Projected spread of Zika virus in the Americas

Qian Zhang[1], Kaiyuan Sun[1], Matteo Chinazzi[1], Ana Pastore-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

Abstract

We use a data-driven global stochastic epidemic model to project past and future spread of the Zika virus (ZIKV) in the Americas. The model has high spatial and temporal resolution, and integrates real-world demographic, human mobility, socioeconomic, temperature, and vector density data. We estimate that the first introduction of ZIKV to Brazil likely occurred between August 2013 and April 2014. We provide simulated epidemic profiles of incident ZIKV infections for several countries in the Americas through December 2016. The ZIKV epidemic is characterized by slow growth and high spatial and seasonal heterogeneity, attributable to the dynamics of the mosquito vector and to the characteristics and mobility of the human populations. We project the expected timing and number of cases of microcephaly assuming three levels of risk associated with ZIKV infection during the first trimester of pregnancy. Our approach represents an early modeling effort aimed at projecting the potential magnitude and timing of the ZIKV epidemic that might be refined as new and more accurate data from the region will be available.

Introduction

The Zika virus (ZIKV) is an RNA virus from the Flaviviridae family, genus Flavivirus [1, 2], first isolated in the Zika Forest of Uganda in 1947 [3]. It results in a mild disease characterized by low

grade fever, rash, and/or conjunctivitis, though only approximately 20% of those infected are symptomatic[4]. Although there have been instances of sexual and perinatal transmission[5, 6, 7] and the potential for transmission by transfusion is present [8], ZIKV is spread primarily through *Aedes* mosquitoes [9, 10].

Until recently, ZIKV was considered a neglected tropical disease with only local outbreaks [11, 12, 13, 14]. In February 2015, the Brazilian Ministry of Health noted an unknown exanthematic disease spreading through six of its northeastern states [15]. The northeastern region of Brazil is characterized by high mosquito densities and dengue virus (DENV) transmission throughout the year. In mid-May 2015, ZIKV was laboratory confirmed in multiple samples, establishing the presence of autochthonous transmission in Brazil. In September 2015, Brazil began to observe an increase in cases of microcephaly, a rare but serious neurologic birth defect resulting in a smaller than normal head [16]. On the basis of the reported microcephaly case cluster in Brazil and the previous association of ZIKV with Guillain-Barré syndrome (a post-infection autoimmune disease that affects the peripheral nervous system), the Director-General of the World Health Organization (WHO) declared a Public Health Emergency of International Concern (PHEIC) in February 2016 [17]. ZIKV has continued to spread throughout the Americas, with 40 countries and territories in the region reporting autochthonous transmission, excluding only Chile and Uruguay [18, 19]. Countries in the Americas reporting the highest incidence of ZIKV disease are also countries with historically endemic dengue transmission. As of recently, other countries with ZIKV outbreaks besides Brazil have reported cases of microcephaly and other birth defects associated with ZIKV infection during pregnancy (Zika Congenital Syndrome) [20].

While enhanced surveillance and new data improve our understanding of ZIKV, many unknowns persist. There is uncertainty surrounding the time of introduction of the virus to the region, though early epidemiological and genetic findings estimate that ZIKV arrived in Brazil between May and December 2013 [21]. Furthermore, little is known about the spread of the virus in 2014 and 2015, prior to the WHO's alert in early 2016. Using a data-driven stochastic and spatial epidemic model, we analyze existing reports to gain insight into the first introduction into the region and early epidemic growth. We use the model to project future spatiotemporal spread in the Americas through December 2016, accounting for seasonal environmental factors and detailed population-level data. We also provide ranges for the expected number of microcephaly cases per country.

Results

Introduction of ZIKV to the Americas. The precise location and start date of the outbreak are unknown. We identified twelve major transportation hubs in areas related to major events, such as the Soccer Confederations Cup in June 2013 and the Soccer World Cup in June 2014, both held in Brazil, and assumed a prior probability of introduction proportional to the daily passenger flow to each hub. We then considered introduction dates between April 2013 and August 2014, including the time frame suggested by phylogenetic and molecular clock analyses [21] through the 2014 Soccer World Cup. Using Latin square sampling over this two-dimensional space, we calculated the likelihood of replicating the observed epidemic peak in Colombia ($\pm$ 1 week), as reported by the Pan-American Health Organization (PAHO) [22], and the resulting posterior density of each location and date combination. The Colombian epidemic was used to calibrate this analysis because of the large number of cases observed and overall consistency in reporting.

In **Fig. 1A** we plot the posterior likelihood as a function of introduction date and location, and in **Figs. 1B** and **1C** we plot the marginal distributions. The greatest posterior density is associated with an introduction to Rio de Janeiro in November 2014. The 90% credible interval for the most likely date extends from August 2013 to April 2014, with its mode in November 2013. The most likely locations of ZIKV introduction, in descending order, are Brasilia (central), Rio de Janeiro (southeast), Salvador (northeast), and Fortaleza (northeast). While Rio de Janeiro experiences the greatest passenger flow, the city also experiences more seasonality in mosquito density, making its likelihood to seed an epidemic sensitive to introduction date. The cities localized in the northeast of Brazil have less passenger flow compared to Rio de Janeiro but they have higher mosquito densities and dengue transmission all year long. Brasilia, in comparison, has little seasonality in mosquito density and high traffic flow, though the area has low DENV transmission.

Spatiotemporal ZIKV spread. Stochastic realizations satisfying the observed peak in Colombia define the model output used to estimate the spatiotemporal pattern of ZIKV spread in the Americas through December 2016. In **Fig. 2** we plot the simulated epidemic profiles of incident ZIKV infections for several countries in the Americas, and in **Tab. 1** we report the associated infection attack rates through February 1, 2016, when the WHO declared a PHEIC, and through December 31, 2016. Estimates for additional countries in the Americas are available at the fol-

lowing webpage: `www.zika-model.org`. The earliest epidemic is observed in Brazil, followed by Haiti, Venezuela and Colombia. The model projects that the epidemics in most of the countries will begin to decline around July 2016.

National infection attack rates are projected to be highest in Haiti, El Salvador, and Honduras and lowest in Colombia and Mexico, though the latter two countries have much larger populations and more heterogeneity in mosquito density and vector borne diseases transmission. For example, nearly half of Colombia's population resides in areas of high altitude where sustained ZIKV transmission is not possible. Even though El Salvador, and Honduras have high infection attack rates, they have comparatively small epidemics on an absolute scale, at least relative to Brazil and Colombia.

Because of the model's fine spatial and temporal resolution, we are able to observe significant variability in the ZIKV reproduction number R_0 across locations and even within the same location at different times. These differences are driven by temperature, the vector presence distribution, and socioeconomic factors, among other variables. (Additional details are provided in the Materials and Methods.) In **Fig. 3** we plot R_0 in a number of areas at different times throughout the year. Equatorial regions experience less seasonality than non-Equatorial regions, where the changes in temperature have a strong impact on the mosquito population and, thus, R_0. Large areas with unexposed populations are visible, such as in high altitude regions of Colombia.

Projected microcephaly cases. Using the epidemic profiles generated by the model, we estimate the number of ZIKV infections in childbearing women and the associated number of microcephaly cases. Assuming that risk is concentrated during the first trimester of pregnancy [23], we adopt three different probabilities of microcephaly estimated in the literature. The first estimate of the risk of microcephaly for ZIKV infected pregnancies is 0.95% (95% CI $[0.34\% - 1.91\%]$) from a retrospective study in French Polynesia [24]. The remaining two are from a study in Bahia, Brazil [25]. Given an overall ZIKV infection attack rate of 28% (95% CI $[26\% - 30\%]$) in Bahia through February 2016, as determined by our model, the estimated first trimester risks are 2.19% (95% CI $[1.98\% - 2.41\%]$), assuming 100% over-reporting of microcephaly cases, and 4.52% (95% CI $[4.10\% - 4.96\%]$), assuming no over-reporting. The estimations do not account for miscarriages or other complications occurring during pregnancy.

In **Fig. 4** we plot the daily number of births through October 2017 from women infected with ZIKV during the first trimester of pregnancy in several countries. The curves closely resemble

the epidemic profiles in **Fig. 2** shifted forward in time around nine months. In **Tab. 1** we report the projected cumulative number of microcephaly cases through February 1, 2016, and through December 31, 2016, assuming three levels of first trimester risk. By the time the WHO declared a PHEIC, Brazil was the only country with a significant (>100) number of ZIKV-attributable microcephaly cases, with cases expected to appear through July 2017. For Colombia, the model projects cases as early as February 2016 with a highest peak in October 2016. In Venezuela, the peak in microcephaly cases is projected to occur in July 2016, continuing through March 2017. Mexico, Puerto Rico, El Salvador and Honduras are projected to experience two peaks of cases, including July-August 2016 and then again starting in 2017.

Sensitivity to mosquito vector. Simulations reported here consider both *Ae. aegypti* and *Ae. albopictus* as competent ZIKV vectors, though less is known about the vectorial capacity of *Ae. albopictus* to transmit. A sensitivity analysis considering *Ae. aegypti* as the only competent vector is provided in the Supplementary Materials with all figures and tables replicated for this scenario. Overall, results are similar as transmission due to *Ae. aegypti* increases to compensate for the absence of the other vector. Differences in the infection attack rates, however, are observed in areas where *Ae. albopictus* is the most common or the only vector. For example, the infection attack rate in Brazil through December 31, 2016, decreases from 15% (95% CI [13% − 16%]) to 13% (95% CI [11% − 14%]) if we consider only *Ae. aegypti*. During the same time period, the infection attack rate in Mexico decreases from 4% (95% CI [2% − 5%]) to 2% (95% CI [1% − 3%]), representing a similar absolute reduction but larger relative reduction than in Brazil.

Model validation. Several approaches were used to independently validate the model with data not used for calibration, as shown in **Fig. 5**. In **Fig. 5A**, we compare model-based projections of the number of ZIKV infections by states in Colombia with observed surveillance data through June 5, 2016. As expected for a typically asymptomatic or mild disease, the model projects a much larger number of infections than that captured by surveillance, suggesting a reporting and detection rate of $1\% \pm 0.8\%$ (from linear regression fit analysis). However, the observed data and model estimates are well-correlated (Pearson's $r = 0.57$, $p = 0.0026$), replicating the often several orders of magnitude differences in infection burden across states within the same country.

In **Fig. 5B** we compare observed data on weekly counts of microcephaly cases reported in Brazil through April 30, 2016 [26] with estimates from the model for each projected level of

microcephaly risk given first trimester ZIKV infection. The three model projection curves vary in magnitude but replicate peaks consistent with the observed data. As the fraction of cases that are confirmed in Brazil is relatively low, it is not possible to identify the most likely level of risk, though the plot suggests that the risk might exceed the lowest estimate of 0.95% [24].

As the computational approach explicitly simulates the number of daily airline passengers traveling globally, the microsimulations allow us to track ZIKV infections imported into countries with no autochthonous transmission. In **Fig. 5C** we plot the number of importations into states in the continental United States (USA) through June 15, 2016, as reported by the Centers for Disease Control and Prevention (CDC) and compare these results with model projections. As the detection rate of ZIKV infections is very low, there are significantly fewer reported infections than projected; we estimate through a linear regression fit that $1.94 \pm 0.36\%$ of imported infections are detected. Nonetheless, the model projections are highly correlated with the observed data (Pearson's $r = 0.96$, $p = 10^{-8}$, as shown in the inset of **Fig. 5C**). A further validation of the model is provided by the reported number of pregnant women infected with ZIKV in the USA. A high detection rate is expected in this closely monitored population. As of June 16, 2016, 265 pregnant women in the USA were laboratory-confirmed for ZIKV, all of whom were imported cases. Since pregnant women comprise approximately 1% of incoming airline traffic flow from the rest of the Americas [27], one can roughly estimate $26,500$, which is consistent with our modeling results projecting $29,898$ (95% CI $[23,643 - 35,663]$) infections imported into the USA by mid-June 2016. These results might be relevant ZIKV risk assessment in USA [28, 29].

Discussion

We use computational modeling to reconstruct the past and project the future spatiotemporal spread of ZIKV in the Americas. To identify the likely date and location of ZIKV's first introduction to the Americas, posterior densities were estimated for twelve major travel hubs in Brazil over a range of dates. The marginal posterior distributions suggest an introduction between August 2013 and April 2014 in a number of potential locations, including Salvador, Rio de Janeiro, Recife and Fortaleza. This date range partially overlaps with that suggested by a recent phylogenetic analysis [21], though ours includes later potential introductions.

We provide epidemic curves for incident ZIKV cases in eight affected countries in the Americas. The results obtained are in good agreement with model-based projections achieved with a

different approach as described elsewhere [30]. Though the initial introduction of ZIKV could date back to August 2013, most countries do not experience the first wave of the epidemic until the early months of 2016. Brazil is the only country that appears to have a well-defined first peak in March of 2015, consistent with reports from the northeast region[31]. The model suggests two epidemic waves in Brazil; the first, occurring between January and July 2015, corresponds with early outbreaks in the northeast region (Maranhao, Bahia, and Rio Grande do Norte) which were not recognized as ZIKV until later; the second wave, between January and May 2016, reflects ZIKV spread throughout the rest of the country [32].

The virus also circulated early on in the Caribbean, with ZIKV samples isolated in Haiti at the end of 2014 and a first peak noted in October 2015 [33]. Colombia first isolated ZIKV in October 2015, at which time it spread rapidly from the Caribbean coast to cities infested with *Ae. aegypti* [34]; the model suggests an introduction to Colombia as early as April 2015, potentially overlapping with the Easter holiday, which is a period of high mobility within and between countries. ZIKV transmission in Venezuela has followed a similar trajectory, first isolated in November 2015 and present in all states by July 2016. In general, we see a time lag between the start of ZIKV transmission and detection of ZIKV by country surveillance systems of around six months. Since March 2016, reported cases have declined in both countries, consistent with our model estimates.

Our model estimates ZIKV transmission in El Salvador and Honduras increasing around July 2015. ZIKV was first detected in El Salvador in November 2015 and in Honduras in December 2015 [35, 36] . Though the first ZIKV infection was confirmed in Puerto Rico in the last week of December, 2015 [37], the model estimates ZIKV transmission in Puerto Rico beginning around August 2015 with two distinct waves (August 2015 to March 2016, and June 2016 to January 2017). In Mexico, the first infection was reported to the surveillance system at the end of November 2015 [38], though circulation may have begun in September 2015.

Another prominent feature emerging from the numerical results is the extreme heterogeneity in the infection attack rates across countries. We find almost a tenfold difference between Honduras and Mexico, exhibiting infection attack rates of 32% (95% CI $[28\% - 35\%]$) and 4% (95% CI $[2\% - 5\%]$), respectively. These large differences in infection attack rates, which are observable at finer geographical resolutions as well, stem from variation in climatic factors, mosquito densities, and socioeconomic variables.

The epidemic has moved slowly and is mostly constrained by seasonality in ZIKV transmissibility. Seasonal drivers result in multiple waves across many countries, as projected for

Honduras, Mexico, and Puerto Rico. Notably, though, incidence rates dramatically decrease by the end of 2016 in all countries described. These results suggest that, after the initial wave in an almost fully susceptible population, ZIKV epidemics could settle into the typical seasonal pattern of mosquito borne diseases like DENV. Transmission may be low for several years with a gradual build-up in susceptibility due to births [39]. Precise projection of long-term ZIKV transmission is a subject of future research.

We project the numbers of births from women who were infected with ZIKV during the first trimester of their pregnancy. There is a well-defined time lag between the epidemic curve and this birth curve. Brazil, who likely experienced their first ZIKV epidemic peak in March 2015, observes a sharp rise in microcephaly cases in September 2015, consistent with what was observed in the field. Venezuela also experiences a comparatively early peak, but it is possible that this has not been reported because case detection and reporting are low due to the political climate. Our model estimates correlate with the situation in Colombia where 18 confirmed cases of congenital Zika syndrome have been observed as of July 19, 2016 [20]. Most of the other countries are not expected to observe an increase in microcephaly cases that is distinguishable from the background rate until May through July 2016. Note that the projected number of microcephaly cases estimated by the model varies considerably depending on the assumed first trimester risk, for which only retrospective estimates are available. We also note that with as high as 80% of ZIKV infections being asymptomatic, most of the ZIKV-infected pregnant women giving birth may not have experienced symptoms during pregnancy; thus, clinicians should be cautious before ruling out ZIKV as the cause of birth defects. The presented results however could be used as a baseline to uncover possible disagreement with real data and highlight the need of additional key evidence for our understanding of the link between ZIKV and neurologic birth defects [40].

Available data on the ZIKV epidemic suffer from several limitations. Although the disease has likely been spreading in the Americas since late 2013, infection detection and reporting began much later and likely increased after the WHO's declaration of a PHEIC in February 2016. Case reporting is inconsistent across countries. Furthermore, comparatively few infections are laboratory-confirmed; this presents an additional challenge as symptomatic cases with other etiologies may be misdiagnosed, and asymptomatic infections are almost entirely missed. Once a reliable ZIKV antibody test is available, seroprevalence studies can help determine the full extent of these outbreaks. For external validation, we compare modeling results with data from Brazil, Colombia, and the USA that were not used to calibrate the model. We are able to

replicate relative trends, though we estimate significantly higher absolute numbers, suggesting reporting and detection rates of 1-2%.

The modeling approach presented here is motivated by the need for a rapid assessment of the ZIKV epidemic, and it contains assumptions and approximations unavoidable due to the sparsity of available data. As a result, transmission is modeled assuming ZIKV behaves similarly to DENV and other mosquito-borne diseases, and further research is needed to provide ZIKV-specific parameter estimates. Mosquito presence maps were available from published data but have limitations as detailed in the literature. Sexual and other modes of transmission were not incorporated in the model. The specific socio-economic features of airline travelers was also not included. Finally, we do not model public health interventions to control the vector population or behavioral changes due to increased awareness. These results may change as more information becomes available from ZIKV affected regions to refine the calibration of the model.

Conclusions

The model presented here provides an early methodological framework for the analysis of the global spread of ZIKV. The model captures the slow dynamic of the epidemic characterized by heterogeneity in the infection attack rate as well as the temporal pattern resulting from local weather, population-level characteristics, and human mobility. Although the modeling results should be interpreted cautiously in light of the assumptions and limitations inherent to the approach, the framework emerging from the numerical results may help in the interpretation of data and provide indications of the magnitude and timing of the epidemic, as well as aid in planning for international and local outbreak response.

Materials and Methods

Model summary. To study spatiotemporal ZIKV spread, we use the Global Epidemic and Mobility Model (GLEAM), a previously described individual-based, stochastic and spatial epidemic model [41, 42, 43, 44, 45, 46]. This model integrates high-resolution demographic, human mobility, socioeconomic [47], and temperature data [48]. It was expanded to incorporate data on *Aedes* mosquito density [49] and the association between socioeconomic factors and population risk of exposure [30, 50]. Similar to previous arbovirus modeling approaches [14], we use a compartmental classification of the disease stages in the human and mosquito populations, assigning plausible parameter ranges based on the available ZIKV literature and assumed simi-

larities between ZIKV and DENV.

Global model for the spreading of vector-borne diseases. The GLEAM model is a fully stochastic epidemic modeling platform that utilizes real world data in order to perform in-silico simulations of the spatial spread of infectious diseases at the global level. GLEAM uses population information obtained from the high-resolution population database of the *Gridded Population of the World* project from the Socioeconomic Data and Application Center at Columbia University (SEDAC) [51]. The model considers geographical cells of 0.25×0.25 degree, corresponding to an approximately 25×25 km square for cells along earth's Equator. GLEAM groups cells into subpopulations defined by a Voronoi-like tessellation of the earth's surface centered around major transportation hubs in different urban areas. The model includes over $3,200$ subpopulations in about 230 different countries (numbers vary by year).

Within each subpopulation, a compartmental model is used to simulate the disease of interest. The model uses an individual dynamic where discrete, stochastic transitions are mathematically defined by chain binomial and multinomial processes. Subpopulations interact through the mechanistically simulated mobility and commuting patterns of disease carriers. Mobility includes global air travel [52], and GLEAM simulates the number of passengers traveling daily worldwide using available data on origin-destination flows among indexed subpopulations.

The transmissibility of vector-borne diseases is associated with strong spatial heterogeneity, driven by variability in vector abundance and characteristics of the exposed populations. Many locations, such as those at high elevation, are not at risk of autochthonous ZIKV transmission simply because the vector is absent. In other locations, the vector may be present but sustained transmission is not possible because of environmental factors that affect the vector's population dynamics, such as temperature or precipitation. Housing conditions, availability of air-conditioning, and socioeconomic factors also contribute significantly to determining the fraction of the population likely exposed to the vector. To extend the GLEAM model to simulate vector-borne diseases, a number of new data sets with high spatial resolution were integrated, including the following:

- *Global terrestrial air temperature data:* The global air temperature data set [48] contains monthly mean temperatures at a spatial resolution of 0.5×0.5 degree. To match the spatial resolution of GLEAM's gridded population density map, the temperature for each population cell is extracted from the nearest available point in the temperature data set. Daily average temperatures were linearly interpolated from each population's monthly

averages.

- *Global Ae. aegypti and Ae. albopictus distribution:* The Global *Ae. aegypti* and *Ae. albopictus* distribution database provides uncertainty estimates for the vector's distribution at a spatial resolution of 5×5 km [49].

- *Geolocalized economic data:* The geophysically scaled economic dataset (G-Econ), developed by Nordhaus *et al* [53], maps the per capita Gross Domestic Product (computed at Purchasing Power Parity exchange rates) at a 1×1 degree resolution. To estimate the per capita Gross Cell Product (GCP) at Purchasing Power Parity (PPP) rates, the amount is distributed across GLEAM cells proportionally to each cell's population size. The data have also been rescaled to reflect 2015 GDP per capita (PPP) estimates.

These databases are combined to model the key drivers of ZIKV transmission, as illustrated in combination with necessary parameters in **Fig. 6**. Temperature impacts many important disease parameters, including the time- and cell-specific values of R_0 whose variation induces seasonality and spatial heterogeneity in the model. Temperature data are also used together with the mosquito presence distribution data to define the daily mosquito abundance (number of mosquitoes per human) in each cell as detailed in Sec. 2 of the Supplementary Materials. Data on mosquito abundance and temperature are used to identify cells where ZIKV outbreaks are not possible because of environmental factors. The human population in these cells is thus considered unexposed to ZIKV and susceptible individuals are assigned an environmental rescaling factor, r_{env}, as described in Sec.3 of the Supplementary Materials. Finally, we use historical data and G-Econ to provide a socioeconomic rescaling factor, r_{se}, reflecting how exposure to the vector is impacted by socioeconomic variables such as availability of air-conditioning; the derivation of this rescaling factor is provided in Sec.3 of the Supplementary Materials.

Once the data layers and parameters have been defined, the model runs using discrete time steps of one full day to run the transmission dynamic model (described in detail below), incorporate human mobility between subpopulations, and partially aggregate the results at the desired level of geographic resolution. The model is fully stochastic and from any nominally identical initialization (initial conditions and disease model) generates an ensemble of possible epidemics, as described by newly generated infections, time of arrival of the infection in each subpopulation, and the number of traveling carriers. The latin square sampling of the initial introduction of ZIKV in Latin America and the ensuing statistical analysis is performed on $150,000$ stochastic epidemic realizations. The sensitivity analysis considers a total of $200,000$ more realizations.

ZIKV transmission dynamics. **Fig. 7A** describes the compartmental classifications used to simulate ZIKV transmission dynamics. Humans can occupy one of four compartments: susceptible individuals S^H who lack immunity against the infection, exposed individuals E^H who have acquired the infection but are not yet infectious, infected individuals I^H who can transmit the infection (may or may not display symptoms), and removed individuals R^H who no longer have the infection. We consider the human population size to be constant, *i.e.* $S^H+E^H+I^H+R^H = N^H$. The mosquito vector population is described by the number of susceptible S^V, exposed E^V, and infectious mosquitoes I^V. The transmission model is fully stochastic. The transitions across compartments, the human-to-mosquito force of infection, and the mosquito-to-human force of infection are described by parameters that take into account the specific abundance of mosquitoes and temperature dependence at the cell level. Exposed individuals become infectious at a rate ϵ_H which is inversely proportional to the mean intrinsic latent period of the infection [54]. These infectious individuals then recover from the disease at a rate μ_H [14] which is inversely proportional to the mean infectious period. The mosquito-to-human force of infection follows the usual mass action law and is the product of the number of mosquitoes per person, the daily mosquito biting rate and specific ZIKV infection transmissibility per day, the mosquito-to-human transmission rate [55], and the number I^V of infected mosquitoes. Exposed mosquitoes transition to the infectious class at a rate ϵ_V that is inversely proportional to the mean extrinsic latent period in the mosquito population [2]. Susceptible, exposed, and infectious mosquitoes all die at a rate that is inversely proportional to the mosquito lifespan, μ_V [56]. The mosquito-to-human force of infection follows the usual mass action law. A full description of the stochastic model and the equations is provided in the Supplementary Materials.

A summary of the parameters defining the disease dynamic is reported in **Fig. 7B**. The empirical evidence related to the ZIKV disease in both human and mosquito populations is fairly limited at the moment. We have performed a review of the current studies of ZIKV and collected plausible ranges for these parameters. As in other studies, we have assumed that the drivers of ZIKV transmission are analogous to those of DENV. In particular, we have considered that mosquito lifespan, mosquito abundance, and the transmission probability per bite depend on the temperature level.

Model calibration. The calibration of the disease dynamic model is performed by a Markov Chain Monte Carlo (MCMC) analysis of data reported from the 2013 ZIKV epidemic in French

Polynesia [14]. Setting the extrinsic and intrinsic latent periods and the human infectious period to reference values and using the average temperature of French Polynesia, we estimate a basic reproduction number (R_0) of 2.33 (95% credible interval (CI) [2.15-2.51]), which is consistent with other analyses of these data [14, 57]. As R_0 depends on the disease serial interval, we report a sensitivity analysis in the Supplementary Materials considering the upper and lower extremes of plausible serial intervals. Briefly, the estimated R_0 values are 2.16 (95% CI [2.00-2.32]) and 3.06 (95% CI [2.80-3.32]) for the shortest and longest serial intervals, respectively. The R_0 values are consistent with those estimated from local outbreaks in San Andres Island ($R_0 = 1.41$) and Girardot, Colombia ($R_0 = 4.61$) [58]. The calibration in French Polynesia provides the basic transmissibility of ZIKV. However, variations in temperature, mosquito abundance and other factors yield varying R_0 in each subpopulation tracked by the model as discussed in the Supplementary Materials.

Acknowledgements

We acknowledge funding from the MIDAS-National Institute of General Medical Sciences U54GM111274 and to the Colombian Department of Science and Technology (Fulbright-Colciencias scholarship to D.P.R.)

References

[1] de Melo Freire CC, Iamarino A, de Lima Neto DF, de Andrade Zanotto PM, et al. (2015) Spread of the pandemic Zika virus lineage is associated with NS1 codon usage adaptation in humans. bioRxiv : 032839.

[2] Hayes EB, et al. (2009) Zika virus outside Africa. Emerg Infect Dis 15: 1347–1350.

[3] Dick G, Kitchen S, Haddow A (1952) Zika virus (I). Isolations and serological specificity. Transactions of the Royal Society of Tropical Medicine and Hygiene 46: 509–520.

[4] Center for Disease Control. Zika Virus: Symptoms, Diagnosis, & Treatment. URL `http://www.cdc.gov/zika/disease-qa.html`.

[5] D'Ortenzio E, Matheron S, de Lamballerie X, Hubert B, Piorkowski G, et al. (2016) Evidence of sexual transmission of Zika Virus. New England Journal of Medicine 374: 2195–2198.

[6] Mlakar J, Korva M, Tul N, Popović M, Poljšak-Prijatelj M, et al. (2016) Zika virus associated with microcephaly. New England Journal of Medicine 374: 951–958.

[7] Gao D, Lou Y, He D, Porco TC, Kuang Y, et al. (2016) Prevention and control of zika fever as a mosquito-borne and sexually transmitted disease. Scientific Reports 6: 28070.

[8] Musso D, Roche C, Robin E, Nhan T, Teissier A, et al. (2015) Potential sexual transmission of Zika virus. Emerging infectious diseases 21: 359.

[9] Chouin-Carneiro T, Vega-Rua A, Vazeille M, Yebakima A, Girod R, et al. (2016) Differential Susceptibilities of Aedes aegypti and Aedes albopictus from the Americas to Zika Virus. PLoS Negl Trop Dis 10: e0004543.

[10] Grard G, Caron M, Mombo IM, Nkoghe D, Ondo SM, et al. (2014) Zika virus in Gabon (Central Africa)–2007: a new threat from Aedes albopictus? PLoS Negl Trop Dis 8: e2681.

[11] Filipe A, Martins C, Rocha H (1973) Laboratory infection with Zika virus after vaccination against yellow fever. Archiv für die gesamte Virusforschung 43: 315–319.

[12] Duffy MR, Chen TH, Hancock WT, Powers AM, Kool JL, et al. (2009) Zika virus outbreak on Yap Island, federated states of Micronesia. New England Journal of Medicine 360: 2536–2543.

[13] Ioos S, Mallet HP, Goffart IL, Gauthier V, Cardoso T, et al. (2014) Current Zika virus epidemiology and recent epidemics. Medecine et maladies infectieuses 44: 302–307.

[14] 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.

[15] European Centre for Disease Prevention and Control Zika virus infection outbreak, Brazil and the Pacific region.

[16] Schuler-Faccini L (2016) Possible Association Between Zika Virus Infection and Microcephaly—Brazil, 2015. MMWR Morbidity and Mortality Weekly Report 65.

[17] World Health Organization. WHO Director-General summarizes the outcome of the Emergency COmmittee regarding clusters of microcephaly and Guillain-Barre syndrome. URL http://www.who.int/mediacentre/news/statements/2016/emergency-committee-zika-microcephaly/en/.

[18] Centers for Disease Control. Countries and Territories in the Americas with Active Zika Virus Transmission. URL http://www.cdc.gov/zika/geo/americas.html.

[19] Pan-American Health Organization. Countries and territories with autochthonous transmission in the Americas reported in 2015-2016. URL http://www.paho.org/hq/index.php?option=com_content&view=article&id=11603&Itemid=41696&lang=en.

[20] Pan-American Health Organization (2016). Home Health Topics Programs Media Center Publications Data Countries and Centers About PAHO Regional Zika Epidemiological Update (Americas) - 14 July 2016. URL http://www.paho.org/hq/index.php?option=com_content&view=article&id=11599:regional-zika-epidemiological-update-americas&Itemid=41691&lang=en.

[21] Faria NR, d S d S Azevedo R, Kraemer MUG, Souza R, Cunha MS, et al. (2016) Zika virus in the Americas: Early epidemiological and genetic findings. Science 352: 345–349.

[22] Pan-American Health Organization (2016). Zika virus infection. URL http://ais.paho.org/phip/viz/ed_zika_epicurve.asp.

[23] Rodrigues LC (2016) Microcephaly and Zika virus infection. The Lancet 387: 2070–2072.

[24] 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.

[25] 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.

[26] França GVA, Schuler-Faccini L, Oliveira WK, Henriques CMP, Carmo EH, et al. (2016). Congenital Zika virus syndrome in Brazil: a case series of the first 1501 livebirths with complete investigation. doi:10.1016/s0140-6736(16)30902-3. URL `http://dx.doi.org/10.1016/s0140-6736(16)30902-3`.

[27] Nelson B, Morrison S, Joseph H, Wojno A, Lash RR, et al. (2016). Travel Volume to the United States from Countries and US Territories with Local Zika Virus Transmission. doi:10.1371/currents.outbreaks.ac6d0f8c9c35e88825c1a1147697531c.

[28] Castro LA, Fox SJ, Chen X, Liu K, Bellan SE, et al. (2016) Real-time zika risk assessment in the united states. bioRxiv .

[29] Manore C, Ostfeld R, Agusto F, Gaff H, LaDeau S (2016) Defining the risk of zika and chikungunya virus transmission in human population centers of the eastern united states. bioRxiv .

[30] Perkins A, Siraj A, Ruktanonchai CW, Kraemer M, Tatem A (2016) Model-based projections of Zika virus infections in childbearing women in the Americas. Technical report. doi:10.1101/039610. URL `http://dx.doi.org/10.1101/039610`.

[31] Cardoso CW, Paploski IA, Kikuti M, Rodrigues MS, Silva MM, et al. (2015) Outbreak of Exanthematous Illness Associated with Zika, Chikungunya, and Dengue Viruses, Salvador, Brazil. Emerging Infectious Diseases 21: 2274–2276.

[32] Pan-American Health Organization (2016). Zika - Epidemiological Report Brazil. URL `http://www.paho.org/hq/index.php?option=com_docman&task=doc_view&gid=35221&Itemid=270&lang=en`.

[33] Lednicky J, De Rochars VMB, El Badry M, Loeb J, Telisma T, et al. Zika Virus Outbreak in Haiti in 2014: Molecular and Clinical Data, volume = 10, year = 2016. PLoS Negl Trop Dis : e0004687.

[34] Pacheco O, Beltrán M, Nelson CA, Valencia D, Tolosa N, et al. Zika Virus Disease in Colombia — Preliminary Report. doi:10.1056/NEJMoa1604037. URL http://dx.doi.org/10.1056/NEJMoa1604037. PMID: 27305043, http://dx.doi.org/10.1056/NEJMoa1604037.

[35] Pan-American Health Organization (2016). Zika - Epidemiological Report El Salvador. URL http://www.paho.org/hq/index.php?option=com_docman&task=doc_view&gid=35225&Itemid=270&lang=en.

[36] Pan-American Health Organization (2016). Zika - Epidemiological Report Honduras. URL http://www.paho.org/hq/index.php?option=com_docman&task=doc_view&gid=35137&Itemid=270.

[37] Pan-American Health Organization (2016). Zika - Epidemiological Report Puerto Rico. URL http://www.paho.org/hq/index.php?option=com_docman&task=doc_view&gid=35231&Itemid=270&lang=en.

[38] Pan-American Health Organization (2016). Zika - Epidemiological Report Mexico. URL http://www.paho.org/hq/index.php?option=com_docman&task=doc_view&gid=35106&Itemid=270.

[39] Ferguson NM, Cucunubá ZM, Dorigatti I, Nedjati-Gilani GL, Donnelly CA, et al. (2016) Countering Zika in Latin America. Science : aag0219.

[40] Declan Butler. Brazil asks wheter Zika acts alone to cause birth defects. URL http://www.nature.com/news/brazil-asks-whether-zika-acts-alone-to-cause-birth-defects-1.20309.

[41] Balcan D, Gonçalves B, Hu H, Ramasco JJ, Colizza V, et al. (2010) Modeling the spatial spread of infectious diseases: The GLobal Epidemic and Mobility computational model. Journal of Computational Science 1: 132–145.

[42] Balcan D, Hu H, Goncalves B, Bajardi P, Poletto C, et al. (2009) Seasonal transmission potential and activity peaks of the new influenza A(H1N1): a Monte Carlo likelihood analysis based on human mobility. BMC Medicine 7: 45.

[43] Tizzoni M, Bajardi P, Poletto C, Ramasco JJ, Balcan D, et al. (2012) Real-time numerical forecast of global epidemic spreading: case study of 2009 A/H1N1pdm. BMC Medicine 10: 165.

[44] Gomes MFC, Pastore-Piontti A, Rossi L, Chao D, Longini I, et al. (2014). Assessing the International Spreading Risk Associated with the 2014 West African Ebola Outbreak. doi:10.1371/currents.outbreaks.cd818f63d40e24aef769dda7df9e0da5. URL http://dx.doi.org/10.1371/currents.outbreaks.cd818f63d40e24aef769dda7df9e0da5.

[45] Poletto C, Gomes M, Pastore-Piontti A, Rossi L, Bioglio L, et al. (2014) Assessing the impact of travel restrictions on international spread of the 2014 West African Ebola epidemic. Eurosurveillance 19: 20936.

[46] Zhang Q, Gioannini C, Paolotti D, Perra N, Perrotta D, et al. (2015) Social Data Mining and Seasonal Influenza Forecasts: The FluOutlook Platform. In: Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Springer International Publishing, pp. 237–240.

[47] William Nordhaus and Xi Chen. Geographically based Economic data (G-Econ). URL http://gecon.yale.edu/.

[48] Kenji Matsuura and Cort J Willmott. Terrestrial Air Temperature: 1900-2010 Gridded Monthly Time Series. URL http://climate.geog.udel.edu/~climate/html_pages/Global2011/GlobalTsT2011.html.

[49] Kraemer MU, Sinka ME, Duda KA, Mylne A, Shearer FM, et al. (2015) The global compendium of Aedes aegypti and Ae. albopictus occurrence. Scientific Data 2.

[50] 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.

[51] Columbia Univerity. Socioeconomic data and Applications center. URL http://sedac.ciesin.columbia.edu/.

[52] OAG, Aviation Worlwide Limited. . URL http://www.oag.com/.

[53] Nordhaus WD (2006) Geography and macroeconomics: New data and new findings. Proceedings of the National Academy of Sciences of the United States of America 103: 3510–3517.

[54] World Health Organization (2016). Zika virus. URL http://www.who.int/mediacentre/factsheets/zika/en/.

[55] Polwiang S (2015) The seasonal reproduction number of dengue fever: impacts of climate on transmission. PeerJ 3: e1069.

[56] Johansson MA, Powers AM, Pesik N, Cohen NJ, Staples JE (2014) Nowcasting the Spread of Chikungunya Virus in the Americas. PLoS ONE 9: e104915.

[57] Chowell G, D HP, Ospina J, Pell B, Tariq A, et al. (2016). Using phenomenological models to characterize transmissibility and forecast patterns and final burden of zika epidemics. doi:10.1371/currents.outbreaks.f14b2217c902f453d9320a43a35b9583.

[58] Rojas DP, Dean NE, Yang Y, Kenah E, Quintero J, et al. (2016) The Epidemiology and Transmissibility of Zika Virus in Girardot and San Andres Island, Colombia. Eurosurveillance 21: 049957.

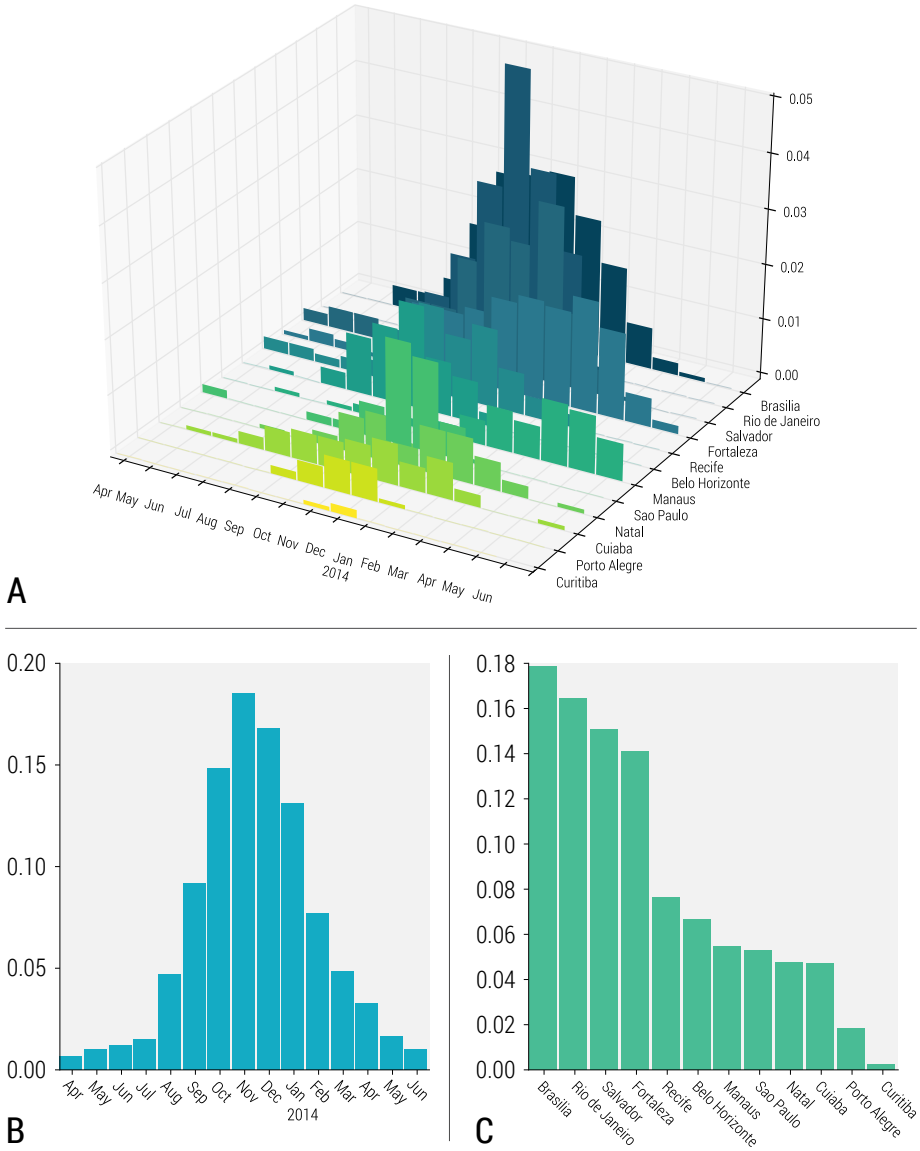


Fig. 1: Posterior distribution for ZIKV introductions in twelve major transportation hubs in Brazil between April 2013 and August 2014, incorporating the likelihood of replicating the observed epidemic peak in Colombia. A) Full posterior distribution as a function of location and time of introduction. B) Marginal posterior distribution for time (month) of introduction. C) Marginal posterior distribution for location of introduction.

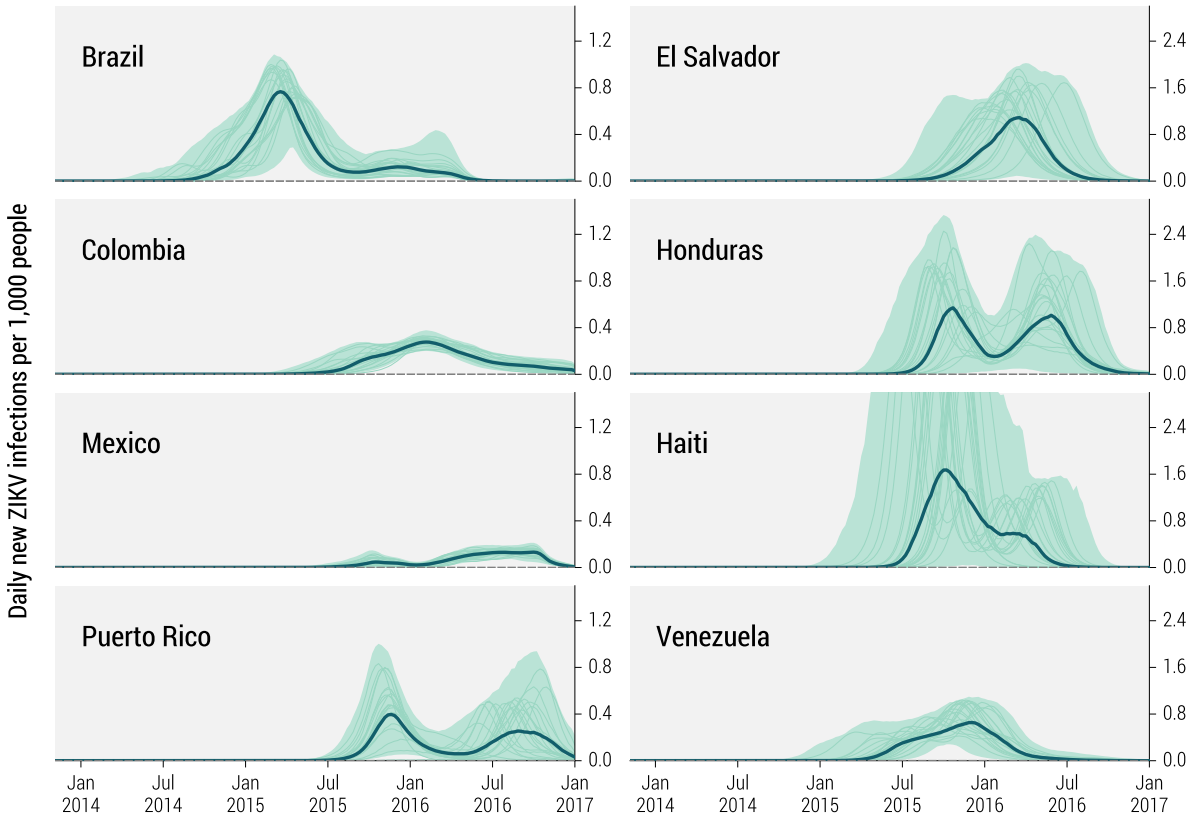


Fig. 2: Estimated daily number of new Zika infections (per 1000 people) in eight affected countries in the Americas between January 2014 and January 2017. The bold line and shaded area refer to the estimated median number of infections and 95%CI of the model projections, respectively. Rates include asymptomatic 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.

	Attack Rate %		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	Dec. 31, 2016	Feb. 1, 2016	Dec. 31, 2016	Feb. 1, 2016	Dec. 31, 2016	Feb. 1, 2016	Dec. 31, 2016
Brazil	14 [11 - 16]	15 [13 - 16]	784 [286 - 1081]	1067 [964 - 1201]	1808 [660 - 2492]	2460 [2222 - 2768]	3732 [1361 - 5144]	5077 [4586 - 5713]
Colombia	4 [2 - 6]	8 [7 - 9]	1 [0 - 15]	107 [70 - 144]	2 [0 - 36]	248 [160 - 332]	5 [0 - 74]	511 [331 - 685]
Mexico	1 [0 - 2]	4 [2 - 5]	0 [0 - 6]	77 [24 - 142]	1 [0 - 14]	178 [55 - 327]	2 [0 - 29]	367 [114 - 675]
Puerto Rico	4 [1 - 10]	10 [6 - 15]	0 [0 - 0]	5 [1 - 11]	0 [0 - 0]	11 [2 - 25]	0 [0 - 1]	23 [5 - 51]
El Salvador	6 [0 - 20]	22 [20 - 24]	0 [0 - 1]	48 [10 - 58]	0 [0 - 3]	110 [23 - 133]	0 [0 - 6]	228 [47 - 274]
Honduras	15 [0 - 29]	32 [28 - 35]	0 [0 - 21]	97 [45 - 126]	0 [0 - 49]	223 [104 - 291]	1 [0 - 101]	460 [215 - 601]
Haiti	38 [10 - 48]	47 [45 - 49]	1 [0 - 192]	294 [215 - 312]	2 [0 - 442]	678 [495 - 720]	5 [0 - 912]	1399 [1021 - 1486]
Venezuela	14 [8 - 16]	16 [15 - 17]	12 [0 - 139]	229 [194 - 247]	28 [1 - 319]	529 [446 - 569]	57 [2 - 659]	1091 [921 - 1175]

Tab. 1: Projected ZIKV attack rates through the time of the WHO declaration of a PHEIC on February 1, 2016, and through December 31, 2016, 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 who 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 31, 2016, in eight affected countries in the Americas. We consider three different microcephaly 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[24]; 2.19% (100% over-reporting) and 4.52% (no over-reporting) first trimester risks, based on a study of Bahia, Brazil[25], given a model-estimated 30% attack rate in Bahia.

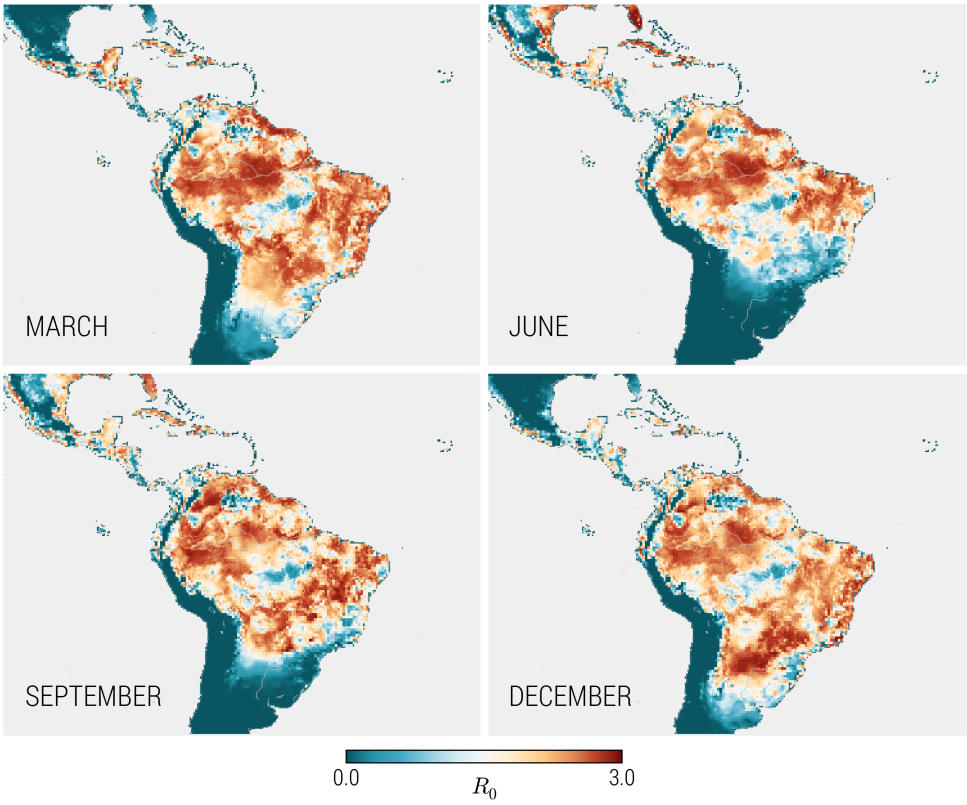


Fig. 3: Monthly seasonality for the time- and location-dependent basic reproduction number. The Equatorial region presents less seasonality than the non-Equatorial regions, where the changes of the season have a strong impact over the temperature and consequently over the basic reproduction number.

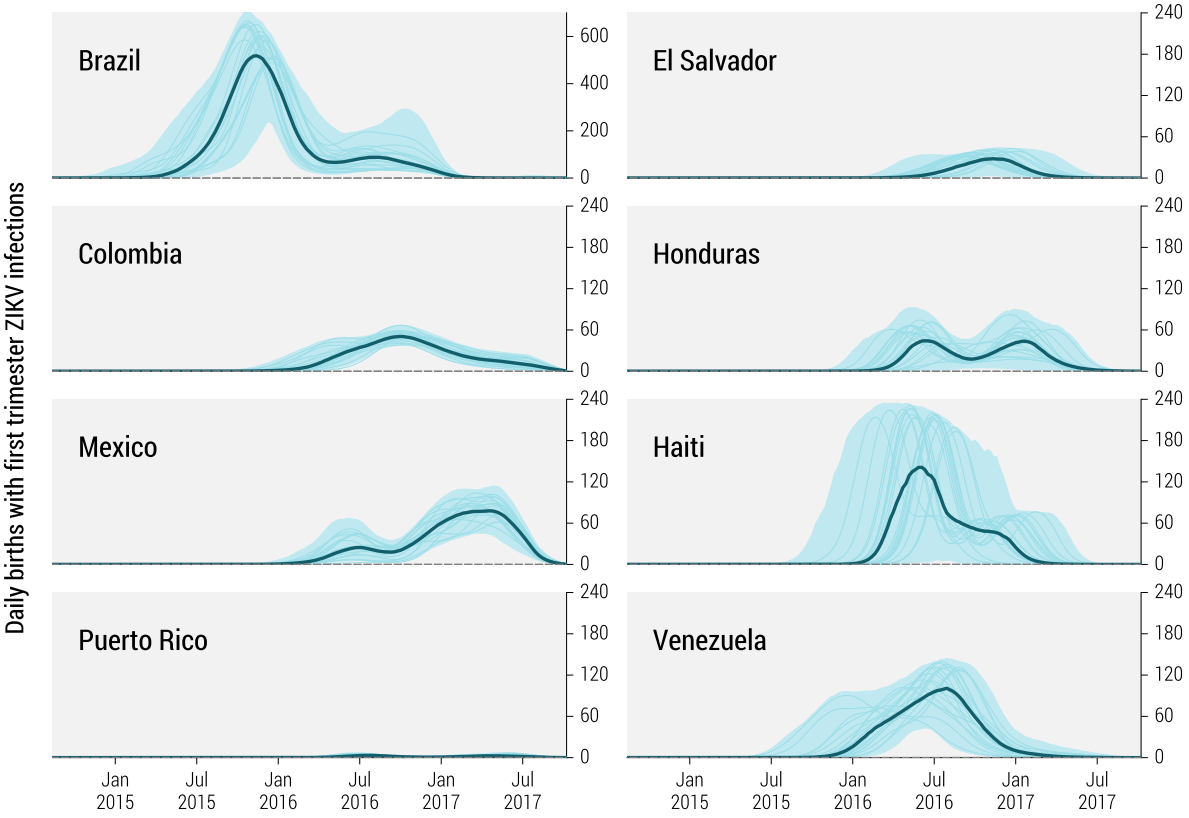


Fig. 4: Estimated daily number of births between October 2014 and October 2017 from women infected with ZIKV during the first trimester of pregnancy in eight affected countries in the Americas. The bold line and shaded area refer to the estimated median number of births and 95%CI of the model projections, 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.

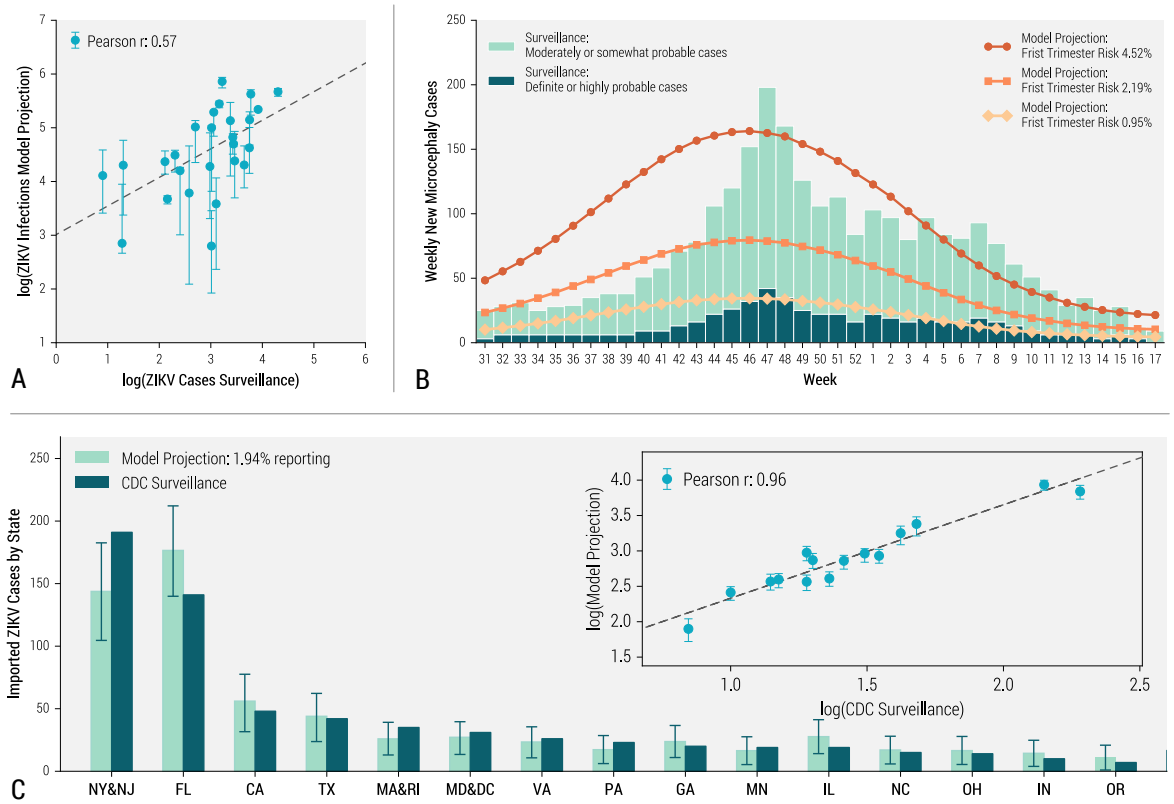


Fig. 5: A) Correlation between the number of ZIKV cases by states in Colombia as reported by surveillance data through June 5, 2016, compared with state-level model projections of infections (median with 95%CI). Pearson's r correlation coefficient is reported for the linear association on the log scale. 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[26]. Line plots indicate estimated weekly new microcephaly cases given three levels of first trimester risk: 4.52% (round) [25], 2.19% (square) [25], and 0.95% (diamond) [24]. C) Bar plot of ZIKV infections imported into the continental USA by state(s) as reported by CDC surveillance through June 15, 2016, and compared to model projections (median with 95% CI) for the same period assuming 1.94% reporting/detection. The insert shows the correlation between CDC surveillance data and model projections (median with 95%CI).

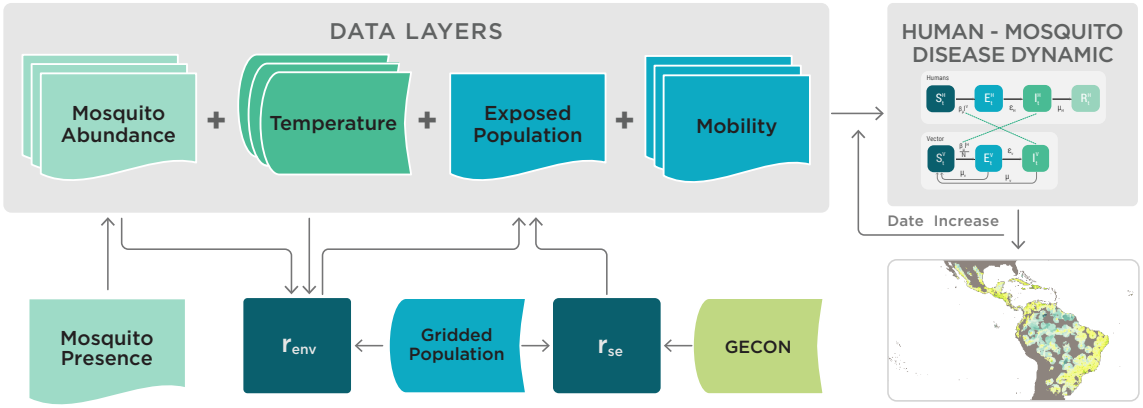
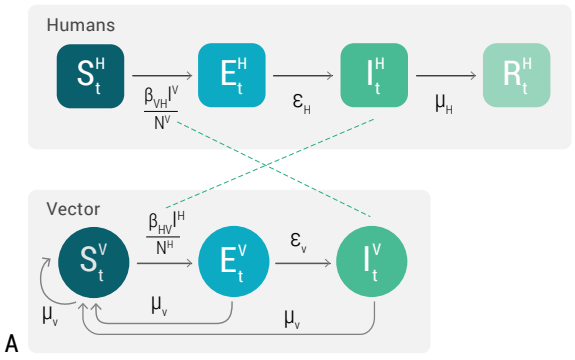


Fig. 6: Schematic representation of the integration of data layers and the computational flow chart defining the Global Epidemic and Mobility (GLEAM) model for ZIKV.



B

Parameter		Value	Range	Ref.
extrinsic latent period	$1/\epsilon_V$	7 days	7 - 10 days	[2]
intrinsic latent period	$1/\epsilon_H$	3.5 days	2 - 7 days	[54]
human infectious time	$1/\mu_H$	4.5 days	4 - 7 days	[14]
proportion of asymptomatic cases	ρ_a	0.8	-	[12]
moquito lifespan	$1/\mu_V$	$T_{dep.}$	-	[56]
mosquitoes per person	k	$T,G_{dep.}$	-	
mosquito-to-human transmission rate	τ_{HV}	$T_{dep.}$	-	[55]
human-to-mosquito transmission rate	τ_{VH}	$T_{dep.}$	-	[55]
reproductive number	R_0	2.33	2.16 - 3.06	

Fig. 7: A) Compartmental classification for ZIKV infection. Humans can occupy one the four top compartments: susceptible, which can acquire the infection through contacts (bites) with infectious mosquitoes; exposed, where individuals are infected but are not able yet to transmit the virus; infectious, where individuals are infected and can transmit the disease to susceptible mosquitoes; and recovered or removed, where individuals are no longer infectious. The compartmental model for the mosquito vector is shown below. B) Summary of the parameters of the model. $T_{dep.}$ denotes parameters that are temperature dependent. $T, G_{dep.}$ denotes parameters that are temperature and geo-location dependent. Specific values for the parameters can be found in Refs. [2, 54, 14, 12, 56, 55].