

Version 1. PLoS Curr. 2018 July 25; 10:

ecurrents.outbreaks.00dd49d24b62731f87f12b0e657aa04c.

Published online 2018 July 25.

doi: 10.1371/currents.outbreaks.00dd49d24b62731f87f12b0e657aa04c:

10.1371/currents.outbreaks.00dd49d24b62731f87f12b0e657aa04c

Research Article

PMCID: PMC6126530

PMID: [30254787](https://pubmed.ncbi.nlm.nih.gov/30254787/)

Identifying Areas at Greatest Risk for Recent Zika Virus Importation — New York City, 2016

[Sharon K. Greene](#), [Sungwoo Lim](#), and [Annie Fine](#)

Sharon K. Greene, New York City Department of Health and Mental Hygiene;

[Contributor Information](#).

[Copyright](#) © 2018 Greene, Lim, Fine, et al

This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Introduction: The New York City Department of Health and Mental Hygiene sought to detect and minimize the risk of local, mosquito-borne Zika virus (ZIKV) transmission. We modeled areas at greatest risk for recent ZIKV importation, in the context of spatially biased ZIKV case ascertainment and no data on the local spatial distribution of persons arriving from ZIKV-affected countries.

Methods: For each of 14 weeks during June–September 2016, we used logistic regression to model the census tract-level presence of any ZIKV cases in the prior month, using eight covariates from static sociodemographic census data and the latest surveillance data, restricting to census tracts with any ZIKV testing in the prior month. To assess whether the model discriminated better than random between census tracts with and without recent cases, we compared the area under the receiver operating characteristic (ROC) curve for each week's fitted model versus an intercept-only model applied to cross-validated data. For weeks where the ROC contrast test was significant at $P < 0.05$, we output and mapped the model-predicted individual probabilities for all census tracts, including those with no recent testing.

Results: The ROC contrast test was significant for 8 of 14 weekly analyses. No covariates were consistently associated with the presence of recent cases. Modeled risk areas fluctuated across these 8

weeks, with Spearman correlation coefficients ranging from 0.30 to 0.93, all $P < 0.0001$. Areas in the Bronx and upper Manhattan were in the highest risk decile as of late June, while as of late August, the greatest risk shifted to eastern Brooklyn.

Conclusion: We used observable characteristics of areas with recent, known travel-associated ZIKV cases to identify similar areas with no observed cases that might also be at-risk each week. Findings were used to target public education and *Aedes* spp. mosquito surveillance and control. These methods are applicable to other conditions for which biased case ascertainment is suspected and knowledge of how cases are geographically distributed is important for targeting public health activities.

Keywords: importation, statistical model, Surveillance, Zika, zika virus

Introduction

The New York City (NYC) population, which includes many travelers and recent immigrants, is at risk for travel-associated communicable diseases, including Zika virus (ZIKV) disease.^{1,2,3} Infected persons who acquire ZIKV while traveling and are viremic in NYC might be bitten by *Aedes albopictus* vectors, introducing a risk of local, mosquito-borne transmission. Any such locally-acquired cases might not be easily recognized, as an estimated 80% of ZIKV infections are asymptomatic,⁴ clinical symptoms when present can be mild and non-specific, and healthcare providers would need to suspect local transmission to order ZIKV testing for patients without a relevant travel history.⁵

The risk of local, mosquito-borne ZIKV transmission in NYC during 2016 was considered low because the primary vector species, *Aedes aegypti*, is not present in NYC. Nevertheless, given limited evidence that *A. albopictus* could transmit ZIKV,^{6,7} a high population density, large numbers of potentially infected travelers, and the serious health consequences of congenital ZIKV infection including microcephaly,⁸ the NYC Department of Health and Mental Hygiene (DOHMH) was concerned about local transmission and aimed to detect such an unlikely occurrence as early as possible to institute measures to interrupt transmission to humans. DOHMH conducted intensive mosquito surveillance to assess the abundance of the mosquito vector and to test mosquitoes for ZIKV.

In parallel, DOHMH also sought to identify locations with higher numbers of persons with ZIKV viremia, as these would be areas where mosquitoes could become infected with ZIKV. There were several challenges. First, the geographic distribution of NYC residents at risk for travel-associated ZIKV infection is unknown. While data on the volume of commercial air passenger arrivals are available by airport,^{9,10} no data are readily available on where or whether these travelers reside within NYC. Second, while acute arboviral infections, including ZIKV infections, are reportable by healthcare providers and laboratories to NYC DOHMH,¹¹ these data are incomplete. Recently reported cases are likely a spatially non-representative sample of all currently viremic persons, given testing and reporting lags and the high proportion of asymptomatic and subclinical ZIKV infections.⁴ Furthermore, not all at-risk NYC patients were equally likely to receive ZIKV testing, given initial barriers to seeking care and receiving testing for ZIKV in areas with high poverty and large numbers of persons born in countries with local ZIKV transmission.^{3,12} Many cases are detected after viremia has ended and are only diagnosed serologically. Thus, mapping the distribution of the residences of re-

ported cases with recent diagnoses can be useful¹³ but is not necessarily sufficient to characterize all areas at highest risk of ZIKV importation. Third, at-risk areas need to be dynamically updated throughout the mosquito-borne disease transmission season to reallocate resources and target new areas as needed. We expected the geographic distribution within NYC of imported ZIKV cases could change over time, reflecting changing patterns of risk to populations living in different areas of NYC, caused by changes in incidence and outbreak dynamics across multiple ZIKV-affected countries.¹⁴

Given that reported ZIKV cases were incomplete and spatially non-representative, we aimed to identify additional areas in NYC potentially at risk for ZIKV importation by using area-level, static census data and the latest data available to DOHMH on ZIKV testing. In this study, we used logistic regression¹⁵ to predict weekly nowcasts¹⁶ throughout the summer of 2016 of census tracts at greatest risk of recent ZIKV importation. Nowcast results were used to inform geographically targeted activities, including performing public education, enrolling additional healthcare facilities in a sentinel surveillance system for detecting local ZIKV transmission, interpreting syndromic surveillance signals suggesting possible ZIKV-like illness, and, when reviewed in conjunction with mosquito surveillance data, informing control of *Aedes* spp. mosquitoes and placement of traps for continued surveillance.^{3,17}

Methods

Data Sources

The NYC population (an estimated >8.5 million persons as of July 2015)¹⁸ was eligible for analysis. The unit of analysis was 2010 census tract (n=2,123 in NYC with >25 residents), i.e., the finest geographic resolution available for all independent variables. We selected small geographic units to prioritize spatial precision in identifying areas at high risk, despite potential instability in estimates for some geographic units. Smaller units have more homogeneous risk factor distributions than larger units, minimizing inferential problems in ecologic analysis.¹⁹ Census tract-level sociodemographic data were obtained from the 2010 U.S. Census and the American Community Survey 2010–2014. De-identified ZIKV-related testing and case data were obtained from the disease surveillance database used by the DOHMH Bureau of Communicable Disease (Maven, Conduent Public Health Solutions, Austin, TX) and aggregated to census tract resolution prior to analysis.

Variables

The dependent variable was the census tract-level presence of any recent ZIKV cases reported to DOHMH. Cases included persons with confirmed or probable laboratory evidence of ZIKV infection or disease²⁰ and persons who tested positive for ZIKV by IgM with pending plaque reduction neutralization test results. A recent case was defined as having an “event date” (illness onset date if available, otherwise specimen collection date) in the 28 days prior to the weekly data extract. A 28-day period was selected to approximately encompass the risk of an imported ZIKV case leading to *A. albopictus* being currently or imminently infectious as of each analysis, considering a 7-day period of viremia in a human post-illness onset and a 14–20-day period before a mosquito becomes infectious.

The input dataset also contained eight independent variables; data for six of these variables were static. We included census tract-level sociodemographic characteristics that we suspected might be associated with the underlying risk of ZIKV importation and/or testing. As the distribution of persons at risk for travel-associated ZIKV infection is unknown, the first and second static variables were used as proxies for the number of travelers from selected areas with local mosquito-borne ZIKV transmission (Mexico, the Caribbean, Central America, and countries in South America):²¹ the number of persons born in these countries [1], grouped by quartile; and the number of persons with ancestry from these countries [2], grouped by quartile. Third, since travel-associated communicable disease incidence is consistently associated with area-based poverty in NYC,²² we included the proportion of the population living below the federal poverty level. Data for those first three variables were from the American Community Survey 2010–2014. We also included the proportion of the population of Hispanic ethnicity, the proportion of the population of women of childbearing age (15–44 years-old), and the quartile of the total population size, per the 2010 Census.

Data for the remaining two independent variables were updated weekly to reflect the most current available ZIKV-related data in the disease surveillance database. Since historical cases might be predictive of more recent cases, we included the number of ZIKV cases with an event date >28 days prior to the weekly data extract. Finally, since ZIKV, dengue virus, and chikungunya virus disease each occur among persons traveling from similar geographic areas, we included the cumulative number of confirmed and probable dengue and chikungunya cases since January 2013 who reported on interview from routine case investigation to have traveled to selected countries [3] with local mosquito-borne ZIKV transmission.²¹

An additional indicator variable, whether any ZIKV tests were ordered for residents of each census tract in the prior month, was updated weekly and used not as an independent variable, but for restriction in model fitting, as described in the next section. For most notifiable diseases, only positive laboratory results are reportable by law. Access to testing data was unusual but possible in this situation because ZIKV testing was coordinated by DOHMH and performed by public health laboratories before commercial testing became available.^{3,23} Time and climatic factors were not included in the model, since our objective was to use the most recent available data to answer a purely spatial (not explicitly spatio-temporal) question, i.e., identifying areas at greatest risk for recent ZIKV importation as of each weekly analysis.

Weekly Nowcasting Statistical Analysis

Preliminary analyses were performed May–June 2016, and 14 weekly analyses were conducted throughout the peak mosquito season for comparison with vector surveillance data, from the week starting June 26 through the week starting September 25, 2016. Each week, we used multivariable logistic regression using Firth’s penalized maximum likelihood estimation to model the presence of any recent ZIKV cases, given all eight covariates described above. We used regression modeling rather than spatial interpolation or smoothing methods because we did not want to inappropriately smooth results across adjacent neighborhoods with heterogeneous sociodemographic characteristics and travel patterns. We fit the model restricting to census tracts with any recent testing, i.e., excluding census tracts where having zero observed recent cases could be attributable to no testing.

Each week, we output the census tract-level leave-one-out cross-validated predicted probabilities²⁴ of any imported ZIKV cases in the prior month, using the OUTPUT statement in SAS (predprobs=cross-validate, SAS Institute, Cary, NC; see Supporting Information). In brief, the data for any given census tract were omitted, the logistic regression model was fit using data from all other census tracts, and then the predicted probability for the given census tract was estimated using the observed values for the given census tract and the modeled parameter estimates. To account for geographic variation in testing uptake, we required census tracts to have any ZIKV testing in the prior month to be included in model fitting. Model performance was assessed using receiver operating characteristic (ROC) curves, which determined the discriminatory accuracy of each week's model to identify census tracts with any recent imported ZIKV cases. An ROC contrast test was used to compare the area under the ROC curve (AUC) for the fitted model compared with an intercept-only model (i.e., random prediction, AUC=0.5) applied to the cross-validated data.²⁵ If the *P*-value from the ROC contrast test was <0.05, then we proceeded to output the individual predicted probabilities of any recent imported ZIKV cases (the “nowcast”) for all census tracts, including those with no recent testing.

We refined this modeling process during and after the mosquito season; thus, we present results of the finalized modeling process as retrospectively applied to weekly archived data extracts. These data extracts represented the data actually available to DOHMH in real-time and were not only incomplete because of reporting lags, but also preliminary because data cleaning was not yet completed. All observations (census tracts) were considered to be independent, ignoring spatial autocorrelation, because we needed unbiased point estimates but were not concerned with the variance of those estimates.²⁶

Any covariates significantly associated with the outcome each week according to the type III analysis of effects were noted. We assessed variation in nowcast results over time using Spearman rank correlation coefficients for the modeled ZIKV importation risk per census tract. In this analysis, the rank values for the census tract-level predicted probabilities of any recent imported ZIKV cases were compared between weeks. A Spearman rank correlation coefficient value of 1 would indicate no changes in the rankings between weeks (such that the same census tracts were always at greatest risk), a value of 0 would indicate no association in the ranks, and a value of -1 would indicate a perfect negative association of ranks. Census tracts in the highest decile of modeled ZIKV importation risk were visualized (ArcMap 10.2.1, Esri, Redlands, CA).

Ethics Statement

This activity involved the use of data collected for non-research purposes, and there were no interactions or interventions with living individuals. All efforts were made to protect individual privacy and anonymity. Data were de-identified before being accessed and used for the purpose of this activity. The scope of this activity was limited to public health practice, and all activities were authorized and conducted by NYC DOHMH, a public health authority that is responsible for such public health matters as part of its official mandate. This activity was categorized as public health surveillance by the NYC DOHMH's Institutional Review Board.

Results

During June–September 2016, 652 cases of ZIKV infection or disease were observed among residents of 475 census tracts in NYC. All reported cases were travel-associated, including travelers returning from ZIKV-affected areas, their sexual contacts, and infants infected in utero. Zero cases were observed in 1,648 (78%) of the 2,123 census tracts with >25 residents. Across the 14 weekly analyses, the median number of census tracts with any recent ZIKV testing and thus used to fit the weekly models was 688 (range: 516–827), and of these, the median number of census tracts with any recent ZIKV cases was 102.5 (range: 51–120).

Across the 14 weekly analyses, no covariates were consistently significantly associated with the presence of recent ZIKV cases (Table 1). One covariate (the proportion of the population of women of childbearing age) was significantly associated with ZIKV cases during 3 of the analysis weeks. Five covariates were each significantly associated with the outcome during 2 analysis weeks, 1 covariate was significantly associated during 1 analysis week, and 1 covariate was never significantly associated. During 4 weekly analyses (July 29, August 3, September 7, and September 13), no covariates were significantly associated with the outcome. Toward the beginning of the peak mosquito season, the covariates most strongly associated with recent ZIKV cases were historical counts of dengue, chikungunya, and ZIKV cases. By August, demographic factors (ethnicity, sex and age, and ancestry) instead were most strongly associated. By late September, only poverty level was significantly associated (Table 1).

According to the ROC contrast test, the model discriminated better than random between census tracts with and without recent cases for 8 of 14 weekly analyses (Table 1). In these weeks, the AUC ranged from 0.57 to 0.66, where AUC=0.5 represents random prediction and AUC=1 represents perfect prediction.

We proceeded to output nowcasts for these 8 weekly analyses, all of which were strongly correlated ($P < 0.0001$). However, variation in the modeled ZIKV importation risk per census tract was observed, with Spearman correlation coefficients ranging from 0.30 (indicating a weak positive association in the ranks of census tracts at risk between August 30 and September 20) to 0.93 (indicating a strong positive association in the ranks of census tracts at risk between August 16 and August 23, Table 2). The variation in Spearman correlation coefficients indicated shifts in areas at risk over time. For example, as of June 29, census tracts in the highest decile of modeled recent ZIKV importation risk were concentrated in the Bronx and upper Manhattan (Fig 1), while as of August 23, the greatest risk had largely shifted to eastern Brooklyn (Fig 1). In one defined Brooklyn neighborhood tabulation area consisting of 33 census tracts, 23 census tracts were in the highest decile of modeled risk as of August 23, but only 10 census tracts had any recent testing, and only 3 census tracts had observed recent cases.

Discussion

Our dynamic process used the most recent available data to nowcast how imported ZIKV cases were likely to be currently distributed in near real-time. While most published efforts to identify areas at risk of ZIKV importation have been at the country-level,^{27,28} local and state health departments with persons at risk for travel-associated diseases require more geographically refined estimates to inform

programmatic activities and to detect and mitigate local transmission.²⁹

By mapping the nowcast output, we visualized shifts in the spatial distribution of risk over time. For instance, as of late June, the greatest risk of recent ZIKV importation was concentrated in the Bronx and upper Manhattan, and the only covariate significantly associated with risk was historical counts of chikungunya and dengue cases, which had been similarly concentrated in these areas.³⁰ In contrast, by late August, the area at greatest risk had largely shifted to eastern Brooklyn, and risk was instead associated with specific population demographic factors, consistent with targeted interventions to increase ZIKV testing at facilities serving this population.^{12,31} Such shifts might reflect changing outbreak dynamics;¹⁴ i.e., as the force of infection peaked in different source countries at different times, different subpopulations of arriving travelers living in geographically distinct communities were therefore affected at different times, changing the geographic distribution of imported ZIKV cases in different parts of NYC. By definition of having a model with better than random discriminatory ability, we expected the nowcast to identify neighborhoods with recent observed cases. We noted situations where the nowcast also identified a neighborhood with a high proportion of at-risk census tracts but only a few observed cases. We also noted situations where the nowcast identified areas (groupings of census tracts) at-risk but with no recent observed cases, as this could inform public health operations, including targeting public education and setting traps for and controlling *Aedes* spp. mosquitoes.

Because ZIKV testing had been arranged primarily by DOHMH in 2016, we were able to restrict the model fitting to census tracts with recent testing. However, with most testing in 2017 done commercially and no access to overall testing data, such data will not be readily available going forward. Public health authorities might consider alternative methods of obtaining geographically resolved testing data, e.g., by partnering with commercial laboratories or regional health information organizations, or by making negative test results reportable. For example, in 2014, the NYC Health Code was amended to mandate reporting of negative hepatitis C nucleic acid test results.³² Even if testing data are not available, this method could still be useful in identifying areas where cases are not ascertained for reasons other than testing biases, e.g., because infections are asymptomatic and not medically attended.

This work is subject to at least three limitations. First, while internal cross-validation determined that the model predicted better than random for 8 of 14 analysis weeks, the maximum AUC across weeks was only 0.66. Model performance would likely improve if census tract-level data on persons arriving in NYC from ZIKV-affected countries were readily available. We displayed census tracts in the highest decile of modeled ZIKV importation risk as of a given week, although this threshold is arbitrary. It is unknown whether census tracts classified as high risk truly had any cases, as external validation would have required special outreach to test persons who would not otherwise have sought ZIKV testing and was not logistically feasible. Nowcasts might also be improved by excluding persons whose viremia resolved prior to arriving in NYC from a ZIKV-affected country. Second, analyses were based on patient residence and did not account for other locations in NYC where patients might have spent time while viremic. Third, the results were subject to the ecologic fallacy, such that inferences drawn from observing census tracts might not necessarily apply to all census tract residents.¹⁹ Nevertheless, the nowcasts were useful for targeting resources geographically.

The weekly nowcasts of modeled ZIKV importation risk were overlaid on vector surveillance data re-

garding mosquito abundance to improve situational awareness for DOHMH leadership and the ability to target preventive measures. No local ZIKV transmission was observed in NYC in 2016. These methods could be applied to other conditions for which biased case ascertainment is suspected and understanding the full geographic distribution of cases at any moment is important for targeting public health activities. Future work can include exploring complementary approaches for nowcasting and forecasting ZIKV importation and local transmission risk at high geographic resolution, e.g., by using agent-based models.³³

[1] Bahamas, Barbados, Belize, Bolivia, Brazil, Colombia, Costa Rica, Cuba, Dominica, Dominican Republic, Ecuador, El Salvador, Grenada, Guatemala, Guyana, Haiti, Honduras, Jamaica, Mexico, Nicaragua, Panama, Puerto Rico, Peru, St. Vincent and the Grenadines, Trinidad and Tobago, U.S. Island Areas, Venezuela, West Indies, and Other Caribbean. From ACS 2010–2014 tables [B05006](#): place of birth for the foreign-born population in the United States, and [B05002](#): place of birth by nativity and citizenship status. Countries not specified in these tables include: Aruba, Bonaire, Curacao, French Guiana, Guadeloupe, Martinique, Paraguay, Saint Barthélemy, Saint Lucia, Saint Martin, Sint Maarten, and Suriname.

[2] Bolivian, Brazilian, Colombian, Costa Rican, Cuban, Dominican, Ecuadorian, Guatemalan, Guyanese, Honduran, Mexican, Nicaraguan, Panamanian, Paraguayan, Peruvian, Puerto Rican, Salvadoran, Venezuelan, West Indian. From ACS 2010–2014 tables [B04006](#): people reporting ancestry, and [B03001](#): Hispanic or Latino origin by specific origin. Not specified in these tables include persons with ancestry from: Aruba, Bonaire, Curacao, French Guiana, Guadeloupe, Martinique, Saint Lucia, Saint Martin, Saint Vincent and the Grenadines, Sint Maarten, and Suriname.

[3] Aruba, Anguilla, Antigua and Barbuda, Barbados, Bahamas, Belize, Bolivia, Bonaire, Saint Eustatius and Saba, Brazil, British Virgin Islands, Cayman Islands, Colombia, Puerto Rico, Costa Rica, Cuba, Curacao, Dominica, Dominican Republic, Ecuador, El Salvador, French Guiana, Grenada, Guadeloupe, Guatemala, Guyana, Haiti, Honduras, Jamaica, Montserrat, Turks and Caicos Islands, Martinique, Mexico, Nicaragua, Panama, Paraguay, Peru, Saint Barthelemy, Saint Kitts and Nevis, Saint Lucia, Saint Martin, Saint Maarten, Saint Vincent and the Grenadines, Suriname, Trinidad and Tobago, U.S. Virgin Islands, and Venezuela.

Supporting Information

SAS code to generate the weekly nowcast of modeled risk of Zika virus importation and assess model performance is available here: <https://github.com/CityOfNewYork/communicable-disease-surveillance-nycdohmh>

Competing Interests

The authors have declared that no competing interests exist.

Funding

S.K.G. was supported by the Public Health Emergency Preparedness Cooperative Agreement (grant NU90TP000546) from the Centers for Disease Control and Prevention. S.L. and A.D.F. were supported by New York City tax levy funds. The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of the Centers for Disease Control and Prevention or the Department of Health and Human Services. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Data Availability

Census tract-level sociodemographic data were obtained from the 2010 U.S. Census and the American Community Survey 2010–2014 and are available from the U.S. Census website (<https://www.census.gov/data.html>). Census tract-level ZIKV testing and case data are not publically available from DOHMH in accordance with patient confidentiality and privacy laws; such data can be made available by contacting BCD_reportable_data@health.nyc.gov and obtaining institutional review board approval and executing a data use agreement approved by the legal departments of participating institutions.

Acknowledgments

Janette Yung wrote the SAS code to generate and refresh the data used in weekly analyses, and Jennifer Baumgartner contributed to reviewing this SAS code. Corinne Thompson assisted in running the weekly analyses. Sally Slavinski, Justin M. Cohen, and members of the Surveillance and Epidemiology Branch of the DOHMH Incident Command System provided useful input on the analysis plan and manuscript.

Contributor Information

Sharon K. Greene, New York City Department of Health and Mental Hygiene.

Sungwoo Lim, Bureau of Epidemiology Services, New York City Department of Health and Mental Hygiene, Queens, NY, USA.

Annie Fine, Bureau of Communicable Disease, New York City Department of Health and Mental Hygiene, Queens, NY, USA.

References

1. Adamson R, Reddy V, Jones L, Antwi M, Bregman B, Weiss D, Phillips M, Horowitz HW. Epidemiology and burden of hepatitis A, malaria, and typhoid in New York City associated with travel: implications for public health policy. *Am J Public Health*. 2010 Jul;100(7):1249-52. PubMed PMID:20466959. [PMCID: PMC2882402] [PubMed: 20466959]
2. Millman AJ, Chamany S, Guthartz S, Thihalolipavan S, Porter M, Schroeder A, Vora NM, Varma JK, Starr D. Active monitoring of travelers arriving from Ebola-affected countries - New York City, October 2014-April 2015. *MMWR Morb*

- Mortal Wkly Rep. 2016 Jan 29;65(3):51-4. PubMed PMID:26820056. [PubMed: 26820056]
3. Lee CT, Vora NM, Bajwa W, Boyd L, Harper S, Kass D, Langston A, McGibbon E, Merlino M, Rakeman JL, Raphael M, Slavinski S, Tran A, Wong R, Varma JK. Zika virus surveillance and preparedness - New York City, 2015-2016. *MMWR Morb Mortal Wkly Rep.* 2016 Jun 24;65(24):629-635. PubMed PMID:27337505. [PubMed: 27337505]
 4. Duffy MR, Chen TH, Hancock WT, Powers AM, Kool JL, Lanciotti RS, Pretrick M, Marfel M, Holzbauer S, Dubray C, Guillaumot L, Griggs A, Bel M, Lambert AJ, Laven J, Kosoy O, Panella A, Biggerstaff BJ, Fischer M, Hayes EB. Zika virus outbreak on Yap Island, Federated States of Micronesia. *N Engl J Med.* 2009 Jun 11;360(24):2536-43. PubMed PMID:19516034. [PubMed: 19516034]
 5. New York City Department of Health and Mental Hygiene (DOHMH). 2016 DOHMH Health Alert #28: Detecting Zika virus transmission in New York City (<http://www1.nyc.gov/assets/doh/downloads/pdf/cd/zika-advisory28.pdf>, accessed November 8, 2016).
 6. Wong PS, Li MZ, Chong CS, Ng LC, Tan CH. *Aedes (Stegomyia) albopictus* (Skuse): a potential vector of Zika virus in Singapore. *PLoS Negl Trop Dis.* 2013;7(8):e2348. PubMed PMID:23936579. [PMCID: PMC3731215] [PubMed: 23936579]
 7. Grard G, Caron M, Mombo IM, Nkoghe D, Mboui Ondo S, Jiolle D, Fontenille D, Paupy C, Leroy EM. Zika virus in Gabon (Central Africa)--2007: a new threat from *Aedes albopictus*? *PLoS Negl Trop Dis.* 2014 Feb 6;8(2):e2681. PubMed PMID:24516683. [PMCID: PMC3916288] [PubMed: 24516683]
 8. Rasmussen SA, Jamieson DJ, Honein MA, Petersen LR. Zika virus and birth defects--reviewing the evidence for causality. *N Engl J Med.* 2016 May 19;374(20):1981-7. PubMed PMID:27074377. [PubMed: 27074377]
 9. Wong DY, Quitugua T, Waters J. Comparison of air passenger travel volume data sources for biosurveillance. *Online Journal of Public Health Informatics.* 2016;8(1):e43 (<http://ojphi.org/ojs/index.php/ojphi/article/viewFile/6457/5264>).
 10. Huff A, Allen T, Whiting K, Breit N, Arnold B. FLIRT-ing with Zika: a web application to predict the movement of infected travelers validated against the current Zika virus epidemic. *PLoS Curr.* 2016 Jun 10;8. PubMed PMID:27366587. [PMCID: PMC4922883] [PubMed: 27366587]
 11. New York City Health Code Article 11: Reportable diseases and conditions (<https://www1.nyc.gov/assets/doh/downloads/pdf/about/healthcode/health-code-article11.pdf>, accessed October 14, 2016).
 12. Lee CT, Greene SK, Baumgartner J, Fine A. Disparities in Zika virus testing and incidence among women of reproductive age — New York City, 2016. *J Public Health Manag Pract.* 2017 [Epub ahead of print]. PubMed PMID: 29084118. [PubMed: 29084118]
 13. Rodriguez-Morales AJ, Patiño-Cadavid LJ, Lozada-Riascos CO, Villamil-Gómez WE. Mapping Zika in municipalities of one coastal department of Colombia (Sucre) using geographic information systems during the 2015-2016 outbreak: implications for public health and travel advice. *Int J Infect Dis.* 2016 Jul;48:70-2. PubMed PMID:27208632. [PubMed: 27208632]
 14. Ogden NH, Fazil A, Safronetz D, Drebot MA, Wallace J, Rees EE, Decock K, Ng V. Risk of travel-related cases of Zika virus infection is predicted by transmission intensity in outbreak-affected countries. *Parasit Vectors.* 2017 Jan 25;10(1):41. PubMed PMID:28122631. [PMCID: PMC5264286] [PubMed: 28122631]
 15. Shmueli G. To explain or to predict? *Statistical Science.* 2010;25(3):289–310.
 16. Codeço CT, Cruz OG, Riback TI, Degener CM, Gomes MF, Villela D, Bastos L, Camargo, Saraceni V, Lemos MCF, Coelho FC. InfoDengue: a nowcasting system for the surveillance of dengue fever transmission. *bioRxiv.* 2016;doi: <http://dx.doi.org/10.1101/046193>.

17. Wahnich A, Clark S, Bloch D, Kubinson H, Hrusa G, Liu D, Rakeman JL, Deocharan B, Jones L, Slavinski S, Stoute A, Mathes R, Weiss D, Connors EE. Surveillance for mosquito-borne transmission of Zika virus, New York City, NY, USA, 2016. *Emerg Infect Dis*. 2018 May;24(5):827-834. PubMed PMID: 29664375 [PMCID: PMC5938798] [PubMed: 29664375]
18. NYC Department of City Planning. New York City population: population facts (<http://www1.nyc.gov/site/planning/data-maps/nyc-population/population-facts.page>, accessed October 31, 2016).
19. Morgenstern H. Uses of ecologic analysis in epidemiologic research. *Am J Public Health*. 1982 Dec;72(12):1336-44. PubMed PMID:7137430. [PMCID: PMC1650553] [PubMed: 7137430]
20. U.S. Centers for Disease Control and Prevention/Council of State and Territorial Epidemiologists. Zika virus disease and Zika virus infection: 2016 case definition, approved June 2016 (<https://wwwn.cdc.gov/nndss/conditions/zika/case-definition/2016/06>, accessed February 14, 2017). 2016.
21. U.S. Centers for Disease Control and Prevention. All countries & territories with active Zika virus transmission (<http://www.cdc.gov/zika/geo/active-countries.html>, accessed May 24, 2016).
22. Greene SK, Levin-Rector A, Hadler JL, Fine AD. Disparities in reportable communicable disease incidence by census tract-level poverty, New York City, 2006-2013. *Am J Public Health*. 2015 Sep;105(9):e27-34. PubMed PMID:26180961. [PMCID: PMC4539797] [PubMed: 26180961]
23. Jacobs-Wingo J, Ezeoke I, Saffa A, Tate A, Lee D, Johnson K, Whittmore K, Illescas A, Collins A, Rand M, Rakeman JL, Varma JK, Vora NM. Using a call center to coordinate Zika virus testing-New York City, 2016. *J Emerg Manag*. 2016 Nov/Dec;14(6):391-395. PubMed PMID:28101877. [PubMed: 28101877]
24. Arlot S, Celisse A. A survey of cross-validation procedures for model selection. *Statistics Surveys*. 2010;4:40-79.
25. SAS Institute, Inc. Usage Note 39724: ROC analysis using validation data and cross validation (<http://support.sas.com/kb/39/724.html>, accessed May 17, 2017) 2015.
26. Dubin RA. Spatial autocorrelation: A primer. *Journal of Housing Economics*. 1998;7(4):304-327.
27. Grills A, Morrison S, Nelson B, Miniota J, Watts A, Cetron MS. Projected Zika virus importation and subsequent ongoing transmission after travel to the 2016 Olympic and Paralympic Games - Country-Specific Assessment, July 2016. *MMWR Morb Mortal Wkly Rep*. 2016 Jul 22;65(28):711-5. PubMed PMID:27442184. [PubMed: 27442184]
28. Nah K, Mizumoto K, Miyamatsu Y, Yasuda Y, Kinoshita R, Nishiura H. Estimating risks of importation and local transmission of Zika virus infection. *PeerJ*. 2016;4:e1904. PubMed PMID:27069825. [PMCID: PMC4824915] [PubMed: 27069825]
29. Imholte S. Public health practice problem definition: Assessing risk for emerging arboviral disease (Arizona Department of Health Services). Available from <https://www.surveillancerepository.org/assessing-risk-emerging-arboviral-disease-use-case>, accessed January 10, 2017: International Society for Disease Surveillance; 2016.
30. Little E, Bajwa W, Shaman J. Local environmental and meteorological conditions influencing the invasive mosquito *Ae. albopictus* and arbovirus transmission risk in New York City. *PLoS Negl Trop Dis*. 2017 Aug;11(8):e0005828. PubMed PMID:28832586. [PMCID: PMC5584979] [PubMed: 28832586]
31. Madad SS, Masci J, Cagliuso NV Sr, Allen M. Preparedness for Zika Virus Disease - New York City, 2016. *MMWR Morb Mortal Wkly Rep*. 2016 Oct 28;65(42):1161-1165. PubMed PMID:27787490. [PubMed: 27787490]
32. Board of Health, New York City Department of Health and Mental Hygiene. Notice of adoption of amendments to Articles

11 and 13 of the New York City Health Code (https://rules.cityofnewyork.us/sites/default/files/adopted_rules_pdf/f-dohmh-06-18-14_a_art._11_13.pdf, accessed May 21, 2017) 2014.

33. Parker J, Epstein JM. A Distributed Platform for Global-Scale Agent-Based Models of Disease Transmission. *ACM Trans Model Comput Simul*. 2011 Dec;22(1):2. PubMed PMID:24465120. [PMCID: PMC3898773] [PubMed: 24465120]

Figures and Tables

Variables†	Jun 29	Jul 7	Jul 11	Jul 22	Jul 29	Aug 3	Aug 9	Aug 16	Aug 23	Aug 30	Sep 7	Sep 13	Sep 20	Sep 27
Number of dengue and chikungunya cases since 2013 with travel to selected countries with local mosquito-borne ZIKV transmission	**	**												
Number of ZIKV cases with event dates >1 month prior to data extract			*	**										
Proportion of the population of Hispanic ethnicity							*		*					
Proportion of the population of women of childbearing age								**	**	**				
Number of persons with ancestry in selected countries with local mosquito-borne ZIKV transmission, grouped by quartile								*	**					
Total population size, grouped by quartile										**				
Proportion of the population living below the federal poverty level													*	**
Model performance														
AUC using cross-validation‡	0.64	0.59	0.60	0.57	0.53	0.54	0.52	0.64	0.66	0.61	0.55	0.55	0.64	0.57
P-value from ROC contrast test§	<.0001	0.005	0.001	0.02	0.24	0.15	0.63	<.0001	<.0001	0.0005	0.20	0.16	0.0002	0.14

* Significant at $0.05 \leq P < 0.1$.

** Significant at $P < 0.05$.

† The following variable was included in every model but was not statistically significant at $P < 0.1$ in any week: number of persons born in selected countries with local mosquito-borne ZIKV transmission, grouped by quartile.

‡ Area under the receiver operating characteristic (ROC) curve (AUC) for fitted model with cross-validation.

§ ROC contrast test comparing AUC for fitted model with cross-validation with an intercept-only model (i.e., $AUC=0.5$).

Table 1: Variables statistically significantly associated with recent Zika virus importation and model performance characteristics, by week of analysis, New York City, 2016

	Jul 7	Jul 11	Jul 22	Aug 16	Aug 23	Aug 30	Sep 20
Jun 29	0.76	0.69	0.72	0.59	0.59	0.48	0.61
Jul 7		0.88	0.67	0.59	0.57	0.53	0.41
Jul 11			0.78	0.69	0.62	0.59	0.38
Jul 22				0.70	0.70	0.59	0.31
Aug 16					0.93	0.80	0.43
Aug 23						0.90	0.40
Aug 30							0.30

*Spearman correlation coefficients; all $P < 0.0001$

Table 2: Correlations* of nowcasts for census tract-level risk of recent Zika virus importation as of eight time points, New York City, 2016

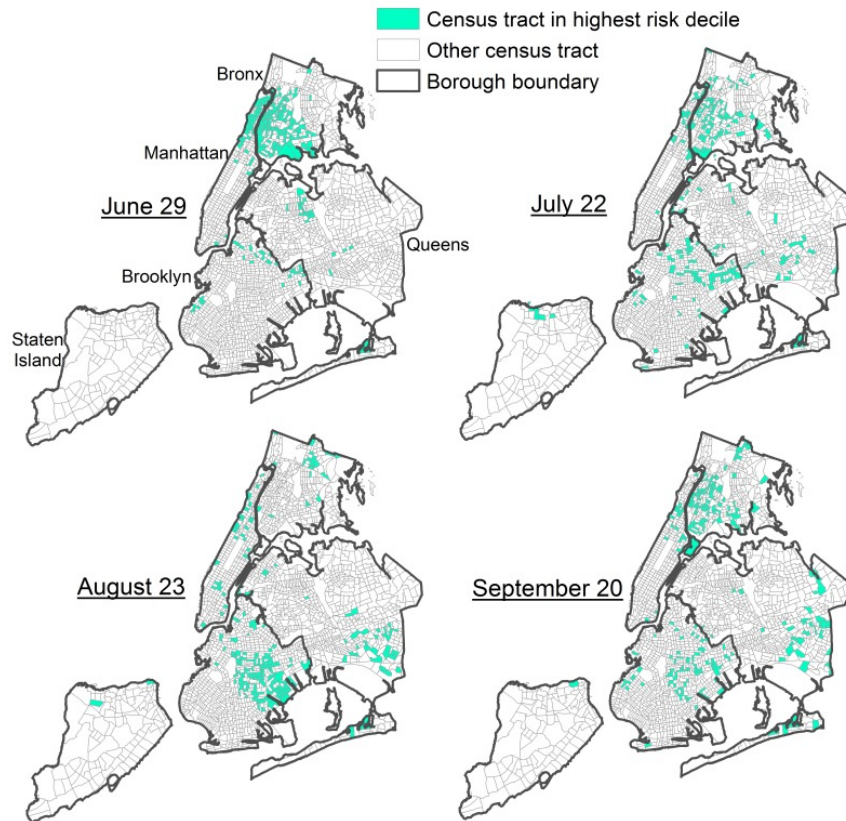


Fig. 1: Census tracts in the highest decile of modeled risk of Zika virus importation as of nowcast for four time points, New York City, 2016.