Epidemiology and Infection

www.cambridge.org/hyg

Original Paper

Cite this article: Dighe A, Khan AI, Bhuiyan TR, Islam MT, Khan ZH, Khan II, Hulse JD, Ahmed S, Rashid M, Hossain MZ, Rashid R, Hegde ST, Gurley ES, Qadri F and Azman AS (2024). Annual risk of hepatitis E virus infection and seroreversion: Insights from a serological cohort in Sitakunda, Bangladesh. *Epidemiology and Infection*, **152**, e52, 1–8
https://doi.org/10.1017/S0950268824000438

Received: 21 October 2023
Revised: 01 February 2024
Accepted: 05 March 2024

Keywords:
hepatitis e virus; serology; risk of infection; antibody waning

Corresponding author:
Amy Dighe;
Email: adighe1@jhmi.edu

Amy Dighe and Ashraful Islam Khan made equal contributions as first authors and Emily S. Gurley, Firdausi Qadri and Andrew S. Azman made equal contributions as last authors.

© The Author(s), 2024. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Annual risk of hepatitis E virus infection and seroreversion: Insights from a serological cohort in Sitakunda, Bangladesh

Amy Dighe[1], Ashraful Islam Khan[2], Taufiqur Rahman Bhuiyan[2], Md Taufiqul Islam[2], Zahid Hasan Khan[2], Ishtiakul Islam Khan[2], Juan Dent Hulse[1], Shakeel Ahmed[3], Mamunur Rashid[3], Md Zakir Hossain[3], Rumana Rashid[3], Sonia T. Hegde[1], Emily S. Gurley[1], Firdausi Qadri[2] and Andrew S. Azman[1,4,5]

[1]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA; [2]Infectious Disease Division, icddr, b, Dhaka, Bangladesh; [3]Bangladesh Institute of Tropical and Infectious Diseases, Chattogram, Bangladesh; [4]Geneva Centre for Emerging Viral Diseases, Geneva University Hospitals, Geneva, Switzerland and [5]Division of Tropical and Humanitarian Medicine, Geneva University Hospitals, Geneva, Switzerland

Abstract

Hepatitis E virus (HEV) is a major cause of acute jaundice in South Asia. Gaps in our understanding of transmission are driven by non-specific symptoms and scarcity of diagnostics, impeding rational control strategies. In this context, serological data can provide important proxy measures of infection. We enrolled a population-representative serological cohort of 2,337 individuals in Sitakunda, Bangladesh. We estimated the annual risks of HEV infection and seroreversion both using serostatus changes between paired serum samples collected 9 months apart, and by fitting catalytic models to the age-stratified cross-sectional seroprevalence. At baseline, 15% (95 CI: 14–17%) of people were seropositive, with seroprevalence highest in the relatively urban south. During the study, 27 individuals seroreverted (annual seroreversion risk: 15%, 95 CI: 10–21%), and 38 seroconverted (annual infection risk: 3%, 95CI: 2–5%). Relying on cross-sectional seroprevalence data alone, and ignoring seroreversion, underestimated the annual infection risk five-fold (0.6%, 95 CrI: 0.5–0.6%). When we accounted for the observed seroreversion in a reversible catalytic model, infection risk was more consistent with measured seroincidence. Our results quantify HEV infection risk in Sitakunda and highlight the importance of accounting for seroreversion when estimating infection incidence from cross-sectional seroprevalence data.

Introduction

Hepatitis E virus (HEV) is an important cause of acute jaundice in populations with limited access to safe drinking water across South Asia and Africa [1–5]. Transmitted from person-to-person via faecal-contaminated water, HEV genotypes 1 and 2 (HEV-1 and 2) carry a high risk of death if contracted during pregnancy [6] and regularly cause outbreaks, with particularly long-lasting transmission documented in displaced persons camps [7]. In Bangladesh, HEV is the leading cause of hospital-attended acute jaundice [3]. Although more recent data are lacking, an analysis of verbal autopsies from 1998 to 2007 estimated that 19–25% of maternal deaths were associated with jaundice, suggesting that HEV may be a key cause of maternal mortality nationally [2].

The risk factors driving HEV infection are not fully understood and attempts to slow transmission through emergency water and sanitation interventions have had limited success [8]. In the absence of effective treatment, vaccination is a promising tool to avert cases and deaths. Whilst an efficacious vaccine exists [9], the lack of reliable burden estimates is one of several barriers preventing the World Health Organisation (WHO) from recommending its routine use [10, 11]. Sparse surveillance data and differing model assumptions mean that estimates of morbidity and mortality attributable to HEV vary widely, rendering them difficult to interpret [4, 5, 12, 13]. Understanding the drivers of HEV infection and improving incidence estimates can greatly support both global and local decision-makers.

Although not a direct measure of disease incidence, serological data can provide important proxy measures of infection. Age-stratified cross-sectional seroprevalence is often used to estimate the rate at which seronegative individuals become infected with a pathogen [14, 15], and has previously been used to estimate HEV infection incidence though past approaches have had several limitations [4, 5]. Despite evidence that antibodies to HEV wane over time [16, 17], models have ignored seroreversion, which cannot always be reliably estimated from cross-sectional seroprevalence [18]. Additionally, such approaches have assumed infection risk does

not vary with age or time, which may not hold for HEV. Collecting serum samples from the same individuals at different time-points can overcome some of these limitations with observed serostatus changes providing a direct measure of seroincidence and seroreversion.

In this study, we aimed to fill several gaps in our understanding of HEV infections by enrolling a population-representative longitudinal serological cohort of 580 households in an HEV-1 endemic region in Bangladesh [19, 20] to explore infection risk factors and estimate the annual risk of infection, and the rate of seroreversion. As a secondary goal, we aimed to compare the concordance of estimates of infection incidence derived from cross-sectional data to those observed in the longitudinal cohort.

Methods

Serosurvey design

We tested serum samples from a population-representative cohort recruited between March 2021 and February 2022 with the original aim of estimating *Vibrio cholerae* O1 seroincidence in the Sitakunda sub-district of Chattogram, Bangladesh (Supplementary Figure S1). Households were recruited through a previously described two-stage sampling process [21]. Briefly, we first divided Sitakunda into 1km^2 grid cells and randomly sampled cells weighted by the number of household structures identified by satellite imagery. We then randomly sampled structures to visit within each grid cell weighted by whether they were single- or multi-story units. For each household, we sought consent from the household head and attempted to enrol all members ≥1 years old. Study staff administered a questionnaire covering household-level infrastructure, assets and sanitation facilities to household representatives, and an individual-level questionnaire on demographics and drinking-water sources to all consenting household members. In addition, ~5 ml of venous blood was collected from each consenting household member (~3 ml for those <5 yrs). Enrolled households were visited at a subsequent timepoint approximately 9 (range 7–11) months from baseline, to ask follow-up questions and repeat blood collection.

Sample testing

Paired serum samples from the two survey rounds were tested for anti-HEV immunoglobulin (IgG) at icddr,b using commercially available Wantai HEV IgG ELISA kits (Wantai Biological, China). Following the manufacturer's instructions, samples with a standardized optical density > 1.1 were considered seropositive, those <0.9 were considered negative, and those between 0.9 and 1.1 borderline. Borderline results were excluded from analyses.

Statistical analysis

Seroprevalence and risk factor analyses

We estimated baseline seroprevalence and 95% confidence intervals, accounting for household sampling survey design, using the Rao–Scott method, implemented in the R package 'survey' [22]. We created smoothed maps of household seroprevalence to visualize heterogeneity across Sitakunda using inverse distance weighting [23], implemented in the R package 'spatstat' [24], and assessed the spatial autocorrelation by estimating the semivariogram for household seroprevalence. We explored the relationship between individual- and household-level variables and baseline seropositivity using mixed-effects logistic regression models to account for household-level random effects. Firstly, we estimated univariate odds ratios of seropositivity for variables pertaining to demography, drinking-water and sanitation, history of jaundice, and livestock keeping that we identified a priori as potentially causally related to HEV exposures. For each potential risk factor, we also estimated an adjusted odds ratio by adjusting for confounders that we identified based on our assumptions represented in the Directed Acyclic Graph (DAG) in Supplementary Figure S2.

Drinking-water sources and sanitation facilities were categorized as improved and unimproved based on definitions from the WHO/UNICEF Joint Monitoring Programme (JMP) service ladders for drinking water [25] and sanitation facilities [26]. Ages groups <5, 5–14, 15–39, and 40–100 were chosen based on common usage of <5 in public health policy, and to reflect our a priori assumption that, as well as determining time spent at risk, age could relate to HEV seropositivity through differences in behaviour by stage of life.

Annual risks of infection and seroreversion

We estimated the annual risk of infection and seroreversion using two methods: (i) using the observed changes in seroprevalence between study visits, and (ii) fitting catalytic models [14, 15] to age-stratified seroprevalence data from a single study visit.

The annual risk of infection was estimated by dividing the number of individuals who seroconverted over the course of the study period, n_{sc}, (i.e., baseline seronegative individuals who became seropositive) by the total number of person-time at risk during the study period (total person-time for those who remained seronegative throughout, pt_{sn}, plus half the person-time for those who seroconverted, pt_{sc}) as described by equation (1). Similarly, the annual risk of seroreversion was estimated by dividing the number of baseline seropositive individuals who became seronegative, n_{sr}, by the total person-time at risk (total person-time for those who remained seropositive throughout pt_{sp}, plus half the person-time for those who seroreverted, pt_{sr}) as described by equation (2). We assumed on average individuals seroconverted or seroreverted at the midpoint of their time in the study.

$$annual\ risk\ of\ infection = \frac{n_{sc}}{pt_{sn} + \frac{1}{2}(pt_{sc})} \quad (1)$$

$$annual\ risk\ of\ seroreversion = \frac{n_{sr}}{pt_{sp} + \frac{1}{2}(pt_{sr})} \quad (2)$$

Alternative estimates of the annual risk of infection were produced by fitting two catalytic models of seroconversion to the cross-sectional age-stratified seroprevalence data collected at baseline. Models were also fit to seroprevalence data collected at follow-up as a sensitivity analysis. In Model 1, we assumed no antibody waning, which has been the traditional approach to estimating HEV infection rates [4, 5]. In Model 2, due to evidence of seroreversion in our empirical data and in previous studies [16, 17], we allowed for seroreversion at a constant rate, ρ. Since previous work has shown that the annual risk of infection and the seroreversion rate often cannot be reliably estimated from cross-sectional seroprevalence data simultaneously due to identifiability issues [18], we used our empirical estimate for the rate of seroreversion in Model 2. For both models, we initially considered infection risk to be constant across age and time, then, to reflect apparent differences in the rate of change of seroprevalence in children and adults in our data, we

repeated model fitting whilst allowing the risk to be different in adults than in children as implemented previously [27, 28]. We varied the age at which risk of infection changed between 12 and 36 years to conservatively cover the transition into adulthood. We then determined the age of change that maximized model fit using leave-one-out cross-validation (LOO-CV) [29] as implemented in the loo R package [30]. Model solutions are presented in the Supplementary Material. The seroprevalence data were assumed to be binomially distributed. Models were fitted within a Bayesian framework using a Hamiltonian Monte Carlo algorithm implemented within rstan which provided a convenient framework within which to propagate uncertainty in the measured rate of seroreversion [31, 32]. We used a uniform prior between 0 and 1 for the per capita annual risk of infection, and between 0 and 10 for the per capita annual risk of seroreversion to conservatively include previous measures of HEV antibody persistence which have ranged from many years down to a matter of months [16, 17, 33, 34].

To estimate the annual number of HEV infections, we multiplied the annual risk of infection, based on observed seroconversions within the cohort, by the estimated population of Sitakunda in 2021. We extrapolated the age-stratified population counts from the 2011 National Census [35], assuming that the population grew by 1.5% each year between 2011 and 2021. Since we did not include <1-year olds in our survey, we subtracted 20% of the population count for 0- to 4-year-olds based on the age distribution presented in the US Census Bureau International Database for Bangladesh in 2021 [36].

Ethical review

Adult study participants provided written informed consent. Parents or guardians of all participants <18 years were asked to provide consent on their behalf, with those 11–17 years old also providing written assent. The protocols for the original study and extension to test samples for HEV antibodies were approved by the icddr,b research and ethics review committee and the Johns Hopkins Bloomberg School of Public Health institutional review board (IRB00014794).

Results

HEV seroprevalence at baseline

Between 27 March and 13 June 2021, 2,337 individuals from 580 households were recruited to the serological cohort, and 2,301 (98%) were tested for anti-HEV IgG antibodies (Supplementary Table S1). At enrolment, 353 (15%; 95% confidence interval [CI] 14–17%; ICC: 0.05; design effect: 1.15) of the sampled population of Sitakunda had antibodies indicating past HEV infection (which we refer to as being 'seropositive'). Seroprevalence was significantly higher in males (20%; 95% CI: 17–22%) than females (12%; 95% CI: 10–14%) (Figure 1a), with the difference becoming more apparent during adulthood (Figure 1b). Seroprevalence was low in children and increased until approximately age 40 (Figure 1b). Household seroprevalence was higher in the relatively urbanized south, near Chattogram city, with notably area of higher seroprevalence in the south-east (Figure 1c). The seroprevalence in individuals sampled in this area was 46% (95%CI: 38–54%), more than 3-fold higher than the average. Compared to just 4% of the overall sample, 20% of individuals within this area reported that their primary drinking water source had been unavailable at least once in the month prior to the baseline survey (Supplementary Table S1). When comparing demographic characteristics, a larger proportion of individuals in the area were male (53% compared to 46%), and more households had a monthly income <10,000TK, shared sanitation facilities, and kept livestock (Supplementary Table S1). Whilst we did detect an area with elevated seroprevalence, we did not detect strong spatial correlation in household seroprevalence across Sitakunda (Supplementary Figure S3).

To further explore the sex differences in seroprevalence and to identify other potential individual- and household-level risk factors for seropositivity, we conducted logistic mixed-effects regression analyses. When using multivariable models to adjust for potential confounders, defined a priori (Table 1, Supplementary Figure S2), we found significantly increased odds of seropositivity among males, those ≥40 years, those reporting that their primary water source was unavailable at least once in the past month and those with business or other occupations outside the home (Table 1). Including both occupation and sex in a multivariable model attenuated the effect size for sex (1.4, 95% CI 1.0–2.0) to the limit of statistical significance.

Self-reported history of jaundice was not significantly associated with an increased odds of seropositivity, but our power to detect a significant difference in this study was limited [37]. Of the 59 people who reported having ever had acute jaundice (lasting <3 months), nine reported having had jaundice between survey rounds but none seroconverted. Reported drinking water sources in week prior to the survey were not associated with significantly different odds of seroprevalence. Only one household kept pigs – a known host of zoonotic HEV genotypes 3 and 4 – meaning that we are not able to investigate whether there is an association between keeping pigs and HEV seropositivity. Keeping mammalian livestock was not associated with significantly different baseline seropositivity.

Evidence of infection during the study period – Empirical estimation of the annual risk of infection

Of the 1,580 individuals who were seronegative at baseline and provided blood at the end of the study, 38 became seropositive, with similar rates of seroconversion in men and women. This translates to an annual risk of infection of 3% (95% CI: 2–5%) for a seronegative individual in Sitakunda, or approximately 12,500 infections in those ≥1 year old in 2021 [35]. The mean annual risk of infection was higher in adults (4%; 95 CI: 3–6%) than in children <18 years old (3%; 95%CI: 1–5%), but this was not statistically significant. Of the 38 seroconverters, seven lived within the high seroprevalence cluster where the annual infection risk was 19% (95%CI: 7–38%). The 38 seroconverters came from 29 households, and among these, five households had >1 seroconverter, including one where five members became seropositive.

Waning of antibodies during the study period – Empirical estimation of the rate of seroreversion

Of the 266 individuals who were seropositive at baseline and for whom we have paired samples, 27 became seronegative during the study. This translates to an annual risk of seroreversion of 15% (95%CI: 10–21%). Seroreversion rates were slightly lower in males (12%, 95%CI: 6–20%) than females (19%, 95%CI: 11–32%). Seroreversion rates were significantly higher in children than in adults (Supplementary Figure S4). Of the five children <10 who were seropositive at baseline, four seroreverted. The mean time to seroreversion for children <10 years was estimated to be 7 months (95%CI: 3–24 months) compared to 8 years (95%CI: 5–13 years) for those ≥10 years.

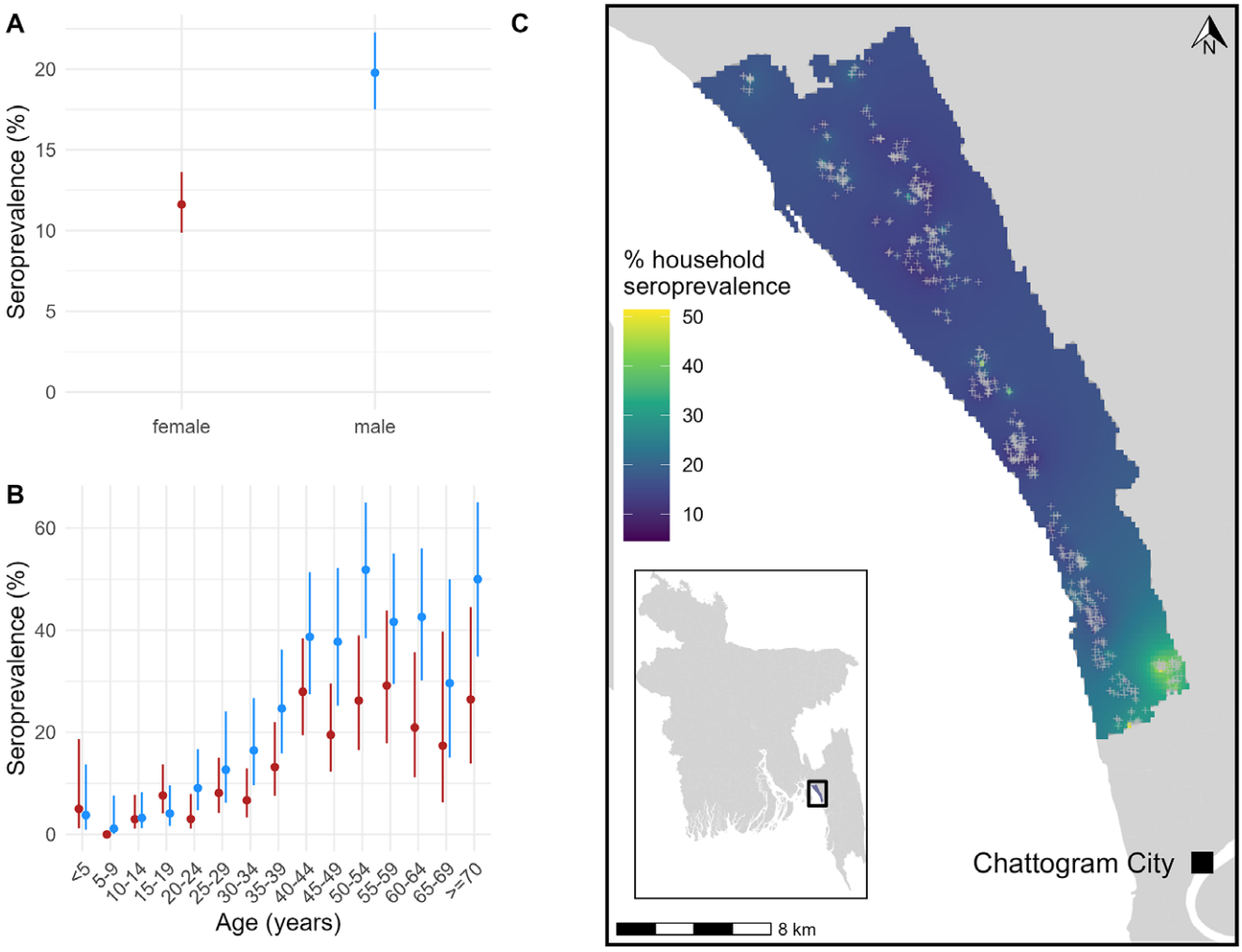

Figure 1. (a) Baseline seroprevalence by sex, with 95% confidence intervals. (b) Baseline seroprevalence stratified by age, with 95% confidence intervals. (c) Smoothed spatial variation in baseline household seroprevalence across Sitakunda. Grey crosses represent the location of sampled households.

Table 1. Potential risk factors for past HEV infection at baseline

	Characteristic		Seronegative n (%)	Seropositive n (%)	OR (95%CI) *Unadjusted*	OR (95%CI) *Adjusted*	Variable adjusted for
Individual level	Sex	Female	1,089 (88%)	143 (12%)	*Reference*		none
		Male	852 (80%)	210 (20%)	1.9 (1.5, 2.4)	NA	
	Age in years	<5	89 (96%)	4 (4%)	*Reference*		none
		5–14	427 (98%)	9 (2%)	0.5 (0.1, 1.5)	NA	
		15–39	932 (91%)	97 (9%)	2.4 (0.8, 6.7)	NA	
		40–100	493 (67%)	243 (33%)	12.9 (4.5, 36.6)	NA	
	Occupation	Homeworker	717 (85%)	127 (15%)	*Reference*		Age, sex
		Business[a]	364 (73%)	138 (27%)	2.2 (1.7, 3)	1.7 (1.1, 2.6)	
		Farmer	53 (66%)	27 (34%)	3.1 (1.8, 5.4)	1.7 (0.9, 3.2)	
		Student	646 (97%)	20 (3%)	0.2 (0.1, 0.3)	0.5 (0.3, 1.0)	
		Child	91 (95%)	5 (5%)	0.3 (0.1, 0.8)	2.5 (0.4, 15.5)	
		None	48 (68%)	23 (32%)	3 (1.7, 5.4)	1.3 (0.7, 2.3)	
		Other[b]	21 (62%)	13 (38%)	3.6 (1.7, 7.9)	3.2 (1.4, 7.4)	

(*Continued*)

Table 1. (*Continued*)

	Characteristic		Seronegative n (%)	Seropositive n (%)	OR (95%CI) *Unadjusted*	OR (95%CI) *Adjusted*	Variable adjusted for
	Travel: time since last leaving the village	> 1 year	1818 (85%)	322 (15%)	*Reference*		Age, sex, household income, occupation
		1 month – 1 year	90 (80%)	22 (20%)	1.4 (0.8, 2.3)	1.1 (0.6, 2)	
		1 week – 1 month	19 (79%)	5 (21%)	1.5 (0.5, 4.4)	2.3 (0.7, 7.3)	
		< 1 week	14 (78%)	4 (22%)	1.6 (0.5, 5.2)	1 (0.3, 3.6)	
	At least once in the last month primary drinking water source was unavailable	No	1871 (85%)	328 (15%)	*Reference*		Household income, occupation
		Yes	69 (74%)	24 (26%)	2.0 (1.2, 3.4)	2.5 (1.4, 4.4)	
	Use of piped water in the past week	No	1,685 (85%)	296 (15%)	*Reference*		Household income, occupation
		Yes	252 (82%)	57 (18%)	1.3 (0.9, 1.9)	1.4 (0.9, 2.0)	
	Use of tubewell in the past week	No	385 (83%)	81 (17%)	*Reference*		Household income, occupation
		Yes	1,552 (85%)	272 (15%)	0.8 (0.6, 1.1)	0.8 (0.6, 1.1)	
	Use of a public tap/standpipe in the past week	No	1790 (85%)	319 (15%)	*Reference*		Household income, occupation
		Yes	147 (81%)	34 (19%)	1.3 (0.9, 2.0)	1.5 (0.9, 2.3)	
	Reported ever having had acute jaundice[c]	No	1,562 (85%)	267 (15%)	*Reference*		none
		Yes	46 (82%)	10 (18%)	1.3 (0.6, 2.7)	NA	
Household level	Type of dwelling	Single house	1,264 (84%)	245 (16%)	*Reference*		Household income
		Several separate structures	271 (87%)	40 (13%)	0.8 (0.5, 1.1)	0.8 (0.5, 1.1)	
		Flat in a multi–story building	175 (83%)	35 (17%)	1.0 (0.7, 1.6)	1.1 (0.7, 1.6)	
		Flat in a single–story building	144 (87%)	22 (13%)	0.8 (0.5, 1.3)	0.8 (0.5, 1.3)	
		Room in a larger dwelling	87 (89%)	11 (11%)	0.7 (0.3, 1.3)	0.6 (0.3, 1.3)	
	Household income	<10,000 BDT[d]	298 (81%)	69 (19%)	1.4 (1.0, 1.9)	1.4 (1.0, 2.0)	Occupation
		>10,000 BDT[d]	1,643 (85%)	284 (15%)	*Reference*		
	Sanitation facility category	Improved private	1,476 (85%)	260 (15%)	*Reference*		Household income, type of dwelling
		Improved shared	453 (83%)	90 (17%)	1.1 (0.9, 1.5)	1.1 (0.8, 1.5)	
		Unimproved	12 (80%)	3 (20%)	1.4 (0.3, 5.6)	1.0 (0.3, 3.9)	
	Keeps mammalian livestock in the household	No	1,381 (85%)	250 (15%)	*Reference*		Household income
		Yes	555 (84%)	103 (16%)	1.0 (0.8, 1.3)	1 (0.8, 1.4)	

[a]Includes responses 'business', 'business man', and business work 'outside the home'.
[b]When asked to specify, common responses included 'service', 'private services', 'teacher', and 'driver'.
[c]Described as an illness during which the eyes or skin turn yellow. This was only asked at the second time point.
[d]10,000 BDT is approximately 100 USD.

As changes in seropositivity could result from small fluctuations in antibody concentration or measurement error around the cut-off, we compared the optical density to cut-off ratios for individuals across rounds. Of the 65 individuals whose serostatus changed, 90% had a considerable change in od:cut-off ratio ranging in magnitude from 2 to 20 (the upper limit of detection) (Supplementary Figure S5A-B). Several individuals classified as seropositive at baseline had a substantial increase in their optical density to cut-off ratio suggesting reinfections (Supplementary Figure S5.C).

Estimating the annual risk of infection from cross-sectional data

The annual risk of HEV infection has previously been estimated by fitting catalytic models of seroconversion to age-stratified cross-sectional seroprevalence data, assuming antibodies do not

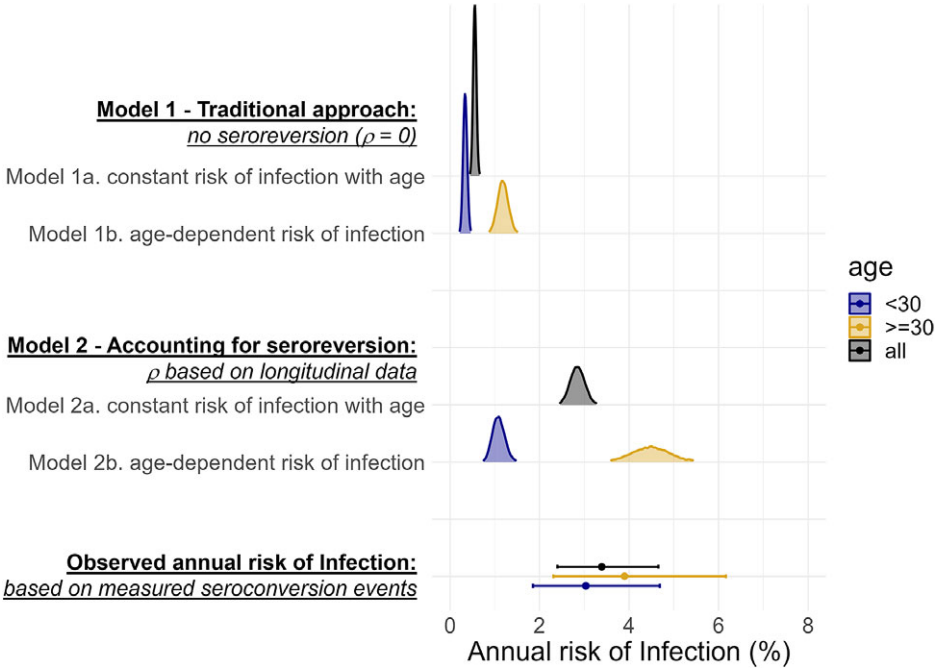

Figure 2. A comparison of posterior estimates of the annual risk of infection from catalytic models 1 and 2 fitted to age-stratified cross-sectional data (smoothed density curves) and the estimates of the annual risk of infection from observed seroconversion events captured in our longitudinal serostatus data (points with bars representing 95%CIs).

wane. When we fitted a model assuming lifelong antibody persistence (Model 1) to our baseline age-stratified seroprevalence data, our estimated annual risk of infection was 5-times lower than what we had measured based on observed seroconversion events (Figure 2). This same magnitude of difference was seen both when we assumed the risk of infection was the same across ages (Model 1a), and when we allowed risk of infection to differ in adulthood (Model 1b, which optimized model fit; Supplementary Table S3, Supplementary Figure S6) (Figure 2).

When we fitted a model (Model 2) assuming antibodies wane below detectable levels at the rate we measured in the cohort (1.8 (95%CI: 0.5–4.6) in <10-year-olds and 0.1 (95%CI: 0.1–0.2) in ≥10-year-olds), the resulting estimates of annual risk of infection were more comparable to what we measured longitudinally (Figure 2). When we assumed that all age classes experienced the same risk of infection (Model 2a), the estimated annual risk of infection was 2.8% (95%CrI: 2.5–3.2%) compared to 3.4% (95%CI: 2.4–4.6%) based on the observed seroconversion events. However, this model fit resulted in overestimates of age-specific seroprevalence for younger age-classes and underestimates for older age-classes when compared to the observed data (Supplementary Figure S6). Allowing age-class-dependent risk of infection (Model 2b) improved model fit, with the best fit achieved when the risk changed at 30-years (Supplementary Figure S7). For individuals >30, the annual risk of infection was 4.5% (95%CrI: 3.7–5.4%), similar to the measured risk for that age group (3.8%, 95%CI: 2.3–6.1%). The annual risk of infection in those <30 years though was lower than what was measured in the cohort (Figure 2). Using the cross-sectional seroprevalence data from follow-up rather than baseline produced very similar estimates of seroprevalence by age and annual risk of infection (Supplementary Figure S8 & Supplementary Table S4).

When we attempted to simultaneously fit both the rate of infection and of seroreversion, our estimates were not compatible with the seroreversion rate we measured in the cohort (Supplementary Figure S9.C) and model fit did not improve (Supplementary Table S3, Supplementary Figure S6). When we assumed infection risk was constant with age, the annual risk of seroreversion approached zero (0.09%, 95%CrI: 0.00–0.35%). When we allowed different infection risks for those over and under 30, the annual risk of seroreversion was 5.7% (95%CrI: 2.1–10.7%), less than half what we observed during the study.

Discussion

In this population-representative longitudinal cohort study, we collected two types of serological data as proxy measures of HEV infection in an endemic region of Bangladesh. We used age-stratified cross-sectional seroprevalence data to explore risk factors associated with past HEV infection and to produce traditional estimates of the annual risk of infection, whilst longitudinal serostatus changes allowed us the opportunity to capture incident infections and seroreversion.

At 15% (95%CI: 14–17%), our estimated baseline seroprevalence was similar to the estimate derived from the 2015 national serosurvey for Sitakunda (19%, 95%CrI: 12–30%) [38]. We estimated that 3% of seronegative people become infected annually in Sitakunda, translating to approximately 15,000 HEV infections in people ≥1 year old which is approximately half that of previous estimates for rural Bangladesh in 2003–2004 (6%, 95%CI: 4–8%). This could reflect lower transmission intensity in Sitakunda than other rural parts of the country or decreasing incidence over time.

Like previous studies in Bangladesh [38, 39], we found male sex to be associated with higher odds of anti-HEV seropositivity in univariate analyses. Occupation may mediate part of the effect of sex on infection risk, with those working in business or other occupations outside the household significantly more likely to be seropositive and male. Collecting data on water source use outside

the household may help us better understand how these consumption patterns are related to HEV infection risk. Despite evidence that HEV outbreaks have been caused by breeches and contamination of municipal piped water supply systems in urban Bangladesh [40], we did not find use of piped water in the week prior to the survey to be associated with significantly higher odds of seropositivity in this study. However, individuals who reported that their primary water source had been unavailable to them at least once in the last month were more than 2 times more likely to be seropositive, suggesting that use of water lines with insufficient supply, or having to use alternatives, may increase risk of infection.

In the absence of longitudinal data, estimating annual HEV infection risk has traditionally relied on fitting catalytic models to cross-sectional seroprevalence, assuming life-long antibody persistence and constant risk of infection. By testing paired samples and capturing seroconversion events, we measured the annual risk of HEV infection in Sitakunda to be 5 times higher than the estimate generated by applying traditional approaches to our baseline seroprevalence data. Using observed seroreversion events to inform a reversible model of seroconversion allowed us to obtain estimates of the annual risk of infection closer to what we measured during our study. Simultaneously fitting the rate of seroreversion and infection risk to cross-sectional seroprevalence data could not eliminate the need for empirical measures of seroreversion, likely due to identifiability issues [18]. We found that allowing for different annual infection risk in those >30 greatly improved the model fit, but it was not possible to determine if this reflects a higher risk in those >30 or lower incidence over the past 30 years.

Our estimated annual risk of seroreversion (15%) is higher than existing estimates which were approximately 2% in rural Bangladesh (20% antibody loss over 10–12 years), 4% in Kashmir (50% over 14 years), and 5% in China (30% over 6.5 years) [16, 17, 34]. We expect this is partly due to the shorter time between paired sample collection in our study compared to previous studies where seroreversion events may have been masked by reinfection. The majority of seroreversion events involved considerable changes in the optical density to cut-off ratio between baseline and follow-up suggesting the higher seroreversion rate we measured cannot be attributed to noise. Even when we exclude the 6 seroreverters with the lowest optical density to cut-off ratio at baseline (<2), we estimate the annual risk of seroreversion to be 12% (95% CI: 7–18%). The seroreversion rate was significantly higher in children <10 than in adults – a trend also observed in the previous study of anti-HEV IgG loss in Bangladesh [17]. HEV antibodies were also measured to wane quickly in a study in children in Egypt, becoming undetectable in a matter of months [33]. In our study, most children who seroreverted started with optical density values near the upper limit of the dynamic range of the assay, suggesting that children may experience faster antibody waning than adults, rather than mounting a lower initial antibody response. Potential reasons for this observed difference in seroreversion rate are not well understood though could include differences in the immune response to HEV in young children, or less antibody persistence due to fewer repeat infections in children compared to adults who have been at risk for longer. A larger sample of seropositive children is needed to investigate these differences in antibody persistence further and to understand the implications for estimating risk of infection in this age group.

Our study has several limitations. Firstly, we used a non-quantitative serological assay and did not use reference serum to allow for generalized comparisons to other studies [41]. We relied on the threshold specified by the kit instructions to classify samples as seropositive and seronegative, assuming perfect assay sensitivity and specificity for detecting infections during the study period. Sensitivity and specificity are estimated to be high [42, 43], but the use of this threshold is unlikely to produce perfect, generalizable classification. Although the majority of serostatus changes were associated with large changes in antibody titres, a small minority of seroconversion and reversion events involved relatively small changes. Without a gold standard assay for comparison, we cannot rule out that some serostatus changes could be due to noise. We also saw several large boosts in antibody measures in seropositive individuals that likely represent uncounted reinfections.

Secondly, changes in infection risk over time could contribute to the discrepancy between the seroincidence measured longitudinally and our estimates from cross-sectional data. For example, an undetected outbreak could have elevated the risk of infection measured during our study, but very low numbers of people self-reporting jaundice, and similar seroprevalence to other studies suggest this did not occur. Endemic transmission of HEV is not known to be strongly seasonal in Bangladesh [44], but if there is a seasonal component, then we could have slightly biased estimates by not spreading equal at-risk person-time across a full year. Spanning 10.5 months, we would not expect the effect to be considerable in our study. Finally, although we were able to estimate the HEV infection rate in Sitakunda, the absence of a serological assay capable of distinguishing between HEV genotypes, with potentially very different clinical consequences, constrains our ability to translate this into estimates of disease burden. To date, genotyped clinical cases in Bangladesh have all been classified as HEV-1 [19, 45], but zoonotic HEV likely circulates in pigs in the country [46] and the incidence of spill over is unknown.

Our results provide evidence of endemic circulation of HEV in Sitakunda, with people who work in occupations outside the home and those reporting their primary water sources to be recently unavailable at higher risk of infection. In the face of widespread under-reporting of hepatitis E cases, estimates of infection incidence from serological data, whilst imperfect, are important for improving our understanding of transmission, risk and burden of hepatitis E. We were able to evaluate estimates produced by traditional approaches against empirical estimates of seroincidence, demonstrating the need to account for the rate of antibody waning and differences in the risk of infection experienced by different age groups, to avoid underestimating incidence of HEV infection. Refining our interpretation of hepatitis E serological data, through both improving analytical methods and collecting new longitudinal data across geographies, will be key to expanding the breadth of our understanding of this vaccine-preventable disease.

Supplementary material. The supplementary material for this article can be found at http://doi.org/10.1017/S0950268824000438.

Data availability statement. The data and code used in these analyses are available here: https://github.com/HopkinsIDD/sitakunda-hev.

Acknowledgements. We thank all the participants in the cohort study for taking part in this research. We thank the field and laboratory staff for their dedicated work.

Author contribution. Conceptualization: A.S.A., F.Q., E.S.G., S.T.H.; Funding acquisition: A.S.A., F.Q., E.S.G.; Project administration: A.S.A.; Supervision: A.S.A., F.Q., E.S.G.; Writing – review & editing: A.S.A., F.Q., A.I.K., E.S.G., I.I.K., J.D.H., M.R., R.R., S.A., S.T.H., M.T.I., T.R.B., Z.H.K., M.Z.H.; Investigation: A.I.K., I.I.K., M.R., R.R., S.A., S.T.H., M.T.I., T.R.B., Z.H.K., M.Z.H.; Data curation: J.D.H.; Formal analysis: T.R.B., A.D.; Methodology: A.D.; Writing – original draft: A.D.

Funding statement. This work was supported by The Bill and Melinda Gates Foundation (grant number: INV-038404).

Competing interest. The authors declare no competing interests exist.

References

[1] **Gerbi GB**, et al. (2015) Hepatitis E as a cause of acute jaundice syndrome in northern Uganda, 2010–2012. *American Journal of Tropical Medicine and Hygiene* **92**, 411–414.

[2] **Gurley ES**, et al. (2012) Estimating the burden of maternal and neonatal deaths associated with jaundice in Bangladesh: Possible role of hepatitis E infection. *American Journal of Public Health* **102**, 2248–2254.

[3] **Paul RC**, et al. (2020) Hepatitis E as a cause of adult hospitalization in Bangladesh: Results from an acute jaundice surveillance study in six tertiary hospitals, 2014–2017. *PLOS Neglected Tropical Diseases* **14**, e0007586.

[4] **Rein DB**, et al. (2012) The global burden of hepatitis E virus genotypes 1 and 2 in 2005. *Hepatology* **55**, 988–997.

[5] **Stanaway JD**, et al. (2016) The global burden of viral hepatitis from 1990 to 2013: Findings from the global burden of disease study 2013. *Lancet* **388**, 1081–1088.

[6] **Amanya G**, et al. (2017) Risk factors, person, place and time characteristics associated with hepatitis E virus outbreak in Napak District, Uganda. *BMC Infectious Diseases* **17**, 451.

[7] **Koyuncu A**, et al. (2021) Setting a course for preventing hepatitis E in low and lower-middle-income countries: A systematic review of burden and risk factors. *Open Forum Infectious Diseases* **8**, ofab178.

[8] **Guthmann J-P**, et al. (2006) A large outbreak of hepatitis E among a displaced population in Darfur, Sudan, 2004: The role of water treatment methods. *Clinical Infectious Diseases* **42**, 1685–1691.

[9] **Zhu F-C**, et al. (2010) Efficacy and safety of a recombinant hepatitis E vaccine in healthy adults: A large-scale, randomised, double-blind placebo-controlled, phase 3 trial. *Lancet* **376**, 895–902.

[10] **Lynch JA**, et al. (2023) Hepatitis E vaccine—Illuminating the barriers to use. *PLOS Neglected Tropical Diseases* **17**, e0010969.

[11] **World Health Organisation** (2015) WHO position paper on hepatitis E vaccines - May 2015. *Weekly Epidemiological Record* **18**, 184–200.

[12] **GBD 2019 Diseases and Injuries Collaborators**, et al. (2020) Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: A systematic analysis for the global burden of disease study 2019. *Lancet* **396**, 1204–1222.

[13] **Li P**, et al. (2020) The global epidemiology of hepatitis E virus infection: A systematic review and meta-analysis. *Liver International: Official Journal of the International Association for the Study of the Liver* **40**, 1516–1528.

[14] **Muench H** (1934) Derivation of rates from summation data by the catalytic curve. *Journal of the American Statistical Association* **29**, 25–38.

[15] **Keiding N** (1991) Age-specific incidence and prevalence: A statistical perspective. *Journal of the Royal Statistical Society: Series A* **154**, 371–396.

[16] **Kmush BL**, et al. (2019) Long-term antibody persistence after hepatitis E virus infection and vaccination in Dongtai, China. *Open Forum Infectious Diseases* **6**, ofz144.

[17] **Kmush BL**, et al. (2018) A 10-year Immunopersistence study of hepatitis E antibodies in rural Bangladesh. *American Journal of Epidemiology* **187**, 1501–1510.

[18] **Arnold BF**, et al. (2019) Enteropathogen antibody dynamics and force of infection among children in low-resource settings. *eLife* **8**, e45594.

[19] **Biswas RSR**, et al. (2020) Viral load and genotype of recent hepatitis E virus outbreak in Chittagong, Bangladesh. *Journal of the Scientific Society* **47**, 144.

[20] **Owada K**, et al. (2022) Epidemiological profile of a human hepatitis E virus outbreak in 2018, Chattogram, Bangladesh. *Tropical Medicine and Infectious Disease* **7**, 170.

[21] **Bhuiyan TR**, et al. (2022) SARS-CoV-2 Seroprevalence before Delta variant surge, Chattogram, Bangladesh, March–June 2021. *Emerging Infectious Diseases* **28**, 429.

[22] **Lumley T** (2023) *survey: Analysis of complex survey samples.* R package version 4.2.

[23] **Shepard D** (1968) A two-dimensional interpolation function for irregularly-spaced data. In *Proceedings of the 1968 23rd ACM National Conference*. New York, NY, USA: Association for Computing Machinery, pp. 517–524.

[24] **Baddeley A and Turner R** (2005) Spatstat: An R package for analyzing spatial point patterns. *Journal of Statistical Software* **12**, 1–42.

[25] **WHO. UNICEF** (n.d.) The WHO/UNICEF Joint Monitoring Programme. *Drinking water.* Available at https://washdata.org/monitoring/drinking-water (accessed 7 June 2023).

[26] **WHO. UNICEF** (n.d.) The WHO/UNICEF Joint Monitoring Programme. *Sanitation.* Available at https://washdata.org/monitoring/sanitation (accessed 7 June 2023).

[27] **Rees EM**, et al. (2021) Estimating the duration of seropositivity of human seasonal coronaviruses using seroprevalence studies. *Wellcome Open Research* **6**, 138.

[28] **Capai L**, et al. (2020) Seroprevalence of hepatitis E virus among blood donors on Corsica, France, 2017. *Eurosurveillance* **25**, 1900336.

[29] **Vehtari A**, **Gelman A and Gabry J** (2017) Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing* **27**, 1413–1432.

[30] **Vehtari A**, et al. (2023) *loo: Efficient Leave-One-Out Cross-Validation and WAIC for Bayesian Models.* R package version 2.6.0.

[31] **Carpenter B**, et al. (2017) Stan: A probabilistic programming language. *Journal of Statistical Software* **76**, 1–32.

[32] **Stan Development Team** (2023) *RStan: the R interface to Stan.* R package version 2.17.3.

[33] **Goldsmith R**, et al. (1992) Enzyme-linked immunosorbent assay for diagnosis of acute sporadic hepatitis E in Egyptian children. *Lancet* **339**, 328–331.

[34] **Khuroo MS**, et al. (1993) Hepatitis E and long-term antibody status. *Lancet* **341**, 1355.

[35] **Bangladesh Bureau of Statistics** (2015) *Bangladesh Population and Housing Census 2011 - Zila Report - Chittagong.* Available at https://bbs.gov.bd/site/page/47856ad0-7e1c-4aab-bd78-892733bc06eb/Population-and-Housing-Census (accessed 20 July 2023).

[36] **U.S. Census Bureau** (n.d.) *International Database (IDB).* Available at https://www.census.gov/programs-surveys/international-programs/about/idb.html (accessed 20 July 2023).

[37] **Lhomme S**, et al. (2020) Clinical manifestations, pathogenesis and treatment of hepatitis E virus infections. *Journal of Clinical Medicine* **9**, 331.

[38] **Azman AS**, et al. (2021) Hepatitis E in Bangladesh: Insights from a national serosurvey. *Journal of Infectious Diseases* **224**, S805–S812.

[39] **Labrique AB**, et al. (2009) Population seroprevalence of hepatitis E virus antibodies in rural Bangladesh. *American Journal of Tropical Medicine and Hygiene* **81**, 875–881.

[40] **Gurley ES**, et al. (2014) Outbreak of hepatitis E in urban Bangladesh resulting in maternal and perinatal mortality. *Clinical Infectious Diseases* **59**, 658–665.

[41] **Zhang J**, et al. (2015) Long-term efficacy of a hepatitis E vaccine. *New England Journal of Medicine* **372**, 914–922.

[42] **Bendall R**, et al. (2010) A comparison of two commercially available anti-HEV IgG kits and a re-evaluation of anti-HEV IgG seroprevalence data in developed countries. *Journal of Medical Virology* **82**, 799–805.

[43] **Yan Q**, et al. (2008) Comparison of two diagnostic reagents to detect anti-hepatitis E virus IgG antibodies. *Chinese Journal of Zoonoses* **24**, 1087–1089.

[44] **Labrique AB**, et al. (2010) Epidemiology and risk factors of incident hepatitis E virus infections in rural Bangladesh. *American Journal of Epidemiology* **172**, 952–961.

[45] **Hoa TN**, et al. (2021) A tightly clustered hepatitis E virus genotype 1a is associated with endemic and outbreak infections in Bangladesh. PLOS ONE *Public Library of Science* **16**, e0255054.

[46] **Haider N**, et al. (2017) Serological evidence of hepatitis E virus infection in pigs and jaundice among pig handlers in Bangladesh. *Zoonoses and Public Health* **64**, 572–577.