Nonsterile immunity to cryptosporidiosis in infants is associated with mucosal IgA against the sporozoite and protection from malnutrition

Mamun Kabir¹, Masud Alam¹, Uma Nayak²,³, Tuahir Arju¹, Biplob Hossain¹, Rubaiya Tarannum¹, Amena Khatun¹, Jennifer A. White⁴, Jennie Z. Ma²,⁴, Rashidul Haque¹, William A. Petri Jr⁴,⁵,⁶* and Carol A. Gilchrist⁴

¹ International Centre for Diarrhoeal Diseases Research, Bangladesh,
² Department of Public Health Sciences, University of Virginia, Charlottesville, Virginia, United States of America
³ Center for Public Health Genomics, University of Virginia, Charlottesville, Virginia, United States of America,
⁴ Department of Medicine, University of Virginia, Charlottesville, Virginia, United States of America
⁵ Department of Microbiology, Immunology and Cancer Biology, University of Virginia, Charlottesville, Virginia, United States of America
⁶ Department of Pathology, University of Virginia, Charlottesville, Virginia, United States of America

Key words: Bangladesh; cohort study; cryptosporidiosis; children; sub-clinical, malnourished

*Corresponding Author

E-mail: wap3g@virginia.edu (WAP)
Abstract

We conducted a longitudinal study of cryptosporidiosis from birth to three years of age in an urban slum of Dhaka Bangladesh. Fecal DNA was extracted from monthly surveillance samples and diarrheal stool samples collected from 392 infants from birth to three years. A pan-Cryptosporidium qPCR assay was used to identify sub-clinical and symptomatic cryptosporidiosis. Anthropometric measurements were collected quarterly to assess child nutritional status. 31% (121/392) of children experienced a single and 57% (222/392) multiple infections with Cryptosporidium. Repeat infections had a lower burden of parasites in the stool (Cq slope = -1.85; p<0.0001) and were more likely to be sub-clinical (Chi square test for trend; p=0.01). Repeat infections were associated with the development of growth faltering (Pearson correlation = -0.18; p=0.0004). High levels of fecal IgA antibodies against the Cryptosporidium Cp23 sporozoite protein at one year of life were associated with a delay in reinfection and amelioration of growth faltering through three years of life (HAZ IgA high responders -1.323 ± 0.932 versus HAZ -1.731 ± 0.984 p=0.0001). We concluded that nonsterile immunity to cryptosporidiosis in young children was associated with high levels of mucosal IgA anti-Cp23 and protection from diarrhea and growth faltering.

Authors Summary

Cryptosporidium is one of the top causes of diarrhea and growth faltering in Bangladesh infants. We discovered that a prior infection resulted in incomplete immunity that protected from diarrhea and growth faltering but not infection and was associated with mucosal IgA against a sporozoite surface protein Cp23. The most important implication of these findings is that a cryptosporidiosis vaccine may not need to achieve complete protection from infection to have a beneficial impact on child health.
Introduction

Cryptosporidium spp. parasites are leading causes of diarrheal disease in infants living in low and middle income countries [1–4]. They are additionally a cause of water-borne outbreaks of diarrhea in high income countries and of chronic diarrhea in people living with HIV infection [5]. There is no vaccine and development will require an understanding of the natural history of cryptosporidiosis [3,6–12]. To this end, a community-based prospective cohort study of cryptosporidiosis was begun in 2014 [13]. The study subjects were born in an urban slum in Dhaka, Bangladesh and enrolled during the first week of life [13–15]. In humans and in animal models vaccination or prior infection resulted in partial protection against reinfection [16–19]. For example we observed that high fecal IgA against the sporozoite protein Cp23 delayed but did not prevent a repeat infection with Cryptosporidium spp. [20,21]. Ajjampur et al observed a decrease in the incidence of diarrhea in reinfected children [22]. In contrast Kattula et al found that while the reinfection frequency was decreased the proportion of symptomatic disease was unchanged [9]. In human volunteer studies second infections were associated with reduced parasite burden and less severe diarrhea [23].

In addition to diarrheal disease cryptosporidiosis is associated with development of malnutrition [8,24–27]. Here we report the natural history of cryptosporidiosis from a longitudinal study of urban slum children from birth through three years of age in Dhaka, Bangladesh, demonstrating that immunity is characterized by protection from diarrhea and growth faltering.
Results

Five hundred children were enrolled within the first week of birth, and of these 392 completed three years of observation. Stool samples were collected monthly and at the time of diarrhea. Successful sample collection and qPCR testing was completed for 96% of monthly surveillance time points and for 84% of the diarrheal cases (Fig 1; S1 Table 1). There were 1336 Cryptosporidium positive samples for analysis by year 3 (Fig 1). Six hundred and ninety eight events met the definition of separate Cryptosporidium infections in the 392 children (Table 1). Of the 698 infections experienced by the 392 infants retained in the study at 3 years of age, 167 were diarrheal and 531 sub-clinical cryptosporidiosis (Table 1). The Cq (cycle of quantification) value of the stool sample in which the parasite was first detected was used as an index of parasite burden.

Fig 1. COHORT diagram. Study subjects, collected samples and new infection numbers

Abbreviations: RT-QPCR quantitative polymerase chain reaction; DS: Diarrheal samples; FU: Follow-up; MS: Monthly Samples, Cq: Cycle Quantification

Table 1 Frequency of Diarrheal Cryptosporidiosis in Repeated Infections

<table>
<thead>
<tr>
<th>Cryptosporidium Infection</th>
<th>Number of infections</th>
<th>Age in days Mean ± SD</th>
<th>Infection phenotype* Mean</th>
<th>Diarrhea Sub-clinical Mean</th>
<th>Diarrhea Frequency** Upper limit</th>
<th>Diarrhea Frequency** Lower limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>343</td>
<td>519 ± 250 [15-1088]</td>
<td>96</td>
<td>247</td>
<td>0.28</td>
<td>0.33 0.23</td>
</tr>
<tr>
<td>2nd</td>
<td>222</td>
<td>758 ± 209 [273-1079]</td>
<td>46</td>
<td>176</td>
<td>0.21</td>
<td>0.27 0.16</td>
</tr>
<tr>
<td>3rd</td>
<td>101</td>
<td>892 ±155 [399-1084]</td>
<td>17</td>
<td>84</td>
<td>0.17</td>
<td>0.26 0.10</td>
</tr>
<tr>
<td>4th</td>
<td>28</td>
<td>939± 129 [639-1095]</td>
<td>6</td>
<td>22</td>
<td>0.21</td>
<td>0.41 0.08</td>
</tr>
</tbody>
</table>
Table 2. Distribution and Clinical Characterization of repeated Cryptosporidium infections

<table>
<thead>
<tr>
<th># of Infections</th>
<th># of Children</th>
<th>Cryptosporidium infections</th>
<th>Total # of infections</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1st DS*</td>
<td>2nd MS**</td>
</tr>
<tr>
<td>0</td>
<td>49</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>121</td>
<td>28</td>
<td>93</td>
</tr>
<tr>
<td>2</td>
<td>121</td>
<td>34</td>
<td>87</td>
</tr>
<tr>
<td>3</td>
<td>73</td>
<td>24</td>
<td>49</td>
</tr>
<tr>
<td>4</td>
<td>25</td>
<td>8</td>
<td>17</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>392</td>
<td>96</td>
<td>247</td>
</tr>
</tbody>
</table>
A) Correlation between parasite burden and the number of *Cryptosporidium* infections. Each symbol represents the first detectable sample of an individual infection. Y-axis represents the quantitative cycle (Cq) of the diagnostic pan-*Cryptosporidium* PCR assay. X-axis shows the number of *Cryptosporidium* infections that had occurred in this child. The line represents the slope (-1.85 ± 0.21) and Y-intercept (26.95 ± 0.44) estimated from the GEE model with the exchangeable correlation structure (p<0.0001)

B) Comparison of single infections (black symbol) with those that are part of a series (gray symbol). Bar graph (indicating data mean ± standard deviation) with individual data points. Each symbol on the box plot represents the first positive sample of an individual infection. X-axis refers to Infection number and, if the first infection, whether a second *Cryptosporidium* infection took place in the 3 years of life. Y-axis represents the quantitative cycle (Cq) of the diagnostic pan-*Cryptosporidium* PCR assay. Horizontal bars represent the result of a non-parametric Kruskal-Wallis test **** indicates p<0.0001

To investigate if an initial high burden infection provided better protection against future infections with the *Cryptosporidium* parasite, we compared the Cq values of the infections in children who only had one infection in the first three years of life vs the Cq values of the first infection in children that had repeated infections (Fig 2B). The mean Cq values were similar in both cases (single infections: Cq 27.6 ± 5.4; 1st infection of multiples: 28.8 ± 6.0) and significantly lower than that in subsequent second infections where infections were >1 (Cq second infection: 31.5 ± 4.9).

We next evaluated if the lower parasite burden in repeat infections was influenced by the age of the children. As most recurrent *Cryptosporidium* infections occurred in older children, (Table 1) we analyzed a subset of the *Cryptosporidium* positive samples corresponding to the first to fifth
infections in children aged between 2.25 and 2.75 years. The parasite burden measured by qPCR remained significantly lower in the recurrent infections (Fig 3). The negative relationship of lower parasite burden with repeated infections was not an artefact of PCR inhibitors in the stool of older children because detection of the Phocine herpesvirus (PhHV) DNA included as internal extraction control [29] was not significantly affected by the number of prior Cryptosporidium infections. We concluded that repeat infections had a lower parasite burden.

**Fig 3 Parasite burden in older children**

The amount of parasite in stool was determined as a function of the number of Cryptosporidium infections in a child by linear regression. The analysis was restricted to children between 2.25 and 2.75 years of age (n=140). Each symbol represents the first detectable sample of an individual infection. Y-axis, quantitative cycle of the diagnostic pan-Cryptosporidium PCR assay (Cq). X-axis, the number *Cryptosporidium* infections that had occurred in each child. Slope: -1.65, R squared value: 0.01125, Significance p<0.0001.

The duration of diarrheal disease was similar in the first infection and later reinfections (single infection: 4.8 ± 2.9 days; primary infection: 5.5 ± 3.9 days; later infections: 5.2 ± 3.6 days), however, the proportion of the diarrhea-associated *Cryptosporidium* infections decreased in the recurrent infections (Chi-squared test for trend p=0.011) (Table 1). We concluded that the repeated *Cryptosporidium* infections were more likely to be sub-clinical.

**Cryptosporidium and growth faltering**

Growth faltering (low height for age; HAZ score) was analyzed from the 3 year old children based on the number of *Cryptosporidium spp.* infections [0 - 3 years] (both diarrhea and sub-
clinical) (Table 2, Fig S5 & 6). The association between cryptosporidiosis and HAZ score at three years was examined by multiple regression in order to account for the effect of the confounding variables previously identified (Table 3) [13,21].

Table 3 Regression Analysis using selected predictors to test the association of Cryptosporidiosis with Height-for-Age Scores at 3 Years

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Effect (95% Confidence Interval)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cryptosporidium Infections</td>
<td>-0.12 (-0.197, -0.043)</td>
<td>0.0024</td>
</tr>
<tr>
<td>Child LAZ at Birth</td>
<td>0.252 (0.159, 0.345)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Maternal Weight</td>
<td>0.017 (0.007, 0.027)</td>
<td>0.0011</td>
</tr>
<tr>
<td>Maternal Height</td>
<td>0.037 (0.018, 0.055)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Maternal Education</td>
<td>0.237 (0.027, 0.448)</td>
<td>0.0273</td>
</tr>
<tr>
<td>Household income</td>
<td>0.001 (0.000, 0.002)</td>
<td>0.1527</td>
</tr>
<tr>
<td>Treated water</td>
<td>0.163 (-0.045, 0.371)</td>
<td>0.1234</td>
</tr>
</tbody>
</table>

Cryptosporidium infection was negatively associated with the HAZ score at 3 years after adjusting for birth length-for-age (LAZ) score and maternal weight and education: each Cryptosporidium infection reoccurrence resulted in a decrease in HAZ score (Δ 0.12) at 3 years (Table 5). No significant relation was found between malnutrition at birth (LAZ score) and total number of Cryptosporidium infections during the follow-up (Fig 4A). However, the total number of Cryptosporidium infections was negatively associated with HAZ score at 3 years (Fig 4B, regression coef=-0.152, p=0.0004) The association between HAZ and Cryptosporidium spp. infections was unaffected by whether the event was a sub-clinical infection or diarrheal disease (Fig S7).

Fig 4 Cryptosporidiosis frequency was associated with growth faltering distinct from the impact of birth nutritional status

A) Relationship between length for age z score (LAZ) at birth (Y-axis) and the total number of Cryptosporidium spp. infections (X-axis). The slope was not significantly different from one.
B) Relationship between height for age z score (HAZ) at 3 years (Y-axis) and the total number of Cryptosporidium spp. infections (X-axis). Slope: -0.152 ± 0.0429, R squared value: 0.0313, significance p=0.0004. Children were defined to be at risk for growth faltering with a LAZ or HAZ score <-1 and malnourished at LAZ or HAZ score <-2. Orange box: birth LAZ or 3 year HAZ score -1 to -2; red box: birth LAZ or 3 year HAZ or score < -2.

Other measurements that are used as indicators of malnutrition were also significantly associated with the number of Cryptosporidium infections. These included mid-upper arm circumference (MUAC) (Fig S8A; MUACZ vs. number of Cryptosporidium infections slope: -0.088; p=0.0123 and weight-for-age (WAZ score) (Fig S8B) (linear regression analysis slope: -0.115 p=0.0093). However, neither BAZ (body-mass- for- age) (Fig S8C), used to measure acute protein-energy malnutrition or wasting (WHZ) were affected by a history of Cryptosporidium infections (Fig S8D).

The Pearson correlations among the number of Cryptosporidium infections, LAZ at birth, diarrheal episodes and HAZ at year 3 are shown in Fig 5A (Cryptosporidium infections: HAZ at year 3: coef = -0.18, p=0.024; Cryptosporidium infections: diarrheal episodes captured (all causes): coef = 0.22, p>0.0001; HAZ at year 3: LAZ at birth: coef = 0.28, p= 0.008). As expected, a significant correlation existed between LAZ at birth and HAZ at year 3 (simple linear regression p<0.0001 Fig 5B).

Enteric pathogens are endemic in the Bangladesh study population [30] and as a consequence, infants enrolled in the study cohort had repeated diarrheal episodes of which only some were associated with infection with the Cryptosporidium parasite. However, while Cryptosporidium infections (diarrheal and sub-clinical) were significantly associated with child HAZ at year 3 (Pearson’s correlation p=0.0004), the number of all-cause diarrheal episodes was not (Fig 5C; S2 Table 3). This result supported our conclusion that this growth shortfall was specifically associated with recurrent cryptosporidiosis.
**Fig 5 Correlates of cryptosporidiosis-associated growth-faltering**

A) Correlation matrix of cryptosporidiosis, all-cause diarrhea, LAZ at birth and HAZ at 3 years, calculated using Pearson r. Bar on the right indicates strength and direction of association. B) Comparison of three year-HAZ with birth LAZ. (Slope: -0.294 ± 0.05; R squared value: 0.08; Significance p < 0.0001). C) Relationship of all-cause diarrhea with HAZ at 3 years of age (p = NS).

**Mucosal IgA against the sporozoite Cp23 protein was associated with protection from growth faltering**

In previous work it was shown in this cohort that a high level (> mean value) of fecal anti-Cp23 IgA at one year of age was associated with an increased resistance to cryptosporidiosis through age three [14,20]. Here we additionally discovered that children with high levels (upper 50th percentile) of fecal anti-Cp23 IgA at one year of age were protected from growth faltering through year 3 (Fig 6). Subgrouping the children into Group 2a (never infected by evidence of anti-Cp23 IgA levels and diagnostic qPCR assays; n=20) and Group 2b (diagnostic qPCR positive only; n=185) versus Group 1 children with high levels of IgA at one year (n=171) did not alter the association with growth faltering (Fig S9A). Analysis of the fecal IgA antibodies against a second sporozoite peptide (Cp17) was also performed. Although a similar trend was observed the difference in year 3 HAZ was not significantly different (Fig S9B). A high level of fecal anti-Cp23 at one year was not associated with any drop in the parasite burden at the next Cryptosporidium infection (first subsequent new infection: Cq of IgA high responders 28.2 ± 5.6 versus 27.6 ± 5.6 p=0.43).

**Fig 6 High anti-Cp23 IgA levels were associated with a reduction in cryptosporidiosis-associated growth-faltering.**
Group 1 and 2 children were in the upper and lower 50th percentile for fecal IgA anti-Cp23 respectively. HAZ are shown for children in both year one and year three of life. Mean ± standard deviation with individual data points. Horizontal bars represent the result of a non-parametric Kruskal-Wallis test ***p<0.001, **p<0.01

Discussion

The key finding of this paper is that naturally acquired immunity protects from Cryptosporidium diarrhea but does not provide sterilizing immunity. The importance of this observation is two-fold: first it indicates that transmission likely occurs in semi-immune populations; and second that continued sub-clinical infections increase the risk of infection-related growth faltering. Encouragingly however, acquired immunity associated with high levels of mucosal IgA against the Cp23 cryptosporidium sporozoite antigen were associated with protection from malnutrition. Many previous studies on cryptosporidiosis have focused on the health impact of diarrhea-associated cryptosporidiosis [7,12,31–34]. However sub-clinical disease, as opposed to infection accompanied by diarrhea, may also have long term effects on child health. The link between sub-clinical cryptosporidiosis and malnutrition is now well known if not yet well understood [8,9,13,32]. In a recent study the global prevalence of cryptosporidiosis in people without diarrheal symptoms was 4.4% (95% confidence interval 2.9 - 6.3)[35]. During the 3 years of this study 212 children (54%) had only sub-clinical Cryptosporidium infections. This longitudinal study allowed us to take an in depth look at the role of sub-clinical reinfections in the exacerbation of growth faltering [13,22,25].

Anthropometric measurements are reliable non-invasive methods to monitor child malnutrition. The most commonly used metrics are a shortfall in child growth (low height for age: HAZ score) a consequence of chronic undernutrition and wasting (exemplified by a low weight for height:
In line with most studies our results show that a history of cryptosporidiosis was associated with a decrease in the HAZ score of children irrespective of infection severity [8,26,36,37]. Here we found that child growth was negatively impacted not only by the first episode of cryptosporidiosis, but both occurred and remained constant in succeeding infections, even though parasite burden and diarrheal disease decreased. This study has, therefore, shown that naturally acquired partial immunity was not effective at preventing growth faltering and that a control strategy focused on only preventing diarrheal cryptosporidiosis may not prevent the stunted growth associated with cryptosporidiosis.

A limitation of the current study is that it was not possible to unambiguously attribute an episode of diarrhea to Cryptosporidium because children in this community were infected with multiple enteropathogens at the same time [26,30]. To mitigate the problem of correctly identifying Cryptosporidium-associated diarrheal infections these were defined as an episode of diarrhea accompanied by a new Cryptosporidium infection (i.e. the immediately preceding surveillance or diarrheal stool sample was negative for Cryptosporidium) [13]. A second limitation was that surveillance stool samples were collected at only monthly intervals which likely missed some subclinical infections, potentially underestimating the impact of cryptosporidiosis on child growth. The study however had notable strengths including most importantly its longitudinal design that combined collection of surveillance and clinical specimens with studies on child growth faltering.

The association of mucosal immunity to Cp23 with protection from growth faltering offers hope that a cryptosporidiosis vaccine could have a measurable impact on child health, even in the absence of absolute protection from infection.
Acknowledgement

We are grateful for the participation of the parents and children in this study as well as the staff of the Emerging Infectious Diseases Division of icddr,b for contributing to this research.

Author Contributions

CG, MK, RH and WP designed, and RH, WP, MA, JAW, CG and MK drafted the study protocols and WP directed the work described in this paper. MK, TA, BH, RT and AK acquired the data used in this manuscript. UN curated the study data. UN, JM, MK and CG analyzed the study data. CG wrote the first draft of the manuscript that was reviewed and amended by all the authors who also approved the final manuscript.

Methods

Child cohort

A total of 500 children were enrolled within one week of birth in an urban slum of Dhaka, Bangladesh beginning in June 2014 through March 2016 and were monitored for diarrheal diseases through bi-weekly home visits by trained field investigators. A monthly stool sample was also collected to evaluate asymptomatic infection and growth was measured every 3 months (“Cryptosporidiosis and Enteropathogens in Bangladesh”; ClinicalTrials.gov identifier NCT02764918). This area (Section 11 of Mirpur Thana) is densely populated with participants in this study having an average of 5.5 people living in 1.6 rooms. Annual median household income of participants was 14,000 Taka or approximately US $164 (Table 4). Anthropometric data was collected as previously described [13]. Each child was weighed on an electronic scale (kilograms, measured with electronic scale; TANITA, HD-314). Child height or length (depending on age) and mid-upper arm circumference were measured to the nearest 0.1 cm using a measuring board and plastic tape (Table 5). The height-for-age z score (HAZ); weight
for age z score (WAZ); weight for height (WHZ); body mass index for age (BAZ); and mid-upper arm circumference for age (MUACZ) were calculated using the World Health Organization Anthro software (version 3.2.2) [13]. Children who had a HAZ score < -1 were defined as ‘at risk for malnutrition’ and HAZ < -2 as malnourished [27,28]. Diarrhea was defined as ≥ 3 loose stools within a 24-hour period as reported by the child’s caregiver with episodes separated by a gap of at least 3 days. This paper reports the data from 392 infants who were followed through three years of age.

### Table 4 Maternal and family demographics

<table>
<thead>
<tr>
<th>Maternal and Family Characteristics</th>
<th>N=392</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Maternal Age, year (SD)</td>
<td>24.63 (4.68)</td>
</tr>
<tr>
<td>Mean Maternal Weight, Kg (SD)</td>
<td>51.82 (10.17)</td>
</tr>
<tr>
<td>Mean Maternal Height, Cm (SD)</td>
<td>149.71 (5.03)</td>
</tr>
<tr>
<td>Mean Maternal BMI, kg/m2 (SD)</td>
<td>23.10 (4.32)</td>
</tr>
<tr>
<td>No Maternal Education, N (%)</td>
<td>87 (22.2)</td>
</tr>
<tr>
<td>Median Household income (BDT*) (IQR)</td>
<td>14,000 (10,000)</td>
</tr>
<tr>
<td>Treated water (Boil), N (%)</td>
<td>390 (74.0)</td>
</tr>
</tbody>
</table>

**Abbreviations**: SD, Standard Deviation; BDT, Bangladesh Taka; IQR, Inter Quartile Range

*1000 BDT is approximately 12 US dollars

### Table 5 Infant demographic characteristics

<table>
<thead>
<tr>
<th>Infant Demographic Characteristics</th>
<th>Year 2 N=421</th>
<th>Year 3 N=392</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender, Female N (%)</td>
<td>229 (54.4)</td>
<td>214 (54.6)</td>
</tr>
<tr>
<td>Mean Infant Age in days, Range</td>
<td>733 (719 - 783)</td>
<td>1099 (1035 - 1134)</td>
</tr>
<tr>
<td>Mean Weight, Kg, (SD)</td>
<td>10.17 (1.31)</td>
<td>11.93 (1.52)</td>
</tr>
<tr>
<td>Mean Height, Cm, (SD)</td>
<td>81.72 (3.18)</td>
<td>89.78 (3.74)</td>
</tr>
<tr>
<td>Mean MUAC, Cm, (SD)</td>
<td>14.87 (1.02)</td>
<td>15.3 (1.00)</td>
</tr>
<tr>
<td>Mean WAZ, (SD)</td>
<td>-1.34 (1.05)</td>
<td>-1.42 (0.99)</td>
</tr>
<tr>
<td>Mean HAZ , (SD)</td>
<td>-1.56 (0.98)</td>
<td>-1.54 (0.98)</td>
</tr>
<tr>
<td>Mean MUACZ , (SD)</td>
<td>-0.15 (0.85)</td>
<td>-0.34 (0.79)</td>
</tr>
</tbody>
</table>
Mean BAZ, (SD) | -0.52 (0.98) | -0.63 (0.91)

| Abbreviations:  | SD, Standard Deviation, MUAC, Mid Upper Arm Circumference, WAZ, Weight-for-age, HAZ, Height-for-age, MUACZ, Mid Upper Arm Circumference-for-age, BAZ, body mass index-for-age |

**Ethics Statement**

The study was approved by the Ethical and Research Review Committees of the International Centre for Diarrhoeal Disease Research, Bangladesh (PR-13092) and by the Institutional Review Board of the University of Virginia (IRB#20388). Informed written consent was obtained from the parents or guardians for the participation of the subjects in the study.

**Sampling and specimen testing**

Fresh stool samples collected in the field were placed on ice and then brought to the lab on the same day and frozen within 6 h of collection (Fig S1). Stool specimens were collected from children every month (monthly surveillance) and during episodes of diarrhea. A modified Qiagen stool DNA extraction protocol with 95°C incubation and a 3-minutes bead-beating step was used to extract DNA [13] (Fig S2). These samples were tested with a multiplex qPCR assay which utilizes pan-Cryptosporidium primers and probes targeting the 18S rDNA gene and primers and probes to detect the Phocine herpesvirus (PhHV) extraction control (obtained from the European Virus Archive Global organization) as previously described (Fig S3). All samples with a cycle threshold of ≤ 40 for cryptosporidium were used in this analysis [13]. In year 3 the diagnostic qPCR assay was not able to be completed on 0.9% of the collected diarrheal and 4.6% of the monthly surveillance samples (Table S1).

Infection with *Cryptosporidium* was defined as detection of *Cryptosporidium* DNA by qPCR from stool. PCR-positive samples were classified as a separate infection if occurring greater than 65 days after the preceding positive sample [13]. The *Cryptosporidium* infection phenotype
(diarrheal or sub-clinical) was based upon symptoms at the time of detection of the first
*Cryptosporidium* - positive stool sample, whether diarrheal stool or monthly surveillance.

**Statistical analysis**

Descriptive statistics were expressed in mean ± standard deviation for continuous variables and as frequencies and proportions for categorical variables. The frequency of repeated *Cryptosporidium* infections in the first 3 years of life was summarized for diarrhea and sub-clinical infections separately and their differences were evaluated with the χ² test. To account for within-child correlations among repeated *Cryptosporidium* infections, the relationship between parasite burden and the number of repeated *Cryptosporidium* infections was evaluated using the Generalized Estimating Equation (GEE) for repeated measurements, assuming an exchangeable correlation structure. Pearson correlation was calculated for univariate association of individual predictors with HAZ at 3 years. Since confounders such as LAZ at birth, maternal weight and height, maternal education, household income and access to treated water were previously shown to impact HAZ [8,13], a multivariable linear regression was performed to evaluate the association between *Cryptosporidium* infection and HAZ at 3 years after adjusting for these factors (Table 3). Similarly, a multiple regression analysis was performed to independently evaluate whether the number of episodes of diarrhea, irrespective of the causative pathogen, was associated with HAZ at 3 years. Analyses were performed using both the GraphPad Prism version 8.4.3 for Mac, (GraphPad Software, San Diego, California USA, ), SAS 9.4 (Raleigh, NC) and R version 3.3.3, 32-bit.
This work was supported by grants NIH R01-043596 to WP and CG, R21-142656 to CG and the Bill & Melinda Gates Foundation OPP1100514 to WP. The governments of Bangladesh, Canada, Sweden, and the UK provide core support to icddr,b. The funders had no role in study design, data collection and analysis or decision to submit for publication.
References


16. Lemieux MW, Sonzogni-Desautels K, Ndao M. Lessons learned from protective immune responses to optimize vaccines against cryptosporidiosis [Internet]. Pathogens. MDPI AG; 2018. doi:10.3390/pathogens7010002


Supplemental Figures and Tables

Table S1. Symptomatic and asymptomatic samples collected during year 3

Table S2. Multivariable analysis of total all cause diarrhea and HAZ at year 3

Fig S1. Flow chart of stool processing and molecular testing

Fig S2. Flow chart of stool TNA extraction procedure using QIAamp Fast DNA Stool Mini Kit from fresh or frozen stool samples.

Fig S3. Flow chart of Multiplex qPCR of Cryptosporidium, Giardia, Entameoba histolytica by targeting the 18S gene

Fig S4. Parasite Burden in diarrheal and sub-clinical infections
Relationship between Parasite Burden and the number of recurrent Cryptosporidium infections. Each symbol represents the first detectable sample of an individual infection. Y-axis, quantitative cycle of the diagnostic pan-Cryptosporidium PCR assay (Cq). X-axis, the total number of Cryptosporidium infections. The infection was designated as either diarrheal (red) or sub-clinical (green) based on the current infection phenotype. The data from diarrheal cases was offset to improve data visualization. To account for within-child correlations among repeated Cryptosporidium infections, the generalized estimating equation (GEE) method for repeated measurements were used with exchangeable correlation structure. As the intercept of the diarrheal and sub-clinical models was not statistically different the common intercept (27.02 ± 0.45) was used. The slope of the data derived from the sub-clinical (1.9 ± 0.2) and diarrheal (1.49 ± 0.31) exchangeable models were not significantly different from each other (p=0.071) although both were statistically different from zero (p<0.0001).

Fig S5. Distribution of repeated Cryptosporidium infections
x axis child age in months; left y-axis child HAZ scores; right y-axis frequency of Cryptosporidium (diarrheal and sub-clinical) infections (shown as the number that occurred per the age of the child in months). All graphs include as a reference the HAZ score of children where no Cryptosporidium infections were detected (green circle and line). Cryptosporidium infections: Light blue triangle dotted blue connection line: infection one; purple circle and dotted line: infection two; light red square and dotted line: infection three; black triangle and dotted line: infection four A) blue symbol and solid line HAZ score of children who had one Cryptosporidium infections by 3 years of age B) purple square and solid line HAZ score of children who had two Cryptosporidium infections by three years of age C) red square and solid line HAZ score of children who had three Cryptosporidium infections by three years of age D) black square and solid line HAZ score of children who had four Cryptosporidium infections by three years of age

Fig S6. Recurrent cryptosporidiosis results in greater growth faltering Each symbol represents a single child. Box plot comparing the height for age z score at 3 years (HAZ) (Y-axis) mean and standard deviation shown Children were considered to be at Risk for malnutrition is they have a HAZ score <-1 and malnourished at HAZ-2: orange box: 3-year HAZ score -1 to -2; red box 3-year HAZ score < -2. X-axis Number of Cryptosporidium infections.
Fig S7. **Comparison of Cryptosporidiosis associated Growth-faltering in diarrheal and sub-clinical infections**

Graphs show results from a simple linear regression with each symbol representing a single child. Black symbols represent children who were never infected or had sub-clinical infections. The blue symbols indicate children who have had one or more than one episodes of diarrhea-associated cryptosporidiosis. Height for age (HAZ) z score at 3 years is shown on the Y-axis. The slope of the diarrheal-associated and sub-clinical groups are identical. Pooled Slope: -0.1545. Children are considered to be at Risk for malnutrition if they have a HAZ score < -1 and malnourished at HAZ -2: orange box: 3-year HAZ score -1 to -2; red box: 3-year HAZ score < -2. X-axis indicates number of *Cryptosporidium* infections.

Fig S8. **Cryptosporidiosis was associated with chronic but not acute malnutrition at year 3**

Graphs show results from a simple linear regression with each symbol representing a single child. X-axis indicates number of *Cryptosporidium* infections A) Y-axis MUACZ circumference of the mid-upper arm (muscle wasting) B) Y-axis WHZ score (low weight for height (wasting) a measure of acute malnutrition C) Y axis WAZ score (low weight for age) a measure of acute and chronic malnutrition and D) BAZ (body mass index for age).
Fig 1 COHORT diagram. Study subjects, collected samples and new infection numbers

- **Enrolled 500**
  - Lost to FU: 79

- **Completed Year 2: 421**
  - Lost to FU: 29
  - Consent withdrawn: 14
  - Mother unreachable for more than 60 days: 8
  - Moved out of study area: 2
  - Missed visit: 2
  - Incorrect year 2 visit age: 1
  - Unknown: 2

- **Completed Year 3: 392**

- **Total Diarrhea recorded (3578)**

- **Diarrheal samples (DS) collected (3004)**
  - DS assayed by RT-QPCR (2978)
  - Samples Crypto Positive (1336)
  - New Crypto Infections (698)

- **Asymptomatic samples (MS) collected (14473)**
  - MS assayed by RT-QPCR (13806)
  - Positive samples defined as < 40 Cq values
  - DS & MS grouped into single infection if within 65 days of previous

Abbreviations: RT-QPCR quantitative polymerase chain reaction; DS: Diarrheal samples; FU: Follow-up; MS: Monthly Samples, Cq: Cycle Quantification
Fig 2

A) Correlation between parasite burden and the number of Cryptosporidium infections. Each symbol represents the first detectable sample of an individual infection. Y-axis represents the quantitative cycle (Cq) of the diagnostic pan-Cryptosporidium PCR assay. X-axis shows the number of Cryptosporidium infections that had occurred in this child. The line represents the slope (-1.85 ± 0.21) and Y-intercept (26.95 ± 0.44) estimated from the GEE model with the exchangeable correlation structure (p<0.0001).

B) Comparison of single infections (black symbol) with those that are part of a series (gray symbol). Bar graph (indicating data mean ± standard deviation) with individual data points. Each symbol on the box plot represents the first positive sample of an individual infection. X-axis
refers to Infection number and, if the first infection, whether a second Cryptosporidium infection took place in the 3 years of life. Y-axis represents the quantitative cycle (Cq) of the diagnostic pan-Cryptosporidium PCR assay. Horizontal bar represents the result of a non-parametric Kruskal-Wallis test **** indicates p<0.0001
Fig 3 Parasite burden in older children
The amount of parasite in stool as a function of the number of cryptosporidia infections in a given child by linear regression. The analysis was restricted to children between 2.25 and 2.75 years of age (n=140). Each symbol represents the first detectable sample of an individual infection. Y-axis, quantitative cycle of the diagnostic pan-Cryptosporidium PCR assay (Cq). X-axis, the number of Cryptosporidium spp. infections that had occurred in each child. Slope: -1.65, R squared value: 0.01125, Significance p<0.0001.
Fig 4 Cryptosporidiosis frequency was associated with growth faltering distinct from the impact of birth nutritional status.

A) Relationship between length for age z score (LAZ) at birth (Y-axis) and the total number of Cryptosporidium spp. infections (X-axis). The slope was not significantly different from one.

B) Relationship between height for age z score (HAZ) at 3 years (Y-axis) and the total number of Cryptosporidium spp. infections (X-axis). Slope: \(-0.152 \pm 0.0429\), R squared value: 0.0313, Significance \(p=0.0004\). Children were defined to be at risk for growth faltering with a LAZ or HAZ score \(<-1\) and malnourished at LAZ or HAZ score \(<-2\). Orange box: birth LAZ or 3 year HAZ score -1 to -2 ; red box: birth LAZ or 3 year HAZ or score < -2.
**Fig 5 Correlates of cryptosporidiosis-associated growth-faltering**

A) Correlation matrix of cryptosporidiosis, all-cause diarrhea, LAZ at birth and HAZ at 3 years, calculated using Pearson r. Bar on the right indicates strength and direction of association. B) Comparison of three year-HAZ with birth LAZ. (Slope: -0.294 ± 0.05; R squared value: 0.08; Significance p < 0.0001) C) Relationship of all-cause diarrhea with HAZ at 3 years of age (p = NS).
B

LAZ at birth

-4 -2 0 2 4

HAZ at year 3

at risk

malnourished

malnourished

at risk
Number of All-Cause Diarrheal Episodes

HAZ at year 3

at risk
malnourished
Fig 6 High anti-Cp23 IgA levels were associated with a reduction in cryptosporidiosis-associated growth-faltering.

Group 1 and 2 children were in the upper and lower 50th percentile for fecal IgA anti-Cp23 respectively. HAZ are shown for children in both year one and year three of life. Mean ± standard deviation with individual data points. Horizontal bars represent the result of a non-parametric Kruskal-Wallis test ***p<0.001, **p<0.01
## Supporting information

### Table S1  Symptomatic and asymptomatic samples collected and assayed by RT-QPCR during year 2 and year 3 follow-up period

<table>
<thead>
<tr>
<th>Sample Type</th>
<th>Year 2</th>
<th>Year 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total Diarrhea Recorded</td>
<td>Total Diarrhea Recorded</td>
</tr>
<tr>
<td>Symptomatic</td>
<td>3763</td>
<td>358</td>
</tr>
<tr>
<td>Sub-clinical*</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>3148 (83.66%)</td>
<td>3148 (83.96%)</td>
</tr>
<tr>
<td></td>
<td>615</td>
<td>574</td>
</tr>
<tr>
<td></td>
<td>3132</td>
<td>2978</td>
</tr>
<tr>
<td></td>
<td>16 (0.5%)</td>
<td>26 (0.9%)</td>
</tr>
<tr>
<td></td>
<td>14638</td>
<td>13806</td>
</tr>
<tr>
<td></td>
<td>674 (4.4%)</td>
<td>667 (4.6%)</td>
</tr>
</tbody>
</table>

*Monthly Stool Samples
Table S2. Regression Analysis using selected predictors to test the association of all cause diarrhea and HAZ at year 3

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Effect (95% Confidence Interval)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>All cause diarrhea*</td>
<td>-0.008 (-0.023, 0.008)</td>
<td>0.3357</td>
</tr>
<tr>
<td>Child LAZ at Birth</td>
<td>0.245 (0.151, 0.339)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Maternal Weight</td>
<td>0.017 (0.007, 0.027)</td>
<td>0.001</td>
</tr>
<tr>
<td>Maternal Height</td>
<td>0.037 (0.019, 0.056)</td>
<td>0.0001</td>
</tr>
<tr>
<td>Maternal Education</td>
<td>0.233 (0.020, 0.446)</td>
<td>0.0323</td>
</tr>
<tr>
<td>Household income</td>
<td>0.001 (0.000, 0.002)</td>
<td>0.079</td>
</tr>
<tr>
<td>Treated water</td>
<td>0.205 (-0.003, 0.413)</td>
<td>0.0537</td>
</tr>
</tbody>
</table>

* diarrheal episodes (all-cause) were not significantly associated with HAZ at year 3
Table S1: Symptomatic and asymptomatic samples collected during year 3
Table S2: Multivariable analysis of total all-cause diarrhea and HAZ at year 3
Fig S1: Flow chart of stool processing and molecular testing
Fig S2: Flow chart of stool TNA extraction procedure using QIAamp Fast DNA Stool Mini Kit from fresh or frozen stool samples.
Fig S3: Flow chart of Multiplex qPCR of Cryptosporidium, Giardia, Entamoeba histolytica by targeting the 18S gene
Fig S4: Parasite burden in diarrheal and sub-clinical infections
Fig S5: Distribution of repeated Cryptosporidium infections
Fig S6: Recurrent cryptosporidiosis results in greater growth faltering
Fig S7: Comparison of cryptosporidiosis associated growth-faltering in diarrheal and sub-clinical infections
Fig S8: Cryptosporidiosis was associated with chronic but not acute malnutrition at year 3
Fig S9: Low anti-Cryptosporidium IgA levels after an infection were associated with a subsequent increase in cryptosporidiosis-associated growth-faltering
Fig S1: **Flow chart of stool processing and molecular testing**

While maintaining the cold chain stool samples were transferred from field to laboratory <6h after sample collection

200mg aliquots of stool frozen at -70°C

PhHV was added as an internal positive control to each sample and fecal DNA isolated using the the modified QIAamp fast stool total nucleic acid extraction protocol

Nucleic acid extraction was done in batches consisting of 23 fecal samples and one extraction blank

After completion of the nucleic acid extraction from > 80 samples a diagnostic qPCR assay was used to detect the common protozoan parasites of *Cryptosporidium, Giardia* and *Entamoeba histolytica*

qPCR preparation and plate loading were performed in a amplicon free area

Each PCR plate contained ≤ 80 unknown fecal DNA; 3 positive controls; extraction blanks, and a no template control

CFX96 real-time detection system (Bio-Rad) with CFX detection software (version 3.1) was used for amplification, detection and data analysis

Amplification consisted of 3 min at 95°C followed by 40 cycles of 10 sec at 95°C, 60 seconds at 60°C
Fig S2: Flow chart of stool TNA Extraction procedure using QIAamp Fast DNA Stool Mini Kit from fresh or frozen stool samples.

180-220 mg of frozen stool samples (thawed at room temperature) were transferred to 2 ml screw cap tubes.

- 370 mg glass bead was added to each sample

1 ml of InhibitEx with 1 μl of PhHV buffer was added to each tube

Bead beating was done for 2.30 minutes at the maximum speed

The suspension was incubated at 95°C for 5 minutes to lyse the sample

Centrifuged at 14000 rpm for 1 minute to pellet the stool particles

600 μl of supernatant was transferred to 2 ml micro-centrifuge tube containing Proteinase K

600 μl of AL buffer was added to the supernatant and incubated at 70°C for 10 minutes

600 μl of ethanol (96-100%) was added to the sample

QIAamp Spin Column was labeled and 600 μl of sample and centrifuged for 1 minute at 14000 rpm and repeated same steps two more times

Column was washed with 500 μl AW1 Buffer by centrifuge for 1 minute at 14000 rpm.

The QIAamp Spin Column was washed with 500 μl AW2 Buffer by centrifuge for 3 minutes at 14000 rpm and again centrifuged for an additional 3 minutes to dry the column.

200 μl of ATE Buffer was added and centrifuged for 1 minute at 14000 rpm to elute the DNA.

Finally, the eluted DNA thus obtained was stored at -80°C
Fig S3: Flow chart of Multiplex qPCR of *Cryptosporidium, Giardia, Entamoeba histolytica* by targeting the 18S gene

23 μl of freshly (12.5 μl Biorad iQ powermix + 7.5 μl primer probe + 3 μl nuclease free water) prepared master mix is added to the 96well containing plate

2.0 μl of template DNA is added to the well containing water and mastermix

Extraction blank, positive and negative control wells positive control materials and negative control (nuclease free water) were added respectively

PCR plate was sealed with optically clear sealing tape

The plate was then centrifuged for 30 seconds to remove bubbles and plated into the thermal cycler

The filter pair for fluorophore was selected and then ran with an automated 40 cycle protocol and the well factors were collected

Amplification consisted of 3 minutes at 95°C followed by 40 cycles of 10 seconds at 95°C, 60 seconds at 60°C
**Fig S4. Parasite burden in diarrheal and sub-clinical infections**

Relationship between Parasite Burden and the number of recurrent *Cryptosporidium* infections. Each symbol represents the first detectable sample of an individual infection. Y-axis, quantitative cycle of the diagnostic pan-*Cryptosporidium* PCR assay (Cq). X-axis, the total number of *Cryptosporidium* infections. The infection was designated as either diarrheal (red) or sub-clinical (green) based on the current infection phenotype. The data from diarrheal cases was offset to improve data visualization. To account for within-child correlations among repeated *Cryptosporidium* infections, the generalized estimating equation (GEE) method for repeated measurements were used with exchangeable correlation structure. As the intercept of the diarrheal and sub-clinical models was not statistically different the common intercept (27.02 ± 0.45) was used. The slope of the data derived from the sub-clinical (1.9 ± 0.2) and diarrheal (1.49 ± 0.31) exchangeable models were not significantly different from each other (p=0.071) although both were statistically different from zero (p<0.0001).
Fig S5. Distribution of repeated *Cryptosporidium* infections

X axis child age in months; left y-axis child HAZ scores; right y-axis frequency of *Cryptosporidium* (diarrheal and sub-clinical) infections (shown as the number that occurred per the age of the child in months). All graphs include as a reference the HAZ score of children where no *Cryptosporidium* infections were detected (green circle and line). *Cryptosporidium* infections: Light blue triangle dotted blue connection line: infection one; purple circle and dotted line: infection two; light red square and dotted line: infection three; black triangle and dotted line: infection four A) blue symbol and solid line HAZ score of children who had one *Cryptosporidium* infections by 3 years of age B) purple square and solid line HAZ score of children who had two *Cryptosporidium* infections by three years of age C) red square and solid line HAZ score of children who had three *Cryptosporidium* infections by three years of age D) black square and solid line HAZ score of children who had four *Cryptosporidium* infections by three years of age
**Fig S6. Recurrent cryptosporidiosis results in greater growth faltering** Each symbol represents a single child. Box plot comparing the height for age z score at 3 years (HAZ) (Y-axis) mean and standard deviation shown. Children are considered to be at Risk for malnutrition if they have a HAZ score < -1 and malnourished at HAZ-2: orange box: 3-year HAZ score -1 to -2; red box 3-year HAZ score < -2. X-axis Number of Cryptosporidium infections. Bar indicates the result of a non-parametric Kruskal-Wallis test for multiple comparisons * indicates $p<0.05$ ** indicates $p<0.01$
Fig S7: **Comparison of cryptosporidiosis associated growth-faltering in diarrheal and sub-clinical infections**

Graphs show results from a simple linear regression with each symbol representing a single child. Black symbols represent children who were never infected or had sub-clinical infections. The blue symbols indicate children who have had one or more than one episodes of diarrhea-associated cryptosporidiosis. Height for age (HAZ) z score at 3 years is shown on the Y-axis. The slope of the diarrheal-associated and sub-clinical groups are identical. Pooled Slope: -0.1545.

Children are considered to be at Risk for malnutrition if they have a HAZ score < -1 and malnourished at HAZ -2: orange box: 3-year HAZ score -1 to -2; red box: 3-year HAZ score < -2. X-axis indicates number of *Cryptosporidium* infections.
Fig S8. *Cryptosporidiosis* was associated with chronic but not acute malnutrition at year 3.

Graphs show results from a simple linear regression with each symbol representing a single child. X-axis indicates number of *Cryptosporidium* infections. A) Y-axis MUACZ (circumference of the mid-upper arm) a measure of muscle wasting. B) Y-axis WAZ score (low weight for age) a measure of acute and chronic malnutrition. C) Y-axis BAZ (body mass index for age) a measure of acute malnutrition. D) Y-axis WHZ score (low weight for height: wasting) a measure of acute malnutrition.
Fig S9. Low anti-Cryptosporidium IgA levels after an infection were associated with a subsequent increase in cryptosporidiosis-associated growth-faltering.

Bar graphs (indicating data mean ± standard deviation) with individual data points. Each symbol on the box plot represents a child. On the X-axis values are shown for children in Groups 1 and 2 in both year one and year 3. Groups on the X-axis refers to the values obtained the end of the first and third years of life. Y-axis represents the growth faltering (HAZ). Horizontal bars represent the result of a non-parametric Kruskal-Wallis test *** indicates p<0.001 ** indicates p<0.01 * p<0.05

A) Group 1 children had higher than average levels of fecal anti-IgA Cp23. Group2a children were negative by diagnostic surveillance by qPCR and anti-Cp23 antibodies at year one. Group2b were positive by qPCR but had nevertheless low levels of anti-Cp23 antibodies.

B) Group 1 children had higher than average levels of fecal anti-IgA Cp17. Group2a children were negative by diagnostic surveillance by qPCR and anti-Cp17 antibodies at year one. Group2b were positive by qPCR but had nevertheless low levels of anti-Cp17 antibodies.