

Chagas Disease Risk in Texas

Sahotra Sarkar^{1,2*}, Stavana E. Strutz¹, David M. Frank², Chissa-Louise Rivaldi¹, Blake Sissel¹, Victor Sánchez-Cordero³

¹ Section of Integrative Biology, University of Texas, Austin, Texas, United States of America, ² Department of Philosophy, University of Texas, Austin, Texas, United States of America, ³ Instituto de Biología, Universidad Nacional Autónoma de México, México City, México

Abstract

Background: Chagas disease, caused by *Trypanosoma cruzi*, remains a serious public health concern in many areas of Latin America, including México. It is also endemic in Texas with an autochthonous canine cycle, abundant vectors (*Triatoma* species) in many counties, and established domestic and peridomestic cycles which make competent reservoirs available throughout the state. Yet, Chagas disease is not reportable in Texas, blood donor screening is not mandatory, and the serological profiles of human and canine populations remain unknown. The purpose of this analysis was to provide a formal risk assessment, including risk maps, which recommends the removal of these lacunae.

Methods and Findings: The spatial relative risk of the establishment of autochthonous Chagas disease cycles in Texas was assessed using a five-stage analysis. 1. Ecological risk for Chagas disease was established at a fine spatial resolution using a maximum entropy algorithm that takes as input occurrence points of vectors and environmental layers. The analysis was restricted to triatomine vector species for which new data were generated through field collection and through collation of post-1960 museum records in both México and the United States with sufficiently low georeferenced error to be admissible given the spatial resolution of the analysis (1 arc-minute). The new data extended the distribution of vector species to 10 new Texas counties. The models predicted that *Triatoma gerstaeckeri* has a large region of contiguous suitable habitat in the southern United States and México, *T. lecticularia* has a diffuse suitable habitat distribution along both coasts of the same region, and *T. sanguisuga* has a disjoint suitable habitat distribution along the coasts of the United States. The ecological risk is highest in south Texas. 2. Incidence-based relative risk was computed at the county level using the Bayesian Besag-York-Mollié model and post-1960 *T. cruzi* incidence data. This risk is concentrated in south Texas. 3. The ecological and incidence-based risks were analyzed together in a multi-criteria dominance analysis of all counties and those counties in which there were as yet no reports of parasite incidence. Both analyses picked out counties in south Texas as those at highest risk. 4. As an alternative to the multi-criteria analysis, the ecological and incidence-based risks were compounded in a multiplicative composite risk model. Counties in south Texas emerged as those with the highest risk. 5. Risk as the relative expected exposure rate was computed using a multiplicative model for the composite risk and a scaled population county map for Texas. Counties with highest risk were those in south Texas and a few counties with high human populations in north, east, and central Texas showing that, though Chagas disease risk is concentrated in south Texas, it is not restricted to it.

Conclusions: For all of Texas, Chagas disease should be designated as reportable, as it is in Arizona and Massachusetts. At least for south Texas, lower than 30°N, blood donor screening should be mandatory, and the serological profiles of human and canine populations should be established. It is also recommended that a joint initiative be undertaken by the United States and México to combat Chagas disease in the trans-border region. The methodology developed for this analysis can be easily exported to other geographical and disease contexts in which risk assessment is of potential value.

Citation: Sarkar S, Strutz SE, Frank DM, Rivaldi C-L, Sissel B, et al. (2010) Chagas Disease Risk in Texas. PLoS Negl Trop Dis 4(10): e836. doi:10.1371/journal.pntd.0000836

Editor: Paul Andrew Bates, Lancaster University, United Kingdom

Received: April 5, 2010; **Accepted:** September 2, 2010; **Published:** October 5, 2010

Copyright: © 2010 Sarkar et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The authors have no support or funding to report.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: sarkar@mail.utexas.edu

Introduction

Chagas disease, a result of infection by the hemoflagellate kinetoplastid protozoan, *Trypanosoma cruzi*, remains an important public health threat in Latin America [1] with an estimated 16–18 million human incidences and 45 000 deaths annually [2]. While the Southern Cone Initiative [3–6] has interrupted the transmission of Chagas disease in several South American countries, and similar efforts are being attempted for other countries of Latin America [5–7], the disease is also endemic in the southern United

States, especially in Texas where it is yet to be designated as reportable [8–13]. Moreover, patterns of human migration into Texas from endemic regions of Latin America may contribute to an increase in the risk of Chagas disease [11,14,15]. Because the disease has a chronic phase that may last for decades, during which parasitaemia falls to undetectable levels [7], the extent of human infection in the southern United States is at present unknown. Based entirely on demographics, Hanford *et al.* [10] provided an extreme estimate of more than 1 million infections for the United States with 267 000 of them being in Texas. However,

Author Summary

Chagas disease is endemic in Texas and spread through triatomine insect vectors known as kissing bugs, assassin bugs, or cone-nosed bugs, which transmit the protozoan parasite, *Trypanosoma cruzi*. We examined the threat of Chagas disease due to the three most prevalent vector species and from human case occurrences and human population data at the county level. We modeled the distribution of each vector species using occurrence data from México and the United States and environmental variables. We then computed the ecological risk from the distribution models and combined it with disease incidence data to produce a composite risk map which was subsequently used to calculate the populations expected to be at risk for the disease. South Texas had the highest relative risk. We recommend mandatory reporting of Chagas disease in Texas, testing of blood donations in high risk counties, human and canine testing for Chagas disease antibodies in high risk counties, and that a joint initiative be developed between the United States and México to combat Chagas disease.

Bern and Montgomery [11] have criticized that estimate for using the highest possible values for all contributory factors; they provide a more credible lower estimate of 300–167 for the entire United States. Infections of zoonotic origin only add to the number of infections of demographic origin and the risk of disease. So far infected vectors or hosts have been found in 82 of the 254 counties of Texas (see Table S1) though only four vector-borne human autochthonous cases have been confirmed [16]. The parasite incidence rate in vectors in Texas has been reported as being >50% [12,16,17] which is higher than the ~41% reported from Phoenix, Arizona [13], but lower than the 81% reported from Guaymas in northwestern México [18]. In contrast to Texas, the disease is reportable in Arizona and Massachusetts even though there has not been an autochthonous human case in either state, compared to the four in Texas. The other autochthonous human cases confirmed for the United States are from California [19], Tennessee [20], and Louisiana [9].

The main human Chagas disease cycle consists of the parasite, *T. cruzi*, being transferred from a mammalian reservoir to a human host through a vector. However, infection through blood transfusion, organ transplants, and the ingestion of infected food are also recognized mechanisms of concern; infections may also occur through congenital transmission [7,21,22]. A large variety of mammal species can serve as reservoirs for *T. cruzi* including humans and dogs [7], which means that a focus on reservoirs would not be effective for disease control. Given that no vaccine exists [23], efforts to control the disease must focus on vector control [7]. Consequently, risk assessment for Chagas disease must focus primarily on the ecology and biogeography of vector species and the incidence of the parasite, besides human social and epidemiological factors [5].

This analysis consists of a five-stage risk assessment for Chagas disease in Texas: (i) an ecological risk analysis using predicted vector distributions; (ii) an incidence-based risk analysis based on parasite occurrence; (iii) a joint analysis of ecology and incidence using formal multi-criteria analysis; (iv) such a joint analysis using a composite risk model; and (v) a computation of the relative expected exposure rate taking into account human population. The purpose of the complete analysis is to argue that there is sufficient widespread risk for Chagas disease in Texas to warrant it to be declared reportable and other measures be taken. The

analysis focuses primarily on the vector distributions but also uses available information on parasite incidence. If the number of human infections in Texas is as high as in the estimates noted earlier [10,11], then humans alone would constitute sufficient reservoirs in disease foci. Moreover, even if the number of human infections is much lower, there is compelling evidence that the disease has established itself in Texas in domestic and peridomestic cycles with canine reservoirs [16,17]. Thus, also given the abundance of wild zoonotic reservoirs in most of the state, including armadillos, coyotes, raccoons, opossums, and rodents of the genus *Neotoma*, the distribution of reservoirs is not likely to limit the occurrence or spread of the disease in Texas. This analysis assumes that competent reservoirs are present everywhere in Texas in sufficient densities to perpetuate or establish the disease cycle. Moreover, the peridomestic cycle makes human exposure to the parasite more likely than what would have been the case with only a sylvatic transmission cycle.

The vectors of Chagas disease are insects from the family Reduviidae, sub-family Triatominae, and in northern México and the United States, restricted to the genus *Triatoma*. Seven *Triatoma* species have been routinely collected in Texas: *Triatoma gerstaeckeri*, *T. sanguisuga*, *T. lecticularia*, *T. protracta*, *T. indictiva*, *T. rubida*, and *T. neotomae* [12]. (One specimen of *T. recurva* was reported from Brewster county in far southwestern Texas on the Mexican border in 1984 [24] but no further specimen has since been found in Texas; available records are restricted to Arizona and northwestern México.)

Using data from new field collections as well as museum records, this analysis begins by constructing species distribution models for the three most widely distributed *Triatoma* species in Texas: *T. gerstaeckeri*, *T. sanguisuga*, and *T. lecticularia*. All three species have been shown to be carriers of *T. cruzi* [12,25]. The other four *Triatoma* species were so rare (collected less than 10 times in total by any researcher in Texas since 2000) that they are presumed not to be important for establishing Chagas disease transmission cycles in the state. The species distribution models were constructed using a maximum entropy algorithm which relies on species occurrence (presence-only) records and environmental layers [26]. Such a modeling strategy, though using a genetic algorithm, has been previously used to model the distribution of *T. gerstaeckeri* in Texas [16], and a variety of triatomine species complexes for North America [27] though at a much coarser spatial resolution than this analysis which used cells with 1 arc-minute edges. The output from these models directly quantify habitat suitability for a species by computing the relative probability of its presence in each cell of the study area. These probabilities establish the potential distribution of a species (and are sometimes interpreted as providing an approximate ecological niche model [28,29]). The predicted distribution is obtained using biological information such as dispersal behavior and other constraints that limit the potential distribution.

These three species' distributions were used to generate a map of the probability of the occurrence of at least one triatomine vector species in each cell. This is the most basic ecological risk map: when these probabilities are low, there is little risk of Chagas disease occurrence through the major vectorial mode of transmission though disease may still occur through contaminated blood transfusion and, less likely, through parasite ingestion. (By "risk," throughout this paper, we will mean *relative risk*, that is, the risk in one cell compared to others throughout the area of interest.) When the ecological (relative) risk is high, other risk factors determine the likelihood of disease, including the abundance of vectors, the incidence of parasites, and anthropogenic features of the habitat, for instance, human behavioral patterns (including

habitation structure) [30,31]. Ecological risk maps of this kind have previously been used for this region to estimate the risk of the spread of leishmaniasis due to climate change [31]. The relevance of that work to the present analysis is that the disease agents for leishmaniasis are also kinetoplastid protozoans which share reservoirs with *T. cruzi* [32–36].

Independently, at the county level (which was the finest resolution at which data were available), a (relative) risk map based on parasite incidence in vectors, canine reservoirs, or humans was constructed using the Bayesian Besag-York-Mollié (BYM) model which is widely used in epidemiology [37]. This map was based on a spatial interpolation of risk from the number of parasite records from each county: it captures the idea that there is spatial correlation between disease incidences. The implications of the incidence-based risk map were combined with those of the basic ecological risk map in two ways: (i) a simple multi-criteria analysis (MCA) [38] was used to find the counties that were most at risk from both suitability for vector species and proximity to locations of parasite incidence; (ii) a multiplicative risk model was used to obtain a composite risk map for Chagas disease in Texas. Both sets of results were used to

prioritize counties for increased surveillance for the occurrence of *T. cruzi*.

Finally, the composite risk map was combined with the relative human population densities of the counties to produce a “relative expected exposure rate” risk map which provides a rough relative measure of potential extent of human exposure to Chagas disease. The entire risk analysis was used to recommend that Chagas disease be made reportable in Texas, that the blood supply be screened in south Texas, and that human and canine serological profiles be investigated in the same region.

Materials and Methods

Study Area

The study area was delimited at the south by the 14.13°N line of latitude along the México-Guatemala border, by the coast of continental México to the east and west, continued by the lines 117.19°W and 86.70°W within the United States and the line 40.61°N at the north, thus enclosing all the species' occurrence points (see Figure 1). It was divided into 1 819 462 cells at a resolution of 1 arc-minute. The average cell area was 2.95 km².

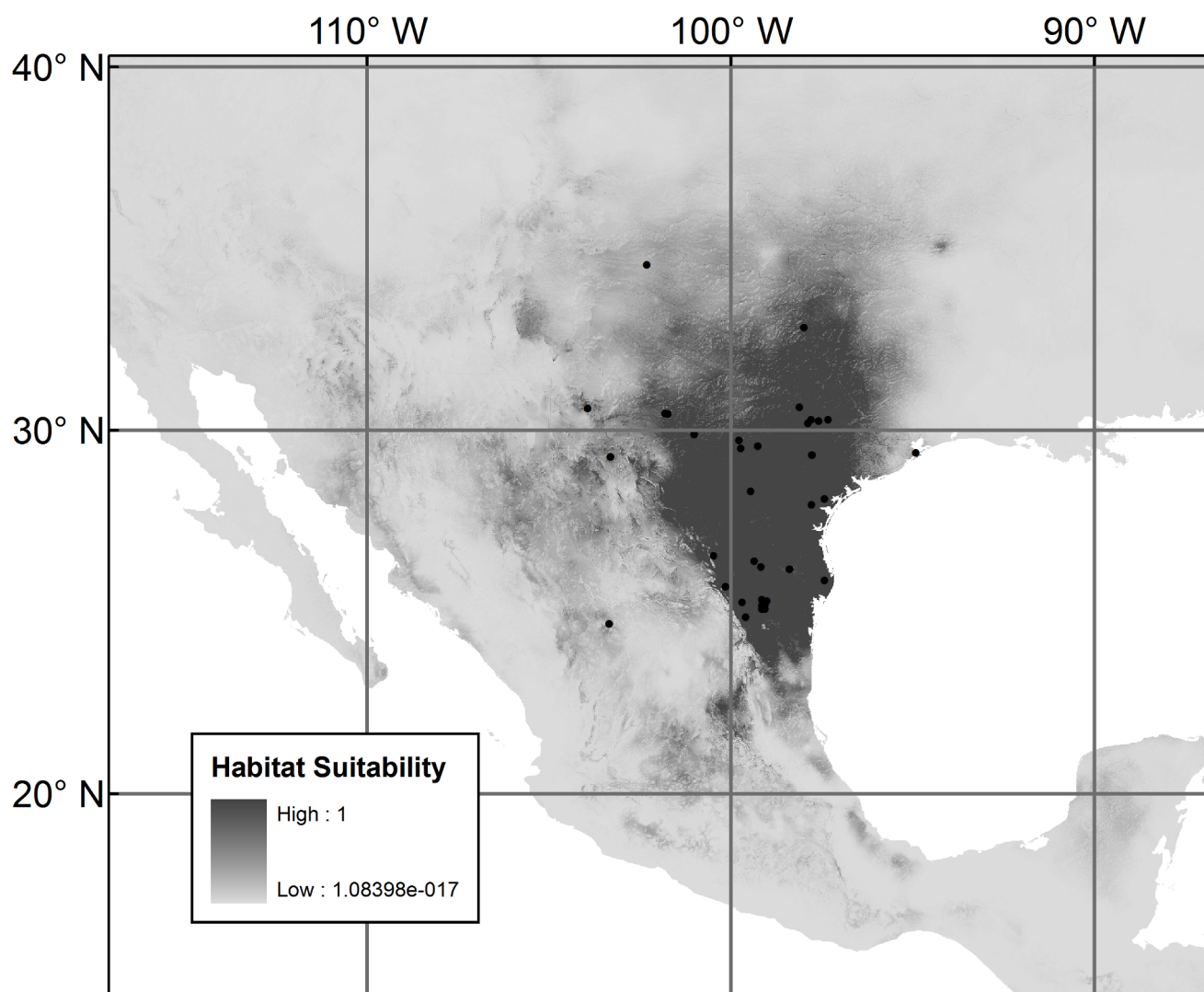


Figure 1. Species distribution model for *Triatoma gerstaeckeri*. The black dots show the occurrence points used for model construction. doi:10.1371/journal.pntd.0000836.g001

Species Distribution Models

Species distribution models were constructed for the three most important triatomine vector species in Texas [12]: *T. gerstaeckeri*, *T. lecticularia*, and *T. sanguisuga*.

Data. Triatomine species occurrence data were obtained from museum collections, other researchers, voluntary collectors (see Acknowledgments for more detailed information on all three categories), as well as organized surveys in Texas and northern México, the results of which will be reported separately in the ecological literature. Species were identified using the key of Lent and Wygodzinsky [39]. All data were entered in the Disease Vectors Database (www.diseasevectors.org; last accessed 28 February 2010; [40]) and were georeferenced using the MaNIS protocol (<http://manisnet.org/GeorefGuide.html>; last accessed 28 February 2010) which has been extensively developed and refined by ecologists for this purpose. (Table S2 shows the number of records that were available for each species.)

For modeling purposes, because of the spatial resolution of the analysis, only records with an estimated error less than 1 arc-minute were retained. Moreover, because the WorldClim environmental layers only average information since 1960, all pre-1960 records were excluded from this analysis. With one exception for *T. lecticularia* and two exceptions for *T. sanguisuga*, all records were post-1980. There were 74 records retained for *T. gerstaeckeri*, 23 for *T. sanguisuga*, and 11 for *T. lecticularia*; these records generated 35, 17, and 11 instances in different cells, respectively, at the spatial resolution of this analysis.

Because only post-1960 triatomine records were used for the species distribution models, parasite incidence records used in this analysis were also restricted to the same period. *T. cruzi* incidence data in Triatoma, canines, and humans were compiled from the literature using the citations of recent reviews [10,12,17] through a backward search of earlier reports until 1960. Records of parasite incidence in vectors and human and canine hosts were used; there was little reliable information on other hosts. (These data are summarized in Table S1.)

Human population data per county were obtained from the Texas State Data Center and Office of the State Demographer (http://txsdc.utsa.edu/tpepp/2008_txpopest_county.php; last accessed 4-March-2010). July 2008 population estimate data were used; these are the most recent estimates available for every county and are based on the 2000 census. Economic data for these counties were obtained from the United States Census Bureau [41].

Model Construction. The species distribution models were constructed from species' occurrence points and environmental layers using a maximum entropy algorithm. The Maxent software package (Version 3.3.4; [26]) was used to construct the models. Maxent has been shown to be robust for modeling species distributions from occurrence (presence-only) records for a large number of taxa [42]. Following published recommendations [26,43,44], Maxent was run with the threshold and hinge features and without duplicates so that there was at most one sample per pixel; linear, quadratic, and product features were used. The convergence threshold was set to a conservative 1.0×10^{-5} . For the AUC, that is, the area under the receiver operating characteristic (ROC) curve, averages over 100 replicate models were computed. For each model the test:training ratio was set to 40:60 following Phillips and Dudík [26] which means that models were constructed using 60% of the data and tested with the remaining 40%.

Two tests were used to assess model performance: (i) A conservative threshold of 0.9 was used for the test AUC. (An optimal model would have an AUC close to 1 while a model that

predicted species occurrences at random would have an AUC of 0.5. Published recommendations suggest using a minimum threshold of 0.7 [42].); (ii) For the eight internal training and test binomial tests performed by Maxent (two each for minimum presence, 10 percentile presence, equal sensitivity and specificity, maximum sensitivity plus specificity), on the average, a *p*-value < 0.05 was required.

The environmental layers used are listed in Table 1. These include four topographical variables (elevation, slope, aspect, and composite topographical index) and 15 bioclimatic variables. The latter were obtained from the WorldClim database (www.worldclim.org; last accessed 28 February 2010; [45]). However, of the standard 19 bioclimatic variables, four were excluded (mean temperatures of the wettest quarter, driest quarter, warmest quarter, and coldest quarter) because the layers contain discontinuities within the study area from Texas. These discontinuities seem to be artefacts introduced during the interpolation used to construct the layers. Elevation was obtained from the United States Geological Survey's Hydro-1K DEM data set (http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/gtopo30/hydro; last accessed 28 February 2010). Slope, aspect, and compound topographical index were derived from the DEM using the Spatial Analyst extension of ArcMap 9.3.

The use of a large number of environmental variables raises the possibility of over-fitting a model due to correlations between the explanatory variables (even though the algorithm in Maxent is designed to counteract such correlations). One sign of such over-fitting is a much lower AUC for the test data compared to the AUC for the training data. To judge the potential occurrence of this problem for the species distribution models, a second set of "simpler" models was constructed using the four topographic variables and only seven bioclimatic variables: the annual mean temperature, mean diurnal range, maximum temperature of the

Table 1. Environmental parameters for species distribution models.

Parameters
Annual Mean Temperature
Mean Diurnal Range
Isothermality
Temperature Seasonality
Maximum Temperature of Warmest Month
Minimum Temperature of Coldest Month
Temperature Annual Range
Annual Precipitation
Precipitation of Wettest Month
Precipitation of Driest Month
Precipitation Seasonality
Precipitation of Wettest Quarter
Precipitation of Driest Quarter
Precipitation of Warmest Quarter
Precipitation of Coldest Quarter
Elevation
Slope
Aspect
Compound Topographic Index

doi:10.1371/journal.pntd.0000836.t001

warmest month, minimum temperature of the coldest month, annual precipitation, precipitation of the wettest month, and precipitation of the driest month, which are all known to be of general ecological relevance. All other model parameters were uniform between the two sets. For each species, and each replicate model, the difference between the training AUC and the test AUC was calculated under each modeling choice resulting in two sets of 100 values for each species, one corresponding to the use of 19 environmental variables and the other to the use of 11 environmental variables. These data were not normally distributed (Shapiro test, $p < 10^{-4}$). For each of the three pairs of 100 models, subsequent use of the Mann–Whitney–Wilcoxon test did not permit distinguishing the mean values of the AUC difference (minimum $p > 0.5$). (All statistical computations were done in R.) Subsequently, models based on all 19 environmental variables were used for the rest of this analysis because they had higher test AUC values.

Probability of Triatomine Presence. The output from Maxent consists of relative suitability values between 0 and 1 which, when normalized, can be interpreted as the probability of occurrence of a species in a landscape cell. The probability that at least one triatomine species is present in a cell was computed as the complement of the probability that none is present. This computation assumed that the probability of the presence of each species is independent of that of the presence of another species. This assumption is reasonable because different species are often found at the same location and there is no evidence of competitive or other interactions between them [27].

Let the probability of the presence of at least one triatomine species in cell k be p_k and that of species i in cell k be p_{ik} . Then:

$$p_k = 1 - \prod_{i=1}^n (1 - p_{ik})$$

where n is the number of species. In this case there were three species, *T. gerstaeckeri*, *T. sanguisuga*, and *T. lecticularia*.

Risk Assessment

Components of Risk. The concept of risk is salient only in those circumstances in which there is a chance of some undesirable event happening. Consequently, two broad components of risk can be distinguished, the *probability* of the event (which is equally applicable to desirable and undesirable events) and its associated *cost* or *harm* or *incidence* (in the case of disease agents) [46]. Risk assessment requires the quantification of both components through adequate choice of parameters. If a variety of scenarios are available, both these parameters are ideally separately computed to produce risk curves and surfaces [46]. However, in the situation being considered here, a portfolio of scenarios was not available. Consequently, the two parameters were combined in a multiplicative model to calculate the relative expected exposure rate (see below).

Both of these components have several (sub-)components themselves. Most importantly, the probability of a disease cycle establishment event will be determined by at least the ecology of the vector, reservoir, or host species, depending on the type of disease (which may make one or more of these elements irrelevant), and on the probability of occurrence of the parasite. Both these parameters were computed separately and then the results compounded in two different ways.

Risk assessment proceeded in five stages:

1. Ecological risk was computed to quantify probable exposure to the parasite (disease factor) due to the ecological suitability of a cell for disease vectors. This process generated an *ecological relative risk map*.
2. Incidence-based risk was computed to quantify probable exposure to the parasite because of physical contact, that is, due to spatial proximity of one cell to another in which a parasite is known to occur. This process generated an *incidence-based relative risk map*.
3. The probability of (human) exposure to a parasite depends on both the ecological risk, which quantifies the probability of vector presence, and the incidence-based risk, which quantifies the likely presence of a parasite. For a more complete risk assessment, even ignoring reservoirs, the effects of these two factors must be jointly analyzed. The first method for this purpose that was used was a multi-criteria analysis in which each factor was taken to be a criterion.
4. The second method of joint analysis was to use a composite risk model that quantitatively combines the ecological and incidence-based risk. This process generated a *composite relative risk map*.
5. The relative expected exposure rate in humans was computed using the composite risk model and the human population in each cell. This process generated a *relative expected exposure rate risk map*.

Ecological Risk. For this analysis, ecological risk was quantified by the probability of the presence of a triatomine vector in each cell, that is, p_k as defined earlier. Since the rest of the analysis had to be performed at the county level, because data at any finer resolution was not available, the average p_k was computed for each of the 254 counties of Texas. In principle, the ecological risk would also incorporate the probability of presence of reservoirs. Such a model of ecological risk has been implicitly [29] and explicitly [30,31] used to define the minimal ecological conditions required for a disease to spread and establish an autochthonous cycle in a region. If the ecological risk is low, such an establishment is highly unlikely. If that risk is high, then other factors, some of which were modeled below, become critical for establishment.

Incidence-Based Risk. It was presumed that incidence-based risk depended on the proximity of a cell to one in which the parasite is present, that is, on spatial correlation. Based on this assumption, incidence-based relative risk was computed using the Besag-York-Mollié (BYM) model [47,48] which has been widely used for this purpose [37]. This is a Bayesian spatial model which assumes a Poisson sampling distribution for the number of incidences, y_i in any area, i . This is appropriate if incidences are rare, as was true in our case. If:

$$y_i \sim \text{Poisson}(S_i),$$

where S_i is the relative risk as measured by incidence, the model assumes a Gaussian Markov random field for $\ln S_i$ [37,49]:

$$\ln S_i = \alpha + u_i + v_i$$

where α is an average level of relative risk, u_i is the correlated heterogeneity, and v_i the uncorrelated heterogeneity. Finally, a conditional autoregressive (CAR) model was used for the u_i [49]. This model was selected because of its superior performance, as measured by the deviance information criterion (DIC) [50], over a range of data sets in a recent review [37]. The two other models with similar superior performance were more complex semi-parametric models which would have been difficult to

parameterize credibly given the lack of more comprehensive data for Chagas disease in Texas.

Model input consisted of an incidence score (0, 1, 2, or 3) for each cell which increased linearly with the number of different types (triatomine vectors, canine hosts, human hosts) in which the parasite was found in a cell (county). Ideally, the exact number of parasites found should be incorporated into the computation but data at that level of detail were not available. Model computations were performed in WinBUGS [51] using code modified from Lawson *et al.* [49]. The CAR model required the specification of a prior, parameterized by the precision, γ , of a multi-variate normal distribution. An uninformative prior (with $\gamma = 5 \times 10^{-4}$) was used because there was no prior information regarding any of the parameters. Model computations were initiated with a “burn-in” of 10 000 iterations followed by a subsequent 200 000 iterations to ensure convergence. Convergence was judged by the lack of autocorrelation after 100 000 and 200 000 iterations as well as inspection of smooth posterior probability densities for all parameters after 100 000 and 200 000 iterations. Model output consisted of a Bayesian posterior probability of relative risk of incidence for each county.

Multi-Criteria Analysis. A wide variety of multi-criteria analysis techniques exist [38]; surprisingly, very few have been used in epidemiological contexts. Since the composite risk model discussed next already quantitatively compounds the ecological and incidence-based risk, both interpreted as probabilities, multi-criteria techniques used here were restricted to those that rely entirely on qualitative (*ordinal* or comparative) rankings [52,53]. Because there was no basis for ordering the two criteria—ecological risk and incidence-based risk—the only method available that is consistent with standard utility theory was *dominance*. One alternative possibility (county, in this case) “dominates” another with respect to risk if it has either higher ecological or incidence-based risk and neither its ecological risk nor its incidence-risk is lower than that of the other alternative (county). The set of non-dominated alternatives is collectively at higher risk than the other alternatives in the sense that none of the other alternatives is worse off than all of the non-dominated alternatives according to every criterion.

However, the technique has well-known problems [52,54]. All counties that have the highest ecological relative risk or the highest incidence-based relative risk are bound to be non-dominated. To ameliorate this problem, this risk assessment was always used in this analysis along with the results of an analysis that quantitatively compounded these two types of risk. All multi-criteria analysis was done using the MultCSync software package [55].

Composite Risk. In contrast to the multi-criteria analysis, the second method of combining ecological risk and incidence-based risk used a multiplicative model to produce a single value of relative risk. Given that what is being computed is the probability component of risk, if both the ecological risk and incidence-based risk are being appropriately interpreted as probabilities (which is reasonable), then, if parasite incidence and vector occurrence are statistically independent, the multiplicative model is appropriate. However, vectors are responsible for introducing the parasite in a cell (even if, as in the case of Chagas disease in Texas, there are other major modes of introduction including migration and transport of contaminated blood [7]). Consequently, quantitative values produced by the multiplicative model must be treated with caution.

Relative Expected Exposure Rate. Because no other source of quantitative data was available, we used only one component contributing to the relative expected exposure rate: the potential population that would be exposed to Chagas disease in a county.

The populations of the 254 counties were normalized on a scale of 0 to 1 (with 1 being the rank of the county with the highest population). This scaled value was then multiplied by the composite risk which was interpreted as the probability of exposure to the parasite. The result, again normalized to lie between 0 and 1, was interpreted as a relative measure of expected exposure rate. Because of the reservations noted above about the composite risk model’s assumption of statistical independence between ecological risk and incidence-based risk, the quantitative estimates produced by this model must be treated with caution. However, it is well-known that the extent to which the housing in an area is built of concrete and similar material (rather than wood, adobe, *etc.*) negatively affects domestic human exposure to triatomines [7,8]. Spatially georeferenced quantitative data on housing construction in Texas was not available. However, there is some correlation between income levels and housing construction, with higher incomes correlated to concrete housing. Moreover, there is also a correlation between poverty and Chagas disease [7,14]. Data on median incomes for each county in Texas from the United States was obtained from the Census Bureau [41] and used to refine the results of the expected exposure rate model.

Results

Triatomine Biogeography

At the county level, our data collection and collation extended the known distribution of the seven triatomine species in Texas [12] in six cases: *T. gerstaeckeri* to Castro, Galveston, Gonzales, Lubbock, Parker, Victoria, Wilson, and Zapata counties, *T. indictiva* to Hays and Kinney counties, *T. lecticularia* to Bastrop, Blanco, Burleson, Lubbock, and Parker counties, *T. protracta* to Andrews, Bexar, and Terry counties, *T. rubida* to Crane and Upton counties, and *T. sanguisuga* to Bastrop and Kaufman counties. For *T. gerstaeckeri* and *T. lecticularia*, these results extend their ranges to northwest Texas for the first time. Over all, triatomines have now been recorded for 10 more counties (Andrews, Burleson, Castro, Crane, Galveston, Kaufman, Parker, Terry, Upton, and Wilson) than what was previously established. (Relevant maps are provided in the supplementary materials.)

Species Distribution Models

Model performance was judged using the test AUC, that is, the area under the receiver operating characteristic (ROC) curve and a set of internal binomial tests in the Maxent software package [26]. All three species produced test AUC values above the threshold of 0.9: averaged over the 100 replicate models, 0.979 for *T. gerstaeckeri*, 0.924 for *T. sanguisuga*, and 0.959 for *T. lecticularia*. On the average, all binomial tests were significant ($p < 0.05$). Because the models for *T. lecticularia* were constructed using only 11 presence records, the fact that its average AUC, besides being high, was greater than that of *T. sanguisuga*, suggests that model predictions are reliable. Moreover, a recent study indicates that models constructed using the Maxent algorithm are reliable so long as there are more than 10 presence records [56].

Figures 1, 2, and 3 show the three species distribution models, respectively. For *T. gerstaeckeri*, four out of 74 occurrence records fell in cells with habitat suitability ≤ 0.5 , for the other species, there was in each case one such record. The presence of a limited number of anomalous points is expected because species are often found in sub-optimal habitats, especially at the geographical margins of their ranges [54,57], as was the case with our points. The model for *T. gerstaeckeri* conforms with what is known about the distribution of the species from field records though it differs from the older model of Beard *et al.* [16] (see Discussion). There is

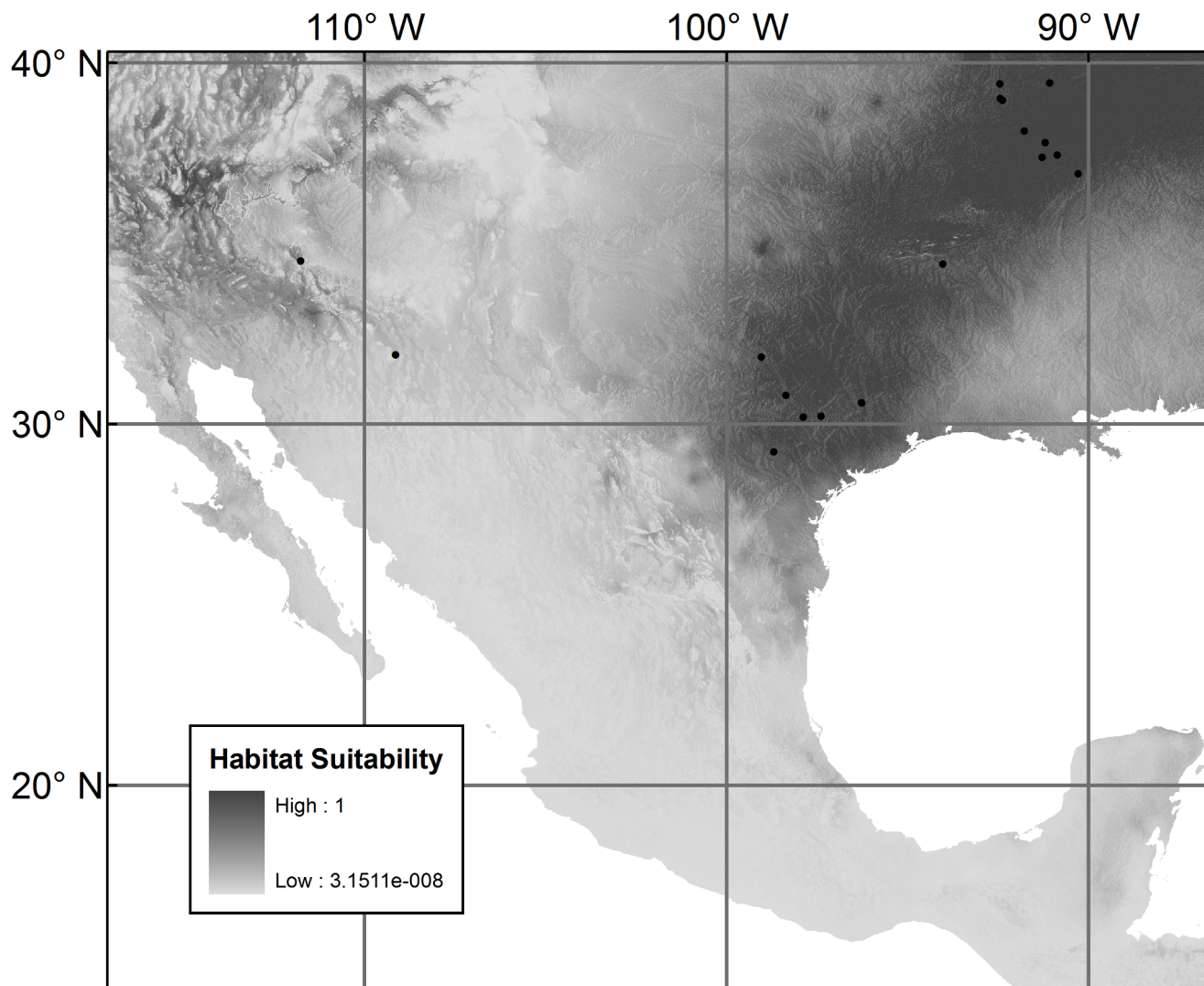


Figure 2. Species distribution model for *Triatoma sanguisuga*. The black dots show the occurrence points used for model construction. Much of the distribution is predicted to be in the eastern United States where the species has been collected from Texas to Florida. However, because of the two occurrence points to the west, a disjoint western distribution is also predicted and merits further investigation, as noted in the text. doi:10.1371/journal.pntd.0000836.g002

a high probability of occurrence (≥ 0.5) in the southern United States, especially in and around Texas, as well as in northeast México. For *T. sanguisuga*, the two occurrence points from the west (obtained from museum collections) have the effect of predicting suitable habitat in the western United States and México where the species has been collected in Arizona, California, and México [8,39,58]. *T. lecticularia* has a widespread predicted distribution along both coasts of North America but remains rare in collections along the western coast where all of our records came from México. Lent and Wygodzinsky [39] included New Mexico in the distribution of *T. lecticularia* but the provenance of those data remains unknown. There appears to be no recent record of the species in New Mexico and predicted highest habitat suitability is only 0.16.

Probabilistic Risk Analysis

Figure 4 shows the (relative) ecological risk map for the region including Texas. Figure 5 shows the incidence-based risk map for Texas, and Figure 6 the composite risk map. Table 2 shows the counties with the highest risk in each of these categories.

Compared to the incidence-based risk map, the composite risk map lowers the relative risk of counties to the far west and north of Texas because, even though *T. cruzi* has been reported in these areas, the habitat suitability for the triatomines remains low.

When we consider ecological risk and incidence-based risk separately in the multi-criteria dominance analysis, instead of compounding them to compute the composite risk, three counties are in the non-dominated set: Cameron, Jim Wells, and Nueces. All of these counties have incidences of *T. cruzi*. When this analysis is restricted to counties with no report as yet of *T. cruzi*, the non-dominated set consists of Goliad, Kenedy, and Wilson counties. This means that these three counties have high suitability for the presence of vector species as well as spatial contiguity to *T. cruzi* occurrences and are foci of special concern for Chagas disease.

When we consider together both non-dominated sets and the top five counties according to the ecological, incidence-based, and composite risk maps, eleven counties are selected (Bee, Bexar, Brooks, Cameron, DeWitt, Goliad, Hidalgo, Jim Wells, Kenedy, Kleberg, and Nueces) and all are in south Texas in an almost contiguous cluster starting at the Mexican border. When we

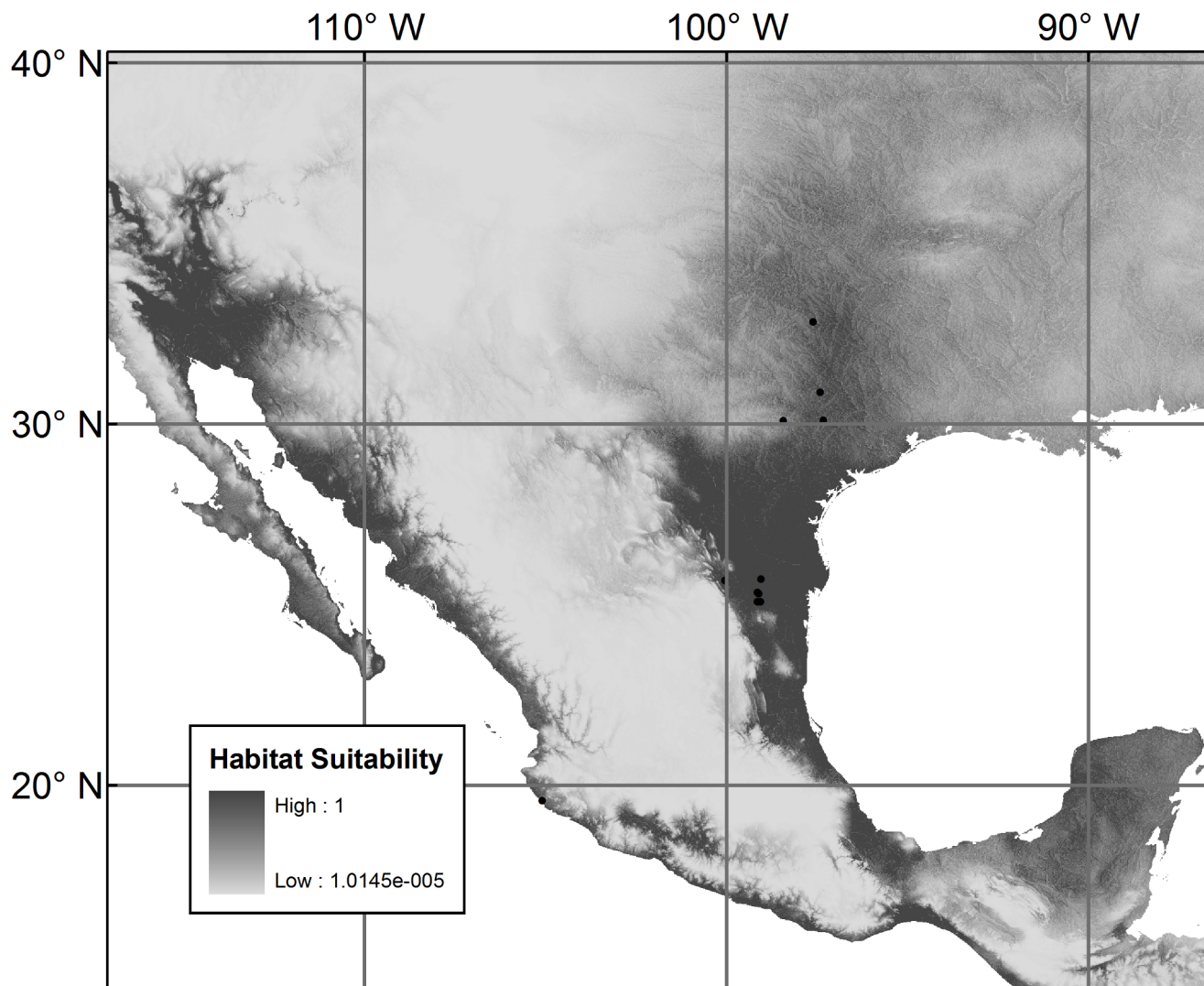


Figure 3. Species distribution model for *Triatoma lecticularia*. The black dots show the occurrence points used for model construction. The distribution is diffusely spread along both the east and west coasts.
doi:10.1371/journal.pntd.0000836.g003

include the top ten counties, an additional nine counties (Bandera, Dimmit, Frio, Guadalupe, Karnes, Live Oak, Medina, San Patricio, and Willacy) are selected; once again, all of these counties are from south Texas.

Relative Expected Exposure Rate Risk Map

Figure 7 shows the relative expected exposure rate at the county level. If the top five counties are added to the list of high risk counties, three counties outside south Texas are added: Dallas (north Texas), Harris (east Texas), and Travis (central Texas), because of the high human populations. If ten such counties are used, three additional counties outside south Texas are included (Collin and Tarrant in north Texas and Williamson in central Texas). Thus, consideration of human population density in a multiplicative model leads to a slightly more widespread attribution of risk than ecological and incidence-based risk. Nevertheless, the focus on south Texas remains strong. Moreover, only two of the high risk counties were ranked very low by median income using 2006 data from the United States Census Bureau [41]—Cameron and Hidalgo, which ranked 228 and 234,

respectively, out of 254 counties. Both of these are in south Texas. Low median income is likely to be indicative of relatively poorer living conditions and possible lack of concrete housing. Thus housing and living conditions, which were not quantitatively modeled, also implicate south Texas as the area of highest risk.

Discussion

For *T. gerstaeckeri*, our model predicted much more highly suitable habitat (high probability of occurrence) in central and east Texas and less in northwest Texas than the earlier model of Beard *et al.* [16] and is more consistent with the distribution map created by Kjos *et al.* [12] on the basis of collection records, including our extension of that distribution map with additional occurrence records (see Figure S1). The better performance of our model is presumably due to the availability of many more occurrence records from the United States for this species. Moreover, our model also predicted more suitable habitat for this species in México than the earlier model. This suggests an enhanced focus on this species for the control of Chagas disease in both Texas and north México.

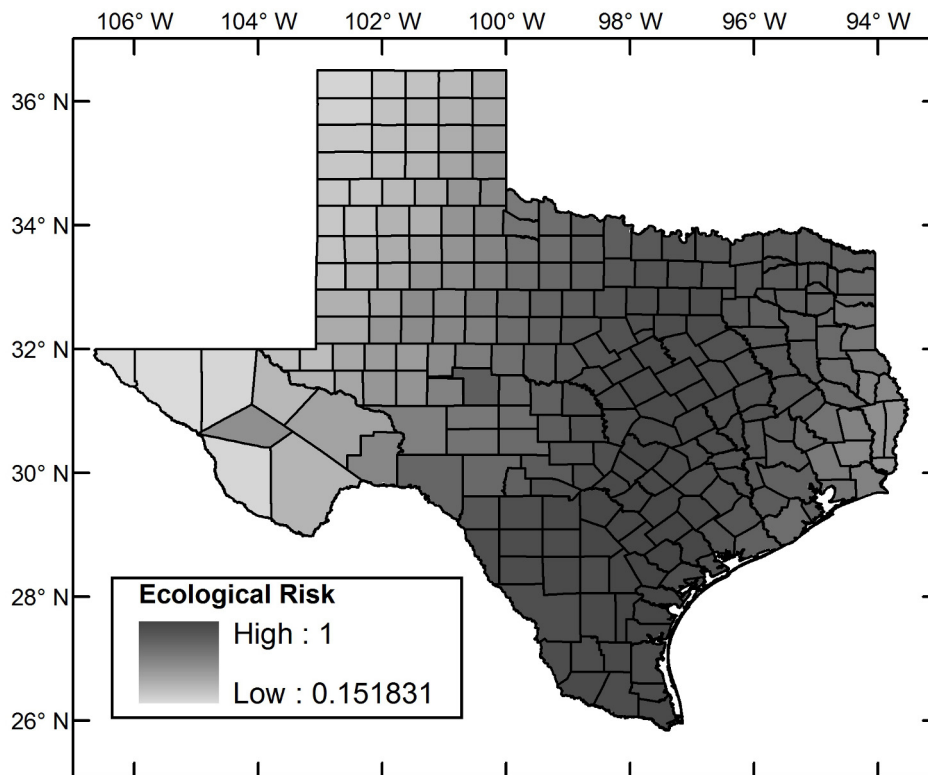


Figure 4. Ecological risk map for Chagas Disease in Texas.

doi:10.1371/journal.pntd.0000836.g004

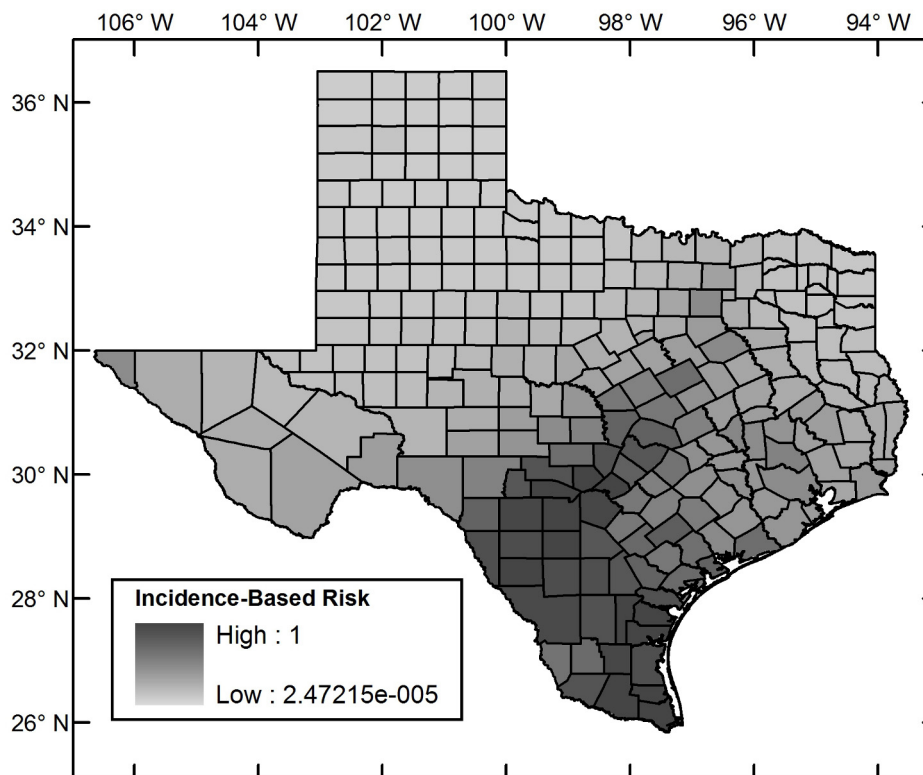


Figure 5. Incidence-based risk map for Chagas Disease in Texas.

doi:10.1371/journal.pntd.0000836.g005

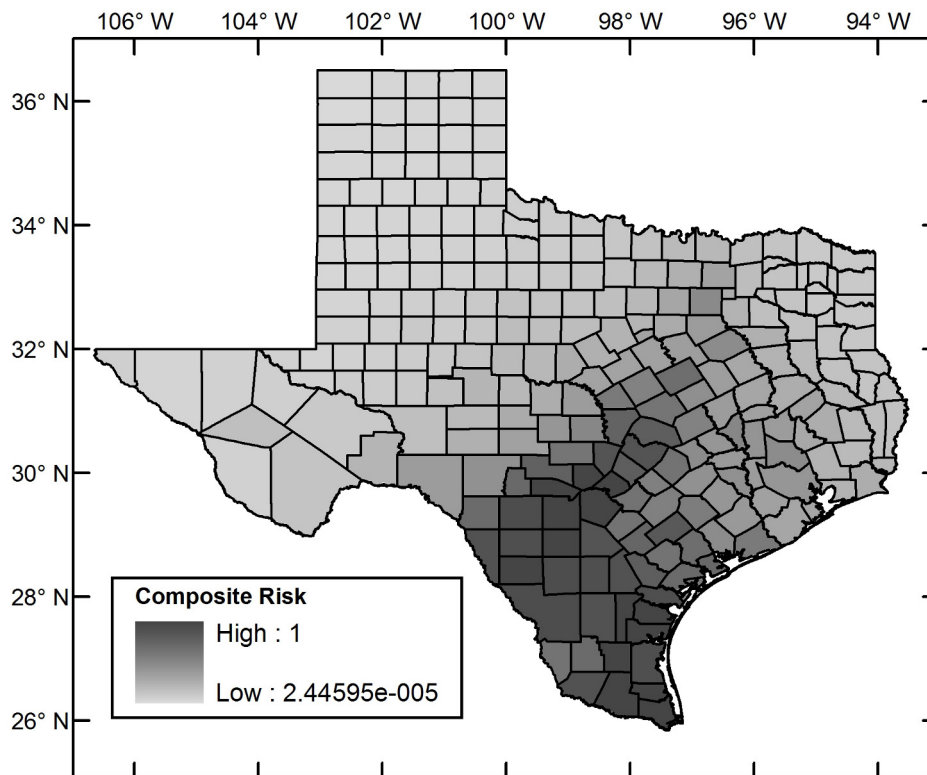


Figure 6. Composite risk map for Chagas Disease in Texas.

doi:10.1371/journal.pntd.0000836.g006

Data collection projects are in place for all triatomine species in the southern United States and in México over the next five years. (See Figures S2–S6 for new occurrence records for *T. indictiva*, *T. lecticularia*, *T. protracta*, *T. rubida*, and *T. sanguisuga*, respectively.) All model predictions will be tested in the field, in particular, the limits of the western distributions of *T. lecticularia* and *T. sanguisuga*. Part of the importance of model construction is to provide testable hypotheses that guide survey design, and the results reported here will be used for that purpose.

All risk maps point to one unsurprising but nevertheless important conclusion: to the extent that there is risk for Chagas disease in the United States, one important focus is south Texas.

Given the relative absence of reported autochthonous disease cases elsewhere (only three such cases have been confirmed outside Texas), it is the most important region of concern.

The methods used in this analysis do not provide a quantitative estimate of absolute risk or expected exposure rate, which is typically hard to produce in any context and the problem is amplified for diseases on which information is not being systematically collected. What it does provide is the relative risk in one unit compared to other spatial units at the county level. Nevertheless, the critical review of Bern and Montgomery [11] of all available data on Chagas disease in the United States strongly suggests that the absolute risk is also high.

The first three recommendations made below are geared towards obtaining the kind of data that would permit quantitative absolute risk assessment. However, the fourth recommendation, requiring the testing of blood donations, presumes that the absolute risk is high, and this needs some justification. Blood transfusion has been etiologically important as a source of Chagas disease along with immigration from areas of high Chagas disease incidence and an autochthonous cycle [11]. Currently, the American Association of Blood Banks (AABB) recommends such tests but does not require them. Testing began in 2007 using a test licensed by the United States Federal Drug Administration, in December 2006. Major laboratories that account for more than 65% of the total blood collected in the United States already carry out such tests (http://www.aabb.org/Content/Programs_and_Services/Data_Center/Chagas; last accessed 28 February 2010). The fourth recommendation is to extend coverage to the remaining 35% for the high risk areas of Texas. There are two arguments against mandatory testing: (i) the added cost; and (ii) the potential for false positive units to be removed from the blood supply. These costs must be compared to the benefits of testing. A

Table 2. High Chagas risk counties in Texas.

Ecological	Incidence-Based	Composite	Expected Exposure
Jim Wells	Cameron	Cameron	Bexar
Bee	Nueces	Nueces	Harris
Goliad	Kleberg	Kleberg	Hidalgo
De Witt	Hidalgo	Hidalgo	Dallas
Nueces	Jim Wells	Jim Wells	Travis
Wilson	Willacy	Willacy	Cameron
San Patricio	Dimmit	Medina	Nueces
Live Oak	Medina	Dimmit	Tarrant
Karnes	Bandera	Frio	Williamson
Guadalupe	Frio	Bandera	Collin

doi:10.1371/journal.pntd.0000836.t002

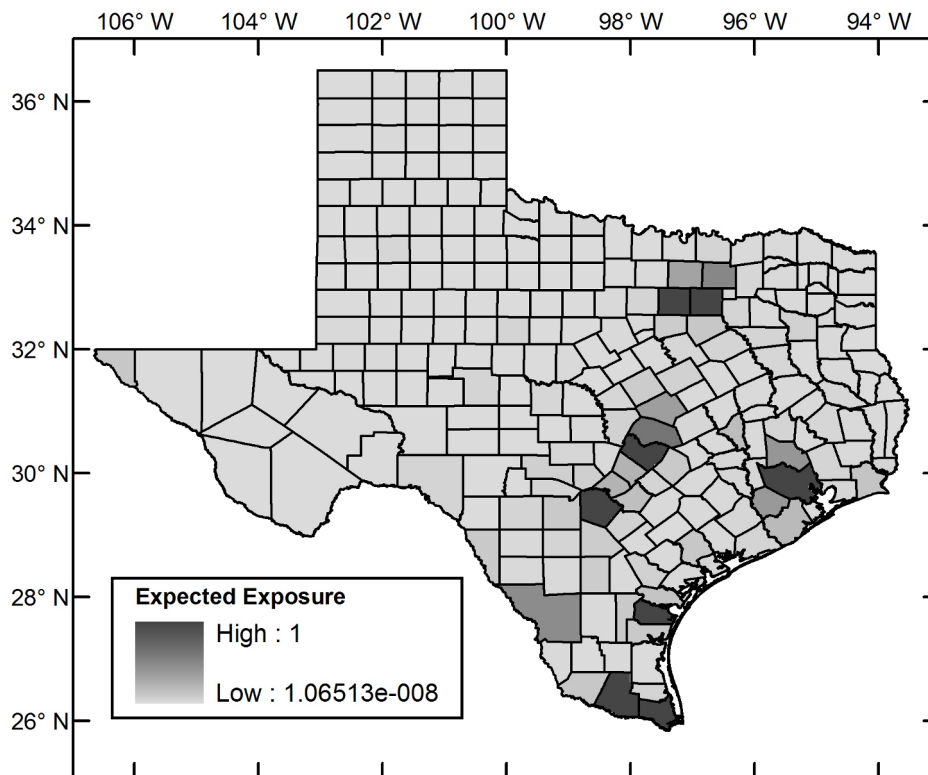


Figure 7. Expected exposure risk map for Chagas Disease in Texas.
doi:10.1371/journal.pntd.0000836.g007

simulation model developed by the Office of Biostatistics and Epidemiology, Center for Biologics Evaluation and Evaluation, United States Food and Drug Administration in 2009 predicted that, with no testing, there would be about 44 cases of transmission-induced Chagas disease in the United States each year (Richard Forshee, personal communication; www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/BloodVaccinesandOther-Biologics/BloodProductsAdvisoryCommittee/UCM155628.pdf). With 65% testing, that reduces to about 15 cases. These numbers are sufficiently high to suggest that areas with high relative risk, which would contribute disproportionately more cases, should have mandatory testing. Moreover, if testing is restricted to only high relative risk areas, rather than the entire blood supply, the cost and the potential loss of false positive test units are lower. Unfortunately, data to quantify these arguments are presently not available.

Recommendations

On the basis of this analysis, we make the following five recommendations:

- Given the risk assessment of this paper, it is imperative that Chagas disease be designated as reportable in Texas as it has been in Arizona since 2007 and Massachusetts since 2008. Additional systematic data acquisition on disease cases will permit more complete risk assessments, including those of absolute risk, in the future which, in turn, can guide the formulation of optimal public health initiatives to prevent disease. This is the most important recommendation from this analysis.
- For the same reasons, the serological status of human and canine populations should be investigated in south Texas, especially in counties south of 30° N of latitude, because of the enhanced relative risk for Chagas disease in that region. This is particularly relevant for the high risk counties identified by the multi-criteria analysis in which *T. cruzi* incidence is yet to be detected: Goliad, Kenedy, and Wilson, since it is highly likely that the parasite is present but unrecorded. Such an investigation has a dual purpose: (i) to prevent the occurrence of Chagas disease and related complications in positive individuals to the extent that it is possible, given the paucity of medical interventions available [7]; and (ii) to help formulate preventive strategies for the establishment (beyond current endemicity) and spread of Chagas disease. Humans and canines must both be monitored because of their competence as reservoirs for *T. cruzi* [7]. In high risk areas, as identified by multi-criteria analysis and by the composite risk model, vector species are very likely present in sufficient numbers; consequently, it is important to monitor reservoir species.
- Similarly, in order to prevent establishment and spread of Chagas disease, wild reservoir species, especially rodents (such as *Neotoma* species, which are confirmed highly competent reservoirs) merit investigation and monitoring in high risk areas.
- The testing of blood donors for antibodies to Chagas disease should be made mandatory at least in high risk areas (once again, at least in the counties south of 30° N of latitude in Texas). The reasons for making this recommendation were explained earlier.
- The Southern Cone Initiative [4,6] has interrupted the transmission of Chagas disease in several South American countries; similar efforts are being attempted for the countries of central America [6,7]. We recommend that a similar international initiative be undertaken by the United States and

México in the trans-border region (and not restricted to Texas within the United States) to control the spread of Chagas disease. Ideally, such an initiative should not be restricted to only Chagas disease. Rather, it should include the entire spectrum of vector-borne diseases capable of spreading across the border including dengue, leishmaniasis, tick-borne diseases, and West Nile virus, besides Chagas, to develop, for instance, integrated strategies of arthropod vector and, in some cases, reservoir control.

Limitations

Finally, beyond those discussed in the Materials and Methods section, eight other limitations of this analysis should be explicitly noted:

1. This analysis only partly incorporated risk from non-autochthonous Chagas disease occurrence and transmission in Texas. (It used data on confirmed parasite incidences of any provenance.) Important mechanisms include blood transfusion and immigration from areas of high Chagas disease incidence [10]. At present it remains impossible to provide reliable quantitative estimates of these risks. What remains important, however, even if the disease is introduced through these mechanisms, its potential to establish itself through an autochthonous cycle will depend on the probability computations reported here, in particular, the ecological risk model.
2. This analysis did not consider all vector species for Chagas disease in Texas. Moreover, it is possible that beyond these seven species that are known to occur in Texas (and besides *T. recurva*), other species found in northern México may also be present in Texas according to recent models [27]. However, given the absence or rarity of these species in the museum collections we investigated, it does not appear very likely that this possibility will be realized. Further, global factors such as climate change may result in species' range shifts. Assessing risk from this possibility was beyond the scope of this paper.
3. This analysis assumed that *T. gerstaeckeri*, *T. lecticularia*, and *T. sanguisuga* are equally competent as vectors of *T. cruzi*. While there is no evidence against such an equivalence, it has not been established through experiment. The assumption was made here in the absence of any alternative.
4. This analysis did not consider well-known differences between *T. cruzi* strains/ types in the etiology of Chagas disease [7,13]. Not enough information was available on the spatial distribution of the different strains anywhere in the study area to assess the significance of these differences.
5. It was assumed here that competent reservoirs for Chagas disease were always present and differences of reservoir occurrence between cells can be ignored. The reason for this assumption was the establishment of Chagas disease in domestic and peridomestic cycles and the abundance of wild reservoirs in non-urban regions of Texas. However, a spatially variable large parasite burden in wild reservoir species would require a revision of the relative risk estimates reported here. At present there is no evidence for the existence of such a factor.
6. The analysis did not consider vector population dynamics which have been shown to be important in the disease cycle [59]. Unlike many regions of South America, there was no information available for this study area to introduce these complications. Moreover, at the level of spatial resolution of this analysis the effects of population dynamics are probably not as important as habitat suitability of the vector species of Chagas disease.
7. The BYM model only takes spatial contiguity into account in the computation of spatial risk. It does not take quantitative spatial information (for instance, distances between points) into account. Unfortunately enough data did not exist to attempt a more sophisticated spatial analysis of risk.
8. Species distribution models only predict species' probable presence or absence, and do not predict abundance. This is an important limitation because epidemiological models typically predict that disease establishment is likely to depend on the abundance of vector (as well as reservoir, host, *etc.*) species. An implicit assumption of the modeling techniques used here is that, beyond a straightforward relationship with probability occurrence, habitat suitability (as predicted by species distribution models) is also correlated with a species' abundance. However, field studies to test this assumption are yet to be reported. For the Yucatán peninsula of México, there have been previous risk assessments based on modeled vector species abundance [60] (using *T. dimidiata*, the most important vector in Mesoamerica [7]) but the required data were not available for this study area.

Finally, one methodological innovation of this analysis should be noted since it is likely to be relevant to other contexts. This is the use of multi-criteria dominance analysis to identify high risk areas. In general, formal decision analysis has been surprisingly sparingly used in epidemiological contexts. However, techniques developed in that field can provide comprehensive decision support whenever complex decisions have to be analyzed. Here, we used one of the simpler multi-criteria techniques, the computation of non-dominated alternatives, to identify counties which are at high risk from Chagas disease even though the parasite has not yet been reported from them. Other, model-based techniques, selected the same region as areas of concern in south Texas. When used together to produce identical or similar results, these strategies lead to a more robust estimation of relative risk than otherwise possible. The strategy is fully general and can be exported to other contexts in which computing and mapping disease relative risk is of interest.

Supporting Information

Figure S1 New counties for *Triatoma gerstaeckeri*. The new counties are shown in dark gray and labeled by name.

Found at: doi:10.1371/journal.pntd.0000836.s001 (0.48 MB TIF)

Figure S2 New counties for *Triatoma indictiva*. The new counties are shown in dark gray and labeled by name.

Found at: doi:10.1371/journal.pntd.0000836.s002 (0.32 MB TIF)

Figure S3 New counties for *Triatoma lecticularia*. The new counties are shown in dark gray and labeled by name.

Found at: doi:10.1371/journal.pntd.0000836.s003 (0.40 MB TIF)

Figure S4 New counties for *Triatoma protracta*. The new counties are shown in dark gray and labeled by name.

Found at: doi:10.1371/journal.pntd.0000836.s004 (0.39 MB TIF)

Figure S5 New counties for *Triatoma rubida*. The new counties are shown in dark gray and labeled by name.

Found at: doi:10.1371/journal.pntd.0000836.s005 (0.34 MB TIF)

Figure S6 New counties for *Triatoma rubida*. The new counties are shown in dark gray and labeled by name.

Found at: doi:10.1371/journal.pntd.0000836.s006 (0.43 MB TIF)

Table S1 *Trypanosoma cruzi* incidence in Texas by county.
Found at: doi:10.1371/journal.pntd.0000836.s007 (0.07 MB PDF)

Table S2 Species records in the Disease Vectors Database.
Found at: doi:10.1371/journal.pntd.0000836.s008 (0.05 MB PDF)

Acknowledgments

This manuscript has substantially benefited from the criticisms of two referees, Richard Forshee and Chris Hall. For providing specimens thanks are due to Andrea Delong Amaya, Tim Brys, John Buckley, Val Bugh, William Calvert, Tammy Dettmann, Brush Freeman, Larry Gilbert, Kathleen O'Connor, Allen Palmer, and Edward Wozniak. For discussions, sharing data, specimen processing, and help with access to collections, thanks are due to John Abbott, Richard Forshee, Elaine Hanford, Patricia Illoldi-Rangel, José Alejandro Martínez-Ibarra, Angel Rodríguez Moreno, Jane O'Donnell, Eduardo A. Rebollar-Tellez, James Reddell, Carolina Reisenman, Ed Riley, Jim Schuermann, Kristin Simpson, Robert W. Sites,

Ophelia Wang, and Benjamin Zhan. For access to private properties for collecting specimens, thanks are due to Stuart Henry. The following institutions and collections were sources of triatomine specimen data: Biological Collections, University of Connecticut; Biospeleology Collection, Texas Natural Science Center (formerly Texas Memorial Museum), University of Texas at Austin; Colecciones Biológicas, Instituto de Biología, Universidad Nacional Autónoma de México; Enns Entomological Collection, University of Missouri; Laboratorio de Entomología Médica, Universidad Autónoma de Nuevo León; Texas A&M University Insect Collection; University of Texas Brackenridge Field Laboratory. Thanks are due to all the curators and researchers who made these resources available.

Author Contributions

Conceived and designed the experiments: SS. Performed the experiments: SS SES DMF. Analyzed the data: SS SES DMF CLR BS VSC. Contributed reagents/materials/analysis tools: SS SES DMF. Wrote the paper: SS SES DMF.

References

- Bastien JW (1998) The Kiss of Death: Chagas' Disease in the Americas. Salt Lake City: University of Utah Press.
- Kirchoff LV (2003) Current public health concerns. In: Tyler KM, Miles MA, eds. American Trypanosomiasis. Dordrecht: Kluwer. pp 157–162.
- Schofield CJ, Dias JCP (1999) The Southern Cone Initiative against Chagas disease. *Advances in Parasitology* 42: 1–27.
- Schmunis GA (1999) Iniciativa del Cono Sur. In: Schofield CJ, Ponce C, eds. Proceedings of the Second International Workshop on Population Biology and Control of Triatominae. Mexico City: INDRE. pp 26–31.
- Schofield CJ, Jannin J, Salvatella R (2006) The future of Chagas disease control. *Trends in Parasitology* 22: 583–588.
- Moncayo A, Silveira AC (2009) Current epidemiological trends for Chagas disease in Latin America and future challenges in epidemiology, surveillance and health policy. *Memórias do Instituto Oswaldo Cruz* 104(S1): 17–30.
- World Health Organization (2002) Control of Chagas Disease: Report of the Second WHO Expert Committee. Geneva: World Health Organization. 109 p.
- Navin T, Roberto RR, Juranek DD, Limpakarnjanarat K, Mortenson E, et al. (1985) Human and sylvatic *Trypanosoma cruzi* infection in California. *American Journal of Public Health* 75: 366–369.
- Dorn PL, Perniciaro L, Yabsley MJ, Roellig DM, Balsamo G, et al. (2007) Autochthonous transmission of *Trypanosoma cruzi*, Louisiana. *Emerging Infectious Diseases* 13: 605–607.
- Hanford EJ, Zhan FB, Lu Y, Giordano A (2007) Chagas disease in Texas: Recognizing the significance and implications of evidence in the literature. *Social Science and Medicine* 65: 60–79.
- Bern C, Montgomery SP (2009) An estimate of the burden of Chagas disease in the United States. *Clinical Infectious Diseases* 49: e52–e54.
- Kjos SA, Snowden KF, Olson JK (2009) Biogeography and *Trypanosoma cruzi* infection prevalence of Chagas disease vectors in Texas, USA. *Vector-Borne and Zoonotic Diseases* 9: 41–50.
- Reisenman CE, Lawrence G, Guerenstein PG, Gregory T, Dotson E, et al. (2010) Infection of kissing bugs with *Trypanosoma cruzi*, Tucson, Arizona, USA. *Emerging Infectious Diseases* 16: 400–405.
- Tarleton RL, Reithinger R, Urbina JA, Kitron U, Gürtler RE (2007) The challenges of Chagas disease: Grim outlook or glimmer of hope? *PLoS Medicine* 4: e332.
- Lambert RC, Kolivras KN, Resler LM, Brewster CC, Paulson SL (2008) The potential for emergence of Chagas disease in the United States. *Geospatial Health* 2: 227–239.
- Beard CB, Pye G, Steurer FJ, Rodriguez R, Campman R, et al. (2003) Chagas Disease in a domestic transmission cycle in southern Texas, USA. *Emerging Infectious Diseases* 9: 103–105.
- Kjos SA, Snowden KF, Craig TM, Lewis B, Ronald N, et al. (2008) Distribution and characterization of canine Chagas disease in Texas. *Veterinary Parasitology* 152: 249–256.
- Paredes EA, Valdéz-Miranda J, Noguera Torres B, Alejandro-Aguilar R, Cannet Romero R (2001) Vectorial importance of triatomine bugs (Hemiptera: Reduviidae) in Guaymas, Mexico. *Revista Latinoamericana de Microbiología* 43: 119–122.
- Schiffler RJ, Mansur GP, Navin TR, Limpakarnjanarat K (1984) Infectious Chagas disease (American trypanomiasis) in California. *Journal of the American Medical Association* 251: 2983–2984.
- Hervaldt BL, Grijalva MJ, Newsome AL, McGhee CR, Powell MR, et al. (2000) Use of polymerase chain reaction to diagnose the fifth reported US case of autochthonous transmission of *Trypanosoma cruzi*, in Tennessee, 1998. *Journal of Infectious Diseases* 181: 395–399.
- Cimo P, Luper W, Scouris M (1993) Transfusion-associated Chagas disease in Texas: Report of a case. *Texas Medicine* 89: 48–50.
- Gürtler EG, Segura EL, Cohen JE (2003) Congenital transmission of *Trypanosoma cruzi* in Argentina. *Emerging Infectious Diseases* 9: 29–32.
- Camargo EP (2009) Perspectives of vaccination in Chagas disease revisited. *Memórias do Instituto Oswaldo Cruz* 104(S1): 275–280.
- Ikenga JO, Richerson JV (1984) *Trypanosoma cruzi* (Chagas) (Protozoa: Kinetoplastida: Trypanosomatidae) in invertebrate and vertebrate hosts from Brewster County in Trans-Pecos Texas. *Journal of Economic Entomology* 77: 126–129.
- Beard CB, Young DG, Butler JF, Evans DA (1988) First isolation of *Trypanosoma cruzi* from a wild-caught *Triatoma sanguisuga* (LeConte) (Hemiptera: Triatominae) in Florida, U.S.A. *Journal of Parasitology* 74: 343–344.
- Phillips SJ, Dudík M (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography* 31: 161–175.
- Ibarra-Cerdana CN, Sánchez-Cordero V, Peterson AT, Ramsey JM (2009) Ecology of North American Triatominae. *Acta Tropica* 110: 178–186.
- Soberón J, Peterson AT (2005) Interpretation of models of ecological niches and species' distributional areas. *Biodiversity Informatics* 2: 1–10.
- Peterson AT (2008) Biogeography of disease: a framework for analysis. *Naturwissenschaften* 95: 483–491.
- Moffett A, Shackelford N, Sarkar S (2007) Malaria in Africa: Vector species niche models and relative risk maps. *PLoS ONE* 2: e824.
- González C, Wang O, Strutz SE, González-Salazar C, Sánchez-Cordero V, et al. (2010) Climate change and risk of leishmaniasis in north america: Predictions from ecological niche models of vector and reservoir species. *PLoS Negl Trop Dis* 4: e585.
- Wood SF (1952) Mammal blood parasite records from southwestern United States and Mexico. *Journal of Pediatrics* 38: 85–86.
- Burkholder J, Allison T, Kelly V (1980) *Trypanosoma cruzi* (Chagas) (Protozoa: Kinetoplastida) in invertebrate, reservoir, and human hosts of the lower Rio Grande valley of Texas. *Journal of Parasitology* 66: 305–311.
- McHugh CP, Grogg M, Kerr SF (1990) Isolation of *Leishmania mexicana* from *Neotoma micropus* collected in Texas. *Journal of Parasitology* 76: 741–742.
- Kerr SF, McHugh CP, Dronen NO Jr. (1995) Leishmaniasis in Texas: Prevalence and seasonal transmission of *Leishmania mexicana* in *Neotoma micropus*. *American Journal of Tropical Medicine Hygiene* 53: 73–77.
- Kerr SF, McHugh CP, Merkelz R (1999) A focus of *Leishmania mexicana* near Tucson, Arizona. *American Journal of Tropical Medicine Hygiene* 61: 378–379.
- Best N, Richardson S, Thompson A (2005) A comparison of Bayesian spatial models for disease mapping. *Statistical Methods in Medical Research* 14: 35–59.
- Moffett A, Sarkar S (2006) Incorporating multiple criteria into the design of conservation area networks: A minireview with recommendations. *Diversity and Distributions* 12: 125–137.
- Lent H, Wygodzinsky P (1979) Revision of the Triatominae (Hemiptera, Reduviidae), and their significance as vectors of Chagas' disease. *Bulletin of the American Museum of Natural History* 163: 123–520.
- Moffett A, Strutz S, Guda N, González C, Ferro MC, et al. (2009) A global public database of disease vector and reservoir distributions. *PLoS Neglected Tropical Diseases* 3: e378.
- United States Census Bureau (2009) Small area income and poverty estimates. URL: <http://www.census.gov/did/www/saiper/index.html>.
- Elith J, Graham CH, Anderson RP, Dudík M, Ferrier S, et al. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29: 129–151.
- Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 190: 231–259.
- Pawar S, Koo MS, Kelley C, Ahmed FM, Choudhury S, et al. (2007) Conservation assessment and prioritization of areas in Northeast India: priorities for amphibians and reptiles. *Biol Conserv* 136: 346–361.
- Hijmans R, Cameron S, Parra J, Jones P, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25: 1965–1978.

46. Kaplan S, Garrick BJ (1981) On the quantitative definition of risk. *Risk Analysis* 1: 11–27.
47. Clayton DG, Kaldor J (1987) Empirical Bayes estimates for age-standardised relative risks for use in disease mapping. *Biometrics* 43: 671–691.
48. Besag J, York J, Mollié A (1991) Bayesian image restoration with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics* 43: 1–59.
49. Lawson AB, Browne WJ, Vidal Rodeiro CL (2003) *Disease Mapping with WinBUGS and MLwiN*. Chichester, UK: Wiley.
50. Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A (2002) Bayesian deviance, the effective number of parameters and the comparison of arbitrarily complex models. *Journal of the Royal Statistical Society B* 64: 583–640.
51. Lunn DJ, Thomas A, Best N, Spiegelhalter D (2000) WinBUGS—a Bayesian modelling framework: Concepts, structure, and extensibility. *Statistics and Computing* 10: 325–337.
52. Sarkar S, Garson J (2004) Multiple criterion synchronization for conservation area network design: The use of non-dominated alternative sets. *Conservation and Society* 2: 433–448.
53. Sarkar S (2005) *Biodiversity and Environmental Philosophy: An Introduction to the Issues*. Cambridge, UK: Cambridge University Press.
54. Margules CR, Sarkar S (2007) *Systematic Conservation Planning*. Cambridge, UK: Cambridge University Press.
55. Moffett A, Garson J, Sarkar S (2005) MultCSync: A software package for incorporating multiple criteria in conservation planning. *Environmental Modelling and Software* 20: 1315–1322.
56. Wisz MS, Hijmans RJ, Peterson AT, Graham CH, Guisan A, et al. (2008) Effects of sample size on the performance of species distribution models. *Diversity and Distributions* 14: 763–773.
57. Stockwell D, Peters D (1999) The GARP modelling system: problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science* 13: 143–158.
58. Grundemann AW (1947) Studies on the biology of *Triatoma sanguisuga* (LeConte) in Kansas, (Reduviidae:Hemiptera). *Journal of the Kansas Entomological Society* 20: 77–84.
59. zu Dohna H, Cecere MC, Gärtler RE, Kitron U, Cohen JE (2009) Spatial re-establishment dynamics of local populations of vectors of Chagas disease. *PLoS Negl Trop Dis* 3: e490.
60. Dumonteil E, Gourbière S (2004) Predicting *Triatoma dimidiata* abundance and infection rate: A risk map for natural transmission of Chagas disease in the Yucatán peninsula of Mexico. *American Journal of Tropical Medicine and Hygiene* 70: 514–519.