Under-Five Mortality during the Covid-19 Outbreak: Evidence from Four Demographic Surveillance Systems in Low-Income Countries

Julio Romero Prieto, Andrea Verhulst, Alam Nurul, Valérie Delaunay, Momodou Jasseh, Sammy Khagayi, Gilles Pison, Michel Guillot, Georges Reniers

Under-five mortality during the COVID-19 outbreak: evidence from four demographic surveillance systems in low-income countries

Julio E Romero Prieto, London School of Hygiene and Tropical Medicine
Andrea Verhulst, University of Pennsylvania
Nurul Alam, International Centre for Diarrhoeal Disease Research, Bangladesh
Valérie Delaunay, Institut de Recherche Pour le Développement, Dakar, Senegal
Momodou Jasseh, Medical Research Council - The Gambia Unit at London School of Hygiene and Tropical Medicine
Sammy Khagayi, Kenya Medical Research Institute - Centre for Global Health Research (KEMRI-CGHR)
Gilles Pison, French Museum of Natural History - French Institute for Demographic Studies
Michel Guillot, University of Pennsylvania - French Institute for Demographic Studies
Georges Reniers, London School of Hygiene and Tropical Medicine

1. Introduction and background

The under-five mortality rate halved during the last 30 years [1], but the progress could be jeopardized by the epidemic of COVID-19. Although, data from high-income countries suggest that young populations are less susceptible to the infection, the percentage of under-five deaths due to COVID-19 is very low, and overall mortality rates did not increase during the lockdown [2, 3, 4, 5]; simulation models anticipate a large increase in maternal and under-five mortality in low- and middle-income countries, due to the indirect effects of the epidemic of COVID-19 [6, 7, 8, 9]. Indirect effects are informed by assumptions that build on the experience of the Ebola outbreak of 2014 in West Africa and the epidemic of severe acute respiratory syndrome of 2003 in Taiwan.

The indirect effects may be due to the interruption of health services, the postponement of diagnosis and treatments, the discontinuity of health interventions and vaccination programs, undernutrition, economic hardship, and lack of social support due to the mortality of parents. Some evidence indeed seems to point at a reduction in the utilization of maternal, neonatal, and child health services; and a reduction in birth related emergencies and the number of births in hospitals [10, 11, 12]. These reductions have been related to travel restrictions and the fear of transmission in hospitals. However, there is still little evidence demonstrating the negative impact of COVID-19 on under-five mortality in the context of less- and middle- income countries.

This paper investigates the effect of the epidemic of COVID-19 on under-five mortality in four Health and Demographic Surveillance Systems (HDDS): Siaya-HDSS in Kenya [13], Basse-HDSS in the Gambia[1], Niakhar-HDSS in Senegal [14], and Matlab-HDSS in Bangladesh [15]. Inasmuch as the number of deaths due to COVID-19 is dependent on testing coverage, variations in case definitions, or the performance of health facility-based surveillance, the effect of COVID-19 on under-five mortality is estimated to be indirect, through a measure of excess mortality [16]. We calculate the excess of mortality as the difference between the probabilities of dying in a calendar month affected by the pandemic and a counterfactual estimate that only depends on the secular trend in mortality of each site before the outbreak of COVID-19. We approach to this aim by forecasting the life tables of under-five mortality and quantifying the effects of the epidemic of COVID-19 on the most relevant indicators of mortality from zero to five years (e.g., the neonatal mortality, the post-neonatal mortality, the infant mortality rate, the child mortality rate, and the under-five mortality rate). Preliminary results did not identify any evidence of excess mortality that can be attributed to the COVID-19 outbreak, and these findings would contradict the predictions of earlier simulation studies.

2. Data

This paper uses data collected in four sites of demographic surveillance: Siaya-HDSS in Kenya (2008-2020), Basse-HDSS in the Gambia (2008-2021), Niakhar-HDSS in Senegal (1997-2020), and Matlab-HDSS in Bangladesh (1978-2021). The Siaya-HDSS has been managed through a collaboration between the Kenya Medical Research Institute (KEMRI) and the Centers for Disease Control (CDC). The total population of the site is approximately 260,000 residents living in 393 villages and the population below 5 years is estimated to be 31,000 individuals. The Basse-HDSS has been founded by the Medical Research Council and, as of 2019, covers a population of approximately 190,000 residents living in 219 settlements. The population below the age of 5 years is estimated to be 30,000 individuals. The Niakhar-HDSS was established by the Institut de Recherche pour le Développement of Senegal and the area of surveillance has an approximated population of 45,000 residents living in 30 villages. The Matlab-HDSS

[1] This site would be part of the analysis if the PIs of Basse-HDSS accept to collaborate in this project.

has been in operation since 1966 and is administered by the International Centre for Diarrhoeal Disease Research. The Matlab HDSS covers a population of approximately 235,000 residents.

Individual records of these sources include the exact date of all life events (i.e., births, deaths, and migration). Information has been collected prospectively, visiting each household three or four times per year. This frequency ensures a short time of retrospection that improves the quality of the reported dates. Compared to other data sources for estimating mortality, HDSS have the advantage that events and exposure are calculated from the same source. In principle, the relatively short periods of retrospection may also contain reporting bias.

Synthetic cohort life tables were constructed for estimating under-five mortality by detailed age groups. First, mortality rates were estimated dividing the number of deaths ${}_nd_x$ by the number of person-years lived in the same interval of age and time ${}_nL_x$. The following exact-age cut-offs were used: 7, 14, 21, and 28 days; 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 15, 18, and 21 months; and 2, 3, 4, and 5 years. Second, cumulative probabilities of dying $q(x)$, were calculated from age-specific mortality rates under the assumption of constant mortality hazards within the age interval, which is a tenable assumption considering that age intervals are small.

Inasmuch as migration is also a source of attrition in these populations and out-migrants can return to the areas of demographic surveillance after some time abroad, mortality and migration rates were calculated in a simultaneous process with two exits (i.e., mortality and out-migration) and two possible entries (i.e., birth and in-migration). In this regard, permanent out-migrants stop contributing events and exposure after leaving the surveillance site (i.e., right censoring); and returning-migrants do not add exposure while they were living outside of the surveillance area (i.e., interval censoring). Therefore, the number of deaths only include those occurring within the borders of demographic surveillance, as well as the mortality experience of the in-migrant population that moved into the surveillance area sometime after birth.

3. Methods

Excess under-five mortality during the COVID-19 epidemic

Excess mortality was estimated as the difference between the probabilities of dying observed in one calendar month affected by the pandemic and a counterfactual that is depending on the secular trends in mortality of each site before the COVID-19 outbreak. Compared to other approaches that can be used for estimating excess mortality, forecasting the probabilities of dying (or the mortality rates) have some clear advantages. First, before the pandemic there was a rapid decline in the under-five mortality of these populations, thus the simple average of the last five years tends is likely to overestimate the expected mortality. Second, compared to the number of deaths, mortality rates and probabilities of dying are less sensitive to changes in the exposure. This is important because data collection was interrupted during lockdown periods and not all households are immediately revisited whenever fieldwork resumed. This attrition –or temporary loss of follow-up–, will depress the observed number of deaths, but mortality rates and probabilities of dying will not necessarily be affected because these metrics also account for the changes in the population at risk of dying.

Forecasting the under-five mortality

Following Lee-Carter's approach [17], life tables of under-five mortality were forecasted from the model of mortality represented by Equation 1. In this case, the natural logarithm of the cumulative probability of dying at the exact age x, and time t (in months), indicated by $q(x,t)$, is a linear combination of the age-specific coefficients $\{a_x, b_x, c_x, d_x\}$, and the value of three time-specific parameters $\{k_1(t), k_2(t), k_3(t)\}$, indicating three independent mortality indices. From this perspective, the forecasted value of the cumulative probability of dying $q(x, t+s)$, will result from a new set of time-specific parameters $\{k_1(t+s), k_2(t+s), k_3(t+s)\}$, keeping constant the age-specific coefficients of the model $\{a_x, b_x, c_x, d_x\}$.

$$\ln[q(x,t)] = a_x + b_x \cdot k_1(t) + c_x \cdot k_2(t) + d_x \cdot k_3(t) + e_{x,t} \quad (1)$$

Life tables of under-five mortality were forecasted after estimating the time parameters and the age coefficients of the model. For each HDSS, all values of the mortality model were estimated in two steps. First, the constant term a_x was estimated as the simple mean of the dependent variable $\ln[q(x)]$. A total of 22 values of a_x were estimated, considering the number of cut-off points of age. Second, the other three sets of age-specific coefficients $\{b_x, c_x, d_x\}$ and the three sets of time-specific parameters were estimated by the Singular Value Decomposition (SVD) of a matrix conformed by the deviations from the mean $\{\ln[q(x,t)] - a_x\}$, at different ages (in rows) and periods (in columns). The values of $\{b_x, c_x, d_x\}$ correspond to the first three columns of the left eigenvectors, while the values of

$\{k_1(t), k_2(t), k_3(t)\}$ are given by the first three rows of the right eigenvectors multiplied by the first three eigenvalues of the SVD.

Considering the properties of the SVD, if all 22 orthonormal eigenvectors were included in the model, the error term in Equation 1 –represented by $e_{x,t}$– would be equal to zero at all ages and periods of time. Hence, the selection of three indices of mortality and three sets of age-specific coefficients rather than one –as typically implemented–, is to maximize the fitting of the model, while allowing a reasonable increase in the complexity of the time dynamic. Finally, after the SVD, the estimated coefficients and the parameters were rescaled to make the sum of the age-specific coefficients $\{b_x, c_x, d_x\}$ equal to one, across ages; and the mean of the mortality indices $\{k_1(t), k_2(t), k_3(t)\}$ equal to zero, over time.

Time series analysis

Following a standard notation and L representing the lag operator (i.e., $k(t-s) = L^s \cdot k(t)$) [18], Equation 2 is assuming that the mortality index $k_1(t)$ is stationary in its first difference. Hence, the current value of $k_1(t)$ is equal to the value of the previous month $k_1(t-1)$, plus a drift μ_1, plus a compounded mortality shock that is following a particular autoregressive structure, which is the main difference compared to the Lee-Carter model to forecast mortality. From our perspective, any mortality shock $\epsilon_1(t)$, will prevail for two additional months –at a proportion given by the coefficients θ_{11} and θ_{12}; and any mortality shock will have an echo effect –at a proportion φ_{11}, on the mortality index of same month but the following year. The aim of this effect is to model the interannual variations of the under-five mortality.

$$k_1(t) = k_1(t-1) + \mu_1 + [1 + \theta_{11} \cdot L + \theta_{12} \cdot L^2] \cdot [1 + \varphi_{11} \cdot L^{12}] \cdot \epsilon_1(t) \quad (2)$$

$$k_2(t) = \mu_2 + [1 + \theta_{21} \cdot L + \theta_{22} \cdot L^2 + \theta_{23} \cdot L^3] \cdot \epsilon_2(t) \quad (3)$$

$$k_3(t) = \mu_3 + [1 + \theta_{31} \cdot L + \theta_{32} \cdot L^2] \cdot \epsilon_3(t) \quad (4)$$

Compared to the first index of mortality, the second and third indices follow a simpler structure, as shown in Equations 3 and 4. Both indices are stationary in levels (i.e., unconditional to the previous value) and the compounded mortality shocks do not assume any seasonal effect. The mortality shock of the second mortality index will prevail for three additional months, while the shocks of the third index will have an impact for the next two months.

Estimation, confidence intervals, and smoothing

The coefficients of Equations 2, 3, and 4 –representing the time dynamic of the under-five mortality of each site, were estimated by Maximum Likelihood, assuming that mortality shocks are independent and follow a normal distribution with a constant variance over time. Since the mortality indices $\{k_1(t), k_2(t), k_3(t)\}$ are unrelated –due to the orthonormal property of the SVD, Equations 2, 3, and 4 were estimated and forecasted independently. Only the life tables pre-dating January 2020 were used for estimation. Given the estimated values of the model and using Equation 1, life tables of under-five mortality were forecasted for the following 24 months, starting in January 2020.

Confidence intervals (CI) were computed by resampling the under-five population 500 times with replacement (bootstrapping) and estimating the model and the forecast for each resample. The aim of the bootstrapped CI is to detect significant differences between the forecasted and the observed mortality in 2020 and part of 2021, that would account for the excess of under-five mortality during the COVID-19 epidemic. The graphical analysis includes smoothed trends and confidence bounds that were calculated from penalized B-Splines [19].

4. Preliminary results

Preliminary results correspond to Siaya-HDSS. The Figure 1 shows the estimated values of the neonatal mortality (i.e., the probability of dying within the first 28 days of life, $q(28d)$) by calendar months, from January 2008 and December 2020. For convenience, all values are plotted on the 15th of the month. While the panel A shows a full time series, the panel B is a close-up to the year 2020 –mostly affected by the epidemic of COVID-19. In addition to the observed values (in black), Figure 1 shows forecasted values to the years 2020 and 2021, depending entirely on the observations before the COVID-19 outbreak. Forecasted values (in red) assume that mortality shocks $\epsilon(t)$ are equal to zero, then the neonatal mortality converges to a long-term foreseeable trajectory. The Figure 1 shows two different approaches to the CIs: first, we indicate the future uncertainty of these mortality shocks (in green); second, we calculate those CIs related to the size of the under-five population and the precision to measure the neonatal mortality (in pink and blue). In the second approach, we are simply resampling the under-five population with replacement and

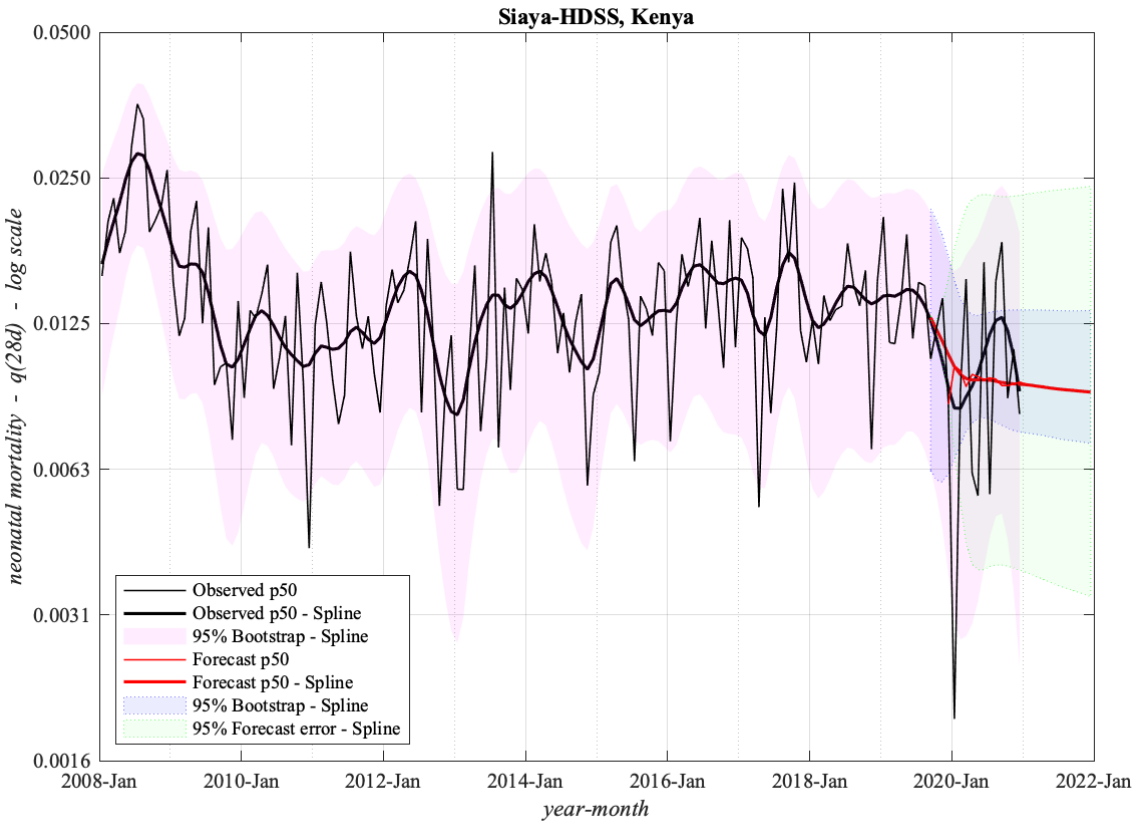
estimating and forecasting under the same assumptions (bootstrapping). Finally, the excess of mortality will result to compare the observed values in 2020 with the foreseeable trajectory of the same mortality indicator.

As shown in Figure 1 (panel B), the neonatal mortality in Siaya-HDSS had a slightly decline during the second half of 2019 and the observed value of January 2020 was very low. As a result, the observed mortality was below the counterfactual, indicating a negative excess of neonatal mortality for the beginning of 2020. During the following months, some of the observed values were above the expected value; and the Figure 1 is suggesting an excess of neonatal mortality for the most part of 2020. However, the CIs of observed and forecasted values overlap, indicating that the excess of neonatal mortality in 2020 are not necessarily significant –from the statistical point of view.

Figure 2 shows similar estimates, but analyzing the under-five mortality rate (i.e., the probability of dying within the first 60 months of life, $q(60m)$). The panel A describes an important decline of this indicator from 2008 to 2020. Compared to the neonatal mortality, the CIs of the under-five mortality are not too much wide, basically, because $q(60m)$ is greater then $q(28d)$. Hence, more power (precision) can be reached using the same number of observations. According to Figure 2, the under-five mortality has a distinctive seasonal pattern in Siaya-HDSS, peaking in the midst of the year. Panel B shows that the pattern did not repeat in 2020 and $q(60m)$ declined during the first part of the year. Although the under-five mortality increased during the second semester of 2020, the observed values were below the counterfactual, indicating a negative excess of under-five mortality. Inasmuch as the CIs of observed a forecasted values overlap, the (negative) excess of mortality during 2020 was not significant.

Figure 1: The neonatal mortality– q(28d)

A. The full period



B. The year 2020

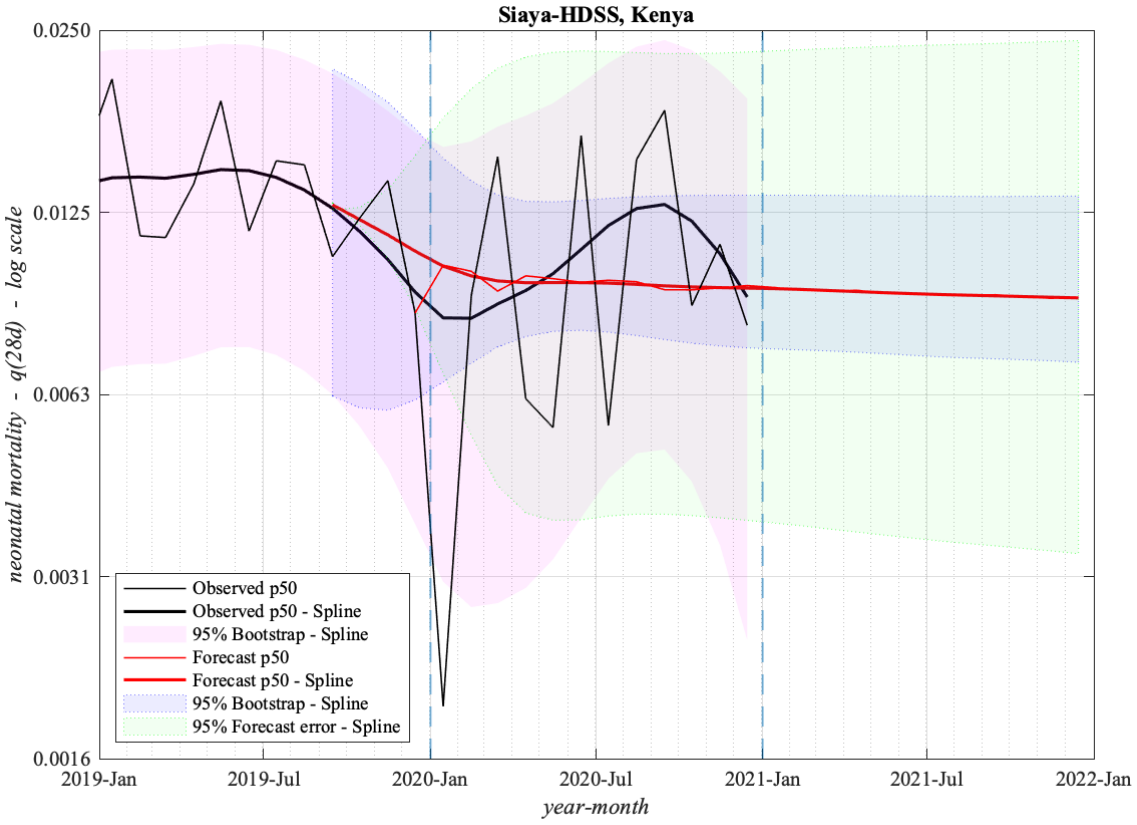
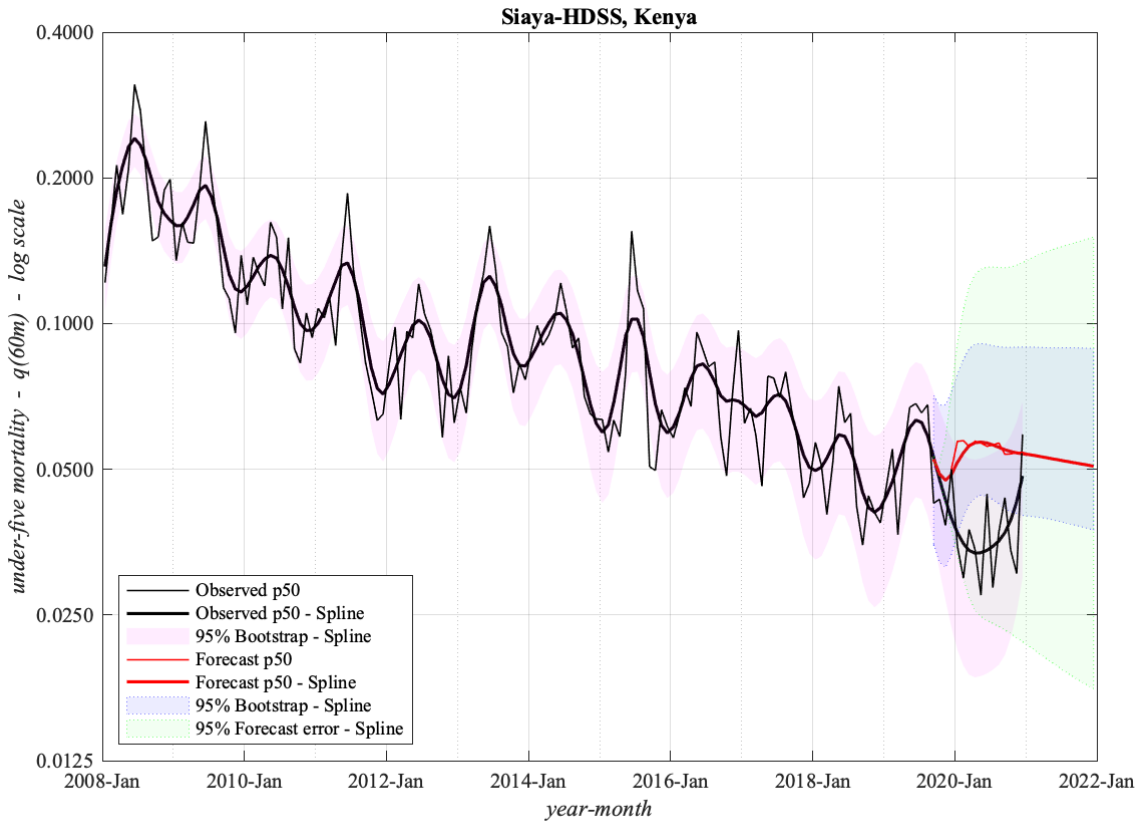
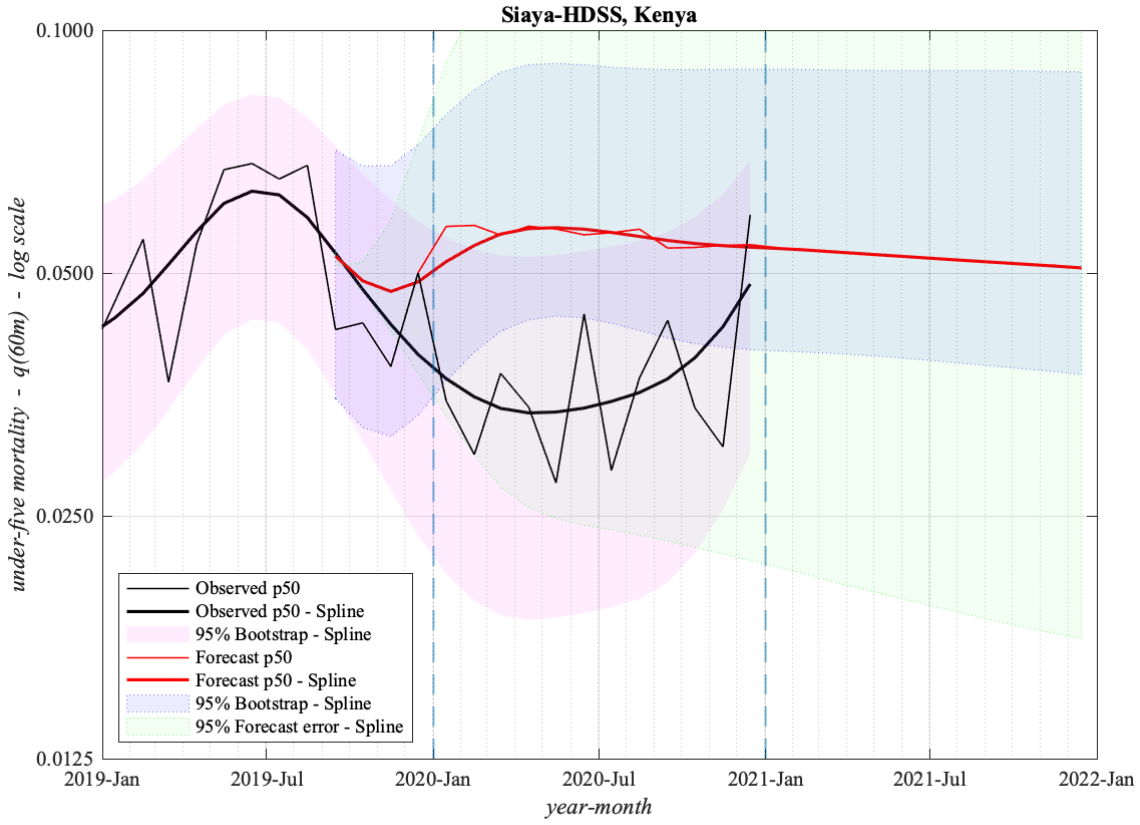


Figure 2: The under-five mortality rate – q(60m)

A. The full period



B. The year 2020



References

[1] L. Liu, S. Oza, D. Hogan, Y. Chu, J. Perin, J. Zhu, J. E. Lawn, S. Cousens, C. Mathers and R. E. Black, “Global, regional, and national causes of under-5 mortality in 2000–15: an updated systematic analysis with implications for the Sustainable Development Goals,” *Lancet,* vol. 388, p. 3027–35, 2016.

[2] N. G. Davies, P. Klepac, Y. Liu, K. Prem, M. Jit, CMMID-COVID-19-working-group and R. M. Eggo, "Age-dependent effects in the transmission and control of COVID-19 epidemics," *Nature Medicine,* vol. 26, no. Letters, p. 1205–1211, 2020.

[3] S. Bhopal, J. Bagaria and R. Bhopal, "Children's mortality from COVID-19 compared with all-deaths and other relevant causes of death: epidemiological information for decision-making by parents, teachers, clinicians and policymakers," *Public Health,* vol. 185, no. Letter to the Editor, pp. 19-20, 2020.

[4] S. S. Bhopal, J. Bagaria, B. Olabi and R. Bhopal, "Children and young people remain at low risk of COVID-19 mortality," *The Lancet Child & Adolescent Health,* vol. 5, no. Correspondence, pp. e12-e13, 2021.

[5] D. Odd, P. Fleming, S. Stoianova, T. Williams, V. Sleap, P. Blair, I. Wolfe and K. Luyt, "Child mortality in England during the COVID-19 pandemic," *Archives of Disease in Childhood,* vol. 0, pp. 1-7, 2021.

[6] T. Roberton, E. D. Carter, V. B. Chou, A. R. Stegmuller, B. D. Jackson, Y. Tam, T. Sawadogo-Lewis and N. Walker, "Early estimates of the indirect effects of the COVID-19 pandemic on maternal and child mortality in low-income and middle-income countries: a modelling study," *The Lancet Global Health,* vol. 8, pp. e901-08, 2020.

[7] W. Mao, O. Ogbuoji, D. Watkins, I. Bharali, E. Nsiah-Boateng, M. M. Diab, D. Dwomoh, D. T. Jamison, P. Kumar, K. K. McDade, J. Nonvignon, Y. Ogundeji, F.-G. Zeng, A. Zimmerman and G. Yamey, "Achieving global mortality reduction targets and universal health coverage: The impact of COVID-19," *PLoS Medicine,* vol. 18, no. 6, p. e1003675, 2021.

[8] S. Osendarp, J. Kweku, R. E. Black, D. Headey, M. Ruel, N. Scott, M. Shekar, N. Walker, A. Flory, L. Haddad, D. Laborde, A. Stegmuller, M. Thoma and R. Heidkamp, "The COVID-19 crisis will exacerbate maternal and child undernutrition and child mortality in low- and middle-income countries," *Nature Food,* vol. 2, p. 476–484, 2021.

[9] S. D. Hillis, H. J. T. Unwin, Y. Chen, L. Cluver, L. Sherr, P. S. Goldman, O. Ratmann, C. A. Donnelly, S. Bhatt, A. Villaveces, A. Butchart, G. Bachman, L. Rawlings, ... and S. Flaxman, "Global minimum estimates of children affected by COVID-19-associated orphanhood and deaths of caregivers: a modelling study," *The Lancet,* vol. 398, p. 391–402, 2021.

[10] T. Ahmed, A. E. Rahman, T. G. Amole, H. Galadanci, M. Matjila, P. Soma-Pillay, B. M. Gillespie, S. E. Arifeen and D. O. Anumba, "The effect of COVID-19 on maternal newborn and child health (MNCH) services in Bangladesh, Nigeria and South Africa: call for a contextualised pandemic response in LMICs," *International Journal for Equity in Health,* vol. 20, no. 77 Commentary, pp. 1-6, 2021.

[11] K. Ashish, R. Gurung, M. V. Kinney, A. K. Sunny, M. Moinuddin, O. Basnet, P. Paudel, P. Bhattarai, K. Subedi, M. P. Shrestha, J. E. Lawn and M. Målqvist, "Effect of the COVID-19 pandemic response on intrapartum care, stillbirth, and neonatal mortality outcomes in Nepal: a prospective observational study," *Lancet Global Health,* vol. 8, p. e1273–e1281, 2020.

[12] M. Mor, N. Kugler, E. Jauniaux, M. Betser, Y. Wiener, H. Cuckle and R. Maymon, "Impact of the COVID-19 pandemic on excess perinatal mortality and morbidity in Israel," *American Journal of Perinatology,* vol. 38, no. 4, pp. 398-402, 2021.

[13] F. O. Odhiambo, K. F. Laserson, M. Sewe, M. J. Hamel, D. R. Feikin, K. Adazu, S. Ogwang, D. Obor, N. Amek, N. Bayoh, M. Ombok, K. Lindblade, M. Desai, F. ter Kuile and P. Phillips-Howard, "Profile: the KEMRI/CDC Health and Demographic Surveillance System - Western Kenya," *International Journal of Epidemiology,* vol. 41, no. 4, pp. 977-87, 2012.

[14] V. Delaunay, L. Douillot, A. Diallo, D. Dione, J.-F. Trape, O. Medianikov, D. Raoult and C. Sokhna, "Profile: The Niakhar Health and Demographic Surveillance System," *International Journal of Epidemiology,* vol. 42, no. 4, p. 1002–1011, 2013.

[15] N. Alam, T. Ali, A. Razzaque, M. Rahman, M. Z. Haq, S. K. Saha, A. Ahmed, A. Sarder, M. M. Haider, M. Yunus, Q. Nahar and P. K. Streatfield, "Health and Demographic Surveillance System (HDSS) in Matlab, Bangladesh," *International Journal of Epidemiology,* vol. 46, no. 3, p. 809–816, 2017.

[16] D. A. Leon, V. M. Shkolnikov, L. Smeeth, P. Magnus, M. Pechholdová and C. I. Jarvis, "COVID-19: a need for real-time monitoring of weekly excess deaths," *The Lancet,* vol. 395, no. 10234 (Correspondence), p. e81, 2020.

[17] R. D. Lee and L. R. Carter, "Modeling and forecasting US mortality," *Journal of the American Statistical Association,* vol. 87, no. 419, pp. 659-671, 1992.

[18] W. Enders, Applied Econometric Time Series (Second Edition), Hoboken, NJ: J Wiley, 2004.

[19] P. H. Eilers and B. D. Marx, "Flexible smoothing with B-splines and penalties," *Statistical Science,* vol. 11, no. 2, pp. 89-121, 1996.