Maternal and infant factors linked with rotavirus vaccine outcome in three continents

Edward P. K. Parker[1]*, Christina Bronowski[2]*, Kulandaipalayam Natarajan C. Sindhu[3]*, Sudhir Babji[3], Blossom Benny[3], Noelia Carmona-Vicente[2], Nedson Chasweka[4], End Chinyama[4], Nigel A. Cunliffe[2], Queen Dube[4], Sidhartha Giri[3], Nicholas C. Grassly[6], Annai Gunasekaran[3], Deborah Howarth[2], Sushil Immanuel[3], Khuzwayo C. Jere[2,4,5], Beate Kampmann[1], Jenna Lowe[2], Jonathan Mandolo[4], Ira Praharaj[3], Bakthavatsalam Sandya Rani[3], Sophia Silas[3], Vivek Kumar Srinivasan[3], Mark Turner[7], Srinivasan Venugopal[3], Valsan Philip Verghese[8], Alistair C. Darby[9†], Gagandeep Kang[3†], Miren Iturriza-Gómara[2†§]

* Joint first authors; † Joint senior authors; § Lead contact

[1] *The Vaccine Centre, Department of Clinical Research, London School of Hygiene and Tropical Medicine, London, WC1E 7HT, UK;*

[2] *Institute of Infection and Global Health, University of Liverpool, Liverpool, L69 7BE, UK;*

[3] *Division of Gastrointestinal Sciences, Christian Medical College, Vellore, Tamil Nadu, 632004, India;*

[4] *Malawi-Liverpool-Wellcome Trust Clinical Research Programme, University of Malawi, Blantyre, PO Box 30096, Malawi;*

[5] *Department of Medical Laboratory Sciences, College of Medicine, University of Malawi, Private Bag 360, Chichiri, Blantyre 3, Malawi;*

[6] *Department of Infectious Disease Epidemiology, Imperial College London, London, W2 1PG, UK;*

[7] *Institute of Translational Medicine, University of Liverpool, Liverpool, L8 7SS, UK;*

[8] *Department of Child Health, Christian Medical College, Vellore, Tamil Nadu, 632004, India;*

[9] *Institute of Integrative Biology, University of Liverpool, Liverpool, L69 7ZB, UK*

Corresponding authors:

Edward Parker (edward.parker@lshtm.ac.uk)

Miren Iturriza-Gómara (M.Iturriza-Gomara@liverpool.ac.uk)

Character count: 44,850

Summary

Identifying risk factors for impaired oral rotavirus vaccine (ORV) efficacy in low-income countries may lead to improvements in vaccine design and delivery. We measured maternal rotavirus antibodies, environmental enteric dysfunction (EED), and bacterial gut microbiota development among infants receiving two doses of Rotarix in India (n = 307), Malawi (n = 119), and the UK (n = 60), using standardised methods across cohorts. ORV shedding and seroconversion rates were significantly lower in Malawi and India than the UK. Maternal rotavirus-specific antibodies in serum and breastmilk were negatively correlated with ORV response in India and Malawi, and this was mediated partly by a reduction in ORV replication. In the UK, ORV replication was not inhibited despite comparable maternal antibody levels. In both India and Malawi, pre-vaccination microbiota diversity was negatively correlated with ORV immunogenicity, suggesting that high early-life microbial exposure may contribute to impaired vaccine efficacy.

Summary word count: 145/150

Key words

Oral vaccines, environmental enteric dysfunction, intestinal microbiota, maternal antibodies, rotavirus, vaccination

Introduction

The roll-out of oral rotavirus vaccine (ORV) has now reached over 100 countries. This initiative – in parallel with economic developments, advances in sanitation infrastructure, and increased use of oral rehydration therapy – has led to substantial declines in the global burden of diarrhoeal mortality (Mokomane et al., 2018). Whereas rotavirus was linked with over 500,000 infant deaths annually at the turn of the century (Tate et al., 2016), this number currently stands at approximately 130,000, and continues to decline year by year (Troeger et al., 2018). In Malawi, introduction of the attenuated monovalent vaccine Rotarix in 2012 has contributed to a decrease in diarrhoea-related mortality by almost a third (Bar-Zeev et al., 2018). Several countries in Latin America have reported comparable declines in diarrhoeal disease after ORV introduction (Paternina-Caicedo et al., 2015), and others are likely to follow as we obtain additional effectiveness data. India is currently undergoing a phased introduction of Rotavac, and post-introduction effectiveness studies are ongoing (Nair et al., 2019).

Yet the potential impact of ORV is constrained by the impaired performance of currently available vaccines in low- and middle-income countries (LMICs). Whereas Rotarix has a 1-year protective efficacy against severe rotavirus-associated gastroenteritis of >95% in Europe (Vesikari et al., 2007), efficacy may fall below 50% in sub-Saharan Africa (Madhi et al., 2010). Whether considering protective efficacy or immunogenicity, ORV is least effective among the infants that need it most (Patel et al., 2013). Moreover, while a variety of interventions to boost ORV performance have been tested (e.g. temporary withholding of breastfeeding), these have generally proven to be either ineffective or of modest benefit (Church et al., 2019), highlighting the need for new strategies informed by a deeper understanding of the mechanisms underlying the vaccine efficacy gap.

Given the variety of factors that may impair oral vaccine performance in LMICs, this issue is unlikely to have a simple solution (Parker et al., 2018a). Passively acquired maternal rotavirus-specific antibodies appear to interfere with ORV immunogenicity (Moon et al., 2016), as documented for a variety of parenteral vaccines (Voysey et al., 2017). However, it remains unclear whether rotavirus-specific antibody concentrations and their inhibitory effect are higher among mothers in LMICs and thus whether they contribute to broad geographic trends in ORV immunogenicity. Environmental enteric dysfunction (EED) – a subclinical condition associated with intestinal inflammation, increased permeability, and blunted intestinal villi – is common among infants living in LMICs and has been cited as a possible cause of oral vaccine failure (Naylor et al., 2015). However, EED is generally measured using faecal or plasma biomarkers as a proxy for mucosal immune status, and studies of these biomarkers at the time of oral vaccine delivery have yielded mixed results (Church et al., 2018).

The diversity and complexity of the bacterial gut microbiota expands continuously over the first few years of life – a process of ecological succession influenced by mode of delivery, breastfeeding practices, antibiotic exposure, and geographic setting (Stewart et al., 2018; Yatsunenko et al., 2012). Microbiota composition shapes and is in turn shaped by the developing infant immune system, and events in early life may have long-term implications for immune-mediated diseases such as asthma (Arrieta et al., 2015). Among infants in Ghana, Rotarix immunogenicity was positively correlated with the relative abundance of *Streptococcus bovis* and negatively correlated with the relative abundance of Bacteroidetes at the time of the first vaccine dose (Harris et al., 2017). However, these trends were not evident among infants in Pakistan (Harris et al., 2018) and no significant discrepancies in microbiota composition were apparent when comparing Rotarix responders with non-responders in India (Parker et al., 2018b). Consistent predictors of ORV failure within the bacterial gut microbiota thus remain elusive.

Much of what we know about the mechanisms shaping oral vaccine performance comes from single-population studies focusing on individual risk factors. Here, we report on a multicentre cohort study of Rotarix response among infants in Malawi, India, and the UK. Using standardised methods across the three cohorts, we measured the effect of maternal antibodies, EED biomarkers, and bacterial gut microbiota development on ORV response.

Results

Study design

A full description of the study sites, enrolment criteria, and protocol has previously been published (Sindhu et al., 2017). Briefly, we enrolled pregnant women during the third trimester in Blantyre (Malawi), Vellore (India), and Liverpool (UK). After delivery, infants received two doses of Rotarix according to the national immunisation schedule (6 and 10 weeks in Malawi and India, 8 and 12 weeks in the UK; **Figure 1A**). We measured rotavirus-specific IgA (RV-IgA) in maternal blood, cord blood, and breastmilk samples collected during or in the week after delivery. In infants, we measured the concentration of RV-IgA pre- and post-vaccination as an indicator of ORV immunogenicity. In Indian participants, all serum samples were also assayed for rotavirus-specific IgG (RV-IgG). Starting in the first week of life, six longitudinal stool samples were collected from each infant and assayed via quantitative PCR (qPCR) for rotavirus shedding, with samples collected 1 week after each ORV dose providing an indicator of vaccine virus replication. As a proxy for bacterial microbiota development in the infant gut, we sequenced the V3–V4 region of the 16S rRNA gene in samples collected at 1 and 4 weeks of age, and before each ORV dose (four samples per infant). Microbiota composition and rotavirus shedding were also measured in maternal stool samples collected in the week after delivery. Myeloperoxidase (MPO; a marker of neutrophil activity) and α1-antitrypsin (α1AT; a marker of protein-losing enteropathy) were measured in stool samples collected before each ORV dose, and α1 acid glycoprotein (α1AG) was measured in pre-vaccination serum samples of Indian and Malawian infants as a marker of systemic inflammation (**Table S1**).

Other routine vaccines were administered according to the routine schedule at each site, including OPV at 0, 6, and 10 weeks of age in Malawi and India. In Malawi, all infants were born after April 2016 – the date of the global switch from trivalent OPV (tOPV) to bivalent OPV (bOPV). In India, 98/307 (32%) infants were born before April 2016. Infants therefore received tOPV (57/307 [19%]), a mixed schedule of tOPV and bOPV (56/307 [18%]), or bOPV (94/307 [31%]). To explore the potential inhibitory effect of OPV on ORV (Church et al., 2020), the 6 and 10 week doses of OPV were replaced with inactivated poliovirus vaccine (IPV) in a sequentially recruited cohort in India (100/307 [33%], all of whom received bOPV at birth). The comparison of OPV and IPV arms is described elsewhere (Babji et al, in preparation). Briefly, while shedding after the first dose of ORV was less common in OPV than IPV recipients, cumulative shedding after both doses was comparable in the two arms, as were seroconversion rates and post-vaccination RV-IgA levels. Given the lack of association between study arm and ORV immunogenicity, we pooled the IPV and OPV arms in the present study.

The primary endpoint (measurement of seroconversion or dose 1 shedding) was reached by 307 infants in India, 119 in Malawi, and 60 in the UK. While the pre- and post-parturition conditions of these disparate cohorts are innumerable, several distinguishing features are highlighted in **Table 1**. In particular, infants in the UK were characterised by a higher birthweight, a greater prevalence of formula feeding (though >75% were partially or exclusively breastfed), and a lower likelihood of antibiotic exposure. Elective caesarean was an exclusion criterion for this cohort. Infants in Malawi were the most likely to be exclusively breastfed throughout the duration of the study. In India, 70/307 (23%) infants were born by caesarean section.

ORV replication and immunogenicity are impaired in Malawian and Indian infants

ORV response was significantly lower in Malawi and India compared with the UK (**Figures 1B** to **1D**). In the UK, we observed near-ubiquitous shedding of the first dose of ORV, with 55/60 (92%) shedding 1 week after vaccination and 24/54 (44%) continuing to shed after 1 month (**Figure S1A**). By contrast, dose 1 shedding was detected in 82/305 (27%) infants in India and 56/101 (55%) in Malawi (**Figure 1B**), and continued shedding after 1 month was much rarer in these cohorts (**Figure S1A**). ORV shedding after the second dose was also more common in the UK than India or Malawi, although the discrepancies were less marked than for the first dose (**Figure 1B**). Shedding following at least one dose was observed in 54/56 (96%) infants in the UK, 151/304 (50%) in India, and 50/72 (69%) in Malawi (Fisher's p values <0.005 for all between-country comparisons after false discovery rate [FDR] adjustment; **Figure 1B**).

We also observed significant geographic discrepancies in ORV immunogenicity. Seroconversion was observed in 85/305 (28%) infants in India, 24/103 (23%) infants in Malawi, and 27/51 (53%) infants in the UK (**Figure 1C**). Geometric mean concentrations (GMCs) of RV-IgA (IU/ml) after vaccination were 20 (95% CI 16–25) in India, 9 (6–12) in Malawi, and 27 (17–45) in the UK (Tukey's post-hoc p values <0.005 for comparisons between Malawi and other cohorts and 0.489 for India vs UK). Whereas the low seroconversion rates in India and Malawi are consistent with recent studies in these settings (Lazarus et al., 2018; Pollock et al., 2018), seroconversion rates in the UK fall notably short of those observed in previous European trials of Rotarix (typically >85%) (Patel et al., 2013). Overall, at least one indicator of ORV response (seroconversion or shedding after at least one dose) was observed in 181/302 (60%) infants in India, 43/61 (70%) in Malawi, and 46/48 (96%) in the UK (**Figure 1D**).

Neonatal G10P[11] rotavirus infection is frequent and highly immunogenic in Indian infants

A unique feature among Indian infants was the high rate of wild-type rotavirus shedding in the first week of life (136/306 [44%] based on qPCR, compared with 6/102 [6%] in Malawi and 1/58 [2%] in the UK; **Figure 1B**). We successfully characterised rotavirus genotype in 104 samples, of which 103 were positive for the strain G10P[11]. These findings corroborate the high rates of G10P[11] rotavirus exposure previously reported in tertiary care hospitals in Vellore (Ramani et al., 2008). Rotavirus shedding was rare at weeks 4 and 6 (**Figure S1A**), suggesting that the majority of pre-vaccination rotavirus exposure occurs shortly after birth. Moreover, of the 170 infants with no detectable week 1 shedding, 31 (18%) were seropositive at 6 weeks (RV-IgA ≥20 IU/ml). Overall, 166/304 (55%) infants in India had some indication of pre-vaccination rotavirus infection (week 1 shedding or baseline RV-IgA seropositivity), whereas the corresponding rates were 10/90 (11%) in Malawi and 2/54 (4%) in the UK. Pre-vaccination infection was more common among infants born in tertiary care facilities in India (relative risk [RR] 1.98 [95% CI 1.72–2.19]; **Table S2)** and among infants delivered by caesarean section as opposed to vaginal delivery (RR 1.31 [95% CI 1.05–1.53]), potentially reflecting enhanced exposure to opportunistic pathogens in the hospital environment (Shao et al., 2019). Notably, all neonatal rotavirus infections were asymptomatic and the longer-term clinical significance of these infections is therefore uncertain.

Consistent with the high prevalence of neonatal rotavirus infection, we observed significantly higher pre-vaccination RV-IgA concentrations among Indian infants compared with those in Malawi and the UK (**Figure 1C**). Prior infection, as defined by week 1 shedding or baseline RV-IgA seropositivity, was associated with a reduced likelihood of shedding ORV after at least one dose (RR 0.65 [0.47–0.83]; **Figure S1B** and **Table S3**). Despite the reduction in ORV shedding, pre-vaccination infection did not influence the likelihood of seroconversion among Indian infants (**Figure S1B** and **Table S4**), and where ORV replication was observed among previously infected infants, this significantly boosted post-vaccination RV-IgA concentrations (**Figure 1E**). Overall, our findings point to the high frequency and immunogenic nature of neonatal rotavirus infection in India. Indeed, as noted above, despite the lower seroconversion rates in India compared with the UK, post-vaccination GMCs did not differ significantly between these countries (**Figure 1C**).

Exclusive breastfeeding, growth, and sanitation are positively correlated with ORV response in India

We examined several baseline health and demographic variables for their potential correlation with ORV response (**Tables S3** to **S5**). In India, seroconversion was positively correlated with exclusive

193 breastfeeding (RR 2.04 [1.07–3.68]) and was more common in infants who received bOPV-only than
194 tOPV-containing schedules (**Table S4**), although infants receiving tOPV were also more likely to be
195 immunised during the dry season (January–May) given the timing of the tOPV–bOPV switch within
196 the study. Similar associations were apparent when considering post-vaccination RV-IgA concentration
197 as an endpoint (**Table S5**). Moreover, RV-IgA was positively correlated with height-for-age Z score at
198 the time of the first ORV dose, and was higher among infants living in houses built from permanent
199 versus temporary or mixed materials (beta 2.03 [95% CI 1.26–3.26], where beta represents the estimated
200 ratio of GMCs) or with access to treated water (beta 1.66 [1.04–2.67]). Baseline health and demographic
201 variables were not strongly associated with ORV response in Malawi or the UK, although fewer
202 covariates were captured in these cohorts.
203
204 *Maternal antibodies inhibit ORV shedding in LMICs*
205 Serum RV-IgA concentrations were significantly higher among mothers in Malawi than in
206 India or the UK (**Figure 2A**), potentially reflecting a greater burden of prenatal rotavirus exposure in
207 this setting and the consequent boosting of rotavirus-specific antibodies (GMCs of 134 [120–150], 340
208 [256–452], and 186 [123–281] in India, Malawi, and the UK, respectively). Interestingly, maternal RV-
209 IgA levels did not differ significantly between India and the UK, while breastmilk RV-IgA
210 concentrations were significantly higher in Malawi and the UK than in India (**Figure 2A**; GMCs of 25
211 [22–29], 97 [84–112], and 100 [68–146] in India, Malawi, and the UK, respectively). Thus, while the
212 degree of maternal antibody exposure was greatest among infants in Malawi, infants in the UK were
213 by no means exempt from this potential ORV inhibitor.
214 Our analysis of the association between maternal antibodies and ORV outcome yielded several
215 notable findings. First, maternal serum RV-IgA concentrations were negatively correlated with infant
216 post-vaccination RV-IgA levels in each cohort (Pearson coefficient [*r*] of -0.29, -0.13, and -0.12 in Malawi,
217 India, and the UK, respectively; **Figure 2B**). Although post-vaccination RV-IgA levels were not
218 significantly correlated with maternal RV-IgA in the UK, this may reflect a lack of statistical power
219 owing to the smaller size of this cohort, and we observed no significant evidence of effect modification
220 between the cohorts (p values of >0.05 for interaction terms between maternal RV-IgA and country).
221 Notably, serum IgA is not transmitted efficiently across the placenta and is therefore unlikely to directly
222 influence ORV. However, we observed a strong correlation between maternal RV-IgA and cord blood
223 RV-IgG in India (*r* 0.41, p <0.001), suggesting that maternal RV-IgA may be a valid (if imperfect) proxy
224 for the degree of transplacental RV-IgG transfer (**Figure 2C**). Second, maternal serum RV-IgA levels

were negatively correlated with ORV shedding in both India and Malawi (**Figure 2C** and **Table S3**), suggesting that maternal antibodies may reduce the replication efficiency of vaccine rotaviruses at the mucosal epithelium. These findings were robust to pre-vaccination rotavirus exposure status in India (**Table S3B**). By contrast, ORV shedding in the UK was ubiquitous despite the fact that maternal antibody levels were equivalent to those in India. Finally, we observed no significant difference in maternal antibody levels according to week 1 rotavirus shedding status in Indian infants (**Table S2**), suggesting that neonatal G10P[11] viruses may not be subject to the same replication inhibition observed for ORV. However, among infants with pre-vaccination infection, we observed a strong negative correlation between maternal RV-IgG levels and infant RV-IgA formation (r of -0.42 for maternal RV-IgG vs infant pre-vaccination RV-IgA; **Figure 2B**), highlighting the potential inhibitory effect of transplacental RV-IgG on antigen processing.

Breastmilk RV-IgA was negatively correlated with ORV immunogenicity and/or shedding in India and Malawi (**Figure 2B** and **Tables S3** to **S5**). Disentangling the relative influence of breastmilk versus transplacental antibodies is challenging given the correlation between the two (**Figure 2C**). However, as noted above, ORV immunogenicity was higher in conjunction with exclusive breastfeeding in India – a finding at odds with a significant inhibitory effect of breastmilk antibodies on ORV response.

Inflammatory biomarker levels differ by geographic setting but not ORV outcome

We observed strong geographic discrepancies in EED biomarkers. Whereas α1AT was highest in Malawi and lowest in the UK, both MPO and α1AG were highest in Indian infants (**Figure 3A**). At the time of the first dose of ORV, GMCs for α1AT (μg/ml) were 734 [676–796], 4,376 [3,136–6,105], and 282 [155–511] in India, Malawi, and the UK, respectively, while GMCs for MPO (ng/ml) were 6,464 [5,446–7,672], 2,966 [2,454–3,585], and 3,767 [2,414–5,878], respectively. Despite their marked variation within and between cohorts, none of the biomarkers were significantly associated with seroconversion, ORV shedding, or post-vaccination RV-IgA (**Figure 3B** and **Tables S3** to **S5**).

Microbiota development is geographically distinct

While composition of the bacterial gut microbiota is known to be shaped by geographic setting (Yatsunenko et al., 2012), the signatures of microbiota development in many regions remain uncharted. Among mother–infant pairs who reached the primary endpoint of this study, we obtained high-quality faecal microbiota profiles in 2,086 samples (142,880 ± 135,860 [mean ± s.d.] reads per sample).

Assessment of microbiota composition was highly consistent across sequencing runs and facilities (explored by independently re-sequencing 10% of samples; **Figure S2**). Hierarchical clustering based on the presence or absence of common genera highlighted the complexity of the maternal microbiota compared with that of infants, with the majority of taxa unique to maternal samples (**Figure S3A**). Consistent with previous studies of early-life microbiota composition, infant samples were dominated by a small number of taxa – the vast majority of observed genera occurred in <20% of infant samples (88–93% depending on the cohort; **Figure S3B**). The trajectory of microbiota development was highly distinct to each cohort, with significant deviations apparent as early as the first week of life. In particular, microbiota diversity and richness were significantly higher in Malawi than both other cohorts (**Figure 4A**). Trajectories of the most common infant genera revealed a greater relative abundance of *Bifidobacterium* and *Streptococcus* among Indian infants, enrichment of *Prevotella* among Malawian infants, and enrichment of *Bacteroidetes* and *Veillonella* in the UK (**Figure 4B** and **Figure S4A**). These discrepancies were also evident in maternal samples (**Figure S3C**), potentially reflecting the importance of vertical transmission events during early microbiota colonisation (Shao et al., 2019).

Longitudinal mixed-effects models based on the abundances of common genera confirmed these trends while revealing more nuanced discrepancies between cohorts. In keeping with their greater diversity, many genera were enriched in Malawi compared with both India (21 genera with FDR p <0.1) and the UK (22 genera with FDR p <0.1; **Figure S4B**). On the other hand, *Bifidobacterium*, *Enterococcus, Staphylococcus,* and *Streptococcus* were enriched in Indian infants compared with both other cohorts, while *Bacteroides, Citrobacter*, *Enterobacter*, *Klebsiella, Haemophilus, Lachnospiraceae,* and *Veillonella* were enriched among infants in the UK. Clustering of samples by country at the time of the first ORV dose was evident based on Bray–Curtis distances, whether weighted by taxon abundance (**Figure 4C**) or unweighted (data not shown). Moreover, using the machine-learning method Random Forests, we were able to accurately classify samples by country at all time-points (mean cross-validation accuracies ranging from 82% to 98%; baseline accuracy 50%; **Figure 4D**). The genera most important to the predictive accuracy of these models overlapped with the discriminant taxa identified by longitudinal models, and differed between cohorts in both their prevalence and abundance (**Figures 4E** and **S5**). Indeed, simple cross-sectional comparisons of taxon prevalence were adept at selecting differential microbiota colonisation patterns between cohorts (**Figure S4C** and **Table S6**).

289 *Microbiota diversity is negatively correlated with ORV response in LMICs*

290 A clear understanding of healthy microbiota development in different geographic settings is
291 presently lacking. Notably, low microbiota diversity is often used as a marker of 'dysbiosis' (Valdes et
292 al., 2018), albeit largely based on health outcomes measured in high-income settings. Interestingly, in
293 both India and Malawi, we observed a significant negative correlation between microbiota diversity
294 and ORV immunogenicity. In India, these discrepancies were apparent in longitudinal models (**Figure**
295 **5A**) and cross-sectional comparisons at the time of the first ORV dose (**Figure 5B** and **Table S4**).
296 Interestingly, stratified analyses revealed the discrepancies to be specific to infants with no rotavirus
297 exposure prior to receiving ORV (**Figure 5A**)**.** A similar trend was evident in Malawian infants, albeit
298 specifically at the time of the first dose of ORV (**Figure 5B** and **Table S4**). By contrast, microbiota
299 diversity was not strongly associated with ORV immunogenicity in the UK. Moreover, neither Shannon
300 index nor richness differed according to ORV shedding status in any cohort (**Figure S6** and **Table S3**).

301 In spite of the significant differences in microbiota diversity, we did not observe clear
302 differences in taxonomic composition according to ORV response. Cross-sectional Random Forests
303 models based on genus abundances failed to accurately predict seroconversion (**Figure 5C**), post-
304 vaccination RV-IgA concentration (**Figure 5D**), or shedding status (**Figure S6C** and **S6D**). While
305 numerous genera either increased or decreased in abundance with age, very few exhibited differential
306 abundance according to ORV outcome when assessed via longitudinal models (**Figure 5E** and **Figure**
307 **S6E**). Exceptions include *Sutterella*, which was negatively correlated with dose 1 ORV shedding in
308 Malawian infants, while *Lactobacillus*, *Corynebacterium*, and *Dolosigranulum* were negatively correlated
309 with ORV shedding and/or final RV-IgA concentration in India (**Table S7**). Overall, while microbial
310 exposure and colonisation appears to be significantly higher in infants who fail to respond to ORV, we
311 did not observe consistent discrepancies in specific bacterial taxa according to vaccine outcome.

312

313 *Integrated analysis of ORV response*

314 Finally, to delineate the relative contribution of different types of risk factor in shaping ORV
315 response, we performed an integrated analysis based on three modules of input data: (i) demographic
316 and baseline health variables (n = 19; see **Table S4** for complete list); (ii) exposure/antibody data (n =
317 12), including inflammatory biomarkers and maternal antibody concentrations; and (iii) pre-
318 vaccination microbiota composition, comprising genus abundances at the time of the first ORV dose
319 (n = 128). We then fit Random Forests models to predict ORV outcome in Indian infants (the largest
320 cohort in terms of number of infants and number of recorded variables) using each module in turn and

321 all modules combined. The prediction of seroconversion and ORV shedding was poor for all modules
322 (mean cross-validation accuracy ≤60%; baseline accuracy 50%; **Figure 6A**). Post-vaccination RV-IgA
323 concentration was predicted with modest accuracy by exposure/antibody data (mean cross-validation
324 R^2 of 0.35), and this module alone offered predictive accuracy commensurate with the integrated
325 analysis combining all input variables (R^2 of 0.37; **Figure 6A**). Consistent with the univariate analyses
326 described above, the most important variables underlying the accuracy of the integrated models for
327 prediction of post-vaccination RV-IgA were pre-vaccination rotavirus infection and maternal RV-IgG
328 concentration (**Figure 6B**), and model accuracy was markedly reduced (R^2 of 0.03) when the analyses
329 were restricted to infants lacking pre-vaccination infection (**Figure 6A**).

330 We complemented the Random Forests models with multivariate regressions. These used the
331 same input variables but with pre-vaccination microbiota diversity measurements in place of genus
332 abundances. Covariates with a p value of <0.05 in univariate analyses were selected for inclusion in the
333 final models (see **Tables S3** to **S5** for complete analyses). Notably, this approach revealed breastfeeding
334 practice, linear growth, pre-vaccination exposure, and polio vaccine schedule to be significantly
335 correlated with post-vaccination RV-IgA (**Figure 6C**). Variables associated with ORV response differed
336 according to pre-vaccination exposure status – whereas breastfeeding practice and maternal antibody
337 levels were strong predictors of RV-IgA among infants with neonatal rotavirus exposure, microbiota
338 diversity, house type, and access to treated water were significantly associated with vaccine
339 immunogenicity in those without prior exposure. Overall, the multivariate regressions offered a
340 predictive accuracy commensurate with or exceeding that of Random Forests (R^2 of 0.43 in the complete
341 cohort and 0.11–0.19 in the stratified analyses for RV-IgA regressions).

342 **Discussion**

343 ORV is unique among oral vaccines in being routinely administered in both high-income
344 countries and LMICs. As such, the vaccine provides a valuable tool for probing the divergent
345 trajectories of the developing immune system in different settings. Here, we report on the extent to
346 which several key determinants of infant immune development – maternal antibody levels, EED, and
347 bacterial microbiota composition – differ among three geographically disparate populations, and the
348 degree to which these factors shape ORV response.

349 As expected, Rotarix response was significantly impaired among infants in Malawi and India
350 compared with the UK. While the later vaccination schedule in the UK (8/12 weeks vs 6/10 weeks)
351 may contribute to these differences by providing more time for passively acquired maternal antibodies
352 to wane, overall shedding rates after both doses in Malawi and India still fell short of the near-
353 ubiquitous dose 1 shedding observed in the UK, highlighting the barrier to ORV replication that
354 emerges within the first months of life in LMICs. Whereas the seroconversion rates in India and Malawi
355 are broadly consistent with those previously reported in LMICs (Bennett et al., 2019; Lazarus et al.,
356 2018), the rate was lower than expected in the UK, particularly when considering the high levels of
357 shedding observed in this cohort (Patel et al., 2013). These findings serve as a reminder of the
358 suboptimal nature of RV-IgA as a correlate of protection. Indeed, while higher RV-IgA concentrations
359 have been shown to correlate with a lower risk of subsequent rotavirus infection, these associations
360 often fail to hold when considering protection against rotavirus-associated gastroenteritis (Lewnard et
361 al., 2017). It is highly possible that vaccine virus replication may be a better marker of vaccine response
362 than seroconversion, and we advocate the further use of this measure in future trials.

363 For a risk factor to account for broad geographic trends in ORV efficacy, it is reasonable to
364 expect it to be more common or observed at higher levels in LMICs than high-income countries. This
365 was not the case for several of the factors examined in this study. Maternal serum RV-IgA levels were
366 comparable in the UK and India (although significantly higher in Malawi), while breastmilk RV-IgA
367 levels in the UK were significantly higher than those in India. These results are somewhat surprising
368 given that recent rotavirus exposure and consequent boosting of antibody levels might be expected to
369 occur more frequently among mothers in LMICs with a high rotavirus disease burden. Moreover, in
370 previous studies, RV-IgA concentrations were significantly higher in India than the USA, with
371 intermediate levels in South Africa (Moon et al., 2013; Moon et al., 2010). Our findings caution against
372 the broader extrapolation of these trends, highlighting the possibility that maternal RV-IgA levels in
373 high-income countries may in some cases equal those in LMICs. This may be due to the high

374 immunogenicity of early rotavirus exposures in high-income countries as well as the subsequent
375 boosting of antibodies by asymptomatic re-infection (Phillips et al., 2010).

376 Nonetheless, our findings confirm the potential inhibitory effect of maternal antibodies on
377 ORV. Whereas previous studies of this phenomenon have typically focused on immunogenicity
378 endpoints (Becker-Dreps et al., 2015; Moon et al., 2016), we observed a reduction in both shedding and
379 immunogenicity, suggesting that the effect of maternal antibodies is mediated in part through a
380 reduction in the replication efficiency of vaccine viruses. This could potentially involve the
381 transudation of transplacental RV-IgG antibodies across the mucosal epithelium or the direct
382 neutralisation of vaccine viruses by RV-IgA in breastmilk. Monoclonal IgG antibodies have been shown
383 to confer protection against lentivirus infection in neonatal macaques, highlighting the potential for
384 passive antibodies to act at mucosal surfaces (Baba et al., 2000). The presence of a leaky mucosal barrier
385 following early-life enteropathogen exposure could feasibly enhance the transfer of maternal RV-IgG
386 to the gut lumen among infants in Malawi and India (Cohen et al., 2014) and contribute to differences
387 in the impact of maternal antibodies in different populations, although this possibility remains
388 untested. If breastmilk RV-IgA is involved in the inhibition of ORV replication, the effect does not
389 appear to be mitigated by the temporary withholding of breastfeeding, which failed to boost ORV
390 response in controlled trials (Ali et al., 2015; Rongsen-Chandola et al., 2014). Our findings are also
391 consistent with the inhibition of downstream antigen processing by transplacental antibodies, as
392 evidenced by the strong inverse correlation between maternal RV-IgG and infant RV-IgA formation in
393 Indian infants with neonatal rotavirus infection. Overall, one could reasonably envision a combination
394 of upstream inhibition of ORV replication (via RV-IgA and/or transudated RV-IgG) and downstream
395 inhibition of antigen processing (via circulating RV-IgG) that synergistically inhibit ORV outcome.
396 However, these mechanisms alone are clearly insufficient to prevent a robust ORV response given that
397 the vaccine was ubiquitously shed and was highly immunogenic in the UK, where maternal antibody
398 exposure levels were high and >75% of infants were exclusively or partially breastfed.

399 We considered several biomarkers as a proxy for EED – a subclinical condition known to occur
400 with a high frequency in LMICs (Lindenbaum et al., 1966). Previous studies have used biomarker levels
401 from healthy adults in high-income settings as reference values. Our reporting of α1AT and MPO levels
402 among healthy infants in the UK therefore serves as a reference point for future studies. Although α1AT
403 concentrations were significantly lower in the UK than both other cohorts, MPO levels were similar in
404 the UK and Malawi, calling into question the extent to which this marker accurately captures
405 geographic differences in the onset of EED. We did not observe a significant relationship between any

406 of the measured EED biomarkers and ORV outcome, which is in keeping with the growing literature
407 in this field (Bucardo et al., 2018; Church et al., 2018). Together, these findings imply either that EED is
408 not a primary driver of oral vaccine failure or that the condition is not accurately captured by the
409 current suite of faecal and systemic biomarkers (quite possibly both).

410 We observed marked geographic discrepancies in composition of the bacterial microbiota.
411 Microbiota diversity was significantly higher among Malawian infants, while Indian infants could be
412 distinguished by their high *Bifidobacterium* abundance. The latter is interesting given that
413 *Bifidobacterium* has previously been suggested to enhance immunological memory responses to a range
414 of vaccine targets (including OPV) in Bangladesh (Huda et al., 2019), yet this genus was least abundant
415 among infants in the UK that were most responsive to ORV and was not significantly correlated with
416 vaccine outcome in any cohort. Similarly, we failed to recapitulate other associations that have between
417 documented between bacterial microbiota composition and ORV outcome, such as the higher
418 abundance of Bacteroidetes among non-responders in Ghana (Harris et al., 2017). However, we
419 observed a significant negative correlation between microbiota diversity and ORV immunogenicity –
420 apparent among Indian infants throughout infancy and in Malawian infants specifically prior to the
421 first vaccine dose. Although this is the first study to find such an association for ORV, we previously
422 reported a negative correlation between microbiota diversity and monovalent type 3 OPV response
423 among 6–11 month-old Indian infants (Praharaj et al., 2019). Several indirect explanations for this trend
424 are plausible. For example, microbiota richness may act as a proxy for other early-life exposures that
425 shape ORV outcome, including viral infections such as non-polio enteroviruses (Taniuchi et al., 2016).
426 The fact that ORV immunogenicity was impaired among Indian infants in households without access
427 to treated water would support this notion. It is also possible that exposure to greater microbial
428 diversity may foster a state of hyporesponsiveness at the mucosal epithelium that impairs oral vaccine
429 outcome. When viewed in this way, impaired vaccine response could be seen as a counterpart to
430 hyperresponsive immune states associated with low early-life microbiota diversity, such as atopic
431 disease (Abrahamsson et al., 2012).

432 Our study has several limitations. First, as noted above, RV-IgA is known to be a suboptimal
433 correlate of protection against rotavirus-associated gastroenteritis. We compensated for this by
434 considering multiple endpoints, including seroconversion and post-vaccination shedding. Second, we
435 cannot rule out the potential contribution of batch effects to the observed discrepancies in maternal
436 antibody levels, EED biomarkers, and microbiota composition. However, we went to considerable
437 lengths to ensure reproducibility and comparability across sites, using the same reagent lots and

438 standards. In the case of microbiota composition, we independently re-sequenced 10% of samples at a
439 separate facility, randomising the order of sample processing, and recapitulated the original microbiota
440 composition with high accuracy. Finally, we measured both EED and microbiota composition using
441 proxies that may fail to accurately capture key mechanisms shaping ORV response. Profiling the
442 virome, proteome, transcriptome, and metabolome can now be achieved using small sample volumes
443 and would provide valuable insight into systemic immune status at the time of vaccination (Lee et al.,
444 2019). These additional layers of OMIC data were beyond the scope of the present study but would be
445 a valuable addition to future work.

446 Notwithstanding these caveats, this study advances our understanding of the potential
447 mechanisms influencing ORV response in several ways. By considering how several risk factors of ORV
448 failure differ both within and between populations, we have been able to place these mechanisms
449 within a broader global context. Our study highlights the immunogenic nature of neonatal rotavirus
450 infection, lending support to the ongoing research into ORV administration at birth. We confirm the
451 inhibitory effect of maternal antibodies on ORV immunogenicity, and highlight for the first time the
452 impaired replication efficiency of vaccine rotavirus replication in infants exposed to higher levels of
453 maternal antibodies. Finally, while specific bacterial taxa do not appear to drive within-population
454 differences in ORV response, high microbial exposure in early life may contribute to the impaired
455 efficacy of this vaccine in LMICs.

Acknowledgements

We thank all members of the clinical study teams in Vellore, Blantyre, and Liverpool, including Falak Diab, Siobhan Holt, and the research midwives at the Liverpool Women's Hospital; Dawn Redman and the team of research nurses at Alder Hey Children's Hospital; Uma Raman, Charlet, Margaret, Jacklin, and the field research assistants at Christian Medical College, Vellore; and James Tamani, Anna Ainani, Amisa Chisale, Bertha Masamba, Carlo Gondwe, and Evelyn Gondwe in Blantyre, Malawi. The UK and Malawi sites were funded by the UK Medical Research Council and the UK Department for International Development (Newton Fund MR/N006259/1). M.I.G. is partly supported by the NIHR HPRU in Gastrointestinal Infections. K.C.J. is funded by a Wellcome International Training Fellowship (number 201945/Z/16/Z). The site in India was funded by the Government of India's Department of Biotechnology. Richard Eccles, Anita Lucaci, Richard Gregory, John Kenny, and other staff at the Centre for Genomic Research (University of Liverpool) provided valuable support for the 16S microbiota sequencing work, as did Laurence Game at the Imperial BRC Genomics Laboratory. Above all, we are grateful to the families involved in the study.

Author contributions

Conceptualisation, M.I.G., N.A.C., A.C.D., N.C.G., and G.K.; Methodology, C.B., E.P.K.P., A.C.D., M.I.G., I.P., and S.B.; Software, E.P.K.P.; Validation, E.P.K.P.; Formal Analysis, E.P.K.P.; Investigation, C.B., E.P.K.P., N.C.-V., D.H., J.L., S.B., I.P., S.G., S.I., V.K.S., B.B., S.S., B.S.R., J.M., and E.C.; Resources Provision, K.N.S., A.G., V.P.V., Q.D., N.C., M.T., and N.A.C.; Data Curation, C.B., E.P.K.P., K.N.S., S.V., J.M., and M.I.G.; Writing – Original Draft, E.P.K.P., C.B., A.C.D., S.B., I.P., A.D., M.I.G., and N.C.G.; Writing – Review & Editing, B.K., Visualisation, E.P.K.P; Supervision, N.A.C., A.C.D., K.C.J., N.C.G., B.K.; Project Administration, C.B., M.I-G., I.P., S.B., K.N.S., K.C.J.; Funding Acquisition, M.I.G., G.K.

Declaration of Interests

M.I.G. has received research grants from GSK and Merck, and has provided expert advice to GSK. K.C.J. and N.A.C. have received investigator-initiated research grant support from GSK. N.A.C. has received research grant support and honoraria for participation in rotavirus vaccine data safety monitoring committee meetings from GSK.

486 **Figure titles and legends**

487 **Figure 1. Study Design and Oral Rotavirus Vaccine Response.** (A) Study design. (B to D) Geographic
488 differences in (B) rotavirus shedding, (C) immunogenicity, and (D) cumulative vaccine response
489 (seroconversion or shedding after either dose). Error bars represent Clopper–Pearson 95% confidence
490 intervals. (E) Association between rotavirus shedding and post-vaccination antibody concentration.
491 Gene targets for shedding assays are indicated in brackets. Groups were compared by Fisher's exact
492 tests with FDR correction (binary outcomes) or ANOVAs with post-hoc Tukey tests (continuous
493 outcomes). The dotted lines at 20 IU/ml indicate the standard cut-off for RV-IgA seropositivity. ns, not
494 significant; ORV, oral rotavirus vaccine; RV, rotavirus; * $p < 0.05$; ** $p < 0.005$; *** $p < 0.0005$.

495

496 **Figure 2. Association between Maternal Antibodies and Oral Rotavirus Vaccine Response.** (A)
497 Maternal antibody concentrations by cohort. Groups were compared by ANOVAs with post-hoc Tukey
498 tests. The dotted lines at 20 IU/ml indicate the standard cut-off for RV-IgA seropositivity. (B)
499 Association between maternal antibodies and RV-IgA formation. Log-transformed concentrations were
500 compared using Pearson's correlation coefficient (*r*). (C) Correlation between antibody concentrations
501 and rotavirus shedding quantity (1/Ct) among Indian infants. Variables were compared using
502 Spearman's rank correlation coefficient (rho). neo+, infected with rotavirus pre-vaccination; neo-,
503 uninfected with rotavirus pre-vaccination; ORV, oral rotavirus vaccine; RV, rotavirus; * $p < 0.05$; ** p
504 < 0.005; *** $p < 0.0005$.

505

506 **Figure 3. Association between Environmental Enteric Dysfunction and Oral Rotavirus Vaccine**
507 **Response.** (A) Environmental enteric dysfunction biomarkers by cohort. Groups were compared by
508 ANOVAs with post-hoc Tukey tests. (B) Association between inflammatory biomarkers and RV-IgA
509 formation. Log-transformed concentrations were compared using Pearson's correlation coefficient (*r*).
510 α1AT, α1-antitrypsin; α1AG, α1 acid glycoprotein; MPO, myeloperoxidase; neo+, infected with
511 rotavirus pre-vaccination; neo-, uninfected with rotavirus pre-vaccination; ORV, oral rotavirus vaccine;
512 RV, rotavirus.

513

514 **Figure 4. Geographic Differences in Microbiota Development.** (A) Longitudinal analysis of alpha
515 diversity. Points represent mean values. Groups were compared using longitudinal mixed-effects
516 models including week as a covariate and study ID as a random effect. Pairwise comparisons between
517 countries were FDR corrected. (B) Longitudinal plot of mean genus abundances. Genera were included

518 if they were present with a mean relative abundance of ≥5% at one or more timepoints. (C) Principal
519 coordinates analysis plot of weighted Bray-Curtis distances for samples collected before the first dose
520 of ORV (week 6 in India and Malawi; week 8 in the UK). (D) Cross-validation accuracy of Random
521 Forests models (mean ± s.d.). (E) The 20 most important genera selected by Random Forests for
522 discriminating infants by country at the time of the first dose of ORV. Mean cross-validation importance
523 scores based on Gini index are depicted alongside the prevalence (≥0.1% abundance threshold) and
524 mean abundance of the corresponding genera in each country. PC, principal coordinate; ns, not
525 significant; * $p < 0.05$; *** $p < 0.0005$.
526
527 **Figure 5. Association between Microbiota Development and Oral Rotavirus Vaccine Response.** (A)
528 Longitudinal analysis of microbiota diversity. Points represent mean values. Groups were compared
529 using longitudinal mixed-effects models including week as a covariate and study ID as a random effect.
530 (B) Association between microbiota diversity and RV-IgA formation. Variables were compared using
531 Pearson's correlation coefficient (*r*). Microbiota richness values (number of genera detectable at ≥0.1%
532 abundance) were log-transformed. (C and D) Random Forests cross-validation accuracy for (C)
533 seroconversion and (D) RV-IgA concentrations based on genus abundance data (mean ± s.d.). (E)
534 Longitudinal zero-inflated negative binomial models of genus abundance. Genera with FDR-adjusted
535 p values of <0.1 are highlighted. neo+, infected with rotavirus pre-vaccination; neo-, uninfected with
536 rotavirus pre-vaccination; nf, not fitted as <10 responders; ORV, oral rotavirus vaccine; RV, rotavirus;
537 ZINB, zero-inflated negative binomial; * $p < 0.05$; ** $p < 0.005$; *** $p < 0.0005$.
538
539 **Figure 6. Integrated Analysis for Prediction of Oral Rotavirus Vaccine Response.** (A) Random Forests
540 models cross-validation accuracy for prediction of vaccine outcome based on demographic,
541 exposure/antibody data, and pre-vaccination (week 6) microbiota data (mean ± s.d.). (B) The top 10
542 variables selected by Random Forests for prediction of RV-IgA concentration based on mean cross-
543 validation importance score. (C) Multivariate regression models for post-vaccination RV-IgA. Variables
544 were included if significantly associated with RV-IgA concentration ($p < 0.05$) during univariate
545 analyses. See **Table S5** for full univariate and multivariate outputs. neo+, infected with rotavirus pre-
546 vaccination; neo-, uninfected with rotavirus pre-vaccination; MSE, mean squared error; ORV, oral
547 rotavirus vaccine; RV, rotavirus.

Table 1. Baseline characteristics of study cohorts.

	India	Malawi	UK
N	307	119	60
Date of birth			
First	03/12/2015	02/11/2016	29/09/2016
Last	15/11/2016	25/04/2018	30/03/2018
Caesarean delivery	70 (22.8)	0 (0.0)	1 (1.7)*
Birthweight (kg)	2.96 (0.42)	3.02 (0.42)	3.68 (0.48)
Female	152 (49.5)	60 (50.4)	24 (40.0)
Polio vaccine schedule			
tOPV	57 (18.6)	–	–
Mixed tOPV/bOPV	56 (18.2)	–	–
bOPV	94 (30.6)	119 (100.0)	–
IPV	100 (32.6)	–	60 (100.0)
Breastfeeding†			
Exclusive	265 (86.3)	108 (90.7)	26 (43.3)
Partial	32 (10.4)	11 (9.2)	20 (33.3)
None	10 (3.3)	0 (0.0)	14 (23.3)
Exposed to antibiotics	84 (27.4)	33 (27.7)	6 (10.0)
HIV status			
Exposed	0 (0.0)	27 (22.7)	-
Unexposed	299 (97.4)	88 (73.9)	-
Unknown	8 (2.6)	4 (3.4)	60 (100.0)

Data are n (%) or mean (s. d.).

* Elective caesarean was an exclusion criterion for this cohort (with 1 delivery by emergency caesarean).

† Data from birth to week 11 in Malawi and India, and birth to week 13 in the UK.

572 **Supplemental figure titles and legends**

573 **Figure S1. Association between pre-vaccination rotavirus exposure and oral rotavirus vaccine**
574 **response.** (A) Rotavirus shedding across all timepoints based on VP6-specific PCR. Groups were
575 compared by Fisher's exact tests with FDR correction. (B) Association between pre-vaccination
576 rotavirus exposure and ORV outcome in Indian infants. Prior exposure was defined by week 1 rotavirus
577 shedding or detection of RV-IgA at baseline (≥20 IU/ml). Exposed and unexposed groups were
578 compared by ANOVA (log antibody concentrations) or Fisher's exact tests (shedding/seroconversion).
579 The cohorts from Malawi and the UK are displayed for reference; statistical comparisons among cohorts
580 are shown in **Figure 1**. Error bars represent Clopper–Pearson 95% confidence intervals. ns, not
581 significant; ORV, oral rotavirus vaccine; RV, rotavirus; *, $p < 0.05$; **, $p < 0.005$; ***, $p < 0.0005$.

582

583 **Figure S2. Replicability of microbiota sequencing across runs and facilities.** Positive controls
584 including an infant stool (BSctrl), maternal stool (MSctrl), and a mock bacterial community (MCctrl)
585 were included in every plate, with up to up to four plates per run. Alpha diversity (A) and beta diversity
586 (B) for these positive controls were highly replicable from run to run, with sample type accounting for
587 the majority of variation among controls based on linear regression (for alpha diversity) and
588 PERMANOVA (for beta diversity). A minimum detection threshold of 0.1% abundance was used for
589 richness and unweighted Bray–Curtis metrics. (C) Proportion of variation associated with run for each
590 sample group. R^2 and statistical significance were determined separately for each sample group by
591 PERMANOVA. Mean R^2 values are indicated by dotted lines. (D–F) For 10% of samples (90 week-1
592 infant samples, 90 week-10/12 infant samples, and 90 maternal samples evenly distributed across the
593 study sites), we validated the microbiota pipeline at a separate sequencing facility in London. PCR,
594 quantification, and pooling all involved the same methods and reagents (ordered separately), although
595 sequencing was performed via Illumina MiSeq rather than HiSeq (used for the majority of runs in
596 Liverpool). The order of samples during the 10% validation was randomised across three plates.
597 Microbiota composition was highly replicable based on alpha diversity (D), beta diversity (E), and
598 relative abundance of major genera (F), with sample ID accounting for >97% of variation based on
599 linear regression or PERMANOVA. In (B), replicates are linked by a line, although often the overlap is
600 so precise that these lines are indistinguishable. The proportion of variation associated with sample
601 code was equivalent when using unweighted distances ($R^2 = 0.996$). PC, principal coordinate; * $p<0.05$.

602

Figure S3. Genus profiles of infant and maternal samples. (A) Hierarchical clustering of stool samples (columns) based on presence/absence of common genera (rows). Genera detected in 1% of samples were included. Rows and columns were clustered using Ward's minimum variance hierarchical clustering method. (B) Genus abundance versus prevalence for infant samples. Samples in each cohort were generally dominated by a small number of taxa with high prevalence and abundance. All infant samples from each cohort were included in the prevalence and abundance calculations (week 1, week 4, week 6/8, and week 10/12). Margins display density plots. (C) Geographic discrepancies in maternal microbiota composition, including comparisons of alpha diversity (left panel) and genus composition (right panel). Diversity metrics were compared using ANOVAs with post-hoc Tukey tests. Only genera from **Figure 4B** are shown. ns, not significant; **, $p < 0.05$; **, $p < 0.005$; ***, $p < 0.0005$.

Figure S4. Discriminant genera by country. (A) Longitudinal relative abundance plots for major genera (mean abundance ≥5% in at least one country for one or more timepoints). Lines show local weighted regression (loess) fits with 95% confidence intervals. (B) Mixed-effects zero-inflated negative binomial models of genus abundance. Genera were included if present in at least 20% of samples from at least one country being compared. Study ID was included as a random effect. Genera with an FDR-adjusted p values of <0.1 are highlighted. (C) Cross-sectional comparisons of genus prevalence. A minimum detection threshold of 0.1% abundance was used to define the presence/absence of a taxon. Genera with an FDR-adjusted p values of <0.1 are highlighted.

Figure S5. Top 20 genera for prediction of country using Random Forests. See **Figures 4E** and **4F** for additional details.

Figure S6. Association between microbiota development and oral rotavirus vaccine response. (A and B) Longitudinal analysis of microbiota diversity in relation to (A) dose 1 shedding status and (B) shedding after either dose. Points represent mean values. Groups were compared using longitudinal mixed-effects models including week as a covariate and study ID as a random effect. (C and D) Random Forests cross-validation accuracy for (C) dose 1 shedding and (D) shedding after either dose based on genus abundance data (mean ± s.d.). (E) Longitudinal zero-inflated negative binomial models of genus abundance. Genera with FDR-adjusted p values of <0.1 are highlighted, with further details provided in **Table S7**. (D) Number of samples per cross-validation fold during cross-sectional Random Forests models. All available samples were included in models for post-vaccination RV-IgA concentration. For

binary outcomes, we randomly selected 50 infants per group in each iteration of cross-validation, or the number of samples in the minority group if this was <50. neo+, infected with rotavirus pre-vaccination; neo-, uninfected with rotavirus pre-vaccination; nf, not fitted as <10 samples in minority group; ORV, oral rotavirus vaccine; RV, rotavirus; ZINB, zero-inflated negative binomial.

Figure S7. Removal of extraction batches with potential cross-contamination. (A) Read counts in no-template and extraction controls. All no-template controls were clear of substantial amplification (left panel), as were pooled extraction controls from Malawi and the UK (middle panel). Four pools of extraction controls from India yielded significant amplification (>10,000 sequences). When the individual extraction controls from these pools were sequenced (right panel), two gave significant amplification (corresponding to extraction batches 14 and 27). (B) Beta diversity plots and (C) genus composition of these extraction controls revealed two clusters consistent with contamination of extraction batches 14 and 27. We therefore removed the samples associated with these extraction batches from the final analysis (n = 8). NTC, no-template PCR control; PC, principal coordinate.

STAR Methods

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Miren Iturriza-Gómara (M.Iturriza-Gomara@liverpool.ac.uk).

Materials availability

This study did not generate new unique reagents.

Data and code availability

The raw sequence data generated during this study have been deposited in the European Nucleotide Archive (accession number PRJEB38948). Processed data and analysis code generated during this study are available on github (https://github.com/eparker12/RoVI).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Pregnant women were enrolled during the third trimester at three study sites: Blantyre (Malawi), Vellore (India), and Liverpool (UK). Women were included if they provided informed written consent to participate in the study and were willing to stay in the study area for 4 months following delivery. Exclusion criteria included: congenital immune deficiency; chronic renal or liver failure; other chronic illnesses which may affect immune function; non-singleton pregnancy; low birthweight or pre-term birth (<34 weeks gestation); congenital anomalies and other neonatal complications requiring prolonged hospitalisation; and delivery by elective caesarean (UK only). Mothers provided a venous blood sample as well as a cord blood sample during birth and a stool sample during the week after delivery. Infants provided two blood samples and four stool samples over the course of the study (**Figure 1**). The study was approved by the Institutional Review Board at the Christian Medical College (CMC) in Vellore, the College of Medicine Research and Ethics Committee in Blantyre, and the North West – Liverpool East Research Ethics Committee in Liverpool.

METHOD DETAILS

Sample processing and storage

Whole blood was collected in anti-coagulation EDTA-tubes (BD) and stored at 4°C for up to 12 hours until collection by laboratory staff. Fractions were separated by centrifugation at 1,500–2,000 g in

681 a benchtop centrifuge (Eppendorf). A disposable plastic transfer pipette was used to aspirate the
682 plasma down to ~1 mm from the red blood cells. The plasma was aliquoted into two screw cap cryo-
683 tubes (Starlab). The buffy coat and red blood cells were subsequently collected into separate cryovials
684 and all samples were stored at -70°C.

685 Stool and breastmilk samples were collected in sterile sample pots by participants and were
686 shipped to the respective laboratory by courier within 24 hours in the UK and within 4 hours in India
687 and Malawi (to account for ambient temperatures). Samples were kept at 4°C for a maximum of 8 hours
688 until processing. Breastmilk samples were stored in 2 ml aliquots in SuperLock tubes (Starlab) at -70°C
689 for maternal RV-IgA analysis. Two 10% stool suspensions were prepared in sterile phosphate-buffered
690 saline (PBS) for assessing rotavirus shedding. Further aliquots of neat stool were stored in 2 ml
691 SuperLock tubes (Starlab) at -70°C for microbiota analysis and inflammatory biomarker measurement.
692 Stool samples were stored at -70°C for a maximum of 2 weeks prior to DNA extraction for microbiota
693 analysis.

694

695 **Rotavirus-specific antibodies**

696 RV-IgA and RV-IgG laboratory assays were conducted at CMC, Vellore, for samples from all
697 three study sites using a custom antibody-sandwich ELISA (Paul et al., 2014). Briefly, 96-well plates
698 (Costar) coated with rabbit hyperimmune serum to rotavirus were incubated with purified cell culture
699 lysates (WC3) or mock-infected MA104 cells. Serial dilutions of standard and test sera were added
700 followed by biotinylated rabbit anti-human IgA (Jackson ImmunoResearch Laboratories) for detection
701 of RV-IgA and biotinylated rabbit anti-human IgG (Vector Laboratories). Absorbance was
702 subsequently read at 492 nm. Background-corrected optical density values from sample wells were
703 compared with the standard curve and IgA or IgG concentration was determined based on derived
704 units of IgA or IgG arbitrarily assigned to the respective standard curve, with a minimum detection
705 limit of 1 IU/ml. Seroconversion was defined as detection of RV-IgA at ≥20 IU/ml in previously
706 seronegative infants or a 4-fold increase in RV-IgA concentration among infants who were seropositive
707 at baseline.

708

709 **Rotavirus shedding**

710 Quantification of rotavirus shedding was performed at each study site. RNA was extracted
711 using the RNeasy mini kit (Qiagen) following the manufacturer's instructions. Briefly, 10% PBS stool
712 suspensions were defrosted on ice, vortexed, and briefly centrifuged (10,000 g for 10 seconds) to pellet

713 larger debris. 250 µl of β-ME in buffer RTL was added to 250 µl of supernatant. The mixture was then
714 vortexed before the addition of 250 µl of 70% molecular-grade ethanol. The RNeasy Mini Handbook
715 (5th Edition, October 2013) was followed for subsequent steps, and RNA was eluted in 50 µl RNase-free
716 water. cDNA was generated immediately after RNA extraction. The master mix for cDNA conversion
717 comprised: 7 µl of 10x Taq DNA PCR buffer (Invitrogen), 7 µl of 50 mM $MgCl_2$ (BIOLINE), 1 µl of
718 Random Primers (Invitrogen), 2 µl of 10 mM dNTP (Invitrogen), 1 µl of DTT (Invitrogen), 2 µl of
719 SuperScript III reverse transcriptase (Invitrogen), and 10 µl of RNase-free water (Invitrogen). 40 µl of
720 eluted RNA was denatured at 95°C for 5 minutes and subsequently cooled on ice for 2 minutes. 30 µl
721 of master mix was added to each RNA sample, mixed gently by pipetting, and centrifuged briefly. The
722 tube was placed in a thermal cycler and incubated at 25°C for 10 minutes, 37°C for 1 hour, then 95°C
723 for 5 minutes. The cDNA was stored at -20°C.

724 Previously validated primers and probes were used for the qPCR-based detection of rotavirus
725 NSP2 (Gautam et al., 2016) and VP6 (Iturriza Gomara et al., 2002) in two separate reactions. cDNA and
726 reagents were thawed on ice. The master mix for qPCR comprised: 12.5 µl of Platinum qPCR supermix-
727 UDG 2x (Invitrogen), 2 µl of 10 µM forward primer (5′-GAACTTCCTTGAA-
728 TATAAGATCACACTGA-3′ [NSP2] or 5′-GACGGVGCRACTACATGGT-3′ [VP6]), 2 µl of 10 µM
729 reverse primer (5′-TTGAAGACGTAAATGCATACCAATTC-3′ [NSP2] or 5′-GTCCAATTCAT-
730 NCCTGGTG-3′ [VP6]), 0.25 µl of 10 µM probe (FAM5′-TCCAATAGATTGAAGTCAGTAACGT-
731 TTCCA-3′MGB [NSP2] or FAM5′-CCACCRAAYATGACRCCAGCNGTA-3′MGB [VP6]), and 6.25 µl
732 of nuclease-free water (ThermoFisher Scientific). In a 96-well plate, 23 µl of master mix was combined
733 with 2 µl of cDNA. Six standards for the target gene as well as a no-template control and a positive
734 control were added to each PCR plate. The plate was sealed with a Microseal 'C' Film (Bio-Rad) and
735 subjected to the following cycles in a Bio-Rad CFX qPCR Instrument (Bio-Rad): initial UDG-incubation
736 50°C for 2 minutes; denaturation for 95°C for 2 minutes; then 40 cycles of 95°C for 15 seconds and 60°C
737 for 1 minute. Data were considered valid if amplification was absent in the negative control and present
738 in the positive control. qPCR standards consisting of TOPO-TA plasmid constructs containing either
739 NPS2 or VP6 gene amplicons (genotype 1) were produced at the University of Liverpool and
740 distributed to the other study sites.

741

742 **Inflammatory biomarkers**

743 For infant stools collected before each dose of ORV, MPO and α1AT concentrations were
744 measured by ELISA at each site. ELISA kits with the same lot number were distributed across the study

745 sites. Stool samples were defrosted on ice and 100 mg then aliquoted into a polypropylene tube. ELISAs
746 were carried out following the manufacturer's instructions (IDK 2018 instructions for MPO and
747 BioVendor RIC6200 for α1AT). Paired samples from a given infant were run on the same plate, and
748 MPO and α1AT assays were carried out in tandem to avoid freeze-thaw cycles. For week 6 serum
749 samples of Indian and Malawian infants, α1AG was measured at CMC, Vellore, using a commercial
750 ELISA kit (abcam) according to the manufacturer's instructions.
751
752 **16S microbiota composition**
753 *DNA extraction from stool*
754 DNA extraction from stool was carried out using the QIAamp DNA stool mini kit (Qiagen)
755 according to the manufacturer's instructions (QIAamp DNA Stool Handbook 06/2012), with several
756 modifications to increase DNA yield. Briefly, the vortexed suspension of ASL buffer and stool was
757 added to 370 mg of 0.1 mm zirconia-silicate beads (BioSpec) with 1.67 µl of lyzozyme (30 mg/ml)
758 (Sigma-Aldrich). Samples were then incubated at 37°C on a shaker (Eppendorf Thermo Mixer Comfort)
759 at 250-300 rpm for 10 minutes. 10 µl of proteinase K (Qiagen), 50 µl of 10% sodium dodecyl sulphate
760 (Sigma-Aldrich), and 20 µl of RNase A (1mg/ml; Sigma-Aldrich) was added before incubation on a
761 heating block (Eppendorf Thermo Mixer Comfort) at 70°C for 10 minutes. After allowing the samples
762 to cool for 3 minutes, bead beating was performed for 5 minutes using a Tissue Lyser II (Qiagen) at 25
763 Hz. Samples were microcentrifuged at 16,100 g for 1 minute and the supernatant was then transferred
764 into a new microcentrifuge tube containing an InhibitEX tablet and vortexed for 1 minute. The
765 suspension was incubated at room temperature for 1 minute, centrifuged at 16,100 g for 5 minutes, and
766 the supernatant (approximately 600 µl) then transferred into a new 2 ml tube. The supernatant was
767 centrifuged at 16,100 g for 3 minutes, then 400 µl added to a 2 ml tube containing 400 µl of buffer AL.
768 400 µl of ethanol (96–100%) was added to the lysate and mixed by vortexing. The manufacturer's
769 instructions were then resumed until elution in 50 µl of AE Buffer. An extraction control was carried
770 out with each batch of DNA extraction at all sites. Extracted DNA was shipped from India and Malawi
771 to the University of Liverpool on dry ice for 16S rRNA sequencing.
772
773 *Amplicon generation for 16S rRNA sequencing*
774 Amplicons spanning 16S rRNA gene variable regions 3 and 4 (primers 309F 5'-overhang-
775 ACTCCTACGGGAGGCAGCAG-3' and 819R 5'-overhang-GGACTACHVGGGTWTCTAAT-3') were
776 produced following the established Illumina protocol (Illumina, 16S Metagenomic Sequencing Library

Preparation Protocol Part # 15044223 Rev. B) with the following amendments: 1 µl of DNA was used as a starting template; NEBNext Q5 Hot Start HiFi PCR Master Mix (NEB) was used for amplicon and indexing PCRs; PCR reaction volumes were halved to 25 µl; primers were used at a concentration of 10 µM; and 20 µg of molecular-grade bovine serum albumin (NEB) was added to the amplicon PCR master mix to mitigate the effect of PCR inhibitors. Cycling conditions for amplicon PCR involved: denaturation at 98°C for 30 seconds; 10 cycles of 98°C for 10 seconds, 55°C for 15 seconds, and 72°C for 40 seconds; and final extension at 72°C for 60 seconds. Cycling conditions for indexing PCR involved: denaturation at 98°C for 3 minutes; 15 cycles of 98°C for 10 seconds, 55°C for 15 seconds, and 72°C for 40 seconds; and final extension at 72°C for 5 minutes. AMPure XP beads were substituted with a custom preparation of Sera-Mag SpeedBeads Protein A/G particles (MERCK/GE Healthcare) throughout the protocol. To determine volumes required for equimolar pooling, amplicon concentrations were determined by Quanti-it (ThermoFisher Scientific) and size distributions by fragment analysis on a 5300 Fragment Analyzer System (Agilent). Amplicons were pooled using a mosquito X1 (TTP Labtech) liquid handling robot. The final library underwent size selection to remove potentially contaminating primer-dimers and genomic DNA using the Pippin Prep (Sage Science) 1.5% Agarose Gel Cassette (Labtech).

Overall, we sequenced amplicons from 2,138 samples across 14 Illumina HiSeq2500 lanes (v2 chemistry, 600 cycles in rapid run mode) and one Illumina MiSeq lane (v3 chemistry, 600 cycles). Since enrolment and sample collection across the three sites were not synchronised, sequencing was batched by geographic origin according to sample availability. Samples from a mother–infant pair were processed on the same PCR plate. Each PCR plate contained a no-template PCR control, stool controls provided by a mother and infant in the UK who were not enrolled in the study, DNA from a mock community (Zymo Research D6306), and a pool of extraction controls corresponding to the samples contained on the PCR plate. Final libraries consisted of up to four PCR plates (384 amplicons). Eight samples were excluded from the analysis owing to the presence of significant amplification from their corresponding extraction controls (**Figure S7**).

Illumina sequencing of 16S rRNA libraries

The quantity and quality of each amplicon pool was assessed by Qubit HS DNA kit (ThermoFisher Scientific), Bioanalyzer High Sensitivity DNA Kit (Agilent), and qPCR using the KAPA Library Quantification Kit (Roche) on a Roche Light Cycler LC480II (Roche) according to the manufacturer's instructions. qPCR data were used to calculate sample molarity. 5 µl of the final pool

809 was denatured for 5 minutes at room temperature using 5 μl of freshly diluted 0.1 N sodium hydroxide
810 (Illumina). The reaction was subsequently terminated by the addition of HT1 hybridization buffer
811 (Illumina) and the library diluted post-denaturation to a final loading concentration of 7–8.5 pM.
812 Libraries were sequenced with 10–20% PhiX (Illumina) using the 2x300 bp paired-end protocol.
813
814 *Independent validation of 16S rRNA sequencing*
815 To validate the robustness of the microbiota protocol, 10% of DNA samples were shipped to
816 Imperial College London and sequenced according to the methods described above. We included 30
817 mother–infant pairs per study site in the validation subset and sequenced two infant stools (from weeks
818 1 and 10/12) and the maternal stool from each pair. Protocol deviations included the use of AMPure
819 XP beads for PCR product purification, fragment analysis via TapeStataion D1000 ScreenTape (Agilent
820 Technologies), concentration quantification via Qubit dsDNA High Sensitivity Assay Kit
821 (ThermoFisher Scientific), and manual pooling of equimolar quantities. The order of samples was
822 randomised across three PCR plates and the pooled libraries were sequenced with 10–20% PhiX
823 (Illumina) on two 2x300 bp paired-end Illumina MiSeq runs (v3 chemistry).
824
825 QUANTIFICATION AND STATISTICAL ANALYSIS
826 **Bioinformatic processing of sequence data**
827 Adapters were trimmed from raw sequences using cutadapt version 1.18 (Martin, 2011).
828 Subsequent steps for quality filtering, denoising, merging of paired reads, and chimera removal
829 followed the DADA2 pipeline (Callahan et al., 2016) implemented in QIIME 2 with default parameters
830 (Bolyen et al., 2019). Forward reads were truncated to 270 bp and reverse reads to 200 bp to account for
831 the fall in sequencing quality scores towards the end of each read. Sequence data were processed
832 separately for each run, and the feature tables and corresponding sequences then merged. All
833 subsequent steps were performed in the programming language R. Taxonomy assignment was
834 performed using the RDP naïve Bayesian classifier trained on the Silva rRNA database (version 132).
835 Sequence variants were included in the analysis if they were 390–440 bp in length, bacterial, detectable
836 at an abundance of ≥0.1% in at least two samples, and passed frequency-based contamination filtering
837 using the *decontam* package (which screens sequences that are more abundant in low-concentration
838 samples) (Davis et al., 2018). Nanodrop concentrations (ng/μl) of extracted DNA served as the basis
839 for contaminant filtering. Samples with at least 25,000 quality-filtered sequences were retained in the
840 final analysis.

841 *Statistical analysis*

842 All analyses were performed in the programming language R. Seroconversion and shedding
843 proportions were compared by country using pairwise Fisher's exact tests with Benjamini–Hochberg
844 FDR correction. Antibody and inflammatory biomarker concentrations were log transformed and
845 compared by country and shedding subgroup (see **Figure 1E**) using ANOVA with post-hoc Tukey tests.
846 To explore covariates associated with ORV outcome, we used either logistic regression (for
847 seroconversion and shedding status) or linear regression and Pearson's *r* (for log-transformed post-
848 vaccination RV-IgA). In Indian mother–infant dyads, Spearman's rank correlation coefficients were
849 calculated between all measured antibody concentrations (RV-IgG and RV-IgA) and rotavirus
850 shedding quantities (reciprocal of qPCR Ct value). A hierarchical clustering analysis of all infant and
851 maternal gut microbiota samples was conducted using Ward's minimum variance hierarchical
852 clustering method based on the presence or absence of common genera (≥1% prevalence at ≥0.1%
853 abundance). Longitudinal mixed-effects models were used to compare microbiota diversity (Shannon
854 index, genus-level richness) by country (pairwise comparisons with FDR correction) and ORV outcome
855 (seroconversion and shedding status), including week as a covariate and study ID as a random effect.
856 We used longitudinal zero-inflated negative binomial models of genus counts to identify taxa that
857 discriminated infants according to country, seroconversion, and shedding status, adjusting for age (by
858 including week as a fixed effect) and including study ID as a random effect. For taxa with a prevalence
859 of >95%, negative binomial models without zero inflation were used. Fisher's exact test was also used
860 for cross-sectional comparisons of genus prevalence (≥0.1% abundance threshold) by country,
861 excluding taxa that were ubiquitous (>95% in both comparison groups) or scarce (<5% in both groups).
862 We report on genera with FDR-adjusted p values of <0.1.

863 We applied the Random Forests algorithm in a series of cross-sectional analyses to predict
864 country, seroconversion, or shedding status (classification approach), or post-vaccination log-
865 transformed RV-IgA concentration (regression approach) based on genus relative abundances. For each
866 analysis, we performed 20 iterations of 5-fold cross-validation. For classification models (of binary
867 outcomes), we standardised the baseline accuracy at 50% by fitting each iteration of 5-fold cross-
868 validation on a random subset of 50 samples per group (or the number of samples in the minority group
869 if this was <50). Models were excluded if there were <10 samples in the minority group. Mean cross-
870 validation variable importance was determined based on the increase in Gini index (classification) or
871 mean squared error (regression) of out-of-bag sample prediction following random permutation. For
872 Indian infants, we assessed additional models based on demography/baseline health measurements

(19 variables), exposure/antibody data (12 variables, encompassing inflammatory biomarkers, wild-type rotavirus shedding, and maternal antibodies), and genus relative abundances (128 variables), and all variables combined.

For technical replicates (including positive controls and samples selected for independent validation), we quantified the proportion of variance explained by sample ID using linear regression (alpha diversity, genus abundances) or PERMANOVA with 999 permutations (weighted and unweighted Bray–Curtis distances). We also used PERMANOVA to quantify the proportion of variation in microbiota composition associated with sequencing run, stratified by age and country.

ADDITIONAL RESOURCES

The trial is registered with the Clinical Trials Registry of India (CTRI/2015/11/006354) and the study protocol has previously been published (Sindhu et al., 2017).

885 **Supplemental table titles and legends**

886 **Table S1. Sample counts for primary outcomes.** α1AG, α1 acid glycoprotein; α1AT, α1-antitrypsin;
887 MPO, myeloperoxidase; neo+, infected with rotavirus pre-vaccination; neo-, uninfected with rotavirus
888 pre-vaccination; RV, rotavirus.

889

890 **Table S2. Cofactors associated with neonatal rotavirus infection.** α1AG, α1 acid glycoprotein; α1AT,
891 α1-antitrypsin; bOPV, bivalent oral poliovirus vaccine; GMC, geometric mean concentration; IPV,
892 inactivated poliovirus vaccine; LRT, likelihood ratio test for fit of model with vs without polio vaccine
893 schedule; MPO, myeloperoxidase; ORV, oral rotavirus vaccine; RR, relative risk; RV, rotavirus; tOPV,
894 trivalent oral poliovirus vaccine; *log-transformed to approximate normality in statistical models.

895

896 **Table S3A. Cofactors associated with rotavirus vaccine shedding.** Data are presented for (A) dose 1
897 shedding and (B) shedding after either dose. Measurements obtained after the first dose of ORV are
898 excluded to focus on potential predictors of ORV shedding rather than shedding–induced changes.
899 α1AG, α1 acid glycoprotein; α1AT, α1–antitrypsin; bOPV, bivalent oral poliovirus vaccine; GMC,
900 geometric mean concentration; IPV, inactivated poliovirus vaccine; LRT, likelihood ratio test for fit of
901 model with vs without polio vaccine schedule; MPO, myeloperoxidase; ORV, oral rotavirus vaccine;
902 RR, relative risk; RV, rotavirus; tOPV, trivalent oral poliovirus vaccine; *log–transformed to
903 approximate normality in statistical models; † excluded to minimise multicollinearity.

904

905 **Table S4. Cofactors associated with rotavirus seroconversion.** α1AG, α1 acid glycoprotein; α1AT, α1-
906 antitrypsin; bOPV, bivalent oral poliovirus vaccine; GMC, geometric mean concentration; IPV,
907 inactivated poliovirus vaccine; LRT, likelihood ratio test for fit of model with vs without polio vaccine
908 schedule; MPO, myeloperoxidase; ORV, oral rotavirus vaccine; RR, relative risk; RV, rotavirus; tOPV,
909 trivalent oral poliovirus vaccine; *log-transformed to approximate normality in statistical models; †
910 excluded to minimise multicollinearity.

911

912 **Table S5. Cofactors associated with post-vaccination RV-IgA concentration.** α1AG, α1 acid
913 glycoprotein; α1AT, α1-antitrypsin; bOPV, bivalent oral poliovirus vaccine; GMC, geometric mean
914 concentration; IPV, inactivated poliovirus vaccine; LRT, likelihood ratio test for fit of model with vs
915 without polio vaccine schedule; MPO, myeloperoxidase; ORV, oral rotavirus vaccine; RV, rotavirus;

tOPV, trivalent oral poliovirus vaccine; *log-transformed to approximate normality in statistical models; † excluded to minimise multicollinearity.

Table S6. Geographic discrepancies in genus prevalence. Data are presented for (A) Malawi vs India, (B) India vs UK, and (C) Malawi vs UK. Genus prevalences were compared using Fisher's exact test, with a 0.1% abundance threshold to define presence/absence. Taxa were excluded if ubiquitous (>95% in both groups) or scarce (<5% in both groups).

Table S7. Taxa correlated with ORV response in longitudinal zero-inflated negative binomial models. FDR, false discovery rate; ORV, oral rotavirus vaccine; RV, rotavirus.

References

Abrahamsson, T.R., Jakobsson, H.E., Andersson, A.F., Bjorksten, B., Engstrand, L., and Jenmalm, M.C. (2012). Low diversity of the gut microbiota in infants with atopic eczema. J Allergy Clin Immunol *129*, 434-440, 440 e431-432.

Ali, A., Kazi, A.M., Cortese, M.M., Fleming, J.A., Moon, S., Parashar, U.D., Jiang, B., McNeal, M.M., Steele, D., Bhutta, Z., *et al.* (2015). Impact of withholding breastfeeding at the time of vaccination on the immunogenicity of oral rotavirus vaccine--a randomized trial. PLoS One *10*, e0127622.

Arrieta, M.C., Stiemsma, L.T., Dimitriu, P.A., Thorson, L., Russell, S., Yurist-Doutsch, S., Kuzeljevic, B., Gold, M.J., Britton, H.M., Lefebvre, D.L., *et al.* (2015). Early infancy microbial and metabolic alterations affect risk of childhood asthma. Sci Transl Med *7*, 307ra152.

Baba, T.W., Liska, V., Hofmann-Lehmann, R., Vlasak, J., Xu, W., Ayehunie, S., Cavacini, L.A., Posner, M.R., Katinger, H., Stiegler, G., *et al.* (2000). Human neutralizing monoclonal antibodies of the IgG1 subtype protect against mucosal simian-human immunodeficiency virus infection. Nat Med *6*, 200-206.

Bar-Zeev, N., King, C., Phiri, T., Beard, J., Mvula, H., Crampin, A.C., Heinsbroek, E., Lewycka, S., Tate, J.E., Parashar, U.D., *et al.* (2018). Impact of monovalent rotavirus vaccine on diarrhoea-associated post-neonatal infant mortality in rural communities in Malawi: a population-based birth cohort study. Lancet Glob Health *6*, e1036-e1044.

Becker-Dreps, S., Vilchez, S., Velasquez, D., Moon, S.S., Hudgens, M.G., Zambrana, L.E., and Jiang, B. (2015). Rotavirus-specific IgG antibodies from mothers' serum may inhibit infant immune responses to the pentavalent rotavirus vaccine. Pediatr Infect Dis J *34*, 115-116.

Bennett, A., Pollock, L., Jere, K.C., Pitzer, V.E., Lopman, B., Parashar, U., Everett, D., Heyderman, R.S., Bar-Zeev, N., Cunliffe, N.A., *et al.* (2019). Infrequent transmission of monovalent human rotavirus vaccine virus to household contacts of vaccinated infants in Malawi. J Infect Dis *219*, 1730-1734.

Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., Alexander, H., Alm, E.J., Arumugam, M., Asnicar, F., *et al.* (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol *37*, 852-857.

Bucardo, F., Nordgren, J., Reyes, Y., Gonzalez, F., Sharma, S., and Svensson, L. (2018). The Lewis A phenotype is a restriction factor for Rotateq and Rotarix vaccine-take in Nicaraguan children. Sci Rep *8*, 1502.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J., and Holmes, S.P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods *13*, 581-583.

Church, J.A., Chasekwa, B., Rukobo, S., Govha, M., Lee, B., Carmolli, M.P., Ntozini, R., Mutasa, K., McNeal, M.M., Majo, F.D., *et al.* (2020). Predictors of oral rotavirus vaccine immunogenicity in rural Zimbabwean infants. Vaccine *38*, 2870-2878.

Church, J.A., Parker, E.P., Kirkpatrick, B.D., Grassly, N.C., and Prendergast, A.J. (2019). Interventions to improve oral vaccine performance: a systematic review and meta-analysis. Lancet Infect Dis *19*, 203-214.

Church, J.A., Parker, E.P., Kosek, M.N., Kang, G., Grassly, N.C., Kelly, P., and Prendergast, A.J. (2018). Exploring the relationship between environmental enteric dysfunction and oral vaccine responses. Future Microbiol *13*, 1055-1070.

Cohen, O.R., Steele, J.A., Zhang, Q., Schmidt, D.J., Wang, Y., Hamel, P.E., Beamer, G., Xu, B., and Tzipori, S. (2014). Systemically administered IgG anti-toxin antibodies protect the colonic mucosa during infection with Clostridium difficile in the piglet model. PLoS One *9*, e111075.

Davis, N.M., Proctor, D.M., Holmes, S.P., Relman, D.A., and Callahan, B.J. (2018). Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. Microbiome *6*.

Gautam, R., Mijatovic-Rustempasic, S., Esona, M.D., Tam, K.I., Quaye, O., and Bowen, M.D. (2016). One-step multiplex real-time RT-PCR assay for detecting and genotyping wild-type group A rotavirus strains and vaccine strains (Rotarix(R) and RotaTeq(R)) in stool samples. PeerJ *4*, e1560.

Harris, V., Ali, A., Fuentes, S., Korpela, K., Kazi, M., Tate, J., Parashar, U., Wiersinga, W.J., Giaquinto, C., de Weerth, C., *et al.* (2018). Rotavirus vaccine response correlates with the infant gut microbiota composition in Pakistan. Gut Microbes *9*, 93-101.

Harris, V.C., Armah, G., Fuentes, S., Korpela, K.E., Parashar, U., Victor, J.C., Tate, J., de Weerth, C., Giaquinto, C., Wiersinga, W.J., *et al.* (2017). Significant correlation between the infant gut microbiome and rotavirus vaccine response in rural Ghana. J Infect Dis *215*, 34-41.

Huda, M.N., Ahmad, S.M., Alam, M.J., Khanam, A., Kalanetra, K.M., Taft, D.H., Raqib, R., Underwood, M.A., Mills, D.A., and Stephensen, C.B. (2019). Bifidobacterium abundance in early infancy and vaccine response at 2 years of age. Pediatrics *143*.

Iturriza Gomara, M., Wong, C., Blome, S., Desselberger, U., and Gray, J. (2002). Molecular characterization of VP6 genes of human rotavirus isolates: correlation of genogroups with subgroups and evidence of independent segregation. J Virol *76*, 6596-6601.

Lazarus, R.P., John, J., Shanmugasundaram, E., Rajan, A.K., Thiagarajan, S., Giri, S., Babji, S., Sarkar, R., Kaliappan, P.S., Venugopal, S., *et al.* (2018). The effect of probiotics and zinc supplementation on the immune response to oral rotavirus vaccine: A randomized, factorial design, placebo-controlled study among Indian infants. Vaccine *36*, 273-279.

Lee, A.H., Shannon, C.P., Amenyogbe, N., Bennike, T.B., Diray-Arce, J., Idoko, O.T., Gill, E.E., Ben-Othman, R., Pomat, W.S., van Haren, S.D., *et al.* (2019). Dynamic molecular changes during the first week of human life follow a robust developmental trajectory. Nat Commun *10*, 1092.

Lewnard, J.A., Lopman, B.A., Parashar, U.D., Bar-Zeev, N., Samuel, P., Guerrero, M.L., Ruiz-Palacios, G.M., Kang, G., and Pitzer, V.E. (2017). Naturally acquired immunity against rotavirus infection and gastroenteritis in children: paired reanalyses of birth cohort studies. J Infect Dis *216*, 317-326.

Lindenbaum, J., Alam, A.K., and Kent, T.H. (1966). Subclinical small-intestinal disease in East Pakistan. Br Med J *2*, 1616-1619.

Madhi, S.A., Cunliffe, N.A., Steele, D., Witte, D., Kirsten, M., Louw, C., Ngwira, B., Victor, J.C., Gillard, P.H., Cheuvart, B.B., *et al.* (2010). Effect of human rotavirus vaccine on severe diarrhea in African infants. N Engl J Med *362*, 289-298.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnetjournal *17(1)*, 10-12.

Mokomane, M., Kasvosve, I., de Melo, E., Pernica, J.M., and Goldfarb, D.M. (2018). The global problem of childhood diarrhoeal diseases: emerging strategies in prevention and management. Ther Adv Infect Dis *5*, 29-43.

Moon, S.S., Groome, M.J., Velasquez, D.E., Parashar, U.D., Jones, S., Koen, A., van Niekerk, N., Jiang, B., and Madhi, S.A. (2016). Prevaccination rotavirus serum IgG and IgA are associated with lower immunogenicity of live, oral human rotavirus vaccine in South African infants. Clin Infect Dis *62*, 157-165.

Moon, S.S., Tate, J.E., Ray, P., Dennehy, P.H., Archary, D., Coutsoudis, A., Bland, R., Newell, M.L., Glass, R.I., Parashar, U., *et al.* (2013). Differential profiles and inhibitory effect on rotavirus vaccines of nonantibody components in breast milk from mothers in developing and developed countries. Pediatr Infect Dis J *32*, 863-870.

Moon, S.S., Wang, Y., Shane, A.L., Nguyen, T., Ray, P., Dennehy, P., Baek, L.J., Parashar, U., Glass, R.I., and Jiang, B. (2010). Inhibitory effect of breast milk on infectivity of live oral rotavirus vaccines. Pediatr Infect Dis J *29*, 919-923.

Nair, N.P., Reddy, N.S., Giri, S., Mohan, V.R., Parashar, U., Tate, J., Shah, M.P., Arora, R., Gupte, M., Mehendale, S.M., *et al.* (2019). Rotavirus vaccine impact assessment surveillance in India: protocol and methods. BMJ Open *9*, e024840.

Naylor, C., Lu, M., Haque, R., Mondal, D., Buonomo, E., Nayak, U., Mychaleckyj, J.C., Kirkpatrick, B., Colgate, R., Carmolli, M., *et al.* (2015). Environmental enteropathy, oral vaccine failure and growth faltering in infants in Bangladesh. EBioMedicine *2*, 1759-1766.

Parker, E.P., Ramani, S., Lopman, B.A., Church, J.A., Iturriza-Gomara, M., Prendergast, A.J., and Grassly, N.C. (2018a). Causes of impaired oral vaccine efficacy in developing countries. Future Microbiol *13*, 97-118.

Parker, E.P.K., Praharaj, I., Zekavati, A., Lazarus, R.P., Giri, S., Operario, D.J., Liu, J., Houpt, E., Iturriza-Gomara, M., Kampmann, B., *et al.* (2018b). Influence of the intestinal microbiota on the immunogenicity of oral rotavirus vaccine given to infants in south India. Vaccine *36*, 264-272.

Patel, M., Glass, R.I., Jiang, B., Santosham, M., Lopman, B., and Parashar, U. (2013). A systematic review of anti-rotavirus serum IgA antibody titer as a potential correlate of rotavirus vaccine efficacy. J Infect Dis *208*, 284-294.

Paternina-Caicedo, A., Parashar, U.D., Alvis-Guzman, N., De Oliveira, L.H., Castano-Zuluaga, A., Cotes-Cantillo, K., Gamboa-Garay, O., Coronell-Rodriguez, W., and De la Hoz-Restrepo, F. (2015). Effect of rotavirus vaccine on childhood diarrhea mortality in five Latin American countries. Vaccine *33*, 3923-3928.

Paul, A., Gladstone, B.P., Mukhopadhya, I., and Kang, G. (2014). Rotavirus infections in a community based cohort in Vellore, India. Vaccine *32 Suppl 1*, A49-54.

1029 Phillips, G., Lopman, B., Rodrigues, L.C., and Tam, C.C. (2010). Asymptomatic rotavirus infections in England:
1030 prevalence, characteristics, and risk factors. Am J Epidemiol *171*, 1023-1030.
1031 Pollock, L., Bennett, A., Jere, K.C., Dube, Q., Mandolo, J., Bar-Zeev, N., Heyderman, R.S., Cunliffe, N.A., and
1032 Iturriza-Gomara, M. (2018). Non-secretor histo-blood group antigen phenotype is associated with reduced risk of
1033 clinical rotavirus vaccine failure in Malawian infants. Clin Infect Dis.
1034 Praharaj, I., Parker, E.P.K., Giri, S., Allen, D.J., Silas, S., Revathi, R., Kaliappan, S.P., John, J., Prasad, J.H.,
1035 Kampmann, B., *et al.* (2019). Influence of nonpolio enteroviruses and the bacterial gut microbiota on oral poliovirus
1036 vaccine response: a study from south India. J Infect Dis *219*, 1178-1186.
1037 Ramani, S., Sowmyanarayanan, T.V., Gladstone, B.P., Bhowmick, K., Asirvatham, J.R., Jana, A.K., Kuruvilla, K.A.,
1038 Kumar, M., Gibikote, S., and Kang, G. (2008). Rotavirus infection in the neonatal nurseries of a tertiary care hospital
1039 in India. Pediatr Infect Dis J *27*, 719-723.
1040 Rongsen-Chandola, T., Strand, T.A., Goyal, N., Flem, E., Rathore, S.S., Arya, A., Winje, B.A., Lazarus, R.,
1041 Shanmugasundaram, E., Babji, S., *et al.* (2014). Effect of withholding breastfeeding on the immune response to a
1042 live oral rotavirus vaccine in North Indian infants. Vaccine *32 Suppl 1*, A134-139.
1043 Shao, Y., Forster, S.C., Tsaliki, E., Vervier, K., Strang, A., Simpson, N., Kumar, N., Stares, M.D., Rodger, A.,
1044 Brocklehurst, P., *et al.* (2019). Stunted microbiota and opportunistic pathogen colonization in caesarean-section
1045 birth. Nature.
1046 Sindhu, K.N., Cunliffe, N., Peak, M., Turner, M., Darby, A., Grassly, N., Gordon, M., Dube, Q., Babji, S., Praharaj,
1047 I., *et al.* (2017). Impact of maternal antibodies and infant gut microbiota on the immunogenicity of rotavirus
1048 vaccines in African, Indian and European infants: protocol for a prospective cohort study. BMJ Open *7*, e016577.
1049 Stewart, C.J., Ajami, N.J., O'Brien, J.L., Hutchinson, D.S., Smith, D.P., Wong, M.C., Ross, M.C., Lloyd, R.E.,
1050 Doddapaneni, H., Metcalf, G.A., *et al.* (2018). Temporal development of the gut microbiome in early childhood
1051 from the TEDDY study. Nature *562*, 583-588.
1052 Taniuchi, M., Platts-Mills, J.A., Begum, S., Uddin, M.J., Sobuz, S.U., Liu, J., Kirkpatrick, B.D., Colgate, E.R.,
1053 Carmolli, M.P., Dickson, D.M., *et al.* (2016). Impact of enterovirus and other enteric pathogens on oral polio and
1054 rotavirus vaccine performance in Bangladeshi infants. Vaccine *34*, 3068-3075.
1055 Tate, J.E., Burton, A.H., Boschi-Pinto, C., Parashar, U.D., and World Health Organization-Coordinated Global
1056 Rotavirus Surveillance, N. (2016). Global, regional, and national estimates of rotavirus mortality in children <5
1057 years of age, 2000-2013. Clin Infect Dis *62 Suppl 2*, S96-S105.
1058 Troeger, C., Blacker, B.F., Khalil, I.A., Rao, P.C., Cao, S., Zimsen, S.R., and al., e. (2018). Estimates of the global,
1059 regional, and national morbidity, mortality, and aetiologies of diarrhoea in 195 countries: a systematic analysis for
1060 the Global Burden of Disease Study 2016. Lancet Infect Dis *18*, 1211-1228.
1061 Valdes, A.M., Walter, J., Segal, E., and Spector, T.D. (2018). Role of the gut microbiota in nutrition and health. BMJ
1062 *361*, k2179.

1063 Vesikari, T., Karvonen, A., Prymula, R., Schuster, V., Tejedor, J.C., Cohen, R., Meurice, F., Han, H.H., Damaso, S.,
1064 and Bouckenooghe, A. (2007). Efficacy of human rotavirus vaccine against rotavirus gastroenteritis during the first
1065 2 years of life in European infants: randomised, double-blind controlled study. Lancet *370*, 1757-1763.
1066 Voysey, M., Kelly, D.F., Fanshawe, T.R., Sadarangani, M., O'Brien, K.L., Perera, R., and Pollard, A.J. (2017). The
1067 influence of maternally derived antibody and infant age at vaccination on infant vaccine responses: an individual
1068 participant meta-analysis. JAMA Pediatr *171*, 637-646.
1069 Yatsunenko, T., Rey, F.E., Manary, M.J., Trehan, I., Dominguez-Bello, M.G., Contreras, M., Magris, M., Hidalgo,
1070 G., Baldassano, R.N., Anokhin, A.P., *et al.* (2012). Human gut microbiome viewed across age and geography.
1071 Nature *486*, 222-227.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit RV-IgG	CMC Vellore, in house	N/A
Biotinylated rabbit anti-human IgA	Jackson ImmunoResearch Laboratories	CAT#309-065-008; RRID:AB_2339673
Biotinylated rabbit anti-human IgG	Vector Laboratories	CAT#BA-3080
Bacterial and Virus Strains		
ZymoBIOMICS DNA microbial community standard	Zymo Research	CAT#D6300
Rotarix	GSK	N/A
Biological Samples		
Maternal, infant and cord blood samples; maternal and infant stool samples; breastmilk samples	UK: Alder Hey Children's Hospital, Liverpool Women's Hospital; India: CMC Vellore; Malawi: Malawi-Liverpool Wellcome Trust	N/A
Chemicals, Peptides, and Recombinant Proteins		
Sera-Mag™ SpeedBeads™ protein A/G particles	Sigma-Aldrich; MERCK (GE Healthcare)	CAT#GE17152104010150
Molecular grade BSA	NEB	CAT# B9000S
Critical Commercial Assays		
QIAamp DNA Stool Mini Kit	Qiagen	CAT#51504
RNeasy Mini Kit (50)	Qiagen	CAT#74104
NEBNext® Q5® Hot Start HiFi PCR Master Mix	NEB	CAT#M0543
Platinum Quantitative PCR Supermix UDG 2x	Invitrogen	CAT#11730-025
Superscript III Reverse Transcriptase	Invitrogen	CAT#10368252
Taq DNA Polymerase PCR Buffer (10x)	Invitrogen	CAT#18067017
HiSeq Rapid SBS Kit v2 (50 cycles)	Illumina	CAT#FC-402-4022
HiSeq Rapid SBS Kit v2 (500 cycles)	Illumina	CAT#FC-402-4023
MiSeq Reagent Kit v3 (600 cycles)	Illumina	CAT#MS-102-3003
PhiX Sequencing Control V3	Illumina	CAT#FC-110-3001
KAPA Library Quantification Kit	Roche	CAT#KK4854 - 07960298001
Agilent High Sensitivity DNA Kit	Agilent	CAT#5067-4627
1.5% Agarose Gel Cassette	Labtech	CAT#CDF1510
Myeloperoxidase (stool and urine) ELISA	IDK - Immunodiagnostic AG	CAT#K6630, LOT#171127
Alpha-1-Anti-Trypsin (A1AT) ELISA	Oxford Biosystems	CAT#RIC IC6200, LOT#171204
Human Alpha 1 acid Glycoprotein ELISA Kit (AGP)	abcam	CAT#ab243675, LOT#GR3200798-1
Deposited Data		
European Nucleotide Archive	Accession number PRJEB38948	N/A

Github repository of code and metadata	https://github.com/eparker12/RoVI	N/A
Experimental Models: Cell Lines		
WC3 cell line	CMC Vellore, in house (original virus lot obtained from Shantha Biotech)	N/A
MA104 cell line	CMC Vellore, in house (original cell line obtained from Estes Laboratory, Baylor)	N/A
Oligonucleotides		
VP6 forward primer (5'-GACGGVGCRACTACATGGT-3')	Sigma-Aldrich	Iturriza Gomara et al., 2002
VP6 reverse primer (5'-GTCCAATTCATNCCTGGTG-3')	Sigma-Aldrich	Iturriza Gomara et al., 2002
VP6 probe (FAM_MGB) (5'-CCACCRAAYATGACRCCAGCNGTA-3')	Life Technologies	EUROROTANET resource (https://www.eurorotanet.com/project-information/documents-and-methods)
NSP2 forward primer (5'-GAACTTCTTGAATATAAGATCACACTGA-3')	Sigma-Aldrich	Gautam et al., 2014
NSP2 reverse primer (5'-TTGAAGACGTAAATGCATACCAATTC-3')	Sigma-Aldrich	Gautam et al., 2014
NSP2 probe (FAM_BHQ1) (5'-TCCAATAGATTGAAGTCAGTAACGTTTC-CA-3')	Sigma-Aldrich	Gautam et al., 2014
V3-V4 319F primer (5'-TCGTCGGCAGCGTCAGATGTGTATAAGA-GACAGACTCCTACGGGAGGCAGCAG-3')	Integrated DNA Technologies	Fadrosh et al., 2014
V3-V4 806R primer (5'-GTCTCGTGGGCTCGGAGATGTGTATAAG-AGACAGGGACTACHVGGGTWTCTAAT -3')	Integrated DNA Technologies	Fadrosh et al., 2014
Nextera XT v2 Index Kit set A	Illumina	CAT#FC-131-2001
Nextera XT v2 Index Kit set D	Illumina	CAT#FC-131-2004
Random primers	Invitrogen	CAT#10646313
Software and Algorithms		
R Statistical Computing Software	The R Foundation	https://www.r-project.org/; RRID:SCR_001905
Qiime2	Bolyen et al., 2019	https://qiime2.org
DADA2	Callahan et al., 2016	https://benjjneb.github.io/dada2/tutorial.html
Cutadapt v.1.18	Martin, 2011	http://gensoft.pasteur.fr/docs/cutadapt/1.9.1/guide.html
Decontam R package	Davis et al., 2018	https://github.com/benjjneb/decontam
Phyloseq R package	McMurdie and Holmes, 2013	joey711.github.io/phyloseq/
ggplot2 R package	Wickham, 2016	https://ggplot2.tidyverse.org
randomForest R package	Liaw and Wiener, 2002	https://cran.r-project.org/web/packages/r

		andomForest/randomForest.pdf
Other		
0.1mm Zirconia/Silica beads	BioSpec USA	CAT#11079101z

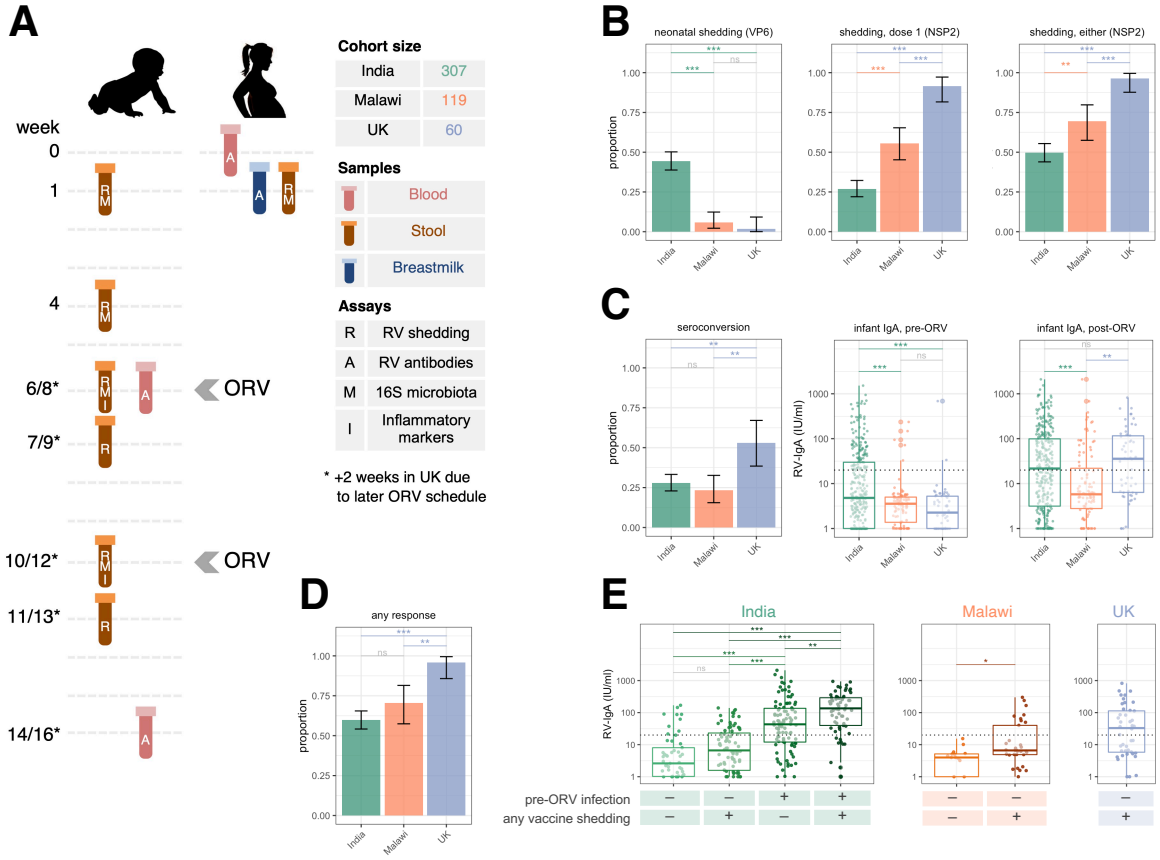

A
week
0
1
4
6/8*
7/9*
10/12*
11/13*
14/16*
ORV
ORV
Cohort size
India 307
Malawi 119
UK 60
Samples
Blood
Stool
Breastmilk
Assays
R RV shedding
A RV antibodies
M 16S microbiota
I Inflammatory markers
* +2 weeks in UK due to later ORV schedule
B
neonatal shedding (VP6)
shedding, dose 1 (NSP2)
shedding, either (NSP2)
proportion
India
Malawi
UK
C
seroconversion
infant IgA, pre-ORV
infant IgA, post-ORV
RV-IgA (IU/ml)
D
any response
E
India
Malawi
UK
pre-ORV infection
any vaccine shedding

Figure 1

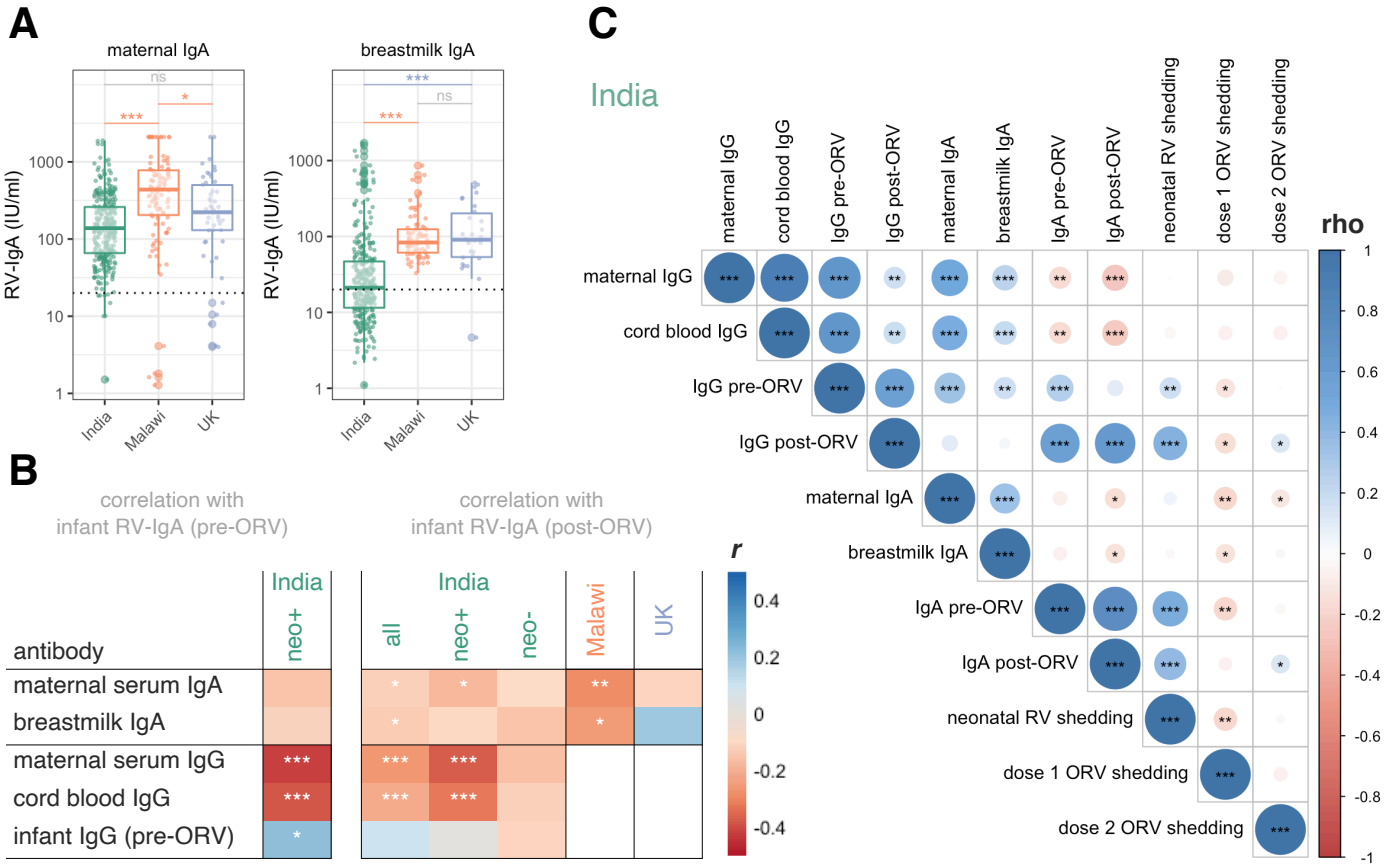

A
maternal IgA
breastmilk IgA
RV-IgA (IU/ml)
1000
100
10
1
ns
***
*
India
Malawi
UK
B
correlation with
infant RV-IgA (pre-ORV)
correlation with
infant RV-IgA (post-ORV)
India
neo+
India
all
neo+
neo-
Malawi
UK
antibody
maternal serum IgA
breastmilk IgA
maternal serum IgG
cord blood IgG
infant IgG (pre-ORV)
r
0.4
0.2
0
-0.2
-0.4
C
India
maternal IgG
cord blood IgG
IgG pre-ORV
IgG post-ORV
maternal IgA
breastmilk IgA
IgA pre-ORV
IgA post-ORV
neonatal RV shedding
dose 1 ORV shedding
dose 2 ORV shedding
rho
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1

Figure 2

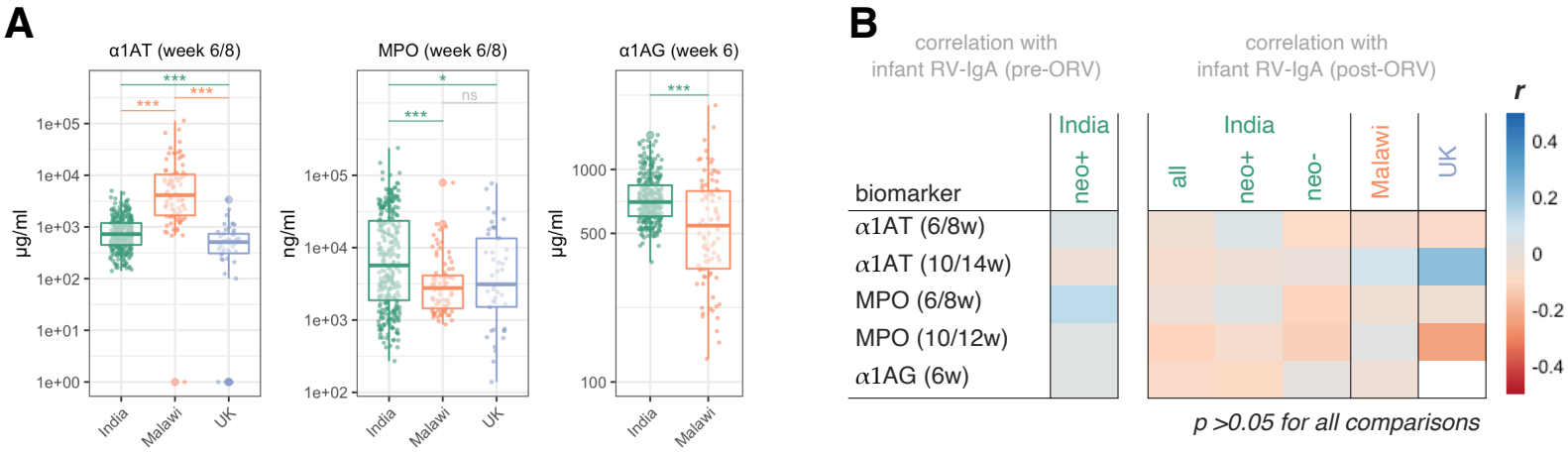

A
α1AT (week 6/8)
MPO (week 6/8)
α1AG (week 6)
μg/ml
ng/ml
μg/ml
1e+05
1e+04
1e+03
1e+02
1e+01
1e+00
1000
500
100
***
*
ns
India
Malawi
UK
B
correlation with
infant RV-IgA (pre-ORV)
correlation with
infant RV-IgA (post-ORV)
biomarker
India
neo+
India
all
neo+
neo-
Malawi
UK
α1AT (6/8w)
α1AT (10/14w)
MPO (6/8w)
MPO (10/12w)
α1AG (6w)
r
0.4
0.2
0
-0.2
-0.4
p >0.05 for all comparisons

Figure 3

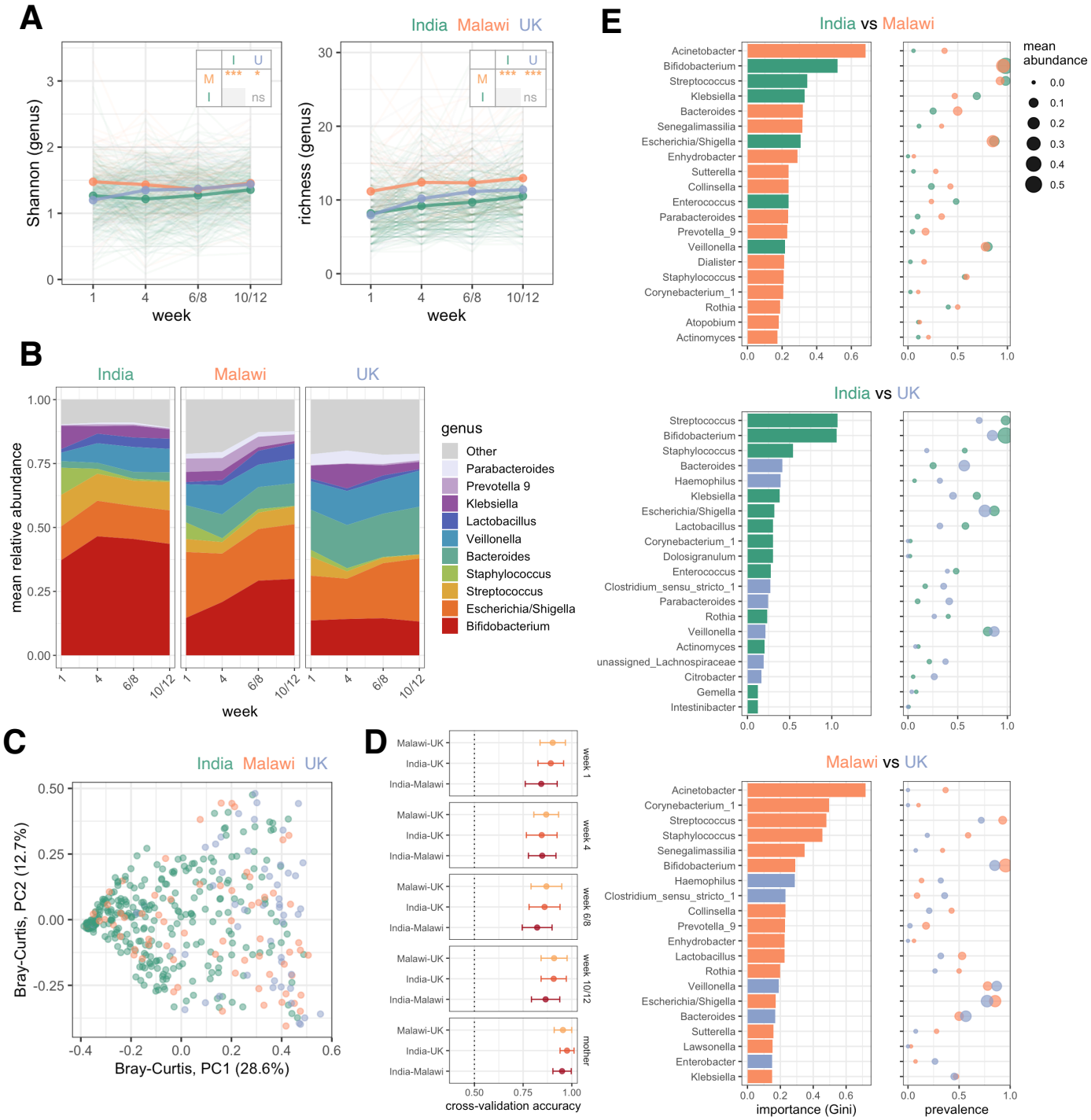
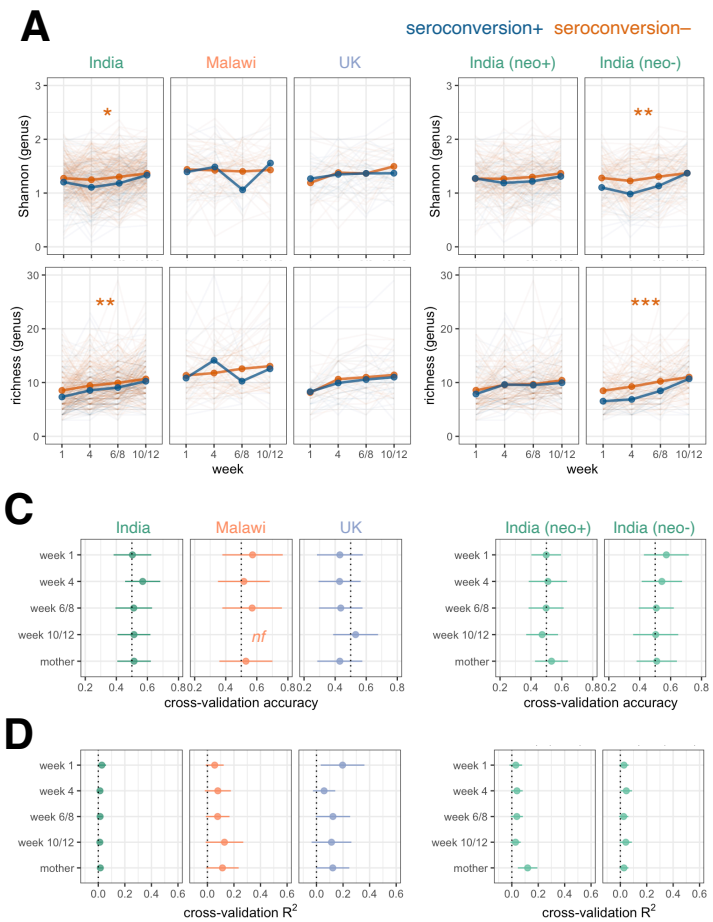
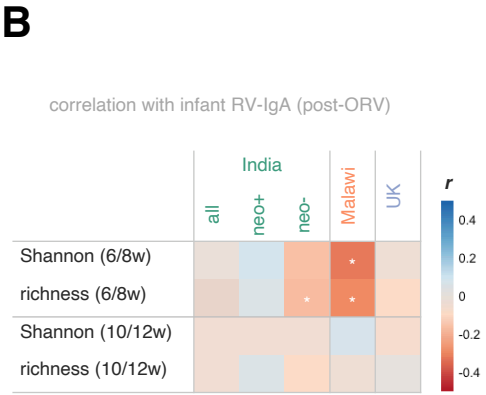

A
India Malawi UK
Shannon (genus)
richness (genus)
week
B
India Malawi UK
mean relative abundance
genus
Other
Parabacteroides
Prevotella 9
Klebsiella
Lactobacillus
Veillonella
Bacteroides
Staphylococcus
Streptococcus
Escherichia/Shigella
Bifidobacterium
C
India Malawi UK
Bray-Curtis, PC2 (12.7%)
Bray-Curtis, PC1 (28.6%)
D
Malawi-UK
India-UK
India-Malawi
week 1
week 4
week 6/8
week 10/12
mother
cross-validation accuracy
E
India vs Malawi
India vs UK
Malawi vs UK
mean abundance
importance (Gini)
prevalence

Figure 4





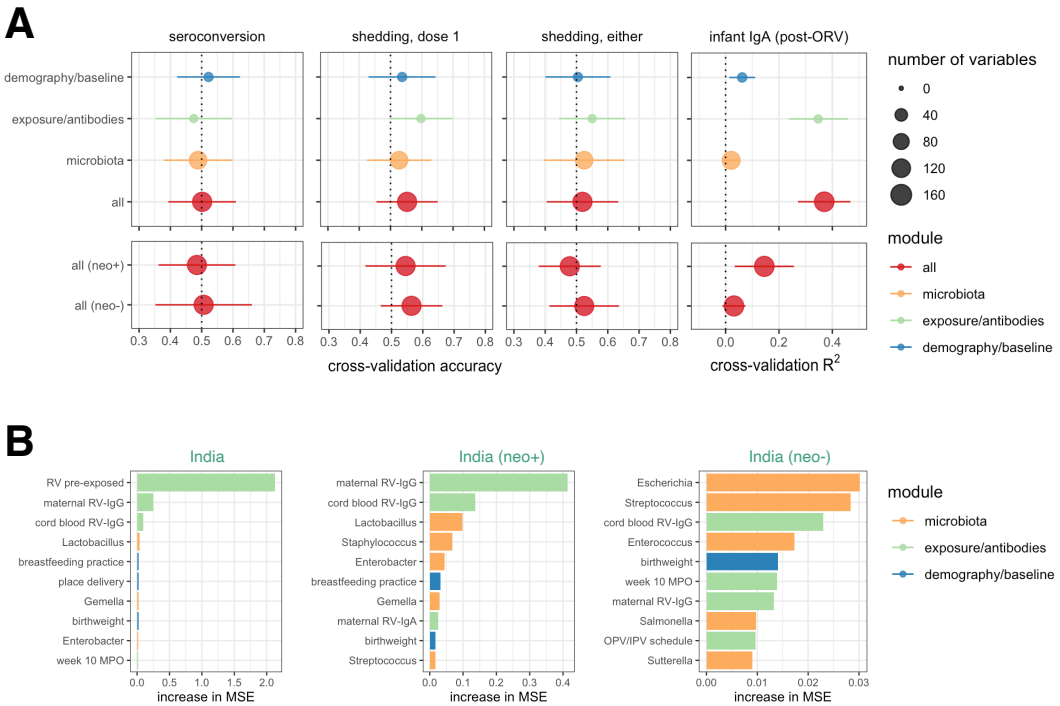
E

longitudinal ZINB models of taxon abundance

	India			Malawi	UK
	all	neo+	neo-		
N genera tested*	25	26	24	41	25
Increases with age	7	8	8	15	4
Decreases with age	6	7	6	12	13
+ correlation with seroconversion	0	0	0	0	0
– correlation with seroconversion	0	0	0	0	0
+ correlation with RV-IgA post-ORV	0	0	0	0	0
– correlation with RV-IgA post-ORV	0	1	0	0	0

** Present in ≥20% of samples*

Figure 5



C

outcome: infant RV-IgA (post-ORV)	India (full, n = 286)		India (neo+, n = 166)		India (neo-, n = 124)	
variable	adjusted beta (95% CI)	p	adjusted beta (95% CI)	p	adjusted beta (95% CI)	p
delivery mode (caesarean vs vaginal)	1.45 (0.93–2.28)	0.102	–	–	1.94 (0.90–4.20)	0.090
delivery place (tertiary vs non-tertiary)	1.08 (0.72–1.64)	0.701	–	–	1.08 (0.55–2.10)	0.831
breastfeeding practice (exclusive vs non-exclusive)	**1.89 (1.11–3.22)**	**0.019**	**3.45 (1.62–7.33)**	**0.001**	–	–
height-for-age Z score (6 weeks)	**1.31 (1.10–1.55)**	**0.002**	1.25 (0.99–1.58)	0.061	–	–
exposed to rotavirus pre-vaccination	**8.48 (5.67–12.67)**	**<0.001**	–	–	–	–
polio vaccine schedule	–	–	–	–	–	–
bOPV only	ref	–	–	–	–	–
bOPV/tOPV mixed	0.48 (0.28–0.85)	0.011	–	–	–	–
tOPV only	**0.43 (0.25–0.75)**	**0.003**	–	–	–	–
IPV	**0.77 (0.49–1.22)**	**0.265**	–	–	–	–
house made of permanent materials (yes vs no)	1.28 (0.86–1.91)	0.228	–	–	**1.69 (1.00–2.85)**	**0.049**
access to treated water (yes vs no)	1.35 (0.93–1.96)	0.120	–	–	**1.82 (1.08–3.06)**	**0.026**
maternal RV-IgG (IU/ml), log-transformed	**0.59 (0.49–0.71)**	**<0.001**	**0.54 (0.41–0.70)**	**<0.001**	–	–
Shannon index (6 weeks)	–	–	–	–	**0.44 (0.23–0.84)**	**0.013**
adjusted R^2	0.434		0.194		0.114	

Figure 6