UC Berkeley
UC Berkeley Previously Published Works

Title

Heterogeneous susceptibility to rotavirus infection and gastroenteritis in two birth cohort studies: Parameter estimation and epidemiological implications

Permalink

Journal

ISSN

Authors

Lewnard, Joseph A
Lopman, Benjamin A
Parashar, Umesh D
et al.

Publication Date

2019

DOI

Peer reviewed

Heterogeneous susceptibility to rotavirus infection and gastroenteritis in two birth cohort studies: Parameter estimation and epidemiological implications
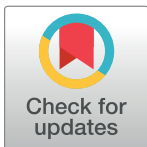
Joseph A. Lewnard[1]*, **Benjamin A. Lopman**[2,3], **Umesh D. Parashar**[3], **Aisleen Bennett**[4,5], **Naor Bar-Zeev**[4,5,6], **Nigel A. Cunliffe**[4,5], **Prasanna Samuel**[7], **M. Lourdes Guerrero**[8], **Guillermo Ruiz-Palacios**[8], **Gagandeep Kang**[7], **Virginia E. Pitzer**[9]

1 Division of Epidemiology and Biostatistics, School of Public Health, University of California, Berkeley, Berkeley, California, United States of America, 2 Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, Georgia, United States of America, 3 Division of Viral Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia, United States of America, 4 Malawi-Liverpool-Wellcome Trust Clinical Research Programme, College of Medicine, University of Malawi, Blantyre, Malawi, 5 Center for Global Vaccine Research, Institute of Infection and Global Health, University of Liverpool, University of Liverpool, Liverpool, United Kingdom, 6 International Vaccine Access Center, Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, United States of America, 7 Department of Gastrointestinal Sciences, Christian Medical College, Vellore, Tamil Nadu, India, 8 Instituto Nacional de Ciences Médicas y Nutrición Salvador Zubirán, Mexico City, Mexico, 9 Department of Epidemiology of Microbial Diseases, Yale School of Public Health, New Haven, Connecticut, United States of America

* jlewnard@berkeley.edu

Abstract

Cohort studies, randomized trials, and post-licensure studies have reported reduced natural and vaccine-derived protection against rotavirus gastroenteritis (RVGE) in low- and middle-income countries. While susceptibility of children to rotavirus is known to vary within and between settings, implications for estimation of immune protection are not well understood. We sought to re-estimate naturally-acquired protection against rotavirus infection and RVGE, and to understand how differences in susceptibility among children impacted estimates. We re-analyzed data from studies conducted in Mexico City, Mexico and Vellore, India. Cumulatively, 573 rotavirus-unvaccinated children experienced 1418 rotavirus infections and 371 episodes of RVGE over 17,636 child-months. We developed a model that characterized susceptibility to rotavirus infection and RVGE among children, accounting for aspects of the natural history of rotavirus and differences in transmission rates between settings. We tested whether model-generated susceptibility measurements were associated with demographic and anthropometric factors, and with the severity of RVGE symptoms. We identified greater variation in susceptibility to rotavirus infection and RVGE in Vellore than in Mexico City. In both cohorts, susceptibility to rotavirus infection and RVGE were associated with male sex, lower birth weight, lower maternal education, and having fewer siblings; within Vellore, susceptibility was also associated with lower socioeconomic status. Children who were more susceptible to rotavirus also experienced higher rates of rotavirus-negative diarrhea, and higher risk of moderate-to-severe symptoms when experiencing

RVGE. Simulations suggested that discrepant estimates of naturally-acquired immunity against RVGE can be attributed, in part, to between-setting differences in susceptibility of children, but result primarily from the interaction of transmission rates with age-dependent risk for infections to cause RVGE. We found that more children in Vellore than in Mexico City belong to a high-risk group for rotavirus infection and RVGE, and demonstrate that unmeasured individual- and age-dependent susceptibility may influence estimates of naturally-acquired immune protection against RVGE.

Author summary

Differences in susceptibility can help explain why some individuals, and not others, acquire infection and exhibit symptoms when exposed to infectious disease agents. However, it is difficult to distinguish between differences in susceptibility versus exposure in epidemiological studies. We developed a modeling approach to distinguish transmission intensity and susceptibility in data from cohort studies of rotavirus infection among children in Mexico City, Mexico, and Vellore, India, and evaluated how these factors may have contributed to differences in estimates of naturally-acquired immune protection between the studies. Given the same exposure, more children were at high risk of acquiring rotavirus infection, and of experiencing gastroenteritis when infected, in Vellore than in Mexico City. The probability of belonging to this high-risk stratum was associated with well-known individual factors such as lower socioeconomic status, lower birth weight, and incidence of diarrhea due to other causes. We also found the risk for rotavirus infections to cause symptoms declined with age, independent of acquired immunity. These findings can, in part, account for estimates of lower protective efficacy of acquired immunity against rotavirus gastroenteritis in high-incidence settings, mirroring estimates of reduced effectiveness of live oral rotavirus vaccines in low- and middle-income countries.

Introduction

Rotavirus is the leading source of gastrointestinal disease burden in children globally, with nearly 10 million severe cases and 193,000 fatalities estimated to occur annually [1]. One decade after their rollout in high-income settings, live oral rotavirus vaccines are currently being introduced to national immunization programs of low- and middle-income countries (LMICs). However, randomized controlled trials and post-licensure studies have reported lower vaccine efficacy and effectiveness against rotavirus gastroenteritis (RVGE) in LMICs compared to higher-income settings [2,3]. Understanding this performance gap is essential to maximizing the impact of rotavirus vaccines where they are needed most.

Recent observational studies have investigated how factors such as oral polio vaccine co-administration [4,5], exposure to breast milk antibodies [6], environmental enteropathy [7], and nutritional status [8,9] influence susceptibility of children to RVGE and performance of oral vaccines. Variation in susceptibility among individuals within and between studies—due to these or other unmeasured risk factors—is well known to influence estimates of vaccine efficacy and effectiveness [10–13]. Differential removal of highly-susceptible individuals to a partially-immune state constitutes a form of frailty bias or effect modification that may persist even in randomized studies [10,14,15]; we use the term bias here in reference to discrepancies between common measures of association, such as hazard ratios and risk ratios, and the per-

exposure biological effect of immunity (from vaccination or natural infection) on infection and/or disease endpoints [16,17]. Demonstrations of the impact of variation in susceptibility have arisen in both experimental and theoretical studies [18,19]. Refinements in our ability to characterize such variation both statistically and experimentally [20–27], together with formalizations of per-exposure measures of intervention efficacy in trials [28–30] and observational studies [16,17], have highlighted the potential for heterogeneity in susceptibility to influence epidemiologic measurements.

While the possibility of such frailty bias in rotavirus vaccine studies has been raised [15,31,32], distinguishing its contribution to variation in estimates of vaccine protection against RVGE has been difficult given the concordance of observed patterns with multiple hypotheses [33]. Importantly, the rate of asymptomatic infections and the distribution of risk factors across settings are not easily measured or compared [34], and individual variation in susceptibility may be only partially attributable to known or measured risk factors.

Similarly-designed birth cohort studies undertaken in socioeconomically-distinct LMIC populations of Mexico City, Mexico and Vellore, India provide an opportunity to characterize heterogeneity in susceptibility to rotavirus infection and RVGE, and to assess its influence on estimates of immune protection [35,36]. While the two studies supplied similar estimates of naturally-acquired immune protection against re-infection, differences in estimates of protection against RVGE reflected discrepancies in estimated vaccine efficacy between Latin America and South Asia [37–39]. Whereas no children in Mexico City experienced moderate-to-severe RVGE after two or more previous infections, two previous infections were associated with only 57% protection against moderate-to-severe RVGE among children in Vellore [35,36]. Paired re-analysis of the studies has provided evidence that differences owe, in part, to the influence of a subset of "high-risk" individuals in the Vellore cohort—who experienced high rates of rotavirus infection as well as high risk for RVGE given infection—and age-dependent risk for RVGE given infection [33].

We revisited data from these studies aiming to better understand and compare the distribution of susceptibility among individuals within the two cohorts, and to explore the implications for epidemiologic analyses. We developed a model to estimate susceptibility of children to rotavirus infection and RVGE, accounting for the natural history of rotavirus and differences between settings in transmission intensity. We conducted statistical inference via kernel-based and Markov chain Monte Carlo inference approaches, recovering near-identical parameter estimates under the two strategies. We used our findings to explore the influence of sources of bias underlying conventional measures of protective immunity.

Results

Cohort monitoring

Incidence of rotavirus infection and RVGE among children enrolled in the two cohorts has been described previously [33,35,36,40]. Briefly, the studies enrolled 200 and 373 unvaccinated Mexican and Indian children who were followed from birth to up to 2 years and 3 years of age, respectively, yielding 3699 and 13,937 child-months of follow-up, and characterized the spectrum of asymptomatic to severe clinical manifestations of each rotavirus infection. In total, 315 rotavirus infections were detected in Mexico City and 1103 were detected in Vellore, with 89 (28% of 315 infections) and 282 (26% of 1103 infections) episodes of RVGE occurring in the two settings, respectively. Incidence was higher in Vellore, such that first infections occurred in 56% and 81% of Indian children by ages 6 months and one year, compared to 34% and 67% of Mexican children, respectively (**Fig 1A and 1B**).
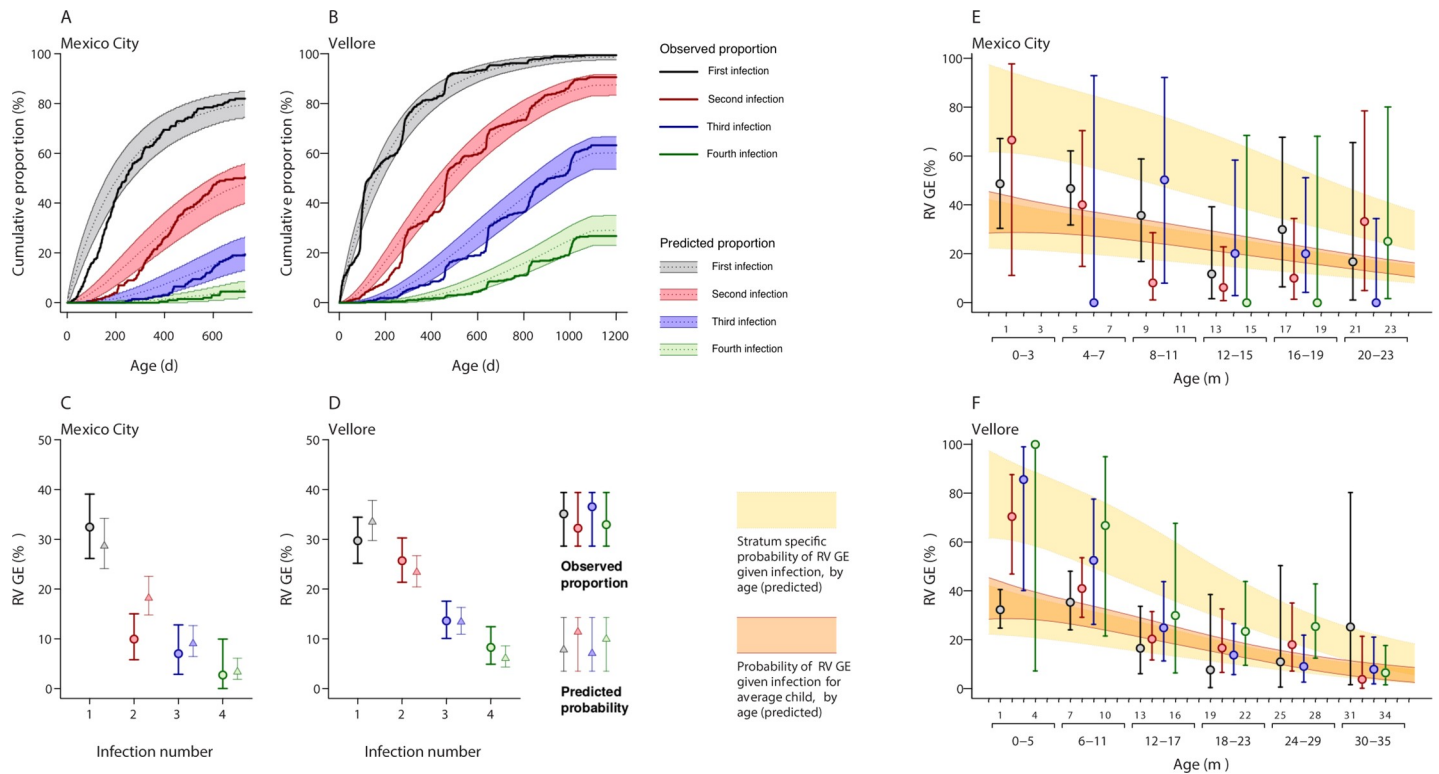
Fig 1. Natural history of rotavirus infection: Observations and model predictions. We plot cumulative incidence of first, second, third, and fourth rotavirus infections by age within the (**A**) Mexico City and (**B**) Vellore cohorts; shaded areas define 95% prediction intervals as fitted by the model, accounting for variation in risk among children and loss to follow-up, with dotted lines indicating median model predictions. (**C** and **D**) Observed proportions (circles, with 95% credible intervals) of infections involving RVGE, and model-predicted probabilities of RVGE given infection (triangles, with 95% prediction intervals), are plotted side by side. (**E** and **F**) We plot age-specific proportions of primary, secondary, third, and later infections involving RVGE (with 95% credible intervals); age strata within which such proportions are calculated correspond to the interval for serological sampling within the two cohorts (four months in Mexico City; six months in Vellore); these age bins are illustrated in the lower half of the x-axis. We superimpose these age-stratified estimates against the continuous decline in risk for infections to cause RVGE estimated by our fitted model (shaded peach); the lower estimate corresponds to the population of a low-risk sub-group (R), while the higher estimate corresponds to a higher-risk subgroup (R^C) whose risk for RVGE, given infection, is increased by a factor ρ (**Table 1**). The intermediate region (shaded orange) denotes risk for the average child within the cohorts. Shaded areas denote 95% prediction intervals around estimates of age-specific risk.

https://doi.org/10.1371/journal.pcbi.1007014.g001

The proportion of infections causing RVGE declined with a higher number of previous infections (**Fig 1C and 1D**). However, analyses stratified by age and previous infection revealed this trend could owe to confounding by age, i.e. declining RVGE risk with older age for each of first, second, and later infections (**Fig 1E and 1F**). At matched ages, RVGE was more common, paradoxically, during second and later infections than first infections in Vellore. In contrast, this trend was not apparent in Mexico City.

Rotavirus natural history

We developed a set of model structures addressing biological hypotheses of rotavirus natural history, based on previous studies of transmission dynamics [41–43] and secondary analysis of the birth-cohort datasets [33]. We estimated the proportion of each cohort belonging to a "high-risk" group, and tested for evidence of variation in susceptibility to infection and/or risk of RVGE given infection among the high-risk group compared to the rest of the cohort (see Materials and Methods). We also tested whether the risk of RVGE given infection varied depending on age at time of infection, and/or the number of previous infections.

Table 1. Parameter estimates and AIC values from models.

Measure		Definition	Estimate (95% CI)	
			Model 1[2]	**Model 2[2]**
α	Mexico City (M)	Proportion of population belonging to R subgroup	0.03 (0.01, 0.38)	0.04 (0.01, 0.46)
	Vellore (V)		0.15 (0.06, 0.39)	0.17 (0.07, 0.49)
$P(\alpha_M > \alpha_V)$		Test for difference between cohorts in prevalence of R subgroup (p value)	$p = 0.0594$	$p = 0.0822$
ϕ		Hazard ratio for infection in R subgroup	1.74 (1.27, 2.46)	1.63 (1.15, 2.37)
ρ		Relative risk for diarrhea (given infection) in R subgroup	2.39 (1.91, 3.14)	2.24 (1.69, 3.12)
β_0		Intercept term, function for RVGE risk given age	-1.05 (-1.31, -0.81)	-1.24 (-1.66, -0.84)
β_1		Linear term, function for RVGE risk given age[1]	-0.16 (-0.52, 0.20)	-0.33 (-0.80, 0.13)
β_2		Quadratic term, function for RVGE risk given age[1]	-0.54 (-1.06, -0.02)	-0.42 (-0.99, 0.14)
Λ	Mexico City (M)	Force of infection (baseline infection rate per child per year in the R^c group)	2.27 (1.54, 3.31)	2.11 (1.42, 3.19)
	Vellore (V)		2.55 (1.82, 3.69)	2.37 (1.67, 3.54)
$P(\Lambda_M > \Lambda_V)$		Test for difference in force of infection (p value)	$p = 0.0496$	$p = 0.0490$
ψ_k	$k = 0$	Hazard ratio for infection given previous exposure	1.00 (ref.)	1.00 (ref.)
	$k = 1$		0.67 (0.59, 0.77)	0.67 (0.59, 0.77)
	$k = 2$		0.50 (0.43, 0.58)	0.50 (0.43, 0.58)
	$k \geq 3$		0.36 (0.30, 0.45)	0.37 (0.30, 0.46)
ω_k	$k = 0$	Relative risk for RVGE during infection given previous exposure	1.00 (fixed)	1.00 (ref.)
	$k = 1$		- -	1.14 (0.91, 1.44)
	$k = 2$		- -	1.15 (0.80, 1.64)
	$k \geq 3$		- -	1.30 (0.83, 2.07)
AIC		Akaike information criterion (penalized measure of model fit)	21000.52	21005.38
$e^{-\Delta AIC/2}$		Relative model likelihood (AIC-penalized)	1	0.088

[1]Linear and quadratic age inputs are centered and scaled to unit variance in the original model.

[2]Model 2 estimates the effect of acquired immune from previous infections on risk of symptoms given reinfection (ω_k) controlling for age at infection; Model 1 assumes that the risk of RVGE given infection only depends on age at infection. A third model—wherein such an effect was limited to the R^C risk group—was not identifiable.

https://doi.org/10.1371/journal.pcbi.1007014.t001

We estimated 33% (95% CI: 23% to 41%), 50% (42% to 57%), and 64% (55% to 70%) reductions in the rates at which children re-acquired rotavirus after one, two, or three or more previous infections (**Table 1**), closely recapitulating estimated protection against re-infection in the original studies (**S2 Table**) [35,36].

Our models also captured declining risk for infections to cause symptomatic RVGE at older ages (**Fig 1B**). The proportion of secondary and subsequent infections causing RVGE in the first year of life in Vellore closely matched expectations among children classified as a "high-risk" subset of the population (detailed below; **Fig 1F**). We compared the fit of models with differing assumptions about acquired immune protection against RVGE given infection (**Table 1**). After accounting for declining risk of RVGE given infection at older ages, we did not identify improvements in fit (based on values of the Akaike Information Criterion [44]) when allowing for acquired immune protection against symptoms during second or later infections (Model 2).

Several salient differences between the two studies were reproduced in model-based predictions. Although we predicted higher-than-observed rates of infection in Mexico City during the first six months of life, predictions accurately reflected between-setting differences in cumulative incidence by the end of the first year (**Fig 1A and 1B**). In addition, fitted parameters recapitulated the observation of significantly lower probabilities of RVGE during second, third, and fourth infections in Mexico City as compared to Vellore (**Fig 1C and 1D**), despite predicting RVGE in a higher-than-observed proportion of second infections in Mexico City.

Variation in susceptibility among children

Our modeling framework partitioned the cohort populations across distinct risk groups (R and R^C) with prevalences α_M and $1-\alpha_M$, respectively, in Mexico City, and α_V and $1-\alpha_V$ in Vellore. Because the size of the risk group and group-specific relative risk for infection and/or disease outcomes are inversely related, the relative susceptibility and prevalence of these two risk groups were not simultaneously identifiable. We therefore estimated conditional between-group differences in susceptibility to infection (hazard ratio ϕ) and RVGE given infection (relative risk ρ) associated with particular values of α_M and α_V. We reconstructed the full distribution of ϕ and ρ from the marginal distributions of $\{\phi,\rho\}|\{\alpha_M,\alpha_V\}$ (see Materials and Methods). Fitting a more complex model (S1 Text) which considered an exhaustive set of risk groups—including children with modified susceptibility to infection only or disease only—allowed us to verify the hypothesis of a linkage between children's susceptibility to infection and disease given infection, which was suggested in previous analyses of the cohort data [33].

This modeling approach enabled us to compare the prevalence of children with particular susceptibility levels between cohorts (Fig 2). We estimated that 3% (1% to 23%) of children in Mexico City would belong to a high-risk-stratum experiencing a ≥50% higher-than-baseline rate of acquiring rotavirus infection, compared to 13% (6% to 29%) of children in Vellore (Fig 2G). A subgroup with over double the baseline rate of infection would include 2% (1% to 8%) of children in Mexico City and 10% (5% to 18%) of children in Vellore, while only 1% (0% to 2%) of children in Mexico City and 6% (5% to 9%) of children in Vellore would belong to a subgroup experiencing rates of infection ≥3-fold higher than the baseline rate. Greater susceptibility to infection was associated with higher risk of experiencing RVGE given infection, regardless of the prevalence of the high-risk group (Fig 2F); fitting both ϕ and ρ, we identified ≥99.99% probability for excess risk of disease given infection within the sub-cohorts defined to have higher rates of acquiring infection (S1 Table). Joint distributions of ϕ and ρ with the size of the risk groups were indistinguishable under the original model specification and a more complex model that allowed either linked or unlinked susceptibility to infection and disease (S1 Text, S2 Fig).

Determinants of susceptibility among individual children

Our modeling approach provided a statistical basis for calculating the probability that each child belonged to the "high-risk" subgroup (see Materials and Methods). To examine the validity of these estimates, we next assessed whether a child's estimated risk of belonging to the "high-risk" subgroup was related to host factors and exposures measured in the original studies that have previously been reported to predict risk for rotavirus infection and RVGE (Tables 2 and 3).

Male children were 26% (4% to 67%) more likely to be among the "high-risk" subgroup than female children (Table 3). Birth weight was also a predictor of being in the "high-risk" subgroup, with each log-kilogram decrease in birth weight conferring 2.05 (1.02 to 5.03)-fold higher probability of belonging to the "high-risk" subgroup. However, we did not detect a significant association between susceptibility and weight at 12 months. In each cohort as well as in the pooled analysis, children without siblings were more likely to belong to the "high-risk" subgroup than children with siblings. In comparison to children whose mothers had completed <5 years of education, children whose mothers had completed ≥10 years of education were 25% (0% to 50%) less likely to belong to the "high-risk" subgroup.

We also identified several factors predicting within-cohort variation in susceptibility that were consistent with findings in primary analyses of the studies [35,36]. In Vellore, children whose household members were involved in producing bidis (indigenous cigarettes)—an
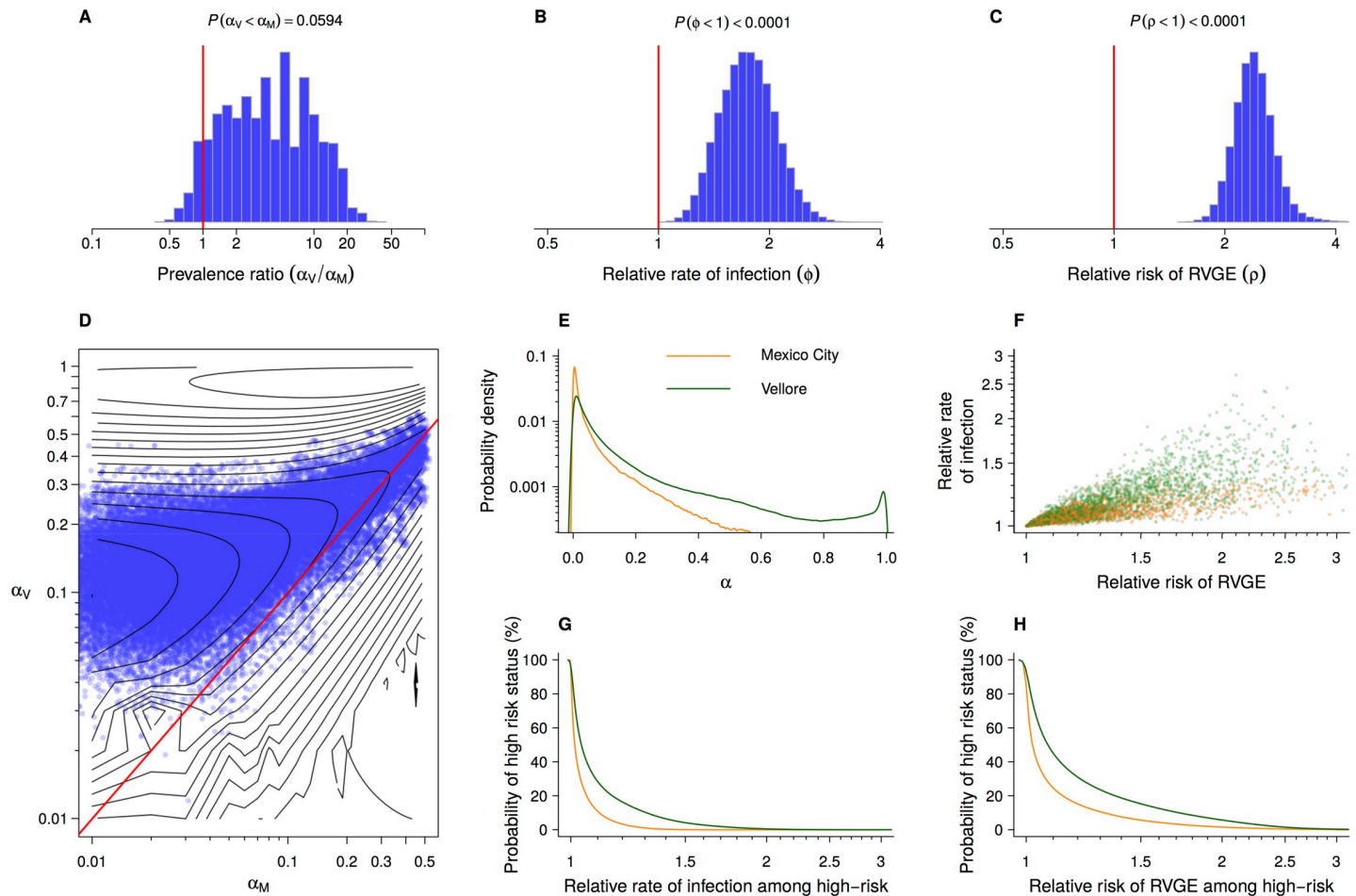
Fig 2. Heterogeneity in susceptibility to rotavirus infection and RVGE within and between cohorts. We plot estimates of (**A**) the relative prevalence of thge "high-risk" subgroup within Vellore (α_V) versus Mexico City (α_M). For the high-risk subgroup within each cohort, we illustrate (**B**) subgroup-specific hazard ratio of infection (ϕ) and (**C**) relative risk of RVGE given infection (ρ). (**D**) Sampled points (blue) illustrate the likelihood surface of the joint prevalence of subgroups within the Vellore and Mexico City cohorts; underlying isoclines illustrate the uneven likelihood surface on a natural log scale, while the preponderance of points above the red diagonal suggests that the high-risk subgroup is more prevalent in Vellore than in Mexico City (as indicated in **A**). Near-identical results arise under a model that allows either linked or unlinked susceptibility to infection and disease (**S1 Text**; **S2 Fig**, panel **E**). We next illustrate (**E**) estimates of the absolute sizes of the high-risk subgroups in Vellore (green) and Mexico City (yellow). Sampling from the joint distributions of α_V, α_M, ϕ, and ρ, (**F**) we generate simulated cohorts of children from Mexico City (orange points) and Vellore (green) to illustrate their relative susceptibility to infection and to RVGE, given infection. We next illustrate (**G**) the probability of "high-risk" status among children in Mexico City and Vellore associated with increasing values of ϕ, the hazard ratio of infection among those who belong to the "high-risk" group. The estimated probability of belonging to the "high-risk" group declines as the group is defined by a more-pronounced risk differential. Similarly, we illustrate (**H**) the declining probability of belonging to the "high-risk" group under increasing values of ρ, the relative risk of RVGE given infection among high-risk children.

https://doi.org/10.1371/journal.pcbi.1007014.g002

indicator of lower household socioeconomic status—were 45% (10% to 122%) more likely than other children to belong to the "high-risk" subgroup. In Mexico City, children with a shorter duration of breastfeeding were more likely to belong to the "high-risk" subgroup, although this association did not reach conventional thresholds of statistical significance in our analysis.

Among children experiencing RVGE in the cohorts, those who experienced moderate-to-severe RVGE symptoms (defined by a Vesikari score ≥11) on at least one episode were 44% (13% to 122%) more likely to belong to the "high-risk" subgroup, suggesting model-based measures of susceptibility to rotavirus infection and (any) RVGE given infection also predicted the severity of rotavirus disease. In addition, children who experienced higher rates of diarrheal episodes caused by pathogens other than rotavirus were more likely to belong to the

Table 2. Characteristics of the two cohorts.

Observation		Mexico City	Vellore
Birth weight	(kg)	3.2 (median; IQR: 2.9, 3.4)	3.0 (median; IQR: 2.7, 3.2)
Weight at 12 mo.	(kg)	10.0 (median; IQR: 9.1, 10.6)	7.9 (median; IQR: 7.3, 8.6)
Sex	Female	94 (47.0%)	186 (49.9%)
	Male	106 (53.0%)	187 (50.1%)
Duration of breastfeeding[1]			
	<3 months	59 (38.1%)	15 (4.1%)
	3–5 months	27 (17.4%)	27 (7.3%)
	6–11 months	31 (20.0%)	56 (15.1%)
	12+ months	38 (24.5%)	272 (73.5%)
Number of siblings			
	0	66 (33.8%)	118 (31.6%)
	1	56 (28.7%)	123 (33.0%)
	2	38 (19.5%)	90 (24.1%)
	3+	35 (17.9%)	42 (11.3%)
Socioeconomic status			
	Lowest	24 (12.3%)	– –
	Intermediate	81 (41.5%)	– –
	Highest	90 (46.2%)	– –
	Not bidi-working household	– –	200 (53.6%)
	Bidi-working household	– –	173 (46.4%)
Maternal education[2]			
	0–4 years	57 (29.2%)	106 (28.4%)
	5–7 years	62 (31.8%)	109 (29.2%)
	8–9 years	46 (23.6%)	90 (24.1%)
	≥10 years	30 (15.4%)	68 (18.2%)
RVGE and symptom severity	Children experiencing any RVGE episode(s)	77 (38.5%)	178 (47.7%)
	Children experiencing RVGE with Vesikari score ≥11	14 (7.0%)	58 (15.5%)
Episodes of rotavirus-negative diarrhea[3]		4 (median; IQR: 2, 7)	4 (median; IQR: 2, 6)
Episodes of respiratory infection		– –	20 (median; IQR: 14, 26)

All percentages are calculated among children from whom data were available for the given variable. IQR: interquartile range.

[1]Calculated among children retained for over 1 year in the study.

[2]Breakpoints for the years of maternal education were defined according to the following standard for Vellore: 0–4, no formal education or primary school incomplete; 5–7, primary school complete; 8–9, middle school complete; 10 or more, high school complete.

[3]Duration of follow-up differs by setting (up to 2 years in Mexico City, versus up to 3 years in Vellore).

https://doi.org/10.1371/journal.pcbi.1007014.t002

"high-risk" subgroup within each cohort. In Vellore, we also found a positive association between the incidence of acute respiratory infections and the likelihood that a child belonged to the "high-risk" subgroup; this information was not available for the Mexico City cohort.

Sources of variation in estimates of naturally-acquired immune protection

We next conducted simulation studies (Fig 3) to assess how variation in transmission intensity and in the susceptibility of children could influence estimates of naturally-acquired immune protection against rotavirus infection, RVGE, and RVGE given infection [45]—as measured by the hazard ratios of infection and RVGE, and relative risk of RVGE given infection—following one, two, or three previous infections, compared to zero previous infections. We compared estimates from *in silico* cohorts with differing prevalence of "high-risk" children (α)

Table 3. Factors associated with belonging to the "high-risk" group.

Risk factor		Relative Risk (95% CrI)		
		Mexico City	**Vellore**	**Pooled analysis**
Birth weight	per log kg	0.68 (0.42, 0.94)	0.84 (0.58, 1.13)	0.47 (0.20, 0.98)
Weight at 12 mo.	per log kg	0.88 (0.74, 1.01)	1.00 (0.85, 1.16)	0.68 (0.23, 1.66)
Sex	Female	ref.	ref.	ref.
	Male	1.05 (0.79, 1.42)	1.38 (1.08, 2.03)	1.26 (1.04, 1.67)
Duration of breastfeeding[1]				
	<3 months	ref.	ref.	ref.
	3–5 months	0.69 (0.41, 1.07)	1.12 (0.434, 3.04)	0.79 (0.50, 1.20)
	6–11 months	0.91 (0.54, 1.44)	1.16 (0.480, 2.86)	0.91 (0.59, 1.37)
	12+ months	0.70 (0.42, 1.05)	1.11 (0.485, 2.51)	0.83 (0.54, 1.17)
Number of siblings				
	0	ref.		ref.
	1	0.88 (0.59, 1.27)	0.72 (0.455, 0.98)	0.76 (0.52, 0.97)
	2	0.69 (0.42, 0.98)	0.79 (0.509, 1.16)	0.75 (0.53, 1.00)
	3+	0.49 (0.28, 0.74)	0.74 (0.413, 1.12)	0.61 (0.40, 0.83)
Socioeconomic status				
	Lowest	ref.	– –	– –
	Intermediate	1.20 (0.78, 1.92)	– –	– –
	Highest	1.17 (0.77, 1.78)	– –	– –
	Not bidi-working household	– –	ref.	– –
	Bidi-working household	– –	1.45 (1.095, 2.22)	– –
Maternal education				
	0–4 years (primary school incomplete)	ref.	ref.	ref.
	5–7 years (primary school complete)	1.21 (0.86, 1.80)	0.89 (0.62, 1.28)	0.99 (0.76, 1.32)
	8–9 years (middle school complete)	1.14 (0.77, 1.76)	0.90 (0.61, 1.35)	0.98 (0.73, 1.34)
	≥10 years (high school complete)	1.13 (0.75, 1.80)	0.62 (0.36, 0.90)	0.75 (0.50, 1.00)
Severity of RVGE episodes	Any episode with Vesikari score ≥11 (among children experiencing RVGE)	1.17 (0.82, 1.97)	1.52 (1.13, 2.55)	1.44 (1.13, 2.22)
Incidence of rotavirus-negative diarrhea	per log episodes/year	1.07 (1.02, 1.16)	1.38 (1.17, 1.72)	1.15 (1.07, 1.28)
Incidence of respiratory infections	per log episodes/year	– –	1.13 (1.10, 1.19)	– –

[1]Only children retained in studies for over 1 year are included in analyses.

exposed to varying forces of infection (Λ). We accounted for susceptibility differences between risk groups by sampling from the joint, unconditional distribution of {ϕ,ρ}, thereby isolating the effect of differences in risk-group prevalence.
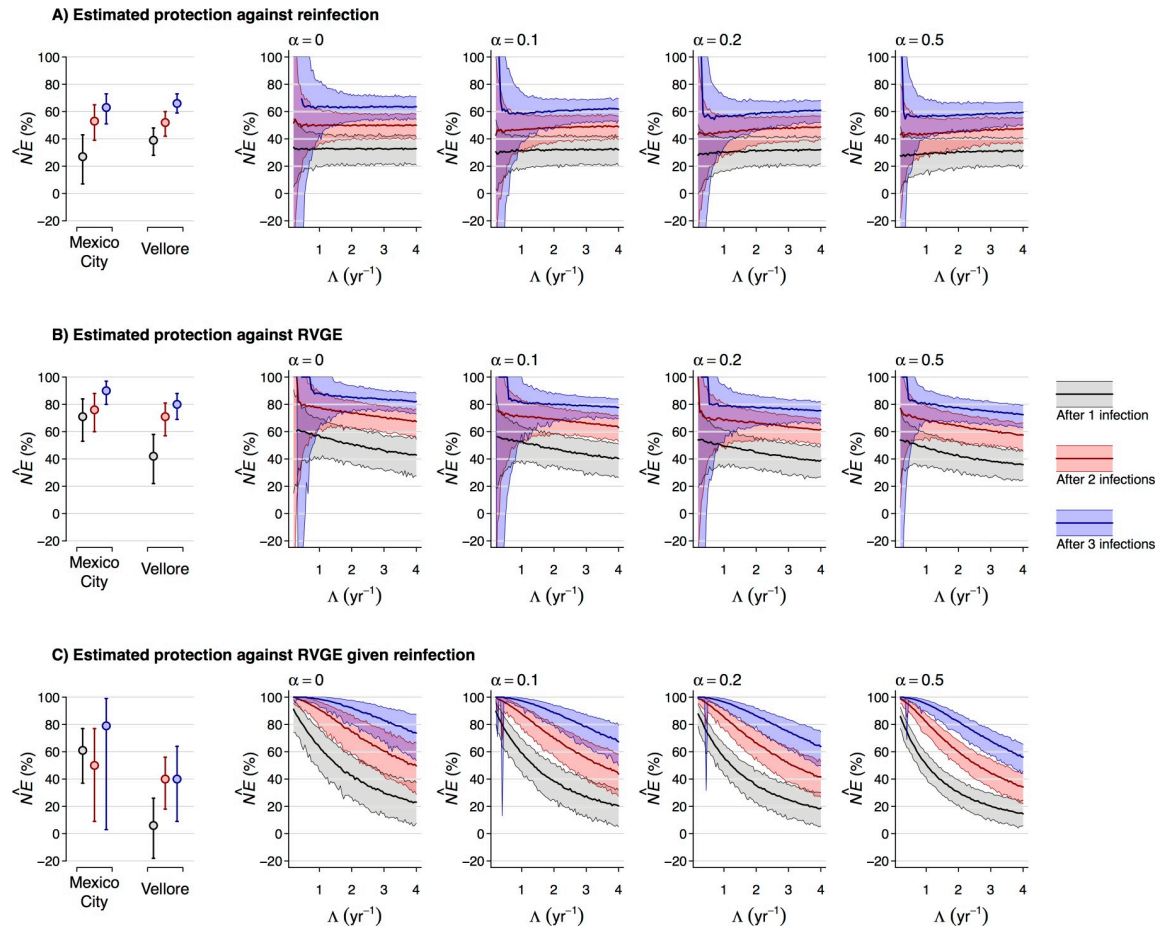
Fig 3. Sources of bias in estimates of naturally-acquired immune protection. We compare "naïve" estimates of naturally-acquired immunity ($\hat{NE}$) against re-infection (**A**), against RVGE incidence (**B**), and against RVGE given re-infection (**C**) following one (grey), two (red), or three (blue) previous infections in simulated cohorts. We parameterize the natural history model with estimates of natural protection against re-infection (ψ), age-specific risk of RVGE given infection (β), and excess infection and RVGE risk within the "high-risk" subgroup (ϕ, ρ) from Model 1 (**Table 1**, **Fig 2**). We vary the force of infection (Λ, measured as infections per susceptible person-year-at-risk) and the proportion of children assumed to belong to the "high-risk" subgroup (α). Consistent with primary analyses of the cohort data, simulated ("naïve") analyses do not account for age-dependent RVGE risk. The plotted values can be interpreted as predictions of estimated natural immunity against rotavirus infection, RVGE, and RVGE given infection, across sites with differing force of infection (x-axes) or prevalence of high-risk individuals (panels left to right). We plot estimates from the original cohort datasets within the left panels of each row (circles, with 95% confidence intervals), revealing that the higher force of infection and prevalence of "high-risk" individuals in Vellore can explain some, although not all, of the difference in estimates of immune protection between the Vellore and Mexico City cohorts.

https://doi.org/10.1371/journal.pcbi.1007014.g003

We did not identify a large of impact of susceptibility differences on estimates of protection against reinfection (i.e. estimates of $\hat{NE}$ were similar across different levels of Λ and α), which may help to explain why these estimates were nearly equal in the original studies [35,36]. However, we found that estimates of protection against RVGE—which were lower in primary analyses of the Vellore cohort—were expected to decline in settings with higher transmission intensity, reflecting acquisition of infection at younger, higher-risk ages. The impacts of heterogeneity in susceptibility were outweighed by the impacts of unaccounted-for age-dependent symptom risk. For a population exposed to transmission intensity on the order of one rotavirus infection per susceptible child-year at risk, increasing the prevalence of the "high-risk" subgroup (α) from 0% 50% reduced the estimate of protection against RVGE conferred

by one previous infection by 9% (–12% to 27%), in absolute terms. Increasing transmission intensity to the equivalent of four infections per susceptible child-year at risk led to a reduction of 14% (–8% to 34%) at $\alpha = 0$ and, similarly, of 12% (–6% to 27%) at $\alpha = 0.5$, in absolute terms.

Discussion

Evidence of naturally-acquired immunity against rotavirus from birth-cohort studies provided an impetus toward the development of live oral rotavirus vaccines, which are now among the most effective strategies for the prevention of severe illness and deaths due to RVGE globally [46]. However, challenges have persisted in understanding and addressing the lower protective efficacy of rotavirus vaccines in high-burden LMIC settings, which mirrors protection derived from naturally-acquired immunity [47,48]. Our analysis suggests that discrepant estimates of protection may in part reflect epidemiological bias, attributable to differences between settings in transmission intensity and differential susceptibility of children to rotavirus infection and RVGE, individually and by age. Lower estimates of protection in settings with high rotavirus burden thus reflect factors other than weaker immunity among children in LMICs.

Accounting for aspects of the natural history of rotavirus enabled us to directly compare the susceptibility of children enrolled in birth cohort studies undertaken in socioeconomically-distinct settings. Although we estimated only modestly higher susceptibility for the average-age child in Vellore as compared to Mexico City, individual variation in susceptibility was considerably greater within the Vellore cohort. We estimated that a higher proportion of children in Vellore, as compared to Mexico City, showed elevated rates of rotavirus infection as well as excess risk for RVGE given infection. This finding can account for several unexpected features of the epidemiology of rotavirus in Vellore. The increasing probability of RVGE in association with first, second, and later infections occurring at matched ages that we identified, particularly in Vellore (**Fig 1F**), reflects high risk for RVGE given infection among individuals susceptible to frequent rotavirus infection. In other words, children who experienced two or more infections before 6 months of age, or three or more infections before 12 months of age, are more likely to belong to a subgroup with pronounced susceptibility to rotavirus infection and disease, given infection. Indeed, our analysis identified that susceptibility to rotavirus infection was positively associated with susceptibility to RVGE given infection among individual children (**S1 Table**, **S2 Fig**). While the proportion of children belonging to a "high-risk" subgroup constituted a source of epidemiologic bias in simulation studies, and was expected to lead to estimates of weaker protection against RVGE in settings with higher transmission intensity such as Vellore, the degree of bias imposed was not large.

Several other studies have recently addressed transmission-dynamic factors that may contribute to the apparent underperformance of rotavirus vaccination in high-transmission settings [32]. Using data from the PROVIDE trial of monovalent rotavirus vaccine in Bangladesh, Rogawski and colleagues demonstrated that acquisition of naturally-acquired immunity may contribute to lower estimates of vaccine efficacy due to earlier and more frequent infection within the control arm; impacts on estimates of protection are most notable in high-transmission settings, and among children in their second year of life [15]. Here we were able to account for the contribution of all previous infections to naturally-acquired immunity, including subclinical infections, and to account for age-specific RVGE risk. Selection bias resulting from variation in individual susceptibility can result in further downward bias [31], underscoring the need for per-exposure estimates of immune effectiveness such as we have sought in this analysis. Directly comparing susceptibility between populations or settings is difficult because determinants of susceptibility are often unknown or unmeasured, and may be imperfectly characterized by measurable epidemiological risk factors. The contributions of susceptibility

and transmission intensity to disease incidence rates are not easily disentangled. Our analysis employed a novel approach to characterize susceptibility of children in two cohorts based on a model that included known aspects of rotavirus natural history, facilitated by access to similar measurements from settings with distinct risk profiles and force of infection.

Our estimates of susceptibility appear externally valid based on their association with previously-reported risk factors for RVGE [49–52]: male children, children with lower birth weight, and children whose mothers had lower educational attainment were more likely to belong to a higher-risk subset of the population in our analysis. In Vellore, children whose households were involved in bidi work—a marker for lower socioeconomic status—were also at higher risk [36], while in Mexico City, we observed a trend toward lower risk associated with longer breastfeeding, consistent with previous studies [53,54]. In addition, we observed higher incidence of diarrhea caused by pathogens other than rotavirus among children who were found to have greater susceptibility to rotavirus. This observation may signify the presence of environmental enteric dysfunction within the cohorts, or other sources of variation in immune status or pathogen exposure. In Vellore, children who we estimated were more susceptible to rotavirus also experienced higher incidence of respiratory infections, as reported previously [36]. While the associations we identify (in particular with time- or age-specific risk factors) do not measure causal effects in either direction, our inferences pertaining to within-cohort susceptibility are supported by the fact that children classified by the model as having "high risk" exhibit risk factors widely believed to be associated with rotavirus infection and RVGE.

These and other host factors associated with susceptibility to rotavirus infection and RVGE have also been reported to predict weaker immune responses to live enteric vaccines such as those against rotavirus. While our model does not address variation in the strength of immune responses among individuals or across settings, 58% of Indian children versus 90% of Mexican children seroconverted after Rotarix immunization in previous studies [47,55]. Nonetheless, near-equal naturally-acquired protection against re-infection was noted among children in the birth cohort studies in Vellore and Mexico City. Our findings demonstrate that some degree of the reported variation in protection against RVGE can be attributed to epidemiological biases resulting from differential transmission intensity and differential susceptibility of children, although we found age-dependent diarrhea risk was a more important contributor to variation in estimates.

The finding that older age diminishes risk for children to experience RVGE given rotavirus infection has been suggested in previous analyses of the cohort datasets [33]. Our simulation study demonstrates that such age-related symptom risk enhances protection against RVGE in low-transmission settings. Deferring infections to later ages significantly reduces the risk for children to experience symptoms upon reinfection. While the mechanisms underlying age-dependent diarrhea risk are not precisely known, the observation has been reported in mouse, rat, rabbit, and gnotobiotic piglet models of rotavirus infection [56–59]. Age-dependent TLR3 expression and host responses to rotavirus enterotoxins contribute to this observation in mice [60,61]. Other aspects of immune maturation, intestinal development, and the establishment of gut microbial communities may further drive associations between age and diarrhea risk in both humans and animals [62]. Furthermore, the greater dehydrating effect of diarrhea in younger children with smaller body volumes may contribute to severity—and thus the reporting and diagnosis—of RVGE in early-life infections [63].

There are several limitations to our analysis. Whereas we assume exponentially-distributed infection times (consistent with a constant hazard of infection), this provides an imperfect fit to the timing of early-life infections, particularly in the Mexico City cohort. The departure between predictions and observations may reflect the protective effect of maternal antibodies, as reported previously [35,64], or the influence of age-specific social mixing patterns on

transmission [65]. Thus, our model tended to overestimate the probability of RVGE associated with second rotavirus infections in Mexico City, although this discrepancy was not sustained for third and fourth infections. Our analyses also do not distinguish between homotypic and heterotypic protection because we lack genotype data for serologically-detected infections, which constitute the majority of infections in both cohorts. Although moderate-to-severe RVGE episodes are the primary endpoint of most studies evaluating vaccine efficacy and effectiveness, our analysis addressed RVGE episodes of any severity. Only 7% of children in Mexico City experienced RVGE episodes with Vesikari score $\geq$11, limiting the statistical power for analyses of moderate-to-severe RVGE. Nonetheless, previous analyses of the studies identified similar risk factors for mild and moderate-to-severe RVGE [33]; moreover, we find that children identified by our method to face higher risk for rotavirus infection and RVGE likewise experienced higher risk for moderate-to-severe manifestations of RVGE episodes. Thus, our findings may inform the interpretation of studies with moderate-to-severe RVGE endpoints.

Our ability to account for variation in susceptibility to infection as well as disease, and indeed to identify a linkage between these traits, is a unique advantage afforded by data describing both clinically-apparent and subclinical infections. While methods exist to account for frailty in time-to-event data [10,66], as may be present in studies with only one class of endpoints (such as serological studies of infection or clinical trials with disease endpoints), susceptibility to disease given infection is also of interest. Importantly, our findings suggest that age, rather than naturally-acquired immunity, determines risk for rotavirus infections to present symptomatically, together with individual-level susceptibility factors. Adaptations of our model to the natural history of other pathogens may facilitate similar studies in other disease-specific contexts. Whereas other models have used continuous distributions to characterize individual susceptibility, this approach has generally relied on the ability to measure or even manipulate exposure intensity at the individual level, for instance by measuring infectious contacts or through controlled-dose challenge experiments [20–30]. As our data presented the opportunity to compare exposure intensity between cohorts but not between individuals, we considered a simpler case of dichotomous risk groups within cohorts, and determined how the sizes of cohort-specific risk strata (α) were jointly distributed with the degree of risk elevation (ϕ,ρ). *A priori* knowledge of risk strata, for instance based on previous estimates of covariate effect sizes, presents an alternative strategy to infer distributions of individual-level susceptibility [67].

Birth-cohort studies have been instrumental to our understanding of the natural history of rotavirus. Uncertainties surrounding differences in the epidemiology of rotavirus in socioeconomically-distinct populations underscore the need for a theoretical basis for comparing outcomes of individual studies. Our approach permitted assessment of how age, acquired immunity, and variation in individual susceptibility independently contributed to infection and disease risk in distinct birth cohorts, helping to resolve discrepancies in estimated protection that arose in primary analyses of the datasets. The modeling framework we introduce here may thus have applicability to studies of other partially-immunizing pathogens.

Materials and methods

Birth cohort data

The two birth cohort studies followed similar protocols that have been described previously [35,36]. Children were enrolled at birth and followed to 24 and 36 months of age in Mexico City and Vellore, respectively. The studies aimed to detect all rotavirus infections, both symptomatic and asymptomatic. Rotavirus infections were detected by three approaches: (1) sera were drawn every 4 and 6 months in Mexico City and Vellore, respectively, and tested for IgA

or IgG titer increases; (2) asymptomatic stool samples were collected weekly in Mexico City and every two weeks in Vellore and tested for rotavirus; and (3) diarrheal stools were collected by field workers every time mothers alerted the study teams of any change in a child's stool pattern (**S2 Table**). Virus detection was performed by ELISA in Mexico City and by ELISA or real-time PCR in Vellore. In Mexico City, 200 children were recruited and retained for 77% of the scheduled follow-up period, while our analysis of the Vellore dataset included the 373 children (83% of 452 enrolled) who completed follow-up. Data were available for 96% (1037/1080) and 99% (2565/2598) of scheduled serum tests; 97% (15,503/16,029) and 93% (26,902/28,906) of scheduled asymptomatic stool tests; and 85% (963/1133) and 99% (1829/1856) of reported diarrheal episodes in Mexico City and Vellore, respectively.

Modeling rotavirus natural history

Summary. We developed a probabilistic model describing rotavirus natural history that allowed us to test how susceptibility to rotavirus infection and RVGE vary innately among individuals and according to age and previous infection. Seeking to account for differences in rotavirus epidemiology in the two settings under a unified model, we assumed parameters reflecting the strength of immune protection and the effect of age on the probability of symptoms given infection were equal in both settings, whereas we permitted the force of infection and the proportion of children belonging to distinct risk strata to vary between settings. We compared the improvement in fit afforded by modifications reflecting distinct hypotheses about the acquisition of protection against symptoms given infection—a source of uncertainty in rotavirus epidemiology [33]. Individual-level susceptibility of children to infection and RVGE was measured from the probability for children to belong to the high-risk stratum, which we also tested for associations with epidemiological risk factors. Last, we simulated from the fitted model to explore how differences in transmission intensity and the prevalence of high-risk individuals influence estimates of protection.

Model. We developed a model for time-to-event data that included asymptomatic and symptomatic rotavirus infection endpoints. Event time distributions were modeled from estimates of the force of infection and from rotavirus natural history parameters modifying the rate at which individuals acquired and re-acquired rotavirus infection, as well as the probability of experiencing RVGE given infection. Observations from previous analyses of the cohort studies suggested individuals differ in their susceptibility: in Vellore, children who experienced diarrhea on second and later rotavirus infections were more likely to have experienced diarrhea on earlier-life infections [33]. We therefore accounted for potential variation in risk among children by defining proportions α_V and α_M of children in Vellore and Mexico City, respectively, belonging to a distinct risk group R, and tested for group-wise differences in susceptibility to rotavirus infection and RVGE in the R stratum relative to the remaining population (R^c, with prevalence $1-\alpha_s$ in setting s). We defined the R stratum as whichever sub-group occupied $\leq 50\%$ of the Mexico City cohort, and thus, $0 \leq \alpha_M \leq 0.5$ and $0 \leq \alpha_V \leq 1$; the designation of the R stratum as lower- or higher-risk depended upon estimates of the relative susceptibility of children to infection and disease (ϕ and ρ, respectively; **Table 1**), the inverse of which would apply for the remaining children. In a supplemental analysis (**S1 Text**), we considered the possibility of additional risk strata, which included all pairwise combinations of modified susceptibility to infection and/or modified susceptibility to RVGE given infection.

Whereas primary analyses of the cohort datasets identified decreasing RVGE incidence rates following previous rotavirus infection, re-analyses suggested this declining risk resulted from naturally-acquired protection against rotavirus infection and older age at time of reinfection [33]. We used the Akaike Information Criterion (AIC) [44] to compare the fit of

alternative, nested models premised on differing assumptions about the influence of age and naturally-acquired immunity on risk of experiencing diarrhea symptoms during rotavirus infection. Model 1 assumed the risk of RVGE given infection did not vary based on the number of previous infections. Model 2 considered both age and previous infection as determinants of risk for RVGE given infection. We also attempted to fit a third model assuming that only children in the R^C (effectively, lower-risk) group acquired protection against symptoms given infection, but found that parameters were not identifiable; thus, only Models 1 and 2 were considered.

We modeled infection at the rate $\lambda_{R^c,s,k} = \Lambda_s \psi_k I_k$ among children in the majority (R^c) risk group, where Λ_s was the force of infection in Mexico City or Vellore, I_k was an indicator that an individual had experienced k previous rotavirus infections, and ψ_k was the hazard ratio for reinfection resulting from naturally-acquired immune protection following k previous infections. Among children in the minority (R) risk group, we modeled $\lambda_{R,s,k} = \phi\lambda_{R^c,s,k}$, where the hazard ratio ϕ conveyed the effect of differential susceptibility to infection.

The probability of RVGE given rotavirus infection, π, was modeled as a function of age and risk group under Model 1. For children of age x in the majority risk group,

$$\pi_{R^c}^{(1)}(x) = \exp(\beta_0 + \beta_1 x + \beta_2 x^2), \tag{1}$$

where the parameters β_{0-2} related RVGE risk to the child's age. Linear, exponential, and higher-order polynomial functions relating risk to age were explored, but found not to improve model fit via AIC. For the remaining children,

$$\pi_R^{(1)}(x) = \rho \exp(\beta_0 + \beta_1 x + \beta_2 x^2), \tag{2}$$

where the relative risk ρ accounted for innate differences in susceptibility to RVGE given infection.

Under Models 2 and 3, we further accounted for variation in symptom risk according to the number of previous infections a child in the R^c stratum has experienced, such that for Model 2,

$$\pi_{R^c k}^{(2)}(x) = \exp(\beta_0 + \beta_1 x + \beta_2 x^2)\omega_k I_k,$$
$$\pi_{Rk}^{(2)}(x) = \rho \exp(\beta_0 + \beta_1 x + \beta_2 x^2)\omega_k I_k, \tag{3}$$

where ω_k indicated the fold change in risk of RVGE given infection after having experienced k previous infections. For Model 3,

$$\pi_{R^c,k}^{(3)}(x) = \exp(\beta_0 + \beta_1 x + \beta_2 x^2)\omega_k I_k, \tag{4A}$$

whereas

$$\pi_{R,k}^{(3)}(x) = \rho \exp(\beta_0 + \beta_1 x + \beta_2 x^2) \tag{4B}$$

for all k.

Likelihood. We estimated model parameters in a likelihood-based framework. Observations included all infections or instances of censoring (indexed by j) for each child (indexed by i). For the jth observation of the ith child, we denoted time since the last observation or birth as Δt_{ij}, indicating whether the child was currently infected with rotavirus with the variable Z_{ij} (= 1 if infected and 0 if censoring occurred) and experiencing RVGE with the variable D_{ij} (= 1 if RVGE occurred upon infection and 0 otherwise). The likelihood contribution of observation j from child i, conditioned on the child belonging to either the R or R^c subgroup (*) and

residing in setting s, was

$$L(i,j|*,s) = f(\Delta t_{ij}|\lambda_{*sk_j})[\pi_{*k_j}(x_{ij})D_{ij} + (1 - \pi_{*k_j}(x_{ij}))(1 - D_{ij})]Z_{ij} + [1 - F(\Delta t_{ij}|\lambda_{*sk_j})](1 - Z_{ij}). \tag{5}$$

The term $f(\Delta t_{ij}|\lambda_{*sk_j})$ represents the probability that the time to infection for individual i was Δt_{ij}, while $1 - F(\Delta t_{ij}|\lambda_{*sk_j})$ was the probability for child i to have escaped infection from the time of the previous observation to the last follow-up visit, if censoring occurred at the jth observation. We assumed an exponentially-distributed time to infection for inference, consistent with an assumption of independent time-to-event observations given the child's exposure, infection history, and susceptibility status.

We obtained the likelihood contribution of each child, $H_s(i)$, via the total probability

$$H_s(i) = \alpha_s \prod_j L(i,j|\mathrm{R},s) + (1 - \alpha_s) \prod_j L(i,j|\mathrm{R}^c,s). \tag{6}$$

Defining $\mathbf{Y}_i$ as the set of observations from each child i, the probability for any particular child to be in the R risk group was

$$P(\mathrm{R}|\mathbf{Y}_i,s) = \frac{P(\mathrm{R} \cap \mathbf{Y}_i|s)}{P(\mathbf{Y}_i|s)} = \frac{\alpha_s P(\mathbf{Y}_i|\mathrm{R},s)}{\alpha_s P(\mathbf{Y}_i|\mathrm{R},s) + (1 - \alpha_s)P(\mathbf{Y}_i|\mathrm{R}^c,s)}$$
$$= \frac{\alpha_s \prod_j L(i,j|\mathrm{R},s)}{\alpha_s \prod_j L(i,j|\mathrm{R},s) + (1 - \alpha_s)\prod_j L(i,j|\mathrm{R}^c,s)}. \tag{7}$$

Our statistical framework assumed that children were sampled at random from the population and that observations $\mathbf{Y}_i$ were randomly-assigned given the setting, age, risk group, and infection history of an individual child i, such that the sample proportion of children in the R group was expected to converge to the population proportion.

We thus defined the overall model likelihood in two components. The first, $G_1^s = \prod_i H_s(i)$, conveyed the direct contributions of exponentially-distributed time-to-event observations $\mathbf{Y}_i$ as described above. The second component, G_2^s, reflected concordance between the population proportion, α_s, and the sample proportion, $\frac{1}{N}\sum_i P(\mathrm{R}|\mathbf{Y}_i,s)$ of children in the R risk group. We calculated G_2^s by evaluating the probability density of α_s under the assumption of random assignment:

$$\alpha_s \sim \mathrm{Beta}\{\sum_i P(\mathrm{R}|\mathbf{Y}_i,s), \sum_i[1 - P(\mathrm{R}|\mathbf{Y}_i,s)]\}. \tag{8}$$

Altogether, the model likelihood was the product of these terms across the two settings, $G_1^M G_1^V G_2^M G_2^V$.

Estimation. Estimated between-group differences in susceptibility to infection and disease (ϕ and ρ, respectively) were dependent upon the size of the underlying proportions α_M and α_V. For efficiency, we used a kernel-based inference approach whereby we reconstructed the global distribution of parameters from their conditional distributions $\theta = \{\phi, \rho, \psi_{1-3}, \omega_{1-3}, \beta_{0-2}, \Lambda_M, \Lambda_V\}$ given fixed values of $\{\alpha_M, \alpha_V\}$, defined "locally" for $\alpha_M = 0, 0.01, \ldots, 0.5$, and $\alpha_V = 0, 0.01, \ldots, 1$, respectively. For each set $\{\alpha_M, \alpha_V\}$, we determined the conditional maximum likelihood parameter estimates $\theta^*|\alpha_M, \alpha_V$ via the Nelder-Mead algorithm, using the optim() function in R:

$$\theta^*|\alpha_M, \alpha_V = \mathrm{argmin}_{\theta|\alpha_M, \alpha_V}(-\ln(G_1^M G_1^V G_2^M G_2^V)). \tag{9}$$

We assumed local parameter distributions at each set $\{\alpha_M, \alpha_V\}$ were multivariate-normal, with means defined by the conditional maximum likelihood estimates, and the asymptotic

covariance matrix defined from the inverse of the negative Hessian, as evaluated by minimizing the negative log likelihood. We pooled draws from these conditional (local) multivariate normal kernels over all $\{\alpha_M, \alpha_V\}$, weighting by their respective likelihoods, to recover the unconditional (global) parameter distributions.

Independent estimation of ϕ and ρ, conditional on $\{\alpha_M, \alpha_V\}$, further enabled us to test the hypothesis of a linkage between susceptibility to infection and disease, as suggested by the finding of higher diarrhea risk among children experiencing high rates of re-infection (**Fig 1E and 1F**) and previously-reported associations [33]. Estimation of these parameters allows us discriminate among hypotheses about susceptibility to infection and disease within the cohorts:

1. That there is no variation in susceptibility to infection or disease ($\phi = 1$ and $\rho = 1$);

2. That excess/diminished susceptibility to infection has no linkage to disease risk ($\phi \neq 1$ and $\rho = 1$), or that excess/diminished susceptibility to disease has no linkage to infection risk ($\rho \neq 1$ and $\phi = 1$);

3. That susceptibility to infection and disease are inversely correlated ($\rho > 1$ and $\phi < 1$, or $\rho < 1$ and $\phi > 1$); or

4. That susceptibility to infection and disease are, in fact, correlated ($\rho > 1$ and $\phi > 1$, or $\rho < 1$ and $\phi < 1$).

We measured the probability of a linkage between these traits, as suggested under Hypothesis 4, as

$$\Pr(\rho > 1 | \phi > 1)\Pr(\phi > 1) + \Pr(\rho < 1 | \phi < 1)\Pr(\phi < 1). \tag{10}$$

Supplemental analysis. We verified parameter estimates from this kernel-based approach in an additional analysis, wherein we sampled from the distribution of parameter estimates via Markov chain Monte Carlo using a Metropolis-Hastings updating procedure. Under this approach, we were able to explore an exhaustive set of risk strata wherein susceptibility to infection and to RVGE given infection could be linked or unlinked. This approach yielded near-identical parameter estimates and inferences in comparison to the original kernel-based approach, as we detail in the supporting files (**S1 Text**, **S1 Table**, **S1 Fig**, **S2 Fig** and **S3 Fig**); the size of the risk groups not considered in the original analysis converged to zero. We identified strong evidence against the null hypothesis of unlinked susceptibility to infection and disease, with posterior probability for the null hypothesis less than 6.7×10^{-6}.

Determinants of individual risk

To better understand variation in susceptibility among children within each cohort, we evaluated associations between individual-level factors (**Table 2**) and the probability for each child to belong to the high-risk (R) subgroup. For each of 10,000 draws of θ, we measured the probability of belonging to the high-risk group, equal to $P(R|\mathbf{Y}_i, s)$ for $\rho > 1$ or $1 - P(R|\mathbf{Y}_i, s)$ for $\rho < 1$, for each child (in all samples, we identified $\phi > 1$ for $\rho > 1$ and $\phi < 1$ for $\rho < 1$). We used least-squares regression to test for associations between covariates and children's log-transformed probability of being in the high-risk group, using estimated regression parameters to measure relative risks (**Table 3**). Models included a setting term to account for differential prevalence of high-risk children. We pooled relative risk estimates across our draws of θ to recover their distribution.

Simulation study

To examine potential bias in conventional estimates of naturally-acquired immune protection, we used our model of the natural history of rotavirus infection to simulate individual histories of infection and RVGE over the first three years of life, sampling from estimated parameters describing the effects of age ($\beta_0, \beta_1, \beta_2$) and previous infection (ψ_1, ψ_2, ψ_3) on susceptibility to RVGE and infection, respectively. We conducted simulations under an external force of infection (Λ) ranging from 0.2 to 4 infections per year, assigning 0% to 50% (α) of children to the high-risk subgroup R; values of ϕ and ρ were drawn independently of α so that we could determine the effect of differences in the proportion of high-risk children on estimates of protection. We sampled exponentially-distributed infection times (calculated from time of birth or previous infection), and defined the occurrence of RVGE for each individual infection as a Bernoulli random variable using the model-predicted probability of RVGE given infection.

For each cohort simulation, we measured the hazard ratio for reinfection and RVGE from the incidence rate (IR_k) of infection and RVGE after one, two, or three previous infections, relative to the IR_0 from birth. We also measured the relative risk of RVGE given reinfection among children who had experienced one, two, or three previous infections, relative to those with no previous infections, calculated from the proportion (p_k) of infections with RVGE. We defined estimates of natural immune efficacy ($\widehat{NE}$) as

$$\widehat{NE}_k = 1 - \frac{IR_k}{IR_0} \tag{11}$$

for protection against infection and RVGE among children who had experienced k previous infections, and

$$\widehat{NE}_k = 1 - \frac{p_k}{p_0} \tag{12}$$

for risk of RVGE given infection among children who had experienced k previous infections.

Supporting information

S1 Fig. Consistency of parameter estimates under the original and MCMC inference approaches (1 of 2). We illustrate parameter estimates under the original kernel-based approach (histograms) and from each of the 10 Markov chain Monte Carlo chains, overlaying their probability densities and presenting thinned draws from the parameter trace plots over 150,000 iterations (after 50,000 burn-in iterations). Parameters include: (**A**) hazard ratios for infection ψ_1, ψ_2, and ψ_3; (**B and C**) setting-specific force of infection for Vellore and Mexico City; and (**D, E, and F**) the polynomial terms describing age-specific diarrhea risk, given infection.
(TIF)

S2 Fig. Consistency of parameter estimates under the original and MCMC inference approaches (2 of 2). We illustrate parameter estimates and their joint distributions under the original kernel-based approach (histograms) and from each of the 10 Markov chain Monte Carlo chains (colored lines, as indicated in **S1 Fig**), overlaying their probability densities and presenting thinned draws from the parameter trace plots over 150,000 iterations (after 50,000 burn-in iterations). Parameters include (**A**) hazard ratio for infection ϕ, (**B**) relative risk of RVGE given infection ρ, (**C**) prevalence of the baseline risk group in Mexico City α_M^{00}, and (**D**) Vellore α_V^{00}. (**E**) We plot samples from the joint distribution of α_M^{11} and α_V^{11}, revealing concordance with the original estimates of the joint distribution of α_M and α_V plotted in **Fig 2D**. We

also illustrate concordance in samples from the joint distribution of the following parameters under the two approaches: (**F**) α_M^{11} and ϕ; (**G**) α_V^{11} and ϕ; (**H**) α_M^{11} and ρ; and (**I**) α_V^{11} and ρ.
(TIF)

S3 Fig. MCMC samples for risk groups. We present samples of the proportion of individuals belonging to the various risk groups from each of the 10 Markov chain Monte Carlo chains, overlaying their posterior distributions and presenting thinned draws from the parameter trace plots over 150,000 iterations (after 50,000 burn-in iterations). The first panels indicate the proportion of individuals belonging to the risk group with modified rates of acquiring infection (i.e., for whom ϕ applies) in (**A**) Vellore and (**B**) Mexico City, while the next two panels (**C** and **D**) illustrate the proportion of these individuals with modified risk of diarrhea given infection (i.e., for whom ρ applies). Among the proportion without modified rates of acquiring infection (i.e., for whom ϕ does not apply), the proportion who also do not experience modified risk of diarrhea given infection (i.e., for whom ρ does not apply) is illustrated in the final two panels (**E** and **F**). Convergence of the parameters α_M^{01}, α_V^{01}, α_M^{10}, and α_V^{10} to zero (see **S3 Table**) results in the concentration of probability mass very close to one across all chains (panels **C**–**F**).
(TIF)

S1 Table. Testing for linkages in susceptibility to infection and disease under Markov Chain Monte Carlo sampling. We present parameter estimates and their 95% credible intervals under the model including strata with distinct susceptibility to infection and disease (either or in combination), as well as posterior probabilities for hypothesis tests.
(DOCX)

S2 Table. Study design, enrollment, and follow-up. [1]Sample restricted to children who completed 3 years of follow-up (83% of initial cohort of 452 children) [2]The total number was not provided in the original study, but was calculated from the information that the 963 tested episodes represented 85% of the total reported episodes. [2]In the Mexico City cohort, the mean age of asymptomatic infections detected by shedding is lower than the mean age of asymptomatic infections detected by seroconversion alone ($p<0.0001$). [3]In the Vellore cohort, the mean age of asymptomatic infections detected by shedding is lower than the mean age of asymptomatic infections detected by seroconversion alone ($p<0.001$).
(DOCX)

S3 Table. Primary estimates of naturally-acquired immune protection. [1]Incidence is measured per 100 child-months at risk [2]The original studies applied differing definitions for moderate-to-severe RVGE; here we consider episodes with Vesikari score ≥ 11 to constitute moderate-to-severe RVGE.
(DOCX)

S1 Text. Supplemental materials. Text including supplemental methods, tables, and figures.
(PDF)

Acknowledgments

Author Contributions

Conceptualization: Joseph A. Lewnard, Benjamin A. Lopman, Umesh D. Parashar, Aisleen Bennett, Naor Bar-Zeev, Nigel A. Cunliffe, Virginia E. Pitzer.

Data curation: Prasanna Samuel, M. Lourdes Guerrero, Guillermo Ruiz-Palacios, Gagandeep Kang.

Formal analysis: Joseph A. Lewnard.

Funding acquisition: Virginia E. Pitzer.

Investigation: Joseph A. Lewnard, Prasanna Samuel, M. Lourdes Guerrero, Guillermo Ruiz-Palacios, Gagandeep Kang.

Methodology: Joseph A. Lewnard, Benjamin A. Lopman, Virginia E. Pitzer.

Project administration: Virginia E. Pitzer.

Supervision: Virginia E. Pitzer.

Visualization: Joseph A. Lewnard.

Writing – original draft: Joseph A. Lewnard.

Writing – review & editing: Benjamin A. Lopman, Umesh D. Parashar, Aisleen Bennett, Naor Bar-Zeev, Nigel A. Cunliffe, Prasanna Samuel, M. Lourdes Guerrero, Guillermo Ruiz-Palacios, Gagandeep Kang, Virginia E. Pitzer.

References

1. Fischer Walker CL, Rudan I, Liu L, Nair H, Theodoratou E, Bhutta Z a., et al. Global burden of childhood pneumonia and diarrhoea. Lancet. 2013; 381: 1405–1416. https://doi.org/10.1016/S0140-6736(13)60222-6 PMID: 23582727

2. Parashar UD, Johnson H, Steele AD, Tate JE. Health impact of rotavirus vaccination in developing countries: Progress and way forward. Clin Infect Dis. 2016; 62: S91–S95. https://doi.org/10.1093/cid/civ1015 PMID: 27059361

3. Jiang V, Jiang B, Tate J, Parashar UD, Patel MM. Performance of rotavirus vaccines in developed and developing countries. Hum Vaccin. 2010; 6: 532–42. https://doi.org/10.4161/hv.6.7.11278 PMID: 20622508

4. Kirkpatrick BD, Colgate ER, Mychaleckyj JC, Haque R, Dickson DM, Carmolli MP, et al. The "Performance of Rotavirus and Oral Polio Vaccines in Developing Countries" (PROVIDE) study: Description of methods of an interventional study designed to explore complex biologic problems. Am J Trop Med Hyg. 2015; 92: 744–751. https://doi.org/10.4269/ajtmh.14-0518 PMID: 25711607

5. Patel M, Steele AD, Parashar UD. Influence of oral polio vaccines on performance of the monovalent and pentavalent rotavirus vaccines. Vaccine. 2012;30. https://doi.org/10.1016/j.vaccine.2011.11.093 PMID: 22520134

6. Ali A, Kazi AM, Cortese MM, Fleming JA, Moon SS, Parashar UD, et al. Impact of withholding breastfeeding at the time of vaccination on the immunogenicity of oral rotavirus vaccine—A randomized trial. PLoS One. 2015; 10. https://doi.org/10.1371/journal.pone.0127622 PMID: 26035743

7. Mohan VR, Ramanujam K, Babji S, McGrath M, Shrestha S, Shrestha J, et al. Rotavirus infection and disease in a multi-site birth cohort: Results from the MAL-ED study. J Infect Dis. 2017; 1–11. https://doi.org/10.1093/infdis/jix199 PMID: 28472348

8. Perez-Schael I, Salinas B, Tomat M, Linhares AC, Guerrero ML, Ruiz-Palacios GM, et al. Efficacy of the human rotavirus vaccine RIX4414 in malnourished children. J Infect Dis. 2007; 196: 537–540. https://doi.org/10.1086/519687 PMID: 17624838

9. Levine MM, Kotloff KL, Nataro JP, Muhsen K. The Global Enteric Multicenter Study (GEMS): Impetus, rationale, and genesis. Clin Infect Dis. 2012;55. https://doi.org/10.1093/cid/cis761 PMID: 23169934

10. Halloran ME, Longini IM, Struchiner CJ. Estimability and interpretation of vaccine efficacy using frailty mixing models. Am J Epidemiol. 1996; 144: 83–97. https://doi.org/10.1093/oxfordjournals.aje.a008858 PMID: 8659489

11. Smith PG, Rodrigues LC, Fine PEM. Assessment of the protective efficacy of vaccines against common diseases using case-control and cohort studies. Int J Epidemiol. 1984; 13: 87–93. https://doi.org/10.1093/ije/13.1.87 PMID: 6698708

12. Greenwood M, Yule GU. An Inquiry into the Nature of Frequency Distributions representative of multiple happenings with particular reference to the occurrence of multiple attacks of disease or of repeated accidents. J R Stat Soc. 1920; 83: 255. https://doi.org/10.2307/2341080

13. Hernán MA. The hazards of hazard ratios. Epidemiology. 2010; 21: 13–15. https://doi.org/10.1097/EDE.0b013e3181c1ea43 PMID: 20010207

14. Morozova O, Cohen T, Crawford FW. Risk ratios for contagious outcomes. J R Soc Interface. 2018; 15. https://doi.org/10.1098/rsif.2017.0696 PMID: 29343627

15. Rogawski ET, Platts-Mills JA, Colgate ER, Haque R, Zaman K, Petri WA, et al. Quantifying the impact of natural immunity on rotavirus vaccine efficacy estimates: A Clinical trial in Dhaka, Bangladesh (PROVIDE) and a simulation study. J Infect Dis. 2018; 217: 861–868. https://doi.org/10.1093/infdis/jix668 PMID: 29514306

16. Halloran ME, Struchiner CJ. Causal inference in infectious diseases. Epidemiology. 1995; 6: 142–151. https://doi.org/10.1097/00001648-199503000-00010 PMID: 7742400

17. O'Hagan JJ, Lipsitch M, Hernán MA. Estimating the per-exposure effect of infectious disease interventions. Epidemiology. 2014; 25: 134–138. https://doi.org/10.1097/EDE.0000000000000003 PMID: 24240656

18. Furumoto WA, Mickey R. A mathematical model for the infectivity-dilution curve of tobacco mosaic virus: Theoretical considerations. Virology. 1967; 32: 216–223. https://doi.org/10.1016/0042-6822(67)90271-1 PMID: 6025875

19. Furumoto WA, Mickey R. Mathematical analyses of the interference phenomenon of tobacco mosaic virus: Experimental tests. Virology. 1970; 40: 322–328. https://doi.org/10.1016/0042-6822(70)90408-3 PMID: 5440441

20. Langwig KE, Wargo AR, Jones DR, Viss JR, Rutan BJ, Egan NA, et al. Vaccine effects on heterogeneity in susceptibility and implications for population health management. MBio. 2017; 8. https://doi.org/10.1128/mBio.00796-17 PMID: 29162706

21. King JG, Souto-Maior C, Sartori LM, Maciel-De-Freitas R, Gomes MGM. Variation in *Wolbachia* effects on *Aedes* mosquitoes as a determinant of invasiveness and vectorial capacity. Nat Commun. 2018; 9. https://doi.org/10.1038/s41467-018-03981-8 PMID: 29662096

22. Pessoa D, Souto-Maior C, Gjini E, Lopes JS, Ceña B, Codeço CT, et al. Unveiling time in dose-response models to infer host susceptibility to pathogens. PLoS Comput Biol. 2014; 10. https://doi.org/10.1371/journal.pcbi.1003773 PMID: 25121762

23. Gomes MGM, Lipsitch M, Wargo AR, Kurath G, Rebelo C, Medley GF, et al. A missing dimension in measures of vaccination impacts. PLoS Pathog. 2014; 10. https://doi.org/10.1371/journal.ppat.1003849 PMID: 24603721

24. Abel S, Abel Zur Wiesch P, Chang HH, Davis BM, Lipsitch M, Waldor MK. Sequence tag-based analysis of microbial population dynamics. Nat Methods. 2015; 12: 223–226. https://doi.org/10.1038/nmeth.3253 PMID: 25599549

25. Abel S, Abel zur Wiesch P, Davis BM, Waldor MK. Analysis of bottlenecks in experimental models of Infection. PLoS Pathogens. 2015. https://doi.org/10.1371/journal.ppat.1004823 PMID: 26066486

26. Gomes MGM, Águas R, Lopes JS, Nunes MC, Rebelo C, Rodrigues P, et al. How host heterogeneity governs tuberculosis reinfection? Proc R Soc B Biol Sci. 2012; 279: 2473–2478. https://doi.org/10.1098/rspb.2011.2712 PMID: 22357260

27. Lewnard JA, Townsend JP. Climatic and evolutionary drivers of phase shifts in the plague epidemics of colonial India. Proc Natl Acad Sci. 2016; 113: 14601–14608. https://doi.org/10.1073/pnas.1604985113 PMID: 27791071

28. Hudgens MG, Gilbert PB. Assessing vaccine effects in repeated low-dose challenge experiments. Biometrics. 2009; 65: 1223–1232. https://doi.org/10.1111/j.1541-0420.2009.01189.x PMID: 19397589

29. Hudgens MG, Gilbert PB, Mascola JR, Wu C-D, Barouch DH, Self SG. Power to detect the effects of HIV vaccination in repeated low-dose challenge experiments. J Infect Dis. 2009; 200: 609–613. https://doi.org/10.1086/600891 PMID: 19591571

30. Long DM, Hudgens MG, Wu C Da. Surrogates of protection in repeated low-dose challenge experiments. Stat Med. 2015; 34: 1747–1760. https://doi.org/10.1002/sim.6436 PMID: 25628249

31. Gomes MGM, Gordon SB, Lalloo DG. Clinical trials: The mathematics of falling vaccine efficacy with rising disease incidence. Vaccine. 2016; 34: 3007–9. https://doi.org/10.1016/j.vaccine.2016.04.065 PMID: 27177948

32. Lopman BA, Pitzer VE. Waxing understanding of waning immunity. J Infect Dis. 2018; 217: 851–853. https://doi.org/10.1093/infdis/jix670 PMID: 29394364

33. Lewnard JA, Lopman BA, Parashar UD, Bar-Zeev N, Samuel P, Guerrero ML, et al. Naturally-acquired immunity against rotavirus infection and gastroenteritis in children: paired re-analyses of birth-cohort studies. J Infect Dis. 2017; https://doi.org/10.1093/infdis/jix310 PMID: 28859432

34. Acosta AM, Chavez CB, Flores JT, Olotegui MP, Pinedo SR, Trigoso DR, et al. The MAL-ED Study: A multinational and multidisciplinary approach to understand the relationship between enteric pathogens, malnutrition, gut physiology, physical growth, cognitive development, and immune responses in infants and children up to 2 years of age in resource-poor environments. Clin Infect Dis. 2014; 59: S193–S206. https://doi.org/10.1093/cid/ciu653 PMID: 25305287

35. Velázquez FR, Matson DO, Calva JJ, Guerrero L, Morrow a L, Carter-Campbell S, et al. Rotavirus infections in infants as protection against subsequent infections. N Engl J Med. 1996; 335: 1022–1028. https://doi.org/10.1056/NEJM199610033351404 PMID: 8793926

36. Gladstone BP, Ramani S, Mukhopadhya I, Muliyil J, Sarkar R, Rehman AM, et al. Protective effect of natural rotavirus infection in an Indian birth cohort. N Engl J Med. 2011; 365: 337–346. https://doi.org/10.1056/NEJMoa1006261 PMID: 21793745

37. Ruiz-Palacios GM, Perez-Schael I, Velazquez FR, Abate H, Breuer T, Clemens SC, et al. Safety and efficacy of an attenuated vaccine against severe rotavirus gastroenteritis. N Engl J Med. 2006; 354: 11–22. https://doi.org/10.1056/NEJMoa052434 PMID: 16394298

38. Zaman K, Anh DD, Victor JC, Shin S, Yunus M, Dallas MJ, et al. Efficacy of pentavalent rotavirus vaccine against severe rotavirus gastroenteritis in infants in developing countries in Asia: A randomised, double-blind, placebo-controlled trial. Lancet. 2010; 376: 615–623. https://doi.org/10.1016/S0140-6736(10)60755-6 PMID: 20692031

39. Bhandari N, Rongsen-Chandola T, Bavdekar A, John J, Antony K, Taneja S, et al. Efficacy of a monovalent human-bovine (116E) rotavirus vaccine in Indian infants: A randomised, double-blind, placebo-controlled trial. Lancet. 2014; 383: 2136–2143. https://doi.org/10.1016/S0140-6736(13)62630-6 PMID: 24629994

40. Raul Velazquez F, Calva JJ, Lourdes Guerrero M, Mass D, Glass RI, Pickering LK, et al. Cohort study of rotavirus serotype patterns in symptomatic and asymptomatic infections in Mexican children. Pediatr Infect Dis J. 1993; 12: 54–61. https://doi.org/10.1097/00006454-199301000-00012 PMID: 8380235

41. Pitzer VE, Viboud C, Simonsen L, Steiner C, Panozzo CA, Alonso WJ, et al. Demographic variability, vaccination, and the spatiotemporal dynamics of rotavirus epidemics. Science. 2009; 325: 290–4. https://doi.org/10.1126/science.1172330 PMID: 19608910

42. Pitzer VE, Atkins KE, de Blasio BF, van Effelterre T, Atchison CJ, Harris JP, et al. Direct and indirect effects of rotavirus vaccination: Comparing predictions from transmission dynamic models. PLoS One. 2012; 7. https://doi.org/10.1371/journal.pone.0042320 PMID: 22912699

43. Atkins KE, Shim E, Carroll S, Quilici S, Galvani AP. The cost-effectiveness of pentavalent rotavirus vaccination in England and Wales. Vaccine. 2012; 30: 6766–6776. https://doi.org/10.1016/j.vaccine.2012.09.025 PMID: 23000223

44. Akaike H. A new look at the statistical model identification. IEEE Trans Automat Contr. 1974; 19: 716–723. https://doi.org/10.1109/TAC.1974.1100705

45. Halloran ME, Struchiner CJ, Longini IM. Study designs for evaluating different efficacy and effectiveness aspects of vaccines. Am J Epidemiol. 1997; 146: 789–803. https://doi.org/10.1093/oxfordjournals.aje.a009196 PMID: 9384199

46. Patel MM, Steele D, Gentsch JR, Wecker J, Glass RI, Parashar UD. Real-world impact of rotavirus vaccination. Pediatr Infect Dis J. 2011; 30: S1–S5. https://doi.org/10.1097/INF.0b013e3181fefa1f PMID: 21183833

47. Patel M, Shane AL, Parashar UD, Jiang B, Gentsch JR, Glass RI. Oral rotavirus vaccines: how well will they work where they are needed most? J Infect Dis. 2009; 200: S39–S48. https://doi.org/10.1086/605035 PMID: 19817613

48. Lopman BA, Pitzer VE, Sarkar R, Gladstone B, Patel M, Glasser J, et al. Understanding reduced rotavirus vaccine efficacy in low socio-economic settings. PLoS One. 2012;7. https://doi.org/10.1371/journal.pone.0041720 PMID: 22879893

49. Dennehy PH, Cortese MM, B??gu?? RE, Jaeger JL, Roberts NE, Zhang R, et al. A case-control study to determine risk factors for hospitalization for rotavirus gastroenteritis in U.S. children. Pediatr Infect Dis J. 2006; 25: 1123–1131. https://doi.org/10.1097/01.inf.0000243777.01375.5b PMID: 17133157

50. Newman RD, Grupp-Phelan J, Shay DK, Davis RL. Perinatal risk factors for infant hospitalization with viral gastroenteritis. Pediatrics. 1999; 103: 1–6. https://doi.org/10.1542/peds.103.1.1

51. Victoria CG, Huttly SRA, Barros FC, Lombardi C, Vaughan JP. Maternal education in relation to early and late child health outcomes: Findings from a Brazilian cohort study. Soc Sci Med. 1992; 34: 899–905. https://doi.org/10.1016/0277-9536(92)90258-R PMID: 1604379

52. Naficy AB, Abu-Elyazeed R, Holmes JL, Rao MR, Savarino SJ, Kim Y, et al. Epidemiology of rotavirus diarrhea in Egyptian children and implications for disease control. Am J Epidemiol. 1999; 150: 770–7. https://doi.org/10.1093/oxfordjournals.aje.a010080 PMID: 10512431

53. Plenge-Bönig A, Soto-Ramírez N, Karmaus W, Petersen G, Davis S, Forster J. Breastfeeding protects against acute gastroenteritis due to rotavirus in infants. Eur J Pediatr. 2010; 169: 1471–1476. https://doi.org/10.1007/s00431-010-1245-0 PMID: 20617343

54. Newburg DS, Peterson JA, Ruiz-Palacios GM, Matson DO, Morrow AL, Shults J, et al. Role of human-milk lactadherin in protection against symptomatic rotavirus infection. Lancet. 1998; 351: 1160–1164. https://doi.org/10.1016/s0140-6736(97)10322-1 PMID: 9643686

55. Patel M, Glass RI, Jiang B, Santosham M, Lopman B, Parashar U. A systematic review of anti-rotavirus serum IgA antibody titer as a potential correlate of rotavirus vaccine efficacy. Journal of Infectious Diseases. 2013. pp. 284–294. https://doi.org/10.1093/infdis/jit166 PMID: 23596320

56. Ciarlet M, Conner ME, Finegold MJ, Estes MK. Group A rotavirus infection and age-dependent diarrheal disease in rats: a new animal model to study the pathophysiology of rotavirus infection. J Virol. 2002; 76: 41–57. https://doi.org/10.1128/JVI.76.1.41-57.2002 PMID: 11739670

57. Ciarlet M, Gilger M a, Barone C, McArthur M, Estes MK, Conner ME. Rotavirus disease, but not infection and development of intestinal histopathological lesions, is age restricted in rabbits. Virology. 1998; 251: 343–60. https://doi.org/10.1006/viro.1998.9406 PMID: 9837799

58. Ramig RF. The effects of host age, virus dose, and virus strain on heterologous rotavirus infection of suckling mice. Microb Pathog. 1988; 4: 189–202. https://doi.org/10.1016/0882-4010(88)90069-1 PMID: 2848173

59. Saif LJ, Ward LA, Yuan L, Rosen BI, To TL. The gnotobiotic piglet as a model for studies of disease pathogenesis and immunity to human rotaviruses. Arch Virol Suppl. 1996. 153–161. https://doi.org/10.1007/978-3-7091-6553-9_17 PMID: 9015112

60. Pott J, Stockinger S, Torow N, Smoczek A, Lindner C, McInerney G, et al. Age-dependent TLR3 expression of the intestinal epithelium contributes to rotavirus susceptibility. PLoS Pathog. 2012; 8. https://doi.org/10.1371/journal.ppat.1002670 PMID: 22570612

61. Ball JM, Tian P, Zeng CQ, Morris AP, Estes MK. Age-dependent diarrhea induced by a rotaviral non-structural glycoprotein. Science. 1996; 272: 101–4. https://doi.org/10.1126/science.272.5258.101 PMID: 8600515

62. Fulde M, Hornef MW. Maturation of the enteric mucosal innate immune system during the postnatal period. Immunological Reviews. 2014. pp. 21–34. https://doi.org/10.1111/imr.12190 PMID: 24942679

63. Cohen MB. Etiology and mechanisms of acute infectious diarrhea in infants in the United States. J Pediatr. 1991;118. https://doi.org/10.1016/S0022-3476(05)81423-4

64. Velázquez FR, Matson DO, Guerrero ML, Shults J, Calva JJ, Morrow AL, et al. Serum antibody as a marker of protection against natural rotavirus infection and disease. J Infect Dis. 2000; 182: 1602–9. https://doi.org/10.1086/317619 PMID: 11069230

65. Mossong J, Hens N, Jit M, Beutels P, Auranen K, Mikolajczyk R, et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS Med. 2008; 5: 0381–0391. https://doi.org/10.1371/journal.pmed.0050074 PMID: 18366252

66. Mcgilchrist CA, Aisbett CW. Regression with frailty in survival analysis. Biometrics. 1991; 4716191: 461–466. https://doi.org/10.2307/2532138

67. Bareinboim E, Pearl J. Causal inference and the data-fusion problem. Proc Natl Acad Sci. 2016; 113: 7345–7352. https://doi.org/10.1073/pnas.1510507113 PMID: 27382148