SUPPLEMENTARY INFORMATION

Transmission model

Transmission model overview

The dynamics of SARS-CoV-2 transmission in each of the 47 Kenyan counties were assumed to follow a dynamic model adapted from a previous model;[1] here we extend the (modified SEIRS type) transmission model structure to include age stratification and vaccination. The epidemiological dynamics in each county are described by the following system of differential equations:

$$\frac{dS_{a,v}}{dt} = -\sigma_a S_{a,v}(1-\pi_v^{sus})\lambda_a + \rho_{a,v-1}S_{a,v-1} - \rho_{a,v}S_{a,v}$$
$$\frac{dE_{a,v}}{dt} = \sigma_a\left(S_{a,v} + \sigma_{\omega 1}W1_{a,v} + \sigma_{\omega 2}W2_{a,v}\right)(1-\pi_v^{sus})\lambda_a - \alpha E_{a,v} + \rho_{a,v-1}E_{a,v-1} - \rho_{a,v}E_{a,v}$$
$$\frac{dA_{a,v}}{dt} = \alpha E_{a,v}(1-\delta_a) - \gamma_A A_{a,v} + \rho_{a,v-1}A_{a,v-1} - \rho_{a,v}A_{a,v}$$
$$\frac{dP_{a,v}}{dt} = \alpha E_{a,v}\delta_a - \alpha_P P_{a,v} + \rho_{a,v-1}P_{a,v-1} - \rho_{a,v}P_{a,v}$$
$$\frac{dD_{a,v}}{dt} = \alpha_P P_{a,v} - \gamma_D D_{a,v} + \rho_{a,v-1}D_{a,v-1} - \rho_{a,v}D_{a,v}$$
$$\frac{dR_{a,v}}{dt} = \gamma_A A_{a,v} + \gamma_M M_{a,v} + \gamma_V V_{a,v} - \omega R_{a,v} + \rho_{a,v-1}R_{a,v-1} - \rho_{a,v}R_{a,v}$$
$$\frac{dW1_{a,v}}{dt} = \omega_1 R_{a,v} - \sigma_\omega\sigma_a W1_{a,v}(1-\pi_v^{sus})\lambda_a - \omega_2 W1_{a,v} + \rho_{a,v-1}W1_{a,v-1} - \rho_{a,v}W1_{a,v}$$
$$\frac{dW2_{a,v}}{dt} = \omega_2 W1_{a,v} - \sigma_a W2_{a,v}(1-\pi_v^{sus})\lambda_a + \rho_{a,v-1}W2_{a,v-1} - \rho_{a,v}W2_{a,v}$$

Equation 1

The state variables are: Susceptible (**S**), latently infected (**E**), asymptomatically infectious (**A**), pre-symptomatic infectious (**P**), symptomatic/diseased infectious (**D**), recovered and temporarily immune (**R**), and previously recovered/immune whose immunity to reinfection has waned (**W1**) before disappearing (**W2**). The indexing variables are age (a) and vaccination dose status (v), see below for further details on index variable structure. The age-specific force of infection was denoted λ_a (see below for details). The rate of progressing through latent uninfected stage (α), pre-symptomatic infectious state (α_P), rate of loss of complete immunity in two stages (ω_1, ω_2), and the decreased susceptibility due to prior infection in stage **W1** (σ_ω), were assumed to be identical for all age groups, vaccination statuses, prior infection events, and eventual severity of the infection episode. The recovery rates (γ_A, γ_D) depended on the severity of the episode, but not other factors. The baseline susceptibility per infectious contact (σ_a) and probability of developing symptoms (δ_a), were assumed to depend on the age of the individuals. π_v^{sus} gives the effectiveness of vaccine status v at blocking transmission per infectious contact (see below). The per-capita rate at which a person in age group a and with vaccine status v transitioned to their next vaccine status was denoted $\rho_{a,v}$ (see below). Supplementary figure 1 gives a visual overview of the model dynamics.

1

1

Age structure for transmission model

In this paper we used a coarse-grained set of age indices, partly to reduce number of model compartments and thereby increase the efficiency of parameter estimation (see below), and also partly because the data used in parameter estimation did not support a finer grained age index. The six age groups corresponding to the age indices $a = 1, \ldots, 6$ were 0-19, 20-49, 50-59, 60-69, 70-79 and 80+ year olds. When using more fine-grained age-specific data we used population weighted average values over the finer age group categories within each of the coarser six age groups used in this model to generate data. The data we used in the transmission model which was translated from a larger number of age groups to the six used in this model was as following:

- Age- and setting- specific contact rates for Kenya from Prem and Jit.[2] These are provided in 16 5-year age groups and 75+ year olds.
- Age-specific relative susceptibility to infection (σ_a). These were accessed from UK focused modelling and analysis.[3] (Supplementary figure 2)
- Age-specific chance of a symptomatic episode given infection (δ_a). These were accessed from UK focused modelling and analysis.[3] (Supplementary figure 2)

Vaccine status and vaccination effects on transmission model dynamics and waning vaccination dynamics

We used five vaccination statuses to index individuals: unvaccinated ($v = 1$), vaccinated with one ($v = 2$) or two doses ($v = 3$) with sufficient time elapsed since dose inoculation that maximum vaccine efficacy had been achieve (assumed to be 14 days), and two stages of waned vaccination ($v = 4,5$). The vaccination waning dynamics follow those described by Keeling et al.[4]

Vaccine status acted on the dynamics of the model in four distinct ways:

1) Decreasing the chance of infection per infectious contact by a factor $1 - \pi_v^{sus}$ where π_v^{sus} was the effectiveness of vaccine dose status v at reducing infection.
2) Decreasing the probability of severe disease after infection by a factor $1 - \pi_v^{dis}$ where π_v^{dis} was the effectiveness of vaccine dose status v against severe disease.
3) Decreasing the probability of death after infection by a factor $1 - \pi_v^{death}$ where π_v^{death} was the effectiveness of vaccine dose status v against death.
4) Decreasing the infectiousness of dosed infecteds by a factor $1 - \pi_v^{inf}$ where π_v^{inf} was the effectiveness of vaccine dose status v against transmitting infection.

We follow the "VE -> 0%" scenario from Keeling et al,[4] where the effectiveness of the vaccine against acquisition of COVID-19 and infectiousness during a COVID-19 episode eventually decreases to zero ($\pi_5^{sus} = 0, \pi_5^{inf} = 0$), whilst the eventual effectiveness against severe disease and death decreases to 70% ($\pi_5^{dis} = \pi_5^{death} = 0.7$). The pre-waned vaccination status ($v = 4$) has the same vaccine effectiveness as full second dose vaccination status ($v = 3$) but is used to give non-exponential waning rates over an average of 430 days for both stages of waning vaccine effectiveness. Therefore, the per-capita transition rates $\rho_{a,v}$ in equation 1 divide into:

$$\rho_{a,v} = \begin{cases} 0, & v = 0 \\ \upsilon_{a,v}, & v = 1,2 \\ \xi_1, & v = 3 \\ \xi_2, & v = 4 \\ 0, & v = 5 \end{cases}. \qquad \text{Equation 2}$$

Where $\upsilon_{a,1}$ and $\upsilon_{a,2}$ are the per-capita daily rates at which individuals in age group a receive their first and second doses of vaccine, and $\xi_1 = \xi_2 = 2/430$ per day are the waning immunity rates of vaccine effectiveness.

Although, this model choice closely follows Keeling et al,[4] it should be noted that unlike Keeling et al we assume that protection from vaccines *and* natural infection combine favorably whereas in Keeling et al it was assumed that protection from natural infection dominated vaccine protection.

Force of infection for transmission model

We define the age-dependent force of infection (λ_a) in three steps. First, we define the effective number of infected in each age group b,

$$I_b = \sum_v (1 - \pi_v^{inf})(\epsilon A_{b,v} + P_{b,v} + D_{b,v}) + I_{ext,b}. \qquad \text{Equation 3}$$

The effective number of infecteds is the total number rescaled by decreased levels of infectiousness, such as the relatively lower infectiousness of asymptomatic infecteds compared to pre- and post-symptomatic infecteds (ϵ). $I_{ext,b}$ represented an external coupling with infectious people external to internal transmission dynamics. We chose $I_{ext,b}$ such that $\sum_b I_{ext,b} = 100$. Second, the rate of infectious contacts from an effective infected in age group b to anyone in age group a was defined as,

$$T_{ab} = \frac{\beta_0(t)\left(\beta_{home} T_{ab}^{home} + \beta_{school}(t)\, T_{ab}^{school} + \beta_{work} T_{ab}^{work} + \beta_{other} T_{ab}^{other}\right)}{N_a}. \qquad \text{Equation 4}$$

Contacts occur in any of four main settings: at home, at school, at work or in some other social setting, each of which has an age-specific contact rate matrix for Kenya ($T_{ab}^{home}, T_{ab}^{school}, T_{ab}^{work}, T_{ab}^{other}$) estimated by Prem and Jit.[2] Google mobility data[5] and previous epidemic modelling in Kenya [1] suggest that contact rates in Kenya had returned to approximately pre-pandemic baseline by January 2021, therefore, we treated the setting specific transmission rates per contact for at home, at work and other social setting ($\beta_{home}, \beta_{work}, \beta_{other}$) as constant. Transmission rate per contact at schools ($\beta_{school}(t)$) was constant during term time, but dropped to zero during Kenyan school holidays (19th March – 10th May, 16th July – 26th July, 1st October – 11th October, 23rd December – 4th January). The baseline transmission rate per contact ($\beta_0(t)$) varied according to the SARS-CoV-2 variant frequency in the county (see below). Third, the force of infection was defined as,

$$\lambda_a = \sum_b T_{ab} I_b. \qquad \text{Equation 5}$$

Alpha, Beta and Delta variant frequency effect on transmission model dynamics

The Alpha, Beta and Delta variants of SARS-CoV-2 have circulated in Kenya,[6] displaying a pattern of sequential dominance. From December 2020, Alpha and Beta variants increased in frequency [7] with

a complex spatial pattern of relative frequency of Alpha vs Beta within Kenya.[7] Then, from May 2021, the frequency of Delta variant increased very rapidly to complete domination across Kenya.

We modelled the effect of variant frequency on the baseline transmission rate per infectious contact $\beta_0(t)$ as a sequence of strain dominations each occurring over a timescale set by a logistic growth curve,

$$\beta_0(t) = \beta_{wt}\left[1 + R_{\alpha\beta} logistic\left(t; r_{\alpha\beta}, T_{\alpha\beta}\right)\right][1 + R_\delta logistic(t; r_\delta, T_\delta)]. \quad \text{Equation 6}$$

Where $logistic(t; r, T) = \frac{\exp(r(t-T))}{1+\exp(r(r-T))}$. In this model, β_{wt} is the baseline transmission rate per infectious contact of the origin strain(s), called wild-type strains, circulating in Kenya before Alpha and Beta variants, $(1 + R_{\alpha\beta})$ is the proportional change in the reproductive number for SARS-CoV-2 after domination by Alpha or Beta variant relative to wild-type strains, and $(1 + R_\delta)$ is the proportional change in the reproductive number for SARS-CoV-2 after domination by Delta variant relative to Alpha or Beta variant. $r_{\alpha\beta}, r_\delta, T_{\alpha\beta}, T_\delta$ set the exponential rates and timing of the logistic growth curves. The implied relative frequencies of wild-type, Alpha/Beta and Delta variants over time are,

$$\begin{gathered} f_{wt}(t) = \left[1 - logistic\left(t; r_{\alpha\beta}, T_{\alpha\beta}\right)\right][1 - logistic(t; r_\delta, T_\delta)], \\ f_{\alpha\beta}(t) = logistic\left(t; r_{\alpha\beta}, T_{\alpha\beta}\right)[1 - logistic(t; r_\delta, T_\delta)], \\ f_\delta(t) = logistic(t; r_\delta, T_\delta). \end{gathered} \quad \text{Equation 7}$$

Transmission model observables: Proportions PCR test and serology test positive

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, age, and county. The chance infected individuals test positive for either type of test depends on the number of days post-infection. Therefore, we coupled the dynamics cumulative infections to transmission model,

$$\begin{gathered} \frac{dF_{a,v}}{dt} = \sigma_a S_{a,v}(1 - \pi_v^{sus})\lambda_a, \\ \frac{dC_{a,v}}{dt} = \sigma_a\left(S_{a,v} + \sigma_\omega W1_{a,v} + W2_{a,v}\right)(1 - \pi_v^{sus})\lambda_a, \\ \frac{d\bar{C}_{a,v}}{dt} = \sigma_a\left(S_{a,v} + 0.5\sigma_\omega\left(1 - \pi_5^{dis}\right)W1_{a,v} + \left(1 - \pi_5^{dis}\right)W2_{a,v}\right)(1 - \pi_v^{sus})\lambda_a. \end{gathered} \quad \text{Equation 8}$$

Where $F_{a,v}(t)$ and $C_{a,v}(t)$ were the cumulative number of people in the county in age group a and vaccine status v infected by time t, respectively split by it being their first infection episode or any infection episode. $\bar{C}_{a,v}(t)$ was the reinfection disease risk-weighted cumulative infection rate; that is the total infections discounted by the decreased risk of severe disease among reinfections (see below for use of this observable). In the absence of other evidence, we follow Keeling et al[4] in assuming that protection against disease due to prior infection is similar to that of vaccination. For the fully waned immunity post-natural infection state (**W2**) we assume that the protection from disease is equivalent to fully waned vaccine protection, and, for the partially waned immunity post-natural infection state (**W1**) that this protection is halved relative to **W2**.

The daily new infections, on each day n starting at $t = n$ are then,

$$\iota_{F,a,v}(n) = F_{a,v}(n+1) - F_{a,v}(n), \;\; n = 1,2,\ldots$$
$$\iota_{C,a,v}(n) = C_{a,v}(n+1) - C_{a,v}(n), \;\; n = 1,2,\ldots \quad \text{Equation 9}$$

The probability that an infected individual would be determined as having been infected τ days after infection *if* tested by either a PCR swab test or a serology test were denoted, respectively, $Q_{PCR}(\tau)$ and $Q_{sero}(\tau)$. We used the same Q_{PCR} and Q_{sero} probabilities as in Brand et al 2021.[1]

By combining the underlying infection processes 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 n in each county with either a PCR test ($P_a^+(n)$), or a serology test ($S_a^+(n)$), was,

$$P_a^+(n) = \sum_{s=1}^{n-1} \sum_v \iota_{C,a,v}(s) Q_{PCR}(n-s),$$
$$S_a^+(n) = \sum_{s=1}^{n-1} \iota_{F,a,1}(s) Q_{sero}(n-s) + p_{FP} \sum_v S_{a,v}(t). \quad \text{Equation 10}$$

where t was the midpoint of day n.

p_{FP} was the false positive rate for the serology assay (see supplementary table 1). Underlying assumptions for equation 10 are: 1) that the PCR test is 100% specific to SARS-CoV-2, 2) that *only* the first infection contributes to the serological status of individuals, but that reinfections contribute to PCR status equally to first infections, and 3) that during the period of transmission parameter inference there were effectively zero vaccinated individuals, and therefore, all seropositivity was evidence of prior nature infection. We do not use any serological data from Kenya after May 2021 when the vaccination rate was very low in Kenya, and, therefore, vaccinations have negligible effect on prevalence of SARS-CoV-2 specific antibodies.

The number of people PCR positive is not observed directly, but rather test positive and negative swab test samples. We consider the *proportion* of these daily samples that are positive to be a potentially biased sample of the true underlying proportion that would be PCR-positive if everyone was tested ($\sum_a P_a^+ / \sum_a N_a$). Therefore, we model the expected proportion PCR test positive (over all age groups) on day n as,

$$\bar{P}^+(n) = \frac{\chi \sum_a P_a^+(n)}{(\chi - 1)\sum_a P_a^+(n) + N}. \quad \text{Equation 11}$$

Where χ is an observed swab sample bias parameter, where $\chi = 1$ indicates unbiased sampling, $\chi < 1$ indicates bias in favour finding PCR negative individuals (i.e., $\bar{P}^+(n) < \sum_a P_a^+ / \sum_a N_a$), and $\chi > 1$ indicates bias in favour of finding PCR positive individuals (i.e., $\bar{P}^+(n) > \sum_a P_a^+ / \sum_a N_a$).

Clinical outcome model

Clinical outcomes of SARS-CoV-2 infections

Severe infections eventually lead to a clinical outcome. We consider three possibilities in this model:

1) Deadly outcome. Deadly infected individuals die after a delay period defined by the probability distribution f_μ. Death, conditional on infection, occurs with probability μ_a^D.
2) Critical outcome. Critically infected individuals require a stay in ICU for a duration defined by the probability distribution f_{ICU}, then move to a general ward in a hospital or health facility,

where they stay for a duration defined by the probability duration f_{hosp}. Critical disease, conditional on infection, occurs with probability μ_a^C.

3) Severe outcome. Severely infected individuals require a stay in a general ward in a hospital or health facility, where they stay for a duration defined by the probability duration f_{hosp}. Severe disease, conditional on infection, occurs with probability μ_a^S.

Clinical outcome model observables: Reported incidence of deaths, occupancy of general wards and Intensive care units

Incidence rate of clinical outcomes. The lag between infection and needing treatment, for those infected individuals who die, was defined as the convolution of two time-duration distributions:

1. The duration of time between infection and symptoms (days), which we assumed was distributed $LogNormal(\widehat{log\mu} = 1.64, \widehat{logstd} = 0.36)$.[8]
2. The duration of time between initial symptoms and severe symptoms (days), sufficient to seek hospitalisation, which we assumed was distributed $U(1,5)$.[9]

We discretized the two distributions to give probability functions f_{IS} for the number of days between infection and symptoms, and $f_{SH}$for the number of days between symptom onset and severe symptom onset. The probability function for the (discrete) number of days between infection and severe or critical disease, for those who died, f_{dis}, was given as a discrete convolution over these probability mass functions:

$f_{dis}(\tau) = [f_{IS} * f_{SH}](\tau)$ for the probability that severe or critical disease leads to seeking medical assistance τ days after infection, conditional on that as outcome. Equation 12

We use this delay distribution to give a rate of people in age group a requiring medical treatment on day n up to some unknown age-dependent and variant-dependent risk-factor, which will fit against available data on reported severe, critical and deadly outcomes,

$$\iota_{dis,a,v,r}(n) \propto \sum_{s<n} f_r(s)\left[\bar{C}_{a,v}(s+1) - \bar{C}_{a,v}(s)\right] f_{dis}(n-s). \quad \text{Equation 13}$$

Where $f_r(s)$ is the relative frequency of variant r on day s (see equation 7). Note that in equation (13) the reduction in risk due to reinfection is already accounted for (see equation 8).

Observation of incidence of deadly outcome of infection. There is likely to be under-reporting of deaths due to COVID-19 in Kenya.[10,11] In this paper, we don't have sufficient data to estimate the true level of under-reporting of deaths in Kenya. However, by assuming that the age-dependent risk of death after infection with SARS-CoV-2 is the same in every county, we can estimate county-specific under-reporting/change in risk *relative* to the capital Nairobi. Concretely, we model the expected number of observed deaths on each day n, in each age group a, and each county c, as,

$$Deaths_{a,c}(n) = \psi_c \mu_a^D \sum_{s<n} \sum_{v} \sum_{r} \psi_r^D (1 - \pi_v^{death}) \iota_{dis,a,v,r}(s) f_\mu(n-s). \quad \text{Equation 14}$$

Where ψ_c is the county-specific under-reporting rate with $\psi_{nairobi} = 1$, ψ_r^D is the relative risk of death by variant with $\psi_{wt}^D = 1$, and the duration of time between needing treatment and death had probability distribution f_μ.

Observation of hospital and ICU occupancy due to severe or critical infections. We don't have access to reports on the incidence of severe and critical cases arriving at hospitals/health facilities and ICUs;

the relevant observables from the clinical outcome model are occupancies of patients overall in Kenya by setting rather than arrival of patients at those settings.

The probability distributions for length of stay in either hospital/health facility or ICU are denoted as f_{hosp} and f_{ICU}, implying upper distribution functions $Q_{hosp}(n) = \sum_{s>n} f_{hosp}(s)$ and $Q_{ICU}(n) = \sum_{s>n} f_{ICU}(s)$ for the probability that a stay in hospital or ICU is longer than n days. The expected total number of people in intensive care units on day n is,

$$ICU(n) = \sum_{a,c} \psi_c \mu_a^C \sum_{s<n} \sum_{v} \sum_{r} \psi_r^C (1 - \pi_v^{dis}) \iota_{dis,a,v,r}(s) Q_{ICU}(n-s). \quad \text{Equation 15}$$

Where ψ_r^C is the relative risk of critical disease by variant with $\psi_{wt}^C = 1$. After a critical case has completed a stay in an ICU, we model them as having a stay in a general ward with the same distribution of length as per a severe case admitted to a general ward (without a stay in ICU). The upper distribution function for the whole stay in ICU *and* general ward is then $Q_{ICUH}(n) = \sum_{s>n} [f_{ICU} * f_{hosp}](s)$, giving the expected number of critical cases in either ICU or general ward as,

$$ICUH(n) = \sum_{a,c} \psi_c \mu_a^C \sum_{s<n} \sum_{v} \sum_{r} \psi_r^C (1 - \pi_v^{dis}) \iota_{dis,a,v,r}(s) Q_{ICUH}(n-s). \quad \text{Equation 16}$$

The expected number of patients occupying general wards is the addition of severe cases who have been admitted directly to general wards, and critical cases who have completed their stay in ICU and are now in general wards,

$$HOSP(n) = [ICUH(n) - ICU(n)] + \sum_{a,c} \psi_c \mu_a^S \sum_{s<n} \sum_{v} \psi_r^S (1 - \pi_v^{dis}) \iota_{dis,a,v,r}(s) Q_{hosp}(n-s). \quad \text{Equation 17}$$

Where ψ_r^S is the relative risk of severe disease by variant with $\psi_{wt}^S = 1$.

Parameter Inference

In this work, we make inferences on two groups of parameters:

1. The parameters of the transmission model in [equations 1-5], and the bias parameter (χ) for the observed versus actual proportion PCR positive in daily swab test [equation (11)].
2. The parameters of the clinical outcome model [equations 12-15]: the relative under-reporting rate by Kenyan county ψ_c, the age-dependent clinical outcome probabilities μ_a^S, μ_a^C, μ_a^D, and the variant specific clinical outcome probabilities ψ_r^D, ψ_r^C, ψ_r^S.

For the transmission model parameters, we used Bayesian inference to infer a joint posterior distribution for the parameters for each county. For the clinical outcome model, we inferred parameters by minimizing the divergence between model prediction of the observables [equations 14-15 and 17] and actual reporting, under the assumption that the rate of people arriving for medical treatment, up to the unknown risk factors [equation 13], was that implied by the posterior mean prediction implied by the Bayesian inference of the transmission model parameters.

A challenge with using the linelist data in Kenya for inference of transmission was that the metadata concerning the reason for receiving a swab test, the levels of symptoms of people who tested positive,

and their healthcare outcomes were often missing. Overall, more than 90% of the people who tested positive in Kenya, and for whom we have a description of their symptoms, reported no symptoms (asymptomatic). Therefore, unlike model-based inference for COVID-19 transmission in high-income countries we didn't use severe outcomes such as hospitalization or death as data sources for inference, e.g.,[12–14] because this data was unreliable. Instead, we concentrated on fitting to the proportion positive of daily swabs test and serological tests jointly with detection rate of cases (see **Transmission model observables: Proportions PCR test and serology test positive** above). It should be noted that this meant that we didn't use age-specific PCR test data, but rather fitted to the aggregate proportion positive over all age groups. However, we did use age-specific seroprevalence data.

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 hierarchy of prior distributions for the parameters, and, 3) the Markov-chain Monte Carlo method used to draw parameter sets from the posterior distribution.

Bayesian inference of transmission model parameters

Data and log-likelihood function for transmission model. Given the daily PCR and serology data for a county,

$$\boldsymbol{\mathcal{D}_c} = \{ObsS_a^+(n), ObsS_a^-(n), ObsP^+(n), ObsP^-(n)\}_{a=1,\ldots,6;\ n=1,2,3,\ldots} \quad \text{Equation 18}$$

The log-likelihood function for the unknown transmission parameters (θ_{TM}) in that county was,

$$l(\theta_{TM}) = P(\boldsymbol{\mathcal{D}_c}|\theta_{TM}) = \sum_n ln\, f_{BB}(ObsP^+(n)\,|\widehat{N_s} = N_{PCR}(n), \hat{p} = \bar{P}^+(n), \widehat{M} = M_{PCR}) + \sum_{n,a} ln\, f_{BB}(\, ObsS_a^+(n)|\widehat{N_s} = N_{sero,a}(n), \hat{p} = \frac{S_a^+(n)}{N_a}, \widehat{M} = M_{sero}). \quad \text{Equation 19}$$

Where $f_{BB}(x|\widehat{N}_s, \hat{p}, \widehat{M})$ is the probability function for a Beta-binomial with sample size $\widehat{N}_s$, expected proportion of successes $\hat{p}$, and effective sample size $\widehat{M}$. This was a convenient reparameterization of the Beta-binomial model under the transformation $\widehat{N}_s = n, \hat{p} = \frac{\alpha}{\alpha+\beta}, \widehat{M} = \alpha + \beta$ from the typical (n, α, β) parameterization. $N_{PCR}(n)$ and $N_{sero,a}(n)$ were the total number of samples (positive and negative) of, respectively, PCR test and serological tests on day n, in age group a. $\bar{P}^+(n)$ was derived from the transmission model for each day n as per above [equations 10-11]. M_{PCR} and M_{sero} were fixed from a previous modelling study.[1] The first day where samples were included in the log-likelihood calculation was 1st January 2021.

Initial conditions and model simulation. We fit to data from 1st January 2021, however, because PCR cases and serological detection are lagged indicators of infections weeks previously, we start the model simulation in each county on 1st December 2020 and use the first month of simulation to allow the simulation to converge onto the epidemic dynamics. Simulation of the model was done by solving the ODE system [equation 1] forwards from the county-specific initial conditions, with county-specific parameter configuration, using an explicit/implicit switching solver provided by the **DifferentialEquations.jl** Julia programming language package.[15]

We reduced the number of unknown parameters for the initial state of the epidemic model by considering only the overall latent infected numbers (E_0) and a scale factor on the proportion exposed

to COVID (i.e. in $R/W1/W2$ epidemiological compartments) relative to a cross-sectional survey done in Nairobi in mid-November 2020[16] which we denote τ. We fix the initial removed and waned immunity numbers as

$$\begin{gathered} R_{a,1}(0) = N_a(0) * P_{sero,a} * \tau * 0.95, \\ W1_{a,1}(0) = N_a(0) * P_{sero,a} * \tau * 0.05, \\ W2_{a,1}(0) = 0. \end{gathered} \quad \text{Equation 20}$$

Where $P_{sero,a}$ was the raw (non-test sensitivity adjusted) seroprevalence estimate for Nairobi in the cross-sectional survey,[16] and τ was an adjustment factor which was added to the set of parameters to be inferred in the set θ_{TM}. Note that we are assuming that 5% of the previously exposed population have lost complete immunity to reinfection by 1st December 2020, and that no previously exposed people had completely lost immunity to reinfection. The adjustment factor τ allowed the model flexibility to represent counties with lower seroprevalence data in 2021 as having had a smaller initially exposed fraction compared to Nairobi, whilst also allowing upwards adjustment to account for the fact that the second wave of cases in Kenya occurred during the Nairobi cross-sectional study.

The age-specific numbers of initially latent infected people were derived from the next-generation matrix $K\left(P_{sero,1}, \dots, P_{sero,6}, \tau, \beta_0, \beta_{school}, \beta_{home}, \beta_{other}, \beta_{work}\right)$, where we have made explicit the parameters being inferred that the next-generation matrix depends upon and its explicit dependence on the baseline seroprevalence estimates. The eigenvector $\boldsymbol{v}$, normalized such that $|\boldsymbol{v}|_1 = 1$, associated with the leading eigenvalue of K, represented the expected distribution of new infections across age groups. Therefore, we specified

$$\boldsymbol{E}(0) = E_0\boldsymbol{v}. \quad \text{Equation 21}$$

Where $\boldsymbol{E}(0) = \left[E_{1,0}(0), \dots, E_{6,0}(0)\right]^T$. The rest of the initial variables were specified as being dependent of the flow out of the latent infected state,

$$\begin{gathered} A_{a,0}(0) = \alpha * E_{a,0}(0) * (1-\delta_a)/(1+\gamma_A), \\ P_{a,0}(0) = \alpha * A_{a,0}(0) * \delta_a/(1+\alpha_p), \\ D_{a,0}(0) = \alpha_p * P_{a,0}(0) * (1-h_a)/(1+\gamma_D). \end{gathered} \quad \text{Equation 22}$$

In every county, every individual in the model was initially unvaccinated.

Priors. In every county we used the following priors for parameter inference:

- $\epsilon \sim Beta(\hat{\alpha} = 50, \hat{\beta} = 50)$.
- $\beta_0 \sim \Gamma\left(\hat{k} = 10, \hat{\theta} = 1.5/10\right)$.
- $\beta_{\text{other}} \sim \Gamma\left(\hat{k} = 10, \hat{\theta} = 1.5/10\right)$.
- $\beta_{\text{home}} \sim \Gamma\left(\hat{k} = 10, \hat{\theta} = 1.5/10\right)$.
- $\beta_{\text{work}} \sim \Gamma\left(\hat{k} = 10, \hat{\theta} = 1.5/10\right)$.
- $\beta_{\text{school}} \sim \Gamma\left(\hat{k} = 10, \hat{\theta} = 1.5/10\right)$.
- $R_{\alpha\beta} \sim \Gamma\left(\hat{k} = 15, \hat{\theta} = \frac{0.4}{15}\right)$.
- $r_{\alpha\beta} \sim \Gamma\left(\hat{k} = 15, \hat{\theta} = \frac{0.15}{15}\right)$.

- $R_\delta \sim \Gamma\left(\hat{k} = 15, \hat{\theta} = \frac{0.6}{15}\right)$.
- $r_\delta \sim \Gamma\left(\hat{k} = 15, \hat{\theta} = \frac{0.2}{15}\right)$.
- $E_0 \sim \Gamma(\hat{k} = 3, \hat{\theta} = \frac{1000}{3})$.
- $\tau \sim \Gamma(\hat{k} = 5, \hat{\theta} = \frac{1.5}{5})$.

The prior for the PCR observation bias parameter χ, differed between counties (see below). For Nairobi and Mombasa we used a prior:

- $\chi \sim \Gamma\left(\hat{k} = 3, \hat{\theta} = \frac{4.5}{3}\right)$.

MCMC draws. We used Hamiltonian MCMC with NUTS[17,18] to perform Bayesian inference by drawing 2,000 samples from the posterior distribution,

$$\theta_{TM}^{(k,c)} \sim P(\theta_{TM}|\boldsymbol{\mathcal{D}_c}) \propto \exp\big(l(\theta_{TM})\big)\pi(\theta_{TM}), \text{ for } k = 1,2,3, \ldots \quad \text{Equation 23}$$

for each county using the NUTS-HMC sampler implemented by the Julia language package *dynamicHMC.jl*. The HMC method required a log-likelihood gradient, $\nabla_\theta(l + \pi)$, which, for our use-case of an ODE system with a comparative low number of parameters (<100 parameters), was most efficiently supplied by forward-mode automatic differentiation 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.[19] The posterior mean (and 95% CIs) for each parameter can also be found in the open code repository.[19]

Approximate county-specific hierarchical model for PCR observation bias. The serological data is important for our inference because it gives information about the proportion of each age group infected at different time points, and, therefore, allows the PCR observation bias parameter (χ) to be identifiable. However, the amount of serological data differs from county to county. To allow cross-inference between counties for the bias parameter we assumed that 1) Nairobi and Mombasa were sufficiently distinct from other counties that the inferred bias parameter for these city/counties was not relevant to other counties, and 2) the other 45 counties had a bias parameter drawn from a common distribution, $\chi_c \sim \Gamma(k_\chi, \theta_\chi)$, where k_χ and θ_χ are the hyperparameters of this hierarchical model. This reflected our underlying belief that despite regional variations in transmission, the observation of data would be similar in all counties outside of the main two urban hubs.

A fully Bayesian approach to inference would involve including $\{\chi_c\}_c$ and the hyperparameters k_χ, θ_χ within a joint log-likelihood over all Kenyan counties (except Nairobi and Mombasa). However, to accelerate inference we used an approximation to this hierarchical model. The 9 counties with the most amount of serological data available, apart from Nairobi and Mombasa, were **Embu**, **Kilifi**, **Kisii**, **Kisumu**, **Kwale**, **Nakuru**, **Nyeri**, **Siaya**, and **Uasin Gishu**. We performed MCMC draws for each of these counties using a prior $\chi \sim \Gamma\left(\hat{k} = 3, \hat{\theta} = \frac{4.5}{3}\right)$, which gathered a set of MCMC draws for χ from the posterior distribution for each county c, $\{\chi^{(k,c)}\}_{k=1,\ldots,2000} \sim P(\chi|\boldsymbol{\mathcal{D}_c})$. We then approximated maximum a-posteriori (MAP) estimates for the hyperparameters k_χ, θ_χ using,

$$\ln P\big(k_\chi, \theta_\chi|\boldsymbol{\mathcal{D}_1}, \ldots \boldsymbol{\mathcal{D}_9}\big) = \ln \pi(k_\chi, \theta_\chi) + \sum_c \ln \sum_k f_\Gamma(\chi^{(k,c)}|\hat{k} = k_\chi, \hat{\theta} = \theta_\chi) + const. \quad \text{Equation 24}$$

Where f_Γ is the density function of a Gamma distribution, and $\ln \pi(k_\chi, \theta_\chi)$ was the log-prior for the hyper-parameters. The hyper-priors used were:

- $k_\chi \sim \exp(\hat{\mu} = 10)$.
- $\theta_\chi \sim \exp(\hat{\mu} = 0.45)$.

The MAP estimation is done under the distributional assumption that the observed outcome data was distributed negative binomially with the same mean as the posterior predictive mean value (that is averaged over the posterior predictive distribution for the infection process), and a negative binomial clustering factor inferred jointly with the risk factors.

We then could use the 9 non-city counties with the most serological data to create MAP estimates $\hat{k}_\chi$, $\hat{\theta}_\chi$ by maximizing equation 24. Equation 24 represented an approximation where we treated the MCMC draws of the χ parameter as "data" for making inference on the hyper-parameters despite using a different prior to generate the MCMC samples.

The second approximation is that for the 36 other counties that were not Nairobi, Mombasa, or one of the 9 listed above, we used a prior for χ generated from these MAP estimates

- $\chi \sim \Gamma\left(\hat{k} = k_\chi/2, \hat{\theta} = 2\theta_\chi\right)$.

The reason for the scaling of 2 was to increase the prior variance for χ to reflect that we are approximating a hierarchical model.

Minimum divergence estimates for the infection outcome model

After performing MCMC we were able to estimate the posterior mean for the rate of diseased incidence, up to a proportionality with unknown risk factors $E[\iota_{dis,a,v,r}(n)|\boldsymbol{\mathcal{D}}]$, for each day n, age group a, variant v, and county c, by solving the ODE system [equation 1] for each set of transmission model parameters drawn from the MCMC and using equation 13. We can then define the posterior expected number of deaths $\overline{Deaths}_{a,c}(n; \psi_c, \psi_r^D, \mu_a^D)$ by replacing the parameter-specific diseased incidence rate in equation 14 with its posterior mean. We define the divergence due to a choice of age dependent mortality rates (μ_a^D), relative reporting rate (ψ_c), relative variant-specific risk of death (ψ_r^D) and clustering factor α_D,

$$Div_D(\psi_c, \psi_r^D, \mu_a^D, \alpha_D) = -2 \sum_{a,c,n} \ln f_{NB}(ObsDeaths_{a,c}(n)|\hat{\mu} = \overline{Deaths}_{a,c}(n; \psi_c, \psi_r^D, \mu_a^D)\ , \hat{\alpha} = \alpha_D). \quad \text{Equation 25}$$

Where $f_{NB}(x|\hat{\mu}, \hat{\alpha})$ is the probability function for a negative binomial with mean $\hat{\mu}$ and clustering factor $\hat{\alpha}$, and $ObsDeaths_{a,c}(n)$ are the daily reported deaths in each age group, each county and on each day. We found a minimum point for equation 25, which we used as estimators, $\hat{\psi}_c$, $\hat{\psi}_v^D$, $\hat{\mu}_a^D$, $\hat{\alpha}_D$ (Supplementary figures 3, 4).

The ICU occupancy data was not available broken down by age or county, therefore, we assumed that the risk of reported critical disease was proportional to the risk of death for everyone and focused on fitting the relative risk of critical disease vs death μ_{rel}^{CD}. We defined the divergence between model prediction of ICU occupancy and observed occupancies due to relative risk of critical disease vs death μ_{rel}^{CD}, variant specific risk of critical disease ψ_r^C, and clustering factor α_C,

$$Div_{ICU}\left(\mu_{rel}^{CD}, \psi_v^C, \alpha_C\right) = -2 \sum_{n} \ln f_{NB}(ObsICU(n)|\hat{\mu} = \sum_{a,c} \overline{ICU}_{a,c}(n; \hat{\psi}_c, \psi_r^C, \mu_{rel}^{CD}\ \hat{\mu}_a^D)\ , \hat{\alpha} = \alpha_C). \quad \text{Equation 26}$$

Where $ObsICU(n)$ was the reported Kenyan National ICU occupancy on day n, and $\overline{ICU}_{a,c}$ is the posterior mean value for the ICU occupancy with COVID. The minimum point for equation 26 gave estimators, $\hat{\psi}_v^C, \hat{\mu}_a^C = \hat{\mu}_{rel}^{CD}\hat{\mu}_a^D, \hat{\alpha}_C$ (Supplementary figures 3, 4).

The general ward occupancy data was also not available broken down by age or county, therefore, we assumed that the risk of reported severe disease was proportional to the risk of death for everyone and focused on fitting the relative risk of severe disease vs death μ_{rel}^{SD}. We defined the divergence between model prediction of general ward occupancy and observed occupancies due to relative risk of severe disease vs death μ_{rel}^{SD} , variant specific risk of severe disease ψ_r^S, and clustering factor α_S,

$$Div_{HOSP}(\mu_{rel}^{SD}, \psi_v^S, \alpha_S) = -2\sum_n \ln f_{NB}(ObsHOSP(n)|\hat{\mu} = \sum_{a,c} \overline{HOSP}_{a,c}(n; \hat{\psi}_c, \psi_r^S, \mu_{rel}^{SD}\,\hat{\mu}_a^D)\ , \hat{\alpha} = \alpha_S). \quad \text{Equation 27}$$

Where $ObsHOSP(n)$ was the reported Kenyan National general ward occupancy with COVID on day n, and $\overline{HOSP}_{a,c}$ is the posterior mean value for the ICU occupancy, with the contribution from patients arriving into general wards from ICU already calculated using the minimum divergence estimates from equation 26. The minimum point for equation 27 gave estimators, $\hat{\psi}_r^S, \hat{\mu}_a^S = \hat{\mu}_{rel}^{SD}\hat{\mu}_a^D, \hat{\alpha}_S$ (Supplementary figures 3,4).

Vaccine scenario projections and immune-escape variant

Vaccination rates. We considered 7 vaccine rollout scenarios starting from 1st September 2021: No vaccination (baseline), 30%, 50%, 70% target coverage of Kenyan over 18s with either an 18 month or 6-month (rapid) time scale (Table 1). In each case we assumed that:

1. The number of vaccines deployed in each county each day was constant over the rollout.
2. Second dose followed first dose after a 56 day lag.
3. Over 50 year olds were offered the vaccine first, but demand saturated at 80% coverage among over 50 year olds (age groups 3-6 in the model) and afterwards the vaccine was offered to 18-49 year olds (age group 2).
4. Vaccines were deployed pro-rata across all disease/infection states; that is there was no dependence on past infection history in seeking vaccines.

Mathematically, this corresponds to this choice for the per-capita vaccination rate (equation 2) for the first dose, that is pro-rata distribution among all unvaccinated groups in stages by age,

$$\upsilon_{a,1}(t) = \frac{NV_{scenario}}{T_{scenario}} \frac{\mathbf{1}(\sum_{a=3,\ldots,6}\sum_{v=2,\ldots,5} N_{a,v} < 0.8\sum_{a=3,\ldots,6} N_a)}{\sum_{a=3,\ldots,6} N_{a,1}},\ \ a = 3,4,5,6$$
$$\upsilon_{2,1}(t) = \frac{NV_{scenario}}{T_{scenario}} \frac{\mathbf{1}(\sum_{a=3,\ldots,6}\sum_{v=2,\ldots,5} N_{a,v} \geq 0.8\sum_{a=3,\ldots,6} N_a)}{N_{2,1}}, \quad \text{Equation 28}$$

$$\text{For } 1st\ sept\ 2021 \leq\ t \leq 1st\ sept\ 2021 + T_{scenario} - 56\ days$$

Where $N_{a,v}$ is the county population by age and vaccine status (summed over disease/infection status). $\mathbf{1}(\cdot)$ is an indicator function enforcing that the vaccination rate among over 50s (age groups 3-6) drops to zero at an 80% coverage. $NV_{scenario}$ is the number of doses implied by the scenario target coverage,

and $T_{scenario}$ is the time in days over which this target coverage is to be achieved in the scenario. The second dose per capita rate is like the first dose but with doses distributed among people who have had their first dose,

$$v_{a,2}(t) = \frac{NV_{scenario}}{T_{scenario}} \frac{\mathbf{1}(\sum_{a=3,\ldots,6} \sum_{v=3,\ldots,5} N_{a,v} < 0.8 \sum_{a=3,\ldots,6} N_a)}{\sum_{a=3,\ldots,6} N_{a,2}}, \quad a = 3,4,5,6$$

$$v_{2,1}(t) = \frac{NV_{scenario}}{T_{scenario}} \frac{\mathbf{1}(\sum_{a=3,\ldots,6} \sum_{v=3,\ldots,5} N_{a,v} \geq 0.8 \sum_{a=3,\ldots,6} N_a)}{N_{2,2}}, \qquad \text{Equation 29}$$

For $1st\ sept\ 2021 + 56\ days \leq\ t \leq 1st\ sept\ 2021 + T_{scenario}$

Uncertainty propagation in vaccination scenarios. In this study we use a mixture of full Bayesian inference, for transmission model parameters, that were specific to each Kenyan county, and minimum divergence estimators, for outcome model parameters, that were specific to Kenyan counties, e.g. the county-specific reporting/disease rate relative to Nairobi ψ_c, or were specific to particular age groups, e.g. the baseline risk of death per infection in each age group μ_a^D.

When generating scenario projections for all Kenya, we solved the transmission model (equation 1) for each of the 47 Kenyan counties and for each of the 2000 MCMC draws of county-specific transmission model parameters. To match the 2000 MCMC draws for transmission parameters per county we drew 2000 replicates of the vaccine effectiveness from reported ranges first and second dose AstraZeneca effectiveness against Delta variant. This generated 2000 expected daily reported death incidence, ICU occupancy and general ward occupancy for each county (equations 14-15, 17) for each county) and for each of the 7 vaccine rollout scenarios.

To account for (1) uncertainty in transmission parameters, (2) unpredictability in reporting, and (3) uncertainty in vaccine effectiveness, the prediction intervals for Kenya as a total were calculated by

1. Augmenting the 2000 projections of expected daily reported observables per day, per age group, and per county into a single group of 2000 projections of expected daily reported observables per day, and per age group by summing across counties.
2. Converting from *expected* daily reported observables to *random* instances by sampling from the negative binomial distribution that minimized divergence between actual observed data and the model projections (equations 25-27) for each day of the 2000 projections.
3. Presenting the daily ensemble average (and 95% ensemble prediction intervals) across the 2000 randomized projections.

Modelling immune escape variant. In this paper we consider an immune escape variant which reduces protection against reinfection by 50% across natural immunity and vaccine protection, and spreads faster through the population with a 30% reduced generation time but is otherwise epidemiologically like the Delta variant. Concretely, we implemented this by assuming that the immune escape variant arrived in Kenya on 15th November 2021, during a period of low infections for other variants, and applying a set of instantaneous effects:

- Relative immune escape variant frequency becomes 100% in all Kenyan counties, reflecting rapid dominance of invading variant.
- All transmission rate parameters (e.g. $\beta_0, \alpha, \alpha_P, \gamma_D, \gamma_A$) increase by a factor 1/0.7.

- 50% of people in fully or partially immune (post-natural infection) categories (**R, W1)** transition instantly to completely waned immunity to reinfection (**W2**).
- Vaccine effectiveness against reinfection and reduction of infectiousness ($\pi_2^{sus}, \pi_3^{sus}, \pi_2^{inf}, \pi_3^{inf}$) decrease by 50%.

Economic Evaluation Equations

Productivity losses

Productivity losses was calculated using the following equation:

$$PL = PL_{morbidity} + PL_{mortality}$$

Equation a

Where: PL=productivity losses; $PL_{morbidity}$=productivity loss due to morbidity; $PL_{mortality}$=productivity loss due to mortality

$$PL_{morbidity} = PL_{asymptomatic} + PL_{mild} + PL_{severe} + PL_{critical}$$

$$PL_{asymptomatic} = Numberof\ asymptomatic\ cases \times testing\ rate \times GDP\ per\ capita \times quarantine\ period \times informal\ sector\ proportion$$

$$PL_{mild} = Number\ of\ mild\ cases \times testing\ rate \times GDP\ per\ capita \times quarantine\ period \times informal\ sector\ proportion$$

$$PL_{severe} = Number\ of\ severe\ cases \times GDP\ per\ capita \times quarantine\ period$$

$$PL_{critical} = Number\ of\ critical\ cases \times GDP\ per\ capita \times duration\ of\ critical\ disease$$

The testing rate was used to apportion the number of asymptomatic and mild cases who are tested. Further the proportion of informal sector is used to apply lost productivity on asymptomatic/mild cases that are in the informal sector, given the assumption that only those in informal sector are likely not to be productive as they isolate. Lastly, the duration of disease is used where the duration of illness is more than the 14 day quarantine period.

$$PL_{mortality} = YLL \times GDP\ per\ capita$$

Where: YLL=Years of life lost (described below)

Disability adjusted life years (DALYs)

Disability adjusted life years (DALYs) was calculated using the equation:

$$\text{DALYs} = \sum_{a}^{j} YLL_a + \sum_{h}^{k} YLD_h$$

Equation b

Where: a is the age at death; j= number of age groups; h are the health states; k =number of health states.

YLL is estimated as [20]:

$$\text{YLL} = \frac{KCe^{ra}}{(r+\beta)^2}\left\{e^{-(r+\beta)(L+a)}[-(r+\beta)(L+a)-1] - e^{-(r+\beta)a}[-(r+\beta)a-1]\right\} + \frac{1-K}{r}(1-e^{-rL})$$

Equation c

Where: K=age weighting modulating factor; C=adjustment constant for age weights; r=discount rate; a=age of death; β=parameter from the age weighting function; L=standard life expectancy at age of death

Years lost due to disability (YLD) is estimated as [20]:

$$\text{YLDs} = \text{D}\{\frac{KCe^{ra}}{(r+\beta)^2}\{e^{-(r+\beta)(L+a)}[-(r+\beta)(L+a)-1] - e^{-(r+\beta)a}[-(r+\beta)a-1]\} + \frac{1-K}{r}(1-e^{-rL})\}$$

Equation (d)

Where: K=age weighting modulating factor; C=adjustment constant for age weights; r=discount rate; a=age of onset of disability; β=parameter from the age weighting function; L=duration of disability; D=disability weight

REFERENCES

1 Brand SPC, Ojal J, Aziza R, *et al.* COVID-19 transmission dynamics underlying epidemic waves in Kenya. *Science (80-)* Published Online First: 7 October 2021. doi:10.1126/science.abk0414

2 Prem K, van Zandvoort K, Klepac P, *et al.* Projecting contact matrices in 177 geographical regions: An update and comparison with empirical data for the COVID-19 era. *PLOS Comput Biol* 2021;**17**:e1009098. doi:10.1371/JOURNAL.PCBI.1009098

3 Moore S, Hill EM, Dyson L, *et al.* Modelling optimal vaccination strategy for SARS-CoV-2 in the UK. *PLOS Comput Biol* 2021;**17**:e1008849. doi:10.1371/JOURNAL.PCBI.1008849

4 Keeling MJ, Thomas A, Hill EM, *et al.* Waning, Boosting and a Path to Endemicity for SARS-CoV-2. *medRxiv* Published Online First: 10 November 2021. doi:10.1101/2021.11.05.21265977

5 Google L.L.C. COVID-19 Community Mobility Reports. 2022.https://www.google.com/covid19/mobility/ (accessed 4 Mar 2022).

6 Freunde von GISAID e.V. GISAID-Tracking of Variants. 2022.https://www.gisaid.org/hcov19-variants/ (accessed 4 Mar 2022).

7 Githinji G, de Laurent ZR, Mohammed KS, *et al.* Tracking the introduction and spread of SARS-CoV-2 in coastal Kenya. *Nat Commun 2021 121* 2021;**12**:1–10. doi:10.1038/s41467-021-25137-x

8 Lauer SA, Grantz KH, Bi Q, *et al.* The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Ann Intern Med* 2020;**172**:577–82. doi:10.7326/M20-0504

9 Moghadas SM, Shoukat A, Fitzpatrick MC, *et al.* Projecting hospital utilization during the COVID-19 outbreaks in the United States. *Proc Natl Acad Sci U S A* 2020;**117**:9122–6. doi:10.1073/PNAS.2004064117/SUPPL_FILE/PNAS.2004064117.SAPP.PDF

10 Tembo J, Maluzi K, Egbe F, *et al.* Covid-19 in Africa. *BMJ* 2021;**372**. doi:10.1136/BMJ.N457

11 Bakilana A, Demombynes G, Gubbins P. Could COVID-19 deaths in Kenya be higher than official records? 2022. https://blogs.worldbank.org/africacan/could-covid-19-deaths-kenya-be-higher-official-records (accessed 10 Mar 2022).

12 Moore S, Hill EM, Tildesley MJ, *et al.* Vaccination and non-pharmaceutical interventions for COVID-19: a mathematical modelling study. *Lancet Infect Dis* 2021;**21**:793–802. doi:10.1016/S1473-3099(21)00143-2/ATTACHMENT/6721682C-734C-4335-A43B-6F5B618390E1/MMC1.PDF

13 Keeling MJ, Hill EM, Gorsich EE, *et al.* Predictions of COVID-19 dynamics in the UK: Short-term forecasting and analysis of potential exit strategies. *PLOS Comput Biol* 2021;**17**:e1008619. doi:10.1371/JOURNAL.PCBI.1008619

14 Van Wees J-D, Osinga S, Van Der Kuip M, *et al.* Forecasting hospitalization and ICU rates of the COVID-19 outbreak: an efficient SEIR model. *World Heal Organ Bull* Published Online First: 2020. doi:10.2471/BLT.20.251561

15 Rackauckas C, Nie Q. DifferentialEquations.jl – A Performant and Feature-Rich Ecosystem for Solving Differential Equations in Julia. *J Open Res Softw* 2017;**5**:15. doi:10.5334/JORS.151

16 Ngere I, Dawa J, Hunsperger E, *et al.* High seroprevalence of SARS-CoV-2 but low infection fatality ratio eight months after introduction in Nairobi, Kenya. *Int J Infect Dis* 2021;**112**:25–34. doi:10.1016/J.IJID.2021.08.062

17 Neal RM. MCMC Using Hamiltonian Dynamics. In: *Handbook of Markov Chain Monte Carlo*. CRC Press 2011.

18 Hoffman MD, Gelman A. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. *J Mach Learn Res* 2014;**15**:1351–81.http://mcmc-jags.sourceforge.net (accessed 10 Mar 2022).

[dataset]19 Orangi S, Ojal J, Brand SP., *et al.* Data from: Epidemiological impact and cost-

effectiveness analysis of COVID-19 vaccination in Kenya. GitHub Repos. 2022.https://github.com/SamuelBrand1/KenyaCoVaccines (accessed 2 Jun 2022).
20 Fox-Rushby JA, Hanson K. Calculating and presenting disability adjusted life years (DALYs) in cost-effectiveness analysis. *Health Policy Plan* 2001;**16**:326–31. doi:10.1093/HEAPOL/16.3.326