Cholera in Africa: Rising Deaths, Shrinking US Aid—Technical note on methodology

Markus Goldstein, Eeshani Kandpal, Charles Kenny, Juan Menendez, Brian Webster

Initially, we compiled all of the publicly available Africa CDC Epidemic Intelligence Reports, which date back through January 2023. These are published weekly, and include cumulative totals of cases and deaths for a variety of disease, presented both at the continental and country level. We developed a scraping tool in Python that compiled an Excel file with the cumulative continental totals for 16 diseases each week that data was available. We used Stata 19 to analyze this dataset. Sorting by disease and date, we calculated the number of cases and deaths that were reported each week by subtracting the previous week from the reported total. Using this data, we identified that cholera deaths were higher in 2025 than at roughly the same time in 2024 and 2023.

Having identified cholera as the investigation focus, we compiled a dataset in Excel of monthly case and death totals at the country level, starting in January 2023 and extending through June 2025. As the country-level data did not lend itself to easy scraping, this data was compiled manually and then checked twice for accuracy. As case and death totals were published cumulatively, the data for any given month was drawn from the first report published in the following month. For example, the data for February 2023 was drawn from the report published on March 4, 2023. The data for September 2024 was drawn from the October 6, 2024 report.

This dataset of cumulative case and death totals for each country recorded on a monthly basis was then loaded into Stata 19. For each country, the actual number of cholera cases and deaths in each month was calculated by subtracting the previous month's cumulative total from the recorded cumulative total. For example, the cumulative case total in Mozambique in March 2025 was 2,324, and in April 2025, 3,028. We calculated that 704 cholera cases had occurred in Mozambique in April 2025. If a country's first data point occurred mid-way through the year, it is assumed that all cases and deaths reported occurred in that month, and that the figure does not include a backlog of data from the rest of the year. Death rates were calculated at the country/month level by dividing the number of deaths in a country each month by the number of cases. Continental totals were calculated by summing all of the country-level case and death data for each month (30 in all). Monthly continental death rates were then calculated by dividing the total number of deaths by total number of cases.

The data demonstrated that 93 percent of the cholera cases reported in 2025 were from four countries; Angola, the Democratic Republic of Congo, Sudan, and South Sudan. In the case of Angola, no cholera activity was reported until January of 2025, indicating a new outbreak.

Aid disbursement data

We downloaded award-level data from usaspending.gov on September 6, 2025, encompassing fiscal years 2023–2025 and three funding and/or awarding agencies: the Agency for International Development, Department of State, and Centers for Disease Control and Prevention. The data query covered 10 primary places of performance, four that were mentioned above and six that showed evidence of spikes in cholera during the 2023–2024 period: Angola, Comoros, DRC, Malawi, Mozambique, Sudan, South Sudan, Tanzania, Zambia, and Zimbabwe. In total, we captured 5,548 entries, including information on project start dates, project end dates, total obligations, total cumulative outlays, and project descriptions.

The data were combined with a list of aid projects compiled by Charles Kenny that tracked the status of individual aid projects through August 20, 2025.

To identify projects related to WASH activities or cholera, we performed a text analysis on the project description fields. We constructed keyword dictionaries for several disease areas (including cholera, malaria, measles, and mpox), as well as an extensive list of WASH-related terms (e.g., “water supply,” “latrine,” “hygiene promotion,” “sewage,” etc.). Using R's *stringr* function with regular expressions and case-insensitive matching, we flagged each award description that included one or more relevant keywords. We then labeled each project according to its focus (e.g., "cholera," "WASH," etc.). A total of 346 projects were initially identified as being related to cholera or WASH activities.

These 346 projects were classified by two award types: “assistance” and “contracts.” A manual audit of the project descriptions revealed that all projects classified as “assistance” were related to aid delivery of some type. However, a high share (85 percent or higher) of “contracts” were for activities unrelated to WASH or cholera projects. Examples include maintenance of embassy facilities and bulk bottled water purchases for USAID staff. Based on this review, only projects listed as “assistance” were included in the analysis.

The final dataset included 219 projects, all of which included key words indicating a WASH project, and two of which specifically mentioning cholera. Of these 219 projects, 22 were identified as being cancelled, 49 as active, and 148 were unspecified status, which we assumed to be active in the absence of evidence to the contrary.

This dataset was then analyzed in Stata 19. Because usaspending.gov only provided consistent data on cumulative outlays and total obligations, as opposed to transaction level data, we estimated monthly spending and obligations based on these totals and the listed

start and end dates of projects. Using the start and end dates, we calculate the total number of days a project was recorded as lasting. We then divided the total obligations by the number of days to find an average daily obligation for each project. We repeated this operation for total outlays, using the date that the data was downloaded as a cutoff. If a contract was identified as cancelled, January 20 2025 was assumed to be the cutoff date as the executive order pausing all foreign assistance spending was signed on that day. Once daily averages for obligations and outlays were calculated, the total number of days each project was active in each month was determined, and outlays and remaining obligations were then estimated based on the average daily outlay and number of days active in each month. If a project was cancelled, daily averages of remaining obligations extending from January 20, 2025 until the end of the project were used to estimate the amount of cancelled spending.

For example, Award ID (fain) 720BHA24IO00201 was recorded as starting October 1, 2024 and ending September 30, 2025—a total of 364 days. Under the award, $6,350,000 were obligated, with a total of $3,744,039 being outlaid as of September 6, 2025 (340 days after the start date). We estimated that on average, $11,011.88 was disbursed under this award daily ($3,744,039 / 340) and that since this award was active every day in January 2025, $341,368 were disbursed in that month ($11,011.88 x 31).

Alternatively, 720BHA22GR00204 extended from July 1, 2022 until December 30 2025 with $18,374,842 in obligations, but was identified as being cancelled. As of September 6, 2025, $14,253,384 were listed as being outlaid in total. It was assumed that the project was active between July 1, 2022 and January 20, 2025—a total of 934 days, and that on average $15,260 were disbursed each day. In January 2025, we estimate that $289,951 were outlaid (15,260 x 19). We assumed that had the contract not been cancelled, an additional $4,121,458 would have been outlaid (18,374,842 – 14,253,384) between January 20, 2025 and December 30, 2025 (344 days). On that basis, we estimated that $371,410 in aid was cancelled for March 2025 ((4,121,458 / 344) x 31).

Using these estimates, the total amount of dollars outlaid and aid cancelled for each month from January 2023 through June 2025 was totaled. These estimates were then combined with the country month data on cholera cases and deaths. Obligations both active and cancelled were then calculated for the remainder of 2025 as well as 2026.