Dengue on islands: a Bayesian approach to understanding the global ecology of dengue viruses

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

Citation

Published Version
doi:10.1093/trstmh/trv012

Citable link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:17295761

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA
Dengue on islands: a Bayesian approach to understanding the global ecology of dengue viruses

Leora R. Feldstein, John S. Brownstein, Oliver J. Brady, Simon I. Hay and Michael A. Johansson*

*Children’s Hospital Informatics Program, Boston Children’s Hospital, 1 Autumn St., Boston, MA 02215, USA; Center for Statistics and Quantitative Infectious Diseases, Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA; Department of Pediatrics, Harvard Medical School, 1 Autumn St., Boston, MA 02215, USA; Spatial Ecology and Epidemiology Group, Department of Zoology, University of Oxford, South Parks Road, Oxford, OX1 3PS, UK; Fogarty International Center, National Institutes of Health, Bethesda, MD, USA; Dengue Branch, Division of Vector-Borne Diseases, CDC, 1324 Calle Canada, San Juan, PR 00920, USA

*Corresponding author: Tel: +1 787 706 2399; E-mail: mjohansson@cdc.gov

Received 7 October 2014; revised 22 January 2015; accepted 29 January 2015

Background: Transmission of dengue viruses (DENV), the most common arboviral pathogens globally, is influenced by many climatic and socioeconomic factors. However, the relative contributions of these factors on a global scale are unclear.

Methods: We randomly selected 94 islands stratified by socioeconomic and geographic characteristics. With a Bayesian model, we assessed factors contributing to the probability of islands having a history of any dengue outbreaks and of having frequent outbreaks.

Results: Minimum temperature was strongly associated with suitability for DENV transmission. Islands with a minimum monthly temperature of greater than 14.8°C (95% CI 12.4–16.6°C) were predicted to be suitable for DENV transmission. Increased population size and precipitation were associated with increased outbreak frequency, but did not capture all of the variability. Predictions for 48 testing islands verified these findings.

Conclusions: This analysis clarified two key components of DENV ecology: minimum temperature was the most important determinant of suitability; and endemicity was more likely in areas with high precipitation and large, but not necessarily dense, populations. Wealth and connectivity, in contrast, had no discernable effects. This model adds to our knowledge of global determinants of dengue risk and provides a basis for understanding the ecology of dengue endemicity.

Keywords: Dengue, Ecology, Epidemiology, Islands, Transmission dynamics

Introduction

Dengue viruses (DENV), vector-borne viruses of four distinct serotypes, have been estimated to cause as many as 280–530 million infections per year1 resulting in a range of manifestations from mild febrile illness to severe disease and death.2,3 Although DENV are widely distributed and are considered to be the most important arboviruses globally, there remains substantial uncertainty about transmission in some locations4 and about the most critical, population-level risk factors that contribute to either sporadic outbreaks or year-round transmission.

The dynamics and geography of DENV transmission have been extensively studied in a number of locations around the world. Since the early 1900s, it has been clear that DENV is transmitted by Aedes aegypti and Ae. albopictus mosquitoes. Populations of these species are limited by environmental conditions, such as temperature, precipitation and humidity.5 Their ability to transmit DENV is further dependent on temperature conditions which support DENV replication and dissemination within the mosquito,6,7 the survival of the adult female mosquito through that process8 and further feeding activity after becoming infectious.9 Socioeconomic factors are also important determinants of DENV transmission, particularly in relation to the frequency of contact between humans and mosquitoes. For instance, in areas with limited infrastructure, water storage practices10–12 and trash accumulation13–17 can provide aquatic habitats for immature mosquitoes in close proximity to homes. Meanwhile, the use of screens on windows and air conditioning can reduce mosquito–human interaction.18 Beyond a suitable environment for transmission, the occurrence of DENV transmission is dependent on the introduction
of the virus which is spread by travelers, and a sufficiently large and dense susceptible human population, a factor somewhat complicated by the mix of short-term heterotypic and long-term homotypic immunity to DENV.19–22

While these climatic and socioeconomic factors are known to be important in some locations, their importance on a global scale remains unclear. Understanding their global importance is key to estimating dengue risk in places where little data is available and in projecting how dengue risk may change in the future as climate and socio-demographic conditions change. To identify factors that mediate DENV transmission risk and endemicity, we focused on DENV transmission on islands. Island populations are relatively isolated, such that transmission dynamics are more likely to be determined by local DENV ecology than by regional dynamics. Using a stratified random sample of inhabited islands throughout the world, we developed a database of dengue indicators from published literature, ministry of health data and informal digital sources. We then collected information on potentially relevant demographic, climate, socioeconomic, and connectivity variables to assess their relative contribution to the presence and endemicity of DENV in those island populations. This information along with data from the dengue database was incorporated into a Bayesian model with three components: the probability of observing a DENV outbreak if one occurred; the probability of climatic suitability for DENV transmission; and the probability of a DENV outbreak actually occurring. This model was fitted and then validated on a separate set of islands.

Methods

Island selection and covariate data

As a basis for selecting islands, we used a list of 1991 islands from the United Nations Environment Programme (UNEP; http://islands.unep.ch/isldir.htm). We collected the area of each island (km²) from the same database and extracted the latitude and longitude for each island by geocoding (www.spatialepidemiology.net). We then excluded all islands with area greater than 100 000 km² and those located more than 45° from the equator, because DENV vector mosquitoes are unlikely to be found at more extreme latitudes. For the remaining 1319 islands, we obtained population size from the most recent national census databases. We then excluded all islands with an unknown population size or less than 100 inhabitants as islands with very sparse populations are numerically insignificant and unlikely to have reliable information on DENV transmission.

For the remaining 728 islands, we collected climate, economic, and connectivity data. We extracted monthly temperature, precipitation, and relative humidity estimates from the NOAA/NCEP Reanalysis (www.esrl.noaa.gov/psd/data/reanalysis). We calculated the average temperature, the average temperature for the coolest month of the year (minimum monthly temperature), average yearly precipitation, and average relative humidity over the past 20 years, 1993–2012. Country-level gross domestic product (GDP) per capita from 2011 (2012 US$) was collected from the United Nations Statistics Division (http://data.un.org/). As an indicator of connectivity, we obtained estimates of travel time from each island to cities ≥50 000 people (by plane or boat) from the European Commission Joint Research Centre Global Environment Monitoring Unit (http://bioval.jrc.ec.europa.eu/products/gam/download.htm).24

Sampling strategy

We stratified the 728 islands by ocean (Atlantic, Indian, and Pacific), median GDP per capita (US$318.13), and median population size (5354), creating a total of 12 categories. We randomly selected 12 islands from each of these 12 categories. We also selected another 12 islands more than 45° removed from the equator to capture characteristics associated with the limits of vector suitability. A total of 142 islands were included in the analysis (the low GDP and low population categories contained less than 12 islands). From each strata, we randomly selected two-thirds of islands for training and one-third of available islands for testing, resulting in 94 and 48 islands, respectively.

Dengue data

We identified relevant published research using PubMed (http://www.ncbi.nlm.nih.gov/pubmed), and Google Scholar (http://www.scholar.google.com) and the search terms ‘dengue’ + country name, ‘dengue’ + island name, ‘dengue’ + country name, and ‘dengue’ + major city on the island (if applicable). We searched Ministry of Health and WHO reports by country using GIDEON (http://www.gideononline.com/) and Ministry of Health websites. We also identified a mix of formal and informal information about dengue outbreaks from ProMED (http://www.promedmail.org/), HealthMap (http://www.healthmap.org/), and Google (http://www.google.com).

We reviewed all the sources for evidence of dengue outbreaks on the selected islands. For the purpose of this study, we defined a local outbreak as any instance when at least 10 confirmed autochthonous cases were reported. For each island, we identified three principal indicators of local dengue transmission: if any outbreak had been recorded, the maximum number of consecutive years with dengue outbreaks with a maximum of ten, and the maximum number of years with outbreaks within a decade. For islands with no history of outbreaks, we also sought evidence of surveillance systems for influenza and vector-borne diseases. The presence of these systems was used as an indicator of the likelihood of dengue being detected should it occur; in a location with no surveillance, detection is unlikely. The collected data and references are available in Supplementary Table 1.

Statistical models

We estimated three different probabilities associated with the occurrence of dengue outbreaks: \( p_{\text{Obs}} \), the probability of observing an outbreak; \( p_{\text{suit}} \), the probability of having a suitable climate for DENV transmission; and \( p_{\text{Out}} \), the yearly probability of having a dengue outbreak.

First, we modeled observation of at least one outbreak as a Bernoulli process dependent solely on \( p_{\text{Obs}} \):

\[
O \sim \text{Bernoulli}(p_{\text{Obs}}),
\]

where \( O \) indicates the presence or absence of at least one reported outbreak for each island \( i \), and \( p_{\text{Obs}} \) is a logistic-linked linear function of the covariate \( X \):

\[
\logit(p_{\text{Obs}}) = \alpha_0 + \alpha_X X_i,
\]

where \( \alpha_0 \) is the intercept and \( \alpha_X \) is the coefficient for covariate \( X \) (e.g., population size). We fitted models using Bayesian Markov

304

L. R. Feldstein et al.
chain Monte Carlo sampling (described in detail below) and compared candidate models using the deviance information criteria (DIC)\textsuperscript{25}.

Next, we fitted a model incorporating both $p_{\text{Obs}}$ and $p_{\text{Suit}}$ as determinants of the observation of at least one outbreak:

\[
O_i \sim \text{Bernoulli}(p_{\text{Obs}}, p_{\text{Suit},i}),
\]

\[
\logit(p_{\text{Obs},i}) = \alpha_0 + \alpha_X X_i,
\]

\[
\logit(p_{\text{Suit},i}) = \beta_0 + \beta_Y Y_i,
\]

where the $X$ covariates are those of the best $p_{\text{Obs}}$ model, the $\alpha$ coefficient priors are the posterior distributions from that model, $\beta_0$ is the intercept and $\beta_Y$ is the coefficient for covariate $Y$ (e.g. average temperature). Using posteriors from the previous step for the $p_{\text{Obs}}$ component in the combined model helps isolate the expected independent effect of $p_{\text{Obs}}$ but does not ensure that it remains significant. Using the $p_{\text{Obs}}$ components from the first stage we fitted alternative $p_{\text{Suit}}$ models, comparing the combined $p_{\text{Obs}}$ and $p_{\text{Suit}}$ models using DIC.

Finally, we defined the product, $p_{\text{Obs}}p_{\text{Suit}}p_{\text{Out}}$, as the yearly probability of an outbreak occurring and being observed. We fitted models using two different observed outcomes. First, the number of consecutive years with dengue outbreaks ($C$) was assumed to come from a negative binomial distribution, i.e., how many years are likely to pass before there is a year without an outbreak:

\[
C_i \sim \text{Negative Binomial}(1 - p_{\text{Obs}}, p_{\text{Suit}}, p_{\text{Out},i}, 1).
\]

Second, the maximum number of outbreak years within a decade ($D$) was assumed to come from a binomial distribution of the same yearly probability and the number of years, 10:

\[
D_i \sim \text{Binomial}(p_{\text{Obs}}, p_{\text{Suit}}, p_{\text{Out},i}, 10).
\]

Again, each probability was treated as a logistic-linked linear function of the covariate:

\[
\logit(p_{\text{Obs},i}) = \alpha_0 + \alpha_X X_i,
\]

\[
\logit(p_{\text{Suit},i}) = \beta_0 + \beta_Y Y_i, \text{ and}
\]

\[
\logit(p_{\text{Out},i}) = \gamma_0 + \gamma_Z Z_i,
\]

where the $X$ and $Y$ covariates are those of the best stage two model, the $\alpha$ and $\beta$ coefficient priors are the posterior distributions from that model, $\gamma_0$ is the intercept and $\gamma_Z$ is the coefficient for covariate $Z$ (e.g. population density). These two outcomes were fit simultaneously and the final model was selected by comparing the DIC of all candidate models.

We initially assigned weakly informative Gaussian priors to each $\alpha$, $\beta$, and $\gamma$ coefficient with zero means and variances based on the potential magnitude of their effects on the OR. For example, we expect a one-log change in population could change the OR by $0$–$50\%$. We thus set the prior SD to 0.5 such that 68\% of the prior distribution fell between a 40\% reduction and a 60\% increase in the OR. The prior SD for the effects of the logged population density was also set to 0.5. For GDP and travel time, we used an SD of 0.1, equivalent to a 10\% change per log change in GDP (in US$) or per log change in travel time (in minutes) to the nearest urban center, respectively. For temperature we used an SD of 0.2 for a 20\% change per 1°C. Relative humidity and precipitation had greater ranges so we used an SD of 0.05, for 5\% change per 1% change in relative humidity or per 1 cm of rain, respectively. Priors for the intercepts were set to zero with an SD of 2, such that there was strong prior coverage across the potential distributions of $p_{\text{Suit}}$, $p_{\text{Obs}}$, and $p_{\text{Out}}$. We fitted each model in sequence using Bayesian Markov chain Monte Carlo sampling in OpenBugs 3.2.2 (http://www.openbugs.net),\textsuperscript{26} R2OpenBUGS (http://cran.r-project.org/web/packages/R2OpenBUGS/index.html)\textsuperscript{27} and R 3.0 (www.r-project.org). We mean-centered all covariates and thinned posterior samples every 10 samples to limit autocorrelation.

Results

Island data

Using a stratified sampling procedure, we selected 142 islands for training (94) and testing (48) datasets. These islands represented a variety of climates, demographics, and socioeconomic conditions (Table 1). We identified at least one dengue outbreak in 57 and 26 of the training and testing islands, respectively. For the remaining islands, we found no reported outbreaks, indicating either the absence of outbreaks or the absence of data document outbreaks. Of the 37 and 22 training and testing islands with no reported outbreaks, only 16 and 12, respectively, had evidence of local surveillance for vector-borne diseases or influenza, indicating the potential capacity to capture cases should they occur.

Model development

First, we analyzed the probability of having observed at least one outbreak on each island in the training dataset. We broke this probability down conceptually into two independent components: the probability of observing an outbreak should one occur; and the climatic suitability of the island for dengue transmission. Should a dengue outbreak occur, the probability of observing it ($p_{\text{Obs}}$) may be reduced where there are very few people or few resources for surveillance. We therefore assessed the role of population size and socioeconomic status on observation. Assuming no difference in climatic suitability across the islands (which we address in the next step), we found that increased population size was associated with a greater probability of observing at least one outbreak, while GDP per capita had no association (Table 2).

We then added a climatic suitability component ($p_{\text{Suit}}$) and assessed how suitability influences DENV transmission. Relative humidity and average annual precipitation had no association with suitability, but both average temperature and the minimum monthly temperature were positively associated, with the minimum monthly temperature model having a lower DIC value (Table 2). Therefore, the optimal model for observing at least one outbreak on an island included an observation component ($p_{\text{Obs}}$) dependent on population size, and a suitability component ($p_{\text{Suit}}$) dependent on the minimum monthly temperature.

Finally, we assessed factors in addition to those driving $p_{\text{Obs}}$ and $p_{\text{Suit}}$ that may contribute to the frequency of outbreaks, specifically the yearly probability of an outbreak occurring ($p_{\text{Out}}$). As outcomes for this component, we used evidence of recurrent dengue outbreaks; the maximum number of years with outbreaks observed in a decade and the maximum number of consecutive years with outbreaks within a decade. The factors potentially

305
influencing the frequency of outbreaks independent of the probability of observation and environmental suitability include demographics, connectivity, socioeconomics, and climate. Population size, population density, and area were all correlated and significantly associated with an increased yearly probability of outbreaks (Table 2). Increased travel time to the nearest city, i.e., decreased connectivity, was associated with decreased risk of outbreaks and higher GDP per capita was associated with increased risk of outbreaks. Of the climate variables, average annual precipitation and relative humidity were both associated with increased yearly probability of outbreaks. The model including population size was the best fitting model with a single covariate, as measured by the DIC (Table 2). Adding precipitation and relative humidity to the population model further reduced the DIC, while other additional covariates did not. The model with the lowest DIC included population and precipitation as covariates for the yearly probability of outbreaks occurring (Table 3).

The final model included three components: the probability of observing an outbreak, the probability of climatic suitability, and the yearly probability of outbreaks. This model included an effect of population on the probability of observation, an effect of minimum monthly temperature on the probability of suitability, and effects of population and precipitation on the yearly probability of outbreaks occurring. For every log increase in population size, the odds of observing an outbreak increased by 92% (95% credible interval [CI] 63–125%), with the probability of observation being greater than 0.5 for populations larger than 3600 (95% CI 1900–5600) (Figure 1A). A 1°C increase in average temperature of

### Table 1. Summary of data for training and testing islands

<table>
<thead>
<tr>
<th>Variable</th>
<th>Training set (n=94)</th>
<th>Median</th>
<th>Testing set (n=48)</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population size</td>
<td>177–23299716</td>
<td>6784</td>
<td>215–18943000</td>
<td>10683</td>
</tr>
<tr>
<td>Area (km²)</td>
<td>0.1–97724.1</td>
<td>178</td>
<td>0.9–84421</td>
<td>252.5</td>
</tr>
<tr>
<td>Population density (per km²)</td>
<td>0.2–9782.4</td>
<td>126.3</td>
<td>0.7–5147.8</td>
<td>80.4</td>
</tr>
<tr>
<td>Relative humidity (%)</td>
<td>29.6–93.4</td>
<td>80.6</td>
<td>64.1–92.3</td>
<td>80.8</td>
</tr>
<tr>
<td>Annual precipitation (cm)</td>
<td>3.3–426.9</td>
<td>154.2</td>
<td>7.4–426.9</td>
<td>172.6</td>
</tr>
<tr>
<td>Yearly average temperature (°C)</td>
<td>−9.1–29.1</td>
<td>26.2</td>
<td>5.5–27.6</td>
<td>26.3</td>
</tr>
<tr>
<td>Minimum monthly average temperature (°C)</td>
<td>−25.2–27.8</td>
<td>24.6</td>
<td>−6.8–27.0</td>
<td>24.8</td>
</tr>
<tr>
<td>GDP per capita (US$)</td>
<td>598–67039</td>
<td>7136.5</td>
<td>665–67039</td>
<td>8191.0</td>
</tr>
<tr>
<td>Travel time to nearest city (minutes)</td>
<td>4.0–5760.0</td>
<td>443.1</td>
<td>6.1–5760</td>
<td>379.5</td>
</tr>
</tbody>
</table>

DIC: deviance information criteria.  
^a log-transformed.

### Table 2. Odds ratios for univariate models

<table>
<thead>
<tr>
<th>Component</th>
<th>Variable</th>
<th>OR</th>
<th>95% Credible Interval</th>
<th>DIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \rho_{\text{obs}} )</td>
<td>Population^a</td>
<td>1.88</td>
<td>1.49–2.49</td>
<td>93</td>
</tr>
<tr>
<td>( \rho_{\text{su}} )</td>
<td>GDP per capita (US$^a)</td>
<td>0.95</td>
<td>0.80–1.14</td>
<td>128</td>
</tr>
<tr>
<td>( \rho_{\text{su}} )</td>
<td>Precipitation (cm)</td>
<td>1.01</td>
<td>1.00–1.03</td>
<td>90</td>
</tr>
<tr>
<td>( \rho_{\text{su}} )</td>
<td>Relative humidity (%)</td>
<td>1.06</td>
<td>0.96–1.14</td>
<td>88</td>
</tr>
<tr>
<td>( \rho_{\text{su}} )</td>
<td>Average temperature (°C)</td>
<td>1.51</td>
<td>1.21–1.95</td>
<td>75</td>
</tr>
<tr>
<td>( \rho_{\text{su}} )</td>
<td>Minimum temperature (°C)</td>
<td>1.51</td>
<td>1.24–1.91</td>
<td>68</td>
</tr>
<tr>
<td>( \rho_{\text{su}} )</td>
<td>Population^a</td>
<td>1.38</td>
<td>1.27–1.50</td>
<td>677</td>
</tr>
<tr>
<td>( \rho_{\text{su}} )</td>
<td>Travel time to nearest city (minutes)^a</td>
<td>0.85</td>
<td>0.76–0.93</td>
<td>719</td>
</tr>
<tr>
<td>( \rho_{\text{su}} )</td>
<td>Area (km²)</td>
<td>1.26</td>
<td>1.17–1.38</td>
<td>705</td>
</tr>
<tr>
<td>( \rho_{\text{su}} )</td>
<td>Population density (per km²)</td>
<td>1.30</td>
<td>1.14–1.50</td>
<td>719</td>
</tr>
<tr>
<td>( \rho_{\text{su}} )</td>
<td>GDP per capita (US$^a)</td>
<td>1.16</td>
<td>1.01–1.33</td>
<td>720</td>
</tr>
<tr>
<td>( \rho_{\text{su}} )</td>
<td>Precipitation (cm)</td>
<td>1.007</td>
<td>1.004–1.010</td>
<td>694</td>
</tr>
<tr>
<td>( \rho_{\text{su}} )</td>
<td>Relative humidity (%)</td>
<td>1.08</td>
<td>1.04–1.13</td>
<td>709</td>
</tr>
<tr>
<td>( \rho_{\text{su}} )</td>
<td>Average temperature (°C)</td>
<td>1.02</td>
<td>0.85–1.16</td>
<td>727</td>
</tr>
<tr>
<td>( \rho_{\text{su}} )</td>
<td>Minimum temperature (°C)</td>
<td>1.04</td>
<td>0.90–1.15</td>
<td>726</td>
</tr>
</tbody>
</table>
the coldest month of the year was associated with a 67% (95% CI 43–95%) higher odds of an island being suitable for DENV transmission. Islands with a minimum monthly temperature greater than 14.8°C (95% CI 12.4–16.6°C) had a probability of suitability greater than 0.5 (Figure 1B). For every log increase in population size, the annual odds of an outbreak occurring increased by 36% (95% CI 27–48%). For every centimeter increase in precipitation, the annual odds of an outbreak occurring increased by 0.56% (95% CI 0.35–0.78%). Figure 1C shows the combined effect of these factors on the annual probability of outbreaks occurring.

Table 3. Coefficients for the best-fit model (DIC=650)

<table>
<thead>
<tr>
<th>Component</th>
<th>Variable</th>
<th>Coefficient</th>
<th>95% Credible Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>pObs</td>
<td>Intercept</td>
<td>−5.3</td>
<td>−6.8, −3.8</td>
</tr>
<tr>
<td></td>
<td>Population²</td>
<td>0.65</td>
<td>0.49, 0.81</td>
</tr>
<tr>
<td>pSuit</td>
<td>Intercept</td>
<td>−7.6</td>
<td>−10.8, −4.7</td>
</tr>
<tr>
<td></td>
<td>Minimum</td>
<td>0.51</td>
<td>0.36, 0.67</td>
</tr>
<tr>
<td></td>
<td>Temperature (°C)</td>
<td>0.51</td>
<td>0.36, 0.67</td>
</tr>
<tr>
<td>pOut</td>
<td>Intercept</td>
<td>−4.4</td>
<td>−5.5, −3.4</td>
</tr>
<tr>
<td></td>
<td>Population²</td>
<td>0.31</td>
<td>0.24, 0.39</td>
</tr>
<tr>
<td></td>
<td>Precipitation (cm)</td>
<td>0.0056</td>
<td>0.0035, 0.0078</td>
</tr>
</tbody>
</table>

DIC: deviance information criteria.
² log-transformed.

Model validation

To assess the model’s ability to identify suitable islands for dengue transmission we compared pSuit predictions to the data for both training and testing islands (Figures 2 and 3A). Of the 74 training islands with evidence of local surveillance or dengue outbreaks, the model predicted 60 to be suitable and 14 to be unsuitable. Three islands classified as suitable had no evidence of dengue transmission (Figure 3A). Two of these, Badu and Cocos, have small population sizes and thus are unlikely to have observed outbreaks (estimated pObs = 0.27 and 0.23, respectively). The third island, Bahrain, was unique in the dataset as it had a large population and high temperatures, but the lowest humidity and precipitation among any of the islands. None of the islands classified as unsuitable had evidence of dengue outbreaks. The overall accuracy of pSuit for the training islands was 96% (57 of 60 islands predicted to be suitable had observed outbreaks and 14 of 14 islands predicted to be unsuitable had no evidence of transmission).

For the 38 testing islands with evidence of outbreaks or local surveillance, 11 had pSuit < 0.5 and 27 had pSuit > 0.5 (Figure 3B). Of the islands predicted to be unsuitable, none of them had evidence of dengue outbreaks. Of the 27 islands predicted to be suitable, only one had no evidence of an outbreak, Groote Eylandt, which is very small and had pObs of 0.37. Overall the model classification for suitability was correct for 97% of the testing islands.

To assess the accuracy of the model to estimate endemicity, we multiplied pSuit and pOut, to consider locations that are suitable and likely to have frequent outbreaks. This index was strongly correlated with increased numbers of outbreaks in consecutive years and within a decade, accounting for 55 and 58% of the variability, respectively (Figure 4A–4B). While four training islands had long-term data indicative of endemicity (yearly outbreaks for a decade or more), 14 islands had less clear data, with 5–9 outbreak years in a decade. These islands and those with fewer outbreaks may be non-endemic areas with frequent
outbreaks or areas with under-recognized transmission, a difference that is critical but difficult to assess. We thus evaluated the model accuracy against a somewhat arbitrary threshold, assuming that islands with at least five consecutive years of transmission were endemic and should have $p_{suitpout} > 0.5$. There were 49 training islands with $p_{suitpout} > 0.5$. Of these, all 49 had less than five consecutive years with outbreaks and 47 had less than five outbreak years in a decade (Figure 3C). Of the 25 training islands with $p_{suitpout} < 0.5$, 12 had at least five consecutive years with outbreaks and 16 had five or more outbreaks observed within a decade. The overall accuracy on the training data was 82% for consecutive years with outbreaks and 85% for outbreak years in a decade.

For the testing islands, $p_{suitpout}$ was also strongly associated with increased numbers of outbreaks in consecutive years and within a decade (accounting for 49% and 47% of the variability, respectively) (Figure 4C–4D). All of the 26 testing islands with $p_{suitpout} < 0.5$ had less than five consecutive years with outbreaks and 16 had five or more outbreaks observed within a decade. Of the 12 testing islands with $p_{suitpout} > 0.5$, three had at least five consecutive years with outbreaks and six had at least five outbreaks in a decade. The overall accuracy for the testing dataset was 76% for consecutive outbreak years and 79% for outbreaks in a decade.

Discussion

Using demographic, socioeconomic, connectivity, climatic and dengue data from a stratified sample of 94 islands, we found that the occurrence of dengue outbreaks and their frequency is strongly associated with climate and population size. Minimum monthly temperature was the strongest determinant of climatic suitability, with locations predicted to be suitable having minimum monthly temperatures of approximately 14.8°C. Minimum temperature has often been found to have a strong association with both temporal and spatial variation in dengue incidence, highlighting the importance of low temperatures as a limiting factor of transmission. It has long been thought that the 10°C winter isotherm roughly defines the geographical limits of Ae. aegypti and DENV transmission. However, this threshold also seems appropriate given that Ae. aegypti infection with at least some viruses is possible at temperatures as low as 10°C. Nonetheless, Ae. aegypti activity is greatly reduced at these temperatures, so transmission, though possible, is generally unlikely. No islands in the training dataset with minimum monthly temperatures below 14.8°C had evidence of past dengue transmission, but other, colder, non-island areas clearly do (e.g., Philadelphia or Nice). These areas are, however, much warmer during some parts of the year, so while areas with minimum temperatures below 14.8°C are unlikely to be hospitable for dengue transmission, there may be times when it can occur.

Precipitation and relative humidity had no discernible effects on climatic suitability across the islands. This may reflect adaptation of both humans and mosquitoes in areas or times of low rainfall where collected water often serves as larval habitat for Ae. Aegypti. Regardless of local climate, wherever there are humans, there is generally some sort of aquatic environment suitable for Ae. aegypti eggs, larvae, and pupae. Increased precipitation, however, was associated with an increased probability of outbreaks. Thus, although rainfall may not be critical for the occurrence of dengue, it may have a strong effect on the frequency of outbreaks.

Despite their biological importance to DENV transmission and the importance of temperature to climatic suitability, relative humidity and average temperature were not significantly associated with the yearly probability of outbreaks. It is still possible, however, that they have local effects in some locations. For example, low humidity may contribute to the absence of dengue outbreaks in Bahrain, the driest island in the dataset, suggesting...
Figure 3. $\hat{p}_{\text{suit}}$ and $\hat{p}_{\text{suit} \cdot \text{out}}$ estimates for training and testing islands. (A) and (B) show the estimated $\hat{p}_{\text{suit}}$ (point) and 95% credible interval (line) for each island, in the training and testing datasets, respectively. White points are islands with no known history of dengue outbreaks, blue points are islands with evidence of at least one outbreak. (C) and (D) show the estimated $\hat{p}_{\text{suit} \cdot \text{out}}$ for each training and testing island, respectively. The colors indicate no outbreaks (white), and an increasing number of consecutive years with outbreaks from 1 (yellow) to 10 (dark green).
that there may be a lower limit on humidity tolerance for DENV transmission that is not characterized here.

The most important determinant of the yearly outbreak probability in suitable environments was population size. Interestingly, population size had a stronger association than density. While high or moderate population density is a well-established risk factor for dengue, our finding suggests that on a population level, it may matter more how many people are functionally living in the same area rather than how closely they reside together. Indeed, increased spatial heterogeneity, which may be associated with slightly less dense populations, is a recognized contributor to pathogen persistence.

After accounting for the effects of population and precipitation, country-level GDP per capita was not associated with differences in the probability of recurring outbreaks, with relatively wealthy islands like Singapore and relatively low-wealth islands, such as many of the Philippine islands, having evidence of frequent outbreaks. On the local level, it is likely that socioeconomics do influence dengue transmission, particularly related to water access, trash removal systems, the usage of screens or air conditioning, and housing density.

Connectivity of islands to cities was associated with increased yearly risk of outbreaks in the univariate analysis. This implies an increased chance of frequent reintroduction or persistence of DENV transmission across regional meta-populations. However, proximity to cities is also related to population size, and connectivity had no significant association when population size was included in the model. Connectivity may certainly have a role in dengue dynamics for the islands considered here, but the related role of population size may have obscured it.

The final risk model included an effect of population size on observability, an effect of minimum temperature on climate suitability, and effects of population size and precipitation on outbreak frequency. Minimum temperature was an excellent predictor of whether outbreaks had ever occurred. Only four islands were predicted to be suitable despite a lack of evidence of outbreaks: three that were predicted to be too small to necessarily have evidence of outbreaks and a fourth, Bahrain, which represents a uniquely dry environment within this dataset where precipitation or humidity may play roles that are not evident elsewhere.

Figure 4. Model fit for $p_{SuitPOut}$ on training and testing data. (A) and (B) show the relationship between $p_{SuitPOut}$ estimates for the training dataset and the number of consecutive years with outbreaks and the maximum number of years with outbreaks within a decade, respectively. ($R^2=0.55$ and $0.58$, respectively). (C) and (D) show the relationship between $p_{SuitPOut}$ for the testing dataset and the number of consecutive years with outbreaks and the maximum number of years with outbreaks within a decade, respectively. ($R^2=0.49$ and $0.47$, respectively).
While islands are not necessarily representative of mainland areas, they represent environments with unique characteristics that may be informative about the large-scale ecology of dengue for islands as well as non-island areas. Minimum temperature alone clearly differentiated suitable and unsuitable environments. Importantly, the data included few islands around the 14.8°C threshold, so the model may underestimate potential risk, especially in areas with strong seasonal variation that is generally reduced on islands. Future work can build on these results, investigating how dengue dynamics in mainland areas are related to these factors and others, such as seasonal weather variation and meta-population dynamics among closely connected locations.

Despite high accuracy in the model predictions, there are clearly other factors that influence DENV transmission, and it is not surprising that a simple global model did not capture all of the variation in the frequency of outbreaks. The areas where there is discordance between the model predictions and dengue data also provide direction for future research. In some areas with characteristics suggesting endemicity there may be unrecognized transmission, and in others, such as Guam, which has no evidence of transmission since 1944, further research may provide insight on how to control dengue.

Conclusions

The key determinants identified here, minimum temperature, precipitation, and population size, have biological and empirical relationships with DENV transmission on a local scale. Although island ecology is not necessarily representative of the ecology of other, non-island locations, the relationships described here provide insight into the current global landscape of dengue and how it might change as climate and demographics change. Effective dengue prevention and control depends on understanding these factors and how they influence the spatiotemporal dynamics of dengue both now and in the future.

Supplementary data

Supplementary data are available at Transactions Online (http://trstmh.oxfordjournals.org/).

Authors' contributions: JSB and MAJ conceived the study; LRF, OJB and MAJ collected the data; LRF and MAJ performed the analysis and drafted the manuscript; all authors critically revised the manuscript for intellectual content. All authors read and approved the final manuscript. MAJ is the guarantor of the manuscript.

Funding: This work is supported by the National Institutes of Health, National Library of Medicine [R01LM010812 to JSB], the Wellcome Trust [#095066 to SIH], the Bill & Melinda Gates Foundation [#OPP1093011 to SIH and JSB], RAPIDD program of the Science & Technology Directorate (to SIH), the Department of Homeland Security, National Institutes of Health (to SIH), the Fogarty International Center (to SIH), and the Biotechnology and Biological Sciences Research Council (to OJB).

Competing interests: None declared.

Ethical approval: Not required.

References


