The spread of Zika virus in the Americas: SI Appendix

Qian Zhang[1], Kaiyuan Sun[1], Matteo Chinazzi [1], Ana Pastore y Piontti [1], Natalie E. Dean [2], Diana Patricia Rojas [3], Stefano Merler [4], Dina Mistry [1], Piero Poletti [5], Luca Rossi [6], Margaret Bray[1], M. Elizabeth Halloran [7,8], Ira M. Longini Jr.[2], Alessandro Vespignani[1,6]

[1]Laboratory for the Modeling of Biological and Socio-technical Systems, Northeastern University, Boston, MA 02115, USA
[2]Department of Biostatistics, College of Public Health and Health Professions, University of Florida, Gainesville, FL 32611, USA
[3]Department of Epidemiology, College of Public Health and Health Professions, University of Florida, Gainesville, FL 32611, USA
[4]Bruno Kessler Foundation, Trento, Italy
[5]Bocconi University, Milan, Italy
[6]Institute for Scientific Interchange Foundation, Turin, Italy
[7]Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA
[8]Department of Biostatistics, University of Washington, Seattle, WA 98195, USA

Contents

1 ZIKV Transmission Dynamics

The human-vector chain-binomial model is based on an SEIR compartmentalization of human populations and SEI compartmentalization of vector populations [1].

Humans can be in one of four compartments: susceptible individuals (S^H) who lack immunity against infection, exposed individuals (E^H) who have acquired infection but are not yet infectious, infected individuals (I^H) who can transmit infection (and may or may not display symptoms), and removed individuals (R^H) who are no longer infected. People in the final compartment may recover and gain immunity or die. As we are considering a timescale of a few years that is relatively short when compared to human demographic dynamics, we treat the total human population size as constant, i.e. $S^H + E^H + I^H + R^H = N^H$.

The transition of people between compartments is performed stochastically, based on various biological factors. Following Ref. [2], susceptible humans transition to the exposed compartment under the force of infection (λ^H) which is proportional to the rate at which a particular human is bitten by the infected mosquitoes (I^V/N^V), the parameter β that accounts for the daily mosquito biting rate and the specific transmissibility of ZIKV, and the temperature dependence of the mosquito-to-human probability of transmission ($\mathcal{T}_{VH}$). By considering the factor k expressing the number of mosquitoes per person we have $N^H = N^V/k$, yielding a force of infection $\lambda^H = k\beta\mathcal{T}_{VH}\frac{I_t^V}{N^V}$. On average individuals stay in the exposed or infectious state for the duration of the mean intrinsic latent period ϵ_H^{-1} and the mean infectious period μ_H^{-1}, respectively.

The vector population is divided into three compartments: susceptible (S^V), exposed (E^V), and infectious (I^V), respectively. The force of infection (λ^V) governing the transition rate from susceptible to exposed individuals among the vector population is proportional to the density of infectious humans (I^H/N^H). On average mosquitoes are in the exposed state the mean extrinsic latent period ϵ_V^{-1}. The average lifetime of mosquitoes in each compartment μ_V^{-1} varies across spatial locations and time of the year [3] as discussed in the next section. The overall mosquito population is rescaled every day as shown in Sec. 2, and it is considered in equilibrium during the daily integration step so that mosquito deaths are replaced by an equal number of new susceptible mosquitoes. Similar to the force of infection from vector to human, the force of infection from human to the vector, λ^V, is a function of β, the temperature dependence of human-to-mosquito transmission ($\mathcal{T}_{HV}$), and the density of infectious humans ($\frac{I_t^H}{N^H}$). We thus have $\lambda^V = \beta\mathcal{T}_{HV}\frac{I_t^H}{N^H}$.

The coupled population equations describing the epidemic time evolution read as:

$$
\begin{aligned}
S_{t+1}^H &= S_t^H - \Delta_{S^H\to E^H} && (1)\\
E_{t+1}^H &= E_t^H + \Delta_{S^H\to E^H} - \Delta_{E^H\to I^H} && (2)\\
I_{t+1}^H &= I_t^H + \Delta_{E^H\to I^H} - \Delta_{I^H\to R^H} && (3)\\
R_{t+1}^H &= R_t^H + \Delta_{I^H\to R^H}, && (4)
\end{aligned}
$$

and

$$S^V_{t+1} = S^V_t - \Delta_{S^V \to E^V} + \Delta_{I^V \to S^V} + \Delta_{E^V \to S^V} \quad (5)$$

$$E^V_{t+1} = E^V_t - \Delta_{E^V \to S^V} - \Delta_{E^V \to I^V} + \Delta_{S^V \to E^V} \quad (6)$$

$$I^V_{t+1} = I^V_t + \Delta_{E^V \to I^V} - \Delta_{I^V \to S^V}. \quad (7)$$

In the above expressions each term $\Delta_{X \to Y}$ represent the number of human or vector individuals transitioning from state X to state Y. Transitions are calculated according to chain binomial processes $\Delta_{X \to Y} = Binomial(X, p_{X \to Y})$, and $p_{X \to Y}$ are transition probabilities determined by the force of infection and the average lifetime of individuals in each compartment. We assume memoryless discrete stochastic transition processes. It is worth stressing that the terms $\Delta_{I^V \to S^V}$, $\Delta_{E^V \to S^V}$ are introduced to model the replenishment of mosquitoes after death.

By using the standard approach of Ref. [2], the basic reproduction number can be expressed as:

$$R_0 = \frac{\epsilon_V}{(\epsilon_V + \mu_V)(\mu_V \mu_H)} k \beta^2 \mathcal{T}_{VH} \mathcal{T}_{HV}. \quad (8)$$

It is worth remarking that the basic reproduction number varies in each location according to the temperature and mosquitoes abundance.

1.1 Temperature/Seasonal Dependent Parameters

- *Mosquito Lifespan:* We base our mosquito lifespan and corresponding mortality rate on temperature. The relationship between mortality rate and temperature is polynomial, taking the form [3]:

  $$\mu_V(T) = 0.3967 - 0.03912T + 2.442 \times 10^{-3} T^2 - 7.479 \times 10^{-5} T^3 + 9.298 \times 10^{-7} T^4. \quad (9)$$

  Considering temperature ranges from 0° C to 40° C, the resulting range of average lifespans goes from just under 1 *day* up to 7.2 *days*. The corresponding minimum and maximum daily mortality rates for mosquitoes are 1 $days^{-1}$ and 0.1389 $days^{-1}$, respectively.

- *Temperature dependence of the transmission probability per bite:* Both the mosquito-to-human and human-to-mosquito probabilities of transmission are temperature dependent for DENV [4]. We thus assume that also for ZIKV the mosquito-to-human transmission probability sharply declines to zero at $T = 28°$C. When $\mathcal{T}_{VH}(T)$ is close to zero, $\mathcal{T}_{HV}(T)$ becomes less relevant. The virus will not continue to circulate if the mosquitoes can no longer infect humans, even if the reverse transmission probability is one. Therefore, for simplicity, we consider $\mathcal{T}_{HV}(T) = \mathcal{T}_{VH}(T)$ and use the expression for $\mathcal{T}_{VH}$ to describe both:

  $$\mathcal{T}_{VH} = 0.001044T(T - 12.286)\sqrt{32.461 - T}. \quad (10)$$

  We also note that in principle the number of bites per day is not constant. We have found reports for Puerto Rico [5], showing a non-statistically relevant association, while there is a mild dependence in Thailand [5]. In addition, the number of blood meals per day seems to be constant across different seasonal cycles. Given our focus on the Americas we decided to assume the results from Puerto Rico. Furthermore blood meal variations appears to be a relatively minor contribution to the many temperature dependent factors affecting the behavior of the model [6].

- *Seasonal variation of mosquito abundance:* For areas with distinct seasonality, the vector abundance may vary significantly from season to season due to temperature, vector life cycle, etc. In the following, we consider a temporal modulation function $k_b(t)$ to describe the relative abundance modulation throughout the year in each subpopulation b considered in the model, as detailed in Sec. 2.

1.2 Spatiotemporal Dependency and Seasonality

The different values of the parameters and mosquitoes per person in each subpopulation considered in the model yield a functional dependence of the basic reproduction number, $R_{0,b}(t)$, in each subpopulation b at time t that reads as:

$$R_{0,b}(t) = \frac{\epsilon_V}{(\epsilon_V + \mu_V(T_{b,t}))\ (\mu_V(T_{b,t})\ \mu_H)} k_b(T_{b,t})\ \beta^2\ \mathcal{T}_{VH}(T_{b,t})\ \mathcal{T}_{HV}(T_{b,t}), \tag{11}$$

where $T_{b,t}$ is the average temperature in subpopulation b at time t. The variable $R_{0,b}(t)$ has distinct temporal and geographical variations as shown in **Fig. 3** of the main article. Therefore, the seasonal and local drivers have the potential to shape both the timing and the magnitude of ZIKV outbreaks.

1.3 MCMC Calibration for ZIKV Transmissibility and Sensitivity Analysis

The calibration of the model is performed using surveillance data from the 2013 ZIKV outbreak in French Polynesia. The dataset is based on weekly situation reports from the Centre d'Hygiene et de Salubrité Publique [1, 7, 8]. The reported number of new weekly suspected ZIKV cases is available for each of the six main regions of French Polynesia: Tahiti, Sous-le-vent, Moorea, Tuamotu, Marquises, and Australes. However, since there are no evident temporal separations between the outbreaks of different regions, the regional data is aggregated to obtain the overall weekly reported number of new suspected ZIKV cases in French Polynesia. We consider a deterministic version of the model, reported in Sec 1 with the same notation.

There are 8 parameters in the infection dynamic model:

- Intrinsic latent period $1/\epsilon_H$,
- Extrinsic latent period $1/\epsilon_V$,
- Human infectious period $1/\mu_H$,
- Mosquito life span $1/\mu_V$,
- Number of mosquitoes per person k,
- ZIKV transmissibility β,
- Human-to-mosquito temperature dependence of the transmissibility $\mathcal{T}_{HV}$, and
- Mosquito-to-human temperature dependence of the transmissibility $\mathcal{T}_{VH}$.

Unfortunately many parameters characterizing ZIKV are surrounded by uncertainty. We set the number of mosquitoes per person as a constant $k = k_{FP}$. A sensitivity analysis using different values of k_{FP} in the range of $[1 - 3]$ has been performed. It must be noted that, all other parameters being equal, variations in k are absorbed by a rescaling of the parameter β. The mosquito life span is temperature-dependent and using equations in Sec. 1.1, along with the typical temperature during 2013 French Polynesia outbreak, we estimated that $1/\mu_V = 7.16$ days.

The parameters $1/\epsilon_H$, $1/\epsilon_V$, $1/\mu_H$, and $1/\mu_V$ define the serial interval of the infection. We have considered different parameter sets that define one reference scenario along with short and long serial interval scenarios which explore the range of parameters reported in the literature. The values of the parameters are reported in Fig. 7b) of the main text. Assuming $\mathcal{T}_{HV} \simeq \mathcal{T}_{VH}$ and utilizing the fact that β and $\mathcal{T}_{HV}(\mathcal{T}_{VH})$ always appear together on both sides of the force of infection, β and $\mathcal{T}_{HV}$ are calibrated together into the overall transmissibility $\tilde{\beta} = \beta\mathcal{T}_{HV}$.

The initial conditions at time $t = 0$ for the number of exposed humans $E^H_{t_0}$, the number of exposed mosquitoes $E^V_{t_0}$, the infected humans $I^H_{t_0}$, and the infected mosquitoes $I^V_{t_0}$ allow us to numerically solve the infection dynamics. The cumulative number of infections C_t can thus be obtained as: $dC_t = \epsilon_H E^H_t dt$. Thus the weekly new incidence c_t is given by $c_t = C_t - C_{t-1}$.

Here we use a negative binomial measurement model [9, 1] with mean ρc_t and variance $\rho c_t(r + \rho c_t)/r$; ρ is the reporting rate, defined as the proportion of infections (symptomatic and asymptomatic) that gets reported as clinical cases; r is the dispersion parameter of the negative binomial distribution used to fit the data. To narrow the parameter space even more, we assume $E^H_{t_0} = I^H_{t_0}$ and $E^V_{t_0} = I^V_{t_0}$. For each scenario, we are left with a total of five unknown parameters that require calibration:

- overall transmission rate $\tilde{\beta}$,
- initial number of infected humans $I^H_{t_0}$,
- initial number of infected mosquitoes $I^V_{t_0}$,
- reporting rate ρ, and
- dispersion parameter r.

A random walk Metropolis-Hastings Markov Chain Monte Carlo (MCMC) algorithm is performed to calibrate the parameters above. We assume no prior information available for these parameters, thus a uniform prior is used. The joint posterior distribution of the parameters was sampled from $200,000$ MCMC iterations, after $100,000$ burn-in steps. The marginal posterior distribution of parameters for each scenario is shown in Fig. S1.

Once calibrated, the 2013 French Polynesia outbreak is used as the reference point to obtain infection parameters in other geo-locations. Specifically, β remains constant and independent of geographical locations, while all other parameters are rescaled in each subpopulation according to the daily temperature data, mosquito presence, and socioeconomic drivers, as shown in the following sections.

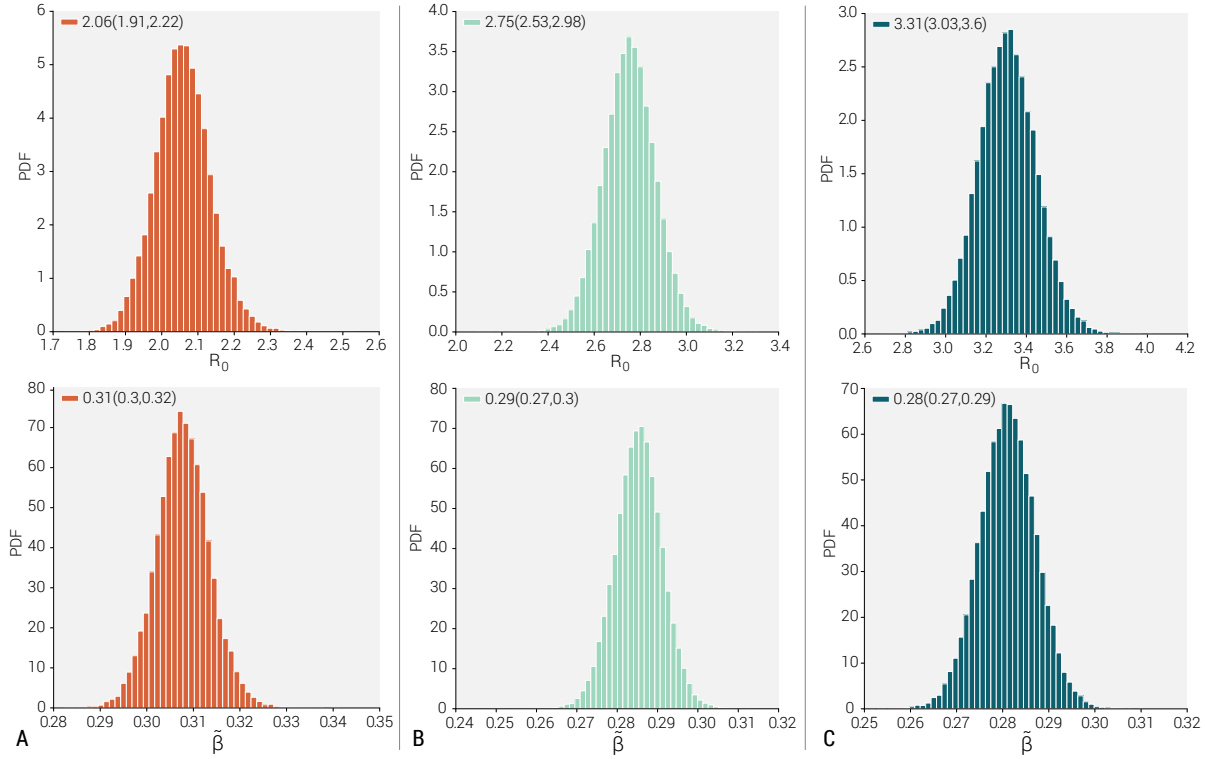


Figure S1: MCMC calibration based on the 2013-2014 French Polynesia outbreak [1]. A) Marginal posterior distribution of the basic reproduction number R_0 and overall transmission rate $\tilde{\beta}$ for the short serial interval scenario. The median R_0 and $\tilde{\beta}$ values are listed in the upper left-hand corners of the top and bottom figures, respectively. B) Marginal posterior distribution of the basic reproduction number R_0 and overall transmission rate $\tilde{\beta}$ for the reference serial interval scenario. C) Marginal posterior distribution of the basic reproduction number R_0 and overall transmission rate $\tilde{\beta}$ for the long serial interval scenario.

2 Spatiotemporal dependence of the vector population

The mosquito abundance, factored into the model through the number of mosquitoes per person, is a quantity that depends on the geographical location and time of the year. Mosquito abundance is crucial in defining the risk of ZIKV outbreaks, as well as temporal patterns. Here we consider the data for *Aedes* mosquitoes presence collected in Ref. [10]. The Global *Aedes aegypti* and *Aedes albopictus* distribution is provided at the fine spatial resolution of 5×5 km cells and yields the uncertainty of *Ae. aegypti/Ae. albopictus* presence in each cell. At 0.25×0.25 degree spatial resolution ($25km \times 25km$ along Earth's Equator), the cell used by GLEAM contains multiple measurements of vector presence uncertainty. A cell of GLEAM contains m measurement of vector presence uncertainty, which are $p_1, p_2, ...p_m$. We thus define an average vector presence

c in each cell of GLEAM, $c = m^{-1} \sum_{i=1}^{m} p_i$. The typical daily commuting range for humans is about the size of a GLEAM population cell. The population within the cell can be considered well-mixed, which means the entire population is exposed to the mosquitoes, but with a relative probability of mosquito explosure c in each cell.

As the mosquito presence distribution does not consider seasonal variation, we have included a monthly modulation function depending on the local temperature in each census area. This function was obtained by simulating a density-dependent stochastic model, which mimics the biological processes driving the developmental cycle of *Ae. albopictus* in a typical breeding site. The model is reported in Sec. 5. The modulation function has the following form:

$$k(t) \propto \exp(-(\tilde{T}(t) - 25)^2/50), \tag{12}$$

where $\tilde{T}(t) = \sum_{i=0}^{78} T(t-i)/79$, t is the time in days and $T(t)$ is the average temperature on day t. To obtain the absolute value of $k(t)$ we need a rescaling constant k_c that provides the variation of mosquitoes per person:

$$k_i(t) = c_i k_c \exp(-(\tilde{T}(t) - 25)^2/50), \tag{13}$$

where c_i accounts for the relative probability of exposure in cell i. Since the model is calibrated on the 2013 ZIKV outbreak in French Polynesia, under the assumption that during the outbreak the effective number of mosquitoes per person in French Polynesia (FP) is equal to the value k_{FP} used in the MCMC calibration, we obtain the following expression for k_c:

$$k_c = \frac{k_{FP}}{c_{FP} \ \exp(-(\tilde{T}(t^*) - 25)^2/50)}, \tag{14}$$

where $T(t^*)$ is the average temperature during the French Polynesia outbreak and c_{FP} is the specific rescaling factor of the number of mosquitoes per person in French Polynesia. The above calibration depends on the number of mosquitoes per person considered in French Polynesia. This value however must be consistent with the MCMC calibration that rescales the vector transmissibility accordingly. Since the MCMC procedure determines the R_0 characterizing the outbreak in French Polynesia, variations of k_{FP} are absorbed in the parameter β; as such, they do not alter the relative reproduction number variation across geographical location and time. We have explicitly considered values of k_{FP} in the range 1 to 3, confirming the invariance of the results under consistent rescaling of all parameters.

3 Population at risk of ZIKV exposure

The GLEAM model integrates the transportation dynamics at the level of subpopulation. Each subpopulation b is defined by a group of cells i that may have different local weather (for example, due to the altitude) and socioeconomic attributes. This implies that only a fraction of individuals belonging to each subpopulation is actually exposed to ZIKV and participates in the global spreading of the infection. In the following we use an approach that bears some resemblance the one used by Perkins et al. [11] in introducing socioeconomic factors, in that we use economic data and correlate with the magnitude of known outbreaks. However, while in

Ref. [11] the analysis aims at rescaling the local reproductive number, we opted for a rescaling of the population effectively exposed to the disease. In order to compute the population at risk, we must exclude from the exposed population, individuals belonging to cells where environmental factors are not favorable to the spreading of ZIKV. In particular, for each cell i, if the average reproductive number during the highest 180 days is less than one, the population is not considered at risk for a self-sustaining outbreak. Thus for each subpopulation, the fraction of population environmentally exposed to ZIKV is:

$$r_{en} = \frac{\sum_{i'} n_{i'}}{\sum_i n_i}, \tag{15}$$

where i' denotes a cell at risk of ZIKV, n_i is the population of the cell considered and the summation is over all cells i included in the subpopulation b.

However, many of the studies suggest that even if the environmental conditions are suitable for arbovirus transmission, the population's risk of exposure to mosquitoes may still vary due to socio-economic heterogeneities [12, 11, 13]. For example, different socio-economic factors, such as improved sanitation facilities, the fraction of the population living in extreme poverty, use of air conditioning in buildings, housing conditions, education level, and level of employment, may alter the arbovirus exposure risk. All of those factors are in general strongly correlated with the level of economic development of the geographical region under study.

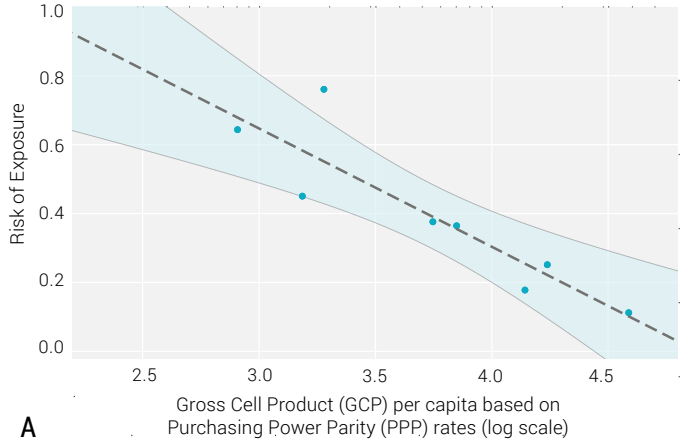
For this reason, we have only considered arbovirus outbreaks in naïve populations for which reliable estimates are available for both the final infection attack (AR_{final}), which is generally obtained through seroprevalence studies, and the ideal infection attack rate (AR_{ideal}), which is computed when only environmental factors are considered. The ratio $r_{se} = AR_{final}/AR_{ideal}$ provides a proxy for the fraction of exposed population that we can associate with the geographically-based version of the per capita Gross Domestic Product based on Purchasing Power Parity rates (GDP per capita, PPP) which is in turn used to capture the socio-economic differences that exist across and within countries. Each cell is then assigned a Gross Cell Product (GCP) by allocating the subpopulation GDP proportionally to the population sizes of the cells within this subpopulation. We find that the above association is well approximated by the relation:

$$\widehat{r}_{se} = \widehat{\alpha} + \widehat{\delta} \log(\text{GCP per capita, PPP}), \tag{16}$$

where $\widehat{\alpha}$ and $\widehat{\delta}$ are estimated using an ordinary least squares (OLS) regression based on the outbreak reported in Fig. S2. The quantity $\widehat{r}_{se}$ is associated with the corresponding value of the Gross Cell Product (GCP) per capita of each GLEAM cell, and it yields the fraction of population actually exposed to ZIKV.

However, in our model simulations, we do not use the point estimate of $\widehat{r}_{se}$. Rather, for each cell, we consider $1,000$ different values as drawn from the 95% prediction interval of the fitted model. By doing so, we control for the fact that our regression model has been calibrated using a limited amount of data and therefore introduce an additional element of stochasticity in our simulations to account for the uncertainty related to our estimates.

In order to derive the fraction of population exposed to ZIKV in each subpopulation b we can consider all cells $i|i \in b$. Let n_i denote the population in cell i, and let $r_{se,i}$ denote the



B

Outbreak Location	Attack Rate	Ref.
Lamu Island Kenya	0.75	[14]
Grange Comore Island	0.63	[15]
Manajary, Madagascar	0.45	[16]
Mayotte	0.37	[12]
Reunion Island	0.35	[17]
Paramaribo, Suriname	0.25	[18]
Saint Martin	0.17	[19]
Castiglione de Cervia, Italy	0.10	[20]

Figure S2: Risk of exposure as function of the Gross Cell Product (GCP) per capita (dashed line and shaded area represent best fit and 95% CI separately). Attack rates of previous chikungunya outbreak can be found in Refs. [14, 15, 16, 12, 17, 18, 19, 20]

fraction of people in cell i exposed to ZIKV for socio-economic reasons. The overall population exposed to ZIKV in the subpopulation b is then:

$$N_b^{exp} = \sum_{\{i'\}} r_{se,i'} n_{i'}, \tag{17}$$

where i' are the cells environmentally exposed to ZIKV.

Within each subpopulation b only exposed individuals N_b^{exp} are considered in the infection transmission dynamic, while the entire population N_b is considered in the mobility process. The baseline level ZIKV infection dynamic works at the homogeneous mixing level, and quantities are thus averaged over the environmentally exposed cells:

$$c_b = \frac{\sum_{\{i'\}} c_{i'} r_{se,i'} n_{i'}}{N_b^{exp}}, \tag{18}$$

and

$$T_b(t) = \frac{\sum_{\{i'\}} T_{i'}(t) r_{se,i'} n_{i'}}{N_b^{exp}}. \tag{19}$$

The remaining spatio-temporal dependent infection parameters at the subpopulation level can be calculated accordingly. In Fig. S3 we show a schematic representation of the process of computing the remaining population in each cell of GLEAM. Starting from the original cell's population in GLEAM, the GECON, and *Aedes* mosquitoes distribution data act like filters for the population at risk through r_{en} and r_{se}.

The spatial heterogeneities of population at risk due to environmental and socio-economic factors also affect the diffusion of disease among subpopulations. A person exposed or infected

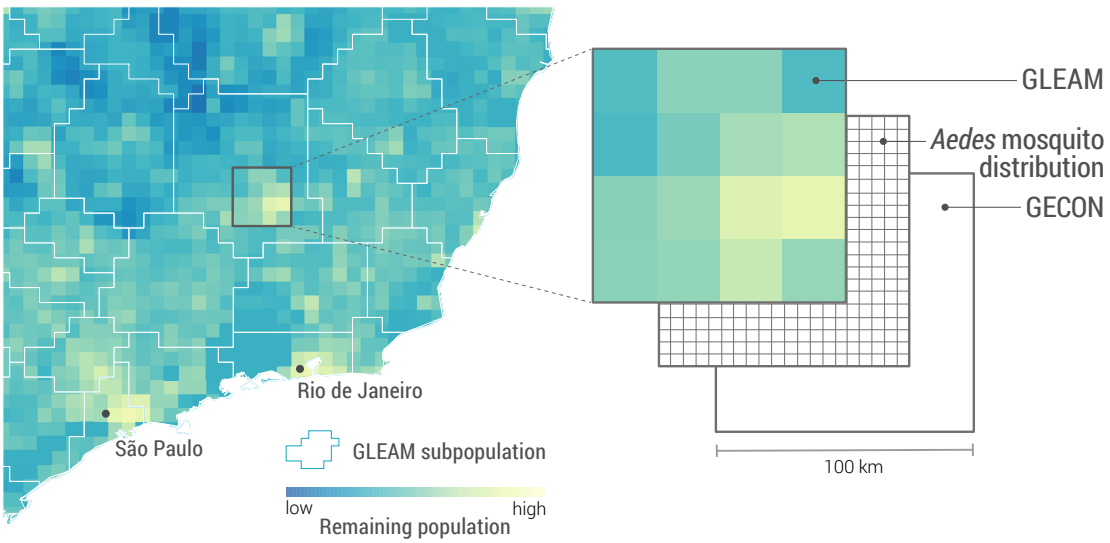


Figure S3: Schematic representation of the process of computing the remaining population for each subpopulation in GLEAM. Starting from the original cell's population in GLEAM, the GECON and *Aedes* mosquito distribution data act like filters for the population at risk through r_{en} and r_{se}.

with ZIKV who travels from a subpopulation experiencing an on-going outbreak will not be able to seed the epidemic in the subpopulation of travel destination, if his or her destination is not at risk of ZIKV due to environmental or socioeconomic factors. Specifically, in the model, assuming the fraction of exposed population in destination subpopulation b is $N_b^{\exp}/N_b$, the probability of a traveling ZIKV carrier entering an area where the population is exposed to ZIKV and participating in the transmission dynamics is $N_b^{\exp}/N_b$. Conversely, the probability of a traveling ZIKV carrier entering the area where the population is not exposed to ZIKV and isolated from further transmission is $1 - N_b^{\exp}/N_b$. It is worth noticing that by focusing on the fraction of effectively exposed population, even in places where economic factors can be extremely favorable it is possible to have smalls outbreaks. A clear example of this situation is the ZIKV outbreak in the US.

4 Microcephaly Projection

The projection of potential microcephaly cases related to ZIKV follows the model proposed in the study of ZIKV-microcephaly association of 2013-2014 French Polynesia outbreak [21]. Specifically, we used a first trimester model: if a woman is infected with ZIKV during the first trimester of her pregnancy, the risk of microcephaly associated with ZIKV is p_m during the first trimester and 0 otherwise. For simplicity, we use a pregnancy model with a fixed duration of pregnancy of 40 weeks; neither miscarriage nor termination of pregnancy is considered. Given

the weekly birth rate r_b [22] and weekly new ZIKV infections $c(t)$ in an administrative area with population N, the number of women beginning a pregnancy in a given week is:

$$n_p = N r_b. \tag{20}$$

The probability of a woman being infected with ZIKV during the first trimester of her pregnancy is:

$$p_z(t) = \frac{\sum_{t'=t}^{t+\Delta t_{trim1}} c(t')}{N}, \tag{21}$$

where $\Delta t_{trim1} = 13\ weeks$ is the length of first trimester.

Thus, the projected number of microcephaly cases of a given week is given by:

$$n_m(t + \Delta t_{preg}) = n_p \times p_z(t) \times p_m = N r_b \times p_m \times \frac{\sum_{t'=t}^{t+\Delta t_{trim1}} c(t')}{N}, \tag{22}$$

where $\Delta t_{preg} = 40\ weeks$ is the duration of pregnancy.

Equation 22 establishes the relationship between number of new ZIKV cases $c(t)$ and projected number of new microcephaly cases $n_m(t)$.

5 Developmental cycle of *Aedes* mosquitoes

We estimate a temperature modulation function that reproduces the seasonal pattern of female adult mosquitoes. The proposed approach is based on a model previously used to estimate the abundance of female adults of *Ae. albopictus* during the 2007 chikungunya outbreak in Emilia Romagna (Italy) [23]. Briefly, the model mimics the biological processes driving the full developmental cycle of *Ae. albopictus* in a typical breeding site, explicitly accounting for egg hatching, pupation, adult emergence, and for the adult life cycle of alternate feeding and laying of eggs (gonotrophic cycle). The developmental rates from one stage to the next, the duration of the gonotrophic cycle, and the mortality rates of different life stages depend on the average daily temperature [24]. The temporal dynamics of eggs (E), larvae (L), pupae (P), and female adults (A) is described by the following equations:

$$\begin{aligned} E' &= n_e \frac{1}{g} A \left(1 - \frac{E}{K}\right) - \mu_e E - d_e E \\ L' &= d_e E - d_l L - \mu_l L \\ P' &= d_l L - d_p P - \mu_p P \\ A' &= \frac{1}{2} d_p P - \mu_a A, \end{aligned} \tag{23}$$

where n_e is the number of eggs laid in one oviposition, g is the duration of gonotrophic cycle, K drives the carrying capacity for the eggs, $\mu_e, \mu_l, \mu_p, \mu_a$ are the death rates associated with different stages of the mosquitoes and d_e, d_l, d_p are the developmental rates driving the transitions of vectors across the different mosquito life stages; the 1/2 term in the last equation accounts for a 1:1 sex ratio of adult mosquitoes.

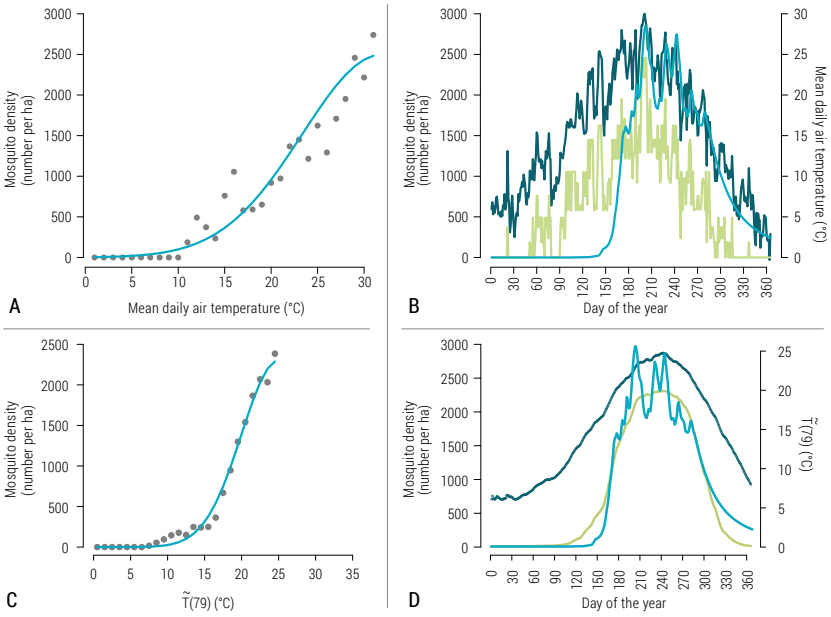


Figure S4: A) Average mosquito density at different temperatures as estimated by model simulation (grey points). Fitted average mosquito density at different temperatures, by assuming the following relationship: $D(T) = 2500exp\{-(T-32)^2/150\}$ (cyan line). B) Mean daily air temperature over time (blue line). Average mosquito density as estimated by model simulation (cyan line). Predicted mosquito density over time, by assuming the following relationship: $D(T) = 2500exp\{-(T-32)^2/150\}$ (green line). C) Average mosquito density at different temperatures as estimated by model simulation (grey points). Fitted average mosquito density at different temperature values of $\tilde{T} = \sum_{i=0}^{78} T(t-i)/79$, by assuming the following relationship: $D(\tilde{T}) = 2300exp\{-(\tilde{T}-25)^2/50\}$ (cyan line). D) Mean temperature values of $\tilde{T}$ over time (blue line). Average mosquito density as estimated by model simulation (cyan line). Predicted mosquito density over time, by assuming the relationship following: $D(\tilde{T}) = 2300exp\{-(\tilde{T}-25)^2/50\}$ (green line).

The simulated average abundance of female adult mosquitoes within a breeding site is displayed in Fig S4 A), along with the average daily air temperature observed in Emilia Romagna during 2007. The expected density of adult mosquitoes at different temperatures were estimated by computing for each degree of temperature T between 0°C–30°C the mean number of female adults predicted among days characterized by an average daily temperature within the range defined by T-0.5°C and T+0.5°C. Since high temperature reduces the mosquito survival rate (especially in adults), and low temperature prevents the development of immature stages into adults, the mosquito density is expected to be lower at both low and high temperature regimes. We therefore fit a Normal density function to the expected number of adult mosquito at different temperatures. Obtained results are shown in Fig S4B. The described procedure provides a modulation function of temperature that is used to approximate the seasonality of the mosquito

abundance.

As shown in Fig S4 B) the proposed approximation overestimates the abundance of the vector in spring (i.e. at the beginning of the breeding season), and underestimates the mosquito density in autumn. In fact, this procedure does not account for two critical factors. The first one is that adult abundance depends on the persistence of favorable temperature conditions during the whole life cycle of the mosquito's development. The second one is that the mosquito density at a given time is influenced by the vector abundance in the preceding generations. We therefore investigated the relationship between the density of adult mosquitoes at a given time t with the mean air temperature recorded between $t - \Delta t$ and t. In particular, we considered different values of Δt, ranging from 1 to 365 days and fit separately for each value of Δt a Normal density function to the expected abundance of vectors at different values of $\tilde{T}(t) = \sum_{i=0}^{\Delta t} T(t-i)/\Delta t$ (see Fig S4C). Results show that the best approximation of the adult mosquito density is obtained when a time window of 79 days is considered. The proposed approach provides a suitable modulation function to reproduce seasonal patterns characterizing the relative abundance of adult mosquitoes over time by using temperature values only (see Fig S4 D).

6 Sensitivity analysis

In the following we report sensitivity analyses that are calibrated according to three different scenarios.

- *Aegypti scenario.* This scenario considers *Ae. aegypti* as the only competent ZIKV vector. The parameters describing the infection are set in the middle of the range of the estimates in the literature.

- *Short serial interval scenario.* This scenario considers both *Ae. aegypti* and *Ae. albopictus* as competent ZIKV vectors. The parameters describing the infection are set in order to explore the shortest serial interval allowed by the range of parameters reported in the literature.

- *Long serial interval scenario.* This scenario considers both *Ae. aegypti* and *Ae. albopictus* as competent ZIKV vectors. The parameters describing the infection are set in order to explore the longest serial interval allowed by the range of parameters reported in the literature.

The three additional scenarios described above, when compared to the reference scenario, provide similar posterior distributions of location and time of introduction in Brazil. The most likely time of introduction is between October and December 2013, and the most likely location of introduction is Rio de Janeiro for all three scenarios. The timing and profile of Zika/births with first trimester ZIKV infections resemble the reference scenario as well. Variations are observed for the country level ZIKV infection ARs (by February 28, 2017). This is clearly due to the change of the reproductive number in the case of longer or shorter serial interval and to the difference in mosquitoes' presence in the case that only *Ae. aegypti* is a competent vector. Figure S5 summarizes the changes in the three additional scenarios when compared to the reference scenario.

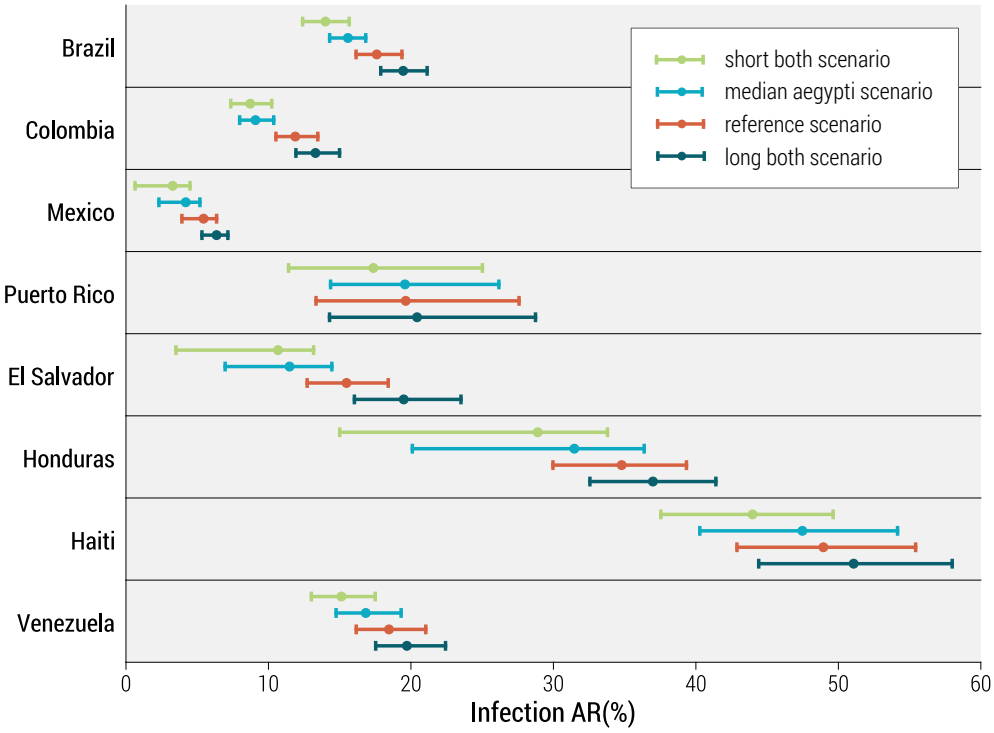


Figure S5: **Comparison of ZIKV infection ARs between reference scenario and three additional scenarios:** Projected ZIKV infection ARs(Median estimates) and 95% CI through February 28, 2017 for the Reference scenario, Aegypti scenario, Short serial interval scenario, Long serial interval scenario, in eight affected countries in the Americas. ZIKV attack rates include asymptomatic infections. The denominator is the entire country population, including regions that are not exposed to the vector.

Aegypti scenario

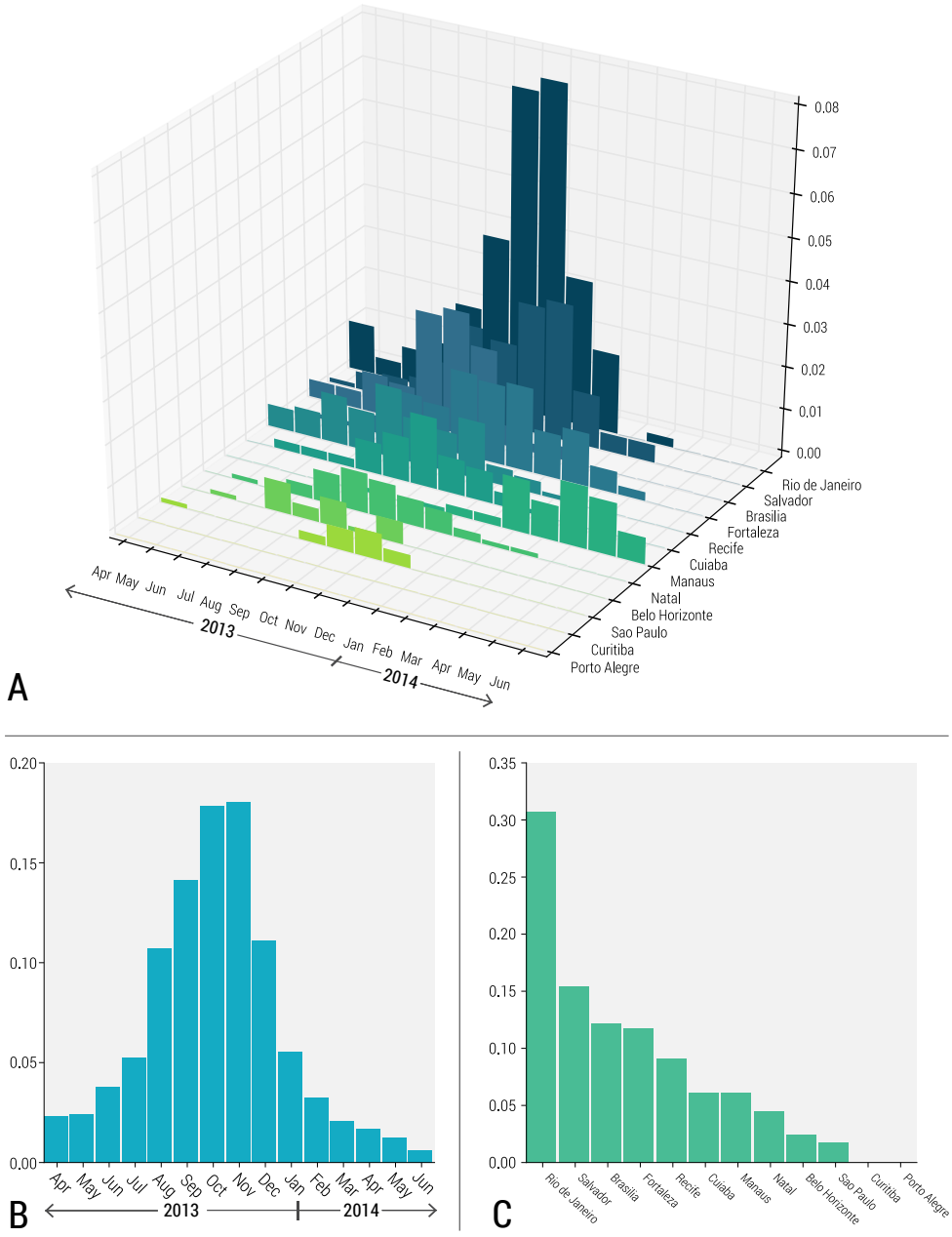


Figure S6: **Aegypti Scenario:** Posterior distribution for ZIKV introductions in twelve major transportation hubs in Brazil between April 2013 and June 2014, incorporating the likelihood of replicating the observed epidemic peak in Colombia. A) Full posterior distribution as a function of location and time of introduction. B) Marginal posterior distribution for time (month) of introduction. C) Marginal posterior distribution for location of introduction.

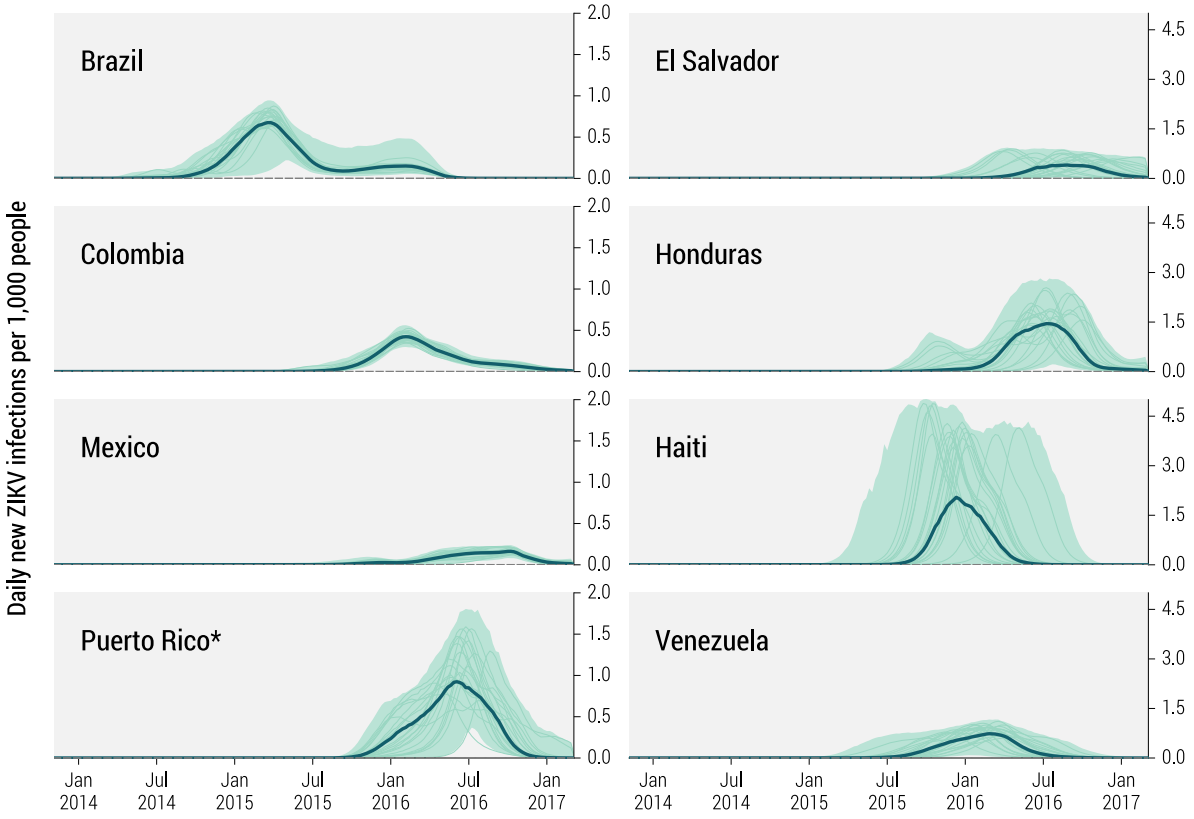


Figure S7: **Aegypti Scenario:** Estimated daily number of new ZIKV infections (per 1000 people) in eight affected countries in the Americas between January 2014 and February 2017. The bold line and shaded area refer to the estimated median number of infections and 95 % CI of the model projections, respectively. Rates include asymptomatic infections. The median incidence is calculated each week from the stochastic ensemble output of the model and may not be representative of specific epidemic realizations. Thin lines represent a sample of specific realizations. *Puerto Rico curves are constrained under the condition that the peak of incidence curve is after March 1, 2016, based on surveillance data [25].

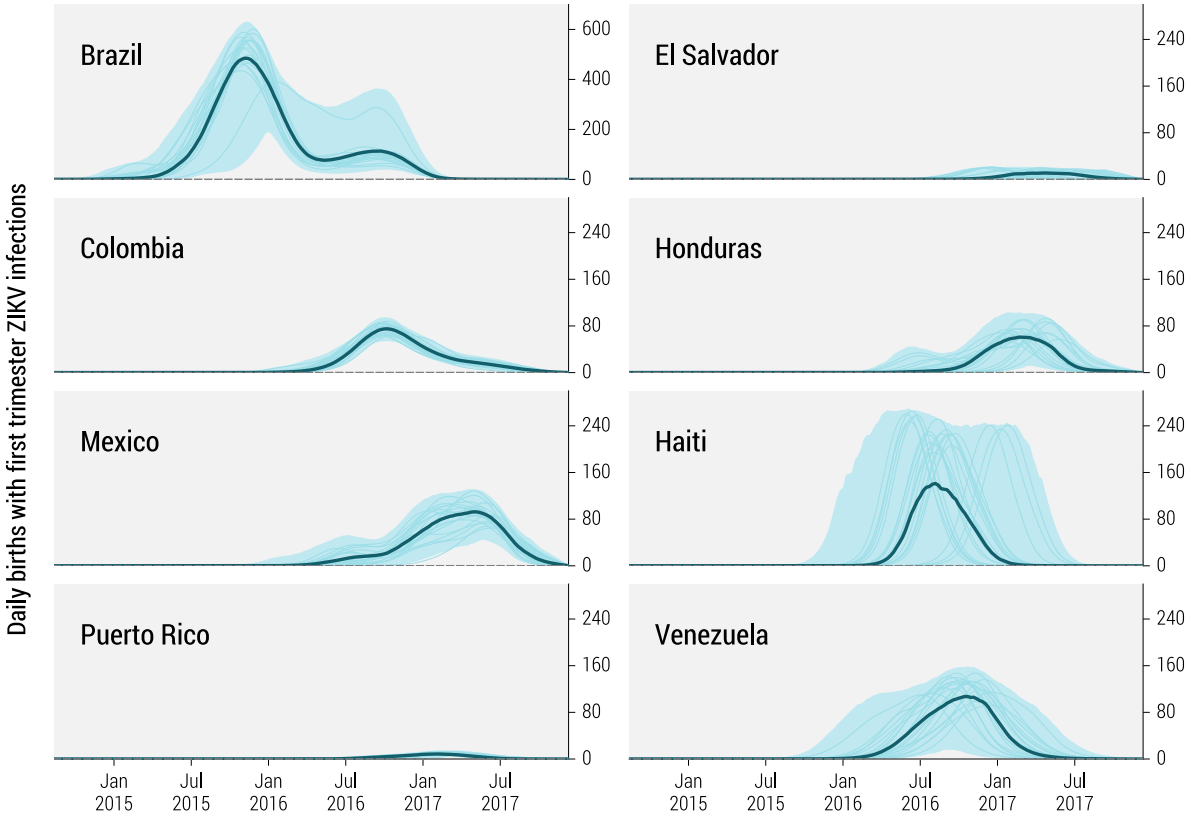


Figure S8: **Aegypti Scenario:** Estimated daily number of births between October 2014 and December 2017 from women infected with ZIKV during the first trimester of pregnancy in eight affected countries in the Americas. The bold line and shaded area refer to the estimated median number of births and 95 % CI of the model projections, respectively. Note that Brazil is plotted with a different scale. The median curve is calculated each week from the stochastic ensemble output of the model and may not be representative of specific epidemic realizations. Thin lines represent a sample of specific realizations.

	Infection AR %		Cumulative Microcephaly Cases (median with 95%CI)					
	(median with 95%CI)		first trimester risk: 0.95%		first trimester risk: 2.19%		first trimester risk: 4.52%	
	Feb. 1, 2016	Feb. 28, 2016	Feb. 1, 2016	Dec. 10, 2017	Feb. 1, 2016	Dec. 10, 2017	Feb. 1, 2016	Dec. 10, 2017
Brazil	15 [12 - 16]	16 [14 - 17]	808 [186 - 1022]	1148 [1054 - 1240]	1863 [430 - 2357]	2647 [2429 - 2859]	3845 [887 - 4864]	5463 [5013 - 5901]
Colombia	3 [2 - 5]	9 [8 - 10]	0 [0 - 5]	167 [147 - 191]	1 [0 - 12]	386 [339 - 440]	1 [0 - 26]	796 [699 - 908]
Mexico	0 [0 - 2]	4 [2 - 5]	0 [0 - 9]	242 [133 - 299]	0 [0 - 21]	557 [307 - 689]	0 [0 - 43]	1151 [634 - 1423]
Puerto Rico*	2 [0 - 6]	20 [14 - 26]	0 [0 - 0]	19 [14 - 25]	0 [0 - 0]	43 [31 - 57]	0 [0 - 0]	88 [65 - 118]
El Salvador	0 [0 - 3]	12 [7 - 15]	0 [0 - 0]	29 [18 - 37]	0 [0 - 0]	67 [41 - 85]	0 [0 - 0]	139 [84 - 175]
Honduras	1 [0 - 15]	32 [20 - 36]	0 [0 - 0]	130 [83 - 151]	0 [0 - 1]	300 [192 - 347]	0 [0 - 2]	620 [396 - 717]
Haiti	37 [0 - 53]	48 [40 - 54]	0 [0 - 120]	306 [260 - 349]	0 [0 - 278]	706 [599 - 805]	0 [0 - 573]	1456 [1235 - 1661]
Venezuela	8 [1 - 18]	17 [15 - 19]	1 [0 - 61]	247 [216 - 283]	2 [0 - 141]	569 [498 - 653]	3 [0 - 291]	1174 [1029 - 1347]

Figure S9: **Aegypti Scenario:** Projected ZIKV infection ARs through the time of the WHO declaration of a PHEIC on February 1, 2016, and through February 28, 2017, in eight affected countries in the Americas. Median estimates and 95 % CIs are provided. ZIKV attack rates include asymptomatic infections. The denominator is the entire country population, including regions that are not exposed to the vector. Cumulative microcephaly cases due to ZIKV infection during the first trimester of pregnancy through the time of the WHO declaration of a PHEIC on February 1, 2016, and through December 10, 2017, in eight affected countries in the Americas. We consider three different risks of microcephaly associated with ZIKV infection during the first trimester: 0.95% first trimester risk based on a study of the 2013-2014 French Polynesian outbreak [21]; 2.19% (100% over-reporting) and 4.52% (no over-reporting) first trimester risks, based on a study of Bahia, Brazil [26], given a model-estimated 29% infection AR in Bahia. *Puerto Rico curves constrained under the condition that the peak of ZIKV incidence curve is after March 1, 2016, based on surveillance data [25].

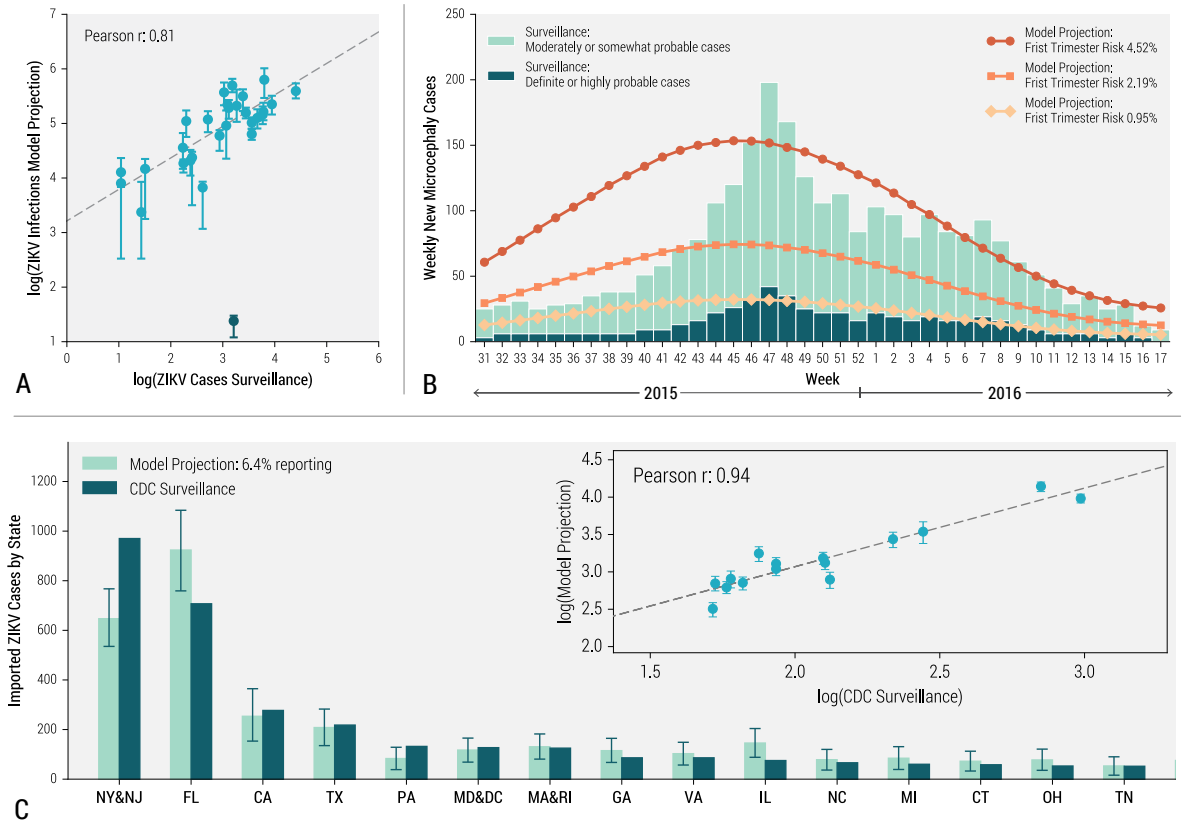


Figure S10: **Aegypti Scenario:** A) Correlation between the number of ZIKV cases by state in Colombia as reported by surveillance data through October 1, 2016 [27], compared with state-level model projections of infections (median with 95 % CI). Pearson's r correlation coefficient is reported for the linear association on the log scale. The outlier (in dark green) excluded from the statistical analysis corresponds to the Arauca region. B) Timeline of microcephaly cases in Brazil though April 30, 2016. Bar plot shows weekly definite (or highly probable cases) and moderately (or somewhat probable cases) from surveillance data [28]. Line plots indicate estimated weekly new microcephaly cases given three levels of first trimester risk: 4.52% (round) [26], 2.19% (square) [26], and 0.95% (diamond) [21]. C) Bar plot of ZIKV infections imported into the continental USA by state(s) as reported by CDC surveillance through October 5, 2016 [29], and compared to model projections (median with 95 % CI) for the same period assuming 5.74% reporting/detection. The insert shows the correlation between CDC surveillance data and model projections (median with 95 % CI).

Short serial interval scenario

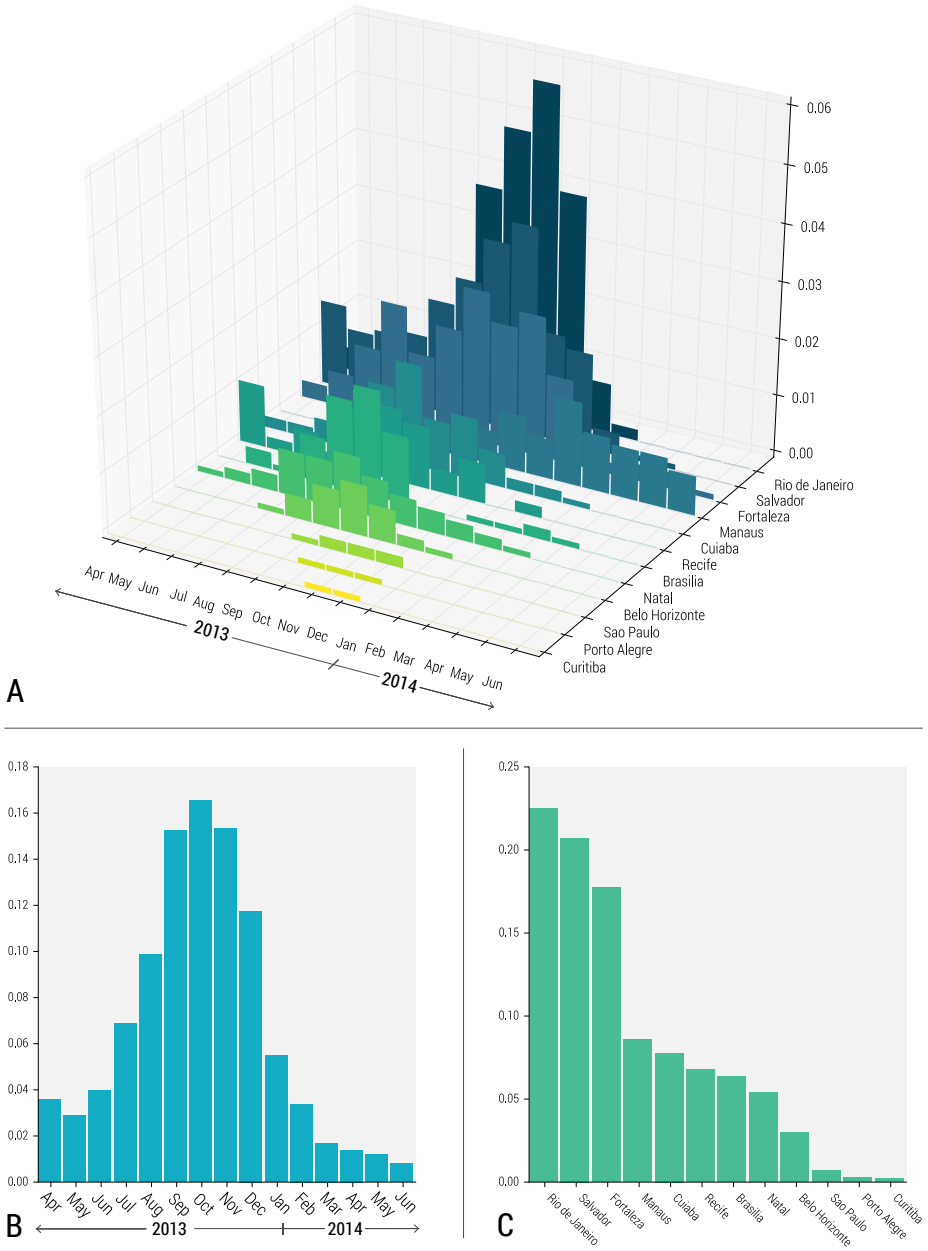


Figure S11: **Short Serial Interval Scenario:** Posterior distribution for ZIKV introductions in twelve major transportation hubs in Brazil between April 2013 and June 2014, incorporating the likelihood of replicating the observed epidemic peak in Colombia. A) Full posterior distribution as a function of location and time of introduction. B) Marginal posterior distribution for time (month) of introduction. C) Marginal posterior distribution for location of introduction.

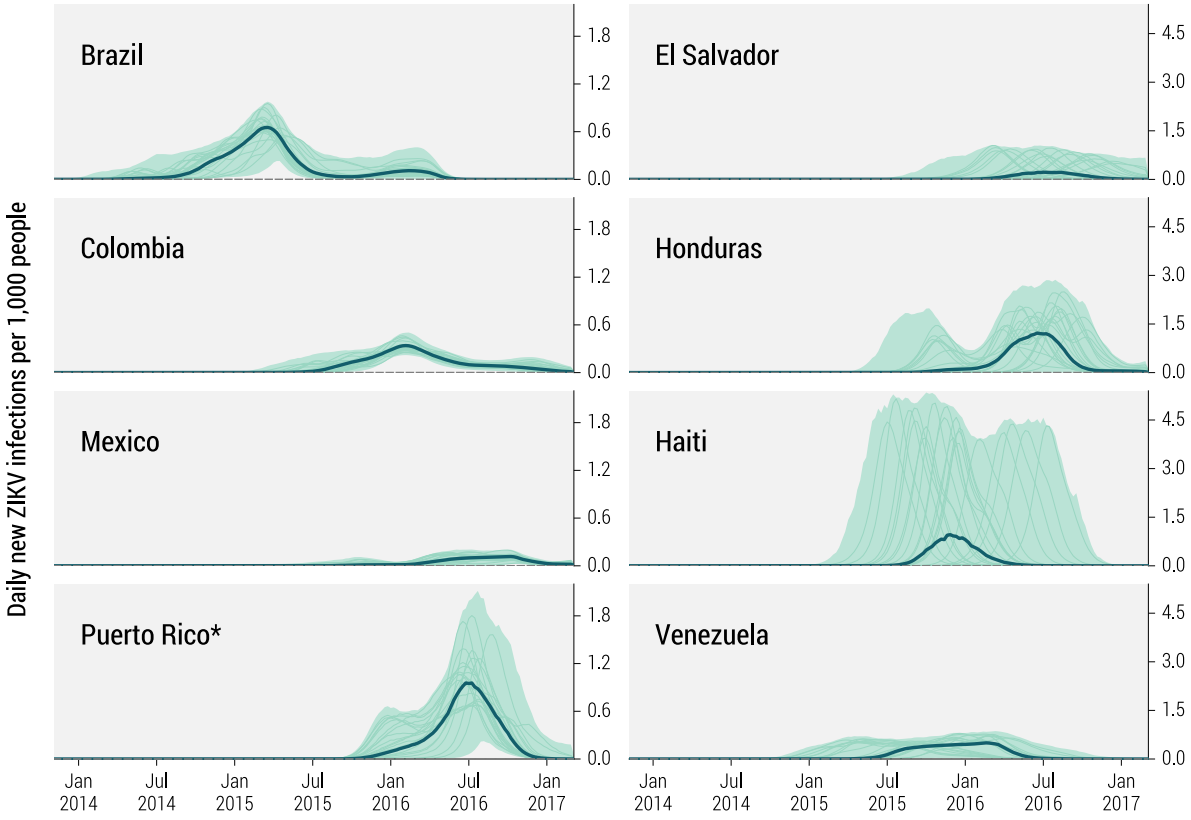


Figure S12: **Short Serial Interval Scenario:** Estimated daily number of new ZIKV infections (per 1000 people) in eight affected countries in the Americas between January 2014 and February 2017. The bold line and shaded area refer to the estimated median number of infections and 95 % CI of the model projections, respectively. Rates include asymptomatic infections. The median incidence is calculated each week from the stochastic ensemble output of the model and may not be representative of specific epidemic realizations. Thin lines represent a sample of specific realizations. *Puerto Rico curves are constrained under the condition that the peak of incidence curve is after March 1, 2016, based on surveillance data [25].

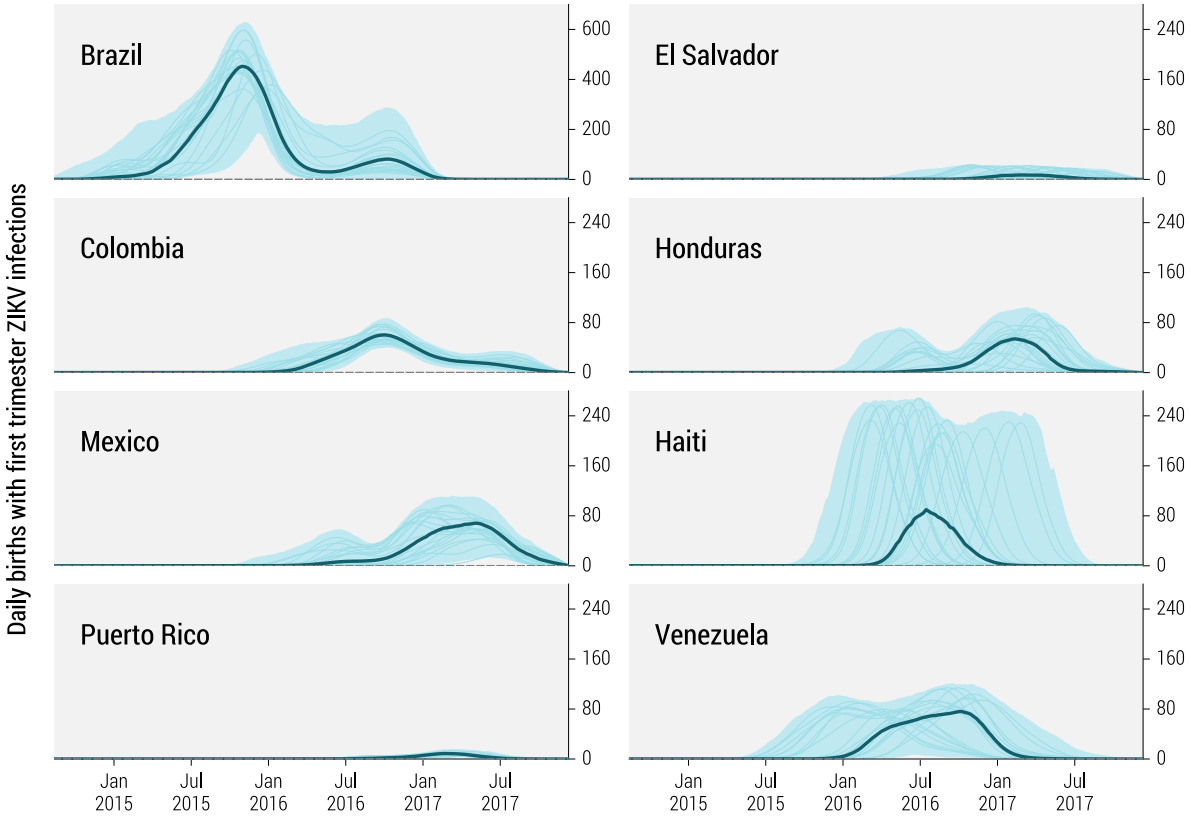


Figure S13: **Short Serial Interval Scenario:** Estimated daily number of births between October 2014 and December 2017 from women infected with ZIKV during the first trimester of pregnancy in eight affected countries in the Americas. The bold line and shaded area refer to the estimated median number of births and 95 % CI of the model projections, respectively. Note that Brazil is plotted with on different scale. The median curve is calculated each week from the stochastic ensemble output of the model and may not be representative of specific epidemic realizations. Thin lines represent a sample of specific realizations.

	Infection AR %		Cumulative Microcephaly Cases (median with 95%CI)					
	(median with 95%CI)		first trimester risk: 0.95%		first trimester risk: 2.19%		first trimester risk: 4.52%	
	Feb. 1, 2016	Feb. 28, 2017	Feb. 1, 2016	Dec. 10, 2017	Feb. 1, 2016	Dec. 10, 2017	Feb. 1, 2016	Dec. 10, 2017
Brazil	13 [10 - 15]	14 [12 - 16]	836 [234 - 1067]	1032 [914 - 1155]	1927 [539 - 2460]	2379 [2107 - 2662]	3978 [1113 - 5077]	4910 [4348 - 5495]
Colombia	4 [2 - 6]	9 [7 - 10]	1 [0 - 25]	161 [135 - 188]	1 [0 - 59]	370 [312 - 434]	3 [0 - 121]	765 [645 - 896]
Mexico	0 [0 - 2]	3 [1 - 4]	0 [0 - 19]	189 [37 - 259]	0 [0 - 43]	436 [85 - 597]	0 [0 - 89]	899 [175 - 1232]
Puerto Rico*	1 [0 - 6]	17 [11 - 25]	0 [0 - 0]	16 [11 - 24]	0 [0 - 0]	38 [25 - 54]	0 [0 - 0]	78 [51 - 112]
El Salvador	0 [0 - 9]	11 [4 - 13]	0 [0 - 0]	27 [9 - 33]	0 [0 - 1]	63 [21 - 77]	0 [0 - 1]	129 [42 - 159]
Honduras	1 [0 - 27]	29 [15 - 34]	0 [0 - 9]	120 [62 - 140]	0 [0 - 21]	276 [143 - 323]	0 [0 - 43]	569 [296 - 666]
Haiti	38 [0 - 49]	44 [38 - 50]	0 [0 - 187]	284 [242 - 320]	0 [0 - 430]	654 [558 - 738]	0 [0 - 888]	1349 [1152 - 1522]
Venezuela	11 [2 - 16]	15 [13 - 18]	3 [0 - 157]	222 [191 - 256]	8 [0 - 362]	511 [440 - 591]	16 [0 - 747]	1054 [908 - 1220]

Figure S14: **Short Serial Interval Scenario:** Projected ZIKV infection ARs through the time of the WHO declaration of a PHEIC on February 1, 2016, and through February 28, 2017, in eight affected countries in the Americas. Median estimates and 95 % CIs are provided. ZIKV attack rates include asymptomatic infections. The denominator is the entire country population, including regions that are not exposed to the vector. Cumulative microcephaly cases due to ZIKV infection during the first trimester of pregnancy through the time of the WHO declaration of a PHEIC on February 1, 2016, and through December 10, 2017, in eight affected countries in the Americas. We consider three different risks of microcephaly associated with ZIKV infection during the first trimester: 0.95% first trimester risk based on a study of the 2013-2014 French Polynesian outbreak [21]; 2.19% (100% over-reporting) and 4.52% (no over-reporting) first trimester risks, based on a study of Bahia, Brazil [26], given a model-estimated 27% infection AR in Bahia. *Puerto Rico curves constrained under the condition that the peak of ZIKV incidence curve is after March 1, 2016, based on surveillance data [25].

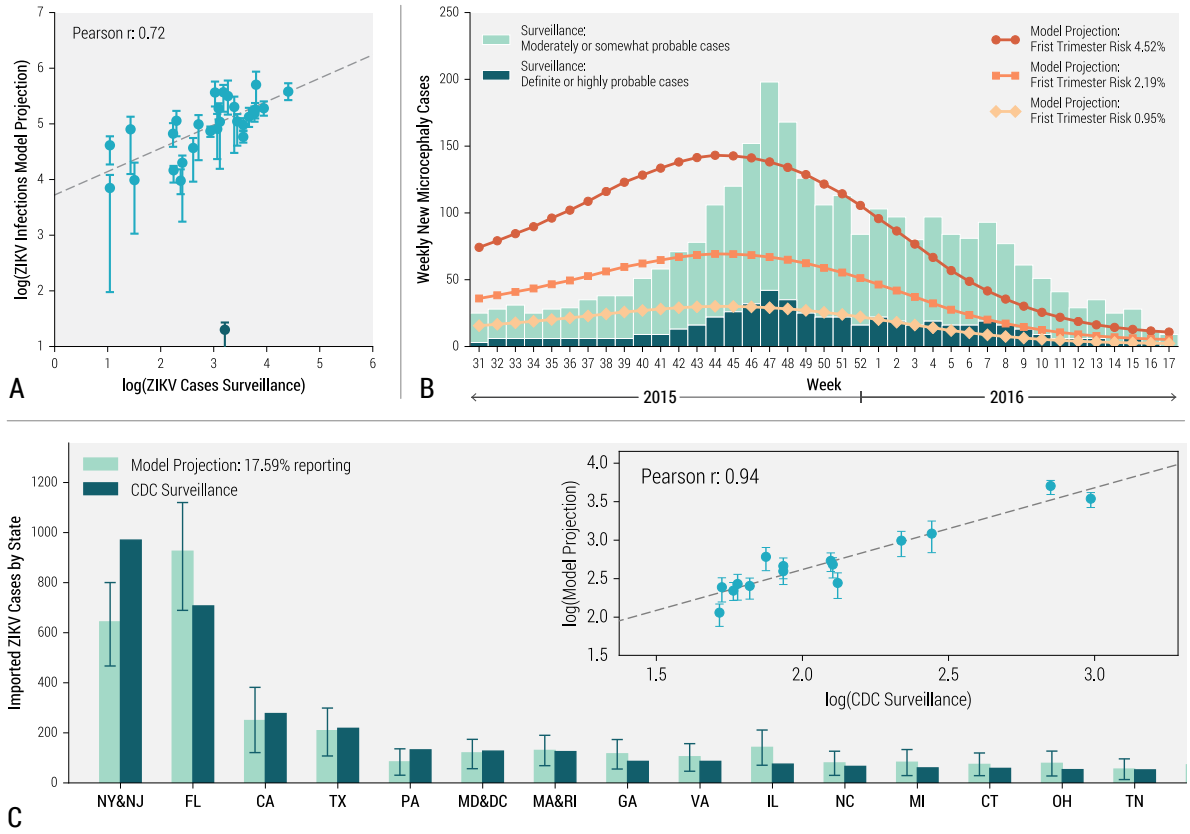


Figure S15: **Short Serial Interval Scenario:** A) Correlation between the number of ZIKV cases by state in Colombia as reported by surveillance data through October 1, 2016 [27], compared with state-level model projections of infections (median with 95 % CI). Pearson's r correlation coefficient is reported for the linear association on the log scale. The outlier (in dark green) excluded from the statistical analysis corresponds to the Arauca region. B) Timeline of microcephaly cases in Brazil though April 30, 2016. Bar plot shows weekly definite (or highly probable cases) and moderately (or somewhat probable cases) from surveillance data [28]. Line plots indicate estimated weekly new microcephaly cases given three levels of first trimester risk: 4.52% (round) [26], 2.19% (square) [26], and 0.95% (diamond) [21]. C) Bar plot of ZIKV infections imported into the continental USA by state(s) as reported by CDC surveillance through October 5, 2016 [29], and compared to model projections (median with 95 % CI) for the same period assuming 5.74% reporting/detection. The insert shows the correlation between CDC surveillance data and model projections (median with 95% CI).

Longserial interval scenario

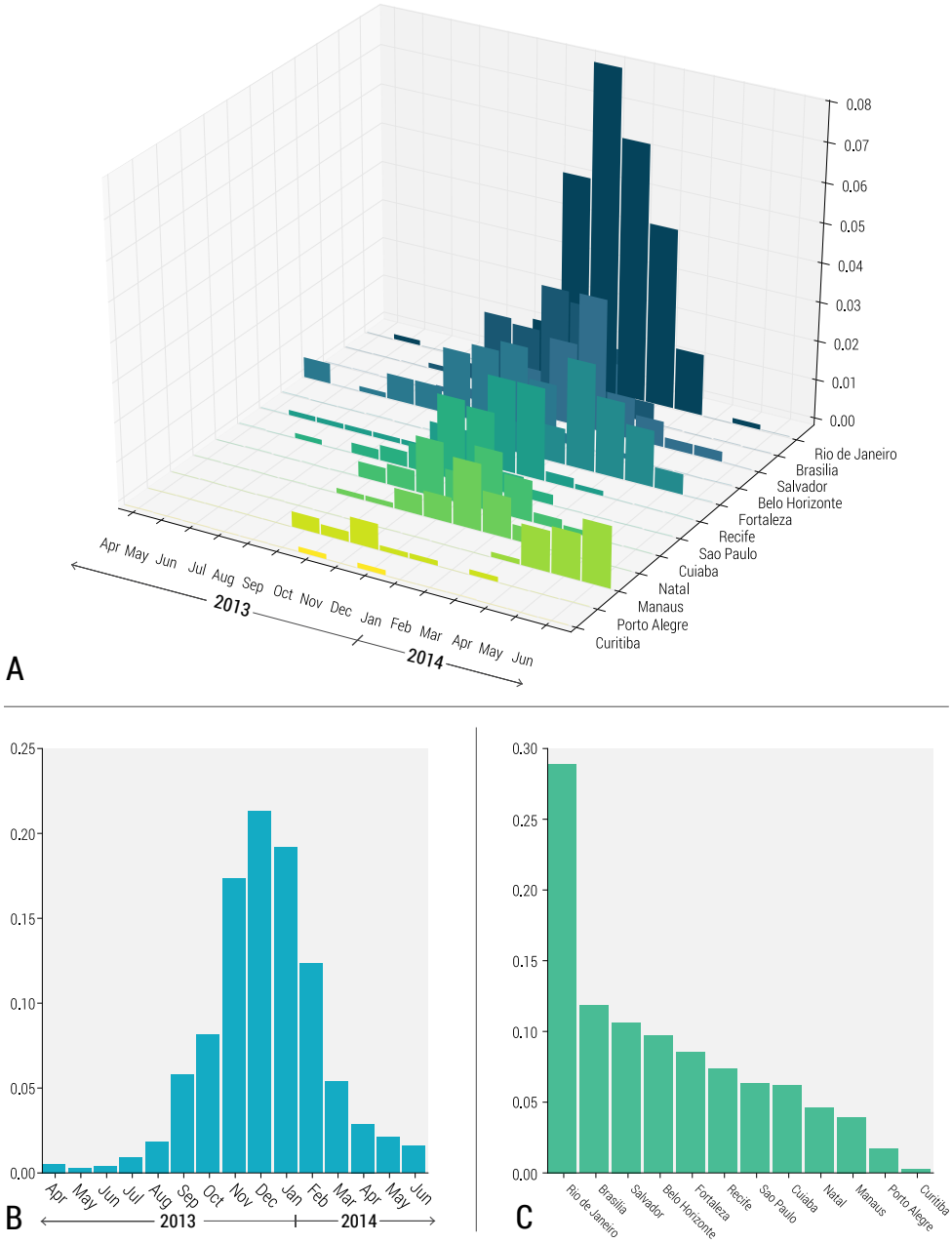


Figure S16: **Long Serial Interval Scenario:** Posterior distribution for ZIKV introductions in twelve major transportation hubs in Brazil between April 2013 and June 2014, incorporating the likelihood of replicating the observed epidemic peak in Colombia. A) Full posterior distribution as a function of location and time of introduction. B) Marginal posterior distribution for time (month) of introduction. C) Marginal posterior distribution for location of introduction.

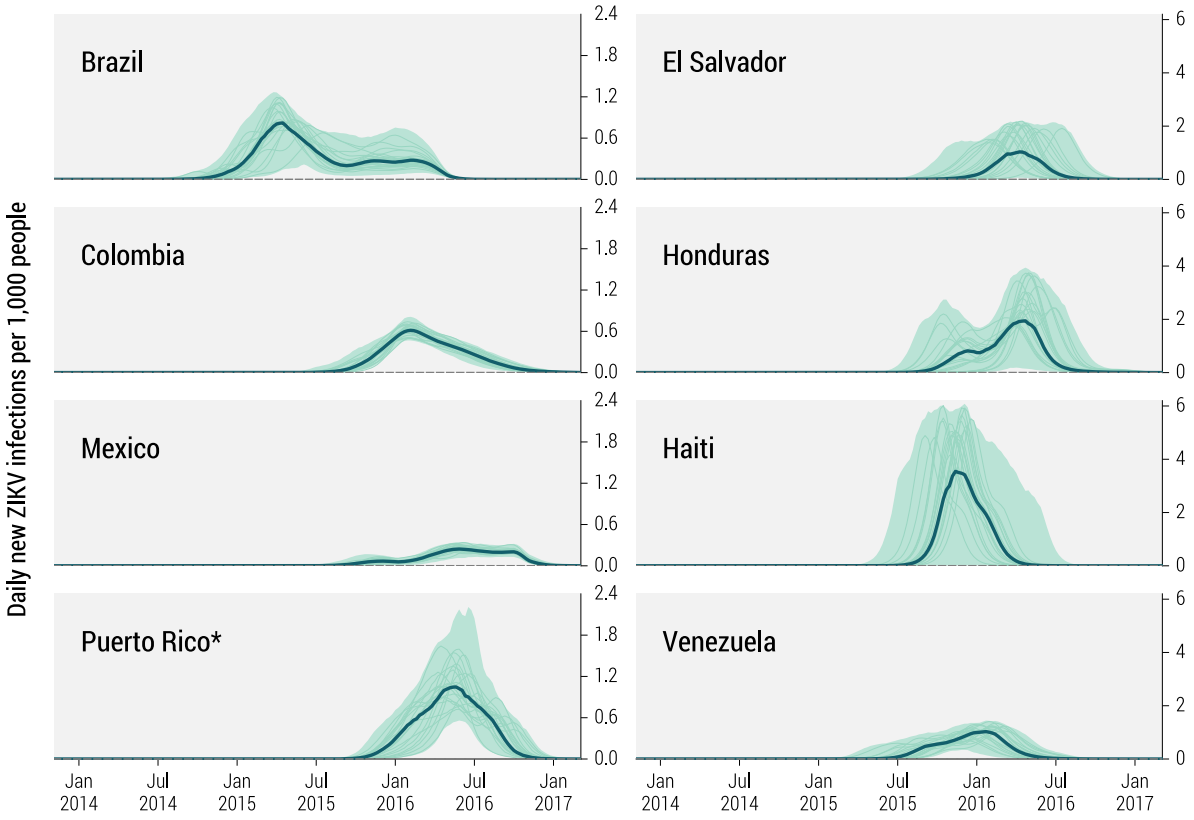


Figure S17: **Long Serial Interval Scenario:** Estimated daily number of new ZIKV infections (per 1000 people) in eight affected countries in the Americas between January 2014 and February 2017. The bold line and shaded area refer to the estimated median number of infections and 95 % CI of the model projections, respectively. Rates include asymptomatic infections. The median incidence is calculated each week from the stochastic ensemble output of the model and may not be representative of specific epidemic realizations. Thin lines represent a sample of specific realizations. *Puerto Rico curves are constrained under the condition that the peak of incidence curve is after March 1, 2016, based on surveillance data [25].

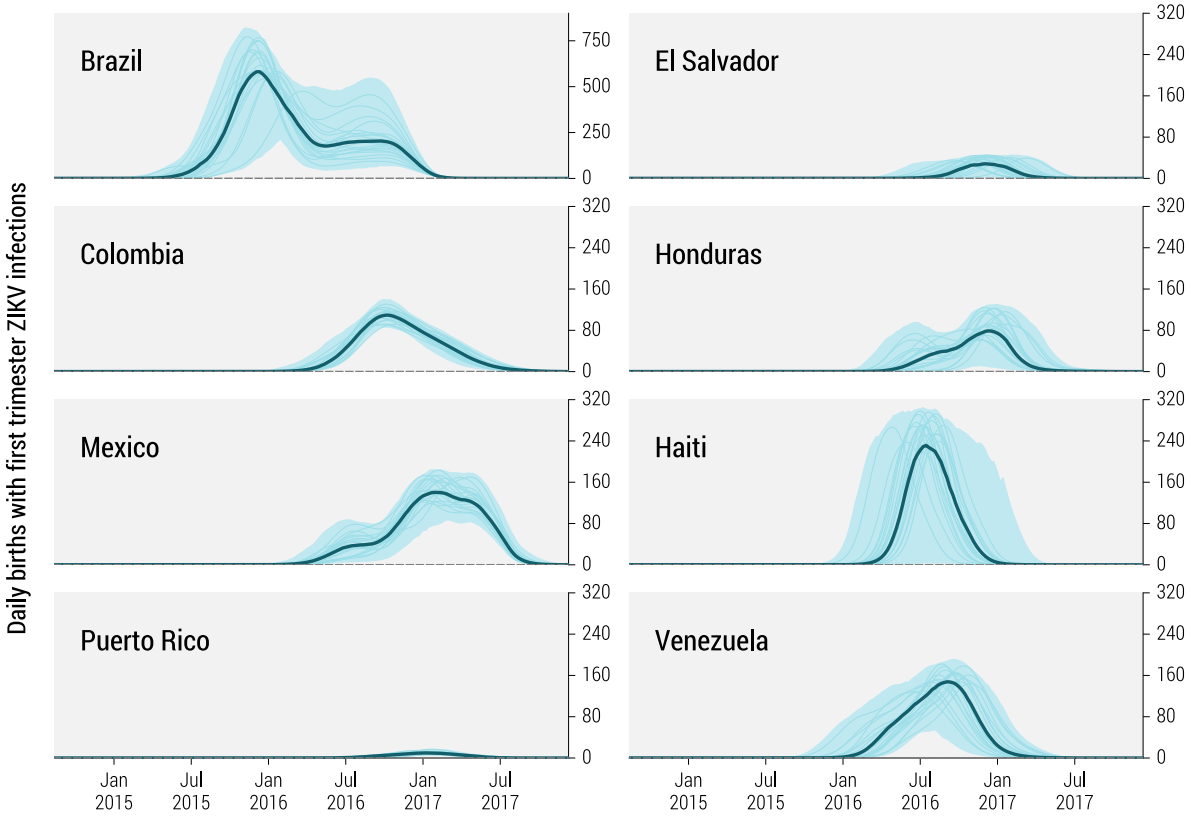


Figure S18: **Long Serial Interval Scenario:** Estimated daily number of births between October 2014 and December 2017 from women infected with ZIKV during the first trimester of pregnancy in eight affected countries in the Americas. The bold line and shaded area refer to the estimated median number of births and 95 % CI of the model projections, respectively. Note that Brazil is plotted with on different scale. The median curve is calculated each week from the stochastic ensemble output of the model and may not be representative of specific epidemic realizations. Thin lines represent a sample of specific realizations.

	Infection AR %		Cumulative Microcephaly Cases (median with 95%CI)					
	(median with 95%CI)		first trimester risk: 0.95%		first trimester risk: 2.19%		first trimester risk: 4.52%	
	Feb. 1, 2016	Feb. 28, 2017	Feb. 1, 2016	Dec. 10, 2017	Feb. 1, 2016	Dec. 10, 2017	Feb. 1, 2016	Dec. 10, 2017
Brazil	17 [15 - 19]	20 [18 - 21]	765 [144 - 1182]	1434 [1318 - 1558]	1763 [332 - 2725]	3306 [3038 - 3591]	3640 [685 - 5623]	6823 [6271 - 7412]
Colombia	5 [3 - 8]	13 [12 - 15]	0 [0 - 3]	245 [220 - 276]	0 [0 - 6]	565 [506 - 636]	1 [0 - 13]	1165 [1045 - 1313]
Mexico	1 [0 - 2]	6 [5 - 7]	0 [0 - 2]	367 [307 - 412]	0 [0 - 5]	845 [708 - 950]	1 [0 - 10]	1744 [1462 - 1961]
Puerto Rico*	2 [0 - 6]	20 [14 - 29]	0 [0 - 0]	19 [13 - 27]	0 [0 - 0]	44 [31 - 63]	0 [0 - 0]	92 [64 - 129]
El Salvador	2 [0 - 17]	20 [16 - 24]	0 [0 - 0]	50 [41 - 60]	0 [0 - 0]	114 [94 - 138]	0 [0 - 1]	236 [194 - 284]
Honduras	10 [0 - 32]	37 [33 - 41]	0 [0 - 1]	153 [135 - 171]	0 [0 - 3]	353 [311 - 395]	0 [0 - 7]	729 [642 - 816]
Haiti	46 [4 - 56]	51 [44 - 58]	0 [0 - 32]	329 [286 - 374]	0 [0 - 74]	759 [660 - 862]	1 [0 - 153]	1567 [1362 - 1779]
Venezuela	14 [6 - 20]	20 [18 - 22]	2 [0 - 61]	289 [257 - 329]	4 [0 - 140]	666 [592 - 758]	8 [0 - 288]	1375 [1222 - 1564]

Figure S19: **Long Serial Interval Scenario:** Projected ZIKV infection ARs through the time of the WHO declaration of a PHEIC on February 1, 2016, and through February 28, 2017, in eight affected countries in the Americas. Median estimates and 95 % CIs are provided. ZIKV attack rates include asymptomatic infections. The denominator is the entire country population, including regions that are not exposed to the vector. Cumulative microcephaly cases due to ZIKV infection during the first trimester of pregnancy through the time of the WHO declaration of a PHEIC on February 1, 2016, and through December 10, 2017, in eight affected countries in the Americas. We consider three different risks of microcephaly associated with ZIKV infection during the first trimester: 0.95% first trimester risk based on a study of the 2013-2014 French Polynesian outbreak [21]; 2.19% (100% over-reporting) and 4.52% (no over-reporting) first trimester risks, based on a study of Bahia, Brazil [26], given a model-estimated 33% infection AR in Bahia. *Puerto Rico curves constrained under the condition that the peak of ZIKV incidence curve is after March 1, 2016, based on surveillance data [25].

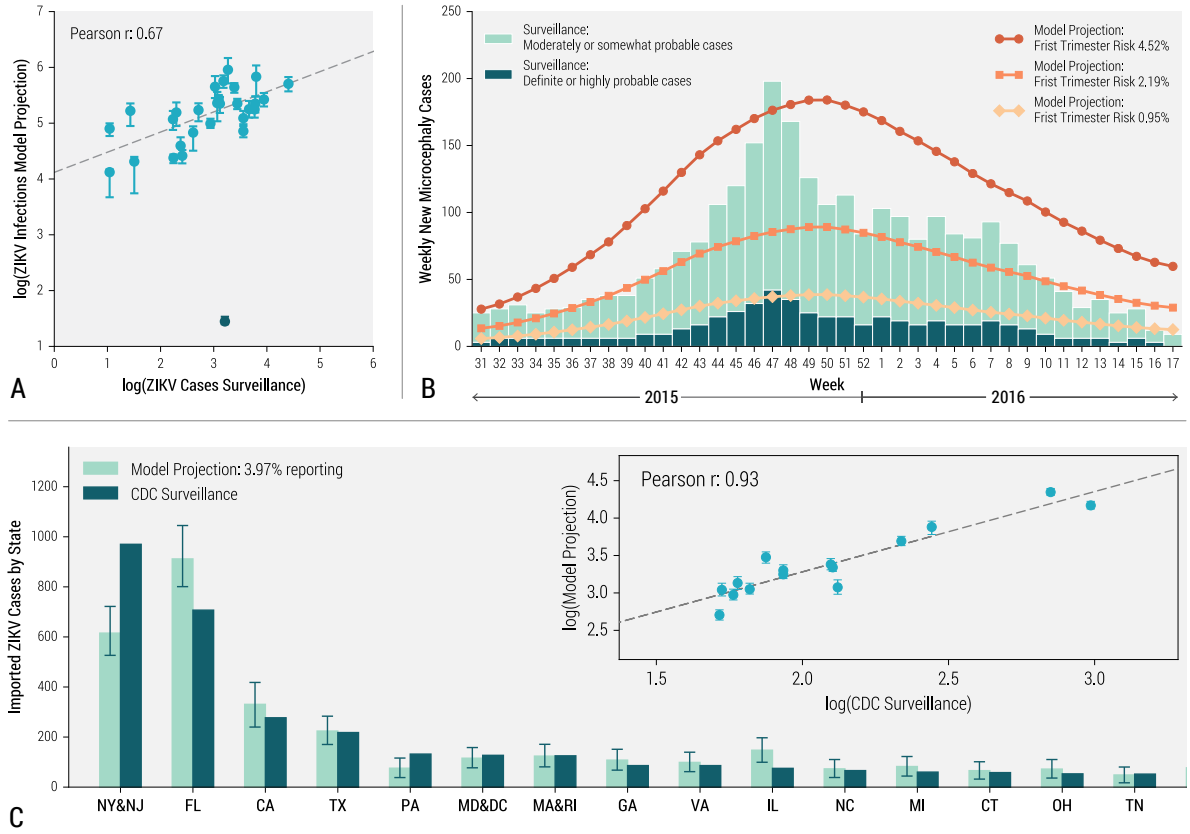


Figure S20: **Long Serial Interval Scenario:** A) Correlation between the number of ZIKV cases by state in Colombia as reported by surveillance data through October 1, 2016 [27], compared with state-level model projections of infections (median with 95% CI). Pearson's r correlation coefficient is reported for the linear association on the log scale. The outlier (in dark green) excluded from the statistical analysis corresponds to the Arauca region. B) Timeline of microcephaly cases in Brazil though April 30, 2016. Bar plot shows weekly definite (or highly probable cases) and moderately (or somewhat probable cases) from surveillance data [28]. Line plots indicate estimated weekly new microcephaly cases given three levels of first trimester risk: 4.52% (round) [26], 2.19% (square) [26], and 0.95% (diamond) [21]. C) Bar plot of ZIKV infections imported into the continental USA by state(s) as reported by CDC surveillance through October 5, 2016 [29], and compared to model projections (median with 95% CI) for the same period assuming 5.74% reporting/detection. The insert shows the correlation between CDC surveillance data and model projections (median with 95% CI).

7 Counterfactual seasonality scenarios

To illustrate how seasonality affects Zika epidemic in terms of both local transmission and global dissemination, we create two counterfactual scenarios with unrealistic seasonal patterns and we compare the ZIKV transmission dynamics with the reference scenario that instead uses real-world seasonality pattern. The detailed settings of the two counterfactual scenarios are as follows:

- **Counterfactual Scenario One (CS1)**: we set the daily temperature pattern of the entire Brazil to be the same as Sao Paulo (Brazil), in which the temperature variation throughout the year significantly limits ZIKV transmission feasibility during winter. The rest of the world maintains its original temperature pattern. This is a lower-bound scenario that illustrates how unsuitable climate in Brazil limits ZIKV epidemics in the Americas in terms of both timing and magnitude of the epidemic. All the other elements of the model are otherwise kept the same as in the Reference Scenario (RS) detailed in the main article.

- **Counterfactual Scenario Two (CS2)**: the daily temperature pattern of the entire world is set to be the same as Fortaleza (Brazil), whose tropical climate allows ZIKV transmission all year long. This is an upper-bound scenario to illustrate how suitable climate facilitate the spread of Zika, providing unrealistic patterns when compared to reported data. All the other elements of the model are otherwise kept the same as in the Reference Scenario (RS) detailed in the main article.

For each counterfactual scenario, a total of 15,000 simulations were performed with the time of introduction in Brazil on November 15, 2013 (in agreement with phylogenetic studies and posterior estimation of the RS) and seeding locations as in the reference scenario. Figure S21 shows the cumulative number of ZIKV infections per 1000 people in Brazil for CS1, CS2 and RS. Here we consider only simulations with outbreaks in Brazil (>1000 cases total in Brazil). CS1 (lower bound scenario, red color in figure) has a slower growth rate at the beginning of the epidemic, and a much lower overall attack rate when compared to the RS. CS2 (upper bound scenario, yellow color in figure), in contrast, has a large growth rate at the beginning of the epidemic and higher overall country attack rate. This is in agreement with the climate settings of the two hypothetical scenarios.

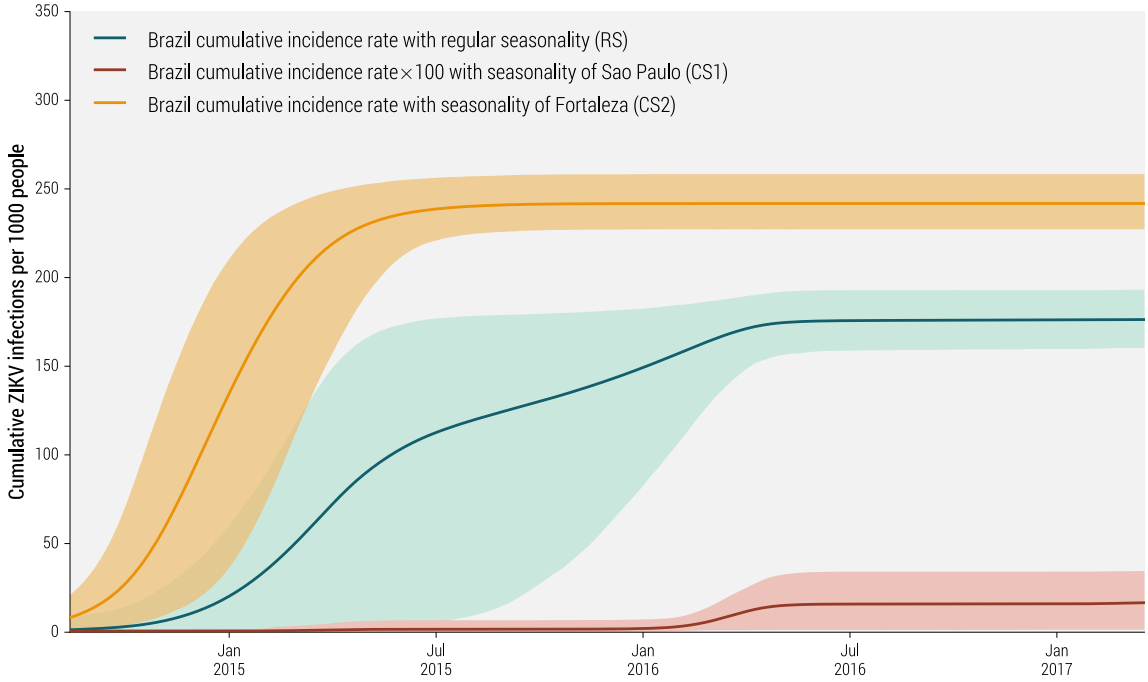


Figure S21: **Conterfactual scenarios:** cumulative number of ZIKV infections per 1000 people in Brazil for CS1, CS2 and RS

8 Additional validation tests

In this section, we provide additional model validation tests based on surveillance data of travel associated ZIKV cases among European countries [30] and state level microcephaly cases in Brazil [31].

Figure S22 shows the correlation between the number of imported ZIKV infections from model projection (reference scenario) and reported travel-associated ZIKV cases from ECDC surveillance by November 2016. The Pearson correlation coefficient is $r = 0.89$ ($p < 0.01$), indicating that numerical results are in good agreement with observations.

Figure S23 shows the correlation of the model-projected number of births with first trimester ZIKV infections and the number of suspected and confirmed microcephaly cases from surveillance data of different states in Brazil. Based on the model projection, birth defects related to Zika have the highest concentration in Northeast region of Brazil, followed by Southeast, North, Central-West, and South. This is in agreement with the spatial distribution of microcephaly cases observed throughout Brazil.

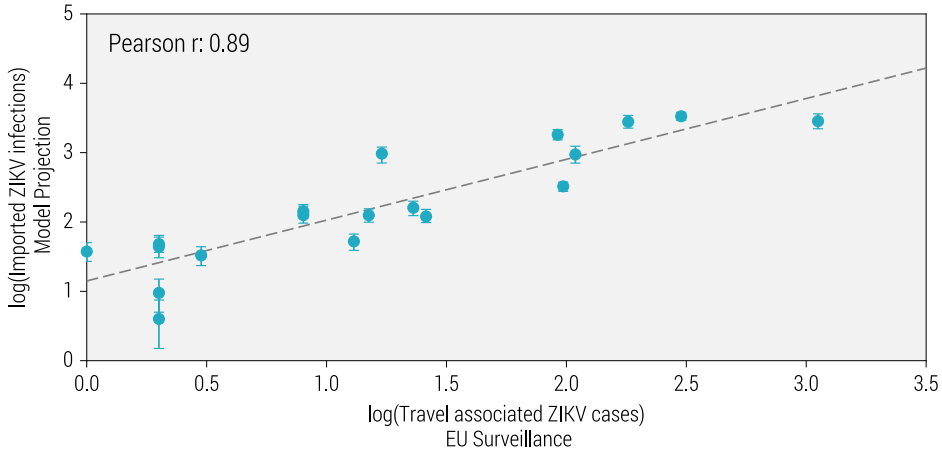


Figure S22: **Travel associated ZIKV cases:** Correlation between imported ZIKV infections (median with 95% CI) projected by model and travel-associated ZIKV cases reported by ECDC surveillance, through November, 2016[30]. Countries with reported travel-associated ZIKV cases includes Austria, Belgium, Czech Republic, Denmark, Finland, France, Greece, Hungary, Ireland, Italy, Luxembourg, Malta, the Netherlands, Norway, Portugal, Romania, Spain, Sweden, the United Kingdom.

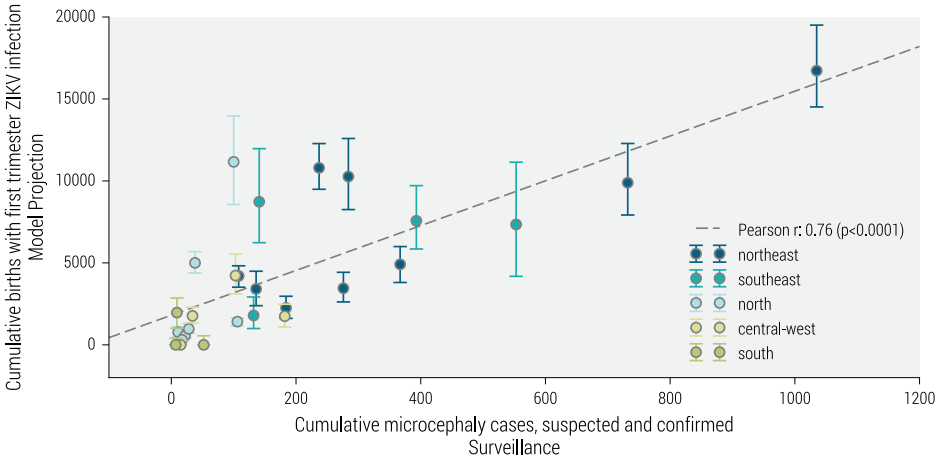


Figure S23: **Microcephaly cases in Brazil:** Correlation of cumulative suspected and confirmed microcephaly cases by state in Brazil as reported by surveillance data [31] through November 19, 2016, compared with state-level model projections of births with first trimester ZIKV infections (median with 95% CI) during the same time window.

9 Incidence map of ZIKV infections

In Figure S24 we provide a spatial projection of the cumulative median number of ZIKV infections, according to the reference scenario, by February 28, 2017 at a spatial resolution of 1×1 km in Latin America and the Caribbean. Each 1×1 km cell is colored according to the median number of ZIKV infections within the cell. It worth noticing the close similarity of our spatial projection with the analogous map obtained by Perkins et al.[11] by using a different methodology.

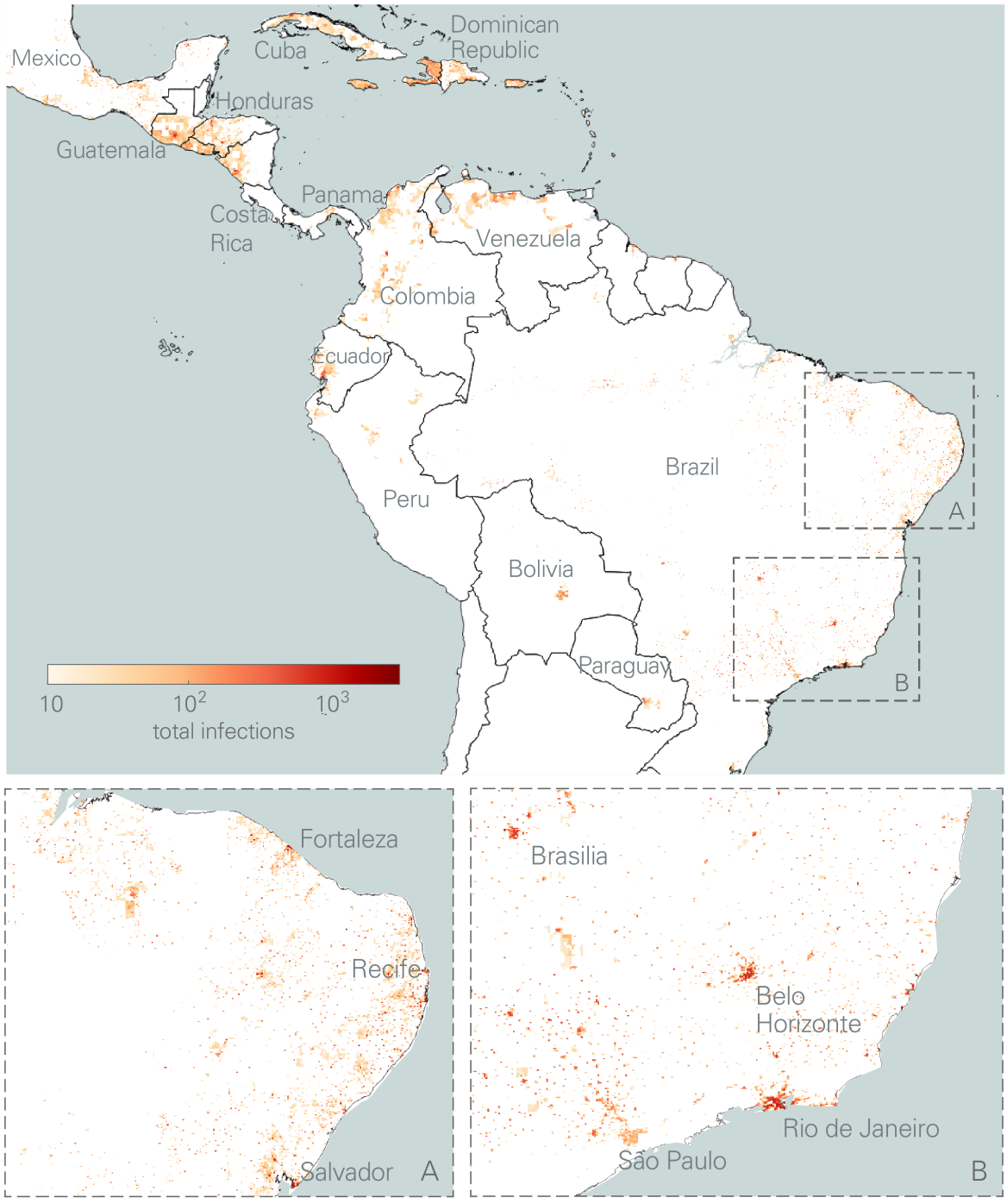


Figure S24: **Incidence map of ZIKV infections:** A spatial projection (reference scenario) of median number of ZIKV infections by February 28, 2017 at a spatial resolution of 1×1 km in Latin America and the Caribbean. In the insets A) and B) we provide detailed projections fro the areas of Recife and Belo Horizonte, Brazil, respectively.

References

[1] Kucharski AJ, Funk S, Eggo RM, Mallet HP, Edmunds WJ, et al. (2016) Transmission Dynamics of Zika Virus in Island Populations: A Modelling Analysis of the 2013–14 French Polynesia Outbreak. PLoS Negl Trop Dis 10: e0004726.

[2] Keeling MJ, Rohani P (2008) Modeling infectious diseases in humans and animals. Princeton University Press.

[3] Johansson MA, Powers AM, Pesik N, Cohen NJ, Staples JE (2014) Nowcasting the spread of chikungunya virus in the Americas. PLoS ONE 9: e104915.

[4] Lambrechts L, Paaijmans KP, Fansiri T, Carrington LB, Kramer LD, et al. (2011) Impact of daily temperature fluctuations on dengue virus transmission by aedes aegypti. Proceedings of the National Academy of Sciences 108: 7460–7465.

[5] Scott TW, Amerasinghe PH, Morrison AC, Lorenz LH, Clark GG, et al. (2000) Longitudinal studies of aedes aegypti (diptera: Culicidae) in thailand and puerto rico: blood feeding frequency. Journal of medical entomology 37: 89–101.

[6] Liu-Helmersson J, Stenlund H, Wilder-Smith A, Rocklöv J (2014) Vectorial capacity of aedes aegypti: effects of temperature and implications for global dengue epidemic potential. PloS one 9: e89783.

[7] Centre d'hygiène et de salubrité publique (2014). Surveillance de la dengue et du zika en polynèsie française. URL `http://www.hygiene-publique.gov.pf/IMG/pdf/bulletin_dengue_28-03-14.pdf`. (Last visited: November 16, 2016).

[8] Mallet HP MD Vial AL (2015). Bilan de l'épidémie a virus zika en polynésie francaise 2013–2014. bulletin d'information sanitaires, epidemiologiques et statistiques. URL `http://www.hygiene-publique.gov.pf/IMG/pdf/no13_-_mai_2015_-_zika.pdf`. (Last visited: November 16, 2016).

[9] Bretó C, He D, Ionides EL, King AA (2009) Time series analysis via mechanistic models. Ann Appl Stat 3: 319–348.

[10] Kraemer M, Sinka M, Duda K, Mylne A, Shearer F, et al. (2015) The global distribution of the arbovirus vectors Aedes aegypti and Ae. albopictus. Elife 4: e08347.

[11] Perkins AT, Siraj AS, Ruktanonchai CW, Kraemer MUG, Tatem AJ (2016) Model-based projections of Zika virus infections in childbearing women in the Americas. Nature Microbiology 1: 16126.

[12] Sissoko D, Moendandze A, Malvy D, Giry C, Ezzedine K, et al. (2008) Seroprevalence and risk factors of chikungunya virus infection in Mayotte Indian Ocean, 2005-2006: A population-based survey. PLoS ONE 3: e3066.

[13] Reiter P, Lathrop S, Bunning M, Biggerstaff B, Singer D, et al. (2003) Texas lifestyle limits transmission of dengue virus. Emerg Infect Dis 9: 86–89.

[14] Sergon K, Njuguna C, Kalani R, Ofula V, Onyango C, et al. (2008) Seroprevalence of chikungunya virus (CHIKV) infection on Lamu Island, Kenya, October 2004. Am J Trop Med Hyg 78: 333-7.

[15] Sergon K, Yahaya A, Brown J, Bedja S, Mlindasse M, et al. (2007) Seroprevalence of chikungunya virus infection on Grande Comore Island, union of the Comoros, 2005. Am J Trop Med Hyg 76: 1189-93.

[16] Schwarz N, Girmann M, Randriamampionona N, Bialonski A, Maus D, et al. (2012) Seroprevalence of antibodies against chikungunya, dengue, and Rift Valley fever viruses after febrile illness outbreak, Madagascar. Emerg Infect Dis 18: 1780-6.

[17] Renault P, Solet J, Sissoko D, Balleydier E, Larrieu S, et al. (2007) A major epidemic of chikungunya virus infection on Reunion Island, France, 2005-2006. Am J Trop Med Hyg 77: 727-31.

[18] van Genderen FT, Krishnadath I, Sno R, Grunberg MG, Zijlmans W, et al. (2016) First chikungunya outbreak in Suriname; clinical and epidemiological features. PLoS Negl Trop Dis 10: e0004625.

[19] Gay N, Rousset D, Huc P, Matheus S, Ledrans M, et al. (2015) Seroprevalence of Asian lineage chikungunya virus infection on Saint Martin island 7 months after the 2013 emergence. American Journal of Tropical Medicine and Hygiene 94: 393–396.

[20] Moro ML, Gagliotti C, Silvi G, Angelini R, Sambri V, et al. (2010) Chikungunya virus in North-Eastern Italy: a seroprevalence survey. American Journal of Tropical Medicine and Hygiene 82: 508–511.

[21] Cauchemez S, Besnard M, Bompard P, Dub T, Guillemette-Artur P, et al. (2016) Association between Zika virus and microcephaly in French Polynesia 2013–15: a retrospective study. The Lancet 387: 2125–2132.

[22] The World Bank (2016). Birth rate, crude (per 1000 people). URL `http://data.worldbank.org/indicator/SP.DYN.CBRT.IN`. (Last visited: November 16, 2016).

[23] Poletti P, Messeri G, Ajelli M, Vallorani R, Rizzo C, et al. (2011) Transmission potential of chikungunya virus and control measures: the case of Italy. PLoS One 6: e18860.

[24] Delatte H, Gimonneau G, Triboire A, Fontenille D (2009) Influence of temperature on immature development, survival, longevity, fecundity, and gonotrophic cycles of Aedes albopictus, vector of chikungunya and dengue in the Indian Ocean. J Med Entomol 46: 33–41.

[25] Health Ministry of Puerto Rico (2017). Informes de Zika por Municipios. `http://www.salud.gov.pr/Estadisticas-Registros-y-Publicaciones/Informes%20Zika%20por%20Municipios/Informe_Municipios%20semana%201-2017.pdf` (Last visited: February 10, 2017).

[26] Johansson MA, y Teran-Romero LM, Reefhuis J, Gilboa SM, Hills SL (2016) Zika and the Risk of Microcephaly. New England Journal of Medicine 375: 1–4.

[27] Instituto Nacional de Salud – Dirección de Vigilancia y Análisis del Riesgo en Salud Pública. BOLETIN EPIDEMIOLOGICO SEMANAL. URL http://www.ins.gov.co/boletin-epidemiologico/Boletn%20Epidemiolgico/2016%20Boletin%20epidemiologico%20semana%2039.pdf. (Last visited: November 16, 2016).

[28] França GVA, Schuler-Faccini L, Oliveira WK, Henriques CMP, Carmo EH, et al. Congenital zika virus syndrome in brazil: a case series of the first 1501 livebirths with complete investigation. The Lancet 388: 891–897.

[29] Center for Disease Control. Case Counts in the US. URL https://www.cdc.gov/zika/geo/united-states.html. (Last visited: November 16, 2016).

[30] European Center for Disease Prevention and Control. Surveillance Atlas of Infectious Diseases, Zika virus infection: reported cases. URL http://atlas.ecdc.europa.eu/public/index.aspx?Instance=GeneralAtlas. (Last visited: January 1, 2017).

[31] Secretaria de vigilância em Saúde. Informe epidemiológico N53. Monitoramento dos casos de microcefalia no Brasil. URL http://portalsaude.saude.gov.br/index.php/o-ministerio/principal/secretarias/svs. (Last visited: February 1, 2017).