Title: Correlates of Protection against Rotavirus, Norovirus, Respiratory Syncytial Virus and Shigella: A Prospective Birth Cohort Study in Dhaka, Bangladesh

Background:

In low- and middle-income countries (LMICs), infectious diseases account for a significant portion of the overall disease burden. In 2019, it was estimated that more than 300 million under-5 children were affected by infectious diseases, measured in terms of disability-adjusted life-years (DALYs) (1) UNICEF stated that approximately 5 million children under the age of five died globally, of which about 30% were attributed to infectious diseases, such as pneumonia and diarrhoea (2). Like many other LMIC countries, Bangladeshi under-5 children are also widely suffering from diarrhoea and acute respiratory infections (3). Diarrhoeal disease is the third leading cause of death in under -5 children and is responsible for the death of more than 400,000 children every year (4). According to the nationally representative Bangladesh Demographic and Health Survey (BDHS) 2022, about 5% under-5 children reported having suffered from diarrhoea in the past 2 weeks (5).

Rotavirus and norovirus are among the most frequently identified viral pathogens in children suffering from diarrhoea. *Shigella* spp. acts as a major bacterial contributor to bloody diarrhoea, though their clinical presentation is not exclusively confined to dysentery (4). Rotavirus remains a principal etiological agent of severe dehydration in under-5 children, accounting for over two million hospitalizations globally each year (6). In Bangladesh, the estimated incidence is approximately 10,000 cases per 100,000 children (7). Annually, rotavirus-associated illness is responsible for an estimated 2,500 to 3,000 deaths among Bangladeshi children within this age group (7).

Norovirus, another leading cause of acute gastroenteritis, is typically associated with a mild and self-limited gastrointestinal illness, but it can lead to severe dehydrating diarrhoea in young children (8, 9). Norovirus is responsible for approximately 20% of all acute diarrhoeal episodes worldwide, accounting for an estimated 685 million cases and 212,000 deaths annually (10-12). Findings from hospital-based surveillance study in Bangladesh indicate that 14% of under-5 children tested positive for norovirus (13). Another case-control study identified the presence of norovirus in 15% of the control group, compared to 9% among the cases (14). The MAL-ED birth cohort study found that children in Bangladesh get infected with diverse norovirus strains, where re-infection is very common (15). The findings suggest that birth cohort studies are critical to understanding cross-protective immunity and advancing the development of norovirus vaccines (15).

In 2016, Shigella was ranked as the second leading cause of diarrhoeal mortality among all ages (16). Shigella was responsible for >212,000 deaths in all ages and >64,000 deaths among children aged less than five years in a year worldwide (16). Shigella can cause disease through the ingestion of as few as 10 viable organisms and thus regarded as a highly infectious agent. Infection typically arises via contaminated water or food, or through faecal-oral route (17). In the year 2016, 60,000 deaths and 74,000,000 cases of diarrhoea were attributed to *Shigella* among under-5 children, approximately 20% of which occurred in this part of the world that is South Asia (18). Fatal outcomes in shigellosis may result from a range of both intestinal and systemic complications. Intestinal manifestations include toxic megacolon, intestinal perforation, and severe dehydration (19). Systemic sequelae encompass sepsis, hyponatremia,

hypoglycemia, seizures and encephalopathy, hemolytic-uremic syndrome, pneumonia, and malnutrition (19). According to the Global Enteric Multicenter Study (GEMS), Shigella ranked as the third leading cause of moderate-to-severe diarrhoea among Bangladeshi children aged 12-23 months, and the second leading cause among those aged 24-59 months (20).

Acute respiratory tract infections (ARI) are one of the commonest causes of death in children in countries like Bangladesh (21). Globally, ARI causes 15% of deaths in children under-5 (22), and the vast majority of deaths occur in LMICs, particularly in Africa and Asia (23). The primary cause of acute respiratory tract infections (ARIs) in children under five is viral infections (24). The most prevalent viruses isolated from lower respiratory tract infections were Respiratory Syncytial Virus (RSV), and in case of upper respiratory tract infections, the second most common detected virus was RSV (24). Every year ARI causes about twenty-five percent of the deaths among under-5 children in Bangladesh (25). Bangladesh has a high rate of RSV infection. Findings from the Respiratory Pathogen Genomic Surveillance conducted in Bangladesh between October 2022 and March 2023 indicate that RSV was the predominant circulating pathogen during this period. The highest prevalence was observed among infants younger than six months, followed by those aged 6 to 11 months (26). This study reported an average prevalence of RSV to be 27.6% in six months among under-5 children who reported with an influenza-like illness (26).

A correlate of protection (CoP) can be defined as an immune marker, like certain antibodies or cellular responses, that can be measured and used to detect whether someone is protected against an infection or disease. Identifying CoPs against key pathogens in paediatric populations is particularly important for research, as it is closely linked to the development of effective vaccines. CoPs that are present during the early years of life are important in guiding the evolution of vaccines and the design of studies (27). Validated CoPs can enable smaller immunogenicity studies to assist in further development of vaccines. Due to the absence of effective correlates of protection against pathogens, vaccine efficacy (VE) clinical trials need a clinical endpoint (pathogen-specific morbidity and/or mortality). These trials represent an important challenge since vaccines have to be introduced in many different settings, placebo-controlled studies are unethical, and comparator assessments for new vaccines with clinical endpoints are very large, complex, and expensive to conduct. The consideration of a CoP as a surrogate endpoint would allow predictions of VE for new vaccines and enable a regulatory pathway, contributing to the more rapid development of a new generation of vaccines, and supporting the guidance of vaccination policies and regulatory decisions (28).

Hypothesis:

Higher levels of antibodies or CoPs are associated with reduced odds of infection (rotavirus, norovirus, RSV, and shigella) among under-2 children in Dhaka, Bangladesh

Study Objectives:

Primary Objective:

- To identify the correlates of protection (CoPs) that are associated with reduced odds of infection in case of Rotavirus, Norovirus, RSV and Shigella infection among under-2 children residing in Dhaka, Bangladesh

Secondary Objective:

- To identify re-infection rates of pathogens of interest and the associated factors related to re-infection among children
- To evaluate the difference in maternal antibody titers among children across various timepoints
- To explore the difference in IgM and IgG antibodies against pathogens across various timepoints
- To investigate the difference in mucosal secretory IgA across various timepoints among re-infected children
- To evaluate incidences of morbidity (diarrhoea, cough, fever, etc.) in the birth cohort among children

Study Design:

This study will adopt a prospective birth cohort design, within which we will implement a nested case-control framework for each targeted pathogen (Rotavirus, Norovirus, RSV and Shigella) to identify correlates of protection (COPs) associated with a reduced risk of disease.

Study Settings:

The participants of our study will be recruited from Bauniabadh and its surrounding areas in the Mirpur subdistrict of Dhaka, Bangladesh. Mirpur represents a densely populated urban area which is situated approximately 7 to 8 kilometers northwest of the icddr,b. This site was purposively selected based on several considerations. Firstly, it encompasses a large proportion of poor and middle-class families. Also, Mirpur's environmental, residential, and sanitary conditions reflect those of typical overcrowded urban settlements in South Asia. Along with these advantages, the investigative team has an established infrastructure in the area with currently ongoing longitudinal research collaborations.

The average household size in this area is approximately 4.5 individuals, with women comprising about 48% of the population. Nearly one-fifth of households live on a monthly income below USD $110, approximately 30% of mothers have received no formal education, and only 3% have attained secondary-level schooling. The occupation of the residents is dominated by day laborers, garment factory workers, and similar low-income jobs. These employment patterns, coupled with low educational attainment, contribute to persistent cycles of poverty and marginalization. It covers only 14.22 square kilometres yet is home to more than half a million people, resulting in an amazing population density of almost 50,000 people per square kilometre. This value well above the averages for the Dhaka district (8,229 persons/km^2) and Bangladesh as a whole (976 persons/km^2). This degree of human concentration brings with it a host of urgent public health concerns, which is precisely why the area serves as such a vital ground for community-based health research (29).

Crucially, icddr,b possesses unparalleled institutional experience and logistical infrastructure in Mirpur, having conducted multiple large-scale longitudinal and intervention studies in this locality over the past two decades. Landmark investigations such as the BEED (Bangladesh Environmental Enteric Dysfunction) study, the MAL-ED (Malnutrition and Enteric Disease) cohort, and the MDCF (Microbiota-Directed

Complementary Food) trial have not only established icddr,b's scientific credibility but also facilitated the development of an extensive surveillance network, trained field personnel, and community rapport essential for sustained cohort engagement and high follow-up rates.

These prior studies have yielded globally influential findings on the microbiological, nutritional, and environmental determinants of child health, and have demonstrated icddr,b's capacity to integrate sophisticated biomarker assessments, dietary interventions, and household-level behavioral data collection within a challenging urban context. Importantly, icddr,b has maintained a longstanding collaboration with local authorities and community leaders in Mirpur, fostering a climate of trust and facilitating informed consent, ethical oversight, and community compliance.

Moreover, the geographic proximity of Bauniabadh to icddr,b's central laboratory and clinical facilities in Mohakhali, approximately 7-8 kilometers, enables timely transportation of biospecimens and rapid clinical referrals, thereby ensuring the scientific integrity and clinical safety of the cohort.

In summary, Bauniabadh, Mirpur offers a fertile and operationally feasible environment for conducting a birth cohort study that aspires to elucidate the complex interplay between environmental exposures, nutrition, infection, and developmental outcomes in early childhood. The convergence of population-level vulnerability, prior research infrastructure, and institutional expertise positions this site as uniquely suited to advance the aims of a comprehensive longitudinal investigation spearheaded by icddr,b.

Study Workflow:

Households at Baunibadh, Mirpur and its adjacent areas will be regularly screened for pregnant mothers by the field workers. Caregivers/mothers who are willing to participate in the study will be followed up to the birth of the child. The child will be enrolled at birth and will be in the cohort study for two years. During this two-year period routine follow-up and sample collection will be done on specific time intervals (at birth, 3 months, 6 months, 9 months, 12 months, 18 months and 24 months). Blood and stool samples will be collected at each time point. Sociodemographic, clinical, and anthropometric data will be collected at enrolment. Regular clinical and anthropometric data will be collected during each follow-up. Food Frequency Questionnaire (FFQ) will be used with a particular focus on breast milk intake. Any child with congenital anomalies will be excluded from the study.

In order to identify any event of illness within 24 hours, mothers will be instructed to contact the field staff immediately if her child shows any symptoms (fever, diarrhoea, cough). Regular phone calls will be also made to contact the mothers. The study physician will examine the child thoroughly to diagnose and provide necessary management. During the illness serum, stool or nasopharyngeal swabs will be collected from the child. Another additional blood sample will be collected two weeks after each sickness to properly investigate associated CoPs.

The children who will get infected first time with our pathogen of interest (rotavirus, norovirus, RSV, and shigella) will be monitored to ascertain any second time reinfection. The children who gets re-infected a second time among this sub-cohort will be considered as cases and who does not get re-infected a second time will be our controls. Using these cases and controls we will conduct a nested case control analysis to find out what levels of CoPs were responsible for providing immunity to the children that did not get re-

infected. Here the CoPs that formed in the body after first time infection will be considered as exposure. Each child will be followed up to 1.5 years to see whether he suffers from a first-time infection. Three separate analytical models will be formulated for each pathogen of interest. A child will be defined as either case or control within 6 months, within 12 months and within 18 months of getting his first infection.

The Dhaka Hospital, icddr,b is the largest diarrhoeal disease hospital in the world, which has been in operation since 1962. Each year, approximately 200,000 patients visit the hospital. Any child in the cohort can be brought to Dhaka Hospital immediately for expert opinion or hospitalization if needed which can be considered an additional support to the study.

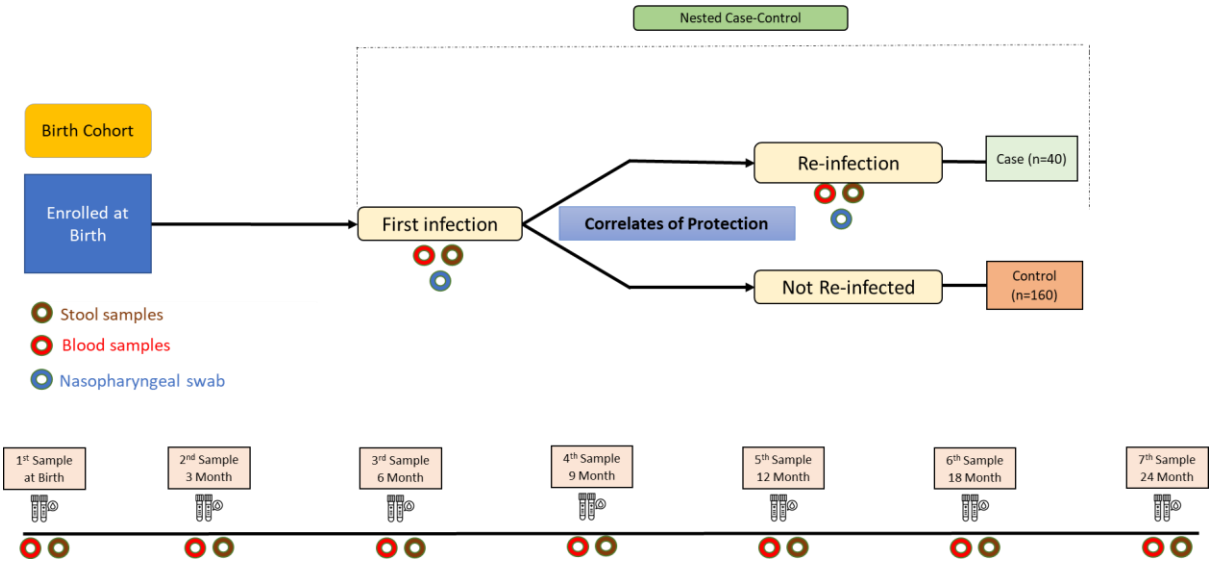

Figure 01: Study workflow

Sample Storage and Laboratory Capacity:

At the International Centre for Diarrhoeal Research, Bangladesh (icddr,b) facility in Dhaka, there are five major laboratory wings, namely: 1) Clinical Laboratories wing, hosting five separate laboratory facilities; 2) Research Laboratories wing, hosting eight individual laboratories ; 3) Environmental Health Laboratory; 4) Food and One Health Laboratory and 5) Genomic Centre containing at icddr,b main campus in Dhaka with the following key equipment: water distillations unit, deep freezer systems, light microscopes, PCR machines, microcentrifuges, vortexes, scales, CO2 / room air incubators, flow cytometry, HPLC, LC-MS, Luminex, ELISA plate readers, microbial culture and detection, VITEK-2, VITEK-MS, 16s rRNA sequencing, Shotgun metagenomics, and whole genome sequencing facilities. icddr,b also has autoclaves to achieve proper sterilization standards. Moreover, there is a central biorepository system with more than 50 pieces of -80°C freezers, more than 25 pieces of -20°C freezers, around 15 pieces of 2-8°C refrigerators, and a temperature monitoring system that sends text alerts to the laboratory team when a temperature excursion occurs. The project staff will maintain two pieces -80°C freezers to preserve the biological samples. This study team will also perpetuate medium-sized and small-sized liquid nitrogen dry shippers along with liquid nitrogen storage tank to transport and preserve stool specimens. The icddr,b team has

well-trained staff renowned for their previous tremendous efforts in lots of clinical trials and cohort studies in collaboration with a number of collaborative universities and institutes worldwide.

Sample size estimation:

In this nested case-control study within a birth cohort, children will be followed prospectively from birth to identify those who experience a documented first infection with a specific pathogen, forming the pathogen-specific sub-cohort. Within this sub-cohort, reinfection serves as the outcome of interest. Children who develop reinfection are designated as cases, while controls are drawn from those who had a first infection but remained reinfection-free at the time the case occurred. Cases and controls will be selected after fixed time intervals from the first infection, ensuring they are at same risk for reinfection during the same period. CoP levels measured during or after the first infection constitute the primary exposure.

1. Rotavirus:

A meta-analysis from Bangladesh reported the rotavirus prevalence rate to be 30% (30). Rotavirus also has an extremely high reinfection rate of 70% (31). So, in order to obtain 40 reinfection cases in this sub-cohort, approximately (40/0.7) = 57 individuals will be needed. However, to achieve a 1:4 ratio for case vs control, we will consider 200 (40+160) children in this sub-cohort. To achieve this sub-cohort size, given that 30% of the birth cohort experiences a first Rotavirus infection, the pathogen-specific cohort would need to be around (200/0.3) = 667 children. Nearly every child is infected with rotavirus at least once before the age of three (32). If we consider that during our two-year follow-up, at least 67% children will get infected with rotavirus, the final cohort size is (667/0.67) = 995 children.

2. Norovirus:

The MAL-ED birth cohort study from Bangladesh reported Norovirus prevalence of 20.4% (15). The study also reported a 46% reinfection rate (15). To get 40 reinfection cases, we would theoretically need a Norovirus-specific sub-cohort of approximately (40/0.46) = 87 children. We will consider at least 200 children in the Norovirus-specific sub-cohort to maintain a case vs control ratio of 1:4. Given that 20.4% of the birth cohort experiences a first Norovirus infection, the initial birth cohort would need to be around (200/0.204) = 980 children.

3. Respiratory Syncytial Virus (RSV):

For RSV, the average first infection prevalence rate in Bangladesh is 27.6% and from another study, we found that the reinfection rate of RSV is 42% (26, 33). To obtain 40 reinfection cases, we would need approximately (40/0.42) = 95 individuals in the RSV-specific sub-cohort. We aim to select 40 cases and 160 controls. We expect nearly all children in the birth cohort will experience at least one respiratory infection in their first two years of life (34, 35). Given that 27.6% of the birth cohort experiences a first RSV infection, the initial birth cohort would need to be around (200/0.276) = 725 children.

4. Shigella:

From the MAL-ED birth cohort study conducted in Mirpur, Bangladesh, it was reported that 75.5% of children were infected with Shigella between 0 and 2 years of age (36). The approximate reinfection rate for *Shigella* is 10% (37). To secure 40 reinfection cases, we'd require approximately (40/0.10) = 400 individuals in the Shigella-specific sub-cohort. So, the initial birth cohort would need to be around (400/0.755) = 530 children. Considering the highest birth cohort size of 995 participants for rotaviral infection and a 20% attrition rate, we will enroll 1250 children.