



OPEN

Combating cholera by building predictive capabilities for pathogenic *Vibrio cholerae* in Yemen

Moiz Usmani¹, Kyle D. Brumfield^{2,3}, Bailey M. Magers¹, Juan Chaves-Gonzalez⁴, Helen Ticehurst⁵, Rosa Barciela⁵, Fergus McBean⁶, Rita R. Colwell^{2,3} & Antarpreet Jutla¹✉

Cholera remains a global public health threat in regions where social vulnerabilities intersect with climate and weather processes that impact infectious *Vibrio cholerae*. While access to safe drinking water and sanitation facilities limit cholera outbreaks, sheer cost of building such infrastructure limits the ability to safeguard the population. Here, using Yemen as an example where cholera outbreak was reported in 2016, we show how predictive abilities for forecasting risk, employing sociodemographical, microbiological, and climate information of cholera, can aid in combating disease outbreak. An epidemiological analysis using Bradford Hill Criteria was employed in near-real-time to understand a predictive model's outputs and cholera cases in Yemen. We note that the model predicted cholera risk at least four weeks in advance for all governorates of Yemen with overall 72% accuracy (varies with the year). We argue the development of anticipatory decision-making frameworks for climate modulated diseases to design intervention activities and limit exposure of pathogens preemptively.

Cholera, is a signature dehydrating diarrheal disease transmitted notably via untreated drinking water carrying the causative agent, *Vibrio cholerae*¹. The disease has plagued humans for thousands of years, with reports of cholera-like symptoms documented in Sanskrit medical texts ~ 500–400 B.C.². While the sporadic seasonal outbreak has been reported throughout history¹, the first cholera pandemic was reported between 1817 and 1823. Globally, cholera has continued to spread, and the ongoing pandemic, the seventh, which began in 1961, is caused by the El Tor biotype of *V. cholerae* O1. While *V. cholerae* non-O1 serogroups are not known to cause epidemics of diarrhea, sporadic cases and small outbreaks of diarrhea and extraintestinal infections have been reported. Early environmental studies of cholera were unsuccessful in detecting reservoirs of *V. cholerae*, such as domestic animals or human carriers³, until the late 1960s, when the bacterium was detected in environmental water samples collected in cholera-free regions^{4,5}. Those *V. cholerae* were subsequently shown to be associated with zooplankton^{6–8}. Thereafter, the environmental source of *V. cholerae* was demonstrated extensively in studies carried out from 1970 to 2000 in countries around the world^{1,9–12}. However, between outbreaks and unfavorable environmental conditions, the bacterium persists in environmental reservoirs, commonly in a viable, but non-culturable state¹³. Warmer sea surface and coastal water temperatures have been identified as drivers of *V. cholerae* prevalence in the environment and are associated with increased numbers of cholera cases^{9,14,15}. Several environmental and climatic variables have been linked to the proliferation of *V. cholerae* and increased incidence of cholera, such as precipitation¹⁶, flooding¹⁷, sea surface temperature and height⁹, river level and freshwater discharge¹⁴, coastal salinity¹⁸, dissolved organic material¹⁹, chlorophyll²⁰, and components of phytoplankton and zooplankton populations²⁰. In addition, recent epidemiological surveillance suggests a link with estuarine ecosystems, namely river and coastal regions²¹.

Based on analysis of cholera records maintained in India from 1823 to 1875, cholera has been defined as occurring in two dominant forms: (1) epidemic, characterized by the sudden and sporadic occurrence of a large

¹GeoHealth and Hydrology Laboratory, Department of Environmental Engineering Sciences, University of Florida, Gainesville, FL, USA. ²Maryland Pathogen Research Institute, University of Maryland, College Park, MD 20742, USA. ³Institute for Advanced Computer Studies, University of Maryland, University of Maryland, College Park, MD 20742, USA. ⁴United Nations Office for the Coordination of Humanitarian Affairs, New York, NY, USA. ⁵Meteorological Office, Exeter, UK. ⁶Foreign, Commonwealth and Development Office, London, UK. ✉email: antar.jutla@essie.ufl.edu

number of cases; and (2) endemic, where cholera cases occur at a baseline level throughout the year with distinct seasonal peaks^{22,23}. Epidemic cholera is hypothesized to be related to elevated air temperatures followed by above-average precipitation in concatenation with insufficient and/or damaged water, sanitation, and hygiene (WASH) infrastructure, placing the human population at high risk of interaction with cholera bacteria and a subsequent disease outbreak²⁴ (Fig. 1). Endemic cholera is associated with the constant occurrence of cholera cases, primarily in regions where coastal or terrestrial water systems create favorable conditions for the growth and proliferation of *V. cholerae*. Under certain environmental conditions, a sustained epidemic mode of cholera can evolve into the endemic form in regions with enhanced and continuing exposure to, and transmission of, *V. cholerae*.

As a pandemic disease, cholera affects millions in vulnerable human populations²⁵ and persists as a dominant water-borne disease in Latin America, sub-Saharan Africa, and Southern Asia^{26,27}. Massive cholera outbreaks are associated with natural and anthropogenic disasters, notably when environmental conditions favor the growth of the bacterium²⁸. The cholera outbreak in Haiti, which occurred during the months following Hurricane Matthew²², serves as a prime example. Since March 2015, Yemen, a coastal Middle Eastern country, has suffered violent surges of civil unrest²⁹, and in October 2016, the country reported a few cholera cases. By the end of 2017, Yemen accounted for *ca.* 80% of cholera cases worldwide since 2015³⁰. During the first six months of the outbreak, cholera in Yemen surpassed the number of reported cases in Haiti over a span of seven years (*ca.* 815,000 cases between 2010 and 2017), now considered historically the largest cholera epidemic³¹.

Cholera is unlikely to be eradicated since the disease-causing agent is autochthonous to aquatic environments and plays a role in those environments' carbon and nitrogen cycles²⁰. Clearly, the ecology of *V. cholerae* must be understood in the context of its natural aquatic habitat and the changing climate, hence a driver of cholera as a potential re-emerging infectious disease.

Disease prediction can be achieved by recognizing that disease progression comprises two components, trigger and transmission, which result in an outbreak and, subsequently, public health emergency^{23,32}. For cholera, "trigger" represents those mechanisms that stimulate the growth, multiplication, and distribution of *V. cholerae* bacteria in the environment. Water insecurity, namely lack of access to safe water and sanitation, enhances the bacterium's interaction with human populations. Per contra, "transmission" comprises pathways that allow the dissemination of *V. cholerae* and involves complex interaction routes between humans and contaminated water²².

In our previous research, protocols were developed to predict cholera in various regions throughout the world^{19,16,33–37}. The hypothesis shown in Fig. 1 was validated using retrospective data from countries in Africa and Asia over a decadal timeframe, most recently in Ukraine³⁸. However, reliable spatial and temporal datasets containing disease prevalence or incidence time series are a major challenge for infectious disease prediction. Without consistent time series datasets, our previous studies^{24,39,40} relied on spatial pattern recognition principles to identify hydroclimatic and environmental processes associated with cholera trigger across geographical regions of interest. Yemen provided a unique opportunity because spatial and temporal cholera prevalence datasets were available from the beginning of the cholera epidemic up to the present, allowing model validation. One of the significant gaps to control cholera is inherent to the absence of knowledge on when and where an outbreak is likely to occur. The World Health Organization (WHO) Global Task Force on Cholera Control (GTFCC) initiative aims to reduce cholera deaths by 90% and eliminate cholera in at least 20 countries by 2030⁴¹. While vaccines⁴² and other interventions, e.g., removal of zooplankton and particulates by filtration^{43,44} and access to WASH infrastructure⁴⁵ are critical to fight cholera disease, a predictive capability and capacity will likely provide an additional toolset for combating cholera outbreaks globally¹. This reflective study presents an overview of the integration of microbiology of cholera used to develop a predictive system that assists aid agencies, namely by reporting cholera risk. Therefore, the objective of this study was to determine the evolution of cholera in Yemen, with the specific aim to validate the cholera trigger using the well-established epidemiological Bradford Hill

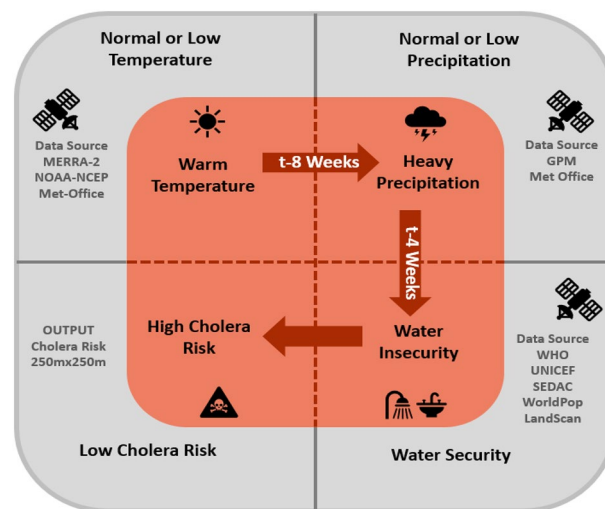


Figure 1. Cholera trigger mechanism. Adapted from “Environmental Factors Influencing Epidemic Cholera”, by Jutla et al., 2013, *The American journal of tropical medicine and hygiene* 89 (3), 597–607.

Criteria (BHC) for causation^{46,47}. The ultimate goal is to provide assistance in enhancing epidemiological and medicine-based decision-making so that future cholera outbreaks can be prevented.

Results and discussion

The objective of this study is to evaluate the performance of a prediction system for cholera, employing near real-time data collected over the past three years by the Yemeni governorates and using Bradford Hill Criteria for the epidemiological association. The first indication of cholera in Yemen was noticed in October 2016. Subsequently, epidemiological incidence data became available in June 2017. The Cholera Risk Model (CRM) trigger module captured the increased risk of a massive cholera outbreak in Yemen (Fig. 2) with *ca.* 92% spatial match between locations where cholera cases were reported, and high risk was computed. The trigger component of the CRM had been validated previously, using historical data for Sudan, Bangladesh, Mozambique, Zimbabwe, Cameroon, and Haiti^{22,24,39}. However, Yemen was unique since epidemiological conditions allowed for assessment of model performance in near real-time. BHC^{46,48}, a set of ten parameters intended to provide epidemiological evidence for causal relationships between a public health outcome and factors influencing said outcome, was used as defining criteria for CRM performance. The following sub-sections detail each parameter of the BHC for model performance.

Strength. Strength is a parameter of BHC that provides epidemiological evidence for the associational relationship between disease prevalence and factors influencing disease outbreaks. Furthermore, the climate has been established as a driver of cholera^{24,49}, and temperature and rainfall facilitate the growth and proliferation of the bacterium and enhance its metabolic activity. Cholera infections have a hallmark seasonal distribution, with most cases occurring during the warmer months when water temperature and salinity are optimal for the growth of *V. cholerae*. In regions of the world where it is endemic, cholera typically occurs as a single annual peak in human disease cases. However, bi-annual peaks in cholera cases are typical for the Bengal Delta region^{26,50}.

To evaluate the strength of CRM, a correlation was calculated, at the governorate level, between cholera prevalence from 2017 to 2019 and risk values computed for the same time period, using the parametric (Pearson) and the non-parametric rank correlation coefficient (Kendall Tau scores). The Pearson method showed a significant ($p < 0.05$) positive correlation for all governorates except Aden (Fig. 3a). Similarly, Kendall Tau values were statistically significant ($p < 0.05$) for all governorates, indicating adequate model strength. The three-year correlation analysis provided evidence of overall model performance. However, it may be argued that if effective intervention strategies were employed, such as robust access to WASH, a decline in model performance over those years would have been observed. The CRM trigger module is designed to capture disease initiation in a region. Therefore, unless a new outbreak(s) occurs in a given area, the model performance should decline over time since the transmission dynamics would dictate the spread of cholera in a human population. Accordingly, correlation analysis was performed on a weekly scale for each year. The Pearson correlation for 19 of 21 (2017), 11 of 21 (2018), and 15 of 20 (2019) governorates, respectively, exhibited a significant ($p < 0.05$) association between computed risk and disease prevalence (Fig. 3b). Using Kendall Tau, 19, 8, and 13 governorates showed a statistically significant association for the same years (Fig. 3c). In 2017, the model detected an increased risk for more than 90% of the governorates, with Aden being the only governorate not determined to be at increased risk. Thus, it can be concluded that the results show correlation analysis can be a useful means of determining the strength of the model. During 2018, the model captured *ca.* 45% (combining both Pearson and Kendall Tau), compared to 2019, where *ca.* 68% of the governorates were captured. The decrease in model performance for 2018 is likely an indication of a change in the definition of cholera-like symptoms and/or the intervention strategies employed to mitigate cholera in the region. However, an increase in the model performance was observed for 2019, suggesting new outbreaks in the region compared to the prior year.

Specificity. Specificity is used here to evaluate the predictive capacity of CRM and is achieved by quantifying causality of model output with cholera prevalence. Causality is quantified using three statistical metrics: accuracy, sensitivity, and specificity, as defined below:

$$\text{Accuracy} = (t_p + t_n) / (t_p + t_n + f_p + f_n) \quad (1)$$

$$\text{Sensitivity} = t_p / (t_p + f_n) \quad (2)$$

$$\text{Specificity} = t_n / (t_n + f_p) \quad (3)$$

Here an increase in risk is considered a true positive (t_p) if it captures the increase in reported cases and a decrease in a true negative (t_n) if it captures the decrease in cases. If the increase in computed risk fails to capture the increase in risk, it is considered a false positive (f_p); and if a decreased risk fails to capture the decrease in cases, it is considered a false negative (f_n). The confusion matrix with these variables is provided in Table S1. As shown in Fig. 4a, the cholera risk model met all three criteria for causality more than 60% of the time. Sensitivity and specificity varied from 55 to 67%, with averages of 60% and 61%, respectively, indicating that the model can detect increases and decreases in cholera cases across a given region. Across nine governorates, where more than 100,000 cholera cases had been reported, the model accuracy varied between 57 and 67%, with an average of 60%.

Biological gradient. Traditionally, the biological gradient is interpreted as a monotonic gradient, indicating direct proportionality of cause of an increase in disease burden with exposure risk. Commensalism of *V.*

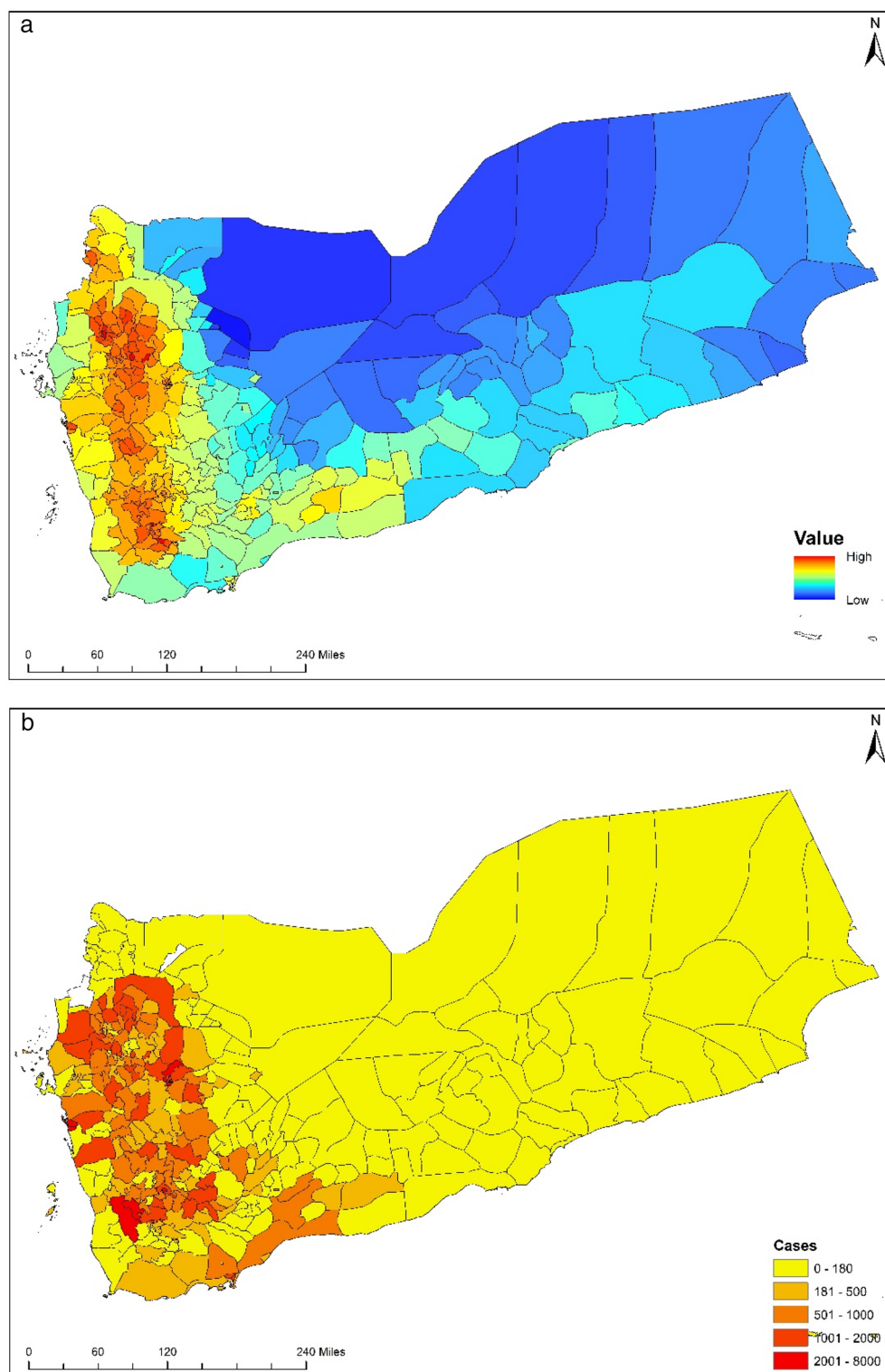


Figure 2. (a) Cholera risk map of Yemen for June, 2017 produced on May 30, 2017; (b) Actual cholera cases observed in June, 2017, maps are generated using ESRI's ArcMap version 10.7 (URL provided in supplementary material).

cholerae with copepods has been established^{16,43,51}. *V. cholerae* attaches to the gut and carapace of copepods, with a single copepod capable of harboring up to 10^4 *V. cholerae* cells⁵¹. Since cholera is dose-dependent, requiring ingestion of *ca.* 10^3 – 10^6 *V. cholerae* cells to induce severe diarrhea symptomatic of cholera, according to human

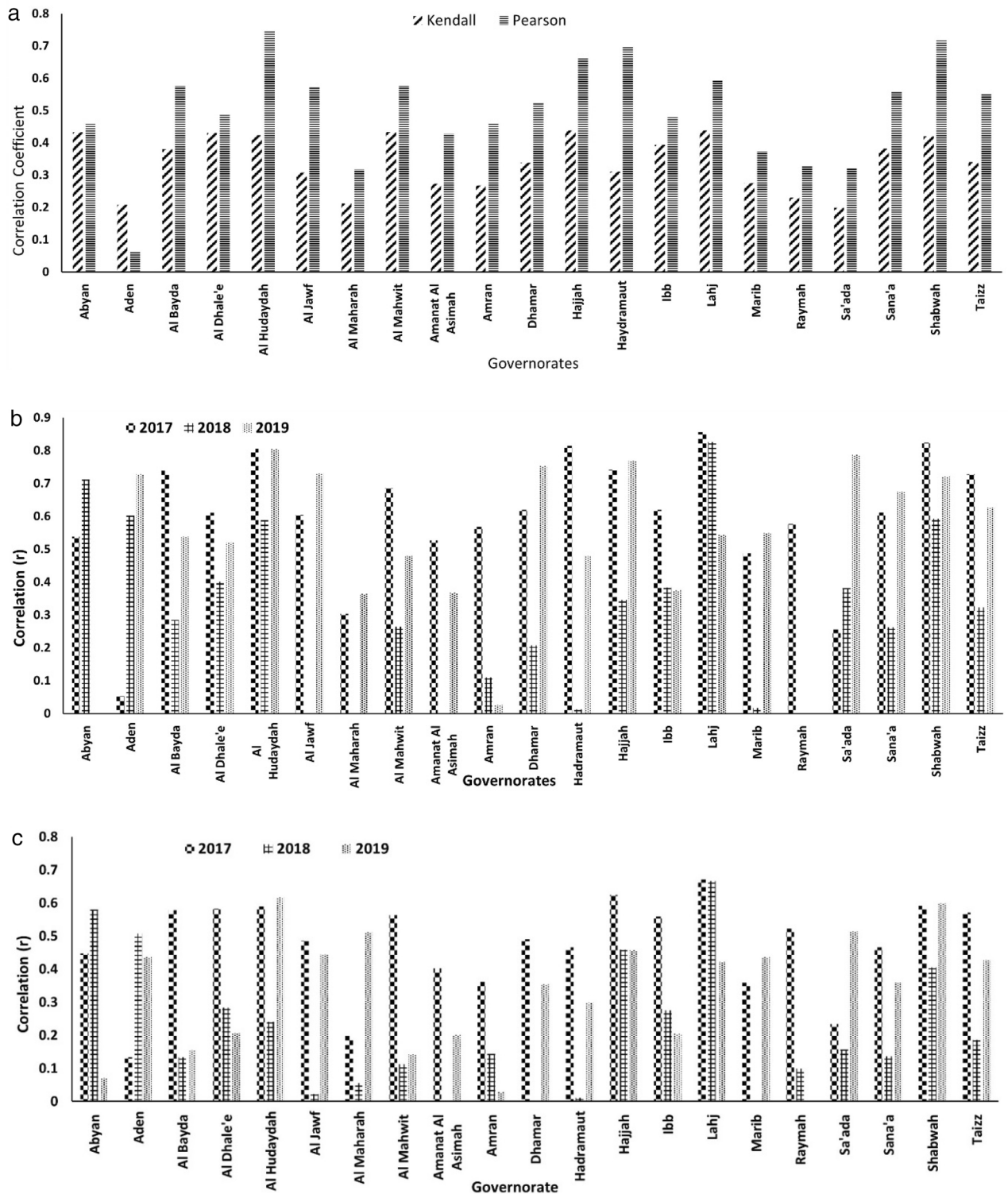


Figure 3. (a) Correlation coefficients between cholera cases and risk values for all governorates, all governorates except for Aden exhibited statistically significant ($p < 0.05$) correlation, (b) Pearson correlation coefficient between cholera cases and risk values for all governorates for individual years (2017, 2018, 2019) and 19 (of 21), 11 (of 21), and 15 (of 20) governorates are statistically significant in 2017, 2018 and 2019 respectively, (c) Kendall Tau correlation coefficient between cholera cases and risk values for all governorates for individual years (2017, 2018, 2019) and 19, 8, and 13 governorates are statistically significant in 2017, 2018 and 2019 respectively.

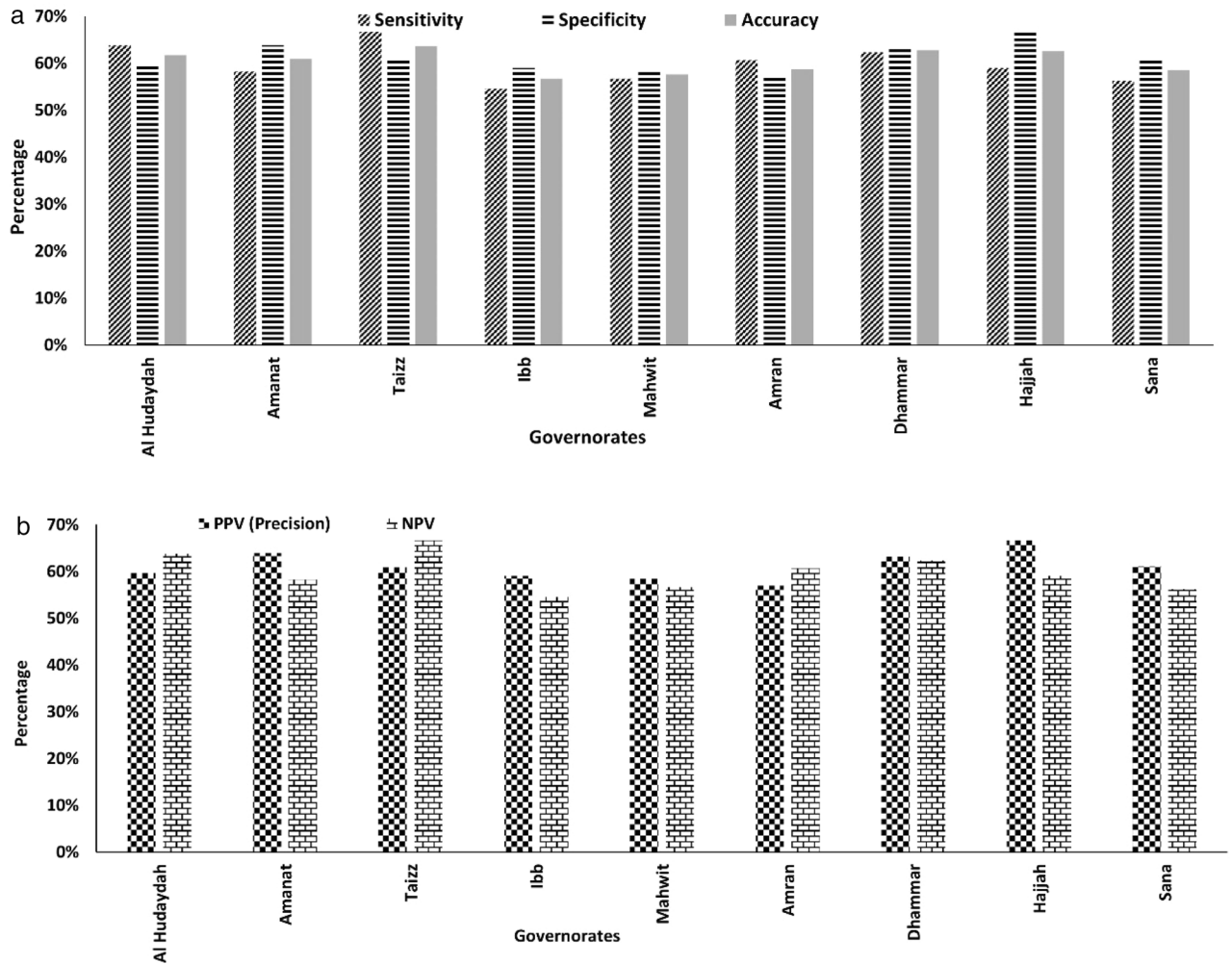


Figure 4. (a) Sensitivity $[t_p/(t_p + f_n)]$, Specificity $[t_n/(t_n + f_p)]$, and Accuracy $[(t_p + t_n)/(t_p + f_p + t_n + f_n)]$ of cholera prediction system. (b) Positive Predictive Value (PPV) or precision $[t_p/(t_p + f_p)]$, and Negative Predictive Value (NPV) $[t_n/(t_n + f_n)]$ of the trigger module of cholera prediction system.

volunteer studies^{52,53}, ingestion of untreated water containing a small number of copepods is sufficient to promote disease. Therefore, if there is contact between human populations and cholera bacteria in the environment, zooplankton blooms can result in numbers of *V. cholerae* well above an infectious dose.

It has been argued that BHC should include non-monotonic and more complex relationships between trigger and transmission, with variation in reported cases⁴⁶. Hence, to evaluate the biological gradient in this study, the relationship between change in computed risk, using the model, with change in reported disease prevalence was explored. To evaluate the incidental (biological) gradient of the model, the positive predictive value (PPV; frequently referred to as precision) and negative predictive value (NPV) of the model outcomes were computed. PPV is a fraction of positive computed risk, which can capture positive change in prevalence. This study used an increase in relative risk and reported disease rather than absolute values. NPV is a fraction of negative (decreased) computed risk, which can capture a negative change in prevalence. These two indicators were determined for those governorates for which specificity (causality) of the model had been calculated and were determined as follows:

$$\text{Positive predictive value} = t_p / (t_p + f_p) \quad (4)$$

$$\text{Negative predictive value} = t_n / (t_n + f_n) \quad (5)$$

Using the three years of data, we determined PPV and NPV for the model (Fig. 4b). PPV and NPV values varied between 57 to 67% and 55% to 67%, respectively, with an average of 60%, suggesting that at least 60% of the time, the model correctly responded to an increase or decrease in the number of reported cholera cases. However, for practical use of the cause-effect relationship, temporality is important, with cause preceding effect and lead time.

Temporality. Epidemiological temporality comprises the duration of exposure and extent of impact in terms of severity or number of incidences. To assess disease risk prediction, lead time is an essential criterion because it provides time to intervene and limit the outbreak's impact. Rainfall has been established as an environmental driver of cholera. In Haiti, increased rainfall was associated with increased cholera risk, with a lag time of up to one month^{1,24,35}. Similarly, in Bangladesh, where rainfall/runoff response from upstream catchment areas is about 3–4 weeks, increased rainfall in July has been associated with increased streamflow of major rivers, causing sediment resuspension and attributed to an increase in cholera cases during the month of August⁵⁴. Risk computed in Yemen using the CRM provides a lead time of four weeks from predicted incidence—providing ample time for intervention and mobilization of resources. The hypothesis presented in Fig. 1 shows cholera cases are generally observed four weeks after a period of anomalous warm temperatures, followed by anomalous high precipitation where there is a significant deviation in population behavior, with respect to water use, caused by damaged WASH infrastructure. Importantly, the hypothesis for the temporality of BHC has been supported in several regions in Africa³⁹, Asia⁴⁰, and Latin America²².

Consistency. Consistency is an essential parameter of the BHC to ensure the reproducibility of findings across different samples and locations. Our key hypothesis in Fig. 1 argues that damaged WASH infrastructure and a combination of hydroclimatic processes favor conditions for an outbreak of cholera. This cause-and-effect relationship has been observed in many studies^{21,24,40}. The hypothesis has been validated for countries in Africa, Asia, and the Americas, reinforcing its reproducibility through its utilization in Yemen. Attributing the cholera outbreak in Peru to El-Niño events in the Central Pacific was one of the earliest precursors to this hypothesis¹. Studies conducted using data from Bangladesh¹⁶ and Haiti²² report a strong relationship between rainfall and the incidence of cholera. In Bangladesh, cholera occurs annually in a bimodal cycle. The first peak occurs in the spring, and a larger peak occurs following the fall monsoon season. Cholera seasonality also coincides with the warmest temperatures of the year and is reduced to sporadic incidence as the temperature decreases in winter²¹. Haiti has been the main focus of cholera research since the 2010–2011 outbreak, which identified rainfall as a critical driver of the disease in that country⁵⁵. Rainfall can significantly impact the water source, e.g., nutrient concentration, salinity, pH, river level, and freshwater discharge, which affect the growth and persistence of *V. cholerae* and its zooplankton host in the environment. Various studies have reported air temperature and precipitation as dominant hydroclimatic variables impacting the occurrence and transmission of cholera in various parts of the world (Table S2).

Plausibility. Biological plausibility under BHC can be used to assess the association between a putative cause and an observed outcome within the context of existing biological and medical knowledge. In aquatic reservoirs, elevated temperatures can cause a density differential amongst layers of the water column, contributing to the stratification of bacterial populations. In addition, stratification promoted by temperature, dissolved oxygen, pH, and other physical/chemical parameters can determine non-uniform microbial community profiles in the water column contributing to environmental conditions enhancing bacterial growth and multiplication. These conditions are generally favorable for the multiplication of zooplankton, namely copepods, shown by Kaneko and Colwell⁶ to host *Vibrio spp.*, including *V. cholerae*. Subsequently, it was shown that employing simple filtration can effectively remove zooplankton and particulate matter, hence attached *V. cholerae*, from drinking water and was used to reduce the number of cholera cases in Bangladeshi villages by more than 50%^{43,44}. Collectively, these studies show that copepods serve as host/vector of *V. cholerae*.

Therefore, ecological parameters enhancing the growth and proliferation of cholera bacteria frame the model developed for risk prediction. Heavy rainfall that follows a period of high air temperature aids explosive growth of bacteria in water bodies that serve communities as drinking water source²⁴. Thus, an inadequate water supply infrastructure exposes a given population to untreated water. A prime example is Yemen, a Middle Eastern country grappling with war and frequently experiencing floods, with the population lacking proper WASH conditions—considered the dominant sociological cause of the continuing cholera epidemic rampant in most Yemen governorates. Hence, identifying and describing the mechanics of the trigger, a rational clarification of the putative 'black box' between the biology and ecology of an infectious agent and its disease epidemiology, is now possible.

Coherence. Under BHC, coherence is the idea that experimental findings in the laboratory support epidemiological observations. The hypothesis used in Yemen to predict the cholera risk had exhibited coherence in its application in Yemen and in laboratory findings. The consistency between the epidemiological data and the predicted risk is exhibited through the correlation analysis as shown in Fig. 3. Year 2017 and 2019 exhibited high associations suggesting coherence between the predicted risk and the epidemiological data. While through laboratory studies we found that in the environment, an increase in *V. cholerae* populations was observed in water and plankton samples collected in a longitudinal, multi-year study carried out in the Chesapeake Bay, Maryland, USA. Results showed when the water temperature rose above 19 °C, *V. cholerae* populations in the water column proliferated with the elevated temperatures⁵⁶. Similarly, water samples collected in estuarine zones of the Bengal Delta yielded similar results, confirming enhanced growth of *V. cholerae* in warmer pond water¹⁹. Huq et al.³⁶ showed a 5 °C increase in water temperature resulted in a 30-fold increased risk of a cholera outbreak, with a lag of six weeks. Over the past decades, extreme heat events in Northern Europe have been linked to an increased number of reported *Vibrio* infections^{57,58}. Archived samples collected by continuous plankton recorder over 60 years and analyzed using molecular genetic methods by Vezzulli and colleagues^{15,59} showed the global warming trend in sea surface temperatures of the North Sea was strongly associated with proliferation of populations of *Vibrio spp.*, including *V. cholerae* and vibriosis in populations inhabiting coastal regions. In fact,

among the environmental variables examined, increased sea surface temperature explained ca. 45% of variance in those studies.

Extreme precipitation can also impact *V. cholerae* populations in the environment, with potential to alter salinity profiles and nutrient availability, as well as sea level in coastal waters, with increased freshwater inflow. For example, strong positive correlations between rainfall patterns and cholera epidemics have been observed in Bangladesh¹⁶, India²⁰, Ghana⁶⁰, Cameroon⁶¹, and numerous other locations in Asia, Africa, and South America⁶². Combined with findings from laboratory experiments conducted in different regions of the world (Table 1), these observations comprise a crucial validation based on experimental evidence that supports the hypothesis of this study (Fig. 1).

Experiment. Various studies have associated hydroclimatic variables with cholera trigger and transmission risk^{20,22,26,39}. As discussed above, laboratory-based investigations showed *V. cholerae* thrives in aquatic environments where the water temperature is between 20 and 45°C⁶³. Experimental studies have shown an increased risk of cholera when air temperatures are between 19 and 28°C^{56,64}, coupled with increasing water entrapment^{23,24}. Temperature and precipitation are primary factors in prediction but do not trigger or control the spread of cholera when each is considered independently⁶⁵. The combination of warm temperature converging with heavy rainfall and inadequate WASH infrastructure^{45,66} has a high probability of leading to an outbreak of the disease^{24,39}. Furthermore, Colwell and colleagues¹ demonstrated resuscitation of VBNC *V. cholerae* to the culturable state in the stool, following ingestion of VBNC cells by human volunteers, evidence that non-culturable *V. cholerae* can cause disease. That is, VBNC cells retain virulence potential. Conversion of antigenic *V. cholerae* serotype O1 to non-O1 (as well as non-O1 to O1) has also been demonstrated in laboratory experiments^{67,68}, with the conclusion that all *V. cholerae*, regardless of serotype, should be considered potentially pathogenic.

Since *V. cholerae* is a ubiquitous and naturally occurring inhabitant of aquatic environments globally^{69,70}, conditions favoring its growth and multiplication suggest that incorporating a single parameter provides, at best, incomplete description of disease trigger and transmission of cholera. The CRM incorporates both temperature and precipitation as hydroclimatic variables, which demonstrated experimentally to have a significant association with the proliferation of the bacterium and cholera.

Analogy. Analogy of BHC indicates similarities between observations resulting in same delivered outcomes. Unfortunately, cholera outbreaks are a regular phenomenon in regions of the world that are subjected to anomalous precipitation that is also associated with anomalous air temperatures (hydroclimatic conditions)^{20,22,24,26}, notably regions with damaged WASH infrastructure^{24,40}. Spatial analyses of data from India, Bangladesh, Nepal, Mozambique, Cameroon, Central African Republic, Congo, and Zimbabwe exhibit a similar hydroclimatic pattern related to cholera outbreaks^{22,26,27,29,71}. Damaged WASH infrastructure accelerates interaction with *V. cholerae* by increased exposure to lack of safe water, sanitation, and hygiene and the likelihood of waterborne disease. In 2015, implementation of sufficient WASH infrastructure was demonstrated in Nepal to have the potential to reduce the extent of an outbreak, even when hydroclimatic conditions favor an outbreak of cholera⁴⁰. Clearly, WASH infrastructure can be highly effective in controlling the spread of cholera in a population and must be considered in the context of public health, as was the case in Yemen in 2018 (Figure S2).

Reversibility. As a final criterion of BHC, if the cause is removed, then the effect should disappear as well. Hence, WASH, included in CRM trigger analysis, allows testing reversibility of the model. After the 2015 earthquake, Nepal presented environmental conditions indicative of a massive cholera epidemic⁴⁰. However, the observed outbreak was less than critical and only a few cholera cases were reported. The comparably low number of cholera cases was attributed to the implementation of extremely effective WASH infrastructure and the availability of WASH facilities. The steps that were taken played a crucial role in preventing a massive cholera outbreak in Nepal. The CRM, by incorporating WASH, precipitation, and air temperature, can test for revers-

Authors	Region
Huq et al. ³⁶	Laboratory environment from the Bay of Bengal
Huq et al. ⁷⁶	Laboratory environment from the Bay of Bengal
Skorupski and Taylor ⁷⁷	Laboratory environment
Louis et al. ⁵⁶	Laboratory environment from the Chesapeake Bay
Schuhmacher and Klose ⁷⁸	Laboratory environment
Singleton et al. ⁷⁹	Laboratory environment
Ravel et al. ⁸⁰	Laboratory environment
Eiler et al. ⁸¹	Laboratory environment in Baltic Sea
Austin and Swing ⁸²	Summary of the impact of temperature from various regions
Stauder et al. ⁸³	Laboratory environment
McCarthy ⁸⁴	Laboratory experiments in USA

Table 1. Evidence of coherence using BHC.

ibility. Every variable failing to demonstrate positive anomalous variability from the long-term average can be interpreted as indicating minimal cholera risk.

Table 2 summarizes the BHC criteria used to evaluate the CRM trigger module, which provides an assessment of cholera risk at least four weeks in advance (based on previously published results³²). The motivation for using risk score, rather than prevalence or incidence, is the ability to circumvent missing data during public health emergencies, a situation common during most major cholera outbreaks. Therefore, the risk score remains associated with reported epidemiological data to evaluate the validity of the CRM. In fact, using the model to respond quickly to outbreaks does not eliminate the need to respond to epidemiological data. Instead, it offers a credible risk prediction tool that aims to anticipate when an early response could assist in changing the shape of an epidemiological curve. A critical observation to use outputs from CRM was that there were no protocols available to make decisions for implementing intervention and mitigation efforts by public health officials. It was noted that the majority of humanitarian response to an outbreak of waterborne disease (such as cholera) falls within the reactionary domain implying that the intervention and mitigation activities are generally initiated after an outbreak has been reported in a region. There needs to be a paradigm shift to the anticipatory decision-making process that prediction-based prevention will likely be more effective in reducing the burden of diseases. However, intervention activities in the anticipatory decision-making domain are likely to be significantly different and unique since it may be geared towards preparedness for an impending cholera outbreak (and not for ongoing outbreak), which remains within the scope of future research. Figure 4a and b show results based on environmental conditions of the cholera trigger module related to access to safe water and sanitation. When information on WASH infrastructure becomes available, values for four metrics (Sensitivity, specificity, accuracy, and precision) will improve significantly. Most of the population-centric governorates (Amanat, Amran, Dhamar, Sana'a) were statistically insignificant relative to cholera trigger risk in 2018, most likely an indication that progressive intervention and mitigation were activated and effective.

Conclusion

This study presents a unique opportunity to reflect upon the future of understanding outbreaks of water-borne diseases where pathogenic growth of *Vibrio spp.* is linked with modalities of water, climate, and environmental processes. Our research results indicate that the solution to water-borne diseases, such as cholera should not be limited within the constraints and bounds of traditional medicinal domains. Rather, a broader approach must be followed with due diligence given to data-driven methodologies aimed at predicting cholera risk in the human population.

We do wish to highlight the importance of using the CRM risk scores rather than the traditional disease model output. The reason is that risk scores are independent of the cholera time series and, therefore, are useful for decision making and developing policies based on the severity of the disease. There are two challenges or limitations of the study: first, the major limitation is inconsistency in the definition of cholera, which was altered in 2018 for Yemen. However, our previous research carried out in Bangladesh^{36,54,72} showed that acute diarrheal disease typically comprises 20 to 30% cholera cases. Second: the availability of global social burden data, including WASH and natural resources. It is important to note that the spatial resolution of the epidemiological data is very coarse, whereas the model output is at 1 km × 1 km. The information available for water and sanitation infrastructure is not efficiently collected; that is, it is not available to be used in our current modeling system.

In summary, the study reported here represents the first to monitor cholera in Yemen with the objective of validating a near real-time cholera prediction model. BHC comprises a series of ten parameters that can be used to provide epidemiological evidence of a causal relationship between public health outcomes and factors

Criteria	Parameter	Fulfillment
Strength	Strong association (correlation) is more causal than weak association	Correlation analysis (Fig. 3)
Consistency	Consistent findings from other studies	Previous studies support the correlation (Huq et al. ²⁴ , Khan et al. ⁴⁰ , Lipp et al. ²¹ , Colwell ¹ , Hashizume et al. ¹⁶ , Khan et al. ²²)
Specificity	Causality of CPS is evaluated through Sensitivity, specificity, and accuracy	Figure 4a
Temporality	Cause occurs before effect	A four-week lead time in hydroclimatic processes was observed to be the cause of cholera
Biological gradient	Higher exposure leads to more public health burden	The gradient analysis was conducted in terms of PPV (precision) and NPV (Fig. 4b)
Plausibility	Mechanism of cause	Previous studies have established precipitation and temperature as the mechanics of survival of cholera bacteria in the environment
Coherence	Epidemiological findings match with laboratory/observational/analytical experiments	Previous studies have determined the presence of cholera bacteria in an aquatic environment (Louis et al. ⁵⁶ , Neogi et al. ¹⁹)
Experiment	Experimental or analytical evidence	Direct dependence of increase in temperature and precipitation with the increase in cholera risk (Hood et al. ⁶⁴ , Louis et al. ⁵⁶ , Huq et al. ²⁴)
Analogy	Are there any similarities/dissimilarities between the observed association to other processes?	A spatial analysis from India, Bangladesh, Nepal, Mozambique, Cameroon, Central African Republic, Congo, Zimbabwe shows a similar pattern of origin of cholera
Reversibility	Do preventative actions lead to alteration of cause-effect or vice versa?	Preventative actions may have a positive cause-effect impact on the reduction of cholera cases in the year 2018 (Fig. S2)

Table 2. Bradford Hill Criteria to evaluate trigger module of the cholera prediction system.

influencing the outcome. BHC was effectively employed to assess the performance of the CRM trigger of the cholera prediction system. When epidemiological data on cholera prevalence becomes available on a finer spatial scale, it will be possible to determine cholera hot spots in locations where the human population density is both high and vulnerable to frequent environmental disturbances in water and sanitation. It is concluded that BHC can be used for model evaluation and performance of CRM with accuracy and as shown here, the results of the BHC evaluations are interpreted as being strongly favorable for CRM.

Methods

We built CRM as an integrated platform, using a heuristic approach, that calculates the risk of the trigger of cholera, model's detail is provided in the supplementary material. Our previous research created a pathway to the formulation of CRM. The hypothesis was first developed over the Indus River Basin²⁴ and defined environmental variables' association with cholera outbreaks. Further, retrospectively, these associations were quantified and validated over Africa³⁹, Asia⁴⁰, and Caribbean²² regions. Here, utilizing those protocols, we developed CRM, a near real-time cholera risk prediction model, to predict the trigger of a cholera outbreak. Trigger represents those mechanisms stimulating cholera bacteria growth, multiplication, and persistence in the environment, after which under given water insecurity conditions, the interaction of the bacteria with the human population occurs. CRM evaluated cholera in Yemen from 2017 to 2019, as shown in Fig. 2a. CRM comprises a trigger^{22,73} module, which uses data for precipitation, temperature, population, and (WASH) infrastructure to compute a risk score with categorical values of high and low risk of cholera in a given region. The trigger algorithm identifies anomalous temperature and rainfall conditions, providing an assessment of cholera for the following four weeks for a given region. Details of model development and algorithmic architecture have been published elsewhere⁴⁰.

Daily and monthly rainfall data at two different resolutions were obtained from the National Aeronautics and Space Administration (NASA). Monthly rainfall data ($0.25^\circ \times 0.25^\circ$) from the Tropical Rainfall Measuring Mission were employed to compute the long-term average. Daily rainfall data ($0.1^\circ \times 0.1^\circ$) were obtained from the Global Precipitation Mission and used to determine precipitation variation from the long-term average at resampled data points. Daily and monthly data for air temperature on the surface ($0.5^\circ \times 0.625^\circ$) were obtained from the NASA Modern-Era Retrospective analysis Research and Application Version 2 and used to determine temperature variation and deviation from long-term averages. LandScan population data ($1 \text{ km} \times 1 \text{ km}$) were obtained from Oak Ridge National Laboratory and used in the model to represent human population (averaged over 24 h) distribution. The CRM outputs were resampled at a spatial resolution of 10 km to provide predicted cholera risk for the following four weeks. Weekly reports of cholera cases at the governorate level, between January 2017 and July 2019, were obtained from the Early Warning, Alert, and Response System and The Assessment Capacities Project⁷⁴. Figure S1 shows total cholera cases across all governorates during 2017, 2018, and the first 28 weeks of 2019.

The trigger model output is a risk score that ranges from high (numerical value of 1) to low (numerical value of 0), inherently different from traditional compartmental disease models where output is usually presented as prevalence or incidence of cholera^{22,75}. Pearson (parametric) and Kendall Tau rank (non-parametric) correlation coefficients were used to establish the association between the CRM risk score and cholera cases for all Yemen governorates. For each time point, trigger risk scores were computed and compared with the total number of cholera cases reported during the following four weeks. That is, the efficacy of forecasted cholera risk (four weeks in advance) was evaluated for the trigger module in near real-time. Complementary analyses of sensitivity, specificity, accuracy, precision, and NPV were used to indicate the association between the changes in model risk scores and the change in the number of cholera cases.

Data availability

All data generated or analyzed during this study are included in this published article.

Received: 14 March 2022; Accepted: 21 October 2022

Published online: 08 February 2023

References

- Colwell, R. R. Global climate and infectious disease: The cholera paradigm. *Science* **274**, 2025–2031 (1996).
- Bhishagratna, K. L. *et al.* An English translation of *The Sushruta Samhita: based on original Sanskrit text*. vol. 2 (author, 1911).
- Pollitzer, R. Cholera studies. 1. History of the disease. *Bull World Health Organ* **10**, 421–461 (1954).
- Mukerjee, S., Basu, S. & Bhattacharya, P. A new trend in cholera epidemiology. *BMJ* **2**, 837–839 (1965).
- Pollitzer, R., Swaroop, S. & Burrows, W. *Cholera*. (World Health Organization, 1959).
- Kaneko, T. & Colwell, R. R. Adsorption of vibrio parahaemolyticus onto chitin and copepods. *Appl. Microbiol.* **29**, 269–274 (1975).
- Kaper, J., Lockman, H., Colwell, R. R. & Joseph, S. W. Ecology, serology, and enterotoxin production of *Vibrio cholerae* in Chesapeake Bay. *Appl. Environ. Microbiol.* **37**, 91–103 (1979).
- Tamplin, M. L., Gauzens, A. L., Huq, A., Sack, D. A. & Colwell, R. R. Attachment of *Vibrio cholerae* serogroup O1 to zooplankton and phytoplankton of Bangladesh waters. *Appl. Environ. Microbiol.* **56**, 1977–1980 (1990).
- Lobitz, B. *et al.* Climate and infectious disease: Use of remote sensing for detection of *Vibrio cholerae* by indirect measurement. *Proc. Natl. Acad. Sci.* **97**, 1438–1443 (2000).
- Bougoudogo, F. Cholera in Africa and in Latin America. *Bull Soc Pathol Exot* **91**, 406 (1998).
- Colwell, R. R. & Huq, A. Environmental reservoir of vibrio cholerae the causative agent of Cholera. *Ann. N.Y. Acad. Sci.* **740**, 44–54 (1994).
- Colwell, R. R. & Huq, A. Global microbial ecology: biogeography and diversity of *Vibrios* as a model. *J. Appl. Microbiol.* **85**, 134S–137S (1998).
- Roszak, D. B. & Colwell, R. R. Survival strategies of bacteria in the natural environment. *Microbiol. Rev.* **51**, 365–379 (1987).
- Akanda, A. S., Griffiths, J. K., Colwell, R., Islam, S. & Jutla, A. S. Warming oceans, phytoplankton, and river discharge: Implications for cholera outbreaks. *Am. J. Trop. Med. Hyg.* **85**, 303–308 (2011).

15. Vezzulli, L. *et al.* Climate influence on *Vibrio* and associated human diseases during the past half-century in the coastal North Atlantic. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E5062–E5071 (2016).
16. Hashizume, M. *et al.* The effect of rainfall on the incidence of cholera in Bangladesh. *Epidemiology* **19**, 103–110 (2008).
17. Koelle, K., Rodó, X., Pascual, M., Yunus, Md. & Mostafa, G. Refractory periods and climate forcing in cholera dynamics. *Nature* **436**, 696–700 (2005).
18. Miller, C., Drasar, B. & Feachem, R. Cholera and estuarine salinity in Calcutta and London. *Lancet* **319**(8283), 1216–1218 (1982).
19. Neogi, S. B. *et al.* Environmental and hydroclimatic factors influencing *Vibrio* populations in the estuarine zone of the Bengal delta. *Environ. Monit. Assess.* **190**(10), 1–16 (2018).
20. de Magny, G. C. *et al.* Environmental signatures associated with cholera epidemics. *Proc. Natl. Acad. Sci.* **105**, 17676–17681 (2008).
21. Lipp, E. K., Huq, A. & Colwell, R. R. Effects of global climate on infectious disease: The cholera model. *CMR* **15**, 757–770 (2002).
22. Khan, R. *et al.* Assessment of risk of cholera in Haiti following Hurricane Matthew. *Am. J. Trop. Med. Hyg.* **97**, 896–903 (2017).
23. Khan, R. *et al.* Long-range river discharge forecasting using the gravity recovery and climate experiment. *J. Water Resour. Plann. Manage.* **145**, 06019005 (2019).
24. Huq, A. *et al.* Environmental factors influencing epidemic cholera. *Am. J. Trop. Med. Hyg.* **89**, 597–607 (2013).
25. Clemens, J. D., Nair, G. B., Ahmed, T., Qadri, F. & Holmgren, J. Cholera. *The Lancet* **390**, 1539–1549 (2017).
26. Jutla, A. S., Akanda, A. S. & Islam, S. Tracking cholera in coastal regions using satellite observations: Tracking cholera in coastal regions using satellite observations. *JAWRA J. Am. Water Resour. Assoc.* **46**, 651–662 (2010).
27. Ali, M., Nelson, A. R., Lopez, A. L. & Sack, D. A. Updated global burden of cholera in endemic countries. *PLoS Negl. Trop. Dis.* **9**, e0003832 (2015).
28. Brumfield, K. D. *et al.* Environmental parameters associated with incidence and transmission of pathogenic *Vibrio* spp. *Environ. Microbiol.* <https://doi.org/10.1111/1462-2920.15716> (2021).
29. Sharp, J. & Salaam-Blyther, T. *Yemen: Cholera Outbreak*. 3 <https://crsreports.congress.gov/product/details?prodcode=IN10729> (2017).
30. WHO. Weekly epidemiological record (Vol. 93). at <https://apps.who.int/iris/bitstream/handle/10665/274654/WER9338.pdf?ua=1> (2018).
31. Lyons, K. Yemen's cholera outbreak now the worst in history as millionth case looms. *The Guardian* (2017).
32. Usmani, M. *et al.* A review of the environmental trigger and transmission components for prediction of cholera. *TropicalMed* **6**, 147 (2021).
33. Islam, M. S., Alam, M. J. & Neogi, P. K. B. Seasonality and toxigenicity of *Vibrio cholerae* non-01 isolated from different components of pond ecosystems of Dhaka City Bangladesh. *World J. Microbiol. Biotechnol.* **8**, 160–163 (1992).
34. Alam, M. *et al.* Toxigenic *Vibrio cholerae* in the aquatic environment of mathbaria Bangladesh. *AEM* **72**, 2849–2855 (2006).
35. Alam, M. *et al.* Viable but nonculturable *Vibrio cholerae* O1 in biofilms in the aquatic environment and their role in cholera transmission. *Proc. Natl. Acad. Sci.* **104**, 17801–17806 (2007).
36. Huq, A. *et al.* Critical factors influencing the occurrence of *Vibrio cholerae* in the environment of Bangladesh. *AEM* **71**, 4645–4654 (2005).
37. Cash, B. A., Rodó, X. & Kinter, J. L. Links between tropical pacific SST and cholera incidence in Bangladesh: Role of the eastern and central tropical pacific. *J. Clim.* **21**, 4647–4663 (2008).
38. Usmani, M. *et al.* Predictive Intelligence for Cholera in Ukraine? *GeoHealth* <https://doi.org/10.1029/2022GH000681> (2022).
39. Jutla, A., Aldaach, H., Akanda, A. S., Huq, A. & Colwell, R. R. Satellite based assessment of hydroclimatic conditions related to cholera in Zimbabwe. *PLOS-One* <https://doi.org/10.1371/journal.pone.0137828> (2015).
40. Khan, R. *et al.* Evaluation of risk of cholera after a natural disaster: Lessons learned from the 2015 Nepal Earthquake. *J. Water Resour. Plann. Manage.* **144**, 04018044 (2018).
41. Legros, D. Global cholera epidemiology: Opportunities to Reduce the burden of cholera by 2030. *J. Infect. Dis.* **218**, S137–S140 (2018).
42. Mosley, J. F., Smith, L. L., Brantley, P., Locke, D. & Como, M. Vaxchora: The first FDA-approved cholera vaccination in the United States. *P T* **42**, 638–640 (2017).
43. Huq, A. *et al.* A simple filtration method to remove plankton-associated *Vibrio cholerae* in raw water supplies in developing countries. *Appl. Environ. Microbiol.* **62**, 2508–2512 (1996).
44. Colwell, R. R. *et al.* Reduction of cholera in Bangladeshi villages by simple filtration. *Proc. Natl. Acad. Sci.* **100**, 1051–1055 (2003).
45. Taylor, D. L., Kahawita, T. M., Cairncross, S. & Ensink, J. H. J. The impact of water, sanitation and hygiene interventions to control cholera: A systematic review. *PLoS ONE* **10**, e0135676 (2015).
46. Fedak, K. M., Bernal, A., Capshaw, Z. A. & Gross, S. Applying the Bradford Hill criteria in the 21st century: How data integration has changed causal inference in molecular epidemiology. *Emerg. Themes Epidemiol.* **12**, 14 (2015).
47. Hill, A. B. The environment and disease: association or causation?. *Proc. R. Soc. Med.* **58**, 295–300 (1965).
48. Lucas, R. M. & McMichael, A. J. Association or causation: evaluating links between “environment and disease”. *Bull. World Health Organ.* **83**, 792–795 (2005).
49. Pascual, M. Cholera dynamics and El Niño-southern oscillation. *Science* **289**, 1766–1769 (2000).
50. Akanda, A. S., Jutla, A. S. & Islam, S. Dual peak cholera transmission in Bengal Delta: A hydroclimatological explanation. *Geophys. Res. Lett.* **36**, L19401 (2009).
51. Huq, A. *et al.* Ecological relationships between *Vibrio cholerae* and planktonic crustacean copepods. *Appl. Environ. Microbiol.* **45**, 275–283 (1983).
52. Cash, R. A. *et al.* Response of man to infection with *Vibrio cholerae*. I. Clinical, serologic, and bacteriologic responses to a known inoculum. *J. Infect. Dis.* **129**, 45–52 (1974).
53. Colwell, R. R. *et al.* Viable but non-culturable *Vibrio cholerae* O1 revert to a cultivable state in the human intestine. *World J. Microbiol. Biotechnol.* **12**, 28–31 (1996).
54. Alam, M. *et al.* Clonal transmission, dual peak, and off-season cholera in Bangladesh. *Infect. Ecol. Epidemiol.* **1**, 7273. <https://doi.org/10.3402/iee.v1i0.7273> (2011).
55. Rinaldo, A. *et al.* Reassessment of the 2010–2011 Haiti cholera outbreak and rainfall-driven multiseason projections. *Proc. Natl. Acad. Sci.* **109**, 6602–6607 (2012).
56. Louis, V. R. *et al.* Predictability of *Vibrio cholerae* in Chesapeake Bay. *AEM* **69**, 2773–2785 (2003).
57. Lukinmaa, S. *et al.* Territorial waters of the Baltic Sea as a source of infections caused by *Vibrio cholerae* non-O1, non-O139: report of 3 hospitalized cases. *Diagn. Microbiol. Infect. Dis.* **54**, 1–6 (2006).
58. Andersson, Y. & Ekdahl, K. Wound infections due to *Vibrio cholerae* in Sweden after swimming in the Baltic Sea, summer 2006. *Weekly Releases (1997–2007)* **11**, 3013 (2006).
59. Vezzulli, L. *et al.* Long-term effects of ocean warming on the prokaryotic community: evidence from the vibrios. *ISME J.* **6**, 21–30 (2012).
60. de Magny, G. C., Cazelles, B. & Guégan, J.-F. Cholera threat to humans in Ghana is influenced by both global and regional climatic variability. *EcoHealth* **3**, 223–231 (2007).
61. Ngwa, M. C. *et al.* Cholera in Cameroon, 2000–2012: Spatial and temporal analysis at the operational (Health District) and sub climate levels. *PLoS Negl. Trop. Dis.* **10**, e0005105 (2016).

62. Asadgol, Z. *et al.* How climate change can affect cholera incidence and prevalence? A systematic review. *Environ. Sci. Pollut. Res.* **27**, 34906–34926 (2020).
63. Martinez, R. M., Megli, C. J. & Taylor, R. K. Growth and Laboratory Maintenance of *Vibrio cholerae*. *Curr. Protoc. Microbiol.* **17**, 6A11–6A17 (2010).
64. Hood, M. A. & Winter, P. A. Attachment of *Vibrio cholerae* under various environmental conditions and to selected substrates. *FEMS Microbiol. Ecol.* **22**, 215–223 (2006).
65. Paz, S. Impact of Temperature Variability on Cholera Incidence in Southeastern Africa, 1971–2006. *EcoHealth* **6**, 340–345 (2009).
66. Sasaki, S., Suzuki, H., Fujino, Y., Kimura, Y. & Cheelo, M. Impact of drainage networks on cholera outbreaks in Lusaka Zambia. *Am. J. Public Health* **99**, 1982–1987 (2009).
67. Colwell, R. R., Huq, A., Chowdhury, M. A. R., Xu, B. & Brayton, P. R. Serogroup conversion of *Vibrio cholerae*. *Can. J. Microbiol.* **41**, 946–950 (1995).
68. Montilla, R., Chowdhury, M. A. R., Huq, A., Xu, B. & Colwell, R. R. Serogroup conversion of *Vibrio cholerae* non-O1 to *Vibrio cholerae* O1: effect of growth state of cells, temperature, and salinity. *Can. J. Microbiol.* **42**, 87–93 (1996).
69. Mishra, A. *et al.* Amplified fragment length polymorphism of clinical and environmental *Vibrio cholerae* from a freshwater environment in a cholera-endemic area India. *BMC Infect. Dis.* **11**, 249 (2011).
70. Grim, C. J. *et al.* RNA colony blot hybridization method for enumeration of culturable *Vibrio cholerae* and *Vibrio mimicus* bacteria. *Appl. Environ. Microbiol.* **75**, 5439–5444 (2009).
71. Nasr-Azadani, F. *et al.* Hydroclimatic sustainability assessment of changing climate on cholera in the Ganges-Brahmaputra basin. *Adv. Water Resour.* <https://doi.org/10.1016/j.advwatres.2016.11.018> (2016).
72. Alam, M. *et al.* Effect of transport at ambient temperature on detection and isolation of *Vibrio cholerae* from environmental samples. *Appl. Environ. Microbiol.* **72**, 2185–2190 (2006).
73. Jutla, A., Khan, R. & Colwell, R. Natural disasters and cholera outbreaks: Current understanding and future outlook. *Curr. Environ. Health Rpt* **4**, 99–107 (2017).
74. ACAPS. *Yemen Complex Crisis*. <https://www.acaps.org/country/yemen/crisis/complex-crisis> (2020).
75. Grad, Y. H., Miller, J. C. & Lipsitch, M. Cholera modeling: Challenges to quantitative analysis and predicting the impact of interventions. *Epidemiology* **23**, 523–530 (2012).
76. Huq, A., West, P. A., Small, E. B., Huq, M. I. & Colwell, R. R. Influence of water temperature, salinity, and pH on survival and growth of toxigenic *Vibrio cholerae* serovar O1 associated with live copepods in laboratory microcosms. *Appl. Environ. Microbiol.* **48**, 420–424 (1984).
77. Skorupski, K. & Taylor, R. K. Control of the ToxR virulence regulon in *Vibrio cholerae* by environmental stimuli. *Mol. Microbiol.* **25**, 1003–1009 (1997).
78. Schuhmacher, D. A. & Klose, K. E. Environmental signals modulate ToxT-dependent virulence factor expression in *Vibrio cholerae*. *J. Bacteriol.* **181**, 1508–1514 (1999).
79. Singleton, F. L., Attwell, R. W., Jangi, M. S. & Colwell, R. R. Influence of salinity and organic nutrient concentration on survival and growth of *Vibrio cholerae* in aquatic microcosms. *Appl. Environ. Microbiol.* **43**, 1080–1085 (1982).
80. Ravel, J., Knight, I. T., Monahan, C. E., Hill, R. T. & Colwell, R. R. Temperature-induced recovery of *Vibrio cholerae* from the viable but nonculturable state: growth or resuscitation?. *Microbiology* **141**, 377–383 (1995).
81. Eiler, A., Gonzalez-Rey, C., Allen, S. & Bertilsson, S. Growth response of *Vibrio cholerae* and other *Vibrio* spp. to cyanobacterial dissolved organic matter and temperature in brackish water: Cyanobacterial DOM, temperature and *Vibrio* growth. *FEMS Microbiol. Ecol.* **60**, 411–418 (2007).
82. *The Biology of Vibrios*. (ASM Press, 2006). <https://doi.org/10.1128/9781555815714>.
83. Stauder, M., Vezzulli, L., Pezzati, E., Repetto, B. & Pruzzo, C. Temperature affects *Vibrio cholerae* O1 El Tor persistence in the aquatic environment via an enhanced expression of GbpA and MSHA adhesins: Temperature affects *V. cholerae* in the aquatic environment. *Environ. Microbiol. Rep.* **2**, 140–144 (2010).
84. McCarthy, S. A. Effects of temperature and salinity on survival of toxigenic *Vibrio cholerae* O1 in seawater. *Microb. Ecol.* **31**(2), 167–175 (1996).

Acknowledgments

The research team acknowledges technical and resource research support from Tim Sumner and Katherine West of the United Kingdom Foreign, Commonwealth and Development Office. AJ and RC acknowledge research support through a grant from UK Foreign, Commonwealth and Development Office and US National Aeronautics and Space Administration (80NSSC20K0814).

Author contributions

M.U. and K.B. did the literature search. M.U., R.C., and A.J. did the study design. M.U., K.B., B.M., F.M., and A.J. analyzed the data. M.U. and A.J. wrote the original draft. All the authors contributed to data interpretation, review and editing of the manuscript. M.U., A.J., F.M. and H.T. accessed and verified the data. All authors had final responsibility for the decision to submit the publication.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-22946-y>.

Correspondence and requests for materials should be addressed to A.J.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023