Real-time Zika risk assessment in the United States

Lauren A Castro*[,1], Spencer J Fox*[,1@], Xi Chen[2], Kai Liu[3], Steve Bellan[4], Nedialko B Dimitrov[2], Alison P Galvani[5,6], Lauren Ancel Meyers[1,7]

Affiliations:

1 - Department of Integrative Biology, The University of Texas at Austin, Austin, TX, USA

2 - Graduate Program in Operations Research Industrial Engineering, The University of Texas at Austin, Austin, TX, USA

3 - Institute for Cellular and Molecular Biology, The University of Texas at Austin, Austin, TX, USA

4 - Center for Computational Biology and Bioinformatics, The University of Texas at Austin, Austin, TX, USA

5 - Center for Infectious Disease Modeling and Analysis, Yale School of Public Health, New Haven, CT, USA

6 - Department of Ecology and Evolution, Yale University, New Haven, CT, USA

7 - The Santa Fe Institute, Santa Fe, NM, USA

* contributed equally to this manuscript

[@] corresponding author: spncrfx@gmail.com

Abstract

Recent Zika Virus (ZIKV) outbreaks in southern Florida have heightened public health concern across the southern United States. As autochthonous (locally-acquired) cases accumulate within the US, policymakers seek early and accurate indicators of self-sustaining transmission to inform intervention efforts in high risk areas. However, given ZIKV's low reporting rates and the geographic variability in both importations and transmission potential, a small cluster of reported cases may reflect diverse scenarios, ranging from multiple self-limiting but independent introductions to a self-sustaining local epidemic. We developed a stochastic model that captures variation and uncertainty in ZIKV case reporting, importations, and transmission, and applied it to assess county-level epidemic risk throughout the state of Texas. For each of the 254 counties, we estimated the future epidemic risk as a function of reported autochthonous cases and evaluated a national recommendation to trigger interventions immediately following the first two reported cases of locally-transmitted ZIKV. Our analysis suggests that the regions of greatest risk for sustained ZIKV transmission include 21 Texas counties along the Texas-Mexico border, in the Houston Metro Area, and throughout the I-35 Corridor from San Antonio to Waco. Variation in vector habitat suitability and importation risk drives epidemic risk variation. Upon detection of two locally transmitted cases in a spatiotemporal cluster, the threat of epidemic expansion depends critically on local vulnerability. For high risk Texas counties, we estimate this likelihood to be 64%, assuming an August 2016 risk projection and a 20% reporting rate. With reliable estimates of key epidemiological parameters, including reporting rates and vector abundance, this framework can help optimize the timing and spatial allocation of public health resources to fight ZIKV in the US.

Author Summary

Given the growing threat of Zika Virus (ZIKV) in the southern US and the critical importance of early intervention, public health decision makers seek early and accurate indicators of local transmission. However, given ZIKV's low reporting rates and the geographic variability in both importations and transmission potential, a small cluster of reported cases may reflect diverse scenarios, ranging from multiple self-limiting but independent introductions to a self-sustaining local epidemic. To support real-time risk assessment of emerging ZIKV outbreaks at a county level, we developed a quantitative framework that estimates current ZIKV burden and future epidemic threat based on recent reported cases and the underlying risks of ZIKV importation and transmission. We assessed ZIKV risk in Texas and found that 21 of the 254 counties are potentially vulnerable to locally-transmitted ZIKV outbreaks. The high risk region includes much of the Texas-Mexico border, the Houston Metro Area, and patches along the I-35 Corridor from San Antonio to Waco. If two locally transmitted and epidemiologically linked cases are detected in one of these counties, we estimate a 64% risk of imminent epidemic expansion.

Introduction

In February 2016, the World Health Organization (WHO) declared Zika virus (ZIKV) a Public Health Emergency of International Concern [1]. As of 18 August of 2016, the WHO confirmed mosquito-transmitted cases in 70 countries and territories, with over 500,000 suspected and confirmed cases in the Americas alone [2,3]. In the US, the 46 reported autochthonous (local) ZIKV cases all occurred in Southern Florida, but the potential range of a primary ZIKV vectors, *Aedes aegypti*, may include over 30 states [4,5]. Texas is a particularly vulnerable state, given its suitable climate, international airports, and geographical proximity to affected countries [4,6–10]. Of the 2,487 imported ZIKV cases in the US that have been identified, 137 have occurred in Texas through the end of August. While these importations have yet to spark autochthonous transmission within the state, Texas has historically sustained several small, autochthonous outbreaks (ranging from 4 - 25 confirmed cases) of another arbovirus vectored by *Ae. Aegypti*—dengue (DENV) [11–13].

As peak mosquito season continues in the US and more cases are introduced via international travelers from the Americas, public health decision makers will face considerable uncertainty in gauging the severity of the threat and in effectively initiating interventions, given the large fraction of undetected ZIKV cases (asymptomatic and symptomatic) as well as the challenge in estimating the economic tradeoff of disease prevention versus disease response [14–17]. Depending on the ZIKV symptomatic fraction, reliability and rapidity of diagnostics, importation rate, and transmission rate, the detection of five autochthonous cases in a single Texas county, for example, may indicate a small chain of cases from a single importation, a self-limiting outbreak, or even a large, hidden epidemic underway (Fig 1). These diverging possibilities have historical precedents. In French Polynesia, a handful of suspected ZIKV cases were reported by October 2013; two months later an estimated 14,000-29,000 individuals had been infected [14,15]. By contrast, Anguilla had 17 confirmed cases from late 2015 into 2016 without a subsequent epidemic, while many surrounding countries experienced large ZIKV epidemics [18]. To address the uncertainty, the CDC recently issued conservative guidelines for state and local agencies; they recommend initiation of public health responses following local reporting of *two* non-familial autochthonous ZIKV cases [19].

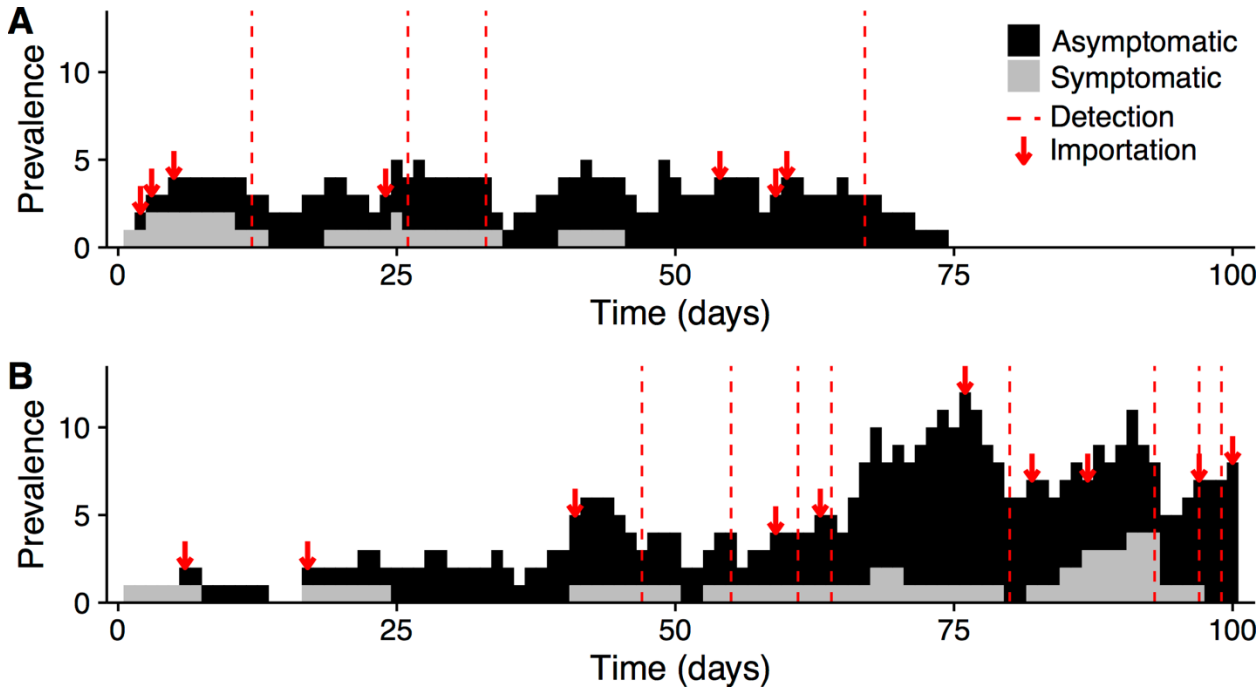


Fig 1. ZIKV emergence scenarios. A ZIKV infection could spark **(A)** a self-limiting outbreak or **(B)** a growing epidemic**.** Cases are partitioned into symptomatic (grey) and asymptomatic (black). Arrows indicate new ZIKV importations by infected travelers and vertical dashed lines indicate case reporting events. On the 75th day, these divergent scenarios are almost indistinguishable to public health surveillance, as exactly three cases have been detected in both. By the 100th day, the outbreak (A) has died out with 21 total infections while the epidemic (B) continues to grow with already 67 total infections. Each scenario is a single stochastic realization of the model with R_0=1.1, reporting rate of 10%, and introduction rate of 0.1 case/day.

Here, we develop a model to support real-time ZIKV risk assessment that accounts for uncertainty regarding ZIKV epidemiology, including importation rates, reporting rates, and local vector population density. This framework can be readily updated as our understanding of ZIKV evolves to provide situational awareness and actionable guidance for public health officials. By simulating ZIKV transmission using a stochastic branching process model [20] based on recent ZIKV data, we estimated ZIKV epidemic risk for each of the 254 counties in Texas, both initially and as a function of the cumulative number of reported autochthonous ZIKV cases. Further, we estimate county-level epidemic risks corresponding to the recommended two-case trigger and demonstrate the design of county-specific

111 triggers indicative of imminent epidemic expansion. Our results suggest that counties along the Texas-
112 Mexico border, in the Houston Metro Area, and throughout the I-35 Corridor from San Antonio to Waco
113 are at highest risk for sustained epidemics.
114
115 **Methods**
116 Risk-assessment methodological overview
117 In our study, we first estimated county-level ZIKV importation rates and sustained transmission
118 risk (R_0) for Texas in August of 2016. For each county, we then used these estimates to simulate 10,000
119 stochastic ZIKV outbreaks. Finally, we analyzed the simulated outbreaks to determine spatiotemporal
120 variation in ZIKV risk and estimate current and future epidemic risk from the number of reported
121 autochthonous cases (Fig S1).
122
123 Estimating County Importation Rates
124 Our analysis assumes that any ZIKV outbreaks in Texas originate with infected travelers
125 returning from regions with ZIKV activity. To estimate the ZIKV importation rate for specific counties,
126 we (1) estimated the Texas statewide importation rate (expected number of imported cases per day), (2)
127 estimated the probability that the next Texas import will arrive in each county, and (3) took the product of
128 the state importation rate and each county importation probability.
129 1. During the first quarter of 2016, 27 travel-associated cases of ZIKV were reported in Texas, with
130 11 occurring in Houston's Harris County [13], yielding a baseline first quarter estimate of 0.3
131 imported cases per day throughout Texas. In 2014 and 2015, arbovirus introductions into Texas
132 were threefold higher during the third quarter than the first quarter of each year, perhaps driven
133 by seasonal increases in arbovirus activity in endemic regions and the approximately 40%
134 increase from quarter 1 to quarter 3 in total number of international travelers to the US [21].
135 Taking this as a *baseline* (lower bound) scenario, we projected a corresponding increase in ZIKV
136 importations to 0.9 cases per day (statewide) for the third quarter. We also considered an *elevated*

importation scenario, in which the first quarter cases (27) in Texas represent only the symptomatic (20%) imported cases, corresponding to a projected third quarter statewide importation rate of 4.5 cases per day.

2. We defined county-specific import risk as the probability that the next imported case in Texas will occur in that county. To build a predictive model for import risk, we fit a probabilistic model (maximum entropy) [22] of importation risk to 183 DENV, 38 CHIKV, and 31 ZIKV Texas importations reported at the county level from 2002 to 2016 and informative socioeconomic, environmental, and travel variables (Supplement §1.1). Given the geographic and biological overlap between ZIKV, DENV and Chikungunya (CHIKV), we used historical DENV and CHIKV importation data to supplement ZIKV importations in the importation risk model, while recognizing that future ZIKV importations may be fueled by large epidemic waves in neighboring regions and summer travel, and thus far exceed recent DENV and CHIKV importations [23]. Currently, DENV, CHIKV, and ZIKV importation patterns differ most noticeably along the Texas-Mexico border. Endemic DENV transmission and sporadic CHIKV outbreaks in Mexico historically have spilled over into neighboring Texas counties. In contrast, ZIKV is not yet as widespread in Mexico as it is in Central and South America, with only one reported ZIKV importation along the border to date (Val Verde County). We included DENV and CHIKV importation data in the model fitting so as to consider potential future importation pressure from Mexico, as ZIKV continues its increasing trend since March 2016 [24]. To find informative predictors for ZIKV importation risk, we analyzed 72 socio-economic, environmental, and travel variables, and removed near duplicate variables and those that contributed least to model performance, based on out-of-sample cross validation [25,26], reducing the original set of 72 variables to 10 (Tables S3-S4).

Estimating County Transmission Risk

The risk of ZIKV emergence following an imported case will depend on the likelihood of mosquito-borne local transmission. For each Texas county, we used the Ross-Macdonald formulation to estimate the ZIKV reproduction number (R_0), which is the average number of secondary infections caused by the introduction of a single infectious individual into a fully susceptible population (Supplement §1.2) [27]. To parameterize the model (Table S5), we used mosquito life history estimates from a combination of DENV and ZIKV studies and estimates of *Ae. aegypti* abundance for each county [6]. For parameters that are sensitive to temperature (i.e., mosquito mortality and the extrinsic incubation period), we adjusted the mosquito parameters using average reported Texas county temperatures for the month of August [25]. Given the uncertainty in using DENV and ZIKV parameters from experimental and field studies from other localities, we have included a sensitivity analysis in Supplement §4 to explore upper and lower bounds of absolute values with changes in the average parameter values.

ZIKV Outbreak Simulation Model

To transmit ZIKV, a mosquito must bite an infected human, the mosquito must get infected with the virus, and then the infected mosquito must bite a susceptible human. We assumed that mosquito-borne transmission would be the main driver of epidemic dynamics, so we did not include sexual transmission in our model. Rather than explicitly model the full transmission cycle, we aggregated the two-part cycle of ZIKV transmission (mosquito-to-human and human-to-mosquito) into a single exposure period where the individual has been infected by ZIKV, but not yet infectious, and do not explicitly model mosquitos. For the purposes of this study, we need only ensure that the model produces a realistic human-to-human generation time of ZIKV transmission.

The resultant model thus follows a Susceptible-Exposed-Infectious-Recovered (SEIR) transmission process stemming from a single ZIKV infection using a Markov branching process model (Fig S2). The temporal evolution of the compartments is governed by daily probabilities of infected individuals transitioning between S, E, I, and R states, and new ZIKV cases arising from importations or autochthonous transmission (Table S6). We treat days as discrete time steps, and the next disease state

188 progression depends solely on the current state and the transition probabilities. We assume that infectious
189 cases cause a Poisson distributed number of secondary cases per day (via human to mosquito to human
190 transmission), but this assumption can be relaxed as more information regarding the distribution of
191 secondary cases becomes available. We also assume infectious individuals are introduced daily according
192 to a Poisson distributed number of cases around the importation rate. Furthermore, *Infectious* cases are
193 categorized into reported and unreported cases according to a reporting rate. We assume that reporting
194 rates approximately correspond to the percentage (~20%) of symptomatic ZIKV infections [16] and occur
195 at the same rate for imported and locally acquired cases. Additionally, we make the simplifying
196 assumption that reported cases transmit ZIKV at the same rate as unreported cases. We track imported
197 and autochthonous cases separately, and conduct risk analyses based on reported autochthonous cases
198 only, under the assumption that public health officials will have immediate and reliable travel histories for
199 all reported cases [28].

200 To accurately model the timing of ZIKV outbreaks, we fit the ZIKV generation time to recent
201 estimates (Supplement §2.4) [29]. The generation time measures the average duration from initial
202 symptom onset to the subsequent exposure of a secondary case, and is estimated to range from 10 to 23
203 days for ZIKV [29]. In our model, the generation time corresponds to the exposure period followed by
204 half of the infectious period. First, we fit the infectious period in our model to human ZIKV estimates for
205 duration of viral shedding, which we assumed to be the length of the infectious period, but may be an
206 underestimate of the total length. Specifically, we solved for transition rates of a Boxcar Model [30] that
207 produce an infectious period with mean duration of 9.88 days (Table S6) [14]. Then, we fit the exposure
208 period to achieve the empirical ZIKV generation time distribution [29], yielding a mean exposure period
209 of 10.4 days (95% CI 6-17) and a mean generation time of 15.3 days (95% CI 9.5-23.5). Given that the
210 exposure period includes human and mosquito incubation periods and mosquito biting rates, this range is
211 consistent with the estimated 5.9 day human ZIKV incubation period [14]. While there is much
212 uncertainty regarding ZIKV transmission, our model framework is flexible to extensions and updates as

213 we learn more about ZIKV epidemiology including reporting rates, transmission probabilities, clinical
214 features, and outside importation pressures.
215
216 Simulations
217 For each county risk scenario, defined by a specified importation rate, transmission rate, and
218 reporting rate, we ran 10,000 stochastic simulations. Each simulation began with a single imported
219 infectious case and terminated either when there were no individuals left in either the *Exposed* or
220 *Infectious* classes or the cumulative number of autochthonous infections reached 2,000. We classified
221 simulations as either epidemics or self-limiting outbreaks; epidemics were all simulations that fulfilled
222 two criteria: reached 2,000 cumulative autochthonous infections and had a maximum daily prevalence
223 (which we defined as the number of current infectious cases) exceeding 50 autochthonous cases (Fig S3).
224 The second criterion distinguishes simulations resulting in large self-sustaining outbreaks (that achieve
225 substantial peaks) from those that accumulate infections through a series of small, independently sparked
226 clusters (that fail to reach the daily prevalence threshold). The latter occurs occasionally under low R_0s
227 and high importation rates scenarios.
228 For simulations resulting in epidemics, the cumulative reported autochthonous infections may
229 include cases from several co-occurring or proximate transmission clusters, as might occur in actual
230 outbreaks. To verify that our simulations do not aggregate cases from clusters with clear temporal
231 separation, we calculated the distribution of times between sequential cases (Fig S4). In our simulated
232 epidemics, almost all sequentially occurring cases occur within 14 days of each other, consistent with the
233 CDC's threshold for identifying local transmission events (based on the estimated maximum duration of
234 the ZIKV incubation period) [19].
235
236 Outbreak Analysis
237 Policymakers must often make decisions in the face of uncertainty, such as when and where to
238 initiate ZIKV interventions. Our stochastic framework allows us to provide real-time county-level risk

assessments as reported cases accumulate. For each county, we found the probability that an outbreak will progress into an epidemic (reach 2,000 cases with a maximum daily prevalence over 50), as a function of the number of reported cases. We call this *epidemic risk*. To solve for epidemic risk in a county following the *x*th reported autochthonous case, we first find all simulations (of 10,000 total) that experience at least *x* reported autochthonous cases, and then calculate the proportion of those that are ultimately classified as epidemics. For example, consider a county in which 1,000 of 10,000 simulated outbreaks reach at least two reported autochthonous cases; the remaining 9,000 simulations dissipate with only one or zero case reports. If only 50 of the 1,000 simulations ultimately fulfill the two epidemic criteria, then the estimated epidemic risk following two reported cases in that county would be 5%. This simple classification scheme performs quite well, only rarely misclassifying a string of small outbreaks as an epidemic, with the probability of such an error increasing with the importation rate. For example, epidemics should not occur when R_0=0.9. If the importation rate is high, however, overlapping series of moderate outbreaks may occasionally meet the two epidemic criteria. Even under the highest importation rate we considered (0.3 cases/day), only 1% of outbreaks were misclassified.

This method can be applied to evaluate universal triggers (like the recently recommended two-case trigger) or derive robust triggers based on local importation and transmission risks as well as the risk tolerance of public health agencies. For example, if a policymaker would like to initiate interventions as soon as the risk of an epidemic reaches 30%, we would simulate local ZIKV transmission and solve for the number of reported cases at which the probability of an epidemic first exceeds 30%. Generally, the recommended triggers decrease as the policymaker threshold for action decreases (for example, policymakers would act sooner (fewer reported cases) for a 10% versus 30% threshold) and as the local transmission potential increases (*e.g.* R_0 = 1.5 versus R_0 = 1.2). A policymaker wishing to trigger interventions early, upon even a low probability of epidemic spread, has a low tolerance for failing to intervene but may waste resources on unnecessary interventions; a policymaker willing to wait longer, has a higher risk tolerance, but may implement interventions too late in the course of the outbreak.

County Uncertainty Analysis

We took two approaches to addressing uncertainty in the model parameters. First, we conducted a sensitivity analysis to address the considerable uncertainty regarding several inputs into our estimation of R_0, including mosquito biology, ZIKV epidemiology, and human-mosquito interactions (Supplement §4). For most factors, the county estimates of R_0 simply scale linearly with changes in the factor. However, county-specific human-mosquito contact rates can change relative county risks based on assumptions regarding the socioeconomic effect on human-mosquito interactions (Fig S6-7), and county risk moves southward as the summer heat subsides (Fig S8). Second, given the considerable uncertainty regarding ZIKV epidemiology, we examined a scenario where the absolute values of both R_0 and importation rate are unknown, but lie within plausible ranges for Texas. To do so, we randomly sampled 10,000 simulations from the high risk Texas county outbreaks (counties with estimates of R_0>1), creating an amalgamous high risk county, and completed the outbreak analysis as we do with individual counties.

Results

To develop a ZIKV risk assessment framework for Texas counties, we first estimate county-level ZIKV importation and transmission rates for August 2016. ZIKV importation risk within Texas is predicted by variables reflecting urbanization, mobility patterns, and socioeconomic status (Table S3), and is concentrated in metropolitan counties of Texas (Fig 2A). The two highest risk counties--Harris, which includes Houston and has an estimated 27% chance of receiving the next imported Texas case, and Travis, which includes Austin and has a 10% chance of the next importation--contain international airports. Other high risk regions include Brazos County, the Dallas and San Antonio metropolitan areas, and several counties along the Texas-Mexico border.

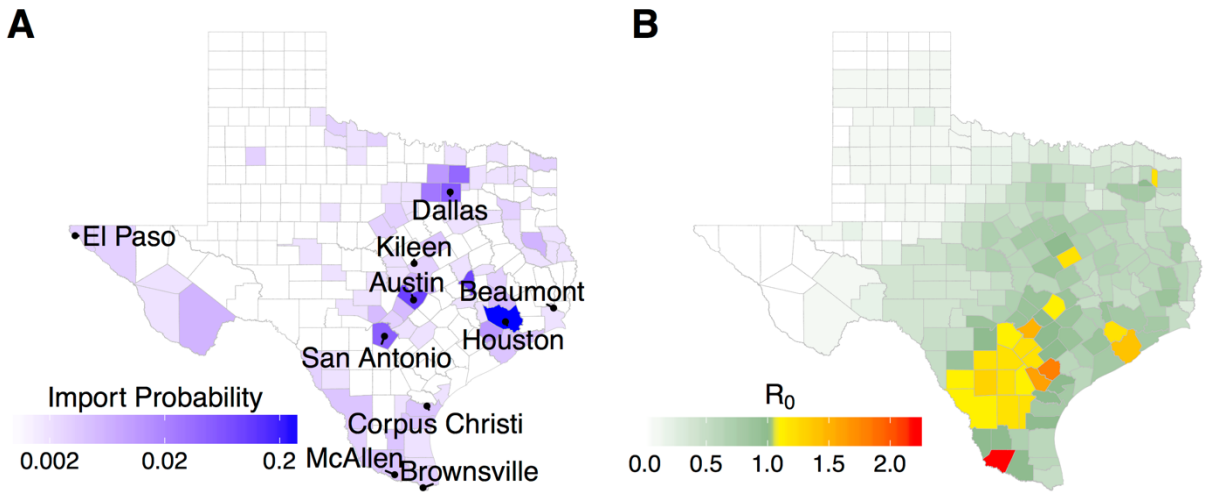


Fig 2. ZIKV importation and transmission risk estimates across Texas. (A) Color indicates the probability that the next ZIKV import will occur in a given county for each of the 254 Texas counties. Probability is colored on a log scale. The 10 most populous cities in Texas are labeled. Houston's Harris County has 2.7 times greater chance than Austin's Travis County of receiving the next imported case. **(B)** Estimated county-level R_0 for ZIKV. Counties with estimated R_0s ≤ 1 are shades of green, and yellow increasing to red counties indicate R_0s > 1 (See Fig S7-11 for sensitivity analysis).

Our county-level estimates of autochthonous ZIKV transmission risk (Fig 2B) suggest that the majority of Texas counties (87%) have an estimated R_0 below one, and thus are unlikely to sustain epidemics. The Southeast region of Texas has the highest estimated transmission risk, driven primarily by high mosquito habitat suitability. These estimates are sensitive to uncertainty in several parameters (Fig S7-11), and can be updated as we learn more about ZIKV. While the average transmission risk may be higher or lower than our baseline assumption, and certainly varies seasonally, the relative high and low risks areas of the state are robust (Fig S7-11), and allow us to conduct plausible case studies and identify at risk areas for enhanced surveillance and preparedness efforts. However given the uncertainty underlying specific county R_0 estimates, we also aggregate the 21 highest absolute estimates into a plausible distribution for any high risk Texas county, ranging from $R_0 = 1.0$ to $R_0 = 2.2$ with a median of 1.1.

Under a single set of epidemiological conditions, wide ranges of outbreaks are possible (Fig 3A). The relationship between what policymakers can observe (cumulative reported cases) and what they wish to know (current underlying disease prevalence) can be obscured by such uncertainty, and will depend critically on both the transmission and reporting rates (Fig 3B). If key drivers, such as R_0, can be estimated with confidence, then the breadth of possibilities narrows, enabling more precise surveillance. For example, under a known high risk R_0 scenario and with a 20% reporting rate, ten linked cumulative autochthonous reported cases corresponds to 6 currently circulating cases with a 95% CI of 1-16; under an unknown but high risk R_0 scenario, the same number of cases corresponds to an expected daily prevalence of 10 cases with a much wider 95% CI of 2-32 (Fig 3B).

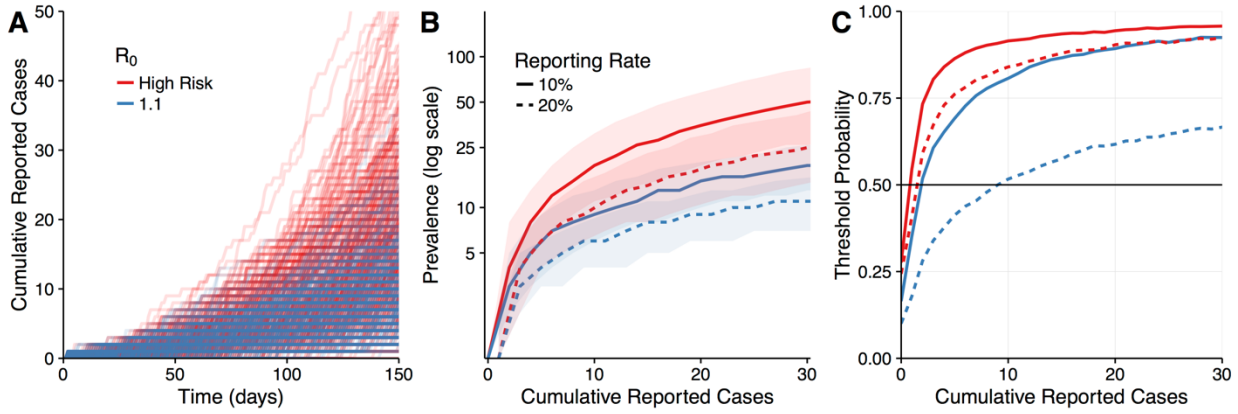


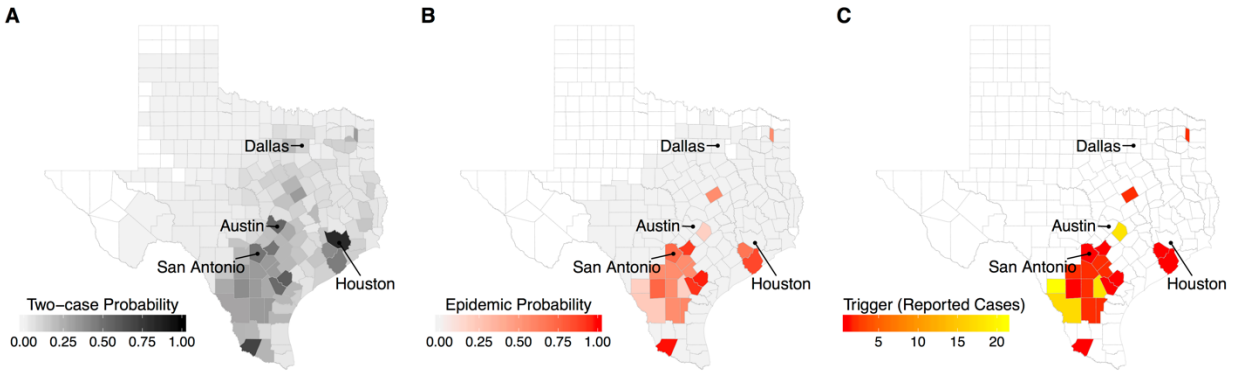
Fig 3. Real-time risk-assessment for ZIKV situational awareness and forecasting. (A) Simulated outbreaks, assuming an importation rate of 0.01 case per day, for a known high risk R_0 (blue) or an unknown high risk R_0 (red). Two thousand randomly sampled simulations are shown for each scenario. (B) Total number of (current) autochthonous cases as a function of the cumulative reported autochthonous cases, assuming an importation rate of 0.01 case per day, for a known R_0 (blue) or an unknown high-risk R_0 (red), and a relatively high (dashed) or low (solid) reporting rate. Ribbons indicate 50% quantiles. (C) The increasing probability of imminent epidemic expansion as reported autochthonous cases accumulate, compared across the high risk (red) and known moderate risk (blue) for a low (solid) and high (dashed) reporting rate. Suppose cases arise in a high risk county and a policymaker wishes to trigger a response as soon as two cases are reported (vertical line). With a 20% reporting rate there is a 64% probability of an ensuing epidemic.

We apply our model to characterize time-varying epidemic risk as cases accumulate in a given county. Under both a known and unknown high risk scenario, we track the probability of epidemic

327 expansion following each additional reported case (Fig 3C). Across the full range of reported cases, the
328 probability of epidemic spread is always higher in the unknown, compared with the known high risk
329 scenario showing more sensitivity to the reporting rate. These curves can support both real-time risk
330 assessment as cases accumulate and the identification of surveillance triggers indicating when risk
331 exceeds a specified threshold. For example, suppose a policymaker wanted to initiate an intervention
332 upon two reported cases and thought the reporting rate was 20%, this would correspond with a 64%
333 probability of an epidemic in the unknown high risk county but only 35% in the known high risk county.
334 Alternatively suppose a policy maker wishes to initiate an intervention when the chance of an epidemic
335 exceeds 50%. In the unknown high risk scenario, they should act immediately following the 1st
336 autochthonous reported case; in the known high risk scenario, the corresponding trigger ranges from two
337 to seven autochthonous reported cases, depending on the reporting rate. As the policymaker's threshold
338 (risk tolerance) increases, the recommended surveillance triggers can be adjusted accordingly.

339 To evaluate a universal intervention trigger of two reported autochthonous cases, we estimate
340 both the probability of a trigger event (two such cases) in each county and the level of epidemic risk at the
341 moment a trigger event occurs (second case reported) in each county. Assuming a baseline importation
342 rate extrapolated from recent importations to August 2016 and a 20% reporting rate, only a minority of
343 counties are likely to experience a trigger event (Fig 4A). While 231 of the 254 counties (91%) have non-
344 zero probabilities of experiencing two reported autochthonous cases, only 63 counties have at least a 10%
345 chance of such an event, with the remaining 168 counties having a median probability of 0.017 (range
346 0.0004 to 0.089). Next, assuming that a second autochthonous case has indeed been reported, we find that
347 the underlying epidemic risk varies widely among the 231 counties, with most counties having near zero
348 epidemic probabilities and a few counties far exceeding a 50% chance of epidemic expansion. For
349 example, two reported autochthonous cases in Starr County, along the Texas-Mexico border, correspond
350 to a 99% chance of ongoing transmission that would proceed to epidemic proportions without
351 intervention. The greater San Antonio metropolitan region appears to be the highest risk metropolitan
352 region with four of its eight counties having a higher than 25% probability of experiencing two reported

353 autochthonous cases; in those four counties, the epidemic risk upon detection of a second case ranges
354 from 19-90%. Houston metropolitan region is also a high risk region with its second (Fort Bend) and
355 fourth (Brazoria) largest counties having a 39% and 45% chance of sustaining two reported
356 autochthonous cases, respectively, with corresponding epidemic risks of 67% and 86% thereafter.
357



358

359 **Fig 4. Texas county ZIKV risk assessment. (A)** Probability of an outbreak with at least two reported autochthonous ZIKV
360 cases. **(B)** The probability of epidemic expansion at the moment the second autochthonous ZIKV case is reported in a county.
361 White counties never reach two reported cases across all 10,000 simulated outbreaks; light gray counties reach two cases, but
362 never experience epidemics. **(C)** Recommended county-level surveillance triggers (number of reported autochthonous cases)
363 indicating that the probability of epidemic expansion has exceeded 50%. White counties indicate that fewer than 1% of the
364 10,000 simulated outbreaks reached two reported cases. All three maps assume a 20% reporting rate and a baseline importation
365 scenario for August 2016 (81 cases statewide per 90 days) projected from historical arbovirus data. (Figure S5 provides
366 corresponding estimates under a worse case elevated importation scenario).
367

368 Given that a universal trigger may signal highly disparate levels of ZIKV risk, policy makers
369 might seek to adapt their triggers to local conditions. Suppose a policymaker wishes to design triggers
370 that indicate a 50% chance of an emerging epidemic (Fig 4C). Under the baseline importation and
371 reporting rates, only 21 of the 254 counties in Texas are expected to reach a 50% epidemic probability,
372 with triggers ranging from one (Starr County) to 21 (Dimmit County) reported autochthonous cases, with
373 a median of two cases. The remaining counties have less than a 1% chance of experiencing sustained
374 ZIKV transmission. Under an elevated importation scenario, assuming that only one fifth of ZIKV

375 importations (the symptomatic proportion) have been observed, we find that the recommended triggers
376 decrease by a mean of 1.5 reported cases (Fig S5) and the size of Texas' population at risk for sustained
377 ZIKV transmission is expected to increase from ~14% to ~30%, largely driven by increased risk in the
378 Houston metropolitan area.
379
380 **Discussion**
381 US public health authorities are responding to ZIKV outbreaks in Southern Florida, and preparing
382 for possible ZIKV outbreaks in other vulnerable regions. A key challenge is knowing when and where to
383 initiate interventions based on potentially sparse and biased ZIKV case reports. Our simple model is
384 designed to address this challenge by providing county-specific epidemic risk estimates as a function of
385 reported cases. We demonstrate its application across the 254 ecologically and demographically diverse
386 counties of Texas, a high risk state for ZIKV [7,8,10]. Based on county-level estimates for ZIKV
387 importation and transmission rates (Fig 2), we expect that most Texas counties are not at risk for a
388 sustained ZIKV epidemic (Fig 4). However, 30% of Texas' population may reside in vulnerable regions,
389 including the cities of Austin, San Antonio, and Waco along the I-35 corridor, Houston, and the Rio
390 Grande Valley. The higher the ZIKV importation rate in these locations, the higher the chance of an
391 epidemic (Fig S5). However, even in the most high risk regions of Texas, we expect far more limited
392 ZIKV transmission than observed in Central America, South America, and Puerto Rico, where R_0 has
393 been estimated to be as high as 11 [27,31,32]. Our analysis is consistent with recent introductions of
394 DENV and CHIKV into Texas that have caused small pockets of local transmission but have failed to
395 spark large epidemics.
396 Surveillance triggers--guidelines specifying situations that warrant intervention--are a key
397 component of many public health response plans. Given the urgency and uncertainty surrounding ZIKV,
398 *universal* recommendations can be both pragmatic and judicious. In choosing an appropriate trigger for an
399 environmentally and socioeconomically diverse region, policymakers must weigh the risks of acting
400 unnecessarily with those of responding too late. A risk averse policymaker will likely select a low trigger

401 (a small number of reported cases), ensuring early detection in the riskiest sites at the cost of false alarms
402 in low risk locations (Fig 4C). In that vein, the CDC recently issued conservative guidelines to state and
403 local public health agencies that suggest a two-case trigger for initiating interventions [28]. To assist
404 Texas policymakers in interpreting this recommendation, we assessed the likelihood and implication of a
405 two-case trigger for each of Texas' 254 counties, under a scenario projected from recent ZIKV data to
406 August 2016. Across counties, there is enormous variation in both the chance of a trigger and the
407 magnitude of the public health threat if and when two cases are reported. If and when two autochthonous
408 ZIKV cases are reported, only 18 of Texas' 254 counties have over a 20% chance of experiencing
409 sustained transmission (under our baseline scenario); in most of the remaining 236 counties, the threat is
410 much lower.

411 Rather than implement a universal *trigger,* which may indicate different threats in different
412 locations, one could design local surveillance triggers that correspond to a universal *risk threshold*. Our
413 modeling framework can readily identify triggers (numbers of reported cases) for indicating any specified
414 epidemic event (e.g., prevalence reaching a threshold or imminent epidemic expansion) with any
415 specified risk tolerance (e.g., 10% or 50% chance of that the event has occurred), given local
416 epidemiological conditions. As a case study, we identify *epidemic expansion* triggers in each of Texas'
417 254 counties, each designed to indicate when the probability of an epidemic exceeds 50%. Across the 21
418 counties with non-negligible probabilities of an epidemic, the recommended triggers range from one to 21
419 reported autochthonous cases, highlighting Texas' spatial risk heterogeneity. These findings apply only to
420 the early, pre-epidemic phase of ZIKV in Texas, when travel from affected regions outside the contiguous
421 US is the primary importation source, and as the epidemiological situation in those affected regions
422 remains constant. If self-sustaining outbreaks emerge within the United States, there may be county-to-
423 county or state-to-state importations, particularly in high risk regions, that are not yet included in the
424 model. For example, if the frequency of travel-related ZIKV importations from Florida increases, like the
425 imported case in El Paso, our model will need to be recalibrated.

426 This simple framework offers a flexible means for bringing current data and expert knowledge to
427 assist critical public health decision making. The design of a trigger--both the event to be detected and the
428 probability threshold upon which to take action--requires extensive public health expertise and
429 deliberation. Our case studies were motivated by formal recommendations and informal discussions with
430 state and national public health agencies. They demonstrate our data-driven approach for relating
431 observed disease activity to underlying risk in the face of great uncertainty, and address the pros, cons and
432 complexity of developing universal versus regional action plans. However, they are not meant to critique
433 or validate any specific policies, and highlight the prudence of recent ZIKV planning efforts.

434 Importantly, our analyses rest on the recent and limited scientific investigations of ZIKV's
435 biology and epidemiology, and should be continually updated as our understanding matures. Specifically,
436 our county-level estimates of R_0 are sensitive to many underlying assumptions, particularly regarding
437 temperature-dependent and socioeconomic factors (Supplement §4). Although the risk of ZIKV
438 transmission may be much higher or lower than assumed in our baseline scenario, the relative
439 vulnerabilities of counties are fairly robust to our assumptions. Thus, counties with the highest estimated
440 risks should likely be prioritized for surveillance and interventions resources. Given the minimal
441 incursions of DENV and CHIKV into Texas, we suspect that, if anything, we may be underestimating the
442 socioeconomic and behavioral impediments to ZIKV transmission in the contiguous US, and thus
443 overestimating the transmission risk of ZIKV across Texas. Given current understanding of ZIKV, we
444 believe that our county-level estimates of R_0 are reasonable, but should be interpreted and applied with
445 full consideration of their underlying uncertainty. As detailed ZIKV outbreak data become available, the
446 research community will continue to improve accuracy and reduce uncertainty in such estimates, which
447 can be readily incorporated into this risk assessment framework.

448 The reporting rate dictates the relationship between observed cases and the underlying outbreak,
449 and its magnitude impacts the timeliness and precision of detection. If only a small fraction of cases are
450 reported, the first few reported cases may correspond to a wide range of underlying epidemiological
451 conditions, from isolated introductions to a growing epidemic. In contrast, if most cases are reported,

452 policymakers can wait longer (in terms of the number of reported cases) to trigger interventions and have
453 more confidence in their epidemiological assessments. ZIKV reporting rates are expected to remain quite
454 low, because an estimated 80% of infections are asymptomatic, and DENV reporting rates have
455 historically matched its asymptomatic proportion [16,33]. Obtaining a realistic estimate of the ZIKV
456 reporting rate is arguably as important as increasing the rate itself, with respect to reliable situational
457 awareness and forecasting. An estimated 8-22% of ZIKV infections were reported during the 2013-2014
458 outbreak in French Polynesia [32]; however estimates ranging from 1 to 10% have been reported during
459 the ongoing epidemic in Columbia [31,34]. While these provide a baseline estimate for the US, there are
460 many factors that could increase (or decrease) the reporting rate, such as ZIKV awareness among both the
461 public and health-care practitioners, or active surveillance of regions with recent ZIKV cases (our
462 analysis assumes continuous passive surveillance). Thus, rapid estimation of the reporting rate should be a
463 high priority. While some methods require extensive epidemiological data not typically available early in
464 an outbreak [35], a new method exploiting early outbreak viral sequence data was introduced during the
465 recent West African Ebola epidemic [36]. However, as of July 2016, there are no US ZIKV sequences
466 available on GenBank and few available from other regions.

467 Unlike other risk estimation methods, this framework is designed to support risk assessments
468 both prior to ZIKV outbreaks and, in real-time, upon the early detection of cases through the lens of
469 potentially sparse and biased surveillance data [37,38]. It can facilitate situational awareness, allowing
470 analysts to translate case counts into estimates of overall prevalence and the potential for future spread,
471 albeit with large error bars. It can also support the development of response plans, by forcing
472 policymakers to be explicit about risk tolerance, that is, the certainty needed before sounding an alarm,
473 and quantifying the consequences of premature or delayed interventions. For example, should ZIKV-
474 related pregnancy advisories be issued when there is only 5% chance of an impending epidemic? 10%
475 chance? 80%? A policymaker has to weigh the costs of false positives--resulting in unnecessary fear
476 and/or intervention--and false negatives--resulting in suboptimal disease control and prevention--
477 complicated by the difficulty inherent in distinguishing a false positive from a successful intervention.

The more risk averse the policymaker (with respect to false negatives), the earlier the trigger should be, which can be exacerbated by low reporting rates, high importation rate, and inherent ZIKV transmission potential. In ZIKV prone regions with low reporting rates, even risk tolerant policymakers should act quickly upon seeing initial cases; in lower risk regions, longer waiting periods may be prudent.

Acknowledgements

We acknowledge the Texas Department of State Health Services (DSHS) for providing historic arbovirus county importation data, and MUG Kraemer and A Perkins for providing Texas mosquito abundance data and technical guidance. We also thank M Johansson and M Meltzer of the Centers for Disease Control and Prevention (CDC) and Dr. John Hellerstedt, commissioner of the Texas DSHS, for critical conversations regarding the modeling framework and its application to Texas and national ZIKV planning efforts. We finally acknowledge the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing High performance computing resources that have contributed to the research results reported within this paper. URL: http://www.tacc.utexas.edu.

Code Availability

The R package, rtZIKVrisk, has been created to accompany the manuscript and can be found here: https://github.com/sjfox/rtZIKVrisk. The package provides functionality to simulate and analyze ZIKV outbreaks according to our framework, and recreate the figures presented in the manuscript.

References

1. Gulland A. Zika virus is a global public health emergency, declares WHO. BMJ. BMJ Publishing Group Ltd; 2016;352. doi:10.1136/bmj.i657
2. PAHO, WHO. Suspected and confirmed Zika cases reported by countries and territories in the Americas, 2015-2016 [Internet]. 2016 p. http://ais.paho.org/phip/viz/ed_zika_epicurve.asp.
3. World Health Organization. Zika situation report: 18 August 2016. 2016.

504 4. About estimated range of Aedes aegypti and Aedes albopictus in the United States, 2016 Maps. In:
505 Atlanta: Centers for Disease Control and Prevention [Internet]. [cited 1 Jan 2016]. Available:
506 http://www.cdc.gov/zika/vector/range.html
507 5. Florida Department of Health. Department of Health Daily Zika Update - August 30, 2016. 2016.
508 6. Kraemer MUG, Sinka ME, Duda KA, Mylne A, Shearer FM, Barker CM, et al. The global
509 distribution of the arbovirus vectors Aedes aegypti and Ae. albopictus. Elife. eLife Sciences
510 Publications Limited; 2015;4: e08347. Available: http://elifesciences.org/content/4/e08347v3
511 7. Messina JP, Kraemer MU, Brady OJ, Pigott DM, Shearer FM, Weiss DJ, et al. Mapping global
512 environmental suitability for Zika virus. Elife. eLife Sciences Publications Limited; 2016;5:
513 e15272. Available: https://elifesciences.org/content/5/e15272
514 8. Gardner LM, Chen N, Sarkar S. Global risk of Zika virus depends critically on vector status of
515 Aedes albopictus. Lancet Infect Dis. Elsevier; 2016;16: 522–523. Available:
516 http://www.thelancet.com/article/S1473309916001766/fulltext
517 9. Peterson AT, Osorio J, Qiao H, Escobar LE. Zika virus, elevation, and transmission risk. PLoS
518 Curr Outbreaks. 2016;9. doi:10.1371/currents.outbreaks.a832cf06c4bf89fb2e15cb29d374f9de
519 10. Monaghan AJ, Morin CW, Steinhoff DF, Wilhelmi O, Hayden M, Quattrochi DA, et al. On the
520 Seasonal Occurrence and Abundance of the Zika Virus Vector Mosquito Aedes Aegypti in the
521 Contiguous United States. PLoS Curr. 2016; 1–31.
522 doi:10.1371/currents.outbreaks.50dfc7f46798675fc63e7d7da563da76
523 11. Texas Department of State Health and Human Services. Arbovirus Activity in Texas 2014
524 Surveillance Report [Internet]. 2014. Available:
525 https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=3&ved=0ahUKEwiE8Oz
526 U0I_OAhVG5SYKHTBTA9QQFggoMAI&url=https://www.dshs.texas.gov/WorkArea/linkit.asp
527 x?LinkIdentifier=id&ItemID=8589999503&usg=AFQjCNGO9cW43_oOrfmhjWIKseBh9Q1n9g
528 &s
529 12. CDC, Texas Department of State Health and Human Services. Dengue Hemorrhagic Fever ---

530 U.S.-Mexico Border, 2005 [Internet]. 2007. Available:
531 http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5631a1.htm
532 13. Zika in Texas. In: Texas Department of State Health Services [Internet]. 2016 [cited 5 Apr 2016].
533 Available: http://www.texaszika.org/
534 14. Lessler JT, Ott CT, Carcelen AC, Konikoff JM, Williamson J, Bi Q, et al. Times to key events in
535 the course of Zika infection and their implications: a systematic review and pooled analysis. Bull
536 World Health Organ. 2016; doi:10.2471/BLT.16.174540
537 15. Lloyd-Smith JO, Funk S, McLean AR, Riley S, Wood JLN. Nine challenges in modelling the
538 emergence of novel pathogens. Epidemics. Elsevier B.V.; 2014;10: 35–39.
539 doi:10.1016/j.epidem.2014.09.002
540 16. Duffy MR, Chen T-H, Hancock WT, Powers AM, Kool JL, Lanciotti RS, et al. Zika Virus
541 Outbreak on Yap Island, Federated States of Micronesia. N Engl J Med. 2009;360: 2536–2543.
542 doi:10.1056/NEJMoa0805715
543 17. Alfaro-Murillo JA, Parpia AS, Fitzpatrick MC, Tamagnan JA, Medlock J, Ndeffo-Mbah ML, et al.
544 A Cost-Effectiveness Tool for Informing Policies on Zika Virus Control. Carabin H, editor. PLoS
545 Negl Trop Dis. 2016;10: e0004743. doi:10.1371/journal.pntd.0004743
546 18. PAHO and WHO. Zika-Epidemiological Report - Anguilla. 2016; 1–2.
547 19. ZIKA CDC Interim Response Plan [Internet]. 2016. Available: http://www.cdc.gov/zika/pdfs/zika-
548 draft-interim-conus-plan.pdf
549 20. Getz WM, Lloyd-Smith JO. Basic methods for modeling the invasion and spread of contagious
550 diseases. Dis Evol Model concepts, data Anal AMS-DIMACS Ser. 2006;71: 87. Available:
551 http://books.google.com/books?hl=en&lr=&id=5uMf4qCmghEC&oi=fnd&pg
552 =PA87&dq=Basic+Methods+for+Modeling+the+Invasion+and+Spread+of+Contagious+Dis
553 eases&ots=M3miy-
554 szrR&sig=Re_k7i7YQoXGHkktCbLilsv41XQ\nhttp://books.google.com/books?h
555 21. Office of Travel & Tourism Industries. US Monthly Arrivals Trend Line [Internet]. 2014.

556 Available: http://travel.trade.gov/view/m-2014-I-001/index.html
557 22. Jaynes ET. Information Theory and Statistical Mechanics. Phys Rev. American Physical Society;
558 1957;106: 620–630. doi:10.1103/PhysRev.106.620
559 23. Musso D, Cao-Lormeau VM, Gubler DJ. Zika virus: following the path of dengue and
560 chikungunya? Lancet. Elsevier Ltd; 2015;386: 243–244. doi:10.1016/S0140-6736(15)61273-9
561 24. WHO P and. Zika-Epidemiological Report - Mexico. 2016.
562 25. Wolsey LA. Integer programming. Wiley; 1998.
563 26. Merow C, Smith MJ, Silander JA. A practical guide to MaxEnt for modeling species'
564 distributions: what it does, and why inputs and settings matter. Ecography (Cop). 2013;36: 1058–
565 1069. doi:10.1111/j.1600-0587.2013.07872.x
566 27. Alex Perkins T, Siraj AS, Ruktanonchai CW, Kraemer MUG, Tatem AJ. Model-based projections
567 of Zika virus infections in childbearing women in the Americas. Nat Microbiol. 2016;1: 16126.
568 doi:10.1038/nmicrobiol.2016.126
569 28. Interim CDC recommendations for Zika vector control in the continental United States. In:
570 Centers for Disease Control and Prevention [Internet]. 2016 [cited 5 Jan 2016]. Available:
571 http://www.cdc.gov/zika/public-health-partners/vector-control-us.html
572 29. Majumder MS, Cohn E, Fish D, Brownstein JS. Estimating a feasible serial interval range for Zika
573 fever. Bull World Health Organ. 2016; doi:10.2471/BLT.16.171009
574 30. Lloyd AL. Realistic Distributions of Infectious Periods in Epidemic Models: Changing Patterns of
575 Persistence and Dynamics. Theor Popul Biol. 2001;60: 59–71. doi:10.1006/tpbi.2001.1525
576 31. Rojas DP, Dean NE, Yang Y, Kenah E, Quintero J, Tomasi S, et al. The Epidemiology and
577 Transmissibility of Zika Virus in Girardot and San Andres Island, Colombia. bioRxiv. 2016;
578 Available: http://biorxiv.org/content/early/2016/04/24/049957.abstract
579 32. Kucharski AJ, Funk S, Eggo RM, Mallet H, Edmunds WJ, Nilles EJ. Transmission dynamics of
580 Zika virus in island populations : a modelling analysis of the 2013 – 14 French Polynesia outbreak.
581 bioRxiv. 2016; 1–15.

582 33. Dechant E, Rigau-Perez J. Hospitalizations for suspected dengue in Puerto Rico, 1991-1995:
583 estimation by capture-recapture methods. The Puerto Rico Association of Epidemiologists. Am J
584 Trop Med Hyg. 1999;61: 574–578. Available: http://www.ajtmh.org/content/61/4/574.short
585 34. Zhang Q, Sun K, Chinazzi M, Pastore-Piontti A, Dean NE, Rojas DP, et al. Projected spread of
586 Zika virus in the Americas. bioRxiv. 2016; 66456. doi:10.1101/066456
587 35. Bhatt S, Gething PW, Brady OJ, Messina JP, Farlow AW, Moyes CL, et al. The global distribution
588 and burden of dengue. Nature. Nature Publishing Group; 2013;496: 504–507.
589 doi:10.1038/nature12060
590 36. Scarpino S V, Iamarino A, Wells C, Yamin D, Ndeffo-Mbah M, Wenzel NS, et al.
591 Epidemiological and viral genomic sequence analysis of the 2014 ebola outbreak reveals clustered
592 transmission. Clin Infect Dis. Oxford University Press; 2015;60: 1079–82. Available:
593 http://cid.oxfordjournals.org/content/early/2015/01/05/cid.ciu1131.full
594 37. Cauchemez S, Epperson S, Biggerstaff M, Swerdlow D, Finelli L, Ferguson NM. Using Routine
595 Surveillance Data to Estimate the Epidemic Potential of Emerging Zoonoses: Application to the
596 Emergence of US Swine Origin Influenza A H3N2v Virus. Peiris JSM, editor. PLoS Med.
597 2013;10: e1001399. doi:10.1371/journal.pmed.1001399
598 38. Blumberg S, Lloyd-Smith JO. Comparing methods for estimating R0 from the size distribution of
599 subcritical transmission chains. Epidemics. 2013;5: 131–145. doi:10.1016/j.epidem.2013.05.002
600
601
602 **S1 Supporting Information.** Supporting information for "Real-time Zika risk assessment in the United
603 States."
604 **S2 Supporting Data.** Raw Texas county arbovirus importation data used for estimating county ZIKV
605 importation risk. Data provided by Texas DSHS.
606 **S3 Supporting Data.** Raw county socioeconomic data used for estimating county ZIKV importation risk.
607 Data accessed from: https://www.census.gov/programs-surveys/acs/data.html

S4 Supporting Data. Texas county ZIKV risk estimates used as inputs for the ZIKV simulation model. Outputs of the importation and transmission risk analysis.