



Spread of Zika virus in the Americas

Qian Zhang^a, Kaiyuan Sun^a, Matteo Chinazzi^a, Ana Pastore y Piontti^a, Natalie E. Dean^b, Diana Patricia Rojas^c, Stefano Merler^d, Dina Mistry^a, Piero Poletti^e, Luca Rossi^f, Margaret Bray^a, M. Elizabeth Halloran^{g,h}, Ira M. Longini Jr.^b, and Alessandro Vespignani^{a,f,1}

^aLaboratory for the Modeling of Biological and Socio-technical Systems, Northeastern University, Boston, MA 02115; ^bDepartment of Biostatistics, College of Public Health and Health Professions, University of Florida, Gainesville, FL 32611; ^cDepartment of Epidemiology, College of Public Health and Health Professions, University of Florida, Gainesville, FL 32611; ^dBruno Kessler Foundation, 38123 Trento, Italy; ^eDondena Centre for Research on Social Dynamics and Public Policy, Università Commerciale L. Bocconi, 20136 Milan, Italy; ^fInstitute for Scientific Interchange Foundation, 10126 Turin, Italy; ^gVaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109; and ^hDepartment of Biostatistics, University of Washington, Seattle, WA 98195

Edited by Alan Hastings, University of California, Davis, CA, and approved March 30, 2017 (received for review December 8, 2016)

We use a data-driven global stochastic epidemic model to analyze the 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 (90% credible interval). We provide simulated epidemic profiles of incident ZIKV infections for several countries in the Americas through February 2017. 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 pregnancies infected with ZIKV during the first trimester and provide estimates of microcephaly cases assuming different levels of risk as reported in empirical retrospective studies. Our approach represents a modeling effort aimed at understanding the potential magnitude and timing of the ZIKV epidemic and it can be potentially used as a template for the analysis of future mosquito-borne epidemics.

Zika virus | computational epidemiology | metapopulation network model | vector-borne diseases

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 generally results in a mild disease characterized by low-grade fever, rash, and/or conjunctivitis, although only ~20% of those infected are symptomatic (4). Although there have been instances of sexual and perinatal/vertical transmission (5–12) and the potential for transmission by transfusion is present (13), ZIKV spreads primarily through infected *Aedes* mosquitoes (14, 15).

Until recently, ZIKV was considered a neglected tropical disease with only local outbreaks (4, 16–18). The association of ZIKV with the reported microcephaly case clusters in Brazil during 2015 (19) led the director-general of the WHO to declare on February 1, 2016, a Public Health Emergency of International Concern (PHEIC) (20) that lasted for nearly 10 mo. During this period, ZIKV spread throughout the Americas, with 47 countries and territories in the region reporting autochthonous transmission (21, 22). Many 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) (23), and the epidemic has been under close scrutiny by all of the major public health agencies around the world.

Although enhanced surveillance and new data have improved our understanding of ZIKV (24–29), many unknowns persist. There is uncertainty surrounding the time of introduction of the virus to the region, although epidemiological and genetic findings estimate that ZIKV arrived in Brazil between May and December 2013 (nextstrain.org/zika; ref. 30). Furthermore, although mathematical and computational models have tackled the characterization of the transmissibility and potential burden of ZIKV (31–35), little is known about the global spread of the virus in 2014 and 2015, before the WHO's alert in early 2016.

Using a data-driven stochastic and spatial epidemic model, we present numerical results providing insight into the first introduction in the region and the epidemic dynamics across the Americas. We use the model to analyze the spatiotemporal spread and magnitude of the epidemic in the Americas through to February 2017, accounting for seasonal environmental factors and detailed population data. We also provide projections of the number of pregnancies infected with ZIKV during the first trimester, along with estimates for the number of microcephaly cases per country using three different levels of risk based on empirical retrospective studies (36, 37).

Results

Introduction of ZIKV to the Americas. We identify 12 major transportation hubs in areas related to major events held in Brazil, such as the Soccer Confederations Cup in June 2013 and the Soccer World Cup in June 2014 and assumed a prior probability of introduction proportional to the daily passenger flow to each hub. We then consider introduction dates between April 2013 and June 2014, including the time frame suggested by phylogenetic and molecular clock analyses (nextstrain.org/zika; ref. 30) through to the 2014 Soccer World Cup. Using Latin square sampling over the two-dimensional space (date–location), we calculated the likelihood of replicating the observed epidemic peak in Colombia (± 1 wk), as reported by Colombia's National Institute of Health (38), and the resulting posterior density of each

Significance

Mathematical and computational modeling approaches can be essential in providing quantitative scenarios of disease spreading, as well as projecting the impact in the population. Here we analyze the spatial and temporal dynamics of the Zika virus epidemic in the Americas with a microsimulation approach informed by high-definition demographic, mobility, and epidemic data. The model provides probability distributions for the time and place of introduction of Zika in Brazil, the estimate of the attack rate, timing of the epidemic in the affected countries, and the projected number of newborns from women infected by Zika. These results are potentially relevant in the preparation and analysis of contingency plans aimed at Zika virus control.

Author contributions: M.E.H., I.M.L., and A.V. designed research; Q.Z., K.S., M.C., A.P.P., S.M., D.M., P.P., L.R., and A.V. performed research; Q.Z., K.S., M.C., A.P.P., D.M., M.B., and A.V. analyzed data; and Q.Z., K.S., M.C., A.P.P., N.E.D., D.P.R., S.M., D.M., P.P., L.R., M.B., M.E.H., I.M.L., and A.V. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

See Commentary on page 5558.

¹To whom correspondence should be addressed. Email: a.vespignani@northeastern.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1620161114/-DCSupplemental.

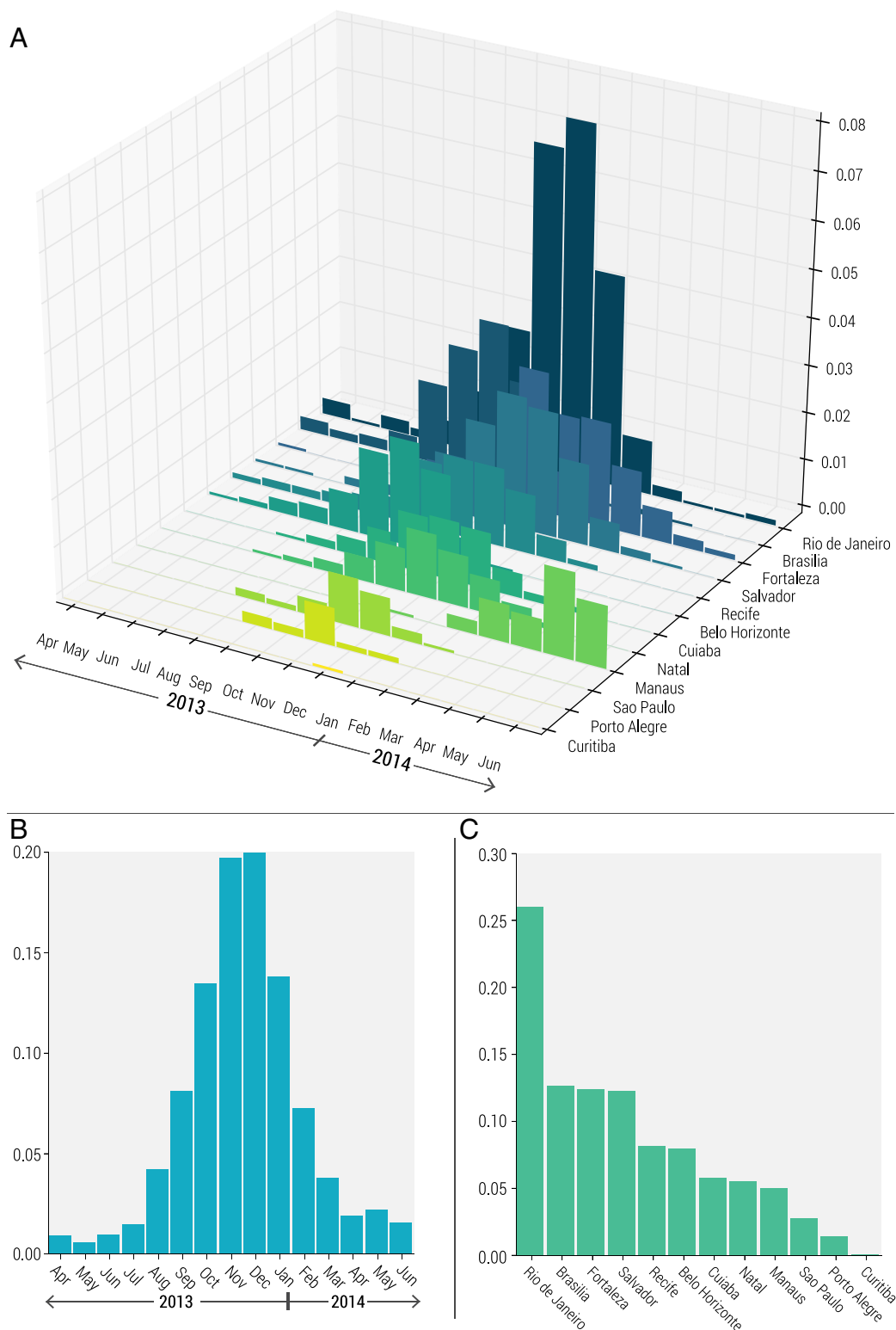


Fig. 1. Posterior distribution for ZIKV introductions in 12 major transportation hubs in Brazil between April 2013 and June 2014, incorporating the likelihood of replicating the observed epidemic peak in Colombia. (A) Full posterior distribution as a function of location and time of introduction. (B) Marginal posterior distribution for time (month) of introduction. (C) Marginal posterior distribution for location of introduction.

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. 1*A* we plot the posterior distribution as a function of introduction date and location, and in Fig. 1*B* and *C* we plot

the marginal posterior distributions of introduction date and location separately. The largest posterior density is associated with an introduction in Rio de Janeiro in December 2013. The 90% credible interval for the most likely date extends from August 2013 to April 2014, with the mode in December 2013, in agreement with

Table 1. 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

	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	16 [13 to 18]	18 [16 to 19]	839 [138 to 1,140]	1,297 [1,190 to 1,428]	1,934 [318 to 2,628]	2,991 [2,744 to 3,291]	3,992 [656 to 5,424]	6,173 [5,664 to 6,792]
Colombia	4 [3 to 7]	12 [11 to 14]	0 [0 to 4]	219 [194 to 248]	0 [0 to 10]	504 [447 to 572]	1 [0 to 20]	1,041 [922 to 1,180]
Mexico	1 [0 to 2]	5 [4 to 6]	0 [0 to 2]	314 [226 to 367]	0 [0 to 5]	723 [522 to 845]	1 [0 to 11]	1,493 [1,077 to 1,744]
Puerto Rico*	2 [0 to 7]	20 [13 to 28]	0 [0 to 0]	19 [13 to 26]	0 [0 to 0]	43 [29 to 60]	0 [0 to 0]	88 [60 to 124]
El Salvador	1 [0 to 13]	16 [13 to 18]	0 [0 to 0]	39 [32 to 47]	0 [0 to 0]	91 [75 to 108]	0 [0 to 1]	187 [154 to 223]
Honduras	8 [0 to 28]	35 [30 to 39]	0 [0 to 1]	144 [124 to 163]	0 [0 to 3]	332 [286 to 376]	0 [0 to 7]	686 [590 to 775]
Haiti	43 [1 to 54]	49 [43 to 55]	0 [0 to 54]	316 [276 to 357]	0 [0 to 124]	728 [637 to 824]	0 [0 to 256]	1,502 [1,315 to 1,700]
Venezuela	13 [5 to 19]	19 [16 to 21]	2 [0 to 96]	271 [237 to 308]	5 [0 to 221]	624 [546 to 711]	9 [0 to 456]	1,288 [1,127 to 1,468]

Median estimates and 95% CIs are provided. ZIKV ARs include asymptomatic infections. The denominator is the entire population of the country, including regions 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 (36) and 2.19% (100% overreporting) and 4.52% (no overreporting) first-trimester risks, based on a study of Bahia, Brazil (37), given a model-estimated 31% 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 (72).

times throughout the year. Equatorial regions experience less seasonality than nonequatorial regions, where 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. It is also worth remarking that maximum R_0 is not the sole determinant of the epidemic magnitude, because seasonality patterns and a small fraction of exposed individuals may not allow large outbreaks to occur.

Projected ZIKV Infections in Childbearing Women and Microcephaly Cases. Using the epidemic profiles generated by the model we project the number of ZIKV infections in childbearing women following the model proposed in the study of ZIKV–microcephaly association for the 2013–2014 French Polynesia

outbreak (36). In Fig. 4 we plot the daily number of births through December 2017 from women infected with ZIKV during their first trimester of pregnancy in several countries. Indeed, the first trimester of pregnancy is when the risk of microcephaly is the highest (36, 37, 39). The curves closely resemble the epidemic profiles in Fig. 2 but shifted forward in time by about 40 wk. We construct our estimates using country-specific birth rates, as detailed in *SI Appendix, section 4*.

To estimate the number of microcephaly cases we adopt three different probabilities, as reported in two empirical retrospective studies (36, 37). The first estimate of microcephaly risk for ZIKV infected pregnancies is 0.95% (95% confidence interval (CI) [0.34 to 1.91%]), from a study in French Polynesia (36). The remaining two estimates come from a study performed in Bahia, Brazil (37). Given an overall ZIKV infection AR of 31% (95%

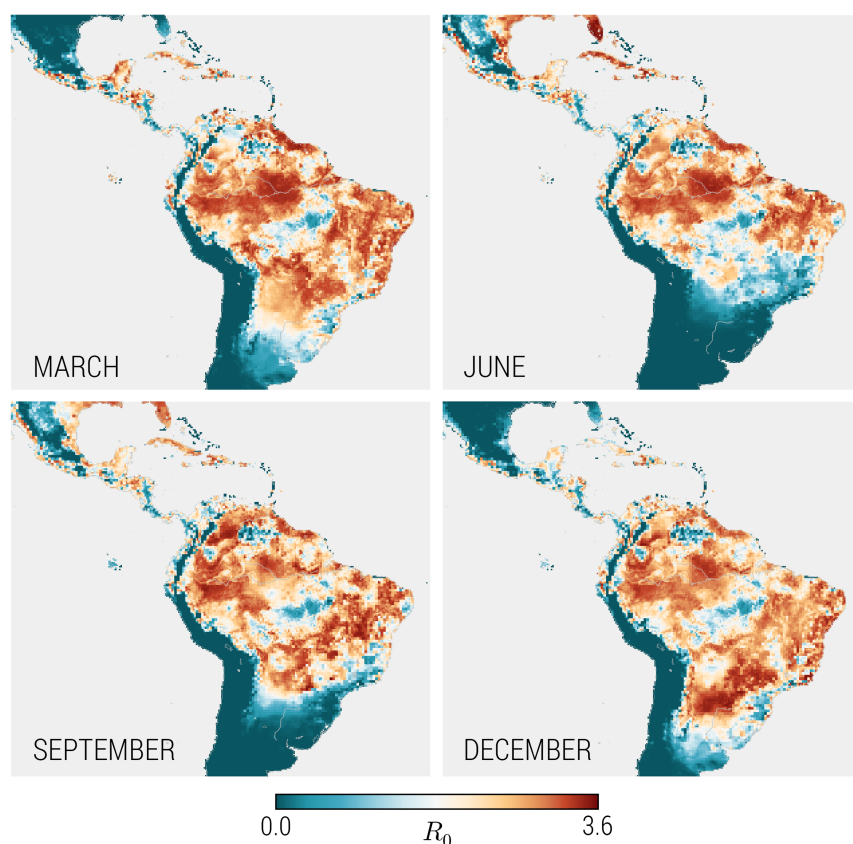


Fig. 3. Monthly seasonality for the time- and location-dependent basic reproductive number, R_0 . The equatorial regions display less seasonality than the nonequatorial regions, where the changes of the season have a strong impact on the temperature and consequently on the basic reproductive number, R_0 .

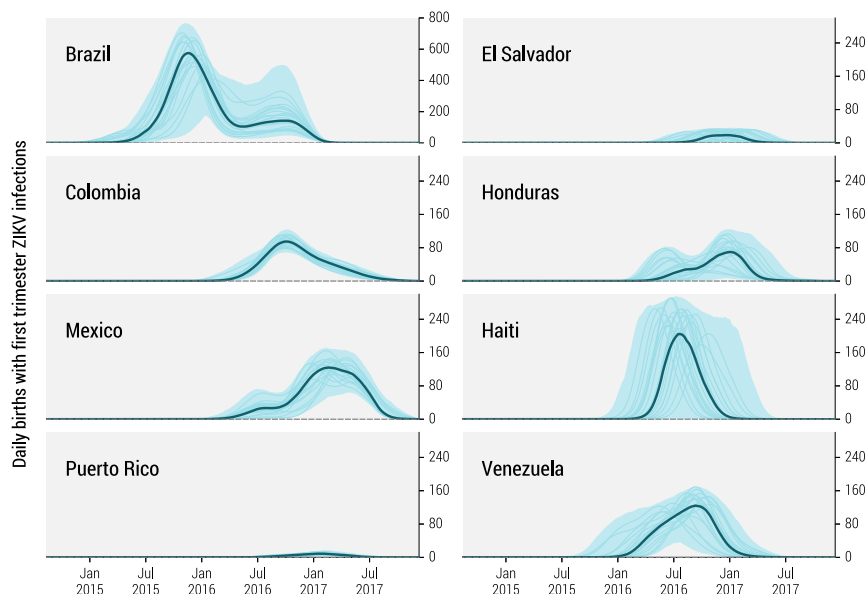


Fig. 4. Estimated daily number of births between October 2014 and December 2017 from women infected with ZIKV during the first trimester of pregnancy in eight affected countries in the Americas. The bold line and shaded area refer to the estimated median number of births and 95% CI of the model projections,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

CI [26 to 36%]) in Bahia through February 2016, as determined by our model, the estimated first trimester microcephaly risks are 2.19% (95% CI [1.98 to 2.41%]), assuming 100% overreporting of microcephaly cases, and 4.52% (95% CI [4.10 to 4.96%]), assuming no overreporting. These estimates do not account for miscarriages or other complications that may occur during pregnancy.

In Table 1 we report the projected cumulative number of microcephaly cases up to February 1, 2016, and December 10, 2017. By the time the WHO declared a PHEIC, Brazil was the only country with a substantial (>100) number of ZIKV-attributable microcephaly cases, with cases expected to appear through July 2017. For Colombia, the model projects a considerable number of new microcephaly cases until March–April 2017. In Venezuela, the peak in microcephaly cases was projected to start in September/October 2016, continuing through February 2017. In Puerto Rico, microcephaly cases were expected to occur mostly from December 2016 to April 2017. It is important to remark, however, that the microcephaly incidence tail extends for most of the countries up to July/August 2017. Along with the microcephaly risk, other birth defects and pregnancy complications are associated with ZIKV infection during pregnancy (36, 37, 39). Although we do not explicitly tabulate here specific projections, they can be calculated from our model by applying the estimated risk for any other complication to our daily number of births from women infected with ZIKV.

Sensitivity to Mosquito Vector. Simulations reported here consider both *Aedes aegypti* and *Aedes albopictus* as competent ZIKV vectors, although less is known about the vectorial capacity of *A. albopictus*. A sensitivity analysis considering *A. aegypti* as the only competent vector is provided in *SI Appendix* with all figures and tables replicated for this scenario. Overall, results are similar because transmission due to *A. aegypti* increases to compensate for the absence of the other vector. Differences in the infection ARs, however, are observed in areas where *A. albopictus* is the most common or the only vector. For example, the infection AR in Brazil up to February 28, 2017, decreases from 18% (95% CI [16 to 19%]) to 16% (95% CI [14 to 17%]) if we consider only *A. aegypti*. During the same time period, the infection AR in Mexico decreases from 5% (95% CI [4 to 6%]) to 4% (95% CI [2 to 5%]). A more thorough analysis of the differences between the two scenarios is reported in *SI Appendix*. At the country scale, in Brazil and other key countries those differences seem small because *A. aegypti* and *A. albopictus* have very similar presence. However, we see noticeable differences in the

infection AR as soon as the country extends to north and south of the equator and we look at specific geographical areas where only *A. albopictus* are present.

Model Validation. Our results have been validated comparing our projections with surveillance data that were not directly used to calibrate the model: the number of ZIKV infections by states in Colombia, the weekly counts of microcephaly cases reported in Brazil, and the number of importations of ZIKV infections in the continental United States (USA), as shown in Fig. 5. In Fig. 5A we compare model-based projections of the number of ZIKV infections for states in Colombia with observed surveillance data through October 1, 2016 (38). 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 $1.02\% \pm 0.93\%$ (from linear regression analysis). However, the observed data and model estimates are well-correlated (Pearson's $r = 0.68$, $P < 0.0001$), replicating the often several-orders-of-magnitude difference 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 (40) 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. Because the fraction of cases confirmed in Brazil is relatively low, it is not possible to identify the most likely level of risk, although the figure suggests that the risk might exceed the lowest estimate of 0.95% (36).

Because 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 USA through October 5, 2016, as reported by the CDC (41) and compare these results with model projections. Because the detection rate of ZIKV infections is very low, there are significantly fewer reported cases than projected; we estimate through a linear regression fit that $5.74\% \pm 1.46\%$ of both symptomatic and asymptomatic imported infections are detected. Nonetheless, model projections are highly correlated with the observed data (Pearson's $r = 0.93$, $P < 0.0001$), as shown in Fig. 5C, *Inset*. A further validation of the model is provided by the reported number of ZIKV cases of pregnant women in the USA. A high detection rate is expected in this closely monitored population. As of September 29, 2016, 837 pregnant women in the USA

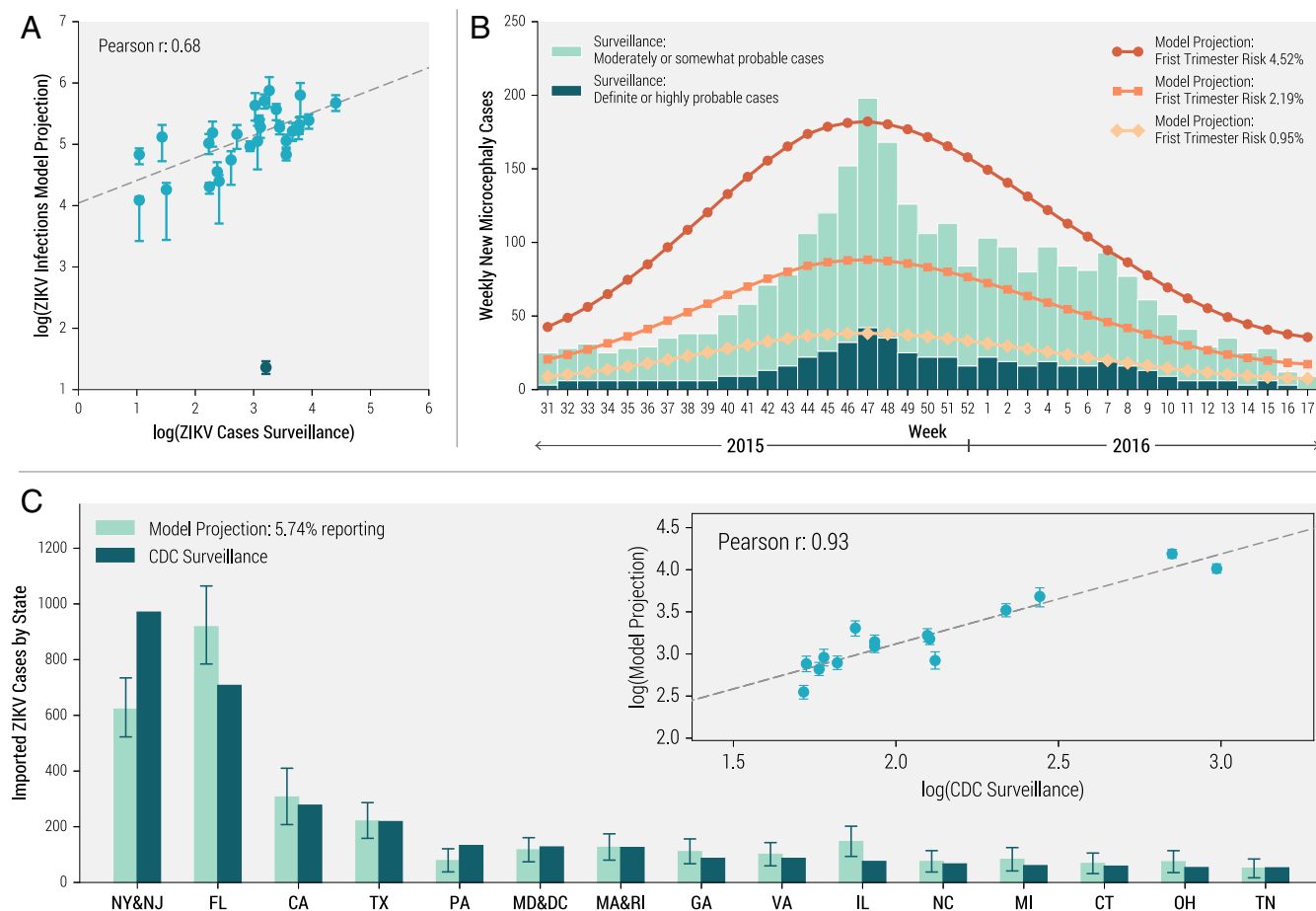


Fig. 5. (A) Correlation between the number of ZIKV cases by state in Colombia as reported by surveillance data through October 1, 2016 (38), compared with state-level model projections of infections (median with 95% CI). Pearson's r correlation coefficient is reported for the linear association on the log scale. The outlier (in dark green) excluded from the statistical analysis corresponds to the Arauca region. (B) Timeline of microcephaly cases in Brazil through April 30, 2016. Bar plots show weekly definite (or highly probable cases) and moderately (or somewhat probable cases) from surveillance data (40). Line plots indicate estimated weekly new microcephaly cases given three levels of first trimester risk: 4.52% (circles) (37), 2.19% (squares) (37), and 0.95% (diamonds) (36). (C) Bar plot of ZIKV infections imported into the USA by state(s) as reported by CDC surveillance through October 5, 2016 (41), and compared to model projections (median with 95% CI) for the same period assuming 5.74% reporting/detection. (Inset) The correlation between CDC surveillance data and model projections (median with 95% CI).

were laboratory-confirmed for ZIKV, all of whom were imported cases. Because pregnant women comprise $\sim 1\%$ of incoming airline traffic flow from the rest of the Americas (42), one can roughly estimate 83,700 infections. Although this is a rough estimate, because of fluctuations in pregnant women traffic flow and testing rates, it is in the ballpark of our modeling results projecting 57,910 (95% CI [50, 138 to 66, 608]) infections imported into the USA by early October 2016. These results are relevant for ZIKV risk assessment in the USA (43, 44). In *SI Appendix* we provide additional validation tests by looking at case reporting in Brazil at the state level, and the detection of travel related cases in European countries.

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 are estimated for 12 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 Rio de Janeiro, Brasilia, Fortaleza, and Salvador. This date range overlaps with that suggested by a recent phylogenetic analysis (nextstrain.org/zika; ref. 30), although our estimate also includes later potential introductions. The model seems to

rule out an introduction concurrent to the Soccer World Cup in June 2014.

The model is able to generate epidemic curves in time for incident ZIKV cases for about two dozen countries in the Americas. Although for the sake of space we report on only eight countries, the full database is publicly available (www.zika-model.org). The results obtained are in good agreement with model-based projections achieved with a different approach developed by Perkins et al. (32) using location-specific epidemic ARs on highly spatially resolved human demographic projections. Although the approach of Perkins et al. (32) does not provide information on the dynamic of the epidemic, it estimates ZIKV infections in the first-wave epidemic in the most-affected countries such as Brazil and Colombia, where the approach projects a median infection AR of 19 and 14%, respectively, which falls within the CI of the results provided here.

Although the initial introduction of ZIKV could date back to August 2013, most countries did not experience the first wave of the epidemic until the early months of 2016. Brazil is the only country that seems to have a well-defined first peak in March 2015, consistent with reports from the northeast region (45). The model suggests two epidemic waves in Brazil. The first wave, occurring between January and July 2015, corresponds to early outbreaks in the northeast region (Maranhao, Bahia, and Rio Grande do Norte) and later on in the rest of the country. This

numbers, suggesting reporting and detection rates ranging from 1% to about 6% depending on the country.

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. This includes the use of some expressions for temperature dependence of transmissibility that are modeled on DENV data. Although this assumption is plausible, more data specific to ZIKV are certainly needed. The model has been calibrated by using data from French Polynesia and the observed epidemic peak in Colombia (± 1 wk), as reported by Colombia's National Institute of Health (38); further research is needed to provide ZIKV-specific parameter estimates and more accurate local calibrations. Mosquito presence/absence maps are available from published data but have limitations as detailed in the literature (32, 34, 58). Sexual and other modes of transmission are not incorporated in the model. The sexual component of the transmission, however, might acquire relevance in areas where the mosquito-borne transmission has a small reproductive number and low incidence (9–12, 59). The specific socioeconomic features of airline travelers are 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 a 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 AR as well as the temporal pattern resulting from local weather, population-level characteristics, and human mobility:

- The model yields a probability distribution for the time and place of introduction of ZIKV in Brazil, generating a comprehensive picture of the past dynamics of the epidemic.
- The numerical simulations allow estimates of the spatiotemporal spread of ZIKV in the Americas through February 2017. In particular, it provides estimates for the infection ARs and epidemic timing in ZIKV affected countries.
- The integration of airline travel data allows the explicit estimation of the number of travel-related cases into the USA and other countries.
- The model allows estimation of the number of newborns from women infected by ZIKV during the first trimester of pregnancy and the potential number of microcephaly cases through 2017 assuming different levels of risk. These projections could be checked against observed data in the future.

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 observed surveillance data and provide indications for the magnitude and timing of the epidemic, as well as aid in planning for international and local outbreak response, and for the planning of phase-III ZIKV vaccine trial sites. The study presented here also provides a computational modeling framework that can potentially be generalized to other *Aedes*-transmitted vector-borne diseases, such as dengue and chikungunya.

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 (60–65). This model integrates high-resolution demographic, human mobility, socioeconomic (geon.yale.edu), and temperature data (climate.geog.udel.edu/~climate/html_pages/Global2011/GlobalTs2011.html); because no human subject research/analysis was performed, IRB approval was not required. Here we expanded

to incorporate data on *Aedes* mosquito density (58) and the association between socioeconomic factors and population risk of exposure (32, 66). Similar to previous arbovirus modeling approaches (18), 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 similarities between ZIKV and DENV.

Global Model for the Spread of Vector-Borne Diseases. The GLEAM model is a fully stochastic epidemic modeling platform that uses real-world data 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.ciesin.columbia.edu). The model considers geographical cells of $0.25^\circ \times 0.25^\circ$, corresponding to an approximately 25-km $\times$ 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 roughly 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

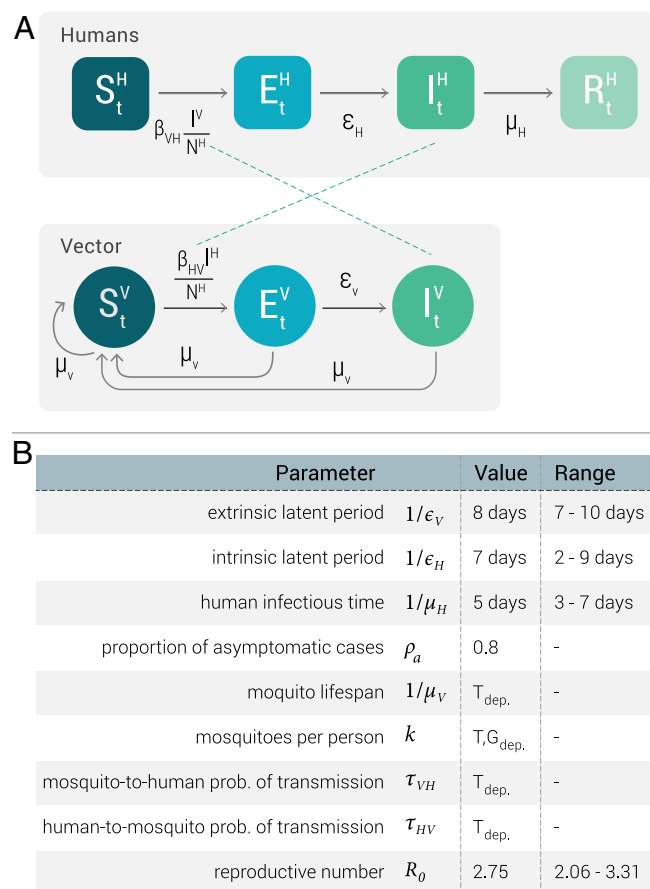


Fig. 7. (A) Compartmental classification for ZIKV infection. Humans can occupy one of the four top compartments: susceptible, which can acquire the infection through contacts (bites) with infectious mosquitoes; exposed, where individuals are infected but are not able yet to transmit the virus; infectious, 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 geolocation-dependent. Specific values for the parameters can be found in refs. 2, 4, 18, 55, and 68–70.

mechanistically simulated mobility and commuting patterns of disease carriers. Mobility includes global air travel (www.oag.com), 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 and seasonality in vector abundance, the temperature dependence modulating the vector competence, and the characteristics of the exposed populations. Many locations, such as those at high elevation, are not at risk for 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 datasets with high spatial resolution are integrated, including the following:

- Global terrestrial air temperature data: The global air temperature dataset (climate.geog.udel.edu/~climate/html_pages/Global2011/GlobalTst2011.html) contains monthly mean temperatures at a spatial resolution of $0.5^\circ \times 0.5^\circ$. 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 dataset. Daily average temperatures are linearly interpolated from each population's monthly averages.
- Global *A. aegypti* and *A. albopictus* distribution: The global *A. aegypti* and *A. albopictus* distribution database provides uncertainty estimates for the vector's distribution at a spatial resolution of $5 \text{ km} \times 5 \text{ km}$ (58).
- Geolocalized economic data: The geophysically scaled economic dataset (G-Econ), developed by Nordhaus et al. (67), maps the per capita Gross Domestic Product [GDP, computed at purchasing power parity (PPP) exchange rates] at a $1^\circ \times 1^\circ$ resolution. To estimate the per capita gross cell product at 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 affects 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 *SI Appendix, section 2*. Data on mosquito abundance and temperature are used to identify cells where ZIKV outbreaks are not possible because of environmental factors. The human populations in these cells are thus considered unexposed to ZIKV and susceptible individuals are assigned an environmental rescaling factor, r_{en} , as described *SI Appendix, section 3*. 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 these rescaling factors is provided in *SI Appendix, section 3*.

Once the data layers and parameters have been defined, the model runs using discrete time steps of one full day to simulate the transmission dynamic model (described in detail below), incorporating human mobility between subpopulations, and partially aggregating 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. From those realizations we find the probability $p(x)$ and $p(x|\theta)$, defined as the probability of the evidence (the epidemic peak in Colombia as from surveillance data) and the likelihood of the evidence given the parameters θ specifying the date and location of introduction of ZIKV in Brazil. From those distributions we

can calculate the posterior probabilities of interest. The sensitivity analysis for the others scenarios considers an additional 200,000 simulations in total.

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 (and may or may not display symptoms), and removed individuals R^H who no longer have the infection and are immune to further ZIKV infection. We consider the human population size to be constant, that is, $S^H + E^H + I^H + R^H = N^H$. 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. 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 (68). These infectious individuals then recover from the disease at a rate μ_H (18), 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 probability of transmission (69), and the number I^V of infected mosquitoes. Exposed mosquitoes transition to the infectious class at a rate ϵ_V , which 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 (70). The mosquito-to-human force of infection follows the usual mass-action law in each subpopulation whose linear extension varies from a few miles to about 50 miles depending on the population density and specific area of the world. A full description of the stochastic model and the equations is provided in *SI Appendix*.

A summary of the parameters defining the disease dynamics is reported in Fig. 7B. The empirical evidence related to the ZIKV infection 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 analysis of data reported from the 2013 ZIKV epidemic in French Polynesia (18). Setting the extrinsic and intrinsic latent periods and the human infectious period to reference values and using average daily temperatures of French Polynesia, we estimate a basic reproduction number at the temperature $T = 25^\circ\text{C}$ for French Polynesia $R_0^{\text{FP}} = 2.75$ (95% CI [2.53 to 2.98]), which is consistent with other ZIKV outbreak analyses (18, 31). Because the reproduction number depends on the disease serial interval, we report a sensitivity analysis in *SI Appendix* considering the upper and lower extremes of plausible serial intervals. Briefly, the estimated R_0^{FP} values are 2.06 (95% CI [1.91 to 2.22]) and 3.31 (95% CI [3.03 to 3.6]) for the shortest and longest serial intervals, respectively. The R_0 values are in the range of those estimated from local outbreaks in San Andres Island ($R_0 = 1.41$) and Girardot, Colombia ($R_0 = 4.61$) (71); however, it is worth recalling that the reproductive number depends on the location and on time through seasonal temperature changes. The calibration in French Polynesia provides the basic transmissibility of ZIKV. However, variations in temperature and mosquito abundance yield varying R_0 in each subpopulation tracked by the model as discussed in *SI Appendix*.

ACKNOWLEDGMENTS. This work was supported by Models of Infectious Disease Agent Study, National Institute of General Medical Sciences Grant U54GM111274, European Commission Horizon 2020 CIMPLEX Grant 641191, and a Colombian Department of Science and Technology Fulbright-Colciencias scholarship (to D.P.R.).

1. de Melo Freire CC, Iamarino A, de Lima Neto DF, de Andrade Zanotto PM (2015) Spread of the pandemic Zika virus lineage is associated with NS1 codon usage adaptation in humans. *bioRxiv* 032839.
2. Hayes EB (2009) Zika virus outside Africa. *Emerg Infect Dis* 15:1347–1350.
3. Dick G, Kitchen S, Haddock A (1952) Zika virus (I). Isolations and serological specificity. *Trans R Soc Trop Med Hyg* 46:509–520.

4. Duffy MR, et al. (2009) Zika virus outbreak on Yap Island, federated states of Micronesia. *N Engl J Med* 360:2536–2543.
5. Besnard M, et al. (2014) Evidence of perinatal transmission of Zika virus, French Polynesia, December 2013 and February 2014. *Euro Surveill* 19:20751.
6. D'Ortenzio E, et al. (2016) Evidence of sexual transmission of Zika virus. *N Engl J Med* 374:2195–2198.

7. Mlakar J, et al. (2016) Zika virus associated with microcephaly. *N Engl J Med* 374: 951–958.
8. Gao D, et al. (2016) Prevention and control of Zika fever as a mosquito-borne and sexually transmitted disease. *Sci Rep* 6:28070.
9. Yakob L, Kucharski A, Hue S, Edmunds WJ (2016) Low risk of a sexually-transmitted Zika virus outbreak. *Lancet Infect Dis* 16:1100–1102.
10. Althaus CL, Low N (2016) How relevant is sexual transmission of Zika virus? *PLoS Med* 13:e1002157.
11. Allard A, Althouse BM, Hébert-Dufresne L, Scarpino SV (2016) The risk of sustained sexual transmission of Zika is underestimated. *bioRxiv* 090324.
12. Moreira J, Lamas CC, Siqueira A (2016) Sexual transmission of Zika virus: Implications for clinical care and public health policy. *Clin Infect Dis* 63:141–142.
13. Musso D, et al. (2014) Potential for Zika virus transmission through blood transfusion demonstrated during an outbreak in French Polynesia, November 2013 to February 2014. *Euro Surveill* 19:20761.
14. Chouin-Carneiro T, et al. (2016) Differential susceptibilities of *Aedes aegypti* and *Aedes albopictus* from the Americas to Zika virus. *PLoS Negl Trop Dis* 10:e0004543.
15. Grard G, et al. (2014) Zika virus in Gabon (Central Africa)—2007: A new threat from *Aedes albopictus*? *PLoS Negl Trop Dis* 8:e2681.
16. Filipe A, Martins C, Rocha H (1973) Laboratory infection with Zika virus after vaccination against yellow fever. *Arch Gesamte Virusforsch* 43:315–319.
17. Ios S, et al. (2014) Current Zika virus epidemiology and recent epidemics. *Med Mal Infect* 44:302–307.
18. Kucharski AJ, 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.
19. Schuler-Faccini L (2016) Possible association between Zika virus infection and microcephaly – Brazil, 2015. Available at [dx.doi.org/10.15585/mmwr.mm6503e2](https://doi.org/10.15585/mmwr.mm6503e2). Accessed April 20, 2017.
20. World Health Organization (2016) WHO director-general summarizes the outcome of the emergency committee regarding clusters of microcephaly and Guillain-Barre syndrome. Available at www.who.int/mediacentre/news/statements/2016/emergency-committee-zika-microcephaly/en/. Accessed April 20, 2017.
21. Centers for Disease Control and Prevention (2016) Countries and territories in the Americas with active Zika virus transmission. Available at www.cdc.gov/zika/geo/active-countries.html. Accessed April 20, 2017.
22. Pan-American Health Organization (2016) Countries and territories with autochthonous transmission in the Americas reported in 2015–2016. Available at www.paho.org/hq/index.php?option=com_content&view=article&id=11603&Itemid=41696&lang=en. Accessed April 20, 2017.
23. Pan-American Health Organization (2016) Regional Zika epidemiological update (Americas) - 14 July 2016. Available at www.paho.org/hq/index.php?option=com_content&view=article&id=11599®ional-zika-epidemiological-update-americas&Itemid=41691&lang=en. Accessed April 20, 2017.
24. Petersen LR, Jamieson DJ, Powers AM, Honein MA (2016) Zika virus. *N Engl J Med* 374:1552–1563.
25. Lessler J, et al. (2016) Assessing the global threat from Zika virus. *Science* 353:aaf8160.
26. Rasmussen SA, Jamieson DJ, Honein MA, Petersen LR (2016) Zika virus and birth defects—reviewing the evidence for causality. *N Engl J Med* 374:1981–1987.
27. Abushouk AI, Negida A, Ahmed H (2016) An updated review of Zika virus. *J Clin Virol* 84:53–58.
28. Plourde AR, Bloch EM (2016) A literature review of Zika virus. *Emerg Infect Dis* 22:1185.
29. Boeuf P, Drummer HE, Richards JS, Scullar MJ, Beeson JG (2016) The global threat of Zika virus to pregnancy: Epidemiology, clinical perspectives, mechanisms, and impact. *BMC Med* 14:112.
30. Faria NR, et al. (2016) Zika virus in the Americas: Early epidemiological and genetic findings. *Science* 352:345–349.
31. Chowell G, et al. (2016) Using phenomenological models to characterize transmissibility and forecast patterns and final burden of Zika epidemics. *PLoS Curr* 8, 10.1371/currents.outbreaks.f14b2217c902f453d9320a43a35b9583.
32. Perkins AT, Siraj AS, Ruktanonchai CW, Kraemer MUG, Tatem AJ (2016) Model-based projections of Zika virus infections in childbearing women in the Americas. *Nat Microbiol* 1:16126.
33. Zinszer K, et al. (2017) Reconstruction of Zika virus introduction in Brazil. *Emerg Infect Dis* 23:91–94.
34. Messina JP, et al. (2016) Mapping global environmental suitability for Zika virus. *Elife* 5:e15272.
35. Bogoch II, et al. (2016) Anticipating the international spread of Zika virus from Brazil. *Lancet* 387:335–336.
36. Cauchemez S, et al. (2016) Association between Zika virus and microcephaly in French Polynesia 2013–15: A retrospective study. *Lancet* 387:2125–2132.
37. Johansson MA, Mier-y-Teran-Romero L, Reefhuis J, Gilboa SM, Hills SL (2016) Zika and the risk of microcephaly. *N Engl J Med* 375:1–4.
38. Instituto Nacional de Salud, Dirección de Vigilancia y Análisis del Riesgo en Salud Pública (2016) Boletín epidemiológico semanal. Available at www.ins.gov.co/boletin-epidemiologico/Boletin%20Epidemiologico%2016%20Boletin%20epidemiologico%20semana%2039.pdf. Accessed April 20, 2017.
39. Rodrigues LC (2016) Microcephaly and Zika virus infection. *Lancet* 387:2070–2072.
40. França GVA, et al. (2016) Congenital Zika virus syndrome in Brazil: A case series of the first 1501 livebirths with complete investigation. *Lancet* 388:891–897.
41. Centers for Disease Control and Prevention (2016) Case counts in the US. Available at <https://www.cdc.gov/zika/geo/united-states.html>. Accessed April 20, 2017.
42. Nelson B, et al. (2016) Travel volume to the United States from countries and US territories with local Zika virus transmission. Available at currents.plos.org/outbreaks/article/travel-volume-to-the-united-states-from-countries-and-u-s-territories-with-local-zika-virus-transmission/. Accessed April 20, 2017.
43. Castro LA, et al. (2016) Real-time Zika risk assessment in the United States. *bioRxiv* 056648.
44. Manore C, Ostfeld R, Augusto 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* 061382.
45. Cardoso CW, et al. (2015) Outbreak of exanthematous illness associated with Zika, Chikungunya, and Dengue viruses, Salvador, Brazil. *Emerg Infect Dis* 21:2274–2276.
46. Pan-American Health Organization (2016) Zika - Epidemiological Report Brazil. Available at www.paho.org/hq/index.php?option=com_docman&task=doc.view&gid=35221&Itemid=270&lang=en. Accessed April 20, 2017.
47. Lednicky J, et al. (2016) Zika virus outbreak in Haiti in 2014: Molecular and clinical data. *PLoS Negl Trop Dis* 10:e0004687.
48. Pacheco O, et al. (2016) Zika virus disease in Colombia — Preliminary report. *N Engl J Med*, 10.1056/NEJMoa1604037.
49. Pan-American Health Organization (2016) Zika - Epidemiological Report Venezuela (Bolivarian Republic of). Available at www.paho.org/hq/index.php?option=com_docman&task=doc.view&gid=35134&Itemid=270. Accessed April 20, 2017.
50. Pan-American Health Organization (2016) Zika - Epidemiological Report El Salvador. Available at www.paho.org/hq/index.php?option=com_docman&task=doc.view&gid=35225&Itemid=270&lang=en. Accessed April 20, 2017.
51. Pan-American Health Organization (2016) Zika - Epidemiological Report Honduras. Available at www.paho.org/hq/index.php?option=com_docman&task=doc.view&gid=35137&Itemid=270. Accessed April 20, 2017.
52. Pan-American Health Organization (2016) Zika - Epidemiological Report Puerto Rico. Available at www.paho.org/hq/index.php?option=com_docman&task=doc.view&gid=35231&Itemid=270&lang=en. Accessed April 20, 2017.
53. Pan-American Health Organization (2016) Zika - Epidemiological Report Mexico. Available at www.paho.org/hq/index.php?option=com_docman&task=doc.view&gid=35106&Itemid=270. Accessed April 20, 2017.
54. Riley S (2016) Epidemiology: Making high-res Zika maps. *Nat Microbiol* 1:16157.
55. Ferguson NM, et al. (2016) Countering the Zika epidemic in Latin America. *Science* 353:353–354.
56. Instituto Nacional de Salud, Dirección de Vigilancia y Análisis del Riesgo en Salud Pública (2017) Boletín epidemiológico semanal. Available at www.ins.gov.co/boletin-epidemiologico/Boletin%20Epidemiologico%2017%20Boletin%20C3%ADn%20epidemiologico%20C3%B3gico%20semana%2010.pdf. Accessed April 20, 2017.
57. Butler D (2016) Brazil asks whether Zika acts alone to cause birth defects. *Nature* 535:475–476.
58. Kraemer MU, et al. (2015) The global compendium of *Aedes aegypti* and *Ae. albopictus* occurrence. *Sci Data* 2:150035.
59. Miller JC (2017) Mathematical models of sir disease spread with combined non-sexual and sexual transmission routes. Available at [dx.doi.org/10.1016/j.idm.2016.12.003](https://doi.org/10.1016/j.idm.2016.12.003). Accessed April 20, 2017.
60. Balcan D, et al. (2010) Modeling the spatial spread of infectious diseases: The Global Epidemic and Mobility computational model. *J Comput Sci* 1:132–145.
61. Balcan D, 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 Med* 7:45.
62. Tizzoni M, et al. (2012) Real-time numerical forecast of global epidemic spreading: Case study of 2009 A/H1N1pdm. *BMC Med* 10:165.
63. Gomes MFC, et al. (2014) Assessing the international spreading risk associated with the 2014 West African Ebola outbreak. *PLoS Curr* 6. Available at [dx.doi.org/10.1371/currents.outbreaks.cd818f63d40e24aef769dda7df9e0da5](https://doi.org/10.1371/currents.outbreaks.cd818f63d40e24aef769dda7df9e0da5). Accessed April 20, 2017.
64. Poletto C, et al. (2014) Assessing the impact of travel restrictions on international spread of the 2014 West African Ebola epidemic. *Euro Surveill* 19:20936.
65. Zhang Q, et al. (2015) Social data mining and seasonal influenza forecasts: The Flu- Outlook platform. *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Lecture Notes in Computer Science (Springer, Cham, Switzerland), Vol 9286, pp 237–240.
66. Sissoko D, 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.
67. Nordhaus WD (2006) Geography and macroeconomics: New data and new findings. *Proc Natl Acad Sci USA* 103:3510–3517.
68. World Health Organization (2016) Zika virus. Available at www.who.int/mediacentre/factsheets/zika/en/. Accessed April 20, 2017.
69. Lambrechts L, et al. (2011) Impact of daily temperature fluctuations on dengue virus transmission by *Aedes aegypti*. *Proc Natl Acad Sci USA* 108:7460–7465.
70. Johansson MA, Powers AM, Pesik N, Cohen NJ, Staples JE (2014) Nowcasting the spread of Chikungunya virus in the Americas. *PLoS One* 9:e104915.
71. Rojas DP, et al. (2016) The epidemiology and transmissibility of Zika virus in Girardot and San Andres Island, Colombia. *Euro Surveill* 21:049957.
72. Health Ministry of Puerto Rico (2017) Informes de Zika por Municipios. Available at www.salud.gov.pr/Estadisticas-Registros-y-Publicaciones/Informes%20Zika%20por%20Municipios/Informe.Municipios%20semana%201-2017.pdf. Accessed April 20, 2017.