Reconstruction of Zika Virus Introduction in Brazil

Kate Zinszer, Kathryn Morrison, John S. Brownstein, Fatima Marinho, Alexandre F. Santos, Elaine O. Nsoesie

We estimated the speed of Zika virus introduction in Brazil by using confirmed cases at the municipal level. Our models indicate a southward pattern of introduction starting from the northeastern coast and a pattern of movement toward the western border with an average speed of spread of 42 km/day or 15,367 km/year.

Autochthonous transmission of Zika virus has been confirmed in 67 countries worldwide and in 46 countries or territories in the Americas (*1*,*2*). It is believed that Zika virus was introduced into the Americas through Easter Island in 2014, after an outbreak in French Polynesia (*3*,*4*). Despite the rapid spread of Zika virus across the Americas and global concerns regarding its effects on fetuses, little is known about the pattern of spread. The risk for local transmission in unaffected regions is unknown but potentially serious where competent Zika virus vectors are present (*5*) and also given the additional complexities of sexual transmission and population mobility (*3*,*6*).

Knowledge of the direction and speed of movement of a disease is invaluable for public health response planning, including timing and placement of interventions. We estimated the speed of Zika virus spread in Brazil by using data on confirmed cases of Zika virus disease at the municipal level and applying an approach used in estimating the speed of Ebola spread across parts of West Africa (*7*).

The Study

Confirmed cases of Zika virus disease were obtained from the Brazil Ministry of Health. Additional reports were also extracted from ProMED mail (*8*) and HealthMap (*9*). We performed the analysis by using 3 dates: 1) date of case registration in the surveillance system of the Brazilian Ministry of Health (model 1); 2) earliest of either date of symptom onset (if available) or registration date (model 2); and 3) earliest of either case registration date, date of symptom onset, or date of case report by other sources (model 3). Surface trend analysis was used to interpolated a continuous estimate of disease spread speed in magnitude and direction (*10*) by using available spatial and temporal information. Time of dispersal was calculated from the start of the epidemic for each model (online Technical Appendix, http://wwwnc.cdc.gov/EID/article/23/1/16-1274-Techapp1.pdf).

Data provided by the Brazilian Ministry of Health on May 31, 2016, indicated that Zika had been confirmed in 316 of 5,564 municipalities in 26 states; 6 additional municipalities were identified from other reporting sources. Contour maps of interpolated temporal trends (Figure 1) indicate a trend of spread into southern and western Brazil, and initial outbreak reports originated from municipalities along the northeastern coast. On the basis of confirmed cases, the earliest location of spread was the northeastern coastal area between the states of Paraíba, Ceará, Bahía, Alagoas, and Rio Grande do Norte. There were also earlier dates of self-reported symptom onset in the northwestern state of Amazonas (January 1, 2015), the west-central state of Matto Grosso (January 4, 2015), and the southeastern coastal state of Rio de Janeiro (January 1, 2015).

Contour maps (Figure 1) indicate slight differences in patterns of dispersion between the models. Model 1 indicates the strongest trend of a southward spread from the northeastern coast toward the populous southeastern coastal states of Rio de Janeiro, Espírito Santo, and São Paulo; the estimated time of dispersal was 22 weeks (Figure 1, panel A). In addition to west to east spread of Zika in southern Brazil, there was a pattern of movement west toward Bolivia.

The dispersal trend for model 2 was more varied but also indicated spread to the southeastern coastal states of Rio de Janeiro, Espírito Santo, and São Paulo (Figure 1, panel B). This model also suggests an initial spread north from the earliest reports in the northeastern region and a spread west toward Bolivia. The model estimates a north to south diffusion of ≈27 weeks. Model 3 suggests a strong southward spread originating from the northeastern coast toward the southeastern coastal states (approximate dispersal time of 29 weeks) and toward the western border and northwestern state of Amazonas (Figure 1, panel C).

Overall, the average speed of diffusion was 42.1 km/day or 15,367 km/year. The minimum speed across all 3 models was 6.9 km/day, and the maximum speed was 634.1 km/day (Figure 2). Municipalities in northeastern and northern regions had the slowest speeds, and municipalities in the west-

Author affiliations: Boston's Children's Hospital, Boston, Massachusetts, USA (K. Zinszer, J.S. Brownstein); McGill University, Montreal, Quebec, Canada (K. Morrison); Harvard Medical School, Boston (J.S. Brownstein); Ministry of Health, Brasilia, Brazil (F. Marinho, A.F. Santos); University of Washington, Seattle, Washington, USA (E.O. Nsoesie)

DOI: http://dx.doi.org/10.3201/eid2301.161274

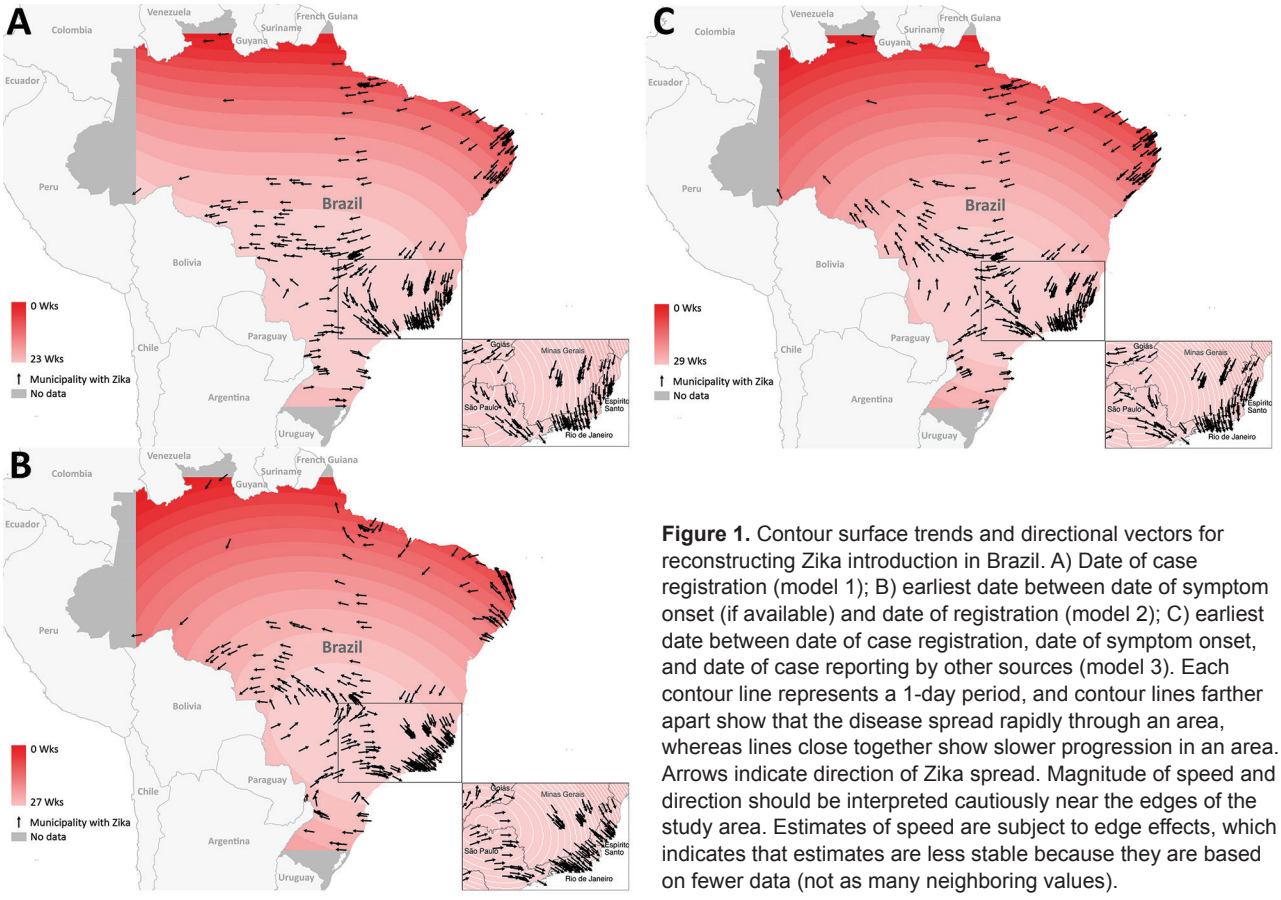

Figure 1. Contour surface trends and directional vectors for reconstructing Zika introduction in Brazil. A) Date of case registration (model 1); B) earliest date between date of symptom onset (if available) and date of registration (model 2); C) earliest date between date of case registration, date of symptom onset, and date of case reporting by other sources (model 3). Each contour line represents a 1-day period, and contour lines farther apart show that the disease spread rapidly through an area, whereas lines close together show slower progression in an area. Arrows indicate direction of Zika spread. Magnitude of speed and direction should be interpreted cautiously near the edges of the study area. Estimates of speed are subject to edge effects, which indicates that estimates are less stable because they are based on fewer data (not as many neighboring values).

central and southeastern regions had the highest speeds. This finding was caused by proximity of cases in time and space. More cases occurred closer in time and over larger areas in southern, southeastern, and west-central regions, which resulted in faster rates of case introduction.

All models were consistent in agreement that Zika dispersal in Brazil followed a general pattern of southward spread toward the populous coastal states (average speed of introduction of 42 km/day), which could be explained by multiple introductory cases into different areas probably caused by movement of viremic persons. We estimate that it took ≈5–6 months for Zika to spread from the northeastern coast to the southeastern coast and western border of Brazil. These findings are supported by the first report of local transmission of Zika virus in Paraguay in late November 2015 (*11*) and in Bolivia in January 2016 (*12*), 7 months after the first registered case in Brazil.

Limitations of this analysis include quality and timeliness of surveillance data that provided the basis for this study. Symptom onset date is subject to error because it is based on self-report, and earlier introductions of Zika in some municipalities might not have been captured by the Ministry of Health surveillance system and supplementary data sources, given the mild and generic nature of Zika symptoms and the high proportion of asymptomatic persons (*3*). The northern region of Brazil had a major dengue outbreak in early 2015, and given symptom similarities between dengue and Zika, it is probable that some suspected dengue cases were in fact early cases of Zika.

Sporadic geographically disparate cases were recorded in various parts of Brazil, which increased the uncertainty associated with speed analysis. These cases, such as those in northwestern Brazil, increased uncertainty in direction and speed estimates, which are also related to edge effects. Edge effects occurred along the boundary of the study area, which in this study were constructed by using fewer data points and are therefore less stable. This effect is shown with directional arrows pointing toward earlier areas of spread versus toward later areas of spread (Figure 1, panels B, C).

Conclusions

The arrival and rapid spread of Zika virus in the Americas resembles that of chikungunya virus, which was introduced into Saint Martin in the Caribbean in 2013 (*13*,*14*). Increased knowledge of the speed of spread and direction of Zika spread can help in understanding its possible future directions and pace at which it travels, which would be essential

for targeted mosquito control interventions, public health messages, and travel advisories. Future work will investigate underlying causes for the southward and westward spread in Brazil by incorporating mobility data and seasonal events, such as movement of persons between northeastern and southeastern regions for vacations, which could have driven the spatial transmission pattern. Furthermore, multicountry analysis is needed to understand continental spatial and temporal patterns of dispersion of Zika virus and co-circulating viruses, such as chikungunya virus.

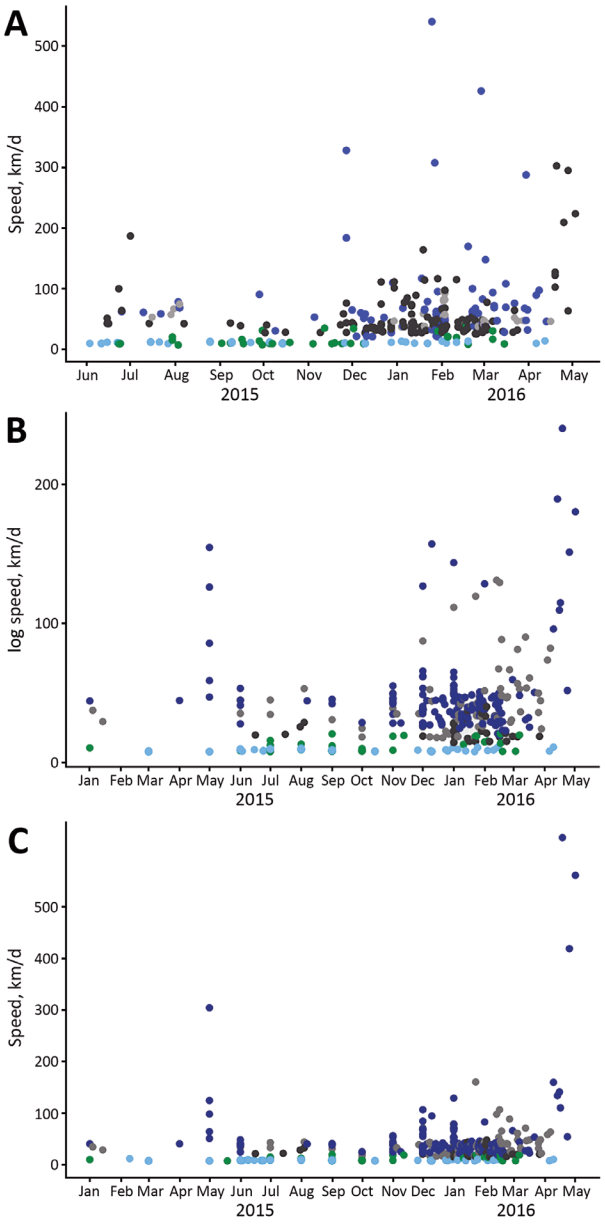

Figure 2. Speed or log speed (km/d) of Zika introduction into municipalities in Brazil. A) June 2015–May 2016; B) January 2015–May 2016; C) January 2015–May 2016. Municipalities are classified by region. Gray circles indicate central–western region, green circles indicate northern region, light blue circles indicate northeastern region, black circles indicate southern region, and dark blue circles indicate southeastern region.

K.Z. was supported by Canadian Institutes of Health Research; J.B. was supported by the National Library of Medicine, National Institutes of Health (grant R01LM010812); and E.N. was supported by the National Institute of Environmental Health Sciences, National Institutes of Health (grant K01ES025438).

Dr. Zinszer is a postdoctoral fellow in the Informatics Program, Boston Children's Hospital, Boston, Massachusetts. Her primary research interests include understanding patterns and determinants of emergence and spread of infectious diseases.

References

1. Centers for Disease Control and Prevention. All countries and territories with active Zika virus transmission [cited 2016 Aug 19]. http://www.cdc.gov/zika/geo/active-countries.html
2. World Health Organization. Zika virus and complications [cited 2016 Aug 19]. http://www.who.int/emergencies/zika-virus/en/
3. Basarab M, Bowman C, Aarons EJ, Cropley I. Zika virus. BMJ. 2016;352:i1049. http://dx.doi.org/10.1136/bmj.i1049
4. Musso D. Zika virus transmission from French Polynesia to Brazil. Emerg Infect Dis. 2015;21:1887. http://dx.doi.org/10.3201/eid2110.151125
5. Messina JP, Kraemer MU, Brady OJ, Pigott DM, Shearer FM, Weiss DJ, et al. Mapping global environmental suitability for Zika virus. eLife. 2016;5:e15272. http://dx.doi.org/10.7554/eLife.15272
6. Broutet N, Krauer F, Riesen M, Khalakdina A, Almiron M, Aldighieri S, et al. Zika virus as a cause of neurologic disorders. N Engl J Med. 2016;374:1506–9. http://dx.doi.org/10.1056/NEJMp1602708
7. Zinszer K, Morrison K, Anema A, Majumder MS, Brownstein JS. The velocity of Ebola spread in parts of West Africa. Lancet Infect Dis. 2015;15:1005–7. http://dx.doi.org/10.1016/S1473-3099(15)00234-0
8. Madoff LC. ProMED-mail: an early warning system for emerging diseases. Clin Infect Dis. 2004;39:227–32. http://dx.doi.org/10.1086/422003
9. Brownstein JS, Freifeld CC, Reis BY, Mandl KD. Surveillance Sans Frontières: Internet-based emerging infectious disease intelligence and the HealthMap project. PLoS Med. 2008;5:e151. http://dx.doi.org/10.1371/journal.pmed.0050151
10. Adjemian JZ, Foley P, Gage KL, Foley JE. Initiation and spread of traveling waves of plague, *Yersinia pestis*, in the western United States. Am J Trop Med Hyg. 2007;76:365–75.
11. World Health Organization. Zika virus infection: Paraguay [cited 2016 May 23]. http://www.who.int/csr/don/03-december-2015-zika-paraguay/en/
12. World Health Organization. Zika virus infection: Bolivia [cited 2016 May 23]. http://www.who.int/csr/don/20-january-2016-zika-bolivia/en/
13. Nsoesie EO, Kraemer MU, Golding N, Pigott DM, Brady OJ, Moyes CL, et al. Global distribution and environmental suitability for chikungunya virus, 1952 to 2015. Euro Surveill. 2016;21:30234. http://dx.doi.org/10.2807/1560-7917.ES.2016.21.20.30234
14. Musso D, Cao-Lormeau VM, Gubler DJ. Zika virus: following the path of dengue and chikungunya? Lancet. 2015;386:243–4. http://dx.doi.org/10.1016/S0140-6736(15)61273-9

Address for correspondence: Kate Zinszer, Computational Epidemiology Group, Boston Children's Hospital, 300 Longwood Ave, #3409, Boston, MA 02215, USA; email: kate.zinszer@mail.mcgill.ca

Article DOI: http://dx.doi.org/10.3201/eid2301.161274

Reconstruction of Zika Virus Introduction in Brazil

Technical Appendix

Suspected Zika cases, on the basis of the Pan American Health Organization and the World Health Organization definition (*1*), were confirmed by using reverse transcription PCR (99% of cases) and virus isolation in Vero or C6-36 cell lines. In addition to Ministry of Health data, Zika case data were obtained from HealthMap (*2*) and ProMED mail (*3*). HealthMap is an internet disease surveillance system based on media and outbreak reports from official public health sources (*2*), and ProMED mail is a moderated communication network for outbreak reporting and information exchange (*3*). Six additional municipalities were identified through the use of HealthMap and ProMED mail. From these data, we identified the registration date (date of entry into the surveillance system) of the first official confirmed Zika case in each municipality, the first symptom onset date for all confirmed cases, and any additional confirmed case dates from either HealthMap or ProMED mail.

Centroids of municipalities in Brazil were taken in meters from shapefiles and used to perform surface trend analysis. These data were geocoded by joining them to shapefiles for the municipalities (Universal Transverse Mercator zone 23 South projection) obtained from the DIVA Geographic System (*4*). Surface trend is a spatial interpolation method used to estimate continuous surfaces from point data. Traditionally, it has been used to model geographic elevation, but it also has been used to generate contour lines for representing disease spread across geographic space (*5*,*6*).

The response variable was time in days from first Zika case for each coordinate, which was January 1, 2015, for the earliest date between symptom onset and registration date (model 2); the earliest date between symptom onset, registration, or other case report (model 3); and June 3, 2015, for the registration date (model 1). The continuous surface of time to infection was estimated by regressing it against the X and Y coordinates. Time was in days and X and Y coordinates were meters. Parameters were estimated by using least squares regression, and if a

simple 2-dimensional plane through the points is insufficient to model the data, high-order polynomials are often used to capture local scale trends (*6,7*). We estimated models beginning with only linear terms by $f(t|X,Y) = \beta_0 + \beta_1 X + \beta_2 Y + \varepsilon$, where $E(\varepsilon) = 0$ (Equation 1).

We explored 10 models by incrementally adding polynomials up until the order 10. Every model beyond the linear model reported by the Pan American Health Organization (*1*) included an interaction term between X and Y:

$$f(t|x,y) = \beta_0 + \beta_1 X + \beta_2 Y + \beta_3 X^2 + \beta_4 Y^2 + \ldots + \beta_{19} X^{10} + \beta_{20} Y^{10} + \beta_{21} XY + \varepsilon \text{ (Equation 2).}$$

The best-fit model was selected by using R^2. The model with polynomial terms of order 3 provided the best fit for the registration date, and polynomial order 2 for the earliest date between symptom onset and registration date, and the earliest date between symptom onset, registration, or other case report.

The residuals for each of the models were assessed for spatial autocorrelation, but no significant correlation was observed in models beyond the linear model in (*1*). The rate of change was obtained by taking the partial derivatives with respect to X and Y, for the best-fit linear model, shown below as order 3 polynomial (*3*):

$$f(t|x,y) = \beta_0 + \beta_1 X + \beta_2 Y + \beta_3 X^2 + \beta_4 Y^2 + \beta_5 X^3 + \beta_6 Y^3 + \beta_7 XY \text{ (Equation 3);}$$

$$\partial f(t|x,y)/\partial x = \beta_1 + 2\beta_3 X + 3\beta_5 X^2 + \beta_7 Y \text{ (Equation 4);}$$

$$\partial f(t|x,y)/\partial y = \beta_2 + 2\beta_4 Y + 3\beta_6 Y^2 + \beta_7 X \text{ (Equation 5).}$$

Equations 4 and 5 provide expressions for a slope vector at a given location (X,Y). The vectors can be converted to express the magnitude and direction of rate of change (in days per kilometer) by finding the inner product of the vector, where magnitude $||xy|| = \sqrt{(x^2 + y^2)}$ and the direction $\theta = \tan^{-1}(y/x)$. Note that care must be used when applying the directions of the vectors (such as for vector field mapping); thus, the correct reference axis is used. The rate we were primarily interested in was speed (kilometers per day), which we obtained by inverting the final magnitude of the slope.

References

1. Pan American Health Organization. Case definitions [cited 2016 Jun 30]. http://www.paho.org/hq/index.php?option=com_content&view=article&id=11117&Itemid=41532&lang=en

2. Brownstein JS, Freifeld CC, Reis BY, Mandl KD. Surveillance Sans Frontières: Internet-based emerging infectious disease intelligence and the HealthMap project. PLoS Med. 2008;5:e151. http://dx.doi.org/10.1371/journal.pmed.0050151

3. Madoff LC. ProMED-mail: an early warning system for emerging diseases. Clin Infect Dis. 2004;39:227–32. http://dx.doi.org/10.1086/422003

4. DIVA GIS. Free spatial data [cited 2016 Mar 17]. http://www.diva-gis.org/Data

5. Berrang-Ford L, Berke O, Abdelrahman L, Waltner-Toews D, McDermott J. Spatial analysis of sleeping sickness, southeastern Uganda, 1970–2003. Emerg Infect Dis. 2006;12:813–20. http://dx.doi.org/10.3201/eid1205.051284

6. Pioz M, Guis H, Calavas D, Durand B, Abrial D, Ducrot C. Estimating front-wave velocity of infectious diseases: a simple, efficient method applied to bluetongue. Vet Res. 2011;42:60. http://dx.doi.org/10.1186/1297-9716-42-60

7. Ball FG. Front-wave velocity and fox habitat heterogeneity. In: Bacon PJ, editor. Population dynamics of rabies in wildlife. New York: Academic Press; 1985. p. 255–89.