

## Original Paper

†The original version of this article was published with the incorrect author name. A notice detailing this has been published and the error rectified in the online PDF and HTML copy. <https://doi.org/10.1017/S0950268818001371>

**Cite this article:** Alarid-Escudero F, Enns EA, MacLehose RF, Parsonnet J, Torres J, Kuntz KM (2018). Force of infection of *Helicobacter pylori* in Mexico: evidence from a national survey using a hierarchical Bayesian model. *Epidemiology and Infection* **146**, 961–969. <https://doi.org/10.1017/S0950268818000857>

Received: 18 December 2017

Revised: 14 March 2018

Accepted: 16 March 2018

First published online: 16 April 2018

**Key words:**

Bayesian hierarchical models; catalytic epidemic models; force of infection; *Helicobacter pylori*; Mexico

**Author for correspondence:**

F. Alarid-Escudero, E-mail: [alarid006@umn.edu](mailto:alarid006@umn.edu)

# Force of infection of *Helicobacter pylori* in Mexico: evidence from a national survey using a hierarchical Bayesian model†

F. Alarid-Escudero<sup>1</sup>, E. A. Enns<sup>1</sup>, R. F. MacLehose<sup>2</sup>, J. Parsonnet<sup>3</sup>, J. Torres<sup>4</sup> and K. M. Kuntz<sup>1</sup>

<sup>1</sup>Division of Health Policy and Management, University of Minnesota School of Public Health, Minneapolis, MN, 55455, USA; <sup>2</sup>Division of Epidemiology and Community Health, University of Minnesota School of Public Health, Minneapolis, MN, 55455, USA; <sup>3</sup>Departments of Medicine and Health Research and Policy, Stanford University School of Medicine, Stanford, CA, 94305-5107, USA and <sup>4</sup>Mexican Social Security Institute, Mexico City, Mexico

**Abstract**

*Helicobacter pylori* (*H. pylori*) is present in the stomach of half of the world's population. The force of infection describes the rate at which susceptibles acquire infection. In this article, we estimated the age-specific force of infection of *H. pylori* in Mexico. Data came from a national *H. pylori* seroepidemiology survey collected in Mexico in 1987–88. We modelled the number of individuals with *H. pylori* at a given age as a binomial random variable. We assumed that the cumulative risk of infection by a given age follows a modified exponential catalytic model, allowing some fraction of the population to remain uninfected. The cumulative risk of infection was modelled for each state in Mexico and were shrunk towards the overall national cumulative risk curve using Bayesian hierarchical models. The proportion of the population that can be infected (i.e. susceptible population) is 85.9% (95% credible interval (CR) 84.3%–87.5%). The constant rate of infection per year of age among the susceptible population is 0.092 (95% CR 0.084–0.100). The estimated force of infection was highest at birth 0.079 (95% CR 0.071–0.087) decreasing to zero as age increases. This Bayesian hierarchical model allows stable estimation of state-specific force of infection by pooling information between the states, resulting in more realistic estimates.

**Introduction**

*Helicobacter pylori* is one of the most prevalent global pathogens, the strongest known biological risk factor for gastric cancer (about a sixfold increase of risk) and is responsible for 80% of gastric ulcers [1]. *H. pylori* is a Gram-negative, microaerophilic bacterium commonly found in the epithelial lining of the human stomach [2].

The mode of transmission of *H. pylori* remains unclear. Transmission appears to occur through close personal contact, such as oral–oral or faecal–oral, particularly within the family and typically in early childhood [1, 3–10]. Once infected with *H. pylori*, individuals will experience life-long infection in the absence of antibiotic treatment [11, 12]. The global burden of disease is substantial, with *H. pylori* present in the stomach of half of the world's population; however, *H. pylori* most heavily burdens low- and middle-income countries (LMICs) where the proportion of people infected is approximately 70% [13, 14]. For example, a prevalence study in Mexico showed that 66% of the Mexican population was infected with *H. pylori*, and that this prevalence increases with age, reaching up to 80% in adults 25 years old and older [15]. This study also found significant variation in the prevalence by age, socio-economic status and geography (e.g. state-level prevalence ranged from 48% to 85%).

The force of infection is the instantaneous rate at which susceptible individuals acquire infection. It is an important epidemiological quantity and a key parameter for mathematical models of disease transmission, which are used to estimate disease burden and the effectiveness and cost-effectiveness of infectious disease treatment and prevention [16, 17]. Like many infectious diseases, it is infeasible to directly measure the force of infection of *H. pylori* [18–20]. However, under certain assumptions, it can be estimated using population-level seroprevalence data [16, 21, 22]. Despite the existence of seroprevalence data of *H. pylori*, it has not been used previously to estimate the force of infection.

In this paper, we estimate the age-specific force of infection of *H. pylori* from national seroprevalence data in Mexico (a middle-income country with a high prevalence of infection). We used data from a nationally representative seroepidemiology survey conducted in Mexico in 1987–1988, prior to widespread antibiotic treatment of *H. pylori*, with rich geographical variation. Given the large heterogeneity in the number of samples obtained in each state, we used a non-linear Bayesian hierarchical catalytic model that allows us to estimate the force of infection in each state in

Mexico while simultaneously estimating the national average force of infection. We demonstrate that this method dramatically stabilises estimation by shrinking the state-specific force of infection towards the national estimate, particularly on states with low data.

**Methods**

*Force of infection as a catalytic model*

The force of infection is the key quantity governing disease transmission within a given population and is defined as the instantaneous per capita rate at which susceptible individuals acquire infection (i.e. the hazard of infection). It reflects both the degree of contact between susceptible and infected individuals and the transmissibility of the pathogen per contact. In cases where contact is age dependent, the force of infection is itself a function of age analogous to the hazard being a function of time in a survival model [21].

Formally, let  $P(a)$  denote the probability that an individual susceptible at birth is still susceptible (i.e. uninfected) at age  $a$ ; this is identical to the cumulative survival function in a survival analysis. Let  $\lambda(a)$  denote the force of infection at age  $a$ , which represents the rate of infection of the susceptible population  $P(a)$ . In terms of a survival model, the force of infection  $\lambda(a)$  represents the instantaneous infection hazard. Therefore, for lifelong infections (or infections conferring lifelong immunity) that do not significantly affect mortality,  $\lambda(a)$  can be defined as:

$$\lambda(a) = -\frac{1}{P(a)} \frac{dP(a)}{da} \tag{1}$$

under the assumption that the population is in dynamic equilibrium (meaning that  $P(a)$  is constant over calendar time and not changing by birth cohort).

Equation (1) specifies a so-called *catalytic epidemic model*, first defined by Muench [22]. The term *catalytic model* derives from its origins in chemistry, where these models were used to study chemical reaction kinetics. Catalytic epidemic models have been widely used to estimate the force of infection of different infectious disease such as measles, rubella, mumps, hepatitis A and yellow fever using seroprevalence epidemiological data from settings where these diseases do not significantly affect mortality [16].

Catalytic models can easily be written in terms of the cumulative probability of infection by age  $a$ ,  $F(a) = 1 - P(a)$ , which more directly corresponds to how seroprevalence data are collected [21, 22]. In terms of  $F(a)$ , the force of infection,  $\lambda(a)$ , becomes:

$$\lambda(a) = \frac{1}{1 - F(a)} \frac{dF(a)}{da} \tag{2}$$

Note that rearranging Equation (2) we can solve for  $F(a)$  as a function of  $\lambda(a)$ , which is equivalent to a survival model where the probability of infection by age  $a$  is the cumulative distribution function of the time to infection:

$$F(a) = 1 - \exp\left[-\int_0^a \lambda(s) ds\right] \tag{3}$$

*Seroprevalence data of H. pylori in Mexico*

In 1987–1988, the National Seroepidemiological Survey was conducted in Mexico as a nationally representative seroprevalence

study of *H. pylori* infection [15, 23]. The survey was designed to represent the country by including all 32 states of Mexico, all ages, all socio-economic levels and both sexes [23]. In total, 32 200 households were surveyed and more than 70 000 serum samples were collected. Of these samples, 11 605 individuals were used for *H. pylori* testing, and out of these, 7 720 (66%) were seropositive for *H. pylori*. *H. pylori* infection was determined by ELISA detection of IgG antibodies to specific *H. pylori* antigens. *H. pylori* prevalence in Mexico varies by state and can be as low as 48% in the state of Colima and as high as 83% in the state of Baja California Sur (Fig. 1).

*Functional form of H. pylori infection in Mexico*

The National Seroepidemiological Survey was collected prior to widespread antibiotic treatment of *H. pylori* in 1991, so we can assume that the population is in dynamic equilibrium. The age-specific prevalence measured nationally is shown in Figure 2 and it is clear that the trend is monotonically non-decreasing.

Based on the shape of these data, we assume that the state-specific cumulative probability of *H. pylori* infection,  $F_i(a)$ , for each state  $i$  follows a modified exponential functional form:

$$F_i(a) = \alpha_i(1 - e^{-\gamma_i a}), \tag{4}$$

where  $0 \leq \alpha_i \leq 1$  and  $\gamma_i \geq 0$  are the state-specific asymptote and rate parameters of the catalytic model. The asymptote parameter  $\alpha$  is often interpreted as the proportion of the population that can be infected (i.e. the susceptible population), if it is assumed that it is the exhaustion of susceptible individuals that leads to the eventual plateau in the proportion infected at older ages [16]. With this interpretation, the rate parameter  $\gamma$  is then the constant rate at which these susceptible individuals become infected. We allow  $\alpha$  and  $\gamma$  to vary by state to reflect the heterogeneity in *H. pylori* prevalence trends resulting from differences in socio-economic and environmental factors. Note that for all states, we assume that no children are infected at birth ( $F(0) = 0$ ) since there is no biological evidence for vertical transmission of *H. pylori* from mother to child during pregnancy, labour or delivery [24].

Based on this functional form for the cumulative probability of infection, we used Equation (3) to solve for the state-specific force of infection,  $\lambda_i(a)$ , in terms of  $\alpha_i$  and  $\gamma_i$ :

$$\lambda_i(a) = \frac{\alpha_i \gamma_i e^{-\gamma_i a}}{1 - \alpha_i(1 - e^{-\gamma_i a})} \tag{5}$$

Note that if the entire population is eventually infected in a given state (i.e.  $\alpha_i = 1$ ),  $\lambda_i(a) = \gamma_i$  and Equation (4) reduces to a traditional exponential model in survival analysis.

*Estimation of the force of H. pylori infection*

We used statistical methods to estimate the parameters  $\alpha$  and  $\gamma$  that best fit the cumulative probability of infection in each state and then calculated the corresponding force of infection. Due to the high degree of variation in sample composition for each state in the National Seroepidemiological Survey, traditional non-linear least-squares estimation yielded imprecise, unstable or unrealistic estimates of  $\alpha$  and  $\gamma$ . This could translate into estimates of force of infection that are implausibly high or low with wide

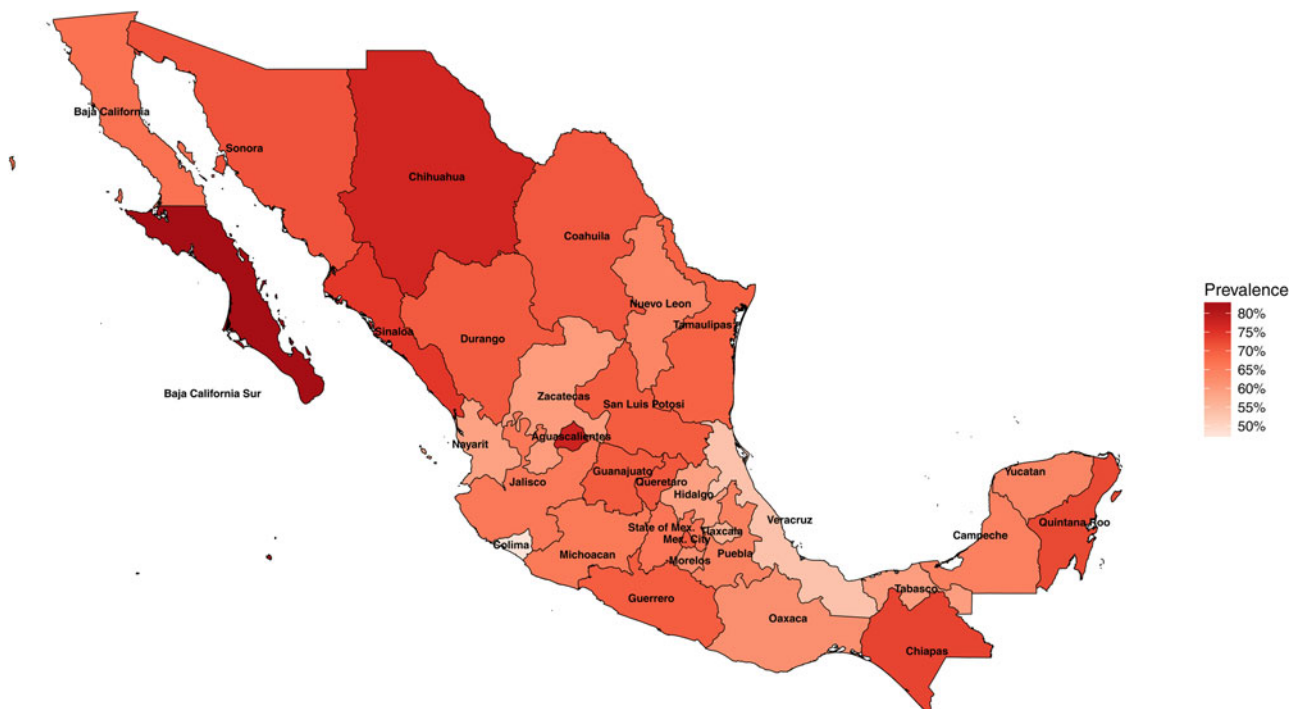


Fig. 1. Map of Mexico with observed prevalence of *Helicobacter pylori* by state in 1987-88.

confidence intervals (CIs) when only a small number of people are sampled (and, for instance, none are infected). For example, some states had individuals sampled at all ages while others had limited data for certain ages. In Figure 3, we show the observed prevalence in four different states with varying amounts of data and the predicted prevalence for these states using the estimates from non-linear least-squares estimation (in blue solid lines). Nuevo Leon and Mexico City have more data and less variability in terms of disease status compared with the states of Baja California Sur and Nayarit. The state of Baja California Sur is a state where all adults sampled older than 20 years old ( $n = 24$ ) are positive for *H. pylori*. The resulting parameter estimates are implausibly high based on our current knowledge of *H. pylori*.

To solve these issues, we instead implemented a hierarchical Bayesian non-linear model that ‘borrows’ information from all states when estimating state-specific effects. Such models have been described in detail elsewhere [25–27] and are known to reduce mean squared error of model estimates in some settings. We are unaware of these models having been applied to catalytic models to estimate force of infection. We first used hierarchical Bayesian non-linear models to estimate the posterior distribution of the  $\alpha$  and  $\gamma$  parameters of model (4) for each of the 32 states. We then estimated the aggregated and state-specific force of infection of *H. pylori* in Mexico.

### Hierarchical non-linear Bayesian model

We assume that for each state  $i = 1, \dots, 32$  the number of infected individuals  $y_i(a)$  at age  $a$  follows a binomial distribution where  $n_i(a)$  is the sample size and  $p_i(a)$  is the cumulative probability of infection (i.e.  $F_i(a)$ ) in state  $i$ :

$$\begin{aligned}
 y_i(a) &\sim \text{Binomial}(n_i(a), p_i(a)) \\
 p_i(a) &= F_i(a) = \alpha_i(1 - e^{-\gamma_i a}) \\
 \begin{bmatrix} \text{logit}(\alpha_i) \\ \text{log}(\gamma_i) \end{bmatrix} &\sim \text{MVN}\left(\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}, \Sigma\right) \\
 \Sigma &\sim \text{IW}\left(\begin{bmatrix} 10^3 & 0 \\ 0 & 10^3 \end{bmatrix}, 2\right)
 \end{aligned} \tag{6}$$

Note that each of the 32 states has its own set of parameters  $\alpha_i$  and  $\gamma_i$ , but these state-specific parameters ‘borrow’ information from each other by assuming that they collectively follow a multivariate normal (MVN) distribution centred at  $\beta_1$  and  $\beta_2$ . To compute the average national-level proportion of susceptibles and infection rate,  $\alpha^0$  and  $\gamma^0$ , we applied a logit and log transformation to the  $\beta_1$  and  $\beta_2$  parameters, respectively.

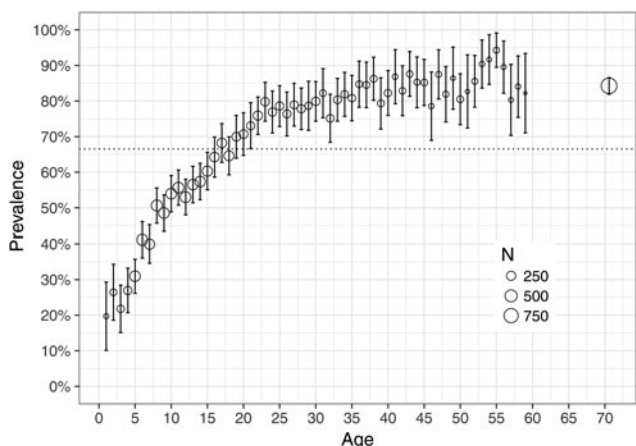
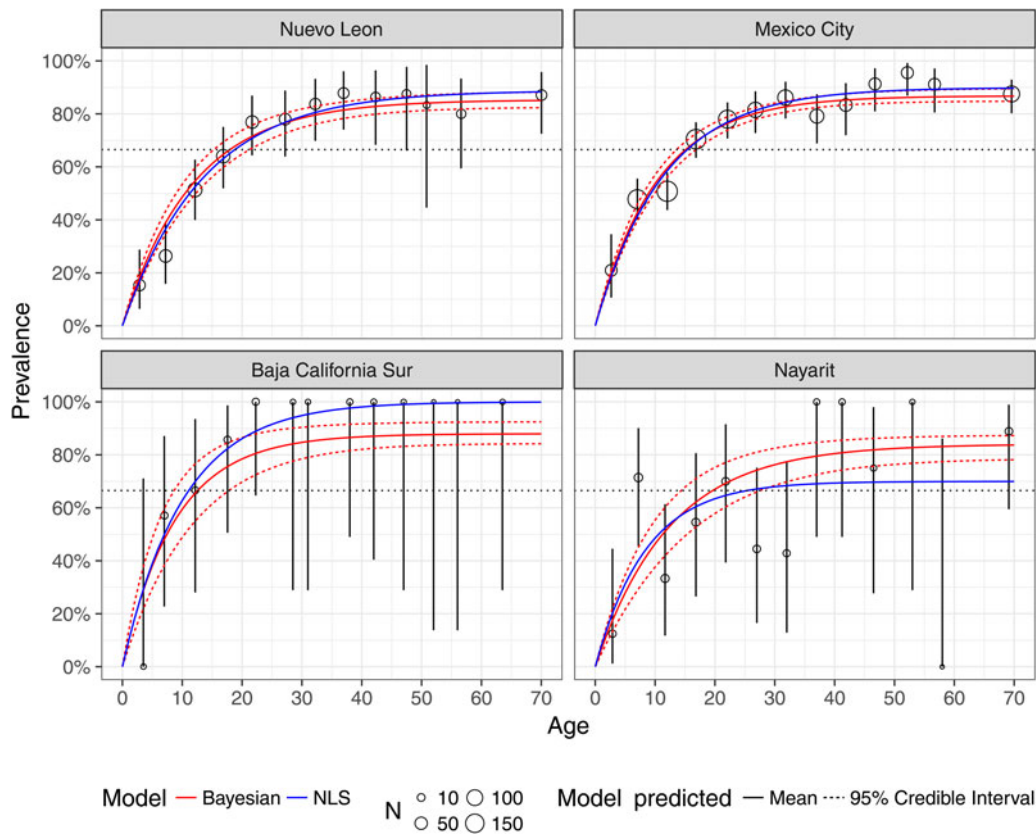


Fig. 2. Empirical prevalence of *Helicobacter pylori* in Mexico by age nationally and in three different states with different amount of data.



**Fig. 3.** Empirical prevalence of *Helicobacter pylori* in Mexico by age in four different states with different amount of data, together with model-predicted prevalence using non-linear least-squares (NLS) and hierarchical Bayesian estimation methods. The upper two reflect states with data for all ages and the lower two reflect states with limited amount of data for some ages. The red solid line denotes the model-predicted posterior mean and the red dashed lines denote the 95% credible bounds. The black dotted horizontal line shows the average national *H. pylori* prevalence.

For the covariance matrix, we used an inverse Wishart (IW) distribution of two degrees of freedom. The IW is a widely used prior distribution for covariance matrices; the large diagonal entries and the zero off-diagonal entries in the matrix represent a weakly informative prior on the variances and no prior correlation between  $\alpha$  and  $\gamma$  parameters [25]. Weakly informative normal prior distributions ( $\mu = 0$  and  $\sigma = 10^3$ ) were specified for  $\beta_1$  and  $\beta_2$  parameters.

Parameters of model (6) were estimated through Gibbs sampling, using Just Another Gibbs Sampler (JAGS) [28]. We ran two Markov chain Monte Carlo (MCMC) chains. For each chain, 50 000 iterations were used to generate a posterior distribution after a 10 000 iterations burn-in period. The mean estimates and 95% credible intervals (CR) were derived from posterior distributions. For each chain, we used starting values computed from the solution of model (4) using non-linear least-squares. We

checked each parameter’s chains to assess convergence. To confirm whether the model provides a reasonable fit of the data, we plotted the observed prevalence by age with 95% CIs, together with the model-predicted prevalence.

The force of *H. pylori* infection was estimated within the MCMC model by evaluating Equation (5) at each parameter set drawn from the posterior distribution, respectively. We then calculated their posterior predicted mean and 95% CR. All statistical modelling was conducted in JAGS [28] and R version 3.2.4 (<http://www.r-project.org>).

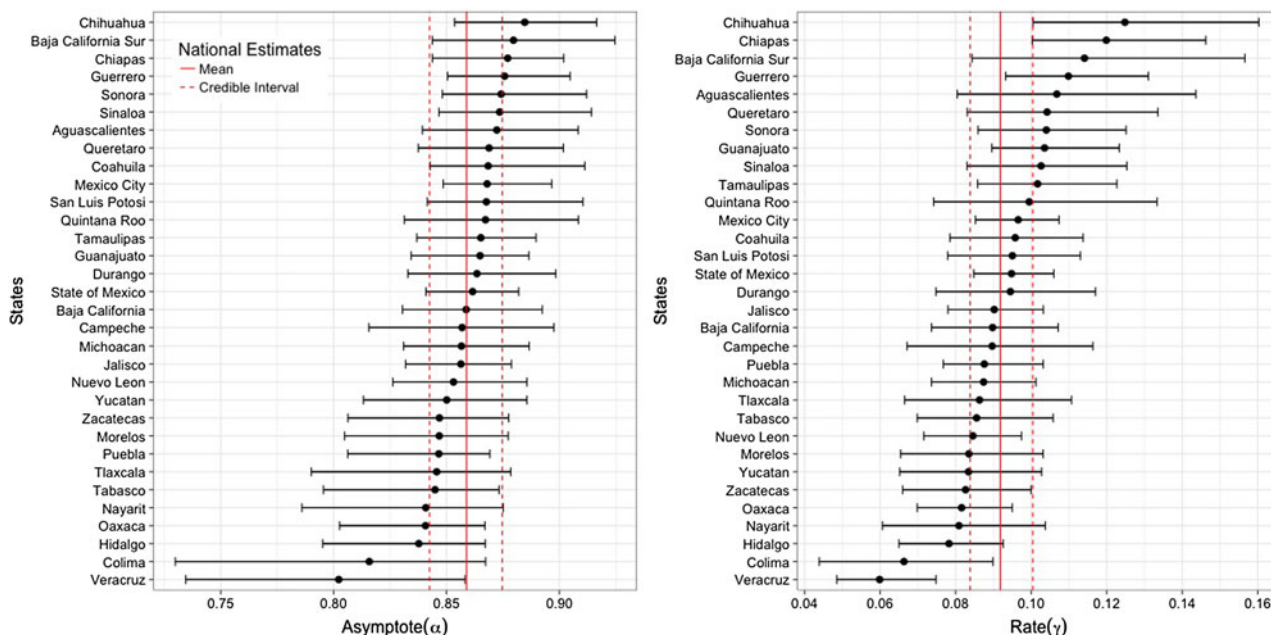
**Results**

Table 1 presents the posterior summaries of the national-level  $\alpha^0$  and  $\gamma^0$  parameters of the hierarchical non-linear Bayesian model. The posterior mean of the asymptote parameter ( $\alpha^0$ ) is 0.859 (95% CR 0.843–0.875). The posterior mean of the rate parameter ( $\gamma^0$ ) is 0.092 (95% CR 0.084–0.100).

Figure 4 shows the state-specific estimates of the asymptote ( $\alpha$ ) and infection rate ( $\gamma$ ) parameters. Parameter estimates exhibit substantial heterogeneity between states. More than a third of the states have posterior means of the asymptote parameter outside the 95% CR of the national average (95% CR 0.843–0.875). Six states (Baja California Sur, Coahuila, Sinaloa, San Luis Potosi, Sonora and Mexico City) have posterior means above the upper bound of the national average, while seven other states (Nayarit, Hidalgo, Tlaxcala, Tabasco, Oaxaca, Puebla and

**Table 1.** Posterior mean estimates, standard deviation (SD) and lower and upper bounds of the 95% credible interval of the national-level proportion of susceptibles ( $\alpha^0$ ) and rate of infection ( $\gamma^0$ ), which correspond to the asymptote and rate parameters of the catalytic epidemic model of *Helicobacter pylori* in Mexico

Parameter	Estimate	S.D.	LB	UB
$\alpha^0$	0.859	0.008	0.843	0.875
$\gamma^0$	0.092	0.004	0.084	0.100

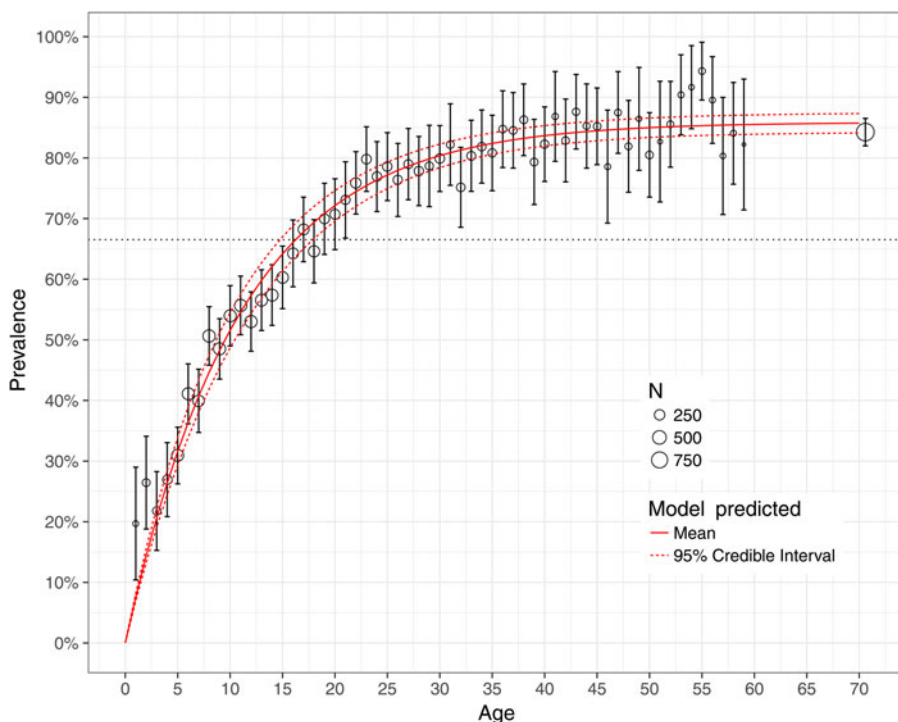


**Fig. 4.** State-specific posterior mean and 95% credible interval for the asymptote and rate parameters. Solid and dashed vertical lines represent the national-level posterior mean and 95% credible interval, respectively.

Veracruz) have estimates below the lower bound of the national average. More than half of the states have posterior means of the rate parameter outside the 95% CR of the estimated national average (95% CR 0.084–0.100). Ten states (Chihuahua, Chiapas, Baja California Sur, Guerrero, Aguascalientes, Queretaro, Sonora, Guanajuato, Sinaloa and Tamaulipas) have posterior means above the upper bound of the national average and eight other states (Morelos, Yucatan, Zacatecas, Oaxaca, Nayarit,

Hidalgo, Colima and Veracruz) have lower estimates than the lower bound of the national average.

Figure 5 shows the observed national prevalence of *H. pylori* by age in Mexico, together with the model-predicted prevalence. All but three model-predicted prevalence lie between the 95% confidence bounds of the empirical data. For 1- and 2-year olds, the model-predicted prevalence lies below the bounds, favouring lower prevalence than those observed for these ages. Figure 3



**Fig. 5.** Model fit to the national prevalence of *Helicobacter pylori* by age in Mexico in 1987–88. Grey circles denote empirical prevalence with size proportional to sample size, and 95% confidence interval per 1-year age group. The red solid line denotes the model-predicted posterior mean and the red dashed lines denote the 95% credible bounds.

shows the observed and model-predicted prevalence of *H. pylori* by age in four different states with varying amounts of data using both the hierarchical Bayesian and the non-linear least-squares estimation methods. The width of the model-predicted credible bounds is a function of the amount of data used to estimate the state-specific prevalence. States for which more data are available have tighter credible bounds than states with limited data. The predicted prevalence from the non-linear least-squares model are sensitive to small sample sizes overestimating the prevalence on ages for which data are skewed towards infection (state of Baja California Sur) and underestimating the prevalence on ages for which data are skewed towards no infection (state of Nayarit). The hierarchical Bayesian model shrinks the estimated prevalence for states with limited data towards the overall mean. The observed and model-predicted prevalence using both the hierarchical Bayesian and the non-linear least-squares estimation methods for all the geographical states in Mexico is shown in Figure 6.

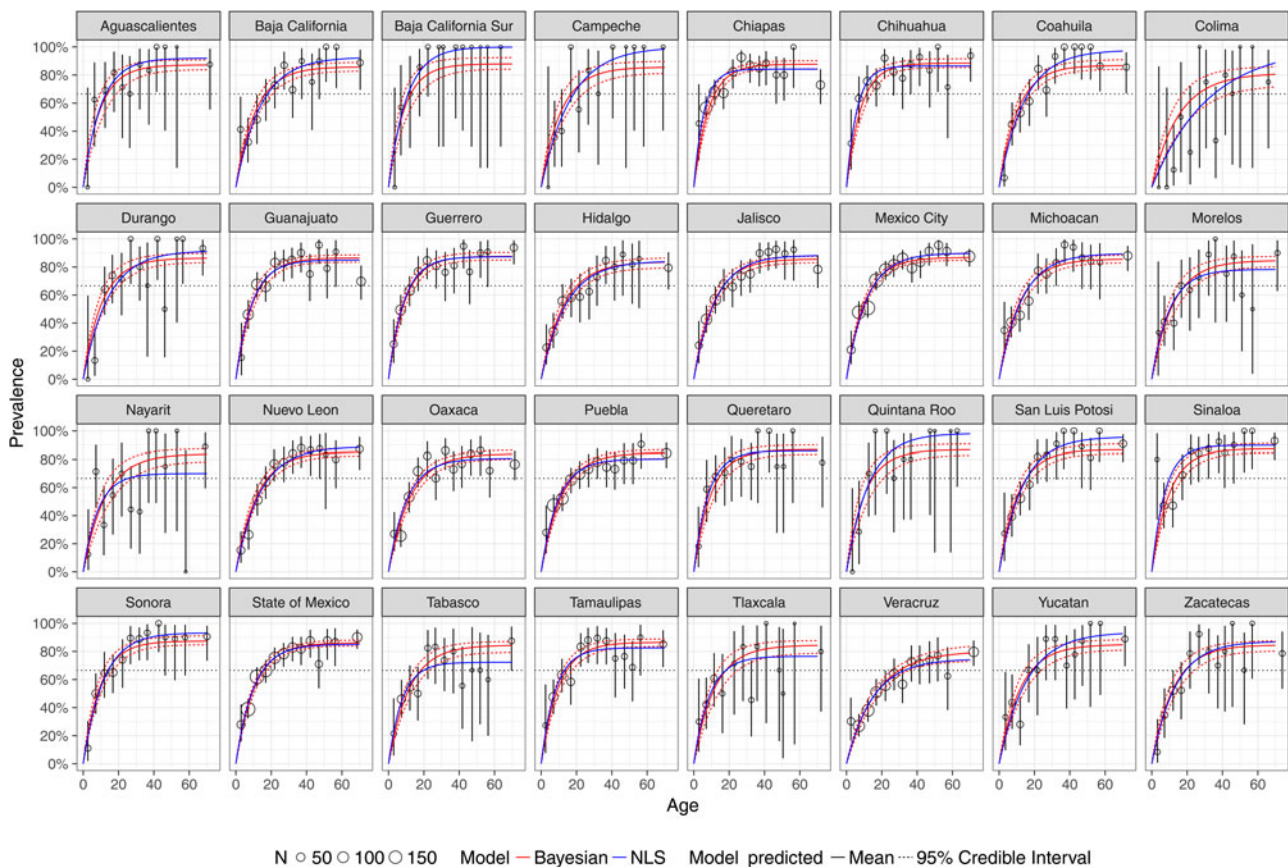
**Force of infection**

Figure 7 shows the model-predicted national force of infection of *H. pylori* by age in Mexico. The force of infection starts at 0.08 right after birth and decreases to zero as age increases. Like  $\alpha$  and  $\gamma$ , the state-specific force of infection varies considerably across states (Fig. 8). At birth, estimates of the force of infection were as high as 0.10 (Chiapas and Chihuahua) and as low as 0.04 (Colima and Veracruz).

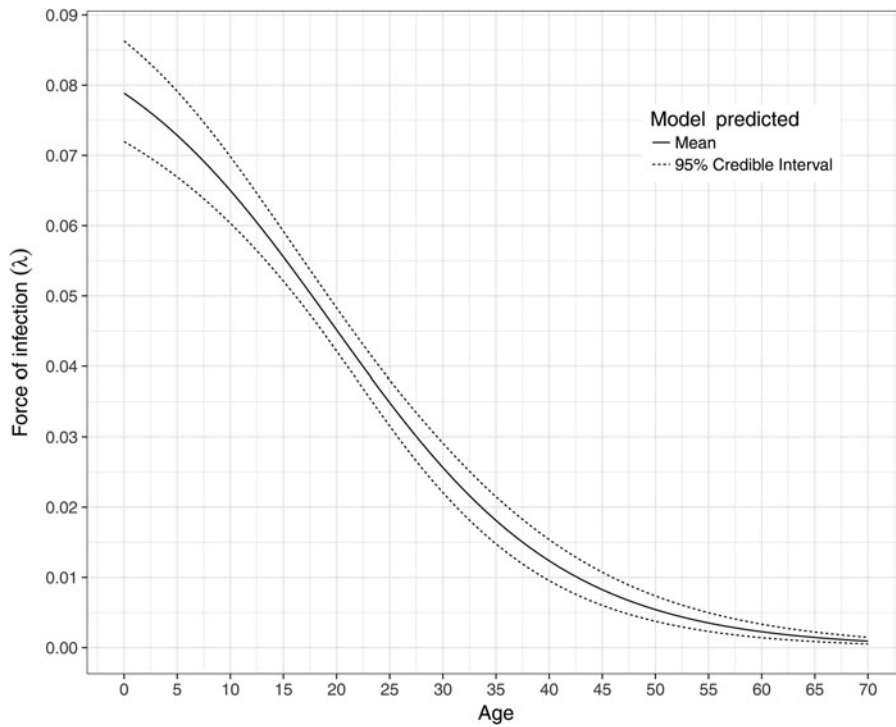
**Discussion**

In this study, we used a catalytic model to estimate the force of infection of *H. pylori* in Mexico at the national level, but also at the state level. Catalytic models have been previously used to estimate the force of infection of other diseases but have not been previously applied to *H. pylori*. We found that age-specific prevalence of *H. pylori* varies greatly by state. The hierarchical Bayesian approach used to estimate the catalytic model of *H. pylori* allowed us to account for state-specific heterogeneity in the force of infection. Specifically, by borrowing information between states, we obtained better inference for undersampled states and quantified the variation across states. In addition, we found great variation in the proportion of the population that can be infected (i.e. the asymptote parameter  $\alpha$ ) and the constant rate of infection among the susceptible population (i.e. the rate parameter  $\gamma$ ). These contrasting differences translate into high variation in the force of infection, which can have implications in the prevention and treatment strategies and highlights the need of state-specific interventions.

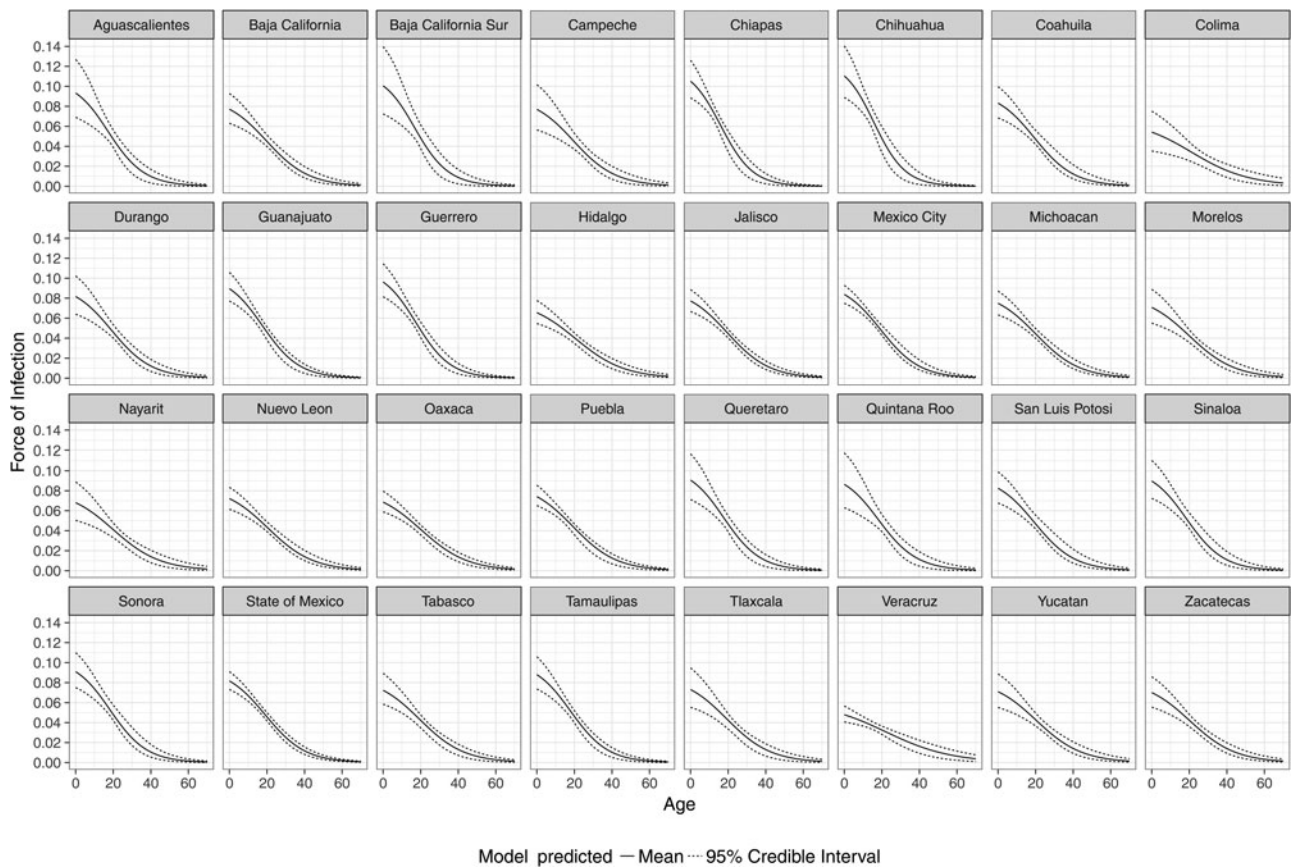
The model-predicted force of infection is comparable with the results seen in previous research. For example, the estimated force of infection at 8 years old is 7% (95% CR 6.4%–7.5%) and at 11 years old is 6.4% (95% CR 6.0%–6.9%), which are similar estimates to the annual incidence rates of 6.56% (95% CI 4.6%–8.53%) in 5–8 years old and 6.01% (95% CI 3.52%–8.5%) in 9–13 years old obtained by observing *H. pylori* incidence over time in Mexican school children who initially tested negative



**Fig. 6.** Model fit to the state-specific prevalence of *Helicobacter pylori* by age in Mexico in 1987–88. Grey circles denote empirical prevalence with size proportional to sample size and 95% confidence interval. The red solid line denotes the model-predicted posterior mean and the red dashed lines denote the 95% credible bounds. The black dotted horizontal line shows the average national *H. pylori* prevalence.



**Fig. 7.** Model-predicted national force of infection of *Helicobacter pylori* by age in Mexico in 1987–88. The solid line denotes the model-predicted posterior mean and the dashed lines denote the 95% credible bounds.



**Fig. 8.** Model-predicted state-specific force of infection of *Helicobacter pylori* by age in Mexico in 1987–88. The solid line denotes the model-predicted posterior mean and the dashed lines denote the 95% credible bounds.

for the infection [29]. Another study estimated an annual infection rate of 19.84% (95% CI 16.43–23.25%) from birth to 2 years old in a cohort of children in El Paso, Texas, and Ciudad Juarez (Chihuahua), which is consistent with our assumption that the force of infection is highest at birth and decreases with age [30].

Our study has several limitations. We imposed a constrained exponential functional form on the cumulative probability of infection, which may have overly restricted our parameter estimates and the predicted force of infection. For example, the model-predicted *H. pylori* prevalence consistently falls below observed prevalence for children under 2 years old, nationally and for many states, which may indicate that the force of infection at these ages is being underestimated. A more flexible catalytic model, such as a piecewise or a spline model, could have provided a better fit to the data [31]. However, by using these more flexible models, we lose the ability to link the model parameters directly to the disease process. We also assumed that the Mexican population is in endemic equilibrium with respect to *H. pylori* infection, which means that the force of infection does not change over time. While antibiotic treatment for *H. pylori* was not widespread prior to 1991 [32, 33], this assumption could still be violated if other factors related to *H. pylori* transmission, such as improved sanitation, have changed prior to 1987 (the year when the National Seroepidemiological Survey started collection). However, if new seroepidemiological data were collected in Mexico, we could apply our model to assess if the force of infection of *H. pylori* has changed over time because of antibiotic treatment, cohort effects or changes in other socio-demographic variables. In addition, we assumed that disease-related mortality is negligible compared with all-cause mortality. This would seem reasonable in the case of *H. pylori*, but because it is a carcinogenic agent, there is a small fraction of the infected population that will develop gastric cancer and face an elevated mortality risk. However, out of the total number of deaths in Mexico between 1987 and 1988 (819 900 deaths), 8143 deaths were attributed to gastric cancer, which only accounted for <1% of all deaths in Mexico in that period [34]. Since the proportion of the population that can be infected and the constant rate at which they become infected vary by state, such variation could reflect some degree of spatial autocorrelation. However, in an exploratory evaluation of spatial autocorrelation using the global Moran's I test [35, 36], we estimated a Moran's I statistic of 0.06 ( $P = 0.18$ ), suggesting little spatial autocorrelation. Furthermore, state-specific risk factors (such as level of sanitation, education, etc.) might be stronger predictors of these state-level variations than the spatial configuration of the states. However, we plan to address these limitations in future work. Despite these limitations, we believe that our estimates of the force of infection are reasonable interpretations of the national seroepidemiological data.

Although we found and quantified the high variation across states, we did not explain potential sources for this variation. Some of the state-level explanatory variables might include level of socio-economic development, sanitation, education and percentage of population living in rural areas. These factors could be used to model both the asymptote ( $\alpha$ ) and the rate of infection ( $\gamma$ ) parameters. Mexico has high socio-demographic variation in its states, so the estimated influence of state-level characteristics on the force of infection might be relevant to other Latin American countries. This, however, is a topic of further research.

While our analysis relied on data specifically from Mexico, there are many other LMICs that face a similar burden of *H. pylori* and

gastric cancer [37, 38]. Many of these countries are in Latin America and likely have similar transmission dynamics [1].

Serological data for other types of infections have been used to estimate the parameters of dynamic transmission models [39, 40] that are then used to estimate the effectiveness and cost-effectiveness of different treatment or vaccination strategies [41, 42]. The methodology and results of our study could be used to estimate the transmission parameters for dynamic transmission models of *H. pylori* to evaluate different screen-and-treat strategies or estimate the benefits of potential vaccines, which are currently under development [43]. Identifying populations at greatest risk of *H. pylori* infection will further the development of appropriately targeted prevention, screening and treatment strategies.

**Acknowledgements.** Financial support for this study was provided in part by a grant from the National Council of Science and Technology of Mexico (CONACYT) and a Doctoral Dissertation Fellowship from the Graduate School of the University of Minnesota as part of Dr Alarid-Escudero's doctoral program. Dr Enns was supported by a grant from the National Institute of Allergy and Infectious Disease of the National Institutes of Health under award no. K25AI118476. The funding agencies had no role in the design of the study, interpretation of results or writing of the manuscript. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing and publishing the report.

**Declaration of Interest.** None.

**Disclaimers.** None.

## References

1. **International Agency for Research on Cancer (IARC)** (2014) *Helicobacter pylori* Eradication as a Strategy for Preventing Gastric Cancer. Lyon, France: WHO Press.
2. **Centers for Disease Control and Prevention** (1998) *Helicobacter pylori: Fact Sheet for Health Care Providers*.
3. **Goodman KJ, et al.** (1996) *Helicobacter pylori* infection in the Colombian Andes: a population-based study of transmission pathways. *American Journal of Epidemiology* **144**, 290–299.
4. **Neale KR and Logan RP** (1995) The epidemiology and transmission of *Helicobacter pylori* infection in children. *Alimentary Pharmacology & Therapeutics* **9**, 77–84.
5. **Koletzko S and Mégraud F.** (2016) *Helicobacter pylori* infection in children. In Backert S, Yamaoka Y (eds). *Helicobacter pylori* Research: From Bench to Bedside. Japan: Springer, pp. 443–467.
6. **Kim N.** (2016) Prevalence and transmission routes of *H. pylori*. In Kim N (ed.). *Helicobacter pylori*. Singapore: Springer, pp. 3–19.
7. **Eusebi LH, Zagari RM and Bazzoli F** (2014) Epidemiology of *Helicobacter pylori* infection. *Helicobacter* **19**, 1–5.
8. **Perry S, et al.** (2006) Gastroenteritis and transmission of *Helicobacter pylori* infection in households. *Emerging Infectious Diseases* **12**, 1701–1708.
9. **Rowland M, et al.** (2006) Age-specific incidence of *Helicobacter pylori*. *Gastroenterology* **130**, 65–72.
10. **Malaty HM, et al.** (2002) Age at acquisition of *Helicobacter pylori* infection: a follow-up study from infancy to adulthood. *Lancet* **359**, 931–935.
11. **Kim SS, et al.** (2011) *Helicobacter pylori* in the pathogenesis of gastric cancer and gastric lymphoma. *Cancer Letters Elsevier Ireland Ltd* **305**, 228–238.
12. **Wroblewski LE, Peek RM and Wilson KT** (2010) *Helicobacter pylori* and gastric cancer: factors that modulate disease risk. *Clinical Microbiology Reviews* **23**, 713–739.
13. **Goss PE, et al.** (2013) Planning cancer control in Latin America and the Caribbean. *The Lancet Oncology* **14**, 391–436.
14. **Torres J, et al.** (2016) Population-based strategies for *Helicobacter pylori*-associated disease management: Latin American perspective. In Backert S, Yamaoka Y (eds). *Helicobacter pylori* Research: From Bench to Bedside. Japan: Springer, pp. 503–517.



15. **Torres J, et al.** (1998) A community-based seroepidemiologic study of *Helicobacter pylori* infection in Mexico. *The Journal of Infectious Diseases* **178**, 1089–1094.
16. **Hens N, et al.** (2010) Seventy-five years of estimating the force of infection from current status data. *Epidemiology and Infection* **138**, 802–812.
17. **Brisson M, Edmunds WJ and Gay NJ** (2003) Varicella vaccination: impact of vaccine efficacy on the epidemiology of VZV. *Journal of Medical Virology* **70**, 31–37.
18. **Anderson RM and May R.** (1991) *Infectious Diseases of Humans: Dynamics and Control*. Oxford, UK: Oxford University Press.
19. **Parsonnet J.** (2003) What is the *Helicobacter pylori* global reinfection rate? *Canadian Journal of Gastroenterology* **17**(Suppl. B), 46–49.
20. **Rupnow MFT, et al.** (2000) A dynamic transmission model for predicting trends in *Helicobacter pylori* and associated diseases in the United States. *Emerging Infectious Diseases* **6**, 228–237.
21. **Farrington CP** (1990) Modelling forces of infection for measles, mumps and rubella. *Statistics in Medicine* **9**, 953–967.
22. **Muench H.** (1959) *Catalytic Models in Epidemiology*. Cambridge, MA: Harvard University Press.
23. **Tapia R, Gutiérrez G and Sepúlveda J** (1992) Metodología de la encuesta nacional seroepidemiológica. *Salud Pública de México* **34**, 124–135.
24. **Espinoza Monje M and Marin Vega A.** (2016) *Helicobacter pylori* and perinatal pathologies: pathogen transmission during childbirth? *Biology and Medicine* **8**, Published online: 2016. doi: 10.4172/0974-8369.1000313.
25. **Gelman A and Hill J.** (2006) *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York, NY: Cambridge University Press.
26. **Greenland S** (2000) Principles of multilevel modelling. *International Journal of Epidemiology* **29**, 158–167.
27. **Witte JS, et al.** (1994) Hierarchical regression analysis applied to a study of multiple dietary exposures and breast cancer. *Epidemiology* **5**, 612–621.
28. **Plummer M.** (2003) JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. In Hornik K, Leish F, Zeileis A (eds). *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*. Vienna, Austria: Achim Zeileis.
29. **Duque X, et al.** (2012) Natural history of *Helicobacter pylori* infection in Mexican schoolchildren: incidence and spontaneous clearance. *Journal of Pediatric Gastroenterology and Nutrition* **55**, 209–216.
30. **Goodman KJ, et al.** (2005) Dynamics of *Helicobacter pylori* infection in a US-Mexico cohort during the first two years of life. *International Journal of Epidemiology* **34**, 1348–1355.
31. **Hens N, et al.** (2012) *Modeling Infectious Disease Parameters Based on Serological and Social Contact Data: A Modern Statistical Perspective*. New York, NY: Springer.
32. **González-Piñera JG, et al.** (1998) Macrólidos. *Acta Medica* **8**, 71–74.
33. **Sacre Hazouri JA** (2006) Artículo de revisión Macrólidos. Efectos anti-inflamatorios e inmunomoduladores. Indicación en enfermedades respiratorias. *Revista Alergia México* **53**, 108–122.
34. **Dirección de Información Epidemiológica** (2011) *Perfil epidemiológico de los tumores malignos en México*. Mexico City, Mexico: Secretaría de Salud.
35. **Moran PAP** (1948) The interpretation of statistical maps. *Journal of the Royal Statistical Society. Series B (Methodological)* **10**, 243–251.
36. **Bivand RS, Pebesma EJ and Gómez-Rubio V.** (2013) *Applied Spatial Data Analysis with R*. In Gentleman R, Hornik K & Parmigiani G. G, Eds., *Use R* (2nd ed., Vol. 1). New York, NY: Springer. doi:10.1007/978-0-387-78171-6.
37. **Peleteiro B, et al.** (2014) Prevalence of *Helicobacter pylori* infection worldwide: a systematic review of studies with national coverage. *Digestive Diseases and Sciences* **59**, 1698–1709.
38. **Pounder RE and Ng D** (1994) The prevalence of *Helicobacter pylori* infection in different countries. *Alimentary Pharmacology & Therapeutics* **9**, 33–39.
39. **Metcalfe CJE, et al.** (2012) Structured models of infectious disease: inference with discrete data. *Theoretical Population Biology Elsevier Inc.* **82**, 275–282.
40. **Brisson M, et al.** (2001) Epidemiology of varicella zoster virus infection in Canada and the United Kingdom. *Epidemiology and Infection* **127**, 305–314.
41. **Brisson M and Edmunds WJ** (2002) The cost-effectiveness of varicella vaccination in Canada. *Vaccine* **20**, 1113–1125.
42. **Edmunds WJ, Brisson M and Rose JD** (2001) The epidemiology of herpes zoster and potential cost-effectiveness of vaccination in England and Wales. *Vaccine* **19**, 3076–3090.
43. **Blanchard TG and Czinn SJ.** (2016) Identification of *Helicobacter pylori* and the evolution of an efficacious childhood vaccine to protect against gastritis and peptic ulcer disease. *Pediatric Research* **81**, 170–176, Published online: 2016. doi: 10.1038/pr.2016.199.