Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Phong Le Date

Symptom and Enteric Pathogen Indicators of Diarrhea After Study Enrollment Among Controls in the Vaccine Impact on Diarrhea in Africa Study

By
Phong Le
Master of Public Health

Department of Epidemiology

[X] _____
Dr. Matthew Freeman PhD, MPH,
Committee Chair

[X] _____
Dr. David Berendes PhD, MSPH,
Committee Member

[X] _____
Graeme Prentice-Mott MSPH,
Committee Member

Symptom and Enteric Pathogen Indicators of Diarrhea After Study Enrollment Among Controls in the Vaccine Impact on Diarrhea in Africa Study

By

Phong Le
BBA., Emory Goizueta Business School, 2018

Thesis Committee Chair: Dr. Matthew Freeman PhD, MPH

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health in the Epidemiology Department 2021

Abstract

Symptom and Enteric Pathogen Indicators of Diarrhea After Study Enrollment Among Controls in the Vaccine Impact on Diarrhea in Africa Study

By Phong Le

There is a high prevalence of pediatric diarrhea in low-to-middle income countries. The Vaccine Impact on Diarrhea in Africa (VIDA) study assessed the attributable fraction of pathogen etiologies on moderate to severe diarrhea among young children. The objective of this study was to assess the prior symptoms and enteric pathogen carriage of Vaccine Impact on Diarrhea in Africa (VIDA) study controls in relation to their onset of diarrhea after study enrollment (DASE). Understanding the background levels of diarrheal disease, enteric pathogen carriage, and symptomatic expression among VIDA reference populations in highly endemic areas would allow for the contextualization of analyses among VIDA cases. This analysis of VIDA controls indicated that there was a high background level of acute diarrhea (11%) and carriage of enteric pathogens (63%) at the Kenya site. Several controls that were free of diarrhea 7 days prior to enrollment may have been actively incubating a diarrheal disease. Norovirus GII, sapovirus, and ST ETEC were enteric pathogens associated with DASE among VIDA controls in this study as well as moderate-to-severe diarrhea (MSD) among VIDA cases in a previous study. As such, previous estimates of the importance of these enteric pathogens for MSD, given these previous estimates were based on a reference population that was not entirely disease-free, may be underestimates. Fever and vomiting presented 7 days prior to enrollment were associated with DASE and could serve as proxies for the expression of acute diarrhea soon after study enrollment. Fever and vomiting 7 days prior to enrollment should be considered as exclusion criteria for disease-free reference populations. These findings related to DASE and enteric pathogen detection in VIDA controls further underscored the need for close examination of disease-free reference populations.

Symptom and Enteric Pathogen Indicators of Diarrhea After Study Enrollment Among Controls in the Vaccine Impact on Diarrhea in Africa Study

By

Phong Le
BBA., Emory Goizueta Business School, 2018

Thesis Committee Chair: Dr. Matthew Freeman PhD, MPH

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health in the Epidemiology Department 2021

1. Introduction

An estimated 1.7 billion cases of diarrhea and 525,000 deaths related to diarrhea are reported every year among children under the age of five [1]. Pediatric diarrhea is caused by a variety of infectious agents including bacteria, viruses, and parasites [6-13]. The main enteric pathogens responsible for diarrheal disease vary by region [2] while mortality and morbidity are concentrated in Africa and South Asia [3]. Enteric pathogens differ in their expression of symptoms, biological mechanisms of infection, incubation period, and duration of disease [14]. WHO has set the standard in defining diarrhea as a 24-hour period of 3 or more loose stools [1], but diarrhea can also be further specified by duration and severity [15-16]. Because of the variety of endemic enteric pathogens, presentations of disease, and definitions of diarrhea, reference populations in studies assessing diarrhea in different countries and regions are not homogenous. Understanding the background levels of diarrheal disease, enteric pathogen carriage, and symptomatic expression in highly endemic areas is important for the contextualization of analyses of pediatric diarrhea. Pediatric diarrhea has been commonly assessed through case-control studies [17-22]. Yet, if reference populations are not disease-free, studies of diarrheal etiologies may underestimate the role of certain pathogens.

Diarrhea can lead to severe morbidities and even mortality among children living in low-to-middle income countries (LMIC) [2-5]. On average, children under the age of 5 develop 2-3 episodes of diarrhea per year [2]. However, children in LMIC can develop up to 3-11 episodes of diarrhea per year [4]. Children in LMICs often have poor water and sanitation conditions which elevates the risk of diarrhea [5] and malnutrition serves as both a risk factor and outcome of diarrhea [1]. Children earlier in their development face the highest risk of exposure to infectious agents of diarrhea [2]. When infants reach developmental milestones, such as crawling and weening, they face higher exposure to enteric pathogens due to mouthing behaviors [3].

Diarrhea can be categorized by its duration and severity. In terms of duration, postprandial diarrhea can be defined as a single loose stool after a meal [23], acute diarrhea can be defined as loose stools lasting hours to days [15] while persistent diarrhea can be defined as loose stools lasting 14 days or longer [24]. The presentation of blood in diarrhea reflects an episode of dysentery [14]. It is not uncommon for a study to assess diarrhea using a definition specific to its needs. For example, the Global Enteric Multicenter Study defined moderate-to-severe diarrhea as the presence of diarrhea in addition to an exhibition of sunken eyes, loss of skin turgor, a need for intravenous rehydration, dysentery, or hospitalization [16]. Criteria for moderate-to-severe diarrhea represented characteristics associated with the highest morbidity and mortality related to pediatric diarrhea. The study of moderate-to-severe diarrhea allows for the prioritization of strategies addressing pediatric diarrhea's most serious complications.

Although most enteric pathogens can cause severe episodes of diarrhea depending on host factors, their presentation of symptoms vary according to the biological mechanism causing disease symptoms [24]. The incubation period for enteric pathogens range from 1-7 days [6-13]. Noninflammatory diarrhea is the result of an enteric pathogen interfering with the absorption of water and nutrients in the gastrointestinal system [14]. Examples of enteric pathogens that cause this include heat-stable enterotoxigenic Escherichia coli (ST ETEC), norovirus, and rotavirus [6,10-12]. Often, loose (watery) stools are the predominant symptom of noninflammatory diarrhea [14]. On the other hand, inflammatory diarrhea occurs through the invasion of the intestinal epithelium and release of cytotoxins in the colon [14]. *Campylobacter* spp., *Salmonella* spp., and *Shigella* spp. can all release this cytotoxin [8-9,17]. Symptoms presented during inflammatory diarrhea include diarrhea as well as abdominal cramping, fever, vomiting, and blood in stool [14-15]. Given the varied presentation of symptoms, it is challenging to define a disease-free individual and demonstrates the need to thoroughly study reference populations.

Correct categorization of children into proper disease states (e.g., diarrhea or disease-free) is critical to appropriately understand dominant etiologies of disease-causing pathogens. Case-control studies are often performed to assess etiological factors associated with severe forms of diarrhea [6-13]. Matching commonly occurred in these studies to ensure a similar number of cases and controls in the various strata

of confounders [25-26]. For studies assessing diarrhea, cases are typically deemed eligible if they exhibited the clinical indicators of diarrhea during a specified amount of time before enrollment [17-22]. Controls are usually matched on known confounders of diarrhea such as age and sex, were then chosen based on the absence of diarrhea prior to enrollment [21]. Controls are enrolled to be disease-free to juxtapose the distribution of exposures among cases [25]. In LMIC, these controls are drawn from reference populations with various levels of background diarrhea and enteric pathogen carriage. Given high levels of enteric infection in LMIC, controls who are free of diarrheal symptoms at and before enrollment may have had an infection with an enteric pathogen still in its incubation period and went on to experience diarrhea soon after enrollment. Diarrhea after enrollment is rarely an exclusion criterion for controls because operationally, it would be difficult to enroll then subsequently unenroll a control [25].

Due to the high prevalence of pediatric diarrhea in LMIC and difficulty in identifying disease-free individuals, the objective of this study was to assess the incidence of diarrhea, enteric pathogen carriage, and symptom presentation within a reference population for the Vaccine Impact on Diarrhea in Africa (VIDA) study. The VIDA study assessed the attributable fraction of pathogen etiologies on moderate-to-severe diarrhea among young children. In VIDA, information regarding loose stools in the days following enrollment was collected. This novel collection of data on controls presented an opportunity to characterize controls by whether they experienced diarrhea soon after enrollment and make inferences on background levels of diarrhea. Understanding the background levels of diarrhea, enteric pathogen carriage, and symptomatic expression among VIDA reference populations in highly endemic areas would allow for the contextualization of analyses among VIDA cases.

2. Methods

We assessed the prior symptoms and enteric pathogen carriage of Vaccine Impact on Diarrhea in Africa (VIDA) study controls in relation to their onset of diarrhea after study enrollment (DASE). VIDA was a 36-month case-control study assessing the causes and burden of childhood moderate-to-severe diarrhea (MSD) in sub-Saharan African sites including Mali, Kenya, and The Gambia [27]. VIDA was a follow-on study of the previous Global Enteric Multicenter Study (GEMS) which also assessed childhood MSD [16]. Unlike GEMS, VIDA encompassed populations that had been vaccinated against rotavirus [27].

Study Site

This analysis focused on the Siaya County, Kenya site. This site is a rural location situated near the western border of Kenya, on the shores of Lake Victoria [28]. This site was chosen for VIDA because of its lower-middle income class and moderate-to-high under 5 mortality rates [29]. Data were collected between 2015 and 2018 among 2,095 controls [29]. The overall study methods are found here [29].

Inclusion Criteria for Controls

Cases were identified via sentinel health centers where children sought care for MSD. MSD was defined as diarrhea in addition to an exhibition of sunken eyes, loss of skin turgor, a need for intravenous rehydration, dysentery, or hospitalization [16]. Children were defined as those 0-59 months old. For each case, 1-3 controls were selected matched on age, sex, residence, and time. Age matching was defined as being within 2 months to the age of the case. Residence matching was defined as being in the same or nearby village. Time matching was defined as identifying a control within 14 days of identifying a case of MSD. Any control that reported having diarrhea in the 7 days prior to enrollment was disqualified. For this analysis, only controls from the Kenya site were included. Any controls in which the memory aid data

was unavailable were excluded from the study. Laboratory data regarding pathogen detection was available for all controls included in the analyses; controls with any missing pathogen data were excluded from this analysis.

Enrollment

At enrollment, an assessment was performed on all cases and controls [29]. Unlike the cases who were assessed at a health center, controls were assessed at their home. Information sought in this assessment include identifying, demographic, clinical, epidemiologic, anthropometric, and vaccine information. From the demographic information, sex and age were assessed. From the clinical information, the presentation of fever or vomiting in the 7 days prior to enrollment was assessed. Vaccination status was confirmed via vaccination card or a verified date of vaccination. A sample of feces was collected at enrollment in the home to test for enteric pathogens.

Enteric Pathogen Detection

Enteric pathogen detection was performed by researchers in Kenya using TaqMan® probe-based real-time reverse transcriptase (RT) polymerase chain reaction (PCR) assays. The TaqMan® Array microfluidic Card (TAC) was found to be specific and sensitive to the detection of enteric pathogens such as rotavirus and is considered a gold standard [30]. Nucleic acids were extracted from a 200 mg aliquot of stool to assess for a TAC panel of 26 enteric pathogens. Enteric pathogens of interest in this study were pathogens previously determined to be significantly associated with MSD among VIDA cases. These enteric pathogens included astrovirus, *Campylobacter* spp., *Cryptosporidium* spp., *H. pylori*, genotype II norovirus (norovirus GII), rotavirus, *Salmonella* spp., sapovirus, *Shigella* spp., and ST ETEC [Unpublished VIDA findings]. These pathogens were considered and referred to throughout this paper as

"MSD pathogens". Pathogen detection was represented as a binary variable based off a cycle threshold value of 35 for detection [29].

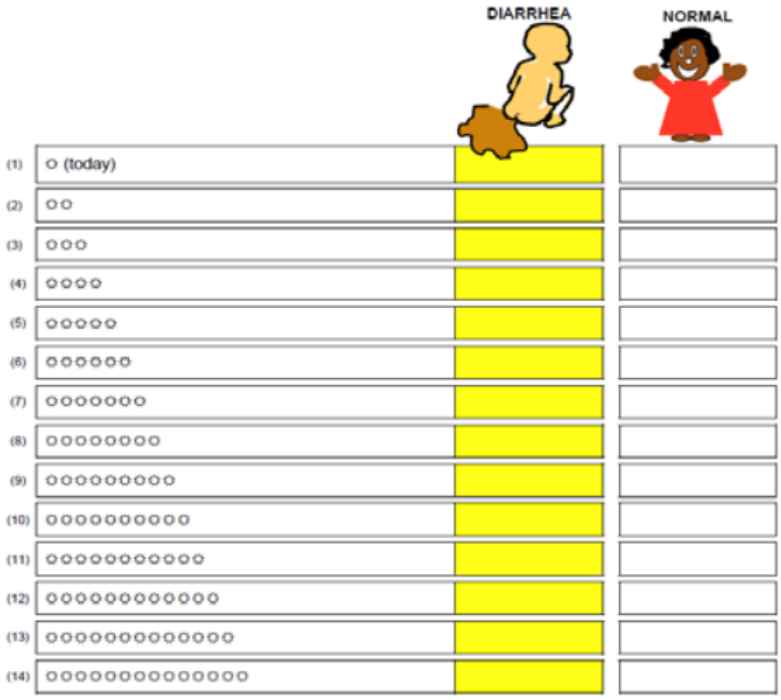
14-Day Memory Aid Form Administration

Memory aid cards were administered at enrollment to record whether the child experienced diarrhea in the 14 days after enrollment (Figure 1). Both cases and controls were provided with a memory aid. Parents were provided writing supplies to ensure they had the capability to fill out the card. Memory aids were illustrated with 28 checkmark boxes (binary choices for 14 days) to determine if a child had loose stools on any respective day. Illustrations were used to ensure that the card was accessible to those that may be illiterate. A day of loose stools was defined as having at least 3 loose stools within the span of 24 hours [1]. Memory aids were collected at 60-day follow up.

Defining the Outcome: Diarrhea After Study Enrollment

Diarrheal episodes were measured by identifying consecutive days of loose stools on the memory aid. When a day of no loose stools was marked, the diarrheal episode was considered to have ended. Diarrhea after study enrollment (DASE) was defined as the occurrence of a diarrheal episode with a length of at least 2 consecutive days of diarrhea within the first week after enrollment. The definition of DASE was specified to meet the requirements of acute diarrhea as defined by WHO [1]. An episode of DASE was also considered an episode of acute diarrhea. Only the first week after enrollment was considered for DASE because the incubation periods for the included enteric pathogens were often less than 7 days [6-13].

Figure 1: Example of memory aid



It was unlikely that diarrhea beginning after the first 7 days of follow up may be attributable to enteric pathogens detected at enrollment. A sensitivity analysis was performed to assess what impact varying definitions of DASE had on associations between enteric pathogen detection and DASE [31]. Alternative definitions of DASE included (a) at least one consecutive day of loose stools and (b) at least three consecutive days of loose stools.

Statistical Analysis

Statistical analyses were performed using R version 4.0.3 [32]. This study treated the controls of the VIDA study as a separate cohort with two time points: (1) at enrollment and (2) 14-days past enrollment. Independent variables analyzed include symptoms experienced within the past 7 days of enrollment and enteric pathogens tested in stool samples at enrollment. The outcome of the analysis was DASE. Descriptive analyses were performed to provide a summary of demographic data from the VIDA control

enrollment surveys as well as to describe the distribution of symptoms reported 7 days prior to enrollment and enteric pathogen detection. Descriptive analyses were also performed to describe the distribution and temporality of diarrheal episodes by day of onset. Risk ratios were estimated via log binomial regression models to measure the associations of symptoms reported 7 days prior to enrollment/ enteric pathogens detected at enrollment and DASE, adjusted for sex and age. Sensitivity analyses were conducted to assess risk ratios using the two defined alternate definitions of DASE to measure the robustness of results. In the sensitivity analysis, separate risk ratios were estimated using presentation of symptoms and enteric pathogen carriage as inputs and varying definitions of DASE as outcomes.

3. Results

Demographics, Symptoms, and Detection of Enteric Pathogens

At the VIDA Kenya site, there were 2,095 controls. Among those, 2,009 (96%) had memory aid data available, and among those, 1,433 (71%) had TAC data available for the MSD pathogens assessed in this study and were included in this analysis. Over half (55%) were male, 38% were between 0-11 months old, 34% were between 12-23 months old, and 29% were between 24-59 months old (Table 1). There were differences across age groups in frequencies of children experiencing DASE. Of the controls that experienced DASE, 45% were between 0-11 months old while only 33% and 22% of controls that experienced DASE were between 12-23 months and 24-59 months respectively (Table 1). At least one MSD pathogen was detected in 63% of controls in this study (Table 1). The most prevalent pathogens detected included campylobacter (20%), ST ETEC (19%), and Shigella (16%) (Table 1). In terms of reported clinical symptoms 7 days prior to enrollment, 27% of controls reported fever and 3% of controls reported vomiting (Table 1).

Table 1: Descriptive analysis of demography, symptoms 7 days prior to enrollment, and enteric pathogen detection stratified by DASE (n = 1433)

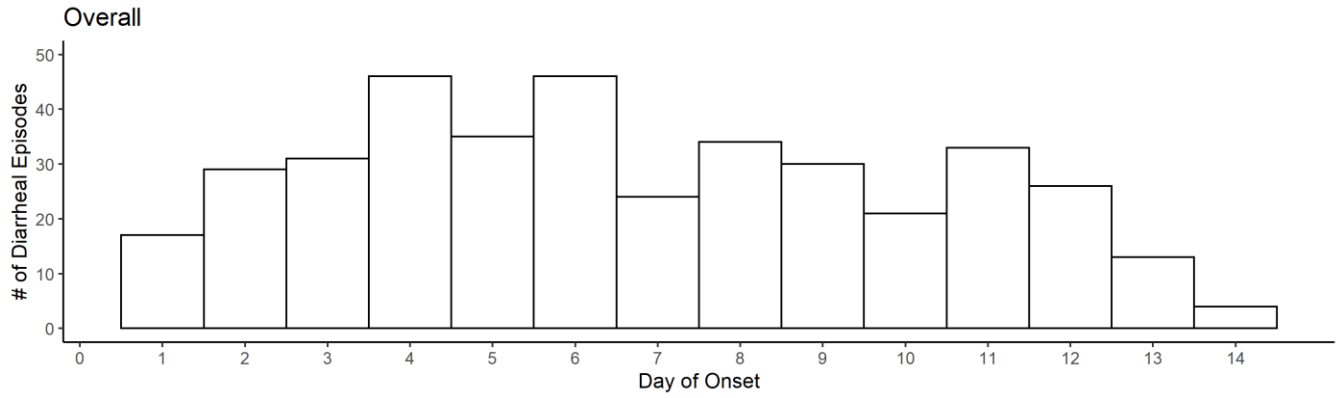
Attribute	No DASE [a]	DASE [a]	Total
Sex			
Male	692 (54.3)	96 (60.4)	788 (55.0)
Female	582 (45.7)	63 (39.6)	645 (45.0)
Age Group			
0-11 Months	465 (36.5)	75 (47.2)	540 (37.7)
12-23 Months	433 (34.0)	51 (32.1)	484 (33.8)
24-59 Months	376 (29.5)	33 (20.8)	409 (28.5)
Symptoms in Past 7 Days			
Fever	311 (24.4)	73 (45.9)	384 (26.8)
Vomit	30 (2.4)	13 (8.2)	43 (3.0)
Pathogen [b]			
Astrovirus	29 (2.3)	2 (1.3)	31 (2.2)
Campylobacter	241 (18.9)	38 (23.9)	279 (19.5)
Cryptosporidium	133 (10.4)	19 (11.9)	152 (10.6)
H Pylori	96 (7.5)	12 (7.5)	108 (7.5)
Norovirus GII	118 (9.3)	29 (18.2)	147 (10.3)
Rotavirus	37 (2.9)	4 (2.5)	41 (2.9)
Salmonella	21 (1.6)	2 (1.3)	23 (1.6)
Sapovirus	96 (7.5)	22 (13.8)	118 (8.2)
Shigella	198 (15.5)	27 (17.0)	225 (15.7)
ST ETEC	235 (18.4)	41 (25.8)	276 (19.3)
Any MSD Pathogen	783 (61.5)	117 (73.6)	900 (62.8)

[a] DASE: Diarrhea After Study Enrollment
[b] Pathogens associated with MSD among VIDA cases [Unpublished VIDA findings]

Analysis of Diarrheal Incidence

Of 1433 controls, 297 (21%) controls experienced DASE. Among controls that experienced DASE, 159 (54%) episodes occurred in the first week while 138 (46%) episodes occurred in the second week. Occurrences of loose stools that started on days 4-6 represented 36% of episodes (Figure 2). Among the controls reporting any loose stools, the average length of the diarrheal episode was 2.41 days and the average number of diarrheal episodes was 1.65 days.

Figure 2: Distribution and temporality analyses of diarrheal incidence 14 days after enrollment



Symptoms by Diarrhea After Study Enrollment

It was found that there was an association between the reporting of fever 7 days prior to enrollment and DASE (Risk Ratio [RR] 2.36, 95% Confidence Interval [CI] 1.77 – 3.15); (Table 3). This association with DASE was also seen between the reporting of vomiting and DASE (RR 2.95, 95% CI 1.73 – 4.48); (Table 3). In the sensitivity analyses, it was found that these associations between fever/vomiting and DASE were robust across varying definitions of DASE (Table 4).

Tables 3: Risk ratios measuring association between self-reported symptoms 7 days prior to enrollment and DASE

	DASE [a]
Symptom	RR [b] (95% CI)
Fever	**2.36 (1.77, 3.15)**
Vomiting	**2.95 (1.73, 4.48)**

Bold indicates significance over 5%
[a] DASE: Diarrhea After Study Enrollment
[b] Adjusted for age and sex

Tables 4: Sensitivity analysis of risk ratios measuring association between self-reported symptoms 7 days prior to enrollment and alternate definitions of DASE [a]

	Consecutive Days of Loose Stools Marked During First 7 Days of Follow Up	
	1+ Days	3+ Days
Symptom	RR [b] (95% CI)	RR [b] (95% CI)
Fever	**1.80 (1.44, 2.24)**	**2.54 (1.68, 3.81)**
Vomiting	**2.19 (1.39, 3.11)**	**2.96 (1.31, 5.49)**

Bold indicates significance over 5%
[a] DASE: Diarrhea After Study Enrollment
[b] Adjusted for age and sex

Enteric Pathogen Detection by Diarrhea After Study Enrollment

There was an association between DASE status and the detection of norovirus GII (RR 1.82, 95% CI 1.24 – 2.57), sapovirus (RR 1.79, 95% CI 1.16 – 2.62), ST ETEC (RR 1.45, 95% CI 1.03 – 1.99), and any of the listed MSD pathogens (RR 1.59, 95% CI 1.15, 2.25) in stool at enrollment (Table 5). Sensitivity analyses revealed no associations between MSD pathogens and DASE were statistically significant across all definitions of DASE (Table 6). In the sensitivity analyses using the definition of DASE measuring at least one consecutive day of loose stools, it was found that the association with the detection of norovirus GII and *Any MSD Pathogen* remained statistically significant (Table 6). There was an association between the detection of campylobacter and at least one consecutive day of loose stools (RR 1.43, 95% CI 1.11 – 1.82); (Table 6). In the other sensitivity analyses using the definition of DASE measuring at least three consecutive days of loose stools, it was found that the association with the detection of sapovirus remained statistically significant (Table 6).

Tables 5: Risk ratios measuring association between enteric pathogen detection at enrollment and DASE

Pathogen	DASE [a] RR [b] (95% CI)
Astrovirus	0.54 (0.09, 1.58)
Campylobacter	1.20 (0.84, 1.66)
Cryptosporidium	1.11 (0.69, 1.69)
H Pylori	1.03 (0.56, 1.70)
Norovirus GII	**1.82 (1.24, 2.57)**
Rotavirus	0.83 (0.27, 1.83)
Salmonella	0.74 (0.13, 2.09)
Sapovirus	**1.79 (1.16, 2.62)**
Shigella	1.15 (0.76, 1.67)
ST ETEC	**1.45 (1.03, 1.99)**
Any MSD Pathogen	**1.59 (1.15, 2.25)**

Bold indicates significance over 5%
[a] DASE: Diarrhea After Study Enrollment
[b] Adjusted for age and sex

Tables 6: Sensitivity analysis of risk ratios measuring association between enteric pathogen detection at enrollment and alternate definitions of DASE [a]

	Consecutive Days of Loose Stools Marked During First 7 Days of Follow Up	
Pathogen	1+ Days RR [b] (95% CI)	3+ Days RR [b] (95% CI)
Astrovirus	1.22 (0.56, 2.14)	1.02 (0.17, 2.99)
Campylobacter	**1.43 (1.11, 1.82)**	1.08 (0.64, 1.73)
Cryptosporidium	1.12 (0.79, 1.54)	1.33 (0.70, 2.29)
H Pylori	0.89 (0.54, 1.35)	0.45 (0.11, 1.18)
Norovirus GII	**1.42 (1.03, 1.89)**	1.32 (0.69, 2.27)
Rotavirus	0.92 (0.41, 1.66)	0.76 (0.13, 2.28)
Salmonella	1.38 (0.60, 2.47)	0.68 (0.04, 2.84)
Sapovirus	1.30 (0.89, 1.80)	**1.99 (1.09, 3.35)**
Shigella	1.16 (0.85, 1.53)	1.01 (0.54, 1.73)
ST ETEC	1.17 (0.89, 1.51)	1.53 (0.95, 2.38)
Any MSD Pathogen	**1.57 (1.22, 2.04)**	1.39 (0.90, 2.23)

Bold indicates significance over 5%
[a] DASE: Diarrhea After Study Enrollment
[b] Adjusted for age and sex

4. Discussion

We conducted an analysis of controls in the VIDA study to better understand background levels of diarrhea by measuring the incidence of DASE and its association with symptom presentation and enteric pathogen carriage. We found that there was a high occurrence of DASE among VIDA controls (11%). Since DASE fulfilled the requirements to be considered acute diarrhea [15], this implied there was a substantial background level of acute diarrhea occurring at the Kenya site among controls. We found considerable background carriage of any MSD pathogens detected in the stool of controls at enrollment (63%). GII noroviruses, sapovirus, and ST ETEC were associated with DASE/ acute diarrhea. These enteric pathogens were also associated with MSD when analyzing VIDA cases [Unpublished VIDA findings] which allowed for the opportunity to contextualize these findings. As a result, previous estimates of the importance of norovirus GII, sapovirus, and ST ETEC for MSD among VIDA cases, given these previous estimates were based on a reference population that was not entirely disease-free, may be underestimates [16].

The understanding of analyses regarding MSD among VIDA cases was further explored by understanding DASE among controls. The high incidence of DASE among controls (11%) suggested that there may have been a large subset of controls that were actively incubating a diarrheal disease at enrollment. DASE could be considered acute diarrhea but it could not be considered MSD as data regarding the clinical presentation of diarrhea was unavailable from memory aids alone [29]. It would be incorrect to state that controls that experienced DASE were misclassified [25].

An objective of the VIDA study was to ensure controls were disease-free rather than free of MSD. That is why VIDA controls that exhibited diarrhea 7 days prior to enrollment were disqualified despite potentially not presenting signs of MSD [29]. Because VIDA controls were meant to be free of disease,

the results of an analysis using controls only free of MSD may bias associations towards the null [33]. Performing a sensitivity analysis where VIDA cases are compared against (1) controls that are free of MSD and (2) controls that are free of MSD and DASE would reveal whether analyses regarding MSD among VIDA cases remain robust while maintaining a disease-free reference group [31].

Our findings regarding symptoms reported 7 days prior to enrollment could aid in the enrollment criteria of future case-control studies assessing acute diarrhea. Use of a memory aid after enrollment to assess the incidence of diarrheal disease among controls may not be feasible for every study [17-22], but it would be feasible to inquire about fever or vomiting prior to enrollment. Unlike detection of enteric pathogens which requires time to transport and test stool specimens in a laboratory setting [34], ascertaining the previous presentation of symptoms provides immediate insight as to whether a control may later develop DASE. The collection of memory aid among VIDA controls provided a unique opportunity to study symptomatic indicators of DASE. We found that there was an association between experiencing fever or vomiting 7 days prior to enrollment and DASE. Both these associations were robust and remained statistically significant in all sensitivity analyses. Although fever and vomiting have been documented to have occurred with no known cause [35], there is a strong case that these symptoms represented inflammatory diarrhea in which fever and vomiting are common [14]. This was an indication that fever and vomiting might serve as a useful proxy for DASE. It would be useful and feasible for future case-control studies studying acute diarrhea to exclude controls experiencing fever or vomiting in the 7 days prior to enrollment.

Associations between enteric pathogen carriage and DASE status may depend on the number of consecutive days of loose stools defining DASE. Sensitivity analyses were performed to understand how sensitive associations with DASE were to varying durations of loose stools [31]. A definition of at least one consecutive day of loose stools was included in the sensitivity analysis since the WHO definition of

acute diarrhea specifies 3 or more loose stools within 24 hours [1]. Only one consecutive day of loose stools would meet this definition. Caretakers were trained to only mark days that met the WHO definition of diarrhea. However, the illustrated memory aid provides no visual reminder that there must be a minimum of three loose stools. Caretakers who did not follow the training definition may have marked a full day on the memory aid for a sudden occurrence of postprandial diarrhea which ended after one bowel movement [23]. As a result, the definition of at least two consecutive days of diarrhea was used as the study's definition of DASE to be more conservative and account for this type of misclassification. The definition of at least three consecutive days of loose stools was included in the sensitivity analysis as the most conservative definition. However, given enteric infections can commonly resolve before the third day, a definition this conservative may have excluded certain infections [15]. The associations between the detection of some viruses in stool and DASE only remained statistically significant under certain definitions of DASE. An association with DASE and norovirus GII was only significant using a definition of 1+ consecutive days. This may be because norovirus infections usually don't result in more than 3 days of diarrhea [12]. When considering sapovirus, an association was only significant using a definition of 3+ consecutive days, which may result from the fact that diarrheal symptoms of sapovirus can last as long as 6 days [13]. The sensitivity of these measures of association suggested the need to align definitions of diarrhea to the specific disease durations of enteric pathogens.

The grouping of MSD pathogens derived from the analysis of VIDA cases drove instances of DASE among VIDA controls. Analyses involving the *Any MSD Pathogen* variable allowed for a high-level understanding of this grouping of pathogens and revealed an association between detection of known MSD enteric pathogens and DASE. The reason this association was statistically significant under a definition of DASE involving 1+ days of loose stools but not 3+ days of loose stools may be because some enteric pathogens in this grouping such as norovirus do not often present diarrhea for more than 3 days

[12]. MSD pathogens were characterized as such because they were less frequent in VIDA controls than VIDA cases. Given this, the association between MSD pathogens and DASE among VIDA controls underscored the idea that the importance between MSD pathogens and VIDA cases may have been underestimated.

This study presented several limitations. The VIDA study controls were not intended to be analyzed as an independent cohort that is representative of a pediatric population in Kenya [25]; therefore, the results should not be seen as generalizable to the Kenyan population [26]. Second, the grouping of MSD pathogens used in this study is specific to the VIDA study's cases [Unpublished VIDA results]. This grouping may not be relevant to all case-control studies on diarrheal disease. Although MSD pathogens were intended to be interpreted as a generalized grouping, different enteric pathogens may have been driving its association with DASE depending on the definition of DASE used. For example, norovirus GII may have driven the association between *Any MSD Pathogen* and DASE defined as at least 1 consecutive day of loose stools while sapovirus may have driven this association between *Any MSD Pathogen* and DASE defined as at least 3 consecutive day of loose stools. This added implications as to how encompassing this *Any MSD Pathogen* variable may have been when it was being driven by a few key enteric pathogens.

In the future, it would provide further insight to set up a range of diarrheal follow-up specific to each enteric pathogen included instead of the overall 1 week follow up in this analysis. This would allow for an analysis that was specifically designed to capture the various incubation periods of those respective enteric pathogens. In this analysis, enteric pathogens were assessed separately while only holding age and sex constant as confounders. However, in order to better understand which pathogens of interest retain or gain associations with DASE in the presence of co-detections with other enteric pathogens, future analyses should consider approaches including all assessed enteric pathogens in the same multivariate model.

Lastly, the enteric pathogens included in this analysis were ones indicated to be significantly associated with MSD among VIDA cases [Unpublished VIDA findings]. It would be interesting in future renditions of this study to group pathogens by whether they cause inflammatory or non-inflammatory diarrhea. This would provide more insight into the associations tied to fever and vomiting prior to enrollment since these symptoms are more characteristic of inflammatory diarrhea [14].

Conclusion

This analysis of VIDA controls found that there was a high background level of acute diarrhea at the Kenya site. Several controls that were free of diarrhea 7 days prior to enrollment may have been actively incubating a diarrheal disease. There were several MSD pathogens associated with DASE among VIDA controls. Therefore, the understanding of their role in severe pediatric diarrhea may have been underestimated. Symptoms presented 7 days prior to enrollment were associated with DASE and could serve as proxies for the expression of acute diarrhea soon after study enrollment [15]. Fever and vomiting 7 days prior to enrollment could be considered as exclusion criteria for controls in future studies. These findings related to DASE and enteric pathogen detection in VIDA controls further underscored the need for close examination of disease-free reference populations.

5. References

1. “Diarrhoeal Disease.” WHO, World Health Organization, 2 May 2017, www.who.int/news-room/fact-sheets/detail/diarrhoeal-disease.
2. Snyder, John D., and Michael H. Merson. "The magnitude of the global problem of acute diarrhoeal disease: a review of active surveillance data." Bulletin of the world health organization 60.4 (1982): 605.
3. Persistent diarrhoea in children in developing countries: memorandum from a WHO meeting. Bull World Health Organ 1988;66:709-17.
4. Lima, A. A., and Richard L. Guerrant. "Persistent diarrhea in children: epidemiology, risk factors, pathophysiology, nutritional impact, and management." Epidemiologic reviews 14 (1992): 222-242.
5. Mata LJ, Guerrant RL. Magnitude and impact of diarrhoeal disease. In: Guerrant RL, ed. Bailliere'sclinical tropical medicine and communicable dis-eases. London, England: Bailliere Tindall, 1988:435-45.
6. “Rotavirus Clinical Information.” Centers for Disease Control and Prevention, Centers for Disease Control and Prevention, 26
7. Mar. 2021, www.cdc.gov/rotavirus/clinical.html.
8. “Campylobacter (Campylobacteriosis).” Centers for Disease Control and Prevention, Centers for Disease Control and Prevention, 14 Apr. 2021, www.cdc.gov/campylobacter/index.html.
9. Kaakoush, Nadeem O., et al. "Global epidemiology of Campylobacter infection." Clinical microbiology reviews 28.3 (2015): 687-720.
10. “Ecoli.” Centers for Disease Control and Prevention, Centers for Disease Control and Prevention, 2 Feb. 2021, www.cdc.gov/ecoli/ecoli-symptoms.html.
11. Isidean, S. D., et al. "A systematic review of ETEC epidemiology focusing on colonization factor and toxin expression." Vaccine 29.37 (2011): 6167-6178.
12. Siebenga, J. Joukje, et al. "Norovirus illness is a global problem: emergence and spread of norovirus GII. 4 variants, 2001–2007." The Journal of infectious diseases 200.5 (2009): 802-812.
13. Oka, Tomoichiro, et al. "Comprehensive review of human sapoviruses." Clinical Microbiology Reviews 28.1 (2015): 32-53.
14. Navaneethan, U., Giannella, R. Mechanisms of infectious diarrhea. Nat Rev Gastroenterol Hepatol 5, 637–647 (2008). https://doi.org/10.1038/ncpgasthep1264
15. Thielman, Nathan M., and Richard L. Guerrant. “Acute Infectious Diarrhea.” New England Journal of Medicine, vol. 350, no. 1, 2004, pp. 38–47., doi:10.1056/nejmcp031534.
16. Kotloff, Karen L, et al. “Burden and Aetiology of Diarrhoeal Disease in Infants and Young Children in Developing Countries (the Global Enteric Multicenter Study, GEMS): a Prospective, Case-Control Study.” The Lancet, vol. 382, no. 9888, 2013, pp. 209–222., doi:10.1016/s0140-6736(13)60844-2.
17. Baqui, A. H., et al. “Enteropathogens Associated with Acute and Persistent Diarrhea in Bangladeshi Children <5 Years of Age.” Journal of Infectious Diseases, vol. 166, no. 4, 1 Oct. 1992, pp. 792–796., doi:10.1093/infdis/166.4.792.
18. Denno, Donna M., et al. “Diarrhea Etiology in a Pediatric Emergency Department: A Case Control Study.” Clinical Infectious Diseases, vol. 55, no. 7, 2012, pp. 897–904., doi:10.1093/cid/cis553.
19. Vu Nguyen, Trung, et al. “Etiology and Epidemiology of Diarrhea in Children in Hanoi, Vietnam.” International Journal of Infectious Diseases, vol. 10, no. 4, July 2006, pp. 298–308., doi:10.1016/j.ijid.2005.05.009.
20. Albert, M. John, et al. “Case-Control Study of Enteropathogens Associated with Childhood Diarrhea in Dhaka, Bangladesh.” Journal of Clinical Microbiology, vol. 37, no. 11, 18 May 1999, pp. 3458–3464., doi:10.1128/jcm.37.11.3458-3464.1999.
21. Gascón J., et al. “Diarrhea in Children under 5 Years of Age from Ifakara, Tanzania: a Case-Control Study.” Journal of Clinical Microbiology, vol. 38, no. 12, 2000, pp. 4459–4462., doi:10.1128/jcm.38.12.4459-4462.2000.
22. Olesen, B., et al. “Etiology of Diarrhea in Young Children in Denmark: a Case-Control Study.” Journal of Clinical Microbiology, vol. 43, no. 8, 2005, pp. 3636–3641., doi:10.1128/jcm.43.8.3636-3641.2005.
23. Money, Mary E., and Michael Camilleri. "Management of postprandial diarrhea syndrome." The American journal of medicine 125.6 (2012): 538-544.
24. Alam, Nure H., and Hasan Ashraf. “Treatment of Infectious Diarrhea in Children.” Pediatric Drugs, vol. 5, no. 3, Mar. 2003, pp. 151–165., doi:10.2165/00128072-200305030-00002.

25. N. E. Breslow (1996) Statistics in Epidemiology: The Case-Control Study, Journal of the American Statistical Association, 91:433, 14-28, DOI: 10.1080/01621459.1996.10476660
26. Pearce, Neil. "Analysis of matched case-control studies." bmj 352 (2016).
27. “Vaccine Impact on Diarrhea in Africa (VIDA).” University of Maryland School of Medicine, Center for Vaccine Development and Global Health, www.medschool.umaryland.edu/CVD/Vaccines/Studies--Projects/Vaccine-Impact-on-Diarrhea-in-Africa-VIDA/.
28. “Republic of Kenya County Government of Siaya.” County Government of Siaya, 19 Sept. 2020, siaya.go.ke/.
29. “Protocol: Vaccine Impact on Diarrhea in Africa Assessment (VIDA).” KEMRI, no. SSC, ser. No.2996, 23 Mar. 2017. No.2996.
30. Jothikumar, N., et al. “Broadly Reactive TaqMan® Assay for Real-Time RT-PCR Detection of Rotavirus in Clinical and Environmental Samples.” Journal of Virological Methods, vol. 155, no. 2, 2009, pp. 126–131., doi:10.1016/j.jviromet.2008.09.025.
31. Saltelli, Andrea. "Sensitivity analysis for importance assessment." Risk analysis 22.3 (2002): 579-590.
32. R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
33. Kristensen, Peter. "Bias from nondifferential but dependent misclassification of exposure and outcome." Epidemiology (1992): 210-215.
34. Alhamlan, Fatimah S., Ahmed Ali Al-Qahtani, and Mohammed N. Ahmed Al-Ahdal. "Recommended advanced techniques for waterborne pathogen detection in developing countries." The Journal of Infection in Developing Countries 9.02 (2015): 128-135.
35. Reimann, Hobart A., John H. Hodges, and Alison H. Price. "Epidemic diarrhea, nausea and vomiting of unknown cause." Journal of the American Medical Association 127.1 (1945): 1-6.