Web Material

Assessing Zika Virus Transmission Within Households
During an Outbreak in Martinique, 2015-2016

Anthony Cousien, Sylvie Abel, Alice Monthieux, Alessio Andronico, Isabelle Calmont, Minerva Cervantes, Raymond Césaire, Pierre Gallian, Xavier de Lamballerie, Cédric Laouénan, Fatiha Najoullah, Sandrine Pierre-François, Mathilde Pircher, Henrik Salje, Quirine A. ten Bosch, André Cabié*, and Simon Cauchemez*

*Joint last authors

Correspondence to Dr Simon Cauchemez, Mathematical Modelling of Infectious Diseases Unit, Institut Pasteur, 28 Rue du Docteur Roux, 75015 Paris, France (e-mail: simon.cauchemez@pasteur.fr)

Table of contents

Web Appendix 1: Model............................................................................................................................... 3
Web Appendix 2: Estimation of the model parameters ................................................................................. 9
Web Appendix 3: Model validation............................................................................................................... 10
Web Appendix 4: Sensitivity analysis – study period was restricted to February - November 2016......... 12
Web Appendix 5: Sensitivity analysis – distribution of household sizes in Martinique.......................... 13
Web Appendix 6: Sensitivity analysis – censorship in the follow-up of households ................................. 14
Web Appendix 7: Frequency-dependent vs. frequency independent probability of within household transmission ................................................................................................................................................ 15
  Web Figure 1........................................................................................................................................... 16
  Web Figure 2........................................................................................................................................... 17
  Web Figure 3........................................................................................................................................... 18
  Web Table 1............................................................................................................................................ 19
  Web Table 2............................................................................................................................................ 20
  Web Table 3............................................................................................................................................ 21
  Web Table 4............................................................................................................................................ 24
Web Appendix 1: Model

We developed a model for Zika virus (ZIKV) transmission within households and in the community to jointly analyze three different datasets documenting ZIKV spread in Martinique: 1- follow-up of symptoms among members of households with a confirmed ZIKV case; 2- a seroprevalence survey among blood donors in March and June 2016 (1); and 3- laboratory results in pregnant women presenting with ZIKV-related symptoms.

ZIKV transmission in household

We used a final-size model for disease transmission in households, derived from a classical chain binomial model (2,3). The model was expanded to account for the fact that households included in the study all had at least one symptomatic confirmed index ZIKV case, that symptoms may not be due to ZIKV infection and that some ZIKV infections remain asymptomatic.

Notations

We denote $N$ the size of the household, $S$ the number of persons within the household reporting Zika-like symptoms, $Z$ the number of ZIKV infections in the household (unobserved) and $A$ the number of asymptomatic ZIKV infections (unobserved).

Basic chain-binomial model

We denote $p_c$ the probability an individual is infected by ZIKV from the community during the course of the epidemic. The probability of within household transmission, i.e. transmission from one household member to another mediated by a mosquito, is denoted $p_H$ and is a function of household size $N$.

The basic chain binomial model provides the probability $P(Z|N, p_H, p_c)$ to observe $Z$ ZIKV infections in a household of size $N$, which is determined by a recursive system (2,3)
\[
\begin{aligned}
P(Z = z|N = n, p_H, p_C) &= \binom{n}{z} p(Z = z|N = z, p_H, p_C)(1 - p_C)^{n-z}(1 - p_H)^z \text{ if } z < n \\
P(Z = n|N = n, p_H, p_C) &= 1 - \sum_{i=0}^{n} P(Z = i|N = n, p_H, p_C) \text{ otherwise}
\end{aligned}
\]  \hspace{1cm} (1)

**Asymptomatic infections**

We denote \( p_A \) the proportion of ZIKV infections that are asymptomatic. The probability to obtain \( A \) asymptomatic ZIKV infections among \( Z \) ZIKV infections is a binomial draw with parameters \( Z \) and \( p_A \)

\[
P(A|Z, p_A) = \binom{Z}{A} p^A (1 - p)^{Z - A},
\]  \hspace{1cm} (2)

**Detection of the household**

All households included in the study had at least one laboratory confirmed symptomatic Zika case. This study design induces a selection bias, as the probability for a household to be recruited depends on the number of members with a symptomatic ZIKV infection. This bias can be addressed by conditioning inference on the fact that the household was recruited.

Denote \( p_D \) the probability that a symptomatic Zika case is detected and invited to participate in the study.

The probability for a household with \( Z - A \) symptomatic Zika cases to be recruited is

\[
P(Recruited|Z, A, p_D) = (1 - (1 - p_D)^{Z - A}) \approx p_D(Z - A)
\]  \hspace{1cm} (3)

This last approximation holds as \( p_D \) is expected to be low (68 index cases were invited to participate in the study which compares to an estimate of around 190,000 ZIKV infections in the island (4)).

**Symptoms not related to ZIKV**

Denote \( p_{NZ} \) the probability to be symptomatic for a cause different than ZIKV infection. Given the total number of ZIKV infections \( (Z) \) and the number of asymptomatic ZIKV infections \( (A) \), the probability to observe a total of \( S \) household members presenting with ZIKV-like symptoms is the probability that \( S -
(Z – A) presented with symptoms not related to ZIKV, among the N – (Z – A) members who do not present with symptoms related to ZIKV. This probability is

\[ P(S|Z, A, N, p_{NZ}) = \binom{N - (Z - A)}{S - (Z - A)} p_{NZ}^S (1 - p_{NZ})^{N - S}, \quad (4) \]

*Household contribution to the likelihood*

Let \( \theta = (p_C, p_H, p_A, p_D, p_{NZ}) \) be the set of parameters introduced above.

For each household, the contribution to the likelihood is

\[ P(N, S|Recruited, \theta) = \frac{P(S, Recruited|N, \theta) \pi(N)}{P(Recruited)} \]

with

- \( \pi \) the distribution of household sizes in Martinique (5). As the distribution \( \pi \) is only given in a truncated form (households with 6 members or more, representing 2% of the total number of households in Martinique, were grouped together), we split this bin uniformly in sizes of 6, 7 and 8. We performed a sensitivity analysis to assess the impact of this assumption on our results (Web Appendix 5).

- \( P(S, Recruited|N, \theta) = \sum_{Z=0}^{N} \sum_{A=0}^{Z} P(S, Z, A, Recruited|N, \theta) \)

- \( P(Recruited|\theta) = \sum_{N=1}^{N_{max}} \sum_{S=0}^{N} \sum_{Z=0}^{S} \sum_{A=0}^{Z} P(S, Z, A, Recruited|N, \theta) \pi(N) \),

with \( N_{max} \) the maximum household size, and \( P(S, Z, A, Recruited|N, \theta) \) that can be expressed according to the previously introduced probabilities:

\[ P(S, Z, A, Recruited|N, \theta) = P(Recruited|Z, A, p_D) \times P(S|Z, A, N, p_{NZ}) \times P(A|Z, p_A) \times P(Z|N, p_H, p_C), \quad (6) \]

Due to the approximation made in (3), \( p_D \) appears in both the numerator and denominator and therefore cancels out.
Seroprevalence survey among blood donors

Gallian et al performed seroprevalence analyses in blood donors at two time points: in 418 donors sampled in early March (9th-23rd), and in 176 donors sampled in early June (6th-13th) (1). Samples were tested using an anti-Zika virus NS1 IgG ELISA kit (anti-Zika Virus ELISA IgG, Euroimmun, Germany) and positives were further processed by microneutralisation in a 96-well format, using Vero cells, 50 pfu of the MRS.OPY.Martinique.PaRi.2015 strain14 and a threshold titer of 40 as recommended by the French National Reference Centre for Arboviruses. The seroprevalence after seroneutralisation was $s_1=13.5\%$ in early March and $s_2=42.2\%$ in early June.

Since it takes about two weeks for an infected individual to seroconvert, we assumed that seroprevalence at the first time point measured the cumulative infection attack rate on $W_1=$week 9 of 2016 (29 Feb – 6 Mar) while the second time point gave cumulative infection attack rate on $W_2=$week 21 of 2016 (23-29 May).

Denote $p_i$ the proportion of the population infected by ZIKV by the end of the epidemic and $q_i$ the proportion of consultations for ZIKV-related symptoms that occurred before or on week $W_i$ ($q_1=29.9\%$ and $q_2=76.1\%$ according to the number of Zika-like cases reported in Martinique in a sentinel network of general practitioners (6)). Under the assumption that the (unobserved) weekly number of ZIKV infections was proportional to the (observed) weekly number of consultations for ZIKV-related symptoms, the proportion of the population infected by week $W_i$ is expected to be equal to $q_ip_i$.

Denoting $M_i$ and $P_i$ the number of blood donors tested and the number of seropositive blood donors in seroprevalence study $i$, the likelihood for the serological data is

$$P(P|M, p_i) = \prod_{i=1,2} \left( \frac{M_i}{P_i} \right) (q_ip_i)^{P_i}(1-q_ip_i)^{M_i-P_i}$$

(7)

In practice, we expressed $p_i$ according to our chain binomial model as
\[ p_I = \sum_{N=1}^{\max(N)} \left[ \pi(N) \times \sum_{Z=0}^{N} ZP(Z|N) \right] \]  

(8)

We checked that the seroprevalence studies of March and June led to similar estimates of the overall seroprevalence \( p_I \) when considered independently. Applying our approach independently to each time point, the final attack rates of the ZIKV epidemic in Martinique was estimated at 45% (95% CI: 35%, 56%) based on the first seroprevalence study and at 55% (95% CI: 46%, 65%) based on the second seroprevalence study. There was no statistical difference between the two estimates (\( p=0.17 \)). We therefore combined the two studies to improve accuracy.

**Proportion of confirmed cases among pregnant women with a suspicion of ZIKV infection**

We also included the lab results for pregnant women presenting with ZIKV-related symptoms in Martinique between February and November 2016 in the likelihood. Let \( W_S \) be the number of women with symptoms and \( W_Z \) the number of women with a confirmed ZIKV infection among them. The contribution to the likelihood of this study is

\[ P(W_Z|W_S, p_{Z|S}) = \binom{W_S}{W_Z} p_{Z|S}^{W_Z} (1 - p_{Z|S})^{W_S - W_Z}, \]  

(9)

where \( p_{Z|S} \) is the probability of being infected by ZIKV given presence of symptoms. This can be expressed as a function of the model parameters \( p_I \) (which depends on \( p_C \) and \( p_H \)), \( p_A \) and \( p_{NZ} \)

\[ p_{Z|S} = \frac{p_I (1 - p_A)}{p_I (1 - p_A) + p_{NZ} - p_I (1 - p_A) p_{NZ}}. \]  

(10)

**Final likelihood**

The final likelihood of the model is

\[ L(\mathcal{F}|p_H, p_C, p_A) = \prod_{\mathcal{F} \in \mathcal{F}} P(N_F,S_F|Recruited) \times P(P|M,p_I) \times P(W_Z|W_S, p_{Z|S}) \]  

(11)
with $\mathcal{F}$ the set of households, $N_F$ the size of the household $F$ and $S_F$ the number of persons with symptoms.
Web Appendix 2: Estimation of the model parameters

We estimated the parameters \( \theta = (p_C, p_H, p_A, p_{NZ}) \) using Markov Chain Monte Carlo (MCMC) sampling (7) with uniform priors on [0,1]. We used log-normal proposal, which were tuned to have an optimal acceptance rate of 24% (8). We used a chain of 100,000 iterations thinned by a factor 10 and removed the first 1,000 iterations as burn-in. Posterior distributions are presented in Web Figure 1, meanwhile Web Figure 2 present the trace plots and show a good chain mixing.
Web Appendix 3: Model validation

To simulate the 500 datasets used to evaluate the model and the inferential approach, we set the value of $p_D$ at 0.0006 to obtain a mean sample size of approximatively 68.

Expected and observed distributions of households by size and case count

The expected and observed distributions of households by size and case count are given in Web Table 5. Expected and observed distributions are close to each other, and the observed value always fall in the 95% confidence interval except for households of size 4 with 1 member reporting Zika-like symptoms and for households of size 4 with 3 members reporting Zika-like symptoms. These small discrepancies can be imputed to the small sample size.

Evaluation of the statistical approach

When we estimated model parameters for the 500 simulated datasets, we obtained the posterior distributions given in Web Figure 3. The parameters estimated were consistent with those used to simulate the samples, and the input value fell in 96.6% of the simulations within the credibility interval (Web Table 6).
Web Appendix 4: Sensitivity analysis – study period was restricted to February - November 2016

In the household transmission study, recruitment of households occurred between December 2015 and October 2016. However, the study documenting ZIKV infections in pregnant women started later, in February 2016. In a sensitivity analysis (SA1), we assessed how our estimates were modified if the analysis was restricted to this shorter time period. Results are presented in Web Table 7.
Web Appendix 5: Sensitivity analysis – distribution of household sizes in Martinique

In existing demographic data (5), households with 6 members or more, that represent 2% of the total number of households in Martinique, are grouped together. In our analyses, we have distributed uniformly these households in sizes of 6, 7 and 8 (see Web Table 1). To assess the impact of this assumption on our results, we also considered a scenario where we modeled it with the tail of a geometric distribution (sensitivity analysis SA2), see Web Table 2.

Results were similar to those obtained in the main analysis (see Web Table 8).
Web Appendix 6: Sensitivity analysis – censorship in the follow-up of households

The follow-up of 11% of households stopped before the end of the epidemic. To assess the impact of this incomplete follow-up on the results, we performed a sensitivity analysis by expressing for each household \( h \) a specific probability of infection during the epidemic \( p_c(h) = p_c \cdot F(h) \), with \( p_c \) the final probability to have been infected by ZIKV from the community with a complete follow-up (as before), and \( F(h) \) the proportion of cases in the epidemic curve with onset before the end of the follow-up of household \( h \). We obtained very similar results to the main analysis (see Web Table 9).
Web Appendix 7: Frequency-dependent vs. frequency independent probability of within household transmission

In our main analysis, we assumed that $p_H$ is independent of household size (frequency-independent probability of within household transmission). However, frequency-dependent relationship, i.e. a transmission probability depending on the household size, has been found in a number of transmission studies looking at respiratory and vector-borne diseases (9–11). In the sensitivity analysis SA3, we used a probability of within household transmission of the form $p_H = 1 - \exp\left(-\frac{\beta}{N}\right)$, with $N$ the household size and a uniform prior for $\beta$ between 0 and 10. Results are presented in Web Table 4. The proportion of infections occurring in households changed from 22% [5%-46%] in the main analysis to 30% [8%-55%] under the assumption of frequency-dependence. However, the DIC was slightly higher than in the main analysis (452 vs. 451 in the main analysis), and the 95% credible intervals were very wide for $p_H$. These results might be explained by the relatively small size of our households sample (68 households). Thus, we used the frequency-independent model in our main analysis.
Web Figure 1 Posterior distributions for the model parameters. A. probability of asymptomatic infection $p_A$; B. probability of infection from the community $P_C$; C. probability of within household transmission; D. probability to present symptoms not related to Zika $p_{NZ}$. 
Web Figure 2 Trace plots for the MCMC algorithm. The dotted line represents the end of the burn-in phase.

A. probability of asymptomatic infection $p_A$; B. probability of infection from the community $P_C$; C. probability of within household transmission; D. probability to present symptoms not related to Zika $p_{NZ}$. 
Web Figure 3 Distribution of the posterior mean of parameters derived from 500 simulated datasets. The input value of the parameter is indicated with the red line. A. probability of asymptomatic infection $p_A$; B. probability of infection from the community $P_C$; C. probability of within household transmission $p_H$; D. probability to present symptoms not related to ZIKV $p_{NZ}$.
**Web Table 1** Distribution of the households per size in the main analysis.

| Household size | Number of households | %    |
|----------------|----------------------|------|
| 1              | 55,931               | 33.9%|
| 2              | 48,939               | 29.7%|
| 3              | 29,556               | 18.0%|
| 4              | 19,244               | 11.7%|
| 5              | 7,279                | 4.4% |
| 6              | 1,218                | 0.74%|
| 7              | 1,218                | 0.74%|
| 8              | 1,218                | 0.74%|
Web Table 2 Distribution of the households per size in the sensitivity analysis.

| Household size | Number of households | %    |
|----------------|----------------------|------|
| 1              | 55931                | 33.9%|
| 2              | 48939                | 29.7%|
| 3              | 29556                | 18.0%|
| 4              | 19244                | 11.7%|
| 5              | 7279                 | 4.4% |
| 6              | 1724                 | 1.5% |
| 7              | 911                  | 0.055%|
| 8              | 481                  | 0.029%|
| 9              | 254                  | 0.015%|
| 10             | 134                  | 0.0081%|
| 11             | 71                   | 0.0043%|
| 12             | 37                   | 0.0025%|
| 13             | 20                   | 0.0012%|
| 14             | 10                   | 0.00061%|
| 15             | 6                    | 0.00036%|
| 16             | 3                    | 0.00018%|
**Web Table 3** Raw data used for the analyses presented in the article. The data used for the analysis are presented below. Each line represents a household in the sample. The mean waiting time between the onset of symptoms in the index case and the onset of symptoms among other household members is 5 days (95% confidence interval (CI): -109-73).

| Number of household members reporting symptoms (including the index case) | Household members not reporting symptoms | Household included after February 2016 (beginning of the study among pregnant women) | Household with complete follow-up (inclusion, day 10, day 21, week 12 or after) |
|---|---|---|---|
| 3 | 0 | 0 | 1 |
| 1 | 0 | 0 | 1 |
| 1 | 0 | 0 | 1 |
| 3 | 0 | 0 | 1 |
| 3 | 2 | 0 | 1 |
| 4 | 0 | 0 | 1 |
| 1 | 0 | 0 | 1 |
| 3 | 0 | 0 | 1 |
| 5 | 0 | 1 | 1 |
| 3 | 3 | 1 | 1 |
| 4 | 0 | 1 | 1 |
| 2 | 0 | 1 | 1 |
| 2 | 2 | 1 | 1 |
| 1 | 3 | 1 | 1 |
| 2 | 1 | 1 | 1 |
| 2 | 0 | 1 | 1 |
| 1 | 0 | 1 | 1 |
| 3 | 0 | 1 | 1 |
| 1 | 1 | 1 | 0 |
| 1 | 0 | 1 | 1 |
| 2 | 0 | 1 | 0 |
| 1 | 3 | 1 | 1 |
| 1 | 1 | 1 | 0 |
| 2 | 1 | 1 | 1 |
| 3 | 0 | 1 | 0 |
| 3 | 3 | 1 | 1 |
| 5 | 2 | 1 | 1 |
| 4 | 1 | 1 | 0 |
| 1 | 3 | 1 | 1 |
| 6 | 0 | 1 | 1 |
| 1 | 1 | 1 | 1 |
| 2 | 0 | 1 | 0 |
| 1 | 0 | 1 | 1 |
| 4 | 1 | 1 | 1 |
| 3 | 2 | 1 | 0 |
| 1 | 1 | 1 | 1 |
| 1 | 2 | 0 | 1 |
| 2 | 1 | 0 | 1 |
| 3 | 0 | 0 | 1 |
| 1 | 2 | 0 | 1 |
| 1 | 0 | 1 | 1 |
| 2 | 1 | 0 | 1 |
| 2 | 1 | 0 | 1 |
| 4 | 0 | 0 | 1 |
| 1 | 3 | 0 | 1 |
| 2 | 1 | 0 | 1 |
| 2 | 2 | 1 | 1 |
| 4 | 2 | 1 | 1 |
| 1 | 0 | 1 | 1 |
| 1 | 1 | 1 | 1 |
| 1 | 2 | 1 | 1 |
| 3 | 0 | 1 | 1 |
| 2 | 0 | 1 | 1 |
| 2 | 0 | 1 | 1 |
| 1 | 6 | 1 | 1 |
| 1 | 1 | 1 | 1 |
| 1 | 4 | 1 | 1 |
| 8 | 0 | 1 | 1 |
| 1 | 3 | 1 | 1 |
| 1 | 3 | 1 | 1 |
| 1 | 2 | 3 | 1 | 1 |
|---|---|---|---|---|
| 2 | 3 | 1 | 1 | 1 |
| 1 | 2 | 1 | 1 | 1 |
| 3 | 4 | 1 | 1 | 1 |
| 1 | 2 | 1 | 1 | 1 |
| 2 | 0 | 1 | 1 | 1 |
| 1 | 3 | 1 | 1 | 1 |
| 1 | 2 | 1 | 1 | 1 |
Web Table 4 Posterior mean and 95% CrI of model parameters in the main analysis and sensitivity analysis SA3, where we used a frequency-dependent model where $p_H = 1 - \exp(-\frac{\beta}{N})$.

| Parameter                                                                 | Main analysis frequency-independent $p_H$ | Frequency-dependent $p_H = 1 - \exp(-\frac{\beta}{N})$ with $N$ the household size |
|---------------------------------------------------------------------------|--------------------------------------------|--------------------------------------------------------------------------------------|
| Proportion of asymptomatic infections $p_A$                              | 40% [23%-56%]                              | 43% [25%-59%]                                                                         |
| Probability of infection from the community $p_c$                        | 39% [27%-50%]                              | 35% [23%-48%]                                                                         |
| 2                                                                         | 48% [13%-97%]                              |                                                                                       |
| 3                                                                         | 36% [9%-90%]                               |                                                                                       |
| Probability of within household transmission (according to the household size) $p_H$ | 21% [5%-51%]                              | 25% [6%-74%]                                                                          |
| 4                                                                         | 48% [13%-97%]                              |                                                                                       |
| 5                                                                         | 36% [9%-90%]                               |                                                                                       |
| 6                                                                         | 21% [5%-51%]                               | 25% [6%-74%]                                                                          |
| 7                                                                         | 19% [4%-62%]                               | 22% [5%-68%]                                                                          |
| 8                                                                         | 17% [3%-57%]                               | 19% [4%-62%]                                                                          |
| Probability of presenting symptoms due to another cause $p_{NZ}$         | 16% [10%-23%]                              | 15% [9%-22%]                                                                          |
| Proportion of infections occurring at household level                     | 22% [5%-46%]                               | 30% [8%-55%]                                                                          |
| Attack rate                                                               | 50% [43%-58%]                              | 50% [43%-58%]                                                                         |
| DIC                                                                       | 451                                        | 452                                                                                   |

DIC=Deviance information criterion
**Web Table 5** Expected (mean [95% confidence interval]) and observed distribution of households by size and number of members reporting symptoms (including the index case) obtained from 500 simulated datasets with a structure similar to the observed one.

| Number of reported symptoms | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   |
|-----------------------------|-----|-----|-----|-----|-----|-----|-----|-----|
| **Observed**                | 8   |     |     |     |     |     |     |     |
| **Expected**                | 8 [3-14] |     |     |     |     |     |     |     |
| **Observed**                | 8   |     |     |     |     |     |     |     |
| **Expected**                | 6 [3-14] | 7 [2-12] |     |     |     |     |     |     |
| **Observed**                | 7   |     |     |     |     |     |     |     |
| **Expected**                | 5 [1-9] | 8 [3-14] | 3 [0-7] |     |     |     |     |     |
| **Observed**                | 6   | 6   | 7   |     |     |     |     |     |
| **Expected**                | 2 [0-5] | 6 [2-11] | 5 [1-9] | 2 [0-4] |     |     |     |     |
| **Observed**                | 2   | 2   | 2   | 1   |     |     |     |     |
| **Expected**                | 1 [0-2] | 2 [0-5] | 3 [0-6] | 2 [0-5] | 0 [0-2] |     |     |     |
| **Observed**                | 1   | 0   | 2   | 1   | 0   | 1   |     |     |
| **Expected**                | 0 [0-1] | 0 [0-2] | 0 [0-2] | 1 [0-2] | 0 [0-2] | 0 [0-1] |     |     |
| **Observed**                | 1   | 0   | 1   | 0   | 1   | 0   | 0   | 0   |
| **Expected**                | 0 [0-1] | 0 [0-1] | 0 [0-2] | 1 [0-3] | 0 [0-2] | 0 [0-1] | 0 [0-1] |     |
| **Observed**                | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 1   |
| **Expected**                | 0 [0-0] | 0 [0-1] | 0 [0-2] | 0 [0-2] | 1 [0-3] | 0 [0-2] | 0 [0-1] | 0 [0-1] |
**Web Table 6** Input parameters and estimates obtained when our statistical approach is applied to 500 simulated datasets with known parameter values.

| Parameter value used in simulation | Mean estimate with 2.5% and 97.5% quantiles | Proportion of datasets for which true value is in 95% credible interval |
|-----------------------------------|---------------------------------------------|---------------------------------------------------------------------|
| Proportion of asymptomatic infections $p_A$ | 40% | 44% [31%-57%] | 97.4% |
| Probability of infection from the community $p_C$ | 39% | 39% [29%-47%] | 99.8% |
| Probability of within household transmission $p_H$ | 21% | 27% [7%-60%] | 99.0% |
| Probability of presenting symptoms due to another cause $p_{NZ}$ | 16% | 16% [11%-21%] | 98.6% |
| All the parameters | | | 96.6% |
**Web Table 7** Posterior mean and 95% CrI of model parameters in the main analysis and sensitivity analysis SA1, where the study period was restricted to February - November 2016.

| Parameter                                                                 | Main analysis Tail of the distribution of households = uniform distribution | SA3 Tail of the distribution of households = geometric distribution |
|----------------------------------------------------------------------------|--------------------------------------------------------------------------------|---------------------------------------------------------------------|
| Proportion of asymptomatic infections \( p_A \)                          | 40% [23%-56%]                                                                | 46% [29%-62%]                                                      |
| Probability of infection from the community \( p_C \)                    | 39% [27%-50%]                                                                | 37% [24%-49%]                                                      |
| Probability of within household transmission \( p_H \)                   | 21% [5%-51%]                                                                 | 25% [7%-72%]                                                      |
| Probability of presenting symptoms due to another cause \( p_NZ \)      | 16% [10%-23%]                                                                | 14% [9%-20%]                                                      |
| Proportion of infections occurring at household level                    | 22% [5%-46%]                                                                 | 26% [7%-52%]                                                      |
| Attack rate                                                              | 50% [43%-58%]                                                                | 50% [43%-58%]                                                      |
**Web Table 8** Posterior mean and 95% CrI of model parameters in the main analysis and sensitivity analysis SA2, where we use a geometric distribution to reconstruct the tail of the household size distribution.

| Parameter                               | Main analysis | SA3          |
|-----------------------------------------|---------------|--------------|
| **Tail of the distribution of households** | uniform distribution | geometric distribution |
| Proportion of asymptomatic infections $p_A$ | 40% [23%-56%] | 41% [23%-56%] |
| Probability of infection from the community $p_C$ | 39% [27%-50%] | 39% [26%-50%] |
| Probability of within household transmission $p_H$ | 21% [5%-51%] | 22% [5%-58%] |
| Probability of presenting symptoms due to another cause $p_{NZ}$ | 16% [10%-23%] | 16% [10%-23%] |
| Proportion of infections occurring at household level | 22% [5%-46%] | 22% [5%-48%] |
| Attack rate                             | 50% [43%-58%] | 50% [43%-58%] |
| DIC                                     | 451           | 453          |

DIC=Deviance information criterion
**Web Table 9** Posterior mean and 95% CrI of model parameters in the main analysis and in sensitivity analysis SA3, where we accounted for the incomplete follow-up of some households.

| Parameter                                                                 | Main analysis assuming complete follow-up | SA2 $p_C(h) = p_C \cdot F(h)$ with $F(h)$ the follow-up duration |
|----------------------------------------------------------------------------|-------------------------------------------|---------------------------------------------------------------|
| Proportion of asymptomatic infections $p_A$                               | 40% [23%-56%]                             | 41% [23%-57%]                                               |
| Probability of infection from the community $p_C$                         | 39% [27%-50%]                             | 38% [26%-50%]                                               |
| Probability of within household transmission $p_H$                       | 21% [5%-51%]                              | 23% [6%-55%]                                               |
| Probability of presenting symptoms due to another cause $p_{NZ}$         | 16% [10%-23%]                             | 16% [10%-22%]                                               |
| Proportion of infections occurring at household level                     | 22% [5%-46%]                              | 24% [6%-48%]                                               |
| Attack rate                                                               | 50% [43%-58%]                             | 50% [43%-58%]                                               |
| DIC                                                                       | 451                                       | 452                                                          |

DIC=Deviance information criterion
References

1. Gallian P, Cabié A, Richard P, Paturel L, Charrel RN, Pastorino B, et al. Zika virus in asymptomatic blood donors in Martinique. Blood. 2017 Jan 12;129(2):263–6.

2. Ludwig D. Final size distribution for epidemics. Math Biosci. 1975 Feb 1;23(1):33–46.

3. Longini IM, Koopman JS. Household and Community Transmission Parameters from Final Distributions of Infections in Households. Biometrics. 1982;38(1):115–26.

4. Andronico A, Dorléans F, Fergé J-L, Salje H, Ghawché F, Signate A, et al. Real-Time Assessment of Health-Care Requirements During the Zika Virus Epidemic in Martinique. Am J Epidemiol. :1–10.

5. Institut national de la statistique et des études économiques. MEN1 - Ménages par taille du ménage et catégorie socioprofessionnelle de la personne de référence en 2013 [Internet]. 2013 [cited 2017 Jun 21]. Available from: https://www.insee.fr/fr/statistiques/2020330?geo=DEP-972&sommaire=2106091

6. Institut National de Veille Sanitaire. Situation épidémiologique du virus Zika aux Antilles. Point au 22 décembre 2016 [Internet]. 2016 [cited 2017 Jun 21]. Available from: http://invs.santepubliquefrance.fr/Publications-et-outils/Points-epidemiologiques/Tous-les-numeros/Antilles/2016/Situation-epidemiologique-du-virus-Zika-aux-Antilles.-Point-au-22-decembre-2016

7. Gilks WR, Richardson S, Spiegelhalter D. Markov Chain Monte Carlo in Practice. CRC Press; 1995. 522 p.

8. Rosenthal, Jeffrey S. Optimal Proposal Distributions and Adaptive MCMC. In: Handbook of Markov Chain Monte Carlo. CRC Press; 2011.

9. Cauchemez S, Carrat F, Viboud C, Valleron AJ, Boëlle PY. A Bayesian MCMC approach to study transmission of influenza: application to household longitudinal data. Stat Med. 2004 Nov 30;23(22):3469–87.

10. Fraser C, Cummings DAT, Klinkenberg D, Burke DS, Ferguson NM. Influenza transmission in households during the 1918 pandemic. Am J Epidemiol. 2011 Sep 1;174(5):505–14.

11. Kuan G, Ramirez S, Gresh L, Ojeda S, Melendez M, Sanchez N, et al. Seroprevalence of Anti-Chikungunya Virus Antibodies in Children and Adults in Managua, Nicaragua, After the First Chikungunya Epidemic, 2014-2015. PLoS Negl Trop Dis. 2016 Jun 20;10(6).