

# The Global Enteric Multicenter Study (GEMS) of Diarrheal Disease in Infants and Young Children in Developing Countries: Epidemiologic and Clinical Methods of the Case/Control Study

Karen L. Kotloff,<sup>1,2</sup> William C. Blackwelder,<sup>2</sup> Dilruba Nasrin,<sup>2</sup> James P. Nataro,<sup>1,2,a</sup> Tamer H. Farag,<sup>2</sup> Annemieke van Eijk,<sup>3</sup> Richard A. Adegbola,<sup>4,b</sup> Pedro L. Alonso,<sup>5,6</sup> Robert F. Breiman,<sup>7</sup> Abu Syed Golam Faruque,<sup>8</sup> Debasish Saha,<sup>4</sup> Samba O. Sow,<sup>9</sup> Dipika Sur,<sup>10</sup> Anita K. M. Zaidi,<sup>11</sup> Kousick Biswas,<sup>12</sup> Sandra Panchalingam,<sup>2</sup> John D. Clemens,<sup>13,c</sup> Dani Cohen,<sup>14</sup> Roger I. Glass,<sup>15</sup> Eric D. Mintz,<sup>16</sup> Halvor Sommerfelt,<sup>17,18</sup> and Myron M. Levine<sup>1,2</sup>

<sup>1</sup>Department of Pediatrics, <sup>2</sup>Department of Medicine, Center for Vaccine Development, University of Maryland School of Medicine, Baltimore; <sup>3</sup>Child and Reproductive Health Group, Liverpool School of Tropical Medicine, University of Liverpool, United Kingdom; <sup>4</sup>Medical Research Council (United Kingdom) Unit, Fajara, The Gambia; <sup>5</sup>Centre de Recerca en Salut Internacional de Barcelona, Hospital Clinic/Universitat de Barcelona, Spain; <sup>6</sup>Centro de Investigação em Saúde de Manhiça, Mozambique; <sup>7</sup>Global Disease Detection Division, Kenya Office of the US Centers for Disease Control and Prevention, Nairobi; <sup>8</sup>International Centre for Diarrhoeal Disease Research, Dhaka, Bangladesh; <sup>9</sup>Centre pour le Développement des Vaccins, Bamako, Mali; <sup>10</sup>National Institute of Cholera and Enteric Diseases, Kolkata, India; <sup>11</sup>Department of Paediatrics and Child Health, The Aga Khan University, Karachi, Pakistan; <sup>12</sup>Department of Veterans Affairs, Cooperative Studies Program Coordinating Center, Perry Point, Maryland; <sup>13</sup>International Vaccine Institute, Seoul, Korea; <sup>14</sup>Department of Epidemiology and Preventive Medicine, School of Public Health, Sackler Faculty of Medicine, Tel Aviv University, Ramat Aviv, Israel; <sup>15</sup>Fogarty International Center, National Institutes of Health, Bethesda, Maryland; <sup>16</sup>Division of Foodborne, Waterborne, and Environmental Diseases, US Centers for Disease Control and Prevention, Atlanta, Georgia; <sup>17</sup>Centre for International Health, University of Bergen, and <sup>18</sup>Division of Infectious Disease Control, Norwegian Institute of Public Health, Oslo, Norway

**Background.** Diarrhea is a leading cause of illness and death among children aged <5 years in developing countries. This paper describes the clinical and epidemiological methods used to conduct the Global Enteric Multicenter Study (GEMS), a 3-year, prospective, age-stratified, case/control study to estimate the population-based burden, microbiologic etiology, and adverse clinical consequences of acute moderate-to-severe diarrhea (MSD) among a censused population of children aged 0–59 months seeking care at health centers in sub-Saharan Africa and South Asia.

**Methods.** GEMS was conducted at 7 field sites, each serving a population whose demography and healthcare utilization practices for childhood diarrhea were documented. We aimed to enroll 220 MSD cases per year from selected health centers serving each site in each of 3 age strata (0–11, 12–23, and 24–59 months), along with 1–3 matched community controls. Cases and controls supplied clinical, epidemiologic, and anthropometric data at enrollment and again approximately 60 days later, and provided enrollment stool specimens for identification and characterization of potential diarrheal pathogens. Verbal autopsy was performed if a child died. Analytic strategies will calculate the fraction of MSD attributable to each pathogen and the incidence, financial costs, nutritional consequences, and case fatality overall and by pathogen.

**Conclusions.** When completed, GEMS will provide estimates of the incidence, etiology, and outcomes of MSD among infants and young children in sub-Saharan Africa and South Asia. This information can guide development and implementation of public health interventions to diminish morbidity and mortality from diarrheal diseases.

<sup>a</sup>Department of Pediatrics, University of Virginia School of Medicine, Charlottesville, Virginia.

<sup>b</sup>GlaxoSmithKline Biologicals, Global Medical Affairs, Wavre, Belgium.

<sup>c</sup>Center for Global Infectious Diseases, University of California, Los Angeles.

Correspondence: Karen L. Kotloff, MD, University of Maryland School of Medicine, Center for Vaccine Development, 685 W Baltimore St, HSF 480, Baltimore, MD 21201 (kkotloff@medicine.umaryland.edu).

**Clinical Infectious Diseases** 2012;55(S4):S232–45

© The Author 2012. Published by Oxford University Press on behalf of the Infectious Diseases Society of America. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1093/cid/cis753

There has been substantial progress toward meeting Millennium Development Goals for child survival during the past 2 decades, such that under-5 mortality rates have decreased in every developing region of the world. Nonetheless, the rates have fallen more sharply in wealthier areas [1], resulting in a large and growing share of deaths in the poorer developing regions. As further declines are made possible by expanding interventions that target the principal causes of death and focus on the most vulnerable children, the availability of accurate, up-to-date assessments at country levels becomes even more important to guide strategic planning and resource allocation. This is especially true for sub-Saharan Africa and South Asia, where 50% and 32%, respectively, of the estimated annual 7.6 million under-5 deaths are now concentrated and where current, systematically collected information on the burden and major causes of child death is lacking [2].

Diarrheal diseases continue to be major causes of childhood mortality in developing countries. The proportion of deaths attributed to diarrhea among children 1–59 months of age is estimated to be 25% in Africa and 31% in South Asia [3]. These estimates were calculated by abstracting studies published between 1980 and 2009 that utilized verbal autopsies (postmortem interviews of family members) to assign cause of death in representative populations. Statistical models were applied to derive estimates of diarrhea-specific mortality and to extrapolate across countries and regions. Modeled estimates of disease burden are an invaluable metric for assessing progress toward achieving health objectives and for estimating the impact of various interventions; however, this approach faces limitations imposed by the quality, scope, age, and consistency of the underlying data. Analyses of the causes of childhood death based on verbal autopsies are subject to misclassification [4, 5], and if they include studies performed over several decades, the results may not reflect the current situation. Without concomitant morbidity assessments, one cannot determine to what extent secular trends in declining disease-specific mortality represent lower disease incidence or diminished case fatality (which can have different determinants and respond to different interventions). Rigorously conducted, prospective, population-based studies can be used to strengthen modeled disease burden estimates [6–8]. Moreover, such studies are essential for providing the detailed information needed to design new and improved interventions to prevent and treat the most life-threatening and disabling episodes, which, in the case of diarrhea, would include knowledge about the etiology, risk factors, nutritional sequelae, and case fatality.

We conducted the Global Enteric Multicenter Study (GEMS), a 3-year, prospective, age-stratified, matched case/control study of moderate-to-severe diarrhea (MSD) among children 0–59 months of age belonging to a censused population and seeking care at hospitals and health centers at 7 sites

located in sub-Saharan Africa and South Asia. A common research protocol with standardized epidemiologic and microbiologic methods was used to facilitate intersite comparisons and allow aggregate estimates of etiology and incidence. This paper describes the study design, including site selection, a surveillance system to characterize the demography and healthcare utilization practices of the catchment population, methods for enrollment, data collection and follow-up of case and control children, and quality control activities. We discuss challenges encountered in the implementation of a large study involving heterogeneous populations located in resource-poor settings.

## METHODS

### Objectives of the Study

The primary objective of GEMS was to measure the population-based burden, microbiologic etiology, and adverse clinical consequences of MSD in developing countries, overall and by age, pathogen, site, and clinical syndrome (simple nonbloody diarrhea, dysentery, or profuse watery diarrhea). The adverse clinical consequences of interest included growth faltering according to World Health Organization (WHO) standards [9], persistent diarrhea lasting  $\geq 14$  days, and death. The secondary objectives were (1) to determine the antigenic and genotypic characteristics of the leading pathogens to guide vaccine development; (2) to elucidate the risk factors attributable to the host, the microorganism, and the environment that are associated with the occurrence and adverse clinical outcomes of MSD; (3) to estimate the public and private financial costs, both direct and indirect, incurred during an episode of MSD; and (4) to create a central repository of well-characterized clinical specimens and isolated etiologic agents that can be shared with other investigators for future research.

### Site Selection Criteria

Seven field sites were selected among countries in sub-Saharan Africa (Kenya, Mali, Mozambique, and The Gambia), and South Asia (Bangladesh, India, and Pakistan) with moderate to high under-five childhood mortality (Table 1). To create a broad view of enteric disease epidemiology, we chose sites that together exemplified a spectrum of child health indicators, with variations in the prevalence of malaria and human immunodeficiency virus (HIV) infection, and a mixture of urban, rural, and periurban settings (Table 1). Sites were required to have access to a population that had been or could undergo a census accompanied by an address system to allow households to be revisited in the future, and to 1 or more healthcare facilities that provide care to children from that population with diarrhea. Infrastructure with the potential for computerized data management, secure freezer storage, at

**Table 1. Selected Child Health Indicators Available in 2005 and Used to Guide Site Selection<sup>a</sup>**

Country	City	Partner	Setting	No. SHCs in CCS	Population <5 y <sup>c</sup>	GNI per Capita (US\$)	U5MR <sup>b</sup> (Country Rank)	% HIV+ (15–49 y) <sup>d</sup>	Malaria Rate <sup>e</sup>	% <5 y Wasted <sup>f</sup>	% <5 y Stunted <sup>f</sup>	National Statistics			
												% Using Improved or Adequate <sup>g</sup>		% <5 y Receiving ORS <sup>h</sup>	% 1 y DPT3 <sup>i</sup>
Mali	Bamako	Centre pour le Développement des Vaccins du Mali (CVD-Mali)	Urban	9	31 768	290	220 (7)	1.9	62.2	11	38	76	59	45	69
The Gambia	Basse	Medical Research Council (MRC)	Rural	5	29 076	310	123 (37)	1.2	ND	9	19	77	46	38	90
Mozambique	Manhiça	Centro de Investigação em Saúde de Manhiça (CISM)	Rural	5	15 380	210	158 (24)	12.2	269.7	4	41	24	14	33	72
Kenya	Nyanza Province	CDC/Kenya Medical Research Institute (KEMRI) Research Station/CDC	Rural	11	21 603	390	123 (37)	6.7	3.9	6	31	46	43	15	73
India	Kolkata, W. Bengal	National Institute of Cholera and Enteric Diseases (NICED)	Urban	2	13 416	530	87 (54)	ND	1.7	16	46	96	58	22	70
Bangladesh	Mirzapur	International Center for Diarrheal Disease Research, Bangladesh (ICDDR,B)	Rural	1	25 560	400	69 (62)	ND	0.4	10	45	72	39	35	85
Pakistan	Karachi (Bin Qasim Town)	Aga Khan University	Peri urban	7	25 659	470	103 (47)	0.1	0.8	13	37	87	35	33	67

Abbreviations: CCS, case/control study; CDC, Centers for Disease Control and Prevention; DPT3, complete coverage with diphtheria-pertussis-tetanus vaccine; GNI, gross national income; HIV, human immunodeficiency virus; ND, no data; ORS, oral rehydration solution; SHC, sentinel health centers where children with moderate-to-severe diarrhea were enrolled in the CCS; U5MR, under-5 mortality rate.

<sup>a</sup> All data pertain to 2003, with the exception of access to improved water and adequate sanitation, which pertain to 2002 [42], and the population <5 years, as described below.

<sup>b</sup> Value is calculated per 1000 live births and ranked out of 192 countries for 2003 [42].

<sup>c</sup> The population <5 years of age represents the median value from sequential demographic surveillance system rounds conducted during the case/control study.

<sup>d</sup> Prevalence of HIV (percentage) among 15- to 49-year-olds, as of end of 2003 [42].

<sup>e</sup> Standardized reported malaria rate per 1000 population, 2003 for all countries but Kenya (2002) [43].

<sup>f</sup> Percentage of children <5 years of age with wasting or stunting graded as moderate or severe [42].

<sup>g</sup> Data shown pertain to urban areas when the study site is urban sites and rural areas when the site is rural. Data for rural areas were considered most appropriate to represent the study site in Pakistan [42].

<sup>h</sup> Percentage of children <5 years of age with diarrhea receiving oral rehydration and continued feeding 1994–2003 [42].

<sup>i</sup> Percentage of children who received DPT3 by 1 year of age [42].

least intermittent internet transmission, and the ability to ship specimens and strains abroad had to be available, with capabilities to perform coprocultures, antigen-detection tests, and nucleic-acid based assays.

### **Establishing a Sampling Frame for the Case/Control Study and Selecting Health Centers for Case Recruitment**

The census at each site will enable population-based estimates of the outcomes of interest. Each census was continually updated using a demographic surveillance system (DSS) in which the households were visited every 4–6 months to record pregnancies, births, deaths, and migrations in and out of the area. Between DSS visits, we enlisted a community reporter from each neighborhood to meet weekly with local leaders (religious figures, political representatives, and elders) and midwives to detect births and deaths among children 0–59 months of age. The reporter visited near-term pregnant women as an additional means of capturing births. Keeping the DSS current was necessary to maintain an accurate sampling frame from which to select matched community controls for the case/control study, and for the timely performance of verbal autopsies, as described below.

In preparation for the case/control study, we performed a Health Care Utilization and Attitudes Survey (HUAS). An age-stratified sample of approximately 1000 children aged 0–59 months per site randomly selected from each updated DSS dataset was visited at home, and parents/primary caretakers were asked whether their child had experienced diarrhea during the previous 14 days. If so, the presence of findings suggestive of MSD was solicited (sunken eyes, wrinkled skin, hospitalization, receipt of intravenous hydration, or dysentery), and source(s) of healthcare were recorded. These data were used to adjust the size of the DSS population at each site as necessary to contribute the requisite number of cases of MSD to each age stratum, and to select 1 or more “sentinel” health centers (SHCs) serving the DSS population at each site (Table 1) as venues for the case/control study based on their potential to capture MSD cases from the DSS.

During the second and third years of the case/control study, an abbreviated HUAS questionnaire (designated “HUAS-lite”) was administered to caretakers of approximately 1000 randomly selected children aged 0–59 months (age-stratified) approximately every 4 to 6 months in association with the DSS interviews. HUAS-lite data were used to refine the selection of SHCs, to estimate the extent to which children with MSD who seek care at SHC are representative of children with MSD in the DSS population (by comparing features of those who do and do not seek care), and to calculate the proportion of children with MSD who sought care at the SHCs at each site (*r* value) as a means of extrapolating the overall and pathogen-specific MSD episodes enumerated at the SHCs to

derive the incidence estimates for the entire DSS population (see Blackwelder, et al, this supplement).

### **Process Development and Training**

Paper case report forms (CRFs) were created and translated into the 4 languages spoken by the interviewers (English, French, Portuguese, and dual Dholuo and English) according to preference of the local study teams. Interviews were always conducted in the native language of the respondent. Initial versions of the CRFs were field tested, then modified as needed at a 4-day study development meeting attended by each site’s senior clinical investigators and study coordinators.

A pilot case/control study was conducted for approximately 3 months followed by the full, 36-month case/control study. Before the pilot and full-study initiations, we conducted a 5-day training program at each site, using interactive adult learning techniques with group participation, role playing, small group practice sessions, and evaluations of competency. We compiled an interviewers’ and a supervisors’ manual of procedures which served as the basis for training sessions. The curriculum covered the principles of human subjects research and elements of good clinical practices [10], how to conduct an interview (eg, issues of privacy, building rapport with the respondent, and asking questions in a nonjudgmental way), perform a focused physical examination, collect, process, and transport stool specimens, and document observations of water and sanitation facilities. The meaning of each question and response choice was discussed. Terms were defined, using pictures and graphics whenever applicable. Participants practiced a standardized format for handwriting English letters and numbers to reduce frequency of errors in data recording and entry. The supervisors’ manual and training focused on supportive supervision techniques, training, handling underperforming staff, quality management, tracking study activities, and performing oversight, spot checks, and reinterviews to ensure the validity and reliability of the data.

We enlisted the assistance of an experienced anthropometrist to train the clinical and field staff at 1 site in Asia and 1 site in Africa in obtaining length/height, weight, and mid-upper arm circumference (MUAC) measurements. These 3-day training sessions also served the purpose of training the Epidemiology Team from the core site at the Center for Vaccine Development (CVD), University of Maryland, Baltimore (K. L. K., D. N., and T. H. F.), who then conducted similar training at the remaining 5 sites. On the third day of training, 10 children (5 aged 0–23 months and 5 aged 24–59 months) participated in a standardization session in which each trainee performed 2 independent measurements of the length/height and MUAC of each child. Intrarater reliability and validity were calculated using the anthropometrist (and later the CVD epidemiologist) as the “gold standard.”

A difference of  $>0.5$  cm was considered unacceptable when comparing a trainee's 2 measurements of the same child or when comparing the trainee's and the gold-standard measurements of the same child. Trainees with unacceptable performance were retrained until competency was achieved. A gold-standard measurer was identified at each site to supervise field measurements and to train all staff including newly hired staff every 4–6 months. Technical error of measurement and average bias will be calculated to assess inter- and intraobserver reproducibility as well as validity of measurements [11].

### Scientific Oversight

We recognized that controversies persist regarding the most appropriate case definitions and detection methods for studying diarrhea, and that the approach chosen would impact the resultant estimates of disease burden [12, 13]. Consequently, as part of a consensus-building process, we assembled a Steering Committee on Epidemiologic and Clinical Issues comprising the lead investigators from each site, and a multinational group of 6 experts in diarrheal disease. The committee vetted the clinical protocol at the initial meeting. Thereafter, the committee was convened annually and on an ad hoc basis as issues arose. Once the study was initiated, external experts were assembled to form a Steering Committee on Nutritional Issues and a Steering Committee on Biostatistical Issues to review the analysis plan and to provide guidance as issues arose. In the final months of the study, an International Strategic Advisory Committee was formed to critically review the methodology and results and to advise the funding agency about the significance of the findings to inform its strategic planning for the future (Farag et al, this supplement).

### Case Definition of MSD and Other Study Outcomes

The initial step in eligibility screening was the selection of children who fulfilled the WHO definition of diarrhea ( $\geq 3$  abnormally loose stools in 24 hours [12]). In subsequent steps we identified cases of MSD, the primary outcome of interest, intending to capture diarrheal illnesses that would not be expected to resolve spontaneously without medical intervention or without sequelae, because these illnesses constitute priorities for development of vaccines and other new or improved preventive and therapeutic strategies. We reasoned that episodes that would qualify as MSD fell into 2 general categories: (1) those accompanied by dehydration to a degree that the child's survival would likely depend on access to life-saving rehydration fluids, and (2) those with evidence of inflammatory destruction of the intestinal mucosa, thereby at increased risk for disabling sequelae (such as persistent diarrhea [14] and stunting [15]) or death [16, 17].

To capture children who had potentially life-threatening diarrheal dehydration, we adapted the WHO definition of

dehydration to our case definition of MSD [18, 19], choosing the most objective signs (sunken eyes more than usual and slow or very slow recoil after an abdominal wall "skin pinch"). In addition, we included the determination by a healthcare provider that the severity of dehydration warranted administration of intravenous fluids. Although not part of our case definition, other signs of dehydration proposed by WHO were documented, including restlessness or irritability and drinking eagerly or appearing thirsty (considered to be signs of "some" dehydration), and lethargy, loss of consciousness, inability to drink, or drinking poorly (as signs of "severe" dehydration). During analysis we will explore the impact on the study findings of including these other signs of dehydration in the definition of MSD. We considered adopting as inclusion criteria elements of systems used widely to define severe illness in rotavirus vaccine trials [20–22]. However, many of the components, such as total duration and maximum severity of diarrhea, vomiting, and fever, can only be determined in retrospect when the episode is resolving or resolved, at which point the decision to include a child in GEMS would already have been made. Instead, our approach has been to collect this information for exploration during analysis.

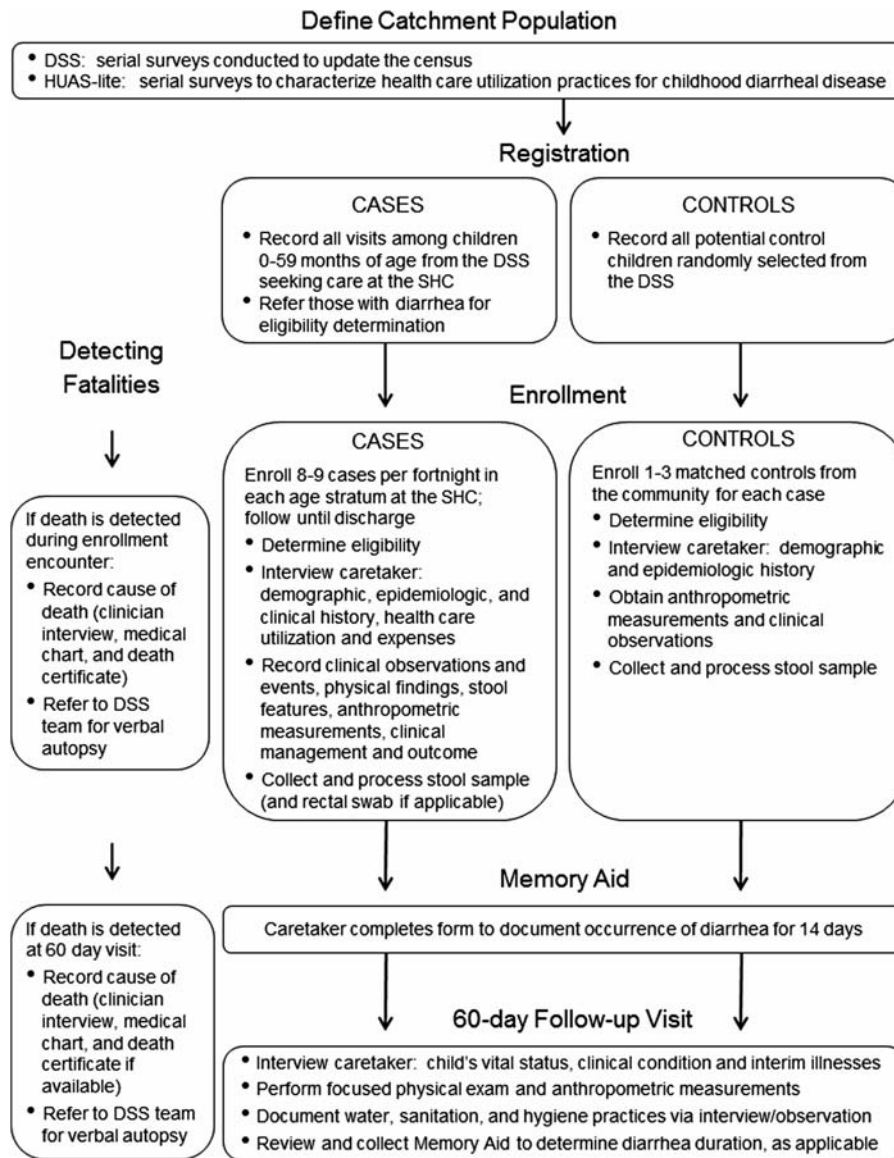
To capture children with evidence of diarrheal diseases caused by inflammation and mucosal injury in the case definition, we enrolled children with dysentery. Because there is no marker to predict which cases of dysentery are likely to experience clinically significant intestinal damage, we included all children with diarrhea who passed at least 1 stool containing visible blood according to either the caretaker or the clinician. Finally, we included children with diarrhea who appeared sufficiently ill to prompt the healthcare provider to recommend overnight admission to the hospital.

We restricted enrollment to children with acute MSD ( $\leq 7$  days' duration) to maximize the opportunity to identify the inciting pathogen and to collect new episodes that can be used together with DSS and HUAS data to estimate annual incidence rates. We defined an episode of diarrhea as days with diarrhea beginning after at least 7 diarrhea-free days and ending when diarrhea is not present for 7 days [23, 24]. Although the WHO definition of a new episode of diarrhea requires only 3 diarrhea-free days [12], we chose a longer interval (as have other investigators [12, 25]) to increase our margin of certainty that the episode was new, recognizing that this approach could underestimate the incidence of MSD.

### Case Ascertainment

Cases of MSD were identified in SHCs (hospital, urgent care facilities, and community clinics) to capture those illnesses that are most severe and that collectively constitute a significant cost in healthcare services, and thus would be targeted for prevention by vaccines and other interventions (Figure 1).





**Figure 1.** Flow diagram illustrating major study activities. Abbreviations: DSS, demographic surveillance system; SHC, sentinel health center; HUAS, Health Care Utilization and Attitudes Survey.

GEMS staff were situated in the intake area at each SHC to complete a registration log documenting each visit made by a child 0–59 months old belonging to the DSS. The GEMS registrar was given access to the DSS database to verify that a child belonged to the DSS, and to record each enrolled child’s unique DSS number as a means for determining, at a later date, who was enrolled into GEMS more than once. Each visit was assigned a unique screening identification number, and the registrar recorded the date and time the child entered the SHC; the child’s age, sex, and village/neighborhood; whether the child had diarrhea; and whether the child was hospitalized. The GEMS registrar referred all children from the DSS

who were aged 0–59 months and had diarrhea to a GEMS clinician. The clinician informed the parent/primary caretaker about the study, determined the child’s eligibility (Table 2), and obtained informed consent. If an eligible child was not enrolled, the reasons for nonenrollment were documented (eg, refusal, missed opportunity, stool sample inadequate or not obtained, 14-day quota filled, or child died before enrollment).

Each site aimed to enroll approximately 220 MSD patients per year into each of 3 age strata: 0–11 months, 12–23 months, and 24–59 months, totaling 1980 cases over 3 years. To ensure even sampling throughout the year, the target was to enroll approximately 8–9 cases per age stratum (25–26

**Table 2. Inclusion Criteria for Cases**

1. Child is 0–59 mo of age
2. Child belongs to the demographic surveillance system population at the site
3. Child is not currently enrolled as a case (meaning previously enrolled and pending 60-day visit) <sup>a</sup>
4. Child meets case definition of diarrhea ( $\geq 3$ abnormally loose stools in the previous 24 h)
5. Diarrhea episode is: <ul style="list-style-type: none"> <li>• Acute (onset within 7 d of study enrollment) and</li> <li>• Represents a new episode (onset after <math>\geq 7</math> diarrhea-free days) [23, 24]</li> </ul>
6. Diarrhea is moderate-to-severe, meaning that the child met at least 1 of the following criteria: <ul style="list-style-type: none"> <li>• Sunken eyes, confirmed by parent/primary caretaker as more than normal</li> <li>• Loss of skin turgor (determined by abdominal skin pinch (slow return [<math>\leq 2</math> s] or very slow return [<math>&gt; 2</math> s]))</li> <li>• Intravenous rehydration administered or prescribed</li> <li>• Dysentery (visible blood in a loose stool)</li> <li>• Hospitalized with diarrhea or dysentery</li> </ul>

<sup>a</sup> A child was eligible to be enrolled as a case irrespective of whether he or she had been included as a case or as a control previously; whereas cases were eligible for reenrollment only after the 60-day follow-up visit had been completed, controls could be enrolled as a case at any time they met the criteria.

cases overall) per fortnight. This strategy prevented the strata from being filled prematurely in seasons with high volume and respected the capacity limitations of the clinical and microbiology personnel, but because all DSS children with MSD were recorded, temporal increases in the case load of MSD and of specific diarrheal pathogens could be measured. Analyses for events that might have seasonal variation will take into account the sampling fraction of MSD for each period.

### Control Selection

For each child with MSD included in the study, we enrolled 1–3 control children without diarrhea from the DSS community (Figure 1) within 14 days of presentation of the index case. Sites tracked their ability to fill each age stratum on a fortnightly basis and followed an algorithm to determine the number of controls to enroll: 1:1 case:control matching if 7–9 cases were enrolled; 1:2 matching if 4–6 cases were enrolled, and 1:3 matching if  $\leq 3$  cases were enrolled. At least 4 children who met the matching criteria (Table 3) were randomly selected from the DSS database as potential controls. A field worker visited the home of selected children sequentially and explained all aspects of the study. If the parent/primary caretaker expressed interest and the child met eligibility criteria (Table 3), informed, written consent was obtained and arrangements were made to collect a stool sample, as described

**Table 3. Inclusion Criteria for Controls**

1. Resides in demographic surveillance system area
2. Matched to the index case as follows: <ul style="list-style-type: none"> <li>• Age:           <ul style="list-style-type: none"> <li>◦ <math>\pm 2</math> mo for cases 0–11 mo</li> <li>◦ <math>\pm 4</math> mo for cases 12–59 mo</li> <li>◦ May not exceed the stratum boundaries of the case, eg, a control for an 11-mo-old case must be between the ages of 9 and 11 mo and a control for a 13-mo-old must be between the ages of 12 and 17 mo</li> </ul> </li> <li>• Same sex</li> <li>• Residence: lives in the same or nearby village/neighborhood as the case<sup>a</sup></li> <li>• Time: enrolled within 14 d of presentation of the case</li> </ul>
3. No diarrhea in the previous 7 d <sup>b</sup>

<sup>a</sup> Each site followed an algorithm beginning with the case's village/neighborhood, and then proceeding to villages/neighborhoods located at an increasing distance from the case's village/neighborhood until a control could be identified.

<sup>b</sup> Control children will be included in the analysis irrespective of whether they developed diarrhea after enrollment.

below. Reasons for not enrolling a selected child were documented (eg, refusal, not found at home after 3 attempts to contact, or failed to produce an adequate stool sample).

### Data Collection at Enrollment From Cases and Controls

Case enrollment interviews took place at the SHC whereas control caretakers were interviewed at home. To facilitate linkage of our results with existing databases, we designed our caretaker interviews to include questions found in the primary sources of population-based data used to estimate child mortality in developing countries, such as the UNICEF-supported Multiple Indicator Cluster Surveys and the US Agency for International Development–supported Demographic and Health Surveys [1]. Demographic information collected about the case or control and his/her household (defined as a group of people who share a cooking fire) included maternal education and household size (including the number of children  $< 5$  years old). Building materials and household possessions were documented (to assess potential risk factors for illness and as indicators for constructing a wealth index for each site [26]). Questions addressed handwashing practices and access and availability of improved water and sanitation facilities [27], animals on the premises, water treatment, sharing sanitation facilities, and disposal of the child's feces. The caretakers were queried about the child's clinical signs and symptoms; how the illness was managed prior to the SHC visit (the reference point was the current illness for cases and the most recent diarrheal illness for controls), for example, use of oral rehydration solutions, zinc, antibiotics, traditional medicines, continued feeding and fluid administration; healthcare seeking

behavior; and breastfeeding practices. The household's direct (out-of-pocket) and indirect (eg, income lost while caring for the sick child) expenditures at home and at the SHC were tabulated. The GEMS staff measured the child's axillary digital temperature, respiratory rate (the average of 2 measures obtained using a rate counter), anthropometric dimensions (described below), and clinical signs of malnutrition (bipedal edema, wasting, flaky skin, and sparse or loose hair).

A clinician examined all cases to document signs of dehydration, including skin pinch return (graded as slow  $\leq 2$  seconds or very slow  $> 2$  seconds), sunken eyes (more than usual confirmed by the parent/primary caretaker), dry mouth (graded as somewhat or very dry), and mental status changes, and examined the child's rectum for signs of prolapse. A member of the clinical team examined the child's stool (if available) for visible blood and recorded any rehydration fluids, zinc, and antibiotics prescribed or administered at the SHC. Cases who remained in the SHC while receiving rehydration fluids were reweighed at 4 hours and again at discharge from the SHC, as applicable, at which points the clinician reassessed the child for signs of dehydration and determined his/her vital status and weight.

#### Collection and Processing of Stool Specimens

To qualify for enrollment, each case and control had to produce a whole stool specimen that weighed at least 3 grams. In one site (Kolkata), stool is routinely collected from hospitalized children by passing a small catheter into the child's rectum and aspirating loose stool using a syringe attached to the other end [28]; at all other sites, whole stool was passed naturally per rectum. Cases were required to provide a whole stool specimen within 12 hours of registration at the center. To collect stool from control children at home, study staff provided the caretaker with a polystyrene foam container containing a cold pack, a culturally accepted stool collection device (such as a plastic "potty"), plastic gloves, a specimen cup, and a scoop. The field worker returned to the household the next morning (or sooner if called by the parent) to retrieve the stool sample and perform the study interview. Because children usually defecate in the morning, and since families frequently used cellular phones to alert the GEMS field team that the child had produced a stool, we were able to fulfill the study requirement of retrieving and processing freshly passed stools from cases and controls within 6 hours of evacuation. Processing involved inserting 2 cotton-tipped swabs into the specimen (if dysentery was present, an area of blood or mucus was swabbed); one swab was placed into modified Cary-Blair transport medium [29], and the other into buffered glycerol saline [30]. Remaining whole stool was retained in an empty vial. The processed sample was placed immediately into either a specimen refrigerator or a polystyrene foam container containing a fresh cold pack, to be delivered to the laboratory and

plated within 24 hours. Stools were evaluated for bacterial pathogens (eg, *Salmonella*, *Shigella*, *Campylobacter*, *Aeromonas*, and *Vibrio* species, and 5 diarrheal pathotypes of *Escherichia coli*), protozoal agents (*Entamoeba histolytica*, *Giardia lamblia*, and *Cryptosporidium* species), and viruses (rotavirus, adenovirus, norovirus, sapovirus, and astrovirus) using microbiologic methods described elsewhere in this supplement (see Panchalingam et al, this supplement).

If antibiotics were to be administered to cases before the whole stool specimen was collected, 2 rectal swabs also were obtained. The cotton tip was moistened with transport media, gently inserted into the child's rectum, rotated 360°, and immediately inserted into transport media, as described above for whole stool swabs. Only swabs stained or covered with fecal material were accepted by the laboratory. This strategy permitted collection of an adequate sample for bacteriology (ie, rectal swabs) prior to antibiotic administration as well as a whole stool for identification of pathogens that are best detected in whole stool but are not expected to be affected by antibiotic administration (see Panchalingam et al, this supplement).

#### Memory Aid for Recording Diarrheal Episodes in Cases and Controls During the 14 Days After Enrollment

We created a memory aid suitable for use by adults regardless of literacy (Figure 2). The data will be used to detect the occurrence of persistent diarrhea in cases and to explore whether the inclusion of control children who developed diarrhea within 7 days after enrollment impacted the association between specific pathogens and MSD. The tool was developed in collaboration with a representative from the Malian Office of Literacy and modified in response to focus groups and field testing at each site. After receiving training at the enrollment visit, each day for the next 14 days the parent/primary caretakers marked whether the child had normal stools only, or diarrhea (passage of  $\geq 3$  abnormally loose stools in the previous 24 hours). The aid was reviewed with the caretaker at the 60-day follow-up visit to resolve missing or unclear markings and then collected. Diarrhea that continued unabated through day 14 will be termed "persistent diarrhea"; diarrheal episodes that continued beyond day 14 (the last day the memory tool collected data) were not systematically tracked.



#### Clinical and Epidemiologic Data Collected at the Single Household Follow-up Visit

GEMS field workers visited the household of each case and control child approximately 60 days after enrollment (acceptable range, 50–90 days). They assessed vital status, recorded interim medical events, took the child's axillary temperature, and performed anthropometric measurements. They directly observed the household's drinking water sources, storage containers, and treatment practices, and tested the water for



Please complete this form every day for each of the next 14 days.

1. Each morning when you wake up, decide whether your child had diarrhea during the previous day. *Diarrhea* means that your child passed 3 or more loose or watery stools that were not normal for him or her on that day.
2. Go to the correct day. "☉" means today. "☉☉" means tomorrow, and so on. A day begins when you wake up in the morning and ends when you wake up the next morning.
3. If your child had diarrhea that day, mark "X" in the dark box for that day [X]. If your child did not have diarrhea, mark "X" in the white box for that day [X]. **Each day, make only one "X"**.
4. If you forget a few days, try to start again on the correct day.
5. Keep this form in a safe place. We will come to your house to collect it in 60 days.

	DIARRHEA	NORMAL
		
(1) ☉ (today)		
(2) ☉☉		
(3) ☉☉☉		
(4) ☉☉☉☉		
(5) ☉☉☉☉☉		
(6) ☉☉☉☉☉☉		
(7) ☉☉☉☉☉☉☉		
(8) ☉☉☉☉☉☉☉☉		
(9) ☉☉☉☉☉☉☉☉☉		
(10) ☉☉☉☉☉☉☉☉☉☉		
(11) ☉☉☉☉☉☉☉☉☉☉☉		
(12) ☉☉☉☉☉☉☉☉☉☉☉☉		
(13) ☉☉☉☉☉☉☉☉☉☉☉☉☉		
(14) ☉☉☉☉☉☉☉☉☉☉☉☉☉☉		

**Figure 2.** Memory aid completed by the caretaker to document the occurrence of diarrhea for 14 days after enrollment of cases and controls.

chlorine if the household reported that they treated it. They examined the sanitation facilities and noted whether fecal contamination was present, and observed hygiene indicators, such as the proximity of soap to the hand washing station.

### Anthropometric Measurements

Weight, length/height, and MUAC were measured for each case and control at enrollment and at the 60-day follow-up visit as previously described [31]. Weight (to the nearest 0.1 kg) was recorded prior to administration of rehydration fluids with the child naked or in light clothing using a digital scale

that was calibrated at least weekly (model 314, Tanita Corp of America, Arlington Heights, Illinois); for children 0–23 months of age, the weight of the mother alone and with the child was recorded, and the child’s weight was computed during analysis. The length of children 0–23 months of age or those who were older but unable to stand unassisted was measured (to the nearest 0.1 cm) in the recumbent position using a board with a fixed head and sliding foot piece (Shorr Productions, Olney Maryland). The same apparatus was used to measure standing height in children 2 years of age and older. A 25-cm paper single-slotted insertion tape was used to

measure MUAC to the nearest 0.1 cm (Shorr Productions). Length/height and MUAC were each measured thrice; the average will be calculated during analysis [32].

### **HIV Substudy**

In appreciation of the importance of HIV infection on the incidence and outcomes from diarrheal disease, including an increased likelihood of dying from an episode of diarrhea compared with HIV-infected children [33, 34], we considered including systematic HIV testing as part of the initial study design but concluded that it was beyond the scope of our capabilities. As the study progressed, national guidelines for provider-initiated counseling and testing were adopted at the Kenya and Mozambique study sites (the only 2 GEMS study sites with high HIV seroprevalence in adults), and home-based counseling and testing has been implemented at the Kenya site. As a result, during the last 2 years of GEMS, we incorporated voluntary HIV testing or the ability to link to existing HIV test results of mothers and children into the study protocol at these 2 sites. Informed consent was obtained to link HIV test results (for participating child and his/her mother) to GEMS data. We will compare frequency, outcomes, and etiologies of episodes of diarrheal diseases among infected and uninfected children born to infected mothers, and among uninfected children born to uninfected mothers.

### **Detection of Deaths and Performance of Verbal Autopsy**

Two parallel systems were in place at all sites to detect deaths. The GEMS team ascertained deaths among children enrolled in the case/control study during the enrollment encounter and at the 60-day follow-up visit. Concomitantly, the DSS teams identified all under-5 deaths regardless of enrollment status. In either case, the DSS team obtained a verbal autopsy using WHO standardized questionnaires with minor modifications [35]. Local customs were followed to respect the mourning period after which a family could be contacted. Whenever possible, information on the cause of death was collected from the medical chart, the healthcare provider, and the death certificate for use as a means to validate the results of the verbal autopsy [36]. A uniform algorithm will be used across all sites to determine the cause of death. The mortality associated with diarrhea and dysentery among enrolled and nonenrolled children will be calculated.

### **Sample Size Considerations**

A sample size at each site of approximately 600 analyzable cases and 1–3 matched controls per stratum was chosen to provide 80% power (2-sided  $\alpha = .05$ ) for site stratum-specific comparison of the proportion of cases and controls in whom a specific enteropathogen is identified, if a specific pathogen is identified in at least 5.8% of cases and 2.5% of controls. In the

event that the proportion of cases with a specific pathogen exceeds 5.8%, the absolute difference between cases and controls needed to achieve statistical significance increases. For example, this sample size will give 80% power to find a significant difference if the proportion is 9.8% in cases vs 5.5% or less in controls. To compensate for dropout, migration, and other losses to follow-up of up to 10%, we planned to enroll a total of 660 cases per stratum per site to achieve the desired number of analyzable cases and controls.

### **Ethical Considerations and Oversight**

GEMS was designed as an observational study that confers minimal risk, and is expected to generate information that can be used by the scientific, public health, policy, and healthcare provider communities to improve the prevention and treatment of diarrheal diseases in the future, both at the site level and globally. Each site was expected to follow WHO guidelines for the clinical diagnosis and management of diarrheal disease, which represent a universal standard of care [18]. We provided supplementary funding for procurement of medical supplies (eg, oral rehydration solution, antibiotics, intravenous cannulae, and fluids) to be used to treat patients with diarrhea at the discretion of each participating SHC. We did not attempt to systematically introduce newer aspects of diarrhea management, such as low-osmolality oral rehydration solution and zinc, in sites lacking national policies to guide usage and trained health providers to administer these products [37]. At the time of study initiation, no site routinely performed the spectrum of assays provided by GEMS to detect potentially treatable pathogens (bacterial culture and immunoassays for protozoa). Therefore, the sites were expected to ensure timely provision of the GEMS results to the clinicians for use in case management. We shared interim results (eg, distribution of pathogens, management of diarrhea, and sequelae) with the investigators and the international community annually at investigators' meetings and at scientific conferences, and each site received a cleaned dataset each year that could be used for more detailed exploration.

The clinical protocol, consent forms, CRFs, and other supporting documents were approved prior to initiation of the study by the ethics committees and applicable scientific review boards at the University of Maryland School of Medicine, and the committees overseeing each site and their collaborating partners from other institutions. Amendments and annual reports underwent ethics committee review. Consent forms were translated into 11 local languages and modified according to the standards at each site; full approval of the University of Maryland ethics committee required back-translation into English and certification by an independent bilingual speaker that the 2 versions were identical. Individual, informed consent was obtained from the parent/primary caretaker of

each participant prior to study activities. When the person supplying consent was illiterate, an impartial third party witnessed the consent process and signed the consent document. Some sites additionally obtained “community consent” from local leaders who were convened in a public forum to discuss the study aims, procedures, potential risks, and benefits.

### **Data Flow, Management, and Analysis**

A data coordinating center (DCC) was responsible for centralized data management as described elsewhere (see Biswas et al, this supplement). In brief, sites transmitted completed CRF pages to the DCC using a variety of electronic formats, but primarily by secure file transfer protocol (SFTP). Although the CRFs were printed in different languages, the structure of the fields was maintained to permit generation of a single database containing data from all sites. The DataFax software system (Clinical DataFax Systems, Hamilton, Ontario, Canada) was used to build and manage the master database and aided in the electronic validation process based on character recognition software. Timelines for transmission of data, data queries, and query resolution were established. A system of security measures and backup procedures preserved the integrity of the data and ensured restoration capabilities.

The GEMS analytic plan (see Blackwelder et al, this supplement) addressed 3 main goals: (1) to determine the major pathogens responsible for MSD, taking into account the prevalence of each pathogen, the frequency of asymptomatic infection in controls, and the presence of multiple pathogens; (2) to determine the pathogen-specific attributable fraction of MSD by age within each site and to extend these estimates to the DSS population; and (3) to identify independent risk factors, including demographic, environmental, and socioeconomic factors as well as pathogens, for MSD and other outcomes of interest (especially death and child growth) using multivariable models.

### **Quality Management**

Activities to ensure high-quality data collection included in-depth training followed by assessment of competency using a variety of techniques, such as written tests and observations (with feedback) during training sessions. To control the quality of data entry, a field supervisor at each site reviewed all completed CRFs daily for legibility, completeness, and consistency. The supervisor’s signature indicating that all discrepancies were resolved was required for submission to the DCC. Quality control at the DCC to detect missing data, missing forms, out-of-range values, and data inconsistencies is described elsewhere (see Biswas et al, this supplement).

Supervision and oversight were maintained for quality assurance purposes. Supervisors utilized growth charts and predefined criteria to identify (in real time) aberrant

measurements that should be repeated, such values that were lower at 60 days compared to enrollment. Clinical and field supervisors performed random reinterviews to ensure the validity of the data collected by the GEMS clinicians and field workers. The CVD epidemiology team visited each site at least twice per year to observe study activities, review the regulatory files, randomly inspect consent forms and CRFs, and retrain as necessary. They provided a written feedback to the site and the CVD investigators. They maintained at least weekly contact with the teams with the use of email, internet calls, and teleconferences. A regulatory affairs specialist at the CVD oversaw the quality, timeliness, and completeness of submissions to each relevant institutional review board, and ensured that each site was compliant with US regulatory requirements. Sites reported all protocol deviations to the CVD team and a corrective action plan was developed jointly.

## **DISCUSSION**

GEMS is the largest and most comprehensive case/control study of acute diarrhea conducted to date, and will, with its complementary components (HUAS-lite and DSS), provide information about the incidence, microbiologic etiology, risk factors, and adverse clinical outcomes of moderate-to-severe diarrheal episodes among infants and young children living in regions of the world where 82% of under-5 deaths occur [2]. GEMS employed standardized data collection instruments and epidemiologic methods across diverse developing country settings that vary with respect to health indicators, access to and quality of affordable healthcare, economic development, and environmental conditions, so the results will be broadly applicable and can be used to augment existing disease burden models and to define the factors likely to influence the outcome of diarrheal disease in the future. The generalizability of the results should be further enhanced by employing case definitions and study methods that were accepted by experts in the field and disclosed in a detailed and transparent way. GEMS will provide a detailed characterization of MSD according to its clinical manifestations and adverse effects on child health. The attributable fraction contributed by each pathogen that is significantly associated with MSD in the case/control study then will be quantified. This process will produce a list of enteropathogens that should be prioritized for public health interventions. In addition, GEMS will provide the serologic, antigenic, and genotypic characteristics of the major etiologic agents, information needed to develop vaccines and other interventions that can be used in decades to come. We have laid the groundwork for building cost-effectiveness models to justify the introduction of selected interventions into countries where GEMS was undertaken by describing the economic burden of diarrheal diseases.

By conducting the GEMS case/control study within a demographically characterized population whose healthcare utilization practices have been characterized, we are also able to derive population-based estimates of MSD incidence and other outcomes of interest. Repeated HUAS-lite interviews of a representative sample of caretakers over a 2-year period provide the proportion of children in the DSS population who seek care at the SHC when they develop MSD (the  $r$  value). Because we are measuring the number of MSD cases from the DSS who visit the SHC each year, we can divide by  $r$  to estimate the number of MSD cases in the DSS population as a measure of incidence. Furthermore, we can use odds ratios from the GEMS case/control study to calculate population attributable fractions for each pathogen found to be associated with MSD by conditional logistic regression adjusted for the presence of other pathogens (see Blackwelder et al, this supplement).

Some limitations of GEMS are noteworthy. First, although designed as an observational study, the resources, infrastructure, training, and frequent visits to the DSS households provided by GEMS may well have altered the natural history of diarrheal diseases in our study populations. These secondary health benefits must be considered in interpreting the outcome and sequelae of diarrheal diseases that are found in GEMS. Another limitation relates to our ability to use the case/control study to derive population-based estimates of MSD incidence and other outcomes of interest. The validity of these estimates will depend on 2 factors. One is the ability of caretakers to accurately report that their child had findings indicative of MSD during the previous 14 days. We attempted to quantify this limitation by conducting a nested study to compare the caretakers' determination of MSD with those of the SHC clinician and found good agreement ( $\kappa$  statistic for interobserver agreement = 0.82 for sunken eyes and 0.64 for wrinkled skin, data not shown). The second factor is the need for a high  $r$  value to achieve precise estimates and to increase the likelihood that the children enrolled in the case/control study indeed represent those with MSD in the DSS population. For this reason, we endeavored to select the SHCs most likely to capture diarrheal diseases at each site; however, despite our efforts, it was generally not possible to achieve high  $r$  values. During analysis, we will take into account the variance of the estimated  $r$ 's in assessing the precision of our disease incidence estimates. Finally, some might argue that our case definition of MSD could bias our pathogen-specific disease estimates by overestimating the importance of agents that cause dysentery. Our decision to include dysentery is supported by published observations indicating that children with dysentery have an increased risk for persistent diarrhea, growth faltering, and death [14, 38, 39]. These concerns can be addressed in the analysis because we

know the proportion of all children seeking care at each SHC with MSD that have dysentery, so adjustments can be made as necessary if children with and without dysentery are not equally sampled. We also know the prevalence of dysentery among children in the DSS population based on serial HUAS-lite rounds and can adjust our population-based estimates accordingly.

An important contribution of GEMS is the single follow-up visit to the homes of both cases and controls approximately 60 days after enrollment, which will allow us to elucidate the outcome of children during the vulnerable period that follows acute MSD. The few cross-sectional health center-based studies in developing countries that have contacted children after discharge suggest that most sequelae are missed if surveillance is limited to the hospital stay [40, 41]. A caveat is that we will be unable to prospectively define interim events that might influence the outcomes, including recurring episodes of diarrhea. Important information will be gleaned nonetheless by evaluating the same outcomes in case and control children and determining the relative risk of adverse clinical outcomes during the 50–90 days following an episode of MSD. The lack of interim contact with participants will also impact the quality of data generated by the memory aid regarding the occurrence of diarrhea during the 14 days after enrollment.

In sum, we have described the design and methods of GEMS and our efforts to achieve scientific rigor while maintaining simplicity and standardization. We presented a candid portrait of the considerations that were entertained in developing the study design, the challenges encountered, and solutions developed along with the potential strengths and limitations of the methods. This level of detail is intended to provide the scientific and public health communities with high-quality data that can be used to update and strengthen diarrheal disease burden models and to guide strategic planning and resource allocation for the future.

## Notes

**Acknowledgments.** We are grateful to those who helped us design and prepare for GEMS, especially the families who participated in the training sessions and focus groups and the investigators and staff at the GEMS sites for their enthusiasm and diligence in mastering the study protocol and procedures, including Ciara O'Reilly, Richard Omoro, Inacio Mandomando, Tacilta Nhamossa, Sozinho Acacio, Doh Sanogo, Suman Kumar Das, Sujit Bhattacharya, Byomkesh Manna, Philip C. Hill, Adebayo Akin-sola, Mitchell Adeyemi, and Farheen Quadri. We thank Paul Stolley for advice in case/control study design, and Karen Ball for support in regulatory affairs. We thank Jan Agosti, Niranjan Bose, Thomas Brewer, and Regina Rabinovich from the Bill & Melinda Gates Foundation for technical support and guidance; Irwin Shorr for providing training in anthropometry; and Cristina Carty, Rebecca Horney, and Barbara Yndo for assistance in developing the data collection and management tools.



**Financial support.** This work was supported by the Bill & Melinda Gates Foundation (grant number 38874).

**Supplement sponsorship.** This article was published as part of the supplement entitled “The Global Enteric Multicenter Study (GEMS),” sponsored by the Bill & Melinda Gates Foundation.

**Potential conflicts of interest.** All authors: No reported conflicts.

All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

## References

1. United Nations Inter-agency Group for Child Mortality Estimation. Levels and Trends in Child Mortality Report 2011. Available at: [http://www.unicef.org/media/files/Child\\_Mortality\\_Report\\_2011\\_Final.pdf](http://www.unicef.org/media/files/Child_Mortality_Report_2011_Final.pdf). Accessed 16 March 2012.
2. You D, Wardlaw T, Salama P, Jones G. Levels and trends in under-5 mortality, 1990–2008. *Lancet* **2010**; 375:100–3.
3. Walker CL, Aryee MJ, Boschi-Pinto C, Black RE. Estimating diarrhea mortality among young children in low and middle income countries. *PLoS One* **2012**; 7:e29151.
4. Anker M. The effect of misclassification error on reported cause-specific mortality fractions from verbal autopsy. *Int J Epidemiol* **1997**; 26:1090–6.
5. Lozano R, Lopez AD, Atkinson C, Naghavi M, Flaxman AD, Murray CJ. Performance of physician-certified verbal autopsies: multisite validation study using clinical diagnostic gold standards. *Popul Health Metr* **2011**; 9:32.
6. Morris SS, Black RE, Tomaskovic L. Predicting the distribution of under-five deaths by cause in countries without adequate vital registration systems. *Int J Epidemiol* **2003**; 32:1041–51.
7. Lawn JE, Wilczynska-Ketende K, Cousens SN. Estimating the causes of 4 million neonatal deaths in the year 2000. *Int J Epidemiol* **2006**; 35:706–18.
8. Cooper RS, Osotimehin B, Kaufman JS, Forrester T. Disease burden in sub-Saharan Africa: what should we conclude in the absence of data? *Lancet* **1998**; 351:208–10.
9. World Health Organization Multicentre Growth Reference Study Group. WHO child growth standards. Methods and development: length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age. Available at: [http://www.who.int/childgrowth/standards/technical\\_report/en/index.html](http://www.who.int/childgrowth/standards/technical_report/en/index.html). Accessed 20 May 2012.
10. World Health Organization. Handbook for good clinical research practice (GCP): guidance for Implementation. Available at: [http://whqlibdoc.who.int/publications/2005/924159392X\\_eng.pdf](http://whqlibdoc.who.int/publications/2005/924159392X_eng.pdf). Accessed 21 May 2012.
11. World Health Organization Multicentre Growth Reference Study Group. Reliability of anthropometric measurements in the WHO Multicentre Growth Reference Study. *Acta Paediatr Suppl* **2006**; 450:38–46.
12. Baqui AH, Black RE, Yunus M, Hoque AR, Chowdhury HR, Sack RB. Methodological issues in diarrhoeal diseases epidemiology: definition of diarrhoeal episodes. *Int J Epidemiol* **1991**; 20:1057–63.
13. Schmidt WP, Arnold BF, Boisson S, et al. Epidemiological methods in diarrhoea studies—an update. *Int J Epidemiol* **2011**; 40:1678–92.
14. Black RE, Brown KH, Becker S, Abdul-Alim ARM, Huq I. Longitudinal studies of infectious diseases and physical growth in rural Bangladesh. II. Incidence of diarrhea and association with known pathogens. *Am J Epidemiol* **1982**; 115:315–24.
15. Black RE, Brown KH, Becker S. Effects of diarrhea associated with specific enteropathogens on the growth of children in rural Bangladesh. *Pediatrics* **1984**; 73:799–805.
16. Butler T, Islam M, Azad AK, Islam MR, Speelman P. Causes of death in diarrhoeal diseases after rehydration therapy: an autopsy study of 140 patients in Bangladesh. *Bull World Health Organ* **1987**; 65:317–23.
17. Dutta P, Mitra U, Rasaily R, et al. Assessing the cause of in-patients pediatric diarrheal deaths: an analysis of hospital records. *Indian Pediatr* **1995**; 32:313–21.
18. World Health Organization. Handbook: IMCI integrated management of childhood illness. Available at: [http://www.who.int/maternal\\_child\\_adolescent/documents/9241546441/en/index.html](http://www.who.int/maternal_child_adolescent/documents/9241546441/en/index.html). Accessed 22 March 2012.
19. UNICEF/World Health Organization. Diarrhoea: why children are still dying and what can be done. Available at: <http://www.who.int/topics/diarrhoea/en/>. Accessed 22 February 2012.
20. Ruuska T, Vesikari T. Rotavirus disease in Finnish children: use of numerical scores for clinical severity of diarrhoeal episodes. *Scand J Infect Dis* **1990**; 22:259–67.
21. Flores J, Perez-Schael I, Gonzalez M, et al. Protection against severe rotavirus diarrhoea by rhesus rotavirus vaccine in Venezuelan infants. *Lancet* **1987**; 1:882–4.
22. Clark HF, Borian FE, Bell LM, Modesto K, Gouvea V, Plotkin SA. Protective effect of WC3 vaccine against rotavirus diarrhea in infants during a predominantly serotype 1 rotavirus season. *J Infect Dis* **1988**; 158:570–87.
23. Ferreccio C, Prado V, Ojeda A, et al. Epidemiologic patterns of acute diarrhea and endemic *Shigella* infections in a poor periurban setting in Santiago, Chile. *Am J Epidemiol* **1991**; 134:614–27.
24. Levine MM, Ferreccio C, Prado V, et al. Epidemiologic studies of *Escherichia coli* infections in a low socioeconomic level periurban community in Santiago, Chile. *Am J Epidemiol* **1993**; 138: 849–69.
25. Wright JA, Gundry SW, Conroy R, et al. Defining episodes of diarrhoea: results from a three-country study in sub-Saharan Africa. *J Health Popul Nutr* **2006**; 24:8–16.
26. Filmer D, Pritchett LH. Estimating wealth effects without expenditure data—or tears: an application to educational enrollments in states of India. *Demography* **2001**; 38:115–32.
27. World Health Organization/UNICEF. WHO/UNICEF Joint Monitoring Programme (JMP) for Water Supply and Sanitation Definition and Methods. Available at: <http://www.wssinfo.org/>. Accessed 6 March 2012.
28. Wang XY, Ansaruzzaman M, Vaz R, et al. Field evaluation of a rapid immunochromatographic dipstick test for the diagnosis of cholera in a high-risk population. *BMC Infect Dis* **2006**; 6:17.
29. Cary SG, Blair EB. New transport medium for shipment of clinical specimens. I. Fecal specimens. *J Bacteriol* **1964**; 88:96–8.
30. Wells JG, Morris GK. Evaluation of transport methods for isolating *Shigella* spp. *J Clin Microbiol* **1981**; 13:789–90.
31. Cogill B. Anthropometric indicators measurement guide. Available at: <http://www.fantaproject.org>. Accessed 19 August 2007.
32. World Health Organization. WHO Anthro for personal computer, version 3, 2009. Available at: <http://www.who.int/childgrowth/software/en/>. Accessed 22 February 2010.
33. Thea DM, St Louis ME, Atido U, et al. A prospective study of diarrhea and HIV-1 infection among 429 Zairian infants. *N Engl J Med* **1993**; 329:1696–702.
34. Madhi SA, Cunliffe NA, Steele D, et al. Effect of human rotavirus vaccine on severe diarrhea in African infants. *N Engl J Med* **2010**; 362:289–98.
35. World Health Organization. Verbal autopsy standards: ascertaining and attributing cause of death. Available at: <http://www.who.int/healthinfo/statistics/verbalautopsystandards/en/>. Accessed 28 March 2012.
36. Anker M, Black RE, Coldham C, et al. A standard verbal autopsy method for investigating causes of death in infants and children (WHO document WHO/CDS/CSR/ISR/99.4). Available at: [http://www.who.int/csr/resources/publications/surveillance/WHO\\_CDS\\_CSR\\_ISR\\_99\\_4/en/](http://www.who.int/csr/resources/publications/surveillance/WHO_CDS_CSR_ISR_99_4/en/). Accessed 22 March 2012.



37. World Health Organization, United Nations Children's Fund, US Agency for International Development, and Johns Hopkins Bloomberg School of Public Health. Implementing the new recommendations on the clinical management of diarrhea: guidelines for policy makers and programme managers. Available at: <http://www.zinctaskforce.org/policy-advocacy/>. Accessed 22 March 2012.
38. Ahmed F, Ansaruzzaman M, Haque E, Rao MR, Clemens JD. Epidemiology of postshigellosis persistent diarrhea in young children. *Pediatr Infect Dis J* **2001**; 20:525–30.
39. Bennish ML, Wojtyniak BJ. Mortality due to shigellosis: community and hospital data. *Rev Infect Dis* **1991**; 13(suppl 4):S245–51.
40. Greenwood BM, Greenwood AM, Bradley AK, Tulloch S, Hayes R, Oldfield FS. Deaths in infancy and early childhood in a well-vaccinated, rural, West African population. *Ann Trop Paediatr* **1987**; 7:91–9.
41. Stanton B, Clemens J, Khair T, Shahid NS. Follow-up of children discharged from hospital after treatment for diarrhoea in urban Bangladesh. *Trop Geogr Med* **1986**; 38:113–8.
42. United Nations Children's Fund (UNICEF). The state of the world's children 2005. Available at: <http://www.unicef.org/sowc05/english/fullreport.html>. Accessed 29 March 2012.
43. World Health Organization. WHO world malaria report 2005. Available at: <http://rbm.who.int/wmr2005/>. Accessed 19 August 2009.