Supplementary Material

Alicia N. M. Kraay, Edward L. Ionides, Gwenyth O. Lee, William F. Cevallos Trujillo, Joseph N.S. Eisenberg

April 13, 2020

S1 Vaccine Data Collection Details

We collected vaccine data from a local health system. Vaccine records were available for 21 communities. The fraction of community members with vaccine records is shown in table S-1 and the coverage of two doses of vaccine among children with records by community is shown in table S-2. For both tables, communities that were included in the case-control analysis are shown in bold. In general, coverage was highest in close villages (see Table S-3). Most children had completed their vaccine dosing by six months of age. Of children who ultimately received only one dose, almost all had received this dose by six months of age and most children who received both doses of vaccine had received both doses by 6 months of age (Table S-4). Based on this information, we calculated vaccine coverage among children with records using the number of children with vaccine records who were six months old in a given case-control study cycle as the denominator and the numerator was the number of age-eligible children who actually received their second dose of vaccine. While the number of children with vaccine records in each study community was small, this constituted a substantial fraction of the population. For the communities included in the case-control analysis, the availability of vaccine records was higher on average, with average record availability of 44%, 58%, 79% and 69% of communities included in the analysis for cycles 8, 9, 10, and 11 respectively (community 2 is not included in the record availability calculation because we did not include cycles in which the number of vaccine records was 0 in the regression analysis).

S2 Vaccine coverage sensitivity analysis

Vaccine records were sparse for some communities and for some cycles which could have resulted in imprecise estimates of coverage and could have resulted in biased relative estimates if coverage was much lower or higher than expected. To better understand the implication of this bias, we ran a sensitivity analysis where we set vaccine coverage to a particular coverage level (0%, 25%, 50%, or 100%) for children without records instead of assuming that coverage was similar among children with or without vaccine records. We then recreated the effect estimates in Table 5 using the new, bias corrected exposure variable. To account for biased reporting, after assigning the new coverage level to children without vaccine records, we reestimated the overall community level coverage using the following formula:

Community	Cycle 8	Cycle 9	Cycle 10	Cycle 11
Close Communities				
1	**58.3 (7/12)**	**45.0 (9/20)**	**44.4 (8/18)**	**81.8 (9/11)**
2	**43.8 (7/16)**	**61.9 (13/21)**	**34.3 (12/35)**	**0 (0/22)**
3	**44.4 (12/27)**	**67.8 (25/37)**	**80.0 (28/35)**	**85.0 (17/20)**
4	0 (0/7)	45.5 (5/11)	50.0 (8/16)	0 (0/14)
5	30.0 (3/10)	40.0 (4/10)	44.4 (8/18)	0 (0/12)
Total	40.2 (29/72)	56.6 (56/99)	52.5 (64/122)	32.9(26/79)
Medium Communities				
6	50.0 (1/2)	100 (2/2)	50.0 (2/4)	100 (1/1)
7	**66.7 (2/3)**	**54.5 (6/11)**	**100 (6/6)**	**60.0 (3/5)**
8	100 (1/1)	100 (3/3)	40.0 (2/5)	0 (0/7)
9	50.0 (1/2)	80.0 (4/5)	66.7 (2/3)	50.0 (1/2)
10	100 (1/1)	66.7 (2/3)	50.0 (2/4)	100 (1/1)
Total	66.7 (6/9)	70.8 (17/24)	63.6 (14/22)	37.5 (6/16)
Far Communities				
11	**19 (3/16)**	**36.4 (4/11)**	**29.4 (5/17)**	**50.0 (3/6)**
12	100 (7/7)	100 (6/6)	66.7 (2/3)	100 (4/4)
13	0 (0/1)	33.3 (1/3)	75.0 (3/4)	33.3 (1/3)
14	**62.5 (5/8)**	**60.0 (15/25)**	**45.0 (9/20)**	**46.2 (6/13)**
15	75.0 (3/4)	0 (0/3)	0 (0/3)	0 (0/8)
16	0 (0/3)	12.5 (1/8)	40.0 (4/10)	71.4 (5/7)
17	76.9 (10/13)	66.7 (12/18)	50.0 (6/12)	75.0 (12/16)
18	0 (0/2)	40.0 (2/5)	50.0 (2/4)	60.0 (3/5)
19	60.0 (3/5)	50.0 (3/6)	25.0 (5/20)	46.2 (6/13)
20	0 (0/4)	0 (0/3)	0 (0/7)	36.4 (4/11)
21	0 (0/12)	18.2 (2/11)	30.0 (6/20)	72.7 (8/11)
Total	41.3 (31/75)	47.4 (46/97)	35.0 (42/120)	53.6(52/97)

Table S-1: Vaccine record availability by community. Each cell is presented as Percent with Vaccine Records (Number with Vaccine Records/Number Eligible). Communities that were included in the rotavirus positivity analysis are shown in bold

Community	Cycle 8	Cycle 9	Cycle 10	Cycle 11
Close Communities				
1	**85.7 (6/7)**	**77.8 (7/9)**	**87.5 (7/8)**	**88.9 (8/9)**
2	**57.1 (4/7)**	**46.2 (6/13)**	**50.0 (6/12)**	**N/A**
3	**100 (12/12)**	**88.0 (22/25)**	**100.0 (28/28)**	**100.0 (17/17)**
4	N/A	20.0 (1/5)	37.5 (3/8)	N/A
5	100 (3/3)	25.0 (1/4)	37.5 (3/8)	N/A
Total	86.2 (25/29)	66.1 (37/56)	73.4 (47/64)	96.2 (25/26)
Medium Communities				
6	0 (0/1)	100 (2/2)	100 (2/2)	100 (1/1)
7	**50.0 (1/2)**	**33.3 (2/6)**	**100 (6/6)**	**100 (3/3)**
8	0 (0/1)	66.7 (2/3)	50.0 (1/2)	N/A
9	100 (1/1)	75.0 (3/4)	50.0 (1/2)	100 (1/1)
10	100 (1/1)	100 (2/2)	100 (2/2)	100 (1/1)
Total	50.0 (3/6)	64.7 (11/17)	85.7 (12/14)	100 (6/6)
Far Communities				
11	**100 (3/3)**	**100 (4/4)**	**100 (5/5)**	**100 (3/3)**
12	57.1 (4/7)	50.0 (3/6)	50.0 (1/2)	75.0 (3/4)
13	N/A	100 (1/1)	66.7 (2/3)	100 (1/1)
14	**100 (5/5)**	**73.3 (11/15)**	**88.9 (8/9)**	**83.3 (5/6)**
15	66.7 (2/3)	N/A	N/A	N/A
16	N/A	100 (1/1)	100 (4/4)	60.0 (3/5)
17	70.0 (7/10)	66.7 (8/12)	66.7 (4/6)	100 (12/12)
18	N/A	50.0 (1/2)	100 (2/2)	33.3 (1/3)
19	100 (3/3)	66.7 (2/3)	60.0 (3/5)	33.3 (2/6)
20	N/A	N/A	N/A	25.0 (1/4)
21	N/A	0 (0/2)	33.3 (2/6)	50.0 (4/8)
Total	77.4 (24/31)	67.4 (31/46)	82.1 (32/39)	67.3 (35/52)

Table S-2: Coverage of two doses of rotavirus vaccine among children with vaccine records. All cells show % Vaccinated (n vaccinated/n with vaccine records). A value of 'N/A' indicates that no vaccine records were available for that cycle.

Remoteness	Cycle 8	Cycle 9	Cycle 10	Cycle 11
Close Communities				
Fraction with Records	40.2% (29/72)	56.6% (56/99)	52.5% (64/122)	32.9% (26/79)
Coverage of two doses	86.2% (25/29)	66.1% (37/56)	73.4% (47/64)	96.2% (25/26)
Medium Communities				
Fraction with Records	66.7% (6/9)	70.8% (17/24)	63.6% (14/22)	37.5% (6/16)
Coverage of two doses	50.0% (3/6)	64.7%(11/17)	85.7% (12/14)	100% (6/6)
Far Communities				
Fraction with Records	41.3% (31/75)	47.4% (46/97)	35.0 % (42/120)	53.6% (52/97)
Coverage of two doses	77.4% (24/31)	67.4% (31/46)	82.1% (32/39)	67.3% (35/52)
Overall Coverage	78.8% (52/66)	66.4% (79/119)	77.8% (91/117)	78.6% (66/84)

Table S-3: Vaccine coverage over time by community remoteness

	Children ultimately receiving one dose (N=47)	Children ultimately receiving two doses (N=264)
Records with date information	76.7% (n=36)	73.8% (n=195)
Vaccinated by six months of age	91.7% (n=43)	88.7% (n=173)
Vaccinated by seven months of age	100% (n=46)	96.2% (n=188)

Table S-4: Fraction of children receiving Rotarix vaccine by age. Column 1 is for children who were considered to have received one dose in the regression model and corresponds to the date by which they received their one and only dose and column 2 is for children who ultimately received two doses and corresponds to the completion date of their second dose of vaccine. These results are shown for all children for whom we had vaccine records, which includes some children who were not in the full regression analysis due to missing data on socioeconomic indicators.

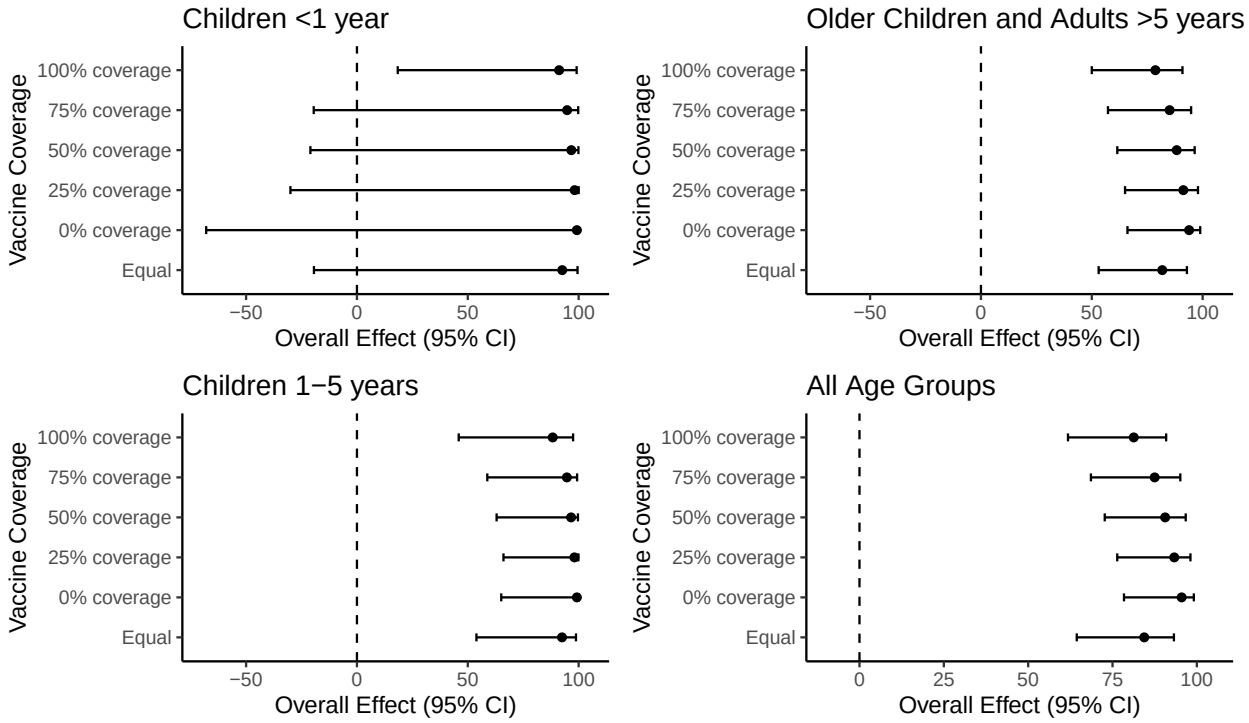

Figure S-1: Total effect of vaccination (100% vs. 0% coverage) on rotavirus infection in the case-control study by level of vaccine coverage among children without records (0%, 25%, 50%, 75%, 100% or equal to the community average).

$$Coverage = p_{record}(Coverage_{record}) + (1 - p_{record})(Coverage_{norecord})$$

Where p_{record} is the fraction of individuals with vaccine records, $Coverage_{record}$ is the coverage of two doses of Rotarix among children with vaccine records (based on data), and $Coverage_{norecord}$ is the coverage of two doses of Rotarix among children without vaccine records (varied systematically: equal to the community mean, 0%, 25%, 50%, 75% or 100%). We then re-ran the fully adjusted model with the new community vaccine coverage as the exposure. Results for rotavirus infection are shown in Figure S-1 and results for all-cause diarrhea are shown in Figure S-2. In general, results were highly similar for rotavirus infection for all coverage levels considered. When the equal coverage assumption was used, the effect was not statistically significant for children $<$1 year of age but the confidence intervals were narrower and became statistically significant when other assumptions were considered. The all-cause diarrhea estimates were more variable with different coverage assumptions, but the statistical significance for associations within each age group was similar. The one exception was the estimate for children $<$1 year, which had a significant protective effect when coverage for children without vaccine records was 100% but non-significant associations for all other coverage assumptions.

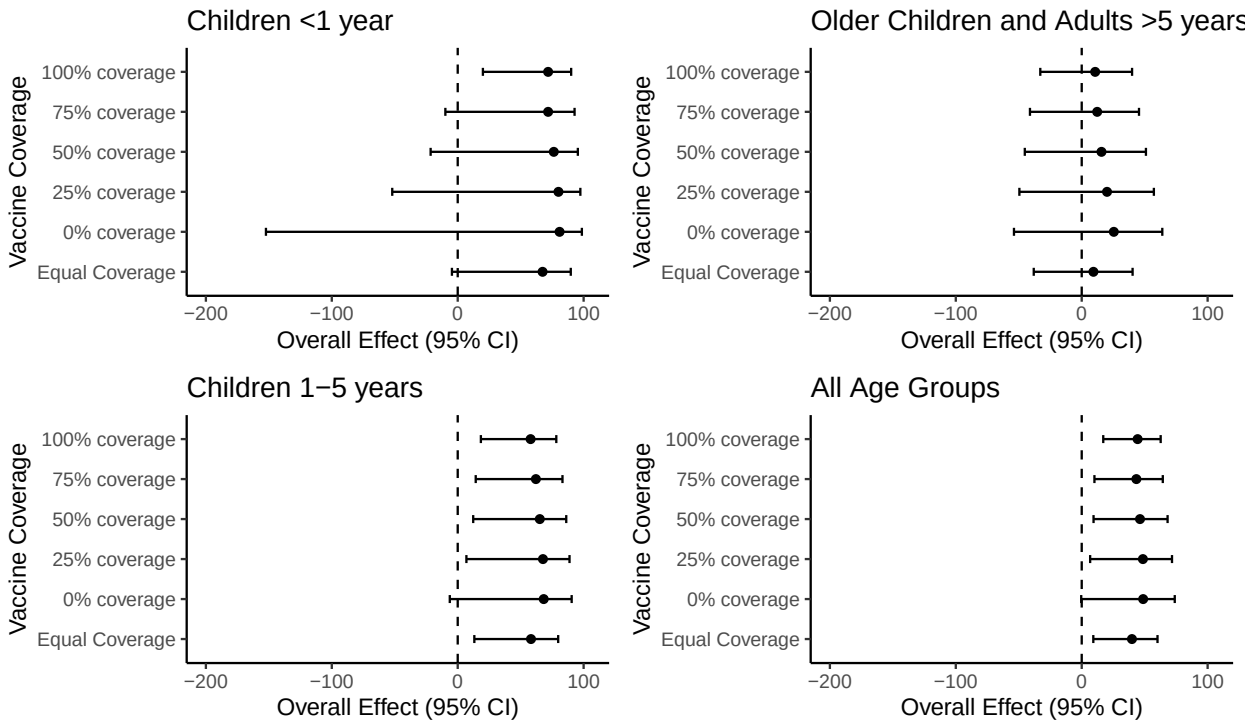

Figure S-2: Total effect of vaccination (100% vs. 0% coverage) on all-cause diarrhea in the case-control study by level of vaccine coverage among children without records (0%, 25%, 50%, 75%, 100% or equal to the community average).

S3 Case control weighting and sensitivity analysis

Weights for the case control analysis were calculated based on probability of sampling. We assume that our study captured all cases of symptomatic diarrhea. Thus, all cases had a weight of 1. Household controls in early cycles were weighted by (total non-cases in the household/total controls in the household). Community controls were weighted by (Community Population Size-Count of Community Cases-Count of Household Controls)/(Count of Community Controls). We changed the control sampling in the second half of the study. We no longer included a household control and we aimed to collect control samples from 10% of the households (selected randomly). Therefore, all controls in the second part of the study were weighted by (Community Population Size – Count of Community Cases)/(Count of controls). Thus, community controls in the first 7 cycles were weighted similarly to all controls in cycles 8-11.

Because household controls were not used later in the study, we re-ran all vaccine coverage models in table 5 excluding data from household controls as a robustness check, and the results were similar.

S4 Assessment of seasonality

Chance confounding could have occurred if certain communities with particularly low/high levels of vaccine coverage were sampled for the case control study only in certain times of the year after the vaccine was introduced. To test for this possibility, we re-ran all models used to generate table 5 with and without adjusting for season as a binary variable (1 = January to May representing the rainy season and 0 = June to December representing the dry season). In these adjusted models, the associations between vaccine coverage and all-cause diarrhea and rotavirus infection changed by less than 5% for almost all age groups and outcomes. The two exceptions were both for symptomatic rotavirus infection. For older children and adults, the season-adjusted symptomatic rotavirus infection effect estimate was 51.7% (95% CI: -9.9%-78.7%, a change of 6.1%) and the season-adjusted symptomatic rotavirus infection estimate for children under 1 was 58.5% (95% CI: -100%-91.4%, a, 10% absolute change). However, in both cases the main conclusions from the model remained the same.

S5 Internal consistency of results between analysis parts

Prior to doing any analysis, we decided which communities would be included in each part of the analysis in an effort to minimize bias in our vaccine effectiveness estimates. Because we used different communities for different parts of this analysis, we compared the results between them to ensure that our analysis was generalizable to all communities across the region and that the associations we found were not purely the result of our sampling process. In general, our results were highly internally consistent throughout the region.

To compare the cohort and case-control study, we subset the data from the overall case-control study and compared the results only among children less than five years of age. Because the all-cause diarrhea estimates from the case control study were highly similar both before and after adjusting for vaccine coverage (i.e., the covariates and vaccine coverage were independent predictors of all-cause diarrhea) we included all

study years in this comparison, including time both before and after the vaccine was introduced. By design, the odds ratio approximates the rate ratio (from the Poisson sensitivity analysis) in the case control study because controls were time matched to cases and because the outcome was rare in all exposure groups. The results are shown in table S-6.

Comparing the case control study with the poisson sensitivity analysis from the cohort, the vaccine efficacy estimate was higher in the case control study, which likely reflects the fact that the case-control study measures overall effects, which combines direct and indirect protection. We also found that older children had a higher hazard of all-cause diarrhea but a lower rate (estimated using Poisson regression, IRR=0.502, 95% CI: 0.361, 0.700). Using Poisson regression also produced an effect estimate of 39.8% for children under 2, which is similar to results presented in vaccine trials, which estimated 39-42% [4, 5]. Older children who had an all-cause diarrhea episode tended to have fewer episodes than younger children.

Another difference is the association between remoteness and all-cause diarrhea. In the case control analysis, far communities had similar rates of disease to close communities whereas in the cohort analysis, far communities were at higher risk of disease. This difference most likely arose because children with vaccine records tended to come disproportionately from close/medium villages. Therefore, it is likely that we are not estimating the rate ratio well for far communities in the cohort analysis. The association with education was also slightly different in the cohort analysis, which is also consistent with the marginally significant association we found comparing children with vaccine records to children without vaccine records.

S6 Rotavirus and diarrhea by age group

In order to determine if rotavirus was a causative diarrheal pathogen for all age groups, we also calculated the odds ratio for diarrhea given rotavirus infection relative to the odds of having diarrhea given no rotavirus infection. The results are shown in table S-7. Rotavirus is strongly associated with diarrheal symptoms in all age groups, including older children and adults (age ≥ 5).

S7 Population level trends in rotavirus infection and diarrhea over time

To calculate confidence intervals for population trends in infection and all-cause diarrhea, we fit intercept-only weighted logistic regression models separately for each outcome (4 outcomes: all-cause diarrhea, rotavirus infection, symptomatic rotavirus infection, and asymptomatic rotavirus infection, non-rotavirus diarrhea) and cycle (11 cycles) of the case control study for a total of 55 models. For rotavirus infection, we were also interested in prevalence by age group, so we subset this model by the three age groups (used in figure 2) to produce a total of 33 models. Then, for each of these models, we used the standard logit transformation to produce confidence intervals for prevalence guaranteed to fall between 0 and 1. Representative equations for rotavirus infection are shown below:

	Case-Control Analysis OR (95% CI)	Cohort Analysis-Cox HR (95% CI)	Cohort Analysis-Poisson (sensitivity) IRR (95% CI)
Male	0.937 (0.622, 1.41)	0.878(0.612, 1.26)	0.983 (0.764, 1.27)
Age			
<2	Ref	Ref	Ref
2–5	0.348 (0.232, 0.523)	2.23 (1.13, 4.42)	0.502 (0.359, 0.7)
Household size	1.00 (0.941, 1.07)	0.995 (0.942, 1.05)	0.992 (0.955, 1.03)
Highest household education	0.917 (0.861, 0.976)	0.976 (0.924, 1.03)	1.01 (0.966, 1.03)
Remoteness			
Close	Ref	Ref	Ref
Medium	0.903 (0.455, 1.80)	1.42 (0.771, 2.60)	1.29 (0.852, 1.96)
Far	1.39 (0.899, 2.15)	2.18 (1.48, 3.22)	1.42 (1.08, 1.86)
Effect of 2 doses (1-effect estimate)$\times$ 100	51.9% (19.0%, 71.5%)	57.1% (16.6%-77.9%)	39.8% (13.7%, 58.0%)

Table S-5: Comparison between the cohort and case-control study. All parameters except the effect of two doses were calculated using all children in the case-control dataset under 5 years of age for the case-control study and all children in the cohort and were estimated in a single model. For the effect of two doses, we subset both data files at 2 years of age and presented calculated vaccine effectiveness using the pre/post for the case control analysis and the two doses for the cohort analysis.

Age group	Unadjusted OR (95% CI)	Adjusted OR (95% CI)
<1	9.48 (2.08, 43.2)	8.76 (2.29, 33.5)
1-5	6.85 (2.48, 18.9)	6.53 (2.37, 18.1)
≥ 5	12.9 (7.7, 21.8)	13.1 (7.6, 22.5)
Total population	11.0 (7.39, 16.5)	10.1 (5.7, 17.8)

Table S-6: Effect of rotavirus infection on diarrhea illness by age. The adjusted odds ratios are adjusted for sex, household size, highest household education, and remoteness. Each odds ratio can be interpreted as the multiplicative increase in odds of diarrhea, given rotavirus infection.

$$\bar{P}(\text{infection}) = \frac{\exp(\beta_0)}{1+\exp(\beta_0)}$$
$$P_{LCL}(\text{infection}) = \frac{\exp(\beta_0 - 1.96 \times SE(\beta_0))}{1+\exp(\beta_0 - 1.96 \times SE(\beta_0))}$$
$$P_{LCL}(\text{infection}) = \frac{\exp(\beta_0 + 1.96 \times SE(\beta_0))}{1+\exp(\beta_0 + 1.96 \times SE(\beta_0))}$$

We found that the prevalence of diarrhea declined over time for both rotavirus and non-rotavirus diarrhea (Figure S-3). For rotavirus, this change was stronger for asymptomatic rotavirus infection than symptomatic rotavirus. However, the confidence intervals overlap and we were not adequately powered to determine if the effect on asymptomatic infection was stronger than on symptomatic infection (Figure S-4).

S8 Supplemental information for the cohort analysis

The regression results presented in the main text are also shown in Table S-8 to illustrate differences between the adjusted and unadjusted models. In this population, none of the socioeconomic indicators were significantly associated with all-cause diarrhea, unlike in the case control study. This difference is probably due to the fact that children with vaccine records were not comparable to children without vaccine records–children with vaccine records had lower household education than children without vaccine records.

While children who were older than 2 years of age had a lower rate of all-cause diarrhea, considering all episodes that occurred during the follow up period, they had a higher hazard of diarrhea. This difference reflects the fact that most diarrheal infections are incompletely immunizing. There as no evidence that the effect of vaccination was stronger among young children (based on the lack of significance of the interaction term).

Supplementary References

[1] Halloran ME, Hudgens MG. Estimating population effects of vaccination using large, routinely collected data. Statistics in Medicine. 2018;37:294–301.

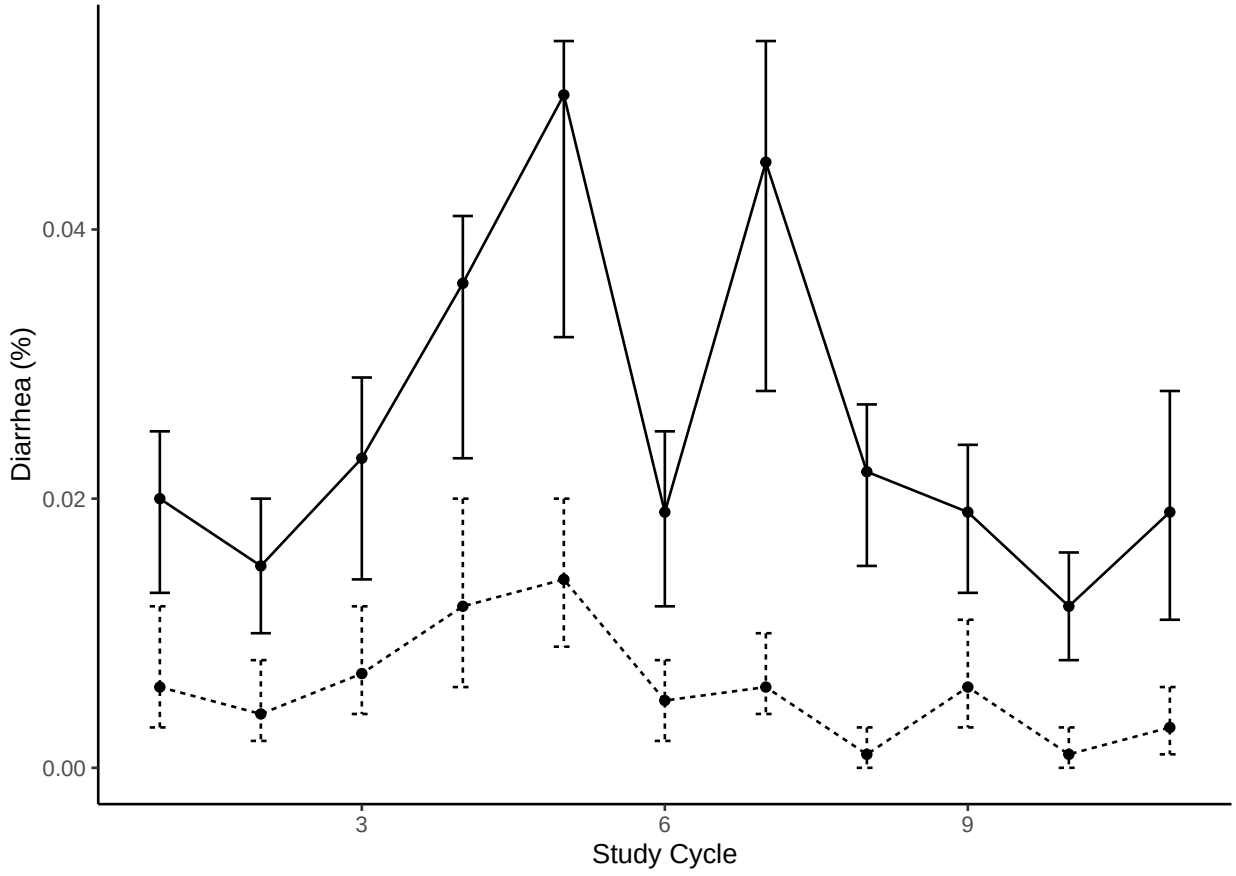

Figure S-3: Prevalence of diarrhea not attributable to rotavirus (solid) and attributable to rotavirus (dashed) over time

Table S-7: Cox Regression Model for the time to first all-cause diarrhea episode among children. Models 2 and 3 are adjusted for household size (kids), highest household education, gender, community remoteness, and BCG vaccination.

	Model 1 Unadjusted HR (95% CI)	Model 2 Adjusted HR (95% CI)	Model 3 Adjusted + Interactions HR (95% CI)
Rotavirus Vaccine			
0 doses	*Ref*	*Ref*	*Ref*
1 dose	1.02(0.552, 1.89)	0.603 (0.290, 1.25)	0.606 (0.269, 1.37)
2 doses	0.749(0.457, 1.23)	0.474 (0.257, 0.874)	0.451 (0.231, 0.882)
Age $\geq$2 years	2.67(1.37, 5.22)	2.35(1.19, 4.66)	1.95 (0.393, 9.68)
Male	0.843 (0.6, 1.18)	0.914 (0.636, 1.31)	0.913 (0.635, 1.31)
BCG Vaccination	1.60 (0.706, 3.64)	2.31 (0.906, 5.88)	2.38 (0.914, 6.19)
Household Size	1.01(0.917, 1.12)	0.977 (0.881, 1.08)	0.976 (0.878, 1.09)
Highest Household Education	0.980 (0.930, 1.03)	0.965 (0.913, 1.02)	0.965 (0.914, 1.02)
Remoteness			
Close	*Ref*	*Ref*	*Ref*
Medium	1.48 (0.818, 2.68)	1.53 (0.830, 2.83)	1.54 (0.835, 2.89)
Far	2.03 (1.41, 2.91)	2.28 (1.53, 3.40)	2.28 (1.53, 3.40)
Age x Vaccine Interaction			
Age($\geq$2) x 1 dose	–	–	1.06 (0.174, 6.45)
Age ($\geq$2) x 2 doses	–	–	1.29 (0.259, 6.43)

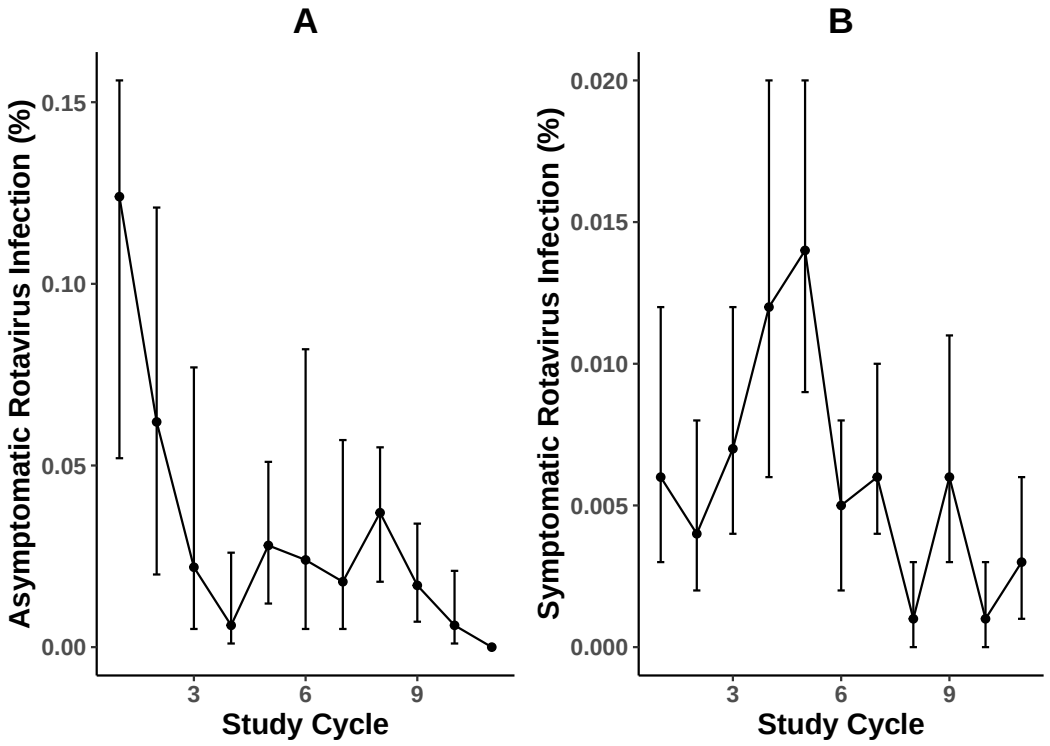

Figure S-4: Prevalence of A) Asymptomatic rotavirus infection and B) Symptomatic rotavirus infection over time. The rotavirus vaccine was introduced in study cycle 8

[2] Panozzo CA, Becker-Dreps S, Pate V, Weber DJ, Jonsson Funk M, Stürmer T, et al. Direct, indirect, total, and overall effectiveness of the rotavirus vaccines for the prevention of gastroenteritis hospitalizations in privately insured US children, 2007–2010. American journal of epidemiology. 2014;179(7):895–909.

[3] Lopman BA, Pitzer VE. Waxing understanding of waning immunity. Journal of Infectious Diseases. 2018;217(6):851–853.

[4] Linhares AC, Velázquez FR, Pérez-Schael I, Sáez-Llorens X, Abate H, Espinoza F, et al. Efficacy and safety of an oral live attenuated human rotavirus vaccine against rotavirus gastroenteritis during the first 2 years of life in Latin American infants: a randomised, double-blind, placebo-controlled phase III study. The Lancet. 2008;371(9619):1181–1189.

[5] Ruiz-Palacios GM, Pérez-Schael I, Veázquez R, Abate H, Breuer T, Clemens SC, et al. Safety and efficacy of an attenuated vaccine against severe rotavirus gastroenteritis. New England Journal of Medicine. 2006;354(1):11–22.