



# Spatial Trends in *Salmonella* Infection in Pigs in Spain

Kendy Tzu-yun Teng<sup>1\*</sup>, Marta Martínez Avilés<sup>2</sup>, María Ugarte-Ruiz<sup>1</sup>, Carmen Barcena<sup>1</sup>, Ana de la Torre<sup>2</sup>, Gema Lopez<sup>3</sup>, Miguel A. Moreno<sup>1,4</sup>, Lucas Domínguez<sup>1,4</sup> and Julio Álvarez<sup>1,4</sup>

<sup>1</sup> VISAVET Health Surveillance Center, Universidad Complutense, Madrid, Spain, <sup>2</sup> Center for Animal Health Research, National Institute of Agricultural and Food Research and Technology, Madrid, Spain, <sup>3</sup> Ministerio de Agricultura, Alimentación y Medio Ambiente (Spain), Madrid, Spain, <sup>4</sup> Department of Animal Health, Faculty of Veterinary Medicine, Universidad Complutense, Madrid, Spain

## OPEN ACCESS

### Edited by:

Francisco Ruiz-Fons,  
Consejo Superior de Investigaciones  
Científicas (CSIC), Spain

### Reviewed by:

Cristina Lanzas,  
North Carolina State University,  
United States  
Andrea Isabel Moreno Switt,  
Pontificia Universidad Católica de  
Chile, Chile

### \*Correspondence:

Kendy Tzu-yun Teng  
kendy.t.teng@gmail.com

### Specialty section:

This article was submitted to  
Veterinary Epidemiology and  
Economics,  
a section of the journal  
Frontiers in Veterinary Science

**Received:** 28 February 2020

**Accepted:** 18 May 2020

**Published:** 23 June 2020

### Citation:

Teng KT, Martínez Avilés M,  
Ugarte-Ruiz M, Barcena C,  
de la Torre A, Lopez G, Moreno MA,  
Domínguez L and Álvarez J (2020)  
Spatial Trends in *Salmonella* Infection  
in Pigs in Spain. *Front. Vet. Sci.* 7:345.  
doi: 10.3389/fvets.2020.00345

*Salmonella* is one of the most important foodborne pathogens worldwide. Its main reservoirs are poultry and pigs, in which infection is endemic in many countries. Spain has one of the largest pig populations in the world. Even though *Salmonella* infection is commonly detected in pig farms, its spatial distribution at the national level is poorly understood. Here we aimed to report the spatial distribution of *Salmonella*-positive pig farms in Spain and investigate the presence of potential spatial trends over a 17-year period. For this, data on samples from pigs tested for *Salmonella* in 2002–2013, 2015, 2017, and 2019 as part of the Spanish Veterinary Antimicrobial Resistance Surveillance program, representing 3,730 farms were analyzed. The spatial distribution and clustering of *Salmonella*-positive pig farms at the province level were explored using spatial empirical Bayesian smoothing and global Moran's *I*, local Moran's *I*, and the Poisson model of the spatial scan statistics. Bayesian spatial regression using a reparameterized Besag-York-Mollie Poisson model (BYM2 model) was then performed to quantify the presence of spatially structured and unstructured effects while accounting for the effect of potential risk factors for *Salmonella* infection at the province level. The overall proportion of *Salmonella*-positive farms was 37.8% (95% confidence interval: 36.2–39.4). Clusters of positive farms were detected in the East and Northeast of Spain. The Bayesian spatial regression revealed a West-to-East increase in the risk of *Salmonella* infection at the province level, with 65.2% (50% highest density interval: 70–100.0%) of this spatial pattern being explained by the spatially structured component. Our results demonstrate the existence of a spatial variation in the risk of *Salmonella* infection in pig farms at the province level in Spain. This information can help to optimize risk-based *Salmonella* surveillance programs in Spain, although further research to identify farm-level factors explaining this pattern are needed.

**Keywords:** *Salmonella*, pigs, spatial analysis, spatial modeling, BYM model, BYM2 model, Stan, Bayesian penalized regression

## INTRODUCTION

*Salmonella* infection is one of the most important foodborne zoonoses worldwide. There were 91,662 confirmed human salmonellosis cases in 2017 in the European Union (EU) (1) and ~1.2 million estimated cases occur every year in the US (2). Salmonellosis, characterized by acute onset of fever, abdominal pain, diarrhea, and nausea, is usually self-limiting. However, sometimes it can be life-threatening, especially in children, elderly and immunosuppressed patients, thus requiring antimicrobial therapy (3). This can be further complicated by the presence of antimicrobial-resistant strains (4).

Poultry is considered a major source of foodborne salmonellosis globally (5), but pork and pork products are also implicated in a large number of cases in many countries, including Spain (6, 7). Two of the most common *Salmonella* serotypes in pigs, *S. Typhimurium* and its monophasic variant (1,4,[5],12:i:-), were among the top serotypes associated with human salmonellosis in Spain (1). Moreover, the European Food Safety Authority (EFSA) baseline reports on *Salmonella* infection in fattening and breeder pig farms in Europe demonstrated that Spain had one of the highest levels of infection in pigs among EU countries, further highlighting their potential role in the occurrence of human salmonellosis in Spain (8, 9).

Still, due to the absence of a national control/monitoring program for *Salmonella* in pigs in Spain, little is known about the spatial distribution of *Salmonella* infection in pigs in Spain. The only available data came from specific studies with a generally limited geographical and temporal scope (10–12). The Spanish Veterinary Antimicrobial Resistance Surveillance Network program (13), starting in 1997, has performed nationwide surveillance of antimicrobial resistance originating from foodborne bacteria, such as *Escherichia coli*, *Campylobacter* spp., and *Salmonella* spp. Active surveillance of antimicrobial resistance in *Salmonella* in healthy pigs has been conducted through this program since 2002. Even though pigs from over 164 farms were sampled every year at the abattoirs since the beginning of the program, the spatial distribution of positive farms has never been evaluated.

Geographic information systems and some spatial statistical analyses have been applied to epidemiological research on *Salmonella* in farm animals. These approaches have allowed the detection of patterns and clusters of infection and prediction of occurrence and risk of infection of *Salmonella* under different situations. *K*-function analysis (14, 15) and Moran's *I* (16, 17) have been used to detect spatial clustering in *Salmonella* infection in farm animals, and a Gaussian kernel function has been used to predict the occurrence of *Salmonella*-infected dairy cattle and pig herds (18, 19). Using human cases, Simpson et al. (21) used the Besag-York-Mollié (BYM) hierarchical model (20) to map the cases of *S. Wangata* and *S. Typhimurium* in New South Wales, Australia (21). However, the usefulness of this method to better understand *Salmonella* infection in pigs seems not to be explored yet.

A BYM model contains two spatial random effects, often called spatially structured and unstructured components. The structured component has an intrinsic conditional autoregressive

(CAR) prior that takes the geographical contiguity into account (correlated heterogeneity). The geographical contiguity is described by the neighborhood relationships between each pair of areas and a full spatial dependency. The unstructured one is a random effect for non-spatial heterogeneity at the same area level as the structured component. Riebler et al. (22) proposed a parameterized BYM model—the BYM2 model to address some limitations of the original BYM model. Briefly, the BYM2 model adopts the penalized complexity framework that favors a model whose parameters have clear interpretations and thus facilitates the use of sensible hyperparameters in the model (23). The BYM2 model combines the two spatial components in the original BYM model into a single spatial component and allows a parameter to describe the proportion of the variance explained by the structured component.

In the current study, we aimed to evaluate the spatial distribution and potential spatial trends of *Salmonella* infection in pig farms in Spain. To do so, we analyzed the data on *Salmonella* detection in samples from pigs across a 17-year period, derived from the Spanish Veterinary Antimicrobial Resistance Surveillance Network program, using several spatial analytical techniques, including a BYM2 model.

## MATERIALS AND METHODS

### Study Population and Data Collection

Data on samples collected for monitoring antimicrobial resistance in *Salmonella* in pigs from 2002 to 2013, 2015, 2017, and 2019 (sampling was conducted every two years since 2015) were derived from the database of the Spanish Veterinary Antimicrobial Resistance Surveillance Network program. In the program, samples from fattening pigs were randomly collected in abattoirs selected based on their slaughter capacity. Each year, selected abattoirs altogether added up to more than 50% of the national slaughter capacity and were located in no less than half of the autonomous communities of the country. The total number of pigs sampled from each abattoir was proportional to their slaughter capacity and was randomly allocated to farm batches being culled on the sampling date. Animal samples consisted of at least 25 g of the content of caecum from two pigs selected at random from those coming from the same farm, except for 2011 when at least 15 g of ileo-caecal lymph nodes of one animal per farm were collected. Samples were collected by trained personnel, put in a clean container, and stored at refrigeration (3–8°C) until being sent to the laboratory within the next 36 h. *Salmonella* culture was performed immediately after reception.

### Bacteriology

*Salmonella* isolation was performed according to ISO 6579:2002/Amd 1:2007, the method recommended by the European Union Reference Laboratory for *Salmonella* in fecal and environmental samples [15]. Briefly, samples were cultured in buffered peptone water (BPW, 1/10 dilution; bioMérieux, Marcy-l'Étoile, France), followed by incubation at  $37 \pm 1^\circ\text{C}$  for  $18 \pm 2$  h. Modified semi-solid Rappaport-Vassiliadis (MSRV; Becton Dickinson France, Le Pont-de-Claix, France) agar plates

were then inoculated with three drops (i.e., 0.1 ml) of BPW culture. Plates were incubated at  $41.5 \pm 1^\circ\text{C}$  for  $24 \pm 3$  h and, if negative, incubated for an additional  $24 \pm 3$  h. Suspected growth of *Salmonella* was confirmed by plating out to both Xylose Lysine Desoxycholate agar (XLD; bioMérieux) and on chrom ID™ *Salmonella* agar (SM ID2; bioMérieux) for incubation during  $24 \pm 3$  h at  $37 \pm 1^\circ\text{C}$ .

Columbia 5% sheep blood agar (bioMérieux) was used for the incubation of colonies of presumptive *Salmonella* that were subcultured for  $24 \pm 3$  h at  $37 \pm 1^\circ\text{C}$ . All *Salmonella* isolates were confirmed by a commercial, biochemical method Enterotube™ II (BD BBL™; Becton Dickinson GmbH, Heidelberg, Germany). Serological typing was performed based on the White-Kauffmann-Le Minor scheme (24).

## Data on Potential Risk Factors (Covariates)

The potential of production-related characteristics to explain at least part of the observed spatial patterns of the risk of *Salmonella* infection in pig farms was assessed in the Bayesian spatial modeling (22 variables, Table 1) at the province level. These included (a) the number, the proportion, and the density (average number per square kilometer) of pigs belonging to different production categories (i.e., piglets, weaners, fattening pigs, gilts, sows, and boars) in each province and (b) the number and density of pig farms in each province. The pig-related information from 2005 to 2019 was collected from the website of the Ministry of Agriculture, Fisheries, and Food of Spain, and averaged over the study period for each province. The information about pig farm distribution was only available for 2016.

## Statistical Analyses

Data cleaning, manipulation, and analyses were performed in Microsoft Excel 2013 (Microsoft Corp.), and R program version 3.5.2 (25) in RStudio interface version 1.2.1330 (26). Descriptive analyses were facilitated by “tidyverse” package (27). The location of all farms from which the sampled pigs originated was available at the province level and used for the following analyses.

The overall, yearly and province-level proportion of *Salmonella*-positive farms was calculated. Empirical Bayesian smoothing was then performed on the proportion of *Salmonella*-positive farms at the province level to incorporate information on the sample size and the proportion of *Salmonella*-positive farms of neighboring provinces by using the “spdep” package (28). Gabriel Graph was used to describe the neighboring relationships between the provinces throughout the study. In addition, the overall and yearly proportion of farms positive to specific *Salmonella* serotypes over the total number of positive farms were calculated for serotypes with >50 isolates over the study period. Empirical Bayesian smoothing, as previously described, was also used to map the proportion of farms positive to these *Salmonella* serotypes over the total number of positive farms at the province level.

The presence of global and local spatial autocorrelation in the spatial distribution of *Salmonella* positive farms was explored using global and local Moran's statistics (29, 30). A global and local Moran's *I*-tests were run on the standardized residuals of a Poisson model using the number of positive farms in each

**TABLE 1 |** Univariable generalized linear regression results for the risk of *Salmonella* infection in pigs at the province level in Spain from 2002 to 2013, 2015, 2017, and 2019.

Name	Mean	Standard deviation	95% posterior probability	Change in risk with every specified unit of increase (95% credible interval)
Number of farms	0.08	0.03	(0.03 to 0.13)	1.010 (1.004 to 1.016) <sup>1</sup>
Density of farms (per km <sup>2</sup> )	0.07	0.02	(0.03 to 0.12)	1.014 (1.005 to 1.023) <sup>2</sup>
Number of fattening pigs	0.06	0.02	(0.03 to 0.10)	1.002 (1.001 to 1.003) <sup>3</sup>
Number of sows	0.06	0.02	(0.01 to 0.10)	1.010 (1.002 to 1.018) <sup>3</sup>
Number of piglets	0.05	0.02	(0.01 to 0.08)	1.002 (1.000 to 1.004) <sup>3</sup>
Number of weaners	0.06	0.02	(0.03 to 0.09)	1.003 (1.001 to 1.005) <sup>3</sup>
Number of gilts	0.05	0.02	(0.01 to 0.09)	1.040 (1.009 to 1.068) <sup>3</sup>
Number of boars	0.04	0.04	(−0.04 to 0.11)	1.262 (0.787 to 1.953) <sup>3</sup>
Total number of pigs	0.06	0.02	(0.02 to 0.10)	1.001 (1.000 to 1.001) <sup>3</sup>
Proportion of fattening pigs	0.08	0.04	(0.00 to 0.15)	1.009 (1.001 to 1.018) <sup>2</sup>
Proportion of sows	−0.11	0.06	(−0.24 to 0.00)	0.989 (0.976 to 1.000) <sup>2</sup>
Proportion of piglets	−0.08	0.03	(−0.15 to −0.02)	0.989 (0.981 to 0.997) <sup>2</sup>
Proportion of weaners	0.09	0.03	(0.02 to 0.15)	1.016 (1.004 to 1.028) <sup>2</sup>
Proportion of gilts	−0.06	0.06	(−0.19 to 0.04)	0.976 (0.929 to 1.017) <sup>2</sup>
Proportion of boars	−0.09	0.06	(−0.22 to 0.03)	0.799 (0.566 to 1.091) <sup>2</sup>
Density of fattening pigs	0.07	0.02	(0.04 to 0.11)	1.003 (1.002 to 1.005) <sup>4</sup>
Density of sows	0.05	0.02	(0.01 to 0.09)	1.011 (1.002 to 1.020) <sup>4</sup>
Density of piglets	0.05	0.02	(0.01 to 0.08)	1.002 (1.000 to 1.004) <sup>4</sup>
Density of weaners	0.08	0.02	(0.04 to 0.12)	1.005 (1.002 to 1.008) <sup>4</sup>
Density of gilts	0.06	0.02	(0.01 to 0.10)	1.050 (1.012 to 1.089) <sup>4</sup>
Density of boars	0.05	0.03	(−0.02 to 0.11)	1.643 (0.813 to 3.211) <sup>4</sup>
Density of pigs	0.07	0.02	(0.03 to 0.11)	1.001 (1.000 to 1.002) <sup>4</sup>

<sup>1</sup>Every 100 farms of change.

<sup>2</sup>Every 1% of change.

<sup>3</sup>Every 10,000 animals of change.

<sup>4</sup>Every 1 unit of change.

province as the outcome and the expected number of positive farms as the offset with the “spdep” package (28). The significance of the global Moran's *I* statistic was estimated through 999 Monte Carlo simulations in which the residuals were randomly shuffled across provinces. For local Moran's *I*, the *P*-values were calculated using the expectation and variance and corrected with

the method described in Benjamini and Hochberg (31). The significance level for all the tests in the current study was set at 0.05.

Additionally, the Poisson model of the scan statistic was also applied to detect the presence of provinces with an increased risk of *Salmonella* infection using the centroid of each province as the point location, facilitated by the “SpatialEpi” package (32, 33). The scan statistic detects the maximum likelihood ratio between the value inside and outside a searching window over the likelihood function under the null hypothesis of complete spatial randomness (32). The search was performed using circular spatial moving windows that contained up to 25% or 50% of the total population alternatively. The pseudo *P*-values of the most likely clusters were generated by comparing the observed risk in the windows with the expected, generated through 999 Monte Carlo simulations in which the risks at each location were randomly allocated.

Bayesian spatial modeling to assess the associations between the risk of *Salmonella* infection at the province level and the available covariates was performed in Rstudio using Stan and associated packages (34). The number of observed *Salmonella*-positive pig farms in different provinces was assumed to follow a Poisson distribution, with the expected number of positive farms in each province as the offset. Regression models were fitted using the “brms” package (35). To explore the directionality of the association between available covariates and the outcome, univariable non-spatial models were first fitted introducing alternatively each of the covariates. For covariate selection, the predictive projection technique proposed by Piironen et al. was implemented using the “projpred” package (36, 37). The selection process contained two steps. First, a Bayesian penalized regression model with a regularized horseshoe prior and all the covariates was constructed as the reference model that warranted a good prediction ability. Penalized regression is a statistical technique designed to avoid overfitting, especially in cases of a large number of covariates (38). This is achieved through the introduction of a penalty term that shrinks small coefficients toward zero while leaving large coefficients large. The implement of penalized regression is rather intuitive within a Bayesian framework as a penalty term can be included as a hyperprior, also called shrinkage prior (38). Many shrinkage priors have been proposed, and a regularized horseshoe prior was chosen for the current study due to the advantages discussed in Piironen and Vehtari (39). Second, the covariates in the best model for each submodel size were identified by decreasing the Kullback–Leibler divergence from the reference model to the projected submodel using a forward stepwise addition. A submodel with the minimal subset of these covariates which had similar predictive power as the reference model, judged by the mean log predictive density and the root mean square error, was selected.

The BYM2 Poisson model including the selected covariates and the spatial components (Equation 1) was then fitted (22). To examine the suitability of alternative distributions to fit the data, another three BYM2 models were fitted with different likelihoods (i.e., zero-inflated Poisson, negative binomial and zero-inflated negative binomial) followed by model selection using Bayesian leave-one-out cross-validation (40, 41). The

default priors specified in the “brms” package were used in the current analyses. Sampling was drawn from four Markov chains with 1,000 iterations. The results final model reported in the next section were sampled from four Markov chains with 3,000 iterations. Half of the iterations were for warm-up (i.e., burn-in) and not used for inference.

Markov chain Monte Carlo diagnostics for the final model were performed with (a) the potential scale reduction statistic ( $\hat{R}$ ) (42), (b) the ratio of the effective sample size to the total sample size drawn from the posterior distribution, and (c) trace plots of Markov chain Monte Carlo generated through the “bayesplot” package (43). Residual check and posterior predictive checks were also performed using the “bayesplot” package (43, 44). The highest density interval of the posterior distribution was estimated using the “bayestestR” package (45).

**Equation 1.** Spatial component in a BYM2 model.

$$T = \frac{1}{\sqrt{T_t}}(\sqrt{p/f} * S + \sqrt{1-p})$$

- *T*: total spatial component,
- $\tau_t$ : precision for the total spatial component;  $\frac{1}{\sqrt{T_t}}$  is the overall standard deviation,
- $p \in [0, 1]$ : the proportion of the variance explained by the spatially structured component,
- *f*: scaling factor,
- *S*: spatially structured component,
- *U*: spatially unstructured component.

For the variables potentially associated with the *Salmonella* detection (see below), the existence of major changes in their spatial distribution over the study period was evaluated using Friedman tests on the province-level yearly data, followed by pairwise Wilcoxon rank sum tests with *P*-values adjusted for multiple comparisons (46).

## RESULTS

Up to 3,730 samples collected over the 15 years in which sampling was conducted, representing the same number of farms, were included in the current study, with an average of 249 (range: 163–384) samples per year. The number of abattoirs where the samples were collected each year, except for 2019 when this information was not available, ranged between 7 and 20. A median of 18 samples (interquartile interval: 11–29, range: 1–60) were collected from each abattoir each year during the study period. Abattoirs were located in 11 out of the 18 autonomous communities in Spain, and 977 (29.2%), 670 (20.0%) and 455 (13.3%) samples came from abattoirs in Cataluña, Castilla La Mancha and Murcia, respectively (**Supplementary File 1**). The sampled farms were located in 43 out of the 52 provinces in Spain; 502 (13.5%) were from Murcia, 371 (9.95%) from Huesca, and 334 (9.0%) from Lleida.

A total of 1,409 of the 3,730 samples were positive, yielding an overall percentage of *Salmonella*-positive farms of 37.8% (95% confidence interval [CI]: 36.2–39.4). This percentage peaked between 2004 (54.2%) and 2006 (53.3%), declined to 29.7% in

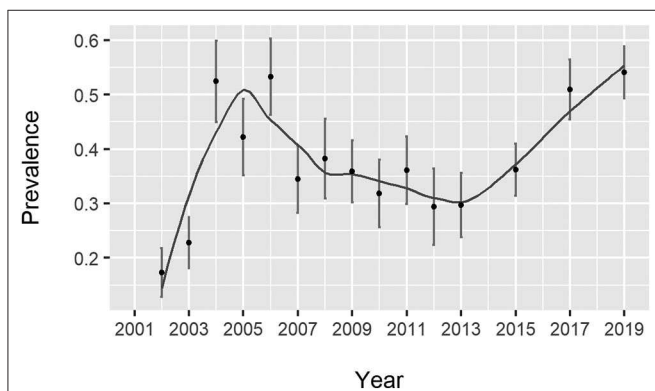
2012, and increased again to 54.1% in 2019 (**Figure 1**). The raw and smoothed proportion of *Salmonella*-positive farms at the province level is shown in **Figure 2**. In the eleven provinces with more than 100 samples, the spatially adjusted proportion ranged from 17.8 % (95% CI: 13.1–22.2) in Toledo to 44.8% (95% CI: 35.4–54.2) in Almeria.

The serotype of 1,360 (96.5%) out of the total 1,409 *Salmonella* isolates recovered was determined, yielding 64 distinct serotypes (**Supplementary Table 1**). The most represented were *S. Rissen* (313, 22.2%), *S. 1,4,[5],12:i:-* (265, 18.8%), *S. Typhimurium* (251, 17.8%), and *S. Derby* (207, 14.7%). The evolution of the proportion of isolates belonging to specific serotypes over the study period varied (**Figure 3**). The proportion of *S. Typhimurium* decreased after 2010, while the proportion of the *S. 1,4,[5],12:i:-* has consistently increased over the years. The proportion of *S. Rissen* remained consistently around 0.3 after 2005, and, for *S. Derby*, a decreasing trend was observed (from slightly lower than 0.3 to <0.2 from 2012 onwards). After the empirical Bayesian smoothing, the proportions of *Salmonella* isolates identified as the *S. 1,4,[5],12:i:-* and *S. Derby*

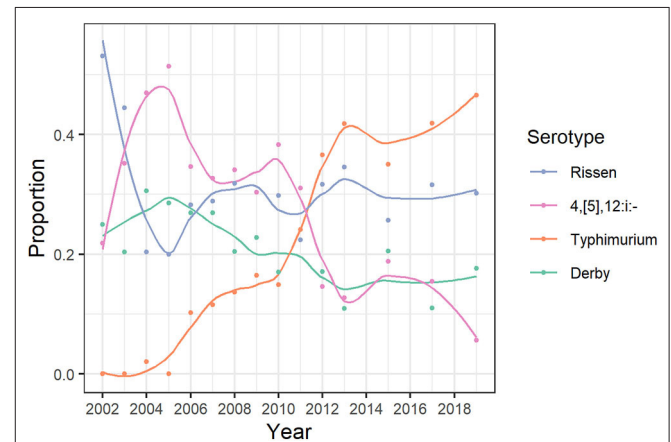
were higher in the provinces of Northeast and South of Spain, respectively (**Figure 4**). Both the proportions of *S. Rissen* and the *S. 1,4,[5],12:i:-* were rather homogeneous across provinces in Spain but slightly lower at the northwest corner and the South of Spain, respectively.

No global (Moran's  $I = -0.02$ ,  $p = 0.458$ ) or local spatial autocorrelation was detected in the standardized residuals of the Poisson model. The spatial scan statistics identified local clusters with an increased risk of *Salmonella* infection in the East and Northeast of Spain (**Figure 5**). The observed-to-expected ratio between inside and outside of the significant clusters identified using a search window of maximum 25 and 50% population was 1.17 ( $P = 0.036$ ) and 1.13 ( $P = 0.001$ ), respectively.

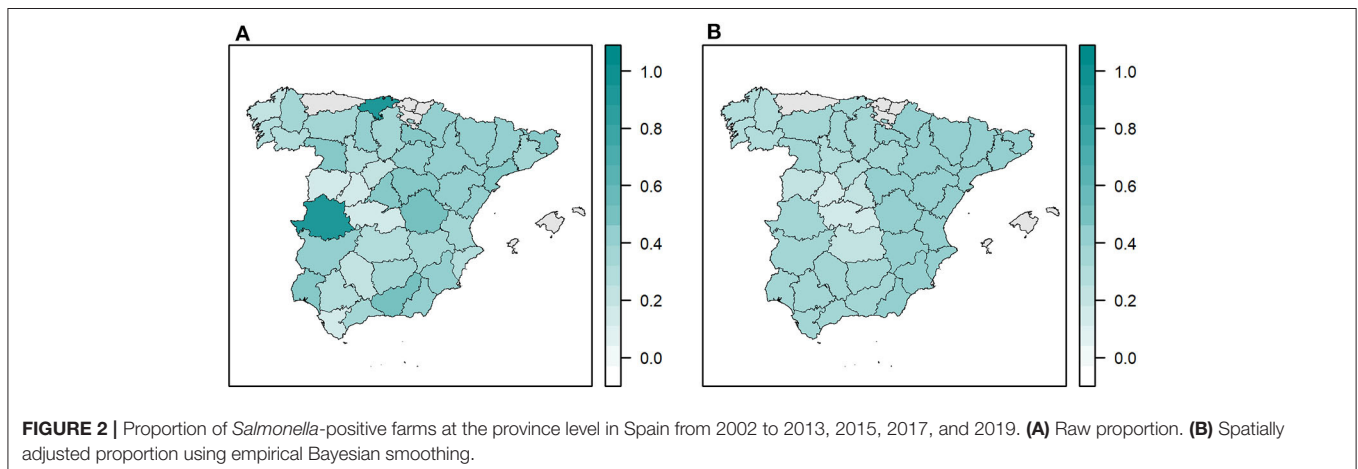
According to the univariable models, provinces with a higher number or density of pig farms showed a higher risk of *Salmonella* infection (**Table 1**). In general, covariates related to the population of fattening pigs, weaners, sows or piglets were



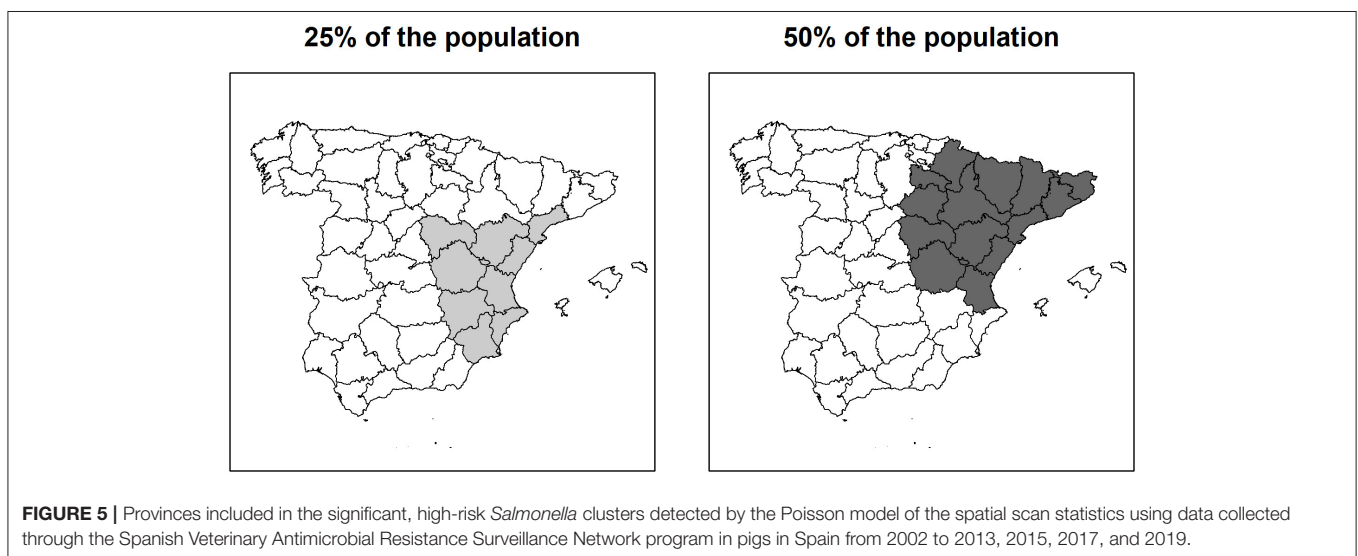
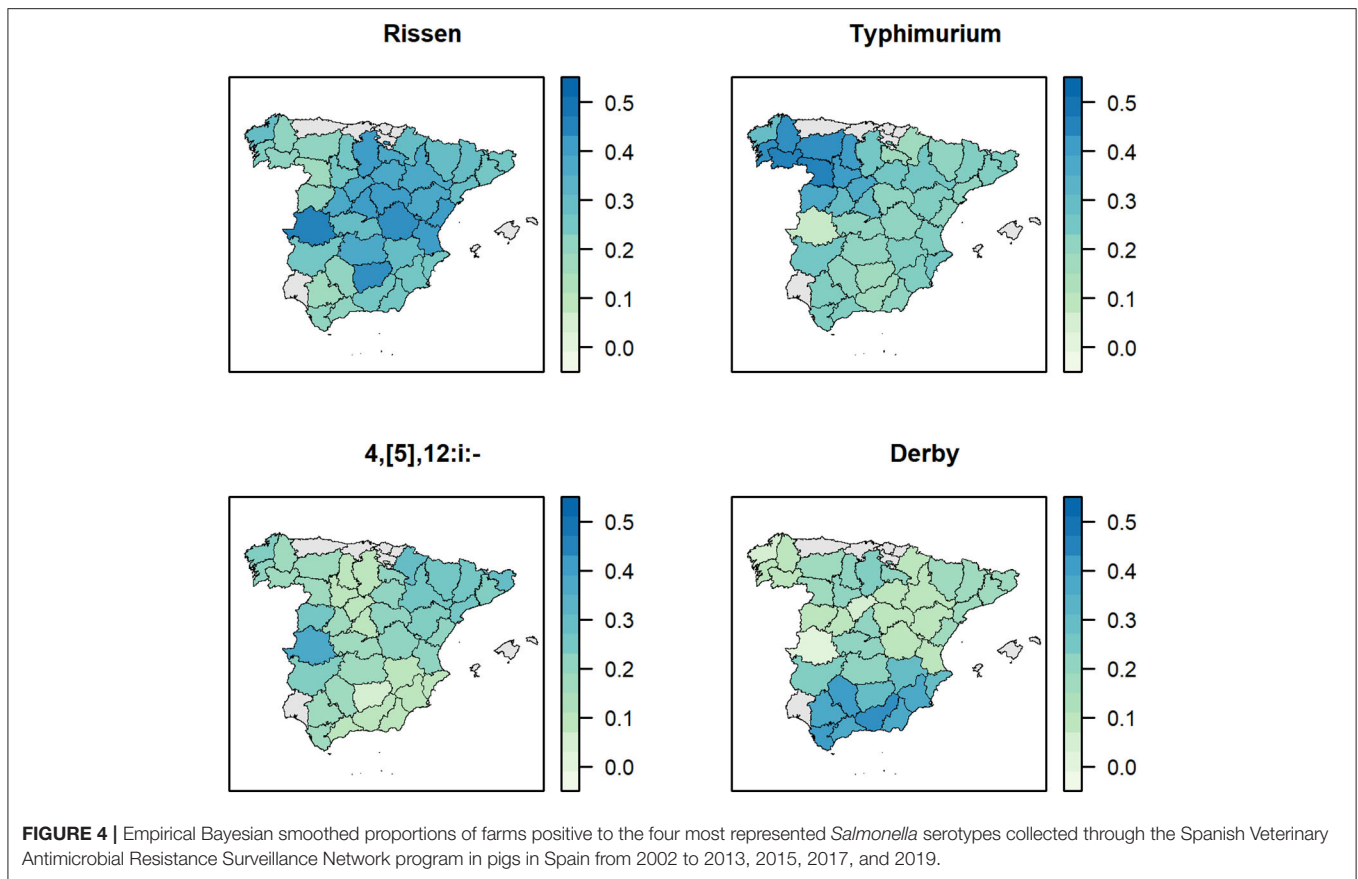
**FIGURE 1** | Annual proportion of *Salmonella*-positive farms in Spain from 2002 to 2013, 2015, 2017, and 2019. The line was smoothed by a locally estimated scatterplot smoothing with a span of 0.5.



**FIGURE 3** | Changes in the proportion of *Salmonella* isolates recovered through the Spanish Veterinary Antimicrobial Resistance Surveillance Network program belonging to one of the four most common serotypes collected in pigs in Spain from 2003 to 2013, 2015, 2017, and 2019. The lines were smoothed by a locally estimated scatterplot smoothing with a span of 0.5.



**FIGURE 2** | Proportion of *Salmonella*-positive farms at the province level in Spain from 2002 to 2013, 2015, 2017, and 2019. **(A)** Raw proportion. **(B)** Spatially adjusted proportion using empirical Bayesian smoothing.



positively associated with the probability of finding positive *Salmonella* farms.

After predictive projection, only one covariate, the density of weaners, was selected to be included in the final Poisson BYM2 model (Table 2). More results of Bayesian penalized regression and predictive projection can be found in Supplementary File 2.

In the Poisson model including only the density of weaners without the BYM2 component, the probability of finding positive *Salmonella* farms at the province level increased by 0.5% (95% credible interval [CrI]: 0.2–0.8%) with every increase in the weaner density. However, the effect of this covariate shrank to close to 0 after the inclusion of the BYM2 component (Table 2).

**TABLE 2 |** Regression results from the final multivariable modeling for risk of *Salmonella* infection in pigs at the province level in Spain from 2002 to 2013, 2015, 2017, and 2019.

Variable	Mean	Standard deviation	95% posterior probability	Change of risk with very 1 unit of increase (95% credible interval)
Intercept	-0.07	0.05	-0.17 to 0.02	-
Density of weaners	-0.01	0.06	-0.13 to 0.10	-0.001 (-0.008 to 0.007)
Standard deviation of the spatial component	0.23	0.05	0.14 to 0.35	-
Proportion explained by the structured component	0.65	0.25	0.10 to 0.99	-

The spatial component in the final BYM2 model suggested a West-East increasing risk of *Salmonella* infection at the province level in Spain. The exponentiated means of the spatial effects are shown in **Figure 6**. The mean of the standard deviation of the spatial component was 0.23 (CrI: 0.14–0.35), and, on average, 65.2% (median: 70.2%, 50% highest density interval: 70–100%) of the variance of the spatial component was explained by the structured component. No specific pattern in the unstructured spatial component was observed.

Markov chain Monte Carlo and model diagnoses are presented in **Supplementary File 3**.

The density of weaners at the province level did not experience major changes over the study period, with higher values reported consistently for provinces in the Northeast corner (**Supplementary File 4**). Still, the Friedman test revealed the existence of significant ( $P < 0.001$ ) differences in the yearly values over time, although *post-hoc* tests revealed that differences were only due to values recorded in 2017 and 2019. The test became non-significant ( $P = 0.071$ ) when values from 2005 to 2015 were used.

## DISCUSSION

Human salmonellosis outbreaks have been linked to pork and pork products in the past (1, 5, 7). In many regions and countries such as Northern and Western Europe and Japan, pork, and pork products are the second most common source for human salmonellosis after eggs and egg products (5). In Spain, pork and pork products have been shown to be one of the top-ranked sources for human salmonellosis (7). However, currently, there is no official control program of *Salmonella* in pig production in Spain (47). In the current study, we used several analytical techniques to explore the spatial distribution of *Salmonella* infection in pigs in Spain and generate information that can be used for surveillance of *Salmonella* in pigs in the future.

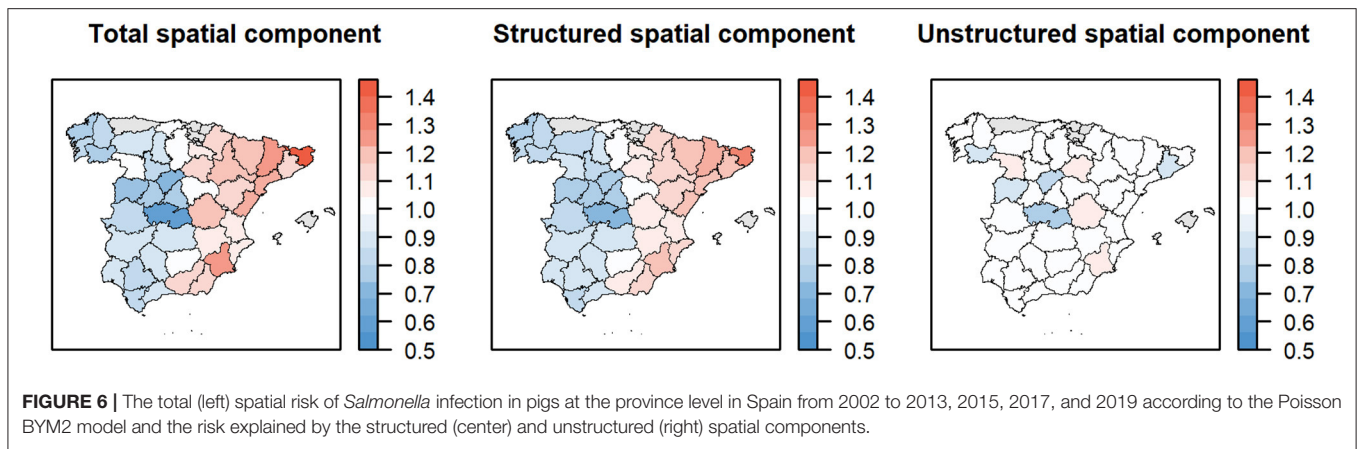
We detected an overall percentage of *Salmonella*-positive farms of 37.8% with great variation across the years. Several

studies have reported even higher values of farm-level *Salmonella* prevalence in Catalonia (77.3% in 2000–2003), northeast Spain [94.1% in 2008–2009 (12)], and the entire country [43.1% in Spain in 2003–2004 (11)]. The lower values suggested by the current study could be partially due to the inclusion of feces from only two pigs per farm. Thus, the probability of detecting the presence of *Salmonella* at the farm level, particularly in farms with a low within-farm prevalence, was not as high as in the aforementioned studies. Nonetheless, the spatial heterogeneity found in the current study agrees with these previous results reporting higher prevalence values in the Northeast of the country, where a large proportion of Spanish pig population is located.

Our results indicated that the recent increase in the percentage of *Salmonella*-positive farms could be related to the increasingly reported S. 1,4,[5],12:i:-. During the study period, the proportion of isolates belonging to this serotype steadily rose and reached 0.47 in 2019, in parallel to the increase in the yearly percentage of *Salmonella*-positive farms (from 29.4 in 2012 to 54.1 in 2019). The importance of the S. 1,4,[5],12:i:- in public health highlights the need for continuous monitoring on its prevalence in pigs in Spain and the EU at large (48).

Several factors may affect the determination of the *Salmonella* status of a farm when samples are collected at the abattoir. Many studies have shown a noticeably higher prevalence of *Salmonella* in samples from slaughtered pigs than the prevalence of the samples from the farm (49, 50). This could be due to the stress generated by the process of harvest, transportation, and retention in the lairage, resulting in the recrudescence of latent carriers and/or an increase in the susceptibility of pigs to new infections (49). Moreover, the long feed withdrawal normally performed before transport might change the gut microbiota and increase the number of *Salmonella* in the fecal content (51). On the other hand, an increase in the diversity of serotypes detected in the abattoir compared to those recovered from the farms of origin has been also described. This could suggest exposure to additional contaminated environments such as trucks and lairages (49, 52). Therefore, the true prevalence of *Salmonella* infection at the farm level may be lower than the detected/apparent prevalence based on samples collected at abattoirs, and the diversity of serotypes could be the result of *Salmonella* from both pig farms and places involved in the process between harvest and slaughtering.

Our results showed a clear pattern suggesting a higher risk of *Salmonella* infection in the (North-)east than in the rest of Spain (**Figure 6**). The BYM2 model showed that, on average, 65.2% of the spatial effect could be explained by the underlying geographical location of the provinces, and the highest density interval included even higher values. We considered several pig population-related covariates in the analysis given that their heterogeneous spatial distribution was in agreement to some extent to the results of the spatial scan statistic (i.e., more pigs and higher densities in the Northeast of Spain). The univariable modeling results showed that many covariates, especially the number of farms and those related to fattening pigs and weaners, were indeed positively associated with the risk. However, only one covariate, the density of weaners, was retained after the



variable selection process. This was expected as many covariates were correlated. Nonetheless, the effect of the density of weaners became close to 0 after the inclusion of the BYM2 component. This suggested that the observed spatial distribution of risk was better explained by the geographical contiguity of the provinces than by the density of weaners and other pig-related factors.

The software used here, Stan, is a highly-expressive probabilistic programming language that allows full Bayesian inference using Hamiltonian Monte Carlo samplers (34), which have been shown more efficient and robust than Gibbs and Metropolis samplers (53). Although many R packages to facilitate Bayesian modeling in Stan have been developed (35, 54), it is yet to be much explored in the field of veterinary epidemiology. With Stan and “brms” package (35), BYM2 model can be performed by regular R users (22).

To our knowledge, the BYM2 model has been little utilized in the field of veterinary science. The BYM2 model has advantages over the original and some of the reparameterized BYM models (22). Firstly, the two components in a BYM2 model can be seen independently from each other, resulting in better estimation for both of them (55). Secondly, BYM2 models facilitate parameters that have clear interpretations and thus the use of sensible hyperparameters (23). Additionally, as the scaling factor in BYM2 models is placed to take into account the underlying neighborhood structure, studies with different neighborhood structure now can use the same hyperprior in the models (22). Therefore, the future application of the BYM2 model in veterinary epidemiology may be encouraged. A BYM2 model can also be performed with INLA (22).

In the current study, we employed a relatively unexplored approach for variable selection in veterinary epidemiology—predictive projection with a Bayesian penalized regression model as the reference model. Shrinkage methods have been recommended when the ratio of the number of observations to the number of variables is  $\leq 10$  (56), and predictive projection is useful in determining the number of variables to be included in the model (36). This technique has several advantages. First, it requires less computational power than cross-validation and is less time-consuming than using either cross-validation or information criteria. Second, selection among many models

using cross-validation may tend to overfit and thus result in choosing a suboptimal model (36).

Here, a number of different spatial analytic tools were applied, offering different results. While the global and local Moran’s *I*-tests ran on the residuals of a Poisson model did not suggest the existence of a spatial pattern in the distribution of the *Salmonella* risk at the province level, the spatial scan statistic and the Bayesian spatial model showed the opposite. This finding suggests that, rather than making a conclusion of spatial independence based on one test, the application of more than one spatial analytic test, based on different hypotheses and assumptions, can provide a more complete picture.

The current study has some limitations. First, as the sample collection was conducted in abattoirs that have high slaughter capacity, the results might not be necessarily representative of the farms that did not (usually) send their pigs to those abattoirs. Second, the current study was conducted using secondary data. Therefore, it may face some common issues of using secondary data, such as out-of-date information, suboptimal sampling procedure for answering specific research questions, insufficient sample size, and lack of information that would be, otherwise, included. For example, the individual or within-farm prevalence of *Salmonella* infection could not be determined in the current study. Also, farm-specific information was not available so could not be included in the models. Farm-specific covariates such as farms types, and management and biosecurity information, will likely be associated with the outcome and allow prediction of the risk. Furthermore, for the multivariable modeling exercise, farm-related information was only available for 2016, and data of pig distribution was averaged across all years under study. Still, the exploration of the evolution of the pig distribution in Spain revealed no significant variation between provinces for at least most of the study period (2002–2015), suggesting that aggregating values across such a period would not result in a major loss of information. Lastly, it has been shown that the inclusion of a spatially-correlated component only after the covariate selection process may affect the results of the covariates in the model (57). This is observed in the current study. Ideally, the predictive projection should be conducted with the BYM2 component included. However, this



is currently unavailable in the “projpred” package and thus was not done.

## CONCLUSION

The current study shows a notable increasing trend in the risks of *Salmonella* infection in pig farms located in provinces from West to East in Spain, evident still after the possible effect of the heterogeneous distribution of pigs in the country was accounted for. The increase in the percentage of *Salmonella* positive farms from 2012 and the *S.* 1,4,[5],12:i:- in Spain demonstrates the usefulness of surveillance to detect changes in the epidemiology of this foodborne pathogen in the animal reservoirs. We demonstrated the usefulness of Stan for various applications that are commonly pursued in a veterinary epidemiological study, such as covariate selection, model selection, and model fit assessment, as well as fitting a BYM2 model. The information generated by the current study can be used for risk-based *Salmonella* antimicrobial resistance surveillance programs in the future, so the probabilities of selecting positive farms and specific serotypes can be optimized. Although some temporal trends in the risk of *Salmonella* are shown in the current study, more data is needed to allow a better understanding of the spatial-temporal distribution and the evolution of *Salmonella* infection in pigs in Spain.

## DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

## ETHICS STATEMENT

The current work did not involve any live animal nor human data. No ethical approval was required as all isolates analyzed

were retrieved through the ongoing Spanish national surveillance program on antimicrobial resistance, performed according to national and EU regulations.

## AUTHOR CONTRIBUTIONS

KT and JA contributed conception and design of the study. KT, MM, MU-R, CB, AT, GL, MAM, LD, and JA participated in the generation, collection, and curation of the data. KT performed the statistical analysis. KT, MM, JA interpreted the results. JA provided supervision. MU-R, GL, and JA facilitated project administration. KT wrote the first draft of the manuscript. MU-R and JA wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

## FUNDING

This research was funded by the European Union's Horizon 2020 Research and Innovation program under Grant Agreement No. 773830 (One Health European Joint Program) and the Community of Madrid project S2013/ABI-2747 for financial support.

## ACKNOWLEDGMENTS

The authors thank Vicente Lopez-Chavarrias for assisting data curation, Maria García-Martin for the logistics, and all the laboratory staff who have performed laboratory work during the 17 years.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fvets.2020.00345/full#supplementary-material>

## REFERENCES

- European Food Safety Authority, European Centre for Disease Prevention and Control. The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2017. *EFSA J.* (2018) 16:262. doi: 10.2903/j.efsa.2018.5500
- Scallan E, Hoekstra RM, Angulo FJ, Tauxe RV, Widdowson M-A, Roy SL, et al. Foodborne illness acquired in the United States-major pathogens. *Emerg Infect Dis.* (2011) 17:7–15. doi: 10.3201/eid1701.P11101
- World Health Organization. *Salmonella (non-typhoidal)*. (2018). Available online at: <https://www.who.int/news-room/fact-sheets/detail/salmonella-non-typhoidal> (accessed July 23, 2019).
- European Food Safety Authority, European Centre for Disease Prevention and Control. The European Union summary report on antimicrobial resistance in zoonotic and indicator bacteria from humans, animals and food in 2017. *Eur Cent Dis Prev Control.* (2019) 17:5598. doi: 10.2903/j.efsa.2019.5598
- Pires SM, Vieira AR, Hald T, Cole D. Source attribution of human salmonellosis: an overview of methods and estimates. *Foodborne Pathog Dis.* (2014) 11:667–76. doi: 10.1089/fpd.2014.1744
- De Knecht LV, Pires SM, Hald T. Attributing foodborne salmonellosis in humans to animal reservoirs in the European Union using a multi-country stochastic model. *Epidemiol Infect.* (2015) 143:1175–86. doi: 10.1017/S0950268814001903
- Martínez M, Novoa I, De La Torre A. Descriptive analysis of *Salmonella* food alerts in Spain (2007–2017). In: *One Health EJP Annual Scientific Meeting*. Dublin (2019).
- European Food Safety Authority. Analysis of the baseline survey on *Salmonella* in breeding pigs in the EU, 2008 - Part A: *Salmonella* prevalence estimates. *EFSA J.* (2009) 7:1–93. doi: 10.2903/j.efsa.2009.1377
- European Food Safety Authority. Report of the task force on zoonoses data collection on the analysis of the baseline survey on the prevalence of *Salmonella* in slaughter pigs, in the EU, 2006–2007- Part A: *Salmonella* prevalence estimates. *EFSA J.* (2008) 135:1–111. doi: 10.2903/j.efsa.2008.135r
- Casanova-Higes A, Marín-Alcalá CM, Andrés-Barranco S, Cebollada-Solanas A, Alvarez J, Mainar-Jaime RC. Weaned piglets: another factor to be considered for the control of *Salmonella* infection in breeding pig farms. *Vet Res.* (2019) 50:45. doi: 10.1186/s13567-019-0666-7
- García-Feliz C, Carvajal A, Collazos JÁ, Rubio P. Herd-level risk factors for faecal shedding of *Salmonella enterica* in Spanish fattening pigs. *Prev Vet Med.* (2009) 91:130–136. doi: 10.1016/j.prevetmed.2009.05.011
- Vico JP, Rol I, Garrido V, Román BS, Grilló MJ, Mainar-Jaime RC. Salmonellosis in finishing pigs in Spain: prevalence, antimicrobial agent susceptibilities, and risk factor analysis. *J Food Prot.* (2011) 74:1070–8. doi: 10.4315/0362-028X.JFP-10-515
- Moreno MA, Dominguez L, Teshager T, Herrero IA, Porrero MC. Antibiotic resistance monitoring: the Spanish programme. *Int J Antimicrob Agents.* (2000) 14:285–90. doi: 10.1016/S0924-8579(00)00138-2

14. Fenton SE, Clough HE, Diggle PJ, Evans SJ, Davison HC, Vink WD, et al. Spatial and spatio-temporal analysis of *Salmonella* infection in dairy herds in England and Wales. *Epidemiol Infect.* (2009) 137:847–57. doi: 10.1017/S0950268808001349
15. Rao S, Kitron U, Weigel RM. Spatial and genotypic clustering of *Salmonella* over time in a swine production unit. *Prev Vet Med.* (2010) 97:90–9. doi: 10.1016/j.prevetmed.2010.09.005
16. Graziani C, Luzzi I, Owczarek S, Dionisi AM, Busani L. *Salmonella enterica* serovar napoli infection in Italy from 2000 to 2013: spatial and spatio-temporal analysis of cases distribution and the effect of human and animal density on the risk of infection. *PLoS One.* (2015) 10:e0142419. doi: 10.1371/journal.pone.0142419
17. de Alwis R, Watson C, Nikolay B, Lowry JH, Thieu NTV, Van TT, et al. Role of environmental factors in shaping spatial distribution of *Salmonella enterica* serovar typhi, Fiji. *Emerg Infect Dis.* (2018) 24:284–93. doi: 10.3201/eid2402.170704
18. Lailler R, Sanaa M, Chadouef J, Fontez B, Brisabois A, Colmin C, et al. Prevalence of multidrug resistant (MDR) *Salmonella* in bovine dairy herds in western France. *Prev Vet Med.* (2005) 70:177–89. doi: 10.1016/j.prevetmed.2005.03.006
19. Benschop J, Stevenson MA, Dahl J, French NP. Towards incorporating spatial risk analysis for *Salmonella* sero-positivity into the Danish swine surveillance programme. *Prev Vet Med.* (2008) 83:347–59. doi: 10.1016/j.prevetmed.2007.09.005
20. Besag J, York J, Mollié A. Bayesian image restoration, with two applications in spatial statistics. *Ann Inst Stat Math.* (1991) 43:1–20. doi: 10.1007/BF00116466
21. Simpson KMJ, Mor SM, Ward MP, Walsh MG. Divergent geography of *Salmonella* wangata and *Salmonella* Typhimurium epidemiology in New South Wales, Australia. *One Health.* (2019) 7:100092. doi: 10.1016/j.onehlt.2019.100092
22. Riebler A, Sørbye SH, Simpson D, Rue H. An intuitive Bayesian spatial model for disease mapping that accounts for scaling. *Stat Methods Med Res.* (2016) 25:1145–65. doi: 10.1177/0962280216660421
23. Simpson DP, Rue H, Martins TG, Riebler A, Sørbye SH. Penalising model component complexity: a principled, practical approach to constructing priors. *Stat Sci.* (2017) 32:1–28. doi: 10.1214/16-STSS576
24. Grimont PA, Weill FX. Antigenic formulae of the *Salmonella* serovars. *WHO Collab Cent Ref Res Salmonella.* (2007) 9:1–66.
25. R Core Team. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing (2019). Available online at: <https://www.R-project.org/>
26. RStudio Team. *RStudio: Integrated development for R.* Boston, MA: RStudio, Inc. (2019). Available online at: <http://www.rstudio.com/>
27. Wickham H. *tidyverse: Easily install and load the “tidyverse.”* (2017). Available online at: <https://CRAN.R-project.org/package=tidyverse>
28. Bivand R, Wong DWS. Comparing implementations of global and local indicators of spatial association. *Test.* (2018) 27:716–48. doi: 10.1007/s11749-018-0599-x
29. Anselin L. Local indicators of spatial association-LISA. *Geogr Anal.* (1995) 27:93–115. doi: 10.1111/j.1538-4632.1995.tb00338.x
30. Moran PAP. Notes on continuous stochastic phenomena. *Biometrika.* (1950) 37:17–23. doi: 10.1093/biomet/37.1-2.17
31. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Methodol.* (1995) 57:289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x
32. Kulldorff M. A spatial scan statistic. *Commun Stat Theory Methods.* (1997) 26:1481–96. doi: 10.1080/03610929708831995
33. Chen C, Kim AY, Ross M, Wakefield J. *SpatialEpi: Methods and Data for Spatial Epidemiology.* R package version (2018). p. 1.
34. Carpenter B, Gelman A, Hoffman MD, Lee D, Goodrich B, Betancourt M, et al. Stan: a probabilistic programming language. *J Stat Softw.* (2017) 76:1–32. doi: 10.18637/jss.v076.i01
35. Bürkner P-C. brms: An R package for Bayesian multilevel models using Stan. *J Stat Softw.* (2017) 80:1–28. doi: 10.18637/jss.v080.i01
36. Piironen J, Paasiniemi M, Vehtari A. Projective inference in high-dimensional problems: prediction and feature selection. (2018) *arXiv:1810.02406 Cs Stat.*
37. Piironen J, Paasiniemi M, Vehtari A, Gabry J, Bürkner P-C. *projpred: Projection Predictive Feature Selection.* (2019). Available online at: <https://CRAN.R-project.org/package=projpred> (accessed September 26, 2019).
38. van Erp S, Oberski DL, Mulder J. Shrinkage priors for Bayesian penalized regression. *J Math Psychol.* (2019) 89:31–50. doi: 10.1016/j.jmp.2018.12.004
39. Piironen J, Vehtari A. Sparsity information and regularization in the horseshoe and other shrinkage priors. *Electron J Stat.* (2017) 11:5018–51. doi: 10.1214/17-EJS1337SI
40. Geisser S, Eddy WF. A predictive approach to model selection. *J Am Stat Assoc.* (1979) 74:153–60. doi: 10.1080/01621459.1979.10481632
41. Vehtari A, Gabry J, Yao Y, Gelman A. *loo: Efficient Leave-One-Out Cross-Validation and WAIC for Bayesian Models.* (2019). Available online at: <https://CRAN.R-project.org/package=loo> (accessed September 26, 2019).
42. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. *Bayesian Data Analysis.* New York, NY: Chapman and Hall/CRC (2013).
43. Gabry J, Mahr T. *bayesplot: Plotting for Bayesian Models.* (2019). Available online at: [mc-stan.org/bayesplot](http://mc-stan.org/bayesplot)
44. Gabry J, Simpson D, Vehtari A, Betancourt M, Gelman A. Visualization in Bayesian workflow. *J R Stat Soc.* (2019) 182:389–402. doi: 10.1111/rssa.12378
45. Makowski D, Ben-Shachar M, Lüdtke D. bayestestR: Describing effects and their uncertainty, existence and significance within the Bayesian framework. *J Open Source Softw.* (2019) 4:1541. doi: 10.21105/joss.01541
46. Holm S. A simple sequentially rejective multiple test procedure. *Scand J Stat.* (1979) 6:65–70.
47. Martínez-Avilés M, Garrido-Esteba M, Álvarez J, de la Torre A. *Salmonella* surveillance systems in swine and humans in Spain: a review. *Vet Sci.* (2019) 6:20. doi: 10.3390/vetsci6010020
48. Bonardi S. *Salmonella* in the pork production chain and its impact on human health in the European Union. *Epidemiol Infect.* (2017) 145:1513–26. doi: 10.1017/S095026881700036X
49. Hurd HS, McKean JD, Griffith RW, Wesley IV, Rostagno MH. *Salmonella enterica* infections in market swine with and without transport and holding. *Appl Environ Microbiol.* (2002) 68:2376–81. doi: 10.1128/AEM.68.5.2376-2381.2002
50. Sanchez J, Dohoo IR, Christensen J, Rajic A. Factors influencing the prevalence of *Salmonella* spp. in swine farms: A meta-analysis approach. *Prev Vet Med.* (2007) 81:148–77. doi: 10.1016/j.prevetmed.2007.04.005
51. Martín-Peláez S, Peralta B, Creus E, Dalmau A, Velarde A, Pérez JF, et al. Different feed withdrawal times before slaughter influence caecal fermentation and faecal *Salmonella* shedding in pigs. *Vet J Lond Engl 1997.* (2009) 182:469–73. doi: 10.1016/j.tvjl.2008.08.002
52. Rostagno MH, Hurd HS, McKean JD, Ziemer C, Gailey J, Leite R. Abattoir holding pens as a source of *Salmonella* for swine. In: *International Conference on the Epidemiology and Control of Biological, Chemical and Physical Hazards in Pigs and Pork.* Leipzig: Iowa State University, Digital Press. (2001). p. 298–300.
53. Hoffman MD, Gelman A. The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *J Mach Learn Res.* (2014) 15:1593–1623.
54. Goodrich B, Gabry J, Ali I, Brilleman S. *rstanarm: Bayesian Applied Regression Modeling via Stan.* (2018). Available online at: <http://mc-stan.org/>
55. MacNab YC. On Gaussian Markov random fields and Bayesian disease mapping. *Stat Methods Med Res.* (2011) 20:49–68. doi: 10.1177/0962280210371561
56. Heinze G, Wallisch C, Dunkler D. Variable selection - a review and recommendations for the practicing statistician. *Biom J.* (2018) 60:431–49. doi: 10.1002/bimj.201700067
57. Hodges JS, Reich BJ. Adding spatially-correlated errors can mess up the fixed effect you love. *Am Stat.* (2010) 64:325–34. doi: 10.1198/tast.2010.10052

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Teng, Martínez Avilés, Ugarte-Ruiz, Barcena, de la Torre, Lopez, Moreno, Dominguez and Alvarez. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.