Identifying high-risk areas for dog-mediated rabies using Bayesian spatial regression

Kaushi S.T. Kanankege\textsuperscript{a,*}, Kaylee Myhre Errecaborde\textsuperscript{a,1}, Anuwat Wiratsudakul\textsuperscript{b,1}, Phrutsamon Wongnak\textsuperscript{c,d}, Chakchalat Yoopatthanawong\textsuperscript{b}, Weerapong Thanapongtharm\textsuperscript{e}, Julio Alvarez\textsuperscript{f,g}, Andres Perez\textsuperscript{a,1}

\textsuperscript{a} College of Veterinary Medicine, University of Minnesota, USA
\textsuperscript{b} Faculty of Veterinary Science, Mahidol University, Nakhon Pathom, Thailand
\textsuperscript{c} Université Clermont Auvergne, INRAE, VetAgro Sup, UMR EPIA, 63122 Saint-Génis-Champanelle, France
\textsuperscript{d} Université de Lyon, INRAE, VetAgro Sup, UMR EPIA, 69280 Marcy l’Etoile, France
\textsuperscript{e} Department of Livestock Development, Ministry of Agriculture and Cooperatives, Thailand
\textsuperscript{f} Centro de Vigilancia Sanitaria Veterinaria (VISAVET), Universidad Complutense, Madrid, Spain
\textsuperscript{g} Departamento de Sanidad Animal, Facultad de Veterinaria, Universidad Complutense, Madrid, Spain

\textbf{ARTICLE INFO}

\textbf{Keywords:}
Stray dogs
Conditional autoregression
Zero-inflated
Spatial epidemiology
Disease mapping
Risk regionalization
One Health

\textbf{ABSTRACT}

Despite ongoing control efforts, rabies remains an endemic zoonotic disease in many countries. Determining high-risk areas and the space-time patterns of rabies spread, as it relates to epidemiologically important factors, can support policymakers and program managers alike to develop evidence-based targeted surveillance and control programs. In this One Health approach which selected Thailand as the example site, the location-based risk of contracting dog-mediated rabies by both human and animal populations was quantified using a Bayesian spatial regression model. Specifically, a conditional autoregressive (CAR) Bayesian zero-inflated Poisson (ZIP) regression was fitted to the reported human and animal rabies case counts of each district, from the 2012–2017 period. The human population was used as an offset. The epidemiologically important factors hypothesized as risk modifiers and therefore tested as predictors included: number of dog bites/attacks, the population of dogs and cats, number of Buddhist temples, garbage dumps, animal vaccination, post-exposure prophylaxis, poverty, and shared administrative borders. Disparate sources of data were used to improve the estimated associations and predictions. Model performance was assessed using cross-validation. Results suggested that accounting for the association between human and animal rabies with number of dog bites/attacks, number of owned and un-owned dogs; shared country borders, number of Buddhist temples, poverty levels, and accounting for spatial dependence between districts, may help to predict the risk districts for dog-mediated rabies in Thailand. The fitted values of the spatial regression were mapped to illustrate the risk of dog-mediated rabies. The cross-validation indicated an adequate performance of the spatial regression model (AUC = 0.81), suggesting that had this spatial regression approach been used to identify districts at risk in 2015, the cases reported in 2016/17 would have been predicted with model sensitivity and specificity of 0.71 and 0.80, respectively. While active surveillance is ideal, this approach of using multiple data sources to improve risk estimation may inform current rabies surveillance and control efforts including determining rabies-free zones, and the roll-out of human post-exposure prophylaxis and anti-rabies vaccines for animals in determining high-risk areas.

\section{1. Introduction}

Rabies remains an endemic and neglected zoonotic disease causing over 59,000 human deaths annually across the world with the majority being developing countries of Africa (36.4%) and Asia (59.6%) [1,2]. The density of reservoir populations, mainly domestic and stray dogs, plays a major role in the persistence of human rabies [3]. Controlling zoonoses such as rabies requires a One Health approach that benefits
from combined veterinary and medical capacity and collaboration. In a One Health solution to rabies, knowledge of risk of animal rabies cases in an area would ideally bring human and animal health authorities to attend disease control and prevention in a collaborative manner. Conversely, detected human cases would trigger the need for animal surveillance in the area. Prompt administration of post-exposure prophylaxis (PEP) upon dog bites is important to mitigate human rabies [4,5]. Effective control of dog-mediated rabies requires sustained mass vaccination of dog population with over 70% coverage, and population control of the stray animals [6–10]. Countries aiming to declare the status of ‘free from dog-mediated rabies’ are required to establish disease surveillance and early warning systems in addition to the preventive measures and dog population control [11].

In pursuit of the goal to eliminate dog-mediated rabies, risk-based regionalization, and application of consistent and targeted approaches are commonly proposed when conducting disease surveillance and control programs in resource-limited environments instead of spreading the resources thinly across an entire geographic region [1,12,13]. In an attempt to support such programmatic efforts, we used a risk-estimation approach that makes use of available data. In this One Health approach, the risk is defined as ‘the likelihood and magnitude of both human and animal populations in each district have contracting dog-mediated rabies.’ The risk was quantified using the association between reported human and animal cases with epidemiologically important variables representing factors that modify the risk of dog-mediated rabies, by fitting a spatial regression model. The geographically varying risk modifying factors include the host population densities primarily human and dog populations [14,15], socio-economic and cultural habits [16], location [17,18], and food source availability influencing the survival of the stray animals [19]. In addition to accounting for these factors, when developing location-based risk maps, accounting for the spatial dependence is critical (i.e. accounting for the unexplained epidemiological similarities or differences of the neighboring districts) [20]. Spatial regression models enable modeling the relationship between the disease outcomes and underlying factors, while accounting for this spatial dependence and improving local estimates [20].

While reported case numbers and host densities may represent convenient means for vaccine and PEP distribution geographically, understanding the factors modifying the risk of disease transmission, population mobility, and taking them into account when strategically distributing prophylactic interventions has been proposed in many studies on disease control strategies [21–23]. Rabies is considered underreported and neglected [1], and we propose that estimation of the disease risk in relation to the underlying factors might improve the potency of surveillance, preventive, and control measures. This study exemplifies the best use of available data to optimize risk regionalization using data available from Thailand. Despite having a government-led control effort to eliminate dog-mediated human rabies deaths by 2030 [2], Thailand continues to struggle with effective program implementation [24–29]. Health authorities of Thailand have been using various criteria to determine the rabies status of the administrative areas, mainly based on the reported numbers of human and animal rabies cases [30]. However, most of the risk estimations are available at large administrative divisions. Moreover, as per our knowledge of the site, the risk regionalization does not account for the distribution of epidemiologically important factors.

Therefore, the specific objective of the study was to identify high-risk areas for dog-mediated rabies using Bayesian spatial regression model and inform policymakers about hotspots of human and dog rabies combined. Furthermore, the study approach highlights the insights gained from the use of a range of data sources regarding the risk of dog-mediated rabies in a country beyond the reported case numbers. Since the risk for humans and animals is combined in this One Health approach, the risk estimates are applicable to both populations. The results from this study, when translated within the context limitations, may complement efforts in rabies control by providing valuable insights on factors to consider during resource allocation and determining high-risk areas.

2. Data and methods

To achieve the objective of improved risk prediction of rabies using multiple data sources, a spatial regression model was fitted to measure the association between reported human and animal rabies cases with epidemiologically important risk modifying factors. The fitted values were used to predict high-risk areas. The model performance was evaluated using cross-validation; i.e. determine how many high-risk districts would have been predicted correctly, had this spatial regression model been used in 2015 (Fig. 1). Additionally, the animal rabies cases from 2017/2018 outbreak were compared against the model-based risk determined.

3. Study area and administrative divisions

Thailand is a South Asian country positioned at latitude 15 $N$ and longitude 100 $E$, and shares borders with four countries: Myanmar, Lao, Cambodia, and Malaysia. The UN estimated human population in Thailand as of 2015 was 68,071,557 and 14 million (20%) of the population is low on the north-western border of the country. Administrative divisions of different granularities in Thailand are comprised of 77 provinces (i.e. changwat), each of which is further divided into 928 districts (i.e. amphoe), and districts are further divided into 7425 subdistricts (i.e. tambons). The land area of the administrative divisions in the central region are relatively smaller and highly populated compared to other areas of the country.

The shapefiles used to map and conduct the analysis which contains administrative divisions of Thailand were downloaded from the database called “The Humanitarian Data Exchange” version v1.34.2 (https://data.humdata.org); the metadata of the shapefile listed the Royal Thai Survey Department (https://www.rtsd.mi.th/main) as the original source of data. Given the variation of granularity (Table 1), a decision was made to aggregate all variables at the administrative districts for the analysis (Table 1), hence the unit of analysis in both spatial and non-spatial models was the districts.

4. Data collection

Government and open-source data available at varying granularities of administrative levels were collected (Table 1). To perform the regression analysis at district level, data available at district or subdistrict level were aggregated by districts ($n = 6$ factors; rabies cases, dog attacks/bites, human population, garbage dumps, Buddhist temples, and neighboring countries). The variables available at the provincial level ($n = 4$) include; total dogs, average poverty, animal rabies vaccines, and yearly average of PEP. These variables were then disaggregated across districts based on the ratio of human population in the district to the human population in the relevant province, similar to the population-based disaggregation by Perez et al., [31]. Epidemiologically important factors hypothesized to modify the risk of dog-mediated rabies in Thailand were recognized based on published literature. Risk modifiers related to the human population, animal population, geographical, socio-cultural, and prevention programming were considered. Data were not readily accessible for dog population control and the number of animal vaccines administered, hence were not included in the analysis. Districts sharing borders with neighboring countries were included as a variable, which was assumed as a proxy for the differences in rabies control efforts implemented by those countries.
5. Data on human and animal rabies cases

Rabies is a notifiable disease in Thailand. Human rabies cases are documented and monitored by the Division of Epidemiology (www.boe.moph.go.th) and the Department of Disease Control of the Ministry of Public Health of Thailand [32] (Table 1). A total of 35 human cases (during 2013–2016) were reported. While it is possible that they acquired the disease from elsewhere, the human cases were assigned to the districts where those were reported. Human cases are reported based on hospitals and healthcare systems, whereas, animal cases are tested and confirmed based on suspected carcass submissions and supplemented by a limited active surveillance effort.

Animal rabies case data were captured from both active and passive surveillance and were obtained from the Department of Livestock Development (DLD). The passive surveillance included animal samples (carcasses or heads) submitted with a case history of suspected rabies to the laboratories by animal owners or veterinarians. There are nine regional DLD laboratories and one laboratory belonging to the Queen Saovabha Memorial Institute providing rabies diagnostic service across the country. As an endemic country for rabies, trained veterinarians responsible for disease detection, reporting, and control are positioned at all provincial DLD offices as well as the district level, throughout the country. The veterinarians communicate and consult with experts from DLD headquarter and regional laboratories in situations of suspected outbreaks. In recent years, a mobile application that enables farmers and animal owners to report disease events is being promoted. The active surveillance conducted by veterinary services entails sampling animal died due to inconclusive symptoms and through a minimum active sampling of animals by collecting at least one sample from each sub-district every year [29]. Further details on current animal rabies status in Thailand is found elsewhere [29]. Animal rabies data included 1877 cases from the 2012–2017 period and; 1706 (90.8%), 117 (6.2%), and 54 (2.8%) of those cases were associated with dogs, cattle, and cats, respectively. All reported animal rabies cases were included in the analysis regardless of the type of animal affected, considering the limited reporting of animal rabies. Therefore, a total of 1912 human and animal rabies cases were used in the analysis (Fig. 2). Among the 928 districts, 621 (67%) reported no cases of human or animal rabies. Human rabies cases, animal rabies cases, and dog bites/attack counts were compared across districts and months regardless of the year of reporting (Fig. 3).

6. Data on epidemiologically important factors

1. Data on dog bites/attacks were available from the Division of Epidemiology (DoE) (www.boe.moph.go.th), and the Department of
Table 1
Epidemiologically important factors/variables hypothesized to modify the risk for dog-mediated rabies in Thailand. The data time period, sources, and the administrative level at which data were available are listed.

| Variable                             | Administrative level* of data availability | Year/s of data availability | Source of data |
|--------------------------------------|-------------------------------------------|-----------------------------|----------------|
| Number of human rabies cases         | Sub-district level                         | 2013–2016                   | Division of Epidemiology, Department of Disease Control, 2017 www.boe.moph.go.th |
| Number of animal rabies cases        | Sub-district level                         | 2012–2017                   | Department of Livestock Development, 2017 http://www.dld.go.th/index.php/th/ |
| Human population                     | Sub-district level                         | 2015                        | Worldpop, 2018 www.worldpop.org |
| Number of dog attacks/bites          | Sub-district level                         | 2013–2017                   | Division of Epidemiology, Department of Disease Control, 2017 www.boe.moph.go.th; DDC, 2018 |
| Number of garbage dumps             | District level                             | 2016                        | Pollution Control Department, 2016 Dhammathai, 2017 http://www.dhannathai.org/index.php |
| Number of Buddhist temples           | District level                             | 2017                        | Department of Livestock Development, 2016 http://www.dld.go.th/index.php/th/ |
| Number of animal vaccines distributed| Provincial level                           | 2016                        | Department of Livestock Development, 2016 http://www.dld.go.th/index.php/th/ |
| Number of human post exposure prophylaxis| Provincial level                          | 2014–2018                   | Division of Epidemiology, Department of Disease Control, 2018 https://apps.boe.moph.go.th/h/index.php/th/ |
| Human poverty index                  | Provincial level                           | 2006–2015                   | National Statistical Office, 2018 http://web.nso.go.th/ |
| Stray animal densities: dogs and cats| Provincial level                           | 2016                        | Department of Livestock Development, 2016 http://www.dld.go.th/index.php/th/ |

* Administrative levels: Thailand is divided into 77 provinces, each of which is further subdivided into districts 928 districts, and these are further divided into 7425 sub-districts.

Disease Control of the Ministry of Public Health of Thailand [32] (Table 1). A total of 2,764,254 dog bites/attacks were reported during 2013–2017 (Table 1). The data indicated a yearly average of 453,086 dog bites/attacks and nearly 7 dog bites/attacks per 1000 people (assuming the total human population of Thailand is susceptible) [33]. The dog bites/attacks were defined according to the WHO guidelines of the possible human exposure to rabies [2].

2. Human population data were obtained from the Worldpop database (www.worldpop.org). The UN-adjusted population count grids consist of estimates of the number of persons per 30 arc-second (~1 km) grid cell, adjusted to match the 2015 revision of the UN World Population Prospects national population estimates, and the population estimates for 2015 were used in the analysis [34].

3. Garbage dump data were collected from the Pollution Control Department of Thailand [35]. The data contained the validity of the garbage dump as of 2016, the input amount per day in tons, land area, and the different systems of garbage disposal which included: composting, controlled dump, controlled dump/incinerator, engineered landfill, incinerator, integrated solid waste disposal, mechanical biological treatment (MBT), sanitary landfill, open dump, and open burning. Considering the potential access by stray dogs and cats [36], the analysis included all garbage dumps per district regardless of the type.

4. Data on number of Buddhist temples: As a practice of compassion toward the homeless animals, many places of worship, especially Buddhist temples tend to shelter stray dogs and cats [16,37,38]. This sociocultural influence makes the presence of Buddhist temples an associated factor to find a multitude of stray or community dogs and cats in the premises. Therefore, the number of Buddhist temples in the district, as of 2017, was added as a variable (Table 1); [39].

5. Shared country borders: Given that rabies could be considered as a transboundary disease [40–43], the neighboring countries were identified for districts sharing a border with another country, using the ‘Polygon Neighbor’ feature of ArcMap 10.6.1 version [44]. For two districts that shared borders with more than one country, namely, 1) Ubon Ratchathani district of Ubon Ratchathani province (Cambodia and Laos) and 2) Mae Sai district of Chiang Rai province (Laos and Myanmar) the country with the longest shared border with the district (Cambodia and Laos, respectively) was assigned to the district. However, data related to rabies status or control measures were not collected from the neighboring countries.

6. Dog and cat populations and vaccines: To estimate the dog and cat populations of Thailand, a survey was conducted in 2016 by the DLD [45]. The collected data included the number of households surveyed, the number of owned and un-owned dogs and cats and their sex. The estimates of animal rabies vaccine distribution by province for 2016 was available from DLD (Table 1). As per the survey, the total dog population in districts ranged between 3367 and 98,689 with an average of 8052.

7. The monthly PEP data during January 2014 through March 2018 were summarized into yearly averages per province and disaggregated based on the ratio of human population across the districts for the analysis.

8. Household income and poverty: Another key factor influencing the decision to vaccinate pet animals against rabies and maintain responsible pet ownership is the household income [10], therefore, poverty index in the units poverty index data available from 2006 through 2015 [45] were collected and summarized by averaging across the years as well as calculating the median of the poverty index for each province. The unit of poverty index at the provincial level was “the number of poor people in 1,000 people,” however, due to the disaggregation of the values at the district level this value should be considered as a mere representation of the estimated level of poverty.

Upon data collection, a descriptive analysis was conducted to eliminate variables from the analysis; specifically, to identify variables that are poorly represented and/or highly correlated using Spearman’s ρ correlation. In parallel to Cohen’s standard for the effect size of Pearson correlation coefficient, the variables with Spearman’s ρ <0.5 were chosen for the regression analysis. However, exceptions to this exclusion criteria were considered when one variable could be used as an auxiliary variable that may help estimate the association with another under-sampled variable. It is important to notice that the ratio based disaggregation of variables involves the assumption that when the human population is high, the poverty index increases [47], more PEP and animal rabies vaccines are needed, and more stray dogs and cats are seen overall [48]. These assumptions were supported in part by studies and data available elsewhere [47–49]. Once all the variables were aggregated at the district level, all numerical variables were tested for correlation using Spearman’s ρ (S1 Table). To take account of possible non-linearity of effects, all continuous scale risk factor variables were...
Fig. 2. Reported cases of human and animal rabies in Thailand. A) Human rabies between 2013 and 2017, B) Animal rabies between 2012 and 2017, and C) summation of both human and animal rabies between 2012 and 2017. Reported cases of animal and human rabies data were collected from Thai Department of Livestock Development (DLD) (http://en.dld.go.th/index.php/en/home-top), Division of Epidemiology (DoE) (http://www.boe.moph.go.th/boed/surdata/), and Department of Disease Control of the Ministry of Public Health. Maps were generated as part of the current study.

| A) | Year | Animal rabies cases | Human rabies cases | Dog bites |
|----|------|---------------------|--------------------|----------|
| 2012 | No data | 5 | No data |
| 2013 | 105 | 5 | 139225 |
| 2014 | 248 | 6 | 644385 |
| 2015 | 327 | 5 | 686750 |
| 2016 | 607 | 14 | 774875 |
| 2017 | 590 | No data | 519019 |

Fig. 3. Comparison of the number of dog bites/attacks (n = 2,265,428), human rabies (n = 35), and animal rabies (n = 1877) cases reported over the period of 2012-2017. Panels: A) available data by year, B) by month, and C) by 928 districts of Thailand (districts are arranged in the ascending order of the postal codes i.e. amphoe).
categorized into two or three categories based on Jenks classification criteria, i.e. natural breaks [50].

7. Regression analysis

The factors were evaluated both individually and collectively to understand their contribution when determining the risk by fitting regression models. Frequentist univariable and multivariable Zero-inflated Poisson (ZIP) regression models were fitted to quantify the association between the sum of human and animal rabies cases (dependent variable) and relevant epidemiological determinants (i.e. the selected independent variables) [21,51,52]. The majority of the 928 districts (n = 621; 67%) did not report any animal or human rabies cases during 2012–2017. This could be due to reporting biases as well as to the true absence of reports of the disease. This skewed distribution of cases led to the decision of using Zero Inflated regression models. The ZIP regression generates two separate models and then combines them. A logit model for the “certain zeros” predicting whether they are districts with truly no cases. Then, a Poisson model is fitted to predict the counts of districts where “uncertain zeros”, i.e. the absence of rabies cases either due to no

### Table 2

Estimated regression model coefficients relevant to the factors hypothesized to modify the risk of animal or human Rabies in Thailand.

| Variable                          | n districts | Univariable models | Finalized Multivariable model | Finalized Multivariable Bayesian CAR model |
|-----------------------------------|-------------|--------------------|-------------------------------|--------------------------------------------|
|                                   |             | Coefficient (CI)    | Exponentiated values         | Median Exponentiated values                |
| Intercept                         |             | 0.0006 (0.0005-0.0007) | 0.00003                      | 0.00001-0.00006                            |
| Dog attacks/bites (total count = 2,265,428) |             |                    |                               |                                            |
| ≤3611 attacks                     | 705         | Ref.               | 0.86 (0.78-0.94)*             | 1.19                                       |
| >3611 attacks                     | 228         | Ref.               | 0.88 (0.79-0.97)*             | 0.78-1.76                                  |
| **Total dogs: owned and un-owned** (total count = 7,472,037) |             |                    |                               |                                            |
| ≤14,582 dogs                      | 816         | Ref.               | 0.55 (0.49-0.62)*             | 0.78                                       |
| >14,582 dogs                      | 112         | Ref.               | 0.61 (0.54-0.68)*             | 0.43-1.37                                  |
| **Un-owned female dogs** (total count = 411,368) |             |                    |                               |                                            |
| ≤1109 dogs                        | 837         | Ref.               | 0.61 (0.55-0.67)*             |                                            |
| >1109 dogs                        | 91          |                    |                               |                                            |
| **Total cats** (total count = 3,035,653) |             |                    |                               |                                            |
| ≤3173 cats                        | 630         | Ref.               | 0.50 (0.46-0.55)*             |                                            |
| >3173 cats                        | 298         |                    |                               |                                            |
| **Buddhist temples** (total count = 32,786) |             |                    |                               |                                            |
| ≤50 temples                       | 706         | Ref.               | 1.14 (1.03-1.25)*             | 1.01                                       |
| >50 temples                       | 222         | Ref.               | 1.10 (0.98-1.24)#             | 0.63-1.62                                  |
| **Average poverty** |             |                    |                               |                                            |
| ≤13                               | 651         | Ref.               | 1.15 (1.04-1.27)*             | 1.14                                       |
| >13                               | 277         | Ref.               | 1.17 (1.04-1.31)*             | 0.72-1.81                                  |
| **Human population of 2015 (68,071,557)** |             |                    |                               |                                            |
| ≤50,000                           | 468         | Ref.               | 1.95 (1.73-2.19)*             |                                            |
| 50,000-200,000                    | 407         |                    | 3.55 (3.09-4.07)*             |                                            |
| >200,000                          | 53          |                    |                               |                                            |
| **Sharing country borders** |             |                    |                               |                                            |
| No international borders          | 799         | Ref.               | 0.85 (0.64-1.15)              | 4.51                                       |
| Myanmar (Burma)                   | 45          |                    | 0.65 (0.48-0.88)*             | 1.05-21.50                                 |
| Malaysia                          | 15          |                    | 1.65 (1.17-2.31)*             | 1.50                                       |
| Cambodia                          | 20          |                    | 1.45 (1.03-2.04)*             | 0.18-10.82                                 |
| Laos                              | 49          |                    | 1.78 (1.14-2.23)*             | 1.50                                       |
| **Garbage dumps** (total count = 2272) |             |                    |                               |                                            |
| ≤6 dumps                          | 854         | Ref.               | 2.42 (2.07-2.84)*             | 3.69                                       |
| >6 dumps                          | 74          |                    | 1.96 (1.65-2.31)*             | 2.35-10.82                                 |
| **Yearly average of PEP** (total count = 51, 296) |             |                    |                               |                                            |
| ≤81                               | 760         | Ref.               | 0.75 (0.66-0.84)*             |                                            |
| >81                               | 168         |                    |                               |                                            |
| **Animal vaccines of 2016** (total count = 907, 400) |             |                    |                               |                                            |
| ≤2153 doses                       | 853         | Ref.               | 0.36 (0.31-0.43)*             |                                            |
| >2153 doses                       | 75          |                    |                               |                                            |
| **Spatial dependence** |             |                    |                               |                                            |
| Omega - intercept                 | N/A         | N/A                | –201.56                       | –645.3 to –12.17                           |
| Tau                               | N/A         | N/A                | 8.48                          | 7.34-11.17                                 |
| Spatial autocorrelation parameter (ρ) | 0.98       | 0.94-0.99          |                               |                                            |
cases or due to underreporting and other factors. The two models were interpreted collectively. The models used data from a 6-year period (2012–2017), which included (1877/1912) 98% of animal rabies cases and (35/1912) 2% human cases; therefore, the predicted case numbers to be interpreted in the same ratio.

While the human population was significantly associated with rabies cases (Table 2), the variable was highly correlated with a number of other variables (S1 Table). Therefore, the district-level human population was used as an offset in these ZIP regression models, although the risk assessed is applicable to both human and animal host populations. Based on the results of the univariable models, variables that were highly correlated with another variable (>0.5 Spearman’s ρ coefficient) and not associated with the outcome (p-value >0.2) were excluded in the multivariable analyses. The variables dog bites and total dog counts were invariably included in the final models given these were considered in the current rabies control programs when disseminating resources [54]. The relative importance and the goodness-of-fit of models for each variable were assessed by comparing the Akaike information criterion (AIC) values for the alternative models including different sets of covariates (Supplementary Table 2) [55,56]. The “pscl” package [57,58] was used to fit the frequentist ZIP models in R statistical software [59]. A simple inflation model in which all zero counts were assumed to have the same probability of belonging to the zero component was specified.

The residuals of the finalized multivariable ZIP model at the district level were subjected to Moran’s I global spatial autocorrelation test to test for spatial dependence between model results at the district level using the ‘spatialreg’ R-statistical package [60,61]. The existence of spatial autocorrelation in the regression residuals was considered evidence of the need of fitting a spatial regression model to account for the spatial dependency of adjacent districts. In these spatial regression models the spatial dependence, i.e. autocorrelation between districts is modeled via random effects [62].

The variables included in the final frequentist multivariable ZIP model were used in a Bayesian Conditional Autoregressive (CAR) ZIP spatial regression model [52,53,62–64]. The human population was used as the offset. The functions available through “spatialreg” and “CARBayes” version 5.2 [65] R statistical packages were used to perform the analyses. Queen contiguity, i.e. districts sharing at least one common edge or vertex, was used to define whether two districts were neighbors. The Bayesian zero-inflated multivariable CAR models were run with 500,000 samples, 10,000 burnings, and 200 thinning. The spatial dependence or the correlation between districts was captured by the conditional autoregressive (CAR) term proposed by Leroux et al. [62], and between variable correlation is captured by a between variable covariance matrix with no fixed structure [65]. The Geweke convergence diagnostics with the threshold set at |Z| >2 were used to confirm the model convergence [66–68]. The fitted values of the CAR model for 2017 were mapped using ArcMap version 10.6.1 [44], and the risk estimates were classified into five based on natural breaks, i.e. Jenks classification criteria [50]. It is important to note that the regression predicted risks are mainly for animal cases given the majority of cases are animal rabies cases. While the transmission pathways are peculiar for animal-to-animal transmission and the zoonotic risk of rabies, here we assumed that having animal cases predicted would indicate a risk for the human population alike.

Temporal cross-validation of the model fits was done by using data from the year 2017 to predict “non-rabies” months in 2016 and 2017. Both multivariable ZIP and Bayesian ZIP CAR models were fitted to the human and animal rabies cases reported between 2012 and 2015 (training data), and fitted value were tested based risk assigned to districts with cases reported between 2016 and 2017 (testing data). Given the data availability for both human and animal rabies varied across the years (Fig. 3: Panel A), 2015 was chosen as the cut off year; which resulted in 701 (37%) cases in the training dataset and 1211 (63%) cases in the testing data. This approach of cross-validation can be considered as if the models were fitted in 2015 and the resulting risk maps were used to identify high-risk districts, how well did the 2015 models would predict the cases reported between 2016 and 2017. The fitted values of the regression models indicated the predicted counts of human and animal rabies cases per district. The fitted values of both multivariable ZIP and Bayesian ZIP CAR models were classified into five threshold classes based on natural breaks, i.e. Jenks classification criteria [50]. The model sensitivity and specificity were calculated for each of the five risk thresholds. Model sensitivity indicates the number of districts that were correctly predicted as ‘at risk’ by the model prediction and the specificity indicates the number of districts that were predicted as negligible risk and did not report any human and animal rabies case during the 2016/17 period. The model performance was estimated using the area under the curve (AUC) of the Receiver Operating Characteristic (ROC). AUC values lower than 0.7 are considered relatively inaccurate because the proportion of false and true positive results is not substantially different, whereas AUC values >0.7 are generally considered appropriate [69]. For further validation, the risk maps developed using the conditional autoregressive (CAR) Bayesian zero-inflated Poisson (ZIP) regression for data from 2012 to 2017 was compared with the animal rabies outbreak data from 2017/2018 period. The percentage of cases observed from each of the five risk categories was summarized.

8. Results

8.1. Regression results

Eleven key variables were included in the univariable analysis (Fig. 4 and Table 2). As per the frequentist univariable ZIP models, dog attacks/bites, total dogs, un-owned female dogs, total cats, Buddhist temples, average poverty, shared country borders, animal vaccination, and yearly average of PEP were resulted as statistically significant to be associated with rabies case counts. The count model results are presented in Table 2. The number of garbage dumps was not statistically significant (p > 0.2; Table 2), hence was not included in the multivariable model. The estimates of animal rabies vaccine distribution by province for 2016 were highly correlated with the survey-based estimates of the total dog population (Spearman’s ρ coefficient = 0.91; Supplementary Table 1) indicating animal vaccines were distributed at provincial level based on the survey. Therefore, animal vaccine data were excluded from the multiple regression. The average PEP was not statistically significant and, therefore removed from the finalized multivariable model. Among the statistically significant variables of the univariable analysis, total dogs were highly correlated with total cats (0.74) and unowned female dogs (0.63) (S1 Table), hence a decision was made to only include the total dogs as an independent variable in the multivariable model. The final multivariable ZIP model with the lowest AIC value included five independent variables (S2 Table and Table 2: Column 4). The Morans’I test performed on the residuals of the multivariable ZIP model were statistically significant indicating spatial autocorrelation between neighboring districts that were not described by the variables included in the finalized ZIP model. Therefore a spatial regression model, i.e. the multivariable Bayesian ZIP CAR model, was fitted.

Among the 928 districts, six were recognized as islands that are not sharing borders with at least one other district and therefore were excluded from the Bayesian ZIP CAR model fit (i.e. only 922 districts were subjected to the spatial regression model). The Bayesian CAR model included the five variables identified by the multivariable ZIP model and accounted for the spatial autocorrelation between districts captured by the conditional CAR term proposed by Leroux et al. [62]. The Geweke convergence diagnostics of the Bayesian ZIP CAR models indicated sufficient convergence of the model (|Z| >2). The posterior median and the 95% credible interval for each of the variable and spatial dependence parameter are presented in Table 2: Column 5. The odds of having an animal or human rabies cases was increased by 19% in those districts with >3611 reported dog bites/attacks. Simply stated, if the
districts with <3611 reported dog bites/attacks have 100 human or animal rabies cases per 10,000 people (the reference category), then the districts with >3611 reported dog bites/attacks would have 119 cases. The odds of reported dog and human rabies cases from districts with >14, 582 total owned and un-owned dogs was reduced by 22% (1–0.78). However, the credible interval was between 0.43 and 1.37 indicating the range of possible increase in the cases up to 37% in districts with >14, 582 dogs. The risk of human or animal rabies cases increased by 1% median (with a maximum of 62% increase) in districts with >50 Buddhist temples, compared to the districts with ≤50 temples. The districts with average poverty categorized as >13 were associated with 14% increase in cases compared to the reference category.

The posterior medians of the Bayesian ZIP CAR model were similar to the multivariable ZIP regression coefficients except for the values for shared county borders (Table 2: columns 4 and 5). The posterior median values for the shared country borders indicated that even after accounting for spatial dependence between districts, the shared borders with adjacent countries indicate an increased risk of dog-mediated rabies. Compared to the districts that are not sharing international borders, a higher risk of rabies was observed in districts sharing borders with Laos (499%), Myanmar (351%), Cambodia (50%), and Malaysia (50%). Risk maps were generated to represent the estimated risk of dog-mediated rabies for human and animal populations at the district level. The *Morans’ I* test performed on the residuals of the multivariable Bayesian ZIP CAR model were not statistically significant (p > 0.05) indicating that the spatial regression model accounted for the spatial dependence (i.e. unexplained epidemiological similarities or differences of the neighboring districts). The fitted values indicating the predicted counts of human or animal rabies by the district using the CAR model for 2017 are illustrated in Fig. 5. The five thresholds of predicted counts included negligible risk: < 1 predicted case (n = 426 districts), low risk: between 1 and 3 cases (n = 350), intermediate-risk: 4–7 cases (n = 65), high-risk: 8–17 cases (n = 58), and very high-risk: >17 cases (n = 23) for 2017.

The temporal cross validation of multivariable ZIP model fitted for 2015 and validated using 2016/17 resulted in 0.68 AUC with model sensitivity and specificity of 0.75 and 0.54, respectively, at the best risk threshold (i.e. the non-spatial ZIP model). Whereas, the cross validation for the spatial regression model (i.e. the Bayesian ZIP CAR model) resulted in 0.81 AUC with model sensitivity and specificity of 0.71 and 0.80, respectively. The comparison indicated the importance of accounting for spatial dependence and the potential improvement of predictions by the Bayesian ZIP CAR model compared to the non-spatial model. The risk threshold chosen was the risk class-2 (i.e. low risk with at least one predicted case for the district). The 1211 cases in the testing data (i.e. cases from 2016 to 2017) were reported from 167 districts and the model sensitivity suggested 119/167 (71%) of these districts were predicted to be at risk. The specificity of 0.80 indicated that 606/705 districts that were predicted as ‘negligible risk’ did not report any case during 2016–2017. The spatial regression model predicted 268 districts to be at risk as of 2015, and only 119 (44%; i.e. the positive predictive value) reported with at least one human or animal rabies case between 2016 and 2017. The false positives at this risk threshold were 149/268 (56%).

In further validation using animal rabies cases reported during the 2017/2018 outbreak [29], 218/224 (97%) districts reported at least one rabies case were correctly predicted by the risk maps (Fig. 5). The breakdown of districts by the risk category comprised of negligible risk (n = 6), low risk (n = 90), intermediate-risk (n = 53), high-risk (n = 54), and very high-risk (n = 21).
9. Discussion

A Bayesian spatial regression model was used to develop risk maps for dog-mediated rabies. The key risk modifiers associated with the case numbers were the number of dog bites/attacks, total number of owned and un-owned dogs, shared country borders, number of Buddhist temples, and poverty level, and the model accounted for spatial dependence between the districts; i.e., accounting for the unexplained epidemiological similarities and potential introduction from neighboring districts. While associations between the case numbers and the risk modifying factors do not indicate causation, the potential of this approach to predict risk areas, we propose that these key factors be considered when allocating resources for rabies prevention and control activities by district. Spatial models indicated fair cross-validation performance, and the compatibility of the risk depicted in Fig. 5 with the 2017/2018 outbreak would improve the model fit and predictions. A field validation using adequate and reliable data originating from surveillance conducted according to the OIE guidelines is ideal to further confirm the model utility. Especially to be used in neglected tropical disease control programs at country-level and determining the data requirements and to evaluate management and fieldwork of the current control programs. Such risk maps can also be used to identify areas with negligible risk and promote to become ‘rabies-free zones’ by prioritizing the implementation of minimum requirements of early warning system to ensure investigation and reporting of animals suspected of being infected, as guided by the OIE terrestrial code.

Sharing borders with neighboring countries was determined as a significant risk modifying factor implying the importance of dog-mediated rabies as a transboundary animal disease for Thailand. While, there are a few contradicting study findings, published literature provides phylogenetic evidences supporting the transboundary movement of the virus via reservoir hosts and anthropogenic introductions. Transboundary importance of an animal/zoonotic diseases emphasizes the need to consider the shared goals of dog vaccination, educational campaigns, and regional perspectives of disease mitigation to prevent reintroduction or exacerbation of the
disease. Interpretation of the importance of rabies as a transboundary disease may denote more efforts on diagnostic, surveillance, and vaccination the bordering districts than the movement restrictions or animal barriers. For example, a modeling approach by Ferguson et al. [78], conducted in the Philippines suggested that spending resources on the restriction of dog movements may not be beneficial in controlling rabies incidences.

The association with the number of Buddhist temples with rabies in Thailand denotes the importance of considering this sociocultural aspect when improving animal welfare, population control, Buddhist temples-based community engagements, awareness programs, and implementing vaccination campaigns to increase local compliance with rabies control. Awareness programs may address both human and animal sides of the disease control related to pet ownership, vaccination, and the importance of seeking medical attention and adherence to PEP regimen for suspected rabies exposure, which has been observed to be overlooked by residents of eastern Thailand [79]. Current educational campaigns organized by the government of Thailand recommends to report dog bites, to perform wound care immediately after, to seek medical attention for timely PEP, and to confine and monitor suspected animal for 10-days if its alive, and to send the carcass to the nearest qualifying laboratory facility to test for rabies [80]. The number of garbage dumps was not significantly associated with the disease, which can be attributed to the obervation that data represented all types of major dump areas and the accessibility of these areas by stray animals was unknown. Poverty was an important factor in both univariate and multivariate analyses, and several studies align with this observation that the socioeconomic status is a key determinant of the decision for animal vaccination; hence, offering free vaccination for the areas with low incomes and mobile vaccine campaigns have been suggested solutions in many instances [16,49].

Limitations of the study include the constraints of data availability, potential underreporting, and analytical assumptions. It is important to recognize that the scalability of this modeling approach require reliable data on reported cases and key risk modifying factors, which may vary by different country/setting. Passive surveillance data on rabies is subjected to the biases of underreporting given the frequency and quality of reporting are often affected by challenges to access healthcare facilities and laboratories, inadequate investment in surveillance, and lack of awareness [81]. Likely the extent of underreporting may vary for human and animal reporting systems. Therefore, it is ideal to have an index measuring the extent of underreporting of both human and animal rabies from districts, which would improve the model predictions similar to the study by Benavides et al. [82]. The potential underreporting may have led to increased number of false positives, which consequently led to lower specificity. Invariably, it is a trade-off between sensitivity, specificity, and several other factors when assigning risk thresholds and defining risk areas. On the bright side, the cross-validation indicated 80% of the districts determined as ‘negligible risk’ in 2015 did not report human or animal rabies cases in 2016 and 2017, which may denote an strength of this modeling approach to determine the reporting areas accurately at the least. We have attempted to account for neighbor effects through the CAR modeling however the thresholds and defining risk areas. On the bright side, the cross-validation indicated 80% of the districts determined as ‘negligible risk’ in 2015 did not report human or animal rabies cases in 2016 and 2017, which may denote an strength of this modeling approach to determine the reporting areas accurately at the least. We have attempted to account for neighbor effects through the CAR modeling however the thresholds and defining risk areas.

Acknowledgements

This project was funded by the Center for Global Health and Social Responsibility of the University of Minnesota. We would like to acknowledge the support of all the study by Benavides et al. [82]. The potential underreporting may have led to increased number of false positives, which consequently led to lower specificity. Invariably, it is a trade-off between sensitivity, specificity, and several other factors when assigning risk thresholds and defining risk areas. On the bright side, the cross-validation indicated 80% of the districts determined as ‘negligible risk’ in 2015 did not report human or animal rabies cases in 2016 and 2017, which may denote an strength of this modeling approach to determine the reporting areas accurately at the least. We have attempted to account for neighbor effects through the CAR modeling however the thresholds and defining risk areas.

Declarations of Competing Interest

Authors declare there is no known conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ohlet.2022.100411.

References

[1] K. Hampson, L. Coudeville, T. Lembo, M. Sambo, A. Kieffer, M. Attlan, et al., Estimating the global burden of endemic canine rabies, PloS Negl. Trop. Dis. 9 (4) (2015), https://doi.org/10.1371/journal.pntd.0003709. PubMed PMID: WOS:000354972200052.
[2] World Health Organization, WHO, Expert Consultation on Rabies, WHO Technical Report Series, Third Report, Geneva [cited 2017 May 2020]1012. ISBN:
