Supplementary Materials for

Revealing the extent of the first wave of the COVID-19 pandemic in Kenya based on serological and PCR-test data

John Ojal1,2\*†, Samuel P. C. Brand1,3,4\*† , Vincent Were5, Emelda A Okiro6,7, Ivy K Kombe1, Caroline Mburu1, Rabia Aziza3,4, Morris Ogero1, Ambrose Agweyu1, George M Warimwe1,7, Sophie Uyoga1, Ifedayo M O Adetifa1,8, J Anthony G Scott1,8, Edward Otieno1, Lynette I Ochola-Oyier1, Charles N Agoti1,9, Kadondi Kasera10, Patrick Amoth10, Mercy Mwangangi10, Rashid Aman10, Wangari Ng’ang’a11, Benjamin Tsofa1, Philip Bejon1,7, Edwine Barasa5,7, Matt. J. Keeling3#, D. James. Nokes1,3,4#.

Correspondence to: John Ojal ([JOjal@kemri-wellcome.org](mailto:JOjal@kemri-wellcome.org)) and/or Samuel Brand ([S.Brand@warwick.ac.uk](mailto:S.Brand@warwick.ac.uk)).

**This document includes:**

Materials and Methods

Supplementary Text

Figs. S1 to S5

Tables S1

Captions for Data S1 to S7

**Other Supplementary Materials for this manuscript include the following:**

Data S1 to S7

Materials and Methods

Data description

In this paper we make inferences of the penetration of SARS-CoV-2 transmission into each Kenyan county using a mechanistic transmission model. Joint posterior distributions for the parameters of the transmission model were inferred for each county using a synthesis of three data sources:

* **Kenyan Ministry of Health National linelist.** We were provided linelisting information about confirmed cases and tests performed, however the period over which data was available differed between cases and tests:
  + Confirmed cases. PCR positive swab samples by collection date, , denoting the number of positive test results collected on day *n* in county *c*. Where collection date was not available lab confirmation date was used. Negative tests were not recorded in the national linelist. We excluded positive swabs that were traced either at entry into Kenya or due to being a contact of an identified case, so as to focus on Kenyan population surveillance trends. Those confirmed cases who died were also recorded by date of death for each county, . The national linelist case data was available over the period we analyse in this paper (13th March to 30th September).
  + National combined laboratory test results. The combined reported tests, by laboratory collection date, across Kenyan laboratories, and , respectively denote the positive and negative test results by confirmation date among all Kenyan laboratories. This data covered the period 16th June to 30th September, however, there is a significant discrepancy between the first few weeks of combined laboratory linelist total daily reported tests and the nationally reported test numbers, collated by *ourworldindata.org* (*22*). We use data from the national combined laboratory linelist from 7th July to 30th September, the period with less discrepancy (Fig. S2).
  + Kemri-Wellcome Trust Research Programme (KWTRP) test result linelist**.** The majority of swab tests performed in the counties of the Coastal Province of Kenya reverted to the KWTRP testing laboratory for confirmation. and , respectively denote the positive and negative test results by collection date in coastal county *c* on day *n* that were confirmed at KWTRP. As with the national linelist we screened out tests due to contact tracing or at entry to Kenya. The KWTRP linelist data was available 13th March to 30th September. The counties covered by the KWTRP linelist were Kilifi, Kwale, Lamu, Mombasa, Taita Taveta and Tana River.
* **KWTRP** **serological surveillance programme (round 1).** Numbers of sero-positive and sero-negative blood samples collected from regional centres of the Kenyan National Blood Transfusion Service (KNBS)on day *n* originating from county *c*. Residual blood samples for serology were obtained from regular blood donors attending 4 regional KNBTS centres (Mombasa, Nairobi, Eldoret and Kisumu) in May and June. The study methodology is fully described in Uyoga et al (*11*).
* **Google mobility data.** Daily estimates of relative human mobility compared to a baseline of the same date in the previous year (2019) derived from Google mobility trends (*23*). We assumed that changes in trends in SARS-CoV-2 transmission in Kenya were due to changes in the underlying population mobility. In particular, by changing frequency of indoor congregations. Therefore, we calculated as the average change in baseline mobility over the “retail and recreation”, “grocery and pharmacy”, “transit stations”, and, “workplaces” settings (Google defined categories), and also over the week prior to day *n*, in order to average over weekend effects. Due to incomplete data, and the likely bias introduced by using a mobility estimate derived from smartphone users in predicting the mobility of semi-rural populations outside of the major urban conurbations in Kenya, we consider only three areas: Nairobi, Mombasa and the pan-Kenyan aggregate (Fig. S1).

For making inference about unknown parameters for each county (see below) we use different sources of PCR swab testing data depending on the county and date according to the following rules:

* For coastal counties the KWTRP linelist was used.
* For non-coastal counties the national case linelist was used between 13th March and 6th July, negative test results are unavailable over this period.
* For non-coastal counties the national laboratory test linelist was used between 7th July and 30th September. During this period both positive and negative test results were available.

The combined linelist data used in this paper is provided (Data S4), as well as the daily serology samples (Data S5).

Transmission model

The dynamics of transmission were assumed to follow a simple SEIR transmission model with an effective population size parameter () (*24*). The parameter accounted for the effect of population heterogeneity in lowering the proportion of the total population that must become immune for incidence to start decreasing, due to depletion of susceptibles rather than increased social distancing, compared to the prediction of a fully homogeneous model. In homogeneous SEIR models both the early exponential growth rate in incidence and the proportion of the total population that must become immune for incidence rate to start decreasing, e.g. “herd-immunity”, can be determined from the basic reproductive number and the mean durations of latency and infectiousness (*25*). For heterogeneous models of transmission, where potentially different at-risk groups are at different risk of contracting the infectious pathogen and have different infectious potential, determining from early growth in incidence aggregated over the different at-risk groups doesn’t give sufficient information to estimate the overall proportion of the population required to become immune before achieving herd immunity. This aspect of heterogeneous models of transmission has been widely investigated, for example, in the context of comparing vaccination coverage thresholds for elimination between uniform and targeted vaccination policies (*26*). In the context of the SARS-CoV-2 pandemic modelling literature, the role of population heterogeneity in lowering the herd-immunity threshold compared to the prediction of a homogeneous population transmission model has again been identified (*27*, *28*). In this study, we have taken a phenomenological approach; the effect of heterogeneity in the population was encoded in the effective population parameter , and this parameter was inferred jointly with .

The rate of infectious contacts per infectious individual, , was given as the time-varying *instantaneous* reproductive number (), that is the number of secondary cases per infected individual assuming *both* a fully susceptible population *and* the time-varying mobility rate being fixed at time *t* (*29*), rescaled by the recovery rate from infection (). We consider two models for :

1. Assuming that instantaneous reproductive number for a county was proportional to the change in the mobility rate:

|  |  |
| --- | --- |
| when *t* was in day *n*. | (1) |

The “area” of the county was Nairobi or Mombasa for those counties, otherwise the Pan-Kenyan mobility estimate was used.

1. Assuming that the instantaneous reproductive number for a county was proportional to an effective relative contact rate for the county :

|  |  |
| --- | --- |
| when *t* was in day *n*. | (2) |

is fixed to be identical to the corresponding for the first 30 days (to avoid joint identification problems with parameter), but is subsequently fitted to epidemic data using maximum a posteriori estimation (see below).

For either model the baseline reproductive number was inferred for each county.

The model dynamics are given as differential equations for susceptible (), latently infected (), actively infectious (), recovered/immune (), and, cumulative infections (),

|  |  |
| --- | --- |
| ,  ,  ,  , | (3) |

All variables and parameters are described in table S1. The transmission model was initialised on 21st February for each county, 21 days before the first confirmed positive test swab and the first day of available Google mobility data with initial state , where and were treated as target parameters for inference.

Equation (3) implies the *effective (instantaneous)* reproductive ratio, that is the instantaneous reproductive ratio where the reduction in susceptibility is also accounted for: .

In total, the transmission model for each county had four unknowns : the baseline reproductive number (), an effective population size scale (), the number of latently infected individuals on 21st Feb (), and the number of actively infectious individuals on 21st Feb ().

Observation model

The underlying transmission of SARS-CoV-2 is not observed, rather we have access to swab tests and serological tests (positive and negative) aggregated by date and county. Therefore, we developed an observation model that connects unobserved daily transmission rates, which depend on the unknown transmission parameters, to a likelihood of the observed test data. There was substantial day-to-day variation in both reported numbers of positive swab tests and percentage of positive tests among all samples confirmed that day (where negative test data is available). The underlying causes of the high level of day-to-day volatility are probable multiple including variation in daily testing rate, as well as in the settings at which swab test were collected, e.g. at hospital, from a walk-up testing facility etc, as the focus of Kenyan public health teams has shifted over the course of the epidemic. Because of the substantial day-to-day variation we use the standard *robust* alternatives to the natural Poisson and Binomial models for count data, the Negative binomial and Beta-Binomial models respectively (*30*). The choice of count data model used on each day depended on whether negative test data was available for that county on that day (in which case we used a Beta-Binomial model), or if negative test data was unavailable for that county on that day (in which case we used a Negative Binomial model). By defining a likelihood-based observation model we gained access to modern techniques in Bayesian inference to infer both the unknown parameters of the underlying transmission process and the unknown parameters of the observation process. We describe the details of the log-likelihood function below.

Underlying infection process. The number of people who would test positive, either as PCR positive, or as sero-converted, on each day *n* depended on: 1) the rate of new incidence on each day *s < n*, and, 2) the probability that someone who was infected on day *s* is detectable by either PCR or serology respectively days later. The daily numbers of new incidence () on each day *n* in each county was predicted by the transmission model,

|  |  |
| --- | --- |
| for each day *n*. | (4) |

The probability that an infected individual would be determined as having been infected days after infection by either a PCR test or a serology test was denoted, respectively, and .

Fitting . We fitted the sensitivity of a PCR swab test on each day *s* post-symptoms, , to data on diagnostic accuracy given in Zhou et al (*16*), the fitted functional form for PCR-detectability more than 5 days after infection was:

|  |  |
| --- | --- |
| for days. | (5) |

Where was the tail distribution function of a Gamma distribution with fitted shape parameter and fitted scale parameter . This aligns with Zhou et al that the median period to become PCR undetectable after symptoms was 20 days with reported interquartile range of 17-24 days (*16*). doesn’t account for the delay between infection and becoming PCR detectable. To account for this delay we assumed that the distribution of delay between infection and maximum detectability followed the same distribution as the delay between infection and onset of symptoms (among those infected individuals who present with COVID-19 symptoms) as reported by Lauer et al (*31*), . Therefore,

|  |  |
| --- | --- |
| . | (6) |

Where is the probability of developing symptoms on day *s* after infection. The true maximum sensitivity of the PCR test in a typical Kenyan setting is absorbed into our observation model via detection probability parameters (see below). is displayed in Fig. S3.

Fitting . The lag between symptoms and maximum detectability by serological assays has been reported as 21 days in a recent metastudy of reported diagnostic sensitivities (*32*). We assumed that, given an additional 5 day lag after infection (mean of onset time distribution), the sensitivity of the serological assay increased linearly from 0 at infection to a maximum of 82.5% (*11*) over 26 days.

|  |  |
| --- | --- |
| for days,  for days. | (7) |

is displayed in Fig. S3.

Observable infection status in each county. By combining the underlying infection process and the delay between infection and observability in our available data sets we find that the number of people who would test positive on each day in each county, with a PCR test , and/or a serology test , is,

|  |  |
| --- | --- |
| ,  ,  where was the midpoint of day *n*. | (8) |

was the false positive rate for the serology assay (see table S1). Underlying equation (8) is an assumption that the PCR test is 100% specific to SARS-CoV-2.

Log-likelihood of observed test data on each day. The total number of swab samples collected, including negative tests, was not available for each county on each day (see above). Therefore, as mentioned above we used two different count data models to account for the log-likelihood of the data conditional on the observed infection status in each county. Both count data models had two unknown parameters per county:

* Negative binomial count data model parameters. The mean detection rate per PCR-detectable individual per day by swab testing (), and the clustering factor[[1]](#footnote-1) of the daily detections (). recovers a Poisson distribution for the number of positive swab tests collected each day, allowed the model to infer much greater variance in daily positive than expected from a Poisson distribution (*33*). The publicly available *ourworldindata.org* coronavirus dataset (*22*) showed that the daily number of tests in Kenya increased from 29th March at an approximate rate of 1.6% per day relative to the overall mean tests per day (Fig. S4). Although the number of tests was not available for every day and county, we incorporated information about the increased national testing rate by considering the linearly increasing function,

|  |  |
| --- | --- |
| . | (9) |

On days where negative swab tests were not available, we connect the observable status of epidemic to the data thus,

|  |  |
| --- | --- |
| for each day *n*. | (10) |

Where is the mean number of PCR positives expected by the model accounting for a potentially biased testing rate .

* Beta-binomial count data model parameters. The relative bias of a PCR-detectable individual being tested compared to a PCR-undetectable individual (), and the effective sample size parameter (). recovers a Binomial distribution for the number of positive PCR tests were observed among the tests conducted that day, allowed the model to infer much greater variance in daily proportion of test positives than would be expected from a Binomial distribution (*30*). On days where negative swab tests were available, we connect the observable status of epidemic to the data thus,

|  |  |
| --- | --- |
| for each day *n*. | (11) |

Where is the total number of PCR swab samples collected on day *n* and is the proportion of tests performed returning positive expected by the model, accounting for bias in the sampling regime. The bias parameter recovers an unbiased sample of PCR positives from the underlying population.

It should be noted that the parameters and effectively absorb PCR swab sample biasing, either for or against selecting individuals who have been exposed to SARS-CoV-2, amongst the swab tested subjects along with the sensitivity of the PCR test as performed in realistic situations.

The reported uncertainty in the maximum sensitivity of serology assay was fairly high: the posterior mean sensitivity was 82.5% (credible interval 69.6-91.2%; (*11*)). The posterior uncertainty in the serological sensitivity influenced the confidence the inference method placed on the serological sample data; if the test sensitivity was known to high precision we would treat each day’s serological samples as a binomial draw from an underlying proportion of seroconverted individuals given by equation (8). Given that the sensitivity of the serological assay was itself an uncertain factor we fitted the posterior uncertainty in the testing sensitivity to a beta distribution: . This implied that the appropriate observation model for the number of positive serological samples on day *n* (), out of the total number of serological samples being collected on day *n,* , was a Beta-binomial distribution,

|  |  |
| --- | --- |
| . | (12) |

Given an underlying realization of the transmission process the mean number positive serological samples on day *n* is . The “total-count” parameter (*30*), , allowed for greater dispersion in the observed seropositive count data than would be allowed by a Binomial model. The parameters of the observation model are .

Parameter inference for each Kenyan County

We use the Bayesian inference to infer a joint posterior distribution for the unknown parameters (both transmission-based parameters and observation-based parameters) for each county (*30*). We describe the three main ingredients for our Bayesian approach below: 1) the log-likelihood function for the data given a set of parameters, 2) the county-specific prior distributions for the parameters, and, 3) the Markov-chain Monte Carlo method used to draw parameter sets from the posterior distribution.

Log-likelihood function. The observation model gives the following log-likelihood function for the unknown parameters given the sampling data for a county, :

|  |  |
| --- | --- |
|  | (13) |

Where and are, respectively, the probability mass functions for the negative binomial and beta-binomial distributions, as described in the observation model subsection. and indicate that the sum is over days where either PCR negative tests were available, or not. The first day where samples were included in the log-likelihood calculation was 12th April, due to testing being even more irregular before that date, and the last day was 30th September.

County-specific priors. We divided the Kenyan counties into three groups: 1) The two main cities **Nairobi** and **Mombasa**, 2) semi-urban counties that either contain significant sized cities or neighbour Nairobi county: **Isiolo, Kajiado, Kiambu, Kilifi, Kisumu, Machakos, Nakuru, Uasin Gishu,** and **Vihiga**, and 3) the rest of the counties where a majority of the population are predominantly rural: **Baringo, Bomet, Bungoma, Busia, Elgeyo Marakwet, Embu, Garissa, Homa Bay, Kakamega, Kericho, Kirinyaga, Kisii, Kitui, Kwale, Laikipia, Lamu, Makueni, Mandera, Marsabit, Meru, Migori, Murang’a, Nandi, Narok, Nyamira, Nyandarua, Nyeri, Samburu, Siaya, Taita Taveta, Tana River, Tharaka Nithi, Trans Nzoia, Turkana, Wajir,** and, **West Pokot** counties.

The prior distributions for , , were the same for each county:

* .
* .
* .
* .
* .

These shared priors reflect *weak* confidence that: 1) the fundamental value in Kenya would be similar to that seen in the early Chinese epidemic i.e. , 2) the most *a priori* likely value of was 1, thus recovering the standard homogeneous SEIR model, 3) PCR positive individuals are *a priori* likely to be over-represented in testing because they are more likely become sick with COVID-like or ILI-like symptoms and get tested, and, 4) both the number of positive tests per day and the proportion of positive tests are *a priori* likely to be significantly overdispersed compared to Poisson or Binomial count models. The county group specific priors were:

* **Urban counties.**
  + .
  + .
* **Semi-urban counties.**
  + .
  + .
* **Rural counties.**
  + .
  + .

The county group specific priors were based on the view that although the most *a priori* likely possibility was that counties had very few infected individuals on 21st February, we might expect the unknown numbers of infecteds to be concentrated in cities. The prior for was based on *a priori* belief of about 1500 tests per day in Nairobi with PCR positives twice as likely as a uniform draw of being selected for testing, with a lower chance of detecting positives outside of the cities.

MCMC draws. We used Hamiltonian MCMC with NUTS (*34*) to perform Bayesian inference by drawing 10,000 samples from the posterior distribution,

|  |  |
| --- | --- |
| , For . | (14) |

for each county using the NUTS-HMC sampler implemented by the Julia language package *dynamicHMC.jl*. Solving the likelihood function for a proposed value of involved solving the ODE system (2), we used the highly performant DifferentialEquations.jl package for ODE solutions (*35*). The HMC method required a log-likelihood gradient, , which, for our use-case of a small ODE system with a low number of parameters, was most efficiently supplied by forward-mode automatic differentiation (*36*) implemented by the package *ForwardDiff.jl*.

The MCMC chain converged for each county (all MCMC chains and MCMC diagnostics can be accessed through the linked open code repository). The posterior mean (and 95% CIs) for each parameter, as well as estimated county-specific IFR, can be found in supplementary data: Data S1 for the model variant where fitted contact rates were used for each county (see below), and, Data S2 for the model variant where contact rates were assumed to follow Google mobility estimates.

Inference for effective contact rates

As mentioned above one approach to modelling is to assume that is proportional to the google mobility data (Fig. S1). However, we can also potentially improve upon this by fitting the effective relative contact rate for each day (recall that for parameter identification reasons we fix the first 30 days of to be their google estimate. When fitting we use the google mobility data as a prior, and estimate each as its marginal posterior mode using the expectation-maximisation (EM) algorithm (*30*).

Priors for counties effective relative contact rates. We fitted the log-relative contact rates for each day of google data (Fig. S1) to a Gaussian process with a standard squared exponential kernel function , and performed optimisation over the kernel hyperparameters: local variance , and covariance length-scale . The package used to perform the Gaussian process regression was GaussianProcesses.jl, and the code for the regression can be found as a notebook on the attach repository. The maximum likelihood estimators for each kernel hyperparameter were similar: for Nairobi,   days, for Mombasa days, and for Kenya overall days. Therefore, we used as a prior for each county a down-weighted version of the Gaussian process fit to the google mobility data

|  |  |
| --- | --- |
| . | (15) |

Where the weight parameter *w* is chosen so that the prior regularises the log-posterior density by 5% of the value of a full *a priori* assumption that is identical to the google estimate on each day. Note that using this prior not only penalises large deviations away the google estimates, but also contact rates that change over time more rapidly than the google mobility trends.

Expectation-maximisation algorithm. The idea of the EM algorithm is to converge some of the inference parameters onto their posterior mode rather than performing computationally intensive/impossible joint inference. In our context, we consider the joint set of parameters , that is the transmission model parameters and the daily effective relative contact rates. The algorithm proceeds:

1. Choose an initial guess of for each day *n*.
2. **E step iteration k.** Determine the Q-function:

|  |  |
| --- | --- |
|  | (16) |

This is the log posterior density of the model parameters and contact rate averaged over the posterior distribution of the parameters. We construct the Q function by performing MCMC draws from the posterior distribution as described above. Note that

|  |  |
| --- | --- |
| + terms that don’t depend on | (17) |

Where is the log-likelihood function, with the dependence on contact rate made explicit, and, is the density function of the Gaussian process fit to the google mobility data (see above).

1. **M step iteration k**. We increased the Q-function in two sub-steps:
   1. Used the default adaptive differential evolution (global) optimizer supplied by BlackBoxOptim.jl to provide an initial improvement to . This was based on 2400 particles, and search steps. Global optimizers are robust to local optima in the maximization method.
   2. was given as the initial guess for the standard gradient-free Nelder-Mead optimizer (supplied by Optim.jl) for an additional search steps. The kth iteration of the M-step was the output of this optimisation step.

Proof of algorithm convergence and more details can be found in (*30*).

Hindcasting for Kenyan counties

The parameter draws from the posterior distribution defined the uncertainty in our model nowcasts and forecasts for each county, since the underlying transmission model was deterministic. Therefore, posterior distributions for epidemic quantities were created by simulating the epidemic for each posterior draws, for example the draws from the distribution of day of peak infection rate were,

|  |  |
| --- | --- |
| . | (18) |

The central estimate is the posterior mean, , and the 95% credible intervals were the 2.5% and 97.5% empirical quantiles of the draws from the day of peak. This approach was used through-out the main document. It should be noted that for the daily positive PCR tests we show the posterior mean and CIs for the mean daily detection rate of new PCR-positive swab tests, . The final posterior mode estimate for the contact rate on each day was used for each of the parameter draws.

Inferring a crude infection fatality ratio

The commonly used infection fatality ratio (IFR) by age estimates from Verity et al (*19*), weighted by the Kenyan population distribution given by the 2019 census, implied a basic IFR prediction in Kenya of ; that is 264 deaths per 10,000 infections. This assumed a uniform attack rate across age-groups in Kenya.

We used the posterior predictions of the underlying daily infections in Kenya counties to infer a crude infection fatality ratio () for each Kenyan county. The lag between infection and death, for those infected individuals who die, was defined as the convolution between three time duration distributions:

1. The duration of time between infection and symptoms (days), which we assumed was distributed (*31*).
2. The duration of time between initial symptoms and severe symptoms (days), sufficient to seek hospitalisation, which we assumed was distributed (*37*).
3. The duration of time between severe symptoms and death estimated from UK hospital data (*38*). This was an empirical distribution with mass function .

We discretized the first two distributions to give probability mass functions for the number of days between infection and symptoms, and for the number of days between symptom onset and severe symptom onset. The probability mass function for the (discrete) number of days between infection and death, for those who died, , was given as a discrete convolution over these probability mass functions:

|  |  |
| --- | --- |
| for the probability that death occurs days after infection. | (19) |

The most likely lag between infection and death was 14 days, however, the distribution was fairly heavy-tailed with mean lag between infection and death 19 days (Fig. S3).

We assumed that the number of deaths observed each day were Poisson distributed, and accounted for the lag between infection and death,

|  |  |
| --- | --- |
|  | (20) |

The conditionality on is given to emphasise that the number of infections per day depended on the unknown parameters for the county. Using as a weak prior for each of the county specific estimates of the crude IFR observed in Kenya, conditional on a realization of , we found that the posterior mean estimator took a simple form (see Posterior distribution of crude infection fatality ratio in supplementary text below):

|  |  |
| --- | --- |
| . | (21) |

We approximated the full posterior distribution of using a set of draws: . The posterior mean estimator and credible intervals for the county-specific values were calculated from this set of draws.

Model validation: Model information scores and posterior predictive P-values

When comparing models we used two different model information scores: the popular in-sample Bayesian model fit score, Deviance information criterion (DIC; (*15*)) defined as , and the log predictive density (LPD; (*15*)) for deaths. The LPD for a model was defined as,

|  |  |
| --- | --- |
|  | (22) |

The LPD measure is the sum log probability of observing the death data for the county, was the probability mass function for a Poisson distribution with mean , the notation emphasises that the number of infections on day *m* depends on the *k*th draw of the parameters. Although the IFR was optimised to fit the death data, we also wish to emphasise that the deaths were not used to infer other model parameters; this justified our description of LPD as an out-of-sample metric.

We validated the overall fit of the model using posterior predictive checking of the LPD value (*30*). Posterior predictive checking is a Bayesian model checking analogy to classical statistical chi-squared tests or G-tests. For each county, after inferring a posterior distribution for the unknown model parameters and the county specific IFR estimate, we sampled 1000 *replicated* death data time series, , each of which had their own LPD value. The observed distribution of LPD values represented the expected distribution of log-predictive densities *if the data was really generated from the model* . The posterior predictive P-value for the model is defined as ; that is the observed proportion of replicated LPD values from the model that were greater than the observed LPD value of the true data.

The average value for the posterior P-value over the counties using the model variant with fitted contact rates was (minimum-maximum spread: 0.19-0.82, Data S1), and, (minimum-maximum spread: 0.18-0.83, Data S2) using the model variant with contact rates assumed to follow Google mobility data. Therefore, for either model variant, the model was: 1) typically *more accurate* at predicting the daily number of deaths actually observed than simulated data, 2) no county had observed deaths in the tails of the model predictions. Therefore, we conclude that the trend in reported deaths is consistent with predictions of either model variant.

Although, both model variants perform reasonably in predicting out-of-sample trends in reported deaths, overall, we found that the model variant with fitted contact rates was better supported by the data. Both the DIC model score over all county fits ( and the LPD model score over all county fits (2) were substantially improved by using the fitted contact rate model variant. Therefore, in the main manuscript we present results from this model variant. However, it should be noticed, that whilst DIC improvements were found for every county by using fitted contact rates, not all counties had improved LPD scores. For completeness, we give a comparison plot of model hindcasting for Nairobi and Mombasa using the model variant where contact rates were assumed to be the same as Google estimates (Fig S5). Model fit scores for each county are given in Data S3.

Supplementary Text

Notation for distributions used in this study

In this study, we have used a number of parameter symbols that are also the most commonly used symbols for various common parametric distributions. Moreover, these parametric distributions are used in the underlying analysis frequently with their distribution parameters defined as functions of underlying transmission model states. To reduce misunderstanding reserve symbols with “hats” as referring to the parameters of a parametric distribution and use “=” to refer to the value of the parameter. Find below the choice of parametrization for the parametric distributions used in the study:

* Exponential distribution. ), with mean .
* Gamma distribution. , with shape parameter and scale parameter .
* LogNormal distribution. , with log-mean parameter and log-standard deviation .
* Negative binomial distribution. , with mean and clustering factor (inverse dispersion parameter) .
* Beta distribution. , with shape parameters (mean ).
* Beta-binomial distribution. , with number of samples , marginal probability per draw , and effective sample size (see for details of this slightly unusual parameterization).
* Poisson distribution. , with mean .

Posterior distribution of crude infection fatality ratio

Here we give details of the posterior distribution of . It is well known that the exponential distribution is a conjugate prior to the mean of a Poisson distribution; that is that given a model where data for , and a prior , then the posterior distribution of is gamma distributed with an analytical solution: We demonstrate that the posterior crude infection fatality ratio, *conditional on some fixed value of the parameters* , , and with the prior distribution is also gamma distributed. This holds despite the mean number of deaths changing daily, and the posterior distribution only depends on the total number of observed deaths and the cumulative number of time delayed infections. First, we define the time delayed infections on each day using the lag distribution [equation (19)], We define the total time delayed infections, , and the total deaths, .

Then, after cancelling terms, the posterior distribution of given the daily deaths is,

|  |
| --- |
| . |

By identifying the denominator as the moment of an exponential distribution we arrive at,

|  |
| --- |
| . |

The right hand side of this expression is the density function of a distribution, that only depends on the total deaths and time-delayed infections. Therefore, the posterior mean estimator for the crude infection fatality ratio conditional on the death data and a fixed value of is:

|  |
| --- |
| . |

Fig. S1.



Fig S1: **Google mobility trends**. The mobility trends used in this study. The curves show a 7-day moving average of the mean relative mobility in the, Google defined setting categories, “retail and recreation”, “grocery and pharmacy”, “transit stations”, and, “workplaces” for Nairobi (*red curve*), Mombasa (*green curve*), and the overall Kenyan trend (*blue curve*).

Fig. S2.



Fig S2: **Difference between *ourworldindata.org* collation of reported total daily national tests in Kenya and Kenyan laboratory linelist**. Data from the Kenyan laboratory linelist is only used in the period after 7th July 2020. This figure was generated by comparing the total linelist tests (positive and negative where available (Data S6) with the reported total tests (Data S7).

Fig. S3.



Fig S3: **Distributions of time between infection and observable events for those infected.** The time-since-infection dependent probability of being detectable by a PCR test *(top left*) or serology test *(top right)* used in this paper. The distribution of time between infection and death, conditional on a death outcome *(bottom)* used in this paper.

Fig. S4.



Fig S4: **Trend in daily testing overall Kenya since 29th March.** The number of tests collected each day reported as a Kenya-wide statistic relative to a mean testing rate of 2482 tests per day across Kenya (*blue* dots; data available from an open source repository (*22*), Data S7). Testing rate increases approximately linearly (*red line*) with a trend of 1.6% increase in testing relative to the mean (40 additional tests per day).



Fig S5: Identical to main manuscript Fig. 1 except using contact rates assumed to follow Google trends rather than fitted from data.

Table S1.

|  |  |
| --- | --- |
| **Parameters and variables** | **Value** |
| **Transmission model parameters** | |
| Number of susceptible people at time t, | Dynamic |
| Number of latently infected people at time , | Dynamic |
| Number of infectious people at time , | Dynamic |
| Number of recovered and immune people at time , | Dynamic |
| Cumulative number of infections at time , | Dynamic |
| Numbers in each infection state on 21st Feb | are **inferred from data.**    Where the county population size N is as reported in the 2019 Kenyan census. |
| Infectious period | 2.4 days. This was chosen to recreate a serial interval of 5.5 days (*39*). |
| Latent period | 3.1 days. The mean incubation period (*31*) was reduced by two days of pre-symptomatic transmission to give a latency period. |
| Baseline reproductive number, . | **Inferred from data** |
| Effective population size scale, | **Inferred from data** |
| Mobility of population in area on day *n* relative to 2019 baseline, . | Data |
| Instantaneous reproductive number, . | for t in day *n*. |
| Effective instantaneous reproductive number, . | for t in day *n*. |
| **Observation model parameters and data** | |
| Number of people who would test PCR positive on day *n*, . | Dynamic |
| Number of people who were observed to test PCR positive on day *n*, . | Data |
| Number of people who would test as sero-converted on day *n*, . | Dynamic |
| Number of people who actually test as sero-converted on day *n*, . | Data |
| Probability that an infected individual would test PCR positive on day after infection, | where was the tail function of a gamma distribution fitted to data given in (*16*), and is the probability function of onset of symptoms post-infection (*31*). |
| Probability that an infected individual would be detectably seropositive on day after infection, | is linearly increasing over 26 days to saturate at 82.5% sensitivity, based on report delay in seroconversion (*32*) and maximum sensitivity of serological assay (*11*). |
| Mean number of PCR tests per capita per day (used when negative tests are unavailable), . | **Inferred from data** |
| Clustering coefficient of daily PCR tests (used when negative tests are unavailable), . | **Inferred from data** |
| Relative bias in favour of selecting a PCR positive individual for testing (used when negative tests are available) () | **Inferred from data** |
| Effective sample size/Clustering coefficient of daily PCR tests (used when negative tests are available), (). | **Inferred from data** |
| **Fatality rate parameters** | |
| Infection fatality ratio, IFR | **Inferred from data** |
| Probability mass function of delay lag between infection and death for those who die, . | Derived as a convolution over the lag from infection to symptom onset (*31*), the lag from symptoms to hospitalisation/severe symptoms (*37*), and the lag between severe symptoms and death (*38*). |

Table S1: **Dynamic and observational model variables and parameters.** “Dynamic”, means that the variable was an output of the transmission and observation model for the county.

Data S1. (separate file)

**Inferred parameters and posterior predictive P-values for each Kenyan county – using fitted values.** The posterior mean and 95% credible intervals for each model parameter (both transmission and observation models), with posterior mean and 95% credible intervals for the peak day of infection rate and county-specific infection fatality ratio. Final column is the proportion of synthetic replicated death time series with greater log-predictive density score than the observed death time series for that county (the posterior predictive P-value for the model).

Data S2. (Separate file)

**Inferred parameters and posterior predictive P-values for each Kenyan county – assuming values follow Google trends.** The posterior mean and 95% credible intervals for each model parameter (both transmission and observation models), with posterior mean and 95% credible intervals for the peak day of infection rate and county-specific infection fatality ratio. Final column is the proportion of synthetic replicated death time series with greater log-predictive density score than the observed death time series for that county (the posterior predictive P-value for the model).

Data S3. (Separate file)

**Model fit scores for each county.** The model fit scores shown are for both fitted values, and when assuming they follow Google trends. Two types of model fit scores are shown: Deviance information criterion (DIC) of county fit to swab and serology testing, and, the log-predictive density for observed deaths in the county.

Data S4. (Separate file)

**The number of positive, and negative where available, PCR-confirmed swab tests for each county by date of sample collection (21st Feb to 30th September).** A “-1” in the negative swab entry indicates that negative swab tests were not available for that county on that date.

Data S5. (Separate file)

**The number of positive and negative serological results for each county by date of sample collection (21st Feb to 6th August).**

Data S6. (Separate file)

**The number of deaths with a PCR-confirmed swab test for each county by recorded date of death (21st Feb to 30thth September).**

Data S7. (Separate file)

**Summary data of Kenyan epidemic, including reported total number of test performed in Kenya.** The total number of tests nationally was used to scale the detection rate when negative test results were unavailable.

1. Here the clustering coefficient is the inverse of the dispersion parameter *k*, a common alternative parameterisation of the negative binomial distribution. [↑](#footnote-ref-1)