributes to the
204 risk of infection for all inhabitants, in addition to other measured individual or household-level explanatory
205 variables.

206 We model the rat abundance data following a similar structure to that previously outlined [33]. Let
207 $R(x)$ denote a spatially continuous stochastic process, representing rattiness. The rat data then consist of
208 a set of outcomes $Y_i = (Y_{i,k} : k = 1, \dots, 5)$, for $i = 1, \dots, N_r$, collected at a discrete set of locations
209 $X = \{x_i : i = 1, \dots, N_r\}$. The outcome variables $Y_k : k = 1, \dots, 5$ are the set of five rat abundance indices that
210 provide information about $R(x)$: traps ($k = 1$), track plates ($k = 2$), number of burrows ($k = 3$), presence of
211 faecal droppings ($k = 4$) and presence of trails ($k = 5$).

212 Human data are collected from N_h households and consist of an infection outcome $Z_{i,j}$ for individual j
213 at household location i , for $i = N_r + 1, \dots, N_r + N_h$, collected at a discrete set of locations $X = \{x_i : i =$
214 $N_r + 1, \dots, N_r + N_h\}$.

215 Let “[.]” be a shorthand notation for “the probability distribution of .” We write $Y = (Y_1, \dots, Y_{N_r})$,
216 $Z = (Z_{N_r+1}, \dots, Z_{N_r+N_h})$ and $R = (R(x_1), \dots, R(x_{N_r+N_h}))$. We assume that the $Y_{i,k} : k = 1, \dots, 5$ and $Z_{i,j}$

217 are conditionally independent given $R(x_i)$, from which it follows that

$$[Y, Z|R] = \prod_{i=1}^{N_r} \prod_{k=1}^5 [Y_{i,k}|R(x_i)] \prod_{i=N_r+1}^{N_r+N_h} \prod_{j=1}^{J_i} [Z_{i,j}|R(x_i)]. \quad (1)$$

218 where $[\cdot]$ is a shorthand notation for “the distribution of” and J_i denotes the number of individuals at household

219 i . This model structure is shown schematically in Figure 2.

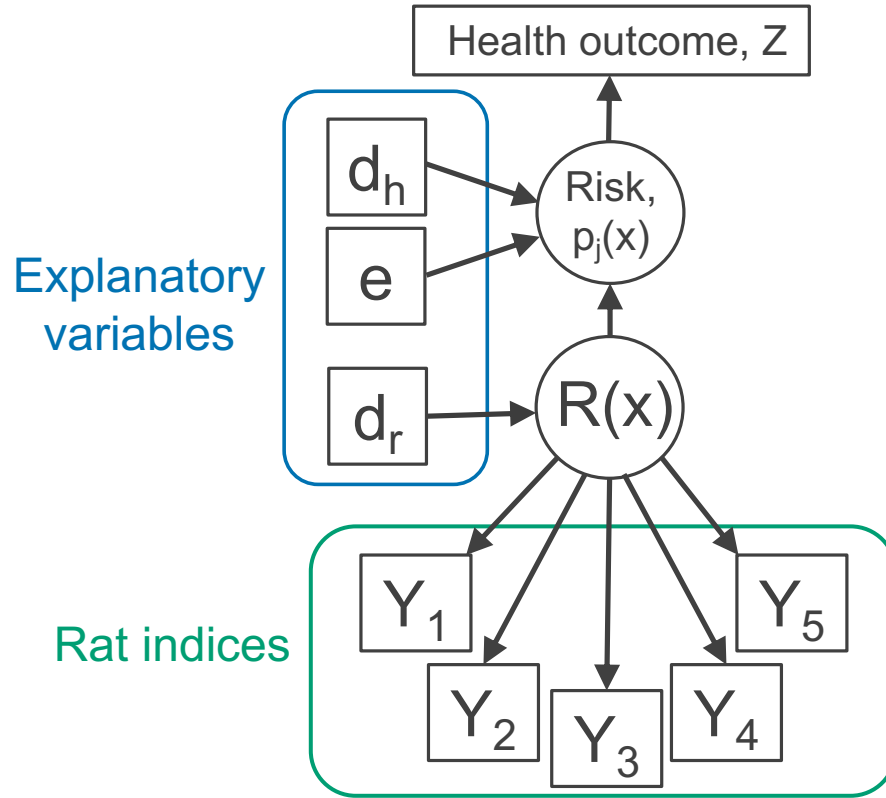


Figure 2: Directed acyclic graph (DAG) of the rattiness-infection model framework. $R(x)$ is the value of a spatially continuous stochastic rattiness process at location x . The outcome variables $Y_k : k = 1, \dots, 5$ are the set of five rat abundance indices that provide information about $R(x)$: traps ($k = 1$), track plates ($k = 2$), number of burrows ($k = 3$), presence of faecal droppings ($k = 4$) and presence of trails ($k = 5$). The outcome variable $Z_{i,j}$ is the observed health outcome, in this case this represents infection status. The terms d_h and d_r represent the sets of spatially continuous explanatory variables which contribute to spatial variation in infection risk in humans and $R(x)$, respectively. The terms d_h and d_r are not mutually exclusive groups of explanatory and the same variables may contribute to both infection risk and $R(x)$. The term e represents a set of individual- and household-level explanatory variables which contribute to variation in infection risk. Square objects correspond to observable variables, and circles to latent random variables.

220 **2.2.1 Rattiness**

We define rattiness at location x as

$$R(x_i) = d_r^T(x_i)\beta_r + \sqrt{\psi} S(x_i) + \sqrt{1-\psi} U_i. \quad (2)$$

221 The terms on the right-hand side of 2 have the following interpretations: $d_r(x_i)$ is a vector of explanatory
222 variables with associated regression coefficients β_r ; U_i is a set of independently and identically distributed
223 zero-mean Gaussian variables with unit variance; $S(x_i)$ is a stationary and isotropic spatial Gaussian process;
224 $\psi \in (0, 1)$ regulates the relative contributions of spatially structured variation, $S(x_i)$, and unstructured random
225 variation, U_i , to $R(x_i)$.

For the Gaussian process, $S(x_i)$, we specify an exponential spatial correlation function:

$$\text{Corr}(S(x), S(x')) = e^{-u/\phi}$$

226 where $u = \|x - x'\|$ is the Euclidean distance between x and x' , and ϕ regulates how fast the spatial correlation
227 decays to zero with increasing distance u .

228

229 **2.2.2 Rat abundance outcomes**

The variable $Y_{i,1}$, conditionally on $R(x_i)$, is a binomial variable representing the number of traps, out of $n_{i,1}$,
in which rats were captured. We assume that the times of rat captures from a trap follow a time-varying
inhomogeneous Poisson process with intensity $t_i\mu_2(x_i)$, where t_i is the time (in days) for which a trap is
operative and $\log\{\mu_2(x_i)\} = \alpha_1 + \sigma_1 R(x_i)$. It follows that the probability of capturing a rat is

$$1 - \exp\{-t_i\mu_2(x_i)\}.$$

230 If a trap is found closed without a rat, we assume that the trap was disturbed and set $t = 0.5$. In all other
231 cases, $t = 1$ day.

232 $Y_{i,2}$, is the number of track-plates, out of $n_{i,2}$, that show presence of rats. We model this as a binomial
233 variable with $n_{i,2}$ trials and probability $\mu_2(x_i)$ where $\log\{\mu_2(x_i)/(1 - \mu_2(x_i))\} = \alpha_2 + \sigma_2 R(x_i)$.

234 $Y_{i,3}$, is the number of active rat burrows found at location x_i . We model this as a Poisson variable with

235 rate $\mu_3(x_i)$ where $\log\{\mu_3(x_i)\} = \alpha_3 + \sigma_3 R(x_i)$.

236 The variables $Y_{i,4}$ and $Y_{i,5}$ are binary indicators taking value 1, if at least one faecal dropping or trail,
237 respectively, was found at location x_i and 0 otherwise. We model the probability of finding a sign of faecal
238 droppings or trails, $\mu_4(x_i)$ and $\mu_5(x_i)$, using logit-linear regressions $\log\{\mu_4(x_i)/(1 - \mu_4(x_i))\} = \alpha_4 + \sigma_4 R(x_i)$
239 and $\log\{\mu_5(x_i)/(1 - \mu_5(x_i))\} = \alpha_5 + \sigma_5 R(x_i)$.

240 2.2.3 Human infection outcome

Conditionally on $R(x_i)$, we model the binary human infection outcome $Z_{i,j}$ as a Bernoulli variable with the probability, $p_j(x_i)$, that individual j at location i is infected. This is modelled with a logit link function and the following linear predictor

$$\log \left\{ \frac{p_j(x_i)}{1 - p_j(x_i)} \right\} = \alpha_h + d_h^T(x_i)\beta_h + e_{i,j}^T\gamma + \xi(x_i)R(x_i) + V_i \quad (3)$$

241 where: $d_h(x_i)$ is a vector of spatially continuous explanatory variables with associated regression coefficients
242 β_h ; $e_{i,j}$ is a vector of household-level and individual-level explanatory variables with associated regression
243 coefficients γ ; V_i is a set of independently and identically distributed zero-mean Gaussian variables with
244 variance σ^2 representing unexplained household-level variation; $\xi(x_i)$ regulates the contribution of rattiness
245 to risk of infection.

246 2.2.4 Parameterising to test for an interaction with elevation

247 To explore variation in the role of local rat populations in transmission within different sections of the study
248 area, ξ was parameterised to test for an interaction between rattiness and household elevation level on human
249 infection risk. This was implemented by dividing the study area into three elevation levels with an equal number
250 of households in each: low (0 – 6.7m from bottom of valley), medium (6.7 – 15.6m) and high (> 15.6m). We
251 then define the set of household locations in each low, medium and high elevation level as x_{low} , x_{med} and
252 x_{high} , respectively. Three values of ξ were then estimated such that:

$$\xi(x_i) = \begin{cases} \xi_{low} & \text{at locations } x_i \in x_{low} \\ \xi_{med} & \text{at locations } x_i \in x_{med} \\ \xi_{high} & \text{at locations } x_i \in x_{high} \end{cases} \quad (4)$$

253 **2.3 Variable selection**

254 **2.3.1 Predictors of rattiness**

255 The exploratory analysis for the rattiness model followed the steps developed and described previously [33].
256 Firstly, we explored the functional form of the relationship between rattiness and three continuous explanatory
257 variables: relative elevation, distance to large refuse piles and land cover type. To do this, we fitted a simplified
258 rattiness model that did not include covariates or account for spatial correlation. Rattiness is consequently
259 modelled purely as unstructured random variation; hence $R(x_i) = U_i$ [33]. We then computed the predictive
260 expectation of this simplified rattiness process, \hat{U}_i , at all locations for which rat index measurements were
261 observed. A generalized additive model (GAM) [50] was then fitted to the \hat{U}_i with the three explanatory
262 variables and the shape of each fitted smooth function was used to assess whether the relationship between
263 each variable and rattiness was linear. Non-linear relationships were modelled using linear splines based on
264 the identified functional form, with knots placed at relative elevations of 8m and 22m, and at a distance from
265 large refuse piles of 50m (see Supplementary file 1). For variable selection, linear models with all combinations
266 of these variables were fitted and ranked by their Akaike Information Criterion (AIC) value [51]. The model
267 with the lowest AIC included all of the variables and their linear splines (Supplementary file 2).

268 Following the methodology outlined previously [33], we fitted the full geostatistical rattiness model using the
269 variables selected in Section 2.3.1. We then plugged in the maximum likelihood estimates and made predictions
270 for rattiness at all human household locations; here, the predictive target is $T(x) = d_r(x)^\top \beta_r + \sqrt{\psi} S(x)$ rather
271 than $R(x)$ as defined by (2) because the predicted value of the spatially uncorrelated $U(x)$ at any location x
272 where rat abundance indices have not been recorded is zero. The expectation of this predictive distribution
273 was then computed to provide an estimate of mean predicted rattiness at all household locations. This was
274 then used as an exploratory covariate in the following section.

275 **2.3.2 Risk factors for human infection**

276 All explanatory variables were grouped into the following four domains: social status, household environment,
277 occupational exposures and behavioural exposures. A group of *a priori* confounding variables was then
278 identified, with age, gender and household per capita income selected based on previous findings [10–12], and
279 valley also included to account for otherwise unmeasured differences between the three valley regions within
280 the study area.

281 The relationship between continuous explanatory variables and infection risk (on the log-odds scale) was

282 assessed for linearity by fitting a GAM while controlling for the four confounders. As before, non-linear
283 relationships were modelled using linear splines based on the identified functional form. Age was modelled with
284 a knot at 30 years old, education at 5 years and relative elevation at 20m (Supplementary file 1). A univariable
285 analysis was conducted to explore the relationship between each explanatory variable and infection risk while
286 controlling for the four *a priori* confounding variables. Crude and adjusted odds ratios were estimated using
287 a mixed effects logistic regression with a random effect to account for unexplained variation at the household-
288 level.

289 For the multivariable model, variable selection was conducted within each domain separately. Mixed-effect
290 logistic regression models were fitted for all combinations of the variables in each domain and were ranked
291 by their Akaike Information Criterion (AIC) value (Supplementary file 2). Variables in the model with the
292 minimum AIC value were selected for each domain. Age, gender, household per capita income and valley
293 were controlled for in all models throughout this process. Then, the variables selected from each domain were
294 combined and the mean predicted rattiness estimate (obtained in Section 2.3.1 at each household location)
295 was included with an interaction with elevation level. This set of variables was reduced once more following
296 the same process and all selected variables were included in the final multivariable model (Supplementary file
297 3).

298 **2.4 Model fitting**

299 All rat and human variables selected in Section 2.3 were then included in the full joint model defined in
300 Equations (2) and (3). We fit this model using the Monte Carlo maximum likelihood (MCML) method [52]
301 as described in Supplementary file 4, and compute 95% confidence intervals by re-fitting the model for 1,000
302 parametric bootstraps.

303 **2.5 Prediction maps**

304 The maximum likelihood parameter estimates were then used to make prediction maps for rattiness and
305 infection risk as follows.

306 To map a general predictive target, $T(x)$ say, we first define $T^* = (T(x_1^*), \dots, T(x_H^*))$, where $X^* =$
307 $\{x_1^*, \dots, x_H^*\}$ is a finely spaced grid of locations to cover the region of interest. We then draw samples from the
308 predictive distribution of T^* , i.e. its conditional distribution given all relevant data. These samples can then
309 be used to compute any desired summary of the predictive distribution. In our analysis, we used as summaries

310 the expectation and 95% prediction interval.

311 Our first predictive target is rattiness, for which $T(x) = d_r^\top(x)\beta_r + \sqrt{\psi}S(x)$. Our second is human infection
312 risk, for which $T(x) = d_h^\top(x)\beta_h + e^\top\gamma + \xi(x_i)R(x) + V_i$. In either case, we first sample from $[R|W; \theta, \omega]$ using
313 the same sampling algorithm as for maximizing the likelihood in Section 2.2, with the parameters θ and ω
314 fixed at their maximum likelihood estimates. After obtaining samples $r_{(b)}$, $b = 1, \dots, B$, we then sample from
315 $[T^*|r_{(b)}]$, which in both cases follows a multivariate Gaussian distribution with mean and covariance matrix
316 easily obtained from their joint Gaussian distribution, $[R, T^*]$. The resulting values, $t_{(b)}(x_h^*)$, $h = 1, \dots, H$; $b =$
317 $1, \dots, B$, constitute b samples drawn from $[T^*|W]$ as required. Note that each $t_{(b)}(x_h^*)$, $h = 1, \dots, H$ is a sample
318 from the joint predictive distribution of the complete surface of $T(x)$ over the whole of the region of interest
319 and can therefore be used to make inferences about spatially aggregated properties of $T(x)$ if required.

320 2.6 Data accessibility

321 Data and code used in this analysis are publicly available at [https://github.com/maxeyre/Rattiness-inf](https://github.com/maxeyre/Rattiness-infection-framework)
322 [ection-framework](https://github.com/maxeyre/Rattiness-infection-framework) and have been published [53]. However, household coordinates and valley ID have been
323 removed from the human data to ensure participant anonymity. The analysis was conducted using R [54] and
324 the following packages: tidyverse [55], mgcv [56], PreVMap [57], MuMIn [58] and lme4 [59].

325 3 Results

326 3.1 Study overview

327 In Pau da Lima, we identified 3,179 eligible residents using a baseline community census, household visits and
328 through other members of the household. Of these, 2,018 (63.4%) individuals consented to join the study
329 and provided a blood sample in the first serosurvey (August-October 2014). As a result of loss to follow-up,
330 only 1,401 (69.4%) of these participants (from 669 households) completed the second serosurvey (January-
331 April 2015). Individuals were lost to follow-up because they could not be found after at least five attempts
332 (44.4%), had moved out of the study area (31.1%) or did not wish to provide a second blood sample (19.8%).
333 An overview of participant recruitment is provided in Figure 3. Individuals lost to follow-up were similar in
334 age to those who remained in the study cohort (mean 29.0 and 28.8 years old, respectively, t-value= -0.37 ,
335 $df = 1288.5$, $p = 0.7$) but were more likely to be male (49.8% male compared to 42.6%, $\chi^2 = 8.5$, $df = 1$,
336 $p < 0.01$). A full description of the study cohort is included in Supplementary file 5.

It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

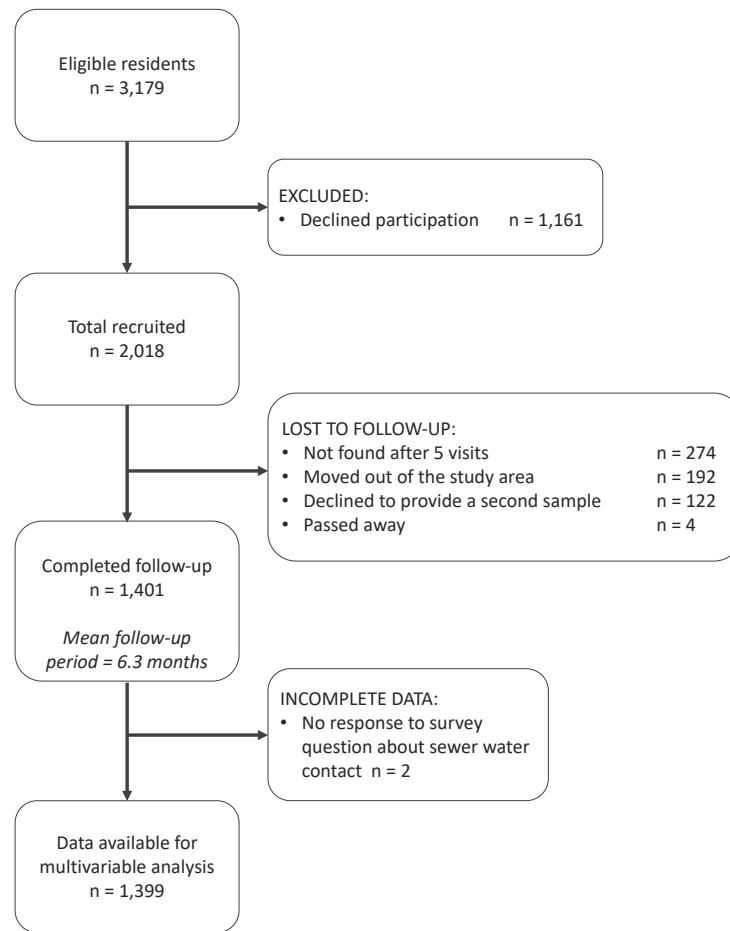


Figure 3: The study participant flow chart in line with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) statement (<http://www.strobestatement.org>)

337 Between the two serosurveys there was serological evidence of 72 leptospiral infections in the cohort, with
338 an overall infection rate of 51.4 (95%CI 40.4, 64.2) infections per 1,000 follow-up events. Valleys 2 and 3
339 had high estimated infection rates with 66.4 (95%CI 47.3, 90.2) infections per 1,000 follow-up events and 49.6
340 (95%CI 33.6, 69.9) infections per 1,000 follow-up events, respectively, compared to 23.2 (95%CI 9.2, 46.9)
341 infections per 1,000 follow-up events in Valley 1. In the rat ecology study: a rat was captured in 129 (9.0%)
342 out of 1,512 trapping-days; 263 (37.4%) out of 703 track plate days had at least one positive plate; 28.5%,
343 19.7% and 25.9% of the 580 sampled locations found at least one sign of active burrows, faecal droppings and
344 trails, respectively.

3.2 Exploratory analysis and model selection

The results from the exploratory multivariable analysis of rattiness are shown in Table 1. The linear splines used were informed by the functional forms shown in Supplementary file 1. The relationship between rattiness and relative elevation demonstrates a trade-off between the high availability of food sources at the bottom of the valley and high risk of flooding which prevents the establishment of burrows. In the lowest elevation areas (0-8m above the bottom of the valley), relative elevation and rattiness were positively associated with an increase of 0.04 (95%CI 0.00, 0.07) rattiness units per 1m; when interpreting the magnitude of effect estimates note that, by definition, rattiness is defined so as to have variance one. Rattiness then peaked at an elevation of 8m before declining with increasing elevation by 0.04 (95%CI -0.09, 0.01) units until an elevation of 22m. Rattiness started to increase again above this elevation by 0.06 (95%CI 0.00, 0.10) units per metre. Rattiness decreased with increasing distance from large refuse piles, a source of food and harbourage, by 0.07 (95%CI -0.13, -0.01) units per 10m distance until a distance of 50m, beyond which there was little effect. Impervious land cover (defined as the proportion of the area within a 30m radius around each sampling location classified as pavement or building) was negatively associated with rattiness, decreasing by -0.05 (95%CI -0.08, -0.01) units for every 10% increase in impervious cover.

Table 1: Multivariable linear regression analysis of predictors for rattiness (note that rattiness is a unit-variance random variable when interpreting the magnitude of effect estimates)

Variable	Estimate (95%CI) ¹
Relative elevation (per 1m increase) ²	
0-8m	0.04 (0.00, 0.07)
8-22m	-0.04 (-0.09, 0.01)
>22m	0.06 (0.00, 0.10)
Distance to large refuse piles (per 10m increase) ²	
0-50m	-0.07 (-0.13, -0.01)
>50m	0.02 (-0.05, 0.09)
Impervious land cover (per 10% increase)	-0.05 (-0.08, -0.01)

¹ CI, Confidence interval;

² The effects of relative elevation and distance to refuse are modelled as broken linear models with transitions at 8m and 22m, and 50m, respectively. This was informed by the relationship described by Generalized Additive Modelling (Supplementary file 1)

Table 2: Univariable mixed effects logistic regression analysis of human risk factors for leptospiral infection

Variable	OR (95%CI) ¹	aOR (95%CI) ¹
Demographic and social status		
Age (per year) ²		
0-30 years old	1.08 (1.03, 1.13)	1.09 (1.04, 1.15)
>30 years old	1.02 (0.96, 1.09)	1.02 (0.95, 1.08)
Male gender	2.22 (1.31, 3.85)	2.78 (1.56, 4.96)
Daily per capita household income (US\$/day)	1.01 (0.89, 1.11)	0.92 (0.80, 1.05)
Valley		
1	REF	REF
2	3.35 (1.33, 10.37)	3.52 (1.23, 10.05)
3	2.39 (0.93, 7.38)	2.53 (0.88, 7.27)
Adult illiteracy	1.34 (0.61, 2.79)	0.66 (0.29, 1.49)
Education (per year of education) ²		
0-5 years	1.05 (0.85, 1.32)	1.14 (0.91, 1.44)
>5 years	0.96 (0.73, 1.27)	0.96 (0.75, 1.26)
Household environment		
Impervious land cover (per 10% increase)	0.87 (0.76, 0.99)	0.82 (0.71, 0.95)
Relative elevation (per 1m increase) ²		
0-20m	0.94 (0.89, 0.99)	0.93 (0.88, 0.99)
>20m	1.12 (0.98, 1.29)	1.12 (0.97, 1.29)
Elevation level		
Low (0-6.7m)	REF	REF
Medium (6.7-15.6m)	0.72 (0.37, 1.39)	0.72 (0.36, 1.44)
High (>15.6m)	0.58 (0.27, 1.20)	0.51 (0.23, 1.11)
Open sewer within 10m	1.60 (0.85, 3.17)	1.69 (0.85, 3.37)
Unprotected from sewer	1.00 (0.55, 1.79)	1.11 (0.61, 2.03)
Live on hillside	0.99 (0.52, 1.86)	0.89 (0.46, 1.71)
Occupational exposures		
Work in construction	1.36 (0.51, 3.21)	0.62 (0.23, 1.67)
Work as travelling salesperson	4.81 (1.12, 18.78)	2.97 (0.71, 12.40)
Work in refuse collection	2.95 (1.04, 7.89)	1.57 (0.56, 4.42)
Work involves contact with floodwater	0.89 (0.04, 5.61)	0.52 (0.05, 4.96)
Work involves contact with sewer water	3.61 (0.45, 20.38)	1.92 (0.29, 12.80)
Behavioural exposures		
Contact with floodwater in last 6 months		
Never/rarely	REF	REF
Sometimes	0.61 (0.27, 1.25)	0.66 (0.30, 1.47)
Frequently	2.14 (0.91, 4.94)	2.84 (1.18, 6.86)
Contact with sewer water in last 6 months		
Never/rarely	REF	REF
Sometimes	0.55 (0.19, 1.31)	0.67 (0.25, 1.78)
Frequently	1.42 (0.51, 3.50)	1.63 (0.61, 4.41)

¹ OR, Odds ratio; aOR, Adjusted odds ratio; CI, Confidence interval; REF, Reference level

² The effect of age, education and relative elevation are modelled as broken linear models with transitions at 30 years old, 5 years of education and an elevation of 20m. This was informed by the relationship described by Generalized Additive Modelling (Supplementary file 1)

It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/) .

362 resident's risk of leptospiral infection (Table 2). Variables with statistically significant associations (at the
363 conventional 5% level) with infection risk were identified in two of the four domains: demographic and social
364 status and behavioural exposures. Within the demographic and social status domain, risk of infection increased
365 with age and was found to be higher for male participants and those living in Valleys 2 and 3. In the behavioural
366 exposures domain, participants who had had frequent contact with floodwater in the last six months were
367 more likely to be infected. Two individuals were excluded from the multivariable analysis (n=1,399) because
368 of missing data for the floodwater exposure survey question.

369 In the exploratory results from the multivariable model there was strong evidence of an interaction between
370 rattiness and household elevation level on human infection risk (see Supplementary file 3 for all parameter
371 estimates for this model). At the high elevation level area, a unit increase in mean predicted rattiness at the
372 household location was estimated to increase the odds of infection by 6.92 (95%CI 1.88, 25.47). In contrast,
373 in the low and medium elevation areas there was no evidence of a relationship between rattiness and infection
374 risk, as shown in Figure 4. Consequently, this interaction effect was also included in the rattiness-infection
375 joint model.

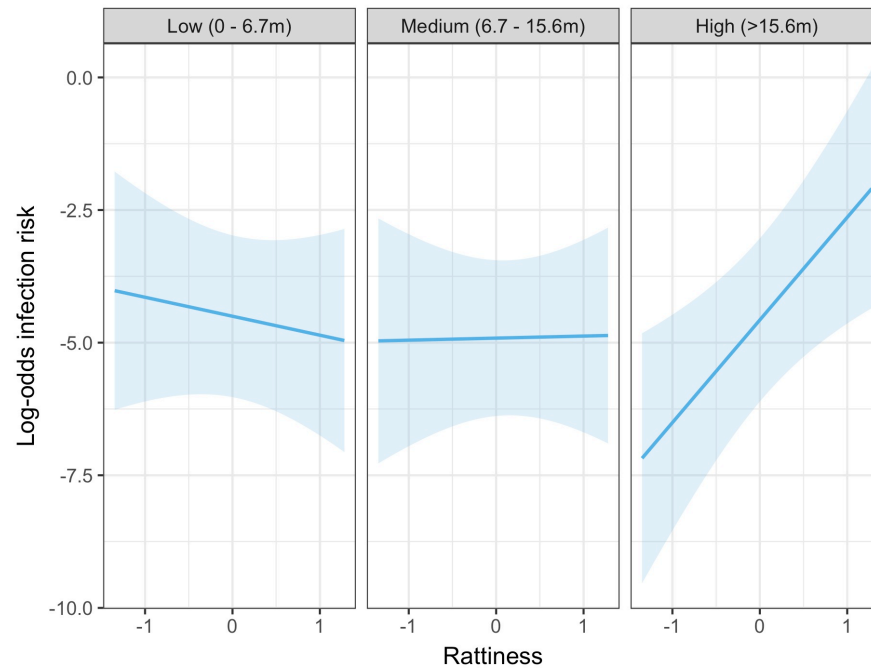


Figure 4: Predicted relationship between rattiness and infection risk from the multivariable mixed effects logistic regression demonstrating evidence of an interaction with elevation level (low, medium and high). Shown on the log-odds scale with shaded areas corresponding to 95% confidence intervals.

376 The explanatory variables selected in the rat and human multivariable analyses were then entered into the

377 full rattiness-infection joint model with the functional forms included in Table 1. To test for residual spatial
378 correlation in the human infection data after controlling for explanatory variables and rattiness, we fitted the
379 joint model with an additional spatial Gaussian process in the human infection linear predictor. The estimated
380 value for the scale of spatial correlation for this Gaussian process was less than 1m and indistinguishable from
381 household-level variation. We consequently fitted the joint model specified in Equation (3) which assumes
382 that there is no residual spatial correlation in the human infection data.

383 3.3 Joint rattiness-infection model

384 Human infection risk factors, rattiness predictors and other model parameters estimated using the joint
385 rattiness-infection model are shown in Table 3. Infection risk was strongly associated with age, with an
386 individual experiencing an increased odds of infection of 1.09 (95%CI 1.04, 1.19) for every year of life up until
387 30 years of age, and 1.02 (95%CI 0.92, 1.09) for each additional year thereafter. Male participants were more
388 likely to be infected than female participants (OR 2.69 95%CI 1.58, 5.89). Compared with individuals living
389 in Valley 1, those living in Valley 2 had a higher estimated odds of infection (OR 2.91 95%CI 1.03, 20.82).
390 Individuals living in the medium (OR 0.77 95%CI 0.31, 1.66) and high (OR 0.67 95%CI 0.11, 1.64) elevation
391 areas had a lower estimated odds of infection relative to those living in the low elevation area where the open
392 sewers are and flooding risk is higher, but these differences were not significant at the conventional 5% level.

393 Infection risk was positively associated with rattiness for households situated in all three elevation levels.
394 However, while the effect size (per unit increase in rattiness) was similar in the low (OR 1.14 95%CI 1.05, 1.53)
395 and medium (OR 1.25 95%CI 1.08, 1.74) elevation areas, in the high elevation area the effect of increasing
396 rattiness on infection risk was significantly stronger (OR 3.27 95%CI 1.68, 19.07). This interaction effect
397 between rattiness and household elevation level on human infection risk was confirmed with a test for evidence
398 against the null hypothesis that $\xi_{low} = \xi_{med} = \xi_{high}$ ($p = 0.026$, $\chi^2 = 7.33$, $df = 2$).

399 Parameter estimates for the rattiness variables were very similar to the estimates from the exploratory
400 linear regression (Table 1), with a slightly higher effect size for the distance to refuse piles and land cover
401 variables. There was evidence of small-scale spatial correlation in rattiness ($\phi = 9.23m$ 95%CI 3.21, 18.24m)
402 corresponding to a spatial correlation range (the distance at which the correlation reduces to 5%) of ap-
403 proximately 28m. The estimate for ψ of about 0.67 (95%CI 0.29, 1.00) indicates that the majority of the
404 unexplained variation in rattiness is spatially structured, with the remainder modelled as a nugget effect.

It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/) .

Table 3: Parameter estimates for the full joint rattiness-infection model

Parameter	Estimate (95%CI)
Human infection risk factors	OR
Age (per year)	
0-30 years old	1.09 (1.04, 1.19)
>30 years old	1.02 (0.92, 1.09)
Male gender	2.69 (1.58, 5.89)
Daily per capita household income (US\$/day)	0.93 (0.74, 1.05)
Valley	
1	REF
2	2.91 (1.03, 20.82)
3	2.28 (0.86, 14.00)
Elevation level	
Low (0-6.7m)	REF
Medium (6.7-15.6m)	0.77 (0.31, 1.66)
High (>15.6m)	0.67 (0.11, 1.64)
Work as travelling salesperson	3.16 (0.38, 20.57)
Contact with floodwater in last 6 months	
Never/rarely	REF
Sometimes	0.62 (0.18, 1.39)
Frequently	2.47 (0.67, 7.41)
<i>Rattiness</i> (per unit rattiness)	
ξ_{low}	1.14 (1.05, 1.53)
ξ_{med}	1.25 (1.08, 1.74)
ξ_{high}	3.27 (1.68, 19.07)
σ^2 (variance of household-level random effect)	1.36 (0.23, 5.35)
Rattiness variables	
Relative elevation (per 1m increase) ²	
0-8m	0.05 (-0.01, 0.13)
8-22m	-0.06 (-0.16, 0.02)
>22m	0.05 (-0.03, 0.14)
Distance to large refuse piles (per 10m increase) ³	
0-50m	-0.10 (-0.21, 0.02)
>50m	0.03 (-0.11, 0.17)
Impervious land cover (per 10% increase)	-0.07 (-0.14, -0.01)
Rattiness parameters	
α_{traps}	-2.94 (-3.27, -2.65)
α_{plates}	-2.06 (-2.50, -1.74)
$\alpha_{burrows}$	-1.41 (-1.67, -1.16)
α_{faeces}	-2.82 (-3.83, -2.32)
α_{trails}	-2.22 (-2.96, -1.76)
σ_{traps}	0.72 (0.45, 0.97)
σ_{plates}	2.37 (2.05, 2.68)
$\sigma_{burrows}$	1.28 (1.08, 1.45)
σ_{faeces}	2.36 (1.80, 3.34)
σ_{trails}	2.43 (1.85, 3.12)
ψ	0.67 (0.29, 1.00)
ϕ	9.23 (3.21, 18.24)

405 3.4 Spatial prediction

406 There was heterogeneous spatial variation in predicted rattiness. The numerous small regions of high rattiness
407 in Figure 5A are indicative of the small-scale spatial correlation in the data. The low elevation areas in the
408 central length of each valley (relative elevation is shown in Figure 5E with the contours marking the low,
409 medium and high elevation areas) had high mean predicted rattiness. High rattiness was also predicted in
410 several high elevation areas, for example the northern tip of Valley 3 and several small hotspots along the
411 three valley's high elevation sides.

412 To illustrate the spatial variation in infection risk within the study area, prediction maps are shown in
413 Figure 5B for a 30-year-old male participant with a household per capita income of USD\$1/day who never or
414 rarely had contact with floodwater in the previous six months and did not work as a travelling salesperson.
415 Infection risk was low across most of Valley 1 (<2.5%), with marginally higher average values found in the
416 central low elevation area (2.5-5%). Risk was consistently higher across most of Valleys 2 and 3 (7.5-15%), with
417 the effect of elevation on risk clearly visible. In areas with higher and more spatially heterogeneous predicted
418 infection risk, for example in the central region of Valley 2, this was driven by high levels of predicted rattiness.
419 The stronger estimated effect of rattiness on infection risk in higher elevation areas was particularly visible in
420 Valleys 2 and 3, as seen in the three hotspots with risk reaching 20% and the moderate risk hotspots along
421 the sides of both valleys. Prediction intervals were relatively narrow across most of the study area (Figures 5C
422 and 5D) with greater uncertainty in the high risk areas.

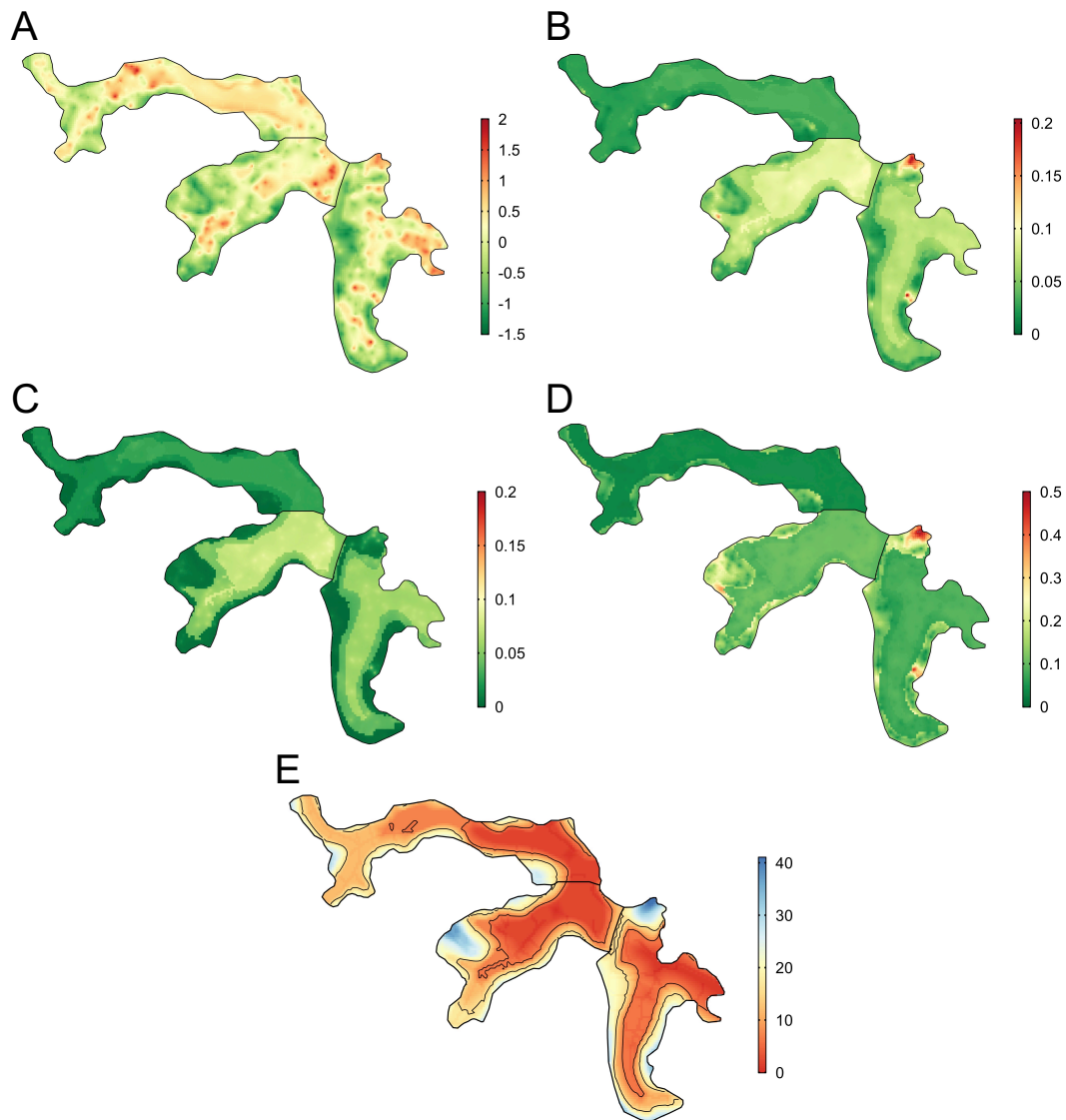


Figure 5: Joint rattiness-infection model predictions. A) Mean predicted rattiness; B) Mean predicted leptospiral infection risk for 30-year-old male participants with a household per capita income of USD\$1/day who never/rarely have contact with floodwater and do not work as a travelling salesperson; C) lower 95% prediction interval for predicted infection risk; D) upper 95% prediction interval for predicted infection risk; E) Elevation (metres) relative to the bottom of valley (low, medium and high levels are shown with contours)

423 4 Discussion

424 We developed and applied a novel framework for joint spatial modelling of disease reservoir abundance and
425 human infection risk to a community-based cohort study and fine-scale rat ecology study. We found that
426 higher levels of rattiness, our proxy for rat abundance, at the household location were associated with a higher

427 risk of leptospiral infection for residents across the entire study area. Importantly, we found that a unit
428 increase in rattiness in high elevation areas was associated with an almost three times higher odds ratio for
429 infection than in low and medium elevation areas. To our knowledge this is the first study to jointly model
430 rodent abundance and human infection data for a rodent-borne zoonosis. The findings provide new insights
431 into how the dominant mechanisms of *Leptospira* transmission within complex urban settings may vary over
432 small distances, as a result of interactions between rats, the environment, geography, and local epidemiology.

433 The finding that rattiness was associated with infection risk indicates that the spatial distribution of rat
434 populations was an important driver of transmission close to the household across the entire study area. This is
435 consistent with a recent study investigating the predictive power of household rat infestation scores for human
436 infection [30]. There was no residual spatial correlation in the infection data after accounting for rattiness
437 in our analysis, possibly suggesting that previously unexplained spatial heterogeneity in risk could be driven
438 by variation in rattiness [12]. Our model also predicted high average rattiness across the low elevation areas
439 where leptospiral transmission is high [10, 12]. This supports the hypothesis that abundant rat populations
440 are responsible for high levels of observed environmental contamination across these lower areas [21, 60], and
441 consequently increased infection risk.

442 The identified interaction between elevation and rattiness on infection risk suggests that relatively small
443 changes in environment and topography can modify transmission pathways within an urban community. The
444 weaker effect of rattiness on infection risk at low and medium elevation areas relative to high elevation areas
445 may be explained by differences in their hydrological profiles. While high rat abundance in low and medium
446 areas results in high leptospiral contamination, these areas are prone to high levels of water runoff and flooding.
447 This disperses the pathogen across low elevation areas. The ability of leptospires to persist in the environment
448 for weeks or months means that this process can significantly increase environmental risk in low elevation areas
449 for long periods. This process disconnects shedding and infection events in space and time [1] and obscures
450 the relationship between infection risk and rattiness in low and medium elevation households.

451 In contrast, high elevation areas have lower levels of water runoff and flooding due to improved drainage
452 and sewage systems, and a smaller upstream catchment area for rainfall. Leptospires are consequently less
453 likely to be washed away from the location at which they were shed, and environmental risk remains more
454 localised and strongly associated with the spatial distribution of rats. This hypothesised role of hydrology
455 in the aggregation and dispersal of leptospires [1] is supported by a recent study in low elevation areas of
456 Pau da Lima which found that soil contamination was not associated with local rat activity [20]. However,
457 our finding that rattiness was associated with infection in low and medium elevation areas suggests that the

458 spatial distribution of environmental risk in these areas is not entirely determined by water dispersal.

459 Interestingly, a previous study of surface waters in Pau da Lima found that the probability of a sample
460 being positive for *Leptospira* was highest in low elevation areas and lowest in medium elevation areas, with
461 no significant difference between low and high elevation areas [21]. This is consistent with our findings and
462 suggests that there may be a ‘washing out’ of locally deposited leptospires in medium elevation areas but not
463 in high elevation areas.

464 This has several implications for disease control strategies which aim to reduce environmental risk. Improv-
465 ing drainage systems at all elevation levels can reduce the dispersal of leptospires from high to low elevation
466 areas. Closure of sewer systems, which generally run through low elevation areas, can protect local residents
467 from exposure and reduce the introduction of additional contamination from upstream sewer water. Paving
468 over soil surfaces can reduce the surface area over which leptospires can persist [6], reducing environmental
469 risk further.

470 A reduction in the dispersal and accumulation of bacteria will result in more localised environmental risk,
471 as was observed in the high elevation areas in this study. Higher risk will then be found in areas with a
472 high abundance of infected rats. This may also reduce environmental exposure for rats, thereby lowering
473 shedding rates and acting as a feedback loop into the *Leptospira* transmission cycle [61]. Given the limited
474 and short-term impact of chemical rodenticide campaigns on Norway rat abundance in these settings [62],
475 longer-term environment management strategies targeted at rattiness hotspots may also be needed to reduce
476 the availability of key predictors of rattiness, such as large refuse piles and vegetation and soil land cover.
477 Funding and political will for large-scale infrastructural interventions is often limited in marginalised urban
478 settings and small-scale community-based interventions which target these mechanisms should be evaluated.

479 Transmission is dynamic in space and time and the alignment of conditions which enable spillover infection
480 can vary over time [1]. Our study was designed to explore the spatial variation in rattiness and infection
481 risk in Pau da Lima during the driest period of the year, and it may not be representative of transmission
482 mechanisms during the rainy season. There is some evidence, however, that this may not necessarily be the
483 case, with two recent studies in Pau da Lima reporting low seasonal variation in both rat abundance [63] and
484 spatial infection risk patterns [12]. Nonetheless, future studies across different time periods are needed to
485 establish the role of rat abundance in *Leptospira* transmission.

486 In this study we used household location to link rattiness to an individual, under the assumption that the
487 majority of their exposure occurs close to home. Given the spatially heterogeneous distribution of rattiness
488 and environmental contamination [21] within the community, future epidemiological studies of leptospirosis

489 and zoonotic spillover could benefit from trying to pinpoint key sources of infection away from the household
490 using GPS mobility data, as has been attempted in a small study previously [64]. The rattiness-infection
491 framework could then be extended to model cumulative environmental exposure to the rattiness surface by
492 integrating along a person's trajectory as they move around the community.

493 Our framework did not account for disease dynamics within rat populations. Given that 80% of rats are
494 estimated to be actively shedding *Leptospira* in Pau da Lima [27, 28] and prevalence in rats is generally high in
495 urban areas globally [22–26], the use of rattiness as a proxy for rat shedding appears reasonable and it may be
496 a useful proxy in other epidemiological studies. However, in other zoonotic spillover systems where pathogen
497 release does not occur at a high and homogeneous rate across the reservoir host population this may not be
498 the case and accounting for spatially heterogeneous or time-varying [65] disease dynamics will be important.

499 The rattiness-infection modelling framework is a flexible tool for exploring the spatial association between
500 reservoir abundance, the environment and human health outcomes. It provides a statistically principled
501 method for joint spatial modelling of infection risk and multiple indices of reservoir abundance, pooling data
502 between indices and directly accounting for uncertainty in their measurement in all parameter estimates and
503 predictions. The framework's geostatistical structure includes spatially continuous predictors for abundance
504 and accounts for spatial correlation, enabling mapping of both infection risk and rattiness. This can be
505 useful for identifying high-risk areas and targeting control. One inherent limitation is its dependence on the
506 availability of spatially continuous environmental variables and abundance data, both of which are prone to
507 high measurement error. This can result in high uncertainty in the model parameter estimates and predictions,
508 as demonstrated by the wider confidence intervals for risk factors in the joint model compared to the standard
509 mixed-effects logistic regression analysis. An additional benefit of the geostatistical structure is that abundance
510 measurements do not have to be taken at the household location, providing some flexibility in the design of
511 eco-epidemiological studies and indices used.

512 In conclusion, we have developed a framework that may have broad applications in delineating com-
513 plex animal-environment-human interactions during zoonotic spillover and identifying opportunities for public
514 health intervention. We demonstrate its potential by applying it to *Leptospira* in an urban setting, finding
515 evidence that the extent to which local rat shedding drives spillover transmission is moderated by elevation,
516 most likely a proxy for water runoff. Future work examining these transmission mechanisms in similar settings
517 and across different time points will be key to establishing how generalisable these results are.

518 **CRedit authorship contribution statement**

519 **Max T. Eyre:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data
520 curation, Visualization, Writing-original draft, Writing - Review and editing.

521 **Fábio N. Souza, Ticiana S. A. Carvalho-Pereira, Nivison Nery Jr., Daiana de Oliveira, Jaqueline**
522 **S Cruz, Kathryn P. Hacker, Gielson A. Sacramento and Elsio A. Wunder Jr:** Investigation, Data
523 curation, Writing - Review and editing.

524 **Hussein Khalil and James E. Childs:** Writing - Review and editing.

525 **Mitermayer G. Reis:** Project administration, Conceptualization, Investigation, Resources, Supervision, Re-
526 view and editing, Funding acquisition.

527 **Mike Begon:** Conceptualization, Supervision, Writing - Review and editing.

528 **Albert I. Ko:** Conceptualization, Funding acquisition, Writing - Review and editing.

529 **Peter J. Diggle and Emanuele Giorgi:** Conceptualization, Methodology, Software, Supervision, Writing -
530 Review and editing.

531 **Federico Costa:** Project administration, Conceptualization, Investigation, Resources, Supervision, Writing
532 - Review and editing, Funding acquisition.

533

534 **Acknowledgements**

535 We thank the residents and community leaders of Pau da Lima community for their support and partici-
536 pation in this study. This work was supported by the Oswaldo Cruz Foundation and Secretariat of Health
537 Surveillance, Brazilian Ministry of Health, the National Institutes of Health of the United States (grant num-
538 bers F31 AI114245, R01 AI052473, U01 AI088752, R01 TW009504 and R25 TW009338); the Wellcome Trust
539 (102330/Z/13/Z), and by the Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB/JCB0020/2016).
540 M.T.E is supported by an MRC doctorate studentship. F.N.S. participated in this study under a FAPESB
541 doctorate scholarship.

542 **Declaration of Competing Interest**

543 The authors declare that they have no known competing financial interests or personal relationships that could
544 have appeared to influence the work reported in this paper.

545 References

- 546 [1] Raina K. Plowright, Colin R. Parrish, Hamish McCallum, Peter J. Hudson, Albert I. Ko, Andrea L.
547 Graham, and James O. Lloyd-Smith. “Pathways to zoonotic spillover”. In: *Nature Reviews Microbiology*
548 15.8 (2017), pp. 502–510. ISSN: 17401534. DOI: 10.1038/nrmicro.2017.45. URL: [http://dx.doi.org/](http://dx.doi.org/10.1038/nrmicro.2017.45)
549 [10.1038/nrmicro.2017.45](http://dx.doi.org/10.1038/nrmicro.2017.45).
- 550 [2] Daniel J. Becker, Alex D. Washburne, Christina L. Faust, Juliet R.C. Pulliam, Erin A. Mordecai,
551 James O. Lloyd-Smith, and Raina K. Plowright. “Dynamic and integrative approaches to understanding
552 pathogen spillover”. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 374.1782
553 (2019), pp. 1–7. ISSN: 14712970. DOI: 10.1098/rstb.2019.0014.
- 554 [3] Pandji W. Dhewantara, Colleen L. Lau, Kathryn J. Allan, Wenbiao Hu, Wenyi Zhang, Abdullah A.
555 Mamun, and Ricardo J. Soares Magalhães. “Spatial epidemiological approaches to inform leptospirosis
556 surveillance and control: A systematic review and critical appraisal of methods”. In: *Zoonoses and Public*
557 *Health* 66.2 (2019), pp. 185–206. ISSN: 18632378. DOI: 10.1111/zph.12549.
- 558 [4] Federico Costa, José E. Hagan, Juan Calcagno, Michael Kane, Paul Torgerson, Martha S. Martinez-
559 Silveira, Claudia Stein, Bernadette Abela-Ridder, and Albert I. Ko. “Global Morbidity and Mortality
560 of Leptospirosis: A Systematic Review”. In: *PLoS Neglected Tropical Diseases* 9.9 (2015), pp. 0–1. ISSN:
561 19352735. DOI: 10.1371/journal.pntd.0003898. URL: [http://dx.doi.org/10.1371/journal.pntd.](http://dx.doi.org/10.1371/journal.pntd.0003898)
562 [0003898](http://dx.doi.org/10.1371/journal.pntd.0003898).
- 563 [5] Paul R. Torgerson, José E. Hagan, Federico Costa, Juan Calcagno, Michael Kane, Martha S. Martinez-
564 Silveira, Marga G.A. Goris, Claudia Stein, Albert I. Ko, and Bernadette Abela-Ridder. “Global Burden
565 of Leptospirosis: Estimated in Terms of Disability Adjusted Life Years”. In: *PLoS Neglected Tropical*
566 *Diseases* 9.10 (2015), pp. 1–14. ISSN: 19352735. DOI: 10.1371/journal.pntd.0004122.
- 567 [6] Emilie Bierque, Roman Thibeaux, Dominique Girault, Marie Estelle Soupé-Gilbert, and Cyrille Goarant.
568 “A systematic review of *Leptospira* in water and soil environments”. In: *PLoS ONE* 15.1 (2020), pp. 1–
569 22. ISSN: 19326203. DOI: 10.1371/journal.pone.0227055.
- 570 [7] Cyrille Goarant. “Leptospirosis: risk factors and management challenges in developing countries”. In:
571 *Research and Reports in Tropical Medicine* Volume 7 (2016), pp. 49–62. ISSN: 1179-7282. DOI: 10.2147/
572 [rrtm.s102543](https://doi.org/10.2147/rrtm.s102543).
- 573 [8] Colleen L. Lau, Lee D. Smythe, Scott B. Craig, and Philip Weinstein. “Climate change, flooding, urban-
574 isation and leptospirosis: Fuelling the fire?” In: *Transactions of the Royal Society of Tropical Medicine*

- 575 *and Hygiene* 104.10 (2010), pp. 631–638. ISSN: 00359203. DOI: 10.1016/j.trstmh.2010.07.002. URL:
576 <http://dx.doi.org/10.1016/j.trstmh.2010.07.002>.
- 577 [9] WHO. *Leptospirosis Burden Epidemiology Reference Group (LERG)*. 2010. URL: [http://www.who.int/
578 zoonoses/diseases/lerg/en/](http://www.who.int/zoonoses/diseases/lerg/en/) (visited on 05/22/2021).
- 579 [10] Renato B. Reis, Guilherme S. Ribeiro, Ridalva D M Felzemburgh, Francisco S. Santana, Sharif Mohr,
580 Astrid X T O Melendez, Adriano Queiroz, Andréia C. Santos, Romy R. Ravines, Wagner S. Tassinari,
581 Marília S. Carvalho, Mitermayer G. Reis, and Albert I. Ko. “Impact of environment and social gradient
582 on *Leptospira* infection in urban slums”. In: *PLoS Neglected Tropical Diseases* 2.4 (2008), pp. 11–18.
583 ISSN: 1935-2735. DOI: 10.1371/journal.pntd.0000228.
- 584 [11] Ridalva D.M. Felzemburgh, Guilherme S. Ribeiro, Federico Costa, Renato B. Reis, José E. Hagan, Astrid
585 X.T.O. Melendez, Deborah Fraga, Francisco S. Santana, Sharif Mohr, Balbino L. dos Santos, Adriano Q.
586 Silva, Andréia C. Santos, Romy R. Ravines, Wagner S. Tassinari, Marília S. Carvalho, Mitermayer G.
587 Reis, and Albert I. Ko. “Prospective Study of Leptospirosis Transmission in an Urban Slum Community:
588 Role of Poor Environment in Repeated Exposures to the *Leptospira* Agent”. In: *PLoS Neglected Tropical
589 Diseases* 8.5 (2014). ISSN: 19352735. DOI: 10.1371/journal.pntd.0002927.
- 590 [12] José E. Hagan, Paula Moraga, Federico Costa, Nicolas Capián, Guilherme S. Ribeiro, Elsie A. Wunder,
591 Ridalva D.M. Felzemburgh, Renato B. Reis, Nivison Nery, Francisco S. Santana, Deborah Fraga, Balbino
592 L. dos Santos, Andréia C. Santos, Adriano Queiroz, Wagner Tassinari, Marília S. Carvalho, Mitermayer
593 G. Reis, Peter J. Diggle, and Albert I. Ko. “Spatiotemporal Determinants of Urban Leptospirosis Trans-
594 mission: Four-Year Prospective Cohort Study of Slum Residents in Brazil”. In: *PLoS Neglected Tropical
595 Diseases* 10.1 (2016), pp. 1–17. ISSN: 19352735. DOI: 10.1371/journal.pntd.0004275.
- 596 [13] Hussein Khalil, Roberta Santana, Daiana de Oliveira, Fabiana Palma, Ricardo Lustosa, Max T. Eyre,
597 Ticiania Carvalho-Pereira, Mitermayer G. Reis, Albert I. Ko, Peter J. Diggle, Yeimi Alzate Lopez, Mike
598 Begon, and Federico Costa. “Poverty, sanitation, and *Leptospira* transmission pathways in residents from
599 four Brazilian slums”. In: *PLOS Neglected Tropical Diseases* 15.3 (2021), e0009256. ISSN: 19352735. DOI:
600 10.1371/journal.pntd.0009256. URL: <http://dx.doi.org/10.1371/journal.pntd.0009256>.
- 601 [14] Christovam Barcellos and Paulo Chagastelles Sabroza. “Socio-environmental determinants of the lep-
602 tospirosis outbreak of 1996 in western Rio de Janeiro: A geographical approach”. In: *International
603 Journal of Environmental Health Research* 10.4 (2000), pp. 301–313. ISSN: 09603123. DOI: 10.1080/
604 0960312002001500.

- 605 [15] Christovam Barcellos and Paulo Chagastelles Sabroza. “The place behind the case: leptospirosis risks
606 and associated environmental conditions in a flood-related outbreak in Rio de Janeiro”. In: *Cadernos de*
607 *Saúde Pública* 17.suppl (2001), S59–S67. ISSN: 0102-311X. DOI: 10.1590/S0102-311X2001000700014.
608 URL: [http://www.scielo.br/scielo.php?script=sci%7B%5C_%7Darttext%7B%5C%7Dpid=S0102-](http://www.scielo.br/scielo.php?script=sci%7B%5C_%7Darttext%7B%5C%7Dpid=S0102-311X2001000700014%7B%5C%7Dlng=en%7B%5C%7Dtlng=en)
609 [311X2001000700014%7B%5C%7Dlng=en%7B%5C%7Dtlng=en](http://www.scielo.br/scielo.php?script=sci%7B%5C_%7Darttext%7B%5C%7Dpid=S0102-311X2001000700014%7B%5C%7Dlng=en%7B%5C%7Dtlng=en).
- 610 [16] Mwanajaa Abdalla Mwachui, Lisa Crump, Rudy Hartskeerl, Jakob Zinsstag, and Jan Hattendorf. “Envi-
611 ronmental and Behavioural Determinants of Leptospirosis Transmission: A Systematic Review”. In: *PLoS*
612 *Neglected Tropical Diseases* 9.9 (2015), pp. 1–15. ISSN: 19352735. DOI: 10.1371/journal.pntd.0003843.
- 613 [17] John Keenan, Genine Ervin, Maung Aung, Gerald Jr McGwin, and Pauline Jolly. “Risk factors for
614 clinical leptospirosis from Western Jamaica.” eng. In: *The American journal of tropical medicine and*
615 *hygiene* 83.3 (Sept. 2010), pp. 633–636. ISSN: 1476-1645 (Electronic). DOI: 10.4269/ajtmh.2010.09-
616 0609.
- 617 [18] S. G. Prabhakaran, S. Shanmughapriya, S. Dhanapaul, A. James, and K. Natarajaseenivasan. “Risk
618 factors associated with rural and urban epidemics of leptospirosis in Tiruchirappalli district of Tamilnadu,
619 India”. In: *Journal of Public Health (Germany)* 22.4 (2014), pp. 323–333. ISSN: 16132238. DOI: 10.1007/
620 [s10389-014-0611-1](https://doi.org/10.1007/s10389-014-0611-1).
- 621 [19] Emily A. Briskin, Arnau Casanovas-Massana, Kyle R. Ryff, Shirley Morales-Estrada, Camila Hamond,
622 Nicole M. Perez-Rodriguez, Kathryn M. Benavidez, Daniel M. Weinberger, Iván Castro-Arellano, Elsie
623 A. Wunder, Tyler M. Sharp, Brenda Rivera-Garcia, and Albert I. Ko. “Seroprevalence, Risk Factors,
624 and Rodent Reservoirs of Leptospirosis in an Urban Community of Puerto Rico, 2015”. In: *The Journal*
625 *of infectious diseases* 220.9 (2019), pp. 1489–1497. ISSN: 15376613. DOI: 10.1093/infdis/jiz339.
- 626 [20] Andrew G Schneider, Arnau Casanovas-Massana, Kathryn P Hacker, Elsie A Wunder Jr, Mike Begon,
627 Mitermayer G Reis, James E Childs, Federico Costa, Janet C Lindow, and Albert I Ko. “Quantification
628 of pathogenic *Leptospira* in the soils of a Brazilian urban slum”. In: *PLoS neglected tropical diseases*
629 12.4 (2018), e0006415.
- 630 [21] Arnau Casanovas-Massana, Federico Costa, Irina N. Riediger, Marcelo Cunha, Daiana de Oliveira, Dio-
631 genes C. Mota, Erica Sousa, Vladimir A. Querino, Nivisson Nery, Mitermayer G. Reis, Elsie A. Wunder,
632 Peter J. Diggle, and Albert I. Ko. “Spatial and temporal dynamics of pathogenic *Leptospira* in surface
633 waters from the urban slum environment”. In: *Water Research* 130 (2018), pp. 176–184. ISSN: 18792448.
634 DOI: 10.1016/j.watres.2017.11.068. URL: <https://doi.org/10.1016/j.watres.2017.11.068>.

- 635 [22] Maysa Pellizzaro, Camila Marinelli Martins, Ana Carolina Yamakawa, Diogo daCunha Ferraz, Vivien
636 Midori Morikawa, Fernando Ferreira, Andrea Pires dos Santos, Alexander Welker Biondo, and Helio
637 Langoni. “Molecular detection of *Leptospira* spp. In rats as early spatial predictor for human disease in
638 an endemic urban area”. In: *PLoS ONE* 14.5 (2019), pp. 1–11. ISSN: 19326203. DOI: 10.1371/journal.
639 pone.0216830.
- 640 [23] Federico Costa, Fleur Helena Porter, Gorete Rodrigues, Helena Farias, Marcus Tucunduva de Faria, Elsio
641 A. Wunder, Lynn M. Osikowicz, Michael Y. Kosoy, Mitermayer Galvão Reis, Albert I. Ko, and James
642 E. Childs. “Infections by *Leptospira interrogans*, Seoul Virus, and *Bartonella* spp. Among
643 Norway Rats (*Rattus norvegicus*) from the Urban Slum Environment in Brazil”. In: *Vector-Borne
644 and Zoonotic Diseases* 14.1 (2014), pp. 33–40. ISSN: 1530-3667. DOI: 10.1089/vbz.2013.1378. URL:
645 <http://online.liebertpub.com/doi/abs/10.1089/vbz.2013.1378>.
- 646 [24] Kenneth Boey, Kanae Shiokawa, and Sreekumari Rajeev. “*Leptospira* infection in rats: A literature
647 review of global prevalence and distribution”. In: *PLoS Neglected Tropical Diseases* 13.8 (2019), pp. 1–
648 24. ISSN: 19352735. DOI: 10.1371/journal.pntd.0007499.
- 649 [25] Muhammad Afif Yusof, Farah Shafawati Mohd-Taib, Siti Nabilah Ishak, Shukor Md-Nor, Shahrul Anuar
650 Md-Sah, Nor Zalipah Mohamed, Nurul Natasya Azhari, Vasanthakumari Neela, and Zamberi Sekawi.
651 “Microhabitat Factors Influenced the Prevalence of Pathogenic *Leptospira* spp. in Small Mammal Host”.
652 In: *EcoHealth* 16.2 (2019), pp. 260–274. ISSN: 1612-9210. DOI: 10.1007/s10393-019-01419-1. URL:
653 <https://doi.org/10.1007/s10393-019-01419-1>.
- 654 [26] L H Krøjgaard, S Villumsen, M D K Markussen, J S Jensen, H Leirs, and A.-C. Heiberg. “High prevalence
655 of *Leptospira* spp. in sewer rats (*Rattus norvegicus*)”. In: *Epidemiology and Infection* 137.11 (2009),
656 pp. 1586–1592. DOI: 10.1017/S0950268809002647.
- 657 [27] Federico Costa, Elsio A. Wunder, Daiana de Oliveira, Vimla Bisht, Gorete Rodrigues, Mitermayer G.
658 Reis, Albert I. Ko, Mike Begon, and James E. Childs. “Patterns in *Leptospira* shedding in Norway rats
659 (*Rattus norvegicus*) from Brazilian slum communities at high risk of disease transmission”. In: *PLoS
660 Neglected Tropical Diseases* 9.6 (2015), pp. 1–14. ISSN: 19352735. DOI: 10.1371/journal.pntd.0003819.
661 URL: <http://dx.doi.org/10.1371/journal.pntd.0003819>.
- 662 [28] Marcos Tucunduva de Faria, Michael S. Calderwood, Daniel A. Athanazio, Alan J.A. McBride, Rudy
663 A. Hartskeerl, Martha Maria Pereira, Albert I. Ko, and Mitermayer G. Reis. “Carriage of *Leptospira*
664 *interrogans* among domestic rats from an urban setting highly endemic for leptospirosis in Brazil”. In:
665 *Acta Tropica* 108.1 (2008), pp. 1–5. ISSN: 0001706X. DOI: 10.1016/j.actatropica.2008.07.005.
666 arXiv: NIHMS150003. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0001706X0800209X>.

- 667 [29] Federico Costa, Guilherme S. Ribeiro, Ridalva D.M. Felzemburgh, Norlan Santos, Renato Barbosa Reis,
668 Andreia C. Santos, Deborah Bittencourt Mothe Fraga, Wildo N. Araujo, Carlos Santana, James E.
669 Childs, Mitermayer G. Reis, and Albert I. Ko. “Influence of Household Rat Infestation on *Leptospira*
670 Transmission in the Urban Slum Environment”. In: *PLoS Neglected Tropical Diseases* 8.12 (2014). ISSN:
671 19352735. DOI: 10.1371/journal.pntd.0003338.
- 672 [30] Federico Costa, Caio Graco Zeppelini, Guilherme S. Ribeiro, Norlan Santos, Renato Barbosa Reis,
673 Ridalva D. Martins, Deborah Bittencourt, Carlos Santana, Jonas Brant, Mitermayer G. Reis, and Albert
674 I. Ko. “Household rat infestation in urban slum populations: Development and validation of a predictive
675 score for leptospirosis”. In: *PLoS Neglected Tropical Diseases* 15.3 (2021), pp. 1–13. ISSN: 19352735. DOI:
676 10.1371/journal.pntd.0009154. URL: <http://dx.doi.org/10.1371/journal.pntd.0009154>.
- 677 [31] Pankaj Bhardwaj, J K Kosambiya, and Vikas K Desai. “A case control study to explore the risk factors
678 for acquisition of leptospirosis in Surat city, after flood.” eng. In: *Indian journal of medical sciences*
679 62.11 (Nov. 2008), pp. 431–438. ISSN: 0019-5359 (Print). DOI: 10.4103/0019-5359.48454.
- 680 [32] Frédéric Bordes, Kim Blasdel, and Serge Morand. “Transmission ecology of rodent-borne diseases: New
681 frontiers”. In: *Integrative Zoology* 10.5 (2015), pp. 424–435. ISSN: 17494877. DOI: 10.1111/1749-4877.
682 12149.
- 683 [33] Max T Eyre, Ticiana S A Carvalho-Pereira, Fábio N Souza, Hussein Khalil, Kathryn P Hacker, Soledad
684 Serrano, Joshua P Taylor, Mitermayer G Reis, Albert I Ko, Mike Begon, et al. “A multivariate geo-
685 statistical framework for combining multiple indices of abundance for disease vectors and reservoirs: a
686 case study of rattiness in a low-income urban Brazilian community”. In: *Journal of the Royal Society*
687 *Interface* 17.170 (2020), p. 20200398.
- 688 [34] Jillian Hurd, Olaf Berke, Zvonimir Poljak, and Martin Runge. “Spatial analysis of *Leptospira* infection
689 in muskrats in Lower Saxony, Germany, and the association with human leptospirosis”. In: *Research in*
690 *Veterinary Science* 114.June (2017), pp. 351–354. ISSN: 15322661. DOI: 10.1016/j.rvsc.2017.06.015.
691 URL: <http://dx.doi.org/10.1016/j.rvsc.2017.06.015>.
- 692 [35] Colleen L. Lau, Conall H. Watson, John H. Lowry, Michael C. David, Scott B. Craig, Sarah J. Wynwood,
693 Mike Kama, and Eric J. Nilles. *Human Leptospirosis Infection in Fiji: An Eco-epidemiological Approach*
694 *to Identifying Risk Factors and Environmental Drivers for Transmission*. 2016. DOI: 10.1371/journal.
695 pntd.0004405.
- 696 [36] Helen J. Mayfield, Carl S. Smith, John H. Lowry, Conall H. Watson, Michael G. Baker, Mike Kama, Eric
697 J. Nilles, and Colleen L. Lau. “Predictive risk mapping of an environmentally-driven infectious disease

- 698 using spatial Bayesian networks: A case study of leptospirosis in Fiji”. In: *PLoS Neglected Tropical*
699 *Diseases* 12.10 (2018), pp. 1–16. ISSN: 19352735. DOI: 10.1371/journal.pntd.0006857.
- 700 [37] C. Rotejanaprasert, A. Lawson, H. Rossow, J. Sane, O. Huitu, H. Henttonen, and V. J. Del Rio Vilas.
701 “Towards integrated surveillance of zoonoses: Spatiotemporal joint modeling of rodent population data
702 and human tularemia cases in Finland”. In: *BMC Medical Research Methodology* 18.1 (2018), pp. 1–8.
703 ISSN: 14712288. DOI: 10.1186/s12874-018-0532-8.
- 704 [38] Matthew C. Nicholson and Thomas N. Mather. “Methods for Evaluating Lyme Disease Risks Using
705 Geographic Information Systems and Geospatial Analysis”. In: *Journal of Medical Entomology* 33.5
706 (1996), pp. 711–720. ISSN: 00222585. DOI: 10.1093/jmedent/33.5.711.
- 707 [39] Anna M. Winters, Bethany G. Bolling, Barry J. Beaty, Carol D. Blair, Rebecca J. Eisen, Andrew M.
708 Meyer, W. John Pape, Chester G. Moore, and Lars Eisen. “Combining mosquito vector and human
709 disease data for improved assessment of spatial west Nile virus disease risk”. In: *American Journal of*
710 *Tropical Medicine and Hygiene* 78.4 (2008), pp. 654–665. ISSN: 00029637. DOI: 10.4269/ajtmh.2008.
711 78.654.
- 712 [40] Elizabeth A. Cromwell, Steven T. Stoddard, Christopher M. Barker, Annelies Van Rie, William B.
713 Messer, Steven R. Meshnick, Amy C. Morrison, and Thomas W. Scott. “The relationship between
714 entomological indicators of *Aedes aegypti* abundance and dengue virus infection”. In: *PLoS Neglected*
715 *Tropical Diseases* 11.3 (2017), pp. 1–22. ISSN: 19352735. DOI: 10.1371/journal.pntd.0005429.
- 716 [41] Elisabeth Fichet-Calvet, Emilie Lecompte, Lamine Koivogui, Barré Soropogui, Amadou Doré, Fodé
717 Kourouma, Oumar Sylla, Stéphane Daffis, Kékoura Koulémou, and Jan Ter Meulen. “Fluctuation of
718 abundance and Lassa virus prevalence in *Mastomys natalensis* in Guinea, West Africa”. In: *Vector-*
719 *Borne and Zoonotic Diseases* 7.2 (2007), pp. 119–128. ISSN: 15303667. DOI: 10.1089/vbz.2006.0520.
- 720 [42] Rebecca J. Eisen and Lars Eisen. “Spatial modeling of human risk of exposure to vector-borne pathogens
721 based on epidemiological versus arthropod vector data”. In: *Journal of Medical Entomology* 45.2 (2008),
722 pp. 181–192. ISSN: 00222585. DOI: 10.1603/0022-2585(2008)45[181:SMOHR0]2.0.CO;2.
- 723 [43] Kathryn P. Hacker, Gielson A. Sacramento, Jaqueline S. Cruz, Daiana de Oliveira, Nivison Nery, Janet
724 C. Lindow, Mayara Carvalho, Jose Hagan, Peter J. Diggle, Mike Begon, Mitermayer G. Reis, Elsio A.
725 Wunder, Albert I. Ko, and Federico Costa. “Influence of Rainfall on *Leptospira* Infection and Disease
726 in a Tropical Urban Setting, Brazil”. In: *Emerging Infectious Diseases* 26.2 (2020), pp. 311–314. ISSN:
727 1080-6040. DOI: 10.3201/eid2602.190102.

- 728 [44] Paul A Harris, Robert Taylor, Robert Thielke, Jonathon Payne, Nathaniel Gonzalez, Jose G Conde, et al.
729 “A metadata-driven methodology and workflow process for providing translational research informatics
730 support”. In: *J Biomed Inform* 42.2 (2009), pp. 377–381.
- 731 [45] Jesús A. Panti-May, Ticiana S. A. Carvalho-Pereira, Soledad Serrano, Gabriel G. Pedra, Josh Tay-
732 lor, Arsinoê C. Pertile, Amanda Minter, Vladimir Airam, Mayara Carvalho, Nivison N. Júnior, Gorete
733 Rodrigues, Mitermayer G. Reis, Albert I. Ko, James E. Childs, Mike Begon, and Federico Costa. “A
734 Two-Year Ecological Study of Norway Rats (*Rattus norvegicus*) in a Brazilian Urban Slum”. In: *Plos*
735 *One* 11.3 (2016), e0152511. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0152511. URL: <http://dx.plos.org/10.1371/journal.pone.0152511>.
- 737 [46] Kathryn P. Hacker, Amanda Minter, Mike Begon, Peter J. Diggle, Soledad Serrano, Mitermayer G. Reis,
738 James E. Childs, Albert I. Ko, and Federico Costa. “A comparative assessment of track plates to quantify
739 fine scale variations in the relative abundance of Norway rats in urban slums”. In: *Urban Ecosystems*
740 19.2 (2016), pp. 561–575. ISSN: 15731642. DOI: 10.1007/s11252-015-0519-8.
- 741 [47] Centers for Disease Control and Prevention (CDC). *Integrated pest management: conducting urban rodent*
742 *surveys*. US Department of Health and Human Services, Atlanta, USA. 2006.
- 743 [48] Robert S Sikes and William L Gannon. “Guidelines of the American Society of Mammalogists for the
744 use of wild mammals in research”. In: *Journal of mammalogy* 92.1 (2011), pp. 235–253.
- 745 [49] Steven L Leary, Wendy Underwood, Raymond Anthony, S Cartner, D Corey, T Grandin, C Greenacre, S
746 Gwaltney-Brant, M A McCrackin, R Meyer, et al. “AVMA guidelines for the euthanasia of animals: 2013
747 edition”. In: *American Veterinary Medical Association Schaumburg, IL*. American Veterinary Medical
748 Association Schaumburg, IL. 2013.
- 749 [50] Trevor Hastie and Robert Tibshirani. “Generalized additive models: some applications”. In: *Journal of*
750 *the American Statistical Association* 82.398 (1987), pp. 371–386.
- 751 [51] Hamparsum Bozdogan. “Model selection and Akaike’s information criterion (AIC): The general theory
752 and its analytical extensions”. In: *Psychometrika* 52.3 (1987), pp. 345–370.
- 753 [52] Ole F Christensen. “Monte Carlo maximum likelihood in model-based geostatistics”. In: *Journal of*
754 *computational and graphical statistics* 13.3 (2004), pp. 702–718.
- 755 [53] Max Eyre, Federico Costa, and Albert Ko. *Linking rattiness, geography and environmental degradation*
756 *to spillover Leptospira infections in marginalised urban settings: Data sources*. Sept. 2021. DOI: 10.
757 17605/OSF.IO/AQZ2Y. URL: osf.io/aqz2y.

- 758 [54] R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2016. URL:
759 <https://www.r-project.org/>.
- 760 [55] Hadley Wickham. “The tidyverse”. In: *R package ver 1.1* (2017), p. 1.
- 761 [56] Simon Wood and Maintainer Simon Wood. “Package ‘mgcv’”. In: *R package version 1* (2015), p. 29.
- 762 [57] Emanuele Giorgi and Peter J Diggle. “PrevMap: An R Package for Prevalence Mapping”. In: *Journal*
763 *of Statistical Software* 78.8 (2017), pp. 1–29. DOI: 10.18637/jss.v078.i08.
- 764 [58] K Barton. *Mu-MIn: Multi-model inference. R Package Version 0.12. 2/r18. 2009. 2020*.
- 765 [59] Douglas Bates, Deepayan Sarkar, Maintainer Douglas Bates, and L Matrix. “The lme4 package”. In: *R*
766 *package version 2.1* (2007), p. 74.
- 767 [60] Arnau Casanovas-Massana, Daiana de Oliveira, Andrew G Schneider, Mike Begon, James E Childs,
768 Federico Costa, Mitermayer G Reis, Albert I Ko, and Elsie A Wunder. “Genetic evidence for a potential
769 environmental pathway to spillover infection of rat-borne leptospirosis”. In: *The Journal of Infectious*
770 *Diseases* (2021).
- 771 [61] Amanda Minter, Peter J. Diggle, Federico Costa, James Childs, Albert I. Ko, and Mike Begon. “A model
772 for leptospire dynamics and control in the Norway rat (*Rattus norvegicus*) the reservoir host in urban
773 slum environments”. In: *Epidemics* 25 (2018), pp. 26–34. ISSN: 18780067. DOI: 10.1016/j.epidem.
774 2018.05.002. URL: <https://doi.org/10.1016/j.epidem.2018.05.002>.
- 775 [62] Eduardo De Masi, Pedro José Vilaça, and Maria Tereza Pepe Razzolini. “Evaluation on the effectiveness
776 of actions for controlling infestation by rodents in Campo Limpo region, São Paulo Municipality, Brazil”.
777 In: *International Journal of Environmental Health Research* 19.4 (2009), pp. 291–304. ISSN: 09603123.
778 DOI: 10.1080/09603120802592723.
- 779 [63] Jesús A. Panti-May, Ticiania S.A. Carvalho-Pereira, Soledad Serrano, Gabriel G. Pedra, Josh Taylor,
780 Arsinoê C. Pertile, Amanda Minter, Vladimir Airam, Mayara Carvalho, Nivison N. Júnior, Gorete Ro-
781 drrigues, Mitermayer G. Reis, Albert I. Ko, James E. Childs, Mike Begon, and Federico Costa. “A
782 Two-Year Ecological Study of Norway Rats (*Rattus norvegicus*) in a Brazilian Urban Slum”. In: *PLoS*
783 *one* 11.3 (2016), e0152511. ISSN: 19326203. DOI: 10.1371/journal.pone.0152511.
- 784 [64] Katharine A. Owers, Juliana Odetunde, Rosan Barbosa de Matos, Gielson Sacramento, Mayara Carvalho,
785 Nivison Nery, Federico Costa, Mitermayer G. Reis, James E. Childs, José E. Hagan, Peter J. Diggle,
786 and Albert I. Ko. “Fine-scale GPS tracking to quantify human movement patterns and exposure to
787 leptospire in the urban slum environment”. In: *PLoS neglected tropical diseases* 12.8 (2018), e0006752.
788 ISSN: 19352735. DOI: 10.1371/journal.pntd.0006752.

It is made available under a [CC-BY 4.0 International license](#) .

- 789 [65] Stephen Davis, E. Calvet, and H. Leirs. “Fluctuating rodent populations and risk to humans from rodent-
790 borne zoonoses”. In: *Vector-Borne and Zoonotic Diseases* 5.4 (2005), pp. 305–314. ISSN: 15303667. DOI:
791 10.1089/vbz.2005.5.305.

Supplementary file 1 - Functional form of continuous explanatory variables

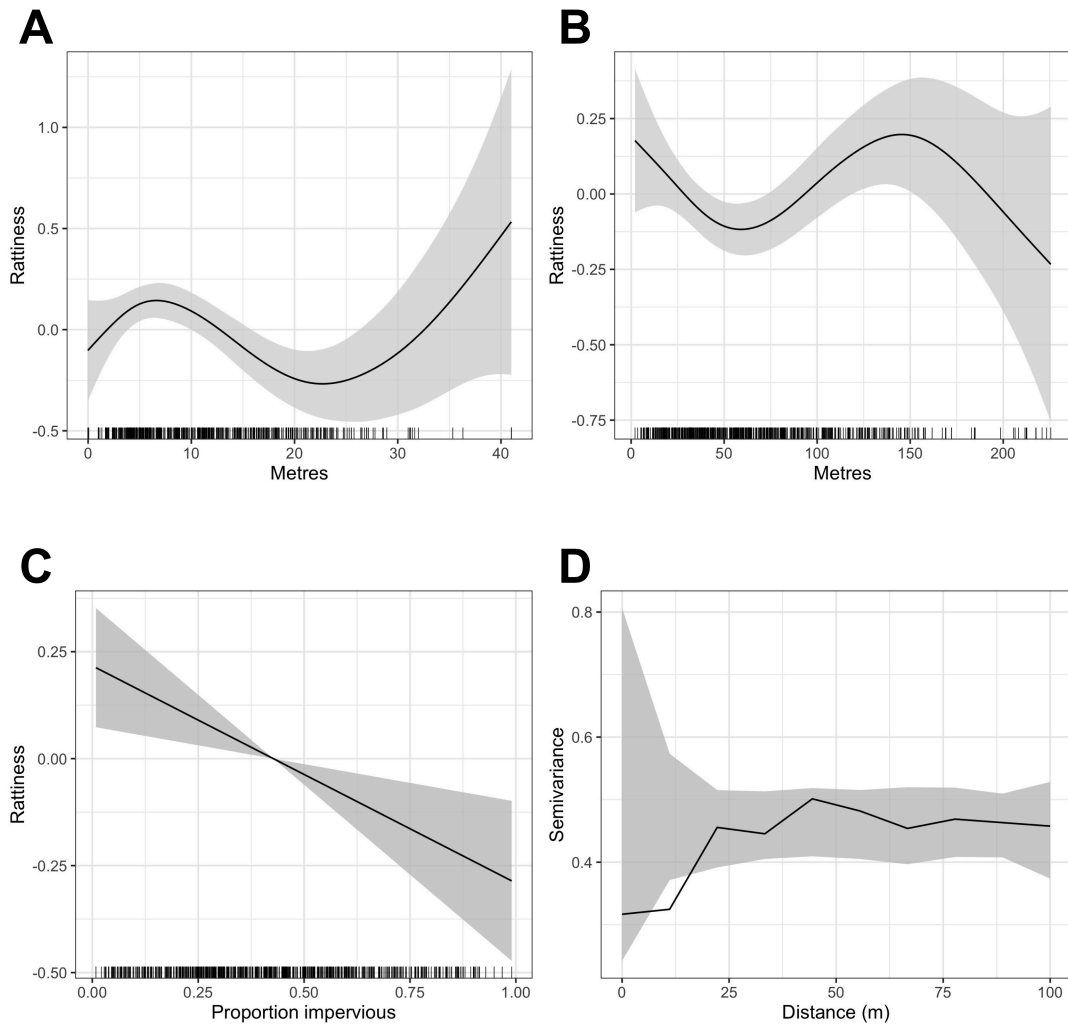


Figure S1: Generalized Additive Model (GAM) partial dependence plots for the unstructured random variation in rattiness, \hat{U}_i , plotted against the continuous explanatory variables considered in the analysis (shaded areas correspond to 95% confidence intervals): A) elevation relative to the bottom of valley, B) distance to large refuse piles, C) impervious land cover in 20m radius buffer around sampling point. \hat{U}_i are estimated using a non-spatial model which excludes all covariates. Panel D) is a variogram computed from \hat{U}_i using a non-spatial model that includes all of the covariates; the dashed lines correspond to 95% confidence intervals under the assumption of spatial independence.

It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

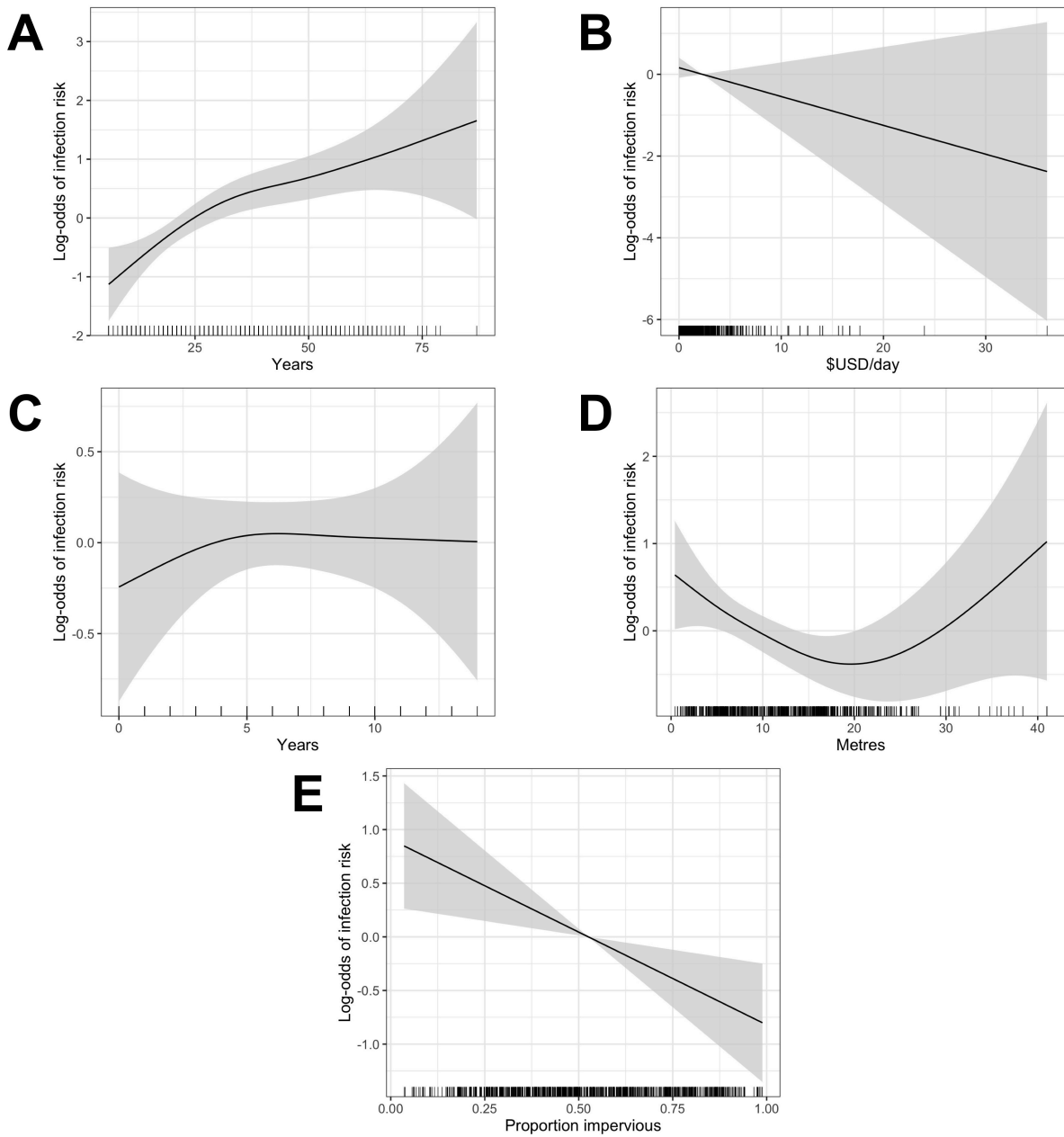


Figure S2: Generalized Additive Model (GAM) partial dependence plots for human infection risk plotted against the continuous explanatory variables considered in this analysis (shaded areas correspond to 95% confidence intervals): A) age, B) household per capita income (in USD), C) years of education, D) household elevation relative to the bottom of valley, E) impervious land cover in 20m radius buffer around household.