Check for updates

STUDY PROTOCOL

REVISED Child health, nutrition and gut microbiota development during the first two years of life; study protocol of a prospective cohort study from the Khyber Pakhtunkhwa, Pakistan

[version 2; peer review: 2 approved]

Muhamamd Shahzad[1,2], Muhammad Ismail[1], Benjamin Misselwitz[3,4], Ahsan Saidal[5], Simon C Andrews[6], Khalid Iqbal[1,7], Hatice Akarsu[3,4,8], Ziad Al Nabhani[3,4]

[1]Institute of Basic Medical Sciences, Khyber Medical University, Peshawar, Pakistan
[2]Faculty of Dentistry, Zarqa University, Zarqa, Jordan
[3]Department of Visceral Surgery and Medicine, Bern University Hospital, Bern, Switzerland
[4]Maurice Müller Laboratories, Department for Biomedical Research, University of Bern, Bern, Switzerland
[5]Institute of Health Sciences, Khyber Medical University, Swat, Pakistan
[6]School of Biological Sciences, University of Reading, Reading, UK
[7]Epidemiological Methods and Etiological Research, Leibniz Institute for Prevention Research and Epidemiology, Bremen, Germany
[8]SIB Swiss Institute of Bioinformatics, Lausanne, Switzerland

v2 **First published:** 07 Nov 2024, **13**:1336
https://doi.org/10.12688/f1000research.158415.1
Latest published: 08 Aug 2025, **13**:1336
https://doi.org/10.12688/f1000research.158415.2

Abstract

Recent evidence suggests that the development of gut microbiota during infancy affects several metabolic, immune, and endocrine pathways in humans. An imbalance in gut microbiota diversity or function, also known as dysbiosis, not only affects early child growth and development, but is also linked with the development of chronic, non-communicable diseases in later life. The **C**hild **H**ealth **A**nd **M**icrobiome Development Study – **P**akistan (CHAMP) study aimed to longitudinally assess gut microbiota development and associated factors (maternal, child, and demographic) during early childhood in populations residing in malnutrition-endemic communities in Pakistan. A prospective cohort of mother-infant pairs (n=70) will be recruited from District Swat, Pakistan, and followed for two years. Complete information about demographic characteristics, anti-natal and post-natal care, dietary intake, feeding practices, and child health will be collected at baseline and 3, 6, 12, 18, and 24 months. Anthropometric measurements (height, weight, mid-upper arm

Open Peer Review

Approval Status ✓ ✓

	1	2
version 2 (revision) 08 Aug 2025	✓ view	✓ view
version 1 07 Nov 2024	? view	

1. **Ryan Marsh**, Northumbria University, Newcastle,, UK

2. **Roshonda B Jones**, North Carolina Agricultural and Technical State University, Greensboro, USA

Any reports and responses or comments on the article can be found at the end of the article.

circumference, and head circumference), dry blood spots, and fecal samples were also collected. Ethical approval for the study was obtained from Khyber Medical University, Pakistan. The study is also registered on clincaltrial.gov (Ref no: NCT05793294). The study findings will help researchers understand gut microbiota development, associated factors, and their impact on longitudinal growth in infants during the first two years of life.

Keywords
Microbiome, Infants, Growth, Vulnerable, Malnutrition

Corresponding authors: Muhamamd Shahzad (shahzad.ibms@kmu.edu.pk), Ziad Al Nabhani (ziad.alnabhani@unibe.ch)

Author roles: Shahzad M: Conceptualization, Funding Acquisition, Methodology, Project Administration, Writing – Original Draft Preparation, Writing – Review & Editing; **Ismail M**: Data Curation, Investigation, Methodology, Writing – Original Draft Preparation; **Misselwitz B**: Formal Analysis, Investigation, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; **Saidal A**: Data Curation, Investigation, Methodology, Project Administration, Writing – Original Draft Preparation; **Andrews SC**: Conceptualization, Funding Acquisition, Investigation, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; **Iqbal K**: Formal Analysis, Investigation, Methodology, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing; **Akarsu H**: Formal Analysis, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; **Al Nabhani Z**: Conceptualization, Funding Acquisition, Methodology, Project Administration, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: The study is funded by National Institute of Health PakistnNHCG, Pakistan via grant number NHCG-22/R2-74/RDC/KMU/214. Z.A.N is funded by Swiss National Science Foundation (IZSEZ0_229993; 310030_215675; 31ND30_213452; CRSII--222781), and European Research Council Starting Grant (WePredict Project Number: 949613).
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2025 Shahzad M *et al*. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Shahzad M, Ismail M, Misselwitz B *et al.* **Child health, nutrition and gut microbiota development during the first two years of life; study protocol of a prospective cohort study from the Khyber Pakhtunkhwa, Pakistan [version 2; peer review: 2 approved]** F1000Research 2025, **13**:1336 https://doi.org/10.12688/f1000research.158415.2

First published: 07 Nov 2024, **13**:1336 https://doi.org/10.12688/f1000research.158415.1

REVISED **Amendments from Version 1**

In the revised manuscript, substantial changes have been made to enhance clarity, scientific rigor, and analytical depth. Notably, the methodology section has undergone significant revision. The sequencing approach has been upgraded from 16S rRNA sequencing to shotgun metagenomic sequencing, a more comprehensive and reliable method that enables higher-resolution taxonomic and functional profiling of microbial communities. Additionally, the data analysis section has been expanded and refined to provide a more detailed and transparent account of the bioinformatics pipelines and statistical methods employed. These improvements aim to strengthen the robustness and reproducibility of the findings.

Any further responses from the reviewers can be found at the end of the article

Introduction

Infancy, the period from birth to around two years of age, is a crucial stage in human development, characterized primarily by rapid physical, mental, and emotional growth.[1] During this period, many environmental factors such as gestational age, sex, ethnicity, dietary intake, nutritional status, and socioeconomic conditions significantly affect child growth, development, and overall health.[2,3] However, to date, these factors, either alone or synergistically, have failed to explain the overall spectrum and variations in children. Recent studies have emphasized that apart from the intersection of environmental factors, the gut microbiome also plays an important role in mediating growth and development, especially during infancy.

The human fetus is believed to develop in a microbe-free environment.[4] However, immediately after birth, the infant gut, especially the distal ileum and colon, is rapidly colonized by different microorganisms (bacteria, fungi, archaea, and viruses) via vertical transmission from maternal body sites such as the vagina, skin, oral cavity, and gut.[5] In newborn infants, the gut microbiota exhibits low diversity, consisting mainly of facultative anaerobic bacteria, such as *Proteobacteria* (e.g., *Escherichia* and *Enterobacter*) and *Firmicutes* (*Staphylococcus*, *Enterococcus*and, and *Streptococcus*).[6] Early colonization has a profound influence on the gut environment and subsequent colonization by obligate anaerobes, such as *Bifidobacterium, Bacteroides* and *Clostridium.* The relative abundance of *Bifidobacteria* gradually increases 3 – 4 days after birth and becomes a predominant bacterial genus when the baby is approximately 1 month old.[7,8] In healthy growing children, the gut microbiota gradually increases in complexity with a highly diverse microbial composition during and after the first two–three years of life.[9] During this period, the patterned ecological assembly of microbes occurs mainly under the influence of delivery mode, type of feeding, complimentary feeding practices, antibiotic use, and geography.[9–12] Gut microbiota development during early life affects several metabolic, immune, and endocrine pathways and is thus intimately linked to child growth and development.[13] Therefore, any disturbances in the host-microbe coevolution, microbial composition, or functional potential of the gut microbiome, commonly referred to as microbial dysbiosis, can impair immune function, growth, and development in children.

In recent years, several studies have reported associations between gut microbiome dysbiosis and impaired growth and development in children from different countries across the word. For example, a causal relationship between gut microbiome immaturity, undernutrition, and impaired growth has been reported in 6- and 18-months old children.[14] Similarly, a Swedish cohort encompassing 471 children has also reported significantly lower microbial diversity and reduced abundance of *Faecalibacterium* and *Ruminococcus* in the gut microbiome of children with slower weight gain compared to normal children.[15] Research studies from developing countries, where malnutrition is common, have also reported impaired or immature gut microbiome in malnourished (stunting, wasting, and underweight) children compared to that in healthy controls. In Bangladesh, children with severe acute malnutrition (SAM) exhibit a considerably less diverse gut microbiome, low relative abundance of *Bacteroidetes* and high relative abundance of *Proteobacteria* especially the pathogen/pathobiont genera such as *Klebsiella, Escherichia, Shigella*, and *Streptococcus*.[16] In another study of a group of SAM infants aged 6–24 months, the absolute abundance of *B. infantis* was considerably lower than that of age-matched healthy infants.[17] These findings were also confirmed by studies conducted in other developing countries in Asia and Africa.[18–20] Altogether, these findings suggest that an immature/poorly developed gut microbiome in infants and young children is associated with malnutrition and its consequences, particularly impaired growth and development. As a result, a nutritional intervention strategy that does not consider the gut microbiome will fail to ameliorate the long-term consequences of malnutrition, including child health and cognitive development.[21] Clinical studies in animals and humans have already demonstrated the positive impact of microbiome-based nutritional interventions on the gut microbiome and ponderal growth.[22,23] These findings may have important implications for low- and middle-income countries such as Pakistan, where malnutrition in children is a crucial public health challenge.

Pakistan is the 5th most populous country in the world (approximately 250 million people)[24] with enormous economic, geopolitical, social, and poverty-related challenges.[25] According to the latest national nutrition survey, the prevalence of

stunting (40.2%), wasting (17.7%), and micronutrient deficiencies in children under five years of age[26] remained unacceptably high. Over the last five years, the country has been placed in the "Serious" category of the Global Hunger Index.[27] The consequences of malnutrition are also economically catastrophic with $7.6 billion/annum (~3% of GDP) lost to malnutrition.[28] Over the last two decades, there has been virtually no improvement in the overall nutritional status of the Pakistani population, despite efforts from the government and developmental organizations. Without coordinated efforts, it is very unlikely that Pakistan will achieve the UN Sustainable Development Goals to end hunger and all forms of malnutrition by 2030.[29] Successful tackling this issue requires cost-effective, culturally relevant, and sustainable nutritional interventions. Keeping in mind the central role of the gut microbiome in mediating malnutrition in children, it is crucial to map the compositional and functional dynamics of the gut microbiome and associated maternal, child, and environment-related factors during early childhood in populations residing in malnourished communities in Pakistan. Therefore, the current study attempts to characterize how the gut microbiome develops in a cohort of children at high risk of malnutrition in the Khyber Pakhtunkhwa province of Pakistan.

Protocol

Objectives

The primary objective of the CHAMP study is to characterize gut microbiota development and longitudinal growth in infants during the first two years of life. This will be done by assessing the microbial diversity, relative abundance of the dominant microbiota, and interindividual variations in microbial diversity and infant growth patterns. The secondary objective of the study is to investigate the impact of (i) sociodemographic characteristics, (ii) mode of delivery, (iii) maternal factors such as diet and anti-natal and post-natal care, (iv) child dietary intake and feeding practices, and (v) nutritional status of the child on gut microbiota composition at different time points from birth until the child is two years old.

Operational definitions

- **Nutritional status:** The nutritional status of the child will be measured in terms of weight for age (WAZ), height for age (HAZ), and weight for height (WHZ) according to the WHO reference range.[30]
- **Underweight:** Defined as weight for age of the child below two standard deviations (-2SD) from the median weight for age of the reference population.
- **Stunting:** Child height for age below minus two standard deviations (-2SD) from the median height for the age of the reference population.
- **Wasting:** Child weight for height below minus two standard deviations (-2SD) from the median weight for height of the reference population.

Study design and setting

The CHAMP study is a population-based, prospective cohort study involving mother-infant pairs (dyads) from District Swat, Pakistan. Swat is located at 34°46′58″ N and 72°21′43″ N in Khyber Pakhtunkhwa province of Pakistan and is home to more than 2.3 million people.[31] Because of its geographic location in the Hindukush–Himalayan region, most of the land is occupied by mountains and thick forests. The area is one of the most marginalized and neglected regions of the province, with livestock, agriculture, horticulture, and tourism as the primary sources of income and livelihood of the local population.[32] Swat is among the most vulnerable districts in Pakistan and is prone to climate change. Since 2010, the district has been hit by at least three major floods that have significantly disrupted essential services, such as healthcare, sanitation, and access to safe drinking water.[33] As a result, the majority of the population, especially children, are at a high risk of malnutrition and its deleterious consequences.

Administratively, the district is divided into seven administrative units known as Tehsil Municipal Administration units (TMAs). Tehsil Matta was chosen as the study area because it is the second most populated tehsil in Swat district, with the highest number of rural and remote communities. Following a preliminary survey, UC Beha was selected as the study site for several reasons. It is located at a high altitude with the longest adjoining borders with district Dir (Upper), and most of the population lives in agropastoral communities with very limited access to education and healthcare facilities. In winter, access to the area is limited by heavy snowfall and landslides.

Sampling technique

District swats are divided into seven administrative units known as Tehsil Municipal Administration units (TMAs). Of these, tehsil Matta was chosen as the study area because it is the second most populated tehsil in Swat district, located at a high altitude with the longest adjoining borders with Dir district (Upper), which has the highest number of

rural and remote communities. The Matta tehsil is further divided into union councils, which are the primary administrative institutions in Pakistan. In rural areas, union councils are referred to as village councils. Of the 13 village councils in the sampling area, three union councils will be randomly selected as the primary sampling unit. Subsequently, the researchers visited the primary health centers and designated vaccination centers within the chosen village councils. Newborn children will be identified using the Extended Program for Immunization (EPI) register. Potential participants will be appraoched with the help of Lady Health Workers (LHWs), community-based health workers linked to local health facilities. All families in the area are registered with a designated LHW who regularly visited them. Important information, such as the child's name, father's name, date of birth, and address, will be recorded, and a list of eligible study participants will be prepared. Once recruited, each participant will be followed up for two years. The study follows ethical principles involving human subjects as outlined in the declaration of Helsinki. Ethical approval for this study was granted by the Ethics Board of Khyber Medical University, Pakistan ref no DIR/KMU-EB/BR/001-03 dated 11/01/2024.

Participants selection

The study population will include 70 mother-infant pairs (dyads), according to the following criteria:

Inclusion criteria

- Apparently healthy infants of both genders aged 0 – 28 days
- Born by natural or caesarean delivery
- Born to parents from district Swat
- Parents/caregivers have no plans to move out of the stud site for at least two years after enrollment in the study.

The exclusion criteria are:

- Child born to an underage (<18 years old) mother.
- Infants born with severe acute or chronic medical conditions that require hospitalization, prolonged use of medication, or both, or diagnosed with enteropathies.
- Weight of the child is <1500 gm.

Ethical approval for the study was granted by the Research Ethics Board of Khyber Medical University. This study was also registered at www.clinicaltrial.gov (NCT05793294).

Recruitment

In the local (Pashtun) culture, it is not possible to directly contact a potential female study participant without prior consent from the husband/male head of the household. Therefore, at each study site, an initial information session was held with the male parents/guardians with the help of elected local representatives and elders of the community. During the session, complete information about the study objectives, eligibility criteria, and data collection process was presented and explained by the team leader in the local (Pashto) language. A participant information sheet containing all the study details and procedures in an easy-to-understand, local (Pashto), and Urdu (national language of Pakistan) were also provided to the parents during the session. All queries were also answered. Following the session, eligible parents/guardians were invited by trained research assistants to enroll their babies in the study. Once they agreed, the parents were asked to sign a written informed consent form in their preferred language.

Participants, funder and public involvement

Community members, funder (NIH Pakistan) or the participants' parents were not directly involved in developing the study design, conduct, and outcome measures. The study protocol was reviewed and approved by the Office of Research, Innovation, and Commercialization (ORIC) Khyber Medical University and National Institute of Health (NIH), Pakistan. Healthcare professionals (LHWs) involved in the study had close liaison with the community and acted as a bridge between the researchers and participants. They also contributed to the design of the data collection questionnaires, informed consent, and helped ensure confidentiality. However, the individual data were not reported to the participants.

Study procedures

Figure 1 provides an overview of the overall study flow.

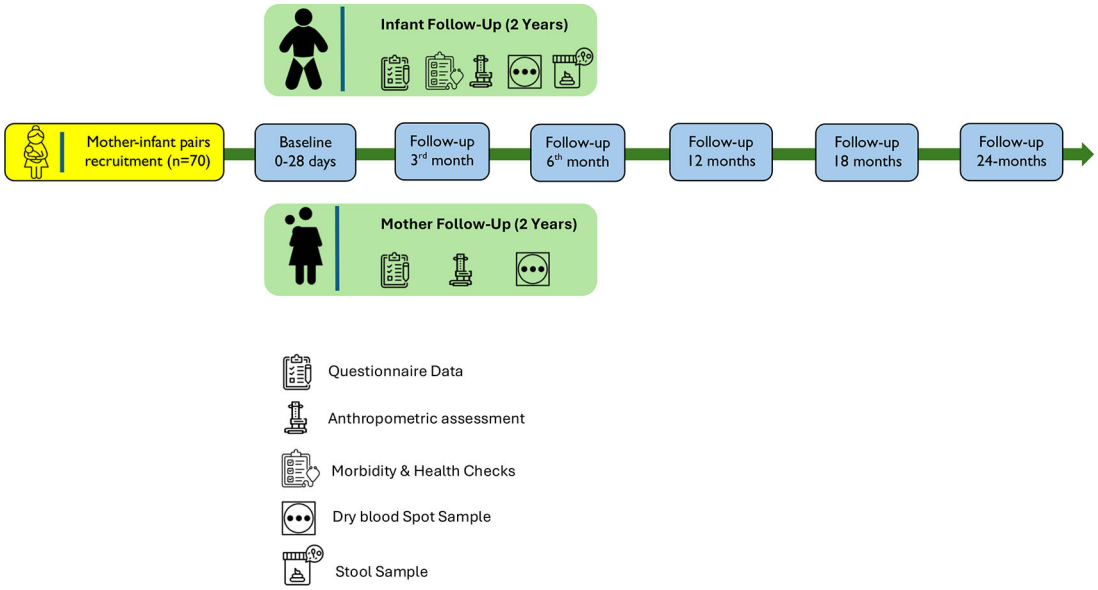


Figure 1. Flowchart of the study.

Data collection

We will collect different data from both mother and infant at baseline and at each time point during the follow-up period. The data collection in this project used an Android version of the KoBoCollect/Toolbox as a data collection instrument. To collect the demographic and socioeconomic information of the study participants, an interviewer administered a structured questionnaire will be used. The questionnaire included information such as name, age, sex, household information, and socioeconomic status. Details about ante-natal and post-natal care will also be recorded using a structured, validated questionnaire adapted from the National Nutrition Survey Pakistan questionnaire.[34] Sociodemographic, ante-natal and post-natal care data will be recorded only at baseline, that is, at the start of the study.

The Dietary Quality Questionnaire (DQQ) will be used to assess the dietary intake of mothers. This method was chosen because, in Pakistan, there is no validated food frequency questionnaire to assess dietary intake. The DQQ is a standardized and easy-to-use tool to collect food group consumption data required for calculating diet quality indicators, such as Minimum Dietary Diversity for Women (MDD-W). The questionnaire has been validated and implemented across more than 100 countries including Pakistan.[35] The DQQ is a list-based method consisting of yes/no responses to 29 food groups consumed by the respondent during the previous day and night (24 hours). The procedure typically requires approximately 5–7 minutes to complete. Information regarding feeding practices of the children will be collected using validated, semi-structured questionnaires based on Infant and Young Child Feeding Practices (IYCF) indicators developed by the Technical Expert Advisory Group on Nutrition Monitoring (TEAM) of the WHO.[36] This set of simple, validated, and reliable indicators is widely used to assess IYCF practices across the globe, especially in low- and middle-income countries. The IYCF indicators were assessed by interviewing the mothers or the main caregivers. The mothers will also be interviewed to collect data about the child's medical history, diarrhea and respiratory tract infection, hospitalization, and use of antibiotics using a structured questionnaire. DQQ, IYCF, and child health record data will be collected at baseline and at each time point during follow-up.

Anthropometric assessment

At baseline and at each time point, the following anthropometric assessment will be performed on each child/infant and their mothers using standard methods.[37]

Height/length: Recumbent length will be measured using an infantometer. Before measuring, the measuring board will be placed on a hard, flat surface such as a ground or sturdy table. The mother will be asked to place the child gently on the board such that the infant's head is aligned against the headboard while an assistant straightens the infant body and feet.

When the child is position is correct, the researcher moves the foot board firmly against the child's heels. The length of the child will be recorded to the nearest 0.1 cm. The measurement will be repeated twice, and the average of the two measurements is recorded as the final length of the child. Mothers' height will be recorded using a stadiometer following standard protocols. Briefly, they will be asked to remove shoes and gently stand on stadiometer. To ensure accurate measurement, the mothers will be instructed to stand up against the board, ensuring the Frankfurt plan position in which the back of the head, shoulder blades and buttocks and heels touch the back of the stadiometer. Once ready, the researcher will gently and firmly slide the measuring board's moveable headpiece down until it touches the crown of the person's head (compresses the hair). The procedure is repeated twice and the average of the two measurements is recorded as the height of the mother.

Weight: To record the weight of the child, a specialized electronic scale (Secca, UK) will be used. Before weighing the child, the parent will be instructed/helped to remove shoes, socks and heavy clothing except dry clean diapers or underpants from the child. When ready, the child is placed on his/her back on the center of the scale pan with the help of a trained researcher. We wait until the infant is still in the pan and the digital display is no more changing. The weight of the child will be recorded to the nearest 0.01 kg. The weight of the mother is measured in kilograms using a calibrated electronic weighing scale. The mothers will be asked to remove shoes, jewellery and extra clothing (dupatta/chaddar) and stand calmly on the scale. The weight will be recorded to the nearest 0.01 kg twice and the average of the two measurement is considered the final weight of the mother.

Mid upper arm circumference (MUAC): For infant MUAC measurements, a non-stretchable, numbered and colored MUAC tape will be used. To prepare for the measurement, the infant mother will be asked to sit on a comfortable chair, place the child in her lap and remove any clothing covering the left hand of the child. The mid-point between the tip of the elbow and the shoulder is located by straightening the arm along the axis of the body and forearm at 90-degree forwards. Once the mid-point is marked with a pen, the arm is straightened and the MUAC tape is placed around the arm at the midpoint and the reading is taken to the nearest mm.

Head circumference: The head circumference of the child will be measured using a non-stretchable measuring tape around the most prominent part of head to the middle of the forehead while the hair and soft tissue is compressed. At least two measurements will be recorded and the average of the two will be taken.

Sample collection

Dry blood spot: Dry blood spots (DBS) are collected from both the mother and the infant at baseline and during the follow-up by skin puncture of the third or fourth finger of the non-writing hand.[38] Before DBS collection, the participants' hands are first warmed, followed by anterograde massaging of the finger in the direction of blood flow towards the puncture site. The DBS collection site, that is, the palmar side of the tip of the distal phalanx, is cleaned with 70% isopropyl alcohol and punctured using a single-use lancet. The first drop of blood is wiped off with a gauze pad, and subsequent drops were transferred to the marked circles on the surface of the filter paper without touching the surface. After drying, the DBS papers will be transferred to a zip-lock bag and stored at -80 °C until further processing in the KMU main lab.

Stool: Stool samples are collected from the diapers of infants at each time point. For this purpose, the parents will be provided diapers, a specimen collection jar, disposable gloves, a zip lock bag, and an instruction sheet one day before sample collection. The parents are instructed to regularly check their child's diaper and remove it immediately after the child passed the stool. Stool samples are transferred immediately to the specimen container, placed in a zip-lock bag, and handed over to the research assistant within 4 h of sample collection. The research assistant transferred 400 mg of stool sample into 2 mL screw-top tubes prefilled with DNA shield reagent (Zymo Research USA) and sent it to the main KMU lab at room temperature, where it will be stored at -80 °C.

Laboratory analysis

Hematological assessment

Dry blood samples will be used to assess the biomarkers of infections, micronutrient deficiencies, and inflammation. For this purpose, commercially available chemiluminscence-based Q-plex Aray kits (Quansys Biosciences) will be used following the manufacturer's instructions. This assay reliably detects and quantifies histidine-rich protein II (biomarker of malarial infections), C-reactive protein, alpha-1-acid glycoprotein (biomarkers of inflammation), ferritin and soluble transferrin receptor (biomarker of iron deficiency), retinol binding protein (biomarker of vitamin A deficiency), and thyroglobulin (biomarker of iodine deficiency).[39]

DNA extraction, library preparation and shotgun metagenomic sequencing

Metagenomic DNA from stool samples will be extracted using MagPure Stool DNA KF Kit B (MAGEN, Guangzhou, China) according to manufacturere instructions. Library preparation is perfored using MGIEasy Universal DNA Library Prep Set (MGI-Shenzhen, China). Genomic DNA will be subjected to fragmentation processing. Fragmented samples are size selected through magnetic beads and converted to blunt-end DNA with end repair reaction. A single Adenosine nucleotide is added on the 3' end of DNA through the A tailing reaction library adapters are connected to the two ends of DNA by adaptor ligation. Finally, the library products are amplified through PCR reaction and subjected to quality control process. The final double strand library products are denatured to generate the single stranded library product. Then, the circularization reaction is set up to get single stranded circularized DNA products. Any single strand linear DNA will be digested to remove. The final single strand circularized library is amplified with phi29 and rolling circle amplification (RCA) to generate the DNA nano ball (DNB) which carries about 300 copies of the initial single stranded library molecule. The DNBs are loaded into the patterned nanoarray and sequencing reads of PE150 bases length are generated with DNBSEQ-G400 platform (BGI-Shenzhen, China).

Monitoring and evaluation strategy

The monitoring and evaluation of the entire project will be conducted at all levels. Data entry will be randomly rechecked for accuracy by re-interviews or re-visits to households participating in the study. The principal investigator oversees the data collectors and enumerators. Bimonthly assessments will be performed by the project team consisting of the project manager, the PI, and the co-PI. In addition, the KMU ORIC team will monitor the overall progress of the project through its own monitoring and evaluation system. Any discrepancy or deviance from the main objectives or lapse in the timeframe and data collection procedures will be handled in close coordination of the project team with the field data collector and support from KMU.

Participants retention strategies

A common issue in longitudinal cohort studies is systemic attrition, which can have a significant impact on the generalizability of the outcome of interest. Study participants may lose interest in the study and therefore decline their participation, especially in long-term cohort studies. Below is a list of the potential reasons for attrition in the current project and mitigation strategies.

1. The first and most important reason might be that some participants left the study district permanently and left the research team unable to track and collect data at different time points. To tackle this issue, we have set our inclusion criteria such that only those participants who are permanent residents of these districts and who confirm that they have no plans to move outside the district permanently will be recruited.

2. Reducing barriers to participation by hiring local research assistants and their training, ensuring anonymity, translating the data collection tools to the local language (Pashto), and pilot testing.

3. A follow-up and reminder strategy was set up by calling the participants before data collection.

4. Tracing via alternative phone numbers and visiting the house physically when participants did not respond.

Data analysis

The raw sequencing data will be processed using metapi pipeline (https://github.com/ohmeta/metapi) followed by quality control process using standards for filtering and trimming the reads (average Phred quality score ≥ 20 and length ≥ 30) using fastp v0.20.1 . Human reads will be removed using Bowtie2 2.4.2 (human genome GRCh38). Taxonomic profiling of the samples will be performed using Kraken2 with standard bacterial, viral, and eukaryotic databases. Diversity assesment will be performed using the MicrobiotaProcess package. Differences in the alpha diversity will be estimated using the Wilcoxon rank sum test. Bidimensional visualization (Beta diversity) will be measurted using non-metric multidimensional scaling and the Bray–Curtis dissimilarity metric implemented in MicrobiotaProcess packages for GNU/R. The ADONIS function was used for obtaining differences in the taxonomic compositions between the groups. Assessment of the differentially abundant species will be done using the metagenomeSeq package.[30] An adjusted p-value ≤ 0.05 will be considered statistically significant.

We will seek to identify any correlations between the collected metadata and microbiome composition, and identify any gutmicrobiota trends associated infant demographic characteristics, dietary intake (IYCF data), nutritional status (weight for height, weight for age, and height for age and BMI for age z-score, MUAC and head cicrumference) health record

(history of diarrhea, upper respiratory tract infections, antibiotic use and duration etc). We will also decorticate the changes in infant gut microbiome in relation to mother dietary intake data (collected though DQQ) and nutritional status (BMI). All postprocessing and statistical analyses will be performed in R.

Dissemination

The results of the study will be disseminated to the public and academia, including our research team and collaborators, scientists, and the broader research community via participation in scientific conferences, meetings, and publications in peer-reviewed journals.

Study status

The study began in 2024 and will continue until May 2026. As of October 2024, 70/70 (100%) were recruited. Data and samples were collected for the baseline (n=70) and first follow-up periods (3 months; n = 66).

Discussion

The CHAMP study aimed to longitudinally characterize gut microbiome development in newborn infants during the first two years of life and included 70 mother-infant pairs (dyads) recruited from remote, rural communities of District Swat, Pakistan. This study involved comprehensive data collection regarding dietary intake, nutritional status, infant growth, and morbidities at different time points. Prospective collection of stool samples will allow gut microbiome assessment over time and identify demographic, maternal, and child-related factors affecting its diversity and abundance. The study findings will contribute towards a better understanding of gut microbiome colonization and development, associated factors, and its impact on growth in infants from malnourished areas of Pakistan.

Microbial colonization of the gut during infancy plays a crucial role in establishing the intestinal barrier and developing the immune system.[40] Numerous studies have reported rapid changes in the composition and diversity of the microbiota during the first two years of life in healthy infants. However, the pattern and extent of microbial colonization are largely influenced by feeding practices and nutritional status of infants. A bidirectional relationship exists between the development of the gut microbiome and nutritional status. Malnutrition can affect gut microbiome development and maturation during early childhood, resulting in dysbiosis. As a result, an immature and altered gut microbiome leads to malnutrition owing to impaired energy production, vitamin biosynthesis, and immune dysfunction.[41,42] These findings are critical, especially in developing countries such as Pakistan, where childhood malnutrition is a public health issue.[34] The CHAMP study will help to dissect the relationships between dietary intake, nutritional status, and gut microbiome phenotypes during infancy and pave the way for developing microbiome-based interventions to tackle the problem of malnutrition in Pakistan.

Our study is the first to characterize gut microbiome development in newborn infants from malnourished areas of Pakistan. The study will greatly help to understand infant growth trajectory in relation to gut microbiome development and provide a conceptual basis for sustainable nutrition recommendations and interventions in the future. The prospective longitudinal nature of the study will help in the collection of detailed data with minimal constraints (six visits in 2 years) and associated risks (no intervention and minimally invasive procedures to collect biological samples). This study will obtain precise data on dietary intake and nutritional status and their short-term and medium-term impacts on gut microbiome development, and vice versa. To gain deeper insights into how these factors, especially dietary intake patterns and socioeconomic status, impact gut microbiome development, we will compare our study results with similar studies conducted in Zimbabwe, a low-income country (The University of Zimbabwe College of Health Sciences (UZ-CHS) Birht Cohort)[43] and Switzerland, a high-income country (The Bern Birth Cohort – BeBiCo study).[44] Both of these studies involved mother-infant pairs and collected neonatal data on demographics, socioeconomic status, and diet at time points corresponding to our study. These studies also collected fecal samples, which will allow us to compare gut microbiome development in Pakistani infants with infants from Switzerland and Zimbabwe, countries with completely different landscapes in terms of environmental, cultural, and dietary intake patterns.

Our study has some limitations. First, the launch and implementation of a cohort study in remote, difficult-to-access rural areas are challenging, anywhere in the world, including Pakistan. As a result, dropout and attrition in the sample size are expected. Second, the study involved multiple assessments, requiring the mother to travel long distances in mountainous areas to reach the sampling site. The process is time-consuming and potentially burdensome to the mother and infants. However, these issues will be partly addressed by providing compensation for travel and sending advance reminders about follow-up visits using phone calls. Third, due to the self-reported and retrospective nature of the data collection during each visit, response and social desirability bias cannot be fully avoided. Finally, we could only identify the association between growth and gut microbiome development, but no causal relationship because of the non-interventional study design.

Ethical considerations
The study follows ethical principles involving human subjects as outlined in the declaration of Helsinki. Ethical approval for this study was granted by the Ethics Board of Khyber Medical University, Pakistan ref no DIR/KMU-EB/BR/001-03 dated 11/01/2024.

Consent statement
In the local (Pashtun) culture, it is not possible to directly contact a potential female study participant without prior consent from the husband/male head of the household. Therefore, at each study site, an initial information session was held with the male parents/guardians with the help of elected local representatives and elders of the community. During the session, complete information about the study objectives, eligibility criteria, and data collection process was presented and explained by the team leader in the local (Pashto) language. A participant information sheet containing all the study details and procedures in an easy-to-understand, local (Pashto), and Urdu (national language of Pakistan) were also provided to the parents during the session. All queries were also answered. Following the session, eligible parents/guardians were invited by trained research assistants to enrol their babies in the study. Once they agreed, the parents were asked to sign a written informed consent form in their preferred language.

Data availability
Underlying data
No data are associated with this article.

Extended data
Muhammad, Shahzad, 2024, "Child health, nutrition and gut microbiota development during the first two years of life; study protocol of a prospective cohort study from the Khyber Pakhtunkhwa, Pakistan", https://doi.org/10.7910/DVN/ROU3VJ[45]

This project contains the following extended data:

- Data Collection Quarionnaire.doc

Data are available under the terms of the Creative Commons Zero "No rights reserved" data waiver (CC0 1.0 Public domain dedication).

References

1. Bornstein MH: **Human Infancy ... and the Rest of the Lifespan.** *Annu. Rev. Psychol.* 2013 Sep 13; **65**: 121–158.
Publisher Full Text
2. Alur P, Holla I, Hussain N: **Impact of sex, race, and social determinants of health on neonatal outcomes.** *Front. Pediatr.* 2024 Apr 9 [cited 2024 Oct 20]; **12**.
PubMed Abstract | Publisher Full Text | Free Full Text
3. Thahir AIA, Li M, Holmes A, *et al.*: **Exploring Factors Associated with Stunting in 6-Month-Old Children: A Population-Based Cohort Study in Sulawesi, Indonesia.** *Nutrients.* 2023 Aug 1; **15**(15): 3420.
PubMed Abstract | Publisher Full Text | Free Full Text
4. Dominguez-Bello MG, Costello EK, Contreras M, *et al.*: **Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns.** *Proc. Natl. Acad. Sci. USA.* 2010 Jun 29; **107**(26): 11971–11975.
PubMed Abstract | Publisher Full Text | Free Full Text
5. Ferretti P, Pasolli E, Tett A, *et al.*: **Mother-to-Infant Microbial Transmission from Different Body Sites Shapes the Developing Infant Gut Microbiome.** *Cell Host Microbe.* 2018 Jul 11; **24**(1): 133–145.e5.
PubMed Abstract | Publisher Full Text | Free Full Text
6. Bäckhed F, Roswall J, Peng Y, *et al.*: **Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life.** *Cell Host Microbe.* 2015 May 13; **17**(5): 690–703.
PubMed Abstract | Publisher Full Text
7. Bittinger K, Zhao C, Li Y, *et al.*: **Bacterial colonization reprograms the neonatal gut metabolome.** *Nat. Microbiol.* 2020 Jun; **5**(6): 838–847.
PubMed Abstract | Publisher Full Text | Free Full Text
8. Henderickx JGE, Zwittink RD, van Lingen RA, *et al.*: **The Preterm Gut Microbiota: An Inconspicuous Challenge in Nutritional Neonatal Care.** *Front. Cell. Infect. Microbiol.* 2019; **9**: 85.
PubMed Abstract | Publisher Full Text | Free Full Text
9. Stewart CJ, Ajami NJ, O'Brien JL, *et al.*: **Temporal development of the gut microbiome in early childhood from the TEDDY study.** *Nature.* 2018 Oct; **562**(7728): 583–588.
PubMed Abstract | Publisher Full Text | Free Full Text
10. Beller L, Deboutte W, Falony G, *et al.*: **Successional Stages in Infant Gut Microbiota Maturation.** *MBio.* 2021 Dec 21; **12**(6): e0185721.
PubMed Abstract | Publisher Full Text | Free Full Text
11. Olm MR, Dahan D, Carter MM, *et al.*: **Robust variation in infant gut microbiome assembly across a spectrum of lifestyles.** *Science.* 2022 Jun 10; **376**(6598): 1220–1223.
PubMed Abstract | Publisher Full Text | Free Full Text
12. Xiao L, Wang J, Zheng J, *et al.*: **Deterministic transition of enterotypes shapes the infant gut microbiome at an early age.** *Genome Biol.* 2021 Aug 24; **22**(1): 243.
PubMed Abstract | Publisher Full Text | Free Full Text
13. Robertson RC, Manges AR, Finlay BB, *et al.*: **The Human Microbiome and Child Growth - First 1000 Days and Beyond.** *Trends Microbiol.* 2019 Feb; **27**(2): 131–147.
PubMed Abstract | Publisher Full Text
14. Blanton LV, Charbonneau MR, Salih T, *et al.*: **Gut bacteria that prevent growth impairments transmitted by microbiota from malnourished children.** *Science.* 2016 Feb 19; **351**(6275): aad3311.
PubMed Abstract | Publisher Full Text | Free Full Text
15. Roswall J, Olsson LM, Kovatcheva-Datchary P, *et al.*: **Developmental trajectory of the healthy human gut microbiota during the first**

5 years of life. *Cell Host Microbe.* 2021 May 12; **29**(5): 765–776.e3.
PubMed Abstract | Publisher Full Text

16. Monira S, Nakamura S, Gotoh K, *et al.*: **Gut microbiota of healthy and malnourished children in bangladesh.** *Front. Microbiol.* 2011; **2**: 228.

17. Barratt MJ, Nuzhat S, Ahsan K, *et al.*: **Bifidobacterium infantis treatment promotes weight gain in Bangladeshi infants with severe acute malnutrition.** *Sci. Transl. Med.* 2022 Apr 13; **14**(640): eabk1107.
PubMed Abstract | Publisher Full Text | Free Full Text

18. Camara A, Konate S, Tidjani Alou M, *et al.*: **Clinical evidence of the role of Methanobrevibacter smithii in severe acute malnutrition.** *Sci. Rep.* 2021 Mar 8; **11**(1): 5426.
PubMed Abstract | Publisher Full Text | Free Full Text

19. Khan Mirzaei M, Khan MAA, Ghosh P, *et al.*: **Bacteriophages Isolated from Stunted Children Can Regulate Gut Bacterial Communities in an Age-Specific Manner.** *Cell Host Microbe.* 2020 Feb 12; **27**(2): 199–212.e5.
PubMed Abstract | Publisher Full Text | Free Full Text

20. Nabwera HM, Espinoza JL, Worwui A, *et al.*: **Interactions between fecal gut microbiome, enteric pathogens, and energy regulating hormones among acutely malnourished rural Gambian children.** *EBioMedicine.* 2021 Nov; **73**: 103644.
PubMed Abstract | Publisher Full Text | Free Full Text

21. Mostafa I, Nahar NN, Islam MM, *et al.*: **Proof-of-concept study of the efficacy of a microbiota-directed complementary food formulation (MDCF) for treating moderate acute malnutrition.** *BMC Public Health.* 2020 Feb 17; **20**(1): 242.
PubMed Abstract | Publisher Full Text | Free Full Text

22. Chen RY, Mostafa I, Hibberd MC, *et al.*: **A Microbiota-Directed Food Intervention for Undernourished Children.** *N. Engl. J. Med.* 2021 Apr 22; **384**(16): 1517–1528.

23. Gehrig JL, Venkatesh S, Chang HW, *et al.*: **Effects of microbiota-directed foods in gnotobiotic animals and undernourished children.** *Science.* 2019 Jul 12; **365**(6449): eaau4732.
PubMed Abstract | Publisher Full Text | Free Full Text

24. Worldometer: **Countires in the world (by population).** 2024 [cited 2024 Oct 21].
Reference Source

25. Afzal M: **Brookings. Pakistan: Five major issues to watch in 2023.** 2023 [cited 2024 Oct 21].
Reference Source

26. UNICEF: **National Nutrition Survey 2018, Key findings report.** 2019.
Reference Source

27. GHI: **Global Hunger Index (GHI) - peer-reviewed annual publication designed to comprehensively measure and track hunger at the global, regional, and country levels. Global Hunger Index 2024 - Pakistan.** 2024 [cited 2024 Oct 21].
Reference Source

28. WFP: **Malnutrition Costs Pakistan US$7.6 Billion Annually, New Study Reveals - Pakistan|ReliefWeb.** 2017 [cited 2024 Oct 21].
Reference Source

29. Kinyoki DK, Osgood-Zimmerman AE, Pickering BV, *et al.*: **Mapping child growth failure across low- and middle-income countries.** *Nature.* 2020 Jan; **577**(7789): 231–234.
Publisher Full Text

30. WHO: *The WHO Child Growth Standards.* 2018 [cited 2024 Oct 21].
Reference Source

31. Pakistan Bureau of Statistics: **Final Results (Census-2017)|Pakistan Bureau of Statistics.** 2017 [cited 2023 Apr 18].
Reference Source

32. Khan SR, Khan SR: **Assessing poverty–deforestation links: Evidence from Swat, Pakistan.** *Ecol. Econ.* 2009 Aug 15; **68**(10): 2607–2618.
Publisher Full Text

33. Bazai NA, Alam M, Cui P, *et al.*: **Dynamics and Impacts of Monsoon-Induced Geological Hazards: A 2022 Flood Study along the Swat River in Pakistan.** *Natural Hazards and Earth System Sciences Discussions.* 2024 Jul 16; pp. 1–27.

34. UNICEF Pakistan: **National Nutrition Survey 2018.** 2018 [cited 2024 Oct 22].
Reference Source

35. Global Diet Quality Project: **Dietary Quality Questionnaire - Pakistan.** 2024 [cited 2024 Oct 22].
Reference Source

36. WHO: **Indicators for assessing infant and young child feeding practices: definitions and measurement methods.** 2021 [cited 2024 Oct 22].
Reference Source

37. USAID: **Food and Nutrition Technical Assistance III Project (FANTA): Module 6 - Protocols and Equipments.** 2018 [cited 2024 Oct 22].
Reference Source

38. Grüner N, Stambouli O, Ross RS: **Dried Blood Spots - Preparing and Processing for Use in Immunoassays and in Molecular Techniques.** *J. Vis. Exp.* 2015 Mar 13; **97**: 52619.
Publisher Full Text

39. Brindle E, Lillis L, Barney R, *et al.*: **Measurement of micronutrient deficiency associated biomarkers in dried blood spots using a multiplexed immunoarray.** *PLoS One.* 2019 Jan 8; **14**(1): e0210212.
PubMed Abstract | Publisher Full Text | Free Full Text

40. Houghteling PD, Walker WA: **Why is initial bacterial colonization of the intestine important to the infant's and child's health?.** *J. Pediatr. Gastroenterol. Nutr.* 2015 Mar; **60**(3): 294–307.
PubMed Abstract | Publisher Full Text | Free Full Text

41. Iddrisu I, Monteagudo-Mera A, Poveda C, *et al.*: **Malnutrition and Gut Microbiota in Children.** *Nutrients.* 2021; **13**(8).
PubMed Abstract | Publisher Full Text | Free Full Text

42. Million M, Diallo A, Raoult D: **Gut microbiota and malnutrition.** *Microb. Pathog.* 2017 May; **106**: 127–138.
Publisher Full Text

43. Duri K, Gumbo FZ, Munjoma PT, *et al.*: **The University of Zimbabwe College of Health Sciences (UZ-CHS) BIRTH COHORT study: rationale, design and methods.** *BMC Infect. Dis.* 2020 Oct 2; **20**(1): 725.
PubMed Abstract | Publisher Full Text | Free Full Text

44. Cecchini L, Barmaz C, Cea MJC, *et al.*: **The Bern Birth Cohort (BeBiCo) to study the development of the infant intestinal microbiota in a high-resource setting in Switzerland: rationale, design, and methods.** *BMC Pediatr.* 2023 Nov 10; **23**(1): 560.
PubMed Abstract | Publisher Full Text | Free Full Text

45. Muhammad S: **Child health, nutrition and gut microbiota development during the first two years of life; study protocol of a prospective cohort study from the Khyber Pakhtunkhwa, Pakistan.** *Harvard Dataverse.* 2024.
Publisher Full Text

Open Peer Review

Current Peer Review Status: ✔ ✔

Version 2

Reviewer Report 19 September 2025

https://doi.org/10.5256/f1000research.186009.r407052

© **2025 Jones R.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

✔ **Roshonda B Jones**
North Carolina Agricultural and Technical State University, Greensboro, North Carolina, USA

The authors of "Child health, nutrition and gut microbiota development during the first two years of life; study protocol of a prospective cohort study from the Khyber Pakhtunkhwa, Pakistan" provide a compelling argument for the need to study the developing gut microbiome in children in Pakistan, given the high rates stunting/wasting/micronutrient deficiencies in this population, which are all directly related to the gut microbiome. To examine this, it is necessary to characterize the gut microbiome in this population to understand how to use nutritional interventions to address malnutrition. Here, they present the study protocol that they will use to characterize the gut microbiome in this population.

Here a few minor comments:
1) This second paragraph of the section "Study design and setting" is repeated in the section "Sampling technique". Maybe remove completely or mention only briefly. As it stands currently, it feels like the reader has to read the same information back to back.

2) Figure 1: As it stands, the figure makes it seem that the follow up measures will only be taken at the 2 year visit. I assume it is also happening at the 3M, 6M, 12M also 18M. I would make this clearer in the Figure. Maybe “Infant Follow-Up (2 years)” can be “Infant Follow up 3M, 6M, 12M, 18M and 24M” and the box spans the entire arrow. Also add in something that states what is being collected at baseline.

3) In the section "**Monitoring and evaluation strategy",** it reads “Data entry will be randomly rechecked for accuracy by re-interviews or re-visits to households participating in the study”. How long after a visit will these random rechecks happen? How will the investigators make sure that differences in answers are due to the responses being taken later in time or that there was a collection error the first time responses were recorded.

4) In the Data Analysis section, when an R package is mentioned make sure that you state "R package" instead of just "package" so as to not assume that the reader knows it is an R package.

Is the rationale for, and objectives of, the study clearly described?
Yes

Is the study design appropriate for the research question?
Yes

Are sufficient details of the methods provided to allow replication by others?
Yes

Are the datasets clearly presented in a useable and accessible format?
Not applicable

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: I am a microbial computational biologist focusing on the developing microbiome in children.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 12 September 2025

https://doi.org/10.5256/f1000research.186009.r403885

© 2025 Marsh R. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Ryan Marsh
Department of Applied Sciences, Northumbria University, Newcastle,, UK

I am pleased to find that the authors have addressed my previous comments. This includes clarity surrounding the study design and data collection, alongside comprehensive detail upon the *now* altered sequencing approach that further accommodates predicted functionality of microbial communities sequenced.

As for other suggestions regarding sampling, I understand the cultural implications as discussed by the authors with respect to various anatomical sites, but applaud the decisions to "*collect oral swab samples from both mothers and infants. These samples are now systematically collected at each scheduled time point.*"

Overall, I am pleased with the changes and eagerly await results from the study - best of luck!

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Gut microbiome research (Cystic Fibrosis), utilising multi-omics.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 05 February 2025

https://doi.org/10.5256/f1000research.173997.r358559

© 2025 Marsh R. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Ryan Marsh
Department of Applied Sciences, Northumbria University, Newcastle,, UK

Dear Authors,

Many thanks for the opportunity to review this work. Please find below my comments to each section of the report.

Is the rationale for, and objectives of, the study clearly described? (YES)

The Study Protocol presented here by Shahzad and colleagues undoubtedly addresses the warranted need to further understand early life microbial dynamics across infants from malnourished areas of Pakistan. Elucidating the relationships between the microbiome and early lifestyle factors will no doubt increase favourable outcomes throughout development in children from such regions.

I have no qualms with the study rationale, and the body of literature cited for this work is of high quality and relatable to the study background.

Is the study design appropriate for the research question? (PARTLY)

Overall, I am happy with the study design, which provides a logical narrative. I understand that sampling may have already begun, but I would double check tense usage throughout just to be consistent and to amend any potential language errors.

I understand a variety of questionnaires will be/are being used by the researchers to collect the relevant demographic information. Some clarity would be appreciated on the following:

Study Procedures, Data collection: First paragraph reads at the end: "These data were recorded only at baseline, that is, at the start of the study". Does this refer only to the original demographic/socio-economic data, and *not* the ante-natal and post-natal care information? If the latter *is* included, I would highly recommend the authors specify examples of what this may encompass, as I believe elements of these may overlap with other questionnaires poised by the authors, such as antibiotic usage.

As for other areas of the sampling:

In-person sampling: It would be good to set or specify if there is a time-standardisation here upon arrival, as the limitations describe the burdensome journey that will have to be made by participants. This would be applicable to dry blood spot testing if this is to commence following this journey, which may skew data on inflammatory outcomes etc.

Sample collection: Given the authors highlight the potential translocation of bacterial taxa from various anatomical sites between mother and newborn, which specifies the gut of the (line 10, Introduction), it would have been very interesting to investigate the microbiota of the mother as close to the infants' baseline sample as possible, to decipher if common seeding patterns that associate with favourable growth outcomes etc. Not sure if this can be feasibly integrated given the original ethical approvals put in place, but food for thought no doubt.

Are sufficient details of the methods provided to allow replication by others? (PARTLY)

Overall, I am happy with the methods provided.

For the proposed approach for investigation into microbial communities utilising 16S rRNA amplicon sequencing on the Illumina Miseq® platform, I would ask for the following information, as to more clearly allow methods to be replicated by others:

DNA extraction and 16S rRNA sequencing: "Next, library preparation was performed via indexing PCR and the reaction was subsequently cleaned up and quantified, following the Illumina kit's guidelines". Can you please elaborate on which kit in particular is used for the second-step PCR and subsequent library clean-up, or was this done manually whilst referring to Illumina's "16S Metagenomic Sequencing Library Preparation" document? If possible, I would also highly recommend the researchers quantify DNA using fluorometric approaches, such as the Qubit fluorometer if manually adjusting and normalising as hinted in the protocol, to ensure adequate sampling read depth across all participants.

DNA extraction and 16S rRNA sequencing: Given that V3-V4 region is the target of interest, I am assuming that this will be performed using 2 × 300 bp, V3 chemistry? I would always recommend highlighting this if possible. From personal experience I have found that the reverse read (R2) can significantly drop-off on these kits. As the background literature cites differences at the genus level across the microbiome across infants from both developed, and developing studies, I might suggest V2 Chemistry targeting the V4 region (2 × 250 bp) to increase robustness of read depth and quality across samples, if genus-level comparisons are sufficient for the authors.

DNA extraction and 16S rRNA sequencing: Mention of sequencing controls (and any positive mock

community controls) would be useful.

As for the wider interpretations and considerations of the results/data:

Data analysis: The authors describe the relationships they are looking to explore between microbiota diversity and community composition with participant demographics they are looking to collect. Within the "data collection" section the protocol, it states: "The mothers will also be interviewed to collect data about the child's medical history, diarrhea and respiratory tract infection, hospitalization, and use of antibiotics using a structured questionnaire". I would anticipate the researchers to also investigate relationships between antibiotic usage and microbiota structure, but there is no specific mention of this (unlike nutritional status, food diversity, height, weight, age, etc..). Perturbation of intestinal microbial communities following antibiotic administration is well documented, which is of heightened importance during the developmentally critical age window (0-2 years) of the infants in this study. Additionally, I understand from the literature, that over-the-counter dispensation of antibiotics can be common in Pakistan which may increase the likelihood of a significant perturbation events to the intestinal microbiota.

General formatting (PARTLY)

Type errors:

"ouncils" should read "councils" – Sampling technique paragraph, line 4.

"measuring tap" should read "measuring tape" – Anthropometric assessment paragraph, line 1.

"Hematological aassessment" – error in title of this section in the Laboratory analysis.

Overall, I would advocate for the acceptance of the study protocol once these minor changes are implemented. I look forward to reading the final indexed study. Great work so far, and good luck moving forward!

Best,
Dr. Ryan Marsh

References

1. Nohri A, Siddiqui M, Usman G, Sarang S, et al.: Antibiotic dispensation without prescription by community pharmacies in Pakistan. *Journal of Medicine, Surgery, and Public Health*. 2024; **2**.
Publisher Full Text

Is the rationale for, and objectives of, the study clearly described?

Yes

Is the study design appropriate for the research question?

Partly

Are sufficient details of the methods provided to allow replication by others?
Partly

Are the datasets clearly presented in a useable and accessible format?
Not applicable

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Gut microbiome research (Cystic Fibrosis), utilising multi-omics.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 01 Aug 2025

Shahzad Muhamamd

Author Response

Aug 1st, 2025

Muhammad Shahzad
Associate Professor, Institute of Basic Medical Sciences, Khyber Medical University, Peshawar, Pakistan
F1000 Peer Review Journal

Ref: Response to Peer Reviewers' Reports - F1000 Research Peer Review "Child health, nutrition and gut microbiota development during the first two years of life; study protocol of a prospective cohort study from the Khyber Pakhtunkhwa, Pakistan".

Dear F1000 Research Editorial Team,
Thank you for facilitating the peer review process for our manuscript. We have carefully addressed the reviewers' comments, and we believe that the revised manuscript now aligns with the journal's requirements. Below, we provide a line-by-line response to the reviewers' comments, with the original comments highlighted in bold.

Is the rationale for, and objectives of, the study clearly described? (YES)

The Study Protocol presented here by Shahzad and colleagues undoubtedly addresses the warranted need to further understand early life microbial dynamics across infants from malnourished areas of Pakistan. Elucidating the relationships between the microbiome and early lifestyle factors will no doubt increase favourable outcomes throughout development in children from such regions.

I have no qualms with the study rationale, and the body of literature cited for this work is of high quality and relatable to the study background.

Our response: The authors thank the reviewer for taking time out to review our manuscript and enhance the quality of our work.

Is the study design appropriate for the research question? (PARTLY)

Overall, I am happy with the study design, which provides a logical narrative. I understand that sampling may have already begun, but I would double check tense usage throughout just to be consistent and to amend any potential language errors.

Our response: Thanks for your comments. In the revised manuscript, tense usage is double checked and corrected accordingly.

I understand a variety of questionnaires will be/are being used by the researchers to collect the relevant demographic information. Some clarity would be appreciated on the following:

Study Procedures, Data collection: First paragraph reads at the end: "These data were recorded only at baseline, that is, at the start of the study". Does this refer only to the original demographic/socio-economic data, and not the ante-natal and post-natal care information? If the latter is included, I would highly recommend the authors specify examples of what this may encompass, as I believe elements of these may overlap with other questionnaires poised by the authors, such as antibiotic usage.

Our response: Thanks for the suggestion. It refers to both sociodemographic data and ante-natal and post-natal care data as clarified in the revised manuscript.

As for other areas of the sampling:

In-person sampling: It would be good to set or specify if there is a time-standardisation here upon arrival, as the limitations describe the burdensome journey that will have to be made by participants. This would be applicable to dry blood spot testing if this is to commence following this journey, which may skew data on inflammatory outcomes etc.

Our response. Thank you for your valuable feedback. Upon arrival at the study site at each scheduled time point, participants first undergo formal identification procedures. Subsequently, stool samples collected at home are processed, followed by comprehensive data collection and anthropometric assessments. These procedures require a minimum of one hour to complete. Thus, participants are provided with adequate time to rest prior to the collection of oral swab and dried blood spot (DBS) samples.

Sample collection: Given the authors highlight the potential translocation of bacterial taxa from various anatomical sites between mother and newborn, which specifies the gut of the (line 10, Introduction), it would have been very interesting to investigate

the microbiota of the mother as close to the infants' baseline sample as possible, to decipher if common seeding patterns that associate with favourable growth outcomes etc. Not sure if this can be feasibly integrated given the original ethical approvals put in place, but food for thought no doubt.

Our response: Thank you for this valuable suggestion. While we fully acknowledge the importance of maternal microbiota, cultural barriers have unfortunately limited our ability to collect microbiota samples from mothers. However, recognizing its significance, we have obtained ethical approval to collect oral swab samples from both mothers and infants. These samples are now systematically collected at each scheduled time point.

Are sufficient details of the methods provided to allow replication by others? (PARTLY)

Overall, I am happy with the methods provided.

For the proposed approach for investigation into microbial communities utilising 16S rRNA amplicon sequencing on the Illumina Miseq® platform, I would ask for the following information, as to more clearly allow methods to be replicated by others:

DNA extraction and 16S rRNA sequencing: "Next, library preparation was performed via indexing PCR and the reaction was subsequently cleaned up and quantified, following the Illumina kit's guidelines". Can you please elaborate on which kit in particular is used for the second-step PCR and subsequent library clean-up, or was this done manually whilst referring to Illumina's "16S Metagenomic Sequencing Library Preparation" document? If possible, I would also highly recommend the researchers quantify DNA using fluorometric approaches, such as the Qubit fluorometer if manually adjusting and normalising as hinted in the protocol, to ensure adequate sampling read depth across all participants.

DNA extraction and 16S rRNA sequencing: Given that V3-V4 region is the target of interest, I am assuming that this will be performed using 2 × 300 bp, V3 chemistry? I would always recommend highlighting this if possible. From personal experience I have found that the reverse read (R2) can significantly drop-off on these kits. As the background literature cites differences at the genus level across the microbiome across infants from both developed, and developing studies, I might suggest V2 Chemistry targeting the V4 region (2 × 250 bp) to increase robustness of read depth and quality across samples, if genus-level comparisons are sufficient for the authors.

DNA extraction and 16S rRNA sequencing: Mention of sequencing controls (and any positive mock community controls) would be useful.

Our response: Thanks for the detailed feedback and suggestions. We have decided to perform shotgun metagenomic sequencing as it is more reliable than 16s sequencing. Please see revised manuscript for detail changes in the methodology and data analysis.

As for the wider interpretations and considerations of the results/data:

Data analysis: The authors describe the relationships they are looking to explore between microbiota diversity and community composition with participant demographics they are looking to collect. Within the "data collection" section the protocol, it states: "The mothers will also be interviewed to collect data about the child's medical history, diarrhea and respiratory tract infection, hospitalization, and use of antibiotics using a structured questionnaire". I would anticipate the researchers to also investigate relationships between antibiotic usage and microbiota structure, but there is no specific mention of this (unlike nutritional status, food diversity, height, weight, age, etc..). Perturbation of intestinal microbial communities following antibiotic administration is well documented, which is of heightened importance during the developmentally critical age window (0-2 years) of the infants in this study. Additionally, I understand from the literature, that over-the-counter dispensation of antibiotics can be common in Pakistan which may increase the likelihood of a significant perturbation events to the intestinal microbiota.

Our response: Infants Gut microbiome diversity and composition will be assessed in relation to the demographic characteristics, dietary intake and nutritional status of both the mother and infant and infants health record including diarrheal episodes, RTI and antibiotic use history and duration. Please see revised manuscript for details.

General formatting (PARTLY)

Type errors:

"ouncils" should read "councils" – Sampling technique paragraph, line 4.
Corrected

"measuring tap" should read "measuring tape" – Anthropometric assessment paragraph, line 1.

"Hematological aassessment" – error in title of this section in the Laboratory analysis.

Corrected

Competing Interests: No

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research