Spatial and Temporal Clustering of Chikungunya Virus Transmission in Dominica

Elaine O. Nsoesie1,2,3*, R. Paul Ricketts4, Heidi E. Brown5, Durland Fish6, David P. Durham7, Martial L. Ndeffo Mbah7, Trudy Christian8, Shalauddin Ahmed4, Clement Marcellin8, Ellen Shelly8, Katharine Owens6, Natasha Wenzel6, Alison P. Galvani7, John S. Brownstein1,2,9

1 Children’s Hospital Informatics Program, Boston Children’s Hospital, Boston, Massachusetts, United States of America, 2 Department of Pediatrics, Harvard Medical School, Boston, Massachusetts, United States of America, 3 Institute for Health Metrics and Evaluation, University of Washington, Seattle, Washington, United States of America, 4 Health Information Unit, Ministry of Health, Roseau, Commonwealth of Dominica, 5 Division of Epidemiology and Biostatistics, Mel and Enid Zuckerman College of Public Health, University of Arizona, Tucson, Arizona, United States of America, 6 Department of Epidemiology of Microbial Disease, Yale School of Public Health, New Haven, Connecticut, United States of America, 7 Center for Infectious Disease Modeling and Analysis, Yale School of Public Health, New Haven, Connecticut, United States of America, 8 Environmental Health Department, Ministry of Health, Roseau, Commonwealth of Dominica, 9 Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Quebec, Canada

* onelaine@vt.edu

Abstract

Using geo-referenced case data, we present spatial and spatio-temporal cluster analyses of the early spread of the 2013–2015 chikungunya virus (CHIKV) in Dominica, an island in the Caribbean. Spatial coordinates of the locations of the first 417 reported cases observed between December 15th, 2013 and March 11th, 2014, were captured using the Global Positioning System (GPS). We observed a preponderance of female cases, which has been reported for CHIKV outbreaks in other regions. We also noted statistically significant spatial and spatio-temporal clusters in highly populated areas and observed major clusters prior to implementation of intensive vector control programs suggesting early vector control measures, and education had an impact on the spread of the CHIKV epidemic in Dominica. A dynamical identification of clusters can lead to local assessment of risk and provide opportunities for targeted control efforts for nations experiencing CHIKV outbreaks.

Author Summary

Chikungunya is a disease transmitted by mosquitoes. Currently, there is an epidemic of chikungunya in several islands and countries in the Americas. Despite efforts at understanding and predicting spread, there have been no studies assessing the spatio-temporal spread of chikungunya in any of the Caribbean islands, mainly due to a lack of data. Here, we present a spatio-temporal analysis of the spread of chikungunya virus in Dominica, an island in the Western Hemisphere, using geo-referenced case data. The findings in this study suggest that females are at higher risk for chikungunya virus transmission in...
Dominica. In addition, there is statistically significant clustering of cases in densely populated areas. Lack of data prevented additional analyses on the impact of mosquito population density, environmental factors and housing conditions on the location and timing of the clusters. This study is relevant for chikungunya control in Dominica, and other regions can use similar methods to assess chikungunya risk at the local level.

Introduction

Chikungunya is an acute febrile illness that can cause incapacitating joint pain, high fever and skin rash. There are no estimates of the global burden of chikungunya, however, country-specific estimates have been as high as 45.26 DALYs (Disability adjusted life years) per million for India, where it is endemic [1]. Chikungunya is caused by chikungunya virus (CHIKV), a mosquito-borne pathogen that is transmitted to humans primarily through the bite of infected Aedes aegypti and Aedes albopictus mosquitoes [2–4]. Symptoms typically appear after an incubation period of 3 to 7 days [2,5,6].

Over the last ten years, CHIKV has emerged and re-emerged in locations including Kenya (2004), Comoros (2004, 2007), Seychelles (2004, 2006), Mauritius (2005), La Reunion (2005, 2007), and India (2005) [4,5,7–9]. In July 2007, the first outbreak in a non-tropical region was reported in the Emilia-Romagna region in Italy [10] and in December 2013, the first autochthonous case of chikungunya in the Western Hemisphere was reported in St. Martin, an island in the Caribbean [11,12]. Due to human movement and abundance of Aedes aegypti mosquitoes in the Americas, an estimated one million people were infected with the virus within one year of its introduction [13].

The first case of CHIKV in Dominica, an island in the Caribbean, involved a 65-year-old woman from Good Hope, on the east coast of Dominica. The affected individual had travelled to St. Martin from December 9th to 19th, 2013 and began experiencing symptoms on December 15th. Laboratory confirmation of diagnosis was received from the Caribbean Public Health Agency (CARPHA) on January 15th 2014. Active surveillance of CHIKV cases began shortly thereafter (January 16th, 2014) [14,15]. Autochthonous transmission of CHIKV in Dominica was confirmed by CARPHA on January 25th, 2014 [14].

We present spatio-temporal analysis of the early spread of CHIKV in a country in the Western Hemisphere using geo-referenced chikungunya case data. We assess the following: (1) distribution of reported cases by sex and age; (2) the presence of statistically significant spatial and spatio-temporal clusters and (3) rate of virus transmission as indicated by distance and date of disease onset between clustered cases. Specifically, we focus on the first 417 cases reported in Dominica. Dynamical assessment of clustering during outbreaks would aid in the identification of high-risk locations for vector control.

Materials and Methods

Study Location

Dominica is a small island nation (750 sq. km) with an estimated population of 71,293 [16]. This volcanic island’s elevation ranges from sea level along the perimeter where the majority of the population resides to an altitude of 1,447 meters (Morne Santé) in the center of the island. The capital, Roseau, is located on the southwestern coast of the island and is the largest community with over 1/5 of the total island population [16]. There are two seasons: a wet season that runs from June to December and a dry season that runs from January to May. Dominica is
administratively divided into ten parishes and is split into two health regions, which are further divided into seven health districts (Grand Bay, La Plaine, Roseau, Castle Bruce, St. Joseph, Marigot and Portsmouth) and fifty-two primary health centers.

Case Definition
At the start of the CHIKV epidemic in Dominica, the following World Health Organization (WHO) case definition was used: 

**Suspected case:** acute onset of fever > 38.5°C and severe arthralgia/arthritis not explained by other medical conditions, and resides or has visited epidemic or endemic areas within 2 weeks prior to the onset of symptoms \[14,15\]. However, once sustained local transmission had been established, the definition changed to: 

**Acute onset of fever (>38°C) and arthralgia/arthritis with or without headache, nausea, vomiting and atypical manifestation** \[14\]. Case confirmation was based on virus detection using real-time PCR, IgM ELISA and plaque-reduction neutralization test, as appropriate. Most infant cases had fever and skin rash and were either born into a family with CHIKV infection or a community with a CHIKV outbreak.

Data Collection
The data provided by the Dominican Ministry of Health (MOH) was de-identified and each case was represented by a unique reference identification code. Information on all suspected cases was collected using a standardized questionnaire, which covered population demographics (age, sex), symptoms, geographic location (e.g., town or village) and occupation (reported for some cases). The geographic location and health district was reported for each case. With the aid of local public health workers and using a Global Positioning System (GPS) receiver, we successfully recorded the geographic coordinates of the home address for 417 of the first 500 cases of the outbreak. The data included cases with symptom onset from December 15th, 2013 to March 11th, 2014. No personal identifiers were present and maps presented in this paper do not identify patients’ houses.

Analysis
We first summarized the distribution of cases across sex and age groups. To quantify the space-time interaction of individual reported chikungunya cases, we used the Knox method \[17–19\]. The method tests for interaction between cases with respect to distance and time, by comparing the observed to the expected number of cases in a specific space-time window. We selected distance and time intervals of 100m and 20 to 30 days range to account for the dispersal distance of \textit{Ae. Aegypti} \[17,20,21\] and maximum sum of the CHIKV incubation period in both the vectors and humans, respectively. The critical chi-square values for the null hypothesis of spatial randomness were estimated based on 999 Monte Carlo simulations.

Based on the results from the Knox method, we applied a space-time permutation model and a Poisson purely spatial model to identify independent high-risk clusters and assess locations and timing of case clusters during the thirteen-week epidemic period. These methods have been used in clustering of cases to guide control programs during other infectious disease outbreaks \[22–24\]. The Poisson spatial model assumes that the number of cases in each town or village is Poisson distributed and under the null hypothesis, the risk and expected number of cases are proportional to the population size. Detailed description of the statistical methodology for the Poisson purely spatial model can be found in \[25\]. The space-time permutation algorithm performed with SaTScan 9.3.1 moves a circular scanning window over the study area, and evaluates thousands of overlapping scanning windows in space \[23,26\]. The height of each cylinder represents a time interval and the base is a geographical region around a centroid.
with a radius ranging from 1% to 50% of the population at risk [27]. The number of observed cases, the number of expected cases, and the Poisson generalized likelihood ratio (GLR) are estimated for each cylinder. The maximum GLR from the observed data is compared to the maximum GLRs from 999 random Monte Carlo simulations under the null hypothesis of no clustering. A p-value was used to indicate the statistical significance of each cluster and significance was assessed at the 0.05 level. The first analysis examined the clustering of cases occurring within a temporal window of 50% of the study period (default setting). The second analysis examined clustering within three day overlapping intervals (1–3 days, 2–4 days, 3–5 days, . . . 28–30 days) and five distances (100m, 200m, 300m, 400m, 500m) to respectively account for uncertainty in the reported date of illness onset, and assess robustness of selected distances [28]. We also included age and sex as covariates. Cases with identical coordinates were represented by a single location, resulting in 353 and 355 unique locations for the models with and without covariates. Statistical analysis and mappings were performed in SaTScan 9.3.1, R 2.10.1 (http://www.r-project.org/) and QGIS v2.4.

Ethics Statement
The Institutional Review Board (IRB) at Boston Children’s Hospital approved this study.

Results
Case Description
Of the 417 cases (Figs 1 and 2), 66 were laboratory confirmed and 250 (60%) were female. The female to male odds ratio was 1.6, implying the odds of a reported case being female was 1.6 times the odds of being male. In addition, the male/female sex ratio was 0.67, which is similar to observations in other CHIKV studies [29–32]. The sex-specific incidence rates were 458.7 and 716.7 per 100,000 persons for males and females, respectively, despite a slightly higher number of males (36,411) than females (34,882) in the population [16]. This preponderance of female cases mostly concerned the age groups of 20–39 years (61.7% female vs. 38.3% male) and 40–59 (60.2% vs. 39.8%). The difference in the younger (19 years and less) and elderly populations (60 years and over) was less pronounced; 58.7% female vs. 41.3% male, and 57.1% vs. 42.9%, respectively. The median age for all cases was 33 years (min: 1, max: 92).

We also disaggregated the data into confirmed and suspected cases (Table 1). The disaggregated data had a male/female case ratio of approximately 0.57 and 0.69 for confirmed and suspected cases respectively. The distribution of cases across the various age groups was slightly different, however, a higher proportion of cases were in the under-40 age groups.

Spatiotemporal Interaction between Cases
The Knox test indicated significant spatiotemporal interaction between cases ($\chi^2 = 158.8; P < 0.001$) with maximum distance and time fixed at 100m and 20 days respectively. As previously stated, we fixed the cluster detection to cases that were close in space (100m) and time (20 to 30 days) to reflect the biology of the system. The Knox test statistics was also statistically significant on all other days (21 to 30) with $\chi^2$ in the range 83.5 to 153.9, and $P < 0.001$.

Spatial Clustering
SaTScan analysis detected two spatial clusters over the outbreak period using the Poisson model: one in the Bath Estate/Elmshall community (log likelihood ratio = 38.067, $P < 0.001$) and one in the Wesley, Woodford Hill community (log likelihood ratio = 35.222, $P < 0.001$) (Fig 3). The clusters contained 78 and 76 cases respectively, and the relative risk compared to
the baseline was estimated at 3.51 and 3.37. Bath Estate/Elmshall and Wesley, Woodford Hill communities are located in St. George and St. Andrew Parishes, respectively.

Spatio-temporal Clustering

The SaTScan space-time permutation model identified three statistically significant clusters (Fig 4). The primary cluster ($P < 0.001$) had 51 cases. The secondary clusters had 10 ($P < 0.001$) and 3 cases ($P = 0.021$). The clusters were located in the St. George, St. Andrew and St. George Parishes, respectively. The time frame for each cluster ranged from a single day to five weeks.

The space-time permutation model was adjusted for age and sex to account for the nonhomogeneous distribution across these groups (Table 1). Age had not been recorded for five patients so these cases were excluded from the analysis. With the addition of age and sex as covariates, two statistically significant clusters were identified. The primary cluster was the same as the cluster identified prior to age and sex adjustment. The secondary cluster ($P < 0.001$) in the St. George Parish was also similar to the previously identified cluster with two fewer cases.

Clustering Rate

We assessed the rate of clustering by varying the temporal interval and distance between cases. For the models without covariates, statistically significant spatial clusters were detected at all temporal intervals (min: 2 days, max: 4 days). The temporal interval with the maximum number of spatial clusters was 1–3 days for cases within 100m. These clusters were independently
observed on January 29–31, 2014 (26 cases, $P < 0.001$), February 10–12, 2014 (6 cases, $P = 0.024$), February 16, 2014 (3 cases, $P = 0.036$), and March 14, 2014 (3 cases, $P = 0.005$). In contrast, the maximum and minimum number of spatial clusters were 3 and 1, respectively, for models with covariates. Three clusters were observed for all spatial distances for temporal intervals 9–12 days; 5–10 days for 300m, 400m, and 500m and 11–14 days interval for cases within 100m. The clustering pattern was relatively consistent for the intervals 14–30 days for the models with and without covariates. The largest clusters for models with/without covariates were observed during the first few weeks of the outbreak.

**Discussion**

This study presents spatio-temporal analysis of the early spread of CHIKV in a country in the Western Hemisphere using geo-referenced data. This constitutes an important step for

![Fig 2. Epidemic curve by week of symptom onset; (a) case classification status and (b) sex. Reports start on epidemiologic week 51 of 2013 and end on week 11 of 2014. doi:10.1371/journal.pntd.0003977.g002](image)

<table>
<thead>
<tr>
<th>Population Characteristics</th>
<th>Total</th>
<th>Confirmed</th>
<th>Suspected</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n (%)</td>
<td>n (%)</td>
<td>n (%)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>167(40.0)</td>
<td>24(36.4)</td>
<td>143(40.7)</td>
</tr>
<tr>
<td>Female</td>
<td>250(60.0)</td>
<td>42(63.6)</td>
<td>208(59.3)</td>
</tr>
<tr>
<td>Age group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19 years and less</td>
<td>104(1.2)</td>
<td>20(30.3)</td>
<td>84(23.9)</td>
</tr>
<tr>
<td>20–39 years</td>
<td>141(11.8)</td>
<td>18(27.3)</td>
<td>123(35.0)</td>
</tr>
<tr>
<td>40–59 years</td>
<td>118(28.3)</td>
<td>15(22.7)</td>
<td>103(29.3)</td>
</tr>
<tr>
<td>60 years and over</td>
<td>49(33.8)</td>
<td>13(19.7)</td>
<td>36(10.3)</td>
</tr>
<tr>
<td>Missing</td>
<td>5(24.9)</td>
<td>0(0)</td>
<td>5(1.4)</td>
</tr>
</tbody>
</table>

| doi:10.1371/journal.pntd.0003977.t001 |
understanding CHIKV spread in the Caribbean, and similar analysis in countries with ongoing CHIKV outbreaks will aid in assessing local CHIKV risk.

Our space-time analyses of the early spread of CHIKV in Dominica identified chikungunya case clusters and demonstrated heterogeneity of spread at the local level. Both the spatial and space-time analyses identified clusters in the St. George and St. Andrew Parishes. While the population density for St. George Parish is the highest at 964 persons per sq. mile, that of St. Andrew is much lower at 137 persons per sq. mile [16]. Population density has been considered a contributor to dengue epidemics [33,34], however, the cluster observations suggest that population density might impact the size but not the occurrence of clustering. Furthermore, two of the three space-time clusters preceded any vector control activities, suggesting some potential impact of vector control programs and education on the spread of CHIKV in Dominica. Additional analyses to investigate the impact of entomological and environmental factors on the cluster locations was not possible due to a lack of detailed entomological data, and unreliable land use and land cover data. Due to the length of the study period, we did not expect environmental covariates such as precipitation and temperature to have a major impact on the timing and location of clusters.

Fig 3. (a) Location of the ten Parishes of Dominica. (b) Spatial clusters identified by SaTScan. The clusters are located in St. Andrew and St. George Parishes.

doi:10.1371/journal.pntd.0003977.g003
The strongest clustering in the temporal interval and spatial cluster analysis was observed for cases within 100m and with 1–3 days between reported symptom onsets. The spatial extent of 100m is consistent with the known dispersal distance of female \textit{Ae. Aegypti} [17,20,21,28]. Additionally, there were 10 unique locations for the 23 cases in the largest cluster suggesting that multiple cases were in the same household. Reports of cases within close proximity over a short time period could be due to transmission by multiple mosquitoes that became infected at about the same time, or a single mosquito feeding on multiple nearby hosts [28,35–37].

The results also suggest that sex and age could also have some impact on the spread of CHIKV in Dominica. The incidence of female cases is much higher than male cases in total, as well as across different age groups. The higher number of female cases could be due to several factors including greater exposure and health seeking behavior [15], and socio-economic factors such as type and location of occupation. Unfortunately data on type and location of occupation was not available for most cases to further investigate these hypotheses.

![Map showing statistically significant space-time clusters identified by SaTScan](doi:10.1371/journal.pntd.0003977.g004)
There are a few factors that could have affected the findings of this study. First, individuals who experience mild illness, and asymptomatic cases (3%-25% [2,38–40]) are less likely to seek medical assistance and subsequently will not be captured by the surveillance system. To mitigate this limitation, active surveillance is needed and differentiation between travel-related and autochthonous transmission for the duration of the epidemic would be useful. In addition, clustering was solely based on patients’ household locations although transmission could have occurred at other locations such as schools and workplaces. Furthermore, the assumed circular shape of the clusters limits the identification of irregular clusters.

As of January 2014, the Environmental Health Department of Dominica has actively performed household inspections for all newly identified cases for vector control purposes, and treated bed nets from CARPHA were distributed starting in February 2014. Additionally, there has been fogging throughout the island and indoor residual spraying (IRS) at the homes of suspected cases [14,15]. Information on CHIKV prevention and vector control was also distributed through pamphlets, television and radio messages, social media, text messages, and community education sessions. These early interventions and surveillance efforts could have had an impact on the spread of CHIKV on the island.

Studies like this could be useful for early evaluation of case distribution and clustering to provide an assessment of risk at a finer geographical scale, identification of locations where vector control is most needed, and parameterization of CHIKV transmission models. Dynamic identification of clusters can lead to targeted local control efforts. However, it is difficult to conduct similar analyses for several of the CHIKV-affected nations in the Americas due to limitations in available data. Chikungunya surveillance can be improved by combining passive notification and active case detection. Consistent surveillance efforts, and detailed data are important to assess and control the spread of CHIKV. Additionally, resource-constrained regions can easily perform similar analysis due to the availability of open source software and geospatial data. Information presented in the paper can be used to guide future epidemiological studies to better understand the emergence of CHIKV in Dominica and elsewhere.

Acknowledgments

We thank the Dominican Ministry of Health for providing the data used in this study.

Author Contributions

Analyzed the data: EON. Contributed materials: RPR SA CM ES TC KO NW DF APG JSB. Drafted the manuscript: EON. Edited the manuscript: EON RPR HEB DF DPD MLNM TC SA JSB. Approved the manuscript: EON RPR HEB DF DPD MLNM TC SA CM ES APG JSB.

References


