# Age-specific case data reveal varying dengue transmission intensity in US states and territories

- 3
  4
  5 Sarah Kada<sup>1</sup>\*, Gabriela Paz-Bailey<sup>1</sup>, Laura Adams<sup>1</sup> & Michael A. Johansson<sup>1</sup>
  6
  7 <sup>1</sup> US Center for Disease Control and Prevention (CDC), Dengue Branch, San Juan, Puerto Rico.
  9
  10 \* Corresponding author
  11 E-mail: sarahkada@yahoo.fr (SK).
- 12 Number of words: 4341

# 13 Abstract

Dengue viruses (DENV) are endemic in the US territories of Puerto Rico, American 14 Samoa, and the US Virgin Islands, with focal outbreaks also reported in the states of 15 Florida and Hawaii. However, little is known about the intensity of dengue virus 16 transmission over time and how dengue viruses have shaped the level of immunity in 17 these populations, despite the importance of understanding how and why levels of 18 19 immunity against dengue may change over time. These changes need to be considered when responding to future outbreaks and enacting dengue management 20 21 strategies, such as guiding vaccine deployment. We used catalytic models fitted to 22 case surveillance data stratified by age from the ArboNET national arboviral surveillance system to reconstruct the history of recent dengue virus transmission in 23 Puerto Rico, American Samoa, US Virgin Islands, Florida, Hawaii, and Guam. We 24 NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice. estimated average annual transmission intensity (i.e., force of infection) of DENV 25

26 between 2010 and 2019 and the level of seroprevalence by age group in each 27 population. We compared models and found that treating all reported cases as 28 secondary infections generally fit the surveillance data better than models 29 considering cases as primary infections. Using the secondary case model, we found that force of infection was highly heterogeneous between jurisdictions and over time 30 within jurisdictions, ranging from 0.00003 (95% CI: 0.00002-0.0004) in Florida to 31 32 0.08 (95% CI: 0.044-0.14) in American Samoa during the 2010-2019 period. For early 2020, we estimated that seropositivity in 10 year-olds ranged from 0.09% 33 (0.02%-0.54%) in Florida to 56.3% (43.7%-69.3%) in American Samoa. In the 34 35 absence of serological data, age-specific case notification data collected through routine surveillance combined with mathematical modeling are powerful tools to 36 monitor arbovirus circulation, estimate the level of population immunity, and design 37 dengue management strategies. 38

# **39** Author Summary

40 Viruses transmitted by *Aedes* mosquitoes are a substantial public health burden 41 globally. In the US, the increasing number of outbreaks in recent decades and the co-circulation of the four dengue viruses present a risk of experiencing sequential 42 43 infections, which can result in more severe disease. However, reported cases only represent a proportion of all infections, so the intensity of dengue circulation in US 44 jurisdictions and the level of population immunity are largely unknown. Combining 45 46 dengue surveillance data stratified by age with mathematical models enables direct estimation of transmission intensity and previous exposure. We found that dengue 47 virus transmission intensity was highest in American Samoa and Puerto Rico. The 48

US Virgin Islands also showed patterns of endemic transmission, while Hawaii and
Guam present more sporadic outbreak profiles. These estimates can help improve
risk assessment in these locations and guide interventions to the populations who
most need them.

#### 53 Introduction

54 Dengue viruses (DENV) re-emerged in the Americas beginning in the 1970s, with the 55 viruses rapidly expanding their range and associated burden of dengue infections [1– 56 3]. However, this expansion has not been uniform. The four antigenically distinct 57 dengue virus serotypes (DENV-1 to DENV-4) emerged at different times with affected 58 areas experiencing different levels of transmission intensity, from sporadic outbreaks 59 to hyperendemic, year-round circulation [4–6].

The varied history of DENV transmission in the Americas directly shapes the current risk. Infection with one serotype generally provides long-term immunity to that serotype and short-term immunity to all serotypes. Secondary exposure to a different serotype from the primary infection has been associated with increased risk of severe dengue [7]. Population immunity is therefore modulated by the circulation of the four serotypes and the intensity of transmission of those viruses.

As population immunity and transmission intensity are intrinsically linked, it is 66 67 challenging to estimate either of those components directly. Existing immunity can be 68 partially measured with age-specific serological surveys, but these are logistically and financially challenging to implement, provide only a snapshot of previous exposure, 69 and differentiating specific serotypes and the sequence of prior serotype exposures is 70 71 very difficult. Transmission intensity can be characterized as the force of infection (FOI), the per capita rate at which susceptible individuals become infected by an 72 infectious disease. A class of models called catalytic models can estimate the FOI, 73 74 providing a measure of changes in exposure to diseases. However, direct estimation 75 of the FOI is also challenging because population susceptibility is generally not

known, many dengue infections are inapparent or unreported, and the proportion of
individuals experiencing inapparent disease is dependent on prior infection [8–11].
Estimating the FOI therefore requires accounting for both underreporting and
historical transmission intensity.

80 To help meet this challenge, age-stratified case data collected through passive surveillance can be used to gather insight into both existing immunity and 81 82 transmission intensity [12]. Due to every age group having different exposure 83 periods, incidence across each age group each year reflects both the age-group specific immunity and the yearly transmission intensity, enabling estimation of yearly 84 85 transmission intensity for both current and past years [13]. Catalytic models that use age-specific case data can thus provide valuable insights into the dynamics of 86 87 arboviruses [12,14,15] by quantifying infection burden in settings where seroprevalence data are unavailable or unrepresentative of the population. 88

89 Here, we analyze reported dengue case data from six US jurisdictions with outbreaks reported since dengue became a nationally reportable disease in 2010: Puerto Rico, 90 91 American Samoa, US Virgin Islands, Florida, Hawaii and Guam. These locations represent diverse historical trends in reported case data, from long-term hyper-92 93 endemicity in Puerto Rico to no reported dengue in decades prior to an outbreak in 94 2019 in Guam (Figure 1). Building on previous work by Rodriguez-Barraquer et al. 95 [12], we developed a generalized model that accounts for variability in transmission intensity and reporting over time and estimates the FOI and age-specific immunity 96 over time. Accounting for yearly variability in FOI and reporting provides critical 97 insights into dengue epidemiology in a wide variety of settings, from hyperendemic to 98 99 sporadic outbreaks, and can be used to better assess dengue burden in the US and

other settings, contribute to epidemic risk assessment, and guide implementation and
evaluation of control strategies such as vaccination.

## 102 Methods

#### 103 Data sources

104 Dengue became a nationally notifiable disease in January 2010 and we analyzed confirmed and probable dengue cases reported to ArboNET for the years 2010-2019 105 for six jurisdictions: Puerto Rico, American Samoa, US Virgin Islands, Hawaii, Florida 106 107 and Guam. Texas was excluded due to low case counts. ArboNET includes information about place of likely acquisition (locally acquired or travel-associated) and 108 disease severity. We analyzed locally acquired cases only, except Florida, where we 109 110 analyzed both categories. For all locations, we combined severe and non-severe 111 cases. In Puerto Rico, where the highest number of severe cases was reported, we also analyzed severe cases (including dengue hemorrhagic fever and dengue shock 112 113 syndrome [16,17]) on their own. Age distribution data for each population was obtained from the 2010 census data provided by the United States Census Bureau 114 (census.gov), except for Florida, Hawaii and Puerto Rico where in addition to the 115 2010 census data, population estimates were available for years 2011–2019. The 116 117 ArboNET surveillance data is available to researchers upon request.

#### **118** Estimating the force of infection

We used a Bayesian catalytic model extending one developed by Rodriguez-119 Barraguer et al., [12] to estimate the time-dependent force of infection (FOI(t)) of all 120 circulating serotypes and derive age-specific seroprevalence estimates. Building on 121 the original model, we (i) added random effects for both reporting and FOI 122 123 probabilities to allow a long-term mean with yearly variability for each of these components. We also (ii) implemented separate models for different reporting 124 125 possibilities (cases reported are primary DENV infections, secondary infections or a 126 combination of both) and (iii) accounted for potential infections in infants that appear more like secondary infections due to possible maternal antibody transfer. Finally, we 127 (iv) generated a full suite of retrospective FOI estimates (prior to 2010, i.e., when 128 129 ArboNET data became available).

Reporting probability. Many DENV infections are mild or asymptomatic, and do not 130 result in a visit to a clinic or being reported as a case [8,9,18]. Building on previous 131 estimates, we set a prior for the probability of reporting a case using a logit-normal 132 distribution (mean = -2.2,  $\beta$  = 0.7) for the long-term mean, which translates into a 133 134 mean value of approximately 10% with a 97.5% bound of 30%. Multiple factors may lead to year-to-year fluctuation of reporting probabilities; the presence of other 135 136 arboviruses infections may influence dengue reporting and testing, and, during 137 epidemic years, doctors may be more likely to recognize and test dengue cases [19-22]. To allow some variability in reporting probabilities, we let it vary between years 138 and partitioned the variance of the reporting prior between uncertainty in the mean 139 and uncertainty in year-to-year variability (see S1). 140

The probability of identifying and reporting dengue cases is also dependent on case severity. In one cohort study, the proportion of inapparent versus symptomatic case was similar between primary and secondary cases, with age, year, and time interval between consecutive infections more likely to affect symptomatic status [23].

Additionally, laboratory confirmation generally relies on virological testing that cannot differentiate primary and secondary infections, so the proportion of cases reported to ArboNET that are primary or secondary is not known. We therefore compared three different models considering either that all reported cases are secondary infections (Model S), that reported cases are a combination of primary and secondary infections (Model PS), and that all reported cases are primary infections (Model P, see Supporting information S1).

152 Severe dengue cases, which are predominantly secondary infections, are also reported to ArboNET. Because substantial numbers of severe cases were reported in 153 154 Puerto Rico during the study time period, we performed a secondary analysis for 155 Puerto Rico limiting the included cases to severe dengue only. While severe cases 156 are more likely to be reported than non-severe cases, severe cases represent only a small proportion of secondary infections. In previous cohort studies, the estimated 157 158 probability of a secondary infection resulting in a reported severe case was lower 159 than the reporting probability of a secondary infection resulting in any reported 160 dengue case (dengue fever or severe dengue), below 5% (less than 1% [23] and 3% 161 [24]). Based on those studies, we used a lower prior for the reporting of severe cases, setting a mean of 1% of all primary or secondary (depending on the model) 162 163 infections being reported as severe cases and a 97.5% bound of 5%.

FOI. The FOI is likely to vary over time depending on the circulating dengue 164 165 serotypes, serotype introduction or re-introduction, and population immunity, but few catalytic models account for this fluctuation [25]. Similar to the reporting probabilities 166 167 (described above), we use a logit-normal prior for long-term average FOI and partitioned the prior variance between uncertainty in the mean and uncertainty in the 168 year-to-year variability. We estimated a specific FOI prior for each location, using the 169 170 average age of infection, life expectancy, and the mean reporting prior (see Supporting information S1, Table S1). 171

**Infant cases.** The presence of maternal antibodies may interfere and potentially predispose some infants to more severe disease, such that infant infections may be more severe and more likely to be reported than expected for a true primary infection [26]. To account for this, we added a proportion parameter,  $\alpha$ , that sets the initial previous exposure for infants to a proportion of the FOI in the previous year rather than zero (see S1, Figure S13).

# 178 Model comparison

Models were compared using leave-one-out (loo) Pareto smoothed importance 179 sampling [27]. Lower loo scores indicate better cross-validation model fit. We also 180 collected yearly log-likelihood estimates to assess how our S, P and PS models fit to 181 the reported case data. For each parameter we extracted the posterior medians and 182 95% Bayesian credible intervals (hereafter 95% CI). Models were run using Bayesian 183 Markov Chain Monte Carlo sampling with a No-U-Turn sampler using the RStan 184 package [28] in R (version 4.1.2, [29]). Detailed information on the methods used can 185 186 be found in Supporting Information S1.

# 187 **Results**

188 Dengue circulation was highly heterogeneous during 2010–2019 across the six jurisdictions (Figure 1). The age distribution of cases also varied markedly, with the 189 most commonly reported age group being 10-14 years in the US Virgin Islands and 190 Guam to 60–64 years for locally acquired cases in Florida (Figure 2). If all age groups 191 192 had the same risk of infection, the age distribution of cases would mirror the 193 population age distribution, e.g., the largest age group is more likely to have the highest number of reported cases. However, immunity from prior infection, especially 194 195 in endemic areas, shifts the age of cases to predominantly younger age groups (e.g., 196 Puerto Rico, American Samoa).

197 We fitted models to each time series of age-specific case data under three different assumptions: (i) that all reported cases are primary infections (Model P), (ii) that all 198 199 reported cases are secondary infections (Model S), and (iii) that reported cases 200 include both primary and secondary cases (Model PS). In Hawaii and Guam, Model 201 P and Model PS had significantly lower loo error than Model S (Figure 3). The fit of 202 Model S was only significantly better for local cases in Florida (i.e., excluding travelassociated cases) though also had lower mean error for the US Virgin Islands and 203 204 severe cases in Puerto Rico.

Comparing the different models across years, age groups and locations, we found
mixed evidence about relative fit. For Puerto Rico, log likelihoods indicated that
Model S tended to have a better fit in years with more reported cases and sometimes
a worse fit in years and age groups with fewer reported cases (S1, Figures S5-S6).
Model S also appeared to fit the distribution of cases across younger age groups

more accurately across most jurisdictions (S1, Figures S7 to S13). Models P and PS
both tend to estimate that a larger number of cases should be reported in the
youngest age group (0-4 years old). Model S qualitatively better matches the distinct
pattern of dengue cases in children across endemic areas. Because of this distinct
advantage and the finding that fit for Model S was only inferior for Hawaii and Guam
(where there was the least evidence of sustained transmission) (S1, Figures S7 to
S13), we used Model S for subsequent analyses.

217 Of the six locations analyzed, Puerto Rico had the second highest long-term mean 218 estimated FOI, an estimated 1.1% (95% CI: 0.9–1.3%) per year with substantial year-219 to-year variation including higher FOI estimates in the early 1990s and 2000s (prior to the dataset analyzed here, Figure 4). The estimated probability of a secondary 220 infection resulting in a reported dengue case was 2.1% (95% CI: 1.4-3.3%). The 221 estimated probability of a secondary infection resulting in a reported severe dengue 222 223 case was 0.1% (95% CI: 0.1–0.3%). Reporting was estimated to be slightly higher in 224 years when more cases were reported (Figure 5). For 2019, we estimated 225 seroprevalence in 5-year-old children to be 10.1% (95% CI: 6.1–16.9%), increasing 226 to 94.5% (95% CI: 92–96.2%) by age 35 (Figure 2A). FOI and seroprevalence 227 estimates using only severe cases were similar to those using all cases in Puerto 228 Rico, although we observe a lower seroprevalence for 2019 in younger age groups 229 and higher in older age groups (Figure 6).

While no cases were reported in American Samoa between 2010 and 2016, the
model estimated low reporting in 2010–2016 (Figure 5) and high uncertainty for FOI
prior to 2010 with the highest long-term average FOI of all locations at 1.4% (95% CI:
1.1–1.8%, Figure 4). The high estimated FOI in 2017–2018 translated to an

estimated 2019 seroprevalence in 2019 for 5-year-olds of 46.5% (95% CI: 33.2–
61.6%, Figure 2B). Thus, the age structure of cases in 2017-2018, indicates that
while this was a large outbreak, there was likely dengue transmission both in 20102016 and in preceding years. The FOI estimate for American Samoa was higher than
the estimate for Puerto Rico, with similar estimated adult seroprevalence (91.9%,
95% CI: 84.8–95.7% for 35-year-olds, Figure 2B).

240 The US Virgin Islands had lower estimated long term average FOI (0.5%, 95% CI:

0.3–0.7%, Figure 4) and more stable estimated reporting (Figure 5). With a low FOI

in recent years, the 5-year-old estimated seroprevalence for 2019 was low (5.1%,

243 95% CI: 2.3–11%, Figure 2C) but cumulative exposure in adults was still high, with

244 35-year-old estimated seroprevalence for 2019 of 86.1% (95% CI: 72.7–94.1%,

245 Figure 2C).

246 Hawaii reported two outbreaks (2011 and 2014-2015) and Guam reported one (2019) during the study period (Figure 1) and both had less distinct age-specific case 247 patterns (Figure 2D, 2F). The estimates for reporting largely reflected the reporting 248 249 prior (Figure 4), indicating that there was little information between the case numbers and case age data to distinguish differences in reporting versus FOI. Nonetheless, 250 251 long-term average FOI estimates were low for both jurisdictions (Hawaii: 0.001%, 252 95% CI: 0.0003–0.004%; Guam: 0.003%, 95% CI: 0.0008–0.013%; Figure 4), with 253 significant increases in the years when the outbreaks were reported. The 2019 254 seroprevalence estimates were low for all age groups in both locations: 6.3% (95% CI: 3.9–12.3%, Figure 2D) and 2.6% (95% CI: 0.8–8.9%, Figure 2F) for 35-year-olds 255 in Hawaii and Guam, respectively. 256

Florida was unique in that 90.5% of cases had reported travel histories. To assess 257 258 local transmission within the state, we therefore only analyzed cases without reported 259 travel (Figure 2E). Local dengue cases were reported in multiple years and had a 260 distinctly different age profile for cases compared to other locations, with no reported cases in younger age groups and many among adults, which resulted in the largest 261 uncertainty for seroprevalence estimates, especially for older age groups (Figure 2E). 262 263 Considering all cases, the uncertainty for estimated seroprevalence in 35-year-olds was very high, with a 95% CI of 3.5% to 98.1%. Among local cases only, uncertainty 264 was still high, but indicated more confidence in lower seroprevalence for younger 265 266 adults, 9.6% (95% CI: 2.0-28.1%) for 35-year-olds. This lower seroprevalence estimate for local cases indicates a relatively low historical transmission intensity 267 268 locally, implying that many imported secondary cases likely also had their primary 269 infections outside of Florida. The long-term mean FOI for Florida considering only 270 local cases was 0.001% (95% CI: 0.0002-0.008%). Similar to Hawaii and Guam, 271 there was insufficient information to further resolve the probability of reporting a case and the estimated reporting probabilities resembled the prior. 272

Across the six locations, estimated long-term average FOI was highest for American
Samoa and Puerto Rico, followed by the US Virgin Islands, with much lower
estimates for Hawaii and Guam. We also assessed the average yearly FOI estimates
for each location and found that for the higher FOI locations (Puerto Rico, American
Samoa, and US Virgin Islands), the yearly average was lower for the period which
included surveillance data (2010–2019) compared to estimates for time periods prior

- to 2010, for which ArboNET data are not available. In contrast, the estimated FOI for
- the 2010–2019 period were higher than the historical period for Hawaii and Guam.

#### 281 **Discussion**

282 Passive surveillance is the most common approach for dengue surveillance globally [30–32]. While passive surveillance can be highly effective for monitoring longitudinal 283 changes in incidence, insight into disease burden is challenging as care seeking 284 285 behavior, access to care, case definitions, reporting, and resources can differ 286 substantially over time and across jurisdictions. Differences in these components can 287 mean that actual increases, decreases, or jurisdictional differences in dengue burden 288 are difficult to disentangle from reporting dynamics. The importance of understanding 289 dengue burden beyond reported incidence has grown in recent years, with the advent of dengue vaccines which can be used to reduce the burden, but also require 290 291 consideration of levels of preexisting population immunity for effective deployment 292 [33]. Here, we extended a previously developed Bayesian model to leverage the age distribution among cases reported via passive surveillance in the US to estimate 293 yearly DENV force of infection, yearly reporting probabilities, and seroprevalence for 294 295 six jurisdictions with very different histories of reported dengue incidence. Across the six jurisdictions, the estimated dengue burden was highly heterogeneous 296

Across the six jurisdictions, the estimated deligue burden was highly heterogeneous
and FOI estimates did not appear be higher over time, except in Hawaii and Guam.
We found evidence of high long-term transmission intensity in Puerto Rico, American
Samoa, and US Virgin Islands (FOI and high seroprevalence in adults), despite
different histories of dengue in each location. Hawaii, Florida, and Guam also differed
in the frequency and magnitude of reported outbreaks, but all had considerably lower
long-term FOIs.

The higher estimated FOI in Puerto Rico, American Samoa, and US Virgin Islands 303 304 reflect different dynamics. Puerto Rico has consistently reported dengue cases for at 305 least three decades ([34,35], timeline by [6]), while dengue reporting in American 306 Samoa and the US Virgin Islands is more sporadic. Outbreaks have been reported in American Samoa in the 1970s, 1990s, 2000s, and 2017–2018 [36–38] and in the US 307 Virgin Islands in the 1990s, 2000s, and in 2012-2013 [6,39]. Here, we found that 308 309 despite these different apparent dynamics, the estimated long-term FOI and seroprevalence in these locations were similar, with a high likelihood that individuals 310 experienced a secondary infection by the time they reached 10 years old (point 311 312 estimates between 42% and 59%). Three seroprevalence studies support these findings. First, a study in the municipality of Patillas in the southeastern part of Puerto 313 314 Rico [40] found a 42.5% seroprevalence in 10–11 year-olds in 2007–2008. For 10-315 year-olds, we estimated a seroprevalence of 57.2% (95% CI: 45.3–69.6%) in 2010 -316 the earliest estimate available for our study — in Puerto Rico. A 2010 seroprevalence study in American Samoa found a seroprevalence of 95.6% (95% CI 93.9–96.8%) for 317 318 18–87 year-olds, with slightly lower estimates for 18–25 year-olds (89.1%, 95% CI: 319 84.0–92.6%) [41]. For the same age groups and year, using only passive surveillance 320 data, we estimated seroprevalence of 95.9% (95% CI: 86.9–98.5%) and 86.9% (95% 321 CI: 63.6–95.6%) for these respective age groups. More recently, CDC's updated 322 traveler risk classifications re-categorized US Virgin Islands from "Sporadic/Uncertain" to "Frequent/Continuous" [42]. Additionally, a cross-sectional 323 seroprevalence study conducted in US Virgin Island schools in April 2022 estimated 324 325 a seroprevalence of 42% (95% CI: 26–60%) and 54% (95% CI: 18–89%) among 10 326 and 13 year-olds respectively [43], while our model estimated a seroprevalence of

34% (95% CI: 18–54%) and 41% (95% CI: 26–61%) in 7 and 13 year-olds for 2019,
3 years prior. No major outbreak has been reported in the US Virgin Islands between
2019 and 2022. The high level of immunity found in the US Virgin Islands with our
model confirms that dengue transmission intensity in this territory is likely higher than
previously suspected.

332 In Hawaii, low FOI and seroprevalence estimates are consistent with several possible 333 histories: small intermittent outbreaks, rare large epidemics, or continuous low 334 intensity transmission. No autochthonous cases had been reported in Hawaii 335 between 1944 and 2001, but recent outbreaks were reported in 2001-2002 [44], 2011, and 2015–2016 [45]. Given some reporting issues during these outbreaks (see 336 337 [46]), the seroprevalence profile in Hawaii should be interpreted with caution. 338 Although the results here do not differentiate between the possible histories, they do indicate that substantial and sustained undetected transmission in recent decades is 339 highly unlikely. Further serological data collection could help decipher between these 340 341 hypotheses.

342 In Florida, we found a low FOI and a low level of immunity, suggesting low circulation of the virus. Frequent introductions of dengue viruses by travelers result in frequent 343 344 opportunities for local transmission, so FOI estimates likely reflect both this relatively 345 high introduction risk and limited local transmission. The travel and migration history 346 of Florida residents may also contribute to a higher seroprevalence among older individuals than what would be expected from local transmission alone. This could 347 explain the higher case numbers in older adults, high uncertainty in seroprevalence 348 estimates for those adults, and the better fit of the secondary case model to the 349 350 Florida data. More detailed data on histories of travel and dengue exposure could

potentially provide more resolution on historical patterns of transmission in Florida as
compared to exposures acquired elsewhere that impact the patterns of dengue seen
in Florida today.

Our FOI and seroprevalence estimates in Guam were lower than for the other US territories with comparable environmental conditions. According to our model, primary infections, or a combination of primary and secondary infections dominated the 2019 outbreak in Guam, which is consistent with other studies that found no evidence of endemic circulation in Guam since World War II [38,47,48], potentially due to successful vector control strategies limiting dengue transmission on this territory [47].

360 Our model estimated time-varying reporting probabilities for all jurisdictions and found important variations across locations and years. In locations with lower long-term FOI 361 estimates, there was little information in the case data to update the broad prior 362 363 assumption for reporting probabilities of approximately 10-30%. However, for locations with higher long-term FOI, reporting was generally below 10% and showed 364 substantial year-to-year variability. Other recent work indicates the possibility of 365 366 similar variation between years [20,22] and particularly between epidemic and nonepidemic years [21], a pattern which we also found in the data and models analyzed 367 368 here. These findings indicate the likely importance of accounting for variability in 369 reporting probabilities over time.

We compared models accounting for reported cases as primary, secondary, or a combination of both primary and secondary infections, finding mixed evidence of fit across jurisdictions. For Florida, the secondary case model clearly fit better by leaveone-out metrics, while it was worse than the other models for Hawaii and Guam

(Figure 3). However, detailed examination of fitted model estimates and log-374 375 likelihoods revealed that models for secondary infections provided better qualitative 376 fits to the age profiles of cases in locations with higher seroprevalence and for years 377 and age groups with higher incidence. Thus, despite equivocal out-of-sample performance, we found that the secondary-only model is likely preferrable in locations 378 379 with high FOI or with high seroprevalence that may have been acquired elsewhere 380 (Florida), while the primary or combined models may perform better in low FOI settings with limited previous exposure. For reporting, this finding implies that most 381 cases reported in Puerto Rico, American Samoa, US Virgin Islands and Florida are 382 383 secondary infections and the majority of reported cases in Hawaii and Guam are likely primary infections. This confirms that in endemic settings primary infections are 384 385 largely unreported, but also suggests this pattern is slightly different in locations with 386 more sporadic transmission. This spatial and temporal variation in reporting is likely 387 to be a more general phenomenon.

388 While we were able to identify clear differences in transmission intensity across 389 locations, this analysis provides limited insight into why those differences exist. Different levels of mobility and interactions with mosquitos between individuals within 390 391 and between households likely lead to some differences. Environmental variables 392 (e.g., temperature and rainfall) also impact the mosquito population and may drive 393 differences in transmission intensity [49,50]. Furthermore, we accounted for previous 394 infections as a condition for reporting and a condition for the force of infection. 395 However, we did not consider how heterotypic and homologous immunity to different serotypes (or other flaviviruses) may substantially impact these dynamics. 396

Some limitations are worth noting. First, ArboNET includes dengue data since 2010 397 398 but in some jurisdictions few cases have been reported in this period, meaning that 399 for some jurisdictions the models rely on relatively few data. While fitting the data in a 400 Bayesian framework with informative priors reduces the possibility of misleading conclusions, it also means that high uncertainty remains around some important 401 402 parameters (e.g., reporting in some jurisdictions). Heterogeneity in dengue exposure 403 and disease across age and gender may result in both age-specific FOI and age-404 specific reporting probabilities. Neither of these possibilities are captured in our model implementation and would be difficult to assess without informative priors for 405 406 that age-specific variation. In early model development, we found that to account for time-variable reporting and FOI, we needed to have informative priors for both 407 408 components. Another possible area for advancing this type of model would be to 409 further differentiate reporting probabilities for primary and secondary infections (see 410 [14]). With informative priors, this could potentially provide further resolution on 411 location specific differences in transmission where reported cases may be dominated 412 by primary infections, secondary infections, or a mix.

Our study shows that analyzing age-specific case notification data with catalytic models can provide invaluable insight into dengue virus transmission dynamics beyond simple case counts, including estimates of transmission intensity over time and of evolving population-level exposure (as an ongoing proxy for seroprevalence). These components can help identify where interventions such as dengue vaccines should be prioritized in the US and elsewhere or what populations may be at higher risk of secondary infections and therefore severe disease. These models may

- 420 provide an important tool for assessing and monitoring dengue transmission risk in
- 421 many locations where age-specific surveillance data already exist.

# 422 Acknowledgments

- 423 This work was supported in part by an appointment to S.K. to the Research
- 424 Participation Program at the Centers for Disease Control and Prevention,
- 425 administered by the Oak Ridge Institute for Science and Education through an
- 426 interagency agreement between the US Department of Energy and CDC. Authors
- 427 would like to thank Aidsa Rivera for support over dataset requests.

# 428 **Disclaimer**

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

# 432 Code availability

- 433 Code to reproduce our analyses is available on GitHub
- 434 at https://github.com/sarahkada/dengue-catalytic-model.
- 435
- 436



438 Figure 1. Confirmed and probable dengue cases reported to ArboNET by state

439 or territory, 2010–2019. Included cases had no travel history and met laboratory

440 criteria for diagnosis. Different y-axis scales were used for each jurisdiction.

441







Figure 2. Age distribution of reported dengue cases and 2019 seroprevalence
estimates by age group for each US state and territory assessed. Age
distribution of all locally reported, confirmed and probable dengue cases in Puerto
Rico (PR, A), American Samoa (AS, B), US Virgin Islands (USVI, C), Hawaii (HI, D),
Florida (FL, E) and Guam (GU, F) in 2010–2019 (colored bar) and age distribution of
the population in these respective territories (grey bars, 2010 census) (left panel) and

- 457 (reported cases are secondary infections only), with corresponding 95% confidence
- intervals (shaded area, right panel). Florida (E) contains the additional 458
- seroprevalence estimates from all cases (imported and local). Detailed tables on 459
- 460 years with reported dengue cases are available in S1. Y-axis scale on the left panels differ for each location.
- 461
- 462
- 463



#### Figure 3. Leave-one-out cross validation model comparison using the 465

- estimated difference in ELPD (expected log pointwise predictive probabilities). 466
- ELPD is a measure of out-of-sample predictive accuracy, as estimated by the 467
- Bayesian leave-one-out cross validation (LOO). We compared model P (cases as 468
- 469 primary) to models PS and S (primary and secondary cases and secondary cases
- only). Models under the horizontal bar at 0 fit the data better than model P 470
- (reference). Vertical bars represent the corresponding standard error of the difference 471 in ELPD. 472
- 473
- 474



Figure 4. Time-varying estimates of the force of infection (FOI) in Puerto Rico
(all cases) and severe cases only (PR severe), American Samoa, US Virgin
Islands, Hawaii, Florida local cases and Guam in 2010–2019 from our model S
(cases as secondary infections). The numbers in the upper left-hand corner of
each panel shows long term FOI estimates for each location. Dark and light shaded
areas represent respectively, 50% and 95% confidence intervals. Different y-axis
scales were used for each jurisdiction.



486

487 Figure 5. Estimates of the yearly probability of reporting cases in Puerto Rico

488 (all cases) and severe cases only (PR severe), American Samoa, US Virgin

489 Islands, Hawaii, Florida and Guam in 2010–2019 from our S model (cases are

490 **considered secondary infections).** Vertical bars represent 95% confidence

intervals. For the severe cases, it is the proportion of all secondary infections that areboth severe and reported.



*dengue cases and severe dengue cases only.* Derived from model S where
499 reported cases are secondary infections only. Shaded areas represent 95%
500 confidence intervals.

# 502 Supporting information Captions

503 S1 Figure. Prior distribution of the reporting hyperprior. Dashed lines represent
 504 median prior value, at 10%.

505 S2 Figure. Age distribution of reported severe and non-severe dengue cases

- (colored bars) in Puerto Rico from 2010 to 2019. The grey bars represent the age
   distribution of the population (US census, 2010).
- 508 S3 Figure. Yearly age distribution of dengue reported cases in Puerto Rico (A),
- 509 American Samoa (B), US Virgin Islands (USVI, C), Hawaii (D), Florida (E), and
- 510 **Guam (F).** Scales may be adjusted for years with fewer dengue cases reported.
- 511 **S4 Figure. Yearly age distribution of locally-acquired and imported dengue**
- 512 **reported cases in Florida.** Scales are adjusted for years with fewer dengue cases 513 reported.
- 514 **S5 Figure. Comparison of log likelihood samples by age group and years of**
- 515 models 1-3 using Puerto Rico data. Comparison of log likelihood samples by

age group and years of models Primary, Primary & Secondary, and Secondary

- 517 **using Puerto Rico data**. Models' log-likelihood values may overlap. The log
- 518 likelihood value in a model is a measure of goodness of fit. The higher the value (i.e.,
- 519 closer to 0), the better.
- 520 S6 Figure. Heatmap comparison between models Primary, Primary and
- 521 Secondary, and Secondary of log likelihood median samples and cases by age
- group and years using Puerto Rico data. The log likelihood value in a model is a
   measure of goodness of fit. Here, models with log likelihood values closest to 0 were
   plotted.
- 525 S7 Figure. Reported cases model fit (all cases and severe cases only) by age

526 group in Puerto Rico from 2010 to 2019. Points represent cases reported to

- 527 ArboNET while lines represent model fit to Primary, Primary & Secondary, and
- 528 Secondary models. Shaded areas represent the best model 95% Cls.
- 529

530 S8 Figure. Reported severe cases model fit (severe cases only) by age group in

- 531 **Puerto Rico from 2010 to 2019.** Points represent cases reported to ArboNET while
- 532 lines represent model fit to Primary, Primary & Secondary, and Secondary models.
- 533 Shaded areas represent the best model 95% Cls.
- 534 S9 Figure. Reported cases model fit by age group in American Samoa from
- 535 **2010 to 2019.** Points represent cases reported to ArboNET while lines represent
- 536 model fit to Primary, Primary & Secondary, and Secondary models.
- 537 Shaded areas represent the best model 95% CIs.
- 538 S10 Figure. Reported cases model fit by age group in US Virgin Islands from
- 539 **2010 to 2019**. Points represent cases reported to ArboNET while lines represent
- 540 model fit to Primary, Primary & Secondary, and Secondary models.
- 541 Shaded areas represent the best model 95% CIs.

### 542 S11 Figure. Reported cases model fit by age group in Hawaii from 2010 to 2019.

543 Points represent cases reported to ArboNET while lines represent model fit to

- 544 Primary, Primary & Secondary, and Secondary models. Shaded areas represent the 545 best model 95% CIs.
- 546 S12 Figure. Reported cases model fit by age group in Florida from 2010 to
- 547 **2019.** Points represent cases reported to ArboNET while lines represent model fit to
   548 Primary, Primary & Secondary, and Secondary models. Shaded areas represent the
- 549 best model 95% Cls.
- 550 S13 Figure. Reported cases model fit by age group in Guam from 2010 to 2019.
- 551 Points represent cases reported to ArboNET while lines represent model fit to
- 552 Primary, Primary & Secondary, and Secondary models. Shaded areas represent the 553 best model 95% CIs.
- 554 S14 Alpha parameter posterior estimates with 95%Cl (vertical bars) in Puerto 555 Rico, Puerto Rico severe (using severe cases only), American Samoa, US
- 556 Virgin Islands, Hawaii, Florida and Guam.
- 557 **Table S1: Force of infection estimates in Puerto Rico, American Samoa, US**
- 558 Virgin Islands, Hawaii, Florida and Guam from 2010 to 2019. Median and 95%
- 559 Bayesian credible intervals.
- 560

# 561 **References**

562

Brady OJ, Hay SI. The global expansion of dengue: how *Aedes aegypti* mosquitoes
enabled the first pandemic arbovirus. Annual Review of Entomology. 2020;65: 191–208.
doi:10.1146/annurev-ento-011019-024918

- 566 2. Monath TP. Dengue: the risk to developed and developing countries. Proc Natl Acad
  567 Sci U S A. 1994;91: 2395–2400.
- Messina JP, Brady OJ, Golding N, Kraemer MUG, Wint GRW, Ray SE, et al. The
  current and future global distribution and population at risk of dengue. Nature Microbiology.
  2019;4: 1508–1515. doi:10.1038/s41564-019-0476-8
- 571 4. Brady OJ, Gething PW, Bhatt S, Messina JP, Brownstein JS, Hoen AG, et al. Refining
  572 the global spatial limits of dengue virus transmission by evidence-based consensus. PLoS
  573 Negl Trop Dis. 2012;6: 1–15. doi:10.1371/journal.pntd.0001760
- 5. Allicock OM, Lemey P, Tatem AJ, Pybus OG, Bennett SN, Mueller BA, et al.
  575 Phylogeography and population dynamics of dengue viruses in the Americas. Molecular
  576 biologue and exclusion 2012;20: 1522-1542 doi:10.1002/malbes/mar220
- 576 biology and evolution. 2012;29: 1533–1543. doi:10.1093/molbev/msr320
- 577 6. Añez G, Rios M. Dengue in the United States of America: a worsening scenario?
  578 BioMed Research International. 2013. p. 13.
- 579 7. Katzelnick LC, Bos S, Harris E. Protective and enhancing interactions among dengue
- viruses 1-4 and Zika virus. Current Opinion in Virology. 2020;43: 59–70.
- 581 doi:https://doi.org/10.1016/j.coviro.2020.08.006
- 582 8. Endy TP, Anderson KB, Nisalak A, Yoon I-K, Green S, Rothman AL, et al.
- 583 Determinants of inapparent and symptomatic dengue infection in a prospective study of
- 584 primary school children in Kamphaeng Phet, Thailand. PLoS Negl Trop Dis. 2011;5: e975.
- 585 9. Yoon I-K, Rothman AL, Tannitisupawong D, Srikiatkhachorn A, Jarman RG, Aldstadt
- J, et al. Underrecognized mildly symptomatic viremic dengue virus infections in rural Thai
   schools and villages. J Infect Dis. 2012;206: 389–398. doi:10.1093/infdis/jis357

588 10. Grange L, Simon-Loriere E, Sakuntabhai A, Gresh L, Paul R, Harris E. Epidemiological risk factors associated with high global frequency of inapparent dengue virus 589 590 infections. Front Immunol. 2014;5. Available: 591 https://www.frontiersin.org/articles/10.3389/fimmu.2014.00280 Clapham HE, Cummings DAT, Johansson MA. Immune status alters the probability of 592 11. 593 apparent illness due to dengue virus infection: Evidence from a pooled analysis across 594 multiple cohort and cluster studies. PLoS Negl Trop Dis. 2017;11: e0005926. 595 doi:doi:10.1371/journal.pntd.0005926 596 Rodriguez-Barraquer I, Salje H, Cummings DA. Opportunities for improved 12. 597 surveillance and control of dengue from age-specific case data. eLife. 2019;8: e45474. doi:10.7554/eLife.45474 598 599 13. Muench H. Catalytic models in epidemiology. Cambridge (MA): Harvard University 600 Press; 1959. doi:https://doi.org/10.4159/harvard.9780674428928 Imai N, Dorigatti I, Cauchemez S, Ferguson NM. Estimating Dengue Transmission 601 14. Intensity from Case-Notification Data from Multiple Countries. PLoS Negl Trop Dis. 602 603 2016;10: e0004833. doi:10.1371/ journal.pntd.0004833 604 O'Driscoll M, Imai N, Ferguson NM, Hadinegoro SR, Satari HI, Tam CC, et al. 15. 605 Spatiotemporal variability in dengue transmission intensity in Jakarta, Indonesia. PLoS Negl 606 Trop Dis. 2020;14: e0008102-. 607 16. CDC NNDSS. Dengue virus infections 2015 case definition. 2015 [cited 30 May 608 2022]. Available: https://ndc.services.cdc.gov/case-definitions/dengue-virus-infections-2015/ 609 17. CDC NNDSS. Dengue virus infections 2010 case definition. 2010 [cited 30 Nov 2022]. Available: https://ndc.services.cdc.gov/case-definitions/dengue-virus-infections-2010/ 610 Burke DS, Nisalak A, Johnson DE, Scott R. A prospective study of dengue infections 611 18. 612 in Bangkok. Am J Trop Med Hyg. 1988;38: 172–180. 19. Shankar MB, Rodríguez-Acosta RL, Sharp TM, Tomashek KM, Margolis HS, Meltzer 613 614 MI. Estimating dengue under-reporting in Puerto Rico using a multiplier model. PLoS Negl Trop Dis. 2018;12: e0006650. 615 Carabali M, Jaramillo-Ramirez GI, Rivera VA, Mina Possu N-J, Restrepo BN, Zinszer 616 20. K. Assessing the reporting of dengue, chikungunya and Zika to the national surveillance 617 system in Colombia from 2014–2017: a capture-recapture analysis accounting for 618 misclassification of arboviral diagnostics. PLoS Negl Trop Dis. 2021;15: e0009014. 619 Vong S, Goyet S, Ly S, Ngan C, Huy R, Duong V, et al. Under-recognition and 620 21. 621 reporting of dengue in Cambodia: a capture-recapture analysis of the National Dengue 622 Surveillance System. Epidemiol Infect. 2012;140: 491–499. 623 doi:10.1017/S0950268811001191 624 Silva MM, Rodrigues MS, Paploski IA, Kikuti M, Kasper AM, Cruz JS, et al. 22. Accuracy of dengue reporting by national surveillance system, Brazil. Emerging infectious 625 626 diseases. 2016;22: 336. 627 Montoya M, Gresh L, Mercado JC, Williams KL, Vargas MJ, Gutierrez G, et al. 23. Symptomatic versus inapparent outcome in repeat dengue virus infections is influenced by the 628 time interval between infections and study year. PLoS Negl Trop Dis. 2013;7: e2357. 629 630 doi:10.1371/journal.pntd.0002357 Halstead SB. Immunological parameters of togavirus disease syndromes. Academic 631 24. Press New York: 1980. 632 Reiner RC, Stoddard ST, Forshey BM, King AA, Ellis AM, Lloyd AL, et al. Time-633 25. 634 varying, serotype-specific force of infection of dengue virus. Proc Natl Acad Sci U S A. 635 2014;111: E2694-E2702. Hens N, Aerts M, Faes C, Shkedy Z, Lejeune O, Van Damme P, et al. Seventy-five 636 26. 637 years of estimating the force of infection from current status data. Epidemiol Infect. 2010;138:

- **638** 802–812.
- 639 27. Vehtari A, Gelman A, Gabry J. Practical Bayesian model evaluation using leave-one640 out cross-validation and WAIC. Statistics and Computing. 2017;27: 1413–1432.
- 641 doi:10.1007/s11222-016-9696-4
- 542 28. Stan Development Team. RStan: the R interface to Stan. 2020. Available: http://mc 643 stan.org/
- 644 29. R Core Team. R: A Language and Environment for Statistical Computing. Vienna,
- 645 Austria: R Foundation for Statistical Computing; 2020. Available: https://www.R-project.org/
- 646 30. PAHO. Dengue and dengue hemorrhagic fever in the Americas. 1997 [cited 19 May
- 647 2022]. Available: https://www.paho.org/en/documents/dengue-and-dengue-hemorrhagic 648 fever-americas-guidelines-prevention-and-control-1997
- 649 31. Runge-Ranzinger S, Horstick O, Marx M, Kroeger A. What does dengue disease
  650 surveillance contribute to predicting and detecting outbreaks and describing trends? Trop Med
  651 Int Health. 2008;13: 1022–1041.
- 652 32. Beatty ME, Stone A, Fitzsimons DW, Hanna JN, Lam SK, Vong S, et al. Best
- 653 practices in dengue surveillance: a report from the Asia-Pacific and Americas Dengue
- 654 Prevention Boards. PLoS Negl Trop Dis. 2010;4: e890. doi:10.1371/journal.pntd.0000890
- 655 33. World Health Organisation. Dengue vaccine: WHO position paper, September 2018 656 Recommendations. Vaccine. 2019;37: 4848–4849.
- 657 34. Dietz V, Gubler DJ, Ortiz S, Kuno G, Casta-Vélez A, Sather GE, et al. The 1986
  658 dengue and dengue hemorrhagic fever epidemic in Puerto Rico: epidemiologic and clinical
  659 observations. P R Health Sci J. 1996;15: 201–210.
- 660 35. Rigau-Pérez JG, Vorndam AV, Clark GG. The dengue and dengue hemorrhagic fever 661 epidemic in Puerto Rico, 1994-1995. Am J Trop Med Hyg. 2001;64: 67–74.
- 662 36. Cotter CJ, Tufa AJ, Johnson S, Matai'a M, Sciulli R, Ryff KR, et al. Outbreak of
  663 Dengue Virus Type 2 American Samoa, November 2016-October 2018. MMWR. 2018;67:
  664 1319–1322. doi:10.15585/mmwr.mm6747a5
- 665 37. Arima Y, Matsui T. Epidemiologic update of dengue in the Western Pacific Region,
- 666 2010. Western Pac Surveill Response J. 2011;2: 4–8. doi:10.5365/WPSAR.2011.2.2.005
- 667 38. Kiedrzynski T, Souares Y, Stewart T. Dengue in the Pacific: an updated story. Pacific
  668 Health Dialog. 1998;5: 129–136.
- 39. Mohammed H, Ramos M, Armstrong J, Muñoz-Jordán J, Arnold-Lewis KO, Ayala A,
  et al. An outbreak of dengue fever in St. Croix (US Virgin Islands), 2005. PLoS One. 2010;5:
  e13729.
- 40. Argüello DF, Tomashek KM, Quiñones L, Beltran M, Acosta L, Santiago LM, et al.
- 673 Incidence of dengue virus infection in school-aged children in Puerto Rico: a prospective
- 674 seroepidemiologic study. Am J Trop Med Hyg. 2015;92: 486–491. doi:10.4269/ajtmh.14675 0231
- 41. Duncombe J, Lau C, Weinstein P, Aaskov J, Rourke M, Grant R, et al. Seroprevalence
  of dengue in American Samoa, 2010. Emerg Infect Dis. 2013;19: 324–326.
- 678 doi:10.3201/eid1902.120464
- 42. Jentes ES, Lash RR, Johansson MA, Sharp TM, Henry R, Brady OJ, et al. Evidencebased risk assessment and communication: a new global dengue-risk map for travellers and
  clinicians. Journal of travel medicine. 2016;23: taw062.
- 43. Mac VV, Wong JM, Volkman HR, Perez-Padilla J, Wakeman B, Delorey M, et al.
- 683 Seroprevalence of previous dengue virus infection among children aged 8–13 Years —
- 684 U.Sjentes. Virgin Islands, April–May 2022. MMWR. : Forthcoming.
- 685 44. Effler PV, Pang L, Kitsutani P, Vorndam V, Nakata M, Ayers T, et al. Dengue fever,
- 686 Hawaii, 2001-2002. Emerg Infect Dis. 2005;11: 742–749. doi:10.3201/eid1105.041063
- 687 45. Lew RJ, Tsai W-Y, Wang W-K. Dengue outbreaks in Hawai'i After WWII A

- 688 Review of Public Health Response and Scientific Literature. Hawai'i Journal of Medicine &
- 689 Public Health. 2018;77: 315–318.
- 690 46. Adalja AA, Sell TK, Bouri N, Franco C. Lessons learned during dengue outbreaks in
- the United States, 2001-2011. Emerging infectious diseases. 2012;18: 608–614.
- 692 doi:10.3201/eid1804.110968
- 693 47. Lambrechts L, Scott TW, Gubler DJ. Consequences of the expanding global
- distribution of *Aedes albopictus* for dengue virus transmission. PLoS Negl Trop Dis. 2010;4:
   e646. doi:doi:10.1371/journal.pntd.0000646
- 696 48. Feldstein LR, Brownstein JS, Brady OJ, Hay SI, Johansson MA. Dengue on islands: a
- Bayesian approach to understanding the global ecology of dengue viruses. Trans R Soc Trop
- 698 Med Hyg. 2015;109: 303–312. doi:10.1093/trstmh/trv012
- 49. Johansson MA, Dominici F, Glass GE. Local and global effects of climate on dengue
   transmission in Puerto Rico. PLoS Negl Trop Dis. 2009;3: e382.
- 50. Mordecai EA, Cohen JM, Evans MV, Gudapati P, Johnson LR, Lippi CA, et al.
- 702 Detecting the impact of temperature on transmission of Zika, dengue, and chikungunya using
- 703 mechanistic models. PLoS Negl Trop Dis. 2017;11: e0005568-.
- 704 doi:doi:10.1371/journal.pntd.0005568
- 705