

# 1 Hepatitis E in Bangladesh: Insights from a National Serosurvey

2 Andrew S Azman<sup>1\*</sup>, Kishor Kumar Paul<sup>2</sup>, Taufiqur Rahman Bhuiyan<sup>2</sup>, Aybüke Koyuncu<sup>1</sup>, Henrik Salje<sup>3,1</sup>,  
3 Firdausi Qadri<sup>2</sup>, Emily S Gurley<sup>1\*</sup>

4 <sup>1</sup> Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, USA

5 <sup>2</sup> icddr, Dhaka, Bangladesh

6 <sup>3</sup> University of Cambridge, Cambridge, UK

7

## 8 Corresponding authors

9 Andrew S Azman

10 [azman@jhu.edu](mailto:azman@jhu.edu)

11

12 Emily Gurley

13 [egurley1@jhu.edu](mailto:egurley1@jhu.edu)

14

15

## 16 **Abstract**

### 17 **Background**

18 Hepatitis E virus, typically genotypes 1 and 2, is a major cause of avoidable morbidity and mortality in  
19 South Asia. Although case fatality risk among pregnant women can reach as high as 25%, a lack of  
20 population-level disease burden data has been cited as a primary factor in key global policy  
21 recommendations against the routine use of licensed hepatitis E vaccines, one of the only effective  
22 tools available for preventing disease and death.

### 23 **Methods**

24 We tested serum from a nationally-representative serosurvey in Bangladesh for anti-HEV IgG. We  
25 estimated the proportion of the population with evidence of historical HEV infection and used  
26 Bayesian geostatistical models to generate high-resolution national maps of seropositivity. We  
27 examined variability in seropositivity by individual-level, household-level, and community-level risk  
28 factors using spatial logistic regression.

### 29 **Results**

30 We tested serum samples from 2924 individuals from 70 communities representing all divisions of  
31 Bangladesh and estimated a national seroprevalence of hepatitis E of 20% (95% CI 17-24%).  
32 Seropositivity increased with age and male sex (OR: 2.2, 95% CI: 1.8–2.8). Community-level  
33 seroprevalence ranged from 0-78% with the seroprevalence in urban areas being higher, including  
34 Dhaka, the capital, with 3-fold (95%CrI 2.3-3.7) higher seroprevalence than the rest of the country.

### 35 **Conclusion**

36 Hepatitis E infections are common throughout Bangladesh, though 90% of women reach reproductive  
37 age without any evidence of previous exposure to the virus, thus likely susceptible to infection and  
38 disease. Strengthening clinical surveillance for hepatitis E, especially in urban areas may help  
39 generate additional evidence needed to appropriately target interventions like vaccines to the  
40 populations most likely to benefit.

### 41 **Keywords**

42 Hepatitis e, hepatitis E virus (HEV), seroprevalence

## 18 Background

19 Hepatitis E is estimated to cause over 3 million cases of acute jaundice each year with more than  
20 70,000 these leading to death and another 3000 still births [1]. Human infections from hepatitis E  
21 viruses, part of the orthohepevirus genus, are caused by four main genotypes (genotypes 1-4), with  
22 only genotypes 1 and 2 known to cause epidemics. HEV genotypes 1 and 2 are associated with self-  
23 limiting acute jaundice in the majority of infections although special populations, like pregnant women,  
24 have particularly poor outcomes with case fatality risk as high as 65% [2,3].

25 While HEV was only identified in 1981 [4], retrospective analyses have identified a number of large  
26 outbreaks which occurred on the Indian subcontinent in the 1970s and 1980s, including India and  
27 Bangladesh [5,6]. In Bangladesh, hepatitis E is endemic with large outbreaks from time-to-time [7–9].  
28 Hepatitis E is the leading cause of acute jaundice in Bangladesh and may be responsible for up to  
29 25% of maternal mortality [7,8,10].

30 While there is no effective treatment for acute hepatitis E and emergency improvements in water and  
31 sanitation have often been unsuccessful in curbing transmission [11], public health workers have very  
32 few effective tools at hand to mitigate the burden of outbreaks. Fortunately, a safe and efficacious  
33 vaccine is licensed in China and Pakistan and efforts are underway for licensure in other countries  
34 and WHO-prequalification [12]. A phase 4 clinical trial is on-going in Bangladesh [13], but no large-  
35 scale vaccination or other HEV-specific prevention efforts are planned, in-part due to our poor  
36 understanding of the burden and geographic distribution of the disease [14].

37 Following infection with hepatitis E, individuals develop long-lasting antibodies that can be measured  
38 through serosurveys to provide detailed insights into the history of infection in a population.  
39 Serosurveys can help us understand the geographic distribution and magnitude of historical HEV  
40 infections, identify risk factors and estimate key epidemiologic parameters related to transmission.

71 Here we use a nationally representative serosurvey in Bangladesh to gain new insights into hepatitis  
72 E and provide critical details needed to target interventions, like vaccines, to areas at the highest risk.

## 73 Methods

### 74 Serosurvey Design

75 This survey was originally conducted as part of an arbovirus study in Bangladesh with two stage  
76 random sampling (community and household) previously described [15]. In brief, 70 communities  
77 from a total of 97,162 in the 2011 national census were selected with probability proportional to each  
78 community's population. In rural areas (around three-quarters of the Bangladeshi population), the  
79 smallest administrative unit is a village, whereas in urban areas they are wards. Within each village or  
80 ward, study staff identified the household where community leaders said the most recent wedding  
81 had taken place and selected the nearest neighbor. From this neighboring household study staff  
82 chose a random direction and counted six households along a transect in that direction in order to  
83 identify the first potential study household. For subsequent households, study staff chose a random  
84 direction and selected the sixth household from the previous household in that community. In each  
85 selected household, study staff identified the household head, described the study, and invited them  
86 to participate in the study. If the household head agreed to participate, all household members older  
87 than 6 months of age were invited to take part. Within each community, study staff visited households  
88 until the day when at least 10 households had been visited with at least 40 serum samples. Within  
89 each household study staff administered structured questionnaires with questions about household-  
90 level infrastructure, wealth and assets in addition to individual data on demographics and travel  
91 history as well as collecting ~5ml venous blood (~3ml from children  $\leq 3$  years) from all consenting  
92 individuals. Data for this survey were collected from 10/2015 through 01/2016.

04 The study was approved by the icddr,b ethics review board (protocol number PR-14058); this  
05 secondary analysis was reviewed and deemed exempt from review by the Johns Hopkins Bloomberg  
06 School of Public Health Institutional Review Board. All adult participants provided written informed  
07 consent to participate in the study. Parents or guardians of all child participants provided written  
08 informed consent on their behalf.

## 09 Laboratory Methods

10 Serum samples were stored at icddr,b at  $-80^{\circ}\text{C}$  before testing. We tested the serum samples for the  
11 presence of anti-HEV IgG using the Wantai immunoassay kit (Wantai HEV IgG [WE-7296] ELISA kit;  
12 Wantai Biological, Beijing, China) following manufacturer's instructions. As suggested in the packet  
13 insert, samples with a standardized optical density  $> 1.1$  were considered positive, those  $< 0.9$  were  
14 considered negative, and those in the range  $0.9\text{--}1.1$  were considered indeterminate.

## 15 Statistical Analyses

16 We estimated the national seroprevalence by including survey design weights and post-stratifying by  
17 age and sex to the 2011 Bangladesh census, with confidence intervals estimated using the Rao-Scott  
18 method implemented in the survey package for R [16,17]. We used the same approach to estimating  
19 seroprevalence by urban/rural locations, sex and age (only post-stratifying by sex for age-group-  
20 specific estimates). We excluded indeterminate results from all primary analyses.

21 We explored the relationship between individual, household and community-level factors and  
22 seropositivity using hierarchical logistic regression models including a spatial random field assuming  
23 a Matern covariance structure using integrated nested Laplace approximations (INLA) as  
24 implemented in the R-INLA package [18]. All individual and household-level data were collected from  
25 the survey questionnaire. Community-level data for population density [19], travel time to the nearest  
26 city [20], distance to a major water body, altitude, and poverty index [21] were collected from publicly

17 available data sources. In the main analyses we included household and community random effects  
18 in addition to the spatial random field, but in sensitivity analyses estimated models with different  
19 combinations of random effects to understand variability in our estimates (Supplement). We explored  
20 univariate models, a fully saturated model, a model with only variables significant in the univariate  
21 analyses and two simplified models selected *a priori* and compared their fit with Wanatabe-Akaike  
22 information criterion (WAIC) [22].

23 Using the same INLA modeling framework we estimated seroprevalence on a 5 km by 5 km grid  
24 across Bangladesh. To do this we assigned each community to a grid-cell by its centroid, estimated  
25 the mean seroprevalence in each cell containing observations and fit spatial regression models to  
26 these data. We then used these fitted models to predict seroprevalence in the unobserved grid-cells.  
27 We fit both a fully saturated model, including population size, distance from a major water body, a  
28 poverty index, travel time to the nearest city, and altitude as linear predictors in addition to spatial  
29 random effects using a Matern covariance structure and a null model with only the spatial random  
30 effects. To quantify the out of sample performance of this approach we used leave-one-community  
31 out cross validation and compared predictions to a 'naive' model that predicted the mean grid-cell  
32 seroprevalence for all but the held-out cells and calculated the mean absolute error.

33 We predicted seropositivity among girls reaching child-bearing age (15 years) by fitting generalized  
34 additive models with penalized cubic splines to age-seroprevalence curves in each first-level  
35 administrative unit (division). We estimated simultaneous 95% confidence intervals by re-sampling  
36 from estimates of the variance-covariance matrix of the fitted model using a simulation-based  
37 approach with 1,000 draws [23].

38  
39 We used the GADM 3.6 spatial database for all administrative boundaries, which does not include  
40 boundary changes made after September, 2015. All analyses were performed in R (version 4.0.2).

Data and source code to reproduce analyses are available at <https://github.com/HopkinsIDD/hepE-bangladesh-national-serosurvey>.

## Results

We tested 2924 individuals from 707 households and 70 communities representing all first-level administrative units (divisions) of Bangladesh. Sampled individuals had a similar age and sex distribution to the population of Bangladesh with the exception of young children, who were under-represented [24].

Overall, 20.9% (610) of individuals tested positive for anti-HEV IgG, 78.8% (2305) were negative and 0.3% (9) were indeterminate. After taking into account the survey design and adjusting for imbalances between the sampled population and that of Bangladesh, we estimated a national seroprevalence of 20.0% (95% Confidence Interval [CI] 16.5-23.9%; design effect = 6.4).

Seroprevalence increased with age, from 2.5% (95%CI 0.6-9.7) in those under 5 years old and reaching a maximum of 40.9% (95%CI 26.0-57.8) among those 70-74 years old. On average, men (24.3%; 95%CI 20.5-28.5) had higher seroprevalence than women (15.9%; 95%CI 12.4-20.3) with this pattern holding across most age groups (Figure 1). Nationally, 90.4% (95%CI 88.4-92.1%) of girls reach the age of 15 (roughly “reproductive age”) without evidence of antibodies and likely susceptible to disease [25].

Seroprevalence varied greatly across communities from 0% to 77.5% across the country (Figure 1). Urban areas (29.7%; 95%CI 21.4-39.6) had 1.8 times higher seroprevalence than rural areas (16.8%; 95%CI 13.5-20.6). Household-level seroprevalence ranged from 0 to 100% and we found no evidence of increasing household seroprevalence with household size.

34 While age, sex and living in an urban area were associated with the risk of being seropositive, other  
 35 community and household-level risk factors may also be important. To explore the relationship  
 36 between these potential risk factors we used a series of univariable and multivariable logistic  
 37 regression models with spatially correlated errors. In univariate models (Table 1), we found significant  
 38 positive associations with age, being male, travel, urbanicity, population density, a community poverty  
 39 index, and protective effects of various indicators of socioeconomic status (e.g., having completed  
 40 primary school compared to having no formal education, having cattle or other animals in the  
 41 household, being a household owner). However, in our primary multivariable model with spatial  
 42 random effects, none of these factors were independently associated with seropositivity except for  
 43 age and sex, though the effect sizes were largely consistent across various models considered (Table  
 44 1 and Supplementary Table 1). Estimates from models with varying assumptions about random  
 45 effects yielded qualitatively similar results.

46 **Table 1. Estimated odds ratios and 95% credible intervals for seropositivity including random**  
 47 **effects for household, community and a spatial random field.**

	Participants (n=2896*)	Univariate model, unadjusted odds ratio (95% CI)	Full model, adjusted odds ratio (95% CI)
<b>Individual-level factors</b>			
Age, years			
0-4 years	131 (0.05)	0.45 (0.14-1.21)	0.44 (0.13-1.24)
5-14 years	661 (0.23)	1.0 (ref)	1.0 (ref)
≥15 years	2104 (0.73)	6.54 (4.67-9.27)	8.73 (6.03-12.81)
Sex			
Male	1385 (0.48)	1.82 (1.50-2.22)	2.19 (1.75-2.76)
Female	1511 (0.52)	1.0 (ref)	1.0 (ref)
Travel history			
No travel in last 6 months	1217 (0.42)	1.0 (ref)	1.0 (ref)

Travel in past week	443 (0.15)	1.75 (1.29-2.38)	1.11 (0.76-1.60)
Travel in past month	589 (0.20)	1.77 (1.35-2.31)	1.11 (0.81-1.53)
Travel in past 6 months	647 (0.22)	1.45 (1.11-1.88)	1.12 (0.83-1.50)

### Household-level factors

#### Household income per month, US\$<sup>†</sup>

<90	308 (0.11)	0.96 (0.67-1.37)	0.91 (0.56-1.47)
91-130	531 (0.18)	1.10 (0.82-1.47)	1.12 (0.77-1.62)
131-261	1094 (0.38)	1.0 (ref)	1.0 (ref)
>261	963 (0.33)	0.97 (0.76-1.24)	0.92 (0.67-1.27)

#### Education (head of household)

No school	897 (0.31)	1.0 (ref)	1.0 (ref)
Primary school	744 (0.26)	0.72 (0.55-0.95)	0.73 (0.51-1.04)
Secondary school	791 (0.27)	0.89 (0.69-1.16)	0.84 (0.60-1.19)
Post-secondary education	464 (0.16)	0.83 (0.61-1.13)	0.79 (0.52-1.20)

Electricity in house	2624 (0.91)	1.18 (0.78-1.81)	1.47 (0.85-2.57)
Owns land	2309 (0.80)	0.94 (0.73-1.22)	0.90 (0.63-1.27)
Owns home	2713 (0.94)	0.61 (0.40-0.94)	0.66 (0.36-1.22)
Owns animals			
Pigs or rabbits	20 (0.01)	0.44 (0.11-1.44)	0.35 (0.06-1.68)
Other animals	2524 (0.87)	0.68 (0.48-0.96)	0.77 (0.48-1.24)

### Community-level factors

Urban	733 (0.25)	2.13 (1.46-3.12)	1.69 (0.88-3.24)
Distance to major water body, per 10km	1.03 (1.58)	0.97 (0.86-1.11)	0.97 (0.81-1.16)
Poverty index	-0.12 (0.60)	1.75 (1.19-2.58)	1.35 (0.55-3.26)
Travel time to nearest city, min	12.42 (14.22)	0.99 (0.98-1.01)	1.01 (0.99-1.03)
Altitude, meters	16.56 (16.05)	0.98 (0.95-1.00)	0.98 (0.95-1.02)
Population density, log	10.72 (1.19)	1.25 (1.05-1.50)	1.09 (0.75-1.58)

Data are n (%) or mean (SD), unless otherwise indicated. The full model includes all covariates shown in the table,

random effects for household and community, in addition to a Matern spatial correlation function. \*Patients with complete

data for all variables. †Categories in Bangladesh Taka (TK) are <7000, 7000–9999, 10000–20000, and >20000;

TK77.6=US\$1 (June, 2015).

32 We fitted Bayesian geostatistical models to make a national map of seroprevalence. Our primary  
33 model demonstrated out-of-sample predictive skill (mean absolute error = 10.6%) with little bias ( $1.85$   
34  $\times 10^{-4}$ ) and moderate correlation of predictions with the true values in cross-validation (Pearson's  
35 correlation) of 0.51 (Figure 2). The seroprevalence map reveals large heterogeneity in  
36 seroprevalence across the country with the highest seroprevalence around Dhaka with some  
37 evidence of higher-than-average risk in two other large cities, Chittagong and Rajshahi. Similar to the  
38 non-spatial analyses, we estimate from these maps that 21.6% (95% CI 19.0-24.3%) of the  
39 population has been infected during their lifetime (35,177,057 people, 95%CrI 30,919,165-  
40 39,527,957). Residents of Dhaka have a 3.0-fold (95%CrI 2.3-3.7) higher seroprevalence than the  
41 mean seroprevalence of the rest of Bangladesh. Alternative models including different combinations  
42 of random effects (household, community and spatial) and covariates led to similar maps.

## 43 Discussion

44 In this nationally representative serosurvey we found that one in five people in Bangladesh had  
45 evidence of prior HEV infection with men having more than 1.5 times higher risk than women.  
46 Seroprevalence was 3 times higher in Dhaka, Bangladesh's capital and largest city than the rest of  
47 the country. Given the lack of specificity of clinical case definitions for hepatitis E (i.e., acute jaundice  
48 syndrome) and limited laboratory diagnostic use across Bangladesh, our approach and results  
49 highlight an important avenue for understanding risk across the country with an aim of targeting  
50 surveillance, prevention and control activities.

51  
52 Although faecal contamination of drinking water is likely the predominant cause of HEV infections in  
53 LMICs [26], measures of socioeconomic status typically associated with household access to clean  
54 water and sanitation (e.g. household income, education level) were not significantly predictive of  
55 seropositivity in our study. Our data are from a study not originally designed to study hepatitis E,

therefore, we did not have data on household water sources and sanitation. Despite this, our findings are consistent with existing literature on risk factors for HEV infection in Bangladesh [27,28] and documented large outbreaks of hepatitis E associated with contaminated municipal water supplies in urban areas of Bangladesh [8]. The higher seroprevalence among males may be due to exposures outside of the home, given their propensity to leave home more often than women in Bangladesh. If this hypothesis is correct, household-level water and sanitation interventions alone may not be sufficient to interrupt HEV transmission. Water and sanitation interventions may also have limited utility in preventing sporadic acute hepatitis cases associated with exposure to blood and animals, which are hypothesized to contribute to the burden of hepatitis E in Bangladesh [29] and other LMICs [30].

Samples from this same serosurvey were previously used to map the annual risk of *Vibrio cholerae* O1 infections across the country [31]. While both *Vibrio cholerae* and hepatitis E are transmitted through fecal contamination of drinking water and food, the spatial distribution of risk of these infections are completely different in Bangladesh. For example, while many *V. cholerae* infections were estimated to occur in Dhaka, inhabitants had lower than average risk overall. In contrast, inhabitants of Dhaka had 3 times higher risk of HEV seropositivity than others in the country. Some of the differences in spatial risk profiles may be due to the fact that cholera estimates capture only a snapshot of transmission (one year) compared to the lifetime exposures captured by HEV antibodies. Additionally, men had significantly higher HEV seroprevalence though we found no significant difference by sex for cholera. These differences might be in part explained by HEV transmission being facilitated through urban water infrastructure (e.g., [8]) and cholera transmission occurring more broadly through fecal oral routes inside households [32].

30 This study comes with a number of limitations. We assumed that these serologic assays had perfect  
31 sensitivity and specificity for detecting historical HEV infections. Previous studies have estimated  
32 sensitivity and specificity of these assays to be high [33,34] however, without a gold standard assay  
33 to compare against these estimates are unlikely to be perfect nor generalisable to all settings.  
34 Furthermore, seropositivity has been shown to decay over time so those infected many years before  
35 the serosurvey may be differentially misclassified as seronegative [35,36]. That seroprevalence does  
36 not significantly increase past the age of ~30 years old is likely due to a combination of seroreversion  
37 and changes in the force of infection over the decades. Future work synthesizing assay validation  
38 data may be valuable for correcting seroprevalence estimates appropriately. As there is only one  
39 HEV serotype, our estimates likely include immunologically meaningful exposures not only to  
40 genotypes 1 and 2, the most concerning for outbreaks, but other genotypes, which have not been  
41 widely described in Bangladesh and lead to different, though still severe, clinical outcomes. While we  
42 present smoothed estimates of the seroprevalence throughout the country, these are based on a  
43 geostatistical model fit to data from only 70 sampled communities throughout the country. These  
44 models assume that risk varies smoothly across space after taking into account covariates though  
45 true HEV risk is likely less smooth over space. Finally, the household sampling approach used to  
46 recruit individuals in the parent study may have systematically excluded migrant populations and  
47 individuals living in informal settlements more likely to have inadequate WASH and more likely to be  
48 seropositive.

49  
50 Our estimates provide a snapshot of cumulative infection risk in Bangladesh in 2015/2016. While this  
51 is useful to understand large-scale differences in risk across the country, it masks important  
52 differences in risk over time and space. The age-stratified patterns of seroprevalence, and in  
53 particular the changes in seroprevalence among the youngest children can be particularly informative  
54 for understanding recent infection risk (i.e., force of infection), which may be more important for

55 guiding policy. Our sample size in each sampled village of children under 5 years old was too small to  
56 permit detailed age-stratified analysis in these young age groups though future serosurveys may  
57 benefit greatly from these individuals. Furthermore, longitudinal or repeated cross-sectional  
58 serosurveys can allow for estimates of seroincidence [28]. Estimates of the contemporary force of  
59 infection can be combined with data on the proportion of infections that become clinically apparent  
60 (and severe) to help estimate the burden of hepatitis E [1].

61  
62 Given that more than 90% of girls reaching child-bearing age are susceptible to infection and the high  
63 case fatality risk associated with hepatitis E in pregnant women [8,26], this group may benefit greatly  
64 from hepatitis E vaccine [37] and existing maternal and child health programs in LMICs such as  
65 Bangladesh may be a cost-effective platform to facilitate delivery. The World Health Organization has  
66 not recommended routine use of this vaccine due to a lack of data on the safety, immunogenicity and  
67 efficacy of the vaccine among pregnant women as well as a lack of epidemiologic data on the burden  
68 of disease in the general population [12]. Fortunately, a clinical trial evaluating the safety,  
69 immunogenicity and effectiveness of hepatitis E vaccines among women of childbearing age,  
70 including those that go on to become pregnant, are underway [13] in Bangladesh. Data from our  
71 study suggest that these vulnerable individuals are at high exposure risk across the country and the  
72 use of safe and effective hepatitis E vaccines among pregnant women in Bangladesh may be  
73 justified.

74  
75 In this study we illustrate how remnant samples from population-based serologic studies not originally  
76 obtained to study hepatitis E can be an effective strategy to generate critical epidemiologic data in  
77 LMICs where surveillance infrastructure is weak or nonexistent. The SARS-CoV-2 pandemic presents  
78 an unprecedented opportunity to leverage the increased number of representative population-based  
79 surveys to improve our understanding of the global burden of hepatitis E [38,39]. Countries currently

30 planning serial cross-sectional serosurveys to monitor trends in SARS-CoV-2 transmission [40]  
31 should consider utilizing remnant samples to generate data that may help quantify hepatitis E risk  
32 over time and accelerate the use of the licensed vaccine.

### 33 **Funding statement**

34 The original data collection was supported by the Centers for Disease Control and Prevention  
35 (5U01GH001207-02). icddr,b acknowledges the commitment of the US CDC to its research efforts and thanks  
36 the Governments of Bangladesh, Canada, Sweden, and the UK for providing core or unrestricted support. US  
37 National Institutes for Health (NIH; R01 AI135115 ASA). Support to FQ and her laboratory also came from the  
38 NIH Fogarty International Center (TW005572) and Emerging Global Leader Award (K43 TW010362). The  
39 funders had no role in study design, data collection and interpretation, or the decision to submit the  
40 work for publication.

### 41 **Acknowledgement**

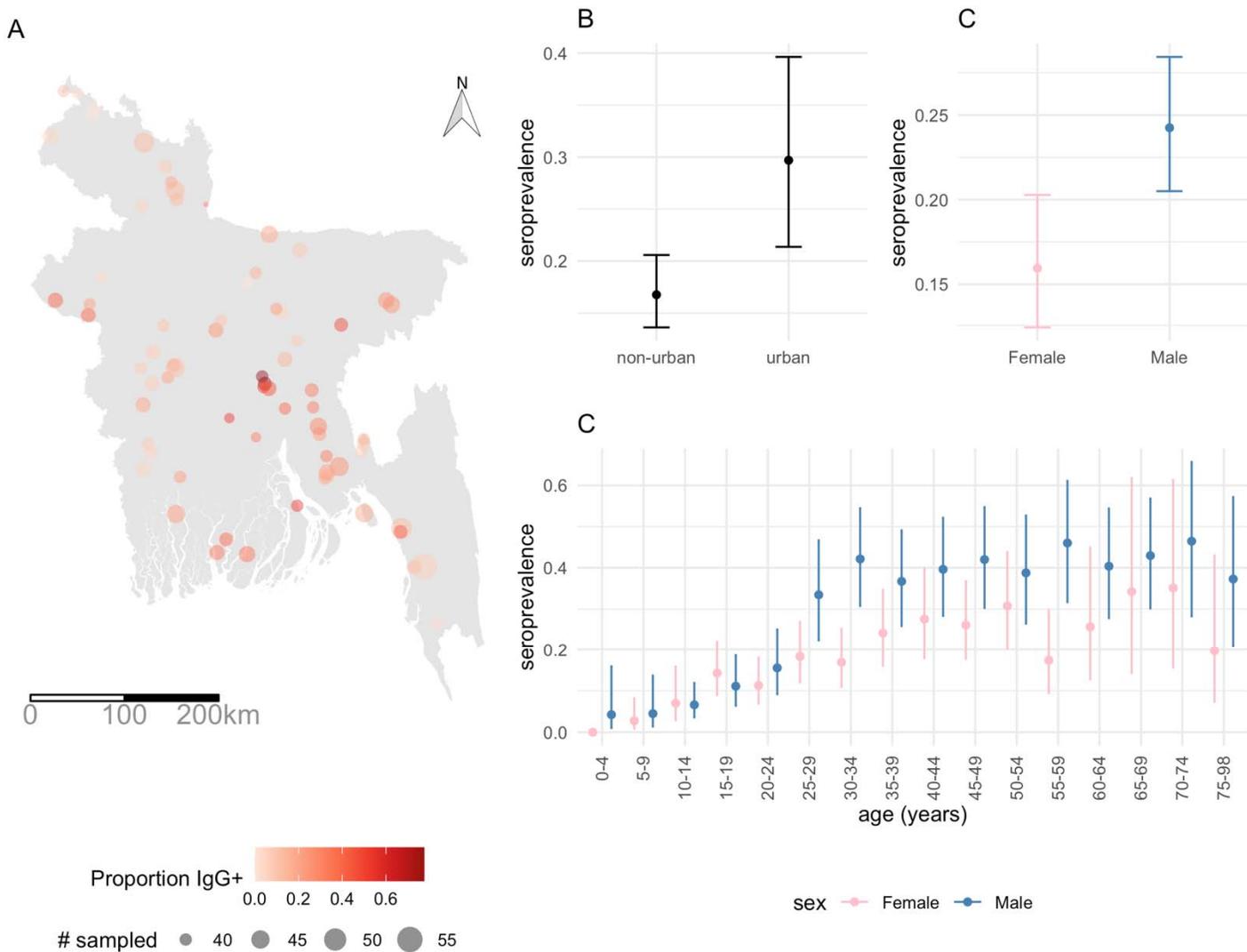
42 We thank all the participants of the study throughout Bangladesh for their willingness to participate in this  
43 research. We also wish to thanks the field and laboratory staff for their dedicated work.  
44

### 45 **Conflicts of interest**

46 The authors have no conflicts of interest to declare.

07

## 08 Figures



09

00 *Figure 1. Overview of sampled individuals and communities and seroprevalence estimates. Panel A illustrates the sampled community*  
01 *locations, number of individuals sampled (size of dots) and the proportion of individuals seropositive (color). Panel B illustrates the*  
02 *adjusted seroprevalence (age/sex) by urban/non-urban classifications and Panels C and D illustrate the sex (adjusted for age) and*  
03 *age/sex-specific seroprevalence.*

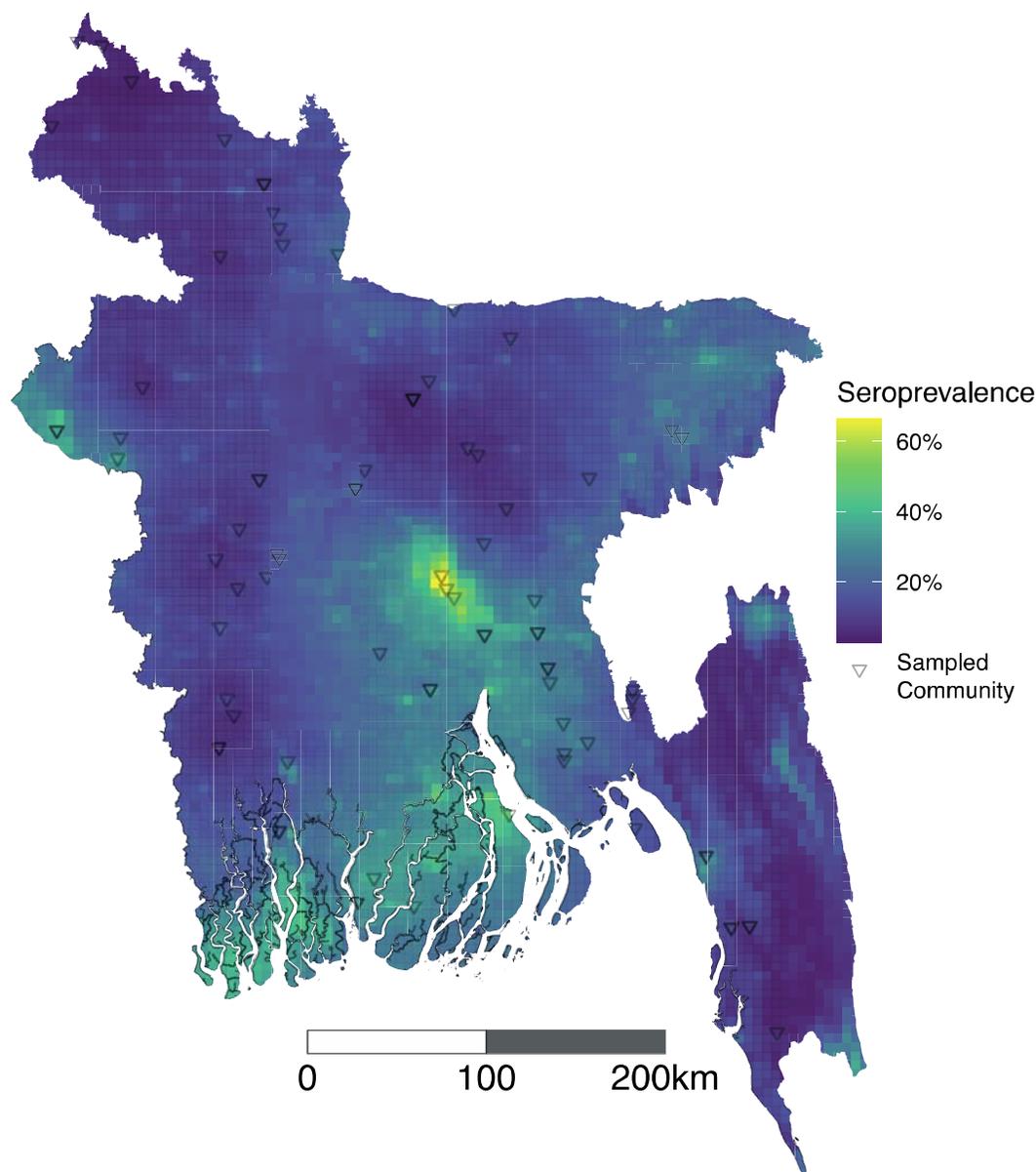


Figure 2. Predicted seroprevalence across Bangladesh, 2015. Predicted seroprevalence from geostatistical model with distance from major water body, population density, altitude, poverty index, and travel time to nearest city.

## References

1. Rein DB, Stevens GA, Theaker J, Wittenborn JS, Wiersma ST. The global burden of hepatitis E virus genotypes 1 and 2 in 2005. *Hepatology*. Wiley; **2012**; 55(4):988–997.
2. Kamar N, Bendall R, Legrand-Abravanel F, et al. Hepatitis E. *Lancet*. **2012**; 379(9835):2477–2488.
3. Amany G, Kizito S, Nabukenya I, et al. Risk factors, person, place and time characteristics

- 13 associated with Hepatitis E Virus outbreak in Napak District, Uganda. *BMC Infect Dis.* **2017**;  
14 17(1):451.
- 15 4. Balayan MS, Andjaparidze AG, Savinskaya SS, et al. Evidence for a virus in non-A, non-B  
16 hepatitis transmitted via the fecal-oral route. *Intervirology.* **1983**; 20(1):23–31.
- 17 5. Khuroo MS. Study of an epidemic of non-A, non-B hepatitis. Possibility of another human  
18 hepatitis virus distinct from post-transfusion non-A, non-B type. *Am J Med.* **1980**; 68(6):818–824.
- 19 6. Hlady WG, Islam MN, Wahab MA, Johnson SD, Waiz A, Krawczynski KZ. Enterically transmitted  
20 non-A, non-B hepatitis associated with an outbreak in Dhaka: epidemiology and public health  
21 implications. *Trop Doct.* **1990**; 20(1):15–17.
- 22 7. Paul RC, Nazneen A, Banik KC, et al. Hepatitis E as a cause of adult hospitalization in  
23 Bangladesh: Results from an acute jaundice surveillance study in six tertiary hospitals, 2014-  
24 2017. *PLoS Negl Trop Dis.* **2020**; 14(1):e0007586.
- 25 8. Gurley ES, Hossain MJ, Paul RC, et al. Outbreak of hepatitis E in urban Bangladesh resulting in  
26 maternal and perinatal mortality. *Clin Infect Dis.* **2014**; 59(5):658–665.
- 27 9. Biswas RSR, Sydu S, Mamun SMH, et al. Viral load and genotype of recent hepatitis E virus  
28 outbreak in Chittagong, Bangladesh. *Journal of the Scientific Society. Medknow Publications and*  
29 *Media Pvt. Ltd.;* **2020**; 47(3):144.
- 30 10. Gurley ES, Halder AK, Streatfield PK, et al. Estimating the burden of maternal and neonatal  
31 deaths associated with jaundice in Bangladesh: possible role of hepatitis E infection. *Am J Public*  
32 *Health.* **2012**; 102(12):2248–2254.
- 33 11. Guthmann J-P, Klovstad H, Boccia D, et al. A large outbreak of hepatitis E among a displaced  
34 population in Darfur, Sudan, 2004: the role of water treatment methods. *Clin Infect Dis.* **2006**;  
35 42(12):1685–1691.
- 36 12. World Health Organization. Recommendations to assure the quality, safety and efficacy of  
37 recombinant hepatitis E vaccines [Internet]. 2018. Available from:  
38 [https://www.who.int/biologicals/expert\\_committee/POST\\_ECBS\\_2018\\_Recommendations\\_HEP\\_](https://www.who.int/biologicals/expert_committee/POST_ECBS_2018_Recommendations_HEP_E_vaccines.pdf)  
39 [E\\_vaccines.pdf](https://www.who.int/biologicals/expert_committee/POST_ECBS_2018_Recommendations_HEP_E_vaccines.pdf)
- 40 13. Zaman K, Dudman S, Stene-Johansen K, et al. HEV study protocol : design of a cluster-  
41 randomised, blinded trial to assess the safety, immunogenicity and effectiveness of the hepatitis  
42 E vaccine HEV 239 (Hecolin) in women of childbearing age in rural Bangladesh. *BMJ Open.*  
43 **2020**; 10(1):e033702.
- 44 14. WHO. Hepatitis E vaccine: WHO position paper, May 2015--Recommendations. *Vaccine.* **2016**;  
45 34(3):304–305.
- 46 15. Salje H, Paul KK, Paul R, et al. Nationally-representative serostudy of dengue in Bangladesh  
47 allows generalizable disease burden estimates. *Elife* [Internet]. **2019**; 8. Available from:  
48 <http://dx.doi.org/10.7554/eLife.42869>
- 49 16. Rao JNK, Scott AJ. On Chi-Squared Tests for Multiway Contingency Tables with Cell Proportions  
50 Estimated from Survey Data [Internet]. *The Annals of Statistics.* 1984. Available from:

- 51 <http://dx.doi.org/10.1214/aos/1176346391>
- 52 17. Lumley T. *Complex Surveys: A Guide to Analysis Using R*. John Wiley & Sons; 2011.
- 53 18. Lindgren F, Rue H, Lindström J. An explicit link between Gaussian fields and Gaussian Markov  
54 random fields: the stochastic partial differential equation approach. *J R Stat Soc Series B Stat*  
55 *Methodol*. Wiley; **2011**; 73(4):423–498.
- 56 19. Tatem AJ. WorldPop, open data for spatial demography. *Sci Data*. **2017**; 4:170004.
- 57 20. Weiss DJ, Nelson A, Gibson HS, et al. A global map of travel time to cities to assess inequalities  
58 in accessibility in 2015. *Nature*. **2018**; 553(7688):333–336.
- 59 21. Steele JE, Sundsøy PR, Pezzulo C, et al. Mapping poverty using mobile phone and satellite  
60 data. *J R Soc Interface [Internet]*. **2017**; 14(127). Available from:  
61 <http://dx.doi.org/10.1098/rsif.2016.0690>
- 62 22. Watanabe K. An alternative view of variational Bayes and asymptotic approximations of free  
63 energy. *Mach Learn*. **2012**; 86(2):273–293.
- 64 23. Ruppert D, Wand MP, Carroll RJ. *Semiparametric Regression*. Cambridge University Press;  
65 2003.
- 66 24. Azman AS, Lauer S, Bhuiyan MTR, et al. *Vibrio cholerae* O1 transmission in Bangladesh:  
67 insights from a nationally-representative serosurvey. medRxiv [Internet]. Cold Spring Harbor  
68 Laboratory Press; **2020**; . Available from:  
69 <https://www.medrxiv.org/content/10.1101/2020.03.13.20035352v1.abstract>
- 70 25. Choi Y, Zhang X, Skinner B. Analysis of IgG Anti-HEV Antibody Protective Levels During  
71 Hepatitis E Virus Reinfection in Experimentally Infected Rhesus Macaques. *J Infect Dis*. **2019**;  
72 219(6):916–924.
- 73 26. Organization WH, Others. The global prevalence of hepatitis A virus infection and susceptibility:  
74 a systematic review. <https://apps.who.int/iris/handle/10665/70513>  
75 [Internet]. World Health Organization; **2010**; . Available from:  
76 <https://apps.who.int/iris/handle/10665/70513>
- 77 27. Labrique AB, Zaman K, Hossain Z, et al. Population seroprevalence of hepatitis E virus  
78 antibodies in rural Bangladesh. *Am J Trop Med Hyg*. **2009**; 81(5):875–881.
- 79 28. Labrique AB, Zaman K, Hossain Z, et al. Epidemiology and risk factors of incident hepatitis E  
80 virus infections in rural Bangladesh. *Am J Epidemiol*. **2010**; 172(8):952–961.
- 81 29. Sazzad HMS, Luby SP, Labrique AB, et al. Risk Factors Associated with Blood Exposure for  
82 Sporadic Hepatitis E in Dhaka, Bangladesh. *Am J Trop Med Hyg*. **2017**; 97(5):1437–1444.
- 83 30. Koyuncu A, Mapemba D, Ciglenecki I, Gurley ES, Azman AS. Setting a course for preventing  
84 hepatitis E in low and lower-middle-income countries: A systematic review of burden and risk  
85 factors. *Open Forum Infect Dis [Internet]*. Oxford University Press; **2021** [cited 2021 May 14]; .  
86 Available from: <https://academic.oup.com/ofid/advance-article/doi/10.1093/ofid/ofab178/6224343>
- 87 31. Azman AS, Lauer SA, Bhuiyan TR, et al. *Vibrio cholerae* O1 transmission in Bangladesh: insights

38 from a nationally representative serosurvey. *Lancet Microbe*. **2020**; 1(8):e336–e343.

- 39 32. Azman AS, Luquero FJ, Salje H, et al. Micro-Hotspots of Risk in Urban Cholera Epidemics. *J*  
40 *Infect Dis*. **2018**; 218(7):1164–1168.
- 41 33. Avellon A, Morago L, Garcia-Galera del Carmen M, Munoz M, Echevarría J-M. Comparative  
42 sensitivity of commercial tests for hepatitis E genotype 3 virus antibody detection. *J Med Virol*.  
43 **2015**; 87(11):1934–1939.
- 44 34. Kodani M, Kamili NA, Tejada-Strop A, et al. Variability in the performance characteristics of IgG  
45 anti-HEV assays and its impact on reliability of seroprevalence rates of hepatitis E. *J Med Virol*.  
46 **2017**; 89(6):1055–1061.
- 47 35. Kmush BL, Zaman K, Yunus M, Saha P, Nelson KE, Labrique AB. A Ten Year  
48 Immunopersistence Study of Hepatitis E Antibodies in Rural Bangladesh. *Am J Epidemiol*  
49 [Internet]. **2018**; . Available from: <http://dx.doi.org/10.1093/aje/kwy044>
- 50 36. Faber M, Willrich N, Schemmerer M, et al. Hepatitis E virus seroprevalence, seroincidence and  
51 seroreversion in the German adult population. *J Viral Hepat*. **2018**; 25(6):752–758.
- 52 37. Zhu F-C, Zhang J, Zhang X-F, et al. Efficacy and safety of a recombinant hepatitis E vaccine in  
53 healthy adults: a large-scale, randomised, double-blind placebo-controlled, phase 3 trial. *Lancet*.  
54 **2010**; 376(9744):895–902.
- 55 38. Chen X, Chen Z, Azman AS, et al. Serological evidence of human infection with SARS-CoV-2: a  
56 systematic review and meta-analysis. *The Lancet Global Health*. Elsevier; **2021**; 9(5):e598–e609.
- 57 39. Mina MJ, Metcalf CJE, McDermott AB, Douek DC, Farrar J, Grenfell BT. A Global Immunological  
58 Observatory to meet a time of pandemics. *Elife*. eLife Sciences Publications, Ltd; **2020**;  
59 9:e58989.
- 60 40. Kumar MS, Bhatnagar T, Manickam P, et al. National sero-surveillance to monitor the trend of  
61 SARS-CoV-2 infection transmission in India: Protocol for community-based surveillance. *Indian J*  
62 *Med Res*. **2020**; 151(5):419–423.