Shigella Isolates From the Global Enteric Multicenter Study Inform Vaccine Development

Sofie Livio,¹ Nancy A. Stockbine,²,3 Sandra Panchalingam,¹ Sharon M. Tennial,¹ Eileen M. Barry,¹ Mark E. Marohn,¹ Martin Antonio,² Anowar Hossain,¹ Inacio Mandomando,⁶ John B. Ochien,³ Joseph O. Uomo,³ Shahida Quareshi,⁸ Thandavaryan Ramamurthy,² Boubou Tamboura,²,³ Richard A. Adegbola,⁴ Mohammed Jahangir Hossain,⁴ Debasish Saha,⁴,⁵ Sunil Sen,¹ Abu Syed Golam Faruque,⁵ Pedro L. Alonso,² Robert F. Breiman,⁷,¹¹ Anita K. M. Zaidi,⁶ Dipika Sur,⁵¹² Samba O. Sow,¹⁰ Lynette Y. Berkeley,¹³¹ Clara E. O’Reilly,¹ Eric D. Mintz,² Kousick Biswas,¹⁴ Dani Cohen,¹⁵ Tamer H. Farag,¹,² Dilruba Nasrin,¹ Yukun Wu,¹ William C. Blackwelder,¹ Karen L. Kotloff,¹ James P. Nataro,¹,³ and Myron M. Levine¹

¹Center for Vaccine Development, University of Maryland School of Medicine, Baltimore; ²Escherichia and Shigella Reference Unit, Enteric Diseases Laboratory Branch, and ³Division of Foodborne, Waterborne and Environmental Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia; ⁴Medical Research Council Unit (United Kingdom), Fajara, The Gambia; ⁵International Centre for Diarrheal Disease Research, Mohakhali, Dhaka, Bangladesh; ⁶Centro de Investigação em Saúde de Manhiça, Manhiça, Mozambique and the Centre de Recerca en Salut Internacional de Barcelona, Hospital Clinic/Universitat de Barcelona, Spain; ⁷Kenya Medical Research Institute/Centers for Disease Control and Prevention, Kisumu, Kenya; ⁸Department of Paediatrics and Child Health, The Aga Khan University, Karachi, Pakistan; ⁹National Institute of Cholera and Enteric Diseases, Kolkata, India; ¹⁰Centre pour le Développement des Vaccins du Mali, Bamako, Mali; ¹¹Global Disease Detection Division, Kenya Office of the US Centers for Disease Control and Prevention, Nairobi, Kenya; ¹²Program for Appropriate Technology in Health (PATH), New Delhi, India; ¹³US Food and Drug Administration, Rockville; ¹⁴Department of Veterans Affairs, Cooperative Studies Program Coordinating Center, Perry Point, Maryland; and ¹⁵Department of Epidemiology and Preventive Medicine, School of Public Health, Sackler Faculty of Medicine, Tel Aviv University, Ramat Aviv, Israel

(See the Editorial Commentary by Van de Verg and Venkatesan on pages 942–3.)

Background. Shigella, a major diarrheal disease pathogen worldwide, is the target of vaccine development. The Global Enteric Multicenter Study (GEMS) investigated burden and etiology of moderate-to-severe diarrheal disease in children aged <60 months and matched controls without diarrhea during 3 years at 4 sites in Africa and 3 in Asia. Shigella was 1 of the 4 most common pathogens across sites and age strata. GEMS Shigella serotypes are reviewed to guide vaccine development.

Methods. Subjects’ stool specimens/rectal swabs were transported to site laboratories in transport media and plated onto xylose lysine desoxycholate and MacConkey agar. Suspect Shigella colonies were identified by biochemical tests and agglutination with antisera. Shigella isolates were shipped to the GEMS Reference Laboratory (Baltimore, MD) for confirmation and serotyping of S. flexneri; one-third of isolates were sent to the Centers for Disease Control and Prevention for quality control.

Results. Shigella dysenteriae and S. boydii accounted for 5.0% and 5.4%, respectively, of 1130 Shigella case isolates; S. flexneri comprised 65.9% and S. sonnei 23.7%. Five serotypes/subserotypes comprised 89.4% of S. flexneri, including S. flexneri 2a, S. flexneri 6, S. flexneri 3a, S. flexneri 2b, and S. flexneri 1b.

Conclusions. A broad-spectrum Shigella vaccine must protect against S. sonnei and 15 S. flexneri serotypes/subserotypes. A quadrivalent vaccine with O antigens from S. sonnei, S. flexneri 2a, S. flexneri 3a, and S. flexneri 6 can provide broad direct coverage against these most common serotypes and indirect coverage against all but 1 (rare) remaining subsertype through shared S. flexneri group antigens.

Keywords. serotyping; Shigella; shigellosis; vaccines.
The Global Enteric Multicenter Study (GEMS) of the burden and etiology of moderate-to-severe diarrheal illness (MSD) in children aged <5 years performed over 3 years at 4 sites in sub-Saharan Africa (Basse, The Gambia; Bamako, Mali; Siaya County, Kenya; Manhiça, Mozambique) and 3 in South Asia (Karachi, Pakistan; Kolkata, India; Mirzapur, Bangladesh) established *Shigella* as 1 of 4 top pathogens [1]. The increased diagnostic yield observed when stool specimens are examined using gel-based or quantitative real-time polymerase chain reaction (PCR) suggests that the burden of disease may be greater than estimated using standard cultures [2, 3]. Although pediatric morbidity from shigellosis remains substantial, mortality has diminished, in part, because of the virtual disappearance worldwide of the highly virulent Shiga toxin-producing *S. dysenteriae* 1 serotype and because World Health Organization guidelines recommend antibiotic treatment for clinical dysentery (diarrhea with gross blood). Regrettably, *Shigella* relentlessly acquires resistance to antibiotics that were previously effective in diminishing disease severity and duration and pathogen excretion [2, 4].

Based on clinical severity, disease burden, and emergence of antimicrobial resistance, *Shigella* is a prime target for vaccine development [2, 4–6]. The 4 species (also called groups or subgroups) of *Shigella* encompass 50 serotypes and subspecies that include the following: *S. dysenteriae* (15 serotypes); *S. flexneri* (15 serotypes and subspecies, including *S. flexneri* 1a, 1b, 2a, 2b, 3a, 3b, 4a, 4b, 5a, 5b, 6, X, and Y and 2 new subspecies 7a and 7b, previously referred to as *Shigella* provisional 88–893, Y394, or “*S. flexneri* 1c” [7]; *S. boydii* (19 serotypes); and *S. sonnei* (1 serotype). The distinct serotypes/subspecies are defined by conformational epitopes of their O polysaccharide antigens [6]. Challenge/rechallenge studies in nonhuman primates [8] and volunteers [9–11], epidemiological field studies [12], and seroepidemiological studies [13, 14] indicate that clinical infection with wild type *Shigella* strains bestows approximately 75% subgroup-specific (and usually subspecies-specific) immunity. Live oral vaccines [15–18] and O-poly saccharide-protein conjugate vaccines [19, 20] that have conferred protection in randomized controlled field trials corroborate the importance of immune responses to *Shigella* O antigens. Most *Shigella* vaccines in clinical development are based on eliciting protection against multiple epidemiologically important serotypes. Accordingly, to rationally guide vaccine formulation, it is imperative to have robust data on the distribution of *Shigella* serotypes associated with shigellosis. GEMS serotype data provide such information from the geographic areas where 80% of deaths due to diarrheal disease among young children occur [1, 21].

**MATERIALS AND METHODS**

Conducted over 3 years, GEMS was an age-stratified, matched case-control study of MSD among children aged 0–59 months residing in censused populations and seeking care at medical facilities serving 7 sites in sub-Saharan Africa and South Asia. Rationale for the GEMS and detailed clinical, epidemiologic, and microbiological methods have been published elsewhere [21–24].

**Epidemiological and Clinical Methods**

The University of Maryland, Baltimore Institutional Review Board and ethics committees at each field site approved the protocol. A censused population provided the sampling frame at each study site where sentinel hospitals or health centers serving the population enrolled cases from 3 age strata: infants (0–11 months), toddlers (12–23 months), and young children (24–59 months) [1, 24]. Age-eligible children from the censused population visiting the centers with diarrhea (≥3 loose stools in the previous 24 hours) were examined for eligibility. To be included, the child’s diarrheal episode had to be new (onset after ≥7 diarrhea-free days), acute (duration <7 days), and had to meet at least 1 of the following criteria defining MSD: clinical evidence of moderate-to-severe dehydration (sunken eyes, loss of skin turgor, or initiation of intravenous fluids based on clinical judgment); dysentery; or clinical judgment that the child with diarrheal illness needed to be hospitalized. For each MSD case, 1–3 (occasionally 4) controls without diarrhea, randomly selected from each site’s census database and matched by age, gender, and residential community, were enrolled within 14 days of the matched index case.

Upon enrollment, each case and matched control provided a stool specimen (≥3 grams) that, within 1 hour of passage, was stored cold until delivered to the laboratory. If antibiotics were to be administered to patients before stool was produced, 2 rectal swabs were obtained for bacterial culture pending passage of whole stool for the remaining assays.

**Bacteriologic Methods**

Stool samples/rectal swabs were introduced into Cary–Blair and buffered glycerol saline (BGS) transport media, the latter to enhance yield of *Shigella* [22, 25]; inoculation onto solid media occurred within 18 hours. To isolate *Shigella*, the BGS swab was plated onto MacConkey and xylose lysine desoxycholate agar. After incubation at 37°C, suspicious colonies were subjected to biochemical tests [22]. *Shigella* isolates at the field sites were serotyped with polyvalent group A, B, C, and D antisera (Denka Seiken Co., Ltd., Tokyo, Japan or Reagensia AB, Solna, Sweden) and shipped to the GEMS Reference Laboratory at the Center for Vaccine Development (CVD) for confirmation and identification of individual *S. flexneri* serotypes/subspecies and *S. dysenteriae* 1. One-third of isolates serotyped at CVD were sent to the Centers for Disease Control and Prevention (CDC) for serotype confirmation.
Chromosomal genes encoding *Shigella* enterotoxin 1 (ShET1) [26, 27] were amplified by PCR using the following primers:

- **set1AForward**: 5'- CAG CGT CTT TCA GCG ACA GTG TTT -3'
- **set1AReverse**: 5'- AGC ATG ATA CTC AAC AGC CAG ACC -3'
- **set1BForward**: 5'- ATA CTG GCT CCT GTC ATT CAC GGT -3'
- **set1BReverse**: 5'- GGA AGT GAC AGG GCA TTT GTG GAT -3' [28, 29].

### Statistical Methods

Distributions of species were compared by \(\chi^2\) test with 3 degrees of freedom. Individual species and subserotypes were compared using \(\chi^2\) test with no continuity correction or 2-sided Fisher exact test. \(P \leq .05\) was considered statistically significant.

### RESULTS

We investigated 1130 *Shigella* isolates from cases (1120 from the matched case-control dataset [1] and 10 from MSD cases for whom there was no control) and 219 isolates from controls without diarrhea; 11 other isolates from the sites (9 cases and 2 controls) were not sent to CVD. The distribution of *Shigella* species among case isolates is shown in Table 1. Only 5.0% of case isolates (N = 56) were *S. dysenteriae* (none were *S. dysenteriae* 1) and 5.4% (N = 61) were *S. boydii*. Overall, 89.6% of case isolates were *S. flexneri* (N = 745; 65.9%) or *S. sonnei* (N = 268; 23.7%). Four serotypes/subserotypes, *S. flexneri* 2a, *S. flexneri* 2b, *S. flexneri* 3a, and *S. flexneri* 6, comprised 581 of the 745 *S. flexneri* isolates (78.0%; Table 1); inclusion of *S. flexneri* 1b raises the total to 89.4% of all *S. flexneri*.

Mirzapur, Bangladesh, where shigellosis exhibits a striking seasonal peak [30], contributed the most *Shigella* cases. Thus, it was important to compare Bangladesh serotype data with data from the other 6 GEMS sites (Table 1). The overall species distributions in Bangladesh and the other sites combined were significantly different (\(P = .015\)), but the absolute differences for individual species were modest. Percentages of the 2 most prevalent species, *S. flexneri* and *S. sonnei*, were similar and not significantly different in Bangladesh vs the combined other sites.

The percentages of *S. flexneri* subserotypes were significantly different only for *S. flexneri* 1b, *S. flexneri* 2b, *S. flexneri* 6, and *S. flexneri* X. The proportion of isolates that were *S. flexneri* plus *S. sonnei* by year of the study at all sites, Bangladesh, and the 6 sites other than Bangladesh is shown in Table 2. Remarkably, there was little variation.

Among 219 isolates from control subjects (Table 1), there was less representation of *S. flexneri* (N = 115; 52.5%) compared with cases, where 65.9% of all isolates were *S. flexneri* (\(P < .001\)). *Shigella flexneri* 2a, *S. flexneri* 2b, *S. flexneri* 3a, and *S. flexneri* 6 accounted for 78.0% of *S. flexneri* case isolates vs 71.3% of control *S. flexneri* isolates. *Shigella flexneri* 2a ranked first in frequency followed by *S. flexneri* 6 among cases; among controls, *S. flexneri* 6 ranked first followed by *S. flexneri* 2a.

There was excellent concordance between CVD and CDC results in serotyping quality control. The 46 *S. dysenteriae* isolates sent by CVD were confirmed by CDC as *S. dysenteriae* serotypes 2 (N = 17), 3 (N = 10), 4 (N = 12), 8 (N = 1), 9 (N = 1), and 12 (N = 5). CDC confirmed 76 of 77 putative *S. boydii* that they serotyped as *S. boydii* 1 (N = 11), 2 (N = 16), 3 (N = 1), 4 (N = 16), 5 (N = 1), 7 (N = 1), 8 (N = 4), 10 (N = 9), 11 (N = 1), 12 (N = 4), 14 (N = 3), 15 (N = 2), 18 (N = 3), and 20 (N = 5); the remaining isolate was *Escherichia albertii*, which is known to share O antigens with *S. boydii* [31]. Of 37 putative *S. sonnei* isolates, 36 were confirmed and 1 was found by CDC to be *S. dysenteriae*-like provisional serotype 96–3162 serotype. CDC confirmed CVD’s subserotyping of 146 of the 147 *S. flexneri* sent, the exception being a strain identified by CVD as *S. flexneri* 1b but shown by CDC to be *S. dysenteriae* 3 (strain mix-up). Finally, 22 *Shigella* isolates sent by CVD were deemed untypeable with available reagents. CDC identified 21 of the 22; 20 were *S. dysenteriae*-like provisional serotype 96–3162 and 1 was *S. boydii*-like provisional serotype 2009C-3081.

ShET1, a classic enterotoxin consisting of 1 enzymatically active A subunit linked to 5 binding B subunits, is encoded by *setAB* [26, 32] located within chromosomal *Shigella* pathogenicity island 1 [26, 32], originally identified in *S. flexneri* 2a. Notably, 85 of 86 GEMS *S. flexneri* 2a tested positive for ShET1, as did all 33 *S. flexneri* 2b strains tested. Only 4 of 134 (3.0%) isolates of other *S. flexneri* serotypes were positive, including 2/2 *S. flexneri* 5b, 1/1 *S. flexneri* Y, and 1/22 *S. flexneri* 3a.

### DISCUSSION

Natural immunity to *Shigella* is largely based on immune responses to O antigens. Follow-up of a Chilean pediatric cohort where *S. sonnei*, *S. flexneri* 2a, and *S. flexneri* 6 were the most prevalent serotypes (79% of cases) indicated that an initial episode of shigellosis conferred approximately 75% protection against subsequent shigellosis due to the same serotype but did not significantly cross protect against illness caused by the other predominant serotypes [12]. Seropidemiological studies from Israel corroborate the importance of preexistent O antibody (denoting prior exposure) in lowering the risk of *S. sonnei* and *S. flexneri* 2a disease [13, 14]. Accordingly, developers of vaccines to prevent shigellosis have designed serotype-based vaccines [4, 6], some of which have conferred significant protection [9, 15–20, 33].
### Table 1. Species and Serotype Distribution by Site of 1130 Shigella Isolates From Children Aged <60 Months With Moderate to Severe Diarrhea in the Global Enteric Multicenter Study and of 219 Isolates From Control Children Without Diarrhea

<table>
<thead>
<tr>
<th>Serogroup, Serotype, or Subserotype</th>
<th>Cases</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All 7 GEMS Sites</td>
<td>6 GEMS Sites</td>
</tr>
<tr>
<td>Total isolates</td>
<td>1130</td>
<td>519</td>
</tr>
<tr>
<td>Shigella dysenteriae</td>
<td>56 (5.0%)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>33 (6.4%)&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>S. boydii</td>
<td>61 (5.4%)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>37 (7.1%)&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>S. flexneri&lt;sup&gt;i&lt;/sup&gt;</td>
<td>268 (23.7%)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>119 (22.9%)&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>S. flexneri&lt;sup&gt;d&lt;/sup&gt; serotypes/subserotypes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1a</td>
<td>3 (0.3%)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1 (0.2%)&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>1b</td>
<td>85 (7.5%)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>55 (10.6%)&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>2a</td>
<td>228 (20.2%)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>101 (19.5%)</td>
</tr>
<tr>
<td>2b</td>
<td>123 (10.9%)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>12 (2.3%)&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>3a</td>
<td>106 (9.4%)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>47 (9.0%)&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>3b</td>
<td>1 (0.1%)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0</td>
</tr>
<tr>
<td>4a</td>
<td>33 (2.9%)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>19 (3.7%)&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>4b</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5a</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5b</td>
<td>3 (0.3%)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>124 (11.0%)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>70 (13.5%)&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>7a&lt;sup&gt;a&lt;/sup&gt;</td>
<td>23 (2.0%)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>13 (2.5%)&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>7b&lt;sup&gt;i&lt;/sup&gt;</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>X</td>
<td>11 (1.0%)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>11 (2.1%)&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Y</td>
<td>5 (0.4%)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1 (0.2%)&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

The distribution of species among the Bangladesh isolates was significantly different (P = .015) from the composite of the other 6 GEMS sites. The percentage of the various S. flexneri serotypes and subserotypes isolated in Bangladesh was significantly different from the percentage at the other 6 sites for the greater percentage of S. flexneri 2b in Bangladesh (P < .0001) and the lower percentages of S. flexneri 1b (P = .0003), S. flexneri 6 (P = .015), and S. flexneri X (P = .0002).

Abbreviation: GEMS, Global Enteric Multicenter Study.

<sup>a</sup> Percent of all 1130 case isolates from the composite of all 7 sites.

<sup>b</sup> Percent of the total 519 case isolates from the 6 GEMS sites other than Bangladesh.

<sup>c</sup> Percent of the total isolates from the individual GEMS site.

<sup>i</sup> Bolded S. flexneri serotypes/subserotypes are those proposed for inclusion, along with S. sonnei, in a quadrivalent broad-spectrum Shigella vaccine.

<sup>x</sup> S. flexneri 7 strains were previously referred to as Shigella "provisional 88–893," "provisional Y394," or "S. flexneri 1c."
Table 2. Prevalence of *Shigella sonnei* and *S. flexneri* Serogroups and Proposed Vaccine Component Serotypes of *S. flexneri* Among *Shigella* Isolates From Global Enteric Multicenter Study Cases by Year of the Study

<table>
<thead>
<tr>
<th>Serogroup, Serotype, or Subserotype</th>
<th>All 7 GEMS Sites</th>
<th>6 GEMS Sites Other Than Bangladesh</th>
<th>Bangladesh</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Year 1</td>
<td>Year 2</td>
<td>Year 3</td>
</tr>
<tr>
<td>Total isolates</td>
<td>457</td>
<td>345</td>
<td>328</td>
</tr>
<tr>
<td><em>S. sonnei</em></td>
<td>94 (20.6%)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>76 (22.0%)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>98 (29.9%)&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td><em>S. flexneri</em></td>
<td>317 (69.4%)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>231 (67.0%)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>197 (60.1%)&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td><em>S. flexneri + S. sonnei</em></td>
<td>411 (89.9%)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>307 (89.0%)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>295 (89.9%)&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td><em>S. flexneri</em> 2a, 3a and 6</td>
<td>183</td>
<td>144</td>
<td>131</td>
</tr>
<tr>
<td>As % of all <em>S. flexneri</em></td>
<td>57.7%&lt;sup&gt;d&lt;/sup&gt;</td>
<td>62.3%</td>
<td>66.5%</td>
</tr>
<tr>
<td>As % of all isolates</td>
<td>40.0%&lt;sup&gt;a&lt;/sup&gt;</td>
<td>41.7%</td>
<td>39.9%</td>
</tr>
<tr>
<td><em>S. sonnei + S. flexneri</em> 2a, 3a and 6</td>
<td>277 (60.6%)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>220 (63.8%)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>229 (69.8%)&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td><em>S. sonnei</em> + all <em>S. flexneri</em> serotypes other than <em>S. flexneri</em> 7a</td>
<td>403 (88.2%)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>296 (85.8%)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>291 (88.7%)&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

Abbreviation: GEMS, Global Enteric Multicenter Study.

<sup>a</sup> Percent of all case isolates for particular study year for all 7 GEMS sites.
<sup>b</sup> Percent of all case isolates for particular study year for 6 GEMS sites other than Bangladesh.
<sup>c</sup> Percent of all case isolates for particular study year for Bangladesh.
<sup>d</sup> Percent of all *S. sonnei* isolates for particular study year for all 7 GEMS sites.
<sup>e</sup> Percent of all *S. sonnei* isolates for particular study year for 6 GEMS sites other than Bangladesh.
<sup>f</sup> Percent of all *S. sonnei* isolates for particular study year for Bangladesh.
One obstacle to developing serotype-based Shigella vaccines is choosing the minimal number from among 50 serotypes. The more serotypes included, the more complex and expensive the vaccine, leading some investigators to pursue common protein vaccines [34]. However, the extensive GEMS serotype data from multiple locations over several years constitutes a hallmark resource to guide vaccine development and provides optimism for serotype-based vaccines.

During GEMS, the fearsome S. dysenteriae 1 serotype, which caused protracted pandemics of severe disease from the 1960s to the 1990s in Central America [35], Asia [36], and Africa [37, 38], was not isolated nor were there S. dysenteriae 1 outbreaks reported during 2006–2013. This suggests one may exclude S. dysenteriae 1 from a multivalent vaccine to prevent endemic shigellosis in developing countries. Nevertheless, public health authorities deem it critical that a S. dysenteriae 1 vaccine be available for future resurgence of pandemic Shiga dysentery, a monovalent vaccine could be stockpiled [39]. Since absence of Shiga disease precludes a controlled field trial, prelicensure efficacy must be demonstrated in alternative ways (eg, volunteer challenges with a nonpathogenic strain) [40].

Although S. dysenteriae has 15 and S. boydii has 19 distinct serotypes, they accounted for only 5.0% and 5.4%, respectively, of all Shigella case isolates. Assuming distributions do not change dramatically over longer time periods, excluding these serotypes from a vaccine would have little impact on breadth of coverage. In contrast, S. flexneri serotypes/subserotypes comprised 65.9% of all Shigella case isolates, making it imperative that coverage be provided against the most important S. flexneri serotypes. Interestingly, a mere 5 of the 15 currently recognized S. flexneri serotypes/subserotypes accounted for 89.4% of S. flexneri isolates, including (in rank order) S. flexneri 2a, S. flexneri 6, S. flexneri 2b, S. flexneri 3a, and S. flexneri 1b. Nevertheless, the relative distribution of S. flexneri serotypes may change over time in various geographic locales. Thus, it will be prudent for Shigella vaccines to provide coverage against all 15 S. flexneri serotypes/subserotypes.

CVD investigators devised a strategy to achieve broad-spectrum coverage against all S. flexneri serotypes except uncommon S. flexneri 7a by presenting to the immune system a mix of the following 3 serotypes: S. flexneri 2a, S. flexneri 3a, and S. flexneri 6 [6, 41]. Shigella flexneri 6 is common, and its O antigen is distinct from other S. flexneri. Indeed, genomic evidence indicates S. flexneri 6 might more appropriately be classified as a S. boydii serotype; however, for historical and practical reasons, it retains designation as a S. flexneri serotype. Shigella flexneri serotypes/subserotypes other than S. flexneri 6 have O antigens that share a common backbone structure that consists of tetrasaccharide repeats of 3 rhamnose residues linked to 1 N-acetylglucosamine [6]. Genes encoding the enzymes that synthesize the tetrasaccharide backbone reside in the S. flexneri

| Table 3. Twelve Serotypes and Subserotypes of Shigella flexneri Not in the Quadrivalent Vaccine and O Group Antigens That They Share With the Vaccine Serotypes |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Serotypes in the Quadrivalent Shigella Vaccine (and their O Group Antigen) | Group | Type | O Antigen | O Antigen | O Antigen | O Antigen | O Antigen | O Antigen |
| S. flexneri 1a (type antigen IIIa and group antigen 3,4) | | | | | | | | |
| S. flexneri 2a (type antigen IIa and group antigen 3,4) | | | | | | | | |
| S. flexneri 3a (type antigen IIIa and group antigen 3,4) | | | | | | | | |
| S. flexneri 6 (type antigen V) | | | | | | | | |
chromosomal rfb locus. *Shigella flexneri* Y’s O antigen consists of tetrasaccharide repeats without further modifications. However, lysogenic bacteriophages that encode enzymes able to decorate the tetrasaccharide backbone at specific sites with O-acetyl or D-glucose moieties create new epitopes or ablate others and result in modified saccharide structures that represent the other *S. flexneri* serotypes [6]. The epitopes created by these O-acetyl and D-glucose groups also constitute group antigens shared among different *S. flexneri* serotypes. If the shared group and type-specific antigens induce cross-protective immunity, the number of subserotypes required for a broadly effective *Shigella* vaccine can be minimized. Thus, a multivalent vaccine that includes *S. flexneri* 2a and *S. flexneri* 3a (Table 3), in addition to cross protecting against *S. flexneri* 2b (via type 2 antigen) and *S. flexneri* 3b (via type 3 antigen), would provide shared group antigens that could elicit cross protection against *S. flexneri* 1a, 3b, 4a, 5a, and Y (via group antigen 3.4); against *S. flexneri* 1b, 3b, 4b, and 7b (via group antigen 6); and against *S. flexneri* 2b, 5b, and X (via group 7.8). No cross protection could accrue against *S. flexneri* 6, as its O tetrasaccharide structure (rhamnose-rhamnose-D-galactose-N-acetylgalactosamine) is distinct and lacks the group antigens shared by other *S. flexneri* serotypes.

Noriega et al [41] used the guinea pig Sereny keratoconjunctivitis model to measure cross-reacting serological responses and cross protection when animals immunized mucosally died from challenge with heterologous serotypes/subserotypes including *S. flexneri* 6. Importantly, like *S. sonnei*, they found no cross protection observed in immunized animals challenged with *S. flexneri* 6 [41].

Following *S. flexneri* 2a illness or vaccination with live oral vaccine expressing *S. flexneri* 2a O antigen, the human immune system mounts cross-reacting antibody responses against other *S. flexneri* serotypes that share type or group antigens [42]. If the cross protection observed in animals can be extrapolated to humans, a multivalent vaccine that includes O antigens of *S. sonnei*, *S. flexneri* 2a, *S. flexneri* 3a, and *S. flexneri* 6 would provide direct coverage against approximately 64% of the GEMS *Shigella* strains, and cross protection could provide up to 88% overall coverage. Indeed, only *S. flexneri* 7a (merely 2.0% of GEMS isolates) lacks any of the mentioned shared group antigens; *S. flexneri* 7b expresses group antigen 6. Excluding *S. flexneri* 7a from a multivalent vaccine would have little impact on global breadth of coverage.

Table 2 displays GEMS serotypes in relation to the proposed quardivalent vaccine composition (*S. flexneri* 2a, *S. flexneri* 3a, *S. flexneri* 6 plus *S. sonnei*) to estimate breadth of coverage and possible variation over time from the perspective of all 7 GEMS sites, the site with the most *Shigella* cases (Bangladesh), and the other 6 sites. Only minimal changes in serotypes are seen from year to year; matching of serotypes between circulating strains and vaccine composition would provide 52%–75% direct protection and, with shared group antigens, 82%–93% coverage can be achieved via cross protection.

One other multicenter study used systematic surveillance to detect *Shigella* cases, quantify the burden of shigellosis, and identify serotypes [2]. Von Seidlein et al [2] used a different definition of diarrheal illness as the eligibility criterion for enrollment, obtained strains from older subjects as well as children aged <5 years, maintained surveillance for different time periods (1–3 years, depending on the site), and worked only in Asia (China, Thailand, Vietnam, Indonesia, Bangladesh, and Pakistan) but used microbiological methods similar to GEMS and similarly included sites in Pakistan and Bangladesh. Among the total 2927 *Shigella* isolates reported by Von Seidlein et al [2], 90% were either *S. flexneri* (68%) or *S. sonnei* (22%), similar to GEMS; 51% of their *S. flexneri* isolates were *S. flexneri* 2a, *S. flexneri* 3a, or *S. flexneri* 6. Importantly, like GEMS, they found no *S. dysenteriae* 1 among their 110 *S. dysenteriae* isolates [2]. Via direct or via shared group antigens, the quadrivalent vaccine would cover at least 84.7% of the 2819 fully serotyped *Shigella* strains isolated by Von Seidlein et al [2].

Preclinical and clinical evidence indicates that ShET1 contributes to the watery diarrhea observed early in *S. flexneri* 2a clinical illness and to diarrheal adverse reactions associated with certain live oral vaccines [26, 27, 43]; deleting *set* and *sen* diminishes vaccine reactogenicity [43, 44]. We confirmed that ShET1 genes are common in *S. flexneri* 2a and 2b isolates (117/119, 98.3%) but rare among other *S. flexneri* subserotypes (4/137, 2.9%) or other species (2/132, 1.5%).

Serotyping of the GEMS *Shigella* isolates offers optimism that a quadrivalent vaccine containing *S. sonnei* and 3 serotype/subserotype of *S. flexneri* (*S. flexneri* 2a, *S. flexneri* 3a, and *S. flexneri* 6) can provide broad coverage against *Shigella*, which causes the majority of endemic pediatric shigellosis in the developing world, and also can provide broad coverage for travelers [45, 46].

### Notes

**Acknowledgments.** The authors acknowledge the exceptional serotyping prowess of Ms Evangeline Sowers at the Centers for Disease Control and Prevention (CDC).

**Disclaimer.** The findings and conclusions in this report are those of the author(s) and do not necessarily represent the views of the CDC.

**Financial support.** This work was supported by grant 38774 from the Bill & Melinda Gates Foundation to M. M. L.

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

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


40. Samandari T, Kotloff KL, Losonsky GA, et al. Production of IFN-gamma and IL-10 to Shigella invasins by mononuclear cells from...


