



Practice of Epidemiology

Reassessing Serosurvey-Based Estimates of the Symptomatic Proportion of Zika Virus Infections

Patrick K. Mitchell, Luis Mier-y-Teran-Romero, Brad J. Biggerstaff, Mark J. Delorey, Maite Aubry, Van-Mai Cao-Lormeau, Matthew J. Lozier, Simon Cauchemez, and Michael A. Johansson*

* Correspondence to Dr. Michael A. Johansson, Dengue Branch, Division of Vector-Borne Diseases, Centers for Disease Control and Prevention, 1324 Calle Cañada, San Juan, Puerto Rico 00920 (e-mail: mjohansson@cdc.gov).

Initially submitted February 14, 2018; accepted for publication August 21, 2018.

Since the 2007 Zika epidemic in the Micronesian state of Yap, it has been apparent that not all people infected with Zika virus (ZIKV) experience symptoms. However, the proportion of infections that result in symptoms remains unclear. Existing estimates have varied in their interpretation of symptoms due to other causes and the case definition used, and they have assumed perfect test sensitivity and specificity. Using a Bayesian model and data from ZIKV serosurveys in Yap (2007), French Polynesia (2013–2014), and Puerto Rico (2016), we found that assuming perfect sensitivity and specificity generally led to lower estimates of the symptomatic proportion. Incorporating reasonable assumptions for assay sensitivity and specificity, we estimated that 27% (95% credible interval (CrI): 15, 37) (Yap), 44% (95% CrI: 26, 66) (French Polynesia), and 50% (95% CrI: 34, 92) (Puerto Rico) of infections were symptomatic, with variation due to differences in study populations, study designs, and case definitions. The proportion of ZIKV infections causing symptoms is critical for surveillance system design and impact assessment. Here, we accounted for key uncertainties in existing seroprevalence data and found that estimates for the symptomatic proportion ranged from 27% to 50%, suggesting that while the majority of infections are asymptomatic or mildly symptomatic, symptomatic infections might be more common than previously estimated.

Markov chain; serological tests; signs and symptoms; Zika virus; Zika virus infection

Abbreviations: CrI, credible interval; IgM, immunoglobulin M; ZIKV, Zika virus.

Zika virus (ZIKV; family *Flaviviridae*, genus *Flavivirus*) has been known to infect humans since the mid-20th century, but was not known to cause outbreaks until 2007, when ZIKV emerged in the Micronesian state of Yap and infected an estimated 73% of the population (1). This epidemic potential was confirmed by an explosive outbreak in French Polynesia in 2013–2014 (2), subsequent spread to other Pacific islands (3), and emergence in the Americas in 2015–2016 (4, 5).

Since the Yap epidemic, it has been clear that many ZIKV infections are asymptomatic, yet severe complications such as birth defects due to congenital infection (6) and Guillain-Barré syndrome (7) also occur. More common clinical manifestations include fever, rash, conjunctivitis, and joint pain (1, 8). A study following the Yap outbreak found that 38% of those with serological evidence of infection had symptoms consistent with Zika disease, as did 19% of those without serological evidence of infection (1). With adjustments for sampling design, the study

estimated that approximately 18% of those infected had symptoms attributable to ZIKV infection, implying that approximately 1 in 5 ZIKV infections were symptomatic. In the same outbreak, however, 108 out of 113 (96%) persons who sought health care with Zika symptoms and had a convalescent serum sample collected were positive for ZIKV RNA or anti-ZIKV immunoglobulin M (IgM). This high rate of positivity indicates that the majority of people with Zika-like symptoms who sought care and were tested were indeed infected with ZIKV, suggesting that many of the serosurvey participants who reported Zika symptoms but had negative IgM test results might in fact have been infected with ZIKV. If large numbers of truly infected symptomatic participants had negative IgM results, estimates of the prevalence of symptoms among those infected could be biased downward.

Two other studies estimated a higher proportion of symptomatic infections than that reported in the Yap study. A serosurvey

at the end of the French Polynesia outbreak estimated that 47%–67% of persons with positive serology reported symptoms suggestive of Zika disease, although this estimate did not account for symptoms from non-ZIKV etiologies (9). While not a population-based study, a cluster investigation in Puerto Rico found that 43% of those with laboratory evidence of ZIKV infection were symptomatic as compared with 13% of those with ZIKV-negative laboratory results (8). Furthermore, each study used a different set of symptoms in their case definition, complicating our understanding of symptomatic disease caused by ZIKV infection.

Understanding the Zika symptomatic proportion, the proportion of ZIKV infections that result in a defined set of symptoms, is essential to Zika epidemiology because it is a key link between observed disease and the true incidence of infection. This link is important to surveillance systems, which generally rely on detecting symptomatic cases, and to risk assessments, which rely on estimates of infection dynamics to support prevention and control activities. Here, we expanded previous estimation approaches by using a Bayesian inference framework to account for imperfect assay performance and for the possibility that reported symptoms could have been caused by both ZIKV infection and by other, unidentified causes over the same reporting period. We applied this more refined model to data from the Yap, French Polynesia, and Puerto Rico studies to compute revised estimates of the Zika symptomatic proportion that account for these effects. We evaluated the modeling framework first by assessing estimates under assumptions of imperfect sensitivity and specificity in the range, 70%–100%. Next, we estimated the Zika symptomatic proportion with reasonable prior estimates for assay sensitivity and specificity based on characteristics of assays for ZIKV and other flaviviruses. Finally, we estimated the symptomatic proportion considering the prior information on assay characteristics and relaxing the assumption that symptoms due to ZIKV infection and other causes were mutually exclusive.

METHODS

Outbreak data

Original, unadjusted serosurvey and symptom data from the outbreaks in Yap (2007), French Polynesia (2013–2014), and Puerto Rico (2016) were used as model inputs (Table 1) (1, 8, 9). We did not stratify by any demographic factors such as age or sex. For French Polynesia, the number of seronegative symptomatic individuals was not previously published and is included here. For the Puerto Rico data set, we used only serological test results, not those obtained through

nucleic acid testing, for comparability to the other 2 studies. We assumed that serological test results did not differ by sex based on evidence of similar seroprevalence by sex for each study considered and no other evidence of differences in antibody duration by sex. The symptomatic case definition in each study was based on a list of symptoms, at least one of which was required to be counted as symptomatic: In the Yap study these included joint pain, rash, or conjunctivitis; in the French Polynesia study, joint pain, rash, conjunctivitis, hand or foot swelling, or fever; and in Puerto Rico, joint pain or rash.

Model

Let Z indicate the event that a person was (Z^+) or was not (Z^-) infected with ZIKV, and T indicate the event that a person tested positive (T^+) or negative (T^-) for ZIKV. Let S indicate that a person reported (S^+) or did not report (S^-) symptoms consistent with the Zika case definition. Finally, among people with symptoms, let C indicate a person with symptoms caused by ZIKV infection and C' a person with symptoms not caused by ZIKV infection. The intersection of C and C' , $C \cap C'$, is the event that a person reported symptoms caused both by ZIKV and by another etiology. We use the parameter γ to represent a binary choice between inclusion ($\gamma = 1$) or exclusion ($\gamma = 0$) of this group, $C \cap C'$. This parameter is included because this possibility was excluded from previous analyses and can be particularly important for seroprevalence studies in which the extended follow-up period allows for multiple episodes of illness. Finally, note that under this model, only values for S and T are observed, with $S^+ = C \cup C'$.

Using these definitions and the respective parameters described in Table 2, we used a Bayesian model and employed Markov chain Monte Carlo methods to estimate: P_Z , the probability that an individual was truly infected by ZIKV; P_{SZ} , the probability that a ZIKV infection resulted in symptoms; and P_{SB} , the baseline probability of symptoms not related to ZIKV infection. The number of people in each category of test result and symptom report, $\mathbf{X} = (X_{T^+S^+}, X_{T^+S^-}, X_{T^-S^+}, X_{T^-S^-})$, was assumed to be the result of a multinomial process $\mathbf{X} \sim \text{mult}(n, \mathbf{p})$, where n is the total number of individuals in the population and \mathbf{p} the vector of probabilities for each specific group as follows:

$$p_{T^+S^+} = sen P_Z (P_{SZ} + P_{SB} - (\gamma P_{SZ} P_{SB})) + (1 - spe)(1 - P_Z) P_{SB}$$

$$p_{T^+S^-} = sen P_Z (1 - P_{SZ} - P_{SB} + (\gamma P_{SZ} P_{SB})) + (1 - spe)(1 - P_Z)(1 - P_{SB})$$

$$p_{T^-S^+} = spe (1 - P_Z) P_{SB} + (1 - sen) P_Z \times (P_{SZ} + P_{SB} - (\gamma P_{SZ} P_{SB}))$$

$$p_{T^-S^-} = spe (1 - P_Z)(1 - P_{SB}) + (1 - sen) P_Z \times (1 - P_{SZ} - P_{SB} + (\gamma P_{SZ} P_{SB})).$$

Sensitivity and specificity of serological testing, denoted by sen and spe , respectively, were assigned either fixed values or

Table 1. Seroprevalence Data From Yap (2007), French Polynesia (2013–2014), and Puerto Rico (2016)

Test Result (T)	Symptoms Reported (S)					
	Yap		French Polynesia		Puerto Rico	
	Yes	No	Yes	No	Yes	No
Positive	156	258	55	42	48	49
Negative	27	116	27	72	35	235

Table 2. Parameter Definitions Used in an Analysis of the Symptomatic Proportion of Zika Virus Infections in Multiple Countries, 2018

Parameter	Description	Notation
P_Z	Probability of ZIKV infection	$P[Z^+]$
P_{SZ}	Probability of having symptoms caused by ZIKV	$P[C Z^+]$
P_{SB}	Baseline probability of symptoms	$P[C' Z^-] = P[C' Z^+]^a$
sen	Sensitivity of the test	$P[T^+ Z^+]$
spe	Specificity of the test	$P[T^- Z^-]$
γ	Indicator for inclusion of outcome $C \cap C'$	$\begin{cases} 0 & \text{do not include } C \cap C' \\ 1 & \text{include } C \cap C' \end{cases}$

Abbreviations: C, symptoms caused by ZIKV; C', symptoms not caused by ZIKV; T⁺, positive diagnostic test result; T⁻, negative diagnostic test result; ZIKV, Zika virus; Z⁺, infected with Zika virus; Z⁻, not infected with Zika virus.

^a We assumed that the rate of symptoms not caused by ZIKV is the same for ZIKV-infected and not-ZIKV-infected persons.

informative prior distributions, as described in Web Appendix 1 (available at <https://academic.oup.com/aje>), for each analysis. P_Z , P_{SZ} , and P_{SB} were considered unknown and were assigned *beta*(1, 1) prior distributions. The Markov chain Monte Carlo simulation was performed using JAGS, version 4.2.0, implemented through R, version 3.3.2 (R Foundation for Statistical Computing, Vienna, Austria), with the rjags, version 4–6, and coda, version 0.19–1, packages (10–12). Simulations were initialized with randomly drawn values between 0 and 1 for all free parameters. For simulations where $\gamma = 0$, the sum of the initial values of P_{SZ} and P_{SB} was restricted to be less than 1, as the sum of these parameters is the total probability of a person with ZIKV infection experiencing symptoms and that probability cannot be greater than 1. For all simulations, 1,000,000 iterations from each of 3 chains were performed with a thinning interval of 1,000 and a burn-in of 10,000 iterations. Convergence was assessed using Gelman and Rubin's diagnostic (13).

Assessing the importance of imperfect sensitivity and specificity

We scanned over a broad range of sensitivity and specificity values to characterize the effect of these parameters on estimates of the symptomatic proportion. For each data set, we performed 49 Markov chain Monte Carlo simulations for each location using fixed sensitivity and specificity values in the range of 0.70–1.00 in intervals of 0.05 and assuming symptoms were not mutually exclusive ($\gamma = 1$).

Estimating the symptomatic proportion

For each outbreak, we estimated P_{SZ} under 3 conditions. For the “original” model, we assumed 100% sensitivity, 100% specificity, and mutually exclusive symptoms ($\gamma = 0$). We then used informative prior distributions for sensitivity and specificity, as described in Web Appendix 1, to estimate P_{SZ} assuming an “imperfect test” and mutually exclusive symptoms ($\gamma = 0$). Sensitivity was assigned a prior distribution with a median value of 0.95 with 95% probability mass in the interval 0.85–0.99, and specificity had a median value of 0.96 with 95% probability mass in the interval 0.80–1.00 (Web Table 1). Finally, we used the same informative priors for sensitivity and specificity but

allowed symptoms not to be mutually exclusive ($\gamma = 1$), “imperfect test with symptom overlap.”

RESULTS

We first estimated the proportion of infections that were symptomatic and the overall proportion infected at fixed values of sensitivity and specificity in the range of 70%–100% for Yap, French Polynesia, and Puerto Rico while allowing for symptoms to be caused by Zika and other illnesses. In general, the estimated proportion of symptomatic infections increased with lower sensitivity or specificity (Figure 1A–1C). As testing accuracy decreases, distinguishing those who were truly infected and those who were not is more difficult. As a result, the observed difference in symptom frequency between those who test positive and those who test negative is relatively small and represents an underestimate of the true symptomatic proportion. When considering the lower sensitivity or specificity explicitly, the estimated symptomatic proportion is therefore generally higher than the observed difference when assay accuracy was not accounted for. This effect, however, might change at boundary conditions. For example, if 85% of symptomatic individuals tested positive (as in the Yap seroprevalence study), and specificity is fixed, the maximum estimated Zika symptomatic proportion occurs at 85% sensitivity. At sensitivities below this threshold, more actual infections can be accounted for only as asymptomatic cases, driving the estimated symptomatic proportion down. In most situations, however, assuming perfect sensitivity and specificity leads to lower estimates of the proportion of infections resulting in symptoms.

The impact of test sensitivity and specificity varied by epidemiologic setting. For Yap, where ZIKV prevalence was high, variation of the symptomatic proportion across specificity values for a given sensitivity was small (Figure 1A). The opposite was true for Puerto Rico, where prevalence in the study population was low and estimates were highly dependent on specificity (Figure 1C). French Polynesia, which had ZIKV prevalence near 50%, was intermediate between Yap and Puerto Rico with respect to the relative effects of sensitivity and specificity (Figure 1B). The estimated proportion infected also varied,

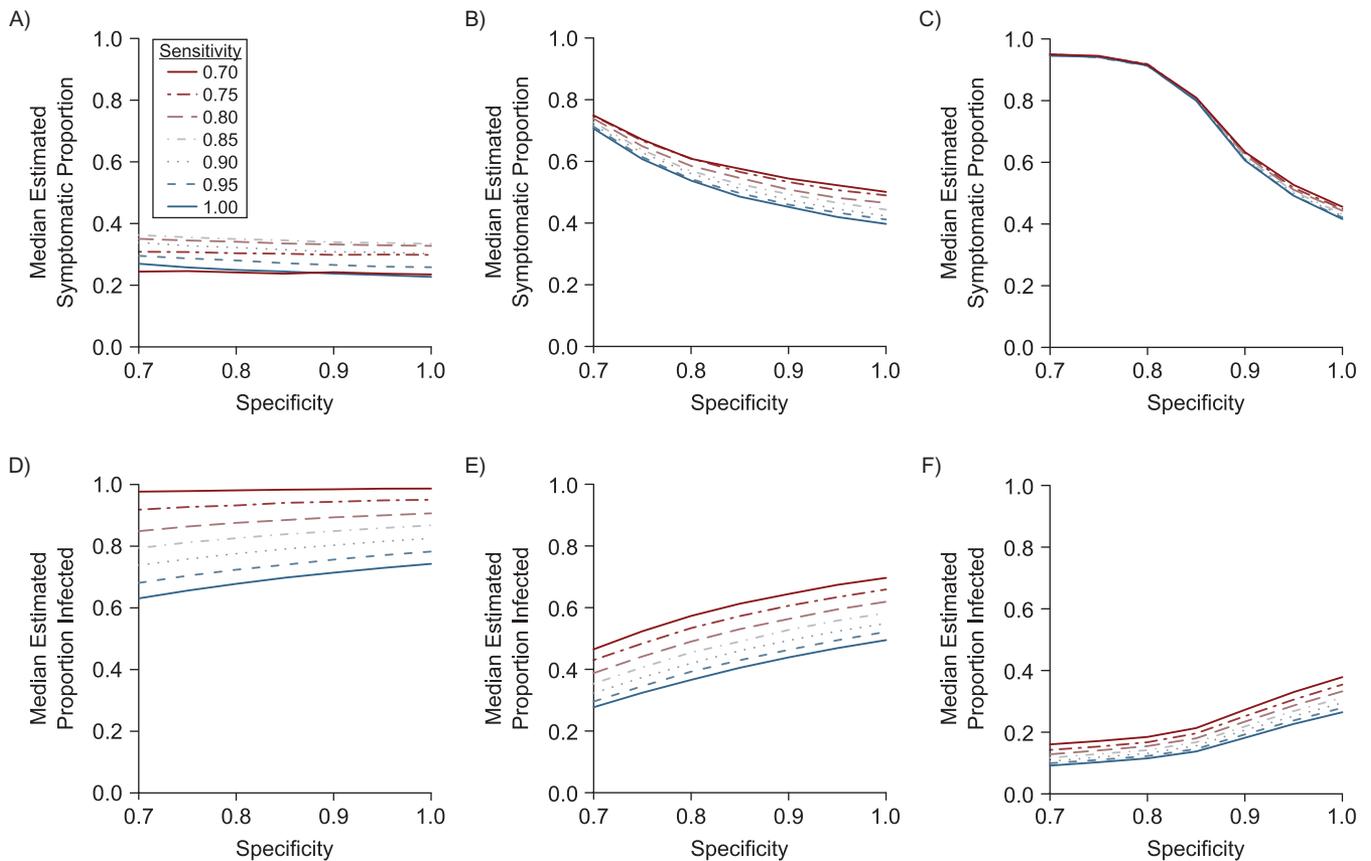


Figure 1. Simulation showing the effect of sensitivity and specificity on estimates of the symptomatic proportion and Zika virus (ZIKV) infection prevalence in Yap (2007), French Polynesia (2013–2014), and Puerto Rico (2016). The top row shows the effect of different values of sensitivity (line style) and specificity (x-axis) on point estimates of the median proportion of ZIKV infections that were symptomatic in studies in Yap (A), French Polynesia (B), and Puerto Rico (C). The bottom row shows the corresponding estimates of infection prevalence for the study populations in Yap (D), French Polynesia (E), and Puerto Rico (F). In each of these simulations, we assumed that symptoms caused by ZIKV and other illnesses could overlap ($\gamma = 1$).

increasing with higher specificity (a higher proportion of test positives are true infections) and decreasing with higher sensitivity (a lower proportion of test negatives are true positives) in each setting (Figure 1D–1F). Again, the strength of the effects was different across locations, with sensitivity being the most important driver for Yap and specificity the most important driver for Puerto Rico.

With sensitivity and specificity both fixed at 100% and assuming that all symptoms were caused either by ZIKV or by some other disease, but not by both simultaneously (“original”), the posterior medians and 95% credible intervals for the estimated proportion of ZIKV infections that caused the symptoms recorded in each study population were 0.18 (95% credible interval (CrI): 0.11, 0.26) for Yap, 0.29 (95% CrI: 0.16, 0.42) for French Polynesia, and 0.37 (95% CrI: 0.25, 0.47) for Puerto Rico (Figure 2A–2C). The estimate for Yap is, as expected, very similar to the previously published estimate of 0.18 (95% CrI: 0.10, 0.27) despite slightly different methods (1). We then incorporated prior distributions for sensitivity and specificity (approximately 85%–99% and 80%–100%, respectively) and reestimated the proportion of symptomatic infections (“imperfect test”). For all locations, the estimates increased

slightly and had higher upper uncertainty bounds: 0.23 (95% CrI: 0.12, 0.36) for Yap, 0.33 (95% CrI: 0.18, 0.52) for French Polynesia, and 0.44 (95% CrI: 0.29, 0.85) for Puerto Rico. This formulation implicitly assumes that the symptoms a person reports result from ZIKV or a non-ZIKV illness but not both. Next, we relaxed this assumption by accounting for the possibility that a single individual could experience symptoms from multiple causes independently (“imperfect test with symptom overlap”), and again the estimates increased. The posterior medians for the proportion of symptomatic infections in this model were 0.27 (95% CrI: 0.15, 0.37) for the study in Yap, 0.44 (95% CrI: 0.26, 0.66) for French Polynesia, and 0.50 (95% CrI: 0.34, 0.92) for Puerto Rico. The estimated proportion of people infected with ZIKV who had symptoms from both ZIKV infection and another cause also varied across studies: 0.04 (95% CrI: 0.01, 0.05) for the study in Yap, 0.11 (95% CrI: 0.07, 0.17) for French Polynesia, and 0.06 (95% CrI: 0.04, 0.12) for Puerto Rico.

Prevalence estimates also varied between these models (Figure 2D–2F). Allowing for imperfect sensitivity and specificity led to a slight prevalence increase in Yap and decrease in Puerto Rico, with increased uncertainty in all 3 settings.

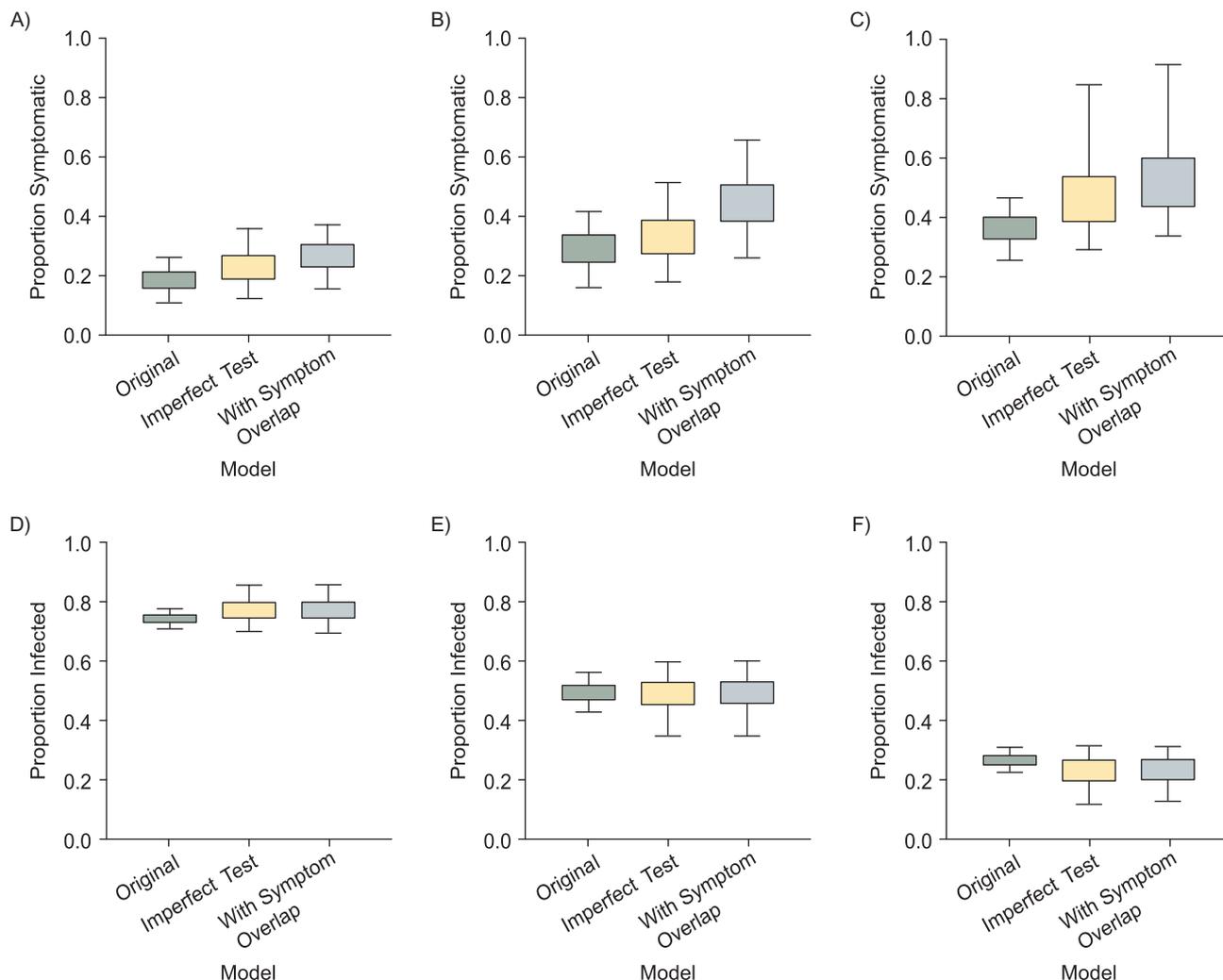


Figure 2. Estimated symptomatic proportion and Zika virus (ZIKV) infection prevalence under different models in Yap (2007), French Polynesia (2013–2014), and Puerto Rico (2016). The top row shows the estimated proportion of ZIKV infections that caused symptoms in studies in Yap (A), French Polynesia (B), and Puerto Rico (C), using 3 different models: “original” model, with a perfect assay and mutually exclusive causes of symptoms; “imperfect test,” with an imperfect assay and mutually exclusive causes of symptoms; and “imperfect test with symptom overlap,” with an imperfect assay and symptoms that could simultaneously result from both infection and other causes. Note that each study used unique case definitions for identifying symptomatic individuals. The bottom row shows the corresponding estimates of infection prevalence for the study populations in Yap (D), French Polynesia (E), and Puerto Rico (F). The boxes and bars show the 50% and 95% credible intervals.

Including potential symptom overlap did not have a substantial effect on prevalence estimates.

Under all models, the estimated baseline prevalence of symptoms not caused by ZIKV during each outbreak corresponded closely to the inclusiveness of the particular symptoms used to define the Zika case definition. Estimates were greatest for the French Polynesia data, which used the broadest definition, and lowest for Puerto Rico, which used the most restrictive set of symptoms (Web Figure 1).

DISCUSSION

Data from the first reported ZIKV outbreak, in Yap, suggested that Zika symptoms occur in approximately 1 in 5 ZIKV infections (1). However, as large ZIKV epidemics spread through the

Pacific and the Americas, higher estimates of the symptomatic proportion emerged (8, 9). We used data on ZIKV outbreaks from Yap, French Polynesia, and Puerto Rico and a Bayesian model to show that, in addition to differences in symptomatic case definitions, assumptions about serological assay sensitivity and specificity and potential symptom overlap between Zika and other diseases can affect estimates of the Zika symptomatic proportion. Using reasonable values for sensitivity and specificity, and assuming symptoms could result from ZIKV infection, other diseases, or both, we estimated that 15%–37% of ZIKV infections caused symptoms in the study in Yap, 26%–66% in French Polynesia, and 34%–92% in Puerto Rico, according to their respective case definitions.

Serosurvey data have been a key element in understanding the spread of ZIKV (1, 9, 14–18), but accounting for the characteristics of the serological assay used could make inferences

more accurate. As shown here, the sensitivity and specificity of diagnostic assays play a key role in estimating the proportion of ZIKV infections that cause symptoms and the proportion of the population infected. However, these values are not well-characterized and likely vary across different assays and populations due to cross-reactivity with other flaviviruses. Additionally, antibody titers vary over time, with IgM titers in particular rising in the first week after symptom onset and decreasing within several months as IgM titers do for dengue (19, 20).

The effects of imperfect sensitivity and specificity depend on the epidemiologic context, specifically the prevalence of infection. Assay sensitivity affects the classification of those who are truly infected, and therefore has the greatest impact in high-seroprevalence settings. Symptomatic infected individuals with false-negative serology inflate the prevalence of symptoms among those thought to be uninfected, leading to a lower perceived difference in the probability of being symptomatic. Meanwhile, specificity had a greater impact on estimates in low-prevalence settings, given that specificity determines how accurately the infection status of the relatively larger uninfected population is classified. In these settings, uninfected asymptomatic individuals with false-positive serology reduce the prevalence of symptoms among those who test positive and are thought to be infected. Location-specific estimates might also benefit from incorporating other available data, such as the Yap data on symptomatic people seeking care, which provide another indicator of the prevalence of ZIKV infection among people with symptoms during that outbreak.

Another challenge in assessing the proportion of infections that are symptomatic is the lack of a consensus clinical case definition. Disparate clinical criteria are used in the case definitions of health organizations, including the US Centers for Disease Control and Prevention and the World Health Organization, each of which have different discriminatory characteristics (21). Each of the 3 studies considered here defined a symptomatic case as a person experiencing at least one of a set of symptoms, with a different set for each study. Between the 2 studies with most similar study designs, Yap and French Polynesia, the latter had a broader case definition and a higher estimated symptomatic proportion. Despite having the most limited case definition, Puerto Rico had the highest estimated symptomatic proportion, possibly due to a different study design. In that study, participants resided near recently infected individuals and might have been recently infected themselves, whereas in the Yap and French Polynesia studies, participants recalled symptoms from up to several months before. The longer follow-up periods might have more recall bias but also more opportunity for participants to experience other diseases; higher baseline symptom risk was evident for both of the longer studies (Web Figure 1). The estimated proportion of ZIKV infections with reported symptoms caused by ZIKV and another cause was highest for French Polynesia, which could indicate that the longer follow-up period in addition to the broader case definition led to observing more symptoms among infected individuals. This possibility is particularly likely as ZIKV and 2 dengue virus serotypes circulated concurrently in French Polynesia (2). While the follow-up period for Yap was similar, Yap had no evidence of dengue or other outbreaks in the population over that period, possibly leading to the relatively lower estimated overlap.

Different case definitions, recall bias, and the local epidemiology of other diseases clearly contribute to uncertainty and variation between studies.

Several other limitations inhibited estimation of an overall symptomatic proportion. We did not account for demographic factors that might be associated Zika symptoms or care-seeking behavior. For example, older age groups and women might be more prone to develop symptoms given ZIKV infection (1, 8, 9). Because women and older individuals were overrepresented in both the Puerto Rico and Yap studies, the resulting estimates of the symptomatic proportion could be inflated for the general population if the probability of symptoms does differ according to sex and age (1, 8). Other factors affecting the baseline health status of a given population could also cause variation in the proportion of symptomatic ZIKV infections. For example, the presence of a concurrent outbreak of another disease could lead to higher baseline symptom rates as described above, but also to immune interactions that could increase or decrease the likelihood of symptoms due to ZIKV infection. Study design elements, such as the diagnostic assays used and timing of serological surveys relative to the actual outbreak, also differ between studies.

Our estimates for the Zika symptomatic proportion, ranging from 15%–37% for the Yap study to 34%–92% for the Puerto Rico study, are in agreement with data from blood-donor studies. In French Polynesia, 26% of asymptomatic ZIKV reverse-transcriptase polymerase chain reaction (RT-PCR)-positive blood donors experienced ZIKV-like symptoms 3–10 days after donation, and 55% reported symptom onset in the week following their donation in Martinique (14, 15). While the blood-donor studies did not account for symptoms due to other causes, the restricted timeframe would substantially reduce the number of individuals potentially experiencing unrelated symptoms. As with the studies considered in our analysis, the role of varying case definitions and study designs must also be considered when interpreting the blood-donor data. Our estimates are also in the same range as other mosquito-borne flaviviruses such as dengue viruses (5%–68% depending on previous infections) and yellow fever virus (26%–63%) (22, 23).

Because most surveillance systems rely on the passive reporting of symptomatic disease, the fraction of symptomatic cases represents an upper bound of the proportion of infections that could be reported. For every reported symptomatic case, many more undetected infections are likely because of limited care-seeking, incomplete reporting, and asymptomatic infections. For example, in the Puerto Rico study, 55% of participants with symptomatic ZIKV infection sought care, and 18.5% of those seeking care were reported to the passive surveillance system (8). Given the substantial number of ZIKV infections invisible to medical care, researchers, and the public, Zika surveillance systems must be designed with the consideration that most infected persons will not experience disease. For example, congenital birth defects might occur because of asymptomatic infection during pregnancy (24). Likewise, screening for symptoms can reduce the number of infected blood donors but will not eliminate that risk without laboratory testing. The symptomatic proportion is also important for estimating population-level infection rates over the past 2 years. Such rates could inform assessment of the possibility for ongoing transmission and new large-scale epidemics (25, 26) as well as planning

for vaccine trials and vector control. Finally, the symptomatic infection proportion is critical for public communication; the low symptomatic proportion is precisely the reason that, in the midst of an epidemic, it can seem as if few people are actually sick.

ACKNOWLEDGMENTS

Author affiliations: Epidemic Intelligence Service, Centers for Disease Control and Prevention, Atlanta, Georgia (Patrick K. Mitchell); Pennsylvania Department of Health, Harrisburg, Pennsylvania (Patrick K. Mitchell); Division of Vector-Borne Diseases, Centers for Disease Control and Prevention, San Juan, Puerto Rico (Luis Mier-y-Teran-Romero, Matthew J. Lozier, Michael A. Johansson); Division of Vector-Borne Diseases, Centers for Disease Control and Prevention, Fort Collins, Colorado (Brad J. Biggerstaff, Mark J. Delorey); Unit of Emerging Infectious Diseases, Institut Louis Malardé, Papeete, Tahiti, French Polynesia (Maite Aubry, Van-Mai Cao-Lormeau); Mathematical Modelling of Infectious Diseases Unit, Institut Pasteur, Paris, France (Simon Cauchemez); Centre National de la Recherche Scientifique, URA 3012, Paris, France (Simon Cauchemez); Center of Bioinformatics, Biostatistics and Integrative Biology, Institut Pasteur, Paris, France (Simon Cauchemez); and Center for Communicable Disease Dynamics, Harvard T.H. Chan School of Public Health, Boston, Massachusetts (Michael A. Johansson).

S.C. acknowledges support from the European Union's Horizon 2020 Programme through ZIKAlliance (grant 734548), the Laboratory of Excellence Integrative Biology of Emerging Infectious Diseases program (grant ANR-10-LABX-62-IBEID), the Models of Infectious Disease Agent Study of the National Institute of General Medical Sciences, and the AXA Research Fund. The serosurvey conducted in French Polynesia was partially funded by the Contrat de Projet Etat-Pays (convention 7331/MSS/DSP du 31/08/12 modifiée) and the French Government's Investissement d'Avenir Programme (Labex Integrative Biology of Emerging Infectious Diseases, grant ANR-10-LABX-62-IBEID). M.A.J. received partial support from the Models of Infectious Disease Agent Study program (Cooperative Agreement IU54GM088558).

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

Conflict of interest: none declared.

REFERENCES

- Duffy MR, Chen TH, Hancock WT, et al. Zika virus outbreak on Yap Island, Federated States of Micronesia. *N Engl J Med*. 2009;360(24):2536–2543.
- Cao-Lormeau VM, Roche C, Teissier A, et al. Zika virus, French polynesia, South pacific, 2013. *Emerg Infect Dis*. 2014;20(6):1085–1086.
- Musso D, Nilles EJ, Cao-Lormeau VM. Rapid spread of emerging Zika virus in the Pacific area. *Clin Microbiol Infect*. 2014;20(10):O595–O596.
- Campos GS, Bandeira AC, Sardi SI. Zika virus outbreak, Bahia, Brazil. *Emerg Infect Dis*. 2015;21(10):1885–1886.
- Pan American Health Organization. Timeline of the emergence of Zika virus in the Americas, April 2016. Washington, DC: World Health Organization; 2016. http://www.paho.org/hq/index.php?option=com_content&view=article&id=11959&Itemid=41711&lang=en. Accessed September 19, 2018.
- Rasmussen SA, Jamieson DJ, Honein MA, et al. Zika virus and birth defects—reviewing the evidence for causality. *N Engl J Med*. 2016;374(20):1981–1987.
- Krauer F, Riesen M, Reveiz L, et al. Zika virus infection as a cause of congenital brain abnormalities and Guillain-Barré syndrome: systematic review. *PLoS Med*. 2017;14(1):e1002203.
- Lozier MJ, Burke RM, Lopez J, et al. Differences in prevalence of symptomatic Zika virus infection, by age and sex—Puerto Rico, 2016. *J Infect Dis*. 2018;217(11):1678–1689.
- Aubry M, Teissier A, Huart M, et al. Zika virus seroprevalence, French Polynesia, 2014–2015. *Emerg Infect Dis*. 2017;23(4):669–672.
- Team RC. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2016.
- Plummer M. rjags: Bayesian Graphical Models using MCMC. 2016.
- Plummer M, Best N, Cowles K, et al. CODA: convergence diagnosis and output analysis for MCMC. *R News*. 2006;6(1):7–11.
- Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. *Stat Sci*. 1992;7(4):457–472.
- Gallian P, Cabié A, Richard P, et al. Zika virus in asymptomatic blood donors in Martinique. *Blood*. 2017;129(2):263–266.
- Musso D, Nhan T, Robin E, et al. Potential for Zika virus transmission through blood transfusion demonstrated during an outbreak in French Polynesia, November 2013 to February 2014. *Euro Surveill*. 2014;19(14):pii=20761.
- Champagne C, Salthouse DG, Paul R, et al. Structure in the variability of the basic reproductive number (R0) for Zika epidemics in the Pacific islands. *ELife*. 2016;5:e19874.
- Kucharski AJ, Funk S, Eggo RM, et al. Transmission dynamics of Zika virus in island populations: a modelling analysis of the 2013–14 French Polynesia outbreak. *PLoS Negl Trop Dis*. 2016;10(5):e0004726.
- Nishiura H, Kinoshita R, Mizumoto K, et al. Transmission potential of Zika virus infection in the South Pacific. *Int J Infect Dis*. 2016;45:95–97.
- WHO Guidelines Approved by the Guidelines Review Committee. *Dengue: Guidelines for Diagnosis, Treatment, Prevention and Control. New Edition*. Geneva, Switzerland: World Health Organization; 2009.
- Oduyebo T, Polen KD, Walke HT, et al. Update: interim guidance for health care providers caring for pregnant women with possible Zika virus exposure—United States (including US Territories), July 2017. *MMWR Morb Mortal Wkly Rep*. 2017;66(29):781–793.
- Chow A, Ho H, Win MK, et al. Assessing sensitivity and specificity of surveillance case definitions for Zika virus disease. *Emerg Infect Dis*. 2017;23(4):677–679.
- Clapham HE, Cummings DAT, Johansson MA. Immune status alters the probability of apparent illness due to dengue virus

- infection: evidence from a pooled analysis across multiple cohort and cluster studies. *PLoS Negl Trop Dis*. 2017;11(9): e0005926.
23. Johansson MA, Vasconcelos PF, Staples JE. The whole iceberg: estimating the incidence of yellow fever virus infection from the number of severe cases. *Trans R Soc Trop Med Hyg*. 2014;108(8):482–487.
24. Reynolds MR, Jones AM, Petersen EE, et al. Vital signs: update on Zika virus-associated birth defects and evaluation of all US infants with congenital Zika virus exposure—US Zika Pregnancy Registry, 2016. *MMWR Morb Mortal Wkly Rep*. 2017;66(13):366–373.
25. Ferguson NM, Cucunubá ZM, Dorigatti I, et al. EPIDEMIOLOGY. Countering the Zika epidemic in Latin America. *Science*. 2016;353(6297):353–354.
26. Zhang Q, Sun K, Chinazzi M, et al. Spread of Zika virus in the Americas. *Proc Natl Acad Sci U S A*. 2017;114(22): E4334–E4343.