

Human *Trypanosoma cruzi* infection risk is driven by eco-social interactions in rural communities of the Argentine Chaco

Maria del Pilar Fernández<sup>1,2,3\*</sup>, Maria Sol Gaspe<sup>1,2</sup>, Paula Sartor<sup>1,4,5</sup>, Ricardo E. Gürtler<sup>1,2</sup>.

<sup>1</sup> Universidad de Buenos Aires. Departamento de Ecología, Genética y Evolución, Facultad de Ciencias Exactas y Naturales, Ciudad Universitaria, C1428EHA, Buenos Aires, Argentina

<sup>2</sup> Consejo Nacional de Investigaciones Científicas y Técnicas. Instituto de Ecología, Genética y Evolución de Buenos Aires (IEGEBA), Ciudad Universitaria, C1428EHA, Buenos Aires, Argentina

<sup>3</sup> Earth Institute, Columbia University. New York, NY, 10025, United States of America.

<sup>4</sup> Ministerio de Salud Pública del Chaco H3500ACC, Resistencia, Chaco, Argentina.

<sup>5</sup> Facultad de Ciencias Exactas, Naturales y Agrimensura, Universidad Nacional del Nordeste W3404AAR, Corrientes, Corrientes.

\*Corresponding author

María del Pilar Fernandez e-mail: [mpf2131@columbia.edu](mailto:mpf2131@columbia.edu)

María del Pilar Fernandez ORCID: 0000-0001-8645-2267

## Abstract

The transmission of *Trypanosoma cruzi* to humans is determined by multiple ecological, socio-economic and cultural factors acting at different scales. Their effects on the human risk of infection with *T. cruzi* have often been examined separately or using a limited set of ecological and socio-demographic variables. Herein, we integrated the ecological and social dimensions of human disease risk with the spatial distribution patterns of human and vector (*Triatoma infestans*) infection with *T. cruzi* in rural communities of the Argentine Chaco composed mainly of indigenous people (90% Qom) and a creole minority. Prior to the implementation of a vector control intervention, the estimated seroprevalence of *T. cruzi* among 1,929 local residents examined in a cross-sectional study was 29.0%, and was twice as large in Qom than creoles. Using generalized linear mixed models, the risk of human infection increased by 60% with each additional infected triatomine and by 40% with each seropositive household co-inhabitant; increased significantly with increasing household social vulnerability (a multidimensional index of poverty), and decreased with increasing host availability in sleeping quarters. A significant negative interaction between household social vulnerability and the relative abundance of infected *T. infestans* indicated that vulnerable household residents were exposed to a higher risk of infection even at low infected-vector abundances. Household mobility within the study area reduced the effects of domiciliary vector abundance, possibly due to less consistent exposures. Nonetheless, the seroprevalence rates of movers and non-movers were not significantly different. Human infection was clustered by household and at a larger spatial scale, with hotspots of human and vector infection matching areas of higher social vulnerability. These results were integrated in a risk map that shows high-priority areas for targeted interventions oriented to suppress house (re)infestations, treat infected children, and thus reduce the burden of future disease.

#### 44 Author summary

45 Chagas disease is one of the main neglected tropical diseases (NTDs) affecting vulnerable communities in  
 46 Latin America where transmission by triatomine vectors still occurs. Access to diagnosis and treatment is  
 47 one of the remaining challenges for sustainable control of Chagas disease in endemic areas. In this study,  
 48 we integrated the ecological and social determinants of human infection with the spatial component to  
 49 identify individuals, households and geographic sectors at higher risk of infection. We found that the risk  
 50 of human infection was higher in indigenous people compared to creoles, and increased with the  
 51 abundance of infected vectors and with household social vulnerability (a multidimensional index of  
 52 poverty). We also found that the social factors modulated the effect of the abundance of infected vectors:  
 53 vulnerable-household residents were exposed to a higher risk of infection even at low infected-vector  
 54 abundance, and human mobility within the area determined a lower and more variable exposure to the  
 55 vector over time. These results were integrated in a risk map that showed high-priority areas, which can  
 56 be used in designing cost-effective serological screening strategies adapted to resource-constrained areas.

# Introduction

The transmission of zoonotic and vector-borne diseases is an “inherently ecological process” involving intraspecific and interspecific interactions between vector, pathogen and host populations [1]. As shown by Ross-Macdonald style mathematical models, the transmission of vector-borne pathogens is determined by the density and survival of insect vectors, vector-host contact rate, host susceptibility, vector and host infectivity, and duration of infection [2,3]. When human populations are implicated, pathogen transmission dynamics also involve socio-economic, cultural, political, psychological and ethical factors, which pertain to the human dimension of disease [4,5]. With a focus on human health, these ecological, biological and social factors affecting the transmission of vector-borne diseases may be classified as intrinsic or extrinsic to the human population [6]. Intrinsic factors are biological in nature (i.e., immune response) and can only be manipulated by advances in medical research and technology [6]. Extrinsic factors include the environmental context, vector ecology and behavior, human activities, socio-economic inequalities and political factors, among other [6]. A thorough understanding of the combined effects of the extrinsic factors is required to design more effective and sustainable disease control interventions, which have traditionally been crafted in a reductionist biomedical approach [7]. Reductionism argues that the sum of information provided by studying each component of the system separately is sufficient to understand disease transmission dynamics [8]. By contrast, the Ecohealth or eco-bio-social approaches focus on the interactions among multiple ecological and social factors and their combined effects on human health as part of a complex socio-ecological system [7].

Chagas disease is a chronic infection caused by the protozoan *Trypanosoma cruzi*, and is mainly transmitted by triatomine vectors in endemic areas [9] in spite of regional and intergovernmental vector control (and elimination) initiatives, such as the INCOSUR (“Iniciativa de Salud del Cono Sur”) launched in the 1990s in the Southern Cone countries in South America [10]. Recent estimates indicate that 62.4% of people infected with *T. cruzi* live in INCOSUR countries, and 30.6% of new cases due to vector-borne transmission occur in Bolivia and Argentina [9], where the Gran Chaco region remains a hotspot for Chagas disease [11,12]. Although in the Argentine Chaco region the overall human seroprevalence of *T.*

*cruzi* has declined over the last 60 years [13–15], it remains high (27.8–71.1%) in rural communities encompassing creole and indigenous populations [16–24].

In the Gran Chaco, human infection with *T. cruzi* usually occurs within sleeping quarters before reaching 15 years of age and is transmitted by *Triatoma infestans* [25–27]. Most studies of human infection with *T. cruzi* in endemic areas have focused on the seroprevalence distribution among demographic subgroups and/or on the effects of vector presence, abundance and *T. cruzi* infection status. Although the association between human *T. cruzi* infection and selected socio-demographic factors has been investigated [24,25,28–31], these studies either did not address the combined effects of ecological and social variables due to limited data availability, or only considered a few socio-demographic variables. Human infection was positively associated with the presence or abundance of domestic animals [24,25,28,29,32]. A less clear association was found between *T. cruzi* human infection and house construction quality (i.e., thatched roofs and cracks in the walls): while in some studies human infection increased in poor-quality housing [25,28,31,32], others did not find such association [29,30].

The joint analysis of the spatial distribution of human and vector infection can shed light into the processes and factors associated with the vector-borne transmission of *T. cruzi*. Integrating the spatial component of disease with household-level and individual-based risk factor analysis is needed to identify transmission hotspots, create risk maps of *T. cruzi* infection, and stratify the affected areas for targeted control [33–35]. For Chagas disease, spatial analysis has been used to investigate the reinfestation process with *T. infestans* after community-wide insecticide spraying [36–40]; identify factors associated with house infestation [41–43]; evaluate the effects of vector control interventions [44], and generate risk maps at regional and continental levels based on vector distribution [45,46]. By contrast, the spatial component of human *T. cruzi* infection has rarely been investigated, and the ones that did it focused on periurban settings [29,47,48].

As part of a broader research and control program on the transmission of *T. cruzi* in rural communities of the Argentine Chaco [24,40–42,49–53], this study investigated the effects of household-level vector indices and demographic factors on human *T. cruzi* infection at individual, household and community

levels. We also considered the household socio-economic status and the interaction between social and ecological variables using indices of social vulnerability (as a measure of socio-economic inequalities) and of domestic host availability, i.e., in human sleeping quarters or domicile [54]. Both indices were positively associated to the abundance of infected domiciliary vectors [54]. Herein, we hypothesize that the interaction between ecological and social factors would enhance the risk of *T. cruzi* transmission to humans, which cannot solely be explained by variation in vector indices. At a community level, we evaluated the spatial correlation between human and vector infection, host availability and social vulnerability. We integrated the evidence provided by the individual-based risk model with the spatial component to derive community-level risk maps of human infection, with an emphasis in children, which represent the target demographic group for *T. cruzi* infection diagnosis and treatment.

## Methods

### ***Study area***

This study was conducted in a rural section of Pampa del Indio municipality (25° 55'S 56° 58'W), Chaco province, Argentina, denominated Area III, which encompassed three contiguous large communities (Cuarta Legua, Pampa Grande and Pampa Chica), and other smaller ones [41]. This area was subjected to a vector control and Chagas disease research program initiated in 2008 with a follow-up period of 7 years as of 2015. In October 2008, a baseline vector survey found that a third of the inhabited houses were infested with *T. infestans*, mainly within human sleeping quarters, and virtually all (93.4%) houses were sprayed with insecticides immediately after [41]. During the 2008-2015 vector surveillance phase we conducted annual triatomine surveys and selectively sprayed with insecticide the few foci detected. This strategy reduced house infestation to <1% during 2008-2012, and no infested house was found in 2015 [40]. Prior to 2008, vector control actions had been sporadic and partial; the last community-wide insecticide spraying campaign took place in 1997-1998, and the last sprays registered prior to the baseline survey were carried out in a few houses in 2006 and July 2008 [41].

The population in the area monotonically increased during the 2008-2015 period, from 2,392 people and 407 inhabited houses in 2008 to 2,548 people and 587 houses in 2015 [54]. The population was composed

of two ethnic groups; 93.1% of the residents self-identified as Qom people and the remaining minority as creoles [54]. A household was defined as all the people who occupy a housing unit, including related and nonrelated family members [55]. The housing unit (or house compound) included one or more domiciles (i.e., sleeping quarters) and associated peridomestic structures.

### ***Study design***

This study complied with STROBE recommendations for observational studies [56] (S1 Text). We integrated vector and epidemiological information collected during the 2008-2015 period during the annual vector and household surveys as described in detail elsewhere [40,41,54]. At each visit, the location of the domicile was georeferenced with a GPS receiver (Garmin Legend) and the demographic information of the household updated. The household surveys included relevant epidemiological data such as number, type and construction materials of the different sites in each housing unit, refuge availability for the vector, the number and resting places of domestic animals, number of house occupants, and their ages and ethnicity. In 2012-2015, we carried out a thorough census of the population, which included several social variables identified as social determinants of health and relevant to NTD transmission, such as educational level, income source, household mobility, access to health services, etc [54]. Response rates were high (>90%) in all surveys [40].

Four years after community-wide insecticide spraying and with a low risk of vector-borne transmission (<1% of infested houses), we conducted a series of serosurveys aimed at total population coverage (older than 9 months old) during the 2012-2015 period. Using a participatory approach, we had meetings with the local population to discuss project tasks and coordinate activities with local healthcare agents and hospital personnel [50]. During 2012-2013, people were recruited passively through the radio, schools and word of mouth; venipuncture was done at schools and at the primary health posts of each large community to draw a blood sample of 3-7 ml. In April 2015, recruitment was done actively and blood draws were conducted coupled to the vector survey. The protocol is described in detail elsewhere [24,50].

## **Ethics statement**

This study complied with the ethical principles included in the Declaration of Helsinki (Ethical Committee "Dr. Carlos A. Barclay", Protocol ref. TW-01-004). All adult subjects provided written informed consent prior to participation in the serosurvey, and a parent or guardian of any child participant younger than 14 years old provided informed consent on the child's behalf. Participants aged between 14 and 18 years old, provided their written informed consent along with consent from a parent or guardian. Informed consent forms were provided in Spanish and translated to Qom language (*Qomlactaq*) orally by primary healthcare agents when needed. Oral consent was obtained for household surveys, including entomological searches, as approved by the IRB and recorded in the data collection sheets as described [41].

## **Serodiagnosis of human infection**

Blood samples were transported to the local hospital where serum samples were obtained by centrifugation at 3,000 rpm for 15 min within the same day. Each serum was allocated in triplicate vials and stored at -20°C. Serum samples were tested using two different ELISA tests based on conventional (Chagatest, Wiener) and recombinant antigens (ELISA Rec V3.0, Wiener) as described elsewhere [50]. Discordant results between both ELISAs were resolved via an indirect immunofluorescence antibody test conducted at the reference center for Chagas disease serodiagnosis (Instituto Nacional de Parasitología Dr M. Fatale Chaben). A person was considered seropositive (i.e., infected) if at least two tests were reactive. The serosurvey results were informed to the participants within 8 weeks of the blood draw, and follow-up of *T. cruzi*-seropositive people was done by physicians from the local hospital as described elsewhere [50], including treatment with benznidazole or nifurtimox of people younger than <18 y.o. and those adults who specifically requested it.

## **Vector-related indices**

All triatomines collected at the baseline vector survey in 2008 were identified taxonomically and the individual infection status with *T. cruzi* was determined by microscope examination of feces [41] or by molecular diagnosis using kDNA-PCR [57], achieving a coverage of 60% of all infested houses. The



occurrence of domiciliary infestation with *T. infestans* was determined by the finding of at least one live triatomine (excluding eggs) through any of the vector collection methods used (i.e., timed-manual searches, bug collections during insecticide spraying operations, and householders' bug collections). The relative abundance of domiciliary *T. infestans* was calculated only for infested houses as the number of live bugs collected by timed-manual searches per 15 min-person per site, as described [41]. The same procedures were used to determine the occurrence and abundance of *T. cruzi*-infected *T. infestans* in the domicile.

### ***Epidemiological data pre-processing***

Human seropositivity data (collected over 2012-2015) were back-corrected to 2008 in order to analyze the risk of human infection prior to the community-wide insecticide spraying campaign when vector-borne transmission was still active, and given the virtual absence of house infestation and of infected triatomines over 2008-2012. This allowed us to compute a more accurate length of exposure to vector-borne infections by subtracting the time elapsed since the baseline vector survey to the age recorded at the serosurvey. Using the demographic data collected in household surveys in 2008-2015 and the location of each household over the different time periods, we identified people's residential location as of 2008. The 2012-2015 detailed census data was also used to determine their mobility pattern, which was classified as: non-movers (individuals whose house remained at the same geographical location), movers (if they moved within the study area), or as in-migrants or out-migrants (to or from outside the study area, respectively, including individuals coming from or leaving to a different section within Pampa del Indio municipality), as described elsewhere [54]. If the in-migrants had resided in the study area prior to 2008, we classified them as return migrants.

Regarding the socio-demographic variables collected in 2012, they were back-corrected to 2008 whenever possible, assuming similar socio-economic conditions prevailed over 2012 and 2008 [41]. Data obtained from the household surveys were used to construct two surrogate indices (the social vulnerability index and the host availability index) by means of multiple correspondence analysis. A detailed description of the indices, their rationale and construction is found elsewhere [54]. Briefly, the social

vulnerability index estimated for 2008 households included characteristics of the domiciles (refuge availability, presence of cardboard roofs and/or mud walls, time since house construction, and domestic area), and household socio-economic and demographic characteristics (overcrowding, goat-equivalent index and educational level). This index aimed at summarizing the multiple interrelated socio-demographic variables associated with socio-economic position. The host availability index in domiciles summarized the number of potential domestic hosts of *T. infestans* (i.e., adult and child residents, total number of dogs, cats, and chickens nesting indoors), and in the case of dogs and cats, whether they rested within or in the proximity of the domicile. A preliminary analysis showed that the household abundance of domestic animal hosts was positively correlated with larger household size, i.e., number of human residents [54].

## **Data analysis**

For each variable we checked whether the missing values were missing completely at random by building a dummy variable (missing and non-missing values) and analyzing the significance of Spearman's correlation coefficient with any another independent variable in the data set, as described elsewhere [41]. Most of the variables with missing values (collected through household surveys) were missing completely at random, except for educational level and overcrowding in 2008, in which the missing data corresponded to households that had moved or out-migrated by 2012 (the year when these data were collected). Missing data for human infection per household were also biased towards adult males (who refused to participate more often than other demographic groups), and movers, as we were not able to assign their residential location in 2008 to all of them. We excluded from the analysis 197 people tested who resided elsewhere but were visiting the study area at the time of the serosurvey. Throughout the text, minors younger than 15 y.o. are referred to as children for simplicity.

For all proportions, 95% confidence intervals (CI<sub>95</sub>) were estimated using the Agresti and Coull method if sample sizes were greater than 50, and the Wilson method for smaller sample sizes [58]. We used  $\chi^2$  tests for bivariate analysis of categorical variables; generalized linear models (GLM) [59] with logit link function when the outcome of interest had a binary distribution, and a negative binomial GLM (link

function: log) when analyzing count data, such as the number of seropositive people. In the case of binary response variables, the relative risk was expressed as Odds Ratios (OR), and in the case of count data as incidence rate ratios (IRR). Negative binomial regression was preferred to Poisson regression given the overdispersed distributions [60]. Mixed-effects models (GLMM) were also considered when individual data was analyzed to account for possible household-related random effects [24,28]. All analysis were implemented in Stata 14.2 [61] and R 3.2.3 (lme4 and car packages) [62].

### ***Infection risk models***

The infection risk model was estimated for the entire population and for children separately. The outcome of interest was their seropositivity status for *T. cruzi* (yes/no). In the model for the total population we included age, gender, ethnicity, and household-level variables: the relative abundance of *T. cruzi*-infected *T. infestans* per unit of catch-effort in the domicile, the social vulnerability index, host availability index, and the number of seropositive co-inhabitants. The model for child infection also included the mother's seropositivity status, and whether any recent insecticide spraying occurred in 2006-2008 prior to the baseline vector survey. We also evaluated the interaction between vector abundance, the social vulnerability and host availability indices, and retained in the model the ones that had a significant effect. We compared a GLMM model (logit link function) considering the household as a random variable and a GLM model in both cases.

We used an information theoretic approach and Akaike's information criterion (AIC) to identify the best-fitting models given the data collected, and a multimodel inference approach to account for model selection uncertainty [47,48] using the MuMin R-package. This approach was used to identify the ecological and social factors associated with house infestation, as described elsewhere [41]. Odds Ratios (ORs) and their 95% confidence intervals were calculated from model-averaged coefficients. The relative importance (RI) of each variable is defined as the sum of Akaike weights in each model in which the variable is present; RI takes values from 0 to 1. Multicollinearity was assessed by the Variance Inflation Factor (VIF), and model fitting for the logistic regression was assessed by the Hosmer-Lemeshow goodness of fit test and the Receiver Operator Curve (ROC), including the area under the ROC (AUC)

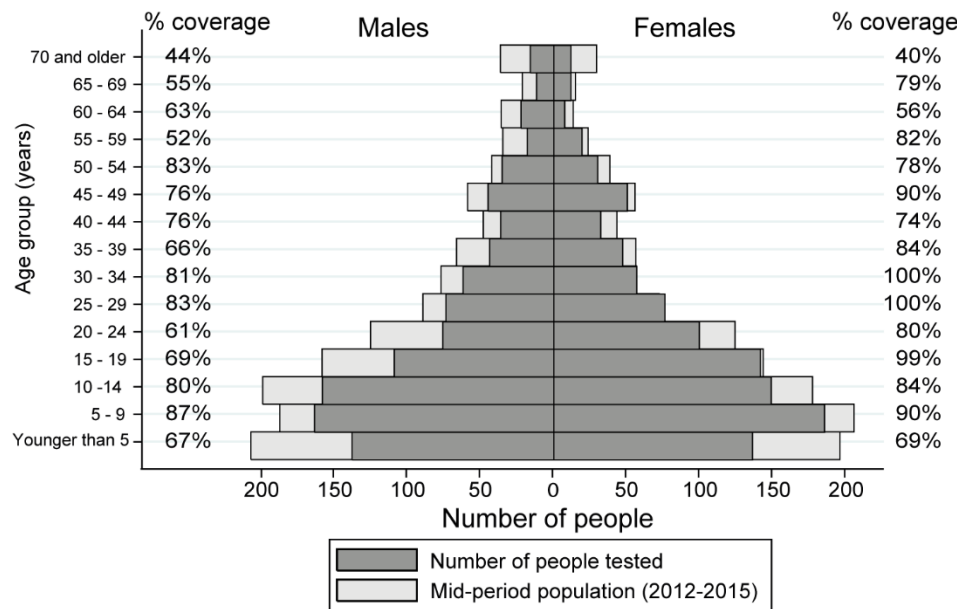
[62]. To assess sensitivity and specificity of the models, we employed an optimal threshold value that minimized the sum of error frequencies [63]. This value was obtained by finding the maximum sum of sensitivity and specificity for all threshold values  $t$  ( $\text{sens}(t) + \text{spec}(t)$ ) using the pROC R-package [64]. The multimodel inference approach does not allow for missing data, therefore individuals with missing data were not considered in the models.

### *Spatial analysis and risk maps*

Global point pattern analysis (univariate and bivariate) were performed for human and vector infection using the weighted K-function implemented in Programita [35]. Random labeling was used to test the null hypothesis of random occurrence of events among the fixed spatial distribution of all households. We used quantitative and qualitative labels for each household as previously described [54]. Monte Carlo simulations ( $n = 999$ ) were performed and the 95% ‘confidence envelope’ was calculated with the 2.5% upper and lower simulations. Additionally, local spatial analysis on the abundance of (infected) vectors were performed using the  $G^*$  statistic implemented in PPA [33]. The selected cell size was 200 m (assuming that each household had at least three neighbors at the minimum distance of analysis), and the maximum distance was set at 6 km (i.e., half of the dimension of the area). We created heatmaps (i.e., density maps) to visualize the spatial aggregation of the variables using a kernel density estimation algorithm within a radius of 200 m as implemented in QGIS 2.18.11. Risk maps were obtained by interpolating the predicted probabilities from the risk model and using the 2008 household location. The geographical coordinates of each housing unit were transformed to preserve the privacy of the households involved in this study, as described elsewhere [41].

### Results

During 2012 and 2015, we tested 1,929 people that resided in the study area at the time of the serosurveys. Passive recruitment during 2012-2013 at local schools and primary health posts led to a coverage of 47.8% of the local population, which increased to 77.0% during 2015, when active recruitment of participants was coupled with vector surveys (S1 Fig). Diagnosis coverage was highest in school children (80.9%) and adult women (76.8%) in general (Fig 1).

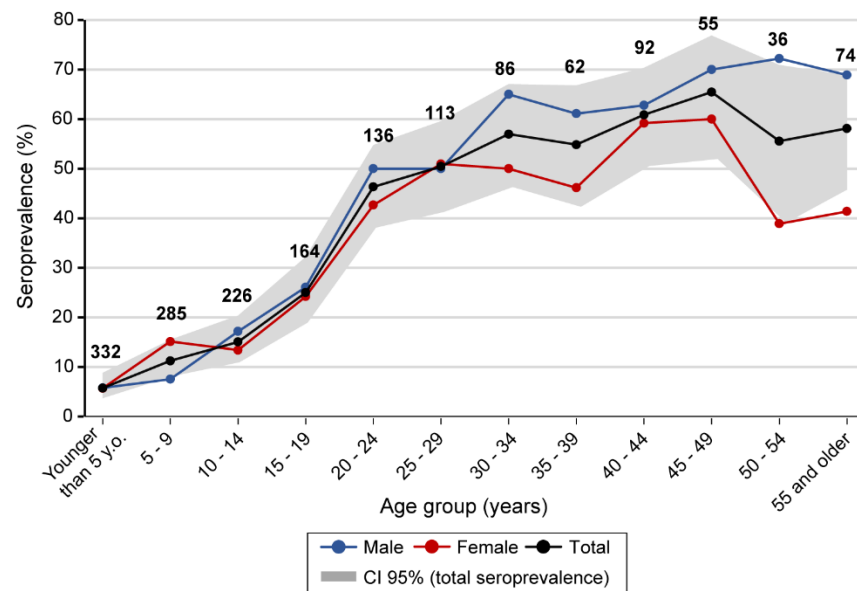


**Fig 1.** The 2012-2015 mid-period total population per 5-year age group and the population serologically tested for *T. cruzi* infection (number and percentage) in Pampa del Indio, Argentina.

The two-tiered ELISA serological testing showed almost perfect agreement between tests (kappa index = 0.9,  $p < 0.001$ ). Seropositive results for *T. cruzi* infection were observed in 25.3% of the samples, whereas 5.5% that were discordant later tested negative at the reference laboratory. Only 1.5% ( $CI_{95} = 4 \times 10^{-4} - 3.0\%$ ) of children born after the community-wide insecticide spray in 2008 were seropositive. All four *T. cruzi*-seropositive children born post-intervention had *T. cruzi*-seropositive mothers and no vectors were collected in their domiciles during the surveillance phase. For the same age group (children  $< 6$  y.o.), the risk of infection was higher if they had been born before the community-wide spraying (logistic regression,  $OR_{2008} = 6.2$ ,  $CI_{95} = 1.2-31.8$ ,  $p = 0.01$ ), after adjusting for age. Only 1.4% of the people tested had been exposed to *T. infestans* due to transient domiciliary reinfestations during the surveillance period, but all 38 also lived in infested houses prior to the 2008 insecticide campaign.

We were able to assign the local residency status as of 2008 (i.e., their house identification code in 2008) to 82.6% of the people tested. The overall seroprevalence estimated for 2008 was 29.0% ( $CI_{95} = 26.7-31.4$ ,  $n = 1373$ ). It increased significantly with age from 5.7% in children younger than 5 y.o. to 25% in teenagers (14 to 19 y.o.), then jumped to 50% in young adults (20 to 29 y.o.), and remained around 60%

in older adults (Fig 2). Although females had a lower overall seroprevalence rate than males (26.5 vs 31.6%;  $\chi^2$  test,  $df=1$ ,  $p=0.04$ ), this difference was more evident in adults (Fig 2) and was not significant after adjusting for age (S1 Table). The seroprevalence for Qom people almost doubled that observed for creoles (29.7 vs 18.7%;  $\chi^2$  test,  $df=1$ ,  $p=0.02$ ) (S1 Table).



**Fig 2.** Seroprevalence of *T. cruzi* infection per 5-year age group and gender for the population in 2008, prior to a community-wide insecticide spraying campaign, Pampa del Indio, Argentina. The numbers above the lines indicate the total number of people serologically tested.

### ***Household infection and domiciliary infestation***

We found at least one *T. cruzi*-seropositive person in 71.3% of the 301 households in which at least one member had been tested (77% of all households existing in 2008). For children, that proportion dropped to 16.7%. The occurrence of at least one seropositive person per household was not associated with the abundance of *T. cruzi*-infected *T. infestans* per unit effort in the domicile (logistic regression, OR = 1.5,  $CI_{95} = 0.8-3.0$ ;  $p=0.2$ ). In contrast, the occurrence of at least one infected child was positively and highly significantly associated with the abundance of infected vectors (logistic regression, OR = 2.8;  $CI_{95} = 1.6-4.9$ ;  $p < 0.001$ ). The household number of seropositive people increased with the abundance of infected vectors (negative binomial regression, IRR = 1.1;  $CI_{95} = 1.1-1.2$ ;  $p < 0.001$ ).

The seroprevalence of *T. cruzi* was higher in local residents who reported past exposure to *T. infestans* at the time of the serosurvey (35.3 vs 15.0%,  $\chi^2$  test, df = 1,  $p < 0.001$ ). There was a fair agreement between the house infestation status assessed in 2008 and reported past exposure to triatomines (kappa index = 0.2;  $p < 0.001$ ). Although 79% of residents in an infested house reported past exposure to the vector, a considerable proportion (54%) of those who inhabited non-infested houses also reported past exposures. Therefore, the individual infection risk adjusted for age, gender and ethnic group was two-fold higher in infested houses as determined in the 2008 (baseline) vector survey (GLMM, OR = 2.3, CI<sub>95</sub> = 1.4-3.9,  $p < 0.001$ ), but the effect was marginally significant when the reported past exposure was considered instead (GLMM, OR = 1.5, CI<sub>95</sub> = 1-1.2,  $p = 0.6$ ) (S2 Table). In children, the infection risk in infested houses was almost four-fold higher (GLMM, OR = 3.7, CI<sub>95</sub> = 1.3-10,  $p = 0.01$ ) even after adjusting for the mother's seropositivity status, while the reported past exposure had no significant effect (GLMM, OR = 0.9, CI<sub>95</sub> = 0.4-2.4,  $p = 0.9$ ) (S2 Table). When vector abundance per unit effort was considered, the infection risk in children increased 42% per 5 additional *T. infestans* collected in the domicile (GLMM, OR = 1.42, CI<sub>95</sub> = 1.2-1.7,  $p = 0.001$ ).

### ***Household mobility and human infection***

Movers (as registered in 2012-2015) had a 30-110% higher risk of residing in a *T. infestans*-positive house in 2008 than non-movers after adjusting for age, gender and ethnic group (logistic regression, OR = 1.7; CI<sub>95</sub> = 1.3-2.1;  $p < 0.001$ ). Out-migrants had a much lower risk (OR = 0.5; CI<sub>95</sub> = 0.3-1.3;  $p < 0.01$ ). These findings proved valid in children, whose infection status is more closely associated with recent domiciliary infestation. Although child infection was not directly associated with their mobility pattern, the effect of domiciliary infestation on infection risk was modified by mobility: in children living in infested houses in 2008, the effect of infestation on risk decreased by 90% in movers compared to non-movers (Table 1). While non-mover children residing in infested houses in 2008 had a six-fold higher risk of infection compared to non-mover children living in non-infested houses (20.0% vs. 4.5%, respectively) (Table 1), mover children living in infested houses had a nearly equal seroprevalence as those living in non-infested houses (11.9% vs. 11.1%) (Fig 3).

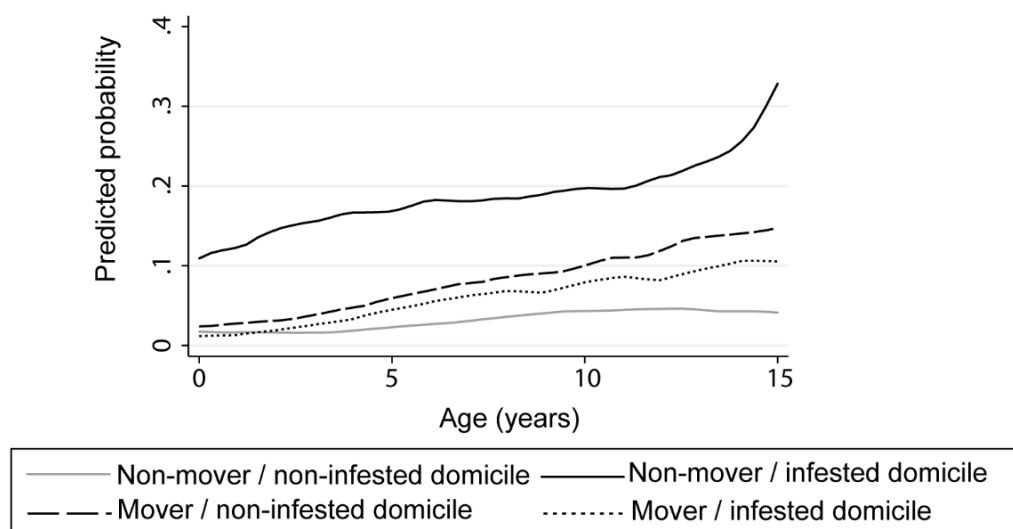
| Variables                             | OR (CI <sub>95</sub> ) | P        |
|---------------------------------------|------------------------|----------|
| <b>Infestation (non-movers)</b>       |                        |          |
| No                                    | 1                      |          |
| Yes                                   | 6.1 (1.9-19.0)         | 0.002*   |
| <b>Age</b>                            | 1.14 (1.03-1.3)        | 0.01*    |
| <b>Gender</b>                         |                        |          |
| Male                                  | 1                      |          |
| Female                                | 1.5 (0.7-3.2)          | 0.3      |
| <b>Ethnic group</b>                   |                        |          |
| Creole                                | 1                      |          |
| Qom                                   | 7.1 (0.4-139.0)        | 0.2      |
| <b>Mother infection status</b>        |                        |          |
| No                                    | 1                      |          |
| Yes                                   | 8.0 (2.5-25.0)         | <0.001** |
| <b>Mobility (2012-2015)</b>           |                        |          |
| Non-movers (non-infested houses)      | 1                      |          |
| Movers (non-infested houses)          | 1.8 (0.4-8.5)          | 0.5      |
| Out-migrants (non-infested houses)    | 2.2 (0.2-20.0)         | 0.5      |
| Return migrants (non-infested houses) | 0                      | 1        |
| <b>Interactions§</b>                  |                        |          |
| Movers*infested houses                | 0.10 (0.01-0.99)       | 0.05*    |
| Out-migrants*infested houses          | 0                      | 1        |
| Return migrants*infested houses       | 0                      | 1        |

§ The reference category is non-movers living in infested houses

OR: Odds ratio; CI<sub>95</sub>: 95% confidence interval

\*\*  $p < 0.001$ ; \*  $0.001 \leq p \leq 0.05$ ;  $\sim 0.5 < p < 1$

**Table 1.** Generalized linear mixed model of child seropositivity for *T. cruzi* infection in 2008 clustered by household (logit link function), Pampa del Indio, Argentina. The odds ratio (OR) of the interaction terms indicate if the effect of infestation for non-movers increased (>1) or decreased (<1) in other mobility categories (out-migrants or return migrants).

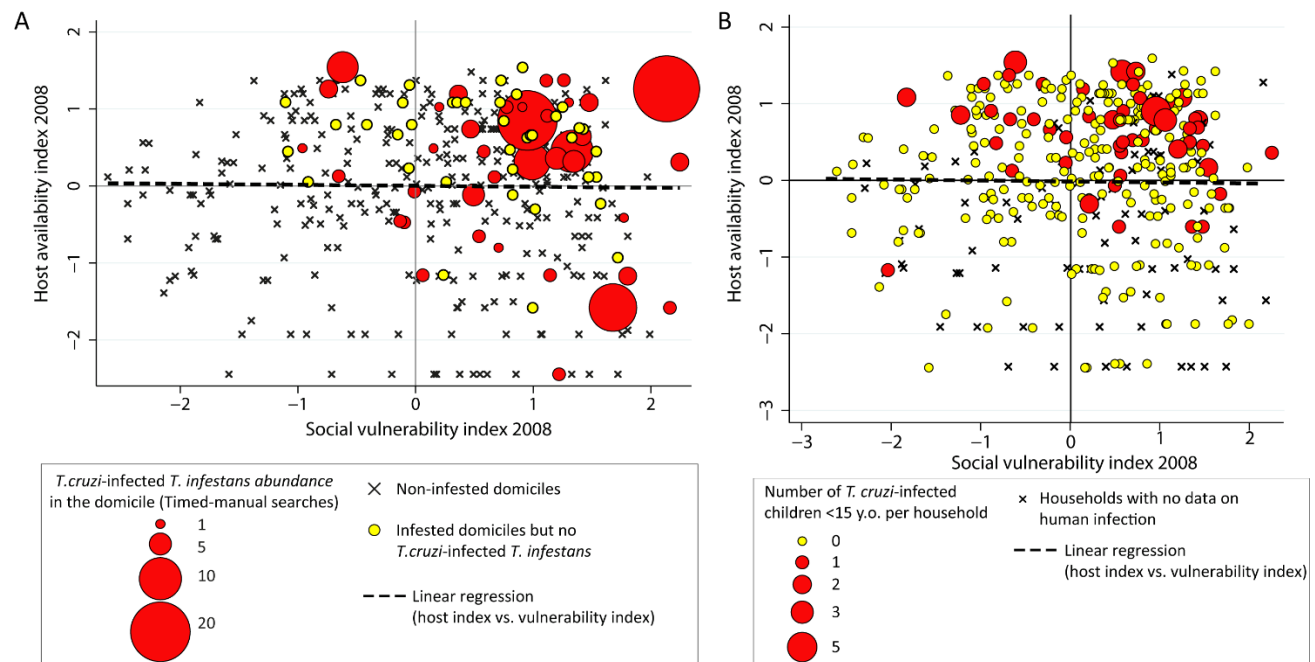




**Fig 3.** Predicted probabilities for the GLMM model of child seropositivity for *T. cruzi* infection regarding the interaction between domiciliary infestation and mobility patterns (non-movers vs. movers), Pampa del Indio, Argentina.

### *Household socio-demographic determinants of human infection*

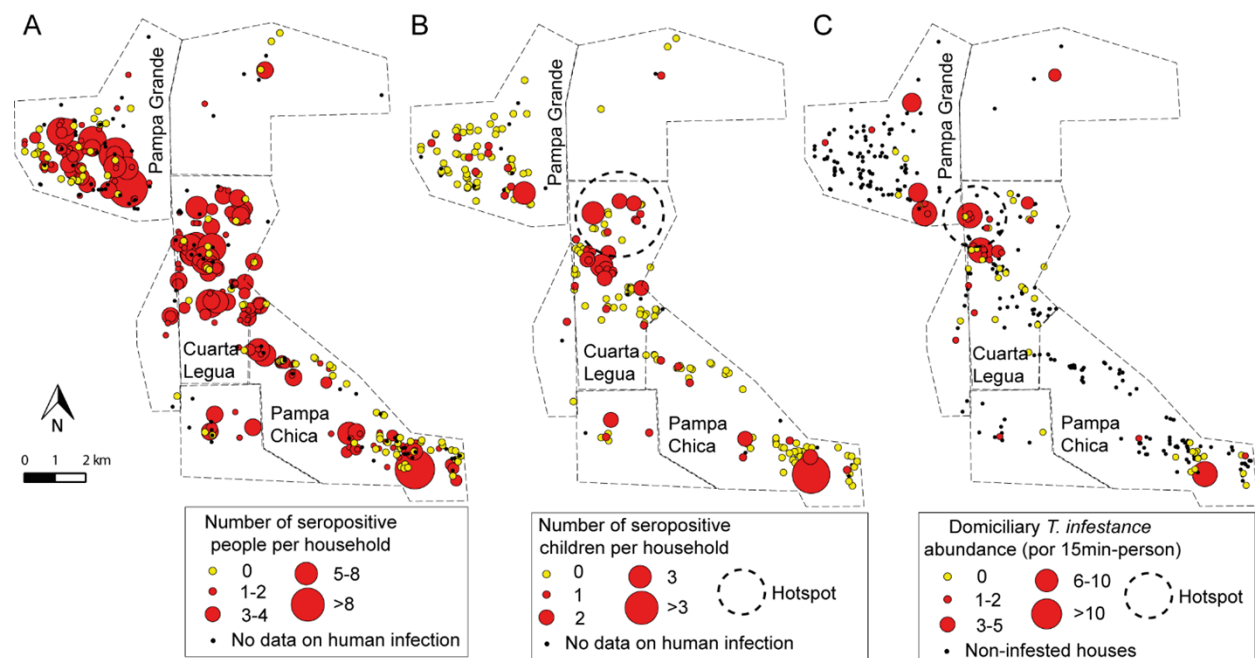
The occurrence of at least one seropositive person per household was positively associated with social vulnerability (OR = 1.5; CI<sub>95</sub> = 1.2-1.8; p = 0.001) and was independent of host availability (OR = 1.3; CI<sub>95</sub> = 0.9-1.7; p = 0.1). The household presence of at least one seropositive child occurred more frequently in houses with high host availability (OR = 2.1; CI<sub>95</sub> = 1.4-3.4; p < 0.001) and high social vulnerability (OR = 1.4; CI<sub>95</sub> = 1.0-2.0; p = 0.05) (Fig 3). In both cases, the effects of the socio-demographic variables were additive, as the interaction terms were not significant. The occurrence and abundance of *T. cruzi*-seropositive children were similar to the distribution of infected domiciliary *T. infestans*, which also increased with host availability and social vulnerability (Fig 4).



**Fig 4.** Multivariate relationship between household-level host availability, social vulnerability and vector indices [54] (A) and child seropositivity for *T. cruzi* infection (B), Pampa del Indio, Argentina.

# **Spatial distribution of human infection and house infestation**

Despite the widespread occurrence of at least one *T. cruzi*-seropositive person per household, human infection was spatially aggregated in 2008 (Fig 5A). The global spatial analysis indicated that households with at least one seropositive person were aggregated at a scale between 2 and 6 km (S2A Fig), implying that up to 2 km (i.e., within each community) the spatial distribution was not significantly different from a random pattern. This result is consistent with the differences among communities in the proportion of households with at least one seropositive person: 84.6% in Cuarta Legua, 69.4% in Pampa Grande and 56.8% in Pampa Chica ( $\chi^2$  test,  $df = 2$ ,  $p < 0.001$ ). When the number of seropositive people per household was considered, we did not find a significant spatial correlation between neighboring houses at any scale. Although the trend shown by the variograms indicated similar numbers of seropositive people between neighboring households, this pattern was not significantly different from a random distribution (S2B Fig).



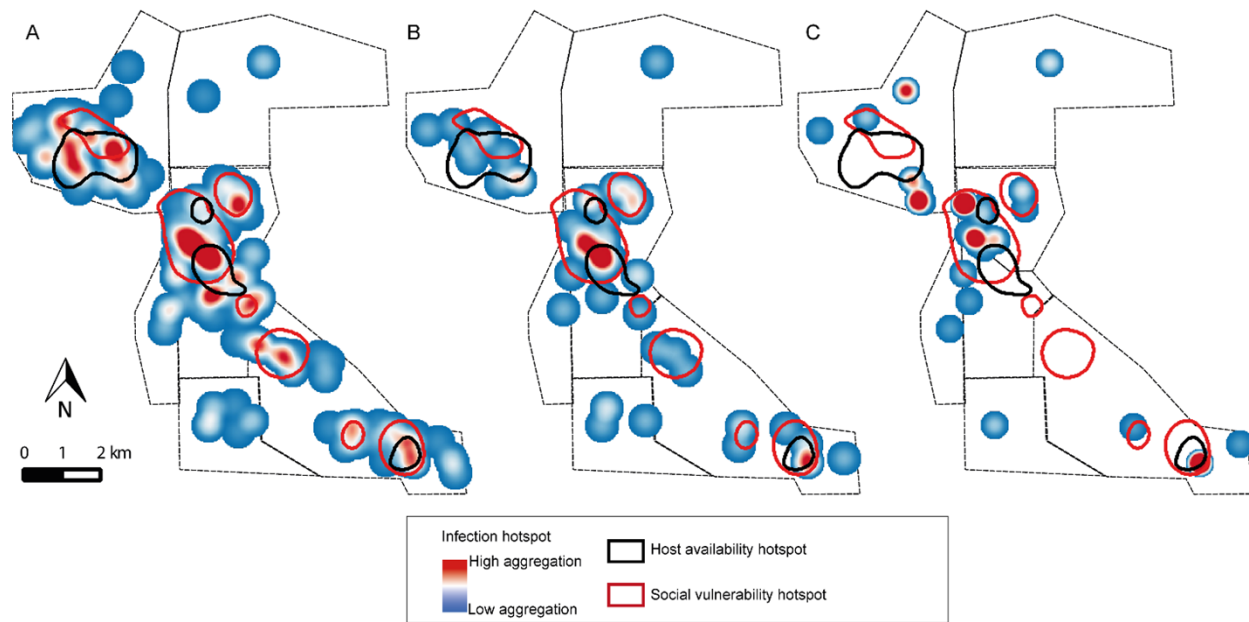
**Fig 5.** Distribution of the household number of *T. cruzi*-seropositive people (A) and children (B), and the relative abundance of infected *T. infestans* per unit of effort (C) prior to community-wide insecticide spraying in 2008, Pampa del Indio, Argentina. The dashed circles represent the hotspot found by local

spatial analysis. The maps were created in QGIS 2.18.11. based on the data collected within the scope of this study (S3 Table).

The proportion of houses with at least one seropositive child in 2008 also varied significantly between communities ( $\chi^2$  test;  $df = 2$ ;  $p < 0.01$ ). The highest proportion was found in Cuarta Legua compared to the other communities (32.9% vs. 14.5-13.8%, respectively), but no spatial aggregation was found globally (S2C Fig). The number of infected children per household did not show a spatial correlation between neighboring households either (S2D Fig). However, local spatial analysis found a hotspot of child infection located in Cuarta Legua at a spatial scale between 1 and 2.6 km (Fig 5B). This hotspot was no longer detected in 2012-2015, nor did the spatial aggregation per community for the total population (S3 Fig).

The occurrence of at least one *T. cruzi*-infected *T. infestans* in the domicile was globally aggregated at all spatial scales when considering all houses (S2E Fig), but it did not differ from a random distribution when only infested houses were considered (S2F Fig). Nonetheless, a hotspot of domiciliary infected vectors was found in Cuarta Legua at a spatial scale between 0.2-1.8 km , which involved 6 houses with at least one infected vector (Fig 5C). This hotspot coincided partially with the child infection hotspot, but no significant correlation was found between human and vector infection patterns (S2G-H Fig).

The heatmaps show that the areas of high social vulnerability coincided with the qualitative and quantitative hotspots of human and vector infection (Fig 6). However, no significant spatial correlation was found between human or vector infection and socio-demographic variables (social vulnerability and host availability) (S4 Fig).



**Fig 6.** Heatmaps of the household number of *T. cruzi*-seropositive people (A) and children (B), and the relative abundance of *T. cruzi*-infected *T. infestans* per unit of effort (B) in 2008, prior to a community-wide insecticide spraying campaign, Pampa del Indio, Argentina. The areas of qualitative aggregation (as determined by heatmaps) of social vulnerability and host availability are also shown. The maps were created in QGIS 2.18.11. based on the data collected within the scope of this study (S3 Table).

### Human infection risk model

Human infection was significantly clustered by household, after adjusting for the abundance of infected vectors, household social vulnerability and host availability ( $\beta_{\text{household}} = 1.3$ ,  $CI_{95} = 1.02-1.7$ , log-likelihood ratio test,  $p < 0.001$ ). When we included the number of seropositive co-inhabitants, the random effect of the household was nil and no significant differences were found between the GLMM and the respective GLM model (log-likelihood ratio test,  $p = 1$ ), indicating that all variables included in the model accounted for the differences among households.

The risk of *T. cruzi* infection increased with age and if they were Qom, but did not vary by gender (Table 2). The ethnic group did not have a significant effect on child infection possibly because only one of the 70 seropositive children was creole. The risk of infection in the total population significantly increased by

60% for each additional infected *T. infestans* collected in the domicile and by 40% with each additional infected co-inhabitant; increased with increasing household social vulnerability, and decreased with increasing host availability index (Table 2).

| <i>Model 1: population model</i>             |     |                        |          | <i>Model 2: children</i>                       |     |                        |        |
|--|-----|------------------------|----------|--|-----|------------------------|--------|
| Variable                                     | RI  | OR (CI <sub>95</sub> ) | P        | Variable                                       | RI  | OR (CI <sub>95</sub> ) | P      |
| <b>Age (years)</b>                           | 1   | 1.07 (1.06-1.08)       | <0.001** | <b>Age (years)</b>                             | 1   | 1.1 (1.04-1.3)         | 0.008* |
| <b>Infected vector abundance (1)</b>         | 1   | 1.6 (1.2-2.3)          | 0.003*   | <b>Infected vector abundance (1)</b>           | 1   | 2.2 (1.001-4.8)        | 0.04*  |
| <b>Social vulnerability (2)</b>              | 1   | 1.5 (1.2-1.8)          | <0.001** | <b>Social vulnerability (2)</b>                | 0.9 | 1.3 (0.8-2.1)          | 0.2    |
| <b>Interaction (1) * (2)</b>                 | 1   | 0.7 (0.6-0.9)          | 0.01*    | <b>Interaction (1) * (2)</b>                   | 0.9 | 0.5 (0.2-0.98)         | 0.04*  |
| <b>Host availability</b>                     | 1   | 0.7 (0.6-0.9)          | 0.003*   | <b>Host availability</b>                       | 0.9 | 0.5 (0.2-0.9)          | 0.02*  |
| <b>Number of seropositive co-inhabitants</b> | 1   | 1.4 (1.3-1.6)          | <0.001** | <b>Seropositive mother</b>                     | 0.9 |                        |        |
| <b>Ethnic group</b>                          | 0.9 |                        |          | No   |     | 1                      |        |
| Creole                                       |     |                        |          | Yes  |     | 3.8 (1.4-11)           | 0.009* |
| Qom  |     | 2.3 (1.1-4.6)          | 0.02*    | <b>Gender</b>                                  | 0.7 |                        |        |
| <b>Gender</b>                                | 0.4 |                        |          | Male   |     | 1                      |        |
| Male   |     | 1                      |          | Female   |     | 2 (0.96-4.1)           | 0.06~  |
| Female                                       |     | 0.8 (0.6-1.1)          | 0.2      | <b>Number of seropositive co-inhabitants</b>   | 0.6 | 1.3 (0.96-1.9)         | 0.08~  |
|  |     |                        |          | <b>Recent insecticide spraying (2006-2008)</b> | 0.5 |                        |        |
|  |     |                        |          | No   |     | 1                      |        |
|  |     |                        |          | Yes  |     | 0.5 (0.2-1.3)          | 0.2    |
|  |     |                        |          | <b>Ethnic group</b>                            | 0.3 |                        |        |
|  |     |                        |          | Creole   |     | 1                      |        |
|  |     |                        |          | Qom  |     | 1.4 (0.2-12.5)         | 0.7    |

\*\*  $p < 0.001$ ; \*  $0.001 \leq p \leq 0.05$ ; ~  $0.5 < p < 1$

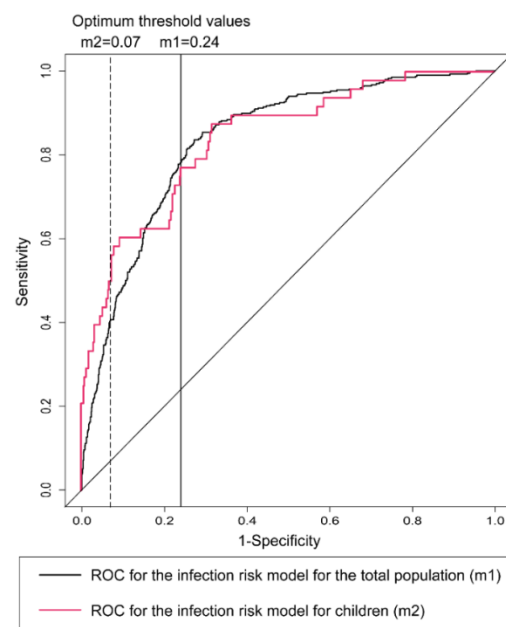
OR: Odds ratio; RI: relative importance of the variable; CI<sub>95</sub>: 95% confidence interval

**Table 2.** Generalized linear model (logit link function) of human seropositivity for *T. cruzi* infection for the total population (model 1) and children (model 2), Pampa del Indio, Argentina.

In the risk model for children, the positive effect of infected domestic vector abundance and the negative effect of host availability still held, but social vulnerability was no longer significant (Table 2). Including the mother's seropositivity status (which was significantly and positively associated with child infection) led to a marginal effect of the number of seropositive co-inhabitants (Table 2). The recent partial insecticide spraying did not modify child infection risk (Table 2); it rather had a protective effect if the abundance of infected vectors was removed from the model (OR = 0.4, CI<sub>95</sub> = 0.1-0.9, p = 0.03),

indicating that recent insecticide treatment reduced the infection risk by decreasing vector abundance. In both models, we observed a significant negative interaction between social vulnerability and infected vector abundance; this indicates that the latter's effect on infection risk was lower in households with higher social vulnerability (Table 2). Thus, people in households with high social vulnerability and low vector abundance have a similar infection risk as those in less vulnerable households but with higher vector abundance.

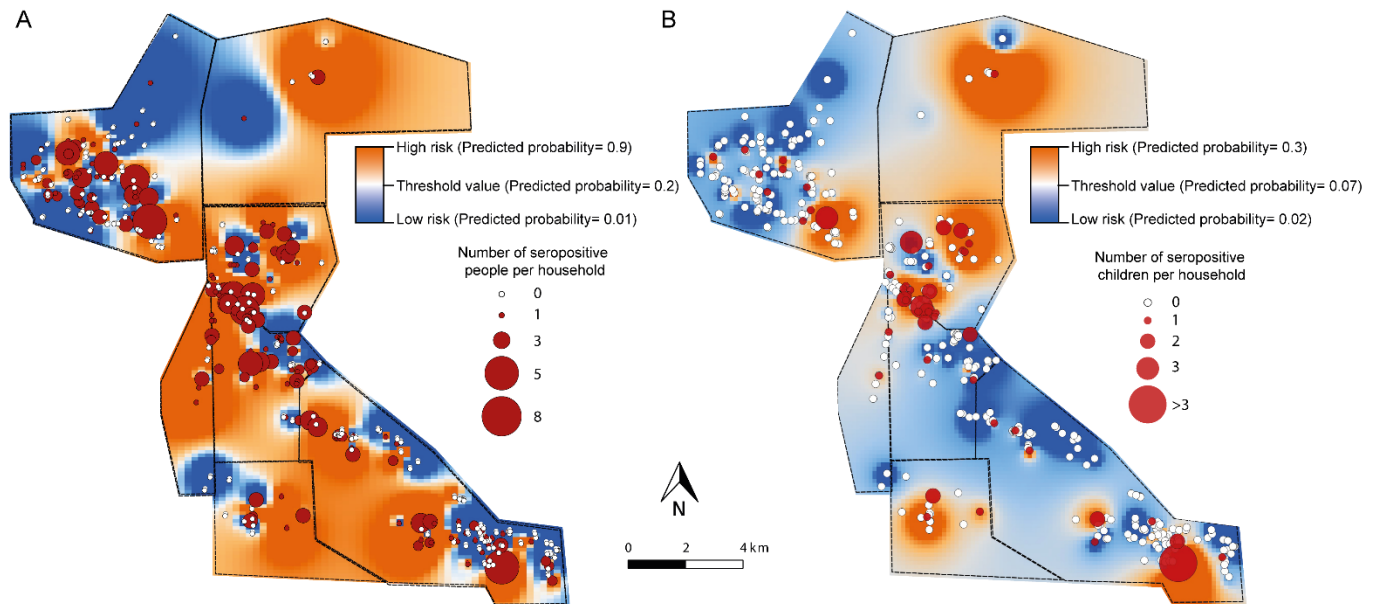
In both models,  $VIF < 2$  indicated no multicollinearity issues. The infection risk model for the total population had  $AUC = 0.83$ , with a sensitivity of 83% and a specificity of 72% (Fig 7). In a post-hoc classification, the model predicted 19.3% of false positives and 4.7% of false negatives. The model fit the data poorly according to the Hosmer-Lemeshow test ( $p < 0.001$ ), perhaps due to the lack of saturation of the model, or non-linear relationship between variables, or limits to the goodness-of-fit test due to a large number of cases. By contrast, the risk model in children had a good fit to the data (Hosmer-Lemeshow test,  $p = 0.6$ ); the AUC was 0.84 (Fig 7), and it had higher sensitivity (87%) and lower specificity (68%) than the previous model. It predicted 28.8% of false positive cases but only 1.1% of false negatives.



**Fig 7.** Receiver operating characteristic curves (ROC) of the *T. cruzi* infection risk models for the total population and children. The vertical lines show the optimum threshold value for classification (i.e. maximizes sensitivity and specificity).



The risk maps derived from the interpolation of the risk model documented the heterogeneous distribution of human infection and the occurrence of high-risk areas, which were more widespread when the total population was considered (Fig 8A) compared to the children risk map (Fig 8B). Although they failed to include a few households with seropositive children, most households with human cases were encompassed in these high-risk areas.



**Fig 8.** Risk maps of *T. cruzi* human infection for the total population (A) and children (B) in 2008, prior to community-wide insecticide spraying, Pampa del Indio, Argentina. The maps were created in QGIS 2.18.11. based on the data collected within the scope of this study (S3 Table).

## Discussion

This study shows the multivariate relationship among the ecological, demographic and socio-economic factors affecting human *T. cruzi* infection in endemic rural communities with active vector-borne transmission. To the best of our knowledge, this may be the first study to comprehensively assess their combined effects and interactions on human *T. cruzi* infection. In addition to corroborating the effects of demographic variables (age, ethnic group, among other) and the relative abundance of infected *T. infestans* on human infection risk [24], we found significant effects of socio-economic position after adjusting for ecological factors. Individual human infection risk significantly increased by 60% for each

additional infected vector collected by unit of effort in the domicile and by 40% with each infected co-inhabitant; increased with increasing household social vulnerability, and decreased with increasing domestic host availability (numbers of indoor-resting bird or mammal hosts). A novel finding was the negative interaction between infected-vector abundance and social vulnerability, which indicates that members of more vulnerable households had an increased infection risk even at low infected-vector abundances. Household mobility also modulated infection risk: internal mobility (within the study area) reduced the effects of domiciliary vector abundance, possibly due to less consistent exposures among movers. Nonetheless, the seroprevalence rates of movers and non-movers were not significantly different. Although transmission was clustered by household, the spatial analysis showed that transmission risk was heterogeneous within the study area, with hotspots of human and vector infection matching areas of higher social vulnerability.

The Gran Chaco eco-region is a hotspot for Chagas disease and other NTDs that disproportionally affect vulnerable rural communities and certain demographic groups [11,65]. Like other NTDs, the highly focal distribution of Chagas disease cases is determined by particular combinations of various ecological and social determinants [65,66]. The baseline (2008) human seroprevalence here estimated (29%) was considerably lower than that registered in adjacent rural communities (40%) of Pampa del Indio [24]; in other indigenous (59-71%) and creole (40-62%) communities living in more disadvantaged, isolated areas of Chaco province known as “The Impenetrable” [17,18,21], and in the Bolivian Chaco (40-80%) [31,67]. Our seroprevalence estimate is similar to other estimates (28-31%) registered elsewhere in the Argentine Chaco [16,19]. This heterogeneous distribution of human infection within Chaco province was evident in the 1980s when the intensity of vector-borne transmission was much higher [68], and remains evident almost three decades later.

The heterogeneous distribution of human infection was also captured within the studied communities through spatial analysis. Human cases, especially among children, were aggregated (Fig 5, S3 Fig) and matched the spatial aggregation of house infestation [41], vector infection and high social vulnerability (this paper). This pattern is consistent with vector-borne transmission mainly occurring at a household



level [28,69], as human infection risk was mostly explained by household-level factors. Household clustering of cases was both revealed when we included the household as a random variable, and when the number of infected co-inhabitants was included in the human risk model. In the child infection model, household clustering was closely linked to the mother's seropositivity status (allowing for putative congenital transmission), and to being under similar vector exposures [29,69].

The strong and positive association between human infection and the abundance of *T. cruzi*-infected vectors indicated active vector-borne transmission prior to the community-wide insecticide spraying we implemented. The occurrence of at least one seropositive child per household was significantly and positively associated with the abundance of infected *T. infestans* in the domicile, consistent with a more recent infection. Compared to the only previous study that quantified the effect of infected-vector domestic abundance on child infection risk, the magnitude we found was half of that estimated in rural communities of the dry Argentine Chaco two decades earlier (two- versus four-fold increase per each additional infected vector) [28]. However, the partial insecticide spraying undertaken within two years before the baseline vector survey, coupled with intense household mobility patterns [54] and slowly improving housing conditions [40], likely deflated the quantitative relationship between human and vector infection. Other studies reported a significant positive association between human infection and the presence of at least one infected domestic triatomine in rural communities [24,30,69], but not in a periurban community where child infection was instead associated with peridomestic infection in *T. infestans* [29]. In our study, because the housing units only had a few peridomestic sites [41], most of the infestations occurred in the domicile where human-vector contact with *T. infestans* is most likely to occur. Host availability is an ecological variable (i.e., number and density of host-feeding sources) closely related to household size. Here, we used a summary index to represent domestic host availability because usually there is a positive and significant correlation between household size and the number of domestic dogs or cats and chicken indoors [e.g., 54]. The occurrence of at least one seropositive child was associated positively with the domestic host availability index, as children were more likely to occur in larger households, and if other dimensions of risk are equal, there would be a greater chance of detecting

at least one infected child in a larger household in which all members were screened. However, when the household number of infected (human) co-inhabitants was added to the individual risk model, the host availability index shifted from exerting positive effects to negative effects on human infection risk (weaker among children). Although there were no multicollinearity issues, understanding the effects of each type of host on the risk of human infection is hindered by the positive correlation between household size and the number of non-human domestic hosts, which pools key reservoir hosts (dogs and cats) with insusceptible hosts (chickens). Even though the host availability index exerted negative effects on the individual risk of infection, this result needs to be interpreted with caution since household size or domestic host abundance sometimes, but not always, were positively associated with larger domestic vector abundance [70,71]. The species identity of non-human hosts is indeed a key factor: one study found that child infection risk increased with the household number of dogs but decreased with the number of chicken indoors [28], whereas in another study human infection risk increased only with 2 or more infected dogs per household [24]. Unfortunately, our study lacks information on the distribution of *T. cruzi* infection in dogs and cats.

The household social vulnerability index used as a surrogate of socio-economic status represents a multidimensional measure of poverty, which we found associated with the domestic abundance of infected *T. infestans* [54]. Herein, we also found significant and positive effects of social vulnerability on human infection risk. To this aim we elaborated a comprehensive measure of socio-economic status; previous studies have focused on a few proxy variables: the goat-equivalent index [24], domicile's characteristics (i.e., construction materials, size, cracks in walls, etc.) [28,31,32], and educational level [72]. The social vulnerability index summarized all of the above and also included household overcrowding, agricultural practices, at least one member with a stable source of income by employment, and the household number of welfare recipients [54]. The social vulnerability index also modified the effects of infected-vector abundance on human infection risk: the magnitude of this effect decreased as social vulnerability increased. This could be explained by several non-exclusive reasons: (i) the social vulnerability index in part may indicate a past, undetected vector infestation prior to the baseline survey

in 2008 (i.e., past vector exposure); (ii) households with higher social vulnerability had greater local mobility [54], suggesting putative vector exposures elsewhere before the insecticide spraying campaign, and (iii) more frequent human-vector contacts given the smaller domicile and greater overcrowding in households with higher social vulnerability. Social vulnerability may also affect household resilience to face house infestation and the consequences of vector exposure. The access of vulnerable households to health services (for diagnosis and treatment, for example) was lower; they were less likely to use domestic insecticides [54] possibly due to economic constraints, and depended on government-sponsored vector control actions.

Household-level variables combined with ethnic group and age determined individual infection risk. In particular, ethnic group was a key demographic factor in stratifying infection risk; Qom people had a two-fold higher risk compared to creoles. Indigenous communities in the Argentine Chaco sometimes displayed higher infection risk than creole communities [17,21], but this pattern was not verified in rural communities in Pampa del Indio with a majority of creoles [24]. The effects of ethnic group on human infection, after adjusting for other social and ecological factors that co-vary with ethnicity, may be explained by differential mobility patterns between Qom and creole households [40,54]. The intense household mobility and the high rates of house turnover among the Qom most likely exerted negative impacts on *T. infestans* populations [41]. Movers were likely exposed to lower and more variable vector abundance over time, and also had higher social vulnerability and a higher risk of occupying an infested house at baseline (preintervention) [54,73]. Complex vector exposure patterns created by human mobility were reflected in the statistically significant interaction between domiciliary infestation and mobility in the child infection risk model. Although infection prevalence was not significantly different between movers residing in infested *versus* non-infested houses at baseline, domiciliary infestation had strong effects on non-movers' infection risk. The predicted infection probability was highest for non-mover children in infested houses, indicating prolonged vector exposure. The impact of human mobility was also evident when the aggregated spatial distribution of human cases at baseline (2008) was compared with the random pattern recorded in 2012-2015. This particular trait of the Qom population needs to be considered

when designing vector control actions and disease screening strategies, given that vector indices alone would only detect a fraction of the *T. cruzi*-seropositive children with internal mobility. Levy et al. (2007) [29] proposed a two-step screening strategy for *T. cruzi*-infected children based on an individual infection risk model adapted to periurban areas of Arequipa (Peru) with a history of recent vector-borne transmission. Their best model included age, *T. infestans* abundance in the domicile, and presence of at least one infected vector. They proposed first to screen children by their age and domestic vector abundance, followed by screening children living in neighboring houses up to the distance of spatial aggregation [29]. The model proposed in this study had a higher diagnostic ability (AUC = 0.84 vs 0.72) and high sensitivity, which would allow to detect most of the infected children. Serological screening strategies that also include household socio-economic status and demographic characteristics, and if possible, the mother's seropositivity status, are expected to perform better at the expense of added cost in obtaining such data. Such costs may be deflated by using the information for an integrated interventions across NTDs.

One limitation of this study was the temporal lag between the baseline vector survey (2008) and the human serosurveys (2012-2015), which required to retrospectively assign the examined individuals to their 2008 residential location. Thus, missing data for human infection increased as a result of considerable human mobility. Assignment errors were minimized by combining data from multiple sources and checking them with local primary healthcare agents, which led to a successful assignment of 80% of the population recorded in 2015. Human mobility patterns were only registered in 2012-2015; we assumed that this pattern was constant over time (i.e., movers in 2012-2015 also behaved as movers in 2008) based on the observation that movers did it more than once [54].

### ***Implications for disease control***

The London Declaration Goals launched in 2012 proposed milestones to eliminate NTDs by 2020, including Chagas disease [74]. Access to diagnosis and treatment is one of the remaining challenges for sustainable control of Chagas disease in endemic areas [50,74]. Chagas disease national estimates, resonated by WHO [9], may fail to approximate the true magnitude of the problem [74]. Integrating the

ecological and social determinants of human infection with the spatial component allowed us to identify individuals, households and geographic sectors at higher risk of infection. This information is key to design cost-effective serological screening strategies in resource-constrained areas for treatment of *T. cruzi*-infected people. Although our analysis took advantage of the detailed data collected during household surveys and census, the same approach could be done using national census data though probably at the coarser spatial resolution that the data would allow. Our infection risk model was also translated into an actionable tool (i.e., risk maps) that identified high-risk areas for local human infection and can point to higher-priority sections for targeted interventions oriented to suppress house (re)infestations, treat infected children, and thus reduce the burden of future disease. This approach may also be useful for integrated control interventions across NTDs.

### Acknowledgments

The authors thank the Chaco and National Chagas control programs for continuing field support and advice; M. Victoria Cardinal, Yael M. Provecho and Natalia P. Macchiaverna for contributions to fieldwork and questionnaire design; Fundación Mundo Sano for hospitality at the study site; local communities, healthcare workers and personnel from the local hospital “Dante Tardelli” for participating in the serosurveys and treatments, granting access to the laboratory, and especially to Drs. Nilda Blanco Acevedo, Omar Lana and Marcelo Wirth. The authors are also very grateful to the local communities for their warm welcome.

### References

1. Keesing F, Holt RD, Ostfeld RS. Effects of species diversity on disease risk. *Ecol Lett.* 2006;9: 485–98. doi:10.1111/j.1461-0248.2006.00885.x
2. Anderson RM, May RM. *Infectious diseases of humans: dynamics and control.* Oxford, U.K.: Oxford Science Publications; 1991.
3. Smith DL, Battle KE, Hay SI, Barker CM, Scott TW, McKenzie FE. Ross, Macdonald, and a theory for the dynamics and control of mosquito-transmitted pathogens. *PLoS Pathog.* 2012;8:

e1002588. doi:10.1371/journal.ppat.1002588

4. Abad-Franch F. Eco-epidemiología de la enfermedad de Chagas. La enfermedad de chagas a la puerta de los 100 años de conocimiento de una endemia americana ancestral. Buenos Aires, Argentina: Organización Panamericana de la Salud - Fundación Mundo Sano; 2007. pp. 203–218.
5. Briceño-León R. La enfermedad de Chagas en las Américas: una perspectiva de ecosalud. *Cad Saúde Pública*. 2009;25: 71–82. doi:S0102-311X2009001300007
6. Ehrenberg JP, Ault SK. Neglected diseases of neglected populations: thinking to reshape the determinants of health in Latin America and the Caribbean. *BMC Public Health*. 2005;5: 119. doi:10.1186/1471-2458-5-119
7. Charron D. Ecohealth reasearch in practice. Innovative applications of an ecosystem approach to health. New York, NY: Springer; 2012. doi:10.1017/CBO9781107415324.004
8. El-Sayed AM, Scarborough P, Seemann L, Galea S. Social network analysis and agent-based modeling in social epidemiology. *Epidemiol Perspect Innov*. BioMed Central Ltd; 2012;9: 1. doi:10.1186/1742-5573-9-1
9. WHO. Chagas disease in Latin America: an epidemiological update based on 2010 estimates. *Wkly Epidemiol Rec*. 2015; 33–44.
10. Silveira AC. El control de la Enfermedad de Chagas en los paises del Cono Sur de América. In: Silveira AC, editor. *El control de la Enfermedad de Chagas en los paises del Cono Sur de América: Historia de una iniciativa internacional 1991/2001*. Organización Panamericana de la Salud; 2002. pp. 15–44.
11. Hotez P. Ten Global “Hotspots” for the Neglected Tropical Diseases. *PLoS Negl Trop Dis*. 2014;8: e2496. doi:10.1371/journal.pntd.0002496
12. Gürtler RE. Sustainability of vector control strategies in the Gran Chaco Region: current

challenges and possible approaches. Mem Inst Oswaldo Cruz. 2009;104: 52–59.

doi:10.1590/S0074-02762009000900009

13. Gürtler RE. Eco-epidemiología regional de la transmisión vectorial: Enfermedad de Chagas en el Gran Chaco. La enfermedad de Chagas A la puerta de los 100 años del conocimiento de una endemia americana ancestral Balance y futuro, 1909-2006 Chagas hacia el siglo XXI. Buenos Aires, Argentina: Organización Panamericana de la Salud - Fundación Mundo Sano; 2007. pp. 137–155.

14. Chuit R, Segura EL. El control de la enfermedad de Chagas en la República Argentina. In: Silveira AC, editor. Rev Fed Arg Cardiol. Organización Panamericana de la Salud; 2012. pp. 46–109.

15. Segura EL, Cura EN, Estani SA, Andrade J, Lansetti JC, de Rissio AM, et al. Long-term effects of a nationwide control program on the seropositivity for *Trypanosoma cruzi* infection in young men from Argentina. Am J Trop Med Hyg. 2000;62: 353–62.

16. Diosque P, Padilla AM, Cimino RO, Cardozo RM, Negrette OS, Marco JD, et al. Chagas disease in rural areas of Chaco Province, Argentina: epidemiologic survey in humans, reservoirs, and vectors. Am J Trop Med Hyg. 2004;71: 590–3.

17. Biancardi MA, Conca Moreno M, Torres N, Pepe C, Altcheh J, Freilij H. Seroprevalence of Chagas disease in 17 rural communities of “Monte Impenetrable”, Chaco Province. Medicina (Buenos Aires). 2003;63: 125–9.

18. Alonso JM, Fabre AR, Galván M, Lucero RH, Brusés BL, Kuc A. La enfermedad de Chagas en poblaciones aborígenes del Noreste de Argentina. Enf Emerg. 2009;11: 115–118.

19. Sosa-Estani S, Dri LL, Touris C, Abalde S, Dell’Arciprete A, Braunstein JJ. Transmission vectorial y congénita del *Trypanosoma cruzi* en Las Lomitas, Formosa. Medicina (Buenos Aires). 2009;69: 424–430.

20. Lucero RH, Brusés BL, Cura CI, Formichelli LB, Juiz N, Fernández GJ, et al. Chagas’ disease in

aboriginal and creole communities from the Gran Chaco region of Argentina: seroprevalence and molecular parasitological characterization. *Infect Genet Evol.* 2016;41: 84–92.  
doi:10.1016/j.meegid.2016.03.028

21. Moretti E, Castro I, Franceschi C, Basso B. Chagas disease: serological and electrocardiographic studies in Wichi and Creole communities of Misión Nueva Pompeya, Chaco, Argentina. *Mem Inst Oswaldo Cruz.* 2010;105: 621–627. doi:10.1590/S0074-02762010000500004

22. Basombrio MA, Segovia A, Esteban E, Stumpf R, Jurgensen P, Winkler MA, et al. Endemic *Trypanosoma cruzi* infection in indian populations of the Gran Chaco territory of South America: performance of diagnostic assays and epidemiological features. *Ann Trop Med Parasitol.* 1999;93: 41–8.

23. Taranto NJ, Cajal SP, De Marzi MC, Fernández MM, Frank FM, Brú AM, et al. Clinical status and parasitic infection in a Wichí aboriginal community in Salta, Argentina. *Trans R Soc Trop Med Hyg.* 2003;97: 554–8.

24. Cardinal MV, Sartor PA, Gaspe MS, Enriquez GF, Colaiani I, Gürtler RE. High levels of human infection with *Trypanosoma cruzi* associated with the domestic density of infected vectors and hosts in a rural area of northeastern Argentina. *Parasit Vectors.* 2018;11: 492. doi:10.1186/s13071-018-3069-0

25. Gürtler RE, Cecere MC, Lauricella M a., Petersen RM, Chuit R, Segura EL, et al. Incidence of *Trypanosoma cruzi* infection among children following domestic reinfestation after insecticide spraying in rural northwestern Argentina. *Am J Trop Med Hyg.* 2005;73: 95–103. doi:73/1/95 [pii]

26. Gürtler RE, Cecere MC, Castañera MB, Canale DM, Lauricella MA, Chuit R, et al. Probability of infection with *Trypanosoma cruzi* of the vector *Triatoma infestans* fed on infected humans and dogs in northwest Argentina. *Am J Trop Med Hyg.* 1996;55: 24–31.

27. Guhl F, Lazdins-Helds J, Grupo de trabajo científico sobre la enfermedad de Chagas. Reporte



sobre la enfermedad de Chagas. World Health Organization; 2007.

28. Gürtler RE, Chuit R, Cecere M, Castanera M, Cohen JEJ, Segura EL, et al. Household prevalence of seropositivity for *Trypanosoma cruzi* in three rural villages in northwest Argentina: environmental, demographic, and entomologic associations. *Am J Trop Med Hyg.* 1998;59: 741–749.
29. Levy MZ, Kawai V, Bowman NM, Waller LA, Cabrera L, Pinedo-Cancino V V, et al. Targeted screening strategies to detect *Trypanosoma cruzi* infection in children. *PLoS Negl Trop Dis.* 2007;1: e103. doi:10.1371/journal.pntd.0000103
30. Alroy KA, Huang C, Gilman RH, Quispe-Machaca VR, Marks MA, Ancca-Juarez J, et al. Prevalence and transmission of *Trypanosoma cruzi* in people of rural communities of the High Jungle of Northern Peru. *PLoS Negl Trop Dis.* 2015;9: e0003779. doi:10.1371/journal.pntd.0003779
31. Samuels AM, Clark EH, Galdos-Cardenas G, Wiegand RE, Ferrufino L, Menacho S, et al. Epidemiology of and impact of insecticide spraying on Chagas disease in communities in the Bolivian Chaco. *PLoS Negl Trop Dis.* 2013;7: e2358. doi:10.1371/journal.pntd.0002358
32. Mott KE, Muniz TM, Lehman JS, Hoff R, Morrow RH, de Oliveira TS, et al. House construction, triatomine distribution, and household distribution of seroreactivity to *Trypanosoma cruzi* in a rural community in northeast Brazil. *Am J Trop Med Hyg.* 1978;27: 1116–22.
33. Fortin M-J, Dale MRT. *Spatial Analysis: A guide for ecologist.* Cambridge, UK: Cambridge University Press; 2005.
34. Kitron U. Risk maps: transmission and burden of vector-borne diseases. *Parasitol Today.* 2000;16: 324–325. doi:10.1016/S0169-4758(00)01708-7
35. Wiegand T, Moloney KA. *Handbook of spatial point-pattern analysis in ecology.* Boca Raton, FL: CRC Press; 2014.

36. Cecere MC, Vazquez-Prokopec GM, Gürtler RE, Kitron U. Spatio-temporal analysis of reinfestation by *Triatoma infestans* (Hemiptera: Reduviidae) following insecticide spraying in a rural community in northwestern Argentina. *Am J Trop Med Hyg.* 2004;71: 803–10.
37. Cecere MC, Vazquez-Prokopec GM, Ceballos LA, Gurevitz JM, Zárate JE, Zaidenberg M, et al. Comparative trial of effectiveness of pyrethroid insecticides against peridomestic populations of *Triatoma infestans* in northwestern Argentina. *J Med Entomol.* 2006;43: 902–9.
38. Vazquez-Prokopec GM, Cecere MC, Kitron U, Gürtler RE. Environmental and demographic factors determining the spatial distribution of *Triatoma guasayana* in peridomestic and semi-sylvatic habitats of rural northwestern Argentina. *Med Vet Entomol.* 2008;22: 273–282.  
doi:10.1111/j.1365-2915.2008.00746.x
39. Gorla DE, Porcasi X, Hrellac H, Catalá SS, Catala SS. Spatial stratification of house infestation by *Triatoma infestans* in la Rioja, Argentina. *Am J Trop Med Hyg.* 2009;80: 405–409.
40. Gaspe MS, Provecho YM, Fernández MP, Vassena C V., Santo Orihuela PL, Gürtler RE. Beating the odds: Sustained Chagas disease vector control in remote indigenous communities of the Argentine Chaco over a seven-year period. *PLoS Negl Trop Dis.* 2018;12: e0006804.  
doi:10.1371/journal.pntd.0006804
41. Gaspe MS, Provecho YM, Cardinal MV, Fernández MP, Gürtler RE. Ecological and sociodemographic determinants of house infestation by *Triatoma infestans* in indigenous communities of the Argentine Chaco. *PLoS Negl Trop Dis.* 2015;9: e0003614.  
doi:10.1371/journal.pntd.0003614
42. Gurevitz JM, Ceballos LA, Gaspe MS, Alvarado-Otegui JA, Enríquez GF, Kitron U, et al. Factors affecting infestation by *Triatoma infestans* in a rural area of the humid Chaco in Argentina: a multi-model inference approach. Turell MJ, editor. *PLoS Negl Trop Dis.* 2011;5: e0001349.  
doi:10.1371/journal.pntd.0001349

43. Levy MZ, Bowman NM, Kawai V, Waller LA, Cornejo Del Carpio JG, Cordova Benzaquen E, et al. Periurban *Trypanosoma cruzi*-infected *Triatoma infestans*, Arequipa, Peru. Emerg Infect Dis. Centers for Disease Control & Prevention (CDC); 2006;12: 1345–1352. doi:10.3201/eid1209.051662
44. Gutfraind A, Peterson JK, Billig Rose E, Arevalo-Nieto C, Sheen J, Condori-Luna GF, et al. Integrating evidence, models and maps to enhance Chagas disease vector surveillance. PLoS Negl Trop Dis. 2018;12: e0006883. doi:10.1371/journal.pntd.0006883
45. Gorla DE. Variables ambientales registradas por sensores remotos como indicadores de la distribución geográfica de *Triatoma infestans* (Heteroptera: Reduviidae). Ecol austral. ASAE; 2002;12: 117–127.
46. Vazquez-Prokopec GM, Spillmann C, Zaidenberg M, Gürtler RE, Kitron U. Spatial heterogeneity and risk maps of community infestation by *Triatoma infestans* in rural northwestern Argentina. PLoS Negl Trop Dis. 2012;6: e1788. doi:10.1371/journal.pntd.0001788
47. Levy MZ, Bowman NM, Kawai V, Plotkin JB, Waller LA, Cabrera L, et al. Spatial patterns in discordant diagnostic test results for Chagas disease: links to transmission hotspots. Clin Infect Dis. 2009;48: 1104–6. doi:10.1086/597464
48. Delgado S, Castillo Neyra R, Quispe Machaca VR, Ancca Juárez J, Chou Chu L, Verastegui MR, et al. A history of chagas disease transmission, control, and re-emergence in peri-rural La Joya, Peru. PLoS Negl Trop Dis. 2011;5: e970. doi:10.1371/journal.pntd.0000970
49. Gaspe MS, Provecho YM, Piccinali RV, Gürtler RE. Where do these bugs come from? Phenotypic structure of *Triatoma infestans* populations after control interventions in the Argentine Chaco. Mem Inst Oswaldo Cruz. 2015;110: 310–318. doi:10.1590/0074-02760140376
50. Sartor P, Colaianni I, Cardinal MV, Bua J, Freilij H, Gürtler RE. Improving access to Chagas disease diagnosis and etiologic treatment in remote rural communities of the Argentine Chaco

- through strengthened primary health care and broad social participation. PLoS Negl Trop Dis. 2017;11: e0005336. doi:10.1371/journal.pntd.0005336
51. Gurevitz JM, Gaspe MS, Enríquez GF, Vassena C V, Alvarado-Otegui JA, Provecho YM, et al. Unexpected failures to control Chagas disease vectors with pyrethroid spraying in northern Argentina. J Med Entomol. 2012;49: 1379–86.
52. Provecho YM, Gaspe MS, Fernández MP, Gürtler RE. House reinfestation with *Triatoma infestans* (Hemiptera: Reduviidae) after community-wide spraying with insecticides in the Argentine Chaco: a multifactorial process. J Med Entomol. 2017;54: 646–657. doi:10.1093/JME/TJW224
53. Enriquez GF, Cardinal M V, Orozco MM, Lanati LA, Schijman AG, Gürtler RE. Discrete typing units of *Trypanosoma cruzi* identified in rural dogs and cats in the humid Argentinean Chaco. Parasitology. 2013;140: 303–8. doi:10.1017/S003118201200159X
54. Fernandez MP, Gaspe MS, Gürtler RE. Inequalities in the social determinants of health and Chagas disease transmission risk in indigenous and creole households in the Argentine Chaco. Parasit Vectors. 2019; *in press*.
55. U.S. Census Bureau. Current Population Survey (CPS)—Definitions. [Internet]. 2017 [cited 10 Nov 2017]. Available: [https://www.census.gov/glossary/#term\\_Household](https://www.census.gov/glossary/#term_Household)
56. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ. 2007;335: 806–8. doi:10.1136/bmj.39335.541782.AD
57. Maffey L, Cardinal M V, Ordóñez-Krasnowski P, Lanati LA, Lauricella MA, Schijman AG. Direct molecular identification of *Trypanosoma cruzi* discrete typing units in domestic and peridomestic *Triatoma infestans* and *Triatoma sordida* from the Argentine Chaco. Parasitology. 2012;139: 1570–1579. doi:10.1017/S0031182012000856

- 795 58. Brown LD, Cai TT, DasGupta A. Interval estimation for a binomial proportion. *Stat Sci.* 2012;16:  
796 101–133.
- 797 59. Gelman A, Hill J. *Data analysis using regression and multilevel/hierarchical models.* Cambridge,  
798 UK; 2007. doi:10.2277/0521867061
- 799 60. Zeileis A, Kleiber C, Jackman S. Regression models for count data in R. *J Stat Softw.* 2008;27: 1–  
800 25.
- 801 61. Stata Corp. *Stata Statistical Software: Release 14.0.* College Station: Stata Corporation; 2017.
- 802 62. R Core Team, Team RC. *R: a language and environment for statistical computing.* Vienna,  
803 Austria., Austria: R Foundation for Statistical Computing; 2018.
- 804 63. Schisterman EF, Perkins NJ, Liu A, Bondell H. Optimal cut-point and its corresponding Youden  
805 Index to discriminate individuals using pooled blood samples. *Epidemiology.* 2005;16: 73–81.
- 806 64. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, et al. pROC: an open-source  
807 package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics. BioMed*  
808 *Central*; 2011;12: 77. doi:10.1186/1471-2105-12-77
- 809 65. Hotez PJ, Bottazzi ME, Franco-Paredes C, Ault SK, Periago MR. The neglected tropical diseases  
810 of Latin America and the Caribbean: a review of disease burden and distribution and a roadmap  
811 for control and elimination. *PLoS Negl Trop Dis.* 2008;2: e300. doi:10.1371/journal.pntd.0000300
- 812 66. Aagaard-Hansen J, Claire LC. Neglected tropical diseases: equity and social determinants. In: Blas  
813 E, Kurup AS, editors. *Equity, social determinants and public health programmes.* Ginebra: World  
814 Health Organization; 2010. pp. 135–157.
- 815 67. Chippaux J-P, Postigo JR, Santalla JA, Schneider D, Brutus L. Epidemiological evaluation of  
816 Chagas disease in a rural area of southern Bolivia. *Trans R Soc Trop Med Hyg.* 2008;102: 578–84.  
817 doi:10.1016/j.trstmh.2008.03.008

68. Gualtieri JM, Baldi EM. Análisis de una experiencia piloto de control de la enfermedad de Chagas. Prensa Med Argent. 1981;68: 209–216.
  69. Mott KE, Lehman JS, Hoff R, Morrow RH, Muniz TM, Sherlock I, et al. The epidemiology and household distribution of seroreactivity to *Trypanosoma cruzi* in a rural community in northeast Brazil. Am J Trop Med Hyg. 1976;25: 552–62.
  70. Cohen JE, Gürtler RE. Modeling household transmission of American trypanosomiasis. Science. 2001;293: 694–698. doi:10.1126/science.1060638
  71. Gürtler RE, Yadon ZE. Eco-bio-social research on community-based approaches for Chagas disease vector control in Latin America. Trans R Soc Trop Med Hyg. 2015;109: 91–98. doi:10.1093/trstmh/tru203
  72. Llovet I, Dinardi G, De Maio FG. Mitigating social and health inequities: Community participation and Chagas disease in rural Argentina. Glob Public Health; 2011;6: 371–384. doi:10.1080/17441692.2010.539572
  73. Stoddard ST, Forshey BM, Morrison AC, Paz-Soldan VA, Vazquez-Prokopec GM, Astete H, et al. House-to-house human movement drives dengue virus transmission. Proc Natl Acad Sci. 2013;110: 994–9. doi:10.1073/pnas.1213349110
  74. Tarleton RL, Gürtler RE, Urbina JA, Ramsey J, Viotti R. Chagas disease and the London declaration on Neglected Tropical Diseases. PLoS Negl Trop Dis. 2014;8: e3219. doi:10.1371/journal.pntd.0003219
- Supporting information captions
- S1 Text.** Checklist of STROBE recommendations for observational studies.
- S1 Table.** Generalized linear mixed model of seropositivity for *T. cruzi* infection vs. demographic variables in 2008, clustered by household (logit link function).

**S2 Table.** Generalized linear mixed model of seropositivity for *T. cruzi* infection vs. demographic and vector indices in 2008, clustered by household (logit link function) for the total population and for children.

**S3 Table.** Individual-level and household-level database including data on human infection with *Trypanosoma cruzi*, domiciliary infestation with *Triatoma infestans*, and environmental and socio-demographic variables in Area III, Pampa del Indio.

**S1 Figure.** Number of people tested in each phase of the serosurvey in 2012-2015 in Area III, Pampa del Indio.

**S2 Figure.** Results from the global spatial point pattern analysis for the human and vector infection in 2008. (A) Univariate global analysis for the occurrence of at least one seropositive person in the household; (B) Variogram of the number of seropositive people per household; (C) Univariate global analysis for the occurrence of at least one seropositive children in the household; (D) Variogram of the number of seropositive children per household; (E) Univariate global analysis for the occurrence of at least one infected *T. infestans* considering all houses; (F) Univariate global analysis for the occurrence of at least one infected *T. infestans* considering only infested houses; (G) Bivariate global analysis for the correlation between children seropositivity and *T. infestans* infection; (H) Variogram for the spatial correlation between the abundance of infected *T. infestans* and the number of seropositive children per household. The lines with red dots indicate the observed values and the solid lines indicate the confidence ‘envelopes’.

**S3 Figure.** Distribution of the household number of *T. cruzi*-seropositive people (A) and children (B) in 2015. The maps were created in QGIS 2.18.11. based on the data collected within the scope of this study (S3 Table).

**S4 Figure.** Variograms from the global spatial point pattern analysis for socio-demographic variables and human or vector infection: the social vulnerability index vs. the number of seropositive people per household (A), vs. the number of seropositive children (B), and vs. the abundance of infected *T. infestans* (C); and the host availability index vs. the number of seropositive people per household (E), vs. the

867 number of seropositive children (F), and vs. the abundance of infected *T. infestans* (G). The lines with red  
868 dots indicate the observed values and the solid lines indicate the confidence ‘envelopes’.



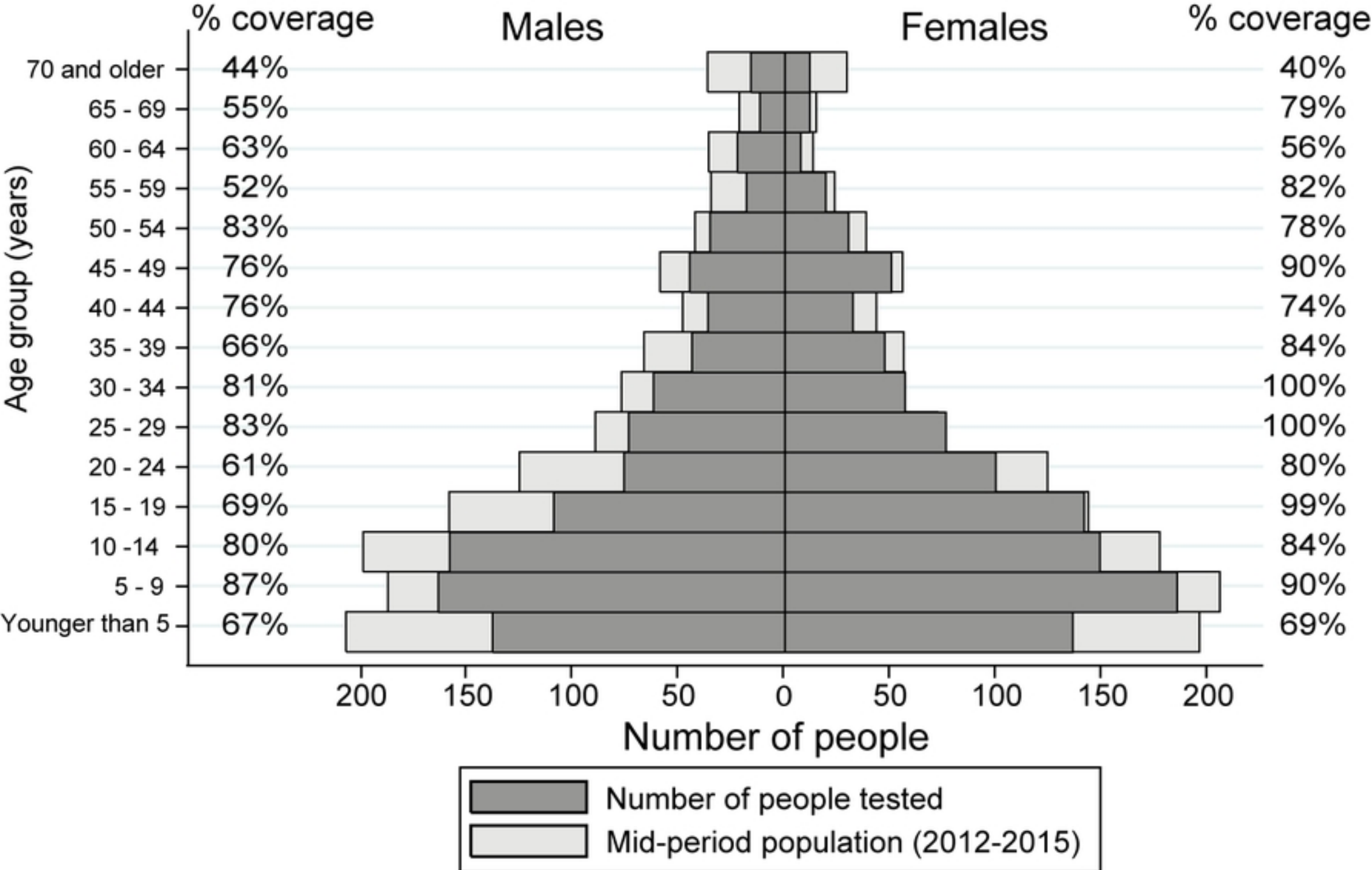


Figure 1

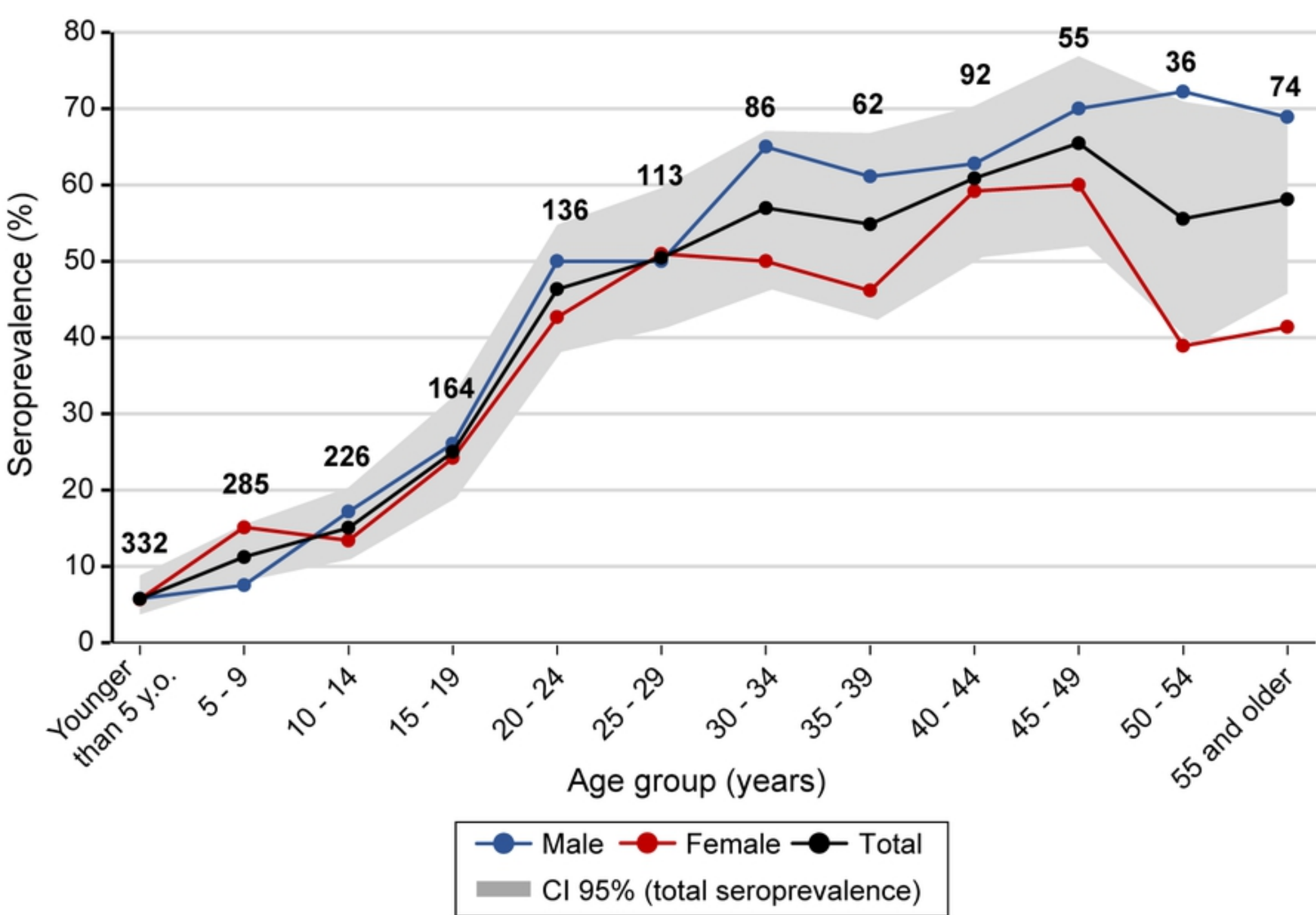
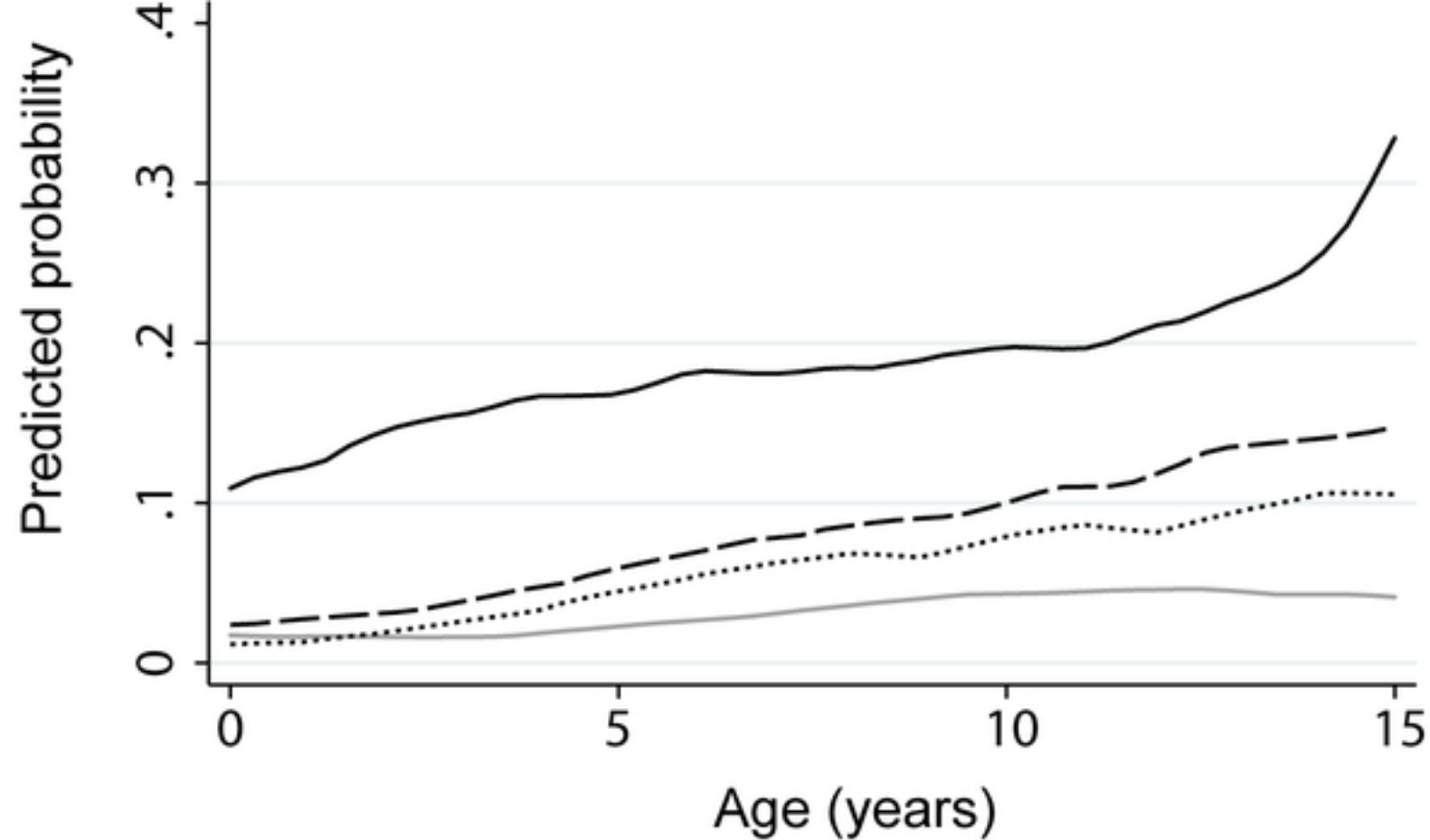


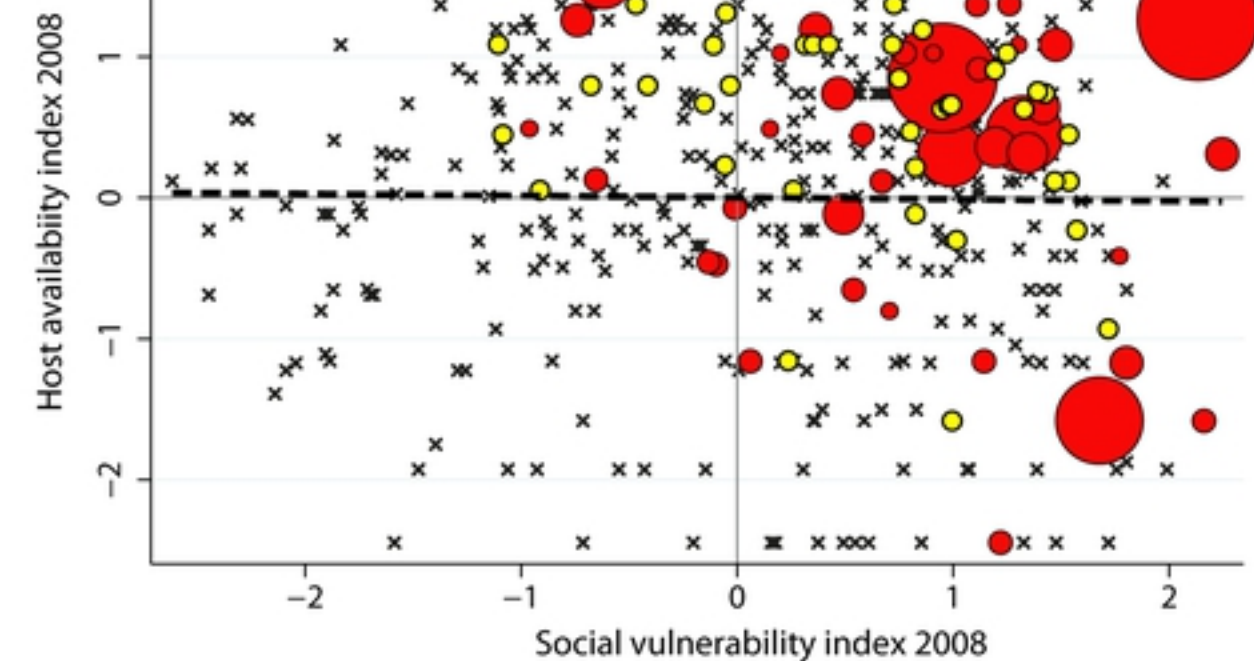
Figure 2



— Non-mover / non-infested domicile — Non-mover / infested domicile  
- - - Mover / non-infested domicile ..... Mover / infested domicile

Figure 3

A



B

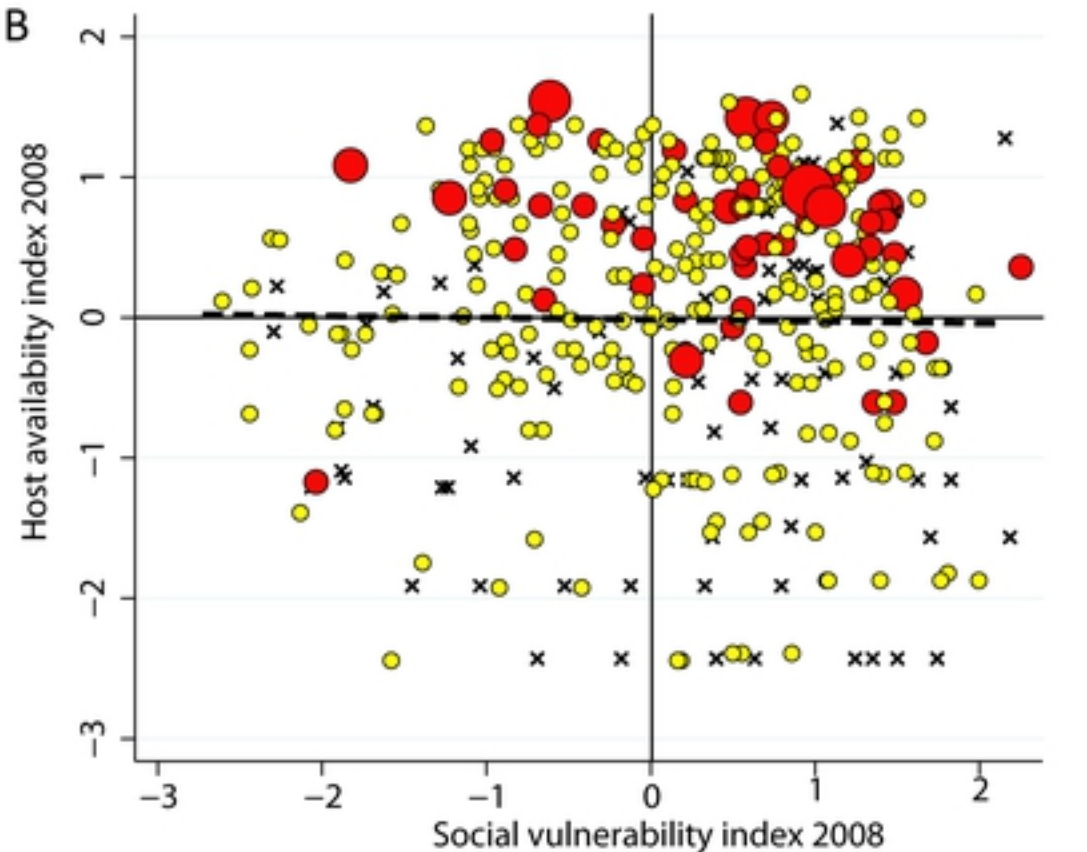


Figure 4

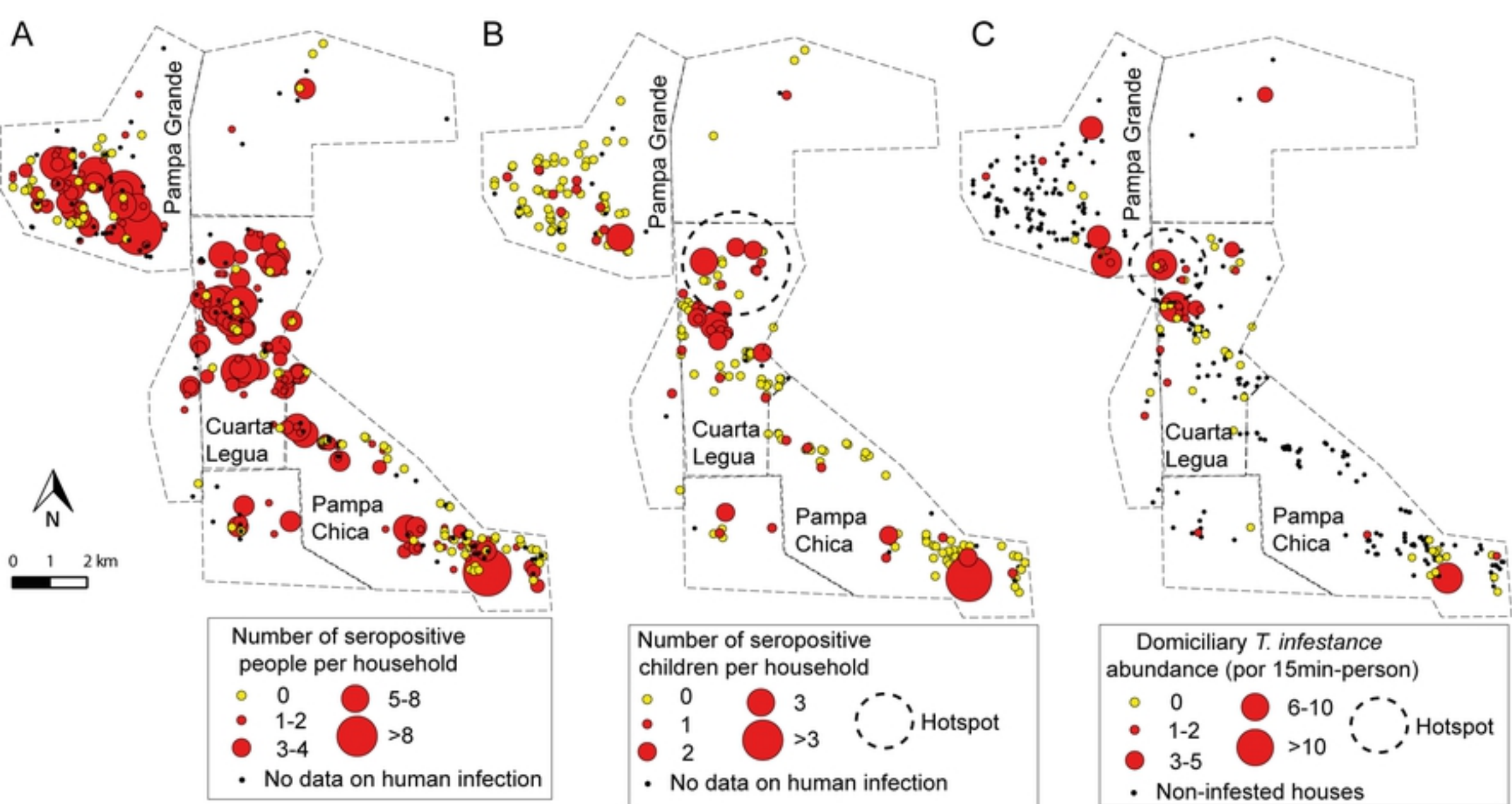


Figure 5



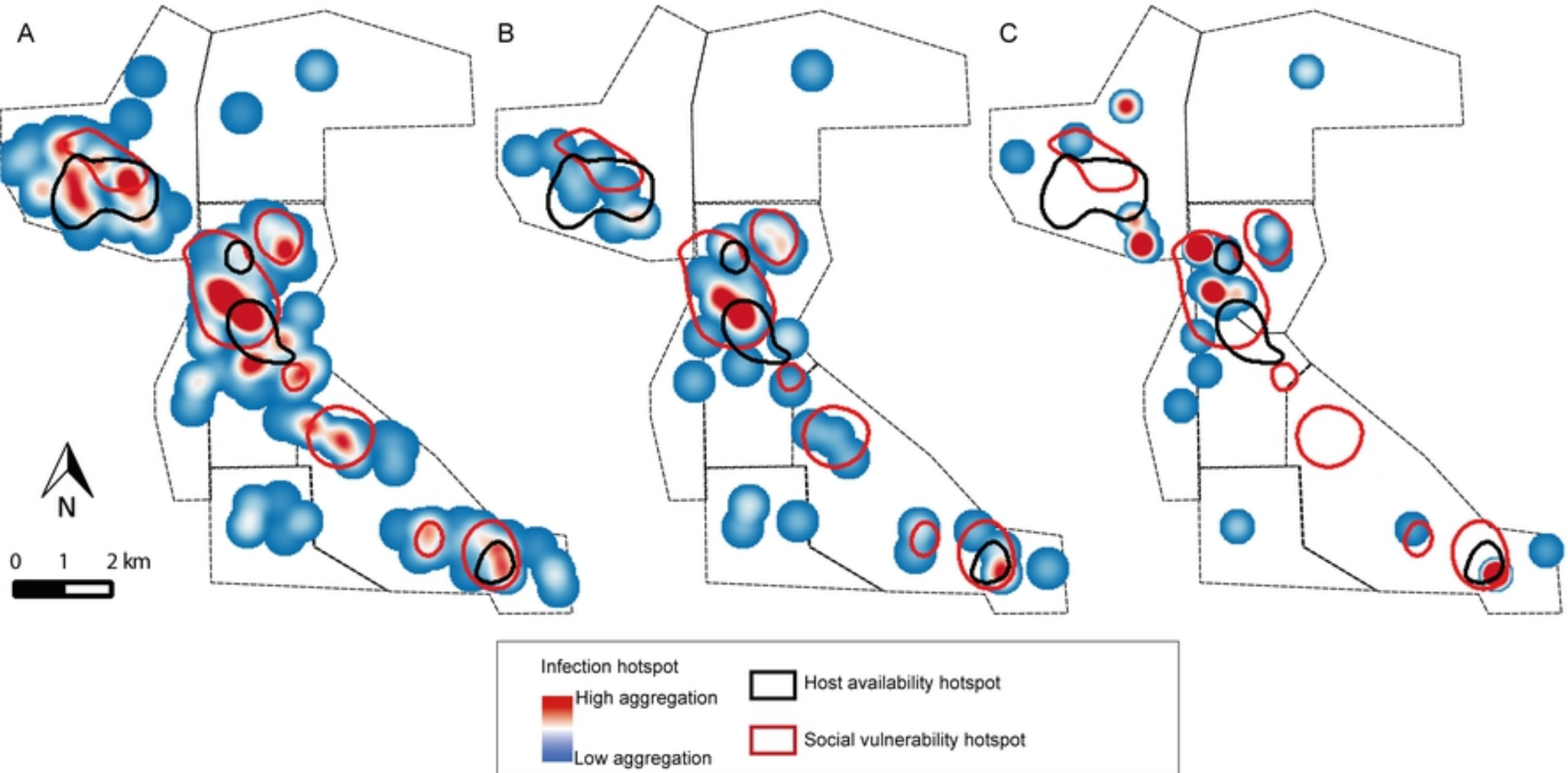
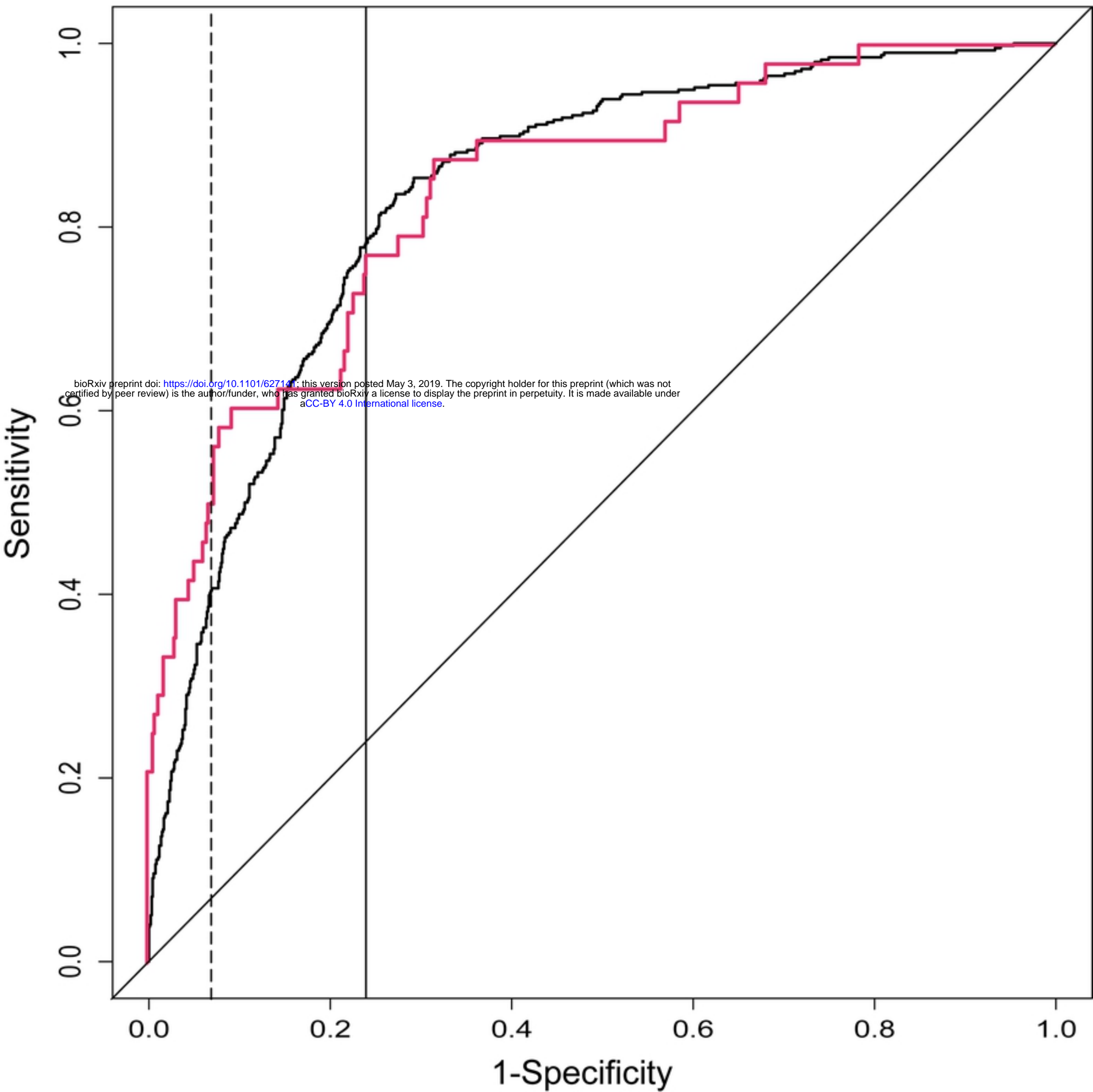


Figure 6

# Optimum threshold values

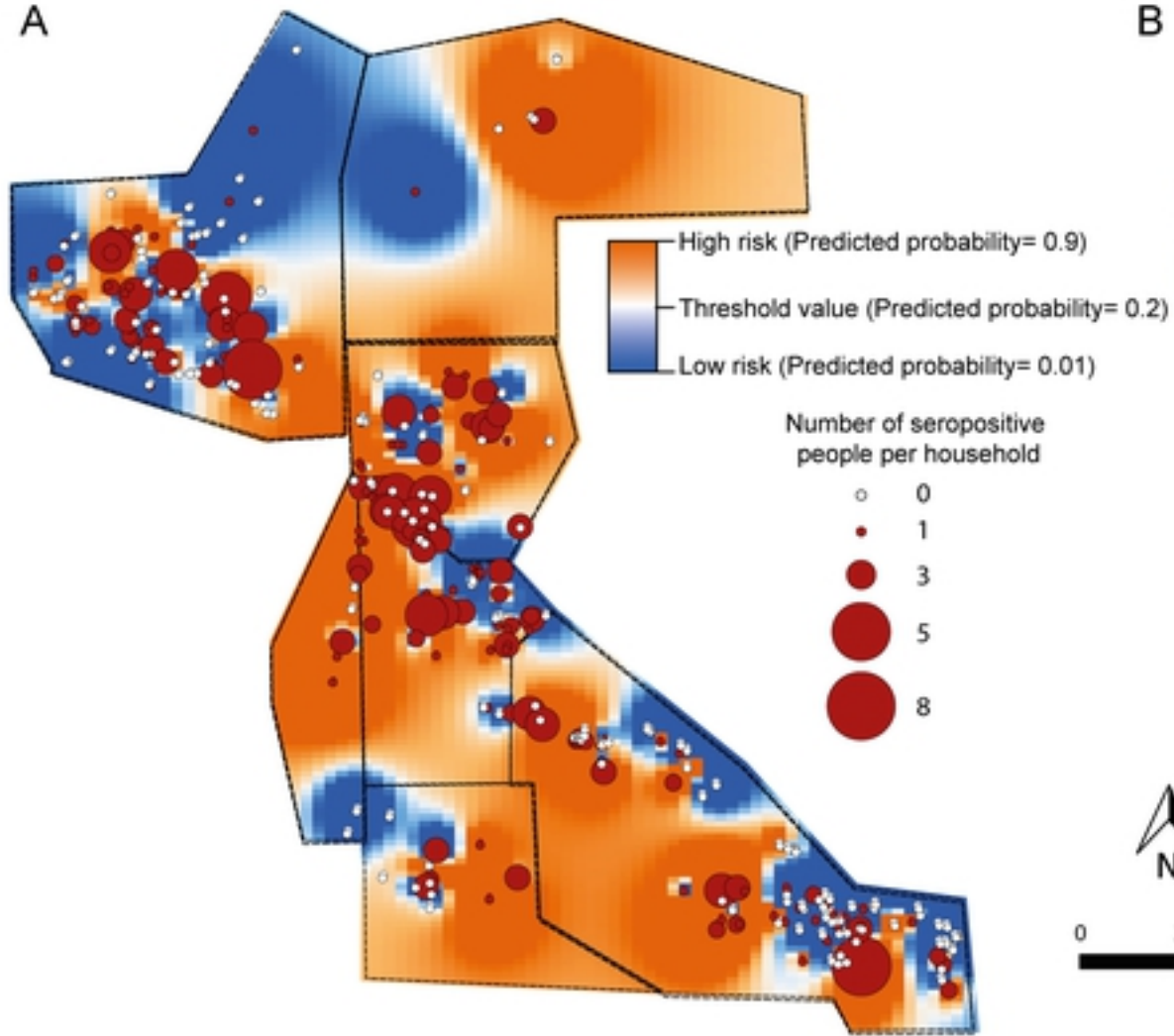
m2=0.07    m1=0.24



- ROC for the infection risk model for the total population (m1)
- ROC for the infection risk model for children (m2)

Figure 7

A



B

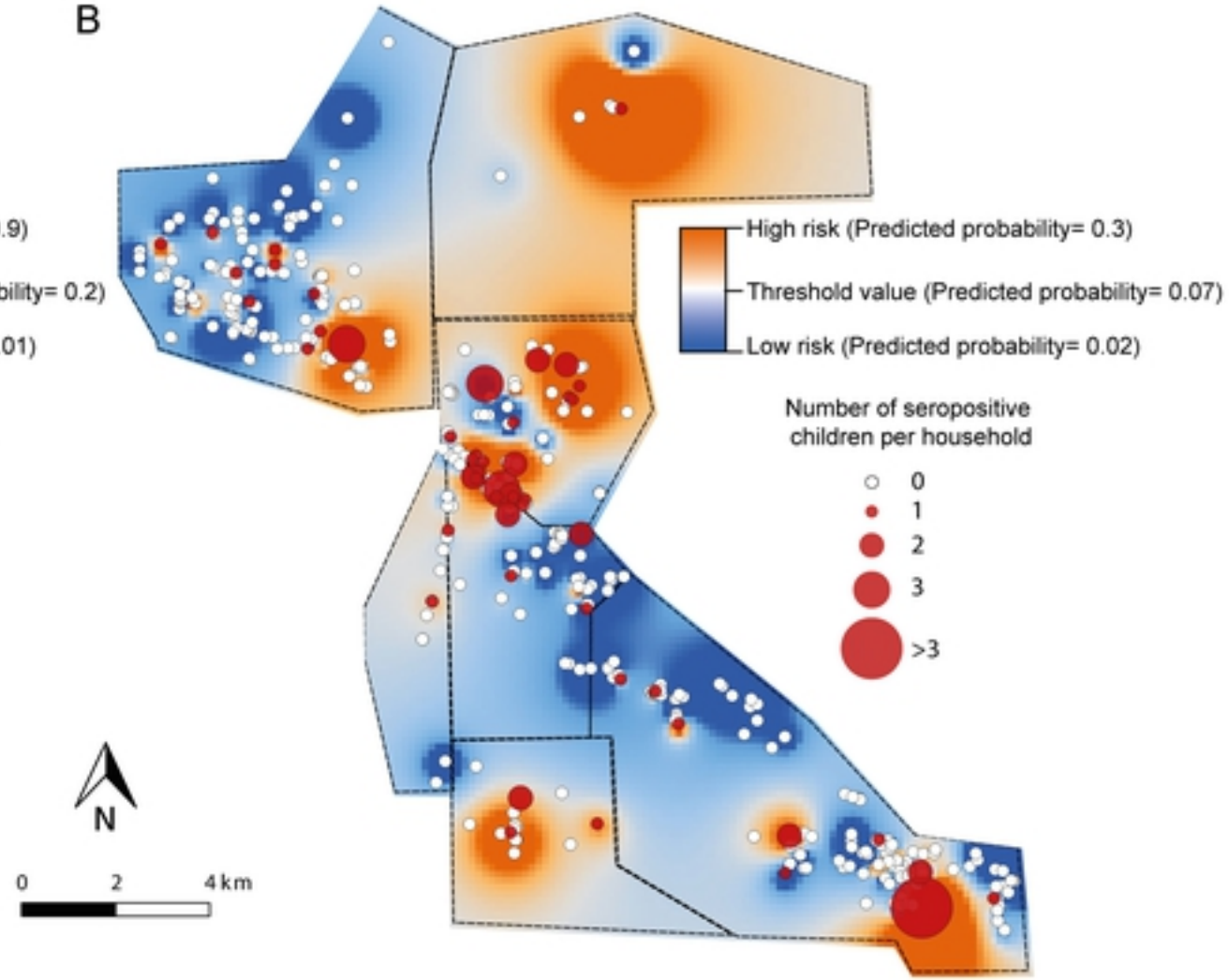


Figure 8