




# BMJ Open Gut microbiome, enteric infections and child growth across a rural–urban gradient: protocol for the ECoMiD prospective cohort study

Gwenyth O Lee <sup>1</sup>, Joseph N S Eisenberg,<sup>1</sup> Jessica Uruchima,<sup>1</sup> Gabriela Vasco <sup>2,3</sup>, Shanon M Smith,<sup>4</sup> Amanda Van Engen,<sup>1</sup> Courtney Victor,<sup>4</sup> Elise Reynolds,<sup>1</sup> Rebecca MacKay,<sup>4</sup> Kelsey J Jesser,<sup>5</sup> Nancy Castro,<sup>6</sup> Manuel Calvopiña,<sup>7</sup> Konstantinos T Konstantinidis,<sup>8</sup> William Cevallos,<sup>9</sup> Gabriel Trueba,<sup>2</sup> Karen Levy <sup>5</sup>

**To cite:** Lee GO, Eisenberg JNS, Uruchima J, *et al.* Gut microbiome, enteric infections and child growth across a rural–urban gradient: protocol for the ECoMiD prospective cohort study. *BMJ Open* 2021;**11**:e046241. doi:10.1136/bmjopen-2020-046241

► Prepublication history for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2020-046241>).

Received 27 October 2020  
Accepted 29 September 2021



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

## Correspondence to

Dr Gwenyth O Lee;  
golee@umich.edu

## ABSTRACT

**Introduction** The functional consequences of the bacterial gut microbiome for child health are not well understood. Characteristics of the early child gut microbiome may influence the course of enteric infections, and enteric infections may change the composition of the gut microbiome, all of which may have long-term implications for child growth and development.

**Methods and analysis** We are conducting a community-based birth cohort study to examine interactions between gut microbiome conditions and enteric infections, and how environmental conditions affect the development of the gut microbiome. We will follow 360 newborns from 3 sites along a rural–urban gradient in northern coastal Ecuador, characterising enteric infections and gut microbial communities in the children every 3 to 6 months over their first 2 years of life. We will use longitudinal regression models to assess the correlation between environmental conditions and gut microbiome diversity and presence of specific taxa, controlling for factors that are known to be associated with the gut microbiome, such as diet. From 6 to 12 months of age, we will collect weekly stool samples to compare microbiome conditions in diarrhoea stools versus stools from healthy children prior to, during and after acute enteric infections, using principal-coordinate analysis and other multivariate statistical methods.

**Ethics and dissemination** Ethics approvals have been obtained from Emory University and the Universidad San Francisco de Quito institutional review boards. The findings will be disseminated through conference presentations and peer-reviewed journals.

## INTRODUCTION

Enteric pathogens are a significant cause of morbidity and mortality among young children in low and middle-income countries (LMICs). However, many infections are either asymptomatic or result in only mild-to-moderate disease.<sup>1 2</sup> In other instances, infection with the same pathogen results in severe diarrhoea, requiring hospitalisation

## Strengths and limitations of this study

- A birth cohort design allows us to examine prospectively how environmental factors impact the gut microbiome over the critical first 2 years of life.
- A longitudinal birth cohort design allows us to examine how early microbiome conditions shape responses to a range of enteric infections.
- A rural–urban gradient allows us to assess how social, environmental and dietary factors impact the relationship between enteric pathogen transmission and the gut microbiome in a population that otherwise is culturally and genetically similar.
- Differences in microbiome diversity may be explained by non-environmental factors that we fail to fully characterise and adjust for (eg, genetics).
- We may be underpowered to examine less common pathogens.

and sometimes leading to death. Cumulative enteric infections early in life, even when asymptomatic, are associated with chronic conditions such as environmental enteric dysfunction (EED),<sup>3</sup> growth faltering<sup>4 5</sup> and cognitive deficits.<sup>6</sup> As a result, understanding the factors that underlie relative susceptibility or resilience has become a critical question in the field of enteric disease research.

In a healthy microbiome, resident microorganisms may reduce the risk of pathogen invasion, while in an unhealthy microbiome, potential pathogens may outcompete commensal microbes.<sup>7</sup> Additionally, the gut microbiome can be adversely impacted by the cycle of infection and malnutrition often experienced by children whose growth is faltering.<sup>8 9</sup> Children with acute diarrhoea have reduced taxonomic microbiome diversity<sup>10 11</sup> and exhibit predictable patterns of

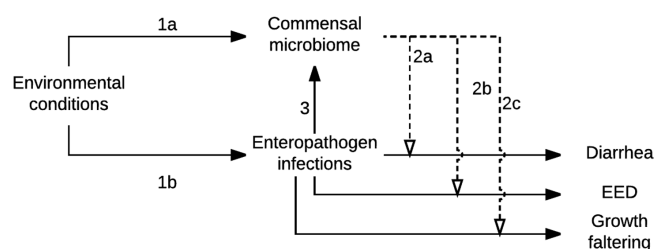
structural and functional gut microbial succession as the microbiome recovers.<sup>12</sup> Severely malnourished children also have a ‘less mature’ microbiome compared with healthy infants,<sup>13</sup> and mice implanted with microbiota from undernourished child donors have less weight gain than mice transplanted with microbiota from healthy child donors.<sup>14 15</sup> A better understanding of how the gut microbiome both protects against and responds to enteric infections may help to mitigate the negative health consequences of these infections.<sup>16</sup>

Across geographic locations, major differences exist in the taxonomic diversity and composition of the gut microbiome.<sup>17–22</sup> Consistently, higher fecal bacterial alpha diversity (within individuals) and lower beta diversity (between individuals) have been reported from remote traditional societies, compared with more ‘westernised’ societies.<sup>17–22</sup> These patterns in bacterial community diversity may be driven by the maternal diet and microbiome,<sup>23</sup> infant dietary habits and nutrition,<sup>24 25</sup> exposure to animals<sup>26</sup> and chemicals<sup>27</sup> or other lifestyle factors.<sup>28–30</sup> Prenatal maternal weight, diet and antibiotic usage have all been shown to impact the infant gut microbiota.<sup>31</sup> Drinking water quality is also likely a contributor to the composition and development of the gut microbiome<sup>32</sup> as well as a source of exposure to enteric pathogens.<sup>33</sup> While these patterns are intriguing, the causal linkage between these population-level determinants, gut microbiome conditions and, ultimately, child health are not well understood. Information about the environmental conditions that shape the early child gut microbiome may help guide the development of appropriate community and household-level interventions that will lead to optimal microbiome development.

Here, we present the protocol for the ‘gut microbiome, enteric infections and child growth across a rural–urban gradient’ or ‘Enteropatógenos, Crecimiento, Microbioma, y Diarrea’ (ECoMiD) study. The primary hypothesis of this study is that the gut microbiome mediates the effect of enteric infections on diarrhoea, EED and growth in the first 2 years of life. Our specific study objectives are to examine: (1) how environmental conditions affect the developing child gut microbiome and enteric pathogen burden, (2) whether the gut microbiome modifies the short-term and long-term health outcomes associated with enteric pathogen infections (ie, diarrhoea, EED and child growth) and (3) how the gut microbiome responds to and recovers from enteric pathogen infection (figure 1).

## METHODS AND ANALYSIS

Longitudinal data are needed to characterise conditions prior to infection and to compare those data to the commensal gut microbiome response following infection, enabling us to better disaggregate cause and effect for interactions between infection and the microbiome. Similarly, longitudinal data are useful in identifying whether alterations to the microbiome precede, accompany or follow from EED and growth faltering. We have,



**Figure 1** Overview diagram of the relationships the study will explore. The numbers on the diagram refer to the specific aims that will address the relationship. Solid lines indicate the main effects that will be tested in models, and dashed lines indicate effect modification. EED, environmental enteric dysfunction.

therefore, developed a community-based birth cohort study of 360 mother–child dyads.

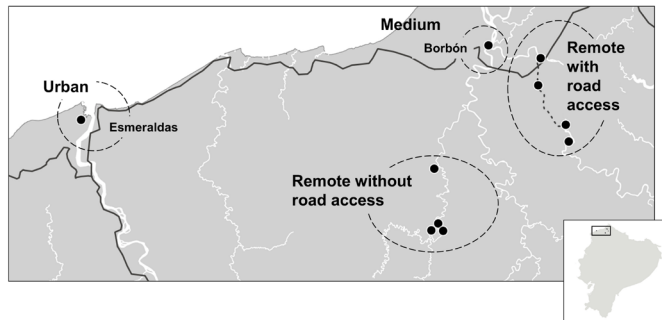
## Study sites

Our study sites span a rural–urban gradient in Ecuador and include the ‘urban’ site of the city of Esmeraldas (population ~1 500 000<sup>34</sup>), the ‘intermediate’ town of Borbón (population ~4500) and several ‘rural’ communities located between 1 and 3 hours travel from Borbón (community populations ~400 to 920) (figure 2). This rural–urban gradient study design provides a population in a high enteric pathogen transmission setting that has variability in microbiome diversity, yet relative similarity in other cultural, social and genetic factors (eg, child feeding practices and race) that might determine host response to infections, thereby allowing us to isolate the effects of the gut microbiome on the health impacts of enteric infections. Similar types of enteric infections are present across the gradient; however, the relative and overall prevalence of these pathogens varies. For example, pathogenic *Escherichia coli*, rotavirus and *Giardia* are more common in less remote areas.<sup>35</sup> We have also observed differences in microbiome composition by community membership as well as by diarrhoea case–control status.<sup>36 37</sup> Previous studies have placed the prevalence of stunting in our study region at around 15%<sup>38–40</sup> and the 2-week prevalence of diarrhoea in children under 5 at around 10%.<sup>41</sup>

## Study design

Women are recruited and enrolled in late pregnancy, from 37 weeks onward, where gestational age is based on the mother’s report. The study began in May 2019, and we expect to complete data collection in early 2024.

Once their children are born, we characterise enteric infections and gut microbial communities in the children at multiple time points over their first 2 years of life. We also comprehensively evaluate environmental conditions that are associated with enteric pathogen exposure and assess other factors that are known to be associated with the infant gut microbiome, including the maternal microbiome in late pregnancy, caesarean versus vaginal delivery,<sup>42</sup> infant diet<sup>24 25</sup> and infant nutritional status.<sup>43</sup>



**Figure 2** Study sites. The urban–rural gradient of the ECoMiD study: the city of Esmeraldas, the town of Borbón and smaller rural communities in the Canton of Eloy Alfaro, including communities with and without road access. ECoMiD, Enteropatógenos, Crecimiento, Microbioma, y Diarrea.

Prior to the cohort, we comprehensively piloted study instruments and sample collection procedures among 75 women with 0 to 2-year-old children at the urban site (city of Esmeraldas).

### Eligibility criteria

Our enrolment strategy is intended to capture healthy pregnancies and minimise the impact of potential confounders related to the development of the gut microbiota. Therefore, study exclusion criteria (table 1) include planned caesarean section, or high-risk pregnancy according to Ministry of Health guidelines, as this is a risk factor for unplanned caesarean section. Women who plan to leave their current community within 6 months or who are under the age of 18 years old are also not eligible. If an enrolled participant experiences an unplanned caesarean section, or any pregnancy complication, they continue to participate in the study. We will perform subsequent analyses to determine how data from these participants should be used.

### Enrolment

We have several recruitment strategies, including recruitment of women attending prenatal services through the local health centre, snowball sampling (participating mothers may nominate their acquaintances) and radio announcements that are broadcast throughout the study area. After potential participants are identified, field staff contact the pregnant woman to confirm criteria before inviting her to participate in the study.

The consent process occurs in-person with the mother, who consents on behalf of herself and her child. This procedure takes place in a private location within the local health centre or in the mother’s home. Study team members, who are community members trained in the ethical conduct of research, are responsible for obtaining consent. To ensure comprehension, the field staff verbally summarise the study using the consent form as a guide and provide a one-page pictorial description of study activities. The staff makes clear that study participation is voluntary, unrelated to access to prenatal or postnatal care and that the mother and her child may withdraw from the study at any time. Written informed consent is sought, unless the mother is illiterate. In this case, the consent form is read to her in the presence of an impartial witness, and a digital impression (thumbprint) is obtained in place of a signature. A copy of the consent form is provided to the mother, and the signed copy is maintained in a secure, locked cabinet in the study offices. Throughout the study, the field team periodically reminds mothers of planned study procedures. Mothers are also offered the opportunity to decline to participate in individual activities.

Benefits of study participation include child growth monitoring, the timely identification of parasitic infections and anaemia, for which appropriate treatment is coordinated through the Ministry of Health, and information about household water quality from environmental testing. We also provide mothers with general

**Table 1** Study inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria	Justification	Additional information
Normal pregnancy	High-risk pregnancy based on the opinion of the attending healthcare professional. HIV + mothers will be counted as high risk. Known multiple pregnancies will also be counted as high risk.	Our aim is to examine the development of the microbiome in healthy children.	Determination of whether a pregnancy is ‘high-risk’ is based on the opinion of the medical professional providing prenatal care to the woman, according to the Ecuadorian Ministry of Health schema, and not by the study team. In the region, 90% of pregnant women receive HIV testing and the estimated prevalence of HIV is 1.13%. <sup>78</sup>
Mother plans to remain in study area at least 6 months	Mother indicates plans to move to a community outside the study catchment area within 6 months of delivery.	This population is highly mobile, so this criterion is included to increase retention.	Participants with planned movement from one rural community to nearby rural communities would still be eligible. Participants moving between established study sites will also be retained (eg, mother recruited in Borbón who moves temporarily or permanently to Esmeraldas).
Planned vaginal delivery	Planned delivery by caesarean section.	Delivery type is known to impact the child’s microbiota.	Mothers who are enrolled who have unplanned caesarean deliveries will continue in the study.

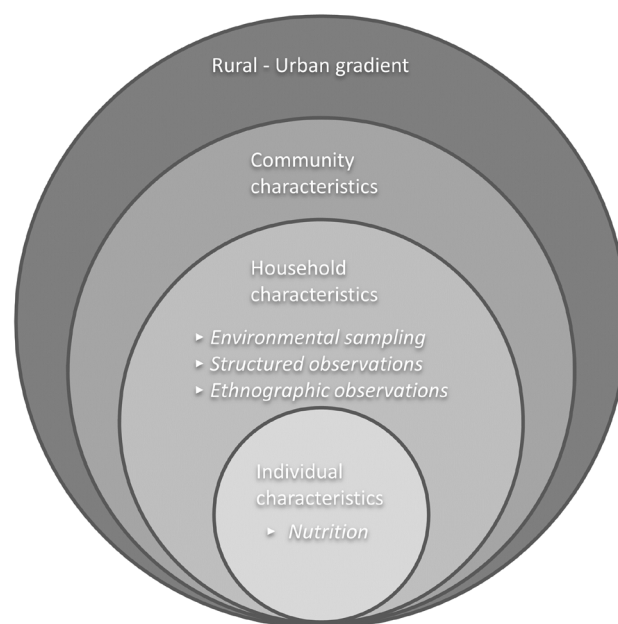
health advice. Study incentives include objects such as soap, baby oil and small toys, of approximately US\$3–5, distributed every 3–6 months.

### Withdrawal

If an enrolled mother–child dyad decides to withdraw or travels permanently outside the study area, the study team documents their exit from the study. Mothers who travel outside the study area and then return may resume study participation, and mothers who travel between study sites remain enrolled. If the mother makes an active decision to withdraw from the study, a supervisor communicates with her to ensure that her reasons for withdrawal are clear to the team.

### Data collection

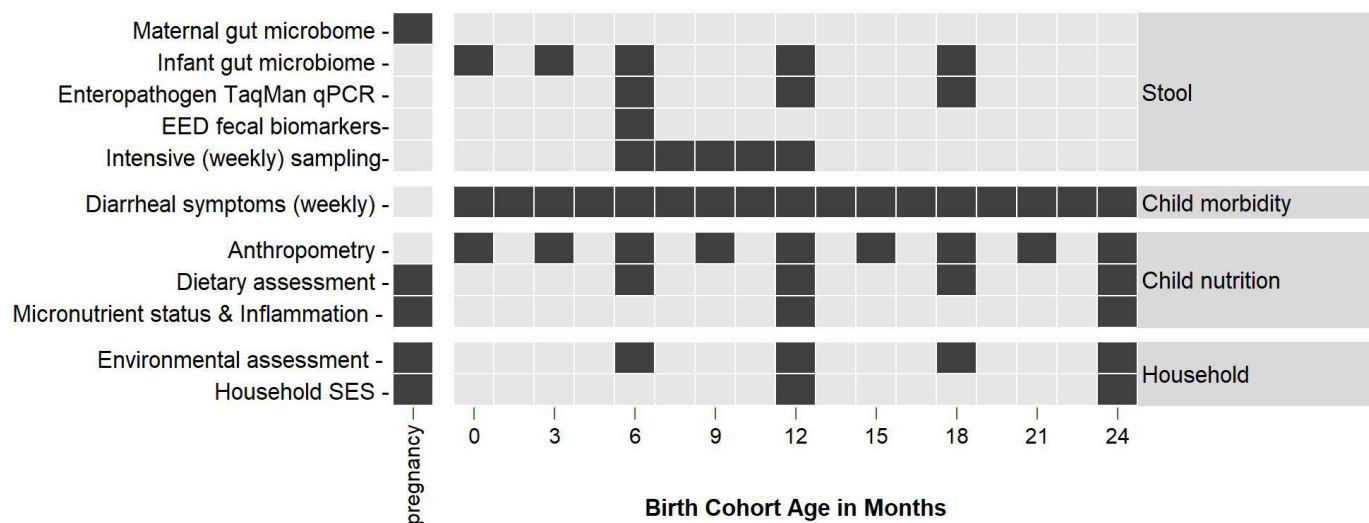
Following enrolment, questionnaires are administered to collect information about prenatal maternal diet, environmental contamination and demographic and socioeconomic factors such as maternal age, parity and household size. A stool sample is collected for determination of the maternal microbiome and a serum sample for determination of maternal anaemia and inflammation. Within 7 days of birth, information about delivery mode (vaginal vs caesarean), gestational age, birth weight and delivery location are recorded based on information recorded on the child's vaccine card, breastfeeding initiation is captured based on maternal report, a child stool sample is collected and child anthropometry is measured. Further stool and serum samples, nutrition data and environmental assessments are then collected at household visits detailed in figure 3 and described in greater detail below. Depending on the analysis in question, these variables may be included in models as exposures of interest, outcomes of interests or as covariates.



**Figure 4** Environmental context of the ECoMiD study. Study data are organised according to a socio-ecological framework. ECoMiD, Enteropatógenos, Crecimiento, Microbioma, y Diarrea.

### Environmental exposures

To understand how environmental conditions affect the developing child gut microbiome and enteric pathogen burden, we comprehensively evaluate environmental conditions using multiple approaches, across different scales (figure 4). At the broadest scale, households are selected across a rural–urban gradient to capture variation associated with urbanisation. We characterise community sanitation coverage using data from prior and ongoing



**Figure 3** Overall participant timeline. Primary study activities are programmed according to the age of the child. Squares indicate a single sampling point, except for intensive samples, which are collected weekly from 6 to 12 months of age (a total of 24 samples), and diarrhoeal symptoms, which are collected weekly for the entire period of the study. EED, environmental enteric dysfunction; SES, socioeconomic status.

**Table 2** Quantitative and categorical results from water and hand-rinse sampling

Colilert result	Petrifilm result	Concentration of <i>E. coli</i> per 100 mL	Category
Absent	0	<1 CFU or MPN/100 mL	Not detected
Present	0	1–99 MPN/100 mL	Low
Present	1–9 CFU/1 mL	100–999 CFU/100 mL	Medium
Present	10–49 CFU/1 mL	1000–4999 CFU/100 mL	High
Present	≥50 CFU/1 mL	≥5000 CFU/100 mL	Very high

We use a combination of 1 mL Petrifilm quantitative tests and 100 mL Colilert presence/absence tests to estimate colony forming units or most probable number. This approach was chosen to optimise the information obtained while staying feasible and within fieldworker capabilities. Combining the 2 tests provides a categorical measure of household contamination. Note that we assume that any sample positive for Petrifilm will also be positive for Colilert, as the sensitivity is much higher for Colilert (Colilert has a lower detection limit of 1 MPN/100 mL, whereas Petrifilms have a lower detection limit of 100 CFU/100 mL). If we detect growth on Petrifilm in the absence of a positive result on the Colilert, we will rely on the Petrifilm result. CFU, colony forming unit; MPN, most probable number.

studies in the region.<sup>35 44</sup> At the household level, we use a combination of surveys, environmental sampling, structured observations and in-depth interviews. These activities are conducted every 6 months (figure 3). Surveys include questions about household water sources, animal ownership and child-related hygiene, and observations about water storage, sanitation and washing and cooking areas. Structured observations are conducted on a subset of children to characterise opportunities for contact between the child and their environment. Environmental sampling consists of rinses of the mother's and children's hands in 101 mL of sterile water and a 101 mL sample of the mother (at the prenatal visit) or children's drinking water. We assay samples for faecal indicator bacteria using both PetriFilms (3M, St. Paul, Minnesota) and the Colilert presence-absence test (IDEXX laboratories, Westbrook, Maine). This approach was chosen to optimise the information obtained while remaining feasible in a remote field setting, limiting sample processing time and using study workers without prior formal environmental microbiology training. The tests are described in table 2. Because of the high variability in water quality results over time,<sup>33</sup> samples are repeated 3 times from the same household within 1 week.

### Stool collection

Maternal stool is collected once, at enrolment. Child stool samples are collected at 1 week, and at 3, 6, 12 and 18 months of age (figure 3). Mothers are given a small container to store stool and asked to store this in a cooler provided by the study until the study team member arrives to collect the sample. In addition, from 6 to 12 months of age, we collect weekly ('intensive') stool samples. After samples are retrieved by the study team member, multiple aliquots of stool are immediately stored at  $-196^{\circ}\text{C}$  in portable liquid nitrogen tanks, which are transported to the Universidad San Francisco de Quito monthly for long-term storage. We collect additional aliquots in Zymo DNA/RNA Shield (Zymo Research, Irving, California). The latter is intended as a backup in case of cold chain failure. The quality of DNA extracted from both frozen

samples and samples preserved in Zymo DNA/RNA shield was previously confirmed during piloting of the laboratory methods (data not shown). After all intensive stool samples have been collected, we will conduct a subanalysis using the banked samples. We will select stool samples from 200 children who experienced an episode of diarrhoea at the time of sample collection, where an episode is defined as an onset of diarrhoea preceded by 7 days of no reported diarrhoea. These samples will be randomly selected among geographic strata to ensure equal representation of children across the rural–urban gradient of the study. A matched stool sample from a child without diarrhoea will be selected based on the criterion that the individual has no reported symptoms at the time of stool collection and for at least 7 days prior and is matched by age ( $\pm 1$  month of the symptomatic child) and by infection with the same pathogen. We will test these samples for enteric infections using a TaqMan array card (TAC; ThermoFisher). Pathogens will be linked to the diarrheal disease episode based on relative cycle threshold values from the TaqMan results.<sup>45</sup> Children may contribute both diarrheal and asymptomatic stool samples to the study at different ages. This repeated sample design will be accounted for in the analysis.

For each selected symptomatic and paired asymptomatic stool sample, we will also assay 3 further stool samples for 16s ribosomal RNA (rRNA) gene amplicon sequencing and enteric infections, using a TaqMan array card (TAC; ThermoFisher) further described below: one collected the week before the diarrheal episode, one during the week of the episode and one 2 or more weeks following the episode.

Beyond analysing the stool to match symptomatic and asymptomatic samples with respect to infection status, downstream analyses planned for the stool samples include: (1) *Bacterial microbiome assessment*: after DNA is extracted from frozen or preserved stool samples, the 16s rRNA gene will be polymerase chain reaction (PCR) amplified and sequenced. In a subset of samples, we will also carry out shotgun metagenome sequencing; (2) *Pathogen*

**assessment:** extracted DNA and RNA will be analysed for multiple viral, bacterial and parasitic enteric pathogens using a TaqMan array card (TAC; ThermoFisher), a singleplex molecular method that consists of nucleic acid amplification of multiple microbial targets.<sup>45–47</sup> The specific pathogens we will test for are enteroaggregative *E. coli*, diarrheagenic *E. coli*, Shiga toxin-producing *E. coli*, enteropathogenic *E. coli*, enterotoxigenic *E. coli*, *Shigella*, enteroinvasive *E. coli*, *Campylobacter jejuni*, *Campylobacter coli*, *Salmonella typhi*, Adenovirus, Astrovirus, Enterovirus, Norovirus GI, Norovirus GII, Rotavirus, Sapovirus, SARS-CoV-2, *Cryptosporidium hominis*, *Cryptosporidium parvum*, *Cyclospora cayentanensis*, *Giardia duodenalis*, *Entamoeba histolytica*, *Ascaris lumbricoides* and *Trichuris trichiura*. TaqMan array card assays will be tested for linearity and matrix inhibition using positive control plasmids<sup>48</sup> as well as limits of detection, repeatability, reproducibility and analytical accuracy using reference strains or genomic DNA/RNA.<sup>45</sup> Phocine herpesvirus and MS2 bacteriophage will be used as extrinsic controls.<sup>45</sup> (3) **EED assessment:** EED will be assessed by analysing the stool for 4 faecal biomarkers of intestinal inflammation and permeability (myeloperoxidase, alpha-1-antitrypsin, neopterin and calprotectin) using commercially available enzyme-linked immunoassay (ELISA) kits.<sup>49 50</sup>

### Serum collection

We collect venous blood samples from mothers at one time point (prenatal) and from children at 2 time points (12 and 24 months). Anaemia is analysed in real time via a haemoglobin test using the HemoCue Hb 201+System. Inflammation is measured by C reactive protein (CRP) and  $\alpha$ -1-acid glycoprotein (AGP), both using a sandwich ELISA. We will define inflammation as CRP > 5 mg/L; and AGP > 1 mg/L.<sup>51–53</sup>

### Diarrhoea incidence

Variables associated with the child's diarrhoea are collected through weekly surveys, administered either through household visits or telephone calls. These include (1) caregiver-reported symptomatic diarrheal disease (defined as 3 or more loose stools in a 24-hour period<sup>54</sup>), (2) caregiver's perception of diarrhoea using locally relevant terms,<sup>55</sup> (3) features of diarrheal severity,<sup>56 57</sup> such as caregiver-reported dysentery, vomiting and fever and (4) antibiotic usage.

### Anthropometry

Field staff measure weight, length and head circumference within 1 week of birth, and quarterly thereafter, using infant scales and recumbent length measuring boards. Measurements are made within 2 weeks (14 days) of the targeted day, or up to 6 weeks after the target day if the family is travelling during the desired window. Length, height and weight will be converted to Z-scores based on World Health Organization reference standards.<sup>58</sup> For quality control, measurements are repeated if length-for-age or head-circumference-for-age Z-scores

are less than  $-6$  or greater than  $6$ , or if weight-for-age or weight-for-length are less than  $-5$  or greater than  $5$ .<sup>58 59</sup> Stunting is defined as 2 SDs below median height for age of the reference population.<sup>58</sup>

### Dietary assessments

Dietary data are collected using 3 questionnaires. First, from 0 to 24 months of age, weekly surveillance visits include 3 questions to characterise the presence or absence of breastfeeding, the intake of non-breast milk liquids and intake of semi solid and solid foods in the past week. Second, on a quarterly basis, a modified dietary diversity questionnaire is administered to characterise (1) dietary diversity, (2) indicators of complementary feeding such as feeding frequency and minimal acceptable diet and (3) the child's usage of common micronutrient supplements (vitamin A, iron or zinc).<sup>60</sup> This questionnaire queries about the presence or absence of food groups consumed in the past 24 hours, with some additional questions that capture common regional complementary foods, such as baby porridges frequently made with and without milk. Third, 24-hour dietary recalls are conducted with the child's primary caregiver at 6, 12 and 18 months of age to assess the child's intake of non-breastmilk macronutrients and micronutrients and to allow for the calculation of nutrient adequacy ratios and summary measures of overall dietary quality.<sup>61 62</sup> At each time point, this consists of 3 recalls conducted on non-consecutive days over a 1-week period. We will calculate macronutrient and micronutrient intakes from complementary foods using a standard reference food composition database for Ecuador.<sup>63</sup> This database has been expanded to include local dishes and recipes specific to the study population. Macro and micronutrient intake data will be used to create covariates for the analyses described below. We will use data reduction techniques to summarise intake data. These techniques may include a priori methods to assess specific characteristics of the diet, such as its inflammatory potential,<sup>64</sup> as well as a posteriori data-driven approaches such as principal component analysis and reduced rank regression.<sup>65</sup>

### Household socioeconomic status

Socioeconomic status (SES) is measured using indices based on household building materials and assets, which have been standardised across prior studies by our group.<sup>39</sup>

### Creation of a sample bank

With the written informed consent of the participant, any stool or plasma samples left over following planned study assays are stored in anticipation of future advancement and standardisation as well as potential follow-on studies. All samples are labelled and stored separately from household or participant identifiers (eg, name, household location) and linkable only by a protected database.

## Data management

Data are captured electronically on Android tablets using Open Data Kit open-source software (opendatakit.org). Tablets are password protected and securely stored when not in use. Identifiable data linking participant and sample ID codes to names and addresses are collected at enrolment and stored in a secure folder, accessible only to the study team. Otherwise, all data are collected and stored without personal identifiers.

Survey forms include built-in checks to reduce missing responses and flag logical inconsistencies. Study data are downloaded monthly, and quality control checks are run to compare recently completed study activities against scheduled activities and to check for missing or implausible values, duplicated records or inconsistencies. Discrepancies are reconciled through communication between the data management team and field supervisors.

## Study monitoring

This observational study does not test an intervention. Therefore, no external data monitoring committee exists.

## COVID-19 pandemic

In March 2020, study enrolment was paused due to the coronavirus pandemic, and activities with already enrolled subjects were modified. All questionnaires that could be conducted by telephone were continued, but biological and environmental sample collection activities, structured observations and anthropometry were paused. As a result, some data continued to be available during this period, but other data are missing. Reinitiation of study activities occurred in a phased manner, beginning in early 2021, as it was considered safe to contact study participants. This started with limited stool sample collection that minimised contact between fieldworkers and study participants. All activities have followed protocols set forth by the Ecuadorian Ministry of Health during the health emergency, declared by ministerial agreement 00126–2020.

## Outcome measures and statistical analysis

### Specific aim 1

Examine the association between household environmental conditions and (a) commensal gut microbiome composition and development and (b) enteric pathogen burden.

We hypothesise that household environmental conditions (water, sanitation and hygiene conditions) are drivers of (a) the composition and development of the child gut microbiome and of (b) the total burden of enteric infections within individual subjects as well as the pathogen profile among those with shared environmental characteristics.

To address SA1a, we will look at the effects of household environmental conditions on gut microbiome community structure at 1 week and 3, 6, 12 and 18 months of age. The outcomes of interest will be measures of alpha (within-sample) diversity of the gut microbiome obtained using

16s rRNA gene amplicon sequencing data. In addition, we will use 16s data to evaluate the influence of environmental conditions using beta (between-sample) diversity measures and will use non-parametric ordinations and multivariate statistical models to examine differences explained by key environmental variables and covariates. Statistical analyses that include effect size corrections will be used to identify specific taxa that differ between individuals with varying environmental exposures. These analyses will be completed using the QIIME2 software platform and other bioinformatics tools.<sup>66</sup>

The outcome of SA1b is the presence or absence of pathogens in children 6–18 months of age. We will examine the effect of household environmental conditions on the burden of enteric infections by developing mixed effect logistic regression models, where the outcome will be total pathogen burden and/or burden of specific pathogens. In addition, we will compare the community similarity of enteric pathogens shed by subjects using similarity matrices and multivariate models. In this case, instead of using the community composition of the gut microbiome as in SA1a, in SA1b, the outcome will be the set of enteropathogens detected in stool samples, partitioning the variance in dissimilarity between the main effects (environmental conditions) and covariates of interest such as child age, sex, delivery mode and nutrition.

### Specific aim 2

Examine whether gut microbiome composition and diversity modifies the acute and chronic effects of enteric infections in children.

We hypothesise that diversity and composition of the child gut microbiome modify the impact of specific enteric infections on child health outcomes, including: (a) diarrhoea incidence, (b) EED and (c) growth faltering.

Using participants from all 3 sites, we will evaluate the effect of cumulative burden of enteric infection (pathogens detected at 6, 12 and 18 months of age) on 3 primary outcome measures at 24 months. The primary goal of this analysis is to evaluate whether gut microbiome alpha diversity measures (described in SA1) modify the enteric infection-health relationship of interest. We will examine whether overall diversity and specific bacterial taxa of interest are effect modifiers. We will also use stool samples collected at 1 week of age to assess whether the initial colonisation of the microbiome influences assembly and subsequent composition at later ages (founder effects). We will develop regression models with random effects for site and individual. We will adjust for other covariates related to each outcome, such as breastfeeding, reported antibiotic usage across 24 months and SES.

### Specific aim 3

Examine how the gut microbiome (a) responds to and (b) recovers from enteric infections in children.

We hypothesise that (a) enteric infections elicit specific changes in the child gut microbiome and that these signature responses will be distinct for symptomatic versus

asymptomatic infections and (b) children with a more diverse gut microbiome prior to infection will recover more quickly from the short-term disturbances of enteric infections than those with a less diverse microbiome.

We will use data collected from children with acute diarrhoea during our intensive stool collection activity, selecting samples that test positive for active enteric infections. The outcomes of interest for SA3a are the (1) alpha diversity, (2) beta diversity and (3) differentially abundant microbiome taxa. Within a given child, we will compare 16s bacterial microbiome communities the week prior to infection with the week during infection. We additionally will compare 16s microbiome communities the week during infection with 2 weeks or more post infection to characterise recovery. We will examine the impact of community, controlling for other covariates,

on microbiome community similarity using ADONIS permutation models, based on Unifrac distances and visualise differences using Non-metric Multi-dimensional Scaling (NMDS) plots. For specific enteric infections, we will evaluate whether any observed changes to gut microbiome composition or abundance are modified by diarrhoea case status. We will also examine similarity measures of the gut microbial communities in the week prior to infection, compared with the week during infection.

To address SA3b, we will carry out similar analyses to those described for SA3a, in this case comparing samples from the week prior to infection to the sample collected 2 weeks or more after infection as an indicator of community recovery from the disturbance caused by the enteric infection.

**Table 3** Sample size and power calculation estimates for study aims

Aim	Exposure used for power calculation	Exposure prevalence	Outcome	Required sample size	Available sample size
1a	Household water access (as an example environmental exposure)	25%–50%*	Difference of 0.25† or more in class-level Chao1 species richness	1268–1800	1800
1b	Household water access (as an example environmental exposure)	5%–50%*	Excess risk of 12% or greater for any specific enteric infection‡	199–880	1080
2a	Any specific enteric infection	10%–50%‡	Excess risk of 8.2% or greater in symptomatic diarrhoea§	306–737	1080
2b	Number of enteric infections from 0 to 24 months	(1) 10%–25%¶ (2) 25%–50%¶	EED score** difference of (1) 0.7 or more; (2) 0.5 or more††	(1) 80–299 (2) 160–308	360
2c	Number of enteric infections from 0 to 24 months	10%–50%¶	Excess risk of 8.1% or greater in growth faltering‡‡ (eg, stunting)	306–737	1080
3a-b	Diarrheal status	50%§§	Difference of 0.40¶¶ or more in class-level Chao1 species richness	388	400 <sup>1</sup>
3a-b	(1) Enteric infection-specific prevalence of 25% or more‡ (2) Any enteric infection***	50%	(1) Difference of 0.80¶¶ or more in class-level Chao1 species richness (2) Difference of 0.45¶¶ or more in class-level Chao1 species richness	(1) 98 (2) 308	(1)~100††† (2)~308†††

\*Prevalence of household piped water access is 47% across the 3 study sites.

†The mean differences in Chao1 species richness across the 3 study sites (most rural to most urban) is 1.1.

‡Prevalence of *E. coli* across the 3 study sites in the EcoZUR study was 24.5%.

§Based on symptomatic diarrhoea prevalence of 10% or more in the cohort—data from the “Ecología, Desarrollo, Salud, y Sociedad” (EcoDeSS) study suggests diarrhoea prevalence of ~12%.<sup>41</sup>

¶“The Etiology, Risk Factors, and Interactions of Enteric Infections and Malnutrition and the Consequences for Child Health and Development Project” (MAL-ED) prevalence of 29% of non-diarrheal stool samples containing 2 or more pathogens.<sup>79</sup>

\*\*Based on the score developed by Kosek *et al.*<sup>49</sup>

††For comparison, George *et al* reported difference on the magnitude of 0.70–1.00 between children with and without specific environmental exposures.

‡‡Based on growth faltering prevalence of 10% or more in the cohort—previously characterised prevalence of stunting of 10%–14% in children <5 in “Ecología, Desarrollo, Salud, y Sociedad” (EcoDeSS) study data.<sup>38</sup>

§§Case-control status selected based on diarrhoea status using banked samples.

¶¶For comparison, in the “E. coli en Zonas Urbanas y Rurales” (EcoZUR) study data, differences between adults with vs without acute diarrhoea had Chao1 differences of approximately 0.5; <sup>1</sup> based on 1:1 case-control design (ie, 200 cases and 200 controls).

\*\*\*The Etiology, Risk Factors, and Interactions of Enteric Infections and Malnutrition and the Consequences for Child Health and Development Project” (MAL-ED) prevalence of ~77% of diarrheal stool samples containing one or more pathogens.<sup>79</sup>

†††Available sample size is dependent on specific pathogen prevalence in the 400 samples tested—we have based our estimate on (i) 24.5% prevalence of pathogenic *E. coli*=49 children infected, for a total sample size of 98 and (ii) 77% prevalence of any pathogen=154 children infected, for a total of 308.

‡‡‡

EED, environmental enteric dysfunction.



## Shotgun metagenomics

To further address SA3, we will carry out shotgun metagenomic sequencing on a subset of at least 225 samples from intensive surveillance to characterise the predominant microbial species and strains present during infection and changes in the presence or absence and relative abundance of functional genes present within the microbial community before, during and following enteropathogen infections. We will compare symptomatic versus asymptomatic individuals with the same type of infection.

We will select samples for metagenome sequencing based on the enteropathogen, or set of enteropathogens, for which we observe the greatest difference in taxonomic composition using 16s sequencing data. Relative to 16s sequencing, shotgun metagenomics will provide improved resolution to study pathogens and pathogen genotypes, measure relative in situ abundance of pathogen populations,<sup>67</sup> and investigate changes in functional and virulence gene abundances in the gut microbiome in response to pathogen presence.<sup>68</sup> The presence or absence and relative abundance of functional genes contributed by microorganisms in the gut before, during and after an enteropathogen infection will be used to test whether the baseline gut microbiome profile is an effect modifier on the resulting outcome of diarrhoea.

Shotgun metagenomes will be sequenced using Illumina sequencing chemistry as previously described.<sup>69</sup> Raw FASTQ reads from the sequencing runs will be quality checked, trimmed, assembled and annotated using tools implemented in the MiGA (Microbial Genomes Atlas) pipeline,<sup>70</sup> which was developed for efficient processing and management of microbial metagenomes. We will use read-based mapping to quantify annotated metagenome features and will normalise abundances based on the sequencing depth of an external spike-in control.<sup>71</sup> Additional open-source or internal software tools will be used to remove human read contamination, annotate gene functions and taxonomy and bin and quality check metagenome-assembled genomes, as we have done with other analyses.<sup>69</sup>

## Sample size

We plan to enrol up to 480 pregnant women (160 per site), with the goal of retaining 360 total mother–child dyads (120 per site) through 24 months of age. Our current enrolment rate is ~18 dyads per month. Sample size calculations were completed to test the hypotheses proposed in aims 1–3 with a  $\beta$  of 80%, an  $\alpha$  of 5% and an intraclass correlation coefficient of 0.1. **Table 3** details the assumptions used to estimate the required sample size for each aim. We used ranges of exposure prevalence in the calculations based on regional variability of the various exposures. For gut microbiome power calculations, we used class-level Chao1 species richness ranges from our previous work in the region.<sup>36</sup> Other microbiome power analyses using more advanced methods (eg, taxonomic based methods<sup>72</sup> or pairwise distances using PERMANOVA<sup>73</sup>) require additional details beyond the

scope of our current data, for example, distances between subjects within each group.<sup>74</sup> However, our sample size is on par with or greater than that of other studies reporting significant differences across gut microbiome samples in LMIC settings.<sup>10–12 17 21 75</sup> For example, household water access and the microbiome will each be measured at least 5 times per child, resulting in an estimated 1800 paired measurements. We similarly estimate that, assuming an intraclass correlation coefficient of 1.1 to account for within-child correlation between measurements, 1268–1800 measurements would be needed to detect differences in the Chao1 richness related to water access if the overall prevalence of access to improved water in the study communities is between 25% and 50%.

## Patient and public involvement

Patients and/or the public were not involved in the design, or conduct, or reporting or dissemination plans of this research.

## Ethics and dissemination

The study protocol has been approved by the institutional review boards of Emory University (IRB00101202) and the Universidad San Francisco de Quito (2018–022M). The University of Michigan and University of Washington formally deferred oversight of the study to Emory University. The study protocol was also reviewed and approved by the Ministry of Health of Ecuador (MSPCURI000253-4).

The results of the study will be published in peer-reviewed journals and presented at international conferences as well as to health officials in Ecuador.

## Discussion

Population-based studies of gut microbiome composition have been largely descriptive, identifying differences between communities living in different regions,<sup>17–22</sup> and most studies of host-pathogen dynamics have been carried out in the laboratory or in animal models.<sup>76 77</sup> To understand the functional role that the commensal microbiome plays in response to enteric infection, a combination of TAC, 16s rRNA gene amplicon sequencing and novel metagenomic bioinformatics tools will be used to characterise differential compositional and functional responses to enteric pathogen infections. This will allow the impacts of pathogen exposure to be better differentiated from factors related to SES and diet, thereby providing new information to understand and reduce the negative consequences of enteric infections for child health.

## Author affiliations

<sup>1</sup>Department of Epidemiology, University of Michigan School of Public Health, Ann Arbor, Michigan, USA

<sup>2</sup>Instituto de Microbiología, Universidad San Francisco de Quito, Quito, Pichincha, Ecuador

<sup>3</sup>Facultad de Ciencias Médicas, Universidad Central del Ecuador, Quito, Ecuador

<sup>4</sup>Department of Environmental Health, Emory University Rollins School of Public Health, Atlanta, Georgia, USA

<sup>5</sup>Department of Environmental and Occupational Health Sciences, University of Washington School of Public Health, Seattle, Washington, USA

<sup>6</sup>Carrera de Nutrición y Dietética, Universidad San Francisco de Quito, Quito, Pichincha, Ecuador

<sup>7</sup>Carrera de Medicina, Universidad de Las Americas Facultad de Ciencias de la Salud, Quito, Pichincha, Ecuador

<sup>8</sup>School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, Georgia, USA

<sup>9</sup>Instituto de Biomedicina, Facultad de Ciencias Médicas, Universidad Central del Ecuador, Quito, Ecuador

**Contributors** GOL contributed to the development of the protocols for fieldwork, data management and analysis and drafted the manuscript. JNSE contributed to the conception and design of the work and critically revised the manuscript. JU contributed to the development of the protocols for fieldwork and data management and critically revised the manuscript. AVE contributed to the development of the protocols for fieldwork and data management and critically revised the manuscript. ER contributed to the development of the protocols for nutritional data collection and dietary data management and critically revised the manuscript. MC contributed to the development of the protocols related to parasitology and critically revised the manuscript. NC contributed to the development of the protocols for nutritional data collection and dietary data management and critically revised the manuscript. GV contributed to the development of the protocols for fieldwork and laboratory analyses and critically revised the manuscript. WC contributed to the development of the protocols for fieldwork and critically revised the manuscript. SMS contributed to the development of the protocols for laboratory analyses and sample size calculations, and critically revised the manuscript. CV, GT, RM, KJJ, KTK: contributed to the development of the protocols for laboratory analyses and critically revised the manuscript. KL contributed to the conception and design of the work, drafted portions of the manuscript text, and critically revised the manuscript.

**Funding** This work is supported by the National Institutes of Health (R01AI137679). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

**Map disclaimer** The inclusion of any map (including the depiction of any boundaries therein), or of any geographic or locational reference, does not imply the expression of any opinion whatsoever on the part of BMJ concerning the legal status of any country, territory, jurisdiction or area or of its authorities. Any such expression remains solely that of the relevant source and is not endorsed by BMJ. Maps are provided without any warranty of any kind, either express or implied.

**Competing interests** None declared.

**Patient consent for publication** Not applicable.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

#### ORCID iDs

Gwenyth O Lee <http://orcid.org/0000-0002-7889-3852>

Gabriela Vasco <http://orcid.org/0000-0002-8318-2465>

Karen Levy <http://orcid.org/0000-0002-0968-9401>

#### REFERENCES

- Praharaj I, Revathy R, Bandyopadhyay R, *et al*. Enteropathogens and gut inflammation in asymptomatic infants and children in different environments in southern India. *Am J Trop Med Hyg* 2018;98:576–80.
- Hellard ME, Sinclair MI, Hogg GG, *et al*. Prevalence of enteric pathogens among community based asymptomatic individuals. *J Gastroenterol Hepatol* 2000;15:290–3.
- McCormick BJJ, Lee GO, Seidman JC, *et al*. Dynamics and trends in fecal biomarkers of gut function in children from 1-24 months in the MAL-ED study. *Am J Trop Med Hyg* 2017;96:465–72.
- MAL-ED Network Investigators. Relationship between growth and illness, enteropathogens and dietary intakes in the first 2 years of life: findings from the MAL-ED birth cohort study. *BMJ Glob Health* 2017;2:e000370.
- MAL-ED Network Investigators. Childhood stunting in relation to the pre- and postnatal environment during the first 2 years of life: the MAL-ED longitudinal birth cohort study. *PLoS Med* 2017;14:1–21.
- MAL-ED Network Investigators. Early childhood cognitive development is affected by interactions among illness, diet, enteropathogens and the home environment: findings from the MAL-ED birth cohort study. *BMJ Glob Health* 2018;3:e000752.
- Kamada N, Chen GY, Inohara N, *et al*. Control of pathogens and pathobionts by the gut microbiota. *Nat Immunol* 2013;14:685–90.
- Robertson RC, Manges AR, Finlay BB, *et al*. The Human Microbiome and Child Growth - First 1000 Days and Beyond. *Trends Microbiol* 2019;27:131–47.
- Rouhani S, Griffin N, Yori P. Diarrhea as a potential cause and consequence of reduced gut microbial diversity among undernourished children in Peru. *J Chem Inf Model* 2019;53:1689–99.
- Lindsay B, Oundo J, Hossain MA, *et al*. Microbiota that affect risk for shigellosis in children in low-income countries. *Emerg Infect Dis* 2015;21:242–50.
- Pop M, Walker AW, Paulson J, *et al*. Diarrhea in young children from low-income countries leads to large-scale alterations in intestinal microbiota composition. *Genome Biol* 2014;15:R76.
- David LA, Weil A, Ryan ET, *et al*. Gut microbial succession follows acute secretory diarrhea in humans. *mBio* 2015;6:e00381–15.
- Subramanian S, Huq S, Yatsunenko T, *et al*. Persistent gut microbiota immaturity in malnourished Bangladeshi children. *Nature* 2014;510:417–21.
- Blanton LV, Charbonneau MR, Salih T, *et al*. Gut bacteria that prevent growth impairments transmitted by microbiota from malnourished children. *Science* 2016;351:aad331.
- Chen RY, Kung VL, Das S, *et al*. Duodenal microbiota in stunted undernourished children with enteropathy. *N Engl J Med* 2020;383:321–33.
- Ducarmon QR, Zwittink RD, Hornung BVH, *et al*. Gut microbiota and colonization resistance against bacterial enteric infection. *Microbiol Mol Biol Rev* 2019;83:1–29.
- Yatsunenko T, Rey FE, Manary MJ, *et al*. Human gut microbiome viewed across age and geography. *Nature* 2012;486:222–7.
- Schnorr SL. The diverse microbiome of the hunter-gatherer. *Nature* 2015;518:S14–15.
- Schnorr SL, Candela M, Rampelli S, *et al*. Gut microbiome of the Hadza hunter-gatherers. *Nat Commun* 2014;5:3654.
- Clemente JC, Pehrsson EC, Blaser MJ, *et al*. The microbiome of uncontacted Amerindians. *Sci Adv* 2015;1:e1500183.
- Martinez I, Stegen JC, Maldonado-Gómez MX, *et al*. The gut microbiota of rural Papua New Guineans: composition, diversity patterns, and ecological processes. *Cell Rep* 2015;11:527–38.
- Obregon-Tito AJ, Tito RY, Metcalf J, *et al*. Subsistence strategies in traditional societies distinguish gut microbiomes. *Nat Commun* 2015;6:6505.
- Maher SE, O'Brien EC, Moore RL, *et al*. The association between the maternal diet and the maternal and infant gut microbiome: a systematic review. *Br J Nutr* 2020;1–29.
- Wu GD, Chen J, Hoffmann C, *et al*. Linking long-term dietary patterns with gut microbial enterotypes. *Science* 2011;334:105–8.
- Albenberg LG, Wu GD. Diet and the intestinal microbiome: associations, functions, and implications for health and disease. *Gastroenterology* 2014;146:1564–72.
- Nermes M, Endo A, Aarnio J, *et al*. Furry pets modulate gut microbiota composition in infants at risk for allergic disease. *J Allergy Clin Immunol* 2015;136:1688–90.
- Lu K, Abo RP, Schlieper KA, *et al*. Arsenic exposure perturbs the gut microbiome and its metabolic profile in mice: an integrated metagenomics and metabolomics analysis. *Environ Health Perspect* 2014;122:284–91.
- Rampelli S, Schnorr SL, Consolandi C, *et al*. Metagenome sequencing of the Hadza hunter-gatherer gut microbiota. *Curr Biol* 2015;25:1682–93.
- Shaughnessy MK, Bobr A, Kuskowski MA, *et al*. Environmental contamination in households of patients with recurrent *Clostridium difficile* infection. *Appl Environ Microbiol* 2016;82:2686–92.
- Miller GE, Engen PA, Gillevet PM, *et al*. Lower neighborhood socioeconomic status associated with reduced diversity of the colonic microbiota in healthy adults. *PLoS One* 2016;11:e0148952.
- Kapourchali FR, Cresci GAM. Early-life gut microbiome—the importance of maternal and infant factors in its establishment. *Nutr Clin Pract* 2020;35:386–405.
- Dias MF, Reis MP, Acurcio LB, *et al*. Changes in mouse gut bacterial community in response to different types of drinking water. *Water Res* 2018;132:79–89.
- Levy K, Hubbard AE, Nelson KL, *et al*. Drivers of water quality variability in northern coastal Ecuador. *Environ Sci Technol* 2009;43:1788–97.

- 34 Instituto Nacional de Estadística y Censos. *Ecuador population and housing census 2010*, 2010.
- 35 Eisenberg JNS, Cevallos W, Ponce K, *et al*. Environmental change and infectious disease: how new roads affect the transmission of diarrheal pathogens in rural Ecuador. *Proc Natl Acad Sci U S A* 2006;103:19460–5.
- 36 Peña-Gonzalez A, Soto-Girón MJ, Smith S, *et al*. Metagenomic signatures of gut infection caused by different *Escherichia coli* pathotypes. *Appl Environ Microbiol* 2019;85:1–18.
- 37 Soto-Girón MJ, Peña-Gonzalez A, Hatt JK, *et al*. Gut microbiome changes with acute diarrheal disease in urban versus rural settings in northern Ecuador. *Am J Trop Med Hyg* 2021. doi:10.4269/ajtmh.20-0831. [Epub ahead of print: 19 Apr 2021].
- 38 Fuller JA, Villamor E, Cevallos W, *et al*. I get height with a little help from my friends: herd protection from sanitation on child growth in rural Ecuador. *Int J Epidemiol* 2016;45:460–9.
- 39 Lopez VK, Dombecki C, Trostle J, *et al*. Trends of child undernutrition in rural Ecuadorian communities with differential access to roads, 2004–2013. *Matern Child Nutr* 2018;14:e12588.
- 40 Lee GO, Gutierrez C, Castro Morillo N, *et al*. Multiple burdens of malnutrition and relative remoteness in rural Ecuadorian communities. *Public Health Nutr* 2021;24:4591–602.
- 41 Bhavnani D, Goldstick JE, Cevallos W, *et al*. Synergistic effects between rotavirus and coinfecting pathogens on diarrheal disease: evidence from a community-based study in northwestern Ecuador. *Am J Epidemiol* 2012;176:387–95.
- 42 Hansen CHF, Andersen LSF, Krych L, *et al*. Mode of delivery shapes gut colonization pattern and modulates regulatory immunity in mice. *J Immunol* 2014;193:1213–22.
- 43 Kane A, Dinh D, Ward H. Childhood malnutrition and the intestinal microbiome malnutrition and the microbiome. *Pediatr Res* 2013;77:1.
- 44 Peña-Gonzalez A V, Soto-Girón MJ, Hatt JK. Disruptions of the human gut microbiome during infectious diarrhea. In: *American society for tropical medicine and hygiene*. Atlanta, GA, 2016.
- 45 Liu J, Gratz J, Amour C, *et al*. A laboratory-developed TaqMan array card for simultaneous detection of 19 enteropathogens. *J Clin Microbiol* 2013;51:472–80.
- 46 Liu J, Platts-Mills JA, Juma J, *et al*. Use of quantitative molecular diagnostic methods to identify causes of diarrhoea in children: a reanalysis of the GEMS case-control study. *Lancet* 2016;388:1291–301.
- 47 Liu J, Gratz J, Amour C, *et al*. Optimization of quantitative PCR methods for enteropathogen detection. *PLoS One* 2016;11:e0158199.
- 48 Kodani M, Winchell JM. Engineered combined-positive-control template for real-time reverse transcription-PCR in multiple-pathogen-detection assays. *J Clin Microbiol* 2012;50:1057–60.
- 49 Kosek M, Haque R, Lima A, *et al*. Fecal markers of intestinal inflammation and permeability associated with the subsequent acquisition of linear growth deficits in infants. *Am J Trop Med Hyg* 2013;88:e2125.
- 50 Kosek M, Guerrant RL, Kang G, *et al*. Assessment of environmental enteropathy in the MAL-ED cohort study: theoretical and analytic framework. *Clin Infect Dis* 2014;59:S239–47.
- 51 Erhardt JG, Estes JE, Pfeiffer CM, *et al*. Combined measurement of ferritin, soluble transferrin receptor, retinol binding protein, and C-reactive protein by an inexpensive, sensitive, and simple sandwich enzyme-linked immunosorbent assay technique. *J Nutr* 2004;134:3127–32.
- 52 Burke RM, Rebolledo PA, Fabiszewski de Aceituno AM, *et al*. Early deterioration of iron status among a cohort of Bolivian infants. *Matern Child Nutr* 2017;13:e12404.
- 53 Burke RM, Suchdev PS, Rebolledo PA, *et al*. Predictors of inflammation in a cohort of Bolivian infants and toddlers. *Am J Trop Med Hyg* 2016;95:954–63.
- 54 Baqui AH, Black RE, Yunus M, *et al*. Methodological issues in diarrhoeal diseases epidemiology: definition of diarrhoeal episodes. *Int J Epidemiol* 1991;20:1057–63.
- 55 Trostle JA, Yépez-Montufar JA, Corozo-Angulo B, *et al*. [Diarrheal illnesses on the Ecuadorian coast: socio-environmental changes and health concepts]. *Cad Saude Publica* 2010;26:1334–44.
- 56 Lee G, Peñataro Yori P, Paredes Olortegui M, *et al*. An instrument for the assessment of diarrhoeal severity based on a longitudinal community-based study. *BMJ Open* 2014;4:e004816.
- 57 Lee GO, Richard SA, Kang G, *et al*. A comparison of diarrheal severity scores in the MAL-ED multisite community-based cohort study. *J Pediatr Gastroenterol Nutr* 2016;63:466–73.
- 58 World Health Organization. *WHO child growth standards: length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: methods and development*. Geneva, Switzerland, 2006.
- 59 WHO Multicentre Growth Reference Study Group. *WHO child growth standards: head circumference-for-age, arm circumference-for-age, triceps skinfold-for-age and subscapular skinfold-for-age: methods and development*. Geneva, 2007.
- 60 Caulfield LE, Bose A, Chandyo RK, *et al*. Infant feeding practices, dietary adequacy, and micronutrient status measures in the MAL-ED study. *Clin Infect Dis* 2014;59:S248–54.
- 61 Hatloy A, Torheim LE, Oshaug A. Food variety—a good indicator of nutritional adequacy of the diet? A case study from an urban area in Mali, West Africa. *Eur J Clin Nutr* 1998;52:891–8.
- 62 Steyn NP, Nel J, Labadarios D, *et al*. Which dietary diversity indicator is best to assess micronutrient adequacy in children 1 to 9 Y? *Nutrition* 2014;30:55–60.
- 63 Ramírez-Luzuriaga MJ, Silva-Jaramillo KM, Belmont P. Tabla de composición de alimentos para Ecuador: Compilación del Equipo técnico de la ENSANUT - ECU 2012. *Nutr J* 2014;8:19.
- 64 Zheng J, Hoffman KL, Chen J-S, *et al*. Dietary inflammatory potential in relation to the gut microbiome: results from a cross-sectional study. *Br J Nutr* 2020;124:931–42.
- 65 Zhao J, Li Z, Gao Q, *et al*. A review of statistical methods for dietary pattern analysis. *Nutr J* 2021;20:1–18.
- 66 Bolyen E, Rideout JR, Dillon MR, *et al*. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* 2019;37:852–7.
- 67 Konstantinidis KT, DeLong EF. Genomic patterns of recombination, clonal divergence and environment in marine microbial populations. *Isme J* 2008;2:1052–65.
- 68 Pickard JM, Zeng MY, Caruso R, *et al*. Gut microbiota: role in pathogen colonization, immune responses, and inflammatory disease. *Immunol Rev* 2017;279:70–89.
- 69 Patin NV, Peña-Gonzalez A, Hatt JK, *et al*. The role of the gut microbiome in resisting norovirus infection as revealed by a human challenge study. *mBio* 2021;12:e02634.
- 70 Rodríguez-R LM, Gunturu S, Harvey WT, *et al*. The microbial genomes atlas (MiGA) Webserver: taxonomic and gene diversity analysis of archaea and bacteria at the whole genome level. *Nucleic Acids Res* 2018;46:W282–8.
- 71 Crossette E, Gumm J, Langenfeld K. Metagenomic quantification of genes with internal standards. *MBio* 2021;12:1–12.
- 72 La Rosa PS, Brooks JP, Deych E, *et al*. Hypothesis testing and power calculations for taxonomic-based human microbiome data. *PLoS One* 2012;7:1–13.
- 73 Kelly BJ, Gross R, Bittinger K, *et al*. Power and sample-size estimation for microbiome studies using pairwise distances and PERMANOVA. *Bioinformatics* 2015;31:2461–8.
- 74 Hanson BM, Weinstock GM. The importance of the microbiome in epidemiologic research. *Ann Epidemiol* 2016;26:301–5.
- 75 Gorkiewicz G, Thallinger GG, Trajanoski S, *et al*. Alterations in the colonic microbiota in response to osmotic diarrhea. *PLoS One* 2013;8:e55817.
- 76 Wallace JG, Gohir W, Sloboda DM. The impact of early life gut colonization on metabolic and obesogenic outcomes: what have animal models shown us? *J Dev Orig Health Dis* 2016;7:15–24.
- 77 Blanton LV, Barratt MJ, Charbonneau MR, *et al*. Childhood undernutrition, the gut microbiota, and microbiota-directed therapeutics. *Science* 2016;352:1533.
- 78 Sánchez-Gómez A, Grijalva MJ, Silva-Aycaguer LC, *et al*. HIV and syphilis infection in pregnant women in Ecuador: prevalence and characteristics of antenatal care. *Sex Transm Infect* 2014;90:70–5.
- 79 Platts-Mills JA, Babji S, Bodhidatta L, *et al*. Pathogen-specific burdens of community diarrhoea in developing countries: a multisite birth cohort study (MAL-ED). *Lancet Glob Health* 2015;3:e564–e57.