Supplementary material

This supplementary material is hosted by *Eurosurveillance* as supporting information alongside the article "Dengue serosurvey after a 2-month long outbreak in Nîmes, France, 2015: was there more than met the eye?" on behalf of the authors who remain responsible for the accuracy and appropriateness of the content. The same standards for ethics, copyright, attributions and permissions as for the article apply. *Eurosurveillance* is not responsible for the maintenance of any links or email addresses provided therein."

Spatial clustering of cases

We quantified spatial clustering at various distance ranges in the community (same household (0), >0-50, >50-100 and >100-200 meters) as the odds ratio (OR) of two cases living within a distance range relative to two non-infected individuals living within that distance range(1). The odds of two cases living within a specific distance range was calculated as the number of case pairs living within range/ number of case pairs living further apart. Likewise, the odds of two non-infected individuals living within that range was calculated as the number of non-case pairs living within range/ number of non-case pairs living further apart.

Statistical transmission model

We used a statistical model to infer the most likely transmission network, estimate the hazard of infection in the community and investigate spatial patterns in the infection hazard. Such models have been previously described in (2, 3).

The force of infection experienced by individual i at time t is

\[
\lambda_i(t) = \sum_{j:i \neq t} \lambda_{j-i}(t|x_j, x_i)
\]

where the hazard that individual j infects individual i with disease onset on day d (d>d_j) at time t is

\[
\lambda_{j-i}(t|x_j, x_i) = \beta(x_i, x_j) \cdot f(t - t_i)
\]

where \(\beta(x_i, x_j)\) is the transmission rate between individual j and i, which in the simplest case is just the overall community transmission rate \(\beta\). \(f(t - t_i)\) is the probabilistic density of the serial interval, which is the distribution of the time between disease onset in a case and disease onset in generated cases.
We investigated two types of spatial dependency in infection hazard. First, we allowed for a change in the force of infection \( \lambda_i(t) \) once a case occurred in the same household e.g., due to shared exposure to an infected mosquito already present in the house (co-primary infections). Second, we fitted a spatial distance kernel for changes in \( \beta(x_i, x_j) \) based on a power law function and the geographic distance between individuals i and j.

For the within household increase, we allowed \( \lambda_i(t) \) to change by \( \Delta \lambda \) once an individual developed disease in the same household

\[
\lambda_i(t) = \sum_{j:t_j \leq t} \lambda_j \cdot \delta(t_j | x_j, x_i) + \Delta \lambda
\]

The increase in \( \lambda_i(t) \) remained for a duration of \( d \lambda \) days since a case in the household occurred. We fixed \( d \lambda \) to 1, 2, 3 or 5 days and compared the model fit.

To investigate changes in the transmission rate by distance, we allowed \( \beta(x_i, x_j) \) to vary by distance between individuals i and j using a power-law distance kernel \( g(x_i, x_j) \) with kernel parameter \( \alpha \), the distance \( d_{ij} \) between individuals i and j in meters, and a maximum distance of 1000 meters

\[
g(x_i, x_j) = \frac{1}{(1 + \frac{d_{ij}}{1000})^\alpha} \cdot A
\]

where \( A \) is a correction factor that ensures that the kernel sums up to 1 over the maximum distance.

\[
A = \frac{1 - \alpha}{1000 \cdot (2^{1-\alpha}) - 1}
\]

\( \beta(x_i, x_j) \) is then obtained by

\[
\beta(x_i, x_j) = \beta \cdot g(x_i, x_j)
\]

The effective reproduction number for individual j was calculated as

\[
R_j = \sum_{i \neq j} \beta(x_i, x_j)
\]

The probability of a household member being a co-primary case was calculated as

\[
P_{\text{co-primary}} = 1 - \exp(-\Delta \lambda \cdot d_{ij})
\]

We estimated model parameters based on a Markov chain Monte Carlo (MCMC) Bayesian framework. Parameters were updated using a Metropolis-Hastings algorithm that runs for 60,000 iterations with a burn-in of 10,000 iterations and every 10\textsuperscript{th} sampled value is stored. Convergence was visually assessed. We provided parameter estimates as the posterior median and a 95% credible interval. Model comparison was done based on the Deviance Information Criterion (DIC).
Deriving the serial interval distribution

We estimated the serial interval distribution from the generation time distribution, which is comparable if onset of symptoms and onset of infectiousness in cases occurs simultaneously. The serial interval is defined as time between symptom onset in a case and symptom onset in generated cases and the generation time as time between infection of the case and infection of the generated case. To estimate the generation time distribution, we used information on different infection stages (incubation period, human to mosquito transmission, and mosquito infectiousness) as described in detail in (3). In brief, we used the following distributions to obtain an overall generation time distribution:

Human incubation period- We used a truncated exponential distribution with a mean incubation period duration of 3 days, a minimum duration of 1 day, and a maximum duration of 1 week.

Human to mosquito transmission- We assumed that symptoms appear simultaneously when individuals become viraemic. We used an exponential distribution with mean viraemia duration of 3 days and a maximum duration of 1 week.

Mosquito infectiousness- To estimate the duration of mosquito infectiousness we drew a mosquito lifespan from a truncated exponential distribution (mean 7.2 days, max 30 days). Subsequently we drew an age by which that mosquito was infected (random draw between 0 and the mosquito lifespan). We then drew the extrinsic incubation period from an exponential distribution with a mean of 2 days. The total mosquito infectiousness period was the mosquito lifespan minus the age of mosquito infection minus the extrinsic incubation period, where values <0 were considered as unsuccessful infections.

We then estimated the empirical serial interval distribution by simulating values for each transmission stage as explained above and weighted the probability of each generation time by the length of viraemia multiplied by the length of mosquito infectiousness to take into account that humans infectious for longer will infect more mosquitoes and mosquitoes infectious for longer will infect more individuals. The derived generation time distribution had a mean of 16 days and a variance of 40 days(4).

Estimated parameters from the transmission model

In the final model, we estimated the transmission rate $\beta$ and the increase in the household force of infection once a case was observed $\Delta_\lambda$, assuming a serial interval distribution with a mean of 16 days and a variance of 40 days and a duration of force of infection increase of $d=2$ (Model 1). The estimated parameters are provided in the Supplementary Table 1. When additionally including a distance kernel (Model 2), we did not find evidence for a decline in the transmission rate by geographic distance from a case. The estimated kernel parameter $\alpha$ was 6.5 (95%CI -6.2; 19.6), where $\alpha=0$ indicates an equal transmission rate at all distances and $\alpha<0$ an increase in the transmission rate by distance.

<table>
<thead>
<tr>
<th>Estimated parameters</th>
<th>Model 1</th>
<th>Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$ (95% CI)</td>
<td>$1.3 \times 10^{-3}$ (0.5 $\times 10^{-3}$; 2.4 $\times 10^{-3}$)</td>
<td>$1.0 \times 10^{-3}$ (0.4 $\times 10^{-3}$; 5.8 $\times 10^{-3}$)</td>
</tr>
</tbody>
</table>
Supplementary Table 1. Estimated transmission model parameters assuming a serial interval distribution with a mean of 16 days and a variance of 40 days and the duration of force of infection increase of $\Delta \lambda = 2$, dengue serosurvey, Nîmes, France, 2015. Model 1 estimated the transmission rate $\beta$ and the increase in the household force of infection once a case was observed $\Delta \lambda$; Model 2 additionally included a distance kernel with parameter $\alpha$.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value (95% CI)</th>
<th>Value (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta \lambda$ (95% CI)</td>
<td>0.34 (8; 93)</td>
<td>0.33 (7; 93)</td>
</tr>
<tr>
<td>$\alpha$ (95% CI)</td>
<td>Not estimated</td>
<td>6.5 (-6.2; 19.6)</td>
</tr>
<tr>
<td><strong>Calculated measures</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R$ (95% CI)</td>
<td>0.8 (0.3; 1.4)</td>
<td>0.6 (0.2; 3.4)</td>
</tr>
<tr>
<td>$P_{co-primary}$ (%) (95% CI)</td>
<td>49 (15; 84)</td>
<td>48 (14; 84)</td>
</tr>
<tr>
<td>DIC</td>
<td>167</td>
<td>168</td>
</tr>
</tbody>
</table>

Sensitivity analysis – Serial interval distribution

To explore whether results were robust to the assumed serial interval distribution, we reconstructed the transmission trees based on two alternative serial interval distributions: i) Weibull distributed with a mean of 14 days and variance of 40 days (Model 3), and ii) Weibull distributed with a mean of 17 days and variance of 37 days (Model 4). The transmission trees are shown in Supplementary Figure 1; estimated parameters are provided in the Supplementary Table 2.

Supplementary Figure 1. Comparing reconstructed transmission trees based on different serial interval distributions (A) mean 16 days and variance 40 days (Model 1), (B) mean 14 days and variance 40 days (Model 3), and (C) mean 17 days and variance 37 days (Model 4). Dengue serosurvey, Nîmes, France, 2015.
Sensitivity analysis – Duration of household increase in the force of infection

We explored how different durations of $d_\lambda$ affect the estimated change in the force of infection $\Delta_\lambda$. The results of the sensitivity analysis are summarized in Supplementary Table 3. The models with $d_\lambda=1$ and $d_\lambda=2$ predicted best the observed proportion of co-primary cases among household members of cases of 40%.

Supplementary Table 3. Comparing models with different assumed durations of $d_\lambda$, dengue serosurvey, Nîmes, France, 2015

<table>
<thead>
<tr>
<th>$d_\lambda$ (days)</th>
<th>$\Delta_\lambda$ (95% CI)</th>
<th>$P_{co-primary}$ (95%CI)</th>
<th>DIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.37 (0.06; 1.27)</td>
<td>31 (6; 72)</td>
<td>181</td>
</tr>
<tr>
<td>2</td>
<td>0.34 (0.08; 0.93)</td>
<td>49 (15; 84)</td>
<td>167</td>
</tr>
<tr>
<td>3</td>
<td>0.25 (0.05; 0.68)</td>
<td>53 (14; 87)</td>
<td>169</td>
</tr>
<tr>
<td>5</td>
<td>0.16 (0.04; 0.43)</td>
<td>55 (18; 88)</td>
<td>171</td>
</tr>
</tbody>
</table>

Sensitivity analysis- Case underreporting and survey participation

We further explored how the estimated effective reproduction number $R$ may be affected by potential underreporting of cases and survey participation. We investigated two scenarios: i) Scenario 1 where 50% of all individuals (cases and non-cases) participated in the surveys (non-participating cases may have been missed) and ii) Scenario 2 where all cases were identified however only 50% of non-cases participated in the surveys. Therefore, we simulated DENV transmission using different parameter sets.

We first simulated DENV transmission among 1,084 individuals (twice the individuals included in the analysis dataset) with $\beta=1.1$ and a spatial kernel with $\alpha=40$ (median transmission distance of 20 meters) resulting in 24 cases. We subsequently i) randomly selected 50% of all individuals (survey scenario 1) or ii) selected all cases and only 50% of non-infected individuals (survey scenario 2). We then compared the reproduction number $R$ of the index case calculated based on the true parameters (as used for the simulations) and...
based on the estimated parameters. Based on true parameters, the reproduction number $R$ of the index case was 0.5, in survey scenario 1 we estimated $R=0.6$ (95%CI 0.3; 1.1) and in survey scenario 2 $R=0.6$ (95%CI 0.4; 0.9). The kernel parameter $\alpha$ was slightly underestimated with a value of 30 (95%CI 19; 43) and 33 (95%CI 23; 45) for survey scenarios 1 and 2, respectively. For both simulated datasets, a significant spatial decline in the transmission rate was detected.

Second, we simulated DENV transmission among 1,084 individuals with $\beta=0.6$ and a spatial kernel with $\alpha=10$ (median transmission distance of 82 meters) resulting in 16 cases. Based on the parameters used for the simulation, the reproduction number $R$ of the index case was 0.8; in survey scenario 1 we estimated $R=1.0$ (95%CI 0.4-1.9) and in survey scenario 2 $R=0.7$ (95%CI 0.4; 1.2). The estimated kernel parameter $\alpha$ was 1.6 (95%CI -8; 13) and 10 (95%CI 3; 18) in survey scenario 1 and 2, respectively. A significant spatial decline in the transmission rate was detected only for survey scenario 2 where all cases were detected. This seems to be mainly due to the number of cases, as when estimating $\alpha$ from a simulated dataset without omitting any individuals (592 individuals, $\beta=1.4$, $\alpha=10$, 10 cases), a significant spatial decline could also not be detected.

**Asymptomatic cases**

No asymptomatic dengue case was identified during the door-to-door surveys, while previous studies demonstrated that depending on the setting 20-94% of dengue cases did not develop symptoms (5). We estimated the maximum proportion of asymptomatic cases that would be consistent with the observation of zero asymptomatic cases out of the nine detected DENV cases. The number of asymptomatic DENV cases $n$ follows a Binomial distribution $n \sim \text{Bin}(N, p_{\text{as}})$, where $N$ is the number of total DENV cases and $p_{\text{as}}$ the probability of a case being asymptomatic. We estimated the upper-bound of $p_{\text{as}}$ that is consistent with the observation of nine symptomatic DENV and zero asymptomatic cases, based on a likelihood ratio test and a type 1 error $\alpha=5\%$. The upper bound for $p_{\text{as}}$ lies between 19- 20%. If $p_{\text{as}}$ was >20%, it would be very unlikely that no asymptomatic cases occurred.

**References**


