METHODS SUPPLEMENT

Data collection and preparation. The monthly number of antenatal care visits and health facility deliveries from January 2016 through December 2020 were collected from each site. For all countries except Mexico, the data came from District Health Information Software 2 (DHIS2), which contains monthly counts for chosen indicators at the facility-level. For Mexico, the data came from electronic medical record systems at the Partners In Health-supported facilities and was collapsed into monthly counts for analysis. Outliers identified through Tukey's rule (larger than 1.5 times the interquartile range) were returned to the country's monitoring and evaluation officer for correction or removal.

Notably, we did not include data on antenatal care visits for Mexico as there was not reliable pre-pandemic data on the use of antenatal care services in primary care clinics because the current electronic medical record system was only introduced in mid-2019. On the other hand, facility-based deliveries occur at the community hospital, where routine data collection began in August 2016 when the Compañeros En Salud patient-centered delivery care model was implemented at the Ángel Albino Corzo Basic Community Hospital. For Sierra Leone PIH-supported health facilities, DHIS2 data were not available prior to 2018.

Baseline model. We fit a time series model for the monthly (1) number of facility-based deliveries (2) number of first antenatal care visits at each PIH-supported health facility. We fit a generalized liner model with negative binomial distribution to estimate the monthly counts:

$$log(E[Y|\ year, t]) = \beta_0 + \beta_1 year_t + \sum_{k=1}^{K} \beta_{2k} cos\left(\frac{2\pi kt}{12}\right) + \beta_{3k} sin(\frac{2\pi kt}{12}) \text{ (Eq. 1)}$$

where *Y* indicates monthly indicator count, *t* indicates the cumulative month number, K indicates the number of harmonic functions to include (we take $K = 3$). The *year* term captures long-term annual trend while the harmonic terms capture seasonality. This mean model was chosen to allow smoothing without imposing strong assumptions on the seasonal behavior. We chose to use a negative binomial distribution (instead of Poisson) to account for overdispersion.

We performed several diagnostic procedures to ensure that the facility-level regression models did not exhibit autocorrelation in the residuals. For the residuals from facility, we conducted the Breusch-Godfrey test for serial correlation and generated plots of residuals by time, autocorrelation functions, and partial autocorrelation functions. There were no signs of residual autocorrelation at any facility or indicator.

Computing predicted counts and deviations. The facility-level models were aggregated using the parametric bootstrap procedure to compute country-level predicted counts during the COVID-19 period (March 2020 through December 2020) with corresponding 95% prediction intervals. In this procedure, we also computed monthly deviations (predicted - observed count) and cumulative deviation over the entire 10-month period with corresponding 95% prediction intervals. The procedure is as follows:

Step 1. For each bootstrap iteration r, values (β^r) are drawn from a multivariate normal distribution with mean $\hat{\beta}$ and variance $\widehat{\Sigma_\beta}$ resulting from the model fit in Eq. 1.

Step 2. Generate predicted counts for all months, $Y_t^r \sim NegBin(\exp(X\beta^r), \frac{\exp(X\beta^r)}{\hat{\phi}-1})$

Step 2a. Generate monthly deviations by $D_t = Y_t^r - Y_t$ where Y_t is the observed count for month *t*

Step 2b. Generate cumulative deviation over a pre-specified period T, $CD_t = \sum_{t \in T} D_t$

Step 3. Repeat Steps 1-2 R=1000 times.

Step 4. Take the median count for the predicted count (or deviation or cumulative deviation) for each month (or ore-specified period) and the 2.5th and 97.5th percentiles as the 95% prediction interval.

This bootstrap procedure is described in Fulcher *et al.* (2021) and has been used in related settings by Weinberger *et al.* (2020) and Lauer *et al.* (2018).

For each country and maternal health indicator, we report: (1) time series plots for the entire study period (January 2016 – December 2020), (2) deviation plots during the COVID-19 period, and (3) cumulative deviation estimates with 95% prediction intervals for the COVID-19 period. All data cleaning and analysis was done using R V3.6.0. Example code can be found on the public GitHub repository: https://github.com/isabelfulcher/global_covid19_response.

Interpretation of the results. To better understand the results observed in the analysis, including the local impact of the pandemic and the strategies implemented by the sites in response, we convened members of the Cross-site COVID-19 Syndromic Surveillance Working Group and invited representatives from each country that work closely with maternal health services. During this meeting, we presented the results for each country and discussed potential reasons for these drops. In addition, we communicated directly with representatives of maternal health services in each country to obtain specific information on strategies to prevent and/or mitigate the effects of the pandemic on the utilization of maternal health services implemented during the COVID-19 pandemic. National information on governmental measures and pandemic evolution was collected through Our World in Data (https://ourworldindata.org/) on October 25, 2021.

Limitations. This time series modeling approach is designed to detect significant deviations based on the baseline models. As such, misspecification of the baseline models will hinder the ability to accurately predict counts “in the absence of COVID-19” thereby precluding valid identification of deviations during the pandemic. We took several steps to ensure the validity of the baseline models. First, we used as much prior data to build the baseline model, which was four years (48 months) for most countries and indicators. Second, we met with monitoring and evaluation site leads to identify outliers and missing data for each indicator and facility. Third, as described above, we performed several diagnostic procedures to ensure that the baseline models fit the data well.

References

Fulcher IR, Boley EJ, Gopaluni A, et al. Cross-site COVID-19 Syndromic Surveillance Working Group. Syndromic surveillance using monthly aggregate health systems information data: methods with application to COVID-19 in Liberia. International journal of epidemiology. 2021 Aug;50(4):1091-102.

Weinberger DM, Chen J, Cohen T, et al. Estimation of excess deaths associated with the COVID-19 pandemic in the United States, March to May 2020. JAMA Internal Medicine. 2020 Oct 1;180(10):1336-44.

Lauer SA, Sakrejda K, Ray EL, et al. Prospective forecasts of annual dengue hemorrhagic fever incidence in Thailand, 2010–2014. Proceedings of the National Academy of Sciences. 2018 Mar 6;115(10):E2175-82.