Estimating time-varying transmission and oral cholera vaccine effectiveness in Haiti and Cameroon, 2021-2023

Erin N. Hulland¹,², Marie-Laure Charpignon²,³, Ghinwa Y. El Hayek², Lihong Zhao⁴, Angel N. Desai²,⁵, Maimuna S. Majumder¹,²

¹ Computational Health Informatics Program, Boston Children’s Hospital & Harvard Medical School, Boston, MA, United States
² Comp Epi Dispersed Volunteer Research Network, Boston, MA, United States
³ Institute for Data, Systems, and Society, Massachusetts Institute of Technology, Cambridge, MA, United States
⁴ Department of Mathematics, Virginia Tech, Blacksburg, VA, United States
⁵ Department of Internal Medicine, Division of Infectious Diseases, University of California, Davis Health Medical Center, Sacramento, CA, United States

Disclaimer: This work is not yet finalized and will be continuously updated. We welcome feedback and suggestions on the analyses and interpretation and any meaningful additions.

June 12, 2024

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Abstract

In 2023, cholera affected approximately 1 million people and caused more than 5000 deaths globally, predominantly in low-income and conflict settings. In recent years, the number of new cholera outbreaks has grown rapidly. Further, ongoing cholera outbreaks have been exacerbated by conflict, climate change, and poor infrastructure, resulting in prolonged crises. As a result, the demand for treatment and intervention is quickly outpacing existing resource availability. Prior to improved water and sanitation systems, cholera, a disease primarily transmitted via contaminated water sources, also routinely ravaged high-income countries. Crumbling infrastructure and climate change are now putting new locations at risk — even in high-income countries. Thus, understanding the transmission and prevention of cholera is critical.

Combating cholera requires multiple interventions, the most common being behavioral education and water treatment. Two-dose oral cholera vaccination (OCV) is often used as a complement to these interventions. Due to limited supply, countries have recently switched to single-dose vaccines (OCV1). One challenge lies in understanding where to allocate OCV1 in a timely manner, especially in settings lacking well-resourced public health surveillance systems. As cholera occurs primarily in low-income and conflict settings, timely, accurate, and openly accessible outbreak data are typically inaccessible for disease modeling and subsequent decision-making.

In this study, we demonstrated the value of open-access data to rapidly estimate disease transmission and vaccine effectiveness. Specifically, we obtained non-machine readable (NMR) epidemic curves for recent cholera outbreaks in two countries, Haiti and Cameroon, from figures published in situation and disease outbreak news reports. We used computational digitization techniques to derive weekly counts of cholera cases. Given these data, we leveraged EpiEstim—an open-source modeling platform—to derive rapid estimates of time-varying disease transmission via the effective reproduction number ($R_t$). We used VaxEstim, a recent extension of this platform allowing the estimation of vaccine effectiveness via the relation between the basic reproductive number ($R_0$), $R_t$, and vaccine coverage, to compare OCV1 effectiveness in the two considered countries. Here, with Haiti and Cameroon as case studies, we demonstrated the first implementation of VaxEstim in low-resource settings.

In the initial phase of the outbreak, weekly rolling average estimates of $R_t$ were elevated in both countries: 2.60 in Haiti [95% credible interval: 2.42-2.79] and 1.90 in Cameroon [1.14-2.95]. Similarly, in both countries, this period of high transmission preceded a longer period during which $R_t$ oscillated around the critical threshold of 1. VaxEstim results suggest that Haiti had higher OCV1 effectiveness than Cameroon (75.32% effective [54.00-86.39%] vs. 56.21% [18.87-84.94%]).

This work underscores the value of combining NMR sources of outbreak case data with computational techniques and the utility of VaxEstim for rapid, inexpensive estimation of vaccine effectiveness in data-poor outbreak settings.

Keywords: Cholera, Haiti, Cameroon, reproduction number, $R_0$, Effective $R$, statistical modeling, computational epidemiology, outbreak and response, vaccination, infectious disease, low-income countries, OCV, data scarcity, vaccine effectiveness

Introduction

Cholera, a disease characterized by severe watery diarrhea caused by toxigenic serogroups O1 or O139 of the bacteria Vibrio cholerae, has ravaged human society through seven deadly pandemics since its first documented international spread in 1817. The seventh pandemic, which began in 1961, is still ongoing. Presently, cholera affects about one quarter of the world’s countries and hundreds of millions of people. Although the disease predominantly impacts low-income countries and conflict settings, most developing countries remain at risk for an outbreak. More alarmingly, since 2022, an increasing number of countries have been reporting new cholera outbreaks. Due to limited resources, worsening political instability and displacement, and poor infrastructure, most outbreaks are continuing. Such crises have been widespread throughout the globe, from Haiti in the Caribbean, to Democratic Republic of Congo, Malawi, and Cameroon in Africa, to Lebanon and Afghanistan in the Middle East, to India and China in Asia.

Cholera spreads through contaminated water by either naturally occurring bacteria or through fecal
transmission in water by an infected individual. Cholera is of particular concern in locations with limited access to improved water and sanitation systems, as outbreaks persist and propagate where water access, sanitation, and hygiene systems are insufficient. Although improving these systems is the most effective way to prevent onward cholera transmission in the long run, infrastructure development is costly and time-intensive, and therefore unsuitable as a primary intervention during an ongoing outbreak. In contrast, treatment of mild to moderate cases is often well-managed with oral rehydration therapy (ORT), a low-cost and rapid intervention. However, both access to and adoption of ORT remain limited in many low-income countries, thus allowing cholera to remain a significant contributor to global morbidity and mortality, with an estimated 1.2 to 4 million cases and an average of 95,000 deaths annually. While high-income countries have been spared in recent decades, they are not entirely immune from the threat cholera represents. In the United States, for example, millions of people lack clean water or complete indoor plumbing. Additionally, most of the water and sanitation infrastructure in the country is outdated and requires immediate replacement or repair to maintain satisfactory performance. Moreover, climate change, which influences water quality, quantity, and access via drought, flooding, and other catastrophic weather events, poses a threat for many existing water and sanitation systems globally. Thus, while the burden of cholera is currently concentrated in low-income and conflict settings, high-income countries are far from immune to a resurgence of cholera as well as other zoonotic and waterborne threats. Hence it is important for epidemiologists and mathematical modelers to consider the additional risks posed by climate change for disease resurgence, both in the US and abroad.

In outbreak contexts, vaccination with killed whole-cell oral cholera vaccines (OCV) is considered a safe and effective intervention when cases are surging. Importantly, vaccination with OCVs generally requires administration of two doses. Three main OCVs are recommended by the World Health Organization for cholera intervention — Shanchol, Euvichol, or Dukoral — but only Shanchol and Euvichol are available for mass vaccination campaigns through the Gavi-supported global stockpile. Both Shanchol and Euvichol have generally shown high effectiveness: field studies in Haiti, Bangladesh, Guinea, Zambia, and India demonstrated that two-year effectiveness of two-dose vaccines ranged from 37% to 97.5%, averaging around 65%. This variability reflects in part the role of each country’s health system and the differing challenges associated with vaccine distribution in urban vs rural areas. However, the global stockpile of OCV — established in 2013 — has recently faced shortages, with demand consistently outpacing supply despite annual increases in procurement. As of October 2022, the International Coordinating Group for vaccine delivery has thus restricted outbreak vaccination campaigns to only single-dose vaccines (OCV1). Studies recently conducted in South Sudan, Bangladesh, and Democratic Republic of the Congo suggest that the short-term effectiveness of OCV1 ranges between 40% and 80% for all cases, but that their effectiveness is lower over longer time periods. In most cases, such short-term estimates largely align with effectiveness observed from two-dose studies, adding evidentiary support to the use of OCV1 in the face of shortages.

Using Haiti and Cameroon as compelling examples of the countries experiencing new or worsening outbreaks in 2022, we sought to investigate cholera transmission dynamics, evaluate vaccine effectiveness, and understand the role of sociopolitical and health systems as well as climate. In both countries, OCV1 campaigns were conducted to aid in reducing cholera transmission, serving as an interesting natural case study to assess one- versus two-dose vaccine effectiveness in each location and further understand differences between them. While Haiti and Cameroon share factors that make them highly vulnerable to cholera, their local climate, health systems, and existing water and sanitation infrastructure differ. Such heterogeneity allows us to formulate hypotheses regarding the impact of these factors on the effectiveness of subsequent vaccination campaigns. By helping us understand how the patterns of cholera transmission and vaccine effectiveness vary by context, this case study can inform future vaccination efforts worldwide.

The current outbreak in Cameroon began in 2021, while Haiti’s first case was not reported until 2022. Disease transmission persists in both countries, despite numerous ebbs and flows in reporting. OCV1 campaigns occurred in February and April 2022 in Cameroon and in December 2022 in Haiti. While Cameroon has had recurrent outbreaks since cholera was first reported in the country in 1971,
Haiti had never had a single confirmed cholera case until 2010, in the aftermath of the deadly earthquake in Port-au-Prince. While the basic and effective reproduction numbers ($R_0$ and $R_e$) – two related but distinct measures of transmissibility – have been used as the key metrics to understand cholera outbreaks, most research has employed mechanistic compartmental models requiring data availability, mathematical modeling expertise, and assumptions about the underlying disease process and corresponding parameterization. Because these are poorly suited to outbreak conditions, we explored an alternative approach, capitalizing on phenomenological models which do not rely on assumptions about the disease process and generally require less parameterization.

A distinct feature of our approach was to start with the digitization of outbreak case data from outbreak and situation reports. Unlike in high-income settings, timely, open-access, machine-readable surveillance data of infectious disease case counts are frequently inaccessible for modeling; thus, the use of alternative data sources is paramount. Upon digitizing published epidemic curves, we obtained the trajectory of cholera case counts in each country. Next, we used an openly available R package (EpiEstim) to estimate the time-varying effective reproduction number $R_e$. This step requires passing four epidemiologic parameters (e.g., the mean serial interval), all of which can be gleaned from published literature. Last, we extended our contribution to also estimate vaccine effectiveness in each country using the canonical relation between $R_0$, $R_e$, vaccine coverage and effectiveness. This step was facilitated by a recent extension of the EpiEstim package called VaxEstim. So far, this extended framework has only been applied within the United States and to readily accessible, high-quality data.

**Methods**

**Basic and effective reproduction numbers**

In epidemiology, the basic reproduction number ($R_0$) measures the average number of secondary infections caused by a primary infection over the course of the infectious period, assuming that the entire population is susceptible. When the population is not fully susceptible, due to either vaccination or natural immunity, the effective reproduction number ($R_e$) is used instead of $R_0$. Using available epidemic curves, $R_e$ can be readily estimated at each time point $t$ via the Bayesian framework proposed by Cori et al. and by leveraging the companion R package, EpiEstim. When data about vaccine coverage are available, estimates of vaccine effectiveness can be derived from the above estimates of $R_e$, along with $R_0$ values obtained from the literature. In this study, we built upon previous uses of EpiEstim and applied VaxEstim to model cholera outbreaks in two distinct countries, Cameroon and Haiti.

**Data sourcing and cleaning**

Without readily-accessible, machine-readable cholera surveillance data, we capitalized on published epidemic curves from situation reports and disease outbreak news reports. Although estimates of OCV1 coverage were reported infrequently in situation and disease outbreak news reports, we extracted every mention for subsequent use in vaccine effectiveness models. For Haiti, we leveraged situation reports released by PAHO and Ministère de la Santé Publique et de la Population (MSPP). For Cameroon, we used disease outbreak news reports from the WHO. In Cameroon, the most recent outbreak of cholera was declared on October 29, 2021. By May 2023 – the last reported disease outbreak news report – over 6,000 confirmed or suspected cases and 370 deaths had been registered. In Haiti, the most recent outbreak was confirmed on October 1, 2022, following a three-year hiatus without any cholera cases. By August 2023 – the last report published by the Pan-American Health Organization (PAHO) at the time of analysis – over 58,000 confirmed or suspected cases and 800 deaths had been attributed to this outbreak.

Information about the serial interval (SI) — the period of time between symptom onset in two consecutive cases — was obtained from previous modeling studies on cholera, as well as from prospective and retrospective epidemic assessments. Based on previous studies, we assumed that the SI followed a Gamma distribution with a mean of 4 days and a standard deviation of 3. Data on the historic $R_0$ for cholera were gleaned from national and regional reports of prior outbreaks and other modeling studies. Given the noted importance of using regional data to estimate $R_0$, we used different priors for Cameroon and for Haiti. In our main analysis, we considered a mean $R_0$ of 2.27 for Haiti, estimated using previous country-specific publications. We considered a mean $R_0$ of 1.95 for Cameroon, estimated using one country-specific study and two regional...
analyses. In Appendix Figure 4, we separately investigated the local sensitivity of our vaccine effectiveness results to the choice of mean value for $R_0$, exploring the full range of values previously observed and reported in literature reviews. Specifically, reported estimates of $R_0$ were similar in the two considered countries, ranging from 1.06 to 3.72 in Haiti and from 1.10 to 3.50 in Cameroon. While eligible priors for $R_0$ were available at the country level in the literature, we did not find country-specific priors for $R_t$. Thus, in our main analysis, we used a Gamma distribution with a mean of 2 and a standard deviation of 0.7 for the prior distribution on $R_0$, consistent with previous research. This prior is used in the EpiEstim Bayesian analysis to estimate $R_t$ within a well-informed, limited boundary. In Appendix Figure 3, we explored the local sensitivity of our effective reproductive number estimates to the choice of priors for $R_0$. All sensitivity analyses are available in the Appendix.

In response to rising case and death counts, OCV1 campaigns were conducted in both countries. In Cameroon, an initial campaign was launched in February 2022, followed by a second campaign in April 2022; in Haiti, the vaccination campaign started in mid-December 2022 and continued through January 2023. In Cameroon, the first campaign targeted the South-West region and reached 85.5% administrative coverage, while the second campaign targeted the Littoral, South, and South-West regions and reached 89.00% administrative coverage. In Haiti, the December campaign reached 76.00% administrative coverage on average, with marked geographical differences (i.e., 69.90% in the Ouest and 100.00% in the Centre region). In sensitivity analyses, we explored the impact of lower and higher values of the estimated OCV1 coverage on our estimates of vaccine effectiveness.

**Digitizing epidemic curves**

Epidemic curves were downloaded from published situation and disease outbreak news reports and digitized using the metaDigitise package in R. Epidemic data for Cameroon were reported weekly. Since the reporting time period (here, 7 days) cannot be longer than the SI (here 5 days), we needed to adjust the weekly incidence into daily incidence; we did so using a uniform distribution over the entire seven-day week. For Haiti, case counts were already reported daily and did not require any additional transformation.

**VaxEstim**

We used the EpiEstim framework and R package to estimate weekly values of the effective reproduction number $R_t$, independently for Haiti and for Cameroon. The procedure outputs a country-specific posterior distribution of $R_t$ at each time point. Using the mean, 0.025 quantile, and 0.975 quantile from these country-specific Gamma distributions, as well as the corresponding reported administrative OCV1 coverage (VC) and estimated $R_0$ from the literature, we then employed the following canonical equation to estimate OCV1 effectiveness ($VE_t$) in Haiti and Cameroon over time:

$$VE_t = \frac{1-(R_t/R_0)}{VC_t}$$

Notably, we removed estimates outside the lower and upper VE limits of 0% and 100%, respectively. The number and percent outside these bounds is available in Appendix Table 2.

**Environmental covariate availability**

In contrast to infectious disease surveillance data, environmental covariates are widely available at high spatial and temporal resolution, consistently across countries. In currently ongoing research, we aim to use environmental covariates and time series ARIMA modeling to identify correlates of higher cholera transmission measured by the effective reproduction number. Because cholera has a strong environmental transmission cycle, we now seek to understand the influence of rainfall, average air temperature, sea surface temperature (SST), and sea surface salinity (SSS) on the resurgence and transmission of cholera. Indeed, each have been linked to cholera incidence in prior work and operate indirectly through flooding, pH changes, bacterial concentrations, or algal blooms. Each of these four covariates—rainfall, air temperature, SST, and SSS—is available daily, the same resolution as our EpiEstim $R_t$ estimates, for our two locations of interest, allowing us to match these two data sources in our time series ARIMA models.

**Data availability and role of the funder**

All analyses were done using R version 4.3.1. This research was supported in part by the National Institute of General Medical Sciences, MIT-Harvard Broad Institute Eric & Wendy Schmidt Center, National Institutes of Health (R35GM146974) and the National Science Foundation (SES2200228 & IIS2229881). The funding sources had no involvement in the study design; in the collection,
Results

In Cameroon, 6,652 confirmed and suspected cases were officially reported between October 10, 2021 and April 24, 2022; in Haiti, 58,230 confirmed and suspected cases were reported between October 1, 2022 and August 5, 2023. In parallel, our digitization of each country’s epidemic curves yielded a total of 6,616 cases in Cameroon and 54,927 cases in Haiti, resulting in a 0.54% and 5.67% overall relative error rate, respectively (Figure 1).

Such aggregate differences between reported and digitized morbidity are quite small and reasonable for the application of EpiEstim methods, particularly in the absence of machine-readable surveillance data. In addition, we performed visual inspection of our digitized and reported epidemic curves in Haiti, where the error rate was higher and the magnitude of the cases digitized was greater than in Cameroon. This inspection (Figure 2) suggests that the case trajectories over time are highly similar — and given that the primary objective is to accurately predict disease spread, these trajectories are considerably more relevant than exact case numbers.

EpiEstim models estimated the $R_t$ for one-week-long rolling windows over the full outbreak periods to date in Cameroon and Haiti (Figure 3). The median initial $R_t$ at the start of the second outbreak week was estimated to be 1.90 (95% credible interval (CI): 1.14 - 2.95) in Cameroon and 2.60 (95% CI: 2.42 - 2.79) in Haiti. Interestingly, while $R_t$ generally remained between 0.8 and 1.2 in Haiti, it was much more variable in Cameroon. Specifically, its trajectory showed a distinct oscillating pattern; $R_t$ rose rapidly to 2.50 (95% CI: 1.97-3.25) in the third week, and then fell steeply to 0.30 (95% CI: 0.15-0.52), followed by one last rise to 1.79 (95% CI: 1.32-2.37) before settling around 1, the threshold above which an outbreak propagates. In both countries, later $R_t$ values hovered around this threshold. In Cameroon, the overall median $R_t$ was 1.06 (95% CI: 0.52 - 1.93), and nearly two-thirds (126 out of 196, or 64.28%) of all median $R_t$ estimates were above 1. In contrast, in Haiti, the overall median $R_t$ was 0.97 (95% CI: 0.80 -1.38) and only about one-third (106 out of 302, or 35.10%) of median estimates were above 1, indicating less potential for an explosive outbreak. However, while the proportion of days with an $R_t$ over 1 was higher in Cameroon, the outbreak lasted much longer in Haiti (29 vs. 44 weeks) and resulted in an overall larger number of cases (58,230 vs. 6,652).

Using published estimates of $R_0$ and administrative statistics for vaccine coverage, we then estimated time-varying OCV1 effectiveness for Cameroon and Haiti separately, following the initiation of vaccination campaigns in each country (Figure 4). In Cameroon, the median OCV1 effectiveness over the period from February 10 to April 30, 2022 was 56.21% (95% CI: 18.87 – 84.94%); it ranged from a minimum of 17.37% on March 22, 2022 to a maximum of 90.45% on April 16, 2022. In Haiti, the median OCV1 effectiveness over the period from December 19, 2022 to August 5, 2023 was 75.32% (95% CI: 54.00 – 86.39%); it ranged from a minimum of 51.26% on May 1, 2023 to a maximum of 96.87% on August 5, 2023.

Median values for OCV1 effectiveness are best understood when contrasted with herd immunity thresholds – the level at which the population is no longer susceptible to disease transmission either through immunization or previous infection. To calculate such country-specific thresholds, we used the standard approximation of $1-(1/R_0)^{t}$ for herd immunity and sourced country-specific $R_0$ values from the literature. We found that median OCV1 effectiveness estimates in Cameroon and Haiti exceeded estimated herd immunity thresholds of 48.72% and 55.95%, respectively. However, the estimated median OCV1 effectiveness in each country dropped below these thresholds at certain times, suggesting periods when sustained transmission occurred. In Cameroon, such periods with a high risk of transmission occurred soon after vaccine delivery, in late February and early-to-mid March. But by the second round of the OCV1 campaign (April 8-12, 2022), OCV1 effectiveness estimates were consistently above the estimated herd immunity threshold. In contrast, in Haiti, these dips below the herd immunity threshold happened well after OCV1 delivery started, in late April to early May and mid-June, likely allowing further disease propagation. Notably, despite a request by MSPP for a second round of OCV1 delivery, no reports of a second vaccination campaign have been identified.

Discussion

To our knowledge, this study is the first to contrast cholera outbreaks in two distinct countries and to assess the relative impact of OCV1 campaigns on local transmission using open-access,
non-machine-readable (NMR) outbreak case data and EpiEstim software. Additionally, our work provides the first application of the VaxEstim extension to a low-resource setting, relying exclusively on NMR case counts published in situation and disease outbreak news reports. Using epidemic curve digitization methods yielded reasonable estimates of cholera cases over time, with an overall discrepancy between actual and inferred case counts of approximately 5% or less. By using data from only situation and disease outbreak news reports along with parameters extracted from the published literature, we demonstrate that EpiEstim and VaxEstim can be employed quickly during outbreaks occurring in data-scarce settings to aid in decision-making and monitoring. Going forward, we encourage other infectious disease researchers and mathematical modelers to consider such tools when data are not immediately accessible in a machine-readable format.

In both Haiti and Cameroon, we found that outbreaks initially caused large surges in the effective reproduction number, $R_t$. Although the point estimate for $R_t$ in Haiti (2.60 (95% CI: 2.42-2.79)) was higher than that in Cameroon (1.90 (95% CI: 1.14-2.95)), the corresponding credible intervals partially overlapped, owing in part to greater uncertainty in $R_t$ estimates for Cameroon. Notably, our results corroborate several patterns of cholera transmission reported in prior studies. First, the order of magnitude we estimated largely aligns with values of $R_t$ previously reported in the literature, whether specifically in Haiti and Cameroon or in other sub-Saharan countries. While $R_0$ and $R_t$ are markedly different metrics, prior studies have relied on estimates of $R_t$ obtained during the first week of an outbreak to generate an “early outbreak reproductive number”. Indeed, such a value can be more closely compared to $R_0$ than later outbreak estimates of $R_t$ because interventions to curb transmission are rarely implemented in the first week following the index case. Second, in both countries, these initial periods — marked by high values of $R_t$ — were followed by longer periods during which $R_t$ oscillated around 1 but was never sufficiently low for transmission to cease completely. This temporal pattern is consistent with prior analyses of $R_t$ dynamics for cholera.

Most importantly, our results suggest that OCV1 campaigns in Cameroon and Haiti were generally effective, corroborating field surveys of one-dose vaccine effectiveness conducted in other countries, including South Sudan, Democratic Republic of the Congo, and Bangladesh. Although our point estimates of OCV1 effectiveness were predominantly higher in Haiti than in Cameroon, the corresponding country-specific credible intervals overlapped. The absence of a statistically significant difference between the two countries suggests that the effectiveness of the vaccine was similar across both contexts, despite differences in population density, ease of implementation, health and water infrastructure, and cholera outbreak history. This key finding indicates that the success of OCV1 campaigns depends less on external factors than do other vaccination campaigns. From a global health equity standpoint, such robustness in vaccine effectiveness across differing country settings is encouraging for implementation of OCV1 campaigns in response to new outbreak locations.

Since one-dose campaigns are a relatively new recommendation, only dating back to October 2022, the literature on OCV1 field studies to date remains sparse. While field-based assessments of ongoing or upcoming OCV1 campaigns will provide additional confirmatory evidence regarding vaccine effectiveness, our work demonstrates a complementary method to on-the-ground surveys — one that requires much less time and funding to conduct. Thus, this approach would be particularly applicable to locations without the resources to implement field studies of vaccine effectiveness. In addition, using digitization and EpiEstim early in an outbreak could provide near real-time information about ongoing transmission, while local applications of VaxEstim could provide dynamic insights into vaccine allocation. However, understanding the local context, drivers of disease transmission, and barriers to efficient vaccine delivery also remain critical when considering such allocation.

$R_t$ — the key marker of disease transmission — depends on the contact rate between people, the infectious disease period, and the probability of infection given contact. Since these factors may differ between Haiti and Cameroon, $R_t$ is a relevant metric to identify the most important mechanisms at play. Notably, Haiti has a much higher population density than Cameroon (415 versus 58 people per km², respectively), allowing for much higher rates of contact between individuals, which can in turn lead to explosive initial rates of transmission. The population in Haiti also has significantly lower access to improved water sources than that in Cameroon (68.29% versus 73.95%), which may increase a
Previous research has suggested additional factors that can influence vaccine effectiveness, from system-level factors like cold chain effectiveness to individual-level factors like the nutrition status of the individual receiving the vaccine or their previous history of cholera infection.\textsuperscript{81–87} Although these factors may explain some of the variability in vaccine effectiveness estimates that we observed between Haiti and Cameroon, it is important to re-emphasize that their credible intervals overlapped, suggesting relatively similar vaccine effectiveness across both contexts. While such similarity brings great promise for the deployment of OCV\textsuperscript{1} vaccines in other low-income and conflict settings affected by cholera outbreaks, additionally considering the aforementioned country-specific factors can still be useful for planning future vaccination campaigns.

Although cholera is much newer to Haiti than it is to Cameroon, the outbreaks that have affected Haiti in the last decade have been much larger and more widespread.\textsuperscript{84–87} While previous cholera infection is not fully protective against future infection, current research suggests that patients may acquire some short-term immunity from prior infection.\textsuperscript{81} Similarly, frequent exposure to cholera does seem to boost the immune response of people living in endemic countries.\textsuperscript{81} In Haiti, where the 2010 outbreak led to over 800,000 cases between 2010 and 2019, we hypothesize that, at the start of the 2022-2023 outbreak, the fraction of the population who was fully susceptible was smaller than that in Cameroon, where more frequent but smaller outbreaks have historically occurred.\textsuperscript{84–87} This assumption is further supported by findings from the most recent outbreak that occurred in Haiti, namely that the largest proportion of reported cases were in children under five who would not have been infected by earlier outbreaks given their age.\textsuperscript{84} Moreover, we found higher overall estimates of \( R_0 \) in Cameroon than in Haiti, supporting our hypothesis that the larger baseline proportion of the population who was susceptible contributed to sustained transmission cycles throughout the country’s outbreak. In addition to some existing immunity, the severe outbreaks that have ravaged Haiti have helped develop the institutional knowledge necessary to prepare for subsequent large-scale outbreaks, further health infrastructure, and fulfill vaccine needs more effectively. Such systemic efforts may also correlate with behavioral health improvements. For instance, we believe that greater awareness about the impact of cholera might also prompt higher adoption of existing treatments such as ORT or uptake of preventive interventions like better handwashing and water practices. This possibility is reflected in recent data on ORT uptake among diarrheal children, estimated at 39.05\% (95\% CI: 29.17-49.41\%) in Haiti versus 19.91\% (95\% CI: 11.95 - 31.05\%) in Cameroon.\textsuperscript{81} In contrast to Haiti, which has had only two full vaccination campaigns and one pilot program to date,\textsuperscript{91,92} Cameroon has had a large number of OCV campaigns since the first one was launched in 2015, particularly within the past 5 years.\textsuperscript{93,94} These successive campaigns may have contributed to higher vaccine coverage in Cameroon than in Haiti, despite lower levels of vaccine confidence.\textsuperscript{93} Ultimately, this resulted in lower vaccine effectiveness in Cameroon than in Haiti, because of its inverse relationship with vaccine coverage.

In contrasting the outbreaks that affected Haiti and Cameroon, we additionally considered environmental drivers of cholera transmission. As previously noted, environmental data on precipitation, air temperature, sea surface temperature, and sea salinity are openly accessible in machine-readable formats for the entirety of the globe on a routine temporal basis. Paired with time-series ARIMA models, these data can provide valuable insights to prepare for future outbreaks in these two countries and be incorporated into disease transmission models to forecast the spread of cholera as a function of evolving environmental conditions. However, if we were to scale that approach to other countries, we would be limited by the availability and quality of infectious disease surveillance data. Indeed, the richness and wide availability of environmental data is in stark contrast to most infectious disease surveillance data, which are often inaccessible, irregularly available, and critically not standardized across contexts or diseases. With emerging and re-emerging infectious diseases on the rise,\textsuperscript{95} a focus on generating high-quality, harmonized infectious disease surveillance data could have profound effects on our ability to model and forecast disease spread around the globe.

**Limitations**

While the digitization of NMR epidemic curves from situation and outbreak reports allows modelers to expand outbreak analyses by leveraging novel data
sources beyond relying solely on traditional surveillance data, their use brings some challenges. First, the digitization of epidemic curves, which provided the daily time series for analyses, can result in estimated daily counts not precisely in line with reported case counts (i.e., a 0.54% and 5.67% discrepancy from reported case count statistics in Cameroon and Haiti, respectively). Although the discrepancy between digitized and reported case counts did not significantly impact our case study, outbreak trajectories may not always be well-captured by the digitization process. Second, detailed information required for mechanistic compartmental models of cholera transmission are not always openly accessible. Examples include the size of the fully susceptible population in each area, records of chains of transmission and individual symptom onset dates, and assumptions about the underlying disease process and corresponding parameterization. Since phenomenological models like those underlying EpiEstim do not rely on assumptions about the disease process and generally require less parameterization, they are generally less precise than mechanistic compartmental models. However, a trade-off exists between the precision of epidemiologic models and their practicality and relevance for decision makers. Here, we demonstrated that using EpiEstim allows estimation of $R_e$, even early in an outbreak, without requiring more information than the SI of a disease and the current epidemic curve. A strong advantage of this built-in R package is its accessibility, allowing practitioners at all levels to make informed, local, and timely decisions.

In addition to using openly accessible outbreak case data, we also relied on published reports of vaccine coverage, which were relatively sparse. We therefore did not have access to the exact location or timing of OCV1 delivery and used the overall coverage as reported instead. Specifically, we assumed a simplified coverage timeline, with instantaneous coverage on the first date of report. Yet, by doing so, we are able to extend the VaxEstim application to places where vaccine coverage surveys are not conducted or where the resulting data are not openly-available. Additionally, in our estimation of vaccine effectiveness, we are combining information at two different geographical resolutions: our vaccine coverage data represent only small regions of each country where campaigns were conducted, in contrast to national-level outbreak case data, which were more representative. In both Haiti and Cameroon, however, OCV1 campaigns targeted the most highly affected regions, likely accounting for a substantial portion of all cases. Notably, in our sensitivity analysis (Appendix Figure 5) of the Ouest department in Haiti, we found slightly higher vaccine effectiveness than the national model presented here, although the difference was not statistically significant. Similarly, our sensitivity analyses using different levels of $R_e$ (Appendix Figure 4) yielded different point estimates for $R_e$ and vaccine effectiveness, although the corresponding credible intervals again overlapped. These additional analyses further emphasize the need for rapidly sharing local data and conducting epidemiologic analyses early in an outbreak for targeted decision making.

References
5. Khoury, P. Cholera is back but the world is looking away. BMJ 380, 141 (2023).


54. Oduro, M. S., Arhin-Donkor, S., Asiedu, L., Kadengye, D. T. & Iddi, S. SARS-CoV-2 incidence monitoring and statistical estimation of the basic and time-varying reproduction number at the early onset of the pandemic in 45 sub-Saharan African countries. *BMC Public Health*


90. World Health Organization. Cholera Vaccination 2016 - 2023: Cameroon. https://app.powerbi.com/view?r=eyJrIjoiYmFmZTBmM2EtYWM3Mi00NWYwLTg3YjgtN2Q0MjM5ZmE1ZjFkliwidCi6ImY2MTBjMGl3LWJkJqNtNGizOS044MTBiLTMkYzI4MGFmYjU5MCIsImQiOiJ9 (2023).


92. World Health Organization. Cholera Vaccination 2016 - 2023: Haiti. https://app.powerbi.com/view?r=eyJrIjoiYmFmZTBmM2EtYWM3Mi00NWYwLTg3YjgtN2Q0MjM5ZmE1ZjFkliwidCi6ImY2MTBjMGl3LWJkJqNtNGizOS044MTBiLTMkYzI4MGFmYjU5MCIsImQiOiJ9 (2023).


Figures

Figure 1: Epidemic curves for cholera outbreaks. Epidemic outbreak case data for Cameroon (in purple) and Haiti.
Haiti (in gold) are shown by date (x-axis). The y-axis shows all daily confirmed and suspected cases.

**Figure 2: Reported and Digitized Epidemic Curves, Haiti.** This figure shows the published suspected and confirmed daily outbreak case data and overall cumulative case count by Ministère de la Sante Publique et de la Population in 2022-2023 in gray, plotted with 25% transparency, overlaid with the digitized daily confirmed and suspected case counts using the metaDigitise package in R in gold.

**Figure 3: Estimated R, over time.** This figure portrays the estimated $R_t$ in weekly rolling windows over the study period of for each Cameroon (purple) and Haiti (gold). The shaded region for each country represents the 95% credible interval and the dashed blue line represents the overall median $R_t$. The black solid line represents the threshold for propagation of 1.0.

**Figure 4: Estimated rolling weekly vaccine effectiveness over time.** This figure portrays the estimated vaccine effectiveness in weekly rolling windows over the study period of for each Cameroon (purple) and Haiti (gold). The shaded region for each country represents the 95% credible interval and the dashed blue line represents the overall median vaccine effectiveness in each country. Estimates start on the first day following the vaccination campaign.