Integration of Phlebotomine Ecological Niche Modelling, and Mapping of Cutaneous Leishmaniasis Surveillance Data, to Identify Areas at Risk of Under-Estimation

Clara B Ocampo[1,2,3], Lina Guzmán-Rodríguez[1], Mabel Moreno[1], María del Mar Castro[1,2], Carlos Valderrama-Ardila[4], Neal Alexander*[1,2]

1. Centro Internacional de Entrenamiento e Investigaciones Médicas, CIDEIM. Cali, Colombia.

2. Universidad Icesi. Cali, Colombia

3. Current address: Ministerio de Ciencia Tecnología e Innovación (Minciencias), Bogotá, Colombia

4. Universidad del Rosario. Bogotá, Colombia

*Corresponding author

Introduction: Passive surveillance systems are thought to under-estimate the true incidence of American cutaneous leishmaniasis (ACL) by two- to five-fold. Ecological niche models based on remotely sensed data can identify environmental factors which favor phlebotomine vectors. Here we report an integrated approach to identifying areas at risk of cutaneous leishmaniasis by applying spatial analysis methods to niche model results, and local surveillance data, in two locations in Colombia with differing vector ecology. The objective was to identify townships in which later phases of the project could implement community-based surveillance to obtain direct estimates of under-reporting.

Materials and methods: The study was carried out in one municipality in each of two departments of the Andean region of Colombia: Pueblo Rico in Risaralda, and Rovira in Tolima. Niche mapping by maximum entropy, based on published and unpublished existing locations of *Pintomyia* (*Pifanomyia*) *longiflocosa* and *Psychodopygus panamensis*, and using variables on land cover, climate and elevation. Field catches were done in each municipality to test predictions of high relative probability of presence. The niche model results were included as a predictor in a conditional autoregressive spatial model, in which the outcome variable was the number of cases by township, as detected by passive surveillance.

Results: Having rarefied 173 geolocated records, 46 of *Pi. longiflocosa* and 57 of *Ps. panamensis* were used for the niche modelling. At the national level, both species had high relative probability of presence on parts of the slopes of the three Andean cordilleras. *Pi. longiflocosa* also has a high relative probability of presence in the higher parts of the Magdalena valley, as does *Ps. panamensis* in some areas close to the Caribbean coast. At the local level, field catches confirmed that *Pi. longiflocosa* was the most abundant species in Rovira, and likewise *Ps. panamensis* in Pueblo Rico. The spatial regression showed that the incidence of ACL, according to surveillance, was positively, but not statistically significantly, associated with the relative probability of presence from the risk model.

Conclusions: These niche maps bring together published and unpublished results on phlebotomine species which are important vectors in Colombia. Maps of the fitted values of incidence were used to guide the selection of townships in which further phases of the study will attempt to quantify the extent of under-estimation of ACL incidence.

Keywords: phlebotomines; niche modelling; cutaneous leishmaniasis; surveillance; under-estimation

Introduction

The leishmaniases are caused by protozoan parasites of the genus *Leishmania* (Kinetoplastida: Trypanosomatidae), which are transmitted by bites of infected female phlebotomine (Diptera: Psychodidae) insect vectors. There are three main forms of leishmaniasis: cutaneous, mucocutaneous, and visceral or kala-azar. They are a complex group of diseases caused by over 20 parasite species and with over 90 recognized phlebotomine vector species, among the approximately 1000 sandflies species described in the world, that transmit the parasites to multiple mammal hosts in different ecological environments (Galati et al., 2017; World Health Organization, 2020). The resulting complex eco-epidemiology makes the disease very difficult to control by local authorities (Alvar et al., 2012). Moreover, the disease is subject to substantial under-estimation, with one study finding that only 14 of 20 countries had five years of surveillance data available on cutaneous leishmaniasis, and that the actual incidence was likely between 2.8 and 4.6 times the reported value (Alvar et al., 2012).

In Colombia, American cutaneous leishmaniasis is the dominant form of the disease, comprising 96% of reported cases (Ramirez et al., 2016). During the 1990s, an average of 6,500 new cases of leishmaniasis were reported per year, a figure that progressively increased to about 20,000 cases annually in 2005 and 2006 (Zambrano, 2007). In 2019, 5,105 cases were reported in the national surveillance system (SIVIGILA) (Agudelo Chivatá, 2019). New World leishmaniasis is considered to be zoonotic (Burza et al., 2018). Nine parasite species have been identified in Colombia, with the most frequent being *Leishmania panamensis* and *L. braziliensis* (Ramirez et al., 2016). Knowledge of mammalian hosts is imperfect, although, in Colombia, they include the opossums *Didelfis marsupialis* and *Gracilinanus marica*, (Roque and Jansen, 2014; Travi et al., 1994)and rodents of the family Muridae: *Oecomys trinitatus*, *Zygodontomys brunneus*, and *Sigmodon hispidus* (Lopez et al., 2021; Ocampo et al., 2012; Roque and Jansen, 2014; Travi et al., 1994). Other hosts may include dogs (Lago et al., 2019; Santaella et al., 2011) and rodents such as *Proechimys* sp. and *Coendou* sp (Lopez et al., 2021). As in other Latin American countries, leishmaniasis in Colombia has been favored by high human movement and an increase of domestic transmission associated with land use changes, and the adaptation of phlebotomine species (Davies et al., 2000; Ferro et al., 2015; Ferro et al., 2011; Valderrama-Ardila et al., 2010).

In Colombia, 163 species of the Phlebotominae subfamily have been described, of which 21 are of medical importance (Ferro et al., 2015). Phlebotomines are distributed in almost all ecological niche environments from sea level to 3500 m (Bejarano et al., 2003). Transmission mainly occurs in rural areas, which are often remote and difficult to access. Disease response in Colombia mainly consists of treatment of confirmed cases and house fumigation when cases appear (Ferro et al., 2011; Instituto Nacional de Salud, 2014), although this largely misses transmission periods. Disease prevention is mainly via the post-outbreak distribution of long-lasting insecticide-treated bednets, but coverage remains very low. Alvar et al. (Alvar et al., 2012) assessed ACL to be under-reported in Colombia by a factor of 2.8-4.6, and the country's WHO resource document notes under-reporting as limitation to its control program (World Health Organization).

Ecological niche models can identify environmental factors which favor phlebotomine vector species. They can use remotely sensed data and may therefore be capable of efficiently identifying areas of potential transmission (Chavy et al., 2019). We report an integrated approach to identifying areas at risk of cutaneous leishmaniasis at the township (*vereda*) level by applying statistical spatial analysis methods to niche model results, and local surveillance data, in two endemic municipalities of Colombia with different vector ecology. The objective was to identify townships to implement methods such as community-based surveillance to obtain direct estimates of under-reporting.

Methodology

Study sites

Two municipalities in the Andean region of Colombia were selected: Pueblo Rico in the department of Risaralda and Rovira in the department of Tolima (Figure 1). These were selected based on prior evidence of domiciliary transmission of cutaneous leishmaniasis with different vectors (Moreno et al., 2015; Moreno et al., 2020). The corresponding studies had established that it was feasible and safe to collect phlebotomines in these municipalities, and used alliances with local public health entities. Two vectors were selected for the ecological niche modeling: *Psychodophygus panamensis* and *Pintomyia* (*Pifanomyia*) *longiflocosa* (Ferro et al., 2011;

Santamaría et al., 2006). Georeferenced collections of these species were obtained from previous studies carried out by ourselves and other Colombian research groups (see acknowledgments), including both unpublished and published data (Bejarano et al., 2007; Bejarano et al., 2015; Ferro et al., 2015; Santamaría et al., 2006; Vivero et al., 2009). The data were obtained in different formats and collated into a Microsoft Excel file (see supplementary information).

Pueblo Rico. This municipality is located in the Department of Risaralda, bordering the Pacific rain forest. It is located on the eastern side of the western Andean cordillera (5.22156, -76.0292) with a mean altitude of 1560 m.a.s.l., range 296-4065 m.a.s.l. and a mean temperature of 20 °C (https://ipt.biodiversidad.co/sib/resource?r=biodiversidad_pueblorico). It has a predominantly humid and hyper-humid tropical forest because of its proximity to the Pacific Choco bioregion. Numerous species of potential vector species — *Nissomyia trapidoi*, *Psychodophygus panamensis* (formerly *L. panamensis*), *Lutzomyia* (*Tricholateralis*) *gomezi*, *Lutzomyia* (*Helcocyrtomyia*) *hartmanni* and *Warileya rotundipennis* — have been recorded in Risaralda and the neighboring department of Caldas (Contreras-Gutiérrez et al., 2014; Ferro et al., 2015).

Rovira. This municipality is located in the Department of Tolima, on the eastern slope of the central mountain range (4.23933, -75.23968) — with an altitudinal range of 591-3,771 m.a.s.l and a mean elevation of 1,109 m.a.s.l. Its topographical complexity gives it a diversity of climatic zones. The highest rainfall is between March-May and September-November (http://atlas.ideam.gov.co/visorAtlasClimatologico.html). Tolima has a history of leishmaniasis transmission and presented one of the country's largest epidemics, with *Pintomyia* (*Pif.*) *longiflocosa* being the most abundant species (Morales Ortegón et al., 2004).

Niche models

Niche models of the areas most suitable for the vector species were developed, and also used as an input to the spatial model of ACL incidence. All data for the niche model were processed in ArcGis 10.6: this includes ArcMap and maximum entropy modelling. Three categories of variables were included: land cover, climate and elevation, with three, four and one variables, respectively.

Satellite images of NDVI with a resolution of 250m and a temporal resolution of 16 days (http://doi.org/10.5067/MODIS/mod13q1.006) for 2012-2014 were obtained from the MODIS repository at USGS Earth Explorer (https://earthexplorer.usgs.gov/). These particular images were selected to lie within the period of the existing entomological data. Data reduction of the 59 images was done by principal component analysis (PCA), with the first three components explaining 49.22% of the variance (PC 1=38.46%, PC 2=6.60%, PC 3=4.16%).

WorldClim layers at a resolution of 1km were obtained from http://worldclim.org/version2 in April 2018. We used Version 2.0 as its data are more current: 1970-2000, as opposed to 1960-1990 in Version 1.4. The following layers were chosen based on our previous work (Pérez et al., 2016) and exploratory analysis for the current project: annual mean temperature, temperature seasonality, annual precipitation, and precipitation seasonality. The digital elevation model (DEM) for Colombia with original resolution of 90m (http://srtm.csi.cgiar.org/srtmdata/) was also used and re-scaled to the same 250m resolution as other data. Each variable was re-scaled in the GIS software to 250m and cut according to the defined area of interest (**Niche modeling**

The number of geolocated records was 82 for *Pi. longiflocosa* and 91 for *Ps. panamensis*. Of the total of 173, 130 were from published and 43 from unpublished sources. They were rarefied to 46 records of *Pi. longiflocosa* and 57 of *Ps. panamensis*. The niche models are shown in Figure 2. The jackknife analysis of the *Pi. longiflocosa* model showed that the model explained approximately 75% of the variation in the data. The most informative variables were annual precipitation (40% explained) and the temperature seasonality (21%). Each of the first three principal components of the NDVI data showed little importance (<10%). The analysis gave an area under the curve (AUC) of 0.858. *Pi. longiflocosa* has zones of high probability of presence on the eastern slopes of the western, central and eastern cordilleras, and also in the higher parts of the valley of the river Magdalena, between the central and eastern cordilleras, and on some parts on the northern Caribbean littoral.

For *Ps. panamensis*, the AUC was 0.819, and jackknife analysis explained at most 51% of the variation in the data. Elevation contributed most (27%), although this was well below its contribution to the *Pi. longiflocosa* model. As before, no principal component of NDVI explained more than 15%. *Ps. panamensis* has high relative probability of presence on smaller sections of the slopes of the cordilleras, and in larger areas to the north of the area of interest, close to the Caribbean coast.

Figure 2).

Geolocated records of presence were included from published sources (Bejarano et al., 2007; Ferro et al., 2015; Santamaría et al., 2006; Vivero et al., 2009) as well as previously unpublished ones from the Instituto Nacional de Salud, and PECET (Programa de Estudio y Control de Enfermedades Tropicales, Universidad de Antioquia). Given the occurrence of several phlebotomine spatial sites in close proximity, to reduce the influence of spatial autocorrelation (Segurado et al., 2006), we rarefied the data (Brown, 2014) to leave a distance of at least 1 km between any two. Within each set of proximal capture sites, one was randomly selected to achieve this. This 1 km radius was chosen bearing in mind phlebotomine flight habits and the flight range of some species of the same family (approximately 10-960m) (Akhoundi et al., 2016; Morrison et al., 1993; Mutinga et al., 1992). This process left 57 sites for *Psychodophygus panamensis*, and 46 for *Pintomyia* (*Pifanomyia*) *longiflocosa*. An initial test was carried with a radius of 500 m for *P. panamensis*, which differed from the 1 km one in only one capture point. The radius of the buffer layer (Brown, 2014) was chosen to seek to cover the entire Andean region of the country.

The SDMToolBox tool (http://sdmtoolbox.org/downloads), an extension for ArcGIS, was used to construct the niche models. This software uses a maximum entropy (MaxEnt) algorithm, and includes different combinations of feature classes and regularization factors that can be tuned by the user (https://biodiversityinformatics.amnh.org/open_source/maxent/). Maximum entropy was chosen since in initial tests, with a subset of the current data, its results were more biologically realistic than those from GARP (Stockman et al., 2006).

MaxEnt is a presence-background method, based on the principle of maximum entropy, that does not require true absence data and estimates a relative probability of presence (Chavy et al., 2019; Elith et al., 2006; Guillera-Arroita et al., 2014; Wang and Stone, 2019). MaxEnt does not use true absence data but rather a 'background' sample of environments in the region of interest (Guillera-Arroita et al., 2014), that we set at 10,000 background points for this study. Estimation of the final model was enhanced by jackknifing across ten replications. MaxEnt generates a map in which the likelihood of species presence is scaled to the range 0.0 to 1.0. This is also called a relative probability of presence (Wang and Stone, 2019). Assumptions of this approach include that the detectability of the species does not vary with the covariates associated with occurrence (Yackulic et al., 2013). The post-processing of the map was done in ArcMap.

Regularization factors 0.5, 1, 1.5, 2 and 3 were tested, bracketing the MaxEnt default value of 1. The feature class combinations used were linear (L), , linear quadratic (LQ), hinge (H), linear quadratic hinge (LQH), and, linear quadratic hinge and product (LQHP). Every regularization factor was combined with each feature class (Elith et al., 2011). Spatial jackknife was done by partitioning the model area into three parts: each model run used two parts and was tested on the third. Each of the five feature class combinations were tested with each of the five regularization factors, with three jackknife partitions and ten replicates, yielding 750 different model runs for each species.

The selection of the most suitable model was carried out following the recommendation of the SDM ToolBox software, which selects the model evaluating the following criteria, in descending order: omission rate (Phillips et al., 2006), area under the curve (AUC) (Elith et al., 2011) and the complexity of the feature class.

Field validation of niche models

To validate the above niche models, a rapid assessment method was used (Moreno et al., 2020), which included phlebotomine catches in five townships in Pueblo Rico and two in Rovira. In each township, houses were randomly selected for phlebotomine sampling. To evaluate the composition and abundance of phlebotomine sand flies, three Centers for Disease Control and Prevention (CDC) incandescent light traps were located in each house over two consecutive nights. One trap was placed indoors, and another two at different points 10 m from the house. All traps were placed at a height of 1.5 m and were active from 18:00 to 06:00 hours. The catches were recovered from the traps early the following morning and immobilized with triethylamine (TEA: 04885-1; Fisher Scientific, Pittsburgh, PA). The collection mesh of the CDC traps was introduced into a black plastic bag for the immobilization of the insects captured using a piece of cotton moistened with 1 ml of TEA for 15 minutes. The phlebotomines were separated from other insects and stored in plastic tubes with 70% ethanol during transportation to the laboratory (Ferro et al., 2011). Each tube was labeled with the code of the house, the date of capture and the location of the trap. The data were recorded on paper forms and later entered into Microsoft Excel.

Species identification

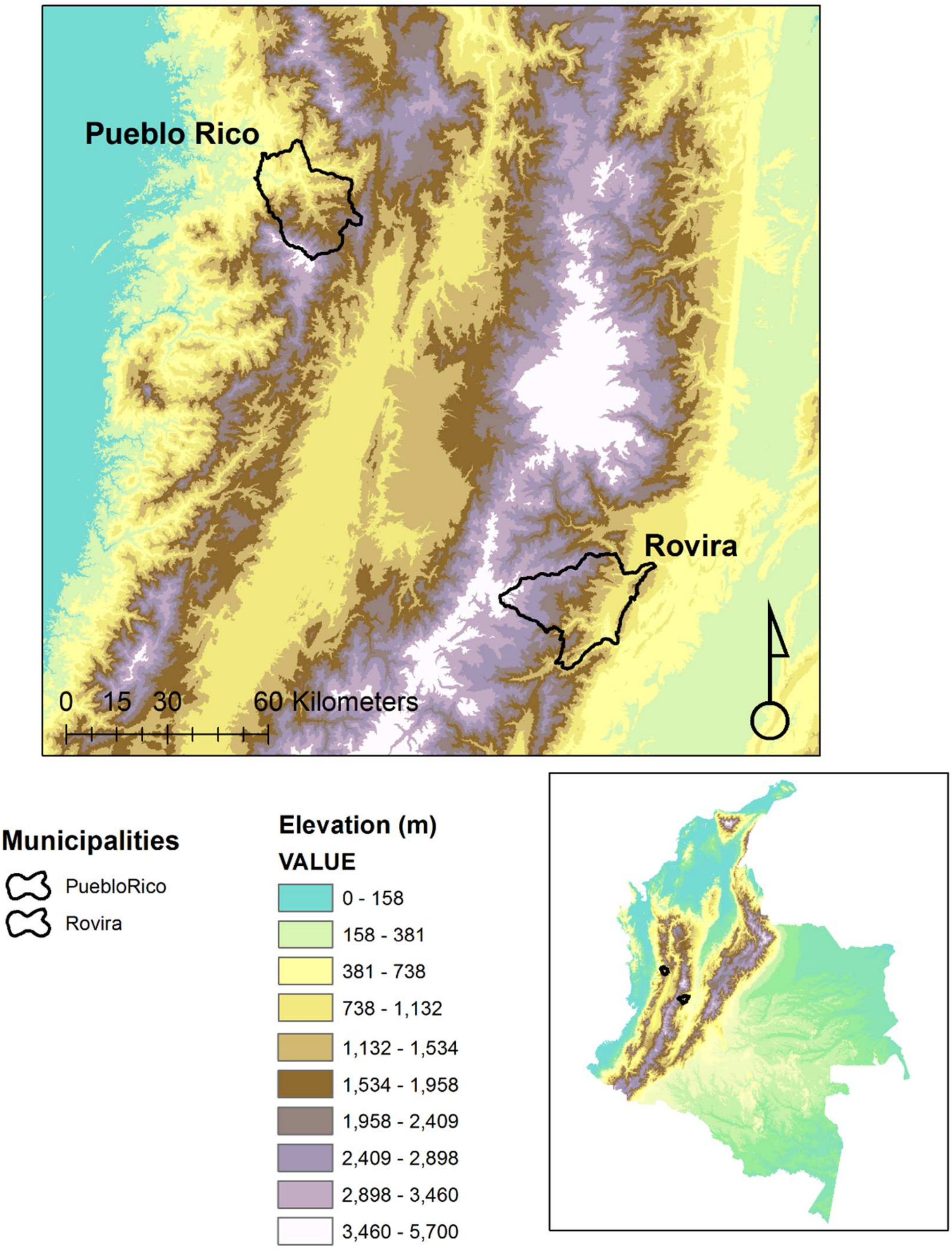
Taxonomic determination of phlebotomines was carried out after clarification in 10% KOH solution at room temperature for 24 hours (Young & Duncan 1994 and Galati 2003). The determination was based on the external morphological characteristics of the male specimens, beginning with the structure of the genitalia of male specimens, the number and distribution of spines and setae in the gonostil and form of gonocoxite, and paramere shape. Additionally, the coloration of the scutum and dorsal region of the head was taken into account, as well as the length ratio of 3rd and 5th palpomeres, presence and form of antennal ascoids, and distribution of teeth in the cibarium, according to the dichotomous keys described above. For female specimens, species determination was based mainly on: the pharyngeal and spermathecal armature, the distribution and number of teeth in the cibarium, length of individual ducts and common duct of the spermathecae, and length of labrum.

Conditional autoregressive (CAR) spatial Poisson regression to identify areas with higher risk of transmission

Numbers of cases of cutaneous leishmaniasis by township were provided by the local Health Department (Secretaría de Salud Municipal) in Microsoft Excel files, as were maps of the townships in ArcGIS (shapefile) format. In Rovira the case totals from 2012 to 2015 were provided, as were the estimated populations of each township. Cases from outside the municipality were removed, and two townships were aggregated to be commensurate with the map. In Pueblo Rico, case totals from 2013 to 2019 were provided, and populations were provided both by the local Health Department and by Sisben (Sistema de Identificación de Potenciales Beneficiarios de Programas Sociales) in Microsoft Excel files. Two townships were merged to be commensurate with the map. Each of two further townships which were in the map, but lacked population information, was merged with the neighboring township with which it shared the longest border. Part of the Tatamá National Natural Park is in this municipality and was excluded because it has no official population. Only rural townships were included, with the main town of each municipality being excluded.

For each municipality, the average value from the niche model was calculated for each township, and this was used as a predictor in a Poisson conditional autoregressive (CAR) spatial model (Lawson, 2013), with the number of cases being the response variable and the logarithm of the township population as an offset. The logarithm of the population density was also included, as this had been found to be associated with the incidence in a previous study (Valderrama-Ardila et al., 2010). This model was fitted by Markov chain Monte Carlo in the OpenBUGS software to the rural *veredas*. The model is similar to that in Moraga's section 6.4.1 (Moraga, 2019), including a spatially structured (CAR) term, and a term for unstructured random variation. The priors for the precision (inverse variance) of the spatial terms was Gamma (0.5, 0.0005), and for the precision of the regression coefficient of the it was Gaussian(0, 10^{-5}), with 10^{-5} again being the precision. Estimation was based on 200,000 iterations, thinned by 10, after a burn-in of 100,000 iterations. Convergence was assessed visually.

Figure 1. Map of Colombia



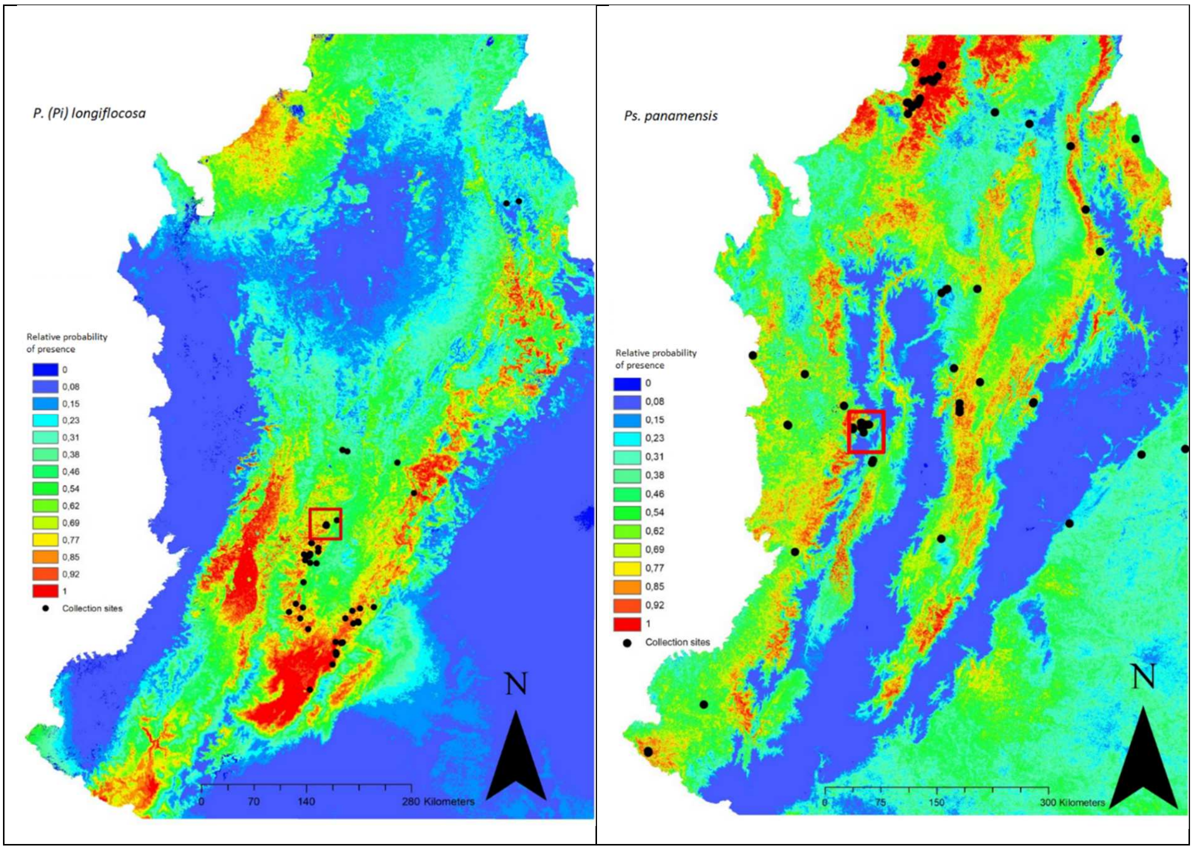
Results

Niche modeling

The number of geolocated records was 82 for *Pi. longiflocosa* and 91 for *Ps. panamensis*. Of the total of 173, 130 were from published and 43 from unpublished sources. They were rarefied to 46 records of *Pi. longiflocosa* and 57 of *Ps. panamensis*. The niche models are shown in Figure 2. The jackknife analysis of the *Pi. longiflocosa* model showed that the model explained approximately 75% of the variation in the data. The most informative variables were annual precipitation (40% explained) and the temperature seasonality (21%). Each of the first three principal components of the NDVI data showed little importance (<10%). The analysis gave an area under the curve (AUC) of 0.858. *Pi. longiflocosa* has zones of high probability of presence on the eastern slopes of the western, central and eastern cordilleras, and also in the higher parts of the valley of the river Magdalena, between the central and eastern cordilleras, and on some parts on the northern Caribbean littoral.

For *Ps. panamensis*, the AUC was 0.819, and jackknife analysis explained at most 51% of the variation in the data. Elevation contributed most (27%), although this was well below its contribution to the *Pi. longiflocosa* model. As before, no principal component of NDVI explained more than 15%. *Ps. panamensis* has high relative probability of presence on smaller sections of the slopes of the cordilleras, and in larger areas to the north of the area of interest, close to the Caribbean coast.

Figure 2. Relative probability of presence from the national-level niche models for *Pintomyia (Pifanomyia) longiflocosa* (left panel) and *Psychodophygus panamensis* (right panel).



Field validation of niche models

Table 1 shows the results of the catches carried out in the two municipalities to validate the niche models. As expected, *Pi. (Pif.) longiflocosa* was found in Rovira but not Pueblo Rico, and the opposite for *Ps. panamensis*.

Spatial modelling of the incidence of cutaneous leishmaniasis

Table 2 and Figure 3 show the results of the spatial (CAR) models in Rovira and Pueblo Rico. In both municipalities, the relative probability of presence from the niche model tended to increase with the incidence of notified cases, although the credible intervals included the null value. In Rovira, some of the highest values from the niche model coincided with townships of high notified incidence, although the niche model values were also high in the higher altitude townships, towards the north and east, which tended to have low reported incidence. In Pueblo Rico, again the high values from the niche model overlapped those of reported incidence, although the zone of high reported incidence extended south into townships at higher altitudes.

The results contributed to the selection of townships to be involved in future phases of the project (community-based surveillance). In Rovira, these results led us to develop future phases of the project in the south east fringe, as well as a smaller area of elevated predicted risk, more to the north and east (Figure 3E). In Pueblo Rico, the results contributed to the selection of townships in the north and north-east. In both sites, local health authorities and site personnel were involved in the selection of townships, since other factors such as geographic barriers (particularly in Pueblo Rico), distance to health posts and safety of the field personnel were also considered.

Discussion

In order to better direct techniques such as capture-recapture analysis of under-estimation of CL (Mosleh et al., 2008; Yadon et al., 2001), we used niche mapping in order to identify areas where ACL is likely to be an appreciable problem, on the rationale that presence of a vector is a necessary but not sufficient condition for transmission to occur. This targeting was necessary because our resources do not allow a large-scale survey such as that carried out in Fars province of Iran (Kazerooni et al., 2018).

We compiled published and unpublished geolocated records of phlebotomine presence in order to estimate niche models. As well as on some parts of the slopes of the three Andean cordilleras, *Pi. longiflocosa* has a high relative probability of presence in the higher parts of the Magdalena valley, and likewise for *Ps. panamensis* to the north of the area of interest, close to the Caribbean coast. We were able to include more records than the previous work of Ferro et al., who also carried out niche modelling of these and other phlebotomine species in Colombia (Ferro et al., 2015). In particular, for *Pi. longiflocosa* we have 46 records (after rarefication) as opposed to 33. In terms of niche modelling of these two species outside of Colombia, we are only aware of the work of Sanchez et al. in Venezuela, which included only *Ps. panamensis* (Sanchez et al., 2015). In fact, in the Global Biodiversity Information Facility, *Pi. longiflocosa* has been registered only for Colombia, while *Ps. panamensis* has been registered also for Ecuador, México, Honduras, Guyana and Venezuela (GBIF Secretariat, 2021). Sanchez et al found that three precipitation WorldClim layers were the most important predictors. By contrast, we found that the most important predictor was elevation, and included NDVI, while Sanchez et al. did not. Association between NDVI and precipitation may explain why the latter did contribute importantly to presence of this species in our model. More work has been done on niche modelling of phlebotomine species in Brazil (Fonseca et al., 2021; Meneguzzi et al., 2016; Peterson and Shaw, 2003), but these are different species and in different ecological conditions to the Andean region of the current study.

These niche models were able to predict the actual presence of the selected species in field catches carried out according to a rapid assessment methodology. This field method may help to optimize surveillance and allow health agencies to target cutaneous leishmaniasis treatment and prevention strategies in Colombia. Limitations of the study include the use of data on vectors and disease incidence which do not completely overlap in time.

The analysis of incidence of American cutaneous leishmaniasis is based on existing surveillance data and is therefore subject to the very limitations on data quality that motivated the project to measure the burden more accurately. The notified cases are likely to be genuinely ACL ("true positives"), since only parasitologically confirmed ACL cases are reported to the surveillance system (Instituto Nacional de Salud, 2017): the greater concern is under-estimation of the incidence in other townships ("false negatives"). The current analysis aimed to identify townships where the vector could reasonably be inferred, while also being relatively close to those with confirmed presence of the disease.

The selection of areas for future work was based only partly on the current results, and also on considerations such as accessibility and community stakeholders interested in taking part. Accessibility here is relative: the candidate townships to the southwest of Rovira are still 4 hours on unpaved road from the capital of the municipality. Accessibility affects the ability to offer adequate treatment to incident ACL cases identified through active surveillance, which usually consists of parenteral drugs, and requires local infrastructure. The selection of areas also took into account knowledge and perspectives of local health authorities, including the vector control programs, and community leaders.

In conclusion, we produced large scale niche maps for two phlebotomine vectors of cutaneous leishmaniasis in Colombia, and the results were used to guide the selection of townships for further study of under-ascertainment and under-reporting. This combination of entomological, ecological and epidemiological methods is applicable to other vector-borne diseases.

Table 1 Summary of field catches

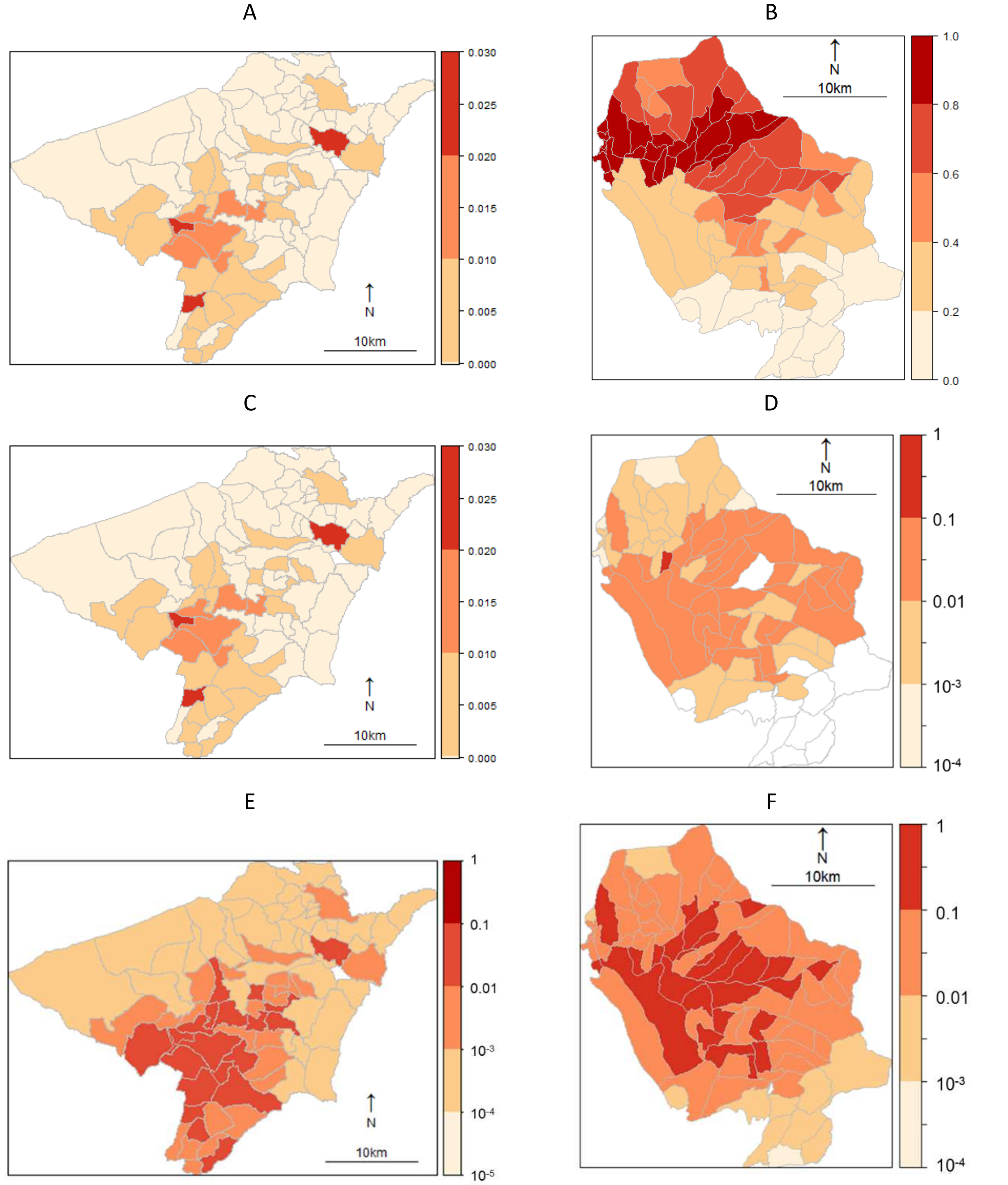
Township	No. houses	Elevation[a] (m)	Latitude (degrees)[a]	Longitude (degrees)[a]	Location	Nights	*Pi.* (*Pif.*) *longiflocosa*			*Ps. panamensis*		
							Niche model[b]	Female	Male	Niche model[b]	Female	Male
Rovira (Tolima)												
Florida	4	1481	4.201	-75.346	Intradomicile	4	0.914	19	-	0.478	-	-
					Peridomicile	4		16	-		-	-
Guadual	3	1444	4.274	-75.314	Intradomicile	3	0.864	-	-	0.408	-	-
					Peridomicile	3		6	1		-	-
Pueblo Rico (Risaralda)												
San Juan	3	465	5.345	-76.096	Intradomicile	3	0.844	-	-	0.889	-	-
					Peridomicile	3		-	-		12	5
Similito	6	686	5.372	-76.073	Intradomicile	6	0.808	-	-	0.915	6	-
					Peridomicile	6		-	-		14	18
Santa Rita	3	473	5.401	-76.100	Intradomicile	3	0.035	-	-	0.903	1	-
					Peridomicile	2[b]		-	-		15	5
Yoraudó	3	351	5.344	-76.204	Intradomicile	3	0.076	-	-	0.936	-	-
					Peridomicile	3		-	-		32	2

[a]Averaged over the houses in each township. [b]Value of the relative probability of presence for the species in question. [c]Trapping was not done at one house on one night.

Table 2 Summary of spatial regression (CAR) analyses

	Rovira	Pueblo Rico
Phlebotomine species of interest	*Pi.* (*Pif.*) *longiflocosa*	*Ps. panamensis*
Number of rural townships	81	84
Total rural population	15,997	10,852
Years	2012-2015	2013-2019
Total number of cases / person-years (rate per thousand person-years)	226/63,988 (3.53)	936/75,964 (12.3)
Incidence rate ratio per increase of 0.1 in the relative probability of presence (95% credible interval)	1.45 (0.49-4.21)	1.29 (0.98-1.66)
Incidence rate ratio per 10-fold increase in population density (95% credible interval)	1.10 (0.89-1.38)	0.65 (0.37-1.14)

Figure 3. Township-level analysis. Top row: relative probability of presence from the niche models for *Pi.* (*Pif.*) *longiflocosa* in Rovira (panel A) and *Ps. panamensis* in Pueblo Rico (panel B). Middle row: annual incidence per person-year of CL from surveillance for Rovira (panel C) and Pueblo Rico (panel D, white areas have zero incidence). Bottom row: risk maps (spatial regression) for Rovira (panel E) and Pueblo Rico (panel F).



Acknowledgments

We thank the health authorities of Risaralda (Secretaría Departamental de Salud de Risaralda) and Tolima (Secretaría Departamental del Tolima) for their support in project activities and key information for planning field activities. In particular, the coordinators of the vector control programs of Risaralda, Shirley Botero and Tolima, Eduardo Lozano. We also thank to the local health authorities of Rovira (Secretaria Municipal de Salud de Rovira, led at the time by Rocio Rodriguez) and the personnel of the Hospital San Vicente; in Pueblo Rico, the local public health team (Dirección Local de Salud) and the personnel of Hospital San Rafael. We also thank the personnel supporting the field catches, including Luis Ernesto Ramirez from CIDEIM, and Nora Vasquez from the Secretaria de Salud of Risaralda, Ludy Marcela Delgado Monroy of Secretaria de Salud of Dosquebradas, Risaralda.

Funding

This study was financed by NIAID-NIH, Award number U19AI129910.

Conflict of interest

The authors state that they have no conflict of interest to declare.

Supplementary material

Table S1. Locations of phlebotomines catches described in existing publications, or previously unpublished.

Bibliography

Agudelo Chivatá, N.J., 2019. Informe de Evento Leishmaniasis Cutánea, Periodo Epidemiológico XIII. Colombia, 2019, Instituto Nacional de Salud, Bogotá.

Akhoundi, M., Kuhls, K., Cannet, A., Votypka, J., Marty, P., Delaunay, P., Sereno, D., 2016. A Historical Overview of the Classification, Evolution, and Dispersion of Leishmania Parasites and Sandflies. PLoS Negl Trop Dis 10, e0004349.

Alvar, J., Velez, I.D., Bern, C., Herrero, M., Desjeux, P., Cano, J., Jannin, J., den Boer, M., WHO Leishmaniasis Control Team, 2012. Leishmaniasis Worldwide and Global Estimates of Its Incidence. PloS ONE 7.

Bejarano, E.E., Sierra, D., Vélez, I.D., 2003. New findings on the geographic distribution of the verrucarum group (Diptera: Psychodidae) in Colombia. Biomédica 23, 341-350.

Bejarano, E.E., Sierra, D., Vélez, I.D., 2007. Two new records of Lutzomyia (Diptera: Psychodidae) for the department of Risaralda, Colombia. Revista Colombiana de Entomología 33, 43-44.

Bejarano, E.E., Uribe, S., Pérez-Doria, A., Egurrola, J., Dib, J.C., Poter, C., 2015. Nuevos hallazgos de flebotomíneos (Diptera: Psychodidae) en la Sierra Nevada de Santa Marta, Colombia. Acta Biológica Colombiana 20, 221-224.

Brown, J.L., 2014. SDM toolbox: a python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. Methods in Ecology and Evolution 5, 694-700.

Burza, S., Croft, S.L., Boelaert, M., 2018. Leishmaniasis. Lancet 392, 951-970.

Chavy, A., Ferreira Dales Nava, A., Luz, S.L.B., Ramirez, J.D., Herrera, G., Vasconcelos Dos Santos, T., Ginouves, M., Demar, M., Prevot, G., Guegan, J.F., de Thoisy, B., 2019. Ecological niche modelling for predicting the risk of cutaneous leishmaniasis in the Neotropical moist forest biome. PLoS Negl Trop Dis 13, e0007629.

Contreras-Gutiérrez, M.A., Vélez, I.D., Porter, C., Uribe, S.I., 2014. An updated checklist of Phlebotomine sand flies (Diptera: Psychodidae: Phlebotominae) from the Colombian Andean coffee-growing region. Biomédica 34, 483-498.

Davies, C.R., Reithinger, R., Campbell-Lendrum, D., Feliciangeli, D., Borges, R., Rodriguez, N., 2000. The epidemiology and control of leishmaniasis in Andean countries. Cad Saude Publica 16, 925-950.

Elith, J., Graham, C.H., Anderson, R.P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J.M., Peterson, A.T., Phillips, S.J., Richardson, K.S., Scachetti-Pereira, R., Schapire, R.E., Soberón, J., Williams, S., Wisz, M.S., Zimmermann, N.E., 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29, 129-151.

Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E., Yates, C.J., 2011. A statistical explanation of MaxEnt for ecologists. Diversity and Distributions 17, 43-57.

Ferro, C., Lopez, M., Fuya, P., Lugo, L., Cordovez, J.M., Gonzalez, C., 2015. Spatial Distribution of Sand Fly Vectors and Eco-Epidemiology of Cutaneous Leishmaniasis Transmission in Colombia. PLoS One 10, e0139391.

Ferro, C., Marin, D., Gongora, R., Carrasquilla, M.C., Trujillo, J.E., Rueda, N.K., Marin, J., Valderrama-Ardila, C., Alexander, N., Perez, M., Munstermann, L.E., Ocampo, C.B., 2011. Phlebotomine vector ecology in the domestic transmission of American cutaneous leishmaniasis in Chaparral, Colombia. Am J Trop Med Hyg 85, 847-856.

Fonseca, E.D.S., Guimaraes, R.B., Prestes-Carneiro, L.E., Tolezano, J.E., Rodgers, M.S.M., Avery, R.H., Malone, J.B., 2021. Predicted distribution of sand fly (Diptera: Psychodidae) species involved in the transmission of Leishmaniasis in Sao Paulo state, Brazil, utilizing maximum entropy ecological niche modeling. Pathogens and global health 115, 108-120.

Galati, E.A.B., Galvis-Ovallos, F., Lawyer, P., Leger, N., Depaquit, J., 2017. An illustrated guide for characters and terminology used in descriptions of Phlebotominae (Diptera, Psychodidae). Parasite 24, 26.

GBIF Secretariat, 2021. Global Biodiversity Information Facility, Copenhagen

Guillera-Arroita, G., Lahoz-Monfort, J.J., Elith, J., 2014. Maxent is not a presence–absence method: a comment on Thibaud et al. Methods in Ecology and Evolution 5, 1192-1197.

Instituto Nacional de Salud, 2014. Protocolo de Vigilancia en Salud Pública: Leishmaniasis. Instituto Nacional de Salud, Grupo de Enfermedades Transmisibles, Bogotá.

Instituto Nacional de Salud, 2017. Protocolo de Vigilancia en Salud Pública: Leishmaniasis. Instituto Nacional de Salud, Equipo de Enfermedades Transmitidas por Vectores, Bogotá.

Kazerooni, P.A., Fararouei, M., Nejat, M., Akbarpoor, M., Sedaghat, Z., 2018. Under-ascertainment, under-reporting and timeliness of Iranian communicable disease surveillance system for zoonotic diseases. Public Health 154, 130-135.

Lago, J., Silva, J.A., Borja, L., Fraga, D.B.M., Schriefer, A., Arruda, S., Lago, E., Carvalho, E.M., Bacellar, O., 2019. Clinical and histopathologic features of canine tegumentary leishmaniasis and the molecular characterization of Leishmania braziliensis in dogs. PLoS Negl Trop Dis 13, e0007532.

Lawson, A., 2013. Bayesian Disease Mapping: Hierarchical Modeling in Spatial Epidemiology, Second ed. Chapman & Hall/CRC, Boca Raton, FL.

Lopez, M., Erazo, D., Hoyos, J., Leon, C., Fuya, P., Lugo, L., Cordovez, J.M., Gonzalez, C., 2021. Measuring spatial co-occurrences of species potentially involved in Leishmania transmission cycles through a predictive and fieldwork approach. Sci Rep 11, 6789.

Meneguzzi, V.C., Santos, C.B., Leite, G.R., Fux, B., Falqueto, A., 2016. Environmental Niche Modelling of Phlebotomine Sand Flies and Cutaneous Leishmaniasis Identifies Lutzomyia intermedia as the Main Vector Species in Southeastern Brazil. PLoS One 11, e0164580.

Moraga, P., 2019. Geospatial Health Data: Modeling and Visualization with R-INLA and Shiny. Chapman & Hall/CRC, Boca Raton.

Morales Ortegón, D.F., Castaño González, C.S., Lozano Guarín, E.A., Vallejo Londoño, H.d.J., 2004. Descripción de la epidemia de leishmaniasis cutánea en Chaparral y San Antonio, Tolima, 2003 y 2004. Inform Quinc J Epidemiol Nac 9, 180-186.

Moreno, M., Ferro, C., Rosales-Chilama, M., Rubiano, L., Delgado, M., Cossio, A., Gomez, M.A., Ocampo, C., Saravia, N.G., 2015. First report of Warileya rotundipennis (Psychodidae: Phlebotominae) naturally infected with Leishmania (Viannia) in a focus of cutaneous leishmaniasis in Colombia. Acta Trop 148, 191-196.

Moreno, M., Guzmán-Rodríguez, L., Valderrama-Ardila, C., Alexander, N., Ocampo, C.B., 2020. Land use in relation to composition and abundance of phlebotomines (Diptera: Psychodidae) in five foci of domiciliary transmission of cutaneous leishmaniasis in the Andean region of Colombia. Acta Tropica 203, 105315.

Morrison, A.C., Ferro, C., Morales, A., Tesh, R.B., Wilson, M.L., 1993. Dispersal of the sand fly Lutzomyia longipalpis (Diptera: Psychodidae) at an endemic focus of visceral leishmaniasis in Colombia. J Med Entomol 30, 427-435.

Mosleh, I.M., Geith, E., Natsheh, L., Abdul-Dayem, M., Abotteen, N., 2008. Cutaneous leishmaniasis in the Jordanian side of the Jordan Valley: severe under-reporting and consequences on public health management. Trop Med Int Health 13, 855-860.

Mutinga, M.J., Kamau, C.C., Basimike, M., Mutero, C.M., Kyai, F.M., 1992. Studies on the epidemiology of leishmaniasis in Kenya: flight range of phlebotomine sandflies in Marigat, Baringo District. East African medical journal 69, 9-13.

Ocampo, C.B., Ferro, M.C., Cadena, H., Gongora, R., Perez, M., Valderrama-Ardila, C.H., Quinnell, R.J., Alexander, N., 2012. Environmental factors associated with American cutaneous leishmaniasis in a new Andean focus in Colombia. Trop Med Int Health 17, 1309-1317.

Pérez, M., Ocampo, C.B., Valderrama-Ardila, C., Alexander, N., 2016. Spatial modeling of cutaneous leishmaniasis in the Andean region of Colombia. Memórias do Instituto Oswaldo Cruz 111, 433-442.

Peterson, A.T., Shaw, J., 2003. Lutzomyia vectors for cutaneous leishmaniasis in Southern Brazil: ecological niche models, predicted geographic distributions, and climate change effects. Int J Parasitol 33, 919-931.

Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190, 231–259.

Ramirez, J.D., Hernandez, C., Leon, C.M., Ayala, M.S., Florez, C., Gonzalez, C., 2016. Taxonomy, diversity, temporal and geographical distribution of Cutaneous Leishmaniasis in Colombia: A retrospective study. Sci Rep 6, 28266.

Roque, A.L., Jansen, A.M., 2014. Wild and synanthropic reservoirs of Leishmania species in the Americas. Int J Parasitol Parasites Wildl 3, 251-262.

Sanchez, I., Liria, J., Feliciangeli, M.D., 2015. Ecological niche modeling of seventeen sandflies species (Diptera, Psychodidae, Phlebotominae) from Venezuela. International Journal of Zoology 2015.

Santaella, J., Ocampo, C.B., Saravia, N.G., Mendez, F., Gongora, R., Gomez, M.A., Munstermann, L.E., Quinnell, R.J., 2011. Leishmania (Viannia) infection in the domestic dog in Chaparral, Colombia. Am J Trop Med Hyg 84, 674-680.

Santamaría, E., Ponce, N., Zipa, Y., Ferro, C., 2006. Presencia en el peridomicilio de vectores infectados con Leishmania (Viannia) panamensis en dos focos endémicos en el occidente de Boyacá, piedemonte del valle del Magdalena medio, Colombia. Biomédica 26, 82-94.

Segurado, P., Araújo, M.B., Kunin, W.E., 2006. Consequences of spatial autocorrelation for niche-based models. Journal of Applied Ecology 43, 433-444.

Stockman, A.K., Beamer, D.A., Bond, J.E., 2006. An evaluation of a GARP model as an approach to predicting the spatial distribution of non-vagile invertebrate species. Diversity and Distributions 12, 81-89.

Travi, B.L., Jaramillo, C., Montoya, J., Segura, I., Zea, A., Goncalves, A., Velez, I.D., 1994. Didelphis marsupialis, an important reservoir of Trypanosoma (Schizotrypanum) cruzi and Leishmania (Leishmania) chagasi in Colombia. Am J Trop Med Hyg 50, 557-565.

Valderrama-Ardila, C., Alexander, N., Ferro, C., Cadena, H., Marin, D., Holford, T.R., Munstermann, L.E., Ocampo, C.B., 2010. Environmental risk factors for the incidence of American cutaneous leishmaniasis in a sub-Andean zone of Colombia (Chaparral, Tolima). Am J Trop Med Hyg 82, 243-250.

Vivero, R.J., Bejarano, E.E., Pérez-Doria, A., Flórez, F., Estrada, L.G., Torres, C., Muskus, C., 2009. Nuevos registros de flebotomíneos (Diptera: Psychodidae), con el hallazgode Lutzomyia longipalpis (Lutz & Neiva, 1912), en los alrededores de la Ciudad de Sincelejo, Colombia. Biota Neotrópica 9, 277.

Wang, Y., Stone, L., 2019. Understanding the connections between species distribution models for presence-background data. Theoretical Ecology 12, 73-88.

World Health Organization, Colombia, Leishmaniasis information resource, Geneva.

World Health Organization, 2020. Leishmaniasis, Fact sheets, Geneva.

Yackulic, C.B., Chandler, R., Zipkin, E.F., Royle, J.A., Nichols, J.D., Campbell Grant, E.H., Veran, S., 2013. Presence-only modelling using MAXENT: when can we trust the inferences? Methods in Ecology and Evolution 4, 236-243.

Yadon, Z.E., Quigley, M.A., Davies, C.R., Rodrigues, L.C., Segura, E.L., 2001. Assessment of Leishmaniasis notification system in Santiago del Estero, Argentina, 1990-1993. Am J Trop Med Hyg 65, 27-30.

Zambrano, P., 2007. Comportamiento de la leishmaniasis en Colombia. Biomédica 27, 83-84.