Supplemental

Dengue virus nested reverse-transcriptase polymerase chain reaction (Nested RT-PCR)

The DENV nested RT-PCR was performed following the method previously described by Klungthong et al. 2015. The method comprised two sequential PCR rounds: the 1st round RT-PCR and the subsequent 2nd round nested PCR. For the 1st round RT-PCR, a pair of universal DENV forward and reverse primers were used. This step involved a single RT-PCR procedure, initiated with a RT step at 42°C for 60 minutes, followed by PCR amplification for 35 cycles with the thermocycling conditions of denaturation at 94°C for 30 seconds, annealing at 55°C for 1 minute, and extension at 72°C for 2 minutes. Following the completion of the 1st round RT-PCR, the PCR products were appropriately diluted and utilized as the template for the subsequent 2nd round nested PCR. The nested PCR assay employed a mixture of five primers, including the universal forward primer used in the 1st round RT-PCR and four DENV type-specific reverse primers. The 2nd round nested PCR was performed using 25 cycles with the same thermocycling conditions as the 1st round PCR step. After the nested PCR amplification, the PCR products were subjected to agarose gel electrophoresis for analysis. The presence of specific DNA bands on the gel enabled the identification of DENV specimens containing types 1, 2, 3, or 4. Specifically, the detection of a DNA band of 482, 119, 290, and 392 base pairs (bp) indicates DENV-1, DENV-2, DENV-3, and DENV-4, respectively. These DNA bands were compared to the amplified DNA from positive controls, which represent the DENV genome. The composition of the PCR buffer mixture and the primers were described previously [1].

Hemagglutination inhibition assay (HAI)

HAI was carried out using goose erythrocytes as previously described [2]. In short, sucrose-acetone extracted DENV1 (Hawaii), DENV-2 (NGC), DENV-3 (H87), DENV-4 (H241), and JEV (JaGAr01) antigens from suckling mouse brain have been used as hemagglutinating antigens. Test sera, positive and negative controls were serial 2-fold diluted starting from 1:10 to 1:20,480 then transferred 250 µL of each dilution into a v-bottom 96-well plate (Thermo Scientific™, US). Then, an equal volume of antigen (8-16 HA units) was added and incubated at 4°C overnight before adding 500 µL of goose red blood cells. After incubation at room temperature for 2 hours, the hemagglutination reaction was observed. A non-inhibitory titer of 1:10 is annotated as <10. Seroconversion was defined by a 4-fold or greater rise in HAI titers for any of the four DENV serotypes and JEV between acute and convalescent sera. A higher seroconversion titer for JEV than for DENV was marked as seroconverted for JEV. Primary DENV infection was defined by a 4-fold rise and HAI titers at equal or lower than 1,280. Secondary DENV infection was defined by a 4-fold rise and HAI titers at equal or higher than 2,560.

Anti-dengue/JE IgM/IgG enzyme immunoassay

Anti-DENV/JEV IgM and IgG capture ELISA was used in this study and performed in duplicate wells of 96-well flat-bottom microplate. Briefly, microplates were coated with 100 µL/well of 1:1,600 dilution of goat anti-human IgM or IgG (KPL, Gaithersburg, MD) in 0.018 M carbonate buffer (pH 9.0). After overnight incubation at 4°C, the plates were washed with phosphate buffered saline (PBS) (pH 7.4) containing 0.5% Tween 20 (PBS-T). Next, 50 µL/well of 1:100

dilution of test serum, negative control (NC), weak positive control (WPC), and strong positive control (SPC) in PBS were added and incubated overnight at 4°C. After washing with PBS-T, 50 µL/well of sucrose acetone extracted suckling mouse brain DENV (pooled DENV antigen: DENV-1 [Hawaii], DENV [NGC], DENV-3 [H87], and DENV-4 [H241]) and JEV (JaGAr01) antigens were added into DENV (IgM/IgG) and JEV (IgM/IgG) plates, respectively. After incubation for 2 h at room temperature, 30 µL/well of human anti-flavivirus IgG–horseradish peroxidase conjugated was added and incubated for 1 hour at 37°C. After washing with PBS-T, 100 µL/well of TMB substrate (KPL, Gaithersburg, MD) was added and incubated for 10–30 min. The reaction was stopped by adding 50 µL/well of 0.2 M sulfuric acid. The absorbance (optical density [OD]) was measured at a wavelength of 450 nm (SoftMax Pro Software, Molecular Devices, San Jose, CA). A valid assay should provide OD values at < 0.100, 0.400–0.600, and > 0.600 for NC, WPC, and SPC, respectively. EIA units of tested serum are equal to 100 × [(ODTest − ODNC)/(ODWPC − ODNC)]; EIA units of IgM ≥ 40 were used as a positive cut-off value. Evidence of dengue infection was classified by a ratio of DENV IgM/JEV IgM ≥ 1.0, and JEV infection when the ratio was < 1.0. Primary DENV infection was interpreted when the ratio of DENV IgM/DENV IgG was ≥ 1.8, and secondary DENV infection was considered when the ratio was < 1.8.

Individual and household related covariate description

Employment status was updated each interval through a yearly questionnaire. We binned occupations by behavior; all farmers were combined into a single classification, all company employees, government officers, general employees, retail merchants, and unspecified occupation being classified as employed, while students and unemployed individuals remained as their own classifications. We then tested a host of household factors. This included factors associated with potential locations for *Aedes* mosquitoes to breed including the number of water containers, plastic water containers, and plastic bottles all of which we analyzed as log10 counts. In addition, indicator covariates associated with the household being near a source of water and whether the household was supplied water by a pipe were included. Covariates for the physical structure of the house were also included. These include whether the house had a zinc roof, was concrete, had door screens, and the number of toilets outside the home. In addition, whether the house was built on poles, was a single unit, or a townhouse was analyzed along with the garbage management system of the home (car collection versus burnt/buried/dumped).

Force of infection

A catalytic model was fit to estimate the force of infection in the population. Baseline seroprevalence of DENV in a subset of individuals enrolled in the study before 2017 were used for this analysis. We fit a model with age only to estimate the general force of infection with a *cloglog* link function [3]. We ran this model on individuals over the age of one to remove the impact of maternal antibodies. We also removed individuals over the age of 30 as the entire population approaches complete seropositivity around this age. If λ is the annual force of infection we are able to translate this to a proportion of susceptible individuals infected per year using the following formula, $1 - exp(-\lambda)$.

Annualized probability of infection

Individuals in the cohort were sampled approximately once a year, however interval lengths varied with 95% being between 229 and 643 days long. Since individuals with longer intervals had more time to get infected we accounted for this by annualizing the probability of infection to be $1 - exp(-365x)$ where x is the daily probability of infection. Due to the interval censoring effect of the yearly sampling process we assumed that if an infection occurred in an interval of length *d* days then the daily probability of infection *x* was *1/d*.

Sensitivity Analysis

We conducted sensitivity analyses on the methods outlined in the *Individual and household level risk* subsection of the methods. When using the four-fold classification rule mentioned in the main text we found the following. Each additional newborn increased infection risk with an aOR of 2.17 (95% CI, 1.65 - 2.85) (Extended Data Figure 4a). We also found that each additional one to five year old increased the odds of infection with an aOR of 1.20 (95% CI, 1.05 - 1.37). In addition, each additional adult reduced the likelihood of infection with an aOR of 0.93 (95% CI 0.88 - 0.98). When stratified by sex we found that each additional male and female newborn increased the odds of infection with an aOR of 2.34 (95% CI, 1.70 - 3.24) and 1.96 (95% CI, 1.41 - 2.73) respectively (Extended Data Figure 4b). Also, each additional male between one and five increased the odds of infection with an aOR of 1.31 (95% CI, 1.12 - 1.53). We conducted this same sensitivity analysis on how the attack rate of the household in the previous interval impacted infection risk and found that high attack rates (>0.2) in the previous interval reduced the likelihood of infection with an aOR of 0.56 (95% CI, 0.44 - 0.72) when compared to individuals coming from a household with no infections the previous year. (Extended Data Figure 4c). We lastly analyzed how average household pre-interval titers impacted risk and we found that both medium (40-66) and high (>66) average household pre-interval titers reduced the likelihood of infection with an aOR of 0.74 (95% CI, 0.61 - 0.90) and 0.66 (95% CI, 0.53 - 0.83) respectively when compared to households with low average pre-interval titers (<40) (Extended Data Figure 4d).

When limiting the analysis to all intervals taken from households with more than 80% of their members sampled we find the following. For the household structure analyses we found that each additional newborn increased the odds of infection with an aOR of 2.27 (95% CI, 1.50 - 3.46) (Extended Data Figure 5a). Each additional adult had a protective but insignificant impact on the likelihood of infection with an aOR of 0.94 (95% CI, 0.86 - 1.02). When stratified by sex we found that both male and female newborns increased the odds of infection with an aOR of 1.70 (95% CI, 1.12 - 2.59) and 2.27 (95% CI, 1.50 - 3.46) respectively (Extended Data Figure 5b). In addition, each additional male between five and eighteen increased the likelihood of infection with an aOR of 1.22 (95% CI, 1.03 - 1.43). No other stratified groups had a significant impact on infection risk. We conducted this same sensitivity analysis on how the attack rate of the household in the previous interval impacted infection risk (Extended Data Figure 5c). We found that high attack rates in the previous interval reduced the likelihood of infection with an aOR of 0.57 (95% CI, 0.41 - 0.79) when compared to individuals coming from a household with no infections the previous year. We found no impact on infection risk of the average titers of

individuals in the household from the previous interval in adjusted analyses (Extended Data Figure 5d).

When limited to seronaive individuals we found the following. Each additional newborn increased infection risk with an aOR of 1.93 (95% CI, 1.10 - 3.37) (Extended Data Figure 6a). We also found that each additional one to five year old and five to 18 year old also increased the odds of infection with an aOR of 1.31 (95% CI, 1.05 - 1.62) and 1.16 (1.01 - 1.33) respectively. When stratified by sex we found that each additional female newborn increased the odds of infection with an aOR of 2.01 (95% CI, 1.03 - 3.90). Also, each additional male between one and five increased the odds of infection with an aOR of 1.49 (95% CI, 1.14 - 1.93) (Extended Data Figure 6b). We conducted this same sensitivity analysis on how the attack rate of the household in the previous interval impacted infection risk and found no significant impact for any level (Extended Data Figure 6c). We found medium (40-66) average household pre-interval titers reduced the likelihood of infection with an aOR of 0.51 (95% CI, 0.35 - 0.76) when compared to households with low average pre-interval titers (<40) (Extended Data Figure 6d).

Percent of household immune

We performed the following analysis to understand how the percent of a household that is immune at the beginning of an interval impacts risk of dengue infection. We define this to be the percentage of the household, not including the individual of interest, who have any of their DENV serotype HAI titers at or above some threshold (20, 40, and 80). We subsequently split this percentage into three ranges, [0,.71], (0.71-1),1 that split the data into approximately thirds. We then ran adjusted analyses as outlined in the methods section adjusting for average individual titers, month, year, and household random effects. Adjusted analyses show significant protective effects for higher proportions of the household that are immune further demonstrating the indirect protective effects of the household on an individual (Figure S4).

Japanese encephalitis virus (JEV)

JEV, another member of the *Flavivirus* genus, is known to have cross-reactivity with DENV after both disease and vaccine induced infections [48]. We wanted to quantify the impact of JEV vaccination, which was introduced nationwide in Thailand in the 1990s, on the predicted outcome of infection. We found that JEV vaccination events occurred in 4.7% of all intervals, but had no significant impact on infection probability in intervals where children under the age of five were administered the vaccine (p=0.63). These results suggest that this cross-reactivity did not significantly impact predictions.

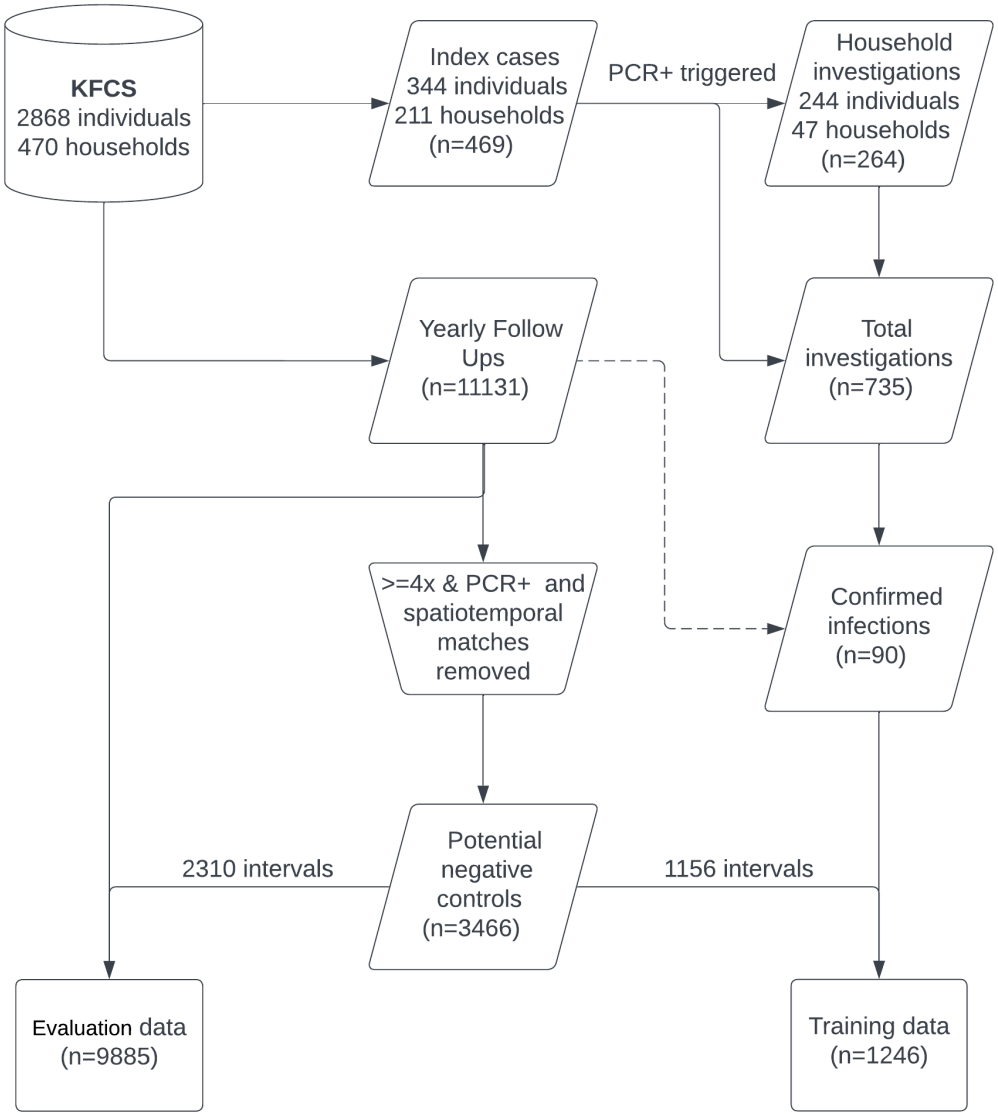

Figure S1. Data flow from original stored data to the training and evaluation data sets utilized for the analysis. The number of intervals found at each step is noted by *n*. A total of 470 households and 2868 individuals were enrolled in our cohort. Index cases were tested for DENV infections using PCR. Positive test results led to household investigations for the remaining household members. This led to 735 total intervals of which 90 had a confirmed DENV infection which formed part of our training data. From the complete dataset we then removed these confirmed DENV infections, their spatiotemporal matches, as well as any intervals that had a maximum DENV ratio at or above a four-fold rise. This led to 3466 total potential negative controls from which one third were randomly chosen to be added to the confirmed infections that would make up our training data (n=1246). The remaining intervals were kept in our evaluation data (n=9885).

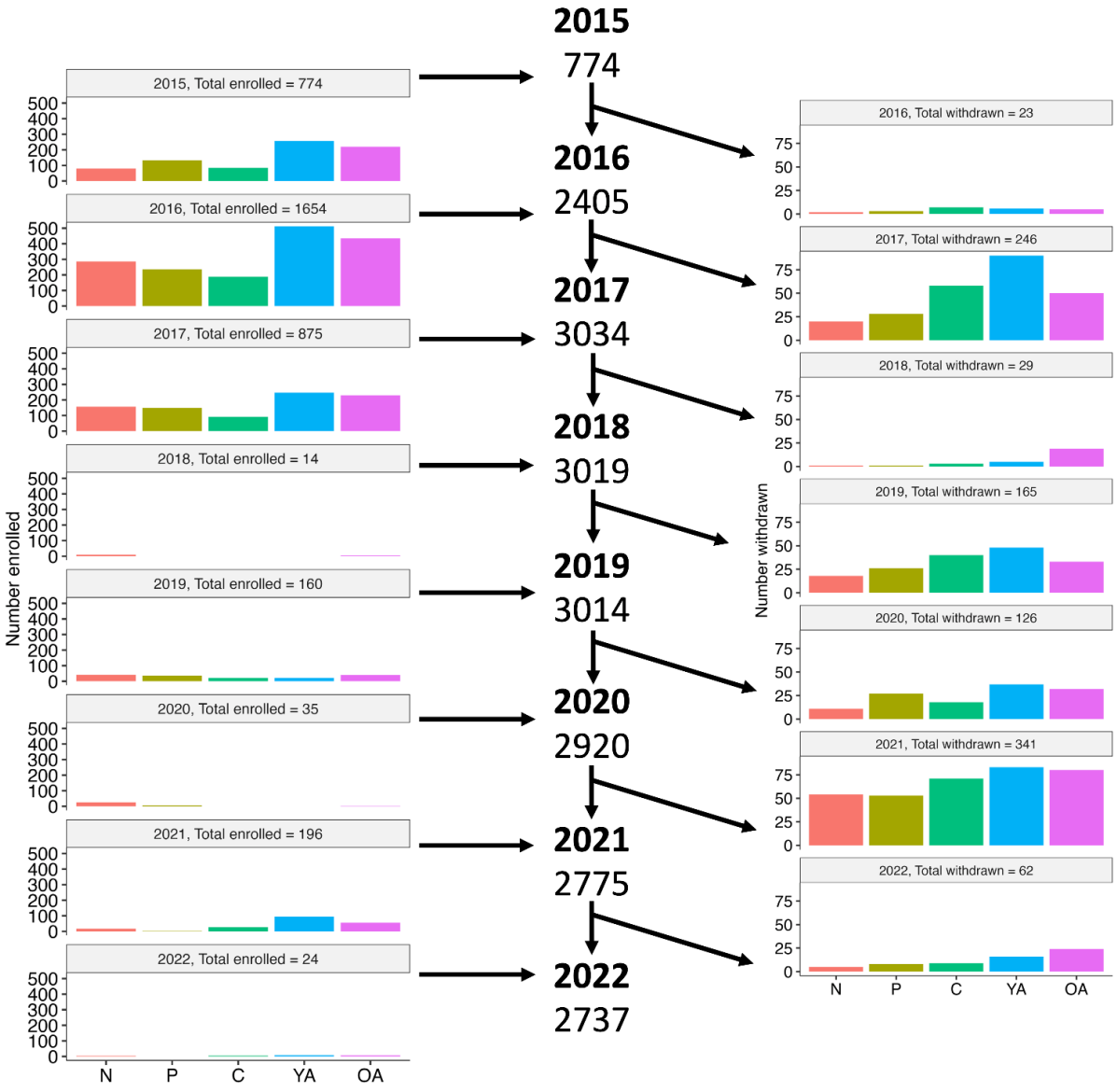

Figure S2. Number of individuals enrolled and withdrawn per year of the analysed study. Each year's total is noted as well as the cumulative sum across time when accounting for individuals gained and lost. The number of newborns (N), mothers (P), children under the age of five (c), young adults between five and eighteen years of age (YA), and other adults (OA) are visualized each year.

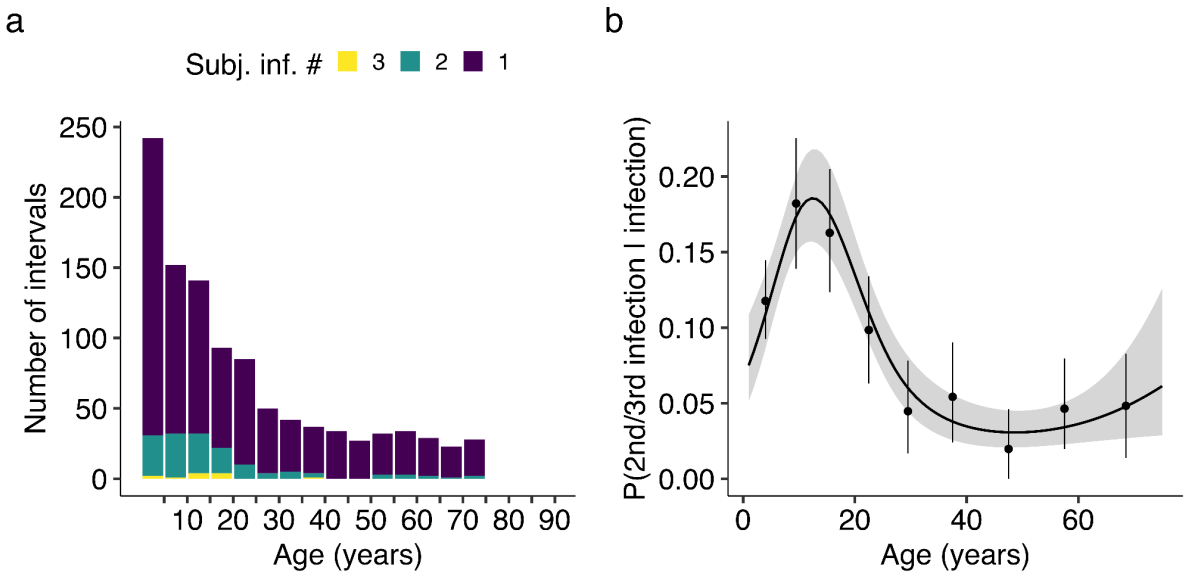

Figure S3. (a) Distribution of the number of intervals predicted to have an infection as a function of age at post–interval follow up date. Intervals are color coded by the number of cumulative infections the subject in said interval has had across the study period (n=1049). (b) Probability of experiencing re-infections during the study by age (n=2298). Points and intervals representing the mean and 95% confidence interval while the shaded region represent the 95% confidence interval for the fit.

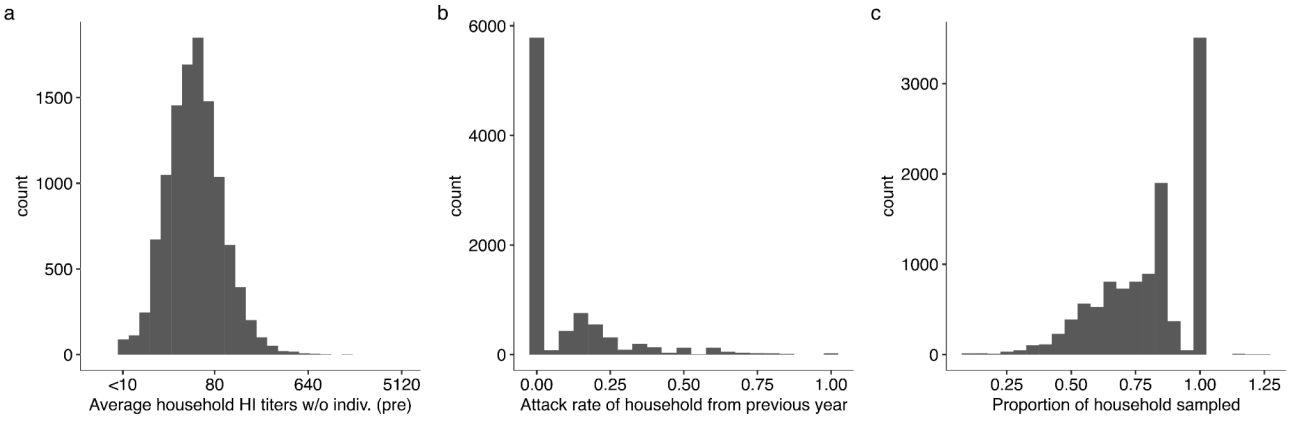

Figure S4. Distribution of household related variables including (a) average household HAI titer not including the individual of interest, (b) attack rate of households from the previous year, and (c) the proportion of a household that is sampled.

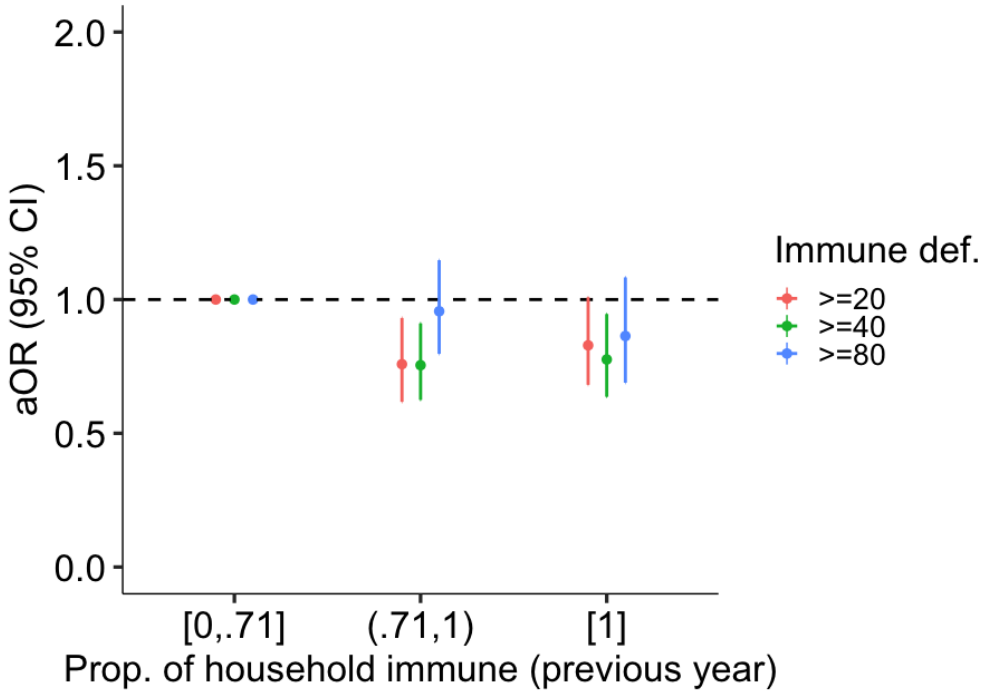

Figure S5. Impact of proportion immune on results for various definitions of immunity (any titer at or above an HAI of 20, 40, or 80 (n=11131). Adjusted model accounts for household random effects, individual pre-interval titers, as well as the year and month of the post-interval sample. Points and intervals represent the mean and 95% confidence interval.

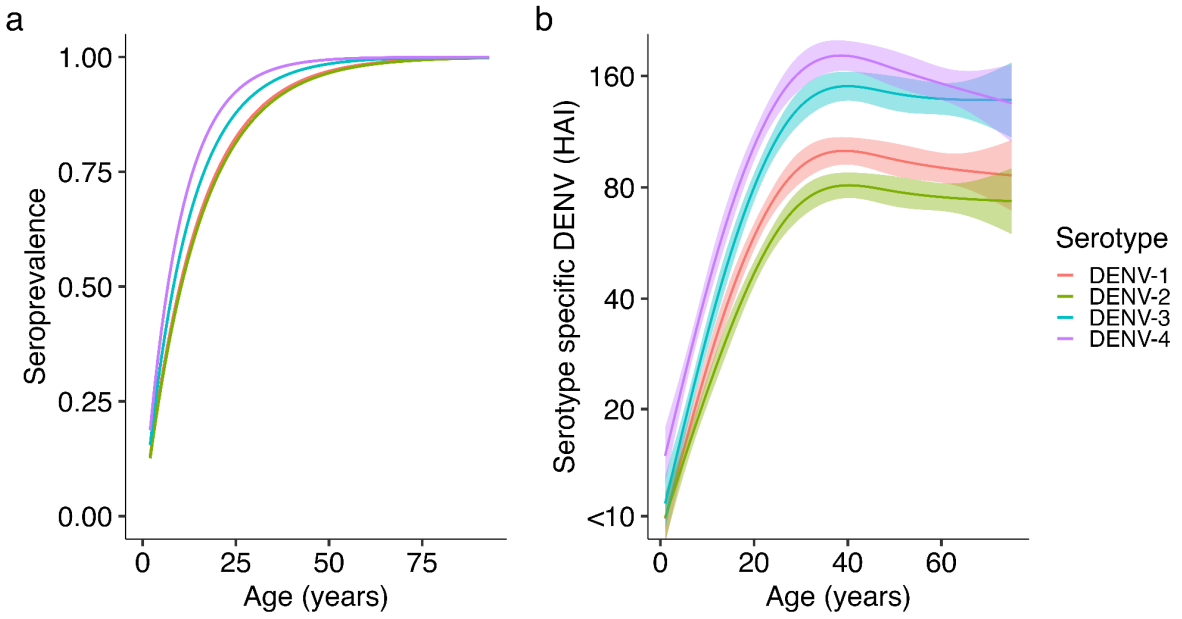

Figure S6. Summary of cohort data separated by DENV serotype. (a) Age-stratified seropositive at enrollment for subjects enrolled before 2017. The resulting fit of the serocatalytic model fit to each serotype is presented, with details found in the Supplemental Information. (b) Average DENV HAI titers at enrollment age binned. 95% confidence intervals are presented.

Table S1. The following predictors are incorporated into the XGboost model for training and prediction.

Predictor	**Units**
Individual pre-interval DENV 1-4 and JEV titers	HAI
Geometric mean of pre-interval DENV 1-4 titers	HAI
Maximum of pre-interval DENV 1-4 titers	HAI
Minimum of pre-interval DENV 1-4 titers	HAI
Variance of pre-interval DENV 1-4 titers	HAI
Individual post-interval DENV 1-4 and JEV titers	HAI
Geometric mean of post-interval DENV 1-4 titers	HAI
Maximum of post-interval DENV 1-4 titers	HAI
Minimum of post-interval DENV 1-4 titers	HAI
Variance of post-interval DENV 1-4 titers	HAI
Serotype and JEV specific post to pre-interval titers	Ratio
Mean of post to pre-interval DENV 1-4 titer ratios	Ratio
Maximum of post to pre-interval DENV 1-4 titer ratios	Ratio
Minimum of post to pre-interval DENV 1-4 titer ratios	Ratio
Variance of post to pre-interval DENV 1-4 titer ratios	Ratio
Serotype and JEV specific post minus pre-interval titers	HAI
Geometric mean of post minus pre-interval DENV 1-4 titers	HAI
Maximum of post minus pre-interval DENV 1-4 titers	HAI
Minimum of post minus pre-interval DENV 1-4 titers	HAI
Variance of post minus pre-interval DENV 1-4 titers	HAI
Sex	NA
Time between samples	Days
Date of pre-interval sample	Days
Date of post-interval sample	Days

Age at post-interval sample	Years
Age at enrollment	Years

Table S2. Odds ratios and adjusted odds ratios for individual and household level covariates of infection risk. Univariate analyses inform which variables are incorporated in the multivariate analyses. Multivariate analyses were conducted using either just household random effects, both household and individual random effects, or household random effects and incorporating individual immunity. Odds ratios and adjusted odds ratios for temporal covariates can be found in Table S3. Abbreviations: aOR, adjusted odds ratio; CI, confidence interval; Nb., number;OR, odds ratio; REF, Reference category.

Covariate		**Univariate**	**Multivariate House res**	**Multivariate indiv., house res**	**Multivariate house res w/ indiv. immunity**
		OR (95% CI)	**aOR (95% CI)**	**aOR (95% CI)**	**aOR (95% CI)**
Individual level					
Sex	Female	REF	–	–	–
	Male	1.11 (0.98 - 1.27)	–	–	–
Age	[1,5)	REF	REF	REF	REF
	[5,18)	1.11 (0.93 - 1.33)	1.16 (0.95 - 1.41)	1.16 (0.95 - 1.41)	1.44 (1.16 - 1.77)
	[18,30)	0.68 (0.55 - 0.84)	0.74 (0.58 - 0.95)	0.74 (0.58 - 0.95)	1.41 (1.06 - 1.89)
	[30,50)	0.46 (0.37 - 0.57)	0.41 (0.30 - 0.57)	0.41 (0.30 - 0.57)	0.93 (0.63 - 1.35)
	50+	0.43 (0.34 - 0.53)	0.41 (0.27 - 0.62)	0.41 (0.27 - 0.62)	0.88 (0.55 - 1.40)
Occupation	Employed	REF	REF	REF	REF
	Farmer	0.57 (0.27 - 1.19)	0.68 (0.31 - 1.48)	0.68 (0.31 - 1.48)	0.63 (0.29 - 1.38)
	Student	1.64 (1.17 - 2.29)	0.78 (0.51 - 1.18)	0.78 (0.51 - 1.18)	0.66 (0.43 - 1.01)
	Unemployed	1.24 (1.02 - 1.50)	0.82 (0.63 - 1.07)	0.82 (0.63 - 1.07)	0.79 (0.60 - 1.04)
Household level					
Nb. Water container (log10)		0.94 (0.78 - 1.14)	–	–	
Nb. Water container - plastic (log10)		1.03 (0.90 - 1.18)	–	–	–
Nb. Plastic bottles (log10)		0.99 (0.86 - 1.14)	–	–	–
House type	Poles	REF	–	–	–
	Single	1.08 (0.91 - 1.28)	–	–	–
	Townhouse	0.97 (0.67 - 1.40)	–	–	–
Garbage management	Burnt / Buried / Dumped	REF	–	–	–
	Car collection	1.11 (0.97 - 1.26)	–	–	–
Concrete house		1.08 (0.92 - 1.26)	–	–	–
Zinc roof		1.00 (0.87 - 1.14)	–	–	–
Nearby source of water		1.01 (0.88 - 1.16)	–	–	–
Water supply by pipe		0.92 (0.74 - 1.13)	–	–	–

Door screens	1.04 (0.96 - 1.13)	–	–	–
Number of toilets outside	1.03 (0.93 - 1.14)	–	–	–

Table S3. Odds ratios and adjusted odds ratios for temporal covariates of infection risk. Univariate analyses inform which variables are incorporated in the multivariate analyses. Multivariate analyses were conducted using either just household random effects, both household and individual random effects, or household random effects and incorporating individual immunity.

Covariate		**Univariate**	**Multivariate House res**	**Multivariate indiv., house res**	**Multivariate house res w/ indiv. immunity**
		OR (95% CI)	**aOR (95% CI)**	**aOR (95% CI)**	**aOR (95% CI)**
Temporal level					
Year	2015	REF	REF	REF	REF
	2016	0.56 (0.36 - 0.88)	0.47 (0.27 - 0.83)	0.47 (0.26 - 0.82)	0.48 (0.27 - 0.85)
	2017	0.77 (0.50 - 1.16)	0.71 (0.42 - 1.20)	0.70 (0.41 - 1.20)	0.73 (0.43 - 1.26)
	2018	1.55 (1.03 - 2.32)	1.21 (0.71 - 2.06)	1.21 (0.71 - 2.08)	1.27 (0.74 - 2.19)
	2019	1.58 (1.05 - 2.37)	1.38 (0.77 - 2.46)	1.39 (0.78 - 2.50)	1.51 (0.84 - 2.72)
	2020	1.43 (0.95 - 2.17)	1.47 (0.77 - 2.79)	1.49 (0.77 - 2.86)	1.58 (0.82 - 3.05)
	2021	0.90 (0.54 - 1.51)	0.58 (0.29 - 1.15)	0.57 (0.29 - 1.14)	0.60 (0.30 - 1.20)
Month	Jan.	2.95 (2.21 - 3.94)	2.13 (1.43 - 3.15)	2.13 (1.43 - 3.15)	2.11 (1.40 - 3.17)
	Feb.	1.80 (1.31 - 2.48)	1.57 (1.02 - 2.43)	1.57 (1.02 - 2.43)	1.58 (1.01 - 2.47)
	Mar.	1.30 (1.01 - 1.67)	0.96 (0.69 - 1.34)	0.96 (0.69 - 1.34)	0.95 (0.68 - 1.33)
	Apr.	1.12 (0.85 - 1.47)	1.05 (0.75 - 1.47)	1.05 (0.75 - 1.47)	1.04 (0.74 - 1.47)
	May	REF	REF	REF	REF
	Jun.	1.37 (1.09 - 1.73)	1.27 (0.95 - 1.71)	1.27 (0.95 - 1.71)	1.26 (0.93 - 1.70)
	Jul.	1.66 (1.24 - 2.22)	0.84 (0.53 - 1.33)	0.84 (0.53 - 1.33)	0.83 (0.52 - 1.32)
	Aug.	1.91 (1.42 - 2.58)	0.90 (0.55 - 1.48)	0.90 (0.55 - 1.48)	0.90 (0.54 - 1.49)
	Sep.	2.17 (1.66 - 2.84)	1.12 (0.71 - 1.79)	1.12 (0.71 - 1.79)	1.18 (0.74 - 1.91)
	Oct.	2.26 (1.65 - 3.08)	1.16 (0.70 - 1.93)	1.16 (0.70 - 1.93)	1.19 (0.71 - 2.00)
	Nov.	1.36 (0.83 - 2.22)	0.48 (0.23 - 1.02)	0.48 (0.23 - 1.02)	0.50 (0.23 - 1.07)
	Dec.	NA	NA	NA	NA

References

1. Klungthong, C. *et al.* Monitoring and improving the sensitivity of dengue nested RT-PCR used in longitudinal surveillance in Thailand. *J. Clin. Virol.* **63**, 25–31 (2015).
2. Clarke, D. H. & Casals, J. Techniques for hemagglutination and hemagglutination-inhibition with arthropod-borne viruses. *Am. J. Trop. Med. Hyg.* **7**, 561–573 (1958).
3. Bjørnstad, O. N. *Epidemics: Models and Data using R*. (Springer, 2018).