Original Contribution

Mass Gatherings and Diarrheal Disease Transmission Among Rural Communities in Coastal Ecuador

Philip A. Collender, Christa Morris, Rose Glenn-Finer, Andrés Acevedo, Howard H. Chang, James A. Trostle, Joseph N. S. Eisenberg, and Justin V. Remais*

* Correspondence to Dr. Justin V. Remais, Division of Environmental Health Sciences, School of Public Health, University of California, Berkeley, 2121 Berkeley Way #5302, Berkeley, CA 94720 (e-mail: jvr@berkeley.edu).

Mass gatherings exacerbate infectious disease risks by creating crowded, high-contact conditions and straining the capacity of local infrastructure. While mass gatherings have been extensively studied in the context of epidemic disease transmission, the role of gatherings in incidence of high-burden, endemic infections has not been previously studied. Here, we examine diarrheal incidence among 17 communities in Esmeraldas, Ecuador, in relation to recurrent gatherings characterized using ethnographic data collected during and after the epidemiologic surveillance period (2004–2007). Using distributed-lag generalized estimating equations, adjusted for seasonality, trend, and heavy rainfall events, we found significant increases in diarrhea risk in host villages, peaking 2 weeks after an event's conclusion (incidence rate ratio, 1.21; confidence interval, adjusted for false coverage rate of ≤0.05: 1.02, 1.43). Stratified analysis revealed heightened risks associated with events where crowding and travel were most likely (2-week-lag incidence rate ratio, 1.51; confidence interval, adjusted for false coverage rate of ≤0.05: 1.09, 2.10). Our findings suggest that community-scale mass gatherings might play an important role in endemic diarrheal disease transmission and could be an important focus for interventions to improve community health in low-resource settings.

crowding; diarrheal disease; environmental determinants; mass gatherings; social dynamics; travel

Abbreviations: CI, confidence interval; FCR, false coverage rate.

Mass gatherings—events that assemble enough people to strain community planning and response resources (1, 2)—have been identified as important drivers of infectious disease transmission in multiple contexts (3–10). Infectious-disease risk factors associated with mass gatherings include increased crowding and contact rates, overextension of sanitation and hygiene resources, and risk-taking behaviors (4, 6, 11). Regional and international travel to participate in such events has been cited as a contributor to the spread of Zika virus (12), cholera (10), influenza (13–15), Ebola (16), and vector-borne diseases (17).

A considerable body of research links mass gatherings to outbreaks of enteric and diarrheal diseases. For example, inadequate hygiene during food preparation, often tied to poor sanitation or excessive demands on food preparers, was implicated in outbreaks of *Shigella*, *Salmonella*, *Staphylococcus aureus*, and enterotoxigenic *Escherichia coli* at gatherings in the United States (18, 19), Spain (20), and Japan (21). Drinking or bathing in water shared with a large concentration of gathering attendees contributed to outbreaks of *Shigella* and *Salmonella* in the United States (22) and Canada (23). Overcrowding and/or insufficient sanitation were linked to cholera outbreaks during Grand Maga de Touba in Senegal (10) and the Hajj in Saudi Arabia (11).

While outbreak studies provide anecdotal evidence for mechanisms of disease transmission during gatherings, the impacts of mass gatherings on the incidence of high-burden, endemic infectious diseases, which might exceed the public health burden of sporadic outbreaks (24), remains understudied. Quantitative studies of associations between gatherings and disease incidence would allow estimation and comparison of disease risks associated with gatherings and help identify key factors that influence disease transmission across multiple gathering contexts.

In the present analysis, we combined epidemiologic data collected via 3 years of weekly, high-coverage, active surveillance among 17 communities in rural coastal Ecuador with gathering exposures derived from ethnographic data on the timing of, and participation in, religious and secular gatherings to examine the association of mass gatherings with subsequent occurrence of diarrheal diseases. The study communities, located in Esmeraldas, Ecuador, comprise small, mainly Afro-Ecuadorian villages along the Cayapas, Santiago, and Onzole rivers in the vicinity of Borbón, the regional population center, of about 5,000 persons (Figure 1). Previous research in this area has investigated effects of road construction (25), social cohesion (26, 27), safe water, sanitation, and hygiene practices (28, 29), and extreme rainfall events (29) on diarrheal disease. Gatherings associated with dozens of secular and religious holidays occur throughout the year in the region. Based on ethnographic information regarding the timing of these events, as well as participation of residents of host communities and visitors from other locales, we explored the association between community-scale gatherings and diarrheal diseases in the region.

METHODS

Epidemiologic outcomes

Population and health data were collected between February 18, 2004, and July 4, 2007, from villages selected by block randomization to ensure that communities throughout the study region were represented (25). As in prior work (29), we considered diarrhea incidence data collected weekly from 19 communities. During weekly follow-up, consenting households in each village (95%) were visited by community health workers under the supervision of a nurse employed by the study. Informed consent was obtained verbally from each participating village and household. Self-identified heads of each household were interviewed about illness experienced by household members in the previous week. Cases of diarrhea were identified following the widely used (30) World Health Organization case definition of 3 or more loose stools in a 24-hour period (31). Two of the 19 communities exhibited unstable residency and low social cohesion during the study period, as evidenced by fluctuations in village population and low mean social-network degree, reported in previous work (27) and recorded in ethnographic commentaries (e.g., "X is not so much a community as a series of houses scattered along the riverbank." A.A., Universidad San Francisco de Quito, unpublished data, 2017), which might have precluded gatherings at these sites. Therefore, we excluded these communities from our models, and we performed sensitivity analyses to determine the impact of this decision. Institutional review board committees at the University of California, Berkeley; University of Michigan; Trinity College; and Universidad San Francisco de Quito approved all survey protocols.

Mass-gathering exposures

Extensive ethnographic data were collected as part of a broader research project in which the present work is embedded (25). Recurrent mass gatherings were catalogued post hoc by a full-time anthropologist with over 20 years of experience in the region, drawing from field notes, participant observation, and interviews with community members (Web Appendix 1, available at https://academic.oup.com/aje). To explore general associations between gatherings and diarrhea incidence, we constructed a binary indicator of whether any gatherings that were hosted within a village concluded within a particular week, reasoning that the conclusion of an event might represent the time point of greatest cumulative exposure.

Based on research indicating the role of travel and population mixing in determining disease risks associated with mass gatherings (4, 6, 10, 11, 18, 23, 32), we formed a priori hypotheses that gathering risks might vary in relation to contact rates determined by the degree of participation in communal activities as well as pathogen importation risks and crowding related to the influx of visitors traveling to the host community. Based on available ethnographic information, we defined a categorical variable indicating the degree of community engagement (low, high, or uncategorized) for each gathering event, representing our qualitative assessment of the concentration of participants associated with each event. Events for which communal activity is limited or absent were categorized as low engagement, reflecting: 1) influence of poor social cohesion within a host community; 2) participation of only a minor segment of the population (e.g., devotees of a specific saint); or 3) the tendency for certain events to be limited to familial gatherings, in which case communal activity is unlikely to extend beyond the household. No events classified as low engagement are associated with substantial inbound travel. In contrast, events where participation encompasses not only the host community, but also large numbers of visitors from elsewhere in the region or further abroad were classified as high engagement. Finally, events were left uncategorized if there was evidence of neither substantial inbound travel nor limited participation in communal activities. Gathering characteristics are presented in Table 1; excerpted commentaries in support of engagement classifications are presented in Web Table 1.

Potential confounders (season, long-term trends, and heavy rainfall events)

We examined a time-varying exposure of interest, and thus we controlled for seasonality and long-term trends in diarrheal incidence using smooth functions of week of the year and week since study inception, respectively. Furthermore, heavy rainfall events, which were previously found to be associated with diarrheal disease in our data set (29), were considered potential confounders of gathering exposures. As in prior work (29), heavy rainfall was defined as maximum daily rainfall in a village-week exceeding the 90th percentile of daily rainfall observed or imputed at 4 gauged locations over the course of the study period, with a 2-week lag between exposure and outcome, and stratified by low (1st tertile), medium (2nd tertile), or high (3rd tertile) cumulative rainfall over the preceding 8 weeks. The result is a 4-level categorical exposure (no heavy rainfall; heavy rainfall/low prior rainfall; heavy rainfall/medium prior rainfall; heavy rainfall/high prior rainfall) (29). Prior to their incorporation
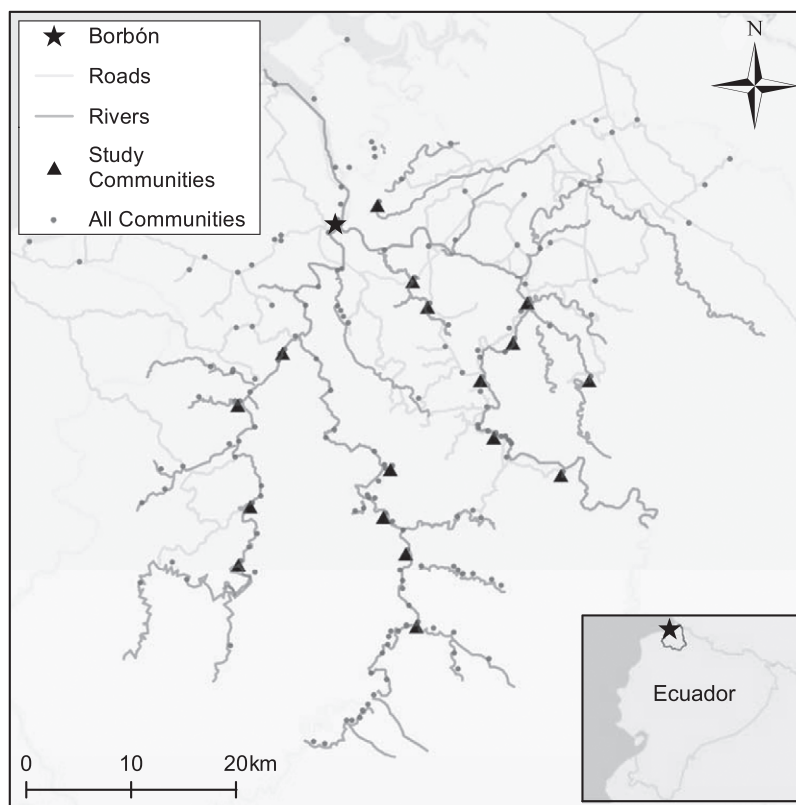
Figure 1.   Map showing villages included for study of recurrent gatherings and diarrhea incidence, northern coastal Ecuador, 2004 to 2007. Also shown are communities not included in the study but potentially participating in gathering events.

into our models, we assessed the potential for confounding between heavy rainfall exposures and gatherings via χ^2 tests of association.

Statistical analyses

Diarrheal outcomes data are correlated at individual, household, and community levels, and exhibit temporal autocorrelation in some communities. We therefore employed Poisson generalized estimating equations for robust inference on correlated outcomes. The lag between gathering exposures and associated diarrhea incidence might vary, due to variable etiologies and incubation periods, discretization error resulting from weekly aggregation of exposures and outcomes, or delayed environmental or secondary transmission. We therefore chose to use a distributed-lag model framework, in which effects are estimated simultaneously at multiple lags to account for associations across the considered timescales.

To efficiently model the temporal structure of association, we used smooth parametric functions to overcome potential collinearity of gathering effects at multiple lags and reduce the number of terms to be estimated (32). Specifically, we constrained gathering effects along the lag dimension to lie on natural cubic splines, with the number of knots selected via optimization of an information criterion developed for

generalized estimating equations (quasi-likelihood under independence model criterion; additional detail in Web Appendix 2 and Pan (33)). To examine any bias potentially introduced by these constraints, we compared constrained gathering effect estimates with those attained from unconstrained models.

We constructed distributed-lag generalized estimating equations for associations of gatherings with disease incidence in host villages at 0- to 3-week lags, controlling for community baseline risks, seasonal and longitudinal trends in incidence, and heavy rainfall exposures:

$$E(Y_{v,t}) = P_{v,t} \exp(\alpha_v + s_{\text{seas}}(t) + s_{\text{long}}(t) + \gamma HRF_{t-2}|CRF_{t-2} + s(G_{t-3:t})$$
$$(1)$$

where $Y_{v,t}$ denotes the count of incident diarrhea cases in village v during week t (sick individuals who were not cases in week $t-1$); $P_{v,t}$ is the population at risk (individuals assessed in week t who were not cases in week $t-1$); α_v is a village effect; $s_{\text{seas}}(t)$ is a seasonal term (week of the year), constructed using a periodic spline function with 4 knots per year; $s_{\text{long}}(t)$ is a longitudinal trend term (week since study inception), constructed using a natural cubic spline with 1 knot per year; $HRF_{t-2}|CRF_{t-2}$ are indicators for occurrence of heavy rainfall (HRF) 2 weeks prior, stratified by low, medium, or high cumulative rainfall (CRF) over the 8 weeks

Table 1. Characteristics of Recurrent Gatherings Hosted Among 17 Communities in Esmeraldas, Ecuador, 2004–2007

Event Name	Typical Duration, days	Typical End Date Within a Year	Participating Villages[a]	Engagement Category[b]
Año Nuevo	2	January 1	All	Low
Día de los Reyes	1	January 6	All	Uncategorized
Carnavales	4	2nd to 4th week of February	All	Low
Semana Santa	7	2nd to 3rd week of April	9, 11, 19	High
	3		All other sites	Uncategorized
Santo Domingo	1	May 14	17	Low
Maria Auxiliadora	1	May 24	15	Uncategorized
San Antonio	2–3	June 14	5, 8, 16, 20–21	Uncategorized
Santisima Trinidad	1	June 15	20	High
Las Cármenes	1	July 16	10, 16	High
			All other sites	Low
Parroquialización/village festivals	1	August 8	4	High
	1	May 14	5	High
	1	October 12	10	Uncategorized
	2	October 5	11	Uncategorized
	1	October 13	17	Low
	1	October 27	19	Low
San Agustin	1	August 28	3	Low
Las Marias	1	September 8	All	Low
Nuestra Señora de las Lajas	1	September 15	17	Uncategorized
Las Mercedes	1	September 25	11, 20	Uncategorized
El Rosario	4	October 8	All	Low
School festival	1	October 26	5	Uncategorized
		October 23	7	
Día de los Difuntos	2	November 3	All	High
La Purisima	1	December 8	11	High
Navidad	2	December 25	All	Uncategorized

[a] Villages are labeled 1–21, following prior work (29).
[b] See Web Appendix 1, Web Table 1, for ethnographic commentaries underlying community engagement.

preceding the heavy rainfall event (29), with coefficient vector γ; and $s(G_{t-3:t})$ is a smooth function of indicators for gathering occurrence 0–3 weeks prior, defined by a natural cubic spline of the lag time, and may be nonzero only for weeks in which a gathering occurred 0–3 weeks prior. Stratified gathering exposures were defined by expanding $s(G_{t-3:t})$ to $s(G_{\text{low } t-3:t}) + s(G_{\text{high } t-3:t}) + s(G_{\text{uncategorized } t-3:t})$, which are natural cubic spline functions, analogous to the unstratified function, of lagged indicators for low-engagement, high-engagement, and uncategorized gatherings, respectively. For consistency and simplicity, all constrained distributed-lag functions featured a single internal knot at 1.5 weeks, selected based on examination of the quasi-likelihood under independence model criterion for unstratified gatherings (Web Appendix 2).

To control for individual-level risk factors and assess their interaction with gatherings, we fitted mixed-effect logistic regressions to individual-level incidence data:

$$E(\pi_{i,v,t}/(1 - \pi_{i,v,t})) = \exp (\alpha_v + \boldsymbol{\beta}\mathbf{X_i} + s(WY)$$
$$+ s(t) + \boldsymbol{\gamma}HRF_{t-2} \,|CRF_{t-2}$$
$$+ s(G_{t-3:t}) + \delta I_{v,t-1} + \epsilon_i)$$

Here, $\pi_{i,v,t}$ is the probability of individual i in village v being an incident case of diarrhea in week t; $\mathbf{X_i}$ is a vector of individual-level traits, namely sex and age group (in years: 0–5, 5–13, 13–45, $\geq$45) with coefficient vector $\boldsymbol{\beta}$; $I_{v,t-1}$ is the incidence rate in village v in the previous week, with coefficient δ; ϵ_i is a random effect for the individual; and other terms are defined as in the village-level model. To explore gathering effect modification by age and sex, we expanded the $s(G_{t-3:t})$ term to $s(G|\text{sex} = \text{male}_{t-3:t}) + s(G|\text{sex} = \text{female}_{t-3:t})$ and $s(G|0 \leq \text{age}<5_{t-3:t})+s(G|5 \leq \text{age} < 13_{t-3:t})+s(G|13 \leq \text{age} <45_{t-3:t})+s(G|45\leq\text{age}_{t-3:t})$ in separate models.

Multiple-testing adjustments and permutation tests

To improve the robustness of our findings to multiple testing across lags and engagement strata, we report confidence intervals at a significance level determined by false-discovery-rate control procedures (34) (Web Appendix 3). We provide a q value for each reported result, which indicates the expected minimum false discovery rate if that result were declared significant. Confidence interval adjustments are made based on the proportion of results declared significant given a target global maximum false discovery rate (Q), such that the proportion of true parameter values not covered by confidence intervals among significant results (false coverage rate) is at or below Q. Following convention, we set $Q = 0.05$.

To further validate our findings, we ran permutation tests to estimate the null distribution of all reported parameter values. For each test, we randomized gathering end dates, maintaining the relative timing of each event year to year, permuted sets of gatherings and corresponding engagement levels across villages, and then estimated pooled and stratified gathering effects. We constructed bidirectional permutation test statistics for each reported parameter, corresponding to the proportion of permutations with parameter values at least as extreme as the true data value. We performed 10,000 permutations at the village level and 1,000 at the individual level.

RESULTS

Study communities, ranging in size from 53 to 872 residents, exhibited generally stable population, age, and sex distributions over the course of the study period (Web Appendix 4, Web Table 2). We identified 21 gatherings recurring in each calendar year (Table 1). Nine annual gatherings occurred in all villages, 4 in multiple villages, and 8 in only 1 village. Most gatherings occurred between May and October and typically lasted 1–2 days, although most months had at least 1 gathering, and some gatherings lasted 4–7 days.

Local participation and inbound travel from nearby and distant communities, the defining characteristics of engagement, varied widely across the gatherings. Some events, such as Año Nuevo or Carnaval, are primarily observed as days of rest, and little communal activity or travel occurs. During more active gatherings, neighboring villagers arrive spontaneously, while members of distant communities might attend upon the host community's invitation. Family members who have dispersed to larger cities in the region or beyond often return for holidays. Semana Santa and Día de los Difuntos are the most important festivals in the region with respect to this pattern of in-migration. Particularly important events, such as Semana Santa in village 11, might also draw tourists from distant cities and countries. Across all village-gathering combinations, 87 annual gatherings were classified as low engagement, 27 as high engagement, and 63 were left uncategorized (Table 1).

We identified risk factors for disease transmission during gatherings, including communal meals, insufficient water and sanitation resources to support visiting populations, changes in sanitation and hygiene practices (e.g., increased reliance on unboiled water to accommodate increased demand during Semana Santa), and behaviors associated with alcohol consumption, particularly during the all-night festivities that take place at the end of many gatherings.

Table 2. Lagged Associations of Diarrhea Incidence With Recurrent Gatherings Hosted Among 17 Communities in Esmeraldas, Ecuador, 2004–2007

Hosting Exposure	IRR	CI[a]	q[b]	$P_{permute}$[c]	No. of Village-Weeks[d]
None[e]	1.00	Referent			1,075
Unadjusted[f]					
Same week	0.92	0.74, 1.14	0.40	0.30	477
1 week prior	1.12	0.97, 1.30	0.11	0.08	469
2 weeks prior	1.19	1.02, 1.39	0.02	0.006	482
3 weeks prior	1.08	0.89, 1.31	0.40	0.35	481
Adjusted					
Same week	0.95	0.76, 1.18	0.59	0.55	477
1 week prior	1.16	0.99, 1.35	0.08	0.03	469
2 weeks prior	1.21	1.02, 1.43	0.02	0.006	482
3 weeks prior	1.09	0.89, 1.34	0.40	0.32	481

Abbreviations: CI, confidence interval; IRR, incidence rate ratio.

[a] Confidence intervals were adjusted to ensure a false discovery rate of ≤0.05 and correspond to a marginal confidence level of 99.25%. See Methods and Web Appendix 3.

[b] q values, which represent the minimum false discovery rate if the threshold for declaring significance includes a parameter, were estimated via the Benjamini and Hochberg procedure (Web Appendix 3).

[c] Proportion of parameter estimates more extreme than the observed value under 10,000 permutations of end dates for each gathering and gathering sets between villages.

[d] Number of village-weeks in exposure category.

[e] The reference group is the set of village-weeks in which no gathering was hosted 0–3 weeks prior.

[f] Unadjusted analysis excluded seasonality, temporal trend, and heavy rainfall events.

Statistical description of exposures and outcomes

A total of 490 gatherings occurred 0–3 weeks prior to any of the 2,711 village-weeks included in our analysis. During the same period, 274 heavy rainfall events were documented, and a total of 2,109 cases of diarrhea were recorded. Heavy rainfall events were associated with gathering exposures at 1-week ($P = 0.02$), 2-week ($P < 0.001$), and 3-week ($P < 0.001$) lags. Community average weekly diarrhea incidence rates ranged from 1.99 to 8.11 per 1,000 (Web Appendix 4, Web Table 3). Seasonal and trend terms fitted in the unstratified analysis indicated modest seasonality, with average diarrhea incidence peaking in April (Web Appendix 5, Web Figure 1) and a trend toward reduced disease incidence from mid-2005 to the end of the study period (Web Appendix 5, Web Figure 2).

Association of mass gatherings with diarrhea incidence at the village scale

Unstratified Poisson distributed-lag generalized estimating equations revealed a gradual peak and decline of disease risks over the 3 weeks subsequent to gatherings, controlling for village effects, seasonality, long-term trend, and heavy rainfall events (Table 2). Evidence of elevated incidence rates was observed at 1- and 2-week lags from the conclusion of gatherings, with the peak incidence rate ratio (compared with weeks in which no event had been hosted) of 1.21 (confidence interval (CI), adjusted for false coverage rate (FCR): 1.02, 1.43) occurring at a 2-week lag. Removing seasonality, trend, and heavy rainfall terms did not substantially alter the effects of gatherings estimated in a crude model. Permutation test statistics confirmed highly significant associations at 2-week lags in the adjusted and crude analyses ($P = 0.006$). Effect estimates were not sensitive to the decision to exclude communities 1 and 12 from analysis (Web Appendix 6, Web Table 4).

Stratification of association by assessed level of community engagement

Distributed-lag Poisson generalized estimating equations for gatherings stratified by community engagement level revealed more pronounced disease risks in association with high-engagement gatherings, which involve high degrees of inbound travel and widespread participation at the host community. Statistically nonsignificant patterns of elevated risk were also present for low-engagement and uncategorized events (Table 3). High-engagement events were associated with diarrheal disease incidence at a 2-week lag, with a peak incidence rate ratio of 1.51 (FCR-adjusted CI: 1.09, 2.10). Permutation test statistics confirmed this association with high significance $P = 0.007$. Effect estimates were not sensitive to the decision to exclude communities 1 and 12 from analysis (Web Appendix 6, Web Table 5).

Association of mass gatherings with diarrhea incidence at the individual scale

Models at the individual scale broadly agreed with estimates at the village level with respect to the temporal scale and magnitude of associations between gatherings and diarrhea, with significant relationships appearing for all gatherings at a 2-week lag (odds ratio = 1.16; FCR-adjusted CI: 1.01, 1.35) and for high-engagement gatherings at lags of 2 weeks (odds ratio = 1.55; FCR-adjusted CI: 1.14, 2.10) and 3 weeks (odds ratio = 1.45; FCR-adjusted CI: 1.00, 2.09) (Web Appendix 7, Web Tables 6 and 7). Under false-discovery-rate control, we found no statistically significant gathering effect modifications by sex or age group (Web Appendix 7, Web Tables 8 and 9).

Comparison of constrained and unconstrained distributed lag estimators

At both village- and individual-level scales, results of constrained and unconstrained distributed-lag models were similar with respect to strength of association and lag structure (Web Appendix 8, Web Tables 10–14, Web Figures 3 and 4). Standard errors of gathering effects at 1- to 2-week lags were notably reduced in the constrained models. This is likely due to mitigation of multicollinearity that we observed among exposures at multiple lags. Goodness of fit was nearly identical for both model specifications, and information criteria consistently favored the more parsimonious constrained models

DISCUSSION

Mass gatherings in our study region were associated with an increase in endemic infectious disease risk in the following weeks. Events characterized as high engagement, involving extensive participation of the host community and visitors from other sites, were associated with higher risks, possibly due to impacts of crowding, communal food sharing, and disease importation by travelers. This association of particularly high risks with gatherings involving travel is consistent with prior research in the study area highlighting travel and hosting out-of-town guests as risk factors for disease (27, 35). While most previous work on mass gatherings and disease highlighted outbreak risks associated with large, often international, mass gatherings, our findings suggest that smaller, community-scale events might have important consequences for endemic infectious disease transmission in low-resource settings.

Diarrhea associated with mass gatherings could be infectious in nature or due to acute noninfectious factors, such as excessive alcohol consumption. Associations observed at 1- to 2-week lags suggest infectious etiologies, because acute noninfectious diarrhea would typically occur within a day of gathering exposures. Common etiologies of infectious diarrhea in the region (and their incubation times) include rotavirus (1–4 days (36)), pathogenic *E. coli* (1–10 days (37)), *Shigella* (1–8 days (38)), and *Giardia* (1–14 days (39)). Other factors that might affect the observed lag structure include timing of symptom onset relative to survey dates and contributions of secondary or environmental transmission.

Our study design was limited in several respects. Gathering exposures are likely subject to some misclassification, having been defined retrospectively from best available knowledge. While the timing of most gatherings in the region is well-established by tradition, the occurrence of specific gatherings

Table 3. Lagged Associations of Diarrhea Incidence With Recurrent Gatherings Hosted Among 17 Communities, Stratified by Community Engagement, Esmeraldas, Ecuador, 2004–2007

Hosting Exposure	IRR	CI[a]	q[b]	$P_{permute}$[c]	No. of Village-Weeks[d]
None[e]	1.00	Referent			1,075
Low-engagement gatherings					
Same week	1.01	0.74, 1.37	0.95	0.93	276
1 week prior	1.13	0.94, 1.34	0.18	0.18	281
2 weeks prior	1.14	0.92, 1.41	0.19	0.15	275
3 weeks prior	1.05	0.75, 1.46	0.73	0.67	276
High-engagement gatherings					
Same week	0.86	0.51, 1.45	0.59	0.40	79
1 week prior	1.29	0.97, 1.71	0.07	0.09	80
2 weeks prior	1.51	1.09, 2.10	0.01	0.007	82
3 weeks prior	1.4	0.85, 2.30	0.18	0.07	81
Uncategorized gatherings					
Same week	0.89	0.56, 1.40	0.59	0.44	122
1 week prior	1.17	0.83, 1.66	0.39	0.22	108
2 weeks prior	1.25	0.88, 1.76	0.19	0.08	125
3 weeks prior	1.07	0.74, 1.54	0.70	0.66	124

Abbreviations: CI, confidence interval; IRR, incidence rate ratio.

[a] Confidence intervals were adjusted to ensure a false discovery rate of ≤0.05 and correspond to a marginal confidence level of 99.25%. See Methods and Web Appendix 3.

[b] q values, which represent the minimum false discovery rate if the threshold for declaring significance includes a parameter, were estimated via the Benjamini and Hochberg procedure (Web Appendix 3).

[c] Proportion of parameter estimates more extreme than the observed value under 10,000 permutations of end dates for each gathering and gathering sets between villages.

[d] Number of village-weeks in exposure category.

[e] The reference group is the set of village-weeks in which no gathering was hosted 0–3 weeks prior.

in a given year—and their duration—can depend on available resources and local interest. Furthermore, our treatment of gatherings as community-wide exposures could result in exposure misclassification if individual participation varied. Our stratified analysis across engagement categories was based on qualitative ethnographic descriptions of participation and travel associated with gatherings, given that no quantitative measures were available. In the case of several gatherings, even qualitative information was unavailable. Furthermore, our outcome measure—occurrence of diarrhea reported by the head of household—could be subject to errors in recall or other reporting biases.

Misclassification of our exposures and outcomes is unlikely to be systematic, and thus the expected impact is a reduction in power and bias of effect estimates towards the null. Exposures were classified independently and at different times from collection and analysis of outcomes data, and ethnographic and epidemiologic data were collected by different staff. Ethnographers contributing to retrospective exposure classifications were not in possession of outcomes data or summaries of temporal variation in outcomes. Regarding outcome misclassification, while we have no means of estimating the extent to which diarrhea reporting biases and gatherings might correlate, it seems unlikely that any such effects would linger over multiple weeks of lag.

This study represents a first effort to examine the contribution of mass gatherings to endemic disease risks. Future studies should be undertaken with formal assessment of gathering characteristics made at the time of the events, which would improve our understanding of the mechanisms that underlie these risks. This information could also aid policy makers and health practitioners by identifying factors associated with increased likelihood of disease transmission. Gatherings often involve extensive planning and preparation by local governments, schools, community organizations, and others, which present opportunities for public health education and risk mitigation. Potential interventions could include enhancing sanitation infrastructure to support increased population concentration, facilitating safe preparation of food, and providing safe water. Gatherings in Canton Eloy Alfaro, in particular, might represent both drivers of endemic disease and excellent opportunities for public health intervention.

REFERENCES

1. World Health Organization. Public health for mass gatherings: key considerations. 2016. http://www.who.int/ihr/publications/WHO_HSE_GCR_2015.5/. Accessed July 3, 2017.
2. World Health Organization. Communicable disease alert and response for mass gatherings. 2008. http://www.who.int/csr/Mass_gatherings2.pdf. Accessed July 3, 2017.
3. le Polain de Waroux O, Saliba V, Cottrell S, et al. Summer music and arts festivals as hot spots for measles transmission: experience from England and Wales, June to October 2016. *Euro Surveill*. 2016;21(44):30390.
4. Elachola H, Gozzer E, Zhuo J, et al. Mass gatherings: a one-stop opportunity to complement global disease surveillance. *J Health Spec*. 2016;4(3):178–185.
5. Gautret P. Religious mass gatherings: connecting people and infectious agents. *Clin Microbiol Infect*. 2015;21(2):107–108.
6. Abubakar I, Gautret P, Brunette GW, et al. Global perspectives for prevention of infectious diseases associated with mass gatherings. *Lancet Infect Dis* 2012;12(1):66–74.
7. Saeed KM, Mofleh J, Rasooly MH, et al. Occurrence of acute respiratory infection, diarrhea and jaundice among Afghan pilgrims, 2010. *J Epidemiol Glob Health*. 2012;2(4):215–220.
8. Gautret P, Steffen R. Communicable diseases as health risks at mass gatherings other than Hajj: what is the evidence? *Int J Infect Dis*. 2016;47:46–52.
9. Botelho-Nevers E, Gautret P. Outbreaks associated to large open air festivals, including music festivals, 1980 to 2012. *Euro Surveill*. 2013;18(11):20426.
10. Finger F, Genolet T, Mari L, et al. Mobile phone data highlights the role of mass gatherings in the spreading of cholera outbreaks. *Proc Natl Acad Sci U S A*. 2016;113(23):6421–6426.
11. Al-Tawfiq JA, Memish ZA. Mass gatherings and infectious diseases: prevention, detection, and control. *Infect Dis Clin North Am*. 2012;26(3):725–737.
12. Petersen E, Wilson ME, Touch S, et al. Rapid spread of Zika virus in the Americas—implications for public health preparedness for mass gatherings at the 2016 Brazil Olympic Games. *Int J Infect Dis*. 2016;44:11–15.
13. Lozano R, Naghavi M, Foreman K, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*. 2012;380(9859):2095–2128.
14. Botelho-Nevers E, Gautret P, Benarous L, et al. Travel-related influenza A/H1N1 infection at a rock festival in Hungary: one virus may hide another one. *J Travel Med* 2010;17(3):197–198.
15. Loncarevic G, Payne L, Kon P, et al. Public health preparedness for two mass gathering events in the context of pandemic influenza (H1N1) 2009—Serbia, July 2009. *Euro Surveill*. 2009;14(31).
16. Blumberg L, Regmi J, Endricks T, et al. Hosting of mass gathering sporting events during the 2013–2016 Ebola virus outbreak in West Africa: experience from three African countries. *Int J Infect Dis*. 2016;47:38–41.
17. Editorial Team. Health risks during the Cricket World Cup in the Caribbean: surveillance and assessment in the French departements. *Euro Surveill*. 2007;12(3):E070308.2.
18. Lee LA, Ostroff SM, McGee HB, et al. An outbreak of shigellosis at an outdoor music festival. *Am J Epidemiol*. 1991;133(6):608–615.
19. Centers for Disease Control and Prevention. Outbreak of salmonellosis associated with consumption of pulled pork at a church festival—Hamilton County, Ohio, 2010. *MMWR Morb Mortal Wkly Rep*. 2014;62(51–52):1045–1047.
20. Camps N, Domínguez A, Company M, et al. A foodborne outbreak of *Salmonella* infection due to overproduction of egg-containing foods for a festival. *Epidemiol Infect* 2005;133(5):817–822.
21. Kitamoto M, Kito K, Niimi Y, et al. Food poisoning by *Staphylococcus aureus* at a university festival. *Jpn J Infect Dis*. 2009;62(3):242–243.
22. Wharton M, Spiegel RA, Horan JM, et al. A large outbreak of antibiotic-resistant shigellosis at a mass gathering. *J Infect Dis*. 1990;162(6):1324–1328.
23. Macey J, Lior L, Johnston A, et al. Outbreak of diarrheal illness in attendees at a Ukrainian dance festival, Dauphin, Manitoba—May 2001. *Can Commun Dis Rep*. 2002;28(17):141–145.
24. Medley GF, Vassall A. When an emerging disease becomes endemic. *Science*. 2017;357(6347):156–158.
25. Eisenberg JN, Cevallos W, Ponce K, et al. Environmental change and infectious disease: how new roads affect the transmission of diarrheal pathogens in rural Ecuador. *Proc Natl Acad Sci U S A*. 2006;103(51):19460–19465.
26. Bates SJ, Trostle J, Cevallos WT, et al. Relating diarrheal disease to social networks and the geographic configuration of communities in rural Ecuador. *Am J Epidemiol*. 2007;166(9):1088–1095.
27. Zelner JL, Trostle J, Goldstick JE, et al. Social connectedness and disease transmission: social organization, cohesion, village context, and infection risk in rural Ecuador. *Am J Public Health*. 2012;102(12):2233–2239.
28. Goldstick JE, Trostle J, Eisenberg JN. Ask when—not just whether—it's a risk: how regional context influences local causes of diarrheal disease. *Am J Epidemiol*. 2014;179(10):1247–1254.

29. Carlton EJ, Eisenberg JN, Goldstick J, et al. Heavy rainfall events and diarrhea incidence: the role of social and environmental factors. *Am J Epidemiol*. 2014;179(3):344–352.

30. Gidudu J, Sack DA, Pina M, et al. Diarrhea: case definition and guidelines for collection, analysis, and presentation of immunization safety data. *Vaccine*. 2011;29(5):1053–1071.

31. World Health Organization, UNICEF, Johns Hopkins Bloomberg School of Public Health, et al. Implementing the new recommendations on the clinical management of diarrhoea: Guidelines for policy makers and programme managers. 2006. http://www.who.int/maternal_child_adolescent/documents/9241594217/en/. Accessed November 4, 2018.

32. Memish ZA, Almasri M, Turkestani A, et al. Etiology of severe community-acquired pneumonia during the 2013 Hajj—part of the MERS-CoV surveillance program. *Int J Infect Dis*. 2014; 25:186–190.

33. Pan W. Akaike's information criterion in generalized estimating equations. *Biometrics*. 2001;57(1):120–125.

34. Benjamini Y, Yekutieli D, Edwards D, et al. False discovery rate: adjusted multiple confidence intervals for selected parameters. *J Am Stat Assoc*. 2005;100(469):71–81.

35. Kraay ANM, Trostle J, Brouwer AF, et al. Determinants of short-term movement in a developing region and implications for disease transmission. *Epidemiology*. 2018;29(1):117–125.

36. Lee RM, Lessler J, Lee RA, et al. Incubation periods of viral gastroenteritis: a systematic review. *BMC Infect Dis*. 2013;13: 446.

37. Nataro JP, Kaper JB. Diarrheagenic *Escherichia coli*. *Clin Microbiol Rev*. 1998;11(1):142–201.

38. Niyogi SK. Shigellosis. *J Microbiol*. 2005;43(2):133–143.

39. Nash TE, Herrington DA, Losonsky GA, et al. Experimental human infections with *Giardia lamblia*. *J Infect Dis*. 1987; 156(6):974–984.